



The 19th National Conference on Computing
and Information Technology

May 18-19 2023

Arnoma Grand Hotel Bangkok, Thailand

2023

Conference
Proceedings



Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
The Association of council of IT Deans

www.nccit.net

TABLE OF CONTENTS

Message from KMUTNB President.....	ii
Message from General Chair	iv
Conference Organizers.....	viii
Conference Organization Committee	x
Steering Committee	xi
Technical Program Committee	xii
Technical Program Contents.....	xiii
Regular Papers	1
Author Index	441

สารจากอธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เทคโนโลยีในโลกปัจจุบันมีการเปลี่ยนแปลงอย่างรวดเร็ว จากยุค Analog ไปสู่ยุค Digital และยุค Robotic จึงทำให้เทคโนโลยีดิจิทัลมีอิทธิพลต่อการดำรงชีวิตและการทำงาน Digital literacy หรือทักษะความเข้าใจและใช้เทคโนโลยีดิจิทัล เทคโนโลยีมีบทบาทสำคัญต่อการศึกษา เพราะเป็นเครื่องมือที่ช่วยสนับสนุนการเรียนการสอนได้อย่างมีประสิทธิภาพและสะดวกสบายมากขึ้น นอกจากนี้ ยังช่วยให้เกิดการเรียนรู้ที่หลากหลายมากขึ้น โดยเฉพาะในสมัยปัจจุบันที่มีการเปลี่ยนแปลงของเทคโนโลยีอย่างรวดเร็ว ตัวอย่างเช่น การใช้เทคโนโลยีในการเรียนรู้ออนไลน์ การใช้แพลตฟอร์มการเรียนรู้ออนไลน์ เว็บไซต์ การเรียนรู้ เป็นต้น ทำให้ผู้เรียนสามารถเรียนรู้ได้ทุกที่ทุกเวลา ไม่ว่าจะอยู่ที่ไหนก็ตาม และสามารถเรียนรู้ได้ตามความต้องการของผู้เรียนเอง

เป็นเรื่องที่น่ายินดีที่เครือข่ายหน่วยงานวิชาการด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศได้ร่วมกันจัดงานประชุมวิชาการด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศอย่างต่อเนื่องทุก ๆ ปี เป็นเวทีวิชาการสำหรับอาจารย์ นักวิจัย และนักศึกษาที่มีผลงานวิจัยนำเสนอและเผยแพร่งานวิจัยดังกล่าวสู่สาธารณะแลกเปลี่ยนความรู้และแสวงหาความร่วมมือ เป็นการส่งเสริมการพัฒนาศักยภาพเพื่อยกระดับคุณภาพการวิจัยด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ พร้อมทั้งจะนำไปประยุกต์ใช้งานจริง ขับเคลื่อนประเทศไปสู่การแข่งขันในภูมิภาค และสากลต่อไป

ท้ายที่สุดนี้ ขอขอบพระคุณกรรมการเครือข่ายหน่วยงานวิชาการ คณะกรรมการจัดงานทุกภาคส่วน โดยเฉพาะคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ที่ได้ร่วมจัดงานประชุมทางวิชาการอย่างต่อเนื่อง ขออวยพรให้การประชุมทางวิชาการครั้งนี้ ประสบความสำเร็จลุล่วงตามวัตถุประสงค์ที่ตั้งไว้ทุกประการ



(ศาสตราจารย์ ดร.สุชาติ เชื้อขนิณ)

อธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

MESSAGE FROM KMUTNB PRESIDENT

Today's world is rapidly moving towards Digital technology and Robotic. This has made digital literacy, or the ability to understand and use digital technology, essential for our work and daily life. Technology plays a crucial role in education as it is a tool that makes learning and teaching more efficient and convenient. Moreover, it helps to promote diverse learning opportunities, especially in the current era of rapid technological change. For example, online learning platforms and websites have made it possible for learners to study anytime and anywhere, according to their own needs and preferences. This is just one of the many ways that technology has transformed education and made learning more accessible and flexible for everyone.

The National Conference on Computing and Information Technology (NCCIT) promotes our Government's policy in digital strategy to support and encourage development of Computer Innovation, Information Technology, and Digital Innovation. NCCIT is a conference for professors, researchers, and students to present and publish their papers, exchange knowledge or seek cooperation. It develops, encourages and enhances Computer and Information Technology research for its applicable usage.

Finally, I would like to thank everyone including the NCCIT collaborative universities, all co-organizers, especially the staff members of the Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology North Bangkok to organize and manage this conference. I wish the conference achieves its purposes.



(Professor Dr.-Ing. habil. Suchart Siengchin)
President, King Mongkut's University of Technology North Bangkok

สารจากประธานจัดงานประชุมวิชาการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ได้เปลี่ยนแปลงวิถีชีวิตของเราไปอย่างมาก เหตุการณ์ดังกล่าวทำให้เราใช้เทคโนโลยีดิจิทัลเพื่อสนับสนุนการดำรงชีวิตประจำวัน ตั้งแต่การอุปโภค บริโภค การศึกษา ไปจนถึงการสื่อสาร การทำงาน การรักษาความสะอาด กิจกรรมทางสังคม ในรูปแบบต่าง ๆ และยังมีผลกระทบอย่างมากต่อภาคธุรกิจและอุตสาหกรรม

งานประชุมวิชาการระดับชาติและระดับนานาชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (NCCIT และ IC2IT) จัดขึ้นอย่างต่อเนื่องเป็นประจำทุกปี ตั้งแต่ พ.ศ. 2548 ในครั้งนี้เป็นการจัดงาน ครั้งที่ 19 เป็นความร่วมมือของหน่วยงานเครือข่ายวิชาการจัดงานประชุมวิชาการ NCCIT และ IC2IT ซึ่งงานประชุมวิชาการนี้เป็นอีกหนึ่งเวทีแห่งการแลกเปลี่ยนแนวความคิด ความรู้ และการวิจัยเพื่อสร้างแรงบันดาลใจ สำหรับการสรรค์สร้างนวัตกรรมดิจิทัลใหม่ ๆ ในปีนี้การประชุมระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ (NCCIT) ครั้งที่ 19 จัดขึ้นระหว่างวันที่ 18-19 พฤษภาคม พ.ศ. 2566 ณ โรงแรมอโนมา แกรนด์ กรุงเทพฯ เล่มรายงานสืบเนื่องฉบับนี้เป็นนำเสนอข้อมูลเชิงหลักการเกี่ยวกับเทคโนโลยีดิจิทัลที่ถูกคัดสรรมาอย่างดีของนักวิจัย นักวิชาการ ที่มีส่วนร่วมในการส่งผลงานเข้าร่วมการประชุมวิชาการในครั้งนี้

ในงานประชุมวิชาการครั้งนี้ต้องขอขอบคุณทุกท่านที่ได้ส่งบทความเพื่อเข้าสู่กระบวนการพิจารณาซึ่งเป็นบทความที่มีคุณภาพสูง ทำให้งานประชุมวิชาการได้รักษาคุณภาพและมาตรฐานของงาน ขอขอบพระคุณคณะกรรมการพิจารณาบทความทุกท่านที่เสียสละเวลาอันมีค่าเพื่อให้ได้ผลงานวิจัยที่มีคุณภาพได้นำเสนอในงานประชุมวิชาการในครั้งนี้ด้วย

ขอขอบคุณมหาวิทยาลัยเครือข่ายทั้ง 6 แห่ง ที่ให้การสนับสนุนเป็นอย่างดีทั้งด้านคณาจารย์ บุคลากรในการร่วมเป็นคณะกรรมการเครือข่าย คณะกรรมการดำเนินงาน และคณะกรรมการพิจารณาบทความ อีกทั้งด้านงบประมาณ ขอขอบพระคุณ ศาสตราจารย์ ดร.สุชาติ เชียงฉิน อธิการบดี มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ สำหรับการสนับสนุนการประชุมวิชาการ NCCIT และ IC2IT อย่างต่อเนื่อง ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์ อาจารย์ ดร.วัชรวิธรรม จิตต์สกุล นางสาวพรพิมล ฝ่ายเทศ และทีมงานจากคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ที่ได้ประสานงานกับนักวิจัยและผู้ที่เกี่ยวข้อง เพื่อรวบรวม ผลงานการวิจัยสำหรับงานประชุมทางวิชาการให้เสร็จสิ้นภายในระยะเวลาที่เหมาะสม

ท้ายนี้หวังเป็นอย่างยิ่งว่างานประชุมวิชาการ NCCIT และ IC²IT จะเป็นเวทีในการนำเสนอผลงานวิจัยและนำไปสู่การพัฒนาตนเองเพื่อให้เป็นนักวิจัยที่ดีและมีคุณภาพสูงยิ่ง ๆ ขึ้นไป NCCIT และ IC²IT จะเป็นแพลตฟอร์มสำหรับการแลกเปลี่ยนความรู้ในด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ สร้างแรงบันดาลใจให้กับนักวิจัยเพื่อคิดค้นนวัตกรรมและพัฒนาองค์ความรู้ใหม่ ๆ เพื่อการพัฒนาชาติอย่างยั่งยืนตลอดไป



(ผู้ช่วยศาสตราจารย์ ดร.สุนันทา สดสี)
คณบดีคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ประธานเครือข่ายงานประชุมวิชาการ NCCIT

MESSAGE FROM GENERAL CHAIR

The COVID-19 pandemic has brought about a transformation in our way of life. It has forced us to utilize digital technology to facilitate our daily activities, from dining and education, to communication, work processes, hygiene, and every social activity. It has also significantly affected businesses and industries.

As a sharing platform, the National Conference on Computing and Information Technology (NCCIT) has been running for nineteen years to serve digital communities in exchanging ideas, knowledge, and research to inspire digital innovations. Herein, the book presents the main contributions of researchers who participated in the 19th National Conference on Computing and Information Technology (NCCIT2023), held in Bangkok, Thailand, on May 18th and 19th of this year. The conference proceedings include carefully selected and reviewed research papers in the areas of machine learning, natural language processing, image processing, robotics, as well as digital technology.

Our heartfelt thanks to our six university partners in Thailand for their financial support and for the contribution of all staff member by serving as members of network committee, executive director committees, and technical program committees.

Huge thanks are extended to Professor Dr.-Ing. habil. Suchart Siengchin, President of King Mongkut's University of Technology North Bangkok (KMUTNB), for his unwavering support for NCCIT & IC²IT and for providing us with the necessary resources.

A very special thank you goes to Assistant Professor Dr. Sakchai Tangwannawit, Dr. Watchareewan Jitsakul, Miss Pornpimon Faythet, and all academic staff of the Faculty of Information Technology and Digital Innovation (ITD), KMUTNB for collaborating, collating and timely production of our quality conference proceedings. Thank you also to NCCIT & IC²IT network universities for providing personnel and budget support.

This conference is made possible with innovative and progressive contributions from the research community at large. Special thanks are due to all our technical committee members for their diligent consideration of all submissions and for maintaining and preserving the high standards for which NCCIT & IC²IT is justifiably renowned. Regrettably, as a consequence of our rigorous peer review process, we were unable to accept some articles for presentation.

Finally, looking to the future, NCCIT & IC²IT will continue to serve as a community platform for research, while raising the bar for high quality academic investigation. It will provide an enduring forum for knowledge exchange in the field of computer and information technology, and it will inspire researchers to develop innovative products, to exchange ideas and to build up a research network geared towards sustainable development.

A handwritten signature in blue ink, reading "Sunantha Sodsee". The signature is written in a cursive style with a blue color.

Assistant Professor Dr. Dr.-Ing. Sunantha Sodsee
Dean, Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok
NCCIT2023 General Chair

CONFERENCE ORGANIZERS



King Mongkut's University of Technology North Bangkok, Thailand
Faculty of Information Technology and Digital Innovation

Fern Universität in Hagen, Germany



Oklahoma State University, USA

Chemnitz University, Germany



Edith Cowan University, Australia

Hanoi National University of Education, Vietnam



Gesellschaft für Informatik (GI), Germany

Kanchanaburi Rajabhat University, Thailand



CONFERENCE ORGANIZERS



Nakhon Pathom Rajabhat University, Thailand

Rajamangala University of Technology Thanyaburi, Thailand
Office of Academic Resources and Information Technology



Mahasarakham University, Thailand
Faculty of Informatics

National Institute of Development Administration, Thailand
Information Technology Management Program,
Graduate School of Applied Statistics



Council of IT Deans of Thailand (CITT), Thailand

CONFERENCE ORGANIZATION COMMITTEE

General Chair: Asst. Prof. Dr. Dr.-Ing. Sunantha Sodsee, KMUTNB

Secretary: Asst. Prof. Dr. Sakchai Tangwannawit, KMUTNB

Assistant Secretary: Dr. Watchareewan Jitsakul, KMUTNB

STEERING COMMITTEE

Assoc. Prof. Dr. Jantima Polpinij, MSU

Asst. Prof. Dr. Photjanee Sukchovna, KRU

Asst. Prof. Dr. Pramote Luenam, NIDA

Asst. Prof. Pongpith Tuenpusa, RMUTT

Dr. Wirat Pinkaew, NPRU

TECHNICAL PROGRAM COMMITTEE

Akkarat Boonyapalanant, KMUTNB	Panana Tangwannawit, PCRU
Alisa Kongthon, KMUTT	Paranya Palawisut, NPRU
Chakadkit Thaenchaikun, PSU	Pattaraporn Warintarawej, PSU
Chalernpol Tapsai, SSRU	Patthanan Tangwannawit, TRU
Chatchawin Namman, UBU	Phayung Meesad, KMUTNB
Jaree Thongkam, MSU	Phollakrit Wongsantisuk, KMUTNB
Jarut Busarathid, PBRU	Pichate Kunakornvong, RMUTT
Jiratti Tengsuthiwat, KMUTNB	Pongpisit Wuttidittachotti, KMUTNB
Kairung Hengproprohm, NPRU	Pramote Kuacharoen, NIDA
Kanchana Viriyapant, KMUTNB	Pudsadee Boonrawd, KMUTNB
Kanita Saengkrajang, PCRU	Sakchai Tangwannawit, KMUTNB
Kriengsak Treeprapin, UBU	Sompong Valuvanathorn, UBU
Kuljaree Tantayakul, PSU	Sucha Smanchat, KMUTNB
Kuntpong Woraratpanya, KMITL	Suchart Khummanee, MSU
Kwankamon Dittakan, PSU	Suksaeng Kukanok, IBASOFT
Laddawan Champa, KRU	Sukuma Uamcharoen, PBRU
Mahasak Ketcham, KMUTNB	Sunantha Sodsee, KMUTNB
Manasawee Kaenampornpan, MSU	Supawadee Hiranpongsin, UBU
Montean Rattanasiriwongwut, KMUTNB	Supoj Hengpraprom, NPRU
Monvorath Phongpaibul, TU	Sutat Gammanee, KRU
Nadh Ditcharoen, UBU	Sutep Tongngam, NIDA
Nalinpat Bhumpenpein, KMUTNB	Suwich Tirakoat, MSU
Narong Phothi, KMUTNB	Tanapon Jentsuttiwetchakul, KMUTNB
Nathaporn Utakrit, KMUTNB	Thanawat Yochanang, KMUTNB
Nattapong Tongtep, PSU	Tongpool Heeptaisong, KMUTNB
Nattavee Utakrit, KMUTNB	Vatinee Nuijian, KMUTNB
Nongluk Promthong, RMUTT	Watchareewan Jitsakul, KMUTNB
Noppon Lertchuwongsa, PSU	Weena Janratchakool, RMUTT
Olarik Surinta, MSU	Wongkot Sriurai, UBU
Paleerat Wongchampa, KMUTNB	

TECHNICAL PROGRAM CONTENTS

Program Schedule

Thursday 18th May 2023	
08:00 - 09:00	Registration
09:00 - 09:15	Report by Assistant Professor Dr. Dr.-Ing. Sunantha Sodsee , <i>Dean, Faculty of Information Technology and Digital Innovation</i> <i>King Mongkut's University of Technology North Bangkok</i>
09:15 - 09:30	Opening Ceremony by Professor Dr. -Ing. habil. Suchart Siengchin , <i>President of King Mongkut's University of Technology North Bangkok</i>
09:30 - 09:45	<i>Group Photo</i>
09:45 - 10:45	Invited Keynote Speech by Professor Dr. Stephan Pareigis , <ul style="list-style-type: none"> • Department Informatik, Professor für Angewandte Mathematik und Technische Informatik, Germany.
10:45 - 11:00	<i>Coffee break</i>
11:00 - 12:00	Invited Keynote Speech by Dr. Winn Voravuthikunchai , <ul style="list-style-type: none"> • PhD in AI from Caen University, France, Founder & CEO of Botnoi Group, Thailand
12:00 - 13:00	<i>Lunch</i>
13:00 - 15:00	Paper Presentation
15:00 - 15:20	<i>Coffee break</i>
15:20 - 16:40	Paper Presentation
18:00 - 22:00	Welcome Dinner and Best Paper Award Presentation

Thursday 18th May 2023		
ROOM: I (Data Mining, Text Mining and Related Fields)		Page
13:00 – 13:20 NCCIT2023-002	Design of machine learning for Lime Classification based upon Thai Agricultural Standard No. TAS 27-2017 <i>Alongkron Klaiklueng and Athakorn Kengpols</i>	1
13:20 – 13:40 NCCIT2023-011	Predicting the Master’s Educational Achievement using Different Models <i>Unhawa Ninrutsirikun, Aphorn Chiawchankaset, Sanit Sirisawatvatana, and Vithida Chongsuphajaisiddhi</i>	8
13:40 – 14:00 NCCIT2023-023	User Feedback Analytics for Supporting Mobile Application Maintainability <i>Sujaree Thairat and Monvorath Phongpaibul</i>	15
14:00 – 14:20 NCCIT2023-029	Alzheimer’s Disease Classification Using Two-Dimensional Variational Mode Decomposition and Principal Component Analysis with Entropy <i>Sittisak Saechueng, Ungsumalee Suttapakti, Aekapop Bunpeng, and Wittawas Puntumchinda</i>	22
14:20 – 14:40 NCCIT2023-041	Development of a Prototype Health Care Recommender Application with Machine Learning Techniques <i>Uraivan Inyaem, Chamaiphon Yeamklin, and Siriwarn Seemo</i>	29
14:40 – 15:00 NCCIT2023-053	A Model for Analysis of Cyber Attack Behaviors Using Artificial Intelligence in the Royal Thai Air Force <i>Natdanai Kongkhunthod, Prasong Praneetpolgrang, and Payap Sirinam</i>	36
	<i>End Session</i>	

Thursday 18th May 2023		
ROOM: II (Data Mining, Text Mining and Related Fields)		Page
13:00 – 13:20 NCCIT2023-001	Application of Bat Search to Design PID Controller for Time Delay System <i>Paitoon Janreung and Danupon Kumpanya</i>	42
13:20 – 13:40 NCCIT2023-004	Automated Non-Functional Properties Testing Tool for GraphQL API <i>Pattadon Prongpreecha and Nuengwong Tuaycharoen</i>	49
13:40 – 14:00 NCCIT2023-054	Forecast of Electric Consumption in Thailand using Data Mining Technique <i>Mintra Mesawong and Tanapon Jentsuttiwetchakul</i>	55
14:00 – 14:20 NCCIT2023-056	Forecasting of Climate Change in Rice Planting Areas in Northeastern Thailand with Time Series Analysis Techniques <i>Kankawin Chamchoi and Nattavee Utakrit</i>	61
14:20 – 14:40 NCCIT2023-065	Study and Development of Business Intelligence System for Forecasting the Demand for Herb Raw Materials, A Case Study of the Chaophya Abhaibhubejhr Hospital Foundation <i>Pasit Thong-on and Montean Rattanasiriwongwut</i>	67
14:40 – 15:00 NCCIT2023-077	Forecasting System for Repayment Through Online Payment Using Time Series Techniques <i>Yongyut Intarpasit and Watchareewan Jitsakul</i>	73
<i>End Session</i>		

Thursday 18th May 2023		
ROOM: III (Image Processing, Internet of Things, Blockchain and Related Fields)		Page
13:00 – 13:20 NCCIT2023-005	Detecting Spammer Groups Using Centrality Measure <i>Natthanicha Suriyamongkol and Sukree Sinthupinyo</i>	79
13:20 – 13:40 NCCIT2023-021	The Development of Artificial Intelligence for Face Mask Detection System in Chantaburi Province <i>Sathit Suwannawach, Thipwan Fufaung, and Khanittha Sittitiamjan</i>	85
13:40 – 14:00 NCCIT2023-025	Performance Enhancement for Environmental Monitoring with Graduated Change Situation in Edge Computing <i>Theerapat Deeprasert, Aphirak Jansang, Chaiporn Jaikaeo, Withawat Tangtrongpairoj, and Anan Phonphoem</i>	91
14:00 – 14:20 NCCIT2023-031	Data Security Using Encryption with Image Steganography <i>Thanawat Yochanang</i>	97
14:20 – 14:40 NCCIT2023-051	Design and Development of a Data Storage and Exchange Systems for the Air Force Command and Control Systems Using Blockchain Technology <i>Thanakrit Pengkian, Prasong Praneetpolgrang, and Payap Sirinam</i>	103
14:40 – 15:00 NCCIT2023-052	Development of IoT System for Cultivation of Cordyceps militaris <i>Samatachai Jantararat, Jedsada Arunruerk, and Patrapee Sunantapot</i>	110
15:00 – 15:20	<i>Break</i>	
15:20 – 15:40 NCCIT2023-061	Applying Image Processing and Natural Language Understanding with LINE Chatbot for Rabbit Caring <i>Preeyarwat Tamawimok and Natavee Utakrit</i>	116
15:40 – 16:00 NCCIT2023-062	Development of Life Jacket to Monitor Water Accident Victim with the Internet of Things <i>Somchoke Kittisitthipong and Mahasak Ketcham</i>	123
	<i>End Session</i>	

Thursday 18th May 2023		
ROOM: IV (Acceptance, Innovation, Adoption and Related Fields)		Page
13:00 – 13:20 NCCIT2023-022	Acceptance Factors for the use of Clubhouse as the Communication Channel and the Expression of Thoughts <i>Rachata Sriaroon and Pramote Luenam</i>	130
13:20 – 13:40 NCCIT2023-024	Factors Affecting the Continuance Intention to Use of QR Payment Services <i>Napatkanok Srikanjanodom and Pramote Luenam</i>	138
13:40 – 14:00 NCCIT2023-055	Factory Affecting User’s Perceive Benefits of Personal Income Tax Payment Through Internet <i>Apisara Onthao and Tanapon Jentsuttiwetchakul</i>	145
14:00 – 14:20 NCCIT2023-064	Factors Analysis Decision Making in Selecting Institutions Faculty of Information Technology and Digital Innovation <i>Chonlada Chuchomchuen and Montean Rattanasiriwongwut</i>	151
14:20 – 14:40 NCCIT2023-075	Factors Affecting the Adoption of Automatic Beverage Dispenser Technology <i>Supichaya Jitarom and Tanapon Jentsuttiwetchakul</i>	158
14:40 – 15:00 NCCIT2023-078	Factors Affecting Attitude of Data Privacy Acceptance of Mobile Banking Users <i>Walailak Gattaisong and Sakchai Tangwannawit</i>	165
	<i>End Session</i>	

Thursday 18th May 2023		
ROOM: V (Information System and Related Fields)		Page
13:00 – 13:20 NCCIT2023-003	Development of an Automatic Tools for Modeling the Relationships of RESTful API Web Service <i>Wipada Klungtes and Nuengwong Tuaycharoen</i>	172
13:20 – 13:40 NCCIT2023-020	Test Automation Report using BDD with Gherkin <i>Jare Sophanpanitchakul and Taratip Suwannasart</i>	179
13:40 – 14:00 NCCIT2023-026	Development of Personalized Programming Tutoring Systems <i>Todsapon Banklongsi and Charun Sanrach</i>	185
14:00 – 14:20 NCCIT2023-030	The Checking the Parking with LINE Application <i>Sittichop Ratrojkij, Pattarapong Jarrusrojoka, Thanaphon Yochana, Orawan Chunhapran, and Rath Burirat</i>	191
14:20 – 14:40 NCCIT2023-033	Application for Training Artificial Intelligence Skills in the Context of Employees in Kanchanaburi Province <i>Sutat Gammanee, Narongdech Rattananonsathien, and Kwannaree Klaprabchone</i>	197
14:40 – 15:00 NCCIT2023-035	User Experience Design of Smart TV Applications for the Elderly in Thailand <i>Sittichoke Sakon and Waraporn Jirapanthong</i>	204
15:00 – 15:20	<i>Break</i>	
15:20 – 15:40 NCCIT2023-037	Automatic User Interface Testing for Web Application with Cypress Framework: A Case Study of an Online Assessment Web Application <i>Tatchanon Srisamudh, Atiwong Suchato, Dittaya Wanvarie, Naruemon Pratanwanich, and Nuengwong Tuaycharoen</i>	212
15:40 – 16:00 NCCIT2023-045	The Development of Personal Data Protection Act Function in Enterprise Resource Planning System <i>Chamaiporn Peesiri and Sakchai Tangwannawit</i>	219
	<i>End Session</i>	

Program Schedule

Friday 19th May 2023	
09:00 - 10:20	Paper Presentation
10:20 - 10:40	<i>Break</i>
10:40 - 12:00	Paper Presentation

Friday 19th May 2023		
ROOM: I (Data Mining, Text Mining, and Related Fields)		Page
09:00 – 09:20 NCCIT2023-049	Text Classification for Automatic Problem Reporting <i>Puwanart Poopankul and Phayung Meesad</i>	226
09:20 – 09:40 NCCIT2023-067	Classification of Packet Pattern for Network Intrusion by Using Data Mining <i>Nichapha Saewiang and Montean Rattanasiriwongwut</i>	232
09:40 – 10:00 NCCIT2023-069	Classifying Infectious Disease Symptoms based on ICD-10 Standards using Deep Learning <i>Duangrudee Phetphirun and Watchareewan Jitsakul</i>	239
10:00 – 10:20 NCCIT2023-071	Deep Learning Techniques Comparison for Analyzing Opinions of Thailand Post Service User <i>Wannakorn Puenguthaisri and Watchareewan Jitsakul</i>	245
10:20 – 10:40	<i>Break</i>	
10:40 – 11:00 NCCIT2023-072	Comparison of Algorithms for Face Recognition <i>Nattapong Jantarasorn and Watchareewan Jitsakul</i>	251
11:00 – 11:20 NCCIT2023-073	Predictive of Outstanding Debts for Housing Loan by Data Mining <i>Kamonchanok Rungjaroen and Watchareewan Jitsakul</i>	257
11:20 – 11:40 NCCIT2023-074	A Comparison of Algorithms for Export Value Forecast of Thai Small and Medium Enterprises <i>Apisak Suno and Watchareewan Jitsakul</i>	263
11:40 – 12:00 NCCIT2023-082	Examination of The Social Security Office's Dental Benefits Disbursement Behavior by Data Mining Techniques <i>Chawarus Kerot and Watchareewan Jitsakul</i>	270
	<i>End Session</i>	

Friday 19th May 2023		
ROOM: II (Data Mining, Text Mining, and Related Fields)		Page
09:00 – 09:20 NCCIT2023-010	A Comparison of Algorithms for Classifying Securities in the Stock Exchange of Thailand <i>Fathip Wanhom and Tanapon Jentsuttiwetchakul</i>	276
09:20 – 09:40 NCCIT2023-034	Forecasting Model for The Number of Patients in The Public Hospital by using Data Mining Techniques <i>Nuttaya Sangnark and Sakchai Tangwannawit</i>	282
09:40 – 10:00 NCCIT2023-057	Development of a Tourist Attraction Recommendation System with Machine Learning Technology <i>Supatsara Krasaesat and Montean Rattanasiriwongwut</i>	288
10:00 – 10:20 NCCIT2023-066	Development of Environmental Quality Measurement Systems via Monitoring and Evaluating with Real-time Notification on Application <i>Marisa Sangnul and Mahasak Ketcham</i>	294
10:20 – 10:40	<i>Break</i>	
10:40 – 11:00 NCCIT2023-086	Classification of Broadband Network Devices Using Text Mining Technique <i>Nattapat Sridoung and Mahasak Ketcham</i>	300
11:00 – 11:20 NCCIT2023-088	Forecasting System of Oil Refinery using Time Series Techniques <i>Ratchaphoom Boonnaka and Watchareewan Jitsakul</i>	307
11:20 – 11:40 NCCIT2023-091	The Comparison of Algorithms Performance to Analysis Hotel Guest Reviews <i>Santi Sukkasem and Watchareewan Jitsakul</i>	313
11:40 – 12:00 NCCIT2023-092	Comparison of Algorithm Performance for Forecasting Thailand's Inflation <i>Onuma Jamuang and Watchareewan Jitsakul</i>	309
	<i>End Session</i>	

Friday 19th May 2023		
ROOM: III (Internet of Things, Network and Related Fields)		Page
09:00 – 09:20 NCCIT2023-063	Co-Farming Platform using Internet of Things Technology <i>Lertsak Hongjan and Chatchawin Namman</i>	326
09:20 – 09:40 NCCIT2023-083	Comparison of Image Classification of Buildings with Deep Learning Techniques <i>Piradate Tolornuwat and Watchareewan Jitsakul</i>	333
09:40 – 10:00 NCCIT2023-084	Forecasting the Number of Victims from Road Accidents in Bangkok Using Time-Series Analysis Techniques <i>Suphannara Siwaratchakun and Mahasak Ketcham</i>	339
10:00 – 10:20 NCCIT2023-089	Development of Chatbot Application Supports Computer Related Tasks in the Organization <i>Perm Ratanaphirom and Montean Rattanasiriwongwut</i>	345
10:20 – 10:40	<i>Break</i>	
10:40 – 11:00 NCCIT2023-093	An Analytic System for Pipe Breakages Using Association Rules and Web Map Services <i>Kantana Pooltong and Pudsadee Boonrawd</i>	353
11:00 – 11:20 NCCIT2023-094	Customer Loyalty Programs Using Decentralized Application <i>Pratin Sonnekorn and Pudsadee Boonrawd</i>	359
11:20 – 11:40 NCCIT2023-095	Laser Marking System Using Optical Character Recognition and Application Program Interface <i>Panumart Inchee and Pudsadee Boonrawd</i>	365
	<i>End Session</i>	

Friday 19th May 2023		
ROOM: IV (Information System and Related Fields)		Page
09:00 – 09:20 NCCIT2023-047	The Development of Guideline for Software Testing by Using Black-Box Techniques for Standard Functions <i>Sarawadee Trakunhun and Sakchai Tangwannawit</i>	371
09:20 – 09:40 NCCIT2023-058	Retirement Mutual Fund Investment Performance Comparison System with Dollar Cost Averaging Strategy and Lump Sum Strategy using Business Intelligence <i>Apichart Kiatyudthachart and Nattavee Utakrit</i>	378
09:40 – 10:00 NCCIT2023-059	A Development of Project Resource Management <i>Nattapol Puninagoon and Nattavee Utakrit</i>	385
10:00 – 10:20 NCCIT2023-060	Development of Purchase and Inventory Management System on Smartphones for a Small Grocery Store <i>Surachat Aunsorn and Nattavee Utakrit</i>	391
10:20 – 10:40	<i>Break</i>	
10:40 – 11:00 NCCIT2023-070	A Game for Dementia Risk Reduction in Elders <i>Sunitta Yentour, Nathee Yongyut, Konggrit Pitanon, and Dechawut Wanichsan</i>	397
11:00 – 11:20 NCCIT2023-076	A Smart LINE Chatbot and Techniques for Auto Correct <i>Jirasak Pipatwarakul and Phayung Meesad</i>	404
11:20 – 11:40 NCCIT2023-090	The Information System for the Strategic Project of Rajabhat University for Local Development: Phuket Rajabhat University <i>Donrat Khongharn and Sakchai Tangwannawit</i>	411
	<i>End Session</i>	

Friday 19th May 2023		
ROOM: V (Network and Related Fields)		Page
09:00 – 09:20 NCCIT2023-006	Continuous Authentication for Library Services Using Facial Recognition <i>Premagorn Thaweehan and Pramote Kuacharoene</i>	417
09:20 – 09:40 NCCIT2023-028	Dynamic Allocation Algorithm for 802.11ah <i>Kanut Injaiea, Aphirak Jansang, Chaiporn Jaikaeo, Withawat Tangtrongpairoj, and Anan Phonphoem</i>	423
09:40 – 10:00 NCCIT2023-050	Cybersecurity Awareness, Readiness and Indicator in the Armed Force Security Center, Royal Thai Armed Forces Headquarters <i>Todsaporn Klearwun, Surasak Mungsing, Chanapatt Pattaramaetakul, Prasan Wongkitisopon, and Prasong Praneetpolgrang</i>	429
10:00 – 10:20 NCCIT2023-085	Development of an ECG Machine to Diagnose Arrhythmias <i>Supeerawong Pothikanit and Mahasak Ketcham</i>	435
<i>End Session</i>		

REGULAR PAPERS

**การออกแบบระบบคัดแยกมะนาวโดยใช้เครื่องจักรเรียนรู้
ตามมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560
Design of machine learning for Lime Classification
based upon Thai Agricultural Standard No. TAS 27-2017**

อลงกรณ์ คล้ายคลึง (Alongkron Klaiklueng)¹ และอรรถกร เก่งพล (Athakorn Kengpol)^{1,2}

¹ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

²ศูนย์ระบบการจัดการขั้นสูงทางวิศวกรรมอุตสาหการ ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

alongkorn.09k@gmail.com, athakorn@kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) คัดแยกมะนาว โดยอ้างอิงตามมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560 และ 2) พัฒนาเทคโนโลยีที่ใช้ในการคัดแยกมะนาวให้มีประสิทธิภาพและแม่นยำมากขึ้น โดยการนำเครื่องจักรมาใช้แทนตาคน ด้วยการออกแบบระบบคัดแยกมะนาว โดยการใช้เครื่องจักรเรียนรู้ตามมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560 และแสดงผลการคัดแยกผ่านหน้าจอ ผู้ใช้งานแบบกราฟิก งานวิจัยนี้ศึกษาเฉพาะมะนาวพันธุ์ แป้น และแบ่งชั้นคุณภาพของมะนาวเป็น 4 ระดับ คือ ชั้นพิเศษ ชั้นหนึ่งและชั้นสอง มะนาวสุก และมะนาวเน่า การออกแบบระบบคัดแยกมะนาวในงานวิจัยนี้ใช้ตัวแบบ โครงข่ายประสาทเทียมแบบคอนโวลูชัน คือ Visual Geometry Group 19: VGG19 ในการคัดแยกมะนาว โดยใช้เทคนิคการประมวลผลภาพ รวมทั้งใช้ตัวแบบ คณิตศาสตร์ในรูปแบบสมการลอจิสติกเชิงอันดับ ในการแสดงความสัมพันธ์ระหว่างค่าคุณลักษณะต่าง ๆ ของมะนาว ได้แก่ ค่าความหนาแน่นของผล ค่าความสว่างของผิว ค่าสีของผิว และค่าระดับรอยตำหนิของผิว ซึ่งส่งผลต่อการจัดชั้นคุณภาพของมะนาว โดยค่าของ คุณลักษณะที่ใช้ในงานวิจัยเป็นค่าที่ได้จากการวัดจริง เมื่อนำระบบคัดแยกมะนาวที่ออกแบบนี้มาทดสอบการใช้งานจริง โดยใช้ข้อมูลที่ระบบนี้ไม่เคยเรียนรู้มาก่อนพบว่า ระบบนี้สามารถคัดแยกมะนาวแทนการใช้ตาคน โดยสามารถแสดงผลการคัดแยก ได้แก่ ภาพและระดับชั้น

คุณภาพของมะนาว ค่าคุณลักษณะของมะนาว ในชั้นคุณภาพนั้น พร้อมกับคำแนะนำ เกี่ยวกับระยะเวลา ในการเก็บรักษาและการบริโภคผ่านหน้าจอผู้ใช้งานแบบ กราฟิกได้ถูกต้องโดยมีค่าความแม่นยำสูงกว่าร้อยละ 90
คำสำคัญ: การคัดแยกมะนาว เครื่องจักรเรียนรู้ โครงข่ายประสาทเทียมแบบคอนโวลูชัน มะนาว

Abstract

This research objectives are 1) to classify limes with reference to Thai Agricultural Standards No. TAS 27-2560 and 2) to develop Technology for lime classification to be more efficiency and accurate by using the machine instead of human eyes. That will be implemented by designing machine learning for lime classification based upon Thai Agricultural Standard No. TAS 27-2017 and displaying the results through a Graphical User Interface screen. This research studied only lime cultivar Paen and the quality of limes were categorized into 4 levels: Extra Class, Class I and Class II, Ripe lime and Rotten lime. The design of lime classification system in this research used a Convolutional Neural Network model, Visual Geometry Group 19: VGG19, to classify limes by using image processing techniques, including used mathematical model in the form of Order Logistic Regression equation was used to show the relationship between the lime characteristic values, such as fruit density, skin brightness, skin color, and skin defects,

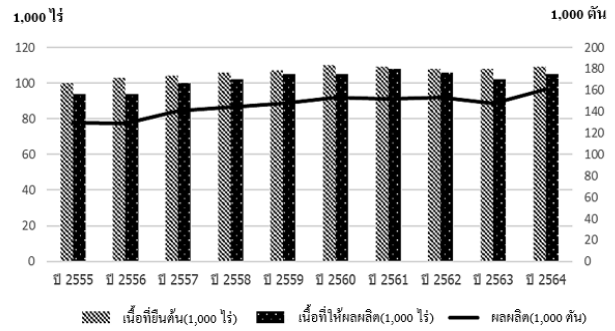
which affected quality classes of lime. The characteristic values used in the research are values derived from actual measurements. When the lime classification system designed in this research was tested for the actual use by using information that system had never learned before, it was found that the system can classify limes instead of human eyes. This system can display the classification results, including pictures and lime quality classes, lime characteristic values in that class with instructions about the retention period and consumption through the graphical user interface accurately with an accuracy more than 90%.

Keywords: Lime Classification, Machine Learning, Convolutional Neural Network, Lime

1. ความเป็นมาและความสำคัญของปัญหา

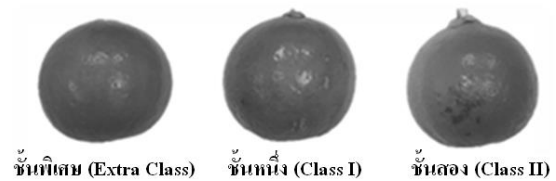
มะนาว (Lime) เป็นพืชเศรษฐกิจที่มีความสำคัญ เพราะสามารถนำไปใช้ประโยชน์ได้หลากหลาย เช่น ใช้ปรุงอาหาร ใช้เป็นส่วนผสมของผลิตภัณฑ์ยา และใช้เป็นวัตถุดิบในการผลิตสินค้า เช่น กรดซิตริก ผลิตภัณฑ์ทำความสะอาด เครื่องสำอาง และอื่น ๆ [1] เป็นต้น

มะนาวเป็นพืชที่ปลูกได้ในทุกภูมิภาคของประเทศไทย โดยเกษตรกรได้มีการขยายเนื้อที่ปลูกมะนาวอย่างต่อเนื่อง ข้อมูลจากสำนักงานเศรษฐกิจการเกษตร [2] ระบุว่า ในปี 2564 ประเทศไทยมีเนื้อที่ปลูกมะนาว 109,000 ไร่ เป็นเนื้อที่ซึ่งให้ ผลผลิตแล้ว 105,000 ไร่ คิดเป็นร้อยละ 96.33 ของเนื้อที่ปลูกมะนาวทั้งหมด โดยมีปริมาณผลผลิตรวม 161,000 ตัน คิดเป็นมูลค่าทั้งสิ้น 5,940 ล้านบาท ดังแสดงในภาพที่ 1 โดยร้อยละ 74.6 ของเนื้อที่ปลูกมะนาวทั้งหมดใช้ปลูกมะนาวแป้น [3] ดังนั้น งานวิจัยนี้จึงเลือกศึกษาเฉพาะมะนาวพันธุ์แป้นเพราะเป็นพันธุ์ที่เกษตรกรปลูกกันอย่างแพร่หลาย และเป็นที่ต้องการของตลาด เนื่องจากประชาชนนิยมบริโภค



ภาพที่ 1: กราฟเนื้อที่และผลผลิตมะนาว ปี 2555-2564

ปัจจุบันผลมะนาวพันธุ์ที่ผลิตเป็นการค้าซึ่งรวมถึงมะนาวพันธุ์แป้นต้องปฏิบัติตามมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560 ซึ่งกำหนดให้แบ่งชั้นคุณภาพของมะนาวดังกล่าวเป็น 3 ระดับ คือ ชั้นพิเศษ (Extra Class) ชั้นหนึ่ง (Class I) และชั้นสอง (Class II) [4] ซึ่งมาตรฐานสินค้าเกษตรนี้ไม่ครอบคลุมผลมะนาวที่ใช้แปรรูปในอุตสาหกรรม



ภาพที่ 2: ตัวอย่างลักษณะมะนาวตามชั้นคุณภาพ

การคัดแยกมะนาวตามมาตรฐานดังกล่าวในปัจจุบันส่วนใหญ่จะคัดแยกโดยใช้การมองด้วยตาของคนเป็นหลัก ซึ่งผู้คัดแยกจะใช้ดุลยพินิจของตนในการคัดแยก ส่งผลให้การคัดแยกไม่เป็นมาตรฐานเดียวกัน และอาจผิดพลาด หรือล่าช้า ดังนั้น จึงจำเป็นต้องมีการพัฒนาเทคโนโลยีที่ใช้ในการคัดแยกมะนาวด้วยการนำเครื่องจักรมาใช้แทนคน เพื่อเพิ่มประสิทธิภาพในการคัดแยก และแก้ปัญหาการคัดแยกมะนาวที่เป็นอยู่

ดังนั้น จึงมีแนวคิดในการออกแบบระบบคัดแยกมะนาวโดยใช้เครื่องจักรเรียนรู้ (Machine Learning) และแสดงผลการคัดแยกผ่านหน้าจอผู้ใช้งานแบบกราฟิก (Graphical User Interface) เพื่อนำระบบนี้มาใช้ในการคัดแยกมะนาวแทนการใช้ตาของคน และเนื่องจาก Convolutional Neural Networks: CNN มีความสามารถในการเรียนรู้ภาพและ

สามารถคัดแยกภาพได้คือ จึงได้รับความสนใจและมีการนำไปประยุกต์ใช้ในการคัดแยกผลไม้โดยใช้เทคนิคการประมวลผลภาพ ประกอบกับการใช้วิธีนี้ได้ผลลัพธ์ที่ดี จึงได้รับความนิยมเป็นอย่างมากนับตั้งแต่ปี 2018 เป็นต้นมา นอกจากนี้ยังมีงานวิจัยที่ใช้ตัวแบบคณิตศาสตร์ในการแสดงความสัมพันธ์ระหว่างคุณลักษณะของผลไม้ ได้แก่ ความหนาแน่นของผล ความสว่างของผิว สีของผิว และค่าระดับรอยตำหนิของผิว กับคุณภาพของผลไม้ที่เหมาะสมกับการบริโภคหรือการขนส่ง จึงนำ Convolutional Neural Networks และตัวแบบคณิตศาสตร์มาใช้ในการออกแบบระบบคัดแยกมะนาวในงานวิจัยนี้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เครื่องจักรเรียนรู้

เครื่องจักรเรียนรู้ (Machine Learning) [5] เป็นการศึกษาอัลกอริทึมของคอมพิวเตอร์ที่มีการพัฒนาโดยการเรียนรู้ของเครื่องจักร [8] เพื่อทำนายองค์ความรู้ โดยปราศจากการป้อนคำสั่งให้กับเครื่องจักรในการประมวลผล

2.2 Convolutional Neural Network

Convolutional Neural Network [6] คือ โครงข่ายประสาทเทียมแบบหนึ่งของการเรียนรู้เชิงลึก (Deep Learning) เหมาะกับการวิเคราะห์ภาพ โดยใช้คอมพิวเตอร์ในการจดจำวัตถุ (Object Recognition) โดย Convolutional Neural Network [7] จะจำลองการมองเห็นออกเป็นพื้นที่ขนาดเล็กหลายส่วน และนำมารวมกันเพื่อแยกว่าสิ่งนั้นคืออะไร โดยในการแยกจะเรียกว่า การแยกคุณสมบัติ (Extract Feature) ของคุณสมบัติในภาพมาวิเคราะห์ร่วมกัน และทำนายว่าภาพนั้น คือ ภาพอะไร [8] โดยงานวิจัยนี้ได้ นำ Convolutional Neural Network มาใช้ในการออกแบบระบบที่ใช้ในการคัดแยกมะนาวเพื่อเพิ่มความสามารถและความแม่นยำในการคัดแยกยิ่งขึ้น

2.3 Ordinal Logistic Regression

Ordinal Logistic Regression [9] เป็นวิธีการทางสถิติที่ใช้อ้างอิงถึงตัวแบบการถดถอยที่ตัวแปรตามมีการแจกแจงแบบพหุนาม (Multinomial Distribution) ที่เป็นอิสระจากกัน และมีลักษณะข้อมูลเป็นลำดับที่ต่อเนื่องกัน โดยระยะห่างของแต่ละลำดับที่ต่อเนื่องกัน ไม่จำเป็นต้องเท่ากัน

แบบเรียงลำดับ โดยมีข้อกำหนดว่า (1) สัมประสิทธิ์ถดถอยของตัวแบบแต่ละกลุ่มจะต้องเท่ากัน (2) ข้อมูลตัวแปรอิสระที่วิเคราะห์มีลักษณะข้อมูลเป็นข้อมูลเชิงคุณภาพ (3) ข้อมูลเชิงปริมาณเป็นตัวแปรร่วม (Covariate Variables) และ (4) เงื่อนไขที่ใช้เป็นเกณฑ์ในการคัดแยกมีมากกว่า 2 เงื่อนไข ดังนั้น วิธีการ Ordinal Logistic Regression จึงเหมาะสมที่จะนำมาใช้ในการศึกษานี้ โดยมีรูปแบบสมการตัวแบบคณิตศาสตร์ ดังนี้

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \alpha_i + \beta x \quad (1)$$

โดย π_i คือ ความน่าจะเป็นสะสมกลุ่มที่ i

β_0 คือ ค่าคงที่กลุ่มตัวแปรตามที่ i

i คือ จำนวนระดับของตัวแปรตาม $1, 2, 3, \dots, j-1$

β_x คือ $\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j$

β คือ จำนวนสัมประสิทธิ์การถดถอย

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ได้ทบทวนงานวิจัยที่ศึกษาเกี่ยวกับการคัดแยกผลผลิตทางการเกษตร โดยใช้คุณลักษณะภายนอก อาทิ รูปทรง สี และตำหนิเป็นเกณฑ์ สามารถสรุปสาระสำคัญได้ ดังนี้

Iqbal และคณะ (2016) [10] ได้ศึกษาการคัดแยกผลไม้ 3 ชนิด ได้แก่ ส้ม มะนาวหวาน และเลมอน ตามระดับความสุกคือ ดิบ (Unripe) สุกปานกลาง (Moderately Ripe) และสุก (Ripe) พบว่า การวิเคราะห์สีจากภาพผลไม้มุมมองเดียว โดยใช้วิธีหาค่าระยะห่างของสี (Color Distance) และการวิเคราะห์การจำแนกเชิงเส้น (Linear Discriminant Analysis) มีความแม่นยำในการคัดแยกผลไม้นี้สูงกว่าร้อยละ 90 แสดงว่า สีของผลไม้เป็นพารามิเตอร์สำคัญที่สามารถนำมาใช้คัดแยกและชี้วัดคุณภาพของผลไม้

Kumar T K (2019) [11] ได้ศึกษาการคัดแยกประเภทผลไม้นตามชั้นคุณภาพ ด้วยวิธีให้เครื่องจักรเรียนรู้ และการถดถอยลอจิสติก (Logistic Regression) ในการฝึกและทดสอบการคัดแยก โดยใช้ภาพจากฐานข้อมูลภาพ Fruit-360 ในการฝึก 490 ภาพ และในการทดสอบ 196 ภาพ พบว่า ผลการฝึกและทดสอบให้ค่าความแม่นยำในการคัดแยกร้อยละ 97.5-100 โดยมีข้อสังเกตว่า ค่าความแม่นยำที่ได้จะลดลงในกรณีที่มีขนาดเล็กลง ดังนั้น จึงควรใช้ภาพที่มีขนาดเหมาะสม

เพื่อป้องกันมิให้การค้นหาลักษณะภาพของตัวแบบได้รับผลกระทบอันเนื่องมาจากภาพมีขนาดเล็กเกินไป

Ngugi และคณะ (2020) [12] ได้เปรียบเทียบความสามารถในการตรวจหาโรคของใบพืชของ Convolutional Neural Network จำนวน 10 ตัวแบบพบว่า Deeper Models ได้แก่ DenseNet201 VGG16 และ ResNet-101 มีความแม่นยำสูง เมื่อเปรียบเทียบกับ Shallower Networks โดย DenseNet201 มีความสามารถดีกว่าตัวแบบอื่นทั้งหมด อย่างไรก็ตาม การนำตัวแบบใดมาใช้ ต้องคำนึงถึงขนาด ความสามารถ และพื้นที่เก็บข้อมูลรวมทั้งระยะเวลาในการฝึกตัวแบบประกอบด้วย

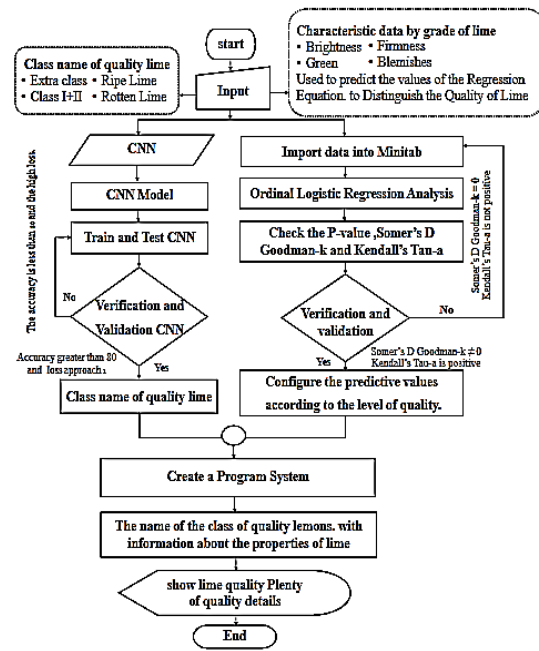
Naranjo-Torres และคณะ (2020) [13] ได้ทบทวนงานวิจัยในช่วงปี 2015-2020 ที่ศึกษาเกี่ยวกับการใช้ Convolutional Neural Network ในการประมวลผลภาพผลไม้ โดยพบว่า Convolutional Neural Network มีความสามารถในการคัดแยกผลไม้สูง และ 6 ตัวแบบที่ใช้ในงานวิจัยดังกล่าวสูง คือ Own Models VGG Alexnet Googlenet Resnet และ LeNet โดยมีสัดส่วนร้อยละ 29.2, 25.0, 12.5, 12.5, 4.2 และ 16.7 ของตัวแบบ Convolutional Neural Network ทั้งหมดที่ใช้ในงานวิจัยดังกล่าวตามลำดับ โดยการนำตัวแบบเหล่านี้มาใช้ในงานวิจัยควรเลือกตัวแบบที่มีความซับซ้อนสอดคล้องกับลักษณะงานและปริมาณข้อมูลของงานวิจัย

การทบทวนงานวิจัยดังกล่าว สรุปได้ว่า การนำตัวแบบ Convolutional Neural Network มาใช้คัดแยกมะนาวในงานวิจัย โดยใช้เทคนิคการประมวลผลภาพควรให้ความสำคัญกับการใช้ภาพที่มีขนาดเหมาะสม ขนาด ความสามารถ และระยะเวลาในการฝึกของตัวแบบ ปริมาณของข้อมูลและพื้นที่ในการเก็บข้อมูล รวมทั้งลักษณะงาน นอกจากนี้ คุณลักษณะของผลไม้เป็นพารามิเตอร์ที่ดีสำหรับใช้คัดแยกและแสดงคุณภาพมะนาวว่าเป็นมะนาวในระดับชั้นคุณภาพใด

3. วิธีการดำเนินงาน

การออกแบบระบบการคัดแยกมะนาวในงานวิจัยนี้สรุปวิธีการดำเนินงานได้ ดังนี้

3.1 กำหนดขั้นตอนการออกแบบระบบคัดแยกมะนาว



ภาพที่ 3: ขั้นตอนการออกแบบระบบคัดแยกมะนาว

กำหนดขั้นตอนการออกแบบระบบคัดแยกมะนาวตามภาพที่ 3 คือ เตรียมข้อมูล เลือกตัวแบบ Convolutional Neural Network ที่เหมาะสม หาตัวแบบคณิตศาสตร์โดยใช้ Order Logistic Regression ทำการประมวลผลและนำผลลัพธ์ที่ได้มาสร้างระบบคัดแยกมะนาว

3.2 แบ่งชั้นคุณภาพมะนาว

งานวิจัยนี้ศึกษาการคัดแยกมะนาวตามมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560 โดยเพิ่มการคัดแยกมะนาวสุกและมะนาวเน่า ดังนั้น จึงแบ่งชั้นคุณภาพมะนาวเป็น 4 ระดับคือ ชั้นพิเศษ ชั้นหนึ่งและชั้นสอง มะนาวสุก และมะนาวเน่าดังภาพที่ 4



ภาพที่ 4: ตัวอย่างคุณภาพของมะนาวในแต่ละกลุ่ม

3.3 จัดกลุ่มมะนาว และถ่ายภาพ

คัดเลือกมะนาวที่ใช้เป็นตัวอย่าง และจัดเป็น 4 กลุ่มตามชั้นคุณภาพแล้วนำไปถ่ายภาพ เพื่อนำไปใช้ในการวิจัย

3.4 เลือกตัวแบบ Convolutional Neural Network

นำตัวแบบ Convolutional Neural Network ที่นิยมใช้ในงานวิจัยที่เกี่ยวข้อง [14] คือ Own Models VGG Alexnet Googlenet Resnet และ LeNet มาใช้เป็นเกณฑ์ในการเลือกตัวแบบ และนำตัวแบบที่เลือกนี้มาทดลองฝึกสอนเครื่องจักรให้เรียนรู้เพื่อเลือกตัวแบบที่มีความแม่นยำสูงสุด 4 อันดับแรกมาใช้ในงานวิจัย ซึ่งตัวแบบดังกล่าว คือ VGG19 Alexnet Googlenet และ Resnet50

3.5 แบ่งภาพมะนาวที่ใช้ในการวิจัย

ภาพที่นำมาใช้มีขนาด 500x500x3 พิกเซล เท่ากัน และมีพื้นหลังสีขาวเหมือนกัน ทำการแบ่งภาพมะนาว ในแต่ละกลุ่มเป็น 2 ชุด ภาพชุดที่ 1 เป็นภาพที่ใช้ฝึกสอนตัวแบบที่ใช้ในการวิจัย ซึ่งมีสัดส่วนร้อยละ 80 ของภาพมะนาวทั้งหมด และภาพชุดที่ 2 เป็นภาพที่ใช้ทดสอบระบบคัดแยกมะนาวที่ออกแบบ ซึ่งมีสัดส่วนร้อยละ 20 ของภาพมะนาวทั้งหมด และเนื่องจากมะนาวชั้นพิเศษ ชั้นหนึ่ง และชั้นสองมีสีที่ใกล้เคียงกัน อีกทั้งกำหนดนิบนิบนิบมะนาวมีความหลากหลาย ดังนั้น จึงต้องใช้ภาพมะนาวจำนวนมากในการฝึกสอนเครื่องจักรให้เรียนรู้ ซึ่งจากการทดลอง พบว่า ต้องใช้ภาพมะนาวกลุ่มละ 1,600 ภาพ ในการฝึกสอนตัวแบบให้คัดแยกภาพมะนาวได้ถูกต้องด้วยความแม่นยำที่ร้อยละ 90 ด้วยเหตุนี้ จึงได้กำหนดจำนวนภาพมะนาวไว้กลุ่มละ 2,000 ภาพ โดยชุดที่ 1 ซึ่งใช้ในการฝึกสอนตัวแบบมีภาพจำนวน 1,600 ภาพ (ร้อยละ 80 ของภาพมะนาวทั้งหมดในกลุ่ม) และชุดที่ 2 ซึ่งใช้ทดสอบความสามารถของระบบคัดแยกมะนาว มีภาพจำนวน 400 ภาพ (ร้อยละ 20 ของภาพมะนาวทั้งหมดในกลุ่ม)

3.6 Training of Convolutional Neural Network

นำภาพมะนาวชุดที่ 1 ของทั้ง 4 กลุ่ม ตามชั้นคุณภาพ เข้าเรียนรู้กับตัวแบบ Convolutional Neural Network เพื่อเลือกตัวแบบที่มีความเหมาะสมมากที่สุดในการนำมาใช้คัดแยกมะนาว ทั้งนี้ ภาพที่ใช้ต้องมีความชัดเจนและถูกต้องตรงตามชั้นคุณภาพ เพราะหากภาพดังกล่าวไม่ชัดเจนหรือไม่ถูกต้องจะไม่สามารถแก้ไขข้อมูลในระบบได้ ซึ่งจะส่งผลให้เกิดการประมวลผลผิดพลาด และขาดความแม่นยำ

3.7 หาค่าคุณลักษณะของมะนาวที่ใช้ในการวิจัย

นำมะนาวจำนวน 300 ตัวอย่าง มาหาค่าคุณลักษณะ โดยทำการวัดค่าความสว่าง ค่าสีของผิว ค่าความหนาแน่นของผล รวมทั้งประเมินระดับรอยตำหนิของผิวตามเกณฑ์ที่กำหนดในมาตรฐานสินค้าเกษตร เลขที่ มกษ. 27-2560

3.8 Ordinal Logistic Regression Analysis

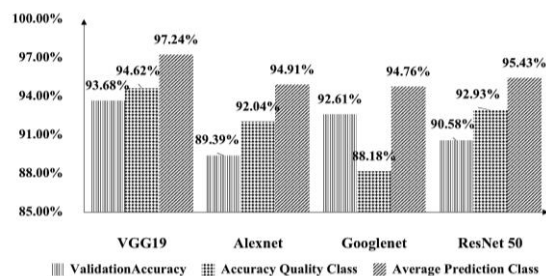
นำสมการ Ordinal Logistic Regression มาใช้หาตัวแบบคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระ คือคุณลักษณะต่าง ๆ ของผลมะนาวที่มีผลต่อตัวแปรตาม คือ ชั้นคุณภาพของมะนาว โดยใช้โปรแกรม Minitab และนำผลลัพธ์ที่ได้ทั้งหมดมาออกแบบระบบคัดแยกมะนาว

3.9 Testing of Convolutional Neural Network

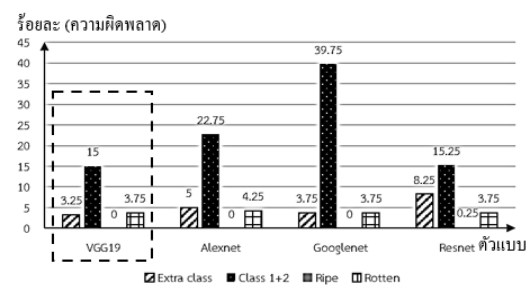
นำภาพมะนาวชุดที่ 2 ของทั้ง 4 กลุ่ม ตามชั้นคุณภาพ มาทดสอบความสามารถของระบบคัดแยกมะนาวที่ออกแบบ โดยใช้ตัวแบบที่เลือก

4. ผลการดำเนินงาน

ตัวแบบ Convolutional Neural Network ที่มีความเหมาะสมมากที่สุดที่จะใช้ในการคัดแยกมะนาว คือ VGG19 เนื่องจากมีความแม่นยำสูงที่สุดดังแสดงในภาพที่ 5 และมีความผิดพลาดในการคัดแยกต่ำกว่าตัวแบบอื่น ดังแสดงในภาพที่ 6



ภาพที่ 5: ผลการทดสอบ VGG19, Alexnet, Googlenet, ResNet50



ภาพที่ 6: ความผิดพลาดของ VGG19, Alexnet, Googlenet, ResNet50

ส่วนการหาตัวแบบคณิตศาสตร์ โดยใช้โปรแกรม Minitab ในการประมวลผล พบว่า ตัวแปรทั้งหมดมีค่า P-value < 0.05 แสดงว่า ตัวแปรที่มีนัยสำคัญทางสถิติ โดยมีค่า Somer's D [15] ค่า Goodman-k เท่ากับ 1 ค่า Kendall's Tau-a เท่ากับ 0.75 และผลลัพธ์ที่ได้มีค่าเป็นบวก แสดงว่า ตัวแปรที่วิเคราะห์มีค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความสัมพันธ์กันสูงและไปในทิศทางเดียวกัน จึงสามารถนำค่า coef (Coefficient) ที่ได้จากการประมวลผล มาสร้างตัวแบบคณิตศาสตร์ ในรูปสมการถดถอยลอจิสติกเชิงอันดับ 3 สมการ ดังนี้

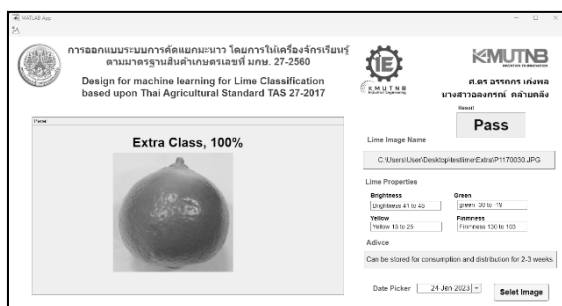
$$\text{Logit}[P(Y \leq 1)] = 180.9 - 1.5(x_1) + 3.5(x_2) - 2.5(x_3) + 128.4(x_4) \quad (3-1)$$

$$\text{Logit}[P(Y \leq 2)] = 375.8 - 1.5(x_1) + 3.5(x_2) - 2.5(x_3) + 128.4(x_4) \quad (3-2)$$

$$\text{Logit}[P(Y \leq 3)] = 407.9 - 1.5(x_1) + 3.5(x_2) - 2.5(x_3) + 128.4(x_4) \quad (3-3)$$

- โดย Y คือ ชั้นระดับชั้นคุณภาพของมะนาว
 x_1 คือ ความสว่างบนผิวมะนาว
 x_2 คือ ค่าสีเขียวบนผิวมะนาว
 x_3 คือ ค่าความหนาแน่น
 x_4 คือ ค่าตำหนิบนพื้นผิวในมะนาว 1 ลูก

จากนั้นจึงนำ VGG19 และสมการถดถอยดังกล่าวมา ออกแบบระบบคัดแยกมะนาว โดยแสดงผลการคัดแยก ได้แก่ ภาพและระดับชั้นคุณภาพของมะนาว ค่าคุณลักษณะของมะนาวในชั้นคุณภาพนั้น พร้อมกับคำแนะนำ เกี่ยวกับระยะเวลาในการเก็บรักษาและการบริโภคผ่านหน้าจอผู้ใช้งานแบบกราฟิก ตามภาพที่ 7 และเมื่อนำระบบคัดแยกมะนาวที่ออกแบบนี้มาทดสอบการใช้งานจริงโดยใช้ข้อมูลที่ระบบนี้ไม่เคยเรียนรู้มาก่อนพบว่า ระบบนี้สามารถคัดแยกมะนาว แทนการใช้ตาคน และสามารถแสดงผลการคัดแยกได้ถูกต้อง โดยมีค่าความแม่นยำ (Accuracy) สูงถึงร้อยละ 94.62



ภาพที่ 7: ตัวอย่างโปรแกรมที่ในงานวิจัย

5. สรุป

การพัฒนาเทคโนโลยีที่ใช้ในการคัดแยกมะนาว โดยการออกแบบระบบคัดแยกมะนาวเพื่อนำมาใช้แทนตาคนในงานวิจัยนี้ นอกจากจะทำให้การคัดแยกมะนาวมีความแม่นยำและเที่ยงตรงเป็นมาตรฐานเดียวกัน แม้ว่าผู้คัดแยกจะมีส่วนบุคคลเดียวกันแล้ว ยังทำให้ผู้ที่ใช้ระบบนี้ได้รับทราบข้อมูลที่จำเป็น ได้แก่ คุณภาพของมะนาว ระยะเวลาในการเก็บรักษา และคำแนะนำในการบริโภค ทั้งนี้ สามารถนำผลการวิจัยนี้ไปประยุกต์ใช้ในการประเมินคุณภาพของผลไม้ หรือสินค้าทางการเกษตรชนิดอื่น เพื่อให้ผู้ผลิตหรือผู้บริโภคสามารถคัดแยกสินค้าที่มีคุณภาพดีตรงตามมาตรฐานตามที่ต้องการได้สะดวกและรวดเร็วยิ่งขึ้น อย่างไรก็ตาม ระบบคัดแยกมะนาวที่ออกแบบในงานวิจัยนี้ยังเป็นเพียงเครื่องมือและอุปกรณ์ที่สร้างขึ้นโดยยังไม่เป็นระบบอัตโนมัติ ทั้งยังมีการประมวลผลภาพในสองมิติ จึงควรมีการพัฒนาให้เป็นระบบอัตโนมัติ และมีการประมวลผลภาพใน 3 มิติ หรือในรูปแบบอื่นเพื่อให้ระบบนี้สามารถคัดแยกมะนาวได้ดียิ่งขึ้น

เอกสารอ้างอิง

- [1] นันทพร ลิขิตไพฑูย์ศิลป์, *การจัดการความรู้ทางบัญชีเพื่อการพัฒนากษตรกรผู้ผลิตมะนาว ในอำเภอบ้านแพ้ว จังหวัดสมุทรสาคร*, มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์, หน้า 27, 2558.
- [2] พัชรารัตน์ ลีเมธีกุล (บ.ก.), *สถิติการเกษตรของประเทศไทย พ.ศ. 2564*, กรุงเทพมหานคร: สำนักงานเศรษฐกิจการเกษตร, เลขที่ 401, หน้า 100, 2564.
- [3] ทวีป หลวงแก้ว, และคณะ, *การเปรียบเทียบสายต้นมะนาวพันธุ์แป้นทะวาย, ศูนย์วิจัยและพัฒนาการเกษตรพิจิตร กรมวิชาการเกษตร*, หน้า 2, 2562.
- [4] มาตรฐานสินค้าเกษตร มท. 27-2560 มะนาว, (28 พฤศจิกายน 2560), ราชกิจจานุเบกษา, เล่ม 134 ตอนพิเศษ 293 ง.
- [5] T. M. Mitchell, *Machine Learning*, New York, 1997.
- [6] ทรงกรด พิมพิศาล, และณัฐวิ ศรีวิบูลย์, "การประมวลผลภาพสำหรับการจำแนกรูปภาพพื้นดัสโดยใช้การเรียนรู้เชิงลึก" *Journal of Information Science and Technology*, ฉบับที่ 10, เล่มที่ 2, กรกฎาคม-ธันวาคม 2563, หน้า 19-25.
- [7] นพรุจ พัฒนสาร, และ ณัฐวิ ศรีวิบูลย์. (2563). *การประมวลผลภาพสำหรับการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึก*. มหาวิทยาลัยกาฬสินธุ์.

- [8] พิสุทธิวัชร กัลป์ชะกันท์, และคณะ. (2564). *การตรวจสอบรอยร้าวบนพื้นผิวคอนกรีตด้วยการเรียนรู้เชิงลึกด้วยสถาปัตยกรรม VGG16*. [วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต]. มหาวิทยาลัยธรรมศาสตร์.
- [9] เกரியง กิจบำรุงรัตน์, “ตัวแบบการถดถอยลอจิสติกเชิงอันดับสำหรับคณะวิชาและการวิเคราะห์เกรดนักศึกษา” *Veridian E-Journal, Science and Technology Silpakorn University*, ปีที่ 4, ฉบับที่ 2, เดือนมีนาคม-เมษายน 2560, หน้า 19-35.
- [10] S. Iqbal, et al., “Classification of Selected Citrus Fruits Based on Color Using Machine Vision System,” *International Journal of Food Properties*, Vol.19, pp. 272-288, 2016.
- [11] G. Kumar T K, “ Logistic Regression to classify Fruits Image Dataset. Indian Institute of Information Technology”, Design & Manufacturing, India Institute of Information Technology, 2019.
- [12] L. C. Ngugi, et al., “Recent Advances in Image Processing Techniques for Automated Leaf Pest and Disease Recognition” *Information Processing in Agriculture*, Vol. 8, pp. 27-51, 2020.
- [13] J. Naranjo-Torres, et al., “A Review of Convolutional Neural Network Applied to Fruit Image Processing,” *Applied Sciences*, Vol. 10, Number 3443, pp. 1-31, 2020.
- [14] A. Göktaş and O. İşçi,, “A Comparison of the Most Commonly Used Measures of Association for Doubly Ordered Square Contingency Tables via Simulation” *Metodološki zvezki*, Vol.8, No.1, pp. 17-37, 2011.

โมเดลทำนายผลสัมฤทธิ์ทางการศึกษาในระดับปริญญาโท

Predicting the Master's Educational Achievement using Different Models

อันฮวา นิลรัตน์ศิริกุล (Unhawa Ninrutsirikun)¹ อภรณ์ เชี่ยวชาญเกษตร (Aphorn Chiawchankaset)²
 สนิท ศิริสวัสดิวัฒนา (Sanit Sirisawatvatana)³ และวิฑิตา จงสุขชัยสิทธิ์ (Vithida Chongsuphajaisiddhi)⁴
 คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
¹unhawa@sit.kmutt.ac.th, ²aphorn@sit.kmutt.ac.th, ³sanit@sit.kmutt.ac.th, ⁴vithida@sit.kmutt.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างโมเดลในการทำนายผลสัมฤทธิ์ทางการศึกษาและเปรียบเทียบประสิทธิภาพของโมเดล โดยใช้ข้อมูลของนักศึกษาในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ จำนวน 1,468 ตัวอย่าง ผู้วิจัยแบ่งการศึกษออกเป็น 2 ส่วน ได้แก่ 1) ศึกษาประสิทธิภาพของปัจจัยที่ใช้เป็นตัวทำนาย และ 2) เปรียบเทียบประสิทธิภาพของอัลกอริทึมที่ใช้ในการทำนายผ่านโมเดลทำนายผลที่ไม่มีการปรับค่าไฮเปอร์พารามิเตอร์ ได้แก่ Naïve Bayes (NB) Multilayer Perceptron (MLP) และ Support Vector Machine (SVM) และโมเดลทำนายผลที่ได้รับการปรับแต่งค่าไฮเปอร์พารามิเตอร์แบบที่ดีที่สุด ผลการศึกษาแสดงให้เห็นว่าโมเดลที่ใช้ปัจจัยมาตรฐานร่วมกับคะแนนเฉลี่ยสะสมของวิชาบังคับและแผนการเรียนที่นักศึกษาเป็นผู้เลือกมีประสิทธิภาพในการทำนายผลสูงกว่าโมเดลที่เพียงใช้ปัจจัยมาตรฐานเพียงอย่างเดียว อัลกอริทึมแบบไม่ปรับแต่งค่าไฮเปอร์พารามิเตอร์มีประสิทธิภาพในการทำนายไม่แตกต่างจากอัลกอริทึมที่ได้รับการปรับแต่งค่าไฮเปอร์พารามิเตอร์ Naïve Bayes โมเดลที่ใช้ปัจจัยมาตรฐานเพียงอย่างเดียว ทำนายผลได้ถูกต้องที่ 64.03 % เมื่อเพิ่มตัวทำนายเป็นคะแนนเฉลี่ยสะสมของวิชาบังคับและแผนการเรียน โมเดลดังกล่าวสามารถทำนายได้ถูกต้องที่ 82.42 % ผลการศึกษาแสดงให้เห็นว่าคะแนนเฉลี่ยสะสมของวิชาบังคับและแผนการเรียนเป็นปัจจัยที่ส่งผลในเชิงบวกต่อประสิทธิภาพในการทำนายของโมเดล

คำสำคัญ: คะแนนเฉลี่ยสะสมของวิชาบังคับ บัณฑิตศึกษา แผนการเรียน โมเดลทำนายผลสัมฤทธิ์ทางการศึกษา

Abstract

The purposes of this research were to create a model to predict educational achievement and to compare the model's efficiency by using the information of 1,468 students in the information technology graduate programs. The researcher divided the study into 2 parts: 1) to study the efficiency of the factors used as predictors and 2) to compare the algorithm performance used in prediction through unadjusted hyperparameter predictive models: Naïve Bayes (NB), Multilayer Perceptron (MLP), and Support Vector Machine (SVM) and a hyperparameter-optimized model. The results showed that the model in which the standard factor was combined with the grade point average (GPA) of the required subjects and the study plan chosen by the student had a higher predictive efficiency than a model that used standard factors alone. The performance of the algorithm without hyperparameters tuning compared to a hyperparameter-optimized model was not different. Naïve Bayes algorithm predicted by standard factors alone was able to predict the result correctly at 64.03 % whereas the model predicted by standard factors together with the cumulative GPA of compulsory subjects and the study plans was able to predict the result correctly at 82.42 %. The results showed that the GPA of required subjects and the study plans were factors that influenced the effectiveness of the predictive model.

Keywords: GPA of required subject, Graduate Studies, Model to Predict Educational Achievement, Study Plans.

1. บทนำ

การศึกษาต่อในระดับปริญญาโท นอกจากจะเป็นการเพิ่มพูนความรู้และศักยภาพให้กับผู้เรียนซึ่งนำไปสู่ความก้าวหน้า และมั่นคงในอาชีพแล้ว [1] ยังเป็นส่วนสำคัญในการพัฒนากำลังคนเชิงคุณภาพ เพิ่มขีดความสามารถของประเทศ ตามแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติฉบับที่ 13 (พ.ศ. 2566 – 2570) [2] ปัจจุบันภาครัฐได้ส่งเสริมให้เกิดสังคมที่พร้อมต่อการเรียนรู้ตลอดชีวิต ส่งเสริมให้สถาบันอุดมศึกษาจัดทำแผนการเรียนที่มีความยืดหยุ่น หลากหลายและเปิดทางเลือกในการเรียนรู้ให้กับผู้เรียนในทุกระดับ สถาบันอุดมศึกษาจึงเพิ่มทางเลือกให้กับผู้เรียนโดยเฉพาะอย่างยิ่งการเรียนต่อในระดับบัณฑิตศึกษา ที่เปิดโอกาสให้บุคคลทั่วไปสามารถเข้ารับการศึกษาระบบรายวิชา กลุ่มวิชา หรือในรูปแบบของหลักสูตรฝึกอบรม [3] โดยผู้เรียนสามารถเก็บสะสมผลการประเมินในรูปแบบของธนาคารสะสมหน่วยกิต (Credit Bank) ได้ยาวนานถึง 5 ปี ซึ่งในช่วงระยะเวลา 5 ปีนี้ผู้เรียนสามารถโอนผลการเรียนที่ได้จากการสะสมหน่วยกิต เข้าสู่การศึกษาในระบบของมหาวิทยาลัยเพื่อขอรับปริญญาบัตรได้ การเรียนในรูปแบบสะสมหน่วยกิตจึงช่วยให้ผู้เรียนได้ทดลองเรียนในวิชาที่สนใจ ได้ค้นหาคำตอบของตนเองรวมถึงวางแผนการศึกษาและตัดสินใจได้ว่าผู้เรียนต้องการเพียงประกาศนียบัตร หรือต้องการโอนหน่วยกิตเพื่อเข้าศึกษาและรับปริญญา

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี เป็นหลักสูตรที่เปิดกว้างให้ผู้ที่มีสนใจเข้าศึกษาต่อที่สำเร็จการการศึกษาในระดับปริญญาตรีทุกสาขาวิชา สามารถสมัครเข้าศึกษาต่อได้ อย่างไรก็ตามนักศึกษาแรกเข้าที่ไม่คุ้นเคยหรือไม่เข้าใจในเทคโนโลยีสารสนเทศอย่างเพียงพอ อาจจะต้องใช้เวลาในการศึกษาและทำความเข้าใจกับศาสตร์นี้มากขึ้น หลักสูตร ฯ จึงมีกลยุทธ์ในการแก้ไขปัญหาโดย การจัดให้มีการสอนความรู้เบื้องต้นด้านเทคโนโลยีสารสนเทศก่อนเริ่มการศึกษาในหลักสูตร (Pre course) และจัดให้มีการสอนเสริมในบางรายวิชา

ซึ่งแสดงให้เห็นว่าหลักสูตร ฯ ให้ความสำคัญกับผลสัมฤทธิ์ทางการศึกษาของผู้เรียนเป็นอย่างมาก นอกจากนี้การเปิดให้ผู้บุคคลทั่วไปที่ต้องการเพิ่มทักษะทางด้านทางด้านเทคโนโลยีสารสนเทศ สามารถเข้าศึกษาในรูปแบบสะสมหน่วยกิต [4] ยังช่วยให้ผู้เรียนได้ทดลองเรียนก่อนแล้วค่อยตัดสินใจที่จะวางแผนการเรียนต่อได้อย่างเหมาะสม

การศึกษาต่อในระดับปริญญาโท มีเพียงวิชาบังคับ (Core Courses) ที่ถือว่าเป็นวิชาพื้นฐานที่ผู้เรียนทุกคนต้องเรียน นอกจากนั้นผู้เรียนสามารถตัดสินใจวางแผนการเรียนได้ด้วยตนเอง ซึ่งส่วนที่สำคัญคือการเลือกแผนการเรียน ซึ่งหลักสูตร ฯ มีแผนการเรียนอยู่ 2 รูปแบบ ได้แก่ 1) แผนการเรียนแบบ แผน ก ผู้เรียนต้องทำวิทยานิพนธ์ (Thesis) จำนวน 12 หน่วยกิต นำเสนอผลงานวิจัยในรูปแบบการประชุมวิชาการ (Conference) หรือตีพิมพ์ในวารสารวิจัย (Journal) 2) แผนการเรียนแบบ แผน ข ผู้เรียนสามารถเลือกทำการศึกษาอิสระ หรือโครงการเฉพาะเรื่อง จำนวน 6 หน่วยกิต ร่วมกับการสอบประมวลความรู้ (Comprehensive Examination) สำหรับอีก 6 หน่วยกิตที่เหลือ ผู้เรียนสามารถเรียนวิชาในหมวดวิชาเลือก ดังนั้นแผนการเรียนรูปแบบ แผน ข จึงเหมาะกับผู้ที่ต้องการเรียนวิชาเลือก เพื่อให้เข้าใจในเทคโนโลยีสารสนเทศเพิ่มมากขึ้น

เพื่อช่วยให้ผู้เรียนในรูปแบบสะสมหน่วยกิต และนักศึกษาที่ผ่านการเรียนในภาคการศึกษาแรก (มีผลการเรียนวิชาบังคับ) สามารถประเมินผลสัมฤทธิ์ทางการเรียนของตนเองได้ งานวิจัยนี้จึงนำเสนอโมเดลทำนายผลสัมฤทธิ์ทางการศึกษาในระดับปริญญาโท ที่ใช้ปัจจัยมาตรฐานร่วมกับคะแนนเฉลี่ยสะสมของวิชาบังคับ และแผนการเรียน เป็นข้อมูลอินพุต เพื่อทำนายว่าผู้เรียนจะสามารถสำเร็จการศึกษาในหลักสูตรนี้ได้หรือไม่

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ปัจจัยที่เกี่ยวข้องกับผลสัมฤทธิ์ทางการศึกษา

งานวิจัยที่ศึกษาเกี่ยวกับปัจจัยหรือตัวแปรที่ส่งผลต่อผลสัมฤทธิ์ทางการศึกษานั้น มีอยู่เป็นจำนวนมาก

ไม่ว่าจะเป็นการศึกษาเกี่ยวกับปัจจัยทางด้านครอบครัว สถานศึกษา ภูมิหลังของผู้เรียน รวมถึงแรงจูงใจในการ บรรลุผลสำเร็จ และความขยันหมั่นเพียร [5], [6] สำหรับบางปัจจัยต้องใช้ระยะเวลาในการเก็บข้อมูลเป็น เวลานาน และ บางปัจจัยการเก็บข้อมูลก็ทำได้ยาก [7] ผู้วิจัยจึงสนใจเฉพาะปัจจัยที่มีความชัดเจน สามารถเก็บ ข้อมูลได้ และนำไปใช้เป็นข้อมูลอินพุตให้กับ โมเดล ทำนายผล ได้ อันได้แก่ ภูมิหลังของผู้เรียน แผนการเรียน ในระดับปริญญาโท และคะแนนเฉลี่ยของวิชาบังคับ ในระดับปริญญาโท

ข้อมูลภูมิหลังของผู้เรียนสามารถแบ่งออกเป็นข้อมูล ส่วนตัวของผู้เรียน (Profile) และ ภูมิหลังทางการศึกษา (Academic Background) งานวิจัยที่ใช้เครื่องมือทางสถิติ และใช้กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาโท เปิดเผย ให้เห็นว่า ผู้เรียนที่มีอายุน้อยมีโอกาสที่จะสำเร็จการศึกษา สูงกว่าผู้เรียนที่มีอายุมาก [8], [9] เพศหญิงมีอัตราการ สำเร็จตามเกณฑ์สูงกว่าเพศชาย [10] งานวิจัยก่อนหน้านี้ที่ ใช้ข้อมูลของนักศึกษาจากแหล่งข้อมูลเดียวกัน [7], [11] ใช้ข้อมูลสาขาวิชาที่เรียนและสถาบันที่สำเร็จการศึกษาใน ระดับปริญญาตรี เพื่อบอกระดับความคุ้นเคย หรือ ความเข้าใจในเทคโนโลยีสารสนเทศ คะแนนเฉลี่ยสะสม ในระดับปริญญาตรีถูกใช้เพื่อระบุภูมิหลังทางการศึกษา ซึ่งพบว่าคะแนนเฉลี่ยสะสมมีความสัมพันธ์กับการสำเร็จ การศึกษา [6], [8], [12] ในส่วนของสถานภาพการทำงาน มีรายงานว่านักศึกษาที่ทำงานอยู่มีโอกาสสำเร็จการศึกษา สูงกว่านักศึกษาที่ไม่ได้ทำงานและสูงกว่านักศึกษา ที่ทำงานส่วนตัว [10] ในงานวิจัยที่ใช้ข้อมูลตัวอย่างจาก นักศึกษาในระดับปริญญาตรี แสดงให้เห็นว่าแผนการเรียน ที่นักศึกษาเลือก มีความสัมพันธ์กับผลสัมฤทธิ์ในการเรียน ระดับรายวิชา [12]

2.2 โมเดลทำนายผล

งานวิจัยในปัจจุบันได้นำเอาเทคนิคทางด้านแมชชีน เลิร์นนิง (Machine Learning) เช่น การใช้สมการถดถอย (Regression) มาใช้ในการทำนายผลการเรียน [13] การนำเทคนิคการจำแนกข้อมูล (Classification) มาใช้ในการทำนายผลการเรียนในรูปแบบของเกรด [12] หรือ

ผลการเรียน ในรูปแบบของการ ผ่าน/ตก [7], [11] รวมถึงทำนายระดับความสามารถของผู้เรียน [14] สำหรับ เทคนิคที่ได้รับความนิยมในการจำแนกข้อมูลได้แก่ การ จำแนกแบบเบย์ (Bayesian Classification) โครงข่าย ประสาทเทียม (Artificial Neural Network :ANN) และ Support Vector Machine (SVM) [14]

Multilayer Perceptron (MLP) เป็นตัวแทนของระบบ โครงข่ายประสาทเทียมที่ได้รับความนิยมในการนำมาใช้ ทำนายผลการเรียน ทั้งในรูปแบบของการทำนายเกรด [12] และทำนายระดับความสามารถของผู้เรียน [14] Naive Bayes (NB) เป็นตัวแทนของเทคนิคการจำแนกแบบ เบย์ อัลกอริทึมนี้ใช้วิธีการทางสถิติ ที่มีพื้นฐานมาจาก ทฤษฎีพื้นฐานของความน่าจะเป็น ซึ่งให้ผลการทำนายใน ระดับที่ใกล้เคียงกับ MLP [13] สำหรับเทคนิค SVM เป็นเทคนิคที่ใช้ Hyper plan ในการสร้างขอบเขต ในการแบ่งข้อมูล (หรือ การจัดกลุ่ม) ซึ่งโครงสร้างแบบ SVM เป็นโครงสร้างที่ให้ผลการทำนายในระดับที่ดีมาก [14]

ปัจจุบัน แมชชีนเลิร์นนิง มีอัลกอริทึมอยู่เป็นจำนวนมากและอัลกอริทึมเหล่านั้นมีค่าไฮเปอร์พารามิเตอร์ (Hyperparameters) เป็นของตัวเอง การตั้งค่าไฮเปอร์ พารามิเตอร์ที่เหมาะสมจะช่วยเพิ่มประสิทธิภาพของ โมเดลให้มีประสิทธิภาพที่สูงมากกว่า การใช้ค่าปกติ (Default Setting) [15] Auto-WEKA เป็นเครื่องมือที่ช่วย ในการค้นหาโมเดลที่ประสิทธิภาพดีที่สุด โดยการค้นหา แบบเต็มรูปแบบ จากการเลือกอัลกอริทึมที่เครื่องมือรู้จัก มาทีละอัลกอริทึมแล้วค้นหาค่าไฮเปอร์พารามิเตอร์ที่ดี ที่สุดให้กับอัลกอริทึมนั้น ในขั้นตอนสุดท้าย Auto-WEKA จะแนะนำ อัลกอริทึมที่ประสิทธิภาพสูงสุด เครื่องมือนี้จึง เหมาะกับผู้ที่ไม่มีความชำนาญ หรือ ผู้ที่ต้องการค้นหา โมเดลที่ดีที่สุด [15], [16]

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้าง โมเดลทำนาย ผลสัมฤทธิ์ทางการศึกษา โดยใช้ข้อมูลตัวอย่างของ นักศึกษาในระดับบัณฑิตศึกษา ใช้เทคนิคแมชชีนเลิร์นนิง ในการสร้าง โมเดลทำนายผล รายละเอียดของข้อมูล

ตัวอย่างที่นำมาใช้และการออกแบบ โมเดลสำหรับที่ใช้ในการทำนายผลมีดังนี้

3.1 ชุดข้อมูลที่ใช้ในงานวิจัย

จากการทบทวนวรรณกรรม พบตัวอย่างงานวิจัยที่ใช้ข้อมูลของนักศึกษาในหลักสูตรวิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี [7, 11] ซึ่งเป็นข้อมูลของนักศึกษาในสาขาเดียวกัน งานวิจัยดังกล่าวเก็บข้อมูลของผู้เข้าศึกษาระหว่างภาคการศึกษาที่ 2/2539 จนถึง 1/2542 ซึ่งมีข้อมูลตัวอย่างเพียง 276 ตัวอย่าง อย่างไรก็ตามงานวิจัยดังกล่าวถือเป็นจุดตั้งต้นที่แสดงให้เห็นว่า เพศ อายุ สถานภาพสมรส ระดับการศึกษาที่จบ สาขาวิชาที่จบ สถาบันการศึกษาที่จบ เกรดเฉลี่ยที่จบ และสถานะภาพการทำงาน เป็นปัจจัยที่มีความสัมพันธ์กับผลสัมฤทธิ์ทางการศึกษา ของผู้เรียนในหลักสูตรนี้ แต่อย่างไรก็ตาม ในปัจจุบันการเก็บข้อมูลส่วนบุคคลมักจะเก็บเฉพาะข้อมูลที่จำเป็นตามความสนใจของผู้ให้ข้อมูล งานวิจัยนี้จึงตัดข้อมูล (หรือ ปัจจัย) ที่ผู้ให้ข้อมูลส่วนใหญ่ไม่ระบุค่า ออกจากปัจจัยที่ใช้เป็นข้อมูล อินพุต และเพิ่มปัจจัยที่เป็นข้อมูลอินพุตใหม่ ได้แก่ ข้อมูลคะแนนเฉลี่ยสะสมของวิชาบังคับ (GPA. of the required subjects) และ แผนการเรียนที่นักศึกษาเลือก

ข้อมูลที่นำมาใช้ในงานวิจัยนี้เป็นข้อมูลของนักศึกษาในสาขาเดียวกันกับงานวิจัยตัวอย่างข้างต้น โดยข้อมูลอินพุตถูกแบ่งออกเป็น 2 กลุ่ม ได้แก่ ข้อมูลที่มักจะถูกใช้เป็นตัวทำนายพื้นฐาน [5-14] ซึ่งในงานวิจัยนี้จะเรียกปัจจัยดังกล่าวว่าปัจจัยมาตรฐาน ซึ่งประกอบด้วย เพศ อายุ สาขาวิชาที่จบ สถาบันการศึกษาที่จบ เกรดเฉลี่ยสะสมที่จบการศึกษา และสถานะทำงาน ข้อมูลอินพุตในกลุ่มที่ 2 เป็นปัจจัยที่งานวิจัยนี้ต้องการศึกษาประสิทธิภาพเมื่อใช้เป็นตัวทำนาย ซึ่งประกอบด้วย คะแนนเฉลี่ยสะสมของวิชาบังคับ และแผนการเรียนที่นักศึกษาเป็นผู้เลือก โดยข้อมูลอินพุตทั้งหมดของงานวิจัยนี้ แสดงไว้ในตารางที่ 1

งานวิจัยนี้เก็บข้อมูลตัวอย่างจากผู้เข้าศึกษาในหลักสูตร ฯ ระหว่างปีการศึกษาที่ 1/2555 จนถึง 2/2562 เมื่อนำข้อมูลทั้งหมดมาทำความสะอาดข้อมูล (Data Cleansing) แล้วคงเหลือข้อมูลตัวอย่างจำนวน 1,468 ชุด

ตารางที่ 1: อินพุตที่ใช้ในการศึกษานี้

แอททริบิวต์	ค่าของแอททริบิวต์
1. เพศ	1: ชาย 2: หญิง
2. อายุ (ปี)	1: 20-25 2: 26-30 3: 31-35 4: 36-40 5: มากกว่า 40
3. สาขาวิชาที่จบ	1: สาขาวิชาทางด้านคอมพิวเตอร์ 2: สาขาวิชาวิทยาศาสตร์และวิศวกรรมศาสตร์แต่ไม่รวมสาขาวิชาทางด้านคอมพิวเตอร์ 3: สาขาวิชาทางการแพทย์ 4: สาขาวิชาทางเศรษฐกิจและบริหารธุรกิจ 5: สาขาวิชาทางด้านศิลปะ สังคม กฎหมาย และครุศาสตร์
4. สถาบันการศึกษาที่จบ	1: สถาบันอุดมศึกษารัฐบาล 2: มหาวิทยาลัยเปิด 3: มหาวิทยาลัยราชภัฏหรือมหาวิทยาลัยราชชมงคล 4: สถาบันอุดมศึกษาเอกชน 5: อื่น ๆ (สถาบันการศึกษาในต่างประเทศหรือที่นอกเหนือจากข้างต้น)
5. GPA ที่จบ	1: 3.50-4.00 2: 3.00-3.49 3: 2.50-2.99 4: 2.00-2.49
6. การทำงาน	1: ทำ 2: ไม่ทำ
7. คะแนนเฉลี่ยสะสมของวิชาบังคับ	1: 3.50-4.00 2: 3.00-3.49 3: น้อยกว่า 3.00
8. แผนการเรียน	1: แผน ก. (ทำวิทยานิพนธ์) 2: แผน ข. (ทำโปรเจค หรือ สัมมนาเชิงปฏิบัติการ)

3.2 การออกแบบโมเดลทำนายผล

งานวิจัยนี้ต้องการตรวจสอบว่า คะแนนเฉลี่ยสะสมของวิชาบังคับและแผนการเรียนที่นักศึกษาเลือกนั้นเป็นปัจจัยที่มีความสัมพันธ์กับผลสัมฤทธิ์ทางการศึกษาของผู้เรียนหรือไม่ ผู้วิจัยจึงออกแบบการทดลองออกเป็น 2 กลุ่มตามข้อมูลอินพุต โดยการทดลองแรกเป็นการสร้างโมเดลทำนายผลโดยใช้ปัจจัยมาตรฐาน (แอททริบิวต์ในลำดับที่ 1-6) เพียงอย่างเดียว และการทดลองที่ 2 จะใช้ข้อมูลอินพุตทั้งหมดในการสร้างโมเดลทำนายผล ทั้งสองการทดลองเลือกใช้ อัลกอริทึมมาตรฐานของโปรแกรม WEKA [14], [15]

ได้แก่ NB Multilayer MLP และ SVM โดยโมเดลเหล่านี้ จะไม่ถูกปรับแต่งค่าไฮเปอร์พารามิเตอร์ ในการทดลองยัง ใช้ Auto-WEKA [15, 16] เพื่อค้นหาโมเดลที่ดีที่สุด โดย Auto-WEKA จะประเมินและเลือกอัลกอริทึม ที่เหมาะสมและทำการปรับแต่งค่าไฮเปอร์พารามิเตอร์ เพื่อให้ได้โมเดลที่มีค่าความถูกต้อง (Accuracy) สูงที่สุด การออกแบบการทดลองนี้ ช่วยให้สามารถเปรียบเทียบผล การทำนาย ระหว่าง โมเดลที่ใช้ปัจจัยจากการศึกษาเดิม กับ โมเดลที่เพิ่ม คະแนนเฉลี่ยสะสมของวิชาบังคับและ แผนการเรียน ซึ่งเป็นปัจจัยใหม่ และยังสามารถ เปรียบเทียบผลการทำนายระหว่าง โมเดลมาตรฐานกับ โมเดลที่ได้รับการปรับแต่งค่าไฮเปอร์พารามิเตอร์

4. ผลการจำแนกข้อมูล

ผลการจำแนกข้อมูลผลสัมฤทธิ์ทางการศึกษาของชุด ข้อมูลตัวอย่างนี้แสดงไว้ในตารางที่ 2 และตารางที่ 3 ประสิทธิภาพของการจำแนกข้อมูลจากทั้ง 2 ตารางแสดง ให้เห็นว่า อัลกอริทึมมาตรฐานที่ไม่ถูกปรับแต่งค่าไฮเปอร์ พารามิเตอร์จะมีประสิทธิภาพใกล้เคียงกับ โมเดลที่ดีที่สุด ที่ถูกปรับแต่งค่าไฮเปอร์พารามิเตอร์

สำหรับการทำนายผลด้วยปัจจัยมาตรฐาน โมเดล ที่สร้างด้วย AttributeSelectedClassifier ให้ประสิทธิภาพ ในการทำนายผลดีที่สุด โดยสามารถจำแนกข้อมูล ได้ถูกต้องที่ 67.37% และเมื่อเพิ่มจำนวนอินพุต โดยใช้แอ ททริบิวท์ทั้งหมด โมเดลที่สร้างด้วย Naïve Bayes ให้ประสิทธิภาพในการทำนายผลดีที่สุดที่ 82.49% นั้นแสดงให้เห็นว่าการเพิ่มจำนวนอินพุตด้วยคะแนนเฉลี่ย สะสมของวิชาบังคับและ แผนการเรียนช่วยเพิ่ม ประสิทธิภาพของการจำแนกข้อมูลได้เป็นอย่างมาก ตัวอย่างเช่น เปรียบเทียบประสิทธิภาพของ โมเดลที่สร้าง ด้วยอัลกอริทึม NB ที่ไม่มีการปรับแต่งค่าไฮเปอร์ พารามิเตอร์ ด้วยรูปแบบการทดสอบแบบ 10-fold Cross Validation หากใช้ข้อมูลอินพุตเพียงปัจจัยมาตรฐาน โครงสร้างนี้จะให้ผลการจำแนกข้อมูลได้ถูกต้องที่ 64.03% และเมื่อเพิ่มจำนวนอินพุต (คะแนนเฉลี่ยสะสมของวิชา บังคับและแผนการเรียน) โมเดลสามารถจำแนกข้อมูลได้

ถูกต้อง 82.43% หรือจำแนกข้อมูลได้ถูกต้องเพิ่มมากขึ้น 18.39%

ตารางที่ 2: ประสิทธิภาพของการจำแนกข้อมูลโดยใช้ปัจจัยมาตรฐาน

	Algorithm	Class Graduate	F-measure	Accuracy
10-Fold Cross Validation	Naïve Bayes	Yes	0.712	64.03
		No	0.522	
	Multilayer Perceptron	Yes	0.712	63.48
		No	0.503	
	Support Vector Machine	Yes	0.733	64.03
		No	0.448	
LMT (Hyperparameters tuning)	Yes	0.739	65.94	
	No	0.511		
5-Fold Cross Validation	Naïve Bayes	Yes	0.712	64.10
		No	0.523	
	Multilayer Perceptron	Yes	0.721	63.96
		No	0.491	
	Support Vector Machine	Yes	0.731	63.49
		No	0.431	
Multilayer Perceptron (Hyperparameters tuning)	Yes	0.767	66.75	
	No	0.419		
Train/Test Split (70/30)	Naïve Bayes	Yes	0.743	62.05
		No	0.277	
	Multilayer Perceptron	Yes	0.697	60.45
		No	0.431	
	Support Vector Machine	Yes	0.736	63.41
		No	0.401	
AttributeSelected Classifier (Hyperparameters tuning)	Yes	0.763	67.37	
	No	0.478		

5. สรุป

ปัจจุบันรูปแบบการศึกษาต่อในระดับบัณฑิตศึกษาได้ เปิดกว้างมากขึ้น โดยผู้เรียนสามารถเข้าศึกษาต่อด้วยรูปแบบเดิม หรือ รูปแบบใหม่ที่เป็นแบบสะสมหน่วยกิต ภายใต้อัตลักษณ์และวิธีการที่สถานศึกษากำหนด แม้ว่าการเรียนในรูปแบบสะสมหน่วยกิต มีวัตถุประสงค์ เพื่อให้ผู้เรียนได้พัฒนาศักยภาพของตนเอง หากผู้เรียน

มีความพร้อมหรือคาดว่าจะสามารถเรียนต่อได้จนจบหลักสูตร ผู้เรียนอาจจะตัดสินใจโอนผลการเรียนเข้าสู่วการศึกษาในระดับบัณฑิตศึกษาได้

งานวิจัยนี้ได้นำเสนอโมเดลที่ใช้ในการทำนายผลสัมฤทธิ์ทางการศึกษาของนักศึกษาในระดับบัณฑิตศึกษา โดยใช้ข้อมูลของนักศึกษาในหลักสูตรวิทยาศาสตร์มหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี โดยขยายขอบเขตการศึกษาจากงานวิจัยเดิม [7], [11] ที่ใช้ข้อมูลจากกลุ่มตัวอย่างที่คล้ายกัน มาเป็นการสร้างโมเดลทำนายผลสัมฤทธิ์ทางการศึกษาของผู้เรียนว่าจะสามารถสำเร็จการศึกษาหรือไม่ โมเดลในการทำนายผลแบ่งออกเป็น 2 กลุ่มคือโมเดลที่ใช้ปัจจัยมาตรฐาน และโมเดลที่เพิ่มคะแนนเฉลี่ยสะสมของวิชาบังคับในหลักสูตร และแผนการเรียนที่นักศึกษาเลือก เมื่อเปรียบเทียบประสิทธิภาพการทำนายผลพบว่าโมเดลที่เพิ่มปัจจัยในการทำนาย ให้ผลการทำนายที่สูงกว่าโมเดลพื้นฐาน (ที่ 18.39%) และให้ผลการทำนายในระดับที่น่าพึงพอใจ (ความถูกต้องที่ 82.43%)

งานวิจัยนี้สามารถนำไปพัฒนาต่อยอดได้ในหลายบริบท ตัวอย่างเช่น ทดลองสร้างโมเดลโดยใช้ต้นไม้ตัดสินใจ (Decision Tree) ซึ่งจะช่วยให้เราสามารถอธิบายความสัมพันธ์ระหว่างข้อมูลอินพุตกับผลสัมฤทธิ์ทางการเรียนได้ การประยุกต์ใช้แนวทางของการทำวิจัยนี้ในการสร้างโมเดลทำนายผลสัมฤทธิ์ทางการศึกษากับข้อมูลของนักศึกษาในหลักสูตรหรือสาขาวิชาอื่น

ตารางที่ 3: ประสิทธิภาพของการจำแนกข้อมูลโดยใช้ปัจจัยทั้งหมด

	Algorithm	Class Graduate	F-measure	Accuracy
10-Fold Cross Validation	Naïve Bayes	Yes	0.861	82.42
		No	0.760	
	Multilayer Perceptron	Yes	0.853	81.40
		No	0.746	
	Support Vector Machine	Yes	0.861	81.74
		No	0.736	
	Naïve Bayes (Hyperparameters tuning)	Yes	0.866	82.49
		No	0.746	
5-Fold	Naïve Bayes	Yes	0.862	82.42
		No	0.758	

	Multilayer Perceptron	Yes	0.851	81.40	
		No	0.746		
	Support Vector Machine	Yes	0.858	81.47	
		No	0.732		
	Naïve Bayes (Hyperparameters tuning)	Yes	0.866	82.49	
		No	0.746		
Train/Test Split (70/30)	Naïve Bayes	Yes	0.840	80	
		No	0.733		
	Multilayer Perceptron	Yes	0.824	78.18	
		No	0.713		
	Support Vector Machine	Yes	0.836	78.63	
		No	0.693		
		Naïve Bayes (Hyperparameters tuning)	Yes	0.866	82.49
			No	0.746	

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ อาสาสมัครผู้ให้ข้อมูลในการทำวิจัย คณาจารย์และเจ้าหน้าที่ได้กรุณารวบรวมข้อมูล และให้คำแนะนำที่เป็นประโยชน์ในการวิจัยครั้งนี้ ทำให้งานวิจัยลุล่วงได้เป็นอย่างดี

เอกสารอ้างอิง

- [1] สำนักงานสภาพัฒนาการเศรษฐกิจและสังคมแห่งชาติ, “แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 13 (พ.ศ. 2566 – 2570)”, ราชกิจจานุเบกษา, 2565
- [2] กระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม, “ธนาคารหน่วยกิตแห่งชาติ (National Credit Bank)”, 1 กุมภาพันธ์, 2565
- [3] มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, “ระเบียบมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ว่าด้วยการศึกษาระดับบัณฑิตศึกษา พ.ศ. 2562”, 6 กุมภาพันธ์, 2562
- [4] คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, “สมัครเรียนแบบสะสมหน่วยกิต”, 28 กุมภาพันธ์, 2561
- [5] ศลิษา เปลี้นนิต, “แรงจูงใจและความคาดหวังของผู้เรียนสาขาวิศวกรรมโยธาเป็นปริญญาโทที่สอง”, *การประชุมวิศวกรรมโยธาแห่งชาติครั้งที่ 27*, เชียงราย, 24-26 สิงหาคม 2565
- [6] A. Neamsaard, et al., “A Study on the Reasons of Decision to Study in a Graduate Level: A case study of Faculty of Technology and Environment, Prince of Songkla University, Phuket Campus” *Journal of*

- Multidisciplinary Academic Research and Development (JMARD)*, vol. 4, no. 4, pp. 1-17, 2022
- [7] K. Lavangnananda and R. Jarassupawat, "Application of Genetic Algorithm for Inductive Learning (GAIL) in Classification", *Thai Journal of Development Administration, NIDA*, Vol. 41, No. 3, pp. 69-82, 2001
- [8] K. Saithanu and J. Mekpanyup. "Modeling to Predict Decision in Continuing Study for Master's Degree.", *Burapha Science Journal* Vol. 16, no. 1, pp. 3-11, 2011
- [9] ชัชชัย ศุภศิษย์, "ปัจจัยที่มีผลต่อผลสัมฤทธิ์ของการเรียนในระดับปริญญาโทของสถาบันบัณฑิตพัฒนบริหารศาสตร์ ปีการศึกษา 2554", พ.ศ. 2556
- [10] โกวิทย์ ประดิษฐ์ผล และ ปรีชา วิจิตรธรรมรส, "การศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามเกณฑ์การศึกษานักศึกษาระดับปริญญาโท ด้วยเทคนิคเหมืองข้อมูลกรณีศึกษา สถาบันบัณฑิตพัฒนบริหารศาสตร์", *วารสารนานาชาติ มหาวิทยาลัยขอนแก่น สาขามนุษยศาสตร์และสังคมศาสตร์*, ปีที่ 6, ฉบับที่ 3, พ.ศ. 2559, หน้า 133-145.
- [11] K. Lavagnananda and R. Poolphol, "A Genetic Algorithm Approach to Partitioning Clustering: A Case Study on M.Sc. Applicants", *In: 2014 13th International Conference on Machine Learning and Applications. IEEE*, p. 535-540, 2014
- [12] P. Naksuwan and K. Saithanu, "Prediction of learning achievement on elementary statistics course of Burapha University undergraduate students by discriminant analysis and neural networks" *Burapha Science Journal*, Vol.17, No.1, 2012, pp. 59-68
- [13] นภาพร อุทยานวุฒิกุล, "ปัจจัยที่มีผลต่อการจำแนกกลุ่มผลสัมฤทธิ์ทางการเรียนวิชาสถิติเพื่อการวิเคราะห์ข้อมูลทางสังคมศาสตร์ของนักศึกษาคณะนิเทศศาสตร์ มหาวิทยาลัยหอการค้าไทย", กองส่งเสริมวิจัยมหาวิทยาลัยหอการค้า, 2552
- [14] U. Ninrutsirikun, et al., "Effect of the Multiple Intelligences in multiclass predictive model of computer programming course achievement," *2016 IEEE Region 10 Conference (TENCON)*, Singapore, pp. 297-300, 2016
- [15] R. Malkawi, et al., "Data mining tools evaluation based on their quality attributes", *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp.13867-13890.
- [16] K. Lars, et al. "Auto-WEKA: Automatic model selection and hyperparameter optimization in WEKA.", *Automated machine learning: methods, systems, challenges* pp. 81-95, 2019

การวิเคราะห์ความคิดเห็นของผู้ใช้งานเพื่อสนับสนุน

การวางแผนบำรุงรักษาโมบายแอปพลิเคชัน User Feedback Analytics for Supporting Mobile Application Maintainability

สุจารี ไทยรัตน์ (Sujaree Thairat)¹ และมนวรรรัตน์ ผ่องไพบูลย์ (Monvorath Phongpaibul)²

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
sujaree.thai@dome.tu.ac.th, pmonvora@dome.tu.ac.th

บทคัดย่อ

บทความนี้นำเสนอแนวคิดในการวิเคราะห์ความคิดเห็นจากผู้ใช้งานโมบายแอปพลิเคชันแบบอัตโนมัติ โดยวิเคราะห์ความคิดเห็นแต่ละความคิดเห็นออกเป็น 2 มุมมอง มุมมองแรกคือมุมมองทางด้านคุณภาพของซอฟต์แวร์ และมุมมองที่สองเป็นมุมมองด้านฟังก์ชันการให้บริการของระบบตามขอบเขต ผลการวิเคราะห์ความคิดเห็นสามารถนำมาใช้ในการวางแผนการปรับปรุงความสามารถของโมบายแอปพลิเคชันในภายหลังได้ ตอบสนองความต้องการของผู้ใช้ได้ครอบคลุมต่อเนื่องและรวดเร็วยิ่งขึ้น ในงานวิจัยนี้ได้นำความคิดเห็นของผู้ใช้งานโมบายแอปพลิเคชันของธนาคาร เป็นกรณีทดสอบ จากผลการทดสอบสามารถวิเคราะห์ความถูกต้องของความคิดเห็นโดยรวมอยู่ที่ 74% สามารถจำแนกความคิดเห็นตามความรู้สึกของผู้ใช้ได้ถูกต้องที่ 92.5% สามารถจำแนกวัตถุประสงค์ของความคิดเห็นได้ถูกต้องที่ 91.5% และจัดความคิดเห็นตามมุมมองได้ถูกต้องที่ 80.5%

คำสำคัญ: โมบายแอปพลิเคชัน โมบายแบงก์กิ้ง วิเคราะห์ความคิดเห็นของผู้ใช้งาน บำรุงรักษาซอฟต์แวร์

Abstract

This study proposed the automated analysis of user feedback in the mobile applications. The feedbacks are analyzed in 2 aspects the software quality aspect and the service functions aspect of that mobile application. The result of the feedback analysis can be used to improve the maintenance plan of mobile

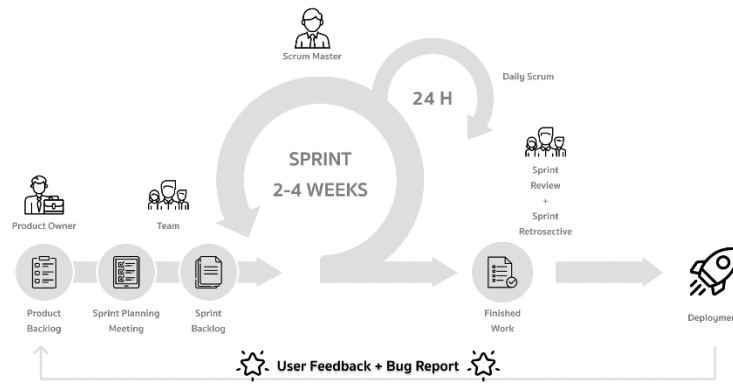
applications to respond continuously and quickly to customers' needs. In this study the user feedbacks of mobile banking applications are used as a case study. We found that our model can accurately analyze user feedback approximately 74%, classify sentiment of user feedbacks correctly at 92.5%, classify the intent of user feedbacks correctly at 91.5% and user feedbacks are grouped correctly at 80.5%.

Keyword: Mobile Applications, Mobile Banking, Review Analytics, Software Maintenance.

1. บทนำ

ปัจจุบันมีแอปพลิเคชันในรูปแบบของเว็บแอปพลิเคชัน (Web Application) และโมบายแอปพลิเคชัน (Mobile Application) เกิดขึ้นมากมาย เพื่อตอบสนองความต้องการของผู้ใช้งาน จากผลสำรวจของสำนักงานสถิติแห่งชาติ ประจำปี 2565 พบว่า คนไทยใช้สมาร์ทโฟน สูงถึง 93.2% และมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องและรวดเร็ว เพื่อให้แอปพลิเคชันที่สร้างขึ้นเป็นที่ต้องการของผู้ใช้งานและสามารถแข่งขันในตลาดได้ ผู้พัฒนาจำเป็นต้องให้ความสำคัญกับการพัฒนาและปรับปรุงความสามารถของแอปพลิเคชันอย่างต่อเนื่อง

สำหรับโมบายแอปพลิเคชันที่มีฐานผู้ใช้งานที่หลากหลายและมีจำนวนมาก ผู้พัฒนาระบบมักใช้กระบวนการพัฒนาแบบบอจิล (Agile) ซึ่งเป็นกระบวนการที่เน้นการพัฒนาเป็นรอบ ๆ โดยใช้ระยะเวลา



ภาพที่ 1: กระบวนการพัฒนาแบบอไจล์ (Agile Process)

ประมาณ 2-4 อาทิตย์ดังภาพที่ 1 เมื่อครบเวลา ฟังก์ชัน จะถูกปล่อยให้กับผู้ใช้งาน โดยเน้นการพัฒนา ความสามารถของแอปพลิเคชันที่มีคุณค่าแก่ผู้ใช้ การแบ่งงานเป็นรอบงานจะไม่เป็นการเพิ่มฟังก์ชัน แต่เป็นการปรับปรุงฟังก์ชันให้พร้อมใช้งาน ทำให้สามารถ ปล่อยแอปพลิเคชันออกมาได้บ่อยขึ้นและรอรับกระแสดอรับ จากผู้ใช้งานเพื่อนำมาปรับปรุงแอปพลิเคชันในภายหลัง [1]

ความคิดเห็นจากผู้ใช้งาน ถือเป็นกระแสดอรับ ที่มีความสำคัญเนื่องจากเป็นความคิดเห็นที่ผ่านการใช้งานจริง และมีอารมณ์การใช้งานประกอบในการแสดงความคิดเห็น ผู้ใช้งานมักแสดงความคิดเห็นทั้งในรูปแบบความพึงพอใจ ประสบการณ์ รายงานปัญหาจากการใช้งาน เปรียบเทียบ แอปพลิเคชันกับคู่แข่ง บอกความสามารถที่ต้องการ ซึ่งแตกต่างจากการรายงานปัญหาที่พบในขณะที่ใช้งานผ่าน ระบบรายงานปัญหา (Bug reporting system) ซึ่งผู้พัฒนา จะทราบถึงแต่ปัญหาการใช้งานระบบเท่านั้น

เพื่อให้ทีมพัฒนาแอปพลิเคชันสามารถนำความคิดเห็น จากผู้ใช้งานมาเป็นส่วนหนึ่งของการปรับปรุง ความสามารถของโมบายแอปพลิเคชันในภายหลัง ผู้วิจัยเสนอแนวคิดในการวิเคราะห์ความคิดเห็นจาก ผู้ใช้งานแบบอัตโนมัติ โดยวิเคราะห์ความคิดเห็นแต่ละ ความคิดเห็นออกเป็น 2 มุมมอง มุมมองแรกคือในมุมมอง ทางด้านคุณภาพของซอฟต์แวร์ ซึ่งเป็นมุมมองที่ทำให้ เข้าใจถึงความคิดเห็นของผู้ใช้ต่อคุณสมบัติทางคุณภาพ ซอฟต์แวร์ เช่น ความคิดเห็นต่อประสิทธิภาพการใช้งาน โมบายแอปพลิเคชัน โดยได้จัดแบ่งเป็นหมวดหมู่ตามหลัก ของ ISO/IEC 25010 ทั้งสิ้น 8 หมวดหมู่ มุมมองที่สอง

เป็นมุมมองด้านฟังก์ชันการให้บริการของระบบตาม ขอบเขตงาน ตัวอย่างเช่น การโอนเงิน การเบิกเงิน ผ่านยูเอทีเอ็ม เป็นการให้บริการของโมบายแอปพลิเคชัน ของระบบธนาคาร มุมมองนี้จะทำให้เข้าใจถึงความคิดเห็น ของผู้ใช้งานต่อฟังก์ชันการทำงานของโมบายแอปพลิเคชัน ซึ่งถูกแบ่งเป็น 7 หมวดหมู่ การวิเคราะห์ความคิดเห็นผ่าน มุมมอง 2 มุมมองนี้ทำให้ผู้พัฒนาระบบเข้าใจถึงรูปแบบ ของความคิดเห็นและสามารถนำมาใช้ในการวางแผนการ ปรับปรุงความสามารถของ โมบายแอปพลิเคชัน ในภายหลังได้ ตอบสนองความต้องการของผู้ใช้ได้ ครอบคลุม ต่อเนื่องและรวดเร็วยิ่งขึ้น

2.วรรณกรรมที่เกี่ยวข้อง

ผู้วิจัยศึกษางานวิจัยเกี่ยวกับกระบวนการพัฒนา แอปพลิเคชันบนมือถือ Flora และ Chande (2013) และ Kirmani (2017) ได้ศึกษาเกี่ยวกับกระบวนการพัฒนา แอปพลิเคชันบนมือถือด้วยวิธีการแบบอไจล์ และเปรียบเทียบกระบวนการแบบอไจล์กับกระบวนการ แบบไล่ระดับ พบว่า กระบวนการแบบอไจล์ถูกพัฒนาขึ้น เพื่อช่วยแก้ไขข้อบกพร่อง เช่น สามารถรองรับ การเปลี่ยนแปลงของลูกค้านี้ได้ตลอดเวลา ลูกค้านำมา เห็นความก้าวหน้าของงานอย่างรวดเร็ว [2], [3] งานวิจัย ของ Mahmud และคณะ (2015) และ Sume (2021) ศึกษาวิธีการแบบอไจล์ที่นำมาปรับใช้ในการพัฒนา แอปพลิเคชันบนมือถือ ซึ่งวิธีการแบบอไจล์เหมาะสมที่จะ นำมาเป็นกระบวนการพัฒนาแอปพลิเคชันบนมือถือ

เนื่องจากทีมงานในการพัฒนามีขนาดใหญ่และแอปพลิเคชันบนมือถือมีลักษณะเฉพาะ [4], [5]

หลังจากผู้วิจัยศึกษาวิธีการพัฒนาแอปพลิเคชันบนมือถือจึงสนใจหาวิธีที่จะนำความคิดเห็นของผู้ใช้งานแอปพลิเคชันกลับเข้ามาในกระบวนการพัฒนาอย่างรวดเร็ว เพื่อที่จะพัฒนาแอปพลิเคชันให้ตอบ โจทย์ และ ตรงตามความต้องการผู้ใช้งาน งานวิจัยเกี่ยวกับวิเคราะห์ความคิดเห็นมีที่ศึกษา เช่น งานวิจัยของพิศัญญ์ บวรเลิศสุธี และวรัทพร ไพโรเกรง (2565) ได้ศึกษาการวิเคราะห์ระดับความรู้สึกของความคิดเห็นต่อสินค้าและบริการออนไลน์ โดยวิเคราะห์รูปแบบและเทคนิคที่นำมาใช้กับความคิดเห็นที่เป็นภาษาไทย ใช้เทคนิคการเรียนรู้ของเครื่อง การทำเหมืองข้อความและเทคนิคการตัดคำ [6] Haruechaiyasak และคณะ (2010) ศึกษาการวิเคราะห์ความคิดเห็นที่เป็นภาษาไทยเกี่ยวกับการรีวิวโรงแรม โดยใช้วิธีการสร้างชุดของรูปแบบจากคลังข้อมูล พบว่ามีบางกรณีที่ไม่สามารถแยกความคิดเห็นได้เนื่องจากความยากของรูปแบบการเขียนความคิดเห็น [7] งานวิจัยของ Chumwatana (2015) ศึกษาการวิเคราะห์ความรู้สึกของความคิดเห็นที่แสดงออกโดยผู้ใช้งานสื่อสังคมในประเทศไทยด้วยวิธีการทำเหมืองข้อมูลและการสกัดคำภาษาไทย [8] Tarar และคณะ (2020) ศึกษาวิธีการช่วยนักพัฒนาในการสรุปรายงานและให้รายละเอียดข้อบกพร่องเพื่อช่วยลดระยะเวลาและมีความแม่นยำในการแนะนำข้อบกพร่อง โดยใช้วิธีการประมวลผลภาษาธรรมชาติ [9] Chiraratanasopha (2021) ศึกษาวิธีการจับคีย์เวิร์ดและวิเคราะห์อารมณ์เพื่อวิเคราะห์แนวโน้มของความรู้สึก โดยใช้วิธีการประมวลผลภาษาธรรมชาติและการจับคีย์เวิร์ดเพื่อวิเคราะห์แนวโน้มของความคิดเห็น [10] Panichella และคณะ (2021) ได้ศึกษาการวิเคราะห์ความคิดเห็นเกี่ยวกับแอปพลิเคชันบน Google Play และ App Store โดยใช้ 3 เทคนิค คือ การประมวลผลภาษาธรรมชาติ, การวิเคราะห์ข้อความ และวิเคราะห์ความรู้สึก ทำให้ผลลัพธ์จากการวิเคราะห์ดีขึ้น 75% [11] จากงานวิจัยที่ศึกษาจะเน้นความคิดเห็นจากผู้ใช้งานเป็นภาษาอังกฤษ ผู้วิจัยจึงสนใจศึกษาการวิเคราะห์ความคิดเห็นของผู้ใช้งานที่เป็นภาษาไทย

3. วิธีการดำเนินการวิจัย

งานวิจัยชิ้นนี้เสนอแนวคิดในการวิเคราะห์ความคิดเห็นจากผู้ใช้งาน โมบายแอปพลิเคชันแบบอัตโนมัติ เพื่อวิเคราะห์และจัดประเภทความคิดเห็นของผู้ใช้ โดยขั้นตอนการทำวิจัยประกอบด้วย 5 ขั้นตอน ได้แก่

3.1 ขั้นตอนการเก็บชุดข้อมูล

งานวิจัยนี้ผู้วิจัยได้เลือกศึกษาความคิดเห็นจากแอปพลิเคชัน โมบายธนาคารในประเทศไทย โดยผู้วิจัยเก็บความคิดเห็นการใช้งานแอปพลิเคชัน โมบายธนาคารที่มียอดดาวน์โหลดสูงสุด 5 อันดับแรก จำนวนทั้งสิ้น 1,200 ความคิดเห็น ซึ่งเป็นความคิดเห็นตั้งแต่ปี 2019 ถึงปี 2021 แบ่งเป็นความคิดเห็นจาก App Store 600 ความคิดเห็นและความคิดเห็นจาก Google Play 600 ความคิดเห็น

ผู้วิจัยแบ่งชุดความคิดเห็นออกเป็น 2 ชุด ได้แก่ ชุดสำหรับสร้างรูปแบบข้อความจำนวน 1,000 ความคิดเห็น และชุดสำหรับทดลองจำนวน 200 ความคิดเห็น

3.2 ขั้นตอนการออกแบบหมวดหมู่

หลังจากผู้วิจัยเก็บความคิดเห็นจาก App Store และ Google Play ผู้วิจัยได้นำตัวอย่างความคิดเห็นมาทำการวิเคราะห์เพื่อออกแบบหมวดหมู่สำหรับจำแนกความคิดเห็น โดยผู้วิจัยได้ออกแบบจำแนกหมวดหมู่ความคิดเห็นของผู้ใช้งานออกเป็น 2 มุมมอง คือ

1. มุมมองทางด้านซอฟต์แวร์ตามหลักของ ISO/IEC 25010 ประกอบด้วย 8 หมวดหมู่ ได้แก่

- ฟังก์ชันทำงานของระบบ (Functionality)
- ประสิทธิภาพการทำงาน (Performance)
- การทำงานของระบบร่วมกับระบบอื่น (Compatibility)
- การใช้งานต่าง ๆ ของระบบ (Usability)
- ความน่าเชื่อถือของระบบ (Reliability)
- ความปลอดภัยของระบบ (Security)
- การบำรุงรักษาระบบ (Maintainability)
- การทำงานข้ามสภาพแวดล้อม (Portability)

2. มุมมองด้านฟังก์ชันการให้บริการของระบบธนาคาร ประกอบด้วย 7 หมวดหมู่ ได้แก่

- ฟังก์ชันทำงานพื้นฐานแอปพลิเคชัน โฆษณา (Operation)
- ฟังก์ชันการทำงานเกี่ยวกับสแตตเมนต์ (Statement)
- ฟังก์ชันการยืนยันตัวตน (Verity)
- ฟังก์ชันเกี่ยวกับกองทุนต่าง ๆ ของธนาคาร (Fund)
- ฟังก์ชันพิเศษของแอปพลิเคชัน โฆษณา (Extra Functions)
- การตรวจสอบเพื่อให้แน่ใจว่าบุคคลที่เข้าถึงบัญชีนั้นเป็นเจ้าของบัญชี (Authentication)
- การแจ้งเตือนต่าง ๆ สำหรับธนาคาร (Notification (Bank Related))

3.3 ขั้นตอนการสร้างโมเดลในการวิเคราะห์

ผู้วิจัยดำเนินการสร้างรูปแบบข้อความ (Pattern) จากความคิดเห็นจำนวน 1,000 ความคิดเห็น ผู้วิจัยพบว่าการเขียนแสดงความคิดเห็นไม่มีรูปแบบที่ชัดเจน ส่งผลให้การจำแนกหมวดหมู่ความคิดเห็นจากรูปแบบข้อความเพียงอย่างเดียวมีความไม่แม่นยำสูง ตัวอย่างเช่น “แอปใช้งานง่าย ทำรายการไม่มีสะดุด โอนเงินสะดวกรวดเร็วมาก” สามารถจำแนกหมวดหมู่ได้ทั้งการใช้งานระบบ (Usability) และฟังก์ชันพื้นฐานแอปพลิเคชัน โฆษณา (Operation)

ผู้วิจัยจึงสร้างโมเดลในการวิเคราะห์เพื่อเพิ่มความถูกต้องในการจำแนกความคิดเห็นตามหมวดหมู่ที่ได้ออกแบบไว้ในขั้นตอนที่ 3.2 โดยมีโมเดลในการวิเคราะห์ 2 รูปแบบ ได้แก่

1. การสร้างกลุ่มของชุดคำสำหรับแทนคำที่มีความหมายเหมือนกันแต่เขียนได้หลากหลาย เพื่อใช้แทนที่คำที่มีความหมายเหมือนกันในรูปแบบข้อความที่สร้างขึ้น ดังแสดงในตารางที่ 1

ตารางที่ 1: คำตัวอย่างที่มีรูปแบบการเขียนที่หลากหลาย

คำที่เขียนถูก	รูปแบบการเขียน	แท็ก
แอป (Application)	แอป, แอพ, แอป, แอปป์, แอปพลิเคชัน, แอปพลิเคชัน	#app

2. การสร้างชุดคำเพื่อที่จะใช้จำแนกความรู้สึกของความคิดเห็นที่เป็นเชิงบวกหรือเชิงลบ ดังตารางที่ 2

ตารางที่ 2: ตัวอย่างคำที่ใช้จำแนกความรู้สึกของความคิดเห็น

ตัวอย่างคำที่ใช้จำแนก	แท็ก
ดี, สวย, ปัง, ทันสมัย, ยอดเยี่ยม, ดีๆ, น่าใช้, ไม่เชย, สะดวก, ง่าย, สุดยอด	#good
ไม่ดี, ห่วย, ไม่แสดง, ไม่ตรง, กาก, บ้าง, ล้ม, น่า, เค็ง, ใช้ไม่ได้, โบราณ, ยาก, แย่	#bad

ตัวช่วยทั้งสองถูกสร้างขึ้นในรูปแบบของแท็กที่นำมาใช้ในการสร้างรูปแบบข้อความ เพื่อให้การวิเคราะห์ข้อความสามารถนำคำในกลุ่มของแท็กดังกล่าวมาแทนแท็กในรูปแบบข้อความที่สร้างขึ้นได้ ดังตารางที่ 3

ตารางที่ 3: ตัวอย่างการสร้างรูปแบบข้อความเพื่อวิเคราะห์

ตัวอย่างความคิดเห็น	รูปแบบข้อความที่ใช้แท็ก
โอนเงินเข้า ยอดเงินไม่ขึ้นค่ะ รอมานิ่งแล้ว สลิปอะไรถูกหมด	<โอนเงิน><ยอดเงิน><ไม่ขึ้น> <#transfer><#money><#bad>

การสร้างรูปแบบข้อความที่ใช้แท็กนี้จะทำให้สามารถวิเคราะห์ความคิดเห็นของผู้ใช้งานที่มีความหมายใกล้เคียงกับตัวอย่างความคิดเห็นได้ เช่น “โอนเงินไปปลายทางแล้วเงินยังไม่เข้า”

3.4 ขั้นตอนการจำแนกหมวดหมู่

การสร้างรูปแบบข้อความจะต้องระบุหมวดหมู่เพื่อใช้จำแนกความคิดเห็น โดยแบ่งออกเป็น 3 กลุ่ม ได้แก่ กลุ่มสำหรับจำแนกความรู้สึกของความคิดเห็น กลุ่มสำหรับจำแนกวัตถุประสงค์ของความคิดเห็น และกลุ่มสำหรับจำแนกหมวดหมู่ของความคิดเห็น ผู้ใช้งานแบ่งไว้ 2 มุมมองตามหัวข้อ 3.2

ตารางที่ 4: ตัวอย่างรูปแบบข้อความเพื่อใช้ในการวิเคราะห์

รูปแบบข้อความ		
<โอนเงิน><ยอดเงิน><ไม่ขึ้น> <#transfer><#money><#bad>		
ความรู้สึก ความคิดเห็น	วัตถุประสงค์ ความคิดเห็น	หมวดหมู่ความคิดเห็น

เชิงลบ	ร้องเรียน	ฟังก์ชันทำงานพื้นฐาน แอปพลิเคชัน โมบาย ธนาคาร
--------	-----------	---

จากตารางที่ 4 เป็นตัวอย่างรูปแบบข้อความที่จะใช้ในการวิเคราะห์ พร้อมระบุหมวดหมู่เพื่อจำแนกความคิดเห็นของผู้ใช้งานเพื่อแสดงให้เห็นว่าเป็นความคิดเห็นของผู้ใช้งานเชิงลบที่ต้องการร้องเรียนในเรื่องของฟังก์ชันทำงานพื้นฐานแอปพลิเคชัน โมบายธนาคาร (Operation)

3.5 ขั้นตอนการวิเคราะห์และประเมินผล

ผู้วิจัยนำชุดข้อมูลสำหรับการทดลองมาวิเคราะห์กับรูปแบบข้อความที่สร้างขึ้น ซึ่งชุดข้อมูลสำหรับการทดลองถูกจำแนกหมวดหมู่ความคิดเห็นโดยผู้ที่มีประสบการณ์ในการพัฒนาระบบมากกว่า 3 ปี

วิธีการประเมินผลจะวัดความถูกต้องและความแม่นยำในการวิเคราะห์ความคิดเห็นของผู้ใช้งาน วิธีการประเมินผลแบ่งออกเป็น 2 ส่วน คือ

1. วัดผลจากการวิเคราะห์ความถูกต้อง 3 เดือน

$$P = \frac{\sum_{i=1}^n A(S,I,T)}{n} \times 100\%$$

โดยที่ $A(S,I,T) = \begin{cases} 1; Model \text{ ทำนาย } S,I \text{ และ } T \text{ ถูกต้อง} \\ 0; \text{ อื่น ๆ} \end{cases}$

2. วัดผลจากการวิเคราะห์ความถูกต้อง 1 เดือน

- วัดผลความคิดเห็นตามความรู้สึกของผู้ใช้

$$P_s = \frac{\sum_{i=1}^n A(S)}{n} \times 100\%$$

โดยที่ $A(S) = \begin{cases} 1; Model \text{ ทำนาย } S \text{ ถูกต้อง} \\ 0; \text{ อื่น ๆ} \end{cases}$

- วัดผลความคิดเห็นตามวัตถุประสงค์ของผู้ใช้

$$P_t = \frac{\sum_{i=1}^n A(I)}{n} \times 100\%$$

โดยที่ $A(I) = \begin{cases} 1; Model \text{ ทำนาย } I \text{ ถูกต้อง} \\ 0; \text{ อื่น ๆ} \end{cases}$

- วัดผลความคิดเห็นตามมุมมอง

$$P_r = \frac{\sum_{i=1}^n A(T)}{n} \times 100\%$$

โดยที่ $A(T) = \begin{cases} 1; Model \text{ ทำนาย } T \text{ ถูกต้อง} \\ 0; \text{ อื่น ๆ} \end{cases}$

เมื่อ S แทน ความคิดเห็นตามความรู้สึกของผู้ใช้
 I แทน ความคิดเห็นตามวัตถุประสงค์ของผู้ใช้
 T แทน ความคิดเห็นตามมุมมองของผู้ใช้
 n แทน จำนวนข้อมูลที่ใช้ทดสอบ

4. ผลลัพธ์ของการทดลอง

จากการนำชุดข้อมูลสำหรับการทดลองมาทดสอบกับรูปแบบข้อความ สามารถสรุปได้ว่า

ตารางที่ 5: ตารางแสดงการวัดผลความถูกต้อง 1 เดือน

การวัดผล	ความถูกต้อง (%)
ความคิดเห็นตามความรู้สึกของผู้ใช้	92.5%
ความคิดเห็นตามวัตถุประสงค์	91.5%
ความคิดเห็นตามมุมมอง	80.5%

จากตารางที่ 5 สามารถสรุปได้ว่า โมเดลสำหรับการวิเคราะห์ความคิดเห็นสามารถจำแนกความคิดเห็นตามความรู้สึกของผู้ใช้ถูกต้อง 92.5% สามารถจำแนกวัตถุประสงค์ของความคิดเห็นถูกต้อง 91.5% และจำแนกความคิดเห็นตามมุมมองได้ถูกต้อง 80.5% การเขียนและสะกดคำผิดของความคิดเห็นทำให้ค่าความถูกต้องลดลง

ตารางที่ 6: ตารางแสดงการวัดผลความถูกต้อง 3 เดือน

การวัดผล	ความถูกต้อง (%)
ความถูกต้องของความคิดเห็นโดยรวม	74%

ตารางที่ 6 แสดงผลการวิเคราะห์ความถูกต้องของความคิดเห็นจาก 3 เดือน อยู่ที่ 74% จะเห็นว่าความถูกต้องในการวิเคราะห์ลดลง เนื่องจากมุมมองความคิดเห็นอาจจะแตกต่างกัน

นอกจากผลความถูกต้องในการวิเคราะห์ โมเดลยังสามารถบอกถึงจำนวนความคิดเห็นในแต่ละหมวดหมู่ได้ดังตารางที่ 7

ตารางที่ 7: ตารางแสดงหมวดหมู่ที่ถูกวิเคราะห์เป็นจำนวนมาก

หมวดหมู่ความคิดเห็น	ความคิดเห็น
ฟังก์ชันทำงานพื้นฐาน	62
ฟังก์ชันการทำงานของระบบ	48
ฟังก์ชันการยืนยันตัวตน	39
อื่น ๆ	51

จากตารางที่ 7 จะเห็นว่า ผู้ใช้งานจะแสดงความคิดเห็นในเรื่องฟังก์ชันทำงานพื้นฐานของแอปพลิเคชัน โมบาย

ธนาคารมากเป็นอันดับแรกและฟังก์ชันการทำงานของระบบเป็นอันดับสอง

นอกจากการวิเคราะห์ความคิดเห็นของผู้ใช้งาน มีความคิดเห็นที่น่าสนใจ นอกเหนือจากรายงานปัญหาของระบบ ผู้วิจัยพบความคิดเห็นเกี่ยวกับการเปรียบเทียบฟังก์ชันการใช้งานระหว่างแอปพลิเคชันแต่ละธนาคาร ความคิดเห็นที่แสดงอารมณ์ของผู้ใช้งานกับแอปพลิเคชันการร้องขอฟังก์ชันเพิ่มเติม เช่น ความคิดเห็นที่แสดงอารมณ์ของผู้ใช้งานกับแอปพลิเคชัน เช่น “แอปนี้ชอบทำเรื่องง่ายให้เป็นเรื่องยาก ระบบช่วยแตก ปากบอกยกระดับๆ ไร้คีย์บอร์ดความห่วย” ซึ่งโมเดลในการวิเคราะห์ความคิดเห็นจะช่วยวิเคราะห์ความคิดเห็นในลักษณะดังกล่าวได้

5. สรุป

งานวิจัยชิ้นนี้ผู้วิจัยเสนอการสร้างโมเดลสำหรับวิเคราะห์ความคิดเห็นของผู้ใช้งานแอปพลิเคชันธนาคารในประเทศไทย โดยใช้วิธีการประมวลผลภาษาธรรมชาติและการจับคีย์เวิร์ดเพื่อวิเคราะห์ความรู้สึกของผู้ใช้งาน ผู้วิจัยกำหนดหมวดหมู่สำหรับการวิเคราะห์ความคิดเห็น 2 มุมมอง ได้แก่ มุมมองทางด้านซอฟต์แวร์ตามหลักของ ISO/IEC 25010 และมุมมองด้านฟังก์ชันการให้บริการของระบบธนาคาร จากผลการทดลองโมเดลมีความถูกต้องในการจำแนก ช่วยให้ทราบถึงความรู้สึกและวัตถุประสงค์ของผู้ใช้ รวมถึงจัดความคิดเห็นตามมุมมอง โดยโมเดลสำหรับการวิเคราะห์ความคิดเห็นสามารถวิเคราะห์ปัญหาของระบบที่เกิดขึ้น ข้อผิดพลาดที่พบ และยังสามารถวิเคราะห์ความคิดเห็นเชิงเปรียบเทียบระหว่างธนาคาร ความคิดเห็นที่แสดงถึงอารมณ์เพื่อนำไปปรับปรุงความสามารถในการให้บริการของแอปพลิเคชัน

เอกสารอ้างอิง

- [1] Pekka Abrahamsson, Outi Salo, Jussi Ronkainen, Juhani Warsta, “Agile software development methods: review and analysis,” VTT Technical report, 2002.
- [2] Harleen K. Flora, Swati V. Chande, “A Review and Analysis on Mobile Application Development Processes Using Agile Methodologies,” *International Journal of Research in Computer Science*. eISSN 2249-8265, vol. 3, no. 4, pp. 9-18, 2013.
- [3] Mudasir M. Kirmani, “Agile Development Method for Mobile applications: A Study,” *International Journal of Advanced Research in Computer Science*. ISSN NO. 0976-5697, vol. 8(5), pp. 586-590, 2017.
- [4] Dewi Mariati Mahmud, Nur Atiqah Sia Abdullah, “Reviews on Agile Methods in Mobile Application Development Process,” *9th Malaysian Software Engineering Conference (MySEC)*, 2015.
- [5] Sume S, Rujuta Hudge, “Mobile Application Development: An Agile Approach,” *International Journal of Scientific Research in Engineering and Management (IJSREM)*, ISSN NO. 2582-3930, vol. 05, 2021.
- [6] พิธิษฐ์ บวรเลิศสุธี และ วรภัทร ไพรเกรง, “การวิเคราะห์ความรู้สึกทางอารมณ์ในบทความแนะนำสินค้าออนไลน์,” *วารสารวิจัย มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย* ปีที่ 14 ฉบับที่ 3 พ.ศ. 2565 หน้า 755-769.
- [7] Choochart Haruechaiyasak, Alisa Kongthon, Pornpimon Palingoon, Chatchawal Sangkeetrakarn, “Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews,” *Proceedings of the 8th Workshop on Asian Language Resources*, pp. 64-71, 2010.
- [8] Todsanai Chumwatana, “USING SENTIMENT ANALYSIS TECHNIQUE FOR ANALYZING THAI CUSTOMER SATISFACTION FROM SOCIAL MEDIA,” *Proceeding of the 5th International Conference on Computing and Informatics, ICOCI 2015*, no. 146, pp. 659-664, 2015.
- [9] M. Irtaza Nawaz Tara, Faizan Ahmed, Wasi Haider Butt, “Automated Summarization of Bug Report to speed-up software development/maintenennance process by using Natural Language Processing (NLP),” *The 15th International Conference on Computer Science & Education (ICCSE 2020)*, pp. 483-488, 2020.
- [10] Boonthida Chiraratanasopha, “Sentimental Analysis and Keyword Extraction from thai Users of Facebook in COVID-19 Period,” *Progress in Applied Science*

and Technology, ISSN NO.2730-3020, vol.11, no.1, 2021.

- [11] Sebastiano Panichella, Andrea Di Sorbo, Emitza Guzman, Corrado A. Visaggio, Gerardo Canfora, Harald C. Gall “How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution,” *ICSME 2015, Bremen, Germany*, pp. 281-290, 2021.

การจำแนกโรคอัลไซเมอร์ด้วยการแยกองค์ประกอบแบบผันแปรร่วมกับ การวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี

Alzheimer's Disease Classification Using Two-Dimensional Variational Mode Decomposition and Principal Component Analysis with Entropy

สิทธิศักดิ์ แซ่จิ่ง (Sittisak Saechueng)¹ อังศุมาลี สุทธภักติ (Ungsumalee Suttapakti)²
เอกภพ บุญเพ็ง (Aekapop Bunpeng)³ และวิทวัส พันธุมจินดา (Wittawas Puntumchinda)⁴

คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

¹64910134@go.buu.ac.th, ²ungsumalee.su@go.buu.ac.th, ³aekapop@buu.ac.th, ⁴wittawas@go.buu.ac.th

บทคัดย่อ

โรคอัลไซเมอร์ทำให้เกิดภาวะสมองเสื่อม ซึ่งมักพบบ่อยในผู้สูงอายุ โรคอัลไซเมอร์ไม่สามารถรักษาให้หายขาดได้ แต่การตรวจพบในระยะเริ่มต้นจะช่วยชะลอความเสื่อมของสมองได้ จึงนำคอมพิวเตอร์วิทัศน์มาช่วยในการวิเคราะห์โรคอัลไซเมอร์จากภาพเอ็มอาร์ไอ วัตถุประสงค์เพื่อเพิ่มประสิทธิภาพของการจำแนกโรคอัลไซเมอร์ งานวิจัยนี้จึงนำเสนอการจำแนกโรคอัลไซเมอร์จากภาพเอ็มอาร์ไอ ด้วยวิธีการแยกองค์ประกอบแบบผันแปรร่วมกับการวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี วิธีการแยกองค์ประกอบแบบผันแปรใช้ในการแยกองค์ประกอบภาพเนื่องจากเป็นวิธีแบบปรับเหมาะที่มีประสิทธิภาพ ส่วนการวิเคราะห์องค์ประกอบหลักร่วมกับเอ็นโทรปีใช้ในการสกัดคุณลักษณะเด่น เนื่องจากเป็นวิธีการลดมิติข้อมูล โดยยังคงลักษณะที่สำคัญของข้อมูล ฐานข้อมูลที่ใช้ในการทดลอง คือ OASIS โดยมีภาพสมองเอ็มอาร์ไอทั้งหมดจำนวน 60 ภาพ ประกอบด้วย ภาพสมองที่ไม่เป็นโรคและเป็นโรค อย่างละเท่ากัน จำนวน 30 ภาพ ผลการทดลองพบว่า ค่าความถูกต้องของการจำแนกโรคของวิธีที่นำเสนอได้เท่ากับ 94.44 เปอร์เซ็นต์ ซึ่งมากกว่าวิธีการแปลงเวฟเลตแบบไม่ต่อเนื่องที่มีค่าความถูกต้องของการจำแนกประเภท เท่ากับ 88.89 เปอร์เซ็นต์

คำสำคัญ: โรคอัลไซเมอร์ ภาพเอ็มอาร์ไอ การแปลงเวฟเลตแบบไม่ต่อเนื่อง การแยกองค์ประกอบแบบแปรผัน

Abstract

Alzheimer's disease is a cause of dementia. This disease is often found in the elderly and cannot be cured. If this disease is diagnosed in the initial phase, it will help slow degeneration of the brain. Thus, the computer vision is applied to analysis Alzheimer's disease from brain MRI images. Objective of this research is to increase the performance of Alzheimer's disease classification. Thus, in this research, we propose Alzheimer's disease classification by using two-dimensional variational mode decomposition (VMD) and principal component analysis with entropy. VMD method is used to decompose the composition of an image because it is an efficient adaptive method. Principal component analysis with entropy method is used to extract prominent features because this method reduces the dimensions of the data while preserving the important characteristics of the data. In the experiment, the OASIS dataset is used to test the effectiveness of the proposed method. The total MRI are 60 images which contain 30 normal and 30 abnormal MRI images. The results showed that the classification accuracy of the variational mode decomposition with PCA combined log energy is highest, 94.44 percent, which is more than the discrete wavelet transform, 88.89 percent.

Keywords: Alzheimer's Disease, MRI Image, Discrete Wavelet Transform, Variational Mode Decomposition.

1. บทนำ

ประชากรผู้สูงอายุมีแนวโน้มที่จะเกิดโรคสมองเสื่อมเพิ่มขึ้นเรื่อย ๆ โดยกระทรวงสาธารณสุขเปิดเผยว่าโรคอัลไซเมอร์เป็นสาเหตุสำคัญของกลุ่มอาการสมองเสื่อมโดยพบได้ถึงร้อยละ 60-70 โดยทุก ๆ 68 วินาทีจะมีผู้ป่วยด้วยโรคอัลไซเมอร์เพิ่มขึ้นหนึ่งรายในโลกลดกว่าทั่วโลกจะมีผู้ป่วยกว่า 50 ล้านคน สำหรับในประเทศไทยพบผู้ป่วยที่เป็นโรคอัลไซเมอร์ ประมาณร้อยละ 3 - 5 โดยจะพบมากขึ้นตามอายุและพบสูงถึงร้อยละ 30 ในผู้สูงอายุที่อายุมากกว่า 85 ปี ทั้งนี้ โรคอัลไซเมอร์ไม่สามารถรักษาให้หายขาดได้ แต่การตรวจพบในระยะเริ่มต้นจะช่วยชะลอความเสื่อมของสมองได้อย่างมีประสิทธิภาพ โดยปกติแพทย์จะวินิจฉัยโรคอัลไซเมอร์จากภาพสมองเอ็มอาร์ไอหรือภาพของสมองด้วยคลื่นแม่เหล็กไฟฟ้า (Magnetic resonance imaging: MRI) โดยลักษณะของภาพเอ็มอาร์ไอแสดงให้เห็นถึงการฝ่อของเนื้อสมองทั้งสมอง ร่วมกับการมีโพรงสมองที่ขยายใหญ่ขึ้น นอกจากนี้ ยังมีการเปลี่ยนแปลงของกลีบสมองบริเวณขมับ (Temporal lobe) ได้ชัดเจนว่ามีการฝ่อของส่วนใดส่วนหนึ่งของสมอง

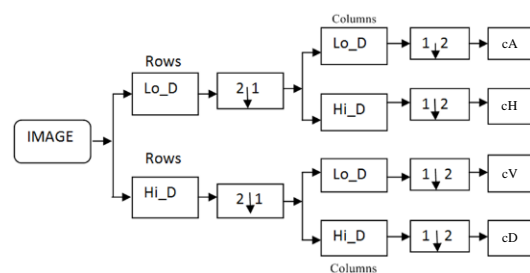
เนื่องจากเทคโนโลยีที่มีความก้าวหน้ารวดเร็วมากขึ้น การนำความรู้ด้านการประมวลผลภาพและการเรียนรู้ด้วยเครื่องจักรมาช่วยในการจำแนกโรคโรคอัลไซเมอร์แบบอัตโนมัติและมีความถูกต้องแม่นยำมากขึ้น ดังนั้น งานวิจัยนี้จึงนำเสนอการจำแนกโรคอัลไซเมอร์แบบอัตโนมัติด้วยการแยกองค์ประกอบแบบผันแปรร่วมกับการวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี เพื่อเพิ่มประสิทธิภาพของการจำแนกประเภทของโรคอัลไซเมอร์จากภาพเอ็มอาร์ไอ วิธีที่นำเสนอนี้มี 3 ขั้นตอน คือ ขั้นตอนที่ 1 เป็นการแยกองค์ประกอบของภาพด้วยวิธีการแยกองค์ประกอบแบบผันแปร (Variational mode decomposition, VMD) เพื่อแยกภาพให้มีองค์ประกอบเด่นสำคัญที่แตกต่างกัน จากนั้นองค์ประกอบที่ได้แต่ละตัวไปสกัดคุณลักษณะเด่นเป็นขั้นตอนที่ 2 ด้วยการวิเคราะห์องค์ประกอบหลัก (Principal component analysis, PCA)

และเอ็นโทรปีแบบ Log energy เพื่อเป็นสัณฐานลักษณะเด่นที่สำคัญและเป็นการลดขนาดของคุณลักษณะเด่นที่เหลือคุณลักษณะที่สำคัญ และขั้นตอนสุดท้ายเป็นการจำแนกประเภทโดยใช้การวัดระยะห่างแบบยูคลิด (Euclidean distance) ซึ่งเป็นวิธีที่ง่ายและมีประสิทธิภาพเพื่อจำแนกคุณลักษณะเด่นที่ได้จากภาพเอ็มอาร์ไอแล้วระบุว่าภาพใดเป็นโรคอัลไซเมอร์หรือไม่เป็นโรค

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การแปลงเวฟเล็ตแบบไม่ต่อเนื่อง (Two-dimensional discrete wavelet transform: DWT)

การแปลงเวฟเล็ตแบบไม่ต่อเนื่อง [1] เป็นเครื่องมือสำหรับแยกองค์ประกอบของสัญญาณในภาพออกเป็นส่วนของความถี่ต่ำและความถี่สูง เมื่อทำการแปลงเวฟเล็ตแบบไม่ต่อเนื่องในภาพ จะถูกแบ่งออกเป็นค่าสัมประสิทธิ์ 4 องค์ประกอบ [2] คือ สัมประสิทธิ์ความถี่ต่ำ (Approximate coefficients, cA) สัมประสิทธิ์ความสูงในแนวนอน (Horizontal detail coefficients, cH) สัมประสิทธิ์ความสูงในแนวตั้ง (Vertical detail coefficients, cV) สัมประสิทธิ์ความสูงในแนวทแยง (Diagonal detail coefficients, cD) ดังภาพที่ 1



ภาพที่ 1: การแยกองค์ประกอบในภาพของวิธี DWT

2.2 การแยกองค์ประกอบแบบแปรผัน (Two-dimensional variational mode decomposition, VMD)

การแยกองค์ประกอบแบบแปรผัน [3] เป็นการแยกองค์ประกอบภาพแบบปรับเหมาะที่มีประสิทธิภาพ และสามารถแยกองค์ประกอบตามข้อมูลของภาพ สามารถเขียนเป็นสมการได้ดังสมการที่ 1

$$\min_{u_k, o_k} \left\{ \sum_{k=1}^K \left\| \varphi \left[\delta(t) + \frac{j}{\pi t} \times v_k(t) \right] e^{j\omega t} \right\|_2^2 \right\}, \text{ s.t. } \sum_{k=1}^K v_k(t) = f \quad (1)$$

โดยที่ ω_k คือโหนดต่าง ๆ, ω_k คือ ความถี่, f คือ สัญญาณ, t คือเวลาที่แตกต่างกัน δ คือ ฟังก์ชันดิเรกเดลตา, k คือจำนวนโหนด และ * คือ การคอนโวลูชัน [4]

2.3 การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis)

การวิเคราะห์องค์ประกอบหลัก [5] เป็นวิธีที่ใช้ในการลดขนาดมิติของข้อมูล โดยการลดมิติข้อมูลที่ไม่สำคัญออกไป แต่ข้อมูลที่ถูกต้องจะต้องไม่สูญเสียความเป็นเอกลักษณ์ของข้อมูล ขั้นตอนการทำงานของ PCA มีดังนี้

- 1) เก็บข้อมูลและนำมาหาค่าเฉลี่ย
- 2) นำค่าเฉลี่ยมาลบกับข้อมูล
- 3) คำนวณ Covariance Matrix หรือเมตริกซ์ความแปรปรวนร่วมโดยใช้สมการที่ 2

$$\text{Cov}(X,Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n - 1)} \quad (2)$$

4) นำเมตริกซ์ความแปรปรวนร่วมที่ได้มาคำนวณหาค่า Eigenvectors และ Eigenvalues จากสมการที่ 3

$$Ax = \lambda x \quad (3)$$

โดยที่ A คือ เมตริกซ์ขนาด $n \times n$ มิติ ส่วน x คือ เวกเตอร์ขนาด $n \times 1$ มิติ และ λ คือ ค่าไอเกน (Eigenvalue)

2.4 การวัดประสิทธิภาพการจำแนกประเภท

การวัดประสิทธิภาพการจำแนกประเภท สามารถวัดได้จากตัวชี้วัดค่าความถูกต้อง (Accuracy) ดังสมการที่ 4

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (4)$$

โดย TP คือ จำนวนข้อมูลที่ทำนายได้ถูกต้อง ในกลุ่มที่ไม่เป็นโรค FP คือ จำนวนข้อมูลที่ทำนายผิด ในกลุ่มที่ไม่เป็นโรค TN คือ จำนวนข้อมูลที่ทำนายได้ถูกต้อง ในกลุ่มที่เป็นโรค และ FN คือ จำนวนข้อมูลที่ทำนายผิดในกลุ่มที่เป็นโรค

2.5 งานวิจัยที่เกี่ยวข้อง

Khagi [6] นำเสนอการสกัดคุณลักษณะเด่น โดยใช้เทคนิคคอนโวลูชัน (Convolutional Neural Network : CNN) และนำไปผ่านการลดจำนวนมิติข้อมูลด้วยเทคนิค PCA และ TSNE จากนั้นนำไปจำแนกประเภทโรคอัลไซเมอร์ด้วยโมเดลแบบ Machine learning เช่น KNN และ

Navies bays ผลการทดลองพบว่าการใช้เทคนิคแบบ Machine learning ให้ผลที่ไม่ค่อยดีนัก แต่ว่ายอยู่ในระดับที่ดีกว่าการใช้โมเดลจำแนกแบบ Softmax ของ CNN

Satapathy และคณะ [7] นำเสนอการใช้การแปลงเวฟเล็ดแบบไม่ต่อเนื่อง (DWT) ร่วมกับการวิเคราะห์องค์ประกอบหลัก (PCA) ในการเปรียบเทียบโมเดลในการจำแนกโรคอัลไซเมอร์ 2 โมเดล ได้แก่ Backpropagation Neural Network (BPNN) และ Extreme Learning Machine (ELM) ผลการทดลอง พบว่า วิธี DWT+PCA+ELM สามารถจำแนกโรคอัลไซเมอร์ได้แม่นยำมากที่สุด

Uma และคณะ[8] นำเสนอการใช้เทคนิคการสกัดคุณลักษณะเด่นแบบ Grey Level Co-occurrence Matrix (GLCM) แบบ Haralick ร่วมกับโมเดล SVM ซึ่งมีผลอยู่ในระดับที่ดีแม้จะใช้คุณลักษณะเด่นเพียง 13 คุณลักษณะ

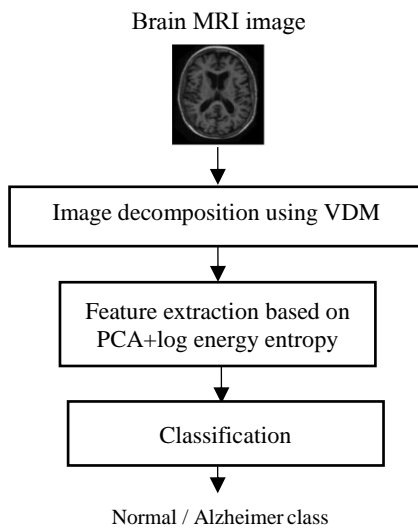
หลายงานวิจัยมีการใช้คอมพิวเตอร์วิทัศน์ (computer vision) และการเรียนรู้ด้วยเครื่อง (Machine learning) มีช่วยในการจำแนกโรคอัลไซเมอร์แบบอัตโนมัติ อย่างไรก็ตาม ยังคงต้องการเพิ่มประสิทธิภาพการจำแนกโรคอัลไซเมอร์ ดังนั้น งานวิจัยนี้จึงนำเสนอวิธีการจำแนกโรคอัลไซเมอร์แบบอัตโนมัติด้วยการแยกองค์ประกอบแบบผันแปรร่วมกับการวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี เพื่อเพิ่มประสิทธิภาพในระบุโรคอัลไซเมอร์จากภาพเอ็มอาร์ไอได้อย่างแม่นยำมากขึ้นในหัวข้อถัดไป

3. วิธีการที่นำเสนอ

ในหัวข้อนี้นำเสนอวิธีการจำแนกโรคอัลไซเมอร์แบบอัตโนมัติด้วยการแยกองค์ประกอบแบบผันแปรร่วมกับการวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี เพื่อเพิ่มประสิทธิภาพในระบุโรคอัลไซเมอร์จากภาพเอ็มอาร์ไอ แนวความคิดหลักของวิธีที่นำเสนอคือ การแยกองค์ประกอบภาพตามข้อมูลแบบปรับเหมาะ (Adaptive decomposition method) ทำให้ได้รับข้อมูลที่สำคัญช่วยติดตามจำแนกโรคได้อย่างแม่นยำมากขึ้น ขั้นตอนการทำงานของวิธีการที่นำเสนอ แสดงดังภาพที่ 2

วิธีการที่นำเสนอประกอบด้วย 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การแยกองค์ประกอบภาพ (Image decomposition) โดยใช้วิธีการแยกองค์ประกอบแบบแปรผัน (Variational mode decomposition, VMD)



ภาพที่ 2: กระบวนการทำงานของวิธีการที่นำเสนอ

โดยใช้สมการที่ 1 จะได้องค์ประกอบได้แก่องค์ประกอบที่ 1-4 (V1, V2, V3, และ V4) แสดงภาพที่ 3

ขั้นตอนที่ 2 การสกัดคุณลักษณะเด่น (Feature extraction) ในขั้นตอนนี้ใช้วิธี PCA และ log energy entropy โดยนำแต่ละองค์ประกอบของภาพที่ได้จากขั้นตอนที่ 1 มาวิเคราะห์ห่องค์ประกอบหลัก (PCA) จากนั้นนำองค์ประกอบหลักที่ได้ไปคำนวณหาค่า log energy entropy โดยใช้สมการที่ 4

$$E(s) = \sum_i \log(s_i^2) \tag{4}$$

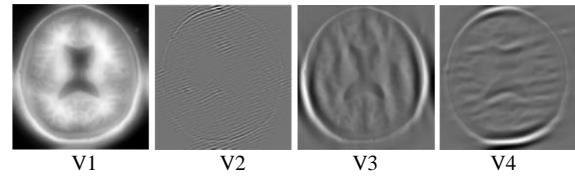
โดยที่ E คือ ค่าเอ็นโทรปี

ขั้นตอนที่ 3 การจำแนกประเภท (Classification) ของโรคอัลไซเมอร์ โดยใช้วิธีวัดระยะห่างแบบยุคลิด เพื่อระบุว่าภาพใดเป็นโรคอัลไซเมอร์หรือไม่เป็นโรคอัลไซเมอร์ ซึ่งการวัดระยะห่างแบบยุคลิด คำนวณได้ ดังสมการที่ 5

$$d(p,q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \tag{5}$$

โดยที่ $d(p, q)$ คือ ระยะห่างจากจุด p ไปยังจุด q ซึ่ง p คือ คุณลักษณะเด่นของข้อมูลชุดฝึกฝน q คือ

คุณลักษณะเด่นของข้อมูลชุดทดสอบ และ n คือ ขนาดของข้อมูล

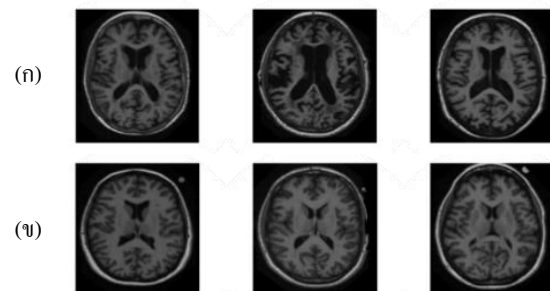


ภาพที่ 3: ภาพองค์ประกอบของที่ได้รับจากวิธี VMD

4. ผลการดำเนินการวิจัย

4.1 ฐานข้อมูลที่ใช้ในการทดสอบ

การเตรียมภาพที่ใช้ในการทดสอบเป็นภาพของสมองที่ได้มาจากฐานข้อมูลชื่อ OASIS [9] ที่เป็นภาพเอ็มอาร์ไอ ความละเอียดของภาพ คือ 208×176 พิกเซล ภาพทั้งหมดจำนวน 60 ภาพ แบ่งเป็นภาพเอ็มอาร์ไอที่เป็นโรคอัลไซเมอร์ และไม่เป็นโรคอัลไซเมอร์ จำนวนอย่างละ 30 ภาพ ตัวอย่างภาพสมองเอ็มอาร์ไอ แสดงดังภาพที่ 4



ภาพที่ 4: ตัวอย่างภาพสมองเอ็มอาร์ไอ โดยที่ (ก) คือ เป็นโรคอัลไซเมอร์ และ (ข) ไม่เป็นโรคอัลไซเมอร์

4.2 การออกแบบการทดลอง

4.2.1 การแปลงเวฟเล็ดแบบไม่ต่อเนื่อง (DWT) ได้ค่าสัมประสิทธิ์ 4 องค์ประกอบ คือ cA, cH, cV และ cD ซึ่งมีการทดสอบ ดังนี้

- 1) การใช้องค์ประกอบเดียว ได้แก่ cA, cH, cV และ cD
- 2) การใช้ 2 องค์ประกอบร่วมกัน ได้แก่ cA+cH, cA+cV, cA+cD, cH+cV, cH+cD และ cV+cD
- 3) การใช้ 3 องค์ประกอบร่วมกัน ได้แก่ cA+cH+cV, cA+cH+cD, cH+cV+cD
- 4) การใช้ทั้ง 4 องค์ประกอบ ได้แก่ cA+cH+cV+cD

4.2.2 การแยกองค์ประกอบแบบแปรผัน (VMD) ในการทดลองใช้วิธี VMD แยกภาพเป็น 4 องค์ประกอบ คือ องค์ประกอบที่ 1, 2, 3 และ 4 (V1, V2, V3, V4) โดยออกแบบการทดลอง ดังนี้

- 1) การใช้องค์ประกอบเดี่ยว ได้แก่ V1, V2, V3 และ V4
- 2) การใช้ 2 องค์ประกอบร่วมกัน ได้แก่ V1+V2, V1+V3, V1+V4, V2+V3, V2+V4 และ V3+V4
- 3) การใช้ 3 องค์ประกอบร่วมกัน ได้แก่ V1+V2+V3, V1+V2+V4, V1+V3+V4, V2+V3+V4
- 4) ใช้ทั้ง 4 องค์ประกอบร่วมกัน ได้แก่ V1+V2+V3+V4

4.2.3 การทดลองวิธีการสกัดคุณลักษณะเด่นที่ต่างกัน ได้แก่ วิธี PCA วิธี Log energy และวิธี PCA ร่วมกับ Log energy (PCA+Log energy)

นอกจากนี้ การแบ่งข้อมูลในการทดสอบได้แบ่งข้อมูลออกเป็นชุดฝึกฝนและชุดทดสอบ เท่ากับ 70 และ 30 เปอร์เซ็นต์ ตามลำดับ

4.3 ผลการทดลอง

ผลการวัดค่าความถูกต้องของการจำแนกประเภทของโรคอัลไซเมอร์ด้วยวิธีการแปลงเวฟเล็ดแบบไม่ต่อเนื่อง (DWT) แสดงดังตารางที่ 1

ตารางที่ 1: ค่าความถูกต้องของการจำแนกประเภทโรคอัลไซเมอร์ด้วยวิธีการแปลงเวฟเล็ดแบบไม่ต่อเนื่อง (DWT)

องค์ประกอบที่	ค่าความถูกต้อง (เปอร์เซ็นต์)		
	PCA	Log energy	PCA+Log energy
cA	72.22	66.67	55.56
cH	66.67	72.22	77.78
cV	77.78	66.67	72.22
cD	61.11	72.22	55.56
cA+cH	55.56	72.22	61.11
cA+cV	61.11	77.78	55.56
cA+cD	61.11	66.67	55.56
cH+cV	77.78	77.78	61.11
cH+cD	72.22	66.67	50.00
cV+cD	61.11	77.78	72.22
cA+cH+cV	61.11	66.67	55.56
cA+cH+cD	72.22	66.67	61.11

องค์ประกอบที่	ค่าความถูกต้อง (เปอร์เซ็นต์)		
	PCA	Log energy	PCA+Log energy
cH+cV+cD	55.56	88.89	50.00
cA+cH+cV+cD	61.11	66.67	50.00

หมายเหตุ ตัวหนา คือ ได้ค่าความถูกต้องมากที่สุด

จากตารางที่ 1 เป็นการจำแนกประเภทโรคอัลไซเมอร์ โดยการแยกภาพด้วยวิธี DWT พบว่า เมื่อใช้องค์ประกอบ cH+cV+cD และสกัดคุณลักษณะเด่นด้วยวิธี Log energy ได้ค่าความถูกต้องมากที่สุด เท่ากับ 88.89 เปอร์เซ็นต์

ผลการวัดค่าความถูกต้องของการจำแนกประเภทของโรคอัลไซเมอร์ด้วยวิธีการแยกองค์ประกอบแบบแปรผัน (VMD) แสดงดังตารางที่ 2

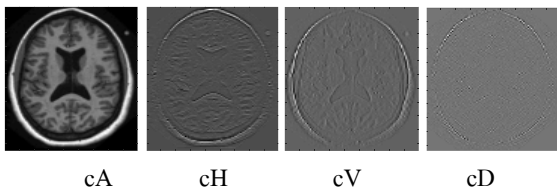
ตารางที่ 2: ค่าความถูกต้องของการจำแนกประเภทโรคอัลไซเมอร์ด้วยวิธีการแยกองค์ประกอบแบบแปรผัน (VMD)

องค์ประกอบที่	ค่าความถูกต้อง (เปอร์เซ็นต์)		
	PCA	Log energy	PCA+Log energy
V1	66.67	66.67	66.67
V2	55.56	61.11	61.11
V3	66.67	55.56	66.67
V4	55.56	61.11	50.00
V1+V2	77.78	72.22	72.22
V1+V3	77.78	66.67	94.44
V1+V4	61.11	61.11	77.78
V2+V3	72.22	66.67	66.67
V2+V4	44.44	61.11	55.56
V3+V4	55.56	55.56	55.56
V1+V2+V3	88.89	72.22	66.67
V1+V2+V4	66.67	61.11	55.56
V2+V3+V4	66.67	61.11	55.56
V1+V3+V4	72.22	61.11	66.67
V1+V2+V3+V4	72.22	66.67	66.67

หมายเหตุ ตัวหนา คือ ค่าความถูกต้องมากที่สุด

จากตารางที่ 2 การจำแนกประเภทโรคอัลไซเมอร์โดยการแยกภาพด้วยวิธี VMD พบว่า เมื่อใช้องค์ประกอบ V1+V3 และสกัดคุณลักษณะเด่นด้วยวิธี PCA+Log energy ได้ค่าความถูกต้องมากที่สุด เท่ากับ 94.44 เปอร์เซ็นต์

เมื่อเปรียบเทียบการแยกภาพทั้ง 2 วิธี (ตารางที่ 1 และ 2) พบว่า การแยกองค์ประกอบของภาพด้วยวิธี VMD ได้ผลดีกว่าวิธี DWT เนื่องจากวิธี VMD เป็นการแยกองค์ประกอบแบบปรับเหมาะตามข้อมูลทำให้การแยกองค์ประกอบภาพได้ข้อมูลที่เหมาะสมมากกว่าการใช้วิธี DWT ซึ่งเป็นการแยกองค์ประกอบแบบคงที่ทุกครั้งจะได้ 4 องค์ประกอบสัมประสิทธิ์ความถี่ต่ำ และความถี่สูงในแนวตั้ง นอน และทแยงเสมอ แสดงดังภาพที่ 5 ซึ่งบางองค์ประกอบอาจไม่เหมาะสมต่อการนำมาใช้ในการจำแนกโรคอัลไซเมอร์



ภาพที่ 5: ภาพองค์ประกอบของที่ได้รับจากวิธี DWT

5. สรุปผล

งานวิจัยนี้แนะนำการจำแนกโรคอัลไซเมอร์แบบอัตโนมัติด้วยการแยกองค์ประกอบแบบผันแปรร่วมกับการวิเคราะห์องค์ประกอบหลักและเอ็นโทรปี เพื่อเพิ่มความถูกต้องของการจำแนกประเภทของโรคอัลไซเมอร์จากภาพเอ็มอาร์ไอ วิธีที่นำเสนอนี้ ประกอบด้วย 3 ขั้นตอนคือ (1) การแยกภาพโดยใช้วิธีการแยกองค์ประกอบแบบผันแปร (VMD) เพื่อแยกภาพให้มีองค์ประกอบเด่นสำคัญที่แตกต่างกัน (2) การสกัดคุณลักษณะเด่นด้วยวิธีการวิเคราะห์องค์ประกอบหลัก (PCA) และเอ็นโทรปีแบบ Log energy เพื่อสกัดคุณลักษณะเด่นที่สำคัญและเป็นการลดขนาดของคุณลักษณะเด่นด้วย และ (3) การจำแนกประเภทโดยใช้การวัดระยะห่างแบบยูคลิด (Euclidean distance) เพื่อจำแนกคุณลักษณะเด่นที่ได้จากภาพเอ็มอาร์ไอแล้วระบุว่าภาพใดเป็นโรคหรือไม่เป็นโรคอัลไซเมอร์ ผลการทดลอง พบว่า การแยกภาพด้วยวิธี VMD โดยใช้ องค์ประกอบที่ 1 และ 3 และสกัดคุณลักษณะเด่นโดยวิธี PCA ร่วมกับ Log energy ได้ค่าความถูกต้องของการจำแนกโรคมามากที่สุดเท่ากับ 94.44% ซึ่งมากกว่าวิธีการอื่น ดังนั้นการใช้วิธีที่นำเสนอช่วยปรับปรุงค่าความถูกต้อง

ของการจำแนกประเภทได้ในอนาคต วิธีที่นำเสนออาจทดลองกับฐานข้อมูลภาพเอ็มอาร์ไออื่นเพิ่มเติมได้นอกจากนี้ ในอนาคตงานวิจัยนี้สามารถนำไปพัฒนาต่อเพื่อให้การเลือกองค์ประกอบเป็นแบบปรับเหมาะ (Adaptive method)

เอกสารอ้างอิง

- [1] วรรมณ กุศลกุลขจร. ตัวแบบการพยากรณ์ข้อมูลอากาศบนพื้นที่สูง ของประเทศไทย โดยระเบียบวิธีโครงข่ายประสาทเทียมร่วมกับการแปลงแบบเวฟเล็ต. [สืบค้นวันที่ 4 กุมภาพันธ์ 2566]. จาก https://www2.mtec.or.th/th/e-magazine/admin/upload/305_36.pdf
- [2] ศรีัญญา ชัยวิสูตร. การตรวจจับจุดออกสัญญาณคลื่นไฟฟ้าหัวใจโดยใช้การแปลงเวฟเล็ต. [สืบค้นวันที่ 4 กุมภาพันธ์ 2566]. จาก <http://kb.psu.ac.th/psukb/bitstream/2016/10271/1/369960.pdf>
- [3] Lahmiri, S., "Image characterization by fractal descriptors in variational mode decomposition domain: Application to brain magnetic resonance," *Physica A: Statistical Mechanics and its Applications*, vol. 456, pp. 235-243, 2016.
- [4] C. Montalvo, C.J. Gavilán, and A. García-Berrocal, "Variational mode decomposition method (VMD) applied to Decay Ratio (DR) calculation for instabilities identification in BWR," *Nuclear Engineering and Design*, Vol. 390, 2022.
- [5] ชัยลจิต สร้อยเพชรเกษม. การวิเคราะห์ส่วนประกอบสำคัญ. [สืบค้นวันที่ 4 กุมภาพันธ์ 2566]. จาก <http://www.edu.tsu.ac.th/major/eva/files/journal/PRINCIPAL.pdf>
- [6] V. Bairagi, "EEG signal analysis for early diagnosis of Alzheimer disease using spectral and wavelet based features," *Int. J. Inf. Technol.* pp. 403–412, 2018.
- [7] P. Satapathy, S. K. Pradhan, and S. Hota, "An Empirical Performance Analysis of Brain Image Classification Models Using Variants of Neural Networks," 2019 International Conference on Applied Machine Learning (ICAML), Bhubaneswar, India, 2019, pp. 87-91.
- [8] K. Uma Rani, S. S. Sharvari, M. G. Umesh, and B. C. Vinay, "Binary Classification of Alzheimer's disease using MRI images and Support Vector Machine," 2021 IEEE Mysore Sub Section International

Conference (MysuruCon), Hassan, India, 2021, pp.
423-426.

[9] OASIS-3 dataset, <https://www.oasis-brains.org/#data>

การพัฒนาต้นแบบแอปพลิเคชันให้คำแนะนำด้านสุขภาพ ด้วยเทคนิคการเรียนรู้ของเครื่อง

Development of a Prototype Health Care Recommender Application with Machine Learning Techniques

อุไรวรรณ อินทร์แหยม (Uraiwan Inyaem)¹ ชไมพร แยมกลิ่น (Chamaiphon Yeamklin)²

และศิริวรรณ สีมอ (Siriwarn Seemo)³

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
¹uraiwan.inyaem@rmutt.ac.th, ²1158109050992@mail.rmutt.ac.th, ³1158109050927@mail.rmutt.ac.th

บทคัดย่อ

ประเทศไทยก้าวเข้าสู่สังคมผู้สูงอายุเป็นเรื่องที่มีการพูดถึงอย่างแพร่หลายในปัจจุบัน ในประเทศไทยกลุ่มประชากรที่มีอายุตั้งแต่ 60 ปีขึ้นไป มีอยู่ 1 ใน 6 ของประชากรไทยทั้งหมด คนไทยมีอายุที่ยืนยาวขึ้น การดูแลสุขภาพจึงเป็นเรื่องที่ต้องให้ความสำคัญเพิ่มขึ้น ประกอบกับเทคโนโลยีที่ทันสมัย สามารถนำมาเป็นเครื่องมือในการสร้างสุขภาพที่ดี และการดำเนินชีวิตที่เหมาะสมให้กับผู้สูงอายุได้ บทความวิจัยนี้นำเสนอแอปพลิเคชันสำหรับให้คำแนะนำด้านสุขภาพบนระบบปฏิบัติการแอนดรอยด์ ที่พัฒนาขึ้นเน้นรูปแบบที่เข้าใจง่าย เน้นการใช้รูปภาพและตัวอักษรที่มีขนาดใหญ่ มีฟังก์ชันการทำงานที่เหมาะสมสำหรับผู้สูงอายุ สามารถบันทึกอาการของสุขภาพในแต่ละวันเพื่อประมวลผลว่าตนเองมีภาวะเสี่ยงที่จะเป็นโรคอะไร สามารถค้นหาข้อมูลของโรคจากอาการที่เป็นอยู่ได้ มีฟังก์ชันสำหรับการค้นหาโรงพยาบาลใกล้เคียง และสำหรับผู้สูงอายุที่อยู่คนเดียวจะมีฟังก์ชันปุ่มฉุกเฉินไว้สำหรับติดต่อกับคนใกล้ชิด สถานีตำรวจ และหน่วยกู้ชีพฉุกเฉินด้วย จากการทดลองแอปพลิเคชันกับผู้ใช้งาน พบว่า สามารถช่วยลดปัญหาสุขภาพ ปัญหาทางใจ ทำให้ผู้สูงอายุรู้สึกว่าได้ถูกมองข้าม ไม่ได้ถูกทอดทิ้งให้โดดเดี่ยว ช่วยให้ผู้สูงอายุมีความสุขกับเทคโนโลยีและการใช้ชีวิตมากยิ่งขึ้น

คำสำคัญ: การแนะนำสุขภาพ ผู้สูงอายุ แอปพลิเคชัน

Abstract

Thailand enters an aging society is a subject that is widely talked about today. In Thailand, one-sixth of population in Thailand will have age more than sixty years old. Today, people have long live, therefore; health care systems and modern technologies are very important in ageing society. This paper proposes a recommender application for health care via android operating system. The application has various functions which are focusing on easy patterns to understand for elderly people, for example; demonstrating with pictures and large font size. For daily recommendation, elderly people can save daily healthy details into the system. It can tell risk conditions to deceases and recommend treatments. Moreover, there are functions for searching nearby hospitals, recuse teams, police stations, and closed relatives. These functions are very helpful for lonely elders. The experimental results from users are shown that it reduces stress and health problems from users. In addition, elderly people are happy in their daily live because of good health care system from the application.

Keywords: Recommender Health Care, Elderly People, Application.

1. บทนำ

ปัจจุบันประเทศไทยก้าวเข้าสู่สังคมผู้สูงอายุ และจำนวนผู้สูงอายุก็จะยังเพิ่มขึ้นอย่างต่อเนื่อง มีการคาดการณ์ว่าจะมีประชากรที่มีอายุตั้งแต่ 60 ปีขึ้นไป อยู่ราว 1 ใน 6 ของประชากรทั้งหมด [1] เนื่องมาจากคนไทย มีอายุที่ยืนยาวขึ้น ปัญหาของผู้สูงอายุกับสุขภาพจึงเป็น สิ่งที่สำคัญ เรามักจะได้ยินเสมอว่า ผู้สูงอายุถูกปล่อยทิ้งไว้ ให้อยู่คนเดียว จึงส่งผลกระทบต่อสุขภาพ สภาพจิตใจ และการใช้ชีวิตของผู้สูงอายุ ทั้งที่จริงแล้วเทคโนโลยี สมัยใหม่สามารถทำให้ผู้สูงอายุติดต่อกับสื่อสารกับคน ในครอบครัวได้มากยิ่งขึ้น อำนวยความสะดวกทาง ด้านสุขภาพ สถานพยาบาลและเป็นช่องทางในการค้นหา ข้อมูลสุขภาพที่ตนเองสนใจได้ ดังนั้นการส่งเสริม ให้ผู้สูงอายุสามารถดูแลตนเองได้ จะช่วยให้ผู้สูงอายุมีชีวิต ที่ยืนยาวขึ้นและใช้ชีวิตอยู่ในสังคมได้อย่างมีความสุข ในทุกวันนี้ไม่ใช่แค่เด็กนักเรียน วัยรุ่น และวัยทำงาน เท่านั้น ที่จะผูกชีวิตอยู่กับสังคมออนไลน์แต่ผู้สูงอายุก็เป็น บุคคลอีกกลุ่มหนึ่งที่เริ่มหันมาใช้แอปพลิเคชันบนสมาร์ตโฟน ถึงแม้ว่าเรื่องของผู้สูงอายุกับแอปพลิเคชันในสมาร์ตโฟน นั้นจะดูเป็นเรื่องที่ไปด้วยกันไม่ได้เป็นอย่างมาก เพราะมี ผู้สูงอายุจำนวนน้อยมากที่จะใช้สมาร์ตโฟน ด้วยเหตุผล ที่ว่าสมาร์ตโฟนยากต่อความเข้าใจในการใช้งาน อายุที่มากขึ้น ทำให้ความคล่องตัวในการฝึกหัดความรู้ใหม่เป็นไปได้ยาก แต่ด้วยวัยของผู้สูงอายุประสิทธิภาพในการทำงานและการ ประกอบอาชีพก็จะลดลง มีการเกษียณอายุงาน ทำให้ผู้สูงอายุเป็นวัยที่มีเวลาว่างมาก ผู้สูงอายุส่วนใหญ่ จึงหันมาสนใจเทคโนโลยี ต้องการใช้ประโยชน์จาก แอปพลิเคชันบนสมาร์ตโฟน ต้องการก้าวทันโลก ไม่ต้องการเป็นผู้สูงอายุที่ล้าหลังอีกต่อไป สำหรับสุขภาพ ของผู้สูงอายุก็เป็นเรื่องที่ใกล้ตัวมาก ถ้าเราทำให้เรื่องของ สุขภาพกับแอปพลิเคชันบนสมาร์ตโฟนเป็นเรื่องที่ง่าย ผู้สูงอายุก็จะเริ่มสนใจเทคโนโลยีมากขึ้นสามารถฝึกการ ใช้เทคโนโลยีผ่านแอปพลิเคชันที่จัดทำขึ้นได้ ซึ่งเป็น ประโยชน์ต่อตัวผู้สูงอายุเป็นอย่างมาก ผู้สูงอายุจะสามารถ วิเคราะห์สุขภาพของตนเองและดูแลตนเองเบื้องต้น ก่อนที่จะไปพบแพทย์ที่โรงพยาบาลได้ จากที่กล่าวมาข้างต้น

คณะผู้วิจัยจึงมีแนวคิดที่จะพัฒนาแอปพลิเคชันขึ้นมา โดยเน้นรูปแบบที่เข้าใจง่าย เน้นการใช้รูปภาพ และตัวอักษรที่มีขนาดใหญ่ มีฟังก์ชันการทำงาน ที่เหมาะสมกับผู้สูงอายุ สามารถบันทึกอาการของสุขภาพ ในแต่ละวันเพื่อประมวลผลว่าตนเองมีภาวะเสี่ยงที่จะเป็น โรคอะไร สามารถค้นหาข้อมูลของโรคจากอาการที่เป็นอยู่ ได้ มีฟังก์ชันสำหรับการค้นหาโรงพยาบาลใกล้เคียง และสำหรับผู้สูงอายุที่อยู่คนเดียว จะมีฟังก์ชันปุ่มฉุกเฉินไว้ สำหรับติดต่อกับคนใกล้ชิด สถานีตำรวจ และหน่วยกู้ชีพ ฉุกเฉินอีกด้วย คณะผู้วิจัยคาดหวังว่าแอปพลิเคชันนี้จะช่วย ลดปัญหาสุขภาพ ปัญหาทางใจ ทำให้ผู้สูงอายุรู้สึกว่า เขา ไม่ได้ถูกมองข้ามไม่ได้ถูกทอดทิ้งให้โดดเดี่ยว ช่วยให้ ผู้สูงอายุมีความสุขกับเทคโนโลยีและการใช้ชีวิตมากขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

งานวิจัยนี้จัดทำขึ้นเพื่อเป็นทางเลือกหนึ่งสำหรับผู้คน ที่รักสุขภาพและต้องการที่จะดูแลสุขภาพตนเองเบื้องต้น ก่อนที่จะไปพบแพทย์ คณะผู้วิจัยได้ศึกษาทฤษฎี ที่เกี่ยวข้องในการพัฒนาแอปพลิเคชัน ซึ่งประกอบไปด้วย หัวข้อดังต่อไปนี้ [2-8]

2.1 การเรียนรู้ของเครื่อง เป็นการศึกษาอัลกอริทึม ของคอมพิวเตอร์ที่มีการพัฒนา หรือที่รู้จักกันในนามของ การเรียนรู้ของเครื่อง (Machine Learning) เพื่อสร้าง แบบจำลองทางคณิตศาสตร์จากข้อมูลชุดตัวอย่าง หรือข้อมูลสอน เพื่อที่จะตัดสินใจหรือคาดการณ์ได้อย่าง ชัดเจน การเรียนรู้ของเครื่องสามารถนำไปประยุกต์ใช้งาน ได้หลากหลาย รวมถึงการจำแนกประเภทข้อมูลด้วย

2.2 อัลกอริทึมต้นไม้ตัดสินใจ ใช้ตัวแบบในรูปแบบต้นไม้ ตัดสินใจ โดยเริ่มจากโหนดรากและไล่ไปยังโหนดลูก จนถึงโหนดใบที่จะบอกคลาสของข้อมูลในการตัดสินใจ เป้าหมายของการสร้างต้นไม้ในการตัดสินใจ คือการ เรียนรู้จากชุดข้อมูลตัวอย่างเพื่อสร้างเป็นต้นไม้ที่ให้ความ ถูกต้องมากที่สุดในการทำงานด้านการจำแนกประเภท ข้อมูล

2.3 อัลกอริทึมนาอ็อบเบย์ หลักการทำงานอาศัย ความน่าจะเป็น ตามทฤษฎีของเบย์ (Bayes Theorem) ซึ่งมี

อัลกอริทึมที่ไม่ในการเรียนรู้ปัญหาที่เกิดขึ้น เพื่อนำมาสร้างเงื่อนไขการแบ่งคลาส ใช้การคำนวณหาความน่าจะเป็นในการทำนายผล เป็นเทคนิคในการแก้ปัญหาแบบที่สามารถคาดการณ์ผลลัพธ์ได้ โดยการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์เหมาะสมกับกรณีของชุดข้อมูลที่มีจำนวนมากและคุณสมบัติของข้อมูลไม่ขึ้นต่อกัน

2.4 อัลกอริทึมเคเนียร์สเนเบอร์ มีหลักการทำงาน โดยการเปรียบเทียบกับข้อมูลตัวอย่างที่เก็บอยู่ในฐานข้อมูล โดยจะค้นหาตัวอย่างข้อมูลในฐานที่มีลักษณะใกล้เคียงข้อมูลใหม่มากที่สุด จำนวน K ตัวมาเข้าร่วมกันตัดสินใจว่าข้อมูลใหม่นี้ควรจะมีความคล้ายคลึงกับประเภทของ Majority Vote ข้อมูลใหม่จะมีประเภทเดียวกับประเภทข้อมูลจำนวนมากที่สุดในข้อมูลเพื่อนบ้าน K ตัวนี้ จะมีการวัดความเหมือนระหว่างข้อมูลใหม่กับข้อมูลตัวอย่างในฐานข้อมูล โดยใช้ตัววัดความเหมือนนั่นเอง

3. วิธีการดำเนินการวิจัย

การดำเนินการวิจัยออกแบบเป็น 2 ส่วน คือ ส่วนแรกจะเป็นการสร้างแบบจำลองด้วยเทคนิคการทำเหมืองข้อมูลสำหรับการแนะนำข้อมูลให้กับผู้ใช้งาน และส่วนที่สองจะเป็นการนำแบบจำลองที่เหมาะสม ไปพัฒนาเป็นแอปพลิเคชัน สร้างความง่ายและสะดวกให้กับผู้ใช้งาน

3.1 การสร้างแบบจำลอง โดยการทำงานเริ่มจากเก็บรวบรวมความต้องการใช้แอปพลิเคชันด้านสุขภาพเพื่อหาฟังก์ชันหลัก ฟังก์ชันรองของแอปพลิเคชัน ในงานวิจัยนี้ คาด้าเซทที่ใช้สำหรับการสร้างแบบจำลองที่เหมาะสม เก็บรวบรวมจากการสัมภาษณ์บุคคลที่มีอายุ 50 ปีขึ้นไป ในบริเวณพื้นที่บ้านคลอง 23 ฝั่งเหนือ หมู่ 12 ตำบลศิระชะกระบือ อำเภอองครักษ์ จังหวัดนครนายก จำนวน 250 คน ข้อมูลที่เก็บรวบรวม ได้แก่ อายุ เพศ อาชีพ น้ำหนัก ส่วนสูง รอบเอว ความดันโลหิต จำนวนชีพจร อาการ และโรคประจำตัว จากนั้นทำการเตรียมข้อมูลที่ได้เก็บรวบรวมมา โดยข้อมูลในคอลัมน์อาการ ได้แก่ อาการกระหายน้ำ กลิ่นปัสสาวะไม่อยู่ กล้ามเนื้อเป็นตะคริว ข้ออักเสบ บวมแดง

กลิ่นได้ ความจำเสื่อม คัดจมูก ค่า cholesterol มากกว่า 200 mg/dL งอขาได้ไม่สุด จุกท้อง เจ็บหน้าอก ใจสั่น ซามือซาเท้า ตัวซีดตัวเหลือง ตาข้างขวมัว ตาข้างซ้ายมัว ตาพร่ามัว ตึงบริเวณหัวเข่า ท้องบวม ท้องอืด นอนไม่หลับ น้ำหนักลดลง แน่นท้อง แน่นหน้าอก เบื่ออาหาร ปวดกระบอกตา ปวดตามข้อแบบรุนแรง ปวดตามข้อหัวเข่า ปวดท้อง ปวดท้องน้อย ปวดบริเวณท้ายทอย ปวดศีรษะ ปวดศีรษะข้างเดียว ปวดศีรษะอย่างรุนแรง ปวดหลัง ปวดบวม ปัสสาวะบ่อย ปัสสาวะมีสีขุ่น แผลอากาศเย็น มองเห็นภาพซ้อน มีก้อนขนาดใหญ่บริเวณคอ มีไข้ มีน้ำมูก มีผื่นขึ้นทั่วร่างกาย มีฝ้าสีขาวในดวงตา มีอาการเรอและมีน้ำรสเปรี้ยว รสขมย้อนออกมาทางปาก มีอาการหลงลืม มีน้ศีรษะมือเท้าบวม มือสั่น รู้สึกเสียวตามข้อเวลาขกของหนักและเดินเยอะ ลมพิษขึ้นเป็นบางครั้ง เล็บเปราะ เลือดกำเดาไหล โลหิตจาง เวลาเคลื่อนไหวจะมีเสียงบริเวณข้อเข่า เวลาตื่นนอนรู้สึกไม่สดชื่น เวลาตื่นนอนจะหุดหุดก สมองไม่ดีสับสน เสียดท้อง แสบร้อนบริเวณหน้าอก หน้ามืด เวียนศีรษะ หัวใจเต้นเร็ว หายใจติดขัดหิวบ่อย เหนื่อยง่าย เหนื่อยล้า ร่างกาย เหยียดขาได้ไม่สุด อ่อนเพลีย อาเจียน โอบ่อย ไอเป็นเลือด ไอเรื้อรัง และคออักเสบ โรคประจำตัวที่เก็บได้มีทั้งหมด 22 โรค ได้แก่ กรดไหลย้อน กระเพาะปัสสาวะอักเสบ กระเพาะอาหาร เก๊าท์ ข้อเสื่อมปวดข้อ ไขมันในเลือดสูง ภาวะสมองจำเสื่อม ความดันโลหิตสูง ต้อกระจก ตับแข็ง ตับอักเสบบี ไตวายเรื้อรัง ไทรอยด์เป็นพิษ นอนไม่หลับ เบาหวาน ภูมิแพ้ มะเร็งปอด ไมเกรน โลหิตจาง วัณโรค หอบหืด หลอดเลือดหัวใจ จากนั้นนำข้อมูลที่ผ่านกระบวนการเตรียมข้อมูลมาทำการสร้างแบบจำลองโดยใช้เครื่องมือสำหรับการทำเหมืองข้อมูล กระบวนการทำงานแบ่งออกเป็น 2 กระบวนการย่อย คือ 1) การฝึกฝนข้อมูลในอดีต และ 2) การทดสอบแบบจำลอง โดยในกระบวนการฝึกฝนข้อมูลแบบมีผู้สอนจากข้อมูลที่เก็บไว้ในอดีต และมีคำตอบ ใช้การทดสอบแบบ Cross Validation กับข้อมูลที่เก็บรวบรวมมา จำนวน 250 รายการ อัลกอริทึมที่ใช้เปรียบเทียบการสร้างแบบจำลอง แบ่งเป็น 3 กลุ่ม ได้แก่ 1) อัลกอริทึมต้นไม้ตัดสินใจ 2) อัลกอริทึมนาอิวเบย์ และ 3) อัลกอริทึมเคเนียร์สเนเบอร์ สำหรับการ

สร้างแบบจำลอง จากนั้นทำการทดสอบประสิทธิภาพ อัลกอริทึมเพื่อคัดเลือกแบบจำลองที่ดีที่สุดมาพัฒนาเป็นแอปพลิเคชัน

3.2 การนำแบบจำลองไปพัฒนาเป็นแอปพลิเคชัน

สำหรับข้อมูลที่ได้จากการสัมภาษณ์ความต้องการของผู้ใช้ ได้นำข้อมูลมาวิเคราะห์เพื่อหาฟังก์ชันของแอปพลิเคชันซึ่งมีหลายข้อมูลที่ยังไม่ชัดเจนในเรื่องของความต้องการ จึงมีการสัมภาษณ์ผู้คนที่เพิ่มเติมแล้วนำมาวิเคราะห์ความต้องการอีกครั้ง จากนั้นก็ทำการออกแบบหน้าจอ โดยเริ่มจากวิเคราะห์การทำงานของแต่ละฟังก์ชัน โดยใช้รูปแบบการวิเคราะห์และออกแบบเชิงวัตถุ โดยแสดงด้วยแผนภาพยูสเคส และรายละเอียดของยูสเคสของแอปพลิเคชัน ซึ่งมีฟังก์ชันดังนี้ ฟังก์ชันส่องสุขภาพ ฟังก์ชันพยากรณ์สุขภาพ ฟังก์ชันโทรฉุกเฉิน ฟังก์ชันค้นหาโรงพยาบาลใกล้เคียง และฟังก์ชันข้อมูลส่วนตัว ส่วนฟังก์ชันพยากรณ์สุขภาพจำแนกแบบจำลองที่ได้จากกระบวนการ 3.1 มาพัฒนาเป็นแอปพลิเคชัน ซึ่งประกอบด้วยฟังก์ชันย่อย คือ พยากรณ์โรค คำนวณดัชนีมวลกาย คำนวณความดันโลหิต คำนวณชีพจร สำหรับการพัฒนาแอปพลิเคชันส่วนติดต่อกับผู้ใช้ พัฒนาขึ้นโดยใช้โปรแกรม Android Studio ใช้ภาษาจาวา ในการพัฒนา แล้วสร้างแอปพลิเคชันตามฟังก์ชันที่ได้วิเคราะห์ตามแผนภาพยูสเคส แอปพลิเคชันที่พัฒนาสมบูรณ์แล้ว จะถูกวัดประสิทธิภาพด้วยแบบสอบถามความพึงพอใจในการใช้งานแอปพลิเคชันแบบมาตราส่วนประมาณค่า 5 ระดับ ได้แก่ มากที่สุด มาก ปานกลาง น้อย และน้อยที่สุด เพื่อวิเคราะห์ข้อมูลพื้นฐานหาค่าเฉลี่ย (Mean) และการหาค่าความเบี่ยงเบนมาตรฐาน (Standard Deviation) ต่อไป

4. ผลการดำเนินงาน

ผลการดำเนินงานจะมี 2 ส่วนที่เกี่ยวข้องกัน คือ ส่วนแรกจะได้แบบจำลองที่เหมาะสมจากเทคนิคการทำเหมืองข้อมูล และส่วนที่สอง นำแบบจำลองที่ได้ไปพัฒนาเป็นแอปพลิเคชัน โดยตลอดการทำวิจัย ได้นำเครื่องมือมาใช้ในการทำงานเพื่อเพิ่มความสะดวกและได้ผลการวิจัยที่

ถูกต้องและมีประสิทธิภาพ ซึ่งประกอบด้วยโปรแกรมเพนซิล ใช้ในการออกแบบหน้าจอแอปพลิเคชันเพื่อติดต่อกับผู้ใช้ โปรแกรมเวก้าใช้สร้างแบบจำลองจากข้อมูลชุดฝึกฝน และนำแบบจำลองไปทดสอบกับข้อมูลชุดทดสอบ เพื่อคัดเลือกแบบจำลองที่เหมาะสมสำหรับไปพัฒนาเป็นแอปพลิเคชัน โปรแกรมไมโครซอฟท์ เอกซ์เซล ใช้ในการเก็บข้อมูลในรูปแบบ CSV สำหรับการทดลองในงานวิจัยนี้ โปรแกรมแอนดรอยด์สตูดิโอ ใช้พัฒนาแอปพลิเคชันในงานวิจัยนี้ แอปพลิเคชันที่พัฒนาขึ้นรองรับสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ ตั้งแต่เวอร์ชัน 4.0 ขึ้นไป

4.1 การหาแบบจำลองที่เหมาะสม เพื่อแนะนำด้านสุขภาพ เราแบ่งการทำงานออกเป็น 2 กระบวนการย่อย คือ 1) การฝึกฝนข้อมูลในอดีต และ 2) การทดสอบแบบจำลอง โดยในกระบวนการฝึกฝนข้อมูลแบบมีผู้สอน จากข้อมูลที่เก็บไว้ในอดีต และมีคำตอบ ใช้การทดสอบแบบ Cross Validation กับข้อมูลที่เก็บรวบรวมมาจำนวน 250 รายการ อัลกอริทึมที่ใช้เปรียบเทียบการสร้างแบบจำลอง แบ่งเป็น 3 กลุ่ม ได้แก่ 1) อัลกอริทึมต้นไม้ตัดสินใจ 2) อัลกอริทึมนาอิวเบย์ และ 3) อัลกอริทึมเคเนียร์สเนเบอร์ สำหรับอัลกอริทึมต้นไม้ตัดสินใจ ได้แก่ อัลกอริทึม DecisionStump, Random Forest, RandomTree และ REPTree ทุกอัลกอริทึมใช้สำหรับการจำแนกประเภทข้อมูลแบบ Nominal class โดยแอตทริบิวต์ของข้อมูลจะเป็นประเภท Nominal attributes ด้วย [9] สำหรับอัลกอริทึมนาอิวเบย์ ได้แก่ BayesNet และ NaiveBayesMultinomial และอัลกอริทึมเคเนียร์สเนเบอร์ ได้แก่ IBK ส่วนผลการวัดประสิทธิภาพแบบจำลอง โดยเปรียบเทียบการสร้างแบบจำลองที่เหมาะสมสำหรับนำไปพัฒนาเป็นแอปพลิเคชัน แสดงไว้ดังตารางที่ 1 ด้านล่างนี้

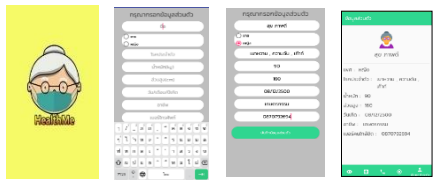
ตารางที่ 1: แสดงผลการวัดประสิทธิภาพแบบจำลอง

Algorithms	Correctly	Precision	Recall	F-Measure	ROC Area
DecisionStump	66.66%	66.70%	66.70%	66.70%	83.30%
Random Forest	95.33%	95.30%	95.30%	95.30%	99.40%
RandomTree	92.00%	92.00%	92.00%	92.00%	94.00%
REPTree	94.00%	94.00%	94.00%	94.00%	96.50%
BayesNet	92.67%	92.70%	92.70%	92.70%	98.00%
NaiveBayesMultinomial	95.33%	95.30%	95.30%	95.30%	98.80%
IBK (K-Nearest Neighbor)	95.33%	95.30%	95.30%	95.30%	96.60%

ผลการวิจัยพบว่า อัลกอริทึมที่ใช้สร้างแบบจำลองได้เหมาะสมกับข้อมูลได้แก่ อัลกอริทึมในกลุ่มต้นไม้ตัดสินใจ คือ Random Forest ที่มีค่าความถูกต้องสูงสุดเท่ากับ 95.33% ค่าความระลึกลับเท่ากับ 95.30% ค่าความแม่นยำเท่ากับ 95.30% ค่าความถ่วงดุลเท่ากับ 95.30% การพิจารณาพื้นที่ใต้กราฟ ROC มีค่าสูงถึง 99.40% ส่งผลให้ค่าความถูกต้องในการทำนายสูงด้วยเช่นกัน ตามมาด้วยอัลกอริทึม NaiveBayesMultinomial และอัลกอริทึมเคเนียร์เซนเนเบอร์ คือ IBK (K-Nearest Neighbor) ตามลำดับ

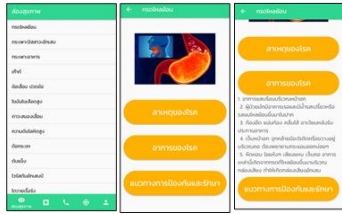
4.2 แอปพลิเคชันที่พัฒนาขึ้นสำหรับให้คำแนะนำด้านสุขภาพบนระบบปฏิบัติการแอนดรอยด์ โดยนำผลการเปรียบเทียบอัลกอริทึมที่ได้จากหัวข้อที่ 4.1 คืออัลกอริทึมในกลุ่มต้นไม้ตัดสินใจ Random Forest สำหรับพยากรณ์โรคมาพัฒนาเป็นแอปพลิเคชันที่มีฟังก์ชันการทำงาน 5 ฟังก์ชัน คือ ฟังก์ชันข้อมูลส่วนตัว ฟังก์ชันส่องสุขภาพ ฟังก์ชันพยากรณ์สุขภาพ ฟังก์ชันโทรฉุกเฉิน และค้นหาโรงพยาบาลใกล้เคียง ซึ่งมีหน้าจอ ดังนี้

1) ฟังก์ชันข้อมูลส่วนตัว: แสดงไอคอน หน้าเริ่มต้นแอปพลิเคชัน หน้าแสดงเมนูหลัก ฟังก์ชันข้อมูลส่วนตัวสามารถกรอกข้อมูลส่วนตัว แก้วไข และแสดงข้อมูลได้ ดังแสดงในภาพที่ 1 ด้านล่างนี้



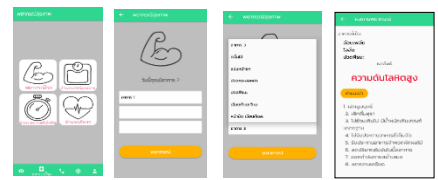
ภาพที่ 1: แสดงหน้าจอฟังก์ชันข้อมูลส่วนตัว

2) ฟังก์ชันส่องสุขภาพ: โดยหน้าแรกของฟังก์ชันแสดงรายการของโรค โดยผู้ใช้งานสามารถเลือกดูข้อมูลของโรค สาเหตุของโรค อาการของโรค และแนวทางการป้องกันและรักษา



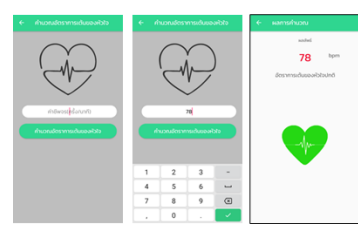
ภาพที่ 2: แสดงหน้าจอฟังก์ชันส่องสุขภาพ

3) ฟังก์ชันพยากรณ์สุขภาพ: มี 4 เมนูหลักให้ใช้งาน ได้แก่ พยากรณ์โรค คำนวณค่าดัชนีมวลกาย คำนวณความดันโลหิต และคำนวณชีพจร สำหรับฟังก์ชันการพยากรณ์โรค ผู้ใช้งานสามารถกรอกอาการ ได้ 3 อย่าง จากนั้น ระบบจะทำการพยากรณ์โรคที่ตรงตามแบบจำลองที่ได้สร้างไว้พร้อมคำแนะนำในการปฏิบัติตนให้กับผู้ใช้งาน ดังแสดงในภาพที่ 3 ด้านล่าง



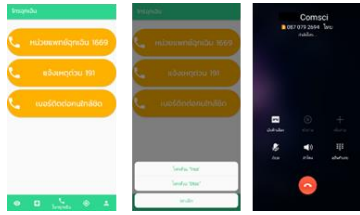
ภาพที่ 3: แสดงหน้าจอฟังก์ชันพยากรณ์สุขภาพ

สำหรับฟังก์ชันคำนวณค่าดัชนีมวลกาย ผู้ใช้งานสามารถกรอกรายละเอียดน้ำหนักเป็นหน่วยกิโลกรัมและส่วนสูงเป็นหน่วยเซนติเมตร ระบบจะคำนวณค่าดัชนีมวลกายเป็นค่า BMI แสดงผลลัพธ์เป็นระดับต่าง ๆ พร้อมคำแนะนำในการปฏิบัติตน ส่วนฟังก์ชันคำนวณความดันโลหิต การใช้งานผู้ใช้งานต้องกรอกรายละเอียดของค่าด้านบนหน่วยเป็น mmHg และค่าด้านล่างหน่วยเป็น mmHg เพื่อคำนวณค่าความดันโลหิตจากข้อมูลที่ผู้ใช้งานกรอกเข้ามา และแสดงผลเป็นระดับความดันพร้อมคำแนะนำในการรับประทานอาหารและออกกำลังกาย ส่วนฟังก์ชันคำนวณชีพจรต้องกรอกข้อมูลค่าชีพจร หน่วยเป็นครั้งต่อนาที เพื่อคำนวณอัตราการเต้นของหัวใจ และแสดงผลเป็นอัตราการเต้นของหัวใจ ดังแสดงในภาพที่ 4 ด้านล่าง



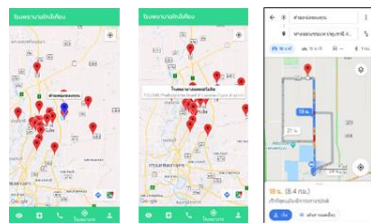
ภาพที่ 4: แสดงหน้าจอการคำนวณการเต้นของหัวใจ

4) ฟังก์ชันโทรฉุกเฉิน: สามารถเลือกการโทรด่วนหาหน่วยแพทย์ฉุกเฉิน 1669 แจ้งเหตุด่วน 191 เบอร์ติดต่อคนใกล้ชิด แสดงดังภาพที่ 5 ด้านล่าง



ภาพที่ 5: แสดงหน้าจอฟังก์ชันโทรฉุกเฉิน

5) ฟังก์ชันค้นหาโรงพยาบาลใกล้เคียง: ระบบจะเชื่อมกับกูเกิลแมพ (Google map) เพื่อแสดงตำแหน่งของผู้ใช้งาน และแสดงตำแหน่งโรงพยาบาลใกล้เคียง ผู้ใช้สามารถเลือกชื่อโรงพยาบาล ระบบจะแสดงรายละเอียดและนำทางไปยังโรงพยาบาลได้เลย ดังแสดงในภาพที่ 6 นี้



ภาพที่ 6: แสดงหน้าจอฟังก์ชันค้นหาโรงพยาบาลใกล้เคียง

5. สรุป

การพัฒนาแอปพลิเคชัน สร้างขึ้นเพื่อให้คำแนะนำด้านสุขภาพกับผู้สูงอายุ การพัฒนาเริ่มจากการเปรียบเทียบเทคนิคการทำเหมืองข้อมูล 3 กลุ่ม ได้แก่ 1) อัลกอริทึมต้นไม้ตัดสินใจ 2) อัลกอริทึม Naïve Bayes และ 3) อัลกอริทึมเคเนียร์เซนเนอร์ กับข้อมูลจำนวน 250 รายการ เพื่อหาแบบจำลองที่เหมาะสมในการพยากรณ์โรค ผลการวิจัยพบว่า อัลกอริทึมที่ใช้สร้างแบบจำลองได้เหมาะสมกับข้อมูล ได้แก่ อัลกอริทึมในกลุ่มต้นไม้ตัดสินใจ Random Forest ที่มีค่าความถูกต้องสูงสุด เพื่อนำไปพัฒนาเป็นต้นแบบแอปพลิเคชัน ที่มีฟังก์ชันการทำงานด้านส่งสุขภาพ ที่ผู้ใช้สามารถดูข้อมูลของโรคได้จำนวน 22 โรค ซึ่งในข้อมูลจะบอก สาเหตุของโรค อาการของโรค และแนวทางการรักษาและป้องกัน ฟังก์ชันพยากรณ์สุขภาพ มี 4 ฟังก์ชันย่อย 1) พยากรณ์โรค ผู้ใช้ต้องกรอกอาการ เพื่อพยากรณ์โรค และแสดงคำแนะนำกับผู้ใช้ 2) จำนวนดัชนีมวลกาย ผู้ใช้ต้องกรอกน้ำหนัก

และส่วนสูง เพื่อคำนวณค่าดัชนีมวลกาย โดยระบบจะส่งค่าดัชนีมวลกายและคำแนะนำออกมา 3) จำนวนความดันโลหิต ผู้ใช้จะต้องกรอกความดันโลหิตตัวล่างและตัวบน เพื่อคำนวณความดันโลหิต โดยระบบจะส่งคำแนะนำออกมา 4) จำนวนชีพจร ผู้ใช้ต้องกรอกอัตราการเต้นของหัวใจ แล้วระบบจะส่งคำแนะนำออกมา ฟังก์ชันโทรฉุกเฉิน ผู้ใช้สามารถโทรออกเบอร์ฉุกเฉินได้ทั้งหมด 3 เบอร์ ได้แก่ เบอร์สถานีตำรวจ 191 เบอร์หน่วยกู้ชีพฉุกเฉิน 1669 และเบอร์โทรคนใกล้ชิดที่ได้บันทึกไว้ในฟังก์ชันข้อมูลส่วนตัว ฟังก์ชันค้นหาโรงพยาบาลใกล้เคียง ผู้ใช้สามารถค้นหาตำแหน่งของโรงพยาบาลใกล้เคียงในรัศมี 25 กิโลเมตรได้ และฟังก์ชันข้อมูลส่วนตัว ผู้ใช้ต้องบันทึกข้อมูลส่วนตัวก่อน ถึงจะสามารถเข้าใช้งานแอปพลิเคชันได้ สำหรับการประเมินความพึงพอใจในการใช้งานแอปพลิเคชันนั้น ได้จัดทำแบบสอบถามสำหรับผู้เชี่ยวชาญ 5 คน และสำหรับผู้ใช้งาน 30 คน ซึ่งให้ผลการประเมินผลจากผู้เชี่ยวชาญได้ค่าเฉลี่ยรวมเท่ากับ 4.13 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.50 และการประเมินจากผู้ใช้งานทั่วไปได้ค่าเฉลี่ยรวมเท่ากับ 4.23 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.49 สรุปได้ว่าแอปพลิเคชันที่ได้ทำการพัฒนาขึ้นมานั้น มีประสิทธิภาพอยู่ในระดับดีสามารถนำแอปพลิเคชันที่พัฒนาขึ้นไปใช้จริงได้

เอกสารอ้างอิง

- [1] สำนักงานปลัด กระทรวงการพัฒนาสังคมและความมั่นคงของมนุษย์. สถิติผู้สูงอายุ. [ออนไลน์]. 2565. แหล่งที่มา: <https://www.dop.go.th/th/know/1> [1 ธันวาคม 2565]
- [2] หทัยรัตน์ เกตุมณีชัยรัตน์. การจำแนกประเภท. [ออนไลน์]. 2565. แหล่งที่มา : <https://www.slideshare.net> [1 ธันวาคม 2565]
- [3] อัจฉราภรณ์ จุฑาผาด. การพัฒนาระบบสารสนเทศเพื่อการพยากรณ์จำนวนนักศึกษาใหม่โดยใช้การจำแนกต้นไม้ตัดสินใจ. วิจัยและนวัตกรรมกับการพัฒนาประเทศ. หน้า 267-278.
- [4] สุรพงศ์ เอื้อวัฒนามงคล. การทำเหมืองข้อมูล, สำนักพิมพ์บัณฑิตพัฒนบริหารศาสตร์: กรุงเทพฯ, 2561.
- [5] Eakasit Pacharawongsakda. *An Introduction to Data Mining Techniques*, Asia Digital Press: Bangkok, 2014.

- [6] Ronen Feldman and James Sanger. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, Cambridge University Press: United States of America, 2007.
- [7] ปริญญา สงวนศักดิ์. *การเรียนรู้ของเครื่อง Machine Learning*, สำนักพิมพ์สถาบันการจัดการปัญญาภิวัตน์ : กรุงเทพฯ, 2562.
- [8] สายชล สนิสมบูรณ์ทอง. *การทำเหมืองข้อมูล Data Mining*, สำนักพิมพ์จามจุรีโปรดักท์: กรุงเทพฯ, 2558.
- [9] Ian Witten. *Data Mining with Weka Course*, Waikato University : New Zealand, 2015.

แบบจำลองการวิเคราะห์พฤติกรรมกรรมการโจมตีทางไซเบอร์โดยใช้ปัญญาประดิษฐ์ ในกองทัพอากาศ

A Model for Analysis of Cyber Attack Behaviors Using Artificial Intelligence in the Royal Thai Air Force

ณัฐคนัย กองขุนทด (Natdanai Kongkhunthod)¹ ประสงค์ ปราณีตพลกรัง (Prasong Praneetpolgrang)²

และพายัพ ศิรินาม (Payap Sirinam)³

หลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีป้องกันประเทศ

^{1,2,3} สำนักบัณฑิตศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

natdanai4@gmail.com¹, prasongspu@gmail.com², p.sirinam@gmail.com³

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อสร้างแบบจำลองการวิเคราะห์พฤติกรรมกรรมการโจมตีทางไซเบอร์โดยใช้ปัญญาประดิษฐ์ในกองทัพอากาศ เครื่องมือที่ใช้ในการวิเคราะห์พฤติกรรมกรรมการโจมตีทางไซเบอร์คือ โปรแกรม RapidMiner Studio โดยใช้อัลกอริทึมการจำแนกข้อมูลการโจมตี 4 อัลกอริทึมด้วยกันคือ อัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees โดยใช้งานร่วมกับชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ (RTAF Dataset) ผลการทดสอบพบว่า อัลกอริทึม Gradient Boosted Trees มีค่าเฉลี่ยความแม่นยำ (Accuracy) และค่าเฉลี่ยความเที่ยง (Precision) เป็นเปอร์เซ็นต์สูงสุด โดยมีค่าเฉลี่ยความแม่นยำเท่ากับ 99.08% ความเที่ยงเท่ากับ 80.28% รวมไปถึงค่าความระลึก (Recall) เท่ากับ 72.82% และค่าความถ่วงดุล (F-Measure) เท่ากับ 76.37% รองลงมาคืออัลกอริทึม Naïve Bayes โดยมีค่าเฉลี่ยความแม่นยำเท่ากับ 98.12% ความเที่ยงเท่ากับ 69.69% ค่าความระลึกเท่ากับ 85.38% และค่าความถ่วงดุลเท่ากับ 76.74% รองลงมาคืออัลกอริทึม Decision Tree โดยมีค่าเฉลี่ยความแม่นยำเท่ากับ 95% ค่าความเที่ยงเท่ากับ 77.18% ค่าความระลึกเท่ากับ 57.84% และค่าความถ่วงดุลเท่ากับ 66.12% และอัลกอริทึม Random Forest โดยมีค่าเฉลี่ยความแม่นยำเท่ากับ 94.98% ค่าความเที่ยงเท่ากับ 77.53% ค่าความระลึกเท่ากับ 56.72% และค่าความถ่วงดุลเท่ากับ 65.51%

ในงานวิจัยนี้แสดงให้เห็นว่าอัลกอริทึม Gradient Boosted Trees มีค่าเฉลี่ยความแม่นยำ (Accuracy) และค่าเฉลี่ยความเที่ยง (Precision) เป็นเปอร์เซ็นต์สูงสุด สำหรับชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ (RTAF Dataset) คำสำคัญ: การโจมตีทางไซเบอร์ ปัญญาประดิษฐ์ แบบจำลอง

Abstract

The purpose of this research is to create a cyber attack behavior analysis model by using artificial intelligence in the Royal Thai Air Force. RapidMiner Studio is the tool for analyzing cyber attack behavior with four attack classification algorithms: Decision Tree, Naïve Bayes, Random Forest and Gradient Boosted Trees. These algorithms operate in conjunction with the Royal Thai Air Force Cyber Attack Dataset. The results showed that the Gradient Boosted Trees algorithm had the highest percentage of average accuracy and average precision; the average accuracy was 99.08%, the average precision was 80.28%, including the recall was 72.82% and the F-Measure was 76.37%. Followed by the Naïve Bayes algorithm with an average accuracy was 98.12%, the average precision was 69.69%, the recall was 85.38%, and the F-Measure was 76.74%. The consequent one was the Decision Tree algorithm with an average accuracy of 95%, the average precision was 77.18%, the recall was 57.84%, and

the F-Measure was 66.12%. The Random Forest algorithm with average accuracy of 94.98%, average precision of 77.53%, the recall was 56.72%, and the F-Measure of 65.51%. In this research, we conclude that the Gradient Boosted Trees algorithm has the highest percentages of average accuracy and average precision for RTAF Dataset.

Keywords: Cyber Threat, Artificial Intelligence, Model.

1. บทนำ

ปัจจุบันภัยคุกคามทางไซเบอร์นั้นวันยังมีจำนวนมากขึ้น และมีความรุนแรงมากขึ้น อีกทั้งยังมีความซับซ้อนมากขึ้น ซึ่งการโจมตีในแต่ละครั้ง ได้ส่งผลกระทบต่อองค์กร ภาครัฐ และภาคเอกชน ไม่ว่าจะด้านความมั่นคงด้านการเงิน การธนาคาร ด้านการแพทย์ และข้อมูลส่วนบุคคลมากขึ้น

ศูนย์ประสานการรักษาความมั่นคงปลอดภัยระบบคอมพิวเตอร์ประเทศไทย (ThaiCERT) รายงานสถิติภัยคุกคามจากเหตุการณ์ที่เกิดขึ้นทั้งหมด 1,884 กรณี ตั้งแต่ช่วงต้นปี 2565 พบว่าปัญหาในด้านช่องโหว่ของการรักษาความมั่นคงปลอดภัยไซเบอร์หรือ Malicious Code เป็นอันดับ 1 ของการโจมตี โดยคิดเป็น 53.82% ของการโจมตีทั้งหมด ในขณะที่ภัยคุกคามในรูปแบบของช่องโหว่โปรแกรมหรือ Vulnerability ตามมาเป็นอันดับ 2

National Cyber Security Index (NCSI) เป็นหน่วยงานที่รวบรวมดัชนีเพื่อใช้วัดความพร้อมของประเทศทั่วโลก ในการป้องกันภัยคุกคามทางไซเบอร์ และจัดการเหตุการณ์ทางไซเบอร์ โดยได้จัดอันดับของประเทศไทยให้อยู่ในอันดับที่ 41 จากประเทศทั้งหมด 161 ประเทศ ซึ่งตกจากอันดับที่ 40 จากรายงานปี 2565

กองทัพอากาศเป็นหน่วยงานด้านความมั่นคงของรัฐ หน่วยงานหนึ่ง ที่ตกเป็นเป้าหมายการโจมตีทางไซเบอร์ อันดับต้น ๆ ของประเทศ ประกอบกับในปัจจุบันเทคโนโลยีที่ได้ก้าวกระโดดไปอย่างรวดเร็ว การที่จะหวังพึ่งพาระบบวิเคราะห์และตรวจจับภัยคุกคามทางไซเบอร์ อาจยังไม่เพียงพอ อันเนื่องจากปัญหาของระบบที่ใช้ในการป้องกันมีข้อจำกัด รวมถึงเทคโนโลยีที่เปลี่ยนไปอย่าง

รวดเร็ว หรือระบบที่ทางหน่วยงานได้ทำการจัดหาอาจไม่ครอบคลุมด้านการเฝ้าระวังภัยคุกคามที่เกิดขึ้นใหม่ ดังนั้นการพัฒนา ระบบปัญญาประดิษฐ์ (Artificial Intelligence: AI) ด้วยอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning: ML) จึงเป็นแนวทางที่ดีที่สุดที่จะสามารถนำมาใช้ในการแก้ไขปัญหาดังกล่าวได้

จากที่กล่าวมาข้างต้นผู้วิจัยได้มีความสนใจ อัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning: ML) ในการสร้างแบบจำลองการวิเคราะห์พฤติกรรมกรมการโจมตีทางไซเบอร์โดยใช้ปัญญาประดิษฐ์ในกองทัพอากาศ โดยผู้วิจัยมีความสนใจ โปรแกรม RapidMiner Studio เพื่อใช้สำหรับการควบคุมและออกแบบการทำงานของแบบจำลองให้เป็นอัตโนมัติ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบตรวจจับการบุกรุก (Intrusion Detection System: IDS)

ระบบตรวจจับการบุกรุก (IDS) คืออุปกรณ์หรือแอปพลิเคชันซอฟต์แวร์ที่คอยทำหน้าที่ตรวจจับการบุกรุกทางไซเบอร์ เมื่อมีการเข้าถึงหรือกิจกรรมที่ถูกระบุว่าเป็นภัยคุกคามหรือการละเมิดนโยบาย หากมีการตรวจจับเหตุการณ์หรือกิจกรรมที่ผิดปกติ (Anomaly) ระบบจะทำหน้าที่แจ้งเตือนไปยังระบบส่วนกลางเพื่อให้ผู้ดูแลทราบ และทำการตรวจสอบทันที

RapidMiner คือ Data Science Platform [1] ใช้สำหรับการออกแบบการวิเคราะห์ข้อมูลและออกแบบอัลกอริทึมที่สามารถทำการจัดการข้อมูล และสร้างอัลกอริทึมในรูปแบบตามที่กำหนด รวมถึงการแสดงผลการทดสอบประสิทธิภาพได้

2.2 การเรียนรู้ของเครื่อง (Machine Learning: ML)

คือ ปัญญาประดิษฐ์ (Artificial Intelligence: AI) ประเภทหนึ่ง [8] ที่ใช้วิธีการฝึกให้คิดเหมือนมนุษย์ โดยอาศัยข้อมูลจากประสบการณ์ที่เคยเกิดขึ้นแล้วในอดีตกลับมาสอนให้คิด และวิเคราะห์ข้อมูลให้มีความฉลาดเหมือนมนุษย์

2.3 อัลกอริทึมที่ใช้จำแนกภัยคุกคามเพื่อวิเคราะห์พฤติกรรมกรมการโจมตีทางไซเบอร์

2.3.1 อัลกอริทึมต้นไม้ตัดสินใจ (Decision Tree) คือแบบจำลองทางคณิตศาสตร์ เพื่อหาแนวทางที่ดีที่สุด ด้วยการนำข้อมูลมาสร้างเป็นโครงสร้างแบบต้นไม้เพื่อทำการพยากรณ์

2.3.2 อัลกอริทึม Naïve Bayes (Naïve Bayes) คืออัลกอริทึมในการจำแนกข้อมูลโดยใช้ความน่าจะเป็น และวิธีการทางสถิติจากกฎของเบย์ เพื่อหาความน่าจะเป็นที่ถูกต้องที่สุด โดยอาศัยข้อมูลก่อนหน้ากับข้อมูลปัจจุบันเพื่อหาสมมติฐานที่ดีที่สุด

2.3.3 อัลกอริทึมการตัดสินใจแบบสุ่ม (Random Forest) เป็นอัลกอริทึมการเรียนรู้ของเครื่องสำหรับแก้ปัญหาการจำแนกประเภทและการถดถอย โดยใช้การนำอัลกอริทึมต้นไม้ตัดสินใจ มาทำการเพิ่มจำนวนต้นไม้หลายๆ ต้นเพื่อเพิ่มประสิทธิภาพการพยากรณ์ [2]

2.3.4 อัลกอริทึม Gradient Boosting คือ อัลกอริทึมพื้นฐานที่มาจาก การสร้าง Decision Tree โดยการสร้าง Decision Tree แบบสุ่มขึ้นมาจำนวนมาก และทำการประเมินผลแต่ละอัลกอริทึมจนกว่าจะได้ Decision Tree ที่สมบูรณ์ที่สุด [3]

2.4 ชุดข้อมูล (Dataset)

ชุดข้อมูลระบบตรวจจับการบุกรุกทางไซเบอร์ (RTAF Dataset) ของศูนย์ไซเบอร์กองทัพอากาศ ที่ได้ทำการรวบรวมมาจาก IDS ที่กองทัพอากาศใช้ในปัจจุบันนั้น จะเป็นชุดข้อมูลที่บันทึกข้อมูลในระบบเครือข่ายคอมพิวเตอร์ในกองทัพอากาศ ซึ่งมีรูปแบบภัยคุกคาม 17 ประเภท จำนวนข้อมูลทั้งสิ้น 638,134 รายการ

2.5 งานวิจัยที่เกี่ยวข้อง

อภิชาติ ปัตถาวะโร (2561) ได้นำเสนอเกี่ยวกับการวิเคราะห์การบุกรุกบนระบบเครือข่ายโดยใช้อัลกอริทึมการเรียนรู้ของเครื่อง โดยนำอัลกอริทึม Machine Learning (ML) แบบไฮบริด มาใช้ในการตรวจสอบการโจมตีบนเครือข่าย โดยใช้อัลกอริทึม Random Forest, Adaboost, Gradient Boosting และ K-Means Clustering กับชุดข้อมูล NSL-KDD ในการทดสอบ [4]

วรารินทร์ ปัญญาวงษ์ (2562) ได้นำเสนอเกี่ยวกับการเปรียบเทียบประสิทธิภาพการตรวจจับมัลแวร์โดยใช้

อัลกอริทึมเหมือนข้อมูล เพื่อทำการสร้างแบบจำลองในการตรวจจับมัลแวร์ โดยใช้อัลกอริทึมเหมือนข้อมูลกับอัลกอริทึม Support vector machines (SVM), Naïve Bayes และ K-Nearest Neighbors (KNN) กับข้อมูลจากเว็บไซต์ MalwareDomainlist (MDL) [5]

3. วิธีการดำเนินงาน

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการวิเคราะห์พฤติกรรมกรมการโจมตีทางไซเบอร์โดยใช้ปัญญาประดิษฐ์ในกองทัพอากาศ โดยมีรายละเอียดดังนี้

3.1 เครื่องมือในการวิจัย

3.1.1 โปรแกรม RapidMiner Studio

3.1.2 อัลกอริทึมที่ใช้ในการจำแนกข้อมูลมี 4 อัลกอริทึม ได้แก่ อัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees

3.2 ขั้นตอนการดำเนินการวิจัย

3.2.1 ทำความเข้าใจเกี่ยวกับการบุกรุกทางไซเบอร์ประเภทต่าง ๆ และศึกษาอัลกอริทึมการจำแนกภัยคุกคามทางไซเบอร์จากชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ เพื่อหาอัลกอริทึมที่เหมาะสมในการจำแนกรูปแบบการโจมตีทางไซเบอร์ โดยศึกษาจากเอกสารและงานวิจัยที่เกี่ยวข้อง

3.2.2 ทำความเข้าใจเกี่ยวกับข้อมูล ผู้วิจัยได้ทำการศึกษาและรวบรวมชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ (RTAF Dataset) ซึ่งเป็นชุดข้อมูลที่มีการเก็บรวบรวมจากระบบตรวจจับการบุกรุก (IDS) ที่เป็นข้อมูลที่ใช้งานจริง มีแอตทริบิวต์ 9 แอตทริบิวต์ แสดงดังตารางที่ 1

ตารางที่ 1: แสดงชื่อและประเภทของแอตทริบิวต์

ลำดับ	ชื่อแอตทริบิวต์	ประเภท
1	Classify	Discrete
2	Datetime	Discrete
3	Des_ip	Discrete
4	Des_port	Discrete
5	Priority	Discrete
6	Protocol	Discrete
7	Source_ip	Discrete

ตารางที่ 1: แสดงชื่อและประเภทของแอททริบิวต์ (ต่อ)

ลำดับ	ชื่อแอททริบิวต์	ประเภท
8	Source_port	Discrete
9	Type	Discrete

อย่างไรก็ตามผู้วิจัยได้ใช้ข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศซึ่งมีจำนวนทั้งสิ้น 638,134 เรคคอร์ด ทั้งนี้ได้แบ่งตามประเภทของการโจมตีไว้แล้ว ดังตารางที่ 2

ตารางที่ 2: ประเภทภัยคุกคามทางไซเบอร์

ประเภทการโจมตี	จำนวน	อัตราส่วน
Normal	4,542	0.712%
Bitcoin	4,048	0.634%
botnet	277	0.043%
compromise	82	0.013%
data_leak	154	0.024%
dropper	26,066	4.085%
Injection	6,947	1.089%
phishing_Email	1,599	0.251%
phishing_web	1,282	0.201%
probe	576	0.090%
R2L	59,741	9.362%
trojan	331	0.052%
U2R	1,073	0.168%
warning	483,584	75.781%
Web_Malware	47,816	7.493%
worm	12	0.002%
XSS	4	0.001%
ALL	638,134	100.00%

3.2.3 การเตรียมข้อมูล ผู้วิจัยได้ทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วนในอัตราเปอร์เซ็นต์ 70:30 คือ ในส่วนอัตรา 70 เปอร์เซ็นต์แรกใช้สำหรับการฝึกสอนเพื่อให้อัลกอริทึมเกิดการเรียนรู้ ส่วนอัตรา 30 เปอร์เซ็นต์ที่สองใช้สำหรับการทดสอบอัลกอริทึม

3.2.4 สร้างแบบจำลองโดยใช้โปรแกรม RapidMiner Studio และเลือกใช้อัลกอริทึม 4 อัลกอริทึมคือ Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees

3.2.5 ประเมินและเปรียบเทียบประสิทธิภาพ ด้วยอัลกอริทึมการจำแนกข้อมูลพฤติกรรมกรมการโจมตีทางไซเบอร์ ด้วยอัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees กับชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ (RTAF Dataset) เพื่อเปรียบเทียบความแม่นยำของแต่ละอัลกอริทึม โดยใช้สมการวัดค่าความแม่นยำ (Accuracy) ค่าความเที่ยง (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-Measure) ได้ตามสมการที่ (1) - (4) ดังนี้

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F-Measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

โดยที่ TP = True Positive คือ ข้อมูลที่ทำนายถูกต้องเมื่อเทียบกับเฉลย

FP = False Positive คือ ข้อมูลที่ทำนายแล้วไม่ถูกต้องเมื่อเทียบกับเฉลย

TN = True Negative คือ ข้อมูลที่อยู่ในเฉลยแต่ไม่มีการทำนาย (ตรงข้ามกับ FN)

Precision คือ ค่าความเที่ยงเกิดจากการนำค่า TP มาเทียบกับ FP

Recall คือค่าความระลึกเกิดจากการนำค่า TP มาเทียบกับ FN

F-Measure คือ ค่าเฉลี่ยของ Precision และ Recall

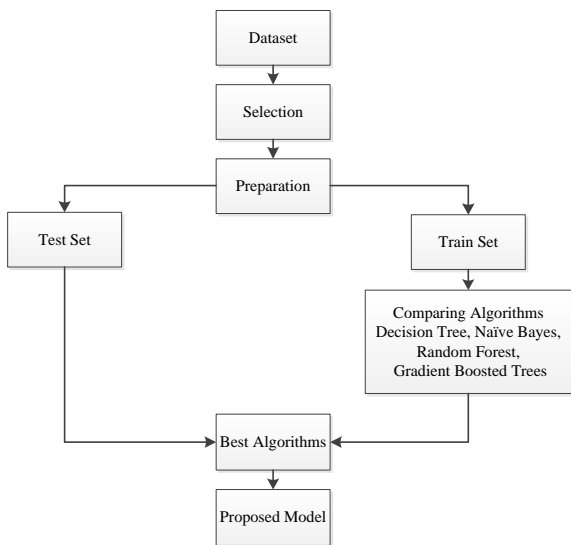
4. ผลการดำเนินงาน

การวิจัยครั้งนี้ ผู้วิจัยได้ใช้ชุดข้อมูลการโจมตีทางไซเบอร์ (RTAF Dataset) ซึ่งเป็นชุดข้อมูลเฉพาะ และเป็นข้อมูลที่เป็นข้อมูลที่ได้มาจากข้อมูลที่เป็นจริง นำเข้ากระบวนการวิเคราะห์พฤติกรรมกรมการโจมตีทางไซเบอร์ด้วยโปรแกรม RapidMiner Studio โดยใช้อัลกอริทึมการจำแนกข้อมูลการโจมตี 4 อัลกอริทึม

คือ อัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees

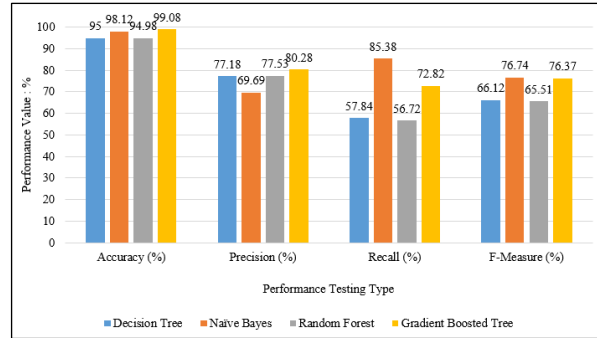
4.1 ผลการพัฒนาแบบจำลอง

ผู้วิจัยได้ดำเนินการพัฒนาแบบจำลองการวิเคราะห์พฤติกรรมจราจรทางไซเบอร์กับชุดข้อมูลการโจมตีทางไซเบอร์ของกองทัพอากาศ (RTAF Dataset) ด้วยอัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees แสดงผลลัพธ์ดังภาพที่ 1



ภาพที่ 1: แบบจำลองการวิเคราะห์พฤติกรรมจราจรทางไซเบอร์

จากภาพที่ 1 แสดงการสร้างแบบจำลองด้วยโปรแกรม RapidMiner Studio โดยเริ่มจากนำชุดข้อมูลเข้าสู่โปรแกรม จากนั้นทำการเลือกแอตทริบิวต์ที่ต้องการ จากนั้นทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วนตามที่ระบุไว้ในขั้นตอนการเตรียมข้อมูลก่อนหน้า จากนั้นทำการทดสอบกับ 4 อัลกอริทึม คืออัลกอริทึม Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees จากนั้นทำการตรวจสอบค่าเฉลี่ยความเที่ยง และค่าเฉลี่ยความแม่นยำ เพื่อทำการคัดเลือกอัลกอริทึมที่มีค่าเฉลี่ยความแม่นยำ และความเที่ยงที่สูงที่สุด เพื่อนำไปเป็นอัลกอริทึมหลักในการสร้างแบบจำลองการวิเคราะห์พฤติกรรมจราจรทางไซเบอร์โดยใช้ปัญญาประดิษฐ์ในกองทัพอากาศ โดยแสดงผลลัพธ์การเปรียบเทียบประสิทธิภาพ ดังภาพที่ 2



ภาพที่ 2: กราฟเปรียบเทียบผลการวิเคราะห์ค่า Accuracy, Precision, Recall และ F-Measure

จากภาพที่ 2 พบการแสดงผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยแสดงเป็นสีที่แตกต่างกัน และสามารถอธิบายผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยแสดงข้อมูลมีค่าเฉลี่ยเป็นเปอร์เซ็นต์ของ Accuracy, Precision, Recall และ F-Measure

จากกราฟสามารถสรุปได้ดังนี้ อัลกอริทึม Gradient Boosted Trees มีค่าเฉลี่ย Accuracy เท่ากับ 99.08% ค่าเฉลี่ย Precision เท่ากับ 80.28% ค่าเฉลี่ย Recall เท่ากับ 72.82% และค่าเฉลี่ย F-Measure เท่ากับ 76.37% อัลกอริทึม Naïve Bayes มีค่าเฉลี่ย Accuracy เท่ากับ 98.12% ค่าเฉลี่ย Precision เท่ากับ 69.69% ค่าเฉลี่ย Recall เท่ากับ 85.38% และค่าเฉลี่ย F-Measure เท่ากับ 76.74% อัลกอริทึม Decision Tree มีค่าเฉลี่ย Accuracy เท่ากับ 95% ค่าเฉลี่ย Precision เท่ากับ 77.18% ค่าเฉลี่ย Recall เท่ากับ 57.84% และค่าเฉลี่ย F-Measure เท่ากับ 66.12% และอัลกอริทึม Random Forest มีค่าเฉลี่ย Accuracy เท่ากับ 94.98% ค่าเฉลี่ย Precision เท่ากับ 77.53% ค่าเฉลี่ย Recall เท่ากับ 56.72% และค่าเฉลี่ย F-Measure เท่ากับ 65.51%

4.2 ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

ผู้วิจัยได้ทำการเปรียบเทียบค่าทางสถิติของอัลกอริทึมทั้ง 4 อัลกอริทึม คือ Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees เปรียบเทียบค่าทางสถิติทั้ง 4 คือ Accuracy, Precision, Recall และ F-Measure แสดงดังตารางที่ 2

ตารางที่ 2: การเปรียบเทียบค่าทางสถิติแบบจำลองทั้ง 4 อัลกอริทึม

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Decision Tree	95.00	77.18	57.84	66.12
Naïve Bayes	98.12	69.69	85.38	76.74
Random Forest	94.98	77.53	56.72	65.51
Gradient Boosted Trees	99.08	80.28	72.82	76.37

จากตารางที่ 2 ผลการเปรียบเทียบค่าทางสถิติการวิเคราะห์พฤติกรรมจราจรทางไฮเบอร์ร่วมกับชุดข้อมูลของกองทัพอากาศของแต่ละอัลกอริทึมพบว่าแบบจำลองที่ใช้ อัลกอริทึม Gradient Boosted Trees มีค่าเฉลี่ยความแม่นยำ (99.08%) และค่าเฉลี่ยความเที่ยง (80.28%) เป็นเปอร์เซ็นต์สูงสุด จากอัลกอริทึมทั้งหมด 4 อัลกอริทึม

5. สรุปผลการวิจัย

งานวิจัยนี้ผู้วิจัยได้ทำการวิเคราะห์พฤติกรรมจราจรทางไฮเบอร์ จากชุดข้อมูลการจราจรทางไฮเบอร์ของกองทัพอากาศ (RTAF Dataset) ด้วยอัลกอริทึม 4 อัลกอริทึม ประกอบด้วย Decision Tree, Naïve Bayes, Random Forest และ Gradient Boosted Trees กับโปรแกรม RapidMiner Studio

ผลการวิจัยพบว่าการทำงานของอัลกอริทึม Gradient Boosted Trees มีค่าเฉลี่ยความแม่นยำ และความเที่ยงสูงสุดสำหรับชุดข้อมูลการจราจรทางไฮเบอร์ของกองทัพอากาศ (RTAF Dataset) จากอัลกอริทึมที่กล่าวมาทั้งหมด

6. อภิปรายผลการวิจัย

จากผลสรุปการวิจัยสามารถอภิปรายผลได้ดังนี้ อัลกอริทึมที่เหมาะสมกับชุดข้อมูลนี้คือ อัลกอริทึม Gradient Boosted Trees โดยค่าเฉลี่ยความแม่นยำ และค่าเฉลี่ยความเที่ยงสูงสุด ซึ่งสอดคล้องกับงานวิจัยต่าง ๆ [3], [4] ดังนั้น ผู้วิจัยจึงได้นำอัลกอริทึม Gradient Boosted Trees ไปสร้างแบบจำลองการวิเคราะห์พฤติกรรมจราจรทางไฮเบอร์โดยใช้ปัญญาประดิษฐ์ในกองทัพอากาศ

นอกจากนี้ ผู้วิจัยยังพบว่าการแก้ไขแอททริบิวต์ของชุดข้อมูลจะส่งผลให้ค่าเฉลี่ยความแม่นยำ และค่าเฉลี่ยความ

เที่ยงลดลง ด้วยเหตุนี้หากมีการปรับเปลี่ยนอัลกอริทึม จะต้องมีการพิจารณาชุดข้อมูลด้วยว่าสามารถเข้ากันได้กับอัลกอริทึมดังกล่าวหรือไม่

เอกสารอ้างอิง

- [1] เอกสิทธิ์ พัชรวงศ์ศักดิ์, *Practical Data Science with RapidMiner Studio เล่ม 1*, บริษัท ไอดีซี พรีเมียร์ จำกัด, กรุงเทพฯ, 2565.
- [2] ธนินท์ธรร วาศพพิสิทธ์ และคณะ, “การแก้ไขปัญหาความไม่สมดุลของข้อมูลสำหรับการพยากรณ์การเสียชีวิตจากอุบัติเหตุ,” *วารสารการประชุมทางวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 18 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ*, หน้า 1138-1146, 2563.
- [3] สกฤตน์ ขุนสูงเนิน และคณะ, “การใช้การวิเคราะห์เชิงทำนายสำหรับการระบุสถานะการจำหน่ายและการรอดชีวิตในผู้ป่วยภาวะติดเชื้อบนพื้นฐานของปัจจัยการผ่าตัดและการทำหัตถการทางการแพทย์,” *วารสารวิจัย มหาวิทยาลัยขอนแก่น (ฉบับบัณฑิตศึกษา)* ปีที่ 19 ฉบับที่ 2, หน้า 117-130, 2562.
- [4] อภิชาติ บัดดาวะโร, “การวิเคราะห์การบุกรุกบนระบบเครือข่ายโดยใช้อัลกอริทึมการเรียนรู้ของเครื่อง,” *วิทยาศาสตร์มหบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ*, 2561.
- [5] วรวิมล ปัญญาวงษ์, “การเปรียบเทียบประสิทธิภาพการตรวจจับมัลแวร์โดยใช้อัลกอริทึมเหมือนข้อมูล,” *วิทยาศาสตร์มหบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม*, 2562.
- [6] A. Tousi et al., “Comparative Analysis of Machine Learning Models for Performance Prediction of the SPEC Benchmarks,” *IEEE Access*, vol. 10, pp. 11994-12011, 2022.
- [7] G. Karatas et al., “Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset,” *IEEE Access*, vol. 8, pp. 32150-32162, 2020.
- [8] L. Liu et al., “Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning,” *IEEE Access*, vol. 9, pp. 7550-7563, 2021.
- [9] S. Ameer et al., “Comparative Analysis of Machine Learning Techniques for Predicting Air Quality in Smart Cities,” *IEEE Access*, vol. 7, pp. 128325-128338, 2019.
- [10] S. V. Mahadevkar et al., “A Review on Machine Learning Styles in Computer Vision—Techniques and Future Directions,” *IEEE Access*, vol. 10, pp. 107293-107329, 2022.

การประยุกต์การค้นหาแบบค้างคาวเพื่อออกแบบตัวควบคุมพีไอดีสำหรับระบบประวิงเวลา

Application of Bat Search to Design PID Controller for Time Delay System

ไพฑูรย์ จันทร์เรือง (Paitoon Janreung)¹ และคณกุล คำปัญญา (Danupon Kumpanya)²

¹สาขาวิชาระบบสารสนเทศและคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ

มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

²สาขาวิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

¹Paitoon.j@rmutsb.ac.th, ²Danupon.k@rmutsb.ac.th

บทคัดย่อ

การค้นหาแบบค้างคาวเป็นหนึ่งในเทคนิคการหาค่าเหมาะที่สุดเชิงอภิศึกษาสำนึกที่ทรงประสิทธิภาพ บทความนี้นำเสนอการประยุกต์การค้นหาแบบค้างคาวเพื่อออกแบบตัวควบคุมพีไอดีสำหรับระบบประวิงเวลา การประวิงเวลาสามารถแทนด้วยการประมาณแบบพาเด ผลการออกแบบตัวควบคุมพีไอดีสำหรับระบบประวิงเวลาด้วยวิธีการค้นหาแบบค้างคาวจะถูกนำไปเปรียบเทียบกับวิธีของซีเกลอร์และนิโคลส์และการค้นหาแบบนกกาเหว่า จากผลการทดสอบพบว่าตัวควบคุมพีไอดีที่ออกแบบด้วยการค้นหาแบบค้างคาวสามารถควบคุมระบบประวิงเวลาได้อย่างน่าพอใจและให้ผลตอบสนองดีกว่าวิธีของซีเกลอร์และนิโคลส์และวิธีการค้นหาแบบนกกาเหว่า

คำสำคัญ: ตัวควบคุมพีไอดี ระบบประวิงเวลา การค้นหาแบบค้างคาวการค้นหาแบบนกกาเหว่าอภิศึกษาสำนึก

Abstract

The bat search (BS) is one of the most efficient meta-heuristic optimization search techniques. The application of the BS to design the PID controller for the time delay system is presented in this paper. The time delay occurred in the system can be substituted by the Pade's approximation. Results obtained by the BS will be compared with those obtained by the Ziegler and Nichols (ZN) method and cuckoo search (CS). As results, it was found that the PID controller designed by the BS can product the satisfactory response superior to ZN and CS methods.

Keywords: PID Controller, Time Delay System, Bat Search, Cuckoo Search, Meta-heuristic.

1. บทนำ

การควบคุมกระบวนการในงานอุตสาหกรรมส่วนใหญ่ เช่น อุตสาหกรรมการผลิตกระดาษ เครื่องปฏิกรณ์เคมี และเครื่องฟอกอากาศ สามารถนำเสนอโดยใช้แบบจำลองที่มีการประวิงเวลา (Time Delay) [1] และมากกว่า 90% ของระบบกายภาพ (Physical) ในการควบคุมกระบวนการแบบจำลองสามารถประมาณด้วยกระบวนการอันดับหนึ่งที่มีการประวิงเวลา (First Order Plus Time Delay หรือ FOPTD) อยู่ที่ประมาณ 60% และกระบวนการอันดับสองที่มีการประวิงเวลา (Second Order Plus Time Delay หรือ SOPTD) อยู่ที่ประมาณ 30% ซึ่งลักษณะโดยทั่วไปของกระบวนการเหล่านี้ก็คือกระบวนการที่มีการประวิงเวลาจะส่งผลกระทบต่อารออกแบบตัวควบคุม

ตัวควบคุมพีไอดี (Proportional Integral Derivative หรือ PID) ได้รับการยอมรับและถูกใช้งานอย่างกว้างขวางในงานอุตสาหกรรม เนื่องจากสร้างง่าย สมรรถนะคงทน ราคาถูกและโครงสร้างอย่างง่าย โดยทั่วไป ตัวควบคุม PID สามารถปรับแต่งได้ด้วยวิธีของซีเกลอร์และนิโคลส์ (Ziegler and Nichols หรือ ZN) [2] ซึ่งเป็นวิธีการปรับแต่งค่าพารามิเตอร์ของตัวควบคุม PID อย่างเหมาะสมสำหรับระบบเชิงเส้นและเสถียรเท่านั้น นอกจากนี้ยังเกิดการแกว่งและมีค่าเกินพิกัดสูงมาก ข้อเสียของวิธีการนี้คือ ถูกจำกัดการทำงานไว้เพียงบางส่วน และมีคุณสมบัติการออกแบบความคงทนต่อการเปลี่ยนแปลงที่ไม่น่าพอใจนัก

จากที่ผ่านมาการหาค่าเหมาะที่สุดเชิงอภิศึกษาสำนึก (Meta-heuristic Optimization) ได้รับความนิยมอย่างมาก และถูกประยุกต์ใช้งานอย่างแพร่หลาย [3] อัลกอริทึมของอภิศึกษาสำนึก (Meta-heuristic Algorithm) จำนวนมากได้รับการพัฒนาสำหรับแก้ปัญหาการค้นหาค่าเหมาะที่สุดเชิงตัวเลขและเชิงการจัด (Combinatorial) และได้ประสบความสำเร็จในการประยุกต์ใช้แก้ปัญหาการหาค่าเหมาะที่สุดที่เป็นปัญหาจริงทางด้านวิศวกรรม หนึ่งในเทคนิคการหาค่าเหมาะที่สุดเชิงอภิศึกษาสำนึกที่มีประสิทธิภาพ ก็คือการค้นหาแบบค้างคาว (Bat Search หรือ BS) ถูกนำเสนอครั้งแรกโดย Yang เมื่อปี ค.ศ. 2010 [4] โดยได้แนวคิดจากพฤติกรรมการปล่อยคลื่นเสียงความถี่สูงของค้างคาว (Sonar) เพื่อใช้ในการนำทางและค้นหาอาหารในที่มืด (Echolocation) บทความนี้นำเสนอการประยุกต์ BS เพื่อออกแบบตัวควบคุม PID สำหรับระบบประวิงเวลา โดยที่ การประวิงเวลาสามารถแทนด้วยการประมาณแบบพาด (Pade's Approximation) อันดับหนึ่ง ผลที่ได้รับจาก BS จะถูกนำไปเปรียบเทียบกับผลที่ได้รับจากวิธีของ ZN และการค้นหาแบบนกกาเหว่า (Cuckoo Search หรือ CS) [5]

2. การค้นหาแบบค้างคาว

การค้นหาแบบค้างคาว (BS) [4] ได้รับแรงบันดาลใจจากพฤติกรรมการปล่อยคลื่นเสียงที่มีความถี่สูงของค้างคาว เพื่อใช้ในการนำทางและค้นหาอาหารในที่มืด จมูกและปากของค้างคาวจะทำหน้าที่ช่วยบังคับคลื่นเสียงที่มีความถี่สูง เพื่อใช้ในการนำทางและค้นหาอาหาร ค้างคาวไม่ได้อาศัยดวงตาในการมองเห็นขณะบินในที่มืด หากแต่จะเคลื่อนที่ขณะบินเข้าหาแหล่งอาหารตามเสียงสะท้อนที่ถูกปล่อยออกมาด้วยคลื่นเสียงที่มีความถี่สูงมาก อย่างเป็นจังหวะและสม่ำเสมอ คลื่นเสียงที่มีความถี่สูง (คลื่นเหนือเสียง) ดังกล่าว หูของมนุษย์ไม่สามารถได้ยิน แต่หูของค้างคาวซึ่งมีความไวสูงมากจะสามารถตรวจจับคลื่นเหนือเสียงที่ส่งออกไปกระทบกับวัตถุใด ๆ แล้วสะท้อนกลับมายังหูของค้างคาวได้ จึงทำให้ค้างคาวสามารถรับรู้ถึงตำแหน่งและระยะห่างของสิ่งกีดขวางจากเสียงที่สะท้อนกลับมา ทำให้บินได้อย่างคล่องแคล่วว่องไว

แม้ยูในที่มืดโดยไม่พุ่งชนสิ่งกีดขวางพร้อมทั้งสามารถตรวจสอบตำแหน่งของสิ่งที่เป็นอาหาร (เหยื่อ) ได้อีกด้วยการค้นหาแบบค้างคาวอาศัยกลไกที่สำคัญดังต่อไปนี้

1) ค้างคาวสามารถค้นหาตำแหน่งของวัตถุโดยใช้เสียงสะท้อนกลับ เพื่อรับรู้ถึงระยะทางความแตกต่างระหว่างอาหาร (เหยื่อ) และสิ่งกีดขวาง กำหนดให้ปริภูมิการค้นหา (Search Space) เป็น d มิติ และค้างคาวตัวที่ i สามารถแทนด้วย x_i ความเร็วของค้างคาวตัวที่ i สามารถแทนด้วย v_i ความถี่พัลส์ของค้างคาวตัวที่ i สามารถแทนด้วย f_i ตำแหน่ง (ผลเฉลย) ใหม่ x'_i และความเร็วใหม่ v'_i ในช่วงเวลา t การปรับปรุงความถี่พัลส์อาศัยสมการที่ (1) ถึงสมการที่ (3) เมื่อ $\beta \in [0,1]$ คือ เวกเตอร์จากการสุ่ม (Random Vector) ซึ่งได้จากการแจกแจงเอกกรุป (Uniform Distribution) x_* คือ ผลเฉลยวงกว้าง (Global Solution) ในรอบปัจจุบัน เมื่อเทียบกับผลเฉลยทั้งหมดในจำนวนค้างคาว n ตัว และ t คือ รอบการค้นหาปัจจุบัน โดยที่ การค้นหาเฉพาะที่ (Local Search) ผลเฉลยใหม่จะถูกสร้างขึ้นจากผลเฉลยวงกว้างในรอบปัจจุบัน ดังสมการที่ (4) เมื่อ $\epsilon \in [-1,1]$ คือ จำนวนการสุ่ม (Random Number) โดยที่ $A' = (A'_i)$ คือ ความดังของเสียงเฉลี่ยของค้างคาวทั้งหมดในช่วงเวลานั้น

$$f_i = f_{\min} + \beta(f_{\max} - f_{\min}) \quad (1)$$

$$v'_i = v_i + f_i(x_i^{t-1} - x_*) \quad (2)$$

$$x'_i = x_i^{t-1} + v'_i \quad (3)$$

$$x_{\text{new}} = x_{\text{old}} + \epsilon A' \quad (4)$$

2) ค้างคาวมีทิศทางการบินอย่างไรแบบแผน (Randomly) ด้วยความเร็ว v_i ในตำแหน่ง x_i ด้วยความถี่คงที่ f_{\min} โดยเปลี่ยนแปลงความยาวคลื่น λ และความดังของเสียง A_0 เพื่อค้นหาอาหารและสามารถปรับเปลี่ยนความยาวคลื่นของพัลส์ที่ปล่อยออกมาและปรับอัตราการปล่อยพัลส์ (Pulse Emission หรือ r) โดยที่ $r \in [0,1]$ โดยอัตโนมัติ จะขึ้นอยู่กับความระยะห่างของเป้าหมาย การปล่อยพัลส์จะเริ่มจากศูนย์และเพิ่มขึ้นเรื่อยๆ เมื่อค้างคาวบินเข้าใกล้อาหาร (เหยื่อ) เพิ่มมากขึ้น ทั้งนี้ อัตราการปล่อยพัลส์ของค้างคาวจะถูกปรับเปลี่ยนไปตามสมการที่ (5)

$$r_i^{t+1} = r_i^0(1 - e^{-\gamma t}) \tag{5}$$

3) ความดังของเสียง (Loudness) สามารถเปลี่ยนแปลงได้หลายรูปแบบ ความดังของเสียงจะเปลี่ยนแปลงจากความดังของเสียงสูงสุด A_0 ไปยังค่าคงที่ต่ำที่สุด A_{min} นั่นคือ ความดังจะลดลงเรื่อยๆ เมื่อค้างคาเข้าใกล้อาหาร (เหยื่อ) เพิ่มมากขึ้น ความดังของเสียงจะถูกปรับเปลี่ยนไปตามสมการที่ (6)

$$A_i^{t+1} = \alpha A_i^t \tag{6}$$

จากกลไกที่สำคัญ 3 ข้อ การค้นหาแบบค้างคา (BS) สามารถแสดงเป็นขั้นตอนการทำงานดังต่อไปนี้

ขั้นตอนที่ 1 กำหนดจำนวนประชากรค้างคา i และรอบการค้นหาสูงสุด t_{max}

ขั้นตอนที่ 2 กำหนดประชากรค้างคาเริ่มต้น x_i พร้อมกับความเร็ v_i ความถี่พัลส์ f_i อัตราการส่งออกพัลส์ r_i และความดังของเสียง A_i ให้กับค้างคาแต่ละตัวโดยวิธีการสุ่มเพื่อใช้เป็นค่าเริ่มต้น

ขั้นตอนที่ 3 สร้างผลเฉลยใหม่โดยการปรับค่าความถี่พัลส์ f_i และปรับปรุงค่าความเร็ว v_i และตำแหน่ง x_i ของค้างคาแต่ละตัวด้วยสมการที่ (1) ถึง (3) และคำนวณหาค่าฟังก์ชันวัตถุประสงค์ (Objective Function หรือ F) ของค้างคาทุกตัว

ขั้นตอนที่ 4 กำหนดตัวเลขการสุ่ม ถ้า $rand < r_i$ ให้เลือกผลเฉลยที่ดีที่สุดตามฟังก์ชันวัตถุประสงค์ F และสร้างผลเฉลยวงแคบเฉพาะที่ (Local Solution) รอบๆ ผลเฉลยที่ดีที่สุดที่เลือกโดยใช้สมการที่ (4)

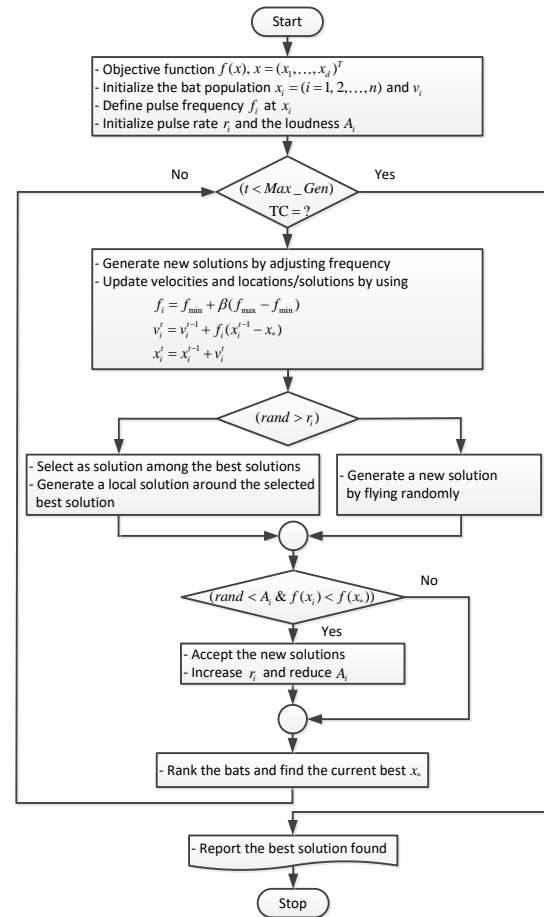
ขั้นตอนที่ 5 สร้างผลเฉลยใหม่ โดยการสุ่มค้างคาใหม่

ขั้นตอนที่ 6 กำหนดตัวเลขการสุ่ม ถ้า $rand < A_i$ และ $f(x_i) < f(x_*)$ ให้ยอมรับผลเฉลยใหม่และเพิ่มอัตราการส่งออกพัลส์ของค้างคาแต่ละตัว r_i และลดความดังเสียงของค้างคาแต่ละตัว A_i โดยใช้สมการที่ (5) และ (6)

ขั้นตอนที่ 7 เรียงลำดับค้างคาตามฟังก์ชันวัตถุประสงค์ F และหาผลเฉลยวงกว้างในรอบปัจจุบัน

ขั้นตอนที่ 8 ทำซ้ำขั้นตอนที่ 3 ถึง 7 จนครบรอบการทำงานซ้ำ

จากขั้นตอนวิธีของ BS สามารถแสดงได้ด้วยแผนภูมิดังภาพที่ 1 เมื่อ Max_Gen คือจำนวนรอบการทำงานสูงสุด และ TC คือ เกณฑ์สิ้นสุดการค้นหา (Termination Criteria)



ภาพที่ 1: แผนภูมิของ BS [4]

3. การค้นหาแบบนกกาเหว่า

การค้นหาแบบนกกาเหว่า (CS) [5] เป็นการเลียนแบบพฤติกรรมการวางไข่ของนกกาเหว่าในรังนกชนิดอื่น (Host Birds) ผสานกับพฤติกรรมการเคลื่อนที่ในลักษณะการสุ่มที่มีการแจกแจงแบบเลวี (Lévy) ซึ่งเป็นพฤติกรรมการบินของนกหรือแมลงบางชนิด โดยในตามธรรมชาติ นกกาเหว่านอกจากจะมีเสียงไพเราะแล้ว ยังมีพฤติกรรมชอบฝากเลี้ยงลูก ซึ่งนกกาเหว่าจะทำการวางไข่และทิ้งไข่ให้อยู่ในรังของนกชนิดอื่น หลังจากนั้นก็จะปล่อยให้แม่นกเจ้าของรัง ซึ่งเป็นนกสายพันธุ์อื่นฟักไข่และเลี้ยงลูกของตนเอง แต่แม่นกเจ้าของรังบางชนิดสามารถแยกแยะไข่ของตนเองในรังได้ว่าเป็นไข่ของตนเองหรือว่าเป็นไข่แปลกปลอม (ไข่ของนกกาเหว่า) ก็จะเขี่ยไข่นั้นทิ้งไปหรือไม่ก็ทิ้งรังแล้วไปสร้างรังใหม่ จากพฤติกรรมดังกล่าว ผวนกับการบินแบบสุ่มที่มีการแจกแจงแบบเลวี สามารถแสดงขั้นตอนวิธีของ CS บนหลักเกณฑ์พื้นฐานดังต่อไปนี้

หลักเกณฑ์ที่ 1 นกกาเหว่าจะวางไข่ครั้งละ 1 ฟองลงในรังของนกชนิดอื่นที่ถูกเลือกแบบสุ่ม

หลักเกณฑ์ที่ 2 รังที่ดีจะให้คุณภาพของไข่ (ผลเฉลย) ที่ดีและจะได้รับการเก็บรักษาเอาไว้

หลักเกณฑ์ที่ 3 จำนวนรังนกจะถูกกำหนดให้คงที่และความสามารถของแม่กกาเหว่าของรังที่จะแยกแยะและค้นพบไข่แปลกปลอม (ไข่ของนกกาเหว่า) ได้ด้วยค่าสัดส่วนความสามารถ $p_a \in [0, 1]$ กรณีที่ไข่ของนกกาเหว่าถูกพบแม่กกาเหว่าของรังอาจจะกำจัดไข่นั้นทิ้งไปหรืออาจจะทิ้งรังเพื่อไปสร้างรังใหม่ในตำแหน่งอื่นต่อไป

จากหลักเกณฑ์ดังกล่าวพบว่า ไข่แต่ละฟองในรังคือผลเฉลยของปัญหาที่น่าสนใจ ในขณะที่ไข่ของนกกาเหว่าคือผลเฉลยใหม่ ซึ่งผลเฉลยที่มีคุณภาพต่ำที่สุดจะถูกแทนที่ด้วยผลเฉลยใหม่ ขั้นตอนวิธีของ CS ในกรณีหาค่าน้อยที่สุด สามารถแสดงได้ดังแผนภูมิของ CS ดังภาพที่ 2

ผลเฉลยใหม่จะถูกสร้างขึ้นโดยอาศัยกระบวนการสุ่มที่มีการแจกแจงแบบเลวี ดังสมการที่ (7) เมื่อ $\alpha > 0$ คือขนาดขั้นการค้นหา (Step size) จากสมการที่ (7) พบว่าเป็นการเดินสุ่ม (Stochastic) ที่มีลักษณะเป็นไปตามห่วงโซ่มาร์คอฟ (Markov chain) นั่นคือ ตำแหน่งใหม่จะขึ้นอยู่กับตำแหน่งเดิมในพจน์แรกของสมการและค่าความน่าจะเป็นในการเปลี่ยนตำแหน่งในพจน์ที่สองสมการ สำหรับสัญลักษณ์ \oplus หมายถึง การคูณแบบแจกแจง (Entry-wise Multiplications) ในขณะที่สัญลักษณ์ Lévy(λ) หมายถึง ค่าสุ่มที่มีการแจกแจงแบบเลวี ซึ่งมีค่าเฉลี่ยและค่าความแปรปรวนเป็นอนันต์ดังสมการที่ (8)

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus \text{Lévy}(\lambda) \quad (7)$$

$$\text{Lévy} \approx u = t^{-\lambda}, \quad (1 < \lambda \leq 3) \quad (8)$$

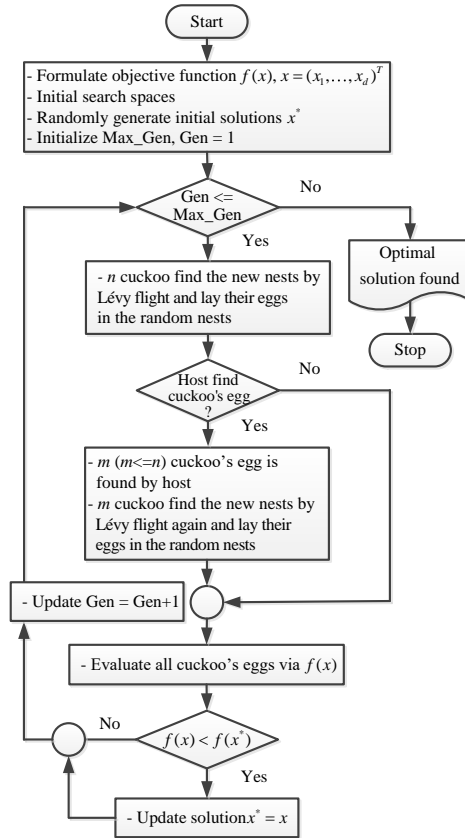
ขนาดขั้นการค้นหา s คำนวณได้จากความสัมพันธ์ในสมการที่ (9) เมื่อ u และ v คือค่าสุ่มที่มีการแจกแจงแบบปกติ (Normal Distribution) ดังสมการที่ (10) ซึ่งค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ของ u และ v แสดงดังสมการที่ (11) ค่า Γ คือฟังก์ชันแกมมามาตรฐานที่คำนวณได้จากความสัมพันธ์ดังสมการที่ (12)

$$s = \frac{u}{|v|^{1/\beta}} \quad (9)$$

$$u \approx N(0, \sigma_u^2), \quad v \approx N(0, \sigma_v^2) \quad (10)$$

$$\sigma_u = \left\{ \frac{\Gamma(1 + \beta) \sin(\pi\beta / 2)}{\Gamma[(1 + \beta) / 2] \beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \quad \sigma_v = 1 \quad (11)$$

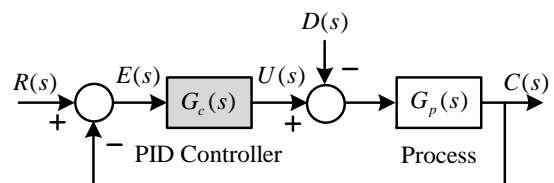
$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \quad (12)$$



ภาพที่ 2: แผนภูมิของ CS [5]

4. การออกแบบตัวควบคุม PID

โครงสร้างของระบบควบคุม PID แสดงได้ดังภาพที่ 3 พบว่า ตัวควบคุม $G_c(s)$ รับสัญญาณคลาดเคลื่อน $E(s)$ มาคำนวณเพื่อผลิตสัญญาณควบคุม $U(s)$ และส่งไปควบคุมกระบวนการ $G_p(s)$ ให้ผลิตผลตอบสนองหรือเอาต์พุต $C(s)$ สอดคล้องกับสัญญาณอินพุตอ้างอิง $R(s)$ พร้อมทั้งกำจัดสัญญาณรบกวน $D(s)$ พร้อมกัน



ภาพที่ 3: โครงสร้างของระบบควบคุม PID

จากภาพที่ 3 ตัวควบคุม PID แสดงดังสมการที่ (13) และฟังก์ชันถ่ายโอนวงปิดของระบบควบคุม PID แสดงดังสมการที่ (14) โดยที่ K_p เป็นอัตราขยายเชิงปรับสัดส่วน (Proportional Gain) K_i เป็นอัตราขยายเชิงปริพันธ์ (Integral Gain) และ K_d เป็นอัตราขยายเชิงอนุพันธ์ (Derivative Gain) ตามลำดับ

$$G_c(s)|_{PID} = K_p + \frac{K_i}{s} + K_d s \quad (13)$$

$$\frac{C(s)}{R(s)} = \frac{\left(K_p + \frac{K_i}{s} + K_d s\right) G_p(s)}{1 + \left(K_p + \frac{K_i}{s} + K_d s\right) G_p(s)} \quad (14)$$

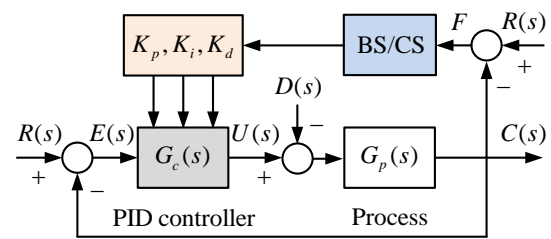
งานวิจัยนี้ได้ใช้แบบจำลองมาตรฐาน โดยที่ K_m คืออัตราขยายของกระบวนการ (Process Gain) T_m คือค่าคงตัวทางเวลา (Time Constant) และ τ_d คือการประวิงเวลา (Time Delay) และเทอมการประวิงเวลาอาศัยการประมาณแบบพาดอันดับหนึ่งแสดงได้ดังสมการที่ (15)

$$G_p(s) = \frac{K_m}{(T_m s + 1)} e^{-\tau_d s}; e^{-\tau_d s} \cong \frac{1 - \tau_d s / 2}{1 + \tau_d s / 2} \quad (15)$$

การจำลองนี้กำหนดให้อัตราขยายของกระบวนการ $K_m = 1.0$ การประวิงเวลา $\tau_d = 1.0$ วินาที และ 10 วินาที โดยให้ค่าคงตัวทางเวลา $T_m = 1.0$ วินาที สัญญาณรบกวน $d(t)$ ที่เข้ามาในระบบควบคุม กำหนดให้เข้าที่ 50 วินาที เป็นสัญญาณขั้นบันได (Unit Step Signal) $d(t) = 1.0$

4. การออกแบบตัวควบคุม PID โดยใช้ BS

การออกแบบตัวควบคุม PID อย่างเหมาะสมสำหรับระบบประวิงเวลาโดยใช้ BS สามารถแสดงได้ภาพที่ 4 โดยกำหนดให้ผลรวมของค่าความคลาดเคลื่อนกำลังสอง (Sum Squared Error หรือ SSE) ระหว่างอินพุต $R(s)$ และ เอาต์พุต $C(s)$ ของระบบเป็นฟังก์ชันวัตถุประสงค์ F แสดงได้ดังสมการที่ (16) โดยที่ค่า F จะถูกป้อนให้กับ BS เพื่อทำให้มีค่าน้อยที่สุดในการหาค่าพารามิเตอร์ K_p , K_i และ K_d ของตัวควบคุม PID อย่างเหมาะสมสำหรับระบบประวิงเวลา โดยที่ M_p คือค่าพุ่งเกินสูงสุด (Maximum Overshoot) และ e_{ss} คือค่าความคลาดเคลื่อนที่สถานะอยู่ตัว (Steady-state Error)



ภาพที่ 4: โครงสร้างของระบบควบคุม PID โดยใช้ BS

$$\begin{aligned} \text{Minimize } F &= \sum_{i=1}^N [r(i) - c(i)]^2 \\ \text{subject to } M_p &\leq 10\%, e_{ss} \leq 0.1\%, \\ 0 \leq K_p &\leq 0.5, 0 \leq K_i \leq 0.5, 0 \leq K_d \leq 0.01 \end{aligned} \quad (16)$$

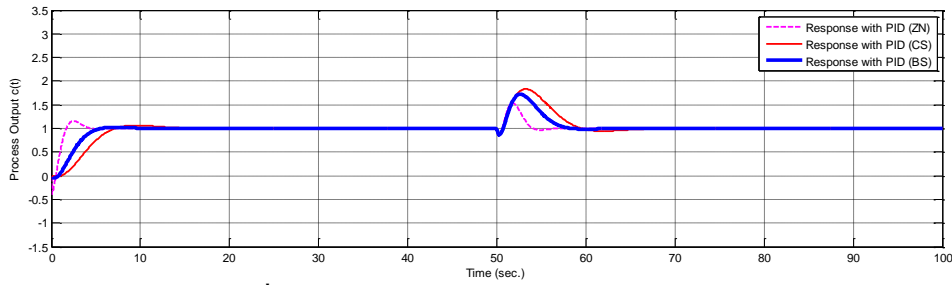
5. ผลการทดสอบและการอภิปรายผล

ขั้นตอนวิธีของ BS/CS ได้รับการพัฒนาด้วยชุดคำสั่งของโปรแกรม MATLAB โดย Yang [4], [5] ถูกดัดแปลงเพื่อออกแบบตัวควบคุม PID สำหรับระบบประวิงเวลาเบื้องต้นปรับตั้งค่าพารามิเตอร์ของวิธี BS และวิธี CS ดังนี้

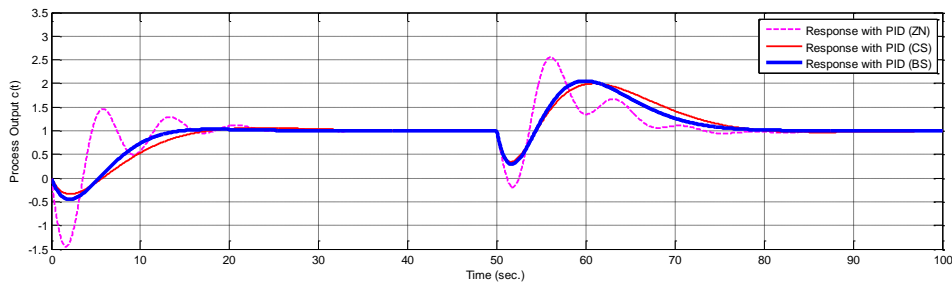
วิธี BS จำนวนของคี่กางคว $n = 20$ ถึง 50 ความถี่ $f_{\min} = 0.0$ และ $f_{\max} = 2.0$ ความดังของเสียง $A_0 = 0.5$ อัตราการส่งออกพัลส์ $r_0 = 0.5$ และ $\alpha = \gamma = 0.9$

วิธี CS จำนวนของรังหรือจำนวนนกกาเหว่า $n = 20$ ถึง 50 และสัดส่วนความสามารถของแม่เหล็กของรังค้นพบไขแปลกล้อม $p_a = 0.25$, $\alpha = 1.0$ และ $\beta = 1.5$

การออกแบบตัวควบคุม PID ด้วยวิธี BS และวิธี CS กำหนดให้จำนวนรอบสูงสุด (Max_Gen) = 1000 เป็น TC ภายในขอบเขตการค้นหาที่กำหนดไว้ในสมการที่ (16) ทำการค้นหาทั้งหมด 40 ตัวอย่าง เพื่อค้นหาพารามิเตอร์ K_p , K_i และ K_d ของตัวควบคุม PID อย่างเหมาะสม โดยผลที่ได้รับจากวิธี BS จะถูกเปรียบเทียบกับผลที่ได้รับจากวิธีของ ZN และวิธี CS ที่การประวิงเวลา 1.0 วินาที และ 10 วินาที หลังจากสิ้นสุดการค้นหาได้ค่าพารามิเตอร์ของตัวควบคุม PID อย่างเหมาะสม ที่การประวิงเวลา $\tau_d = 1.0$ วินาที แสดงดังตารางที่ 1 และผลตอบสนองของระบบควบคุมแสดงดังภาพที่ 5 ในขณะที่การประวิงเวลา $\tau_d = 10$ วินาที แสดงดังตารางที่ 2 และผลตอบสนองของระบบควบคุมแสดงดังภาพที่ 6 โดยที่ t_r คือช่วงเวลายื่น (Rise Time) t_s คือช่วงเวลาเข้าที่ (Settling Time) และ t_d คือช่วงเวลากำจัดสัญญาณรบกวน (Disturbance Rejection)



ภาพที่ 5: ผลตอบสนองของระบบควบคุม เมื่อ $\tau_d = 1.0$ วินาที



ภาพที่ 6: ผลตอบสนองของระบบควบคุม เมื่อ $\tau_d = 10$ วินาที

ตารางที่ 1: สมรรถนะของระบบควบคุม เมื่อ $\tau_d = 1.0$ วินาที

Methods	Parameters			System Responses				
	K_p	K_i	K_d	t_r (sec.)	M_p (%)	t_s (sec.)	e_{ss} (%)	t_d (sec.)
ZN	1.2284	1.1100	0.2775	1.6750	15.88	4.1040	0.00	5.6420
CS	0.0013	0.2755	0.0056	7.1825	6.63	13.4069	0.00	15.4520
BS	0.3215	0.3947	0.0026	5.6270	2.28	7.9620	0.00	7.8002

ตารางที่ 2: สมรรถนะของระบบควบคุม เมื่อ $\tau_d = 10$ วินาที

Methods	Parameters			System Responses				
	K_p	K_i	K_d	t_r (sec.)	M_p (%)	t_s (sec.)	e_{ss} (%)	t_d (sec.)
ZN	0.7200	0.1515	0.8550	4.5667	46.56	29.5169	0.00	34.7950
CS	0.3842	0.0931	0.0001	17.6425	5.95	33.6850	0.00	42.1920
BS	0.4986	0.1008	0.0014	14.6110	3.18	22.2189	0.00	29.4250

จากตารางที่ 1 และภาพที่ 5 กับตารางที่ 2 และภาพที่ 6 เมื่อ $\tau_d = 1.0$ วินาที และ 10 วินาที พบว่า วิธี BS จะให้ผลตอบสนองของระบบควบคุมสำหรับระบบประวิงเวลา มีช่วงเวลานั้น t_r มีช่วงเวลาที่เข้าที่ t_s มีค่าพุงเกิน M_p และมีช่วงเวลากำจัดสัญญาณรบกวน t_d น้อยกว่าวิธีของ CS โดยไม่มีค่าความคลาดเคลื่อนที่สถานะอยู่ตัว e_{ss} ในขณะที่ วิธีของ ZN มีการแกว่ง และมีค่าพุงเกินสูงสุด M_p สูงมาก

6. สรุป

การประยุกต์ BS เพื่อออกแบบตัวควบคุม PID สำหรับระบบประวิงเวลาที่นำเสนอในบทความนี้ วิธี BS จะถูกนำไป

เปรียบเทียบกับวิธีของ ZN และวิธี CS จากผลการทดสอบพบว่า ตัวควบคุม PID ที่ถูกออกแบบด้วย BS สามารถควบคุมระบบประวิงเวลาได้อย่างน่าพอใจ และให้ผลตอบสนองดีกว่าวิธีของ ZN และวิธี CS ในขณะที่ วิธีของ ZN จะพบว่า มีการแกว่งและมีค่าพุงเกินสูงสุดที่สูงมากจนไม่สามารถอนุรักษ์ได้จริงในทางปฏิบัติ

เอกสารอ้างอิง

- [1] Kuo, B. C., and F. Golnaraghi, Automatic Control Systems, 8th ed., John Wiley & Sons, 2003.
- [2] Ziegler, J. G., and N. B. Nichols, Optimum Settings for Automatic Controllers, In Trans. of the ASME, Vol. 64, 1942, p. 759–768.
- [3] Yang, X. S., Engineering Optimization, An Introduction with Metaheuristic Applications, John Wiley & Sons, 2010.
- [4] Yang, X. S., A new metaheuristic bat-inspired an algorithm, in: Nature Inspired Cooperative Strategies for Optimization (NICSO 2010) (Eds. J. R. Gonzalez et al.), Studies in Computational Intelligence, Springer Berlin, 284, Springer, 2010, p. 65–74.
- [5] Yang, X. S., and S. Deb, Cuckoo search via Lévy flights, 2009 World Congress on Nature &

Biologically Inspired Computing (NaBIC 2009),
2009, p. 210–214.

เครื่องมือทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่แบบอัตโนมัติสำหรับส่วนต่อประสาน ประยุกต์ชนิดกราฟคิวแอล

Automated Non-Functional Properties Testing Tool for GraphQL API

พัทธดนย์ โปร่งปรีชา (Pattadon Prongpreecha) และเนืองวงศ์ ทวยเจริญ (Nuengwong Tuaycharoen)

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

6470237521@student.chula.ac.th, nuengwong.t@chula.ac.th

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอเครื่องมือทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่แบบอัตโนมัติสำหรับ GraphQL API ที่สามารถครอบคลุมการทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ 3 รายการ ได้แก่ ประสิทธิภาพ ความพร้อมใช้งาน และความปลอดภัย เพื่ออำนวยความสะดวกและลดระยะเวลาของกระบวนการทดสอบ โดยเมื่อนำไปทดสอบกับ API ที่เปิดให้ใช้งานแบบสาธารณะจำนวน 4 API และ API ที่ผู้วิจัยพัฒนาขึ้นเองจำนวน 1 API พบว่า ได้ผลลัพธ์เป็นค่าเฉลี่ยความเร็วการตอบกลับ API uptime และช่องโหว่ทางด้านความปลอดภัย ซึ่งเป็นตัวชี้วัดที่แสดงให้เห็นคุณสมบัติเชิงไม่ใช้หน้าที่ของ GraphQL API ในแต่ละรายการตามที่ได้คาดหวัง

คำสำคัญ: GraphQL API คุณสมบัติเชิงไม่ใช้หน้าที่ ประสิทธิภาพ ความพร้อมใช้งาน ความปลอดภัย

Abstract

This paper presents an automated tool for testing non-functional properties of GraphQL API to facilitate testing process and reduce testing time. The tool is capable of testing 3 non-functional properties: performance, availability, and security. The results from testing with 4 public APIs and 1 API developed by researchers show an average response time, API uptime, and security vulnerabilities which represent non-functional properties of GraphQL API as expected.

Keywords: GraphQL, API, Non-Functional Properties, Performance, Availability, Security.

1. บทนำ

GraphQL [1] คือเทคโนโลยีใหม่ที่ถูกนำเสนอโดยบริษัท Facebook ในปี ค.ศ. 2015 เพื่อเป็นอีกทางเลือกหนึ่งของการพัฒนา API (Application Programming Interface) [2] นอกเหนือจาก REST API [3] ที่กำลังได้รับความนิยม ในปัจจุบันมีบริษัทขนาดใหญ่หลายแห่งได้เลือกใช้ GraphQL เป็นทางเลือกสำหรับพัฒนา API ใหม่ ๆ เช่น Facebook, Shopify และ GitLab เป็นต้น ประโยชน์ของ GraphQL คือ เครื่องลูกข่าย (Client) สามารถร้องขอเฉพาะข้อมูลที่ต้องการจากเครื่องแม่ข่าย (Server) ได้ รวมถึงสามารถร้องขอข้อมูลจากหลายทรัพยากร (Resource) ได้จากการส่งคำร้อง (Request) เพียงครั้งเดียว ดังนั้น GraphQL จึงเป็นทางเลือกที่เหมาะสมในการพัฒนา API ที่ถูกเรียกใช้โดยอุปกรณ์ขนาดเล็กที่สามารถเชื่อมต่อกับเครือข่ายอินเทอร์เน็ตได้ เช่น โทรศัพท์มือถือ และอุปกรณ์ IoT (Internet of Thing) เป็นต้น ภาพที่ 1 คือตัวอย่างคำร้องของ GraphQL API ในขณะที่ภาพที่ 2 คือตัวอย่างการตอบกลับ (Response) ของ GraphQL API

```
query {  
  countries {  
    code,  
    name,  
    capital,  
    currency  
  }  
}
```

ภาพที่ 1: ตัวอย่างคำร้องของ GraphQL API

```

{
  "data": {
    "countries": [
      {
        "code": "AD",
        "name": "Andorra",
        "capital": "Andorra la Vella",
        "currency": "EUR"
      },
      ...
      {
        "code": "ZW",
        "name": "Zimbabwe",
        "capital": "Harare",
        "currency": "USD, ZAR, BWP, GBP, AUD, CNY, INR, JPY"
      }
    ]
  }
}

```

ภาพที่ 2: ตัวอย่างการตอบกลับของ GraphQL API

เมื่อก้าวถึงการพัฒนาซอฟต์แวร์ใด ๆ ก็ตาม การทดสอบนับว่าเป็นกระบวนการที่จำเป็นที่ต้องทำ ทั้งการทดสอบทางด้านฟังก์ชันการทำงานและคุณสมบัติเชิงไม่ใช้หน้าที่ [4] ในรายการต่างๆ เพื่อเป็นการประกันคุณภาพของซอฟต์แวร์ มีผู้วิจัยบางกลุ่มได้ให้ความสนใจในการทดสอบ GraphQL API เช่น Stefan Karlsson และคณะ [5] ที่ได้มีการนำเสนอเครื่องมืออัตโนมัติสำหรับทดสอบฟังก์ชันการทำงานของ GraphQL API แต่อย่างไรก็ตามเครื่องมือดังกล่าวไม่สามารถทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ของ GraphQL API ได้ เช่น ประสิทธิภาพ ความพร้อมใช้งาน และความปลอดภัย เป็นต้น

ผู้วิจัยได้สังเกตเห็นช่องว่างนี้จึงมีความคิดที่จะนำเสนอเครื่องมือที่สามารถทดสอบคุณสมบัติไม่ใช้หน้าที่ของ GraphQL API แบบอัตโนมัติได้ โดยจะครอบคลุมคุณสมบัติเชิงไม่ใช้หน้าที่ 3 รายการ ได้แก่ ประสิทธิภาพ ความพร้อมใช้งาน และความปลอดภัย เพื่อเป็นการอำนวยความสะดวกและลดระยะเวลาของการทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ของ GraphQL API

2. งานวิจัยที่เกี่ยวข้อง

2.1 An OpenAPI-based Testing Framework to Monitor Non-Functional Properties of REST APIs

Steven Bucaille และคณะ [6] ได้นำเสนอเครื่องมือที่ชื่อว่า Gadolinium ซึ่งสามารถทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ของ REST API ได้แบบอัตโนมัติจากการนำเข้าข้อมูลไฟล์ OpenAPI Specification โดยครอบคลุมการทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ 2 รายการคือ ประสิทธิภาพ และความพร้อมใช้งาน สำหรับการทดสอบประสิทธิภาพจะวัดผลจาก Latency Time ที่คำนวณจากระยะห่างของเวลานับตั้งแต่เครื่องถูกข่าส่งคำสั่งไปยังเครื่องแม่ข่ายจนกระทั่งมีการตอบกลับ ในขณะที่การทดสอบความพร้อมใช้งาน (Availability) จะวัดผลจาก API uptime (UT) และ Downtime (DT) ซึ่งสามารถคำนวณได้จากสมการที่ 1 ดังนี้

$$Availability(\%) = \frac{UT}{UT + DT} * 100 \quad (1)$$

2.2 Automatic Property-based Testing of GraphQL APIs

Stefan Karlsson และคณะ [5] ได้นำเสนอวิธีในการสร้าง Query แบบอัตโนมัติจาก GraphQL Schema เพื่อทดสอบฟังก์ชันการทำงานของ GraphQL API และตรวจจับข้อผิดพลาด (Fault) ที่เกิดขึ้น โดยได้ไปทำการทดสอบกับ API 2 ประเภทคือ API ที่มีการใช้ในอุตสาหกรรมจริงอย่าง GitLab API และ API ที่ผู้วิจัยได้พัฒนาขึ้นเอง ซึ่งได้มีการใส่ข้อผิดพลาดไว้ทั้งหมด 15 จุด จากผลการทดสอบพบว่าวิธีการที่ได้นำเสนอนั้นสามารถครอบคลุม Schema Coverage ของ GitLab API ได้ทั้งหมด 100% และสามารถตรวจจับข้อผิดพลาดของ API ที่ผู้วิจัยพัฒนาขึ้นเองได้ 11 จุดจากทั้งหมด 15 จุดหรือคิดเป็น 73%

2.3 Performance Measurement of GraphQL API

in Home ESS Data Server

Euggi Lee และคณะ [7] ได้นำเสนอขั้นตอนการวัดประสิทธิภาพ GraphQL API จากระบบ Home Energy Management System ที่ผู้วิจัยและคณะได้พัฒนาขึ้นเอง โดยจะวัดจากความเร็วในการตอบกลับ (Response Time) และขนาดของข้อมูลในการตอบกลับ (Response Size) พร้อมทั้งเปรียบเทียบผลลัพธ์ที่ได้กับ REST API ของระบบเดียวกันที่ถูกพัฒนามาก่อนหน้านี้ เพื่อเป็นการเปรียบเทียบประสิทธิภาพของ REST API และ GraphQL API

3. แนวคิดและวิธีการดำเนินงาน

3.1 การกำหนดตัวแปรสำหรับการทดสอบ

ก่อนเริ่มกระบวนการทดสอบผู้ทดสอบจำเป็นจะต้องกำหนดตัวแปรที่ต้องใช้ในการทดสอบ ได้แก่

- Endpoint ของ API ที่ต้องการทดสอบ
- ความถี่สำหรับการทดสอบประสิทธิภาพและความพร้อมใช้งานในหน่วยวินาที
- จำนวนครั้งสำหรับการทดสอบประสิทธิภาพและความพร้อมใช้งานในหน่วยครั้ง
- ระยะเวลาสูงสุดที่จะรอการตอบกลับสำหรับการทดสอบความพร้อมใช้งานในหน่วยวินาที
- จำนวน Query ที่ซ้ำกันสำหรับการทดสอบความปลอดภัยในประเด็น Batching Attack
- ระดับชั้นของ Object ที่ซ้อนกัน ใน Query สำหรับการทดสอบความปลอดภัยในประเด็น Denial of Service Queries

3.2 การเข้าถึงข้อมูล GraphQL Schema

เพื่อที่จะทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ของ GraphQL API แบบอัตโนมัติ จำเป็นจะต้องมีการส่งคำร้องไปยัง Endpoint ของ API ที่ต้องการทดสอบ ซึ่งจะประกอบไปด้วยการดำเนินการ (Operation) ที่ API นั้นสามารถรองรับได้ โดยที่เครื่องมือของงานวิจัยนี้จะทำการสร้างคำร้องตามการดำเนินการทั้งหมดที่ได้ระบุไว้ใน GraphQL Schema [8] ซึ่งคือเอกสารที่ใช้อธิบายการดำเนินการทั้งหมดที่ API สามารถรองรับได้

ดังตัวอย่างที่แสดงในภาพที่ 3 ด้วยการส่ง Introspection Queries ไปยัง Endpoint ของ API ที่ต้องการทดสอบ ซึ่งถือว่าเป็นคุณสมบัติหนึ่งของ GraphQL API ที่อนุญาตให้เข้าถึงข้อมูลของ GraphQL Schema ได้

```

type BlogPost {
  id: ID!
  title: String!
  content: String!
  hasAuthor: Author!
}

type Author {
  id: ID!
  name: String!
  hasBlog: [BlogPost]
}

```

ภาพที่ 3: ตัวอย่างของ GraphQL Schema

3.3 การทดสอบประสิทธิภาพ

สำหรับการทดสอบประสิทธิภาพจะใช้ตัวชี้วัดเป็นความเร็วในการตอบกลับของแต่ละการดำเนินการทั้งหมดที่ API นั้นสามารถรองรับได้ตามที่ได้มีการระบุไว้ใน GraphQL Schema ซึ่งคล้ายกันกับ [5] จากนั้นบันทึกข้อมูลความเร็วในการตอบกลับในหน่วยมิลลิวินาทีเพื่อนำไปคำนวณหาค่าเฉลี่ยความเร็วการตอบของแต่ละการดำเนินการต่อไป

3.4 การทดสอบความพร้อมใช้งาน

สำหรับการทดสอบความพร้อมใช้งานจะใช้ API uptime ตามสมการในหัวข้อที่ 2.1 ทำให้ต้องมีการบันทึกข้อมูลระยะเวลาที่ระบบทำงานได้และระยะเวลาที่ระบบไม่สามารถทำงานได้ โดยที่การตรวจสอบว่าระบบนั้นสามารถทำงานได้หรือไม่จะตรวจสอบด้วยการใช้วิธีการส่งคำร้องไปยัง Endpoint ของ API ที่ต้องการทดสอบ โดยจะสุ่มเลือกจากการดำเนินการทั้งหมดที่ API นั้นรองรับ หากพบว่าระบบสามารถตอบกลับได้โดยที่ไม่มีข้อผิดพลาดจะสรุปได้ว่า ณ ช่วงเวลาดังกล่าวระบบสามารถทำงานได้ แต่ถ้าหากระบบไม่ตอบกลับภายในระยะเวลาที่ได้กำหนดไว้ในตัวแปรของการทดสอบ หรือตอบกลับพร้อมกับข้อผิดพลาดจะสรุปได้ว่า ณ ช่วงเวลาดังกล่าวระบบไม่สามารถทำงานได้

3.5 การทดสอบความปลอดภัย

เพื่อเป็นการทดสอบความปลอดภัยของ GraphQL API เครื่องมือที่พัฒนาขึ้นครอบคลุมการทดสอบด้านความปลอดภัย 3 ประเด็น โดยได้อ้างอิงจากคำแนะนำของเว็บไซต์ OWASP [9] (Open Web Application Security Project) ที่กล่าวถึงประเด็นของช่องโหว่ทางด้านความปลอดภัยต่างๆ ของ GraphQL API ที่ควรจะถูกทดสอบ ได้แก่ Injection, Batching of Attack และ Denial of Service Queries

ในการทดสอบประเด็นของ Injection ที่ประกอบด้วย SQL Injection และ Cross-Site Scripting จะใช้ชุดข้อมูลทดสอบที่มาจาก GitHub ของ SecList [10] เพื่อเป็นพารามิเตอร์สำหรับสร้างคำร้องในแต่ละการดำเนินการของ API ในขณะที่การทดสอบในประเด็น Batching of Attack และ Denial of Service Queries จะใช้วิธีการส่งคำร้องแต่ละรูปแบบจากตัวแปรที่ได้มีการกำหนดไว้ตั้งแต่ก่อนการทดสอบ หากพบว่าหลังจากที่ได้ส่งคำร้องไปยัง API แล้วระบบตอบกลับพร้อมกับข้อผิดพลาดจะถือว่าผ่านการทดสอบ แต่ถ้าหากรบบตอบกลับโดยที่ไม่มีข้อผิดพลาดจะถือว่าไม่ผ่านการทดสอบ

3.6 การรายงานผลการทดสอบ

เมื่อการทดสอบคุณสมบัติทั้ง 3 รายการเสร็จสิ้นแล้วผลลัพธ์ที่ได้จากการทดสอบจะถูกรายงานตามรูปแบบที่แตกต่างกันตามคุณสมบัติในแต่ละรายการ โดยผลลัพธ์ของการทดสอบประสิทธิภาพจะถูกรายงานในรูปแบบของแผนภูมิแท่ง ผลลัพธ์ของการทดสอบความพร้อมใช้งานจะถูกรายงานในรูปแบบของแผนภูมิวงกลม ในขณะที่ผลลัพธ์ของการทดสอบความปลอดภัยจะถูกรายงานในรูปแบบของตาราง

3.7 การพัฒนาเครื่องมือทดสอบ

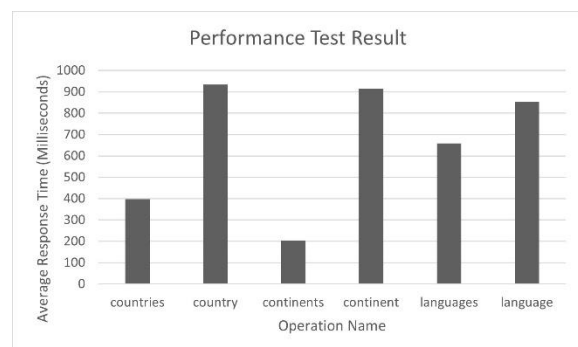
การพัฒนาเครื่องมือทดสอบจะถูกแบ่งออกเป็น 2 ส่วน คือ ส่วนที่เป็น Backend ที่พัฒนาด้วยภาษาโปรแกรม Node.js สำหรับคำร้องไปยัง GraphQL API เพื่อทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่และบันทึกผลการทดสอบลงในฐานข้อมูล MongoDB อีกส่วนหนึ่งคือ Frontend สำหรับ

แสดงผลการทดสอบผ่านทาง User Interface ซึ่งพัฒนาด้วยเฟรมเวิร์ค AngularJS

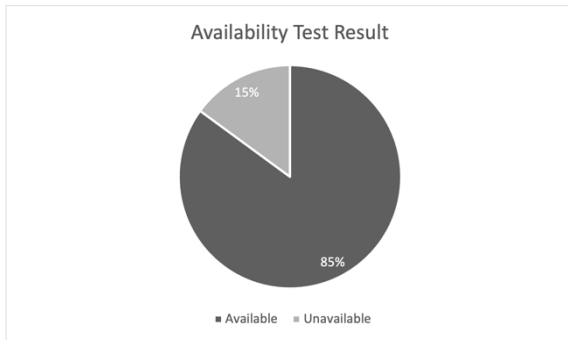
4. ผลการดำเนินการวิจัย

GraphQL API ที่ถูกทดสอบโดยเครื่องมือของงานวิจัยนี้ประกอบด้วย API 2 ประเภทคือ API ที่ผู้วิจัยพัฒนาขึ้นเองจำนวน 1 API เพื่อเป็นการตรวจสอบความถูกต้องของการทดสอบความปลอดภัย และ API ที่มีการเปิดให้ใช้งานแบบสาธารณะจำนวน 4 API [11] ผลลัพธ์จากการทดสอบพบว่าสามารถหาค่าเฉลี่ยความเร็วการตอบกลับของการดำเนินการทั้งหมดที่ปรากฏใน GraphQL Schema ได้ ดังตัวอย่างในภาพที่ 4 และ 5 ซึ่งเป็นผลลัพธ์ที่ได้จากการทดสอบประสิทธิภาพและความพร้อมใช้งานของ API สาธารณะสำหรับค้นคืนข้อมูลของประเทศต่างๆ

ภาพที่ 4 คือผลลัพธ์ที่ได้จากการทดสอบประสิทธิภาพที่แสดงในรูปแบบของแผนภูมิแท่ง โดยที่แกน y แสดงค่าเฉลี่ยความเร็วการตอบกลับในหน่วยมิลลิวินาที ส่วนแกน x แสดงชื่อของแต่ละการดำเนินการ ในขณะที่ภาพที่ 5 คือผลลัพธ์ที่ได้จากการทดสอบความพร้อมใช้งานที่แสดงในรูปแบบของแผนภูมิวงกลม โดยพื้นที่วงกลมจะแบ่งออกเป็น 2 ส่วนคือ 85% หมายถึง API uptime หรืออัตราส่วนที่ API สามารถเรียกใช้งานได้ และอีก 15% คืออัตราส่วนที่ API ไม่สามารถเรียกใช้งานได้



ภาพที่ 4: ตัวอย่างผลลัพธ์จากการทดสอบประสิทธิภาพของ API สาธารณะ



ภาพที่ 5: ตัวอย่างผลลัพธ์จากการทดสอบความพร้อมใช้งานของ API สาธารณะ

ในขณะที่ผลลัพธ์ของการทดสอบความปลอดภัยจากการนำเครื่องมือไปทดสอบกับ API ที่ผู้วิจัยพัฒนาขึ้นเองพบว่าสามารถตรวจจับช่องโหว่ทางด้านความปลอดภัยของแต่ละการดำเนินการทั้งหมดที่ปรากฏใน GraphQL Schema ได้ทั้ง 3 ประเด็น ได้แก่ Injection ที่ประกอบด้วย SQL Injection และ Cross-Site Scripting, Batching of Attack และ Denial of Service Queries ดังตารางที่ 1 ที่แสดงให้เห็นว่าผลลัพธ์ของการทดสอบในแต่ละประเด็นจะมีการดำเนินการทั้งที่ผ่านและไม่ผ่านการทดสอบ

ตารางที่ 1: ผลลัพธ์จากการทดสอบความปลอดภัยในแต่ละประเด็นของ API ผู้วิจัย

Operation	Injection		Batching of Attack
	SQLi	XSS	
users	ไม่ผ่าน	ผ่าน	ผ่าน
createUser	ผ่าน	ไม่ผ่าน	ผ่าน
stores	ผ่าน	ผ่าน	ไม่ผ่าน
orders	ผ่าน	ผ่าน	ผ่าน
products	ผ่าน	ผ่าน	ผ่าน
createProduct	ผ่าน	ผ่าน	ผ่าน

มีเพียงการทดสอบในประเด็นของ Denial of Service Queries ที่ผลลัพธ์ของการดำเนินการจะเหมือนกันทั้งหมด เนื่องจาก API ที่ผู้วิจัยพัฒนาขึ้นเองนั้นได้มีการป้องกันช่องโหว่ในประเด็นของ Denial of Service Queries ด้วยการกำหนดระดับชั้นของ Query ของทุกการดำเนินการไว้สูงสุดที่ 5 ชั้น ไม่สามารถแจกแจงตามแต่ละการดำเนินการได้ ผู้วิจัยจึงได้ใช้วิธีกำหนดค่าตัวแปรสำหรับการทดสอบคือระดับชั้น Object ของ Query

ที่แตกต่างกันเพื่อให้ได้ผลการทดสอบทั้งแบบที่ผ่านและไม่ผ่าน ดังผลลัพธ์ที่ได้แสดงในตารางที่ 2

ตารางที่ 2: ผลลัพธ์จากการทดสอบในประเด็น Denial of Service

Queries ของ API ผู้วิจัย	
ระดับชั้นของ Object ใน Query ที่ใช้ทดสอบ	ผลการทดสอบของทุกการดำเนินการ
3	ไม่ผ่าน
6	ผ่าน

หลังจากตรวจสอบความถูกต้องของการทดสอบความปลอดภัยด้วย API ของผู้วิจัย จึงนำเครื่องมือนี้ไปทดสอบความปลอดภัยของ API สาธารณะอีก 4 API โดยตารางที่ 3 คือตัวอย่างผลลัพธ์ที่ได้จากการทดสอบพบว่าทุกการดำเนินการของ API นี้ไม่ผ่านการทดสอบในประเด็น Batching of Attack และมี 4 การดำเนินการที่ไม่ผ่านการทดสอบในประเด็น Denial of Service Queries ได้แก่ countries, country, continents และ continent ส่งผลให้เกิดช่องโหว่ที่ผู้ไม่หวังดีสามารถส่งคำร้องที่ Brute Force พารามิเตอร์ด้วยการส่งคำร้องเพียงครั้งเดียว และยังสามารถส่งคำร้องที่มี Object ซ้อนกันหลายระดับชั้นมาที่เครื่องแม่ข่าย เมื่อมีการประมวลผลเกิดขึ้นอาจทำให้ทรัพยากรของเครื่องแม่ข่ายทำงานหนักมากเกินไป

ตารางที่ 3: ตัวอย่างผลลัพธ์จากการทดสอบความปลอดภัยของ API สาธารณะ

Operation	Injection		Batching of Attack	DoS Queries
	SQLi	XSS		
countries	ผ่าน	ผ่าน	ไม่ผ่าน	ไม่ผ่าน
country	ผ่าน	ผ่าน	ไม่ผ่าน	ไม่ผ่าน
continents	ผ่าน	ผ่าน	ไม่ผ่าน	ไม่ผ่าน
continent	ผ่าน	ผ่าน	ไม่ผ่าน	ไม่ผ่าน
languages	ผ่าน	ผ่าน	ไม่ผ่าน	ผ่าน
language	ผ่าน	ผ่าน	ไม่ผ่าน	ผ่าน

5. สรุป

งานวิจัยนี้ได้นำเสนอเครื่องมือทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่สำหรับ GraphQL API แบบอัตโนมัติ

โดยครอบคลุมการทดสอบคุณสมบัติเชิงไม่ใช้หน้าที่ทั้งหมด 3 รายการ ได้แก่ ประสิทธิภาพ ความพร้อมใช้งาน และความปลอดภัย โดยได้มีการนำเครื่องมือนี้ไปทดสอบกับ API ที่เปิดใช้งานแบบสาธารณะจำนวน 4 API และ API ที่ผู้วิจัยได้พัฒนาขึ้นเองจำนวน 1 API พบว่าสามารถหาค่าเฉลี่ยความเร็วการตอบกลับของแต่ละการดำเนินการของ API เพื่อแสดงให้เห็นถึงประสิทธิภาพของระบบ และหา API uptime เพื่อแสดงให้เห็นถึงความพร้อมใช้งานของระบบ รวมถึงสามารถหาช่องโหว่ทางด้านความปลอดภัยทั้ง 3 ประเด็น ได้แก่ Injection, Batching of Attack และ Denial of Service Queries

สำหรับการต่อยอดงานวิจัยนี้ต่อไปในอนาคต ผู้วิจัยมีความคิดที่จะทำให้เครื่องมือนี้สามารถครอบคลุมการทดสอบทั้งฟังก์ชันการทำงานและคุณสมบัติเชิงไม่ใช้หน้าที่ในรายการอื่น ๆ เช่น การขยายตัว (Scalability) เป็นต้น รวมถึงทดสอบกับ API จำนวนมากขึ้น เพื่อให้สามารถครอบคลุมการทดสอบความปลอดภัยในประเด็นอื่นที่เหลือ ได้แก่ Authorization ต้องทดสอบกับ GraphQL API ที่มีการจำกัดสิทธิ์ของการเข้าถึงข้อมูล Exposure of Underlying API ต้องทดสอบกับ GraphQL API ที่คืนข้อมูลจาก API รูปแบบอื่น เช่น REST API เป็นต้น ในขณะที่การทดสอบในประเด็นของ Detailed Error Message ต้องอาศัยการวิเคราะห์ข้อมูลที่ได้จาก API ว่ามีโอกาสเป็นข้อมูลที่ไม่ควรเปิดเผยหรือไม่ เช่น ข้อผิดพลาดภายในระบบ (Internal Error Details) หรือ การกำหนดค่าของแอปพลิเคชัน (Application Configurations) เป็นต้น

เอกสารอ้างอิง

- [1] What is GraphQL? Available online at <https://www.redhat.com/en/topics/api/what-is-graphql>.
- [2] What is an API? Available online at <https://aws.amazon.com/th/what-is/api>.
- [3] What is RESTful API? Available online at <https://aws.amazon.com/th/what-is/restful-api>.
- [4] D. Mairiza, D. Zowghi and N. Nurmuliani, An investigation into the notion of non-functional requirements, Proceedings of the 2010 ACM Symposium on Applied Computing (SAC). 2010.
- [5] S. Karlsson, A. Čaušević and D. Sundmark, Automatic Property-based Testing of GraphQL APIs, IEEE/ACM International Conference on Automation of Software Test (AST). 2021.
- [6] S. Bucaille, J. Luis Cánovas Izquierdo, H. Ed-douibi and J. Cabot, An OpenAPI-Based Testing Framework to Monitor Non-functional Properties of REST APIs, International Conference on Web Engineering. 2020.
- [7] E. Lee, K. Kwon and J. Yun, Performance Measurement of GraphQL API in Home ESS Data Server, International Conference on Information and Communication Technology Convergence (ICTC). 2020.
- [8] E. Wittern, A. Cha, James C. Davis, G. Baudart and L. Mandel, An Empirical Study of GraphQL Schemas, International Conference on Service-Oriented Computing. 2019.
- [9] Testing GraphQL. Available online at https://owasp.org/www-project-web-security-testing-guide/latest/4-Web_Application_Security_Testing/12-API_Testing/01-Testing_GraphQL.
- [10] SecLists. Available online at <https://github.com/danielmiessler/SecLists>.
- [11] Public GraphQL APIs. Available online at <https://github.com/IvanGoncharov/graphql-apis>.

การพยากรณ์การใช้ไฟฟ้าของประเทศไทยด้วยเทคนิคเหมืองข้อมูล

Forecast of Electric Consumption in Thailand using Data Mining Technique

มินตรา มีสาวงษ์ (Mintra Mesawong)¹ และธนพล เจนสุทธิเวชกุล (Tanapon Jentsuttiwetchakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856020@email.kmutnb.ac.th, ²tanapon.j@itd.kmutnb.ac.th

บทคัดย่อ

บทความวิจัยฉบับนี้เป็นกรนำเสนอการพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทย โดยทำการเปรียบเทียบประสิทธิภาพจำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึมการถดถอยเชิงเส้น อัลกอริทึมโครงข่ายประสาทเทียม และอัลกอริทึมการเรียนรู้เชิงลึก โดยข้อมูลนักท่องเที่ยวต่างชาติที่เดินทางเข้าประเทศไทย ผลิตภัณฑ์มวลรวมของประเทศ อัตราการใช้กำลังการผลิตจำนวนประชากร จำนวนบ้านที่อยู่อาศัย และจำนวนรถยนต์ไฟฟ้า (BEV: Battery Electric Vehicle) และข้อมูลปริมาณการใช้ไฟฟ้าของประเทศไทย ปี พ.ศ. 2559 - 2565 จำนวน 84 เดือน มาเพื่อศึกษาคุณลักษณะของตัวแปรที่มีผลต่อปริมาณการใช้ไฟฟ้า จากการประเมินประสิทธิภาพของอัลกอริทึม พบว่า อัลกอริทึมโครงข่ายประสาทเทียมให้ประสิทธิภาพการพยากรณ์ที่ดีที่สุด โดยค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) อยู่ที่ 4.453

คำสำคัญ: อัลกอริทึมการถดถอยเชิงเส้น อัลกอริทึมโครงข่ายประสาทเทียม อัลกอริทึมการเรียนรู้เชิงลึก

Abstract

This research aims to develop a model for forecasting of electric consumption in Thailand. Performances of three algorithms, namely linear regression, artificial neural network (ANN) and deep learning were compared. Data form various aspects such as number of foreign tourists, GDPs, capacity utilization rate, population, number of residential houses, number of electric Vehicles were use with the

electric consumption of Thailand in the year 2016 – 2022 in developing forecasting model. The characteristics of variables affecting electric consumption are considered. The result indicated that artificial neural network (ANN) provided the best performance gave the Mean Absolute Percentage Error (MAPE) predictive performance of 4.453.

Keywords: Linear Regression, Neural network, Deep Learning.

1. บทนำ

ในปี 2564 ระบบไฟฟ้าของ กฟผ. มีความต้องการพลังงานไฟฟ้าสูงสุดสุทธิเท่ากับ 30,135.30 เมกะวัตต์ ซึ่งเกิดขึ้นเมื่อวันที่ 3 เมษายน 2564 เวลา 21.03 น. มีค่าเพิ่มขึ้นจากความต้องการพลังงานไฟฟ้าสูงสุดสุทธิของปี 2563 เป็นจำนวน 1,498.60 เมกะวัตต์ หรือเพิ่มขึ้นร้อยละ 5.23 สำหรับความต้องการพลังงานไฟฟ้าสุทธิของระบบ กฟผ. ในปี 2564 มีค่าเท่ากับ 194,868.69 ล้านกิโลวัตต์-ชั่วโมง มีค่าเพิ่มขึ้นจากความต้องการพลังงานไฟฟ้าสุทธิของปี 2563 ที่มีค่าเท่ากับ 191,934.55 ล้านกิโลวัตต์-ชั่วโมง เป็นจำนวน 2,934.14 ล้านกิโลวัตต์-ชั่วโมง หรือคิดเป็นร้อยละ 1.53[1]

ไฟฟ้าได้กลายเป็นปัจจัยสำคัญในการพัฒนาเศรษฐกิจและสังคมของทุกประเทศทั่วโลก การบริหารจัดการไฟฟ้าจึงเป็นนโยบายหลักของรัฐบาลในแต่ละประเทศให้มีความสำคัญ เพราะนั่นหมายถึง ชีวิตความเป็นอยู่ที่ดีของประชาชน จึงจำเป็นต้องวางแผนการใช้ “เชื้อเพลิง” ให้เหมาะสม ‘ตามบริบทของแต่ละประเทศ’ ไม่ว่าจะเป็นสภาพเศรษฐกิจ สภาพภูมิศาสตร์ และความต้องการใช้ไฟฟ้า ซึ่งหน่วยงานที่รับผิดชอบมีส่วนสำคัญอย่างยิ่งในการกำหนด

นโยบายให้ประชาชนมีไฟฟ้าใช้อย่างเพียงพอ มีคุณภาพที่ดี มีราคาที่เป็นธรรม ควบคู่ไปกับการดูแลรักษาสิ่งแวดล้อม ดังนั้นแล้ว ต้องคำนึงถึงข้อดี-ข้อจำกัดของแต่ละเชื้อเพลิงผลิตไฟฟ้าอย่างรอบด้านและครบทุกมิติ [2]

ดังนั้นในงานวิจัยนี้ ผู้วิจัยจึงมีแนวคิดในการพยากรณ์ ปริมาณการใช้ไฟฟ้าของประเทศ เพื่อช่วยให้การบริหารจัดการปริมาณไฟฟ้ามีประสิทธิภาพมากขึ้น โดยได้นำ ข้อมูลนักท่องเที่ยวชาวต่างชาติที่เดินทางเข้าประเทศไทย ผลิตภัณฑ์มวลรวมของประเทศ อัตราการใช้กำลังการผลิต จำนวนประชากร จำนวนบ้านที่อยู่อาศัย และจำนวนรถยนต์ไฟฟ้า (BEV: Battery Electric Vehicle) จำนวน 84 เดือน ระหว่างปี พ.ศ 2559 – 2565 มาใช้ในการสร้างโมเดลในการพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศ โดยนำเทคนิคเหมืองข้อมูลด้วยวิธีการพยากรณ์ ประกอบด้วย การถดถอยเชิงเส้น โคจรข่ายประสาทเทียม และการเรียนรู้เชิงลึก มาทำการเปรียบเทียบประสิทธิภาพ และเลือกวิธีที่เหมาะสมมากที่สุด มาเป็นแบบต้นแบบในการพยากรณ์ ปริมาณการใช้ไฟฟ้าของประเทศไทย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล (Data Mining)

เป็นเทคนิคการวิเคราะห์ข้อมูลเพื่อค้นหารูปแบบ (Pattern) ความสัมพันธ์ ของข้อมูลที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติและการเรียนรู้ของเครื่องคอมพิวเตอร์แล้วนำผลวิเคราะห์ที่ได้ไปใช้ประโยชน์เพื่อการตัดสินใจในการเรียนรู้ของเครื่องคอมพิวเตอร์เพื่อให้ใกล้เคียงกับสมองมนุษย์ มีชนิดของการเรียนรู้ข้อมูล 2 แบบคือ (1) การเรียนรู้แบบมีการสอน (Supervised learning) เป็นการเรียนรู้จากข้อมูลที่มีผู้สอน หรือกล่าวได้ว่าข้อมูลนั้นมีการกำหนดคลาสหรือค่าเฉลี่ย และ (2) การเรียนรู้แบบไม่มีการสอน (Unsupervised learning) เป็นการเรียนรู้แบบไม่มีผู้สอน หรือกล่าวได้ว่าข้อมูลนั้น ไม่มีการกำหนดคลาสซึ่งส่วนใหญ่เป็นข้อมูลธรรมชาติที่พบได้โดยทั่วไป การทำเหมืองข้อมูลใช้โมเดลในการเรียนรู้ จึงต้องนำอัลกอริทึมของการเรียนรู้ของเครื่อง (Machine learning) มาใช้โดยชนิดของอัลกอริทึม ในการทำเหมืองข้อมูล (Data Mining Algorithms) แบ่งได้เป็นสองกลุ่มใหญ่ [3]

2.1 ปัจจัยที่ใช้ในการทำนายปริมาณการใช้ไฟฟ้า

2.1.1 ปริมาณการใช้ไฟฟ้า

เป็นคลาสเป้าหมายในการพยากรณ์ของงานวิจัยนี้ โดยทำการรวบรวมข้อมูลสถิติการใช้ไฟฟ้าของทั้งประเทศ พ.ศ.2559–2565 จากระบบบัญชีข้อมูลกลาง (Government Data Catalog) จำแนกตามสาขาการใช้ไฟ 7 รายการที่แสดงผลรวมทุกสาขา มีหน่วยเป็นจำนวน “GWh” ดังนี้ (1) Residential (2) Free of Charge (3) Government & Non-Profit (4) Business (5) Industrial (6) Agriculture (7) Other

2.1.2 สถิตินักท่องเที่ยวชาวต่างชาติที่เดินทางเข้าประเทศไทย

หลังสถานการณ์โควิด-19 ธุรกิจต่างๆ เช่น โรงแรม ร้านค้าตามแหล่งท่องเที่ยวต่างๆ กลับมาเปิดให้บริการ ส่งผลให้มีการใช้ไฟฟ้าเพิ่มมากขึ้นในการทำกิจกรรมทางธุรกิจ ในงานวิจัยนี้จึงได้นำปัจจัยดังกล่าวมาเป็นปัจจัยในการพยากรณ์ โดยทำการรวบรวมจากเว็บไซต์กระทรวงการท่องเที่ยวและกีฬา ข้อมูล ซึ่งเป็นข้อมูลนักท่องเที่ยวจากทุกประเทศที่เดินทางเข้ามายังประเทศไทย ตั้งแต่ปี พ.ศ. 2559 – 2565

2.1.3 สถิติผลิตภัณฑ์มวลรวมของประเทศ (Gross Domestic Product)

เป็นตัวเลขที่แสดงมูลค่าสินค้าและบริการขั้นสุดท้ายทั้งหมดที่ผลิตขึ้นในประเทศภายในช่วงระยะเวลาหนึ่ง เช่น 1 ไตรมาส และ 1 ปี เพื่อบอกมูลค่าเศรษฐกิจที่ได้รับความนิยมมากที่สุดในปัจจุบัน อีกทั้งยังใช้ในการเปรียบเทียบกับ GDP ในอีกช่วงเวลาหนึ่งเพื่อใช้ออกการเติบโตหรือการหดตัวของเศรษฐกิจ [4] ซึ่งผู้วิจัยได้ทำการศึกษางานวิจัยที่เกี่ยวข้อง [5] พบว่า ปริมาณการใช้ไฟฟ้ามีอิทธิพลหรือเป็นสาเหตุให้เกิดการเปลี่ยนแปลงในผลิตภัณฑ์มวลรวมภายในประเทศภาคบริการอย่างมีนัยสำคัญทางสถิติ ในขณะที่เดียวกันผลิตภัณฑ์มวลรวมภายในประเทศภาคบริการก็มีอิทธิพลหรือเป็นสาเหตุให้เกิดการเปลี่ยนแปลงของปริมาณการใช้ไฟฟ้าของภาคบริการทั้งในภาพรวมและสาขาย่อยเช่นเดียวกัน แสดงว่าทั้งปริมาณการใช้ไฟฟ้ากับผลิตภัณฑ์มวลรวมภายในประเทศภาคบริการ มีลักษณะเป็นทั้งสาเหตุและเป็นทั้งผลซึ่งกันและกัน

2.1.4 อัตราการใช้กำลังการผลิต

เป็นเครื่องชี้ระดับการผลิตของภาคอุตสาหกรรม โดยเปรียบเทียบการผลิตจริงกับกำลังการผลิตสูงสุดหรือ

ศักยภาพของเครื่องจักร โดยความสำคัญของอุตสาหกรรมแต่ละกลุ่มแตกต่างกันตามสัดส่วนค่าถ่วงน้ำหนักของแต่ละอุตสาหกรรม วิธีการคำนวณ (สำหรับชุดข้อมูลสถิติ) ใช้สูตรลาสเปร์รี่ โดยนำข้อมูลสถิติรายสินค้าที่ได้มาหาผลรวมของปริมาณผลผลิตของสินค้า ณ เดือนปัจจุบันหารกับปริมาณกำลังการผลิตสูงสุดของสินค้า ณ ปีฐาน คูณค่าถ่วงน้ำหนักมูลค่าเพิ่ม ณ ปีฐาน แล้วคูณ 100

$$\text{อัตราการใช้จ่ายการผลิต}^t = \sum \frac{Q_{it}^p}{Q_{it}^c} \times W_i^0 \times 100$$

โดยที่ อัตราการใช้จ่ายการผลิต^t แทน อัตราการใช้จ่ายการผลิต ณ เวลา t

Q_{it}^p แทนปริมาณผลผลิตจริงของสินค้าที่ i ณ เวลา t

Q_{it}^c แทนปริมาณกำลังการผลิตสูงสุดของสินค้าที่ i ณ เวลา t

W_i^0 แทนค่าถ่วงน้ำหนักของกลุ่มอุตสาหกรรมที่ i ในปีฐาน

จากการศึกษาพบว่า ไฟฟ้าเป็นปัจจัยด้านเทคโนโลยีที่เข้ามาเปลี่ยนแปลงกระบวนการผลิต ดังนั้นอัตรากำลังการผลิต เป็นตัวสะท้อนการใช้จ่ายไฟฟ้าของภาคอุตสาหกรรม

2.1.5 สถิติจำนวนประชากร

เนื่องจากกิจกรรมการใช้จ่ายไฟฟ้าต่างๆ มาจากการดำเนินกิจกรรมของประชากรในประเทศ ดังนั้นจำนวนประชากรจึงเป็นตัวสะท้อนความต้องการในการใช้จ่ายไฟฟ้ายาน วิจัยนี้ได้ทำการรวบรวมข้อมูลสถิติประชากรศาสตร์จากสำนักบริหารการทะเบียน กรมการปกครอง โดยเป็นข้อมูลจำนวนประชากรรายเดือน

2.1.6 สถิติจำนวนบ้านที่อยู่อาศัย

เนื่องจากในแต่ละบ้านที่อยู่อาศัยจะมีเครื่องใช้ไฟฟ้า เพื่อคอยอำนวยความสะดวกในการดำเนินชีวิต ซึ่งมีผลต่อปริมาณการใช้จ่ายไฟฟ้าด้วย โดยงานวิจัยนี้ได้ทำการรวบรวมข้อมูลสถิติจำนวนบ้าน จากสำนักบริหารการทะเบียน กรมการปกครอง เป็นข้อมูลบ้านทั่วประเทศแยกเป็นรายเดือน มีหน่วยเป็น “หลัง”

2.1.7 สถิติจำนวนรถยนต์ไฟฟ้า (BEV: Battery Electric Vehicle)

แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 13 ตั้งเป้าหมายที่ 3 ว่า “ไทยเป็นฐานการผลิตยานยนต์ไฟฟ้าที่สำคัญของโลก” โดยมุ่งเน้นการพัฒนายานยนต์ที่ปล่อยมลพิษเป็นศูนย์ (Zero Emission Vehicle: ZEV) ตั้งเป้าหมายการ

ใช้ยานยนต์ไฟฟ้า ในปี 2573 จำนวน 440,000 คัน และเป้าหมายการผลิตจำนวน 725,000 คัน เพิ่มสถานีชาร์จให้เพิ่มขึ้น 5,000 หัวจ่าย และส่งเสริมแรงงานด้านยานยนต์ไฟฟ้าไม่น้อยกว่า 30,000 คน [6] ดังนั้นผู้วิจัยจึงให้ความสนใจและได้นำจำนวนรถยนต์ไฟฟ้า (BEV: Battery Electric Vehicle) รายเดือนมาเป็นปัจจัยในการพยากรณ์ปริมาณการใช้จ่ายไฟฟ้า

2.2 อัลกอริทึมที่ใช้ในการพยากรณ์

2.2.1 การวิเคราะห์การถดถอย (Linear Regression)

เป็นวิธีการทางสถิติที่ใช้ในการตรวจสอบลักษณะของความสัมพันธ์ระหว่างตัวแปร ตั้งแต่ 2 ตัวแปรขึ้นไป โดยแบ่งเป็นตัวแปรอิสระ (Independent variable) และตัวแปรตาม (Dependent variable) เพื่อศึกษาจะทำให้ทราบถึงความสัมพันธ์ระหว่างตัวแปรอิสระ ที่มีต่อตัวแปรตาม และแบบจำลองความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม[7]

2.2.2 นิวรอนเน็ตเวิร์ก (Artificial Neural Network)

คือโมเดลทางคณิตศาสตร์ ที่จำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ โดยสมองของมนุษย์นั้นจะมีหน่วยประมวลผลขนาดเล็กรวมกันอยู่มากมาย ซึ่งมีความสามารถในการเรียนรู้การจดจำแบบรูป (Pattern Recognition) และการอุปมาความรู้ (Knowledge Deduction) ซึ่งมีการปรับเปลี่ยนตัวเองต่อการตอบสนองของอินพุตตามกฎของการเรียนรู้ (Learning Rule) หลังจากที่โครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายนั้นจะสามารถทำงานที่กำหนดไว้ได้ ซึ่งโครงข่ายประกอบด้วย Input Layer, Hidden Layer และ Output Layer ภายในแต่ละ Layer จะประกอบด้วย โหนด (Node) ซึ่งความซับซ้อนของจำนวน Layer และ Node ขึ้นอยู่กับการออกแบบและความเหมาะสมในการทำงาน [8] สามารถนำข้อมูลทางสถิติเดิมที่มีอยู่ในระบบมาใช้คาดการณ์หรือพยากรณ์ข้อมูลในอนาคตได้และยังเกิดข้อผิดพลาดได้ยาก

2.2.3 การเรียนรู้เชิงลึก (Deep Learning)

คือเทคนิคที่ถูกสร้างขึ้นมาเพื่อเรียนรู้ของเครื่องจักร (Machine Learning) หรือเครื่องคอมพิวเตอร์ มีพื้นฐานมาจากโครงข่ายประสาทเทียม (Artificial Neural Network : ANN) โดยชุดคำสั่งนี้จะทำให้ตัวเครื่องจักรสามารถประมวลผลข้อมูลจำนวนมากด้วยการจำลองเครือข่ายประสาทแบบเดียวกับในสมองของมนุษย์

ที่พยายามเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพเช่น การเรียนรู้จดจำใบหน้าหรือการรู้จำการแสดงออกทางสีหน้า โดย Deep Learning มีประสิทธิภาพดีในชุดข้อมูลขนาดใหญ่

2.3 การวัดประสิทธิภาพ และเปรียบเทียบแบบจำลอง

ค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) เป็นค่าความคลาดเคลื่อนที่แปลงให้อยู่ในรูปแบบของเปอร์เซ็นต์ในการวัดค่าความแม่นยำ จากวิธีการนี้ยิ่งค่าที่ได้มีค่าน้อยแสดงว่าโมเดลที่ได้จะมีความแม่นยำมาก โดยมีสมการดังนี้

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| 100 * \frac{Y_t - \hat{Y}_t}{Y_t} \right|$$

โดยที่ MAPE คือค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน

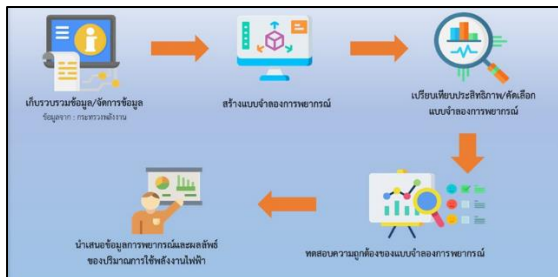
n คือจำนวนข้อมูลที่ได้

Y_t คือค่าจริงที่เวลา t ใดๆ

\hat{Y}_t คือค่าที่ได้จากการพยากรณ์ที่เวลา t ใดๆ

3. การดำเนินงานวิจัย

การกำหนดกรอบแนวคิดการดำเนินงานวิจัย ในการพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทย แสดงดังภาพที่ 1 โดยมีขั้นตอนการดำเนินงานดังต่อไปนี้



ภาพที่ 1 : กรอบแนวคิดการดำเนินงานวิจัย

3.1 การเก็บรวบรวมข้อมูล และเตรียมข้อมูล

การดำเนินงานเก็บรวบรวมข้อมูล และเตรียมข้อมูล นับเป็นขั้นตอนที่สำคัญมากในการทำวิจัย ได้ทำการเก็บรวบรวมข้อมูลการใช้ไฟฟ้าของทั้งประเทศรายเดือน สถิตินักท่องเที่ยวต่างชาติที่เดินทางเข้าประเทศไทย สถิติผลิตภัณฑ์มวลรวมของประเทศ อัตราการใช้กำลังการผลิต สถิติจำนวนประชากร สถิติจำนวนบ้านที่อยู่อาศัย สถิติจำนวนรถยนต์ไฟฟ้า ดังตารางที่ 1 โดยใช้โปรแกรม Microsoft Excel ในการ

จัดเก็บข้อมูลตรวจสอบความสมบูรณ์ของข้อมูล จะได้ข้อมูลทั้งหมด 84 รายการ (จำนวน 7 ปี หรือ 84 เดือน)

ตารางที่ 1 : บัญชีที่นำเข้าที่ใช้ในการสร้างแบบจำลอง

ข้อมูล	ชื่อ Attribute
การใช้ไฟฟ้าของทั้งประเทศ (จำแนกตามสาขา)	Electric_Total
สถิตินักท่องเที่ยวต่างชาติที่เดินทางเข้าประเทศไทย	Tourism
สถิติ GDP (Gross Domestic Product) หรือ สถิติผลิตภัณฑ์มวลรวมของประเทศ	GDP
อัตราการใช้กำลังการผลิต	Product
สถิติจำนวนประชากร	Population
สถิติจำนวนบ้านที่อยู่อาศัย	Home
สถิติจำนวนรถยนต์ไฟฟ้า (BEV: Battery Electric Vehicle)	BEV

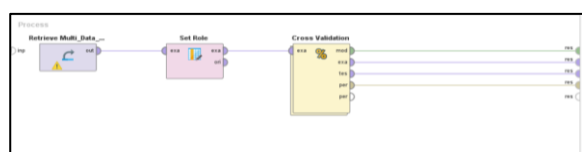
3.2 การเตรียมข้อมูล

3.2.1 การทำความสะอาดข้อมูลโดยการตรวจสอบและแก้ไขข้อมูลที่ไม่สมบูรณ์ เช่น ข้อมูลไม่ครบถ้วน ข้อมูลมีค่าผิดปกติ ได้แก่การจัดการกับสถิติของ GDP ข้อมูลจะมาในรูปแบบของไตรมาส จึงจำเป็นต้องทำให้อยู่ในรูปแบบของรายเดือนเพื่อให้สามารถใช้ในการพยากรณ์ได้เหมือนกับข้อมูลชนิดอื่น ๆ

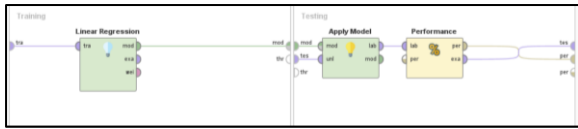
3.2.2 การคัดเลือกข้อมูล ควรเลือกข้อมูลตัวแปรอิสระให้มีความสอดคล้องกับตัวแปรตาม พิจารณาจากปัจจัยที่ส่งผลจากงานวิจัยที่เกี่ยวข้องโดยสามารถคัดเลือกได้ 12 แอททริบิวต์ แสดงดังตารางที่ 1 กำหนด Electric Total เป็นค่าเป้าหมาย

3.3 การพัฒนาแบบจำลอง

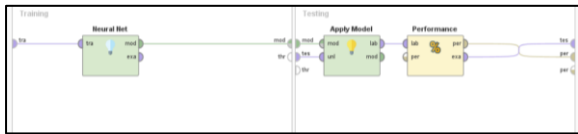
สร้างแบบจำลองด้วยโปรแกรม RapidMiner Studio โดยใช้ข้อมูลการใช้ไฟฟ้าของทั้งประเทศตั้งแต่ปี 2559 – 2565 โดยเป็นการเปรียบเทียบ อัลกอริทึมจำนวน 3 อัลกอริทึม ได้แก่ 1) อัลกอริทึมการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) 2) อัลกอริทึมโครงข่ายประสาทเทียม (Neural Network) 3) อัลกอริทึมการเรียนรู้เชิงลึก (Deep Learning) เพื่อหาประสิทธิภาพและหาค่าเฉลี่ยของความผิดพลาดน้อยที่สุดของการพยากรณ์ จึงใช้เครื่องมือ Cross Validation ดังภาพที่ 2 ถึง 5



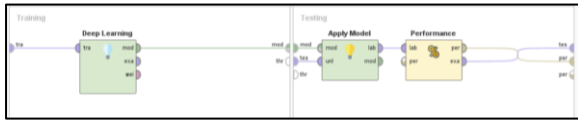
ภาพที่ 2 : การใช้เครื่องมือ Cross Validation



ภาพที่ 3 : การหาค่าและทำสอบพารามิเตอร์ด้วย Linear Regression



ภาพที่ 4 : การหาค่าและทำสอบพารามิเตอร์ด้วย Neural Network



ภาพที่ 5 : การหาค่าและทำสอบพารามิเตอร์ด้วย Deep Learning

3.4 การวัดประสิทธิภาพและเปรียบเทียบแบบจำลอง

ในงานวิจัยนี้ทำการวัดประสิทธิภาพโดยทำการวัดค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE)

4. ผลการดำเนินงาน

4.1 ผลการพัฒนาแบบจำลองการถดถอยเชิงเส้น (Linear Regression)

โดยการกำหนดค่า Min Tolerance อยู่ในช่วง 0.1 ถึง 1 จะมีค่า ค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) แสดงดังตารางที่ 2

ตารางที่ 2 : ผลเปรียบเทียบประสิทธิภาพ การวิเคราะห์การถดถอยเชิงเส้น

Min Tolerance	MAPE
0.1	5.058663
0.2	5.058663
0.3	5.058663
0.4	5.098281
0.5	5.060496
0.6	4.998543
0.7	4.998543
0.8	5.106853
0.9	5.113964
1.0	5.153654

4.2 ผลการพัฒนาแบบจำลองการวิเคราะห์อัลกอริทึมโครงข่ายประสาทเทียม (Neural Network)

โดยกำหนดค่าพารามิเตอร์ Learning rate อยู่ในช่วง 0.1 ถึง 0.7 ค่า Momentum อยู่ในช่วง 0.1 ถึง 0.7 และค่า Training cycle ที่ 200,400,600,800 จะมีค่า ค่าสัมบูรณ์ของเปอร์เซ็นต์

ของความคลาดเคลื่อน (Mean Absolute Percentage Error: MAPE) แสดงดังตารางที่ 3

ตารางที่ 3 : ตัวอย่างผลเปรียบเทียบประสิทธิภาพ การวิเคราะห์โครงข่ายประสาทเทียม

Learning rate	Momentum	Training cycle			
		200	400	600	800
0.1	0.1	4.902	4.672	4.682	5.028
0.1	0.2	4.875	4.586	4.808	4.772
0.1	0.3	4.755	4.672	4.800	4.778
0.1	0.4	4.724	4.773	4.8	4.717
0.1	0.5	4.86	4.853	4.9	4.928
0.1	0.6	4.829	5.076	5.372	5.812
0.1	0.7	4.945	4.904	4.991	5.502
0.2	0.1	4.638	5.179	5.220	5.258
0.2	0.2	4.638	4.453	4.636	4.894
0.2	0.3	4.709	4.984	5.205	5.301
0.2	0.4	4.829	5.153	5.236	5.444
0.2	0.5	5.111	4.975	5.290	5.641
0.2	0.6	5.059	5.032	5.446	5.388
0.2	0.7	5.512	5.327	5.285	5.348

4.3 ผลการพัฒนาแบบจำลองการวิเคราะห์อัลกอริทึมการเรียนรู้เชิงลึก (Deep Learning)

โดยกำหนดค่าพารามิเตอร์ Epochs 1-10 จะมีค่า ค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) แสดงดังตารางที่ 4

ตารางที่ 4 : ผลเปรียบเทียบประสิทธิภาพ การเรียนรู้เชิงลึก

Epochs	MAPE
1	5.026488
2	4.796251
3	5.393768
4	5.016073
5	5.175651
6	5.012928
7	4.937738
8	5.190643
9	4.973781
10	4.590921

4.4 ผลการเปรียบเทียบและคัดเลือกแบบจำลองพยากรณ์

จากผลการวิจัยจึงนำการวิเคราะห์โครงข่ายประสาทเทียม เนื่องจากได้ค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) ต่ำสุด แสดงดังตารางที่ 5

ตารางที่ 5 : ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

Linear Regression	Neural Network	Deep Learning
4.998	4.453	4.591

5. สรุปผลและอภิปราย

งานวิจัยครั้งนี้ ผู้วิจัยได้นำเสนอการเปรียบเทียบประสิทธิภาพสำหรับการพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทย โดยใช้แบบจำลอง 3 อัลกอริทึม ได้แก่ 1) อัลกอริทึมการวิเคราะห์การถดถอยเชิงเส้น 2) อัลกอริทึมโครงข่ายประสาทเทียม 3) อัลกอริทึมการเรียนรู้เชิงลึก โดยงานวิจัยนี้ได้เลือกใช้อัลกอริทึมที่สอดคล้องกับกับงานวิจัยระบบธุรกิจอัจฉริยะสำหรับการพยากรณ์ราคาสินค้าเกษตรที่นำข้อมูลราคาสินค้าเกษตรรายวัน ระหว่างปี พ.ศ. 2559 – 2564 ซึ่งเป็นข้อมูลที่มีรูปแบบคล้ายกันกับข้อมูลที่ผู้วิจัยกำลังศึกษา ผู้วิจัยจึงเลือกใช้อัลกอริทึมดังกล่าวเพื่อศึกษาเปรียบเทียบประสิทธิภาพ โดยการศึกษาจากปัจจัยที่มีผลต่อปริมาณการใช้ไฟฟ้าของประเทศไทยจำนวน 7 ตัวแปร ระหว่างปี พ.ศ. 2559 – 2565 เปรียบเทียบแบบจำลองด้วยค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน (Mean Absolute Percentage Error : MAPE) ที่ต่ำที่สุดโดยมีค่า MAPE อยู่ที่ 4.453 จากผลการพยากรณ์พบว่า การพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทยด้วยอัลกอริทึมโครงข่ายประสาทเทียม มีความแม่นยำมากที่สุด เนื่องจากมีค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อนต่ำสุด ดังนั้นการวิเคราะห์โครงข่ายประสาทเทียม ให้ประสิทธิภาพและค่าเฉลี่ยของความผิดพลาดการพยากรณ์ดีกว่าอัลกอริทึมอัลกอริทึมการเรียนรู้เชิงลึกและอัลกอริทึมการถดถอยเชิงเส้น ในขณะที่งานวิจัย [9] ใช้วิธีบ็อกซ์-เจนกินส์ โดยให้ค่า MAPE ต่ำที่สุดเท่ากับร้อยละ 6.32 เท่านั้น ซึ่งเป็นค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อนที่มากกว่าวิธี โครงข่ายประสาทเทียม ดังนั้นสำหรับงานวิจัยที่ได้ศึกษาครั้งนี้สามารถนำไปเป็นแนวทางเพื่อให้สามารถวางแผนการหาพลังงานทดแทนในการผลิตไฟฟ้าของประเทศไทยได้ และต่อยอดในการพัฒนาเครื่องมือสำหรับทำการวิเคราะห์และช่วยสนับสนุนการวางแผนในอนาคตได้

เอกสารอ้างอิง

- [1] การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย, “รายงานประจำปี 2564 การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย” (2564). [online], [สืบค้นเมื่อวันที่ 27 กุมภาพันธ์ 2566]. จาก https://www.egat.co.th/home/wp-content/uploads/2022/06/EGAT-Annual-2021_2022-06-22.pdf
- [2] การไฟฟ้าฝ่ายผลิต, “เชื้อเพลิงเพื่อการผลิตไฟฟ้า” [Online], [สืบค้นเมื่อวันที่ 11 กุมภาพันธ์ 2566]. จาก <https://www.egat.co.th/home/fuel/>
- [3] วราวุฒิ นาคบุญนำ, ‘การพยากรณ์ภาวะซึมเศร้าด้วยวิธีการเหมืองข้อมูล’, ปริญญาปรัชญาดุษฎีบัณฑิต, มหาวิทยาลัยมหาสารคาม, 2564.
- [4] FINNOMENA, “GDP คือ อะไร? ตัวเลขสำคัญทางเศรษฐกิจที่นักลงทุนควรรู้จัก” [ออนไลน์], [สืบค้นเมื่อวันที่ 26 กุมภาพันธ์ 2566]. จาก <https://www.finnomena.com/getwealthissoon/what-is-gdp/>
- [5] นิรมล ดันดี ชนบริบูรณ์, ‘การศึกษาความสัมพันธ์ระหว่างปริมาณการใช้ไฟฟ้ากับผลิตภัณฑ์มวลรวมภายในประเทศของ ภาคบริการในประเทศไทย’, มหาวิทยาลัยศรีนครินทรวิโรฒ, 2561.
- [6] “หุ้น EV ในตลาดหุ้นไทย ตัวไหนน่าจับตามองบ้าง?” [Online], [สืบค้นเมื่อวันที่ 28 กุมภาพันธ์ 2566]. จาก <https://www.moneybuffalo.in.th/stock/ev-stocks>
- [7] พรพรรณ เจริญศิลป์, (2564) “ระบบธุรกิจอัจฉริยะสำหรับการพยากรณ์ราคาสินค้าเกษตร” การค้นคว้าอิสระวิทยาศาสตร์มหาบัณฑิต สาขาวิชาการบริหารสารสนเทศเพื่อการจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- [8] พิพัฒน์ หล่อทรัพย์คง, ‘การคาดการณ์ผลตอบแทนในอนาคตของพอร์ตโพลีโอบของตราสารทุนหมวดพาณิชย์และเทคโนโลยีสารสนเทศและการสื่อสาร โดยการใช้โครงข่ายประสาทเทียม’, สารนิพนธ์, วิทยาลัยการจัดการ มหาวิทยาลัยมหิดล, 2562
- [9] ญาดา. มงคลลีลาไพบูลย์, ‘การพยากรณ์ปริมาณการใช้ไฟฟ้าสำหรับกิจการเฉพาะรายรายเดือนกรณีศึกษา: พื้นที่อำเภอเมือง จังหวัดนครราชสีมา’, December 10, 2021.

การพยากรณ์การเปลี่ยนแปลงสภาพอากาศในพื้นที่เพาะปลูกข้าว ทางภาคตะวันออกเฉียงเหนือของประเทศไทยโดยใช้เทคนิคการวิเคราะห์อนุกรมเวลา

Forecasting of Climate Change in Rice Planting Areas in Northeastern Thailand with Time Series Analysis Techniques

กัญญ์กวิน แชมช้อย (Kankawin Chamchoi)¹ และณัฐวี อุดกฤษณ์ (Nattavee Utakrit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856071@email.kmutnb.ac.th, ²nattavee.u@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อทำการศึกษาคัดเลือกแบบจำลองที่เหมาะสมสำหรับการพยากรณ์การเปลี่ยนแปลงสภาพอากาศในพื้นที่เพาะปลูกข้าวทางภาคตะวันออกเฉียงเหนือของประเทศไทยโดยใช้เทคนิคการวิเคราะห์อนุกรมเวลา ประกอบด้วยวิธีทางสถิติ 4 วิธี ได้แก่ วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์เชิงบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์เชิงคูณ วิธีการวิเคราะห์การถดถอยเชิงเส้น และวิธีบ็อกซ์-เจนกินส์ ใช้เกณฑ์ในการคัดเลือกด้วยค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) โดยศึกษาจากข้อมูลอุณหภูมิ ปริมาณน้ำฝน และความชื้นในอากาศ ของจังหวัดอุบลราชธานี นครราชสีมา ร้อยเอ็ด สุรินทร์ และศรีสะเกษ จำนวน 84 เดือน โดยข้อมูลชุดแรกจำนวน 72 เดือน ตั้งแต่เดือนมกราคม 2559 - ธันวาคม 2564 สำหรับใช้ในการทดสอบแบบจำลองการพยากรณ์ทั้ง 4 วิธี และใช้ข้อมูลชุดที่ 2 จำนวน 12 เดือน ตั้งแต่เดือนมกราคม - ธันวาคม 2565 เพื่อทำการทวนสอบความถูกต้องของแบบจำลอง จากผลการวิจัยพบว่า ข้อมูลอุณหภูมิ ปริมาณน้ำฝน และความชื้นในอากาศ ที่นำมาพยากรณ์ด้วยวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์เชิงบวก ให้ประสิทธิภาพในการพยากรณ์สูงที่สุด โดยมีค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ยอยู่ที่ 0.86, 102.71 และ 6.87 ตามลำดับ

คำสำคัญ: เทคนิคการวิเคราะห์อนุกรมเวลา วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ การวิเคราะห์การถดถอยเชิงเส้น บ็อกซ์-เจนกินส์ การเปลี่ยนแปลงสภาพอากาศ

Abstract

The objective of this research is to study and select a suitable modeling approach for Forecasting of Climate Change in Rice Planting Areas in Northeastern Thailand using Time Series Analysis, including four statistical methods: Holt-Winters exponential smoothing with additive method, Holter-Winters exponential smoothing with multiplicative method, linear regression analysis method, and Box-Jenkins method. The selection criteria are based on the mean absolute error (MAE) and root mean square error (RMSE), which are evaluated using data on temperature, rainfall, and relative humidity in the provinces of Ubon Ratchathani, Nakhon Ratchasima, Roi Et, Surin, and Si Sa Ket, over a period of 84 months. The training data consists of 72 months from January 2016 to December 2021 and is used to test all four prediction models. The test data, consisting of 12 months from January to December 2022, is used to validate the accuracy of the selected models. The results of the research showed that the Holt-Winters exponential smoothing with additive method adjustments provided the highest prediction accuracy

for temperature, rainfall, and relative humidity with root mean square error (RMSE) equal to 0.86, 102.71, and 6.87, respectively.

Keywords: Time Series Analysis, Holt-Winters Exponential Smoothing, Linear Regression Analysis, Box-Jenkins, Climate Change.

1. บทนำ

ประเทศไทยเป็นประเทศที่มีการทำเกษตรกรรมเป็นจำนวนมาก ทั้งด้านการปลูกพืชผล ด้านการปศุสัตว์ ด้านการประมงและด้านการเกษตรแบบผสมผสาน ซึ่งแต่ละด้านก็สร้างผลผลิตรวมทั้งทำรายได้ให้กับประเทศไทยมาโดยตลอด และผลผลิตหลักที่สำคัญก็คือข้าว ในปี พ.ศ. 2564 ประเทศไทยมีพื้นที่การปลูกข้าวรวมทั่วประเทศจำนวน 53.12 ล้านไร่ โดยพื้นที่เพาะปลูกข้าวมากที่สุดคือ ภาคตะวันออกเฉียงเหนือ พื้นที่ 34.63 ล้านไร่ คิดเป็นร้อยละ 65.2 ของทั้งประเทศ จังหวัดที่มีพื้นที่เพาะปลูกข้าวมากที่สุด 5 อันดับแรก ได้แก่ จังหวัดอุบลราชธานี นครราชสีมา ร้อยเอ็ด สุรินทร์ และศรีสะเกษ [1] ปัจจัยที่สำคัญต่อการปลูกข้าวก็คือฤดูกาล เพราะการปลูกข้าวปีจะทำในช่วงฤดูฝน จะเริ่มตั้งแต่เดือนพฤษภาคมถึงเดือนตุลาคม และเก็บเกี่ยวไม่เกินเดือนกุมภาพันธ์ การทำนาปีต้องอาศัยน้ำจากฝนเพื่อใช้ทำนา ซึ่งฤดูฝนในแต่ละปีก็อาจจะมีความแตกต่างกันออกไป สาเหตุมาจากการเปลี่ยนแปลงของสภาพอากาศที่มีความรุนแรงมากกว่าในอดีต ในบางปีอาจมีช่วงของฤดูฝนที่สั้นกว่าปกติ บางปีก็อาจจะมีช่วงของฤดูฝนที่ยาวนานกว่าปกติ อีกทั้งยังมีเรื่องของสภาวะความเครียดของข้าวที่มาจากสิ่งไม่มีชีวิต เช่น อุณหภูมิ ความชื้นในอากาศ แสง เป็นต้น ซึ่งเป็นปัจจัยสำคัญอย่างหนึ่งที่ทำให้ต้นข้าวได้รับความเสียหายและตายในที่สุด ทำให้ส่งผลกระทบต่อปริมาณของผลผลิตได้ [2]

ปัจจุบันเทคโนโลยีต่าง ๆ ได้มีการพัฒนาและมีความก้าวหน้าเป็นอย่างมาก ทำให้เกิดการนำเทคโนโลยีเข้ามาประยุกต์ใช้ในการแก้ปัญหาในหลาย ๆ เรื่อง หากการพยากรณ์ให้ค่าความแม่นยำสูงจะสามารถช่วยให้เพิ่มประสิทธิภาพในการวางแผนการดำเนินงานได้มากขึ้น

โดยมีงานวิจัยที่ผ่านมาที่มีการวิจัยในการพยากรณ์ต่าง ๆ เช่น ตัวแบบการพยากรณ์มูลค่าการส่งออกอัญมณีและเครื่องประดับของประเทศไทย [3] ตัวแบบพยากรณ์ราคามังคุดคละ [4] การพยากรณ์ปริมาณน้ำฝนของจังหวัดนครราชสีมา [5] การเปรียบเทียบวิธีพยากรณ์จำนวนบัญชีผู้ใช้บริการธนาคารทางอินเทอร์เน็ตด้วยวิธีของ บอซ-เจนกินส์ และวิธีของโฮลต์ [6] เป็นต้น

จากการศึกษาข้างต้น จึงทำให้ผู้วิจัยมีความคิดที่จะนำข้อมูลสภาพภูมิอากาศในประเทศไทยต่าง ๆ ตัวอย่างเช่น อุณหภูมิ ความชื้นในอากาศ และปริมาณน้ำฝน มาศึกษาและวิเคราะห์เพื่อพยากรณ์การเปลี่ยนแปลงของสภาพภูมิอากาศในพื้นที่เพาะปลูกข้าวทางภาคตะวันออกเฉียงเหนือของประเทศไทย สำหรับวางแผน ตัดสินใจในด้านของการเพาะปลูกข้าวของเกษตรกรทางภาคตะวันออกเฉียงเหนือของประเทศไทยได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีการพยากรณ์โดยเทคนิคการวิเคราะห์อนุกรมเวลา

อนุกรมเวลา คือ ลำดับของข้อมูลที่ถูกเรียงตามช่วงเวลา โดยข้อมูลจะมีความต่อเนื่องกันตามช่วงเวลา เกิดขึ้นในอดีตมาจนถึงเวลาปัจจุบันและเรียงลำดับข้อมูลตามเวลาที่เกิดขึ้น ซึ่งข้อมูลประเภทนี้จะให้ความสำคัญกับลำดับเวลาของข้อมูลที่เกิดขึ้น ข้อมูลที่มีลักษณะเป็นอนุกรมเวลามีอยู่เป็นจำนวนมาก เช่น ในด้านของการเกษตรที่มีการเก็บข้อมูลของผลผลิตและราคาขายในด้านของวิศวกรรมที่มีการเก็บข้อมูลของสัญญาณเสียง สัญญาณไฟฟ้า และแรงดันไฟฟ้า ในด้านของอุตุนิยมวิทยาที่มีการเก็บข้อมูลของความเร็วลม อุณหภูมิ และปริมาณน้ำฝน เป็นต้น การวิเคราะห์อนุกรมเวลา จึงหมายถึงการวิเคราะห์ข้อมูลต่าง ๆ ที่เป็นอนุกรมเวลาโดยใช้เทคนิคทางสถิติ [7]

2.2 วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโฮลต์-วินเทอร์เชิงบวก (Holt-Winters exponential smoothing additive Method)

วิธีการปรับให้เรียบแบบโฮลต์-วินเทอร์เชิงบวก เป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีของแนวโน้ม

และฤดูกาลที่ไม่เปลี่ยนแปลงตามเวลาที่เปลี่ยนแปลงไป เป็นวิธีการที่พัฒนามาจากวิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียล มีการใช้ค่าปรับให้เรียบ α β และ γ มีค่าอยู่ในช่วงระหว่าง 0 ถึง 1 [8]

$$\begin{aligned}\hat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\ l_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta * (l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}\end{aligned}$$

โดย \hat{y}_{t+h} = ค่าพยากรณ์ที่ช่วงเวลา $t+h$
 l_t = ค่าระดับของข้อมูลที่ช่วงเวลา t
 b_t = ค่าแนวโน้มของข้อมูลที่ช่วงเวลา t
 s_t = ค่าฤดูกาลของข้อมูลที่ช่วงเวลา t

2.3 วิธีการปรับให้เรียบเอกซ์โพเนนเชียลแบบโฮลท์-วินเทอร์เชิงคูณ (Holt-Winters exponential smoothing Multiplicative Method)

วิธีการปรับให้เรียบแบบโฮลท์-วินเทอร์เชิงคูณ เป็นวิธีการที่พัฒนามาจากวิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียล เป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีแนวโน้มและฤดูกาลที่มีการเพิ่มขึ้นตามเวลาที่เปลี่ยนแปลงไป มีการใช้ค่าปรับให้เรียบ α β และ γ มีค่าอยู่ในช่วงระหว่าง 0 ถึง 1 [8]

$$\begin{aligned}\hat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\ l_t &= \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma \frac{y_t}{(l_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m}\end{aligned}$$

โดย \hat{y}_{t+h} = ค่าพยากรณ์ที่ช่วงเวลา $t+h$
 l_t = ค่าระดับของข้อมูลที่ช่วงเวลา t
 b_t = ค่าแนวโน้มของข้อมูลที่ช่วงเวลา t
 s_t = ค่าฤดูกาลของข้อมูลที่ช่วงเวลา t

2.4 วิธีการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Method)

วิธีการวิเคราะห์การถดถอยเชิงเส้นถูกนำมาใช้ในการวิเคราะห์การถดถอยที่ตัวแปรอิสระส่วนใหญ่เป็นตัวแปรเชิงปริมาณ และตัวแปรตามจะต้องเป็นตัวแปรเชิงปริมาณเท่านั้น ความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามสามารถใช้สมการทางคณิตศาสตร์ที่เป็นเชิงเส้น [8]

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

โดย β_0, β_1 = เป็นค่าคงที่และสัมประสิทธิ์หน้าตัวแปรอิสระ
 x = ค่าของตัวแปรอิสระในรูปคะแนนดิบ และคะแนนมาตรฐาน

ε = ค่าความคลาดเคลื่อน

2.5 วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins Method)

ในการพยากรณ์ระยะสั้น วิธีบ็อกซ์-เจนกินส์ ให้ความสำคัญในการพยากรณ์ค่อนข้างสูงกว่าวิธีอื่น ๆ ซึ่งมีส่วนประกอบของแนวโน้มและค่าความผันแปรตามฤดูกาล มีตัวแบบในรูปทั่วไป คือ SARIMA(p, d, q)(P, D, Q)_s แต่ถ้ากรณีที่ไม่มีความผันแปรตามฤดูกาล จะสามารถลดรูปให้เหลือแค่ ARIMA(p, d, q) [8]

2.6 งานวิจัยที่เกี่ยวข้อง

คชินทร์ [9] ได้ทำการศึกษาการพยากรณ์ราคาขายทองคำแท่ง เพื่อเปรียบเทียบประสิทธิภาพการวิเคราะห์อนุกรมเวลาจำนวน 5 วิธี ได้แก่ วิธี ARIMA, Simple Exponential Smoothing, Holt Exponential Smoothing, Brown's Exponential Smoothing และวิธี Damped Trend Exponential Smoothing ข้อมูลที่ใช้เป็นข้อมูลราคาทองคำ ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2562 วัดประสิทธิภาพการพยากรณ์ด้วยค่า RMSE และค่า MAPE ผลการวิจัยพบว่า วิธี ARIMA ให้ประสิทธิภาพสูงสุด รองลงมา คือ วิธี Simple Exponential Smoothing

ธนกร และสมชาย [10] ได้ทำการศึกษาวินิจฉัย การพยากรณ์ปริมาณน้ำฝนรายเดือนในจังหวัดพัทลุง ณ ศูนย์อุทกวิทยาภาคใต้ โดยทำการเปรียบเทียบด้วยเทคนิคการพยากรณ์ 2 วิธี คือ วิธีบ็อกซ์-เจนกินส์ และวิธีการทำให้เรียบแบบเอกซ์โพเนนเชียลที่มีฤดูกาลอย่างง่าย ซึ่งพิจารณา จากค่า MAPE และค่า RMSE ที่ต่ำที่สุด ศึกษาข้อมูลปริมาณน้ำฝนรายเดือนในจังหวัดพัทลุง ปี พ.ศ. 2546-2548 เพื่อพยากรณ์ปริมาณน้ำฝนรายเดือนตั้งแต่เดือนมกราคม - กันยายน 2549 ผลการวิจัยพบว่า วิธีการทำให้เรียบแบบเอกซ์โพเนนเชียลที่มีฤดูกาลอย่างง่าย เป็นวิธีที่เหมาะสมกับข้อมูลชุดนี้มากที่สุด

3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อการศึกษา เปรียบเทียบ และคัดเลือกแบบจำลองที่เหมาะสมกับการพยากรณ์

การเปลี่ยนแปลงสภาพอากาศในพื้นที่เพาะปลูกข้าวทางภาคตะวันออกเฉียงเหนือของประเทศไทย โดยมีรายละเอียดและขั้นตอนการดำเนินงานดังนี้

3.1 การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล

ในการศึกษาค้นคว้าได้ทำการเก็บรวบรวมข้อมูลสภาพอากาศแบบรายเดือน จากเว็บไซต์กรมอุตุนิยมวิทยา ประกอบด้วย อุณหภูมิ ปริมาณน้ำฝน และความชื้นในอากาศ ตั้งแต่เดือนมกราคม 2559 - ธันวาคม 2565 รวม 84 เดือนของจังหวัดอุบลราชธานี นครราชสีมา ร้อยเอ็ด สุรินทร์ และศรีสะเกษ โดยมีการจัดการกับข้อมูลที่มีค่าว่างในบางเดือนด้วยค่าเฉลี่ยของเดือนเดียวกันในปีอื่น ๆ แล้วจึงปรับรูปแบบของข้อมูลให้เหมาะสมกับแบบจำลองที่จะใช้ในการทดสอบ จากนั้นทำการแบ่งข้อมูลออกเป็น 2 ชุด โดยชุดแรกจะเป็นข้อมูลตั้งแต่เดือนมกราคม 2559 - ธันวาคม 2564 สำหรับใช้ในการสร้างแบบจำลองพยากรณ์และเปรียบเทียบประสิทธิภาพในการพยากรณ์ ส่วนชุดที่สองจะเป็นข้อมูลตั้งแต่เดือนมกราคม - ธันวาคม 2565 สำหรับใช้ในการทดสอบความถูกต้องของแบบจำลอง

3.2 การสร้างแบบจำลองพยากรณ์

นำข้อมูลสภาพอากาศในช่วงตั้งแต่เดือนมกราคม 2559 ถึงเดือนธันวาคม 2564 (Training Set) มาสร้างแบบจำลองโดยใช้เทคนิคการวิเคราะห์อนุกรมเวลา (Time Series Analysis) 4 วิธี ได้แก่ วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงคูณ วิธีการวิเคราะห์การถดถอยเชิงเส้น และวิธีบอซซ์-เจนกินส์

3.3 การเปรียบเทียบประสิทธิภาพและคัดเลือกแบบจำลอง

เมื่อได้ผลจากการสร้างแบบจำลองแล้ว ก่อนที่จะนำแบบจำลองไปใช้ จะต้องมีการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 4 วิธี และจะเลือกใช้แบบจำลองที่ให้ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) ที่ต่ำที่สุด

3.4 การทดสอบความถูกต้องของแบบจำลอง

ทดสอบความถูกต้องของแบบจำลองที่เลือกโดยนำข้อมูลสภาพอากาศตั้งแต่เดือนมกราคม - ธันวาคม 2565 มา

พิจารณาตารางที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)

3.5 การนำเสนอข้อมูลการพยากรณ์

นำเสนอข้อมูลการพยากรณ์การเปลี่ยนแปลงสภาพอากาศในพื้นที่เพาะปลูกข้าวทางภาคตะวันออกเฉียงเหนือของประเทศไทย ในรูปแบบ Dashboard ด้วยโปรแกรม Power BI Desktop

4. ผลการดำเนินงาน

จากการดำเนินการตามขั้นตอนแล้ว จะได้แบบจำลองที่เหมาะสมกับการพยากรณ์การเปลี่ยนแปลงสภาพอากาศในพื้นที่เพาะปลูกข้าวทางภาคตะวันออกเฉียงเหนือของประเทศไทย แสดงผลรายละเอียดการวิจัยได้ดังนี้

4.1 ผลการสร้างแบบจำลองพยากรณ์

ผลการสร้างแบบจำลองโดยเทคนิคการวิเคราะห์อนุกรมเวลา จำนวน 4 วิธี ได้แก่ วิธีการถดถอยเชิงเส้น วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงคูณ วิธีการถดถอยเชิงเส้น และวิธีบอซซ์-เจนกินส์ ผลการวิจัยพบว่าค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) มีดังนี้

ตารางที่ 1: ผลประสิทธิภาพวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงบวก (HWA)

HWA	Temperature		Precipitation		Humidity	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
นครราชสีมา	0.67	0.96	43.86	62.60	2.26	2.52
ร้อยเอ็ด	0.77	0.99	16.16	29.68	3.78	4.10
ศรีสะเกษ	0.71	0.98	51.38	87.27	3.65	4.45
สุรินทร์	1.87	2.04	71.35	117.66	9.07	12.94
อุบลราชธานี	0.83	1.05	45.81	51.68	1.28	2.08
Avg.	0.97	1.20	45.71	59.78	4.01	5.22

ตารางที่ 2: ผลประสิทธิภาพวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เชิงคูณ (HWM)

HWM	Temperature		Precipitation		Humidity	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
นครราชสีมา	0.69	0.96	75.95	96.57	3.46	3.94
ร้อยเอ็ด	0.86	1.06	25.36	30.65	6.50	7.80
ศรีสะเกษ	0.79	1.02	60.81	95.43	4.34	5.53
สุรินทร์	1.99	2.16	77.09	104.58	10.57	14.96

อุบลราชธานี	0.85	1.07	125.03	178.32	1.49	2.10
Avg.	1.04	1.25	72.85	101.11	5.27	6.86

ตารางที่ 3: ผลประสิทธิภาพวิธีการวิเคราะห์การถดถอยเชิงเส้น (LR)

LR	Temperature		Precipitation		Humidity	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
นครราชสีมา	1.89	2.27	74.22	93.15	6.62	7.89
ร้อยเอ็ด	2.19	2.62	85.96	97.91	8.47	10.66
ศรีสะเกษ	1.98	2.45	132.37	144.09	7.90	9.82
สุรินทร์	1.63	2.13	97.89	107.89	12.74	15.39
อุบลราชธานี	2.12	2.50	112.82	129.63	7.09	8.07
Avg.	1.96	2.39	100.65	114.54	8.57	10.37

ตารางที่ 4: ผลประสิทธิภาพวิธีบอซ-เจนกินส์ (ARIMA)

ARIMA	Temperature		Precipitation		Humidity	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
นครราชสีมา	0.74	1.03	44.78	64.11	3.19	3.55
ร้อยเอ็ด	1.08	1.20	40.24	60.99	6.48	7.62
ศรีสะเกษ	0.85	1.08	47.26	85.76	4.86	5.72
สุรินทร์	4.21	4.30	80.87	137.85	10.42	14.90
อุบลราชธานี	0.94	1.18	73.08	105.85	1.49	2.39
Avg.	1.56	1.76	57.24	90.91	5.29	6.84

จากตารางที่ 1-4 ซึ่งเป็นผลการวิเคราะห์แบบจำลองทั้ง 4 วิธีของทั้ง 5 จังหวัดที่ทำการศึกษา จากนั้นผู้วิจัยได้หาค่าเฉลี่ยรวมในแต่ละวิธีเพื่อใช้สำหรับเปรียบเทียบและคัดเลือกแบบจำลองที่เหมาะสมต่อไป

4.2 ผลการเปรียบเทียบประสิทธิภาพและคัดเลือกแบบจำลอง

จากผลการวิจัย นำค่า RMSE ของวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ เชิงบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ เชิงคูณ วิธีการวิเคราะห์การถดถอยเชิงเส้น และวิธีบอซ-เจนกินส์ มาเปรียบเทียบเพื่อหาแบบจำลองที่เหมาะสม โดยพิจารณาจากค่าความคลาดเคลื่อนที่ต่ำที่สุด รายละเอียด แสดงดังตารางที่ 5

ตารางที่ 5: การเปรียบเทียบความแม่นยำของแบบจำลองพยากรณ์

Methods	Temperature		Precipitation		Humidity	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
HWA	0.97	1.20	45.71	59.78	4.01	5.22
HWM	1.04	1.25	72.85	101.11	5.27	6.86
LR	1.96	2.39	100.65	114.54	8.57	10.37
ARIMA	1.56	1.76	57.24	90.91	5.29	6.84

จากตารางที่ 5 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองพยากรณ์ทั้ง 4 วิธี พบว่า แบบจำลองการปรับ

ให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์เชิงบวก ให้ประสิทธิภาพในการพยากรณ์สูงสุด โดยมีค่า MAE และค่า RMSE ต่ำที่สุดเมื่อเทียบกับวิธีอื่น

4.3 การทดสอบความถูกต้องของแบบจำลอง

ผลการทดสอบความถูกต้องของแบบจำลอง โดยใช้ข้อมูลตั้งแต่เดือนมกราคม - ธันวาคม 2565 รวม 12 เดือนของจังหวัดอุบลราชธานี นครราชสีมา ร้อยเอ็ด สุรินทร์ และศรีสะเกษ จากการพิจารณาค่าความคลาดเคลื่อนของแบบจำลองพยากรณ์ที่ให้ผลลัพธ์ที่ดีที่สุด ซึ่งก็คือ วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ เชิงบวก พบว่า การพยากรณ์อุณหภูมิ ปริมาณน้ำฝน และความชื้นในอากาศ มีค่าความคลาดเคลื่อนเฉลี่ยอยู่ที่ 0.86, 102.71 และ 6.87 ตามลำดับ รายละเอียดแสดงดังตารางที่ 6

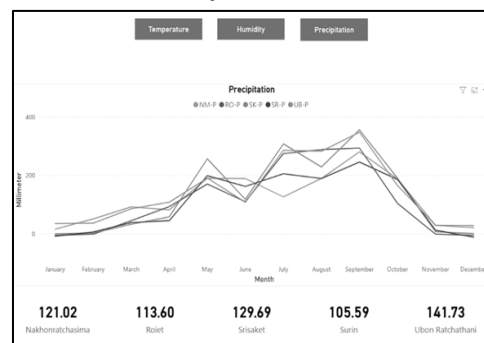
ตารางที่ 6: ผลการทดสอบความถูกต้องของแบบจำลอง

Province	Temperature	Precipitation	Humidity
นครราชสีมา	0.72	76.22	4.45
ร้อยเอ็ด	0.89	77.60	6.32
ศรีสะเกษ	0.86	56.00	5.64
สุรินทร์	0.91	142.90	11.70
อุบลราชธานี	0.93	160.82	6.23
RMSE	0.86	102.71	6.87

จากตารางที่ 6 แสดงให้เห็นถึงค่าความคลาดเคลื่อนของข้อมูลปริมาณน้ำฝนของจังหวัดสุรินทร์ และจังหวัดอุบลราชธานีที่มีค่าสูงกว่าจังหวัดอื่น และจังหวัดสุรินทร์ยังมีค่าความคลาดเคลื่อนของข้อมูลความชื้นในอากาศสูงสุด

4.4 การนำเสนอข้อมูล

นำเสนอผลการพยากรณ์ที่ได้จากแบบจำลองการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ เชิงบวก โดยใช้โปรแกรม Power BI Desktop สำหรับใช้ในการแสดงผลการพยากรณ์ในรูปแบบรายงาน แสดงดังภาพที่ 1



ภาพที่ 1: ตัวอย่างแผนภูมิแสดงข้อมูลสภาพอากาศ

จากภาพที่ 1 แสดงตัวอย่างการเปรียบเทียบข้อมูลสภาพอากาศ (ปริมาณน้ำฝน) ตั้งแต่เดือนมกราคม - ธันวาคม 2565 รวม 12 เดือน กับข้อมูลการพยากรณ์ที่ได้มาจากแบบจำลองการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ เซิงบวก

5. สรุปผลและข้อเสนอแนะ

งานวิจัยนี้ได้ศึกษาการพยากรณ์โดยใช้เทคนิคการวิเคราะห์อนุกรมเวลาเพื่อให้ได้วิธีการพยากรณ์ที่เหมาะสมในการพยากรณ์ข้อมูลสภาพอากาศของจังหวัดอุบลราชธานี นครราชสีมา ร้อยเอ็ด สุรินทร์ และศรีสะเกษ ผู้วิจัยได้เลือกวิธีการพยากรณ์ จำนวน 4 วิธี ได้แก่ วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เซิงบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เซิงคุณ วิธีการวิเคราะห์การถดถอยเชิงเส้น และวิธีบอซซ์-เจนกินส์ ทดสอบประสิทธิภาพของแบบจำลองพยากรณ์ด้วยค่า MAE และ RMSE จากการวิจัยพบว่า แบบจำลองการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์เซิงบวก มีความแม่นยำที่สุด เนื่องจากมีค่าความคลาดเคลื่อนน้อยที่สุด และหลังจากนำข้อมูลที่เกิดขึ้นจริงเปรียบเทียบกับข้อมูลที่ได้ออกมาจากการพยากรณ์ พบความคลาดเคลื่อนเล็กน้อยในการพยากรณ์ของข้อมูลอุณหภูมิ ปริมาณน้ำฝน และความชื้นในอากาศอยู่ที่ 0.86, 102.71 และ 6.87 ตามลำดับ ผลการวิจัยนี้สามารถนำไปเป็นแนวทางในการวางแผนด้านเกษตรกรรม และต่อขั้นตอนในการพัฒนาระบบสำหรับการสนับสนุนการตัดสินใจในอนาคตได้

จากการวิจัยครั้งนี้ข้อมูลที่นำมาใช้สร้างแบบจำลองพยากรณ์ มีจำนวน 72 เดือน และในบางจังหวัดมีข้อมูลที่ไม่สมบูรณ์ หากได้ข้อมูลที่สมบูรณ์และมีจำนวนมากขึ้น จะช่วยให้สามารถสร้างแบบจำลองพยากรณ์ที่มีความแม่นยำสูงขึ้นได้ ส่วนข้อมูลปริมาณน้ำฝนที่มีค่าความคลาดเคลื่อนสูงเนื่องจากมีเดือนที่ฝนตกและฝนไม่ตก จึงทำให้มีความแตกต่างของข้อมูลสูงที่ส่งผลกระทบต่อพยากรณ์ ในอนาคตอาจใช้วิธีการพยากรณ์แบบผสมในการพยากรณ์ข้อมูลปริมาณน้ำฝนเพื่อให้มีความแม่นยำมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] สำนักงานเศรษฐกิจการเกษตร. (2562). เนื้อที่เพาะปลูก ผลผลิต และผลผลิตต่อไร่ รายจังหวัด ปีเพาะปลูก 2564/65 สืบค้นเมื่อวันที่ 20 สิงหาคม 2565 จาก <https://www.oae.go.th/view/1/ตารางแสดงรายละเอียดจำนวนปี/TH-TH>.
- [2] อินทิตรา ชูคแก้ว. (2560). ความเครียดของพืช (plant stress) สืบค้นเมื่อวันที่ 20 สิงหาคม 2565 จาก <http://biology.ipst.ac.th/?p=3361>.
- [3] สมฤดี พงษ์เสนา, กัญญา บวรโชคชัย และอรุวรรณ รุ่งทอง. “ตัวแบบการพยากรณ์มูลค่าการส่งออกอัญมณีและเครื่องประดับของประเทศไทย.” *Journal of Business Management, Burapha University* ปีที่ 2 ฉบับที่ 9 ก.ค.-ธ.ค. 2563 หน้า 66-85.
- [4] วรางคณา เรียงสุทธิ. “ตัวแบบพยากรณ์ราคามังคุดคลด.” *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยอุบลราชธานี* ปีที่ 19 ฉบับที่ 2 พ.ค.-ส.ค. 2560 หน้า 31-42.
- [5] วราฤทธิ พานิชกิจโกศลกุล. “การพยากรณ์ปริมาณน้ำฝนของจังหวัดนครราชสีมา.” *วารสารวิจัยและพัฒนา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี* ปีที่ 28 ฉบับที่ 2 เม.ย.-มิ.ย. 2548 หน้า 155-167.
- [6] กัญญา บวรโชคชัย, สมฤดี พงษ์เสนา และฑาลิสานี นิยมณี. “การเปรียบเทียบวิธีการพยากรณ์จำนวนบัญชีผู้ใช้บริการธนาคารทางอินเทอร์เน็ตด้วยวิธีของ บอซซ์-เจนกินส์ และวิธีของไฮลท์.” *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม* ปีที่ 38 ฉบับที่ 4 มิ.ย.-ส.ค. 2562 หน้า 437-443.
- [7] William W. S. Wei, *Time Series Analysis Univariate and Multivariate Methods*, Pearson Addison Wesley: Boston, 2006.
- [8] Hyndman, R.J., & Athanasopoulos, G. *Forecasting: principles and practice, 2nd edition*, OTexts: Melbourne, 2018.
- [9] ลชินทร์ โกลนุทนกรณ์. “การเปรียบเทียบวิธีการพยากรณ์ราคาขายทองคำแท่ง.” *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏยะลา* ปีที่ 5 ฉบับที่ 1 ม.ค.-มิ.ย. 2563 หน้า 1-9.
- [10] ธนากร โชติช่วง และสมชาย เล็กเจริญ. “การพยากรณ์ปริมาณน้ำฝนรายเดือนในจังหวัดพัทลุง ณ ศูนย์อุทกวิทยา ภาคใต้” *การประชุมวิชาการระดับชาติ วิทยาลัยนครราชสีมา* ปีที่ 5 มี.ค. 2561 หน้า 206-271.

**การพัฒนาระบบธุรกิจอัจฉริยะเพื่อพยากรณ์ความต้องการวัตถุดิบสมุนไพร
กรณีศึกษามูลนิธิโรงพยาบาลเจ้าพระยาอภัยภูเบศรฯ
Study and Development of Business Intelligence System for
Forecasting the Demand for Herb Raw Materials, A Case Study of
the Chaophya Abhaibhubejhr Hospital Foundation**

ภษิต ทองอ่อน (Pasit Thong-on)¹ และมมเชียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857239@email.kmutnb.ac.th, ²montean.r@itd.kmutnb.ac.th

บทคัดย่อ

บทความวิจัยฉบับนี้เป็นการนำเสนอการพัฒนาระบบธุรกิจอัจฉริยะเพื่อพยากรณ์ความต้องการวัตถุดิบสมุนไพรกรณีศึกษามูลนิธิโรงพยาบาลเจ้าพระยาอภัยภูเบศรฯ โดยทำการเปรียบเทียบประสิทธิภาพจำนวน 3 เทคนิค ได้แก่ เทคนิค (1) การเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (2) การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว และ (3) การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังสองครั้ง โดยข้อมูลที่ใช้เป็นข้อมูลข้อมูลปริมาณการใช้วัตถุดิบสมุนไพร ตั้งแต่เดือน มกราคม พ.ศ. 2556 ถึงเดือน ธันวาคม พ.ศ. 2565 แยกเป็น ผงฟ้าทะลายโจร ข้อมูล 1975 แถว มะขามป้อมขึ้นแห้ง ข้อมูล 2865 แถว อัญชันขึ้นแห้ง ข้อมูล 2011 แถว จำนวน 17 คุณลักษณะ เพื่อศึกษาคุณลักษณะของตัวแปรที่มีผลต่อการพยากรณ์ปริมาณการใช้วัตถุดิบ ประเมินประสิทธิภาพของเทคนิค โดยพิจารณาค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD) ค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE) ทั้งนี้ผลการพยากรณ์พบว่า เทคนิคการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว ให้ประสิทธิภาพการพยากรณ์สูงที่สุด ผงฟ้าทะลายโจร MAD 0.0928 MSE 0.0324 มะขามป้อมขึ้นแห้ง MAD 0.1166 MSE 0.0244 อัญชันขึ้นแห้ง MAD 0.1459 MSE 0.0371

คำสำคัญ: เทคนิคการเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก

การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว

การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังสองครั้ง

Abstract

This research article presents the development of an intelligent business system for predicting the demand for herbal raw materials, using the case study of the Chao Phya Abhaibhubejhr Hospital Foundation. The efficiency of three techniques is compared: (1) weighted moving average, (2) single exponential smoothing, and (3) double exponential smoothing. The data used is the quantity of herbal raw materials used from January 2013 to December 2022, divided into three types: 1,975 rows of Fa Thalai Jone powder data, 2,865 rows of dried Makhom Pom, and 2,011 rows of dried Anchan data, with 17 characteristics to study the variables affecting the prediction of the quantity of raw materials used. The effectiveness of the techniques is evaluated by considering the average absolute deviation (MAD) and mean square error (MSE). The results show that the technique of single exponential smoothing provides the highest prediction efficiency. For Fa Thalai Jone powder, the MAD is 0.0928 and the MSE is 0.0324. For dried Makhom Pom, the MAD is 0.1166 and the MSE is 0.0244. For dried Anchan, the MAD is 0.1459 and the MSE is 0.0371.

Keywords: Weighted Moving Average, Single Exponential Smoothing, Double Exponential Smoothing.

1. บทนำ

มูลนิธิโรงพยาบาลเจ้าพระยาอภัยภูเบศรฯ เป็นผู้ผลิตและจำหน่าย ผลิตภัณฑ์สมุนไพรที่มีความหลากหลายทั้งยาสมุนไพรแผนโบราณเครื่องสำอางซึ่งทุกผลิตภัณฑ์ล้วนแล้วแต่ใช้วัตถุดิบสมุนไพรทั้งสิ้น ส่งผลให้ต่อไปจะมีการรับเข้าวัตถุดิบสมุนไพรเป็นจำนวนมาก และวัตถุดิบแต่ละชนิดมีระยะเวลาในการเก็บเกี่ยวที่ต่างกัน อีกทั้งปริมาณการใช้วัตถุดิบแต่ละชนิดที่ไม่เท่ากัน จึงทำให้วัตถุดิบบางชนิดไม่เพียงพอต่อการนำไปผลิตเป็นสินค้าสำเร็จรูป มูลนิธิโรงพยาบาลเจ้าพระยาอภัยภูเบศรฯ จึงตระหนักถึงความสำคัญในการใช้ระบบเทคโนโลยีสารสนเทศเข้ามาช่วยจัดการพยากรณ์ความต้องการวัตถุดิบสมุนไพร เพื่อช่วยในการตัดสินใจของผู้บริหารระดับกลางและระดับสูง

การวิจัยครั้งนี้ผู้วิจัยทำการศึกษาและรวบรวมข้อมูลงานวิจัยเกี่ยวกับการพยากรณ์ทั้งการพยากรณ์ความต้องการของวัตถุดิบ การพยากรณ์ราคา การพยากรณ์ความต้องการสินค้า รวมถึงการใช้อัลกอริทึมเพื่อให้การพยากรณ์มีความถูกต้องแม่นยำมากขึ้น เช่น (1) วิธีบ็อกซ์-เจนกินส์ (Box-Jenkins) (2) วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังที่ไม่มีแนวโน้มแบบเต็มพิกัด (Damped Trend Non-Seasonal) (3) วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังที่มีแนวโน้มผลบวกแบบเต็มพิกัด (Damped Trend Seasonal Additive) (4) วิธีการปรับเรียบด้วยเส้นโค้งเลขชี้กำลังที่มีแนวโน้มผลคูณแบบเต็มพิกัด (Damped Trend Seasonal Multiplicative) (5) วิธีปรับเรียบด้วยการถ่วงเฉลี่ยเคลื่อนที่อย่างง่าย (Simple Moving Average) (6) วิธีปรับเรียบด้วยการถ่วงเฉลี่ยเคลื่อนที่สองครั้ง (Double Moving Average) (7) วิธีปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว (Single Exponential Smoothing) (8) วิธีปรับเรียบด้วยเส้นโค้งเลขชี้กำลังสองครั้ง (Double Exponential Smoothing) (9) วิธีฤดูกาลแบบผลบวก (Seasonal Additive) (10) วิธีฤดูกาลแบบผลคูณ (Seasonal Multiplicative) (11) วิธีปรับเรียบด้วยเส้นโค้งเลขชี้กำลังของโฮลต์วินเทอร์แบบบวก (Holt-Winters'

Additive) (12) วิธีการเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (Weighted Moving Average) เป็นต้น

จากปัญหาที่กล่าวมาข้างต้น ผู้วิจัยมีแนวคิดในการพยากรณ์ข้อมูลความต้องการวัตถุดิบสมุนไพร โดยใช้ อัลกอริทึม (1) การเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (Weighted Moving Average) (2) การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว (Single Exponential Smoothing) (3) การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังสองครั้ง (Double Exponential Smoothing) จากนั้นทำการทดสอบความแม่นยำในการพยากรณ์ แล้วคัดเลือกอัลกอริทึมที่ให้ค่าความแม่นยำที่ดีที่สุดไปใช้ และนำข้อมูลพยากรณ์มาออกแบบและพัฒนาาระบบสารสนเทศในการพยากรณ์ความต้องการวัตถุดิบสมุนไพร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 วิธีการพยากรณ์

2.1.1 การเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (Weighted Moving Average) เป็นการปรับปรุงจาก Moving Average ให้สะท้อนการแปรผัน ได้ใกล้เคียงความเป็นจริงมากขึ้น โดยให้น้ำหนักแต่ละระยะเวลาที่แตกต่างกันไป โดยมีสูตรดังนี้

$$WMA = \frac{\sum[(\text{weight for period } n) \times (\text{demand in period } n)]}{\sum \text{weights}}$$

2.1.2 การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังครั้งเดียว (Single Exponential Smoothing) เป็นวิธีที่ใช้หลักการเดียวกันกับการพยากรณ์แบบค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก นั่นคือให้ความสำคัญกับข้อมูลชุดใหม่ที่สุดมากที่สุด (ค่าถ่วงน้ำหนักสูงสุด) และค่อยๆ ลดค่าถ่วงน้ำหนักลง วิธีเอกซ์โปเนนเชียลปรับเรียบจะทำการพยากรณ์ โดยนำค่าพยากรณ์ของช่วงเวลาที่ผ่านมามวกเข้ากับอัตราส่วนความแตกต่างระหว่างข้อมูลจริงกับค่าพยากรณ์ ณ ช่วงเวลานั้น ๆ โดยมีสูตรดังนี้

$$s_t = \alpha x_t + (1 - \alpha)s_{t-1}$$

2.1.3 การปรับเรียบด้วยเส้นโค้งเลขชี้กำลังสองครั้ง (Double Exponential Smoothing) การพยากรณ์โดยวิธีการปรับเรียบเอกซ์โปเนนเชียลซ้ำ 2 ครั้ง เป็นการใช้หลักการของ Exponential Smoothing มาทำซ้ำกันสองครั้งเหมาะกับข้อมูลที่มีเฉพาะความไม่แน่นอน (Random) เพียงอย่างเดียว (ไม่มีแนวโน้มและฤดูกาล) โดยมีสูตรดังนี้

$$s_t = \alpha x_t + (1 - \alpha)(s_{t-1} + b_{t-1})b_t \\ = \beta(s_t - s_{t-1}) + (1 - \beta)b_{t-1}$$

2.2 งานวิจัยที่เกี่ยวข้อง

ศีลวัชร และประจวบ (2564) ทำการศึกษาการเปรียบเทียบวิธีการพยากรณ์ความต้องการแอลกอฮอล์สำหรับการผลิตกระแสไฟฟ้ากรณีศึกษาโรงไฟฟ้าชีวมวลแห่งหนึ่ง กลุ่มตัวอย่างที่ใช้ศึกษาคือ ข้อมูลปริมาณการสั่งซื้อแอลกอฮอล์ของโรงไฟฟ้าชีวมวลกรณีศึกษาสำหรับผลิตกระแสไฟฟ้าปี 2559-2563 อัลกอริทึมที่ใช้ในการวิเคราะห์มีทั้งหมด 4 ชนิดได้แก่ (1) การหาค่าเฉลี่ยเคลื่อนที่ (Moving Average Method) (2) การปรับเรียบแบบเอกซ์โปเนนเชียลแบบง่าย (Single Exponential Smoothing Method) (3) การปรับเรียบเอกซ์โปเนนเชียลสองเท่า (Double Exponential Smoothing Method) (4) การพยากรณ์แบบถดถอยเชิงเส้นตรง (Linear Regression Analysis) วัดความแม่นยำของแต่ละวิธีโดยเปรียบเทียบ (1) ค่าเฉลี่ยการเบี่ยงเบนสัมบูรณ์ (Mean absolute deviation: MAD) (2) ค่าเฉลี่ยความผิดพลาดยกกำลังสอง (Mean Squared Error: MSE) (3) ค่าเฉลี่ยความผิดพลาดร้อยละสัมบูรณ์ (Mean absolute percentage error: MAPE) ผลที่ได้คือ (1) การหาค่าเฉลี่ยเคลื่อนที่ (Moving Average Method) ได้ค่า MAD 1547, MSE 3479826, MAPE 32.20 (2) การปรับเรียบแบบเอกซ์โปเนนเชียลแบบง่าย (Single Exponential Smoothing Method) ได้ค่า MAD 1489, MSE 3017037, MAPE 28 (3) การปรับเรียบเอกซ์โปเนนเชียลสองเท่า (Double Exponential Smoothing Method) ได้ค่า MAD 1283, MSE 2600611, MAPE 24 (4) การพยากรณ์แบบถดถอยเชิงเส้นตรง (Linear Regression Analysis) ได้ค่า MAD 1534, MSE 3314002, MAPE 27.30 โดยที่การวิเคราะห์ด้วยเทคนิคการปรับเรียบเอกซ์โปเนนเชียลสองเท่า (Double Exponential Smoothing Method) ให้ค่าดีที่สุด

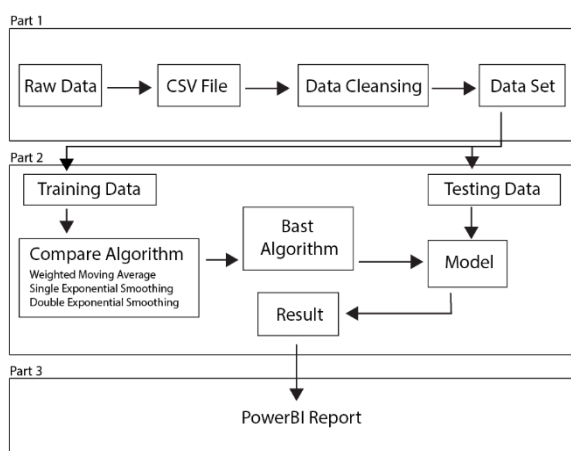
ภาสุระ และคณะ (2561) ทำการศึกษาการเปรียบเทียบวิธีการพยากรณ์ สำหรับโรงงานผลิตชุด กลุ่มตัวอย่างที่ใช้ศึกษาคือ ข้อมูลยอดขายของปี พ.ศ. 2559 (ตั้งแต่เดือนมกราคม-มิถุนายน) อัลกอริทึมที่ใช้ในการวิเคราะห์มีทั้งหมด 4 ชนิดได้แก่ (1) วิธีถัวเฉลี่ยเคลื่อนที่ (Moving Average) (2) วิธีปรับเรียบเอกซ์โปเนนเชียลอย่างง่าย (Simple Exponential Smoothing) (3) วิธีพยากรณ์แบบฤดูกาลของวินเตอร์ (Winter's Linear and Seasonal Exponential Smoothing) (4) วิธีการพยากรณ์โดยการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Regression Analysis) วัดความแม่นยำของแต่ละวิธีโดยเปรียบเทียบ (1) ค่าเฉลี่ยการเบี่ยงเบนสัมบูรณ์ (Mean absolute deviation: MAD) (2) ค่าเฉลี่ยความผิดพลาดยกกำลังสอง (Mean Squared Error: MSE) (3) ค่าเฉลี่ยความผิดพลาดร้อยละสัมบูรณ์ (Mean absolute percentage error: MAPE) ผลที่ได้คือ (1) วิธีถัวเฉลี่ยเคลื่อนที่ (Moving Average) ค่าที่ได้แบ่งได้เป็น 4 ช่วง เฉลี่ยเคลื่อนที่ 3 เดือน MAD 324, MSD 160872, MAPE 58 เฉลี่ยเคลื่อนที่ 5 เดือน MAD 338, MSD 186482, MAPE 70 เฉลี่ยเคลื่อนที่ 7 เดือน MAD 312, MSD 172430, MAPE 80 และเฉลี่ยเคลื่อนที่ 9 เดือน MAD 333, MSD 163772, MAPE 98 (2) วิธีปรับเรียบเอกซ์โปเนนเชียลอย่างง่าย (Simple Exponential Smoothing) ได้ค่า MAD 335, MSD 161543, MAPE 67 (3) วิธีพยากรณ์แบบฤดูกาลของวินเตอร์ (Winter's Linear and Seasonal Exponential Smoothing) ได้ค่า MAD 353, MSD 171245, MAPE 75 (4) วิธีการพยากรณ์โดยการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Regression Analysis) ได้ค่า MAD 422, MSD 317759, MAPE 67 โดยที่วิธีการพยากรณ์แบบถัวเฉลี่ยเคลื่อนที่ 3 เดือน ให้ค่าพยากรณ์ดีที่สุด

ศุภกานต์ และกัปตัน (2564) ทำการศึกษาการพยากรณ์ความต้องการสินค้าเพื่อการวางแผนและการจัดซื้อวัตถุดิบบริษัท โขว้า ออโต้พาร์ท (ประเทศไทย) จำกัด กลุ่มตัวอย่างที่ใช้ศึกษาคือข้อมูลคำสั่งซื้อลูกค้าปี 2562-2563 และแบ่งช่วงเวลาวิเคราะห์เป็น 2 ช่วง คือ 3 เดือนและ 5 เดือน อัลกอริทึมที่ใช้ในการวิเคราะห์มีทั้งหมด 3 ชนิดได้แก่ (1) การเฉลี่ยเคลื่อนที่อย่างง่าย (Simple Moving

Average) (2) การปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) (3) การพยากรณ์แบบวิเคราะห์การถดถอย (Regression Analysis Method) วัดความแม่นยำของแต่ละวิธีโดยเปรียบเทียบ (1) ค่าเฉลี่ยการเบี่ยงเบนสัมบูรณ์ (Mean absolute deviation: MAD) (2) ค่าเฉลี่ยความผิดพลาดยกกำลังสอง (Mean Squared Error: MSE) (3) ค่าเฉลี่ยความผิดพลาดร้อยละสัมบูรณ์ (Mean absolute percentage error: MAPE) ผลที่ได้คือ (1) การเฉลี่ยเคลื่อนที่อย่างง่าย (Simple Moving Average) ค่าเฉลี่ยที่ 3 เดือน ได้ค่า MAD 223.33, MSE 84144.44, MAPE 18.85 ค่าเฉลี่ยที่ 5 เดือน ได้ค่า MAD 252.29, MSE 97516, MAPE 20.05 (2) การปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) ค่าเฉลี่ยที่ 3 เดือน ได้ค่า MAD 239.35, MSE 86344.35, MAPE 21.70 ค่าเฉลี่ยที่ 5 เดือน ได้ค่า MAD 244.55, MSE 101027.27, MAPE 22.77 (3) การพยากรณ์แบบวิเคราะห์การถดถอย (Regression Analysis Method) ค่าเฉลี่ยที่ 3 เดือน ได้ค่า MAD 258.42, MSE 107254.42, MAPE 24.09 ค่าเฉลี่ยที่ 5 เดือน ได้ค่า MAD 196.43, MSE 48119.75, MAPE 17.27 โดยที่วิธีการพยากรณ์แบบวิเคราะห์การถดถอยให้ค่าพยากรณ์ดีที่สุด

3. วิธีการดำเนินการวิจัย

การพยากรณ์ความต้องการวัตถุดิบสมุนไพร ผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากภาพที่ 1 กรอบแนวคิดการทำวิจัยพยากรณ์ความต้องการวัตถุดิบสมุนไพร ประกอบด้วย 3 ส่วนในการทำงานดังนี้

3.1 การรวบรวมข้อมูล (Data Gathering)

การรวบรวมชุดข้อมูลสำหรับการพยากรณ์ ข้อมูลที่ใช้ในการทดสอบเป็นข้อมูลปริมาณการใช้วัตถุดิบสมุนไพร ณิธิโรงพยาบาลเจ้าพระยาอภัยภูเบศรฯ ตั้งแต่เดือนมกราคม พ.ศ. 2556 ถึงเดือนธันวาคม พ.ศ. 2565 แยกเป็นผงฟ้าทะลายโจร มะขามป้อมขึ้นแห้ง อัญชันขึ้นแห้ง ประกอบด้วยคุณลักษณะ จำนวน 3 คุณลักษณะ แสดงดังตารางที่ 1

ตารางที่ 1: ตัวอย่างคุณลักษณะข้อมูล

คุณลักษณะ	ตัวอย่างข้อมูล
วันที่	1/1/2021
ชื่อวัตถุดิบ	มะขามป้อมขึ้นแห้ง
จำนวนที่ใช้	1000

3.2 การเตรียมข้อมูล (Data Preprocessing)

กระบวนการเตรียมข้อมูลเพื่อนำเข้าสู่แบบจำลองประกอบด้วย 3 ขั้นตอน

3.2.1 การทำความสะอาดข้อมูล ได้แก่ การจัดการกับข้อมูลสูญหายหรือค่าว่างโดยการแทนค่าข้อมูลจะใช้วิธีการหาค่าเฉลี่ยปริมาณการใช้วัตถุดิบในปีที่มีค่าว่างและแทนค่าเฉลี่ยที่ได้ลงไป

3.2.2 การคัดเลือกแอททริบิวต์ ทำการคัดเลือกโดยอ้างอิงจากแอททริบิวต์ของงานวิจัยที่เกี่ยวข้องได้จำนวน 3 แอททริบิวต์ แยกเป็น 3 วัตถุดิบ ได้แก่ ผงฟ้าทะลายโจร มีจำนวนข้อมูล 1975 แถว มะขามป้อมขึ้นแห้ง มีจำนวนข้อมูล 2865 แถว อัญชันขึ้นแห้ง มีจำนวนข้อมูล 2011 แถว และมีแอททริบิวต์ ได้แก่ วันที่ ชื่อวัตถุดิบ จำนวนที่ใช้ และลบแอททริบิวต์ที่ไม่ใช้งานออกไป

3.2.3 การแปลงข้อมูล โดยแปลงข้อมูลที่ได้อให้อยู่ในรูปแบบของเดือน และนำข้อมูลมาปรับค่าสัมประสิทธิ์สหสัมพันธ์

3.3 การสร้างแบบจำลอง (Modeling)

อัลกอริทึมที่ใช้ในการสร้างแบบจำลองพยากรณ์ ประกอบด้วย 3 อัลกอริทึม ได้แก่ DES, ES และ WMA และทำการแบ่งข้อมูลตามช่วงเวลาออกเป็น Training 50 เปอร์เซ็นต์ 2556 – 2560 และ Testing 50 เปอร์เซ็นต์ 2561 – 2565 เท่ากัน โดยการดำเนินงานเป็นดังนี้

3.3.1 การกำหนดพารามิเตอร์ที่ใช้สำหรับการทดสอบอัลกอริทึมทั้ง 3 อัลกอริทึม และทำการปรับค่าพารามิเตอร์ตามที่กำหนดไว้ในแต่ละอัลกอริทึม และเปรียบเทียบผลลัพธ์ แสดงดังตารางที่ 2

ตารางที่ 2: พารามิเตอร์สำหรับการทดสอบอัลกอริทึม

อัลกอริทึม	พารามิเตอร์
DES	Alpha, Beta
ES	Alpha
WMA	Weighted

3.4 การประเมินผล (Evaluation)

ทำการวัดประสิทธิภาพแบบจำลองการพยากรณ์ด้วย

- ค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD)
- ค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE)

$$MAD = \frac{\sum_{t=1}^n |A_t - F_t|}{n}$$

$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n}$$

4. ผลการดำเนินงาน

ผลการดำเนินงานแสดงดังตารางที่ 3 ถึง 6 โดยมีรายละเอียดดังนี้

ตารางที่ 3: ผลลัพธ์การวัดค่าความถูกต้องของอัลกอริทึม ES

	ฟังก์ชันหลายโจร		มะขามป้อมขึ้นแห้ง		อัญชันขึ้นแห้ง	
	MAD	MSE	MAD	MSE	MAD	MSE
ES (a,0.3)	0.0933	0.0322	0.1176	0.0246	0.1459	0.0371
ES (a,0.4)	0.0928	0.0324	0.1166	0.0244	0.1471	0.0383
ES (a,0.6)	0.0949	0.0347	0.1171	0.0255	0.1504	0.0409
ES (a,0.9)	0.1030	0.0431	0.1255	0.0310	0.1600	0.0469

จากตารางที่ 3 (1) ฟังก์ชันหลายโจรและ (2) มะขามป้อมขึ้นแห้ง ค่าพารามิเตอร์ alpha = 0.4 ให้

ประสิทธิภาพที่ดีที่สุด ที่ค่า (1) MAD = 0.0928 MSE = 0.0324 , (2) MAD = 0.1166 MSE = 0.0244 ตามลำดับ และ (3) อัญชันขึ้นแห้ง ค่าพารามิเตอร์ alpha = 0.3 ให้ประสิทธิภาพที่ดีที่สุด ที่ค่า (3) MAD = 0.1459 MSE = 0.0371

ตารางที่ 4: ผลลัพธ์การวัดค่าความถูกต้องของอัลกอริทึม DES

	ฟังก์ชันหลายโจร		มะขามป้อมขึ้นแห้ง		อัญชันขึ้นแห้ง	
	MAD	MSE	MAD	MSE	MAD	MSE
DES (a,0.3 b,0.9)	0.1804	0.1043	0.2083	0.0981	0.2369	0.1247
DES (a,0.4 b,0.6)	0.1596	0.0912	0.1840	0.0826	0.2202	0.1065
DES (a,0.6 b,0.4)	0.1429	0.0763	0.1611	0.0646	0.2043	0.0873
DES (a,0.9 b,0.3)	0.1426	0.0754	0.1684	0.0595	0.2107	0.0804
DES (a,0.2 b,0.4)	0.2787	0.3041	0.2910	0.2999	0.3307	0.3128
DES (a,0.4 b,0.2)	0.1990	0.1698	0.2216	0.1637	0.2522	0.1787

จากตารางที่ 4 (1) ฟังก์ชันหลายโจรและ ค่าพารามิเตอร์ alpha = 0.9 beta = 0.3 ให้ประสิทธิภาพที่ดีที่สุด ที่ค่า (1) MAD = 0.1426 MSE = 0.0754 และ (2) มะขามป้อมขึ้นแห้ง (3) อัญชันขึ้นแห้ง ค่าพารามิเตอร์ alpha = 0.6 beta = 0.4 ให้ประสิทธิภาพที่ดีที่สุด ที่ค่า (2) MAD = 0.1611 MSE = 0.0646 (3) MAD = 0.2043 MSE = 0.0873

ตารางที่ 5: ผลลัพธ์การวัดค่าความถูกต้องของอัลกอริทึม WMA

	ฟังก์ชันหลายโจร		มะขามป้อมขึ้นแห้ง		อัญชันขึ้นแห้ง	
	MAD	MSE	MAD	MSE	MAD	MSE
WMA(0.1,0.2 ,0.3)	0.0957	0.0352	0.1184	0.0249	0.1534	0.0428
WMA(0.1,0.2 ,0.4)	0.0968	0.0361	0.1185	0.0255	0.1529	0.0426
WMA(0.1,0.2 ,0.5)	0.0981	0.0369	0.1189	0.0262	0.1530	0.0427
WMA(0.1,0.2 ,0.6)	0.0991	0.0378	0.1194	0.0267	0.1533	0.0430

จากตารางที่ 5 (1) ฟังก์ชันหลายโจรและ (2) มะขามป้อมขึ้นแห้ง ค่าพารามิเตอร์ Weighted = 0.1,0.2,0.3 ให้ประสิทธิภาพที่ดีที่สุด ที่ค่า (1) MAD = 0.0957 MSE = 0.0352 , (2) MAD = 0.1184 MSE = 0.0249 ตามลำดับ และ(3) อัญชันขึ้นแห้ง ค่าพารามิเตอร์ Weighted = 0.1,0.2,0.4 ให้ประสิทธิภาพที่ดีที่สุด ที่ค่า (3) MAD = 0.1529 MSE = 0.0426

จากผลการทดสอบประสิทธิภาพแบบจำลอง 3 แบบจำลอง พบว่า อัลกอริทึม ES ให้ผลลัพธ์ที่ดีที่สุด แสดงดังตารางที่ 6

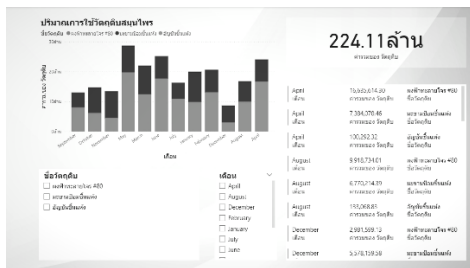
ตารางที่ 6: ตารางเปรียบเทียบผลการพยากรณ์ทั้ง 3 แบบจำลอง

	ผงฟ้าทะเลใจ			มะขามป้อมขึ้นแห้ง	
	MAD	MSE		MAD	MSE
WMA(0.1,0.2,0.3)	0.0957	0.0352	WMA(0.1,0.2,0.3)	0.1184	0.0249
ES (a,0.4)	0.0928	0.0324	ES (a,0.4)	0.1166	0.0244
DES (a,0.9 b,0.3)	0.1426	0.0754	DES (a,0.6 b,0.4)	0.1611	0.0646

	อัญชันแห้ง	
	MAD	MSE
WMA(0.1,0.2,0.4)	0.1529	0.0426
ES (a,0.3)	0.1459	0.0371
DES (a,0.6 b,0.4)	0.2043	0.0873

จากตารางที่ 6 พบว่า การพยากรณ์พยากรณ์ความต้องการวัตถุดิบสมุนไพร ทั้ง 3 ชนิดด้วยเทคนิค ES มีค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD) และค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE) ดีที่สุด

จากผลการพยากรณ์แนวโน้มความต้องการวัตถุดิบสมุนไพร ผู้วิจัยได้ทำการออกแบบรายงานบนโปรแกรม Power BI เพื่อแสดงผลการพยากรณ์แนวโน้มความต้องการวัตถุดิบสมุนไพร แสดงผลดังนี้



5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการเปรียบเทียบประสิทธิภาพแบบจำลองการพยากรณ์ความต้องการวัตถุดิบสมุนไพร จำนวน 3 แอททริบิวต์ แยกเป็น 3 วัตถุดิบ ได้แก่ ผงฟ้าทะเลใจ มีจำนวนข้อมูล 1975 แถว มะขามป้อมขึ้นแห้ง มีจำนวนข้อมูล 2865 แถว อัญชันขึ้นแห้ง มีจำนวนข้อมูล 2011 แถว เลือกเทคนิคที่ใช้จำนวน 3 เทคนิค ได้แก่ เทคนิค ES, DES และ WMA ผลการทดสอบประสิทธิภาพของแบบจำลองด้วยเทคนิค ES มีประสิทธิภาพการพยากรณ์ความต้องการวัตถุดิบสมุนไพรสูงที่สุด (1) ผงฟ้าทะเลใจ ได้ค่า MAD = 0.0928 MSE = 0.0324 (2) มะขามป้อมขึ้นแห้ง ได้ค่า MAD = 0.1166 MSE = 0.0244 และ (3) อัญชันขึ้นแห้ง ได้ค่า MAD = 0.1459 MSE = 0.0371

เอกสารอ้างอิง

- [1] ศीलวัชร แก้วพิจิตร และประจวบ ก่อมจิตร (2564). การเปรียบเทียบวิธีพยากรณ์ความต้องการแลกเปลี่ยนสำหรับการผลิตกระแสไฟฟ้า กรณีศึกษาโรงไฟฟ้าชีวมวลแห่งหนึ่ง. การประชุมวิชาการเสนองานวิจัยระดับชาติ ด้านวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏจันทรเกษม ครั้งที่ 4, 1(1), 282-288.
- [2] ภาสุระ อังกุลานนท์, ชนิกา นิเวศานนท์, ลักขณา ฤกษ์เกษม, ปิยะชัย สุทธิเข้ม, ตนะพันธ์ ภูธานนท์, วิชัย แหวนเพชร และอดิพิงษ์ เทพทับทิม (2561). การเปรียบเทียบวิธีการพยากรณ์ สำหรับโรงงานผลิตชุด. วารสารวิจัยเทคโนโลยีอุตสาหกรรม มจร. พระนคร, 1(1), 37-46.
- [3] สุกกานต์ ช่างสากล และกัปตัน ภูเบศ อยู่สุข (2564). การพยากรณ์ความต้องการสินค้า เพื่อการวางแผนและการจัดซื้อวัตถุดิบ บริษัทโซว่า ออโต้พาร์ท (ประเทศไทย) จำกัด. คณะโลจิสติกส์ มหาวิทยาลัยบูรพา.
- [4] Barzenji, H. (2021). Sentiment analysis of Twitter texts using machine learning algorithms. Academic Platform Journal of Engineering and Science, 9(3), 460-471.
- [5] Frenda, F. (2022). Forecasting System Prototype Amount of Printing Raw Material Inventory Needs Using the Autoregressive Integrated Moving Average (Arima) Method. Jurnal Mantik, 6(2), 1964-1970.
- [6] George C. G., S., and Sumathi, B. (2021). A novel deep learning approach of convolutional neural network and random forest classifier for fine-grained sentiment classification. International Journal on Electrical Engineering and Informatics, 13(2), 465-476.
- [7] Dangi, D., Dixit, D. K., and Bhagat, A. (2022). Sentiment analysis of COVID-19 social media data through machine learning. Multimedia Tools and Applications.
- [8] Kasthuri S., and Nisha Jebaseeli A. (2020). An efficient Decision Tree Algorithm for analyzing the Twitter Sentiment Analysis. Journal of critical reviews, 7(4), 1010-1018.
- [9] Liu, H., Li, C., Shao, Y., Zhang, X., Zhai, Z., Wang, X., Qi, X., Wang, J., Hao, Y., Wu, Q., and Jiao, M. (2020). Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011-2019 using the Seasonal Autoregressive Integrated moving average (sarima) and exponential smoothing (ETS) models. Journal of Infection and Public Health, 13(2), 287-294.

ระบบพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา

Forecasting System for Repayment Through Online Payment Using Time Series Techniques

ยงยุทธ อินทรประสิทธิ์ (Yongyut Intarparasit)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857140@kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบการสร้างระบบการพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ ผ่านการคัดเลือกเทคนิคการวิเคราะห์อนุกรมเวลาโดยแบ่งส่วนออกเป็น 2 ขั้นตอน ขั้นตอนที่ 1 การคัดเลือกแบบจำลองสำหรับการพยากรณ์จำนวน 2 วิธี ได้แก่ วิธีอาร์มา และวิธีโฮลท์-วินเทอร์ ทำการวัดประสิทธิภาพแบบจำลองพยากรณ์ที่มีความแม่นยำที่สุดโดยมีค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุด ข้อมูลที่ใช้เป็นข้อมูลรายการชำระเงินผ่านช่องทางออนไลน์ระหว่างเดือนมีนาคม พ.ศ. 2564 ถึงเดือนพฤศจิกายน พ.ศ. 2565 จำนวน 250 วัน และมีจำนวน 4 ช่องทาง ได้แก่ ช่องทางชำระเงินผ่านเว็บไซต์ ช่องทางชำระเงินด่วน ช่องทางชำระเงินผ่านแอปพลิเคชัน และช่องทางชำระเงินผ่านใบแจ้งหนี้อิเล็กทรอนิกส์ ผลการวิจัยพบว่า แบบจำลองการพยากรณ์ด้วยวิธีโฮลท์-วินเทอร์ ให้ประสิทธิภาพการพยากรณ์แม่นยำสูงที่สุด โดยมีค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดทั้ง 4 ช่องทางข้อมูลคือ 0.104, 0.102, 0.087, 0.159 ตามลำดับ หลังจากนั้นนำวิธีโฮลท์-วินเทอร์ไปใช้ในขั้นตอนที่ 2 คือ การสร้างระบบการพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์

คำสำคัญ: เทคนิคการวิเคราะห์อนุกรมเวลา วิธีอาร์มา
วิธีโฮลท์-วินเทอร์

Abstract

The objective of this research is to develop a process for creating a debt payment prediction system through online payment channels, by selecting time series analysis techniques in two steps. Step 1 involves selecting a predictive model, with two methods available: the ARIMA method and the Holt-Winters method. The most accurate predictive model with the lowest absolute average error is determined. The data used are online payment records from March 2021 to November 2021, totaling 250 days, and there are four payment channels: website payments, express payments, app payments, and e-invoice payments. The research found that the Holt-Winters method provided the highest accuracy for debt payment prediction, with the lowest absolute average error of all four datasets: 0.104, 0.102, 0.087 and 0.159, respectively. Step 2 involves implementing the Holt-Winters method to develop a debt payment prediction system through online payment channels.

Keywords: Time Series, ARIMA, Holt-Winters.

1. บทนำ

ในปัจจุบันเทคโนโลยีเข้ามามีบทบาทสำคัญในการดำเนินกิจการของบริษัท อันเนื่องมาจากผู้บริโภคที่เป็นลูกค้าของบริษัทได้ปรับเปลี่ยนการใช้ชีวิตไปตามเทคโนโลยี ทำให้พฤติกรรมของลูกค้าที่มีการเปลี่ยนแปลง

ส่งผลต่อการดำเนินกิจการของบริษัท เทคโนโลยีเข้ามามีบทบาทต่อการทำธุรกรรมทางการเงิน ทำให้ทางบริษัทพัฒนาช่องทางให้บริการทางการเงิน [1] เช่น การทำธุรกรรมการเงินผ่านเว็บไซต์ การทำธุรกรรมผ่านทางช่องทางควอนตัมออนไลน์ การทำธุรกรรมผ่านทางระบบสมาชิก การทำธุรกรรมผ่านทางแอปพลิเคชัน เป็นต้น ส่งผลให้ทางบริษัทจำเป็นต้องพัฒนาเทคโนโลยีเพื่อตอบสนองความต้องการของลูกค้า และการจัดการระบบงานวางแผนบริหารงานทำงานล่วงหน้าได้อย่างมีประสิทธิภาพ จึงจำเป็นต้องมีการพยากรณ์จำนวนการชำระหนี้

การพยากรณ์ด้วยเทคนิคอนุกรมเวลา เป็นการรวบรวมข้อมูลตามช่วงเวลาที่เกิดขึ้นในอดีตที่ต่อเนื่องกัน แบบวันเดือนปี มาทำการวิเคราะห์แนวโน้มที่จะเกิดขึ้นในอนาคต โดยการพยากรณ์ของอนุกรมเวลามีหลายเทคนิค เช่น การพยากรณ์และการวัดประสิทธิภาพความแม่นยำของราคาหลักทรัพย์ในประเทศไทย โดยวิธี ARIMA [2] การสร้างแบบจำลองและการพยากรณ์แนวโน้มปริมาณน้ำฝนตามข้อมูลย้อนหลังใน Bungoma County, Western Kenya โดยใช้วิธี Holt Winters [3] เป็นต้น

จากปัญหา และงานวิจัยดังกล่าว ผู้วิจัยจึงมีแนวคิดในการสร้างระบบการพยากรณ์ด้วยเทคนิคอนุกรมเวลา และศึกษาทฤษฎีการพยากรณ์ที่เหมาะสมที่สุด มาใช้ในการสร้างระบบพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ เพื่อวางแผนบริหารงาน และวางกลยุทธ์ในการพัฒนาช่องทางชำระเงินตามความต้องการใช้งานของลูกค้าเพื่อสร้างความพึงพอใจ และประสิทธิภาพในการทำงานของผู้ปฏิบัติงานทางด้านบริการ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 ทฤษฎีการพยากรณ์

2.1.1 วิธีอาร์มา (ARIMA) อาศัยข้อมูลในอดีตเพื่อกำหนดรูปแบบของข้อมูลและพยากรณ์ข้อมูลในอนาคตแบบจำลอง ARIMA ประกอบด้วย 3 ส่วน Autoregressive Process, Integrated, Moving Average Process [4]

2.1.2 วิธีโฮลท์-วิทเทอร์ (Holt-Winters) คือการใช้โมเดลเพื่อการคาดการณ์ในชุดข้อมูลที่มีแนวโน้ม (Trend) และฤดูกาล (Seasonal) ซึ่งเหมาะสำหรับการพยากรณ์ระยะสั้นและระยะปานกลาง โดยมีการปรับค่าพารามิเตอร์ α , β และ γ เพื่อควบคุมความสำคัญของแนวโน้ม ค่าความลาดชัน และอิทธิพลของฤดูกาลในการพยากรณ์แต่ละช่วงเวลา โดยพารามิเตอร์ทั้งสามมีค่าระหว่าง 0 ถึง 1 โดยค่า α คือค่าปรับน้ำหนักสำหรับแนวโน้ม ค่า β คือค่าปรับน้ำหนักสำหรับค่าความลาดชันและค่า γ คือค่าปรับน้ำหนักสำหรับอิทธิพลของฤดูกาลในการพยากรณ์ [5]

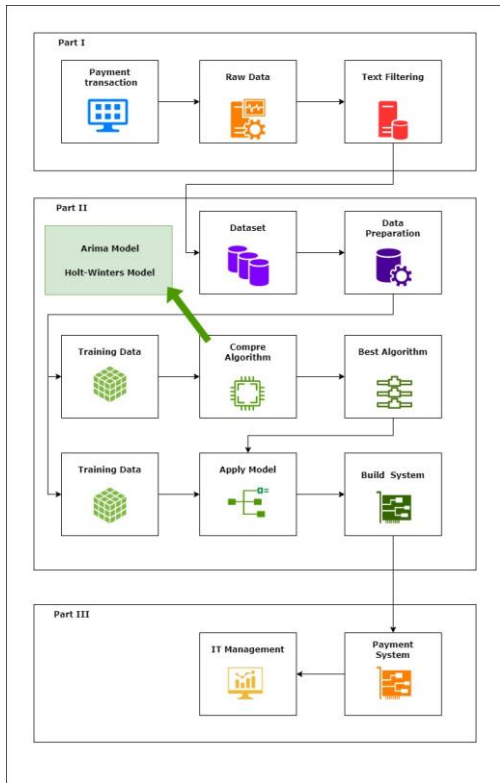
2.2 งานวิจัยที่เกี่ยวข้อง

ปณิษฐา [4] ได้ทำการศึกษาตัวแบบการพยากรณ์ราคาซื้อขายมันเส้นในประเทศไทย โดยข้อมูลที่ใช้ในการพยากรณ์รวบรวมมาตั้งแต่เดือนมกราคม 2559 ถึง เดือน เมษายน 2565 จำนวน 76 เดือน ซึ่งข้อมูลดังกล่าวมีส่วนประกอบของแนวโน้ม โดยนำเสนอเทคนิคการพยากรณ์ทั้ง 3 วิธีได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลพารามิเตอร์เดียวของบราวน์ และวิธีบ็อกซ์-เจนกินส์ ผลการวิจัยพบว่าการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุดใกล้เคียงกันได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งของโฮลท์ ด้วยร้อยละ 0.70 และ 0.90 และวิธีบ็อกซ์-เจนกินส์ ด้วยตัวแบบพยากรณ์ ARIMA (1,0,1) ส่วนวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลพารามิเตอร์เดียวของบราวน์ โดยที่ค่าเท่ากับ 0.70 มีค่า MAPE สูงสุด

วริศ และคณะ [6] ได้ทำการศึกษาการเปรียบเทียบวิธีพยากรณ์เพื่อเพิ่มความแม่นยำในการพยากรณ์ยอดขายกรณีศึกษาบริษัทผลิตขวดน้ำพลาสติกแห่งหนึ่งในประเทศไทย ใช้ข้อมูลเดือนสิงหาคม 2561 ถึง กรกฎาคม 2563 มาใช้ในการทดลอง โดยแบ่งข้อมูลออกเป็น 2 ปีวัดค่าประสิทธิภาพการพยากรณ์ด้วย MAPE ผลการวิจัยพบว่าวิธีการปรับให้เรียบแบบโฮลท์-วิทเทอร์ เป็นวิธีที่ดีที่สุดโดยมีค่า MAPE เท่ากับ 0.10 โดยต่ำสุดอยู่ที่ 0.047 และเมื่อเปรียบเทียบกับวิธีพยากรณ์ยอดขายเดิมของบริษัท

3. วิธีการดำเนินการวิจัย

ระบบพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ ผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากภาพที่ 1 กรอบแนวคิดการวิจัย สามารถแบ่งขั้นตอนการดำเนินงานวิจัยเป็น 3 ส่วน ดังนี้

3.1 การรวบรวมข้อมูลและเตรียมข้อมูล

การรวบรวมชุดข้อมูลสำหรับการพยากรณ์ เป็นข้อมูลรายการชำระหนี้ผ่านช่องทางออนไลน์ 4 ช่องทาง และทำการเตรียมข้อมูลโดยการ Query Data สรุปรายการในแต่ละวันตั้งแต่เดือน มีนาคม 2564 ถึง พฤศจิกายน 2564 ได้จำนวน 250 วัน ตัวอย่างแสดงคุณลักษณะทั้ง 4 ช่องทางดังตารางแสดงที่ 1 ถึง 4 โดยมีรายละเอียดดังนี้

ตารางที่ 1: ตัวอย่างคุณลักษณะข้อมูลช่องทางชำระหนี้ผ่านเว็บไซต์

คุณลักษณะ	ตัวอย่างข้อมูล
วันที่ทำรายการ	2022-05-31
ช่องทางการชำระหนี้	TOTCOTH
ผลรวมการรับชำระหนี้	170,771.85

ตารางที่ 2: ตัวอย่างคุณลักษณะข้อมูลช่องทางชำระหนี้ผ่านแอปพลิเคชัน

คุณลักษณะ	ตัวอย่างข้อมูล
วันที่ทำรายการ	2022-05-31
ช่องทางการชำระหนี้	EASYLIFE
ผลรวมการรับชำระหนี้	235,619.67

ตารางที่ 3: ตัวอย่างคุณลักษณะข้อมูลช่องทางชำระหนี้ผ่าน

คุณลักษณะ	ตัวอย่างข้อมูล
วันที่ทำรายการ	2022-05-31
ช่องทางการชำระหนี้	QUICKPAY
ผลรวมการรับชำระหนี้	156,958.44

ตารางที่ 4: ตัวอย่างคุณลักษณะข้อมูลช่องทางชำระหนี้ผ่านใบแจ้งหนี้อิเล็กทรอนิกส์

คุณลักษณะ	ตัวอย่างข้อมูล
วันที่ทำรายการ	2022-05-31
ช่องทางการชำระหนี้	PAY-SR
ผลรวมการรับชำระหนี้	194,871.54

3.2 ขั้นตอนการเปรียบเทียบ

นำข้อมูลที่ได้จากการเตรียมข้อมูลมาทำการสร้างแบบจำลองการพยากรณ์ RapidMiner โดยใช้เทคนิคการวิเคราะห์อนุกรมเวลา จำนวน 2 วิธี ได้แก่ วิธี ARIMA และวิธี Holt-Winters โดยพิจารณาการวัดค่าประสิทธิภาพรากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error : RMSE) ซึ่งในแต่ละเทคนิคกำหนดเลือกพารามิเตอร์ และทำการทดสอบเพื่อหาค่าพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลที่ดีที่สุด โดยกำหนดค่าพารามิเตอร์ของแต่ละเทคนิค แสดงดัง ตารางที่ 5

ตารางที่ 5: แสดงค่าพารามิเตอร์ของแต่ละวิธี

วิธี	พารามิเตอร์	ตัวแปร
ARIMA	p	1 - 10
	d	0 - 5
	q	1 - 10

วิธี	พารามิเตอร์	ตัวแปร
Holt-Winters	alpha	0.1 – 1.0
	beta	0.0 – 1.0
	gamma	0.0 – 1.0
	period	1- 7

4. ผลการดำเนินงาน

4.1 ผลการดำเนินงานสร้างแบบจำลองพยากรณ์

ผลทดสอบด้วยวิธี ARIMA แสดงดังตารางที่ 6 ถึง 9 โดยมีรายละเอียดดังนี้

ตารางที่ 6: ผลลัพธ์ข้อมูลช่องทางชำระเงินผ่านเว็บไซต์

รอบ	p	d	q	RMSE
7	7	0	1	0.156
8	8	0	1	0.165
9	9	0	1	0.169
10	10	0	1	0.201
11	1	1	1	0.760
12	2	1	1	0.464

จากตารางที่ 6 ค่าพารามิเตอร์ที่ดีที่สุดคือ $p = 7, d = 0, q = 1$ ให้ค่า RMSE = 0.156

ตารางที่ 7: ผลลัพธ์ข้อมูลช่องทางชำระเงินผ่านแอปพลิเคชัน

รอบ	p	d	q	RMSE
61	1	0	2	0.151
66	6	0	2	0.189
123	3	0	3	0.179
126	4	0	3	0.189
130	10	0	3	0.244
187	7	0	4	0.232

จากตารางที่ 7 ค่าพารามิเตอร์ที่ดีที่สุดคือ $p = 7, d = 0, q = 2$ ให้ค่า RMSE = 0.151

ตารางที่ 8: ผลลัพธ์ข้อมูลช่องทางชำระเงินผ่าน

รอบ	p	d	q	RMSE
182	2	0	4	0.137
185	5	0	4	0.189
186	6	0	4	0.193
189	9	0	4	0.269
241	1	0	5	0.149
361	1	0	7	0.338

จากตารางที่ 8 ค่าพารามิเตอร์ที่ดีที่สุดคือ $p = 2, d = 0, q = 4$ ให้ค่า RMSE = 0.137

ตารางที่ 9: ผลลัพธ์ข้อมูลช่องทางชำระเงินผ่านใบแจ้งหนี้อิเล็กทรอนิกส์

รอบ	p	d	q	RMSE
7	7	0	1	0.188
8	8	0	1	0.203
9	9	0	1	0.211
10	10	0	1	0.239
11	1	1	1	0.301
12	2	1	1	0.343

จากตารางที่ 8 ค่าพารามิเตอร์ที่ดีที่สุดคือ $p = 7, d = 0, q = 1$ ให้ค่า RMSE = 0.188

ผลการทดสอบด้วยวิธี Holt-Winters แสดงดังตารางที่ 10 ถึง 13 โดยมีรายละเอียดดังนี้

ตารางที่ 10: ผลลัพธ์การทดสอบของวิธี Holt-Winters ข้อมูลช่องทางชำระเงินผ่านเว็บไซต์

รอบ	alpha	beta	gamma	period	RMSE
17319	0.46	0.1	0	7	0.104
17329	0.37	0.2	0	7	0.105
17318	0.37	0.1	0	7	0.105
17320	0.55	0.1	0	7	0.105
17330	0.46	0.2	0	7	0.105
17340	0.37	0.3	0	7	0.106

จากตารางที่ 6 ค่าพารามิเตอร์ที่ดีที่สุดคือ $\alpha = 0.46, \beta = 0.1, \gamma = 0.1, \text{period} = 7$ ให้ค่า RMSE = 0.104

ตารางที่ 11: ผลลัพธ์การทดสอบของวิธี Holt-Winters ข้อมูลช่องทางชำระเงินผ่านแอปพลิเคชัน

รอบ	alpha	beta	gamma	period	RMSE
17319	0.46	0.1	0	7	0.102
17320	0.55	0.1	0	7	0.103
17321	0.64	0.1	0	7	0.104
17322	0.73	0.1	0	7	0.105
17323	0.82	0.1	0	7	0.107
17324	0.91	0.1	0	7	0.108

จากตารางที่ 6 ค่าพารามิเตอร์ที่ดีที่สุดคือ $\alpha = 0.46, \beta = 0.1, \gamma = 0.1, \text{period} = 7$ ให้ค่า RMSE = 0.102

ตารางที่ 12: ผลลัพธ์การทดสอบของวิธี Holt-Winters ข้อมูล

ช่องทางชำระเงินด่วน

รอบ	alpha	beta	gamma	period	RMSE
17319	0.46	0.1	0	7	0.087
17320	0.55	0.1	0	7	0.087
17321	0.64	0.1	0	7	0.087
17322	0.73	0.1	0	7	0.088
17323	0.82	0.1	0	7	0.089
17324	0.91	0.1	0	7	0.091

จากตารางที่ 6 ค่าพารามิเตอร์ที่ดีที่สุดคือ alpha = 0.46, beta = 0.1, gamma = 0.1, period = 7 ให้ค่า RMSE = 0.087

ตารางที่ 13: ผลลัพธ์การทดสอบของวิธี Holt-Winters ข้อมูล

ช่องทางชำระเงินผ่านใบแจ้งหนี้อิเล็กทรอนิกส์

รอบ	alpha	beta	gamma	period	RMSE
8123	0.46	0.1	0.1	7	0.159
8244	0.46	0.1	0.2	7	0.160
8365	0.46	0.1	0.3	7	0.161
8486	0.46	0.1	0.4	7	0.161
8607	0.46	0.1	0.5	7	0.167
8728	0.46	0.1	0.6	7	0.173

จากตารางที่ 6 ค่าพารามิเตอร์ที่ดีที่สุดคือ alpha = 0.46, beta = 0.1, gamma = 0.1, period = 7 ให้ค่า RMSE = 0.159

4.2 การเปรียบเทียบและคัดเลือกแบบจำลองพยากรณ์

การเปรียบเทียบและคัดเลือกแบบจำลองพยากรณ์ โดยนำค่าความคลาดเคลื่อนของเทคนิคที่ใช้ในการพยากรณ์ ได้ค่า RMSE ของวิธี ARIMA และวิธี Holt-Winters มาเปรียบเทียบเพื่อหาแบบจำลองพยากรณ์ที่เหมาะสมกับข้อมูลการชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ที่มีความคลาดเคลื่อนต่ำที่สุด พบว่า วิธี Holt-Winters เป็นเทคนิคที่มีค่าความคลาดเคลื่อนต่ำที่สุด ซึ่งเหมาะสมที่จะใช้สร้างระบบในการพยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์ รายละเอียดแสดงดังตารางที่ 14

ตารางที่ 14: การเปรียบเทียบค่าความคลาดเคลื่อน

ชุดข้อมูล	RMSE	
	ARIMA	Holt-Winters
TOTCOTH	0.156	0.159
EASYLIFE	0.151	0.102
QUICKPAY	0.137	0.087
PAY-SR	0.188	0.159

จากตารางที่ 5 เมื่อพิจารณาค่าความคลาดเคลื่อน พบว่าเทคนิค Holt-Winters มีค่าความคลาดเคลื่อน RMSE น้อยที่สุดในทุกชุดข้อมูล

4.2 การสร้างระบบพยากรณ์

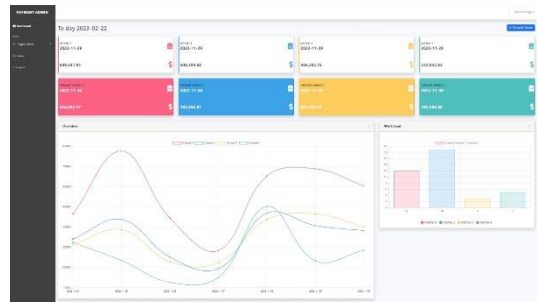
จากการเปรียบเทียบสำหรับการสร้างระบบพยากรณ์วิธีที่เหมาะสมที่สุดคือ วิธี Holt-Winters โดยสมการที่นำไปใช้การสร้างระบบพยากรณ์มีดังนี้

$$l_t = \alpha(Y_t - s_{[t-m]}) + (1 - \alpha)(l_{[t-1]} + b_{[t-1]}) \quad (1)$$

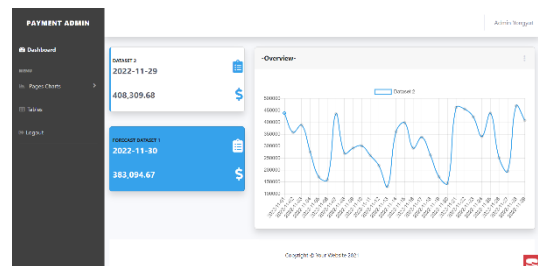
$$b_t = \beta(l_t - l_{[t-1]}) + (1 - \beta)b_{[t-1]} \quad (2)$$

$$s_t = \gamma(Y_t - l_t) + (1 - \gamma)s_{[t-m]} \quad (3)$$

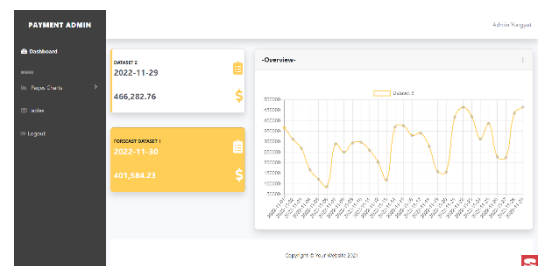
จากนั้นนำสมการข้างต้นมาสร้างระบบพยากรณ์โดยใช้ Software Framework ในการพัฒนารูปแบบเว็บไซต์โดยเชื่อมต่อกับฐานข้อมูลแสดงดังภาพที่ 2 ถึง 4



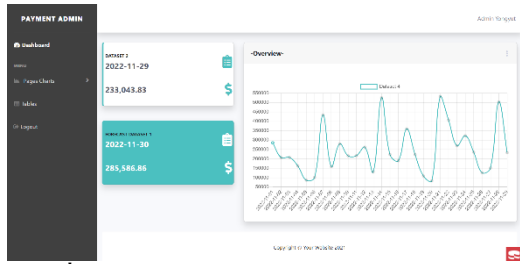
ภาพที่ 2: ระบบแสดงผลการพยากรณ์ภาพรวม



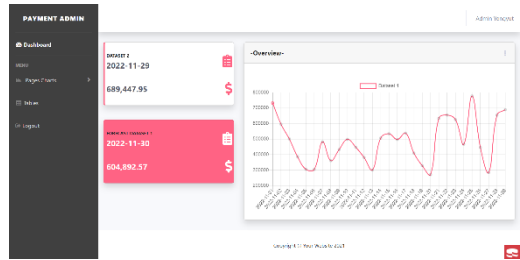
ภาพที่ 3: ระบบแสดงผลการพยากรณ์ตามช่องทางชำระเงิน



ภาพที่ 4: ระบบแสดงผลการพยากรณ์ตามช่องทางชำระเงิน



ภาพที่ 5: ระบบแสดงผลการพยากรณ์ตามช่องทางชำระเงิน



ภาพที่ 6: ระบบแสดงผลการพยากรณ์ตามช่องทางชำระเงิน

จากภาพที่ 2 เป็นการแสดงผลการพยากรณ์ภาพรวม การชำระเงินทั้ง 4 ช่องทาง ส่วนภาพที่ 3 ถึงภาพที่ 6 เป็นผลการพยากรณ์ที่เลือกโดยจะทำการแสดงผลข้อมูล การชำระเงินที่ผ่านมามีจำนวน 29 วัน และผลการพยากรณ์ การชำระเงินในวันถัดไปอีก 1 วัน

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้ศึกษาการพยากรณ์ด้วยเทคนิค อนุกรมเวลาเพื่อให้ได้วิธีการพยากรณ์ที่ดี และเหมาะสม ที่สุดสำหรับการพัฒนาระบบการพยากรณ์การชำระหนี้ ผ่านช่องทางชำระเงินออนไลน์ ผู้วิจัยได้คัดเลือกวิธีการ พยากรณ์ที่ได้รับความนิยม จำนวน 2 วิธี ได้แก่ วิธี ARIMA และวิธี Holt-Winters ทดสอบประสิทธิภาพ ของแบบจำลองการพยากรณ์ด้วยค่า RMSE ผลวิจัยพบว่า เทคนิค Holt-Winters มีความเหมาะสมที่สุด สำหรับชุด ข้อมูลค่า RMSE เท่ากับ 0.104, 0.102, 0.087, 0.159 ตามลำดับ เป็นค่าที่มีความคลาดเคลื่อนต่ำที่สุดทั้ง 4 ชุด ข้อมูลในแบบจำลองการพยากรณ์ และได้นำสมการของวิธี Holt-Winters ไปพัฒนาระบบเว็บไซต์ สรุปได้ว่า การ พยากรณ์การชำระหนี้ผ่านช่องทางชำระเงินออนไลน์โดย สร้างแบบจำลองการพยากรณ์ด้วยวิธี Holt-Winters มีความแม่นยำที่สุด

เอกสารอ้างอิง

- [1] วริศรา วรจารุงพงศ์, การพัฒนารูปแบบธุรกิจระบบการชำระ เงินแบบอิเล็กทรอนิกส์ SYe-Pay (สารนิพนธ์ปริญญาการ จัดการมหาบัณฑิต วิทยาลัยการจัดการ มหาวิทยาลัยมหิดล, 2565)
- [2] กนกกาญจน์ เสน่ห์ นมะหุด, “การเปรียบเทียบประสิทธิภาพ ในการทำนายผลการปรับความไม่สมดุลของข้อมูลในการ จำแนกด้วยอัลกอริทึมการทำเหมืองข้อมูล” วารสาร วิทยาศาสตร์และเทคโนโลยี, ปีที่ 8 ฉบับที่ 6 (พฤศจิกายน - ธันวาคม): 565-584
- [3] Vanessa A. Otieno and Samson W. Wanyonyi, “Modeling and Forecasting of Rainfall Trends based on Historical Data in Bungoma County, Western Kenya using Holt Winters Method” Asian Journal of Probability and Statistics, 17 (4). pp. 38-44. ISSN 2582-0230.
- [4] ปณิษฐา มีชอบธรรม, “ตัวแบบการพยากรณ์ราคาซื้อขายมัน เส้นในประเทศไทย” ปีที่ 11 ฉบับที่ 1 2022: วารสารวิชาการ ฉบับวิทยาศาสตร์และเทคโนโลยีสมาคมสถาบันอุดมศึกษา เอกชน แห่งประเทศไทย
- [5] พรรณีภา คุ่มสิน และสมศรี บัณฑิตวิไล, “การเปรียบเทียบตัว แบบการพยากรณ์ยอดขายเครื่องปรับอากาศโดยวิธีปรับให้ เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์, วิธีบ็อกซ์-เจน กิ้นส์ และวิธีโครงข่ายประสาทเทียม” วารสารพัฒนาเทคนิค ศึกษา ปีที่ 26 ฉบับที่ 3 พฤษภาคม - มิถุนายน พ.ศ. 2561 หน้า 364 - 376.
- [6] วริศ ลีมล่าววัลย์ และคณะ, “การเปรียบเทียบวิธีพยากรณ์เพื่อ เพิ่มความแม่นยำในการพยากรณ์ยอดขาย: กรณีศึกษาบริษัท ผลิตขวดน้ำพลาสติกแห่งหนึ่งในประเทศไทย” วารสารสุทธิ ปริทัศน์ ปีที่ 36 ฉบับที่ 2 เมษายน-มิถุนายน 2565

การตรวจจับกลุ่มผู้สร้างสแปมด้วยค่าความเป็นศูนย์กลาง

Detecting Spammer Groups Using Centrality Measure

ณัฐณิชา สุริยะมงคล (Natthanicha Suriyamongkol)¹ และสุกรี สินธุภิญโญ (Sukree Sinthupinyo)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹6370087621@student.chula.ac.th, ²sukree.s@chula.ac.th

บทคัดย่อ

การจ้างผู้ชำนาญเขียนรีวิวลดอบบนอีคอมเมิร์ซ (e-Commerce) แพลตฟอร์มเพื่อดึงดูดลูกค้าเป็นที่พบเจอมากขึ้นในยุคที่อีคอมเมิร์ซแพลตฟอร์มนั้นถูกใช้เป็นส่วนช่องทางหลักในการซื้อสินค้าและบริการ ด้วยปริมาณรีวิวลดอบที่มากขึ้นส่งผลให้ประสิทธิภาพในการตรวจจับของผู้เชี่ยวชาญ (fraud/spam specialist) ลดน้อยลง การวิจัยนี้จึงได้นำเสนอวิธีการตรวจจับกลุ่มผู้สร้างสแปม เพื่อช่วยลดปริมาณผู้ต้องสงสัยและเพิ่มประสิทธิภาพในการตรวจสอบให้แก่ผู้เชี่ยวชาญ โดยการวิจัยนี้ได้นำ XGBoost สำหรับจัดหมวดหมู่ (classification) มาใช้ในการตรวจจับการสร้างสแปมแบบบุคคล และใช้วิธีการหาค่าความเป็นศูนย์กลาง (centrality measure) และ Structural Clustering Algorithm for Network (SCAN) มาช่วยในการตรวจจับกลุ่มผู้สร้างสแปม จากผลการวิจัยแสดงให้เห็นว่าวิธีที่นำเสนอสามารถลดจำนวนผู้ต้องสงสัยที่สร้างสแปมในการส่งตรวจ อีกทั้งยังมีค่าความถูกต้อง (accuracy) และค่าความแม่นยำ (precision) ที่มากกว่างานวิจัยก่อนหน้าเมื่อพิจารณาจากกลุ่มที่มีความน่าสงสัยในการเป็นสแปมสูงสุดหนึ่งร้อยกลุ่มแรก

คำสำคัญ: สแปมรีวิวลดอบ, สแปมกลุ่ม, การตรวจจับกลุ่มผู้เขียนสแปม, พฤติกรรมผู้เขียนสแปม

Abstract

The practice of hiring professional spammers to create fake reviews on e-commerce platforms in order to attract customers can easily be found on the internet. As a result, the number of suspicious spammers has rapidly increased, hindering the

performance of fraud/spam specialists' investigations.

This study proposes a solution to mitigate the overwhelming number of investigative cases and improve investigation performance. The study utilizes XGBoost for classification to detect individual spammers; and centrality measure and the Structural Clustering Algorithm for Networks (SCAN) to identify spammer groups. Experimental evaluations show that this approach effectively reduces the number of spammers and increases accuracy and precision compared to previous studies, particularly concerning the hundred highest suspicious groups.

Keywords: Spam Review, Fraud Review, Spam Group, Spammer Group Detection, Spammer Behavioral.

1. บทนำ

อีคอมเมิร์ซ (e-Commerce) มีการเติบโตอย่างรวดเร็วในช่วงเหตุการณ์การแพร่ระบาดเชื้อไวรัสโควิด (COVID-19) [1] ร้านค้าจำนวนมากหันมาขายสินค้าผ่านช่องทางออนไลน์ ทำให้จำนวนสินค้าบนอีคอมเมิร์ซนั้นมีปริมาณเพิ่มมากขึ้นตามปริมาณผู้บริโภคที่มากขึ้น [2] ดังนั้นรีวิวลดอบ (review) จึงเป็นปัจจัยหลักของผู้บริโภคที่ใช้ในการตัดสินใจซื้อสินค้า หากร้านค้าหรือสินค้ามีคะแนนสูงก็จะดึงดูดลูกค้าได้มากขึ้น [3] จากสถิติแสดงว่า 90% ของผู้บริโภคอ่านรีวิวลดอบเพื่อประกอบการตัดสินใจก่อนซื้อสินค้า และ 70% ของผู้บริโภคเชื่อเนื้อหา รีวิว [4], [5] ผลที่ตามมาคือ ร้านค้าเลือกที่จะจ้างคนเขียนรีวิวลดอบ (spammer) หรือกลุ่มผู้สร้างรีวิวลดอบ (spammer group) เพื่อเพิ่มคะแนนหรือแม้แต่ทำลายชื่อเสียงของร้านคู่แข่ง [6]

กลุ่มผู้สร้างสแปม หมายถึง กลุ่มคนที่ร่วมกันเขียนรีวิพลอมจำนวนมากเพื่อเพิ่มคะแนนหรือลดคะแนนสินค้าที่ถูกอ้าง [7] โดยกลุ่มผู้สร้างสแปมจะพยายามอำพรางตัวตนด้วยการใช้หลายบัญชี (reviewer ID) [8] กลุ่มผู้สร้างสแปมนั้นสามารถที่จะสร้างความรุนแรงได้มากกว่าผู้สร้างสแปมเพียงหนึ่งหรือสองคน [9] เนื่องจากปริมาณสแปมที่มากนั้นสามารถที่จะปรับเปลี่ยนหรือควบคุมทิศทาง และมุมมองที่มีต่อสินค้านั้น ๆ ได้ [10], [11]

ปัญหาหลักในการตรวจจับสแปม คือ ผู้ตรวจนั้นไม่สามารถเชื่อถือผลการทำนายหรือวิเคราะห์จากโมเดลการเรียนรู้ของเครื่อง (Machine Learning model) ได้ทั้งหมด เนื่องจากผู้ปลอมรีวิวนั้นมีความชำนาญในการปลอมได้เหมือนของจริง ดังนั้นจึงต้องอาศัยผู้เชี่ยวชาญด้านการปลอมรีวิหรือสแปมในการยืนยันผล แต่สแปมปริมาณมหาศาลส่งผลให้คุณภาพในการตรวจสอบของผู้เชี่ยวชาญลดลง ดังนั้นการเลือกตรวจกลุ่มผู้สร้างสแปมนั้นจึงเป็นทางออกในการลดปริมาณผู้ต้องสงสัยเพื่อเพิ่มประสิทธิภาพการตรวจ และลดค่า Full-time equivalent (FTE) หรือหน่วยนับภาระงานของผู้เชี่ยวชาญต่อชั่วโมงการทำงาน โดยจะเน้นความสำคัญไปยังกลุ่มผู้สร้างสแปมที่มีผลกระทบรุนแรงเป็นอันดับแรก

แม้ว่าในปัจจุบันจะมีงานวิจัยในการตรวจจับกลุ่มผู้สร้างสแปมเป็นจำนวนมาก แต่ก็มีพื้นที่ที่ยังไม่ได้ถูกสำรวจ ซึ่งการวิจัยนี้ได้พัฒนาต่อจากงานวิจัยของ [12] โดยได้มีการต่อยอดด้วยวิธีการหาค่าความเป็นศูนย์กลาง อีกทั้งการวิจัยนี้ยังเป็นงานวิจัยแรกจากข้อมูลเท่าที่ผู้เขียนศึกษาที่นำปัจจัยด้านภาษาในการเขียนรีวิ พหุกรรม ผู้เขียนแบบบุคคล และแบบกลุ่มมาใช้ในการตรวจจับกลุ่มผู้สร้างสแปม ผลที่ได้คือ การวิจัยนี้สามารถลดจำนวนผู้ต้องสงสัยที่สร้างสแปม โดยมีค่าความถูกต้อง และความแม่นยำมากกว่างานวิจัย [12] ด้วยการใช้ชุดข้อมูลเดียว (YelpNYC dataset)

2. วรรณกรรมที่เกี่ยวข้อง

การศึกษาและพัฒนาระบบตรวจจับสแปมรีวิแบบบุคคล (Individually Spammer Detection) บนอีคอมเมิร์ซนั้น

มีต่อเนื่องมาอย่างยาวนาน โดย [13] เป็นงานวิจัยแรกที่นำชุดข้อมูลรีวิจากเชลท์ (Yelp.com) มาพัฒนาตรวจจับรีวิพลอม โดยใช้เมทาดาทา (metadata) ที่เป็นข้อความ วันที่คะแนน และเครือข่ายการรีวิ โดยงานวิจัยนี้มีประสิทธิภาพการคำนวณที่โดดเด่นในปรับขยายได้ตามขนาดของชุดข้อมูล และ [14] ได้ต่อยอดจากการวิเคราะห์รีวิด้วยข้อความ โดยการเพิ่มการวิเคราะห์ลักษณะการใช้ภาษาเฉพาะตัวของผู้ปลอมรีวิด้วยการใช้ LIWC, รูปแบบประโยคที่สามารถอ่านได้เข้าใจ, ความน่าเชื่อถือของคำที่ใช้, คำกริยาและกริยาช่วยในประโยค และ POS (Part of Speech) Tagging และยังมีกรนำพฤติกรรมของผู้รีวิมาใช้กับ Hierarchical Model ในการจับผู้ปลอมรีวิ

ไม่เพียงการตรวจจับสแปมแบบบุคคลเท่านั้นที่มีการวิจัยเป็นจำนวนมาก หากแต่การตรวจจับสแปมแบบกลุ่ม (Spammer Group Detection) นั้นก็เป็นที่ยอมรับในหมู่นักวิจัยเช่นกัน โดย [12] นั้นเป็นงานวิจัยนั้นมีจุดเด่นในเรื่องการนำวิธีการจับคู่ผู้รีวิ (co-reviewer) มาใช้พิจารณาและจับคู่ผู้รีวิที่มีพฤติกรรมและสินค้าที่รีวิที่เหมือนกัน จากนั้นก็ทำการจัดผู้รีวิเข้ากลุ่มด้วยวิธีการ SCAN และวิเคราะห์พฤติกรรมของผู้รีวิแบบบุคคลในกลุ่ม และพฤติกรรมของกลุ่ม แต่ [12] นี้ยังสามารถเพิ่มค่าความถูกต้องและความแม่นยำได้มากยิ่งขึ้นด้วยการใส่การวิเคราะห์ลักษณะของการเขียนรีวิ เนื่องจากลักษณะการเขียนเปลี่ยนแปลงได้ยากและแสดงถึงตัวผู้รีวิได้เป็นอย่างดี [14]

ในทางตรงกันข้ามการงานวิจัยด้านวิเคราะห์หาผู้เขียนรีวิที่มีความน่าเชื่อถือมากที่สุด [15] ได้นำเอาวิธีการการหาค่าความเป็นศูนย์กลางมาใช้ในการคำนวณหาคะแนนความน่าเชื่อถือ (Centrality-based Reputation) โดยดูจากคะแนนความเป็นประโยชน์จากผู้รีวิอื่นให้ (Helpfulness-based Reputation) และใช้ Gradient Boosting Regression ในการหาความสัมพันธ์ของคะแนนเพื่อระบุผู้รีวิที่น่าเชื่อถือ

การวิจัยที่น่าเสนอนี้ได้ต่อยอด [12] ด้วยการนำจุดเด่นที่ใช้จากงานวิจัยข้างต้นไม่ว่าจะเป็นคุณลักษณะ (feature), วิธีการตรวจจับสแปมแบบบุคคล, แบบกลุ่ม และวิธีการหาค่าความเป็นศูนย์กลางมาใช้ต่อยอด เพื่อพัฒนาการ

ตรวจจับของ [12] ที่มุ่งเน้นวิเคราะห์แค่พฤติกรรมผู้รีวิว และกลุ่ม แต่ยังคงเก็บจุดเด่นด้านวิธีการจับกลุ่มไว้

3. ขั้นตอนวิธีการดำเนินการวิจัย

การวิจัยนี้ได้นำเสนอการตรวจจับกลุ่มผู้สร้างสแปมที่พัฒนาต่อออกมาจาก [12] โดยได้มีการใช้ชุดข้อมูลเดียวกัน คือ YelpNYC ที่มีการจัดเก็บข้อความการรีวิวที่เป็นโรงแรม และร้านอาหารในรัฐนิวยอร์ก ประเทศสหรัฐอเมริกา ระหว่างช่วงปี ค.ศ. 2004 ถึง 2015 โดยชุดข้อมูลนี้มีจำนวน 359,052 รีวิว และ 160,225 ผู้รีวิว เมื่อคำนวณแล้วข้อมูลมีสแปมรีวิวอยู่ 10.27% และ ผู้เขียนสแปมอยู่ 17.79%

3.1 แนวคิด และ การจัดเตรียมข้อมูล

การวิจัยนี้ประยุกต์ใช้แนวคิดการค่าความเป็นศูนย์กลางของผู้เขียนรีวิว โดยพิจารณาจากสินค้าที่ผู้รีวิวไปเขียนรีวิวหรือให้คะแนนเป็นวิธีการหลักหรือจุดเด่นที่ใช้ปรับปรุงความถูกต้องและแม่นยำของงานวิจัยก่อนหน้านี้ [12] ซึ่งการวิจัยนี้ได้เปรียบผู้ปลอมรีวิวเป็นเสมือนผู้ป่วย ดังนั้นเมื่อผู้ปลอมรีวิวหรือผู้ป่วยไปรีวิวสินค้าใด สินค้านั้นก็เปรียบได้กับพาหะนำโรค และยังสินค้านั้นถูกรีวิวด้วยผู้ป่วยจำนวนมากเท่าไร ตัวสินค้านั้นก็จะมีค่าความเข้มของเชื้อโรคสูงตามไปด้วย กล่าวโดยสรุป คือ หากผู้เขียนรีวิวไปเขียนรีวิวสินค้าที่มีความเข้มของเชื้อสูงก็มีโอกาสที่จะติดเชื้อสูงตามลำดับหรือมีโอกาที่จะเป็นสแปมสูง

ดังนั้นการวิจัยนี้จึงเริ่มต้นด้วยหาตัวผู้เขียนสแปมแบบบุคคลเป็นลำดับแรกหรือหาผู้ป่วย โดยเริ่มจากลบผู้เขียนรีวิวที่มีจำนวนรวิวน้อยกว่าสองรีวิว ลบสินค้าที่มีจำนวนรวิวน้อยกว่าสามรีวิว ลบรีวิวที่จำนวนค่าน้อยกว่าห้าคำ ลบเครื่องหมายวรรคตอน (punctuation), ช่องว่าง (whitespace), และ คำฟุ่มเฟือย (stopword) โดยเมื่อเสร็จกระบวนการนี้แล้วจะเหลือรีวิวจำนวน 358,957 รีวิว

3.2 การตรวจจับผู้สร้างสแปมแบบบุคคล

เมื่อรีวิวผ่านกระบวนการจัดเตรียมตามขั้นตอนข้างต้นแล้วก็จะนำมาสกัดคุณลักษณะ (Feature Engineering) เพื่อสร้างคุณลักษณะทางด้านภาษา เนื่องจากสไลด์การเขียนนั้นยากที่จะเปลี่ยนแปลงในช่วงเวลาสั้น ๆ จึงจัดเป็นคุณลักษณะที่สำคัญในการระบุตัวตน [14], [16]

และพฤติกรรมของผู้เขียนรวิวนั้นก็มีความสำคัญไม่แพ้กัน ในการช่วยจับผู้ปลอมรีวิวในหลายงานวิจัย [11]-[12], [17]-[19] โดยแต่ละคุณลักษณะมีรายละเอียดดังนี้

1. LIWC แสดงถึงคะแนนจากการวิเคราะห์มุมมองทางด้านภาษาทั้ง 11 ด้าน
2. สามารถอ่านได้เข้าใจ (readability) วิธีการนี้ถูกนำมาใช้ในการวิเคราะห์ภาษาเก่า 80 ปีแล้ว
3. POS Tagging เป็นการวิเคราะห์คำกริยา (verb), กริยาวิเศษ (adverb), คำคุณศัพท์ (adjective) และ คำคุณศัพท์ขั้นสุด (superlative)
4. ข้อเท็จจริง (evidentiality) เป็นการประเมินข้อเท็จจริงของผู้รีวิวในรูปประโยคด้วยการพิจารณากริยาช่วย (modal adverb), เลกซิเคิล (lexical), กริยานุเคราะห์ (auxiliary verbs) และ กลุ่มคำคุณศัพท์ (epistemic adjectives) [20]
5. ไทม์เบอสทีเนส (time burstiness) ช่วงระยะเวลาที่ผู้ปลอมรีวิวเขียนรีวิวเป็นจำนวนมากเพื่อบรรลุประสงค์
6. จำนวนรีวิวที่โพสต์มากที่สุดในหนึ่งวัน
7. ค่าเบี่ยงเบนของคะแนนรีวิวจากค่าเฉลี่ย (average rating deviation)

จากนั้นจะนำรีวิวที่ผ่านการสร้างคุณลักษณะข้างต้นไปสอน XGBoost สำหรับจัดหมวดหมู่ [21] และระบุตัวผู้ต้องสงสัยหรือผู้ป่วยที่เป็นผู้เขียนสแปมแบบบุคคล

3.3 การคำนวณค่าความเป็นศูนย์กลาง

เมื่อทราบผู้ต้องสงสัยแบบบุคคลหรือผู้ป่วยจาก XGBoost แล้ว การวิจัยนี้จะใช้ Eigenvector Centrality (EC) [22] เพื่อคำนวณหาค่าความเป็นศูนย์กลางของสินค้าที่ผู้เขียนสแปมไปรีวิว หรือเป็นการหาว่าสินค้าไหนที่เป็นพาหะนำโรค โดยความเข้มของเชื้อหรือค่าความเป็นศูนย์กลางจะสูงตามจำนวนผู้ป่วยที่มาเขียนรีวิวบนสินค้านั้น ๆ จากนั้นค่าความเข้มของโรคจะถูกนำกลับไปคำนวณให้กับผู้รีวิวทุกคนในเครือข่ายเพื่อดูการแพร่กระจายของเชื้อโรค เช่น หากผู้รีวิวเขียนรีวิวให้สินค้าที่เป็นพาหะนำโรคเป็นจำนวนมากก็จะมีค่าเชื้อโรคสูงหรือมีโอกาสสูงที่จะเป็นผู้สร้างสแปม

3.4 การสร้างกลุ่มผู้เขียนสแปมด้วย SCAN

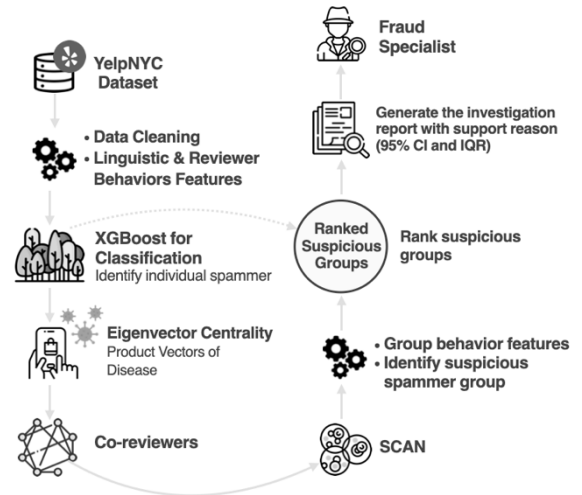
การสร้างกลุ่มผู้เขียนสแปมนั้นมีส่วนประกอบหลักคือ กราฟคู่ผู้เขียนรีวิวก (co-reviewer graph) [12], [23] หรือกราฟที่เชื่อมสองผู้เขียนรีวิวที่มีความคล้ายกันในเรื่อง ได้แก่ 1) วันที่รีวิวบนสินค้าเดียวกันนั้นห่างกันไม่เกินสองวัน หรือ 2) ให้คะแนนรีวิวต่างกันไม่เกินหนึ่งคะแนน และ 3) มีค่าเชื้อโรคที่ได้จาก EC ต่างกันไม่เกิน 50% และ 4) รีวิวสินค้าเดียวกันอย่างน้อยหนึ่งชิ้น จากนั้นจะนำกราฟคู่ผู้เขียนรีวิวมาระบุว่าเป็นศูนย์กลางของกลุ่มและค่อยๆ ขยายกลุ่มด้วยการเพิ่มสมาชิกเข้ามาด้วยการใช้คะแนนความคล้ายคลึง (similarity score) ตามขั้นตอนใน SCAN [24] จากนั้นก็นำกลุ่มทั้งหมดที่ได้มาสกัดคุณลักษณะทางด้านพฤติกรรมกลุ่มเพิ่มเติม ได้แก่

1. ความคล้ายคลึงของรีวิว (review tightness) จะพิจารณาจำนวนรีวิว และ สินค้าของสมาชิกในกลุ่ม
2. ความแปรปรวนของคะแนนรีวิว (rating variance) วิเคราะห์ความเหมือนของคะแนนรีวิวของสมาชิกในกลุ่ม

จากนั้นการวิจัยนี้จะระบุกลุ่มที่ต้องสงสัยที่จะเป็นผู้สร้างสแปมด้วยการใช้ค่าเรซโฮลด์ (threshold) คัดแบ่งคะแนนจากคุณลักษณะทางด้านพฤติกรรมกลุ่มข้างต้นตามงานวิจัยก่อนหน้านี้ [12]

3.5 การสรุปผลการและจัดทำรายงาน

การวิจัยนี้จะใช้สัดส่วนเปอร์เซ็นต์ (percent) ของผู้ปลอมรีวิวแบบบุคคลที่ได้จาก XGBoost มาช่วยจัดลำดับกลุ่มผู้สร้างสแปมต้องสงสัย โดยจะเรียงตามลำดับเปอร์เซ็นต์จากมากไปน้อย กลุ่มที่มีเปอร์เซ็นต์สูงคือกลุ่มที่มีโอกาสเป็นกลุ่มผู้สร้างสแปมสูง เพื่อลดเวลาและอำนวยความสะดวกให้แก่ผู้เชี่ยวชาญในการตรวจสอบกลุ่มที่มีโอกาสเป็นสแปมสูงก่อน พร้อมทั้งแนบเหตุผลสนับสนุนว่าแต่ละกลุ่มนั้นมีความน่าสงสัยอย่างไรด้วยการใช้ 95% ช่วงความเชื่อมั่น (confidence interval) และ ค่าพิสัยระหว่างควอไทล์ (interquartile range, IQR) โดยการใช้ค่าจากคุณลักษณะทั้งสาม ได้แก่ การใช้ภาษาหรือสไตล์การเขียนพฤติกรรมของผู้เขียนทั้งแบบบุคคล และกลุ่มเป็นเหตุผล



ภาพที่ 1: โครงสร้างของการตรวจจับกลุ่มผู้สร้างสแปมของงานวิจัยที่นำเสนอ

3.5 การเปรียบเทียบกับงานวิจัยก่อนหน้า

หากสรุปแล้วการวิจัยนี้เป็นการนำวิธีการของการวิจัยก่อนหน้านี้ [12] มาปรับให้เหมาะกับวัตถุประสงค์ของการวิจัยนี้ ซึ่งการวิจัยก่อนหน้านี้ได้นำข้อมูลสกัดคุณลักษณะพฤติกรรมของผู้เขียนรีวิวเพียงอย่างเดียว การวิจัยนี้จึงได้มีการศึกษาและเพิ่มการสกัดคุณลักษณะสไตล์การเขียนเข้าไป จากนั้นก็จะสร้างกราฟคู่ผู้เขียนรีวิว, SCAN และทำการสกัดคุณลักษณะทางด้านพฤติกรรมกลุ่ม และใช้การกำหนดค่าเรซโฮลด์เพื่อใช้ระบุกลุ่มผู้สร้างสแปม ซึ่งทั้งสี่ขั้นตอนนี้ การวิจัยนี้ไม่มีได้มีการเปลี่ยนแปลงใดๆ จากงานวิจัยก่อนหน้านี้ แต่สิ่งการวิจัยที่นำเสนอมีการเพิ่มเติมคือการประยุกต์ใช้ EC และทำการระบุผู้สร้างสแปมแบบบุคคลเพื่อเพิ่มความถูกต้อง, ความแม่นยำ (precision) และยังคงที่ค่าเรียกคืน (recall) ไว้ รวมถึงการแนบเหตุผลสนับสนุนเพื่ออำนวยความสะดวกในการตรวจสอบ

4. ผลการดำเนินงาน

การวิจัยนี้ได้มีการนำวิธีการหาค่าความเป็นศูนย์กลาง และนำเอาคุณลักษณะทั้งด้านภาษา พฤติกรรมแบบบุคคล และกลุ่มมาใช้ในการตรวจจับเป็นครั้งแรก ซึ่งผลลัพธ์ที่ได้นั้นแสดงถึงความถูกต้องและแม่นยำที่มากกว่างานวิจัยก่อนหน้านี้ [12] จากการใช้ชุดข้อมูลเดียวกันตามตารางที่ 1 โดยจำนวนกลุ่มสแปมหมายถึงถึง

กลุ่มที่มีความเสี่ยงสูงสุด เช่น จำนวนกลุ่มสแปม 30 กลุ่ม คือ 30 กลุ่มแรกที่มีความเสี่ยงเป็นในการสแปมสูงสุด

ตารางที่ 1: ผลการเปรียบเทียบค่าความแม่นยำ และค่าเรียกคืนของกลุ่มสแปม และค่าความถูกต้อง

# กลุ่ม สแปม	งานวิจัยก่อนหน้า			งานวิจัยที่นำเสนอ		
	ความ แม่นยำ	เรียกคืน	ความ ถูกต้อง	ความ แม่นยำ	เรียกคืน	ความ ถูกต้อง
30	0.07	1.00	0.07	0.29	0.75	0.43
50	0.06	1.00	0.06	0.25	0.80	0.48
70	0.06	1.00	0.06	0.19	0.80	0.47
90	0.07	1.00	0.07	0.15	0.80	0.49
110	0.07	1.00	0.19	0.15	0.82	0.52
130	0.07	0.88	0.31	0.15	0.82	0.59
150	0.07	0.88	0.39	0.15	0.82	0.65
170	0.07	0.78	0.45	0.15	0.82	0.69

เมื่อพิจารณาจากวัตถุประสงค์ของการวิจัยที่ต้องการลดค่าปริมาณงานของผู้เชี่ยวชาญในการตรวจสอบหรือลดค่า FTE เพื่อเพิ่มคุณภาพของการตรวจสอบ ตารางที่ 2 จึงแสดงการเปรียบเทียบปริมาณของผู้สร้างสแปมรีวิวก่อนและหลังการตรวจสอบ

ตารางที่ 2: ผลการเปรียบเทียบจำนวนผู้รีวิวในกลุ่มผู้สร้างสแปม

#กลุ่ม สแปม	งานวิจัยก่อนหน้า		งานวิจัยที่นำเสนอ	
	#ผู้เขียน รีวิว	#ผู้ต้อง สงสัย	#ผู้เขียน รีวิว	#ผู้ต้อง สงสัย
30	239	239	227	115
50	390	390	389	178
70	614	614	574	243
90	799	799	758	290
110	943	835	927	335
130	1107	835	1121	335
150	1236	835	1231	335
170	1396	835	1380	335

5. สรุป

งานวิจัยนี้ได้นำเสนอวิธีการใหม่ในการสร้างการตรวจสอบจับกลุ่มผู้สร้างสแปม โดยพัฒนาต่อจากงานวิจัยของ [12] ซึ่งผลลัพธ์ที่ได้แสดงว่าการวิจัยนี้สามารถตรวจจับกลุ่มผู้สร้างสแปมด้วยค่าความถูกต้องและแม่นยำ

ที่มากขึ้น และมีการลดปริมาณงานที่ผู้เชี่ยวชาญต้องตรวจสอบตามวัตถุประสงค์ที่กล่าวไว้ข้างต้น แต่อย่างไรก็ตามปัญหาที่ทำให้ค่าความถูกต้องและแม่นยำไม่มากนักเนื่องจากชุดข้อมูลที่ใช้นี้ในงานวิจัยนี้ [13] นั้นไม่มีเฉลย (ground-truth label) ของกลุ่มผู้สร้างสแปม จึงทำให้ผู้เชี่ยวชาญทำการสร้างเฉลยกลุ่มโดยกำหนดว่าหากกลุ่มใดที่มีเฉลยสแปมบุคคลมากกว่า 50% จะพิจารณาว่าเป็นกลุ่ม สแปม โดยในอนาคตผู้เชี่ยวชาญวางแผนที่จะสร้างความร่วมมือกับผู้เชี่ยวชาญด้านสแปมในการสร้างเฉลยแบบกลุ่มสแปมที่แก้ไขปัญหาข้างต้น

เอกสารอ้างอิง

- [1] Brewster, Mayumi. "Annual Retail Trade Survey Shows Impact of Online Shopping on Retail Sales During COVID-19 Pandemic." *Ensus.Gov*. <https://www.census.gov/library/stories/2022/04/ecommerce-sales-surged-during-pandemic.html>, 2022.
- [2] Coppola, Daniela. "E-Commerce Worldwide - Statistics & Facts." *Statista.Com*. https://www.statista.com/topics/871/online-shopping/#topicHeader__wrapper, 2022.
- [3] Tang, Xiaoya, Tiejun Qian, and Zhenni You. "Generating Behavior Features for Cold-Start Spam Review Detection with Adversarial Learning." *Information Sciences* 526: 274–88, 2020.
- [4] Luca, Michael. "Reviews, Reputation, and Revenue: The Case of Yelp.Com." *Harvard Business School*. https://www.hbs.edu/ris/Publication_Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf, 2016.
- [5] Mintel Press. "Seven In 10 Americans Seek Out Opinions Before Making Purchases." *Mintel.com*. <https://www.mintel.com/press-centre/social-and-lifestyle/seven-in-10-americans-seek-out-opinions-before-making-purchases>, 2015.
- [6] Xu, Chang, and Jie Zhang. "Towards Collusive Fraud Detection in Online Reviews." In *2015 IEEE International Conference on Data Mining*, 1051–56. IEEE, 2015.
- [7] Mukherjee, Arjun, Bing Liu, Junhui Wang, Natalie Glance, and Nitin Jindal. "Detecting Group Review Spam." In *Proceedings of the 20th International*

- Conference Companion on World Wide Web - WWW '11*, 93. New York, New York, USA: ACM Press, 2011.
- [8] Vidanagama, Dushyanthi U., Thushari P. Silva, and Asoka S. Karunananda. “Deceptive Consumer Review Detection: A Survey.” *Artificial Intelligence Review* 53 (2): 1323–52, 2020.
- [9] Li, Jiandun, Pin Lv, Wei Xiao, Liu Yang, and Pengpeng Zhang. “Exploring Groups of Opinion Spam Using Sentiment Analysis Guided by Nominated Topics.” *Expert Systems with Applications* 171 (June): 114585, 2021.
- [10] Mukherjee, Arjun, Bing Liu, and Natalie Glance. “Spotting Fake Reviewer Groups in Consumer Reviews.” In *Proceedings of the 21st International Conference on World Wide Web*, 191–200. New York, NY, USA: ACM, 2012.
- [11] Wang, Zhuo, Songmin Gu, Xiangnan Zhao, and Xiaowei Xu. “Graph-Based Review Spammer Group Detection.” *Knowledge and Information Systems* 55 (3): 571–97, 2018.
- [12] Hussain, Naveed, Hamid Turab Mirza, Abid Ali, Faiza Iqbal, Ibrar Hussain, and Mohammad Kaleem. “Spammer Group Detection and Diversification of Customers’ Reviews.” *PeerJ Computer Science* 7 (April): e472, 2021.
- [13] Rayana, Shebuti, and Leman Akoglu. “Collective Opinion Spam Detection.” In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 985–94. New York, NY, USA: ACM, 2015.
- [14] Le, Thi-Kim-Hien, Yi-Zhen Li, and Sheng-Tun Li. “Do Reviewers’ Words and Behaviors Help Detect Fake Online Reviews and Spammers? Evidence From a Hierarchical Model.” *IEEE Access* 10: 42181–97, 2022.
- [15] Meo, Pasquale De, Katarzyna Musial-Gabrys, Domenico Rosaci, Giuseppe M. L. Sarnè, and Lora Aroyo. “Using Centrality Measures to Predict Helpfulness-Based Reputation in Trust Networks.” *ACM Transactions on Internet Technology* 17 (1): 1–20. <https://doi.org/10.1145/2981545>, 2017.
- [16] Ott, Myle, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock. “Finding Deceptive Opinion Spam by Any Stretch of the Imagination.” *ACL-HLT 2011 - Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies 1*: 309–19, 2011.
- [17] Choo, Euijin, Ting Yu, and Min Chi. “Detecting Opinion Spammer Groups Through Community Discovery and Sentiment Analysis.” In , 170–87, 2015.
- [18] D’Onfro, Jillian. “No Title.” Insider. <https://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9>, 2013.
- [19] Zhang, Lu, Zhiang Wu, and Jie Cao. “Detecting Spammer Groups From Product Reviews: A Partially Supervised Learning Model.” *IEEE Access* 6: 2559–68, 2018.
- [20] Su, Qi, Chu-Ren Huang, and Helen Kai-yun Chen. “Evidentiality for Text Trustworthiness Detection.” In *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, 10–17. NLP LING ’10. USA: Association for Computational Linguistics, 2010.
- [21] Chen, Tianqi, and Carlos Guestrin. “XGBoost.” In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. New York, NY, USA: ACM, 2016.
- [22] Eigenvector centrality. “Eigenvector Centrality - Wikipedia.” *Wikipedia*. https://en.wikipedia.org/wiki/Eigenvector_centrality, 2022.
- [23] Kaghazgaran, Parisa, Majid Alfifi, and James Caverlee. “Wide-Ranging Review Manipulation Attacks.” In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 981–90. New York, NY, USA: ACM, 2019.
- [24] Xu, Xiaowei, Nurcan Yuruk, Zhidan Feng, and Thomas A. J. Schweiger. “SCAN.” In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '07*, 824. New York, New York, USA: ACM Press, 2007.

การพัฒนาระบบปัญญาประดิษฐ์สำหรับการตรวจจับผู้ไม่สวมใส่หน้ากากอนามัย ในจังหวัดจันทบุรี

The Development of Artificial Intelligence for Face Mask Detection System in Chantaburi Province

สาธิต สุวรรณเวช (Sathit Suwannawach)¹ ทิพย์วรรณ ฟูเฟื่อง (Thipwan Fufaung)²

และขนิษฐา สิริทิเทียมจันทร์ (Khanittha Sittitiamjan)³

^{1,2,3}คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏรำไพพรรณี

¹sathit.s@rbru.ac.th, ²thipwan.f@rbru.ac.th and ³khanittha.s@rbru.ac.th

บทคัดย่อ

เทคโนโลยีปัญญาประดิษฐ์ได้เข้ามามีบทบาทในการวิเคราะห์ข้อมูลจำนวนมากแทนคนมากขึ้นอีกทั้งมีความแม่นยำสูงในการประยุกต์ใช้ในชีวิตประจำวันและการศึกษาวิจัยด้านปัญญาประดิษฐ์มีปริมาณมากขึ้นตามในการวิจัยที่มวิจัยได้นำเสนอการประยุกต์ใช้โมเดล YOLOv5 มาใช้ในการฝึกอบรมชุดข้อมูลและวิเคราะห์ข้อมูลโดยมีวัตถุประสงค์การวิจัยเพื่อพัฒนาระบบปัญญาประดิษฐ์สำหรับการตรวจจับผู้ไม่สวมใส่หน้ากากอนามัยในจังหวัดจันทบุรี และมีวิธีการดำเนินการวิจัยมีขั้นตอน 5 ขั้นตอนประกอบด้วย 1) การสำรวจพื้นที่ที่จัดเก็บข้อมูลในการวิจัย 2) การรวบรวมข้อมูลในการวิจัย 3) การออกแบบและพัฒนาระบบ 4) การทดลองระบบ และ 5) การวิเคราะห์และสรุปผลการทดลองและแบ่งชุดข้อมูลออกสองส่วนคือชุดแรกใช้ในการฝึกอบรมและชุดที่สองสำหรับการทดสอบผลการฝึกอบรม โดยมีขนาดภาพที่ได้ปรับขนาด 640x640 พิกเซล ซึ่งในกลุ่มตัวอย่างที่ใช้ของระบบที่ติดตั้งอุปกรณ์ไว้เพื่อให้ง่ายต่อการตรวจจับทิศทางการเดินทางได้กำหนดระยะห่างไม่เกิน 5 เมตรและพื้นที่ว่างให้เหมาะสมกับระยะความสูงของกล้องตรวจจับวิดีโอที่ 480 พิกเซลและสอดคล้องกับความเร็วในการประมวลผลจากความสามารถในการประมวลผลที่มีอย่างจำกัดจากวิดีโอที่ใช้ในงานวิจัยและตรวจวัดประสิทธิภาพของระบบหลังจากดำเนินการทดลองและประเมินผลการตรวจจับพบว่าระบบมีความแม่นยำแบบเรียลไทม์ 96%

คำสำคัญ: เทคโนโลยีปัญญาประดิษฐ์ ความแม่นยำ โมเดล ประสิทธิภาพ ชุดข้อมูล

Abstract

Artificial intelligence technology has played a role in analyzing large amounts of data instead of more people and has high accuracy in everyday life applications. And research on artificial intelligence is increasing accordingly. In this research, the research team presents the application of the YOLOv5 model used to train the dataset and analyze the data. The research objective is to develop an artificial intelligence system for detecting people who do not wear masks in Chanthaburi Province. There are 5 steps in the research process, consisting of 1) exploring the research storage area, 2) collecting research data, 3) designing and developing the system, 4) testing the system, and 5) analyzing. And summarized the experimental results and divided the data set into two parts, the first set was used for training and the second set was used for testing the training results. with an image size that has been resized to 640x640 pixels, in which the sample used of the system that is installed to make it easier to detect the direction of walking has a distance of no more than 5 meters and free space to suit the height The camera detects video at 480 pixels and is consistent with the processing speed due to the limited processing capacity from the video used in research and measurement of system performance

after conducting experiments and detection evaluations.
The system has a real-time accuracy of 96%.

Keywords: Artificial Intelligence, Accuracy, Model, Efficiency, Dataset.

1. บทนำ

เทคโนโลยีปัญญาประดิษฐ์หรือ AI ย่อมาจาก Artificial Intelligence เป็นโปรแกรมที่ถูกเขียนสร้างขึ้นให้มีระดับความสามารถใกล้เคียงกับมนุษย์เป็นเทคโนโลยีที่ทันสมัยที่สามารถแก้ปัญหาที่ซับซ้อนเกินกว่าที่มนุษย์จะสามารถรับมือได้ โดยเฉพาะความสามารถในการคิดเองได้ หรือทำเองได้ โปรแกรมจะพัฒนาให้มีความฉลาด มีความสามารถในการคิด วิเคราะห์ วางแผน และตัดสินใจ ในปัจจุบันยังพบว่าระบบ AI มีความสำคัญต่อธุรกิจ เช่น ใช้ในด้านความปลอดภัย การตรวจจับใบหน้าของบุคคล เพื่อเก็บเป็นหลักฐานข้อมูล ทางข้อมูลราชการตำรวจ ข้อมูลพนักงาน หรือแม้กระทั่งตรวจจับ บันทึกข้อมูล บุคคลอันตรายได้ ทางทีมผู้วิจัยจึงนำความสำคัญของเทคโนโลยีปัญญาประดิษฐ์มาปรับใช้ในการพัฒนาระบบโดยระบบจะทำหน้าที่ทดแทนเจ้าหน้าที่ที่คอยยืนตรวจสอบคนที่ไม่สวมใส่หน้ากากในจุดตรวจวัดอุณหภูมิ หรือจุดคัดกรองบุคคล จากนั้นเมื่อตรวจจับพบคนไม่สวมใส่หน้ากากอนามัย ระบบจะส่งแจ้งเตือนไปยังเจ้าหน้าที่ที่เกี่ยวข้องหรือผู้ดูแลระบบให้สามารถเรียกบุคคลดังกล่าวไปเตือนให้ใส่หน้ากากอนามัยได้ลดระยะสัมผัสใกล้ชิดกับกลุ่มคนลงได้ วิธีการที่นำเสนอนี้สามารถรวมเข้ากับจุดสังเกตบนใบหน้าด้วยการมาสก์หน้าโดยใช้ Multi-Task Cascaded Convolutional Neural Network ซึ่งเป็นวิธีการตรวจจับใบหน้าที่แม่นยำที่สุด (MTCNN) [1]

เทคโนโลยีการตรวจหาวัตถุ (Object Detection) เป็นเทคนิคหนึ่งของเทคโนโลยีปัญญาประดิษฐ์ ที่ใช้กับกล้องวงจรปิด สามารถค้นหาสิ่งของโดยใช้ AI มาวิเคราะห์ข้อมูล จากการมองเห็นของคอมพิวเตอร์ (Computer Vision) และการประมวลผลภาพ (Image Processing) เพื่อตรวจจับวัตถุที่อยู่ในรูปหรือวิดีโอ เช่น มนุษย์ สัตว์ และสิ่งของ เป็นต้น ซึ่งมีทั้งชุดข้อมูลที่อยู่ในรูปแบบภาพนิ่งหรือวิดีโอโดยมีตัวตรวจหาวัตถุ (Object Detector)

ที่สำคัญในโมเดล เช่น R-CNN (Regions with CNN) [5], SSD (Single Shot MultiBox Detector) และ YOLO (You Only Look Once)

YOLOv5 เป็นเทคโนโลยีการตรวจจับวัตถุแบบเรียลไทม์ซึ่งเป็นเทคนิคที่ใช้ในการตรวจจับวัตถุจากกล้องวิดีโอโดย YOLO การตรวจจับวัตถุแบบเรียลไทม์ (Real Time) ที่ใช้กันอย่างแพร่หลายในปัจจุบันซึ่งเวอร์ชันใหม่ รุ่นที่ 5 นี้เขียนด้วยนวัตกรรม PyTorch ซึ่งข้อดีนั้นมีความเร็วในการประมวลผลได้สูงถึง 140 เฟรมต่อวินาที โดยมีความแม่นยำและมีขนาดเล็กกว่า YOLO v4 ถึง 90% ซึ่งนั่นหมายความว่าจะสามารถนำไปฝังหรือประยุกต์กับอุปกรณ์ขนาดเล็กได้ดีกว่า ดังนั้นงานวิจัยนี้จึงได้เสนอการใช้เทคโนโลยีตรวจหาวัตถุมาใช้กับปัญหาการตรวจจับ หน้ากากอนามัย โดยใช้โมเดล YOLOv5 มาฝึกฝนชุดข้อมูลภาพถ่ายหน้ากากอนามัยสำหรับผู้ที่สวมใส่เพื่อตรวจหาผู้ไม่สวมใส่หน้ากากอนามัยเข้าสู่พื้นที่ที่เคร่งครัด

2. วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาระบบปัญญาประดิษฐ์สำหรับการตรวจจับผู้ไม่สวมใส่หน้ากากอนามัยในจังหวัดจันทบุรี

3. งานวิจัยที่เกี่ยวข้อง

เทคนิคสำหรับการตรวจหาวัตถุแบบเวลาจริง (Realtime Object Detection) ที่สำคัญ เทคนิคที่ได้รับ ความนิยมมาก YOLO เป็นเทคนิคในกลุ่มการตรวจหาวัตถุแบบขั้นตอนเดียว (One-Stage) [2] ในงานวิจัยที่ใช้ การเปรียบเทียบความแม่นยำระหว่าง YOLOv4 และ YOLOv3 พบว่า YOLOv4 มีความแม่นยำเฉลี่ยมากขึ้น เมื่อเทียบกับ YOLOv3 ซึ่งมีค่าเฉลี่ยเท่ากับ 78.94% และ YOLOv4 มีความแม่นยำเฉลี่ยเท่ากับ 93.51% ซึ่งแสดงให้เห็นว่า YOLOv4 มีค่าความแม่นยำและความสามารถในการตรวจจับวัตถุแบบเวลาจริงได้ผลสูง [3] สำหรับงานวิจัยวิธีการวิเคราะห์ภาพวิดีโอคอนและสัมภาระสำหรับการตรวจจับวัตถุที่ปราศจากเจ้าของ และตรวจจับบุคคลที่มี การถือกระเป๋าประเภทต่าง ๆ เนื่องจากเหตุการณ์วางทิ้ง กระเป๋าเป็นเหตุการณ์ที่มีความเสี่ยงและเกิดขึ้นได้จริง

โดยเฉพาะพื้นที่ สาธารณะที่ต้องการเฝ้าระวังเป็นพิเศษ เช่น สถานีรถไฟ สนามบิน หรือสถานที่สำคัญภายในอาคาร เป็นต้น การตรวจจับเพื่อระบุตำแหน่งคนและกระเป๋าสานที่ใช้วิธีการเรียนรู้เชิงลึกที่ถูกฝึกสอนด้วยภาพจำนวน 12,000 ภาพ ที่ประกอบด้วยภาพคนและกระเป๋าสาน เช่น กระเป๋าเป้ กระเป๋าถือ และกระเป๋าเดินทาง โดยในงานวิจัยนี้ใช้แบบจำลองชุดข้อมูลด้วยเทคนิคโมเดล YOLOv3 สำหรับการตรวจจับวัตถุซึ่งประมวลผลได้แบบเรียลไทม์ และมีประสิทธิภาพความถูกต้อง (Accuracy) ถึง 98% [4]

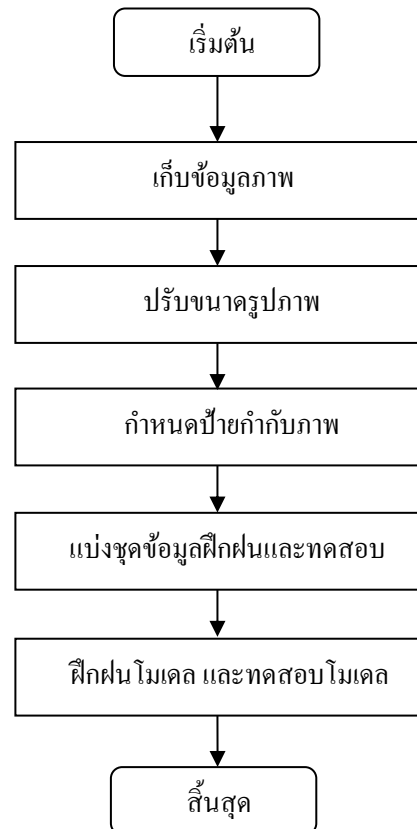
รามณรงค์ และคณะ (2564). การตรวจจับการสวมหน้ากากอนามัยโดยประยุกต์ใช้โครงข่ายประสาทเทียมแบบเชิงลึกและโครงข่ายประสาทเทียมแบบคอนโวลูชัน โดยมีการดำเนินงานของการวิจัยบรรลุวัตถุประสงค์ที่ได้กำหนดไว้คือเพื่อพัฒนาระบบตรวจจับการสวมใส่หน้ากากภาพ การพัฒนาระบบตรวจสอบการสวมหน้ากากอนามัยมีขั้นตอนการพัฒนา 3 ขั้นตอน ประกอบด้วย 1) การเก็บรวบรวมข้อมูล 2) การสร้างแบบจำลอง และ 3) การประยุกต์ใช้แบบจำลอง ทีมผู้วิจัยได้เก็บข้อมูลใบหน้าภาพแบบสวมหน้ากาก 1,919 ภาพ และไม่สวมหน้ากาก 1,916 ภาพ มีวลดหลาย ก้มหน้า เอียงหน้า และสิ่งปกปิดอัตลักษณ์บนใบหน้าทั้งหมด 3,835 ภาพ ใช้เครื่องมือคือ confusion matrix ทำให้ผลการวิจัยแบบ Per Frame จาก การทดลองมีผลลัพธ์ของค่าความถูกต้องมากที่สุดอยู่ที่ร้อยละ 97 และการทดลองแบบ Real Time ใช้เครื่องมือคือ Intersection over Union (IoU) ความถูกต้องอยู่ที่ร้อยละ 98 และแบบค่าความถูกต้องแบบตามเหตุการณ์อยู่ที่ร้อยละ 96 ใช้ MobileNetV2 ร่วมกับกระบวนการ โครงข่ายประสาทเทียม ANN (Artificial Neuron Network) [5]

3. วิธีการดำเนินการวิจัย

กระบวนการพัฒนาระบบตรวจจับผู้ไม่สวมใส่หน้ากากอนามัยมีขั้นตอน 5 ขั้นตอน ประกอบด้วย 1) การสำรวจพื้นที่ที่จัดเก็บข้อมูลในการวิจัย 2) การรวบรวมข้อมูลในการวิจัย 3) การออกแบบและพัฒนาระบบ 4) การทดลองระบบ และ 5) การวิเคราะห์และสรุปผลการทดลอง

3.1 วิธีการดำเนินการวิจัย

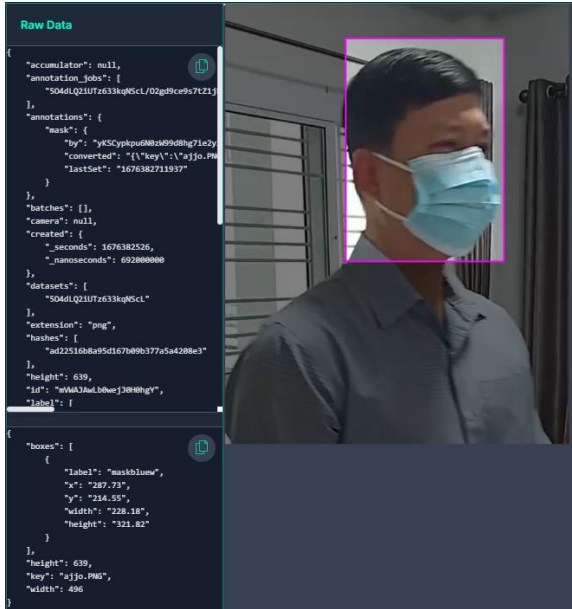
การสำรวจพื้นที่ที่จัดเก็บข้อมูลในการวิจัย การรวบรวมข้อมูลในการวิจัย ได้เลือกพื้นที่ในเมืองจังหวัดจันทบุรีและเจาะจงพื้นที่ที่ผู้คนจะสัญจรไปมาจำนวนมากในแต่ละวัน และการออกแบบและพัฒนาระบบจากนั้นผู้วิจัยจึงได้นำโมเดล YOLOv5 ดังกล่าวมาประยุกต์ใช้ในการรู้จำ ดังนั้นขั้นตอนการดำเนินงานวิจัยในภาพที่ 1



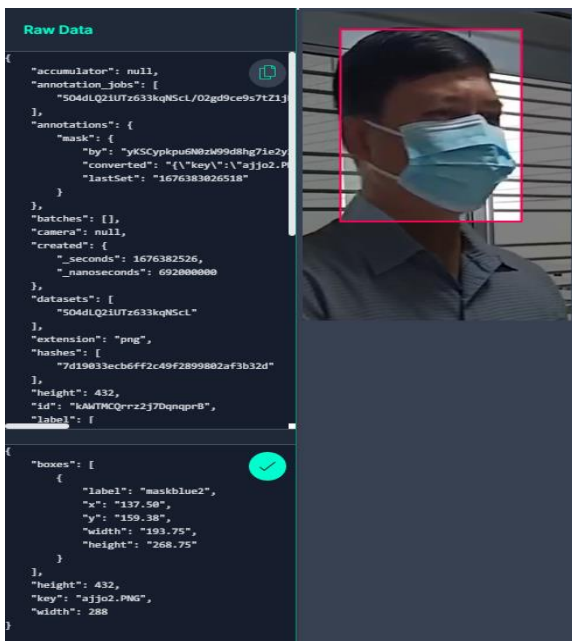
ภาพที่ 1: ขั้นตอนการดำเนินงานวิจัย

จากขั้นตอนการดำเนินงานวิจัยในภาพที่ 1 การเก็บรวบรวมชุดข้อมูลประชากรในการสัญจรยังพื้นที่ทำการทดลองระบบนั้นจะต้องได้รับการอนุญาตจากเจ้าของข้อมูลก่อนดำเนินการดังนั้นทางทีมผู้วิจัยจึงได้รวบรวมข้อมูลที่เป็นชุดข้อมูลตัวอย่างมาเพื่อฝึกฝนข้อมูลเบื้องต้นดังภาพที่ 2 จากการรวบรวมข้อมูลภาพถ่ายจำนวนชุดข้อมูลฝึกจะถูกนำไปปรับขนาดภาพถ่ายที่จะใช้โดยมีขนาดภาพถ่ายเท่ากับ 640x640 พิกเซลเท่ากันทุกภาพ จากนั้นทำป้ายกำกับรูปภาพ (Label) ดังภาพที่ 3 เมื่อได้ภาพมาแล้วในกระบวนการต่อไปคือการทำป้ายโดยเป็น

ขั้นตอนของการกำหนดตำแหน่งและขนาดของวัตถุในภาพ
 ขั้นตอนนี้จะใช้หน่วยความจำหลักของเครื่องคอมพิวเตอร์
 ที่กำลังดำเนินการและหน่วยประมวลผลกลางที่สูง ซึ่งอาจ
 ใช้เวลานานในขั้นตอนนี้ดังภาพตัวอย่าง



ภาพที่ 2: ตัวอย่างภาพที่ใช้ในการฝึกฝน



ภาพที่ 3: ตัวอย่างภาพทดสอบการฝึกและทำป้ายข้อมูล

การเขียนป้ายกำกับในภาพที่จัดทำกำหนดตำแหน่ง
 และขนาดดังตัวอย่างชุดคำสั่งด้านล่างนี้

```
{
  "boxes": [
    {
      "label": "maskblue2",
      "x": "137.50",
      "y": "159.38",
      "width": "193.75",
      "height": "268.75"
    }
  ],
  "height": 432,
  "key": "ajjo2.PNG",
  "width": 288
}
```

4. ผลดำเนินการวิจัย

การดำเนินการทดลองระบบได้กำหนดขอบเขตของ
 สภาวะแวดล้อมเป็นบริเวณเฉพาะที่คนเดินผ่านเท่านั้นใน
 การทดลองระบบ โดยเป็นการเดินแบบที่มีทิศทางเดิน
 ในแนวตั้งของวิดีโอ ในส่วนของระบบที่ตั้งอุปกรณ์ไว้
 เพื่อให้ง่ายต่อการตรวจจับทิศทางเดิน ได้กำหนด
 ระยะห่างไม่เกิน 5 เมตรและพื้นที่ว่างให้เหมาะสมกับระยะ
 ความสูงของกล้องตรวจจับวิดีโอที่ 480 พิกเซล และ
 สอดคล้องกับความเร็วในการประมวลผลจาก
 ความสามารถในการประมวลผลที่มีอย่างจำกัดจากวิดีโอที่
 ใช้ในงานวิจัยมีความยาวทั้งสิ้นไม่เกิน 3 นาที ซึ่งจาก
 กระบวนการฝึกเป็นขั้นตอนที่ต้องใช้มุมมองภาพให้ได้ครบทุก
 มุมเพื่อความมีประสิทธิภาพในการตรวจจับ จึงได้ผลตาม
 ภาพตัวอย่างที่ 4



ภาพที่ 4: ตัวอย่างภาพการตรวจจับผู้สวมใส่หน้ากากอนามัย

โดยมีบุคคลตัวอย่างเดินคนเดียวเพื่อตรวจสอบความ
 แม่นยำในระยะของการตรวจจับวัตถุบุคคลและหน้ากาก
 อนามัย และการเดินพร้อมกัน 2 คน ผลรวมจำนวนนับคน

30 คน และตรวจวัดประสิทธิภาพของระบบโดยแบ่งได้ดังนี้ ประสิทธิภาพด้านความแม่นยำในการนับจำนวนโดยเปรียบเทียบจำนวนบุคคลที่นับได้จากระบบกับจำนวนบุคคลทั้งหมดที่เดินจริงทั้งหมด อัลกอริทึม YOLO ตรวจจับบุคคลจากวิดีโอที่บันทึกมาจากกล้องเว็บแคม (Webcam) โดยผลลัพธ์ของการตรวจจับมีลักษณะเป็นข้อมูลชนิดของวัตถุและตำแหน่งของวัตถุ จากนั้นพิจารณาเฉพาะวัตถุที่เป็นบุคคล โดยคำนวณจากการเปลี่ยนแปลงของตำแหน่งของจุดกึ่งกลางวัตถุ จากการคำนวณจะได้ทิศทางเคลื่อนที่ของวัตถุและเมื่อวัตถุเคลื่อนที่ผ่าน เมื่อนำข้อมูลการทดลองมาวิเคราะห์ผลแสดงดังตารางที่ 1 โดยมีการกำหนดการทดสอบจำนวน 10 ครั้ง ในแต่ละเหตุการณ์ซึ่งเหตุการณ์แรกทดสอบตรวจจับผู้ไม่สวมใส่แบบเดี่ยวเมื่อได้ผลการทดลองแล้ว ให้ดำเนินการในเหตุการณ์ตรวจจับผู้สวมใส่หน้ากากอนามัยเดี่ยว เมื่อครบจำนวนครั้งที่ได้กำหนดให้ทดลองไม่สวมใส่หน้ากากอนามัยผู้สวมใส่หน้ากากอนามัยและผู้ไม่สวมใส่หน้ากากอนามัยปนในกลุ่ม

ตารางที่ 1: การหาค่าความถูกต้องตามเหตุการณ์ตรวจจับหน้ากากอนามัยแบบเรียลไทม์

เหตุการณ์ การ ตรวจจับ ของระบบ	การทดสอบครั้งที่										ค่าเฉลี่ย ความ แม่นยำ (%)
	1	2	3	4	5	6	7	8	9	10	
ไม่สวมใส่ หน้ากาก อนามัย เดี่ยว	/	/	/	/	/	/	/	/	/	/	100
สวมใส่ หน้ากาก อนามัย เดี่ยว	/	/	/	/	/	/	/	/	/	/	100

ตารางที่ 2: การหาค่าความถูกต้องตามเหตุการณ์ตรวจจับหน้ากากอนามัยแบบเรียลไทม์ (ต่อ)

เหตุการณ์ การ ตรวจจับ ของระบบ	การทดสอบครั้งที่										ค่าเฉลี่ย ความ แม่นยำ (%)
	1	2	3	4	5	6	7	8	9	10	
ไม่สวมใส่ หน้ากาก อนามัยผู้	/	/	/	/	/	/	/	x	/	/	90
สวมใส่ หน้ากาก อนามัยผู้	/	/	/	/	/	/	/	/	/	/	100
ไม่สวมใส่ หน้ากาก อนามัยปน ในกลุ่ม	x	/	/	/	/	/	/	/	/	/	90
ค่าเฉลี่ยแบบเรียลไทม์											96

5. สรุป

จากการศึกษาข้อมูลในตารางที่ 1 พบว่าชุดข้อมูลที่ใช้ประกอบด้วยรูปภาพของคนที่มีหน้ากากและไม่มีหน้ากากที่ดึงมาจากเครื่องมือระบบมีการคำนวณประสิทธิภาพของการตรวจจับหน้ากากหลังจากดำเนินการตรวจสอบความถูกต้องแม่นยำ 96% โดยใช้เทคนิคโมเดล YOLOv5 แต่เนื่องจากความละเอียดของภาพเมื่อมีการปรับขนาดแล้ว การกำหนดตำแหน่งจากระยะการตรวจจับจึงน้อยกว่าการใช้เทคนิคแบบ K-cross อัลกอริทึมทั้งหมดให้ความแม่นยำสูงกว่า 90% โดย CNN นั้นสูงที่สุดด้วยความแม่นยำ 98.6% [6] แนวทางการพัฒนาศึกษาต่อหากพบบุคคลที่ไม่สวมหน้ากาก เจ้าหน้าที่ที่เกี่ยวข้องจะได้รับข้อความและข้อความเสียงเตือนให้ "สวมหน้ากาก" ให้กับบุคคลนั้น ชุดข้อมูลที่รวบรวมจากแหล่งต่างๆ ประกอบด้วย ภาพบุคคลที่สวมหน้ากากและไม่สวมหน้ากาก สิ่งนี้จะใช้ในการฝึกอบรมสถาปัตยกรรมการเรียนรู้เชิงลึก [7] ในการตรวจจับหน้ากากโดยอัตโนมัติ จำเป็นต้องมีระบบกล้องวงจรปิดที่มีน้ำหนักเบา คุ่มค่า ทนทาน และระบบวิดีโอตามเทคนิคอัลกอริทึมแบบ Convolution Neural Networks (CNN) ซึ่งงานวิจัยนี้นำเสนอระบบอัตโนมัติที่มีอุปกรณ์

น้ำหนักเบา และประหยัดต้นทุนสำหรับการตรวจจัดการ
จำแนกประเภทใบหน้าและหน้ากากโดยโมเดล
MobileNetV2 ซึ่งเป็นตัวตรวจจับแบบ single-shot (SSD)
ที่ใช้ตัวแยกประเภทแบบไบโนรีถูกนำมาใช้ในระบบ
อัตโนมัติเครื่องตรวจจับใบหน้าและตัวแยกประเภท
หน้ากากแบบต่างๆ ตลอดจนความแตกต่างระหว่างรุ่น
เหล่านี้ความแม่นยำ ใช้เพื่อประเมินความสำเร็จของระบบ
อัตโนมัติสูงถึง 98% และความแม่นยำ 95.85 % [8]
สอดคล้องกับงานวิจัยที่ทดลองความแม่นยำในการ
ฝึกรูปแบบสูงสุดได้รับในกรณีของ MTCNN ความแม่นยำ
ในการทดสอบสูงสุดคือ 98.6% [9]

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จได้ด้วยดีเนื่องจากได้รับทุน
สนับสนุนจากสถาบันวิจัยและพัฒนา มหาวิทยาลัยราชภัฏ
รำไพพรรณีที่ได้สนับสนุนให้ทีมผู้วิจัยในการให้คำแนะนำ
การเก็บรวบรวมข้อมูลและทำการวิจัยร่วมกับท้องถิ่นได้
เป็นอย่างดี

เอกสารอ้างอิง

- [1] M.N. Kavitha et al. "Face Mask Detection Using Deep Learning". *International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. (2022). DOI: 10.1109/ICAIS53314.2022.9742825
- [2] Zhu, X., Lyu, S., Wang, X., and Zhao, Q. (2021). TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. ArXiv, abs/2108.11539.
- [3] Redmon, J. and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR-2018)*, pp. 1-6.
- [4] ปริญญา ดันทวิวัฒน์ และคณะ. "การวิเคราะห์ภาพคนและสัมภาระสำหรับการตรวจจับวัตถุที่ปราศจากเจ้าของ", *Journal of Engineering and Digital Technology (JEDT)*, Vol.9 No.2, pp.49-60, 2021.
- [5] รามศรัวี ศิวรักษา และคณะ, "การตรวจจัดการสวมหน้ากากอนามัยโดยประยุกต์ใช้โครงข่ายประสาทเทียมแบบเชิงลึกและโครงข่ายประสาทเทียมแบบคอนโวลูชัน" *รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติ สำหรับนักศึกษา*

มหาวิทยาลัยราชภัฏกำแพงเพชร ครั้งที่ 1, 22 กุมภาพันธ์
2564, หน้า 671-680.

- [6] K. Anirudh et al, "Face Mask Detection Using Machine Learning". *International Students' Conference on Electrical, Electronics and Computer Science (SCEECS).IEEE*, (2022). DOI: 10.1109/SCEECS54111.2022.9740913
- [7] A. Mashyal et al, "Facial Mask Detection and Alert System". *International Conference on Intelligent Technologies (CONIT). IEEE*. (2021). DOI: 10.1109/CONIT51480.2021.9498278.
- [8] M. R. K. Sujon et al, "Real-time face mask detection for COVID-19 prevention". *Annual Computing and Communication Workshop and Conference (CCWC). IEEE* (2022). DOI: 10.1109/CCWC54503.2022.9720764
- [9] R. Gorhekarat et al, "A Survey on Covid-19 Face-Mask Detection Techniques". *International Conference on Computing, Power and Communication Technologies (GUCON). IEEE*. (2021). DOI: 10.1109/GUCON50781.2021.9573738

การปรับปรุงประสิทธิภาพการเฝ้าระวังสถานะสิ่งแวดล้อมที่เปลี่ยนแปลงน้อย ในการประมวลผลที่ขอบ

Performance Enhancement for Environmental Monitoring with Graduated Change Situation in Edge Computing

ธีรภัทร ดีประเสริฐ (Theerapat Deeprasert)¹ อภิรักษ์ จันทร์สร้าง (Aphirak Jansang)² ชัยพร ใจแก้ว (Chaiporn Jaikaeo)³

วิรัช ตั้งตรงไพโรจน์ (Withawat Tangtrongpairoj)⁴ และอนันต์ ผลเพิ่ม (Anan Phonphoem)⁵

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์(วิทยาเขตบางเขน)

¹theerapat.de@ku.th, ²aphirak.j@ku.ac.th, ³chaiporn.j@ku.ac.th, ⁴withawat.t@ku.ac.th, ⁵anan.p@ku.ac.th

บทคัดย่อ

ในการเก็บข้อมูลสิ่งแวดล้อมในที่ห่างไกล การใช้พลังงานอย่างประหยัดเป็นสิ่งสำคัญเพื่อให้อุปกรณ์ใช้งานได้ยาวนานขึ้นในที่ที่ไม่มีไฟฟ้า การส่งข้อมูลสถานะแวดล้อมที่มีการเปลี่ยนแปลงน้อยมีการส่งข้อมูลซ้ำและสิ้นเปลืองพลังงานโดยไม่จำเป็น งานวิจัยนี้นำเสนอวิธีการส่งข้อมูลโดยให้อุปกรณ์ส่งเฉพาะข้อมูลที่จำเป็นเพื่อลดการใช้พลังงานของอุปกรณ์ จากการทดลองเมื่อลดจำนวนการส่งแพ็คเกจข้อมูลที่ไม่จำเป็น ส่งผลให้มีการใช้พลังงานลดลง

คำสำคัญ: การประมวลผลขอบ การเฝ้าสังเกตสิ่งแวดล้อม

Abstract

Efficient power usage is crucial when collecting environmental data from remote locations where power may be limited. However, excessive data transmission can lead to unnecessary power consumption in environments where minor changes occur. This research proposes a method for transmitting data that involves sending only the essential information to minimize the device's power consumption. The experiments indicate that reducing the number of unnecessary data packets sent results in lower energy consumption.

Keywords: Edge Computing, Environment Monitoring.

1. บทนำ

การประมวลผลขอบได้ถูกประยุกต์เข้าไปในงานวิจัยต่าง ๆ เนื่องจากการประหยัดพลังงานที่มากจากการส่งการประมวลผลไปที่ส่วนอื่นของระบบและดีเลย์ที่มีน้อยเนื่องจากข้อมูลที่ประมวลผลอยู่ใกล้กับส่วนที่ใช้งาน ในงานที่ใช้การประมวลผลขอบการเฝ้าสังเกตสิ่งแวดล้อมเป็นหนึ่งในกระบวนการที่ได้รับประโยชน์ เนื่องจากการสังเกตสิ่งแวดล้อมเซนเซอร์โดยส่วนใหญ่ถูกติดตั้งในที่ห่างไกลและไม่ได้ถูกเชื่อมต่อกับแหล่งไฟฟ้า พลังงานไฟฟ้าจึงเป็นทรัพยากรที่สำคัญ การเปลี่ยนแปลงของสภาพแวดล้อมบางประเภทมีการเปลี่ยนแปลงน้อยในระยะเวลาอันยาวนาน หากใช้วิธีการเก็บค่าข้อมูลปกติเซนเซอร์จะส่งข้อมูลเดิมซ้ำและสิ้นเปลืองพลังงาน

ในงานวิจัยนี้นำเสนอการส่งข้อมูลสิ่งแวดล้อมอย่างมีประสิทธิภาพในสภาพแวดล้อมที่มีการเปลี่ยนแปลงน้อยบนระบบการประมวลผลขอบ โดยประเมินข้อมูลที่รับค่าเข้ามาล่าสุดหากค่าความแตกต่างไม่เกินค่าที่กำหนดไว้เซนเซอร์จะไม่ส่งข้อมูลและแอคเซสพอยต์จะคำนวณว่าค่าข้อมูลไม่มีการเปลี่ยนแปลงและใช้ค่าเดิม แต่ถ้าหากมีการเปลี่ยนแปลงมากกว่าค่าที่กำหนดเซนเซอร์จะส่งข้อมูลตามปกติ

2. วรรณกรรมที่เกี่ยวข้อง

ในงานวิจัยเกี่ยวกับการประมวลผลขอบสิ่งที่เป็นปัญหาหลักนั้นคือพลังงาน หลายงานวิจัยพยายามที่จะหาวิธีในการลดการใช้พลังงานด้วยวิธีต่าง ๆ

โครงสร้างการทำงานของกรอบการประมวลผลที่ขอบ (Framework) สามารถถูกเปลี่ยนให้มีประสิทธิภาพมากขึ้น เช่น การเปลี่ยนวิธีการเชื่อมต่อของโหนดเพื่อให้ใช้งานกับโหนดที่มีการเคลื่อนที่ไกล [1], งานวิจัย [2] จัดการกับวิธีการเลือกช่องสัญญาณการสื่อสารให้มีการใช้พลังงานน้อยลง

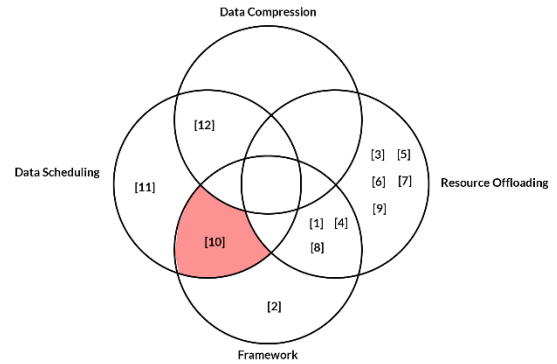
การส่งการประมวลผลไปที่ส่วนอื่น (Offloading) ทำให้โหนดไม่ต้องใช้พลังงานกับการประมวลผลงานวิจัย [3] เสนอวิธีการ Offloading เหมาะกับวิธีการคำนวณข้อมูลเพื่อที่จะประหยัดพลังงานได้มากที่สุด ได้ถูกนำเสนอ, การเอา Cache เข้ามาช่วยในการส่งการประมวลผล [4], การ Offload ข้อมูลใน Network ที่มีการใช้งาน Device-to-Device [5], วิธีการ Offloading บน Network ประเภท Non-Orthogonal Multiple Access ให้มีการใช้พลังงานน้อยลง [6], การใช้งาน Deep learning ในการช่วย Offloading [7], การเปลี่ยนโหนดระหว่างคำนวณข้อมูลเองและส่งข้อมูลไปประมวลผลที่อื่น [8], การเลือกวิธีการ Offloading โดยมีพลังงานเป็นตัวแปรการตัดสินใจ [9]

การตั้งเวลาการตื่นหลับของโหนด (Scheduling) ลดการใช้พลังงานของโหนด โดยให้โหนดหลับในตอนที่ไม่ทำงานลดการใช้พลังงานที่ไม่จำเป็นลงได้ วิธีการ Scheduling ใหม่ถูกนำเสนอเช่น วิธีการที่จะทำให้โหนดอยู่ในโหมดครั้งหลับเพื่อจะประหยัดพลังงานรักษา Quality of Service ได้โดยให้โหนดตื่นอยู่ตลอดเวลาการส่งข้อมูลนั้นทำใน Virtual Machine ซึ่งอยู่ในโหนดอีกที [10], การเลือกโหนดที่หลับโดยใช้ข้อมูลการประมวลผลเป็นตัวแปรการตัดสินใจ [11]

นอกจากนี้ปริมาณของข้อมูลส่งผลกระทบต่อพลังงานที่ใช้ในการส่งข้อมูล การย่อข้อมูล (Compression) จึงเป็นการประหยัดพลังงานได้อีกอย่างหนึ่ง งานวิจัย [12] เสนอวิธีการย่อข้อมูลโดยการลดจำนวนการส่งลงแล้วใช้การรวมข้อมูล

ที่ลดลงนั้นให้อยู่ในรูปแบบของ Polynomial Regression แล้วส่งเพียงแค่ค่าสัมประสิทธิ์เพื่อให้ส่วนกลางสร้างค่าของโหนดขึ้นมาใหม่เอง

แนวโน้มในการวิจัยเรื่องการประมวลผลขอบเพื่อลดการใช้พลังงานสามารถจัดให้อยู่เป็นกลุ่มและแสดงพื้นที่ที่งานวิจัยนี้พยายามที่จะแก้ไขในพื้นที่สีแดงได้ดังภาพที่ 1



ภาพที่ 1: แผนภาพแสดงวัตถุประสงค์ของงานวิจัยตัวอย่างที่มีเป้าหมายในการลดการใช้พลังงานของการประมวลผลขอบ

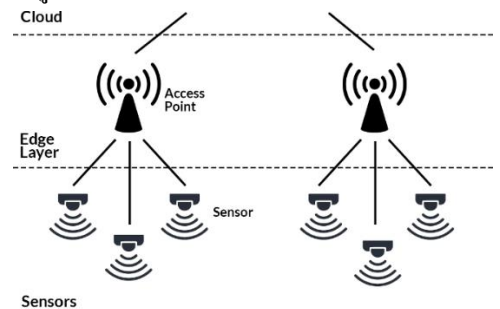
3. วิธีการดำเนินการวิจัย

3.1 คำชี้แจงปัญหา

ในสภาพสิ่งแวดล้อมที่มีการเปลี่ยนแปลงน้อย การส่งข้อมูลแบบทั่วไปจะส่งข้อมูลซ้ำซ้อนบ่อยทำให้เสียพลังงานโดยเปล่าประโยชน์

3.2 ส่วนที่เปลี่ยนแปลงในระบบ

งานวิจัยนี้ต้องการลดจำนวนการส่งข้อมูลในสภาพแวดล้อมที่มีการเปลี่ยนแปลงน้อยจึงเปลี่ยนวิธีการไหลของข้อมูลของเซนเซอร์และแอคเซสพอยต์ดังนี้

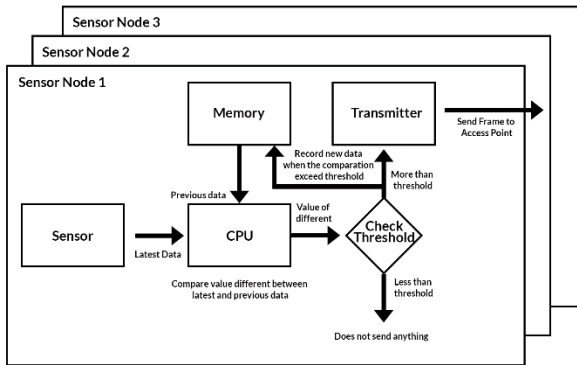


ภาพที่ 2: รูปแบบการเชื่อมต่อของระบบการประมวลผลขอบ

(a.) โหนดเครื่องรับค่าข้อมูล (Sensor nodes)

ติดตั้งไว้ที่ปลายสุดของระบบเพื่อเก็บข้อมูลดังภาพที่ 2 เซนเซอร์ตรวจรับค่าอุณหภูมิและนำมาเปรียบเทียบกับอุณหภูมิที่ถูกส่งไปครั้งล่าสุด หากข้อมูลที่ตรวจรับเข้ามา

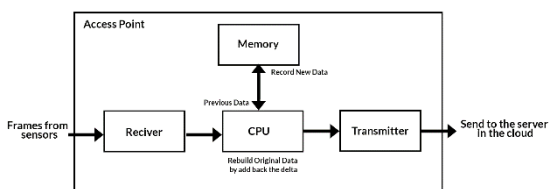
ใหม่มีความต่างกันมากกว่า Threshold เซนเซอร์จะบันทึกข้อมูลอุณหภูมิที่ได้รับใหม่แทนข้อมูลเดิมและทำการส่งข้อมูลอุณหภูมินั้นต่อไปให้แอคเซสพอยต์ แต่ถ้าหากว่าข้อมูลที่ตรวจรับมีความต่างน้อยกว่า Threshold จะไม่มีการดำเนินการเกิดขึ้น วิธีการไหลของข้อมูลสามารถอธิบายเป็นรูปการไหลของข้อมูลได้ดังภาพที่ 3



ภาพที่ 3: การไหลของข้อมูลในโหนดรับค่าข้อมูล

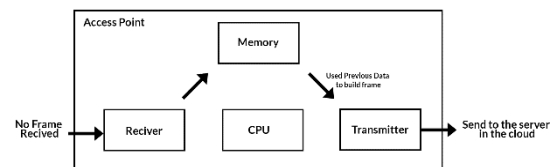
(b.) โหนดจุดเชื่อมต่อ (Access Point)

โหนดจุดเชื่อมต่อรอให้มีข้อมูลถูกส่งมาจากเซนเซอร์และส่งต่อข้อมูลไปที่ฐานผ่าน Cloud ดังภาพที่ 2 ข้อมูลเพื่อการวิเคราะห์ต่อไปและบันทึกข้อมูลที่ได้รับมาล่าสุดไว้ดังภาพที่ 4



ภาพที่ 4: การไหลของข้อมูลในโหนดจุดเชื่อมต่อเมื่อโหนดได้รับข้อมูล

ในกรณีที่โหนดจุดเชื่อมต่อไม่ได้รับข้อมูลจากเซนเซอร์ในเวลาที่กำหนด โหนดจุดเชื่อมต่อจะสร้างข้อมูลซ้ำจากข้อมูลที่บันทึกไว้ล่าสุดแล้วส่งข้อมูลนั้นต่อไปให้กับฐานข้อมูล นับว่าข้อมูลที่รับมาเป็นข้อมูลซ้ำดังภาพที่ 5



ภาพที่ 5: การไหลของข้อมูลในโหนดจุดเชื่อมต่อเมื่อโหนดไม่ได้รับข้อมูล

3.3 การเตรียมการเพื่อทดสอบ

ในการทดสอบเพื่อบันทึกผลประสิทธิภาพของวิธีการที่นำเสนอเซนเซอร์จำนวนสามโหนดถูกกำหนดขึ้นเพื่อจำลองกลุ่ม เซนเซอร์โหนดในพื้นที่และให้ทั้งสาม

โหนดส่งข้อมูลไปที่แอคเซสพอยต์หนึ่งโหนดปลายทางเพื่อบันทึกผล

3.4 ขั้นตอนการทดสอบ

กำหนดค่า Threshold ที่ใช้ในการทดสอบแล้วจึงกำหนดข้อมูลอุณหภูมิให้โหนดเซนเซอร์ตัดสินใจว่าจะทำการส่งข้อมูลหรือไม่ หลังจากข้อมูลส่งไปจะมีการจำลองว่า แอคเซสพอยต์ได้รับข้อมูลหรือไม่จากค่าการตกหล่นของข้อมูลที่กำหนด จนครบข้อมูลอุณหภูมิ หลังจากนั้นทำการทำซ้ำโดยกำหนดค่า Threshold และ ค่าการตกหล่นของข้อมูลใหม่เพื่อเปรียบเทียบผล

3.5 รายละเอียดข้อมูลที่ใช้ในการทดสอบ

ตารางที่ 1: ค่าตัวแปรเพื่อจำลองการส่งข้อมูลในการทดสอบ

Platform การทดสอบ	NS-3 Network Simulator
ระยะเวลาระหว่าง Packet	1 วินาที
ขนาดของ Packet	200 ไบต์
ระยะห่างระหว่าง โหนด	100 เมตร
พลังงานเริ่มต้นของโหนด	1000 จูล

การทดสอบเพื่อนำผลมาวิเคราะห์ในงานวิจัยนี้ได้ทำการสร้างรูปแบบการเชื่อมต่อตามที่กล่าวไว้ในหัวข้อที่ 3.3 บน NS-3 Network Simulator โดยค่าตัวแปรสำคัญในการทำงานของการทดสอบถูกตั้งค่าตามตารางที่ 1

ข้อมูลภูมิอากาศที่ใช้ในการทดสอบนี้เป็นข้อมูลสภาพอากาศรวมตามภูมิศาสตร์รายชั่วโมงของทวีปยุโรปของโครงการ Open Power System Data ซึ่งเป็นการใช้ข้อมูลของโครงการ NASA MERRA-2 อีกทีหนึ่ง [13]

ในการทดสอบ ข้อมูลที่ใช้เป็นข้อมูลสภาพอากาศรวมตามภูมิศาสตร์รายชั่วโมงของประเทศออสเตรเลียระหว่างวันที่ 1 ถึง 15 สิงหาคม ปี ค.ศ. 2019

4. ผลการดำเนินงาน

4.1 ความซับซ้อนของอัลกอริทึมงานวิจัย

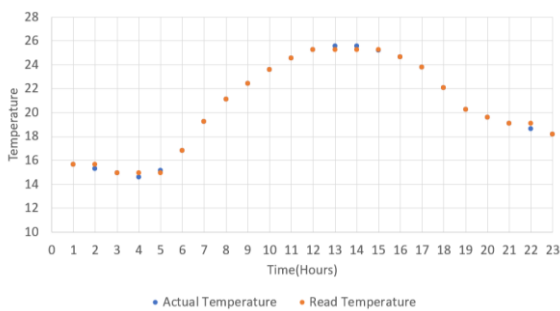
ความซับซ้อนของอัลกอริทึมคือ $O(n)$ ซึ่ง n คือจำนวนครั้งที่เซนเซอร์ตรวจจับวัดอุณหภูมิ

4.2 ข้อมูลความแม่นยำของอุณหภูมิ

เพื่อที่จะสังเกตความเปลี่ยนแปลงของความแม่นยำเมื่อเพิ่มขนาดของค่า Threshold การทดสอบจึงทดลองบนค่า Threshold ที่ค่า 0.5, 1, 1.5, และ 2 องศาเซลเซียส

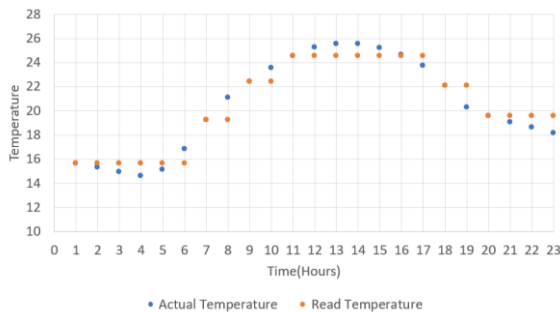
หลังจากการทดสอบ ความแม่นยำของข้อมูลที่เซนเซอร์ส่งเมื่อเทียบกับข้อมูลอากาศจริงนั้นลดลงเมื่อขนาดของ Threshold เพิ่มขึ้น

เพื่อให้เห็นความแตกต่างของการส่งข้อมูลระหว่าง Threshold ขนาดเล็กและใหญ่ กราฟต่อไปนี้จะเป็นการเปรียบเทียบข้อมูล อุณหภูมิจริงในหนึ่งวันและอุณหภูมิที่เซนเซอร์ส่งไปที่แอสเซสพอยต์ของค่า Threshold 0.5 และ 2 องศา



ภาพที่ 6: การเปรียบเทียบระหว่างอุณหภูมิจริงและอุณหภูมิ

ณ ฐานข้อมูลโดยมีค่า Threshold 0.5°C



ภาพที่ 7: การเปรียบเทียบระหว่างอุณหภูมิจริงและอุณหภูมิ

ณ ฐานข้อมูลโดยมีค่า Threshold 2°C

ค่า Threshold ขนาดใหญ่ทำให้ความเปลี่ยนแปลงเล็กน้อยไม่ถูกนับว่าเป็นการเปลี่ยนแปลงที่สำคัญพอทำให้ค่าข้อมูลจริงกับข้อมูลที่ส่งแตกต่างกันดังภาพที่ 6 และ 7

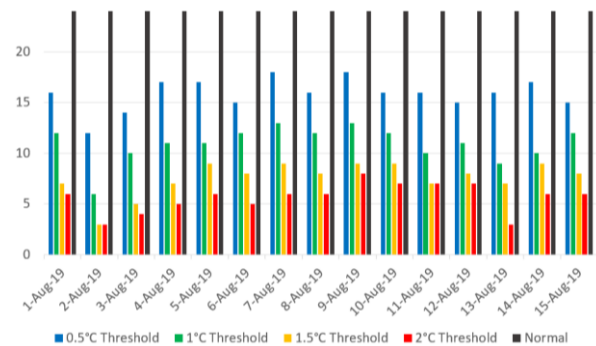
ค่าความผิดพลาดถูกคำนวณโดยนำเอาค่าความแตกต่างของแต่ละชั่วโมงตั้งแต่วันที่ 1 ถึง 15 มาเฉลี่ยและได้ผลดังตารางที่ 2

ตารางที่ 2: ค่าความผิดพลาดของการส่งข้อมูลอุณหภูมิที่มี การกำหนดค่า Threshold ต่างกัน

0.5 °C Threshold's Error	1 °C Threshold's Error	1.5 °C Threshold's Error	2 °C Threshold's Error
±0.0866°C	± 0.2872°C	±0.5017°C	0.6817°C

4.3 จำนวนข้อมูลที่ส่ง

ด้วยวิธีการส่งข้อมูลที่งานวิจัยนี้นำเสนอเซนเซอร์จะส่งเฉพาะข้อมูลที่มีความแตกต่างมากพอที่จะเรียกได้ว่ามีความสำคัญ ทำให้จำนวนข้อมูลที่ส่งลดลงอย่างเห็นได้ชัด

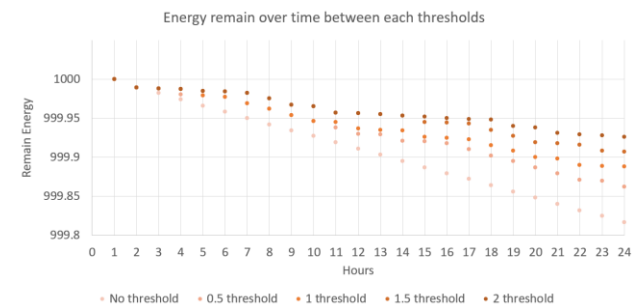


ภาพที่ 8: จำนวนข้อมูลที่ส่งในแต่ละวันของ Threshold 0.5, 1, 1.5, 2, และการส่งแบบไม่คัดกรอง

จากภาพที่ 8 ทำให้เห็นว่าในการทดลอง ใช้ Threshold ขนาดเล็กสามารถลดปริมาณการส่งของข้อมูลให้อยู่ในปริมาณเฉลี่ย 15 ครั้งต่อวันจากการส่งข้อมูลแบบไม่มีการใช้ Threshold อยู่ที่ 24 ครั้งต่อวัน และสามารถลดปริมาณลงได้อีกเมื่อขนาดของ Threshold เพิ่มขึ้น

4.4 พลังงานที่ใช้ในการทำงาน

พลังงานที่เหลือในโหมดเซนเซอร์จากการทดสอบในค่า Threshold ที่ต่างกันถูกแสดงบนกราฟเปรียบเทียบดังนี้



ภาพที่ 9: พลังงานที่โหนดเหลือในหนึ่งวันของค่าความสำคัญ 0.5, 1, 1.5, 2, และการส่งแบบไม่คัดกรอง

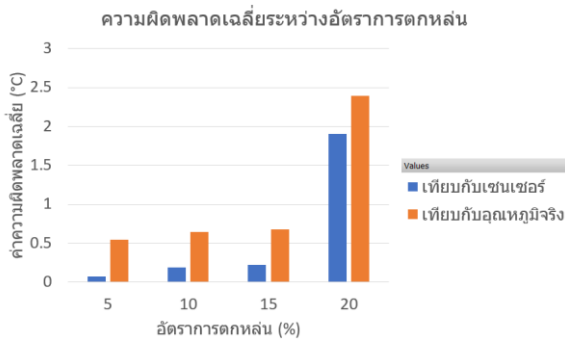
จากภาพที่ 9 การใช้พลังงานของการส่งข้อมูลแบบไม่ใช้ค่า Threshold มีลักษณะเส้นตรงและอยู่ต่ำที่สุดจากทุกประเภทการทดลองเนื่องจากการส่งข้อมูลทุกชั่วโมง เส้นกราฟที่เหลือเป็นพลังงานที่เหลือจากการส่งข้อมูลแบบมีการใช้ Threshold ซึ่งมีการส่งน้อยกว่าหนึ่งครั้งต่อชั่วโมง

จากหัวข้อที่ 4.3 Threshold ขนาดใหญ่ขึ้นทำให้จำนวนข้อมูลที่ส่งน้อยลง ซึ่งสอดคล้องกับกราฟในภาพที่ 9 การส่งข้อมูลที่มี Threshold ใหญ่จะทำให้พลังงานเหลือมีมากกว่าวิธีการส่งที่ใช้ Threshold เล็ก

4.5 การสูญหายของข้อมูล

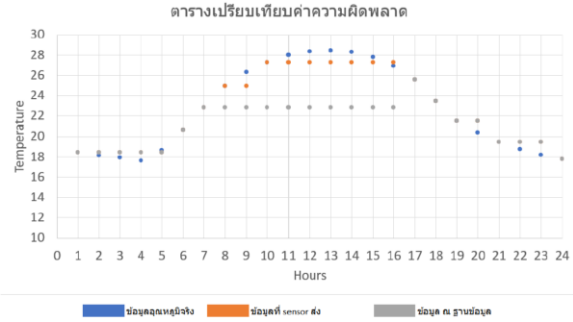
เพื่อทดสอบการทำงานของวิธีการส่งที่นำเสนอในสภาพแวดล้อมที่มีการตกหล่นของข้อมูลจึงได้มีการบันทึกผลการส่งข้อมูลโดยจำลองการตกหล่นของข้อมูลโดยมีโอกาสดังกล่าวของข้อมูลที่ 5, 10, 15, และ 20 เปอร์เซ็นต์และบันทึกผลความแม่นยำ

โดยความแม่นยำของข้อมูล ณ ฐานข้อมูลถูกเปรียบเทียบกับข้อมูลที่เซนเซอร์และข้อมูลจริง โดยใช้ข้อมูลอากาศทุกชั่วโมงเป็นเวลา 10 วัน และใช้ค่า Threshold อยู่ที่ 1.5 องศา



ภาพที่ 10: ค่าความผิดพลาดเฉลี่ยเมื่ออัตราการตกหล่นของข้อมูลเปรียบเทียบกับเซนเซอร์และอุณหภูมิจริง

จากข้อมูลบนภาพที่ 10 ทำให้เห็นว่าค่าความผิดพลาดเพิ่มสูงขึ้นตามอัตราการตกหล่นของข้อมูล โดยคาดว่าเวลาที่อัตราการตกหล่นเพิ่มขึ้นจะทำให้การส่งข้อมูลตกหล่นตรงกับช่วงที่ค่าการเปลี่ยนแปลงของข้อมูลน้อยทำให้เกิดความผิดพลาดที่มีจำนวนมาก



ภาพที่ 11: ค่าข้อมูลอุณหภูมิจริง, อุณหภูมิที่โหนดส่ง, และอุณหภูมิที่ฐานข้อมูลได้รับโดยมีอัตราการตกหล่น 20%

กราฟในภาพที่ 11 เป็นข้อมูลอุณหภูมิหนึ่งวันของการทดสอบที่มีอัตราการตกหล่นของข้อมูลที่ 20% ข้อมูลวันดังกล่าวมีการตกหล่นของข้อมูล ณ ชั่วโมงที่ 8 และ 10 ซึ่งตรงกับช่วงที่มีการเปลี่ยนแปลงของอุณหภูมिन้อย จึงคาดว่าเป็นเหตุผลที่การทดลองที่มีเงื่อนไขข้างต้นมีค่าความผิดพลาดมากกว่าการทดลองครั้งอื่นมาก

5. สรุป

ในสภาพสิ่งแวดล้อมที่มีการเปลี่ยนแปลงน้อย การส่งข้อมูลแบบทั่วไปจะส่งข้อมูลซ้ำซ้อนบ่อยทำให้สิ้นเปลืองพลังงาน งานวิจัยนี้ได้นำเสนอวิธีการให้ข้อมูลที่ถูส่งจากเซนเซอร์ส่งเฉพาะข้อมูลที่มีความเปลี่ยนแปลงที่นับว่ามีความสำคัญเท่านั้น ทำให้การส่งข้อมูลของเซนเซอร์น้อยลงและลดการใช้พลังงานของโหนดลงไปด้วย จาก การทดสอบทำให้เห็นว่าวิธีที่นำเสนอสามารถลดจำนวนข้อมูลที่ส่งได้จริงและถ้าหากเพิ่มขนาดของ Threshold จำนวนข้อมูลที่ส่งก็จะลดลงไปอีก แต่จะลดความแม่นยำของอุณหภูมิ ณ ฐานข้อมูลลง

จากปัญหาที่เกิดขึ้นในหัวข้อที่ 4.5 ทำให้เห็นว่าในสภาพแวดล้อมที่มีการตกหล่นของข้อมูลสูงจะทำให้มีโอกาสที่ข้อมูล ณ ฐานข้อมูลมีความผิดพลาดสูงเนื่องจากข้อมูลไม่ได้รับข้อมูลการเปลี่ยนแปลงจากเซนเซอร์เมื่อมีการตกหล่นของข้อมูลตรงกับช่วงที่สภาพแวดล้อมมีการเปลี่ยนแปลงน้อย การสื่อสารระหว่างเซนเซอร์และแอคเซสพอยต์เพื่อตรวจสอบความถูกต้องของข้อมูล ณ ฐานข้อมูลจึงเป็นขั้นตอนต่อไปที่จะพัฒนางานวิจัยนี้ให้มีความผิดพลาดน้อยลงจากการตกหล่นของข้อมูล

นอกจากนี้ค่า Threshold เพียงค่าเดียวอาจไม่เหมาะสมกับสภาพแวดล้อม การปรับเปลี่ยนค่า Threshold ตามตัวแปรอื่น (Adaptive Threshold) จะสามารถที่จะลดความผิดพลาดของข้อมูลลงโดยปรับตัวให้เข้ากับข้อมูลสภาพสิ่งแวดล้อมให้ได้มากที่สุด ขึ้นต่อไปจึงเป็นการทดสอบว่าควรใช้ตัวแปรใดในการปรับเปลี่ยน Threshold ให้เข้ากับข้อมูลของสภาพแวดล้อม เช่น ข้อมูลความแตกต่างของข้อมูลจริงและข้อมูลในหน่วยความทรงจำของเซนเซอร์สามารถใช้เป็นเกณฑ์เพื่อเปลี่ยนค่า Threshold ในเวลาจริงได้ หรือ กำหนดจำนวนของปริมาณการส่งที่ต้องการแล้วจึงเปลี่ยนค่า Threshold ให้มีปริมาณการส่งข้อมูลใกล้เคียงกับปริมาณที่กำหนดไว้โดยใช้ข้อมูลรอบก่อนหน้ามาเป็นตัวอ้างอิง

เอกสารอ้างอิง

- [1] Han, Q., G. Zheng, and C. Xu. Energy-Efficient Resource Allocation for Mobile Edge Computing in NOMA-Enabled Small Cell Networks. in 2020 IEEE 20th International Conference on Communication Technology (ICCT). 2020.
- [2] Kuo, W. and Y. Wang. An Energy-Saving Edge Computing and Transmission Scheme for IoT Mobile Devices. in 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE). 2019.
- [3] Chouhan, S. Energy Optimal Partial Computation Offloading Framework for Mobile Devices in Multi-access Edge Computing. in 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM). 2019.
- [4] Dong, L., et al. Edge Cache-Aided Computation Offloading for Mobile Cloud Computing. in 2021 IEEE/CIC International Conference on Communications in China (ICCC). 2021.
- [5] Cai, Y., et al. Optimal Task Allocation Policy for Energy Efficient D2D-enabled Mobile Edge Computing. in 2021 13th International Conference on Wireless Communications and Signal Processing (WCSP). 2021.
- [6] Xu, C., G. Zheng, and X. Zhao, Energy-Minimization Task Offloading and Resource Allocation for Mobile Edge Computing in NOMA Heterogeneous Networks. IEEE Transactions on Vehicular Technology, 2020. 69(12): p. 16001-16016.
- [7] Zhang, J., et al. Deep Reinforcement Learning for Offloading and Shunting in Hybrid Edge Computing Network. in 2021 IEEE International Conference on Communications Workshops (ICC Workshops). 2021.
- [8] Wu, X., et al. A Hybrid Mode Task Offloading Algorithm for Edge Computing. in 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). 2021.
- [9] Zaw, C.W., et al., Energy-Aware Resource Management for Federated Learning in Multi-Access Edge Computing Systems. IEEE Access, 2021. 9: p. 34938-34950.
- [10] Alnoman, A. and A. Anpalagan. QoS-aware Energy Saving Scheme and Traffic Management in Mobile Edge Computing Networks. in 2021 International Wireless Communications and Mobile Computing (IWCMC). 2021.
- [11] Alnoman, A. and A. Anpalagan, Computing-Aware Base Station Sleeping Mechanism in H-CRAN-Cloud-Edge Networks. IEEE Transactions on Cloud Computing, 2021. 9(3): p.958-967.
- [12] Rivet, F., L. Foucaud, and G. Ferré. Edge Computing Technique for a 87% Energy Saving for IoT Device Dedicated to Environmental Monitoring. in 2021 IEEE 12th Latin America Symposium on Circuits and System (LASCAS). 2021.
- [13] Global Modeling and Assimilation Office (GMAO) (2015), inst3_3d_asm_Cp: MERRA-2 3D IAU State, Meteorology Instantaneous 3-hourly (p-coord, 0.625x0.5L42), version 5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), Accessed Enter User Data Access Date at doi: 10.5067/VJAFPLIICSIV.

ความมั่นคงปลอดภัยข้อมูลโดยใช้การเข้ารหัสลับร่วมกับ การซ่อนข้อมูลในภาพ Data Security Using Encryption with Image Steganography

ธนวรรธน์ โยชนะนัง (Thanawat Yochanang)

ภาควิชาการบริหารเครือข่ายดิจิทัลและความมั่นคงปลอดภัยสารสนเทศ

คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

thanawat.y@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มุ่งเน้นการศึกษาความมั่นคงปลอดภัยข้อมูลโดยใช้เทคนิคการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพเพื่อป้องกันการดักจับข้อมูลจากบุคคลที่ไม่ได้รับอนุญาตให้เข้าถึง ในงานวิจัยนี้ผู้วิจัยได้นำเสนอการใช้เทคนิคการเข้ารหัสลับด้วยวิธีการมาตรฐานการเข้ารหัสลับขั้นสูง (Advanced Encryption Standard : AES) ร่วมกับการซ่อนข้อมูลในภาพดิจิทัลด้วยวิธีการบิตที่มีนัยสำคัญน้อยที่สุด (Least Significant Bit : LSB) โดยการซ่อนข้อมูลลับในภาพดิจิทัลนั้นจะไม่ทำให้ภาพดูเปลี่ยนแปลงและยังคงคุณภาพของภาพเดิมไว้ได้ ซึ่งข้อมูลที่ถูกซ่อนจะถูกเข้ารหัสลับก่อนที่จะถูกซ่อนลงในภาพดิจิทัล ผลการทดสอบหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าประมาณ 0 ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) มีค่าสูง และค่าดัชนีความคล้ายคลึงกันของโครงสร้าง (SSIM) เท่ากับ 1 แสดงให้เห็นว่าการใช้เทคนิคการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพนี้สามารถป้องกันการดักจับข้อมูลจากบุคคลที่ไม่ได้รับอนุญาตให้เข้าถึงได้อย่างมีประสิทธิภาพ

คำสำคัญ: การเข้ารหัสลับ การซ่อนข้อมูลในภาพ
ความมั่นคงปลอดภัยของข้อมูล

Abstract

This research focuses on studying data security by using encryption techniques and image steganography to prevent unauthorized access by individuals. In this research, the researchers proposed using

encryption techniques with the Advanced Encryption Standard (AES) combined with image steganography using the Least Significant Bit (LSB) method. Hiding confidential information in digital images does not change the appearance of the image, and the original image quality is preserved. The hidden data is encrypted before being concealed in the digital image. The Mean Squared Error (MSE) is approximately 0, the Peak Signal-to-Noise Ratio (PSNR) is high, and the Structural Similarity Index (SSIM) is 1, indicating that using encryption techniques combined with image steganography can effectively prevent unauthorized access to data.

Keywords: Encryption, Image Steganography, Data Security.

1. บทนำ

ในยุคปัจจุบันที่เทคโนโลยีสารสนเทศและการสื่อสารมีบทบาทสำคัญในชีวิตประจำวันของเราอย่างมาก การใช้งานอินเทอร์เน็ตเพื่อสื่อสารและส่งข้อมูลกันระหว่างบุคคลหรือองค์กรมีปริมาณที่มากขึ้นตลอดเวลา อย่างไรก็ตามการสื่อสารและการส่งข้อมูลนั้นมีความเสี่ยงที่จะถูกดักจับข้อมูลซึ่งไม่ปลอดภัย ดังนั้นในงานวิจัยนี้จึงได้ศึกษาการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพเป็นการป้องกันข้อมูลจากการถูกเข้าถึงโดยบุคคลที่ไม่ได้รับอนุญาตและการป้องกันการดักจับข้อมูล โดยการซ่อนข้อมูลในภาพจะทำให้ข้อมูลที่ถูกส่งออกไปนั้นไม่มีการเปลี่ยนแปลงใด ๆ และผู้รับสามารถถอดรหัสข้อมูลที่ถูกซ่อนไว้ได้โดยใช้เครื่องมือเฉพาะ

ผู้วิจัยได้ทดลองใช้โปรแกรมเข้ารหัสลับร่วมกับโปรแกรมซ่อนข้อมูลในภาพ ทดสอบหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) การวัดดัชนีความคล้ายคลึงกันของโครงสร้าง (SSIM) และอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) หลังจากนั้นสกัดข้อมูลและถอดรหัสลับ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความมั่นคงปลอดภัยข้อมูล (Data Security)

หมายถึงการป้องกันข้อมูลไม่ให้ถูกเข้าถึงโดยไม่มีความจำเป็นหรือไม่ได้รับอนุญาตจากเจ้าของข้อมูล โดยความมั่นคงปลอดภัยข้อมูลมีรูปแบบของการป้องกันข้อมูลจากการสูญหาย การเข้าถึงโดยไม่ได้รับอนุญาต การปลอมแปลงข้อมูล และการทำลายข้อมูล วิธีการป้องกันความมั่นคงปลอดภัยข้อมูลมีหลายวิธี ซึ่งรวมถึงการใช้เทคโนโลยีเข้ารหัสลับ (Encryption) เพื่อปกปิดข้อมูลและป้องกันการถูกอ่านโดยไม่ได้รับอนุญาต การควบคุมการเข้าถึง (Access Control) เพื่อจำกัดการเข้าถึงข้อมูลให้เฉพาะผู้ที่ได้รับอนุญาตเท่านั้น การสำรองข้อมูล (Data Backup) เพื่อป้องกันการสูญหายของข้อมูลในกรณีที่เกิดความเสียหาย และการเข้ารหัสแบบสองปัจจัย (Two-Factor Authentication) เพื่อเพิ่มความมั่นคงปลอดภัยในการเข้าถึงข้อมูล โดยมีการตรวจสอบตัวตนของผู้ใช้งานด้วยสิ่งที่มีผู้ใช้ เช่น ตัวเลขลับ การตอบคำถาม และสิ่งที่มีผู้ใช้รู้ เช่น รหัสผ่าน การประยุกต์ใช้วิธีการเหล่านี้จะเหมาะสมจะช่วยให้ความมั่นคงปลอดภัยของข้อมูลสูงขึ้น และลดความเสี่ยงที่จะเกิดการบุกรุกจากผู้ไม่หวังดี [1]

2.2 วิทยาการเข้ารหัสลับ (Cryptography)

เป็นสาขาวิชาในด้านความมั่นคงปลอดภัยของข้อมูล โดยมีวัตถุประสงค์เพื่อป้องกันข้อมูลไม่ให้ถูกเข้าถึงโดยผู้ไม่ได้รับอนุญาต โดยการใช้เทคนิคการเข้ารหัสลับ (Encryption) ใช้อัลกอริทึมเพื่อแปลงข้อมูลปกติ (Plaintext) เป็นข้อมูลที่ถูกเข้ารหัสลับ (Ciphertext) เมื่อผู้รับได้รับข้อมูลที่เข้ารหัส จะทำการถอดรหัส (Decryption) เพื่อแปลงให้เป็นข้อมูลปกติ โดยมีวัตถุประสงค์เพื่อป้องกันการอ่านและการแก้ไขข้อมูลโดย

ไม่ได้รับอนุญาต วิทยาการเข้ารหัสลับจะใช้กุญแจ (Key) โดยนำกุญแจไปใช้กับอัลกอริทึมการเข้ารหัสลับเพื่อกำหนดวิธีการเข้ารหัสลับและการถอดรหัสลับ การเลือกกุญแจที่ปลอดภัยมีความสำคัญอย่างยิ่ง เนื่องจากกุญแจที่ไม่ปลอดภัยอาจทำให้ข้อมูลถูกแฮ็ก (Hacking) หรือถูกเข้าถึงโดยไม่ได้รับอนุญาต วิทยาการเข้ารหัสลับถูกนำมาใช้ในหลายสถานการณ์ เช่น การส่งข้อมูลผ่านอินเทอร์เน็ต การจัดเก็บข้อมูลสำคัญ เป็นต้น [2]

2.3 การซ่อนข้อมูล (Steganography)

เป็นกระบวนการซ่อนข้อมูลซึ่งจะไม่ใช่ที่เห็นได้โดยปกติในสื่อต่าง ๆ ได้แก่ ภาพถ่ายดิจิทัล ไฟล์เสียง และไฟล์วิดีโอ เป็นต้น โดยมีที่จะใช้เทคนิคการแปลงข้อมูลอย่างไม่สามารถสังเกตได้ เช่น การเปลี่ยนแปลงสีพิกเซลในรูปภาพ การเปลี่ยนแปลงความถี่ หรือความละเอียดของเสียงในไฟล์เสียง และการเปลี่ยนแปลงเฟรมในไฟล์วิดีโอ เพื่อซ่อนข้อมูลภายในสื่อต่าง ๆ เหล่านี้ การซ่อนข้อมูลมักนิยมใช้เพื่อการรักษาความลับหรือความปลอดภัยของข้อมูล โดยส่วนมากจะนำมาใช้ในการสื่อสารที่ต้องการความปลอดภัยสูง เช่น การสื่อสารทางการทหาร การสื่อสารทางธุรกิจ หรือการสื่อสารระหว่างสมาชิกขององค์กรที่ต้องการความลับ และความปลอดภัยของข้อมูล อย่างไรก็ตาม การซ่อนข้อมูลไม่ได้ทำให้ข้อมูลปลอดภัยอย่างสมบูรณ์แบบ เนื่องจากสามารถนำเครื่องมือเฉพาะกลุ่มมาใช้สกัดข้อมูลที่ซ่อนออกมาได้ ดังนั้นการซ่อนข้อมูลเป็นเพียงส่วนหนึ่งของกระบวนการรักษาความปลอดภัยของข้อมูล และควรจะใช้ร่วมกับเทคโนโลยีอื่น ๆ เช่น การเข้ารหัสลับข้อมูล หรือการตรวจจับการเข้าถึงข้อมูลโดยไม่ได้รับอนุญาต [3]

2.4 งานวิจัยที่เกี่ยวข้อง

Manjula ได้ทำการวิจัยเรื่อง “Enhanced Secure Image Steganography using Double Encryption Algorithm” โดยการใช้การเข้ารหัสแบบไฮบริด (Hybrid encryption) ซึ่งเรียกว่าวิธี AES-ECC ใช้ระบบการเข้ารหัสแบบสมมาตร (Symmetric cryptosystem) เพื่อรวมกุญแจของระบบการเข้ารหัสแบบไม่สมมาตร (Asymmetric cryptosystem) เข้าด้วยกัน [4]

Kumar ได้ทำการวิจัยเรื่อง “Enhanced Digital Image and Text Data Security Using Hybrid Model of LSB Steganography and AES Cryptography Technique” ซึ่งเป็นแบบจำลองการใช้เทคนิคการซ่อนข้อมูลในบิตที่มียุทธศาสตร์น้อยที่สุด (LSB) ร่วมกับเทคนิคการเข้ารหัสลับขั้นสูง (AES) แบบไฮบริดเพื่อเพิ่มความปลอดภัยของภาพดิจิทัลและข้อความ ที่ยากจะถูกดักจับโดยบุคคลที่ไม่ได้รับอนุญาตให้เห็น ระดับความปลอดภัยของข้อมูลลับถูกประเมินโดยใช้ MSE และ PSNR เพื่อให้การซ่อนดีขึ้นโดยต้องมีค่า MSE ต่ำและ PSNR สูง [5]

Biswas ได้ทำการวิจัยเรื่อง “An Efficient Algorithm for Confidentiality, Integrity and Authentication Using Hybrid Cryptography and Steganography” ในการเข้ารหัสแบบไฮบริดนี้ ให้ความสำคัญแบบสมมาตรที่ใช้ในการเข้ารหัสข้อความจะถูกเข้ารหัสเพิ่มเติมเพื่อให้มีความปลอดภัยมากขึ้น คุณสมบัติเพิ่มเติมของงานวิจัยนี้คือการสร้างลายเซ็นดิจิทัลโดยการเข้ารหัสค่าแฮชของข้อความ [6]

2.5 สถิติที่ใช้ในงานวิจัย

2.5.1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) เป็นหนึ่งในวิธีการวัดความคลาดเคลื่อนของค่าที่คำนวณได้จากการทำนาย โดยสามารถนำมาใช้ในการประเมินประสิทธิภาพของแบบจำลองในการทำนาย และการประมวลผลสัญญาณ เป็นต้น MSE คำนวณจากความแตกต่างเฉลี่ยของค่าจริงกับค่าที่ทำนายได้โดยนำค่าต่างระหว่างค่าจริงและค่าที่ทำนายได้ของแต่ละตัวอย่างยกกำลังสองแล้วหาค่าเฉลี่ยของค่าต่างๆ ดังสมการที่ 2-1

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2-1)$$

เมื่อ

n คือ จำนวนตัวอย่างทั้งหมด

y_i คือ ค่าจริงของตัวอย่างที่ i

\hat{y}_i คือ ค่าที่ทำนายได้ของตัวอย่างที่ i

2.5.2 ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (Peak Signal-to-Noise Ratio : PSNR) เป็นวิธีการวัดคุณภาพของภาพหรือสัญญาณ โดยใช้ค่าส่วนต่างของสัญญาณต้นฉบับกับสัญญาณที่ถูกประมวลผล จากนั้นจึง

เปลี่ยนค่านี้ให้เป็นหน่วย dB และสามารถนำมาใช้ในการวัดความเปลี่ยนแปลงของคุณภาพภาพหรือสัญญาณที่ถูกประมวลผล PSNR คำนวณจากสัญญาณสูงสุดที่สามารถแสดงได้จากภาพหรือสัญญาณต้นฉบับกับค่าสัญญาณรบกวนหรือความคลาดเคลื่อนที่เกิดขึ้นในการประมวลผล โดยใช้สูตรดังสมการที่ 2-2

$$PSNR = 10 \log \left(\frac{MAX^2}{MSE} \right) \quad (2-2)$$

เมื่อ

MAX คือ ค่าสัญญาณสูงสุดที่เป็นไปได้ในภาพหรือสัญญาณต้นฉบับ

MSE คือ ค่าคลาดเคลื่อนกำลังสองเฉลี่ย

2.5.3 ค่าดัชนีความคล้ายคลึงกันของโครงสร้าง (Structural Similarity Index : SSIM) เป็นวิธีการวัดคุณภาพของภาพหรือสัญญาณด้วยการสังเคราะห์คุณภาพภาพที่มีการเปลี่ยนแปลงเป็นช่วงๆ โดยใช้วิธีการเปรียบเทียบความคล้ายคลึงระหว่างภาพหรือสัญญาณต้นฉบับและภาพหรือสัญญาณที่ถูกประมวลผล วิธีการนี้ใช้สมการทางคณิตศาสตร์เพื่อคำนวณค่าความคล้ายคลึงระหว่างสองภาพหรือสัญญาณโดยคำนวณค่าความคล้ายคลึงใน 3 ด้าน คือ ความแตกต่างทางความสว่างของสี ความแตกต่างทางความเข้มของสี และความแตกต่างทางโครงสร้าง ดังสมการที่ 2-3

$$SSIM(x, y) = [l(x, y)^{\alpha} c(x, y)^{\beta} s(x, y)^{\gamma}] \quad (2-3)$$

เมื่อ

x และ y เป็นภาพหรือสัญญาณที่ถูกประมวลผล

$l(x, y)$ เป็นค่าความแตกต่างทางความสว่างของสี

$c(x, y)$ เป็นค่าความแตกต่างทางความเข้มของสี

$s(x, y)$ เป็นค่าความแตกต่างทางโครงสร้าง

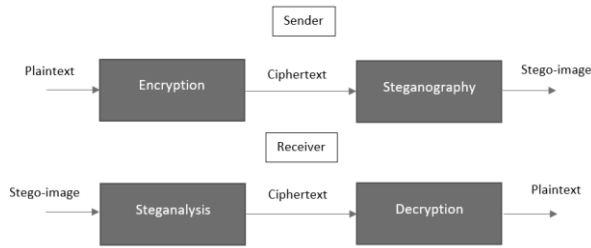
α, β และ γ เป็นค่าที่กำหนดน้ำหนักของแต่ละด้าน

โดยค่าเริ่มต้นคือ 1

3. วิธีการดำเนินงานวิจัย

กรอบแนวความคิดงานวิจัยนี้แสดงดังภาพที่ 1 โดยเริ่มจากฝั่งผู้ส่งนำข้อมูลต้นฉบับมาเข้ารหัสลับ หลังจากนั้นนำข้อมูลที่เข้ารหัสลับไปซ่อนในภาพดิจิทัล เมื่อฝั่งผู้รับได้รับภาพดิจิทัลจึงทำการสกัดข้อมูลในภาพดิจิทัล

หลังจากนั้นนำข้อมูลที่ถูกรหัสลับไปถอดรหัสลับเป็นข้อมูลต้นฉบับ

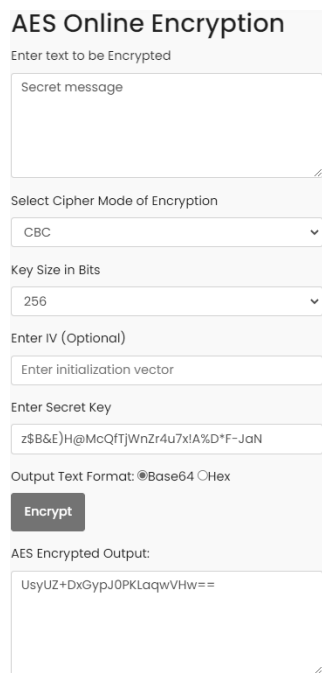


ภาพที่ 1: กรอบแนวคิดงานวิจัย

3.1 ศึกษาเครื่องมือที่ใช้ในการเข้ารหัสลับข้อมูลในงานวิจัยนี้ใช้ AES calculator [7]

3.2 ศึกษาเครื่องมือที่ใช้ในการซ่อนข้อมูลในภาพในงานวิจัยนี้ใช้ MATLAB

3.3 นำข้อมูลมาทำการเข้ารหัสลับดังภาพที่ 2 โดยใช้กุญแจจาก All keys generator [8] ร่วมกับการใช้ AES calculator โดยกำหนดโหมดแบบไซเฟอร์บล็อกเชนนิ่ง (Cipher Block Chaining : CBC) ขนาดกุญแจ 256 บิต



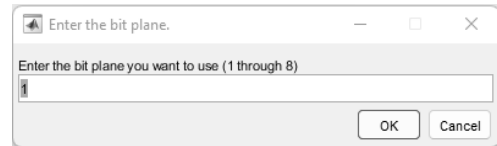
ภาพที่ 2: การเข้ารหัสลับของข้อมูล

3.4 นำข้อมูลที่ผ่านการเข้ารหัสลับมาทำการซ่อนข้อมูลในภาพดิจิทัลดังภาพที่ 3 โดยการกำหนดตำแหน่งของบิตที่ต้องการซ่อนข้อมูลซึ่งในงานวิจัยนี้กำหนดให้ซ่อนในบิต

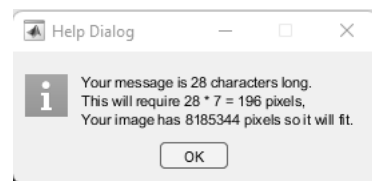
ที่ 1 ดังภาพที่ 4 และคำนวณจำนวนพิกเซลที่ใช้ซ่อนข้อมูลดังภาพที่ 5



ภาพที่ 3: การใส่ข้อมูลที่ถูกรหัสลับ



ภาพที่ 4: การกำหนดตำแหน่งบิตที่ใช้ซ่อนข้อมูล



ภาพที่ 5: การคำนวณจำนวนพิกเซลที่ใช้ซ่อนข้อมูล

3.5 ทำกระบวนการซ้ำในข้อที่ 3.3-3.4 จำนวน 10 รอบ โดยไม่มีการเปลี่ยนแปลงข้อมูลที่เข้ารหัสลับ แต่เปลี่ยนภาพดิจิทัลที่ใช้ในการซ่อนข้อมูล ซึ่งเราใช้ภาพดิจิทัลจำนวน 10 ภาพ เรียงตามลำดับ ดังภาพที่ 6



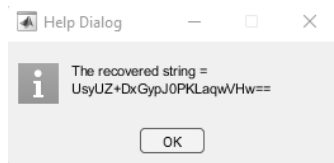
ภาพที่ 6: ภาพดิจิทัลที่ใช้ทดสอบ



ภาพที่ 6: (ต่อ)

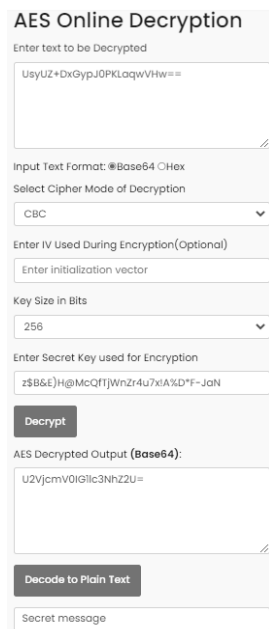
3.6 ทดสอบหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด และค่าดัชนีความคล้ายคลึงกันของโครงสร้าง ของภาพที่ผ่านการซ่อนข้อมูลกับภาพต้นฉบับทั้งหมด 10 ภาพ

3.7 ทำการสกัดข้อมูลที่ซ่อนในภาพดังภาพที่ 7



ภาพที่ 7: การสกัดข้อมูลจากภาพ

3.8 ทำการถอดรหัสข้อมูลดังภาพที่ 8



ภาพที่ 8: การถอดรหัสลับข้อมูล

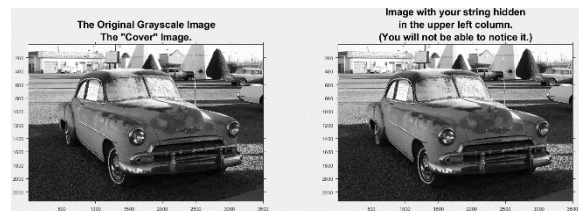
3.9 ประเมินผลการทดลอง

3.10 สรุปผลการทดลอง

4. ผลการดำเนินงานวิจัย

4.1 ผลการทดสอบการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพ

ผลทดสอบการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพ ปรากฏว่าภาพดิจิทัลก่อนและหลังการซ่อนข้อมูลที่ถูกเข้ารหัสลับไม่มีการเปลี่ยนแปลงใด ๆ จนเป็นที่น่าสังเกตได้ด้วยตาเปล่า ดังแสดงในภาพที่ 9 ภาพซ้ายมือเป็นภาพต้นฉบับก่อนการซ่อนข้อมูล และภาพขวามือคือภาพที่ผ่านการซ่อนข้อมูล ซึ่งภาพทั้งสองมีความเหมือนกัน

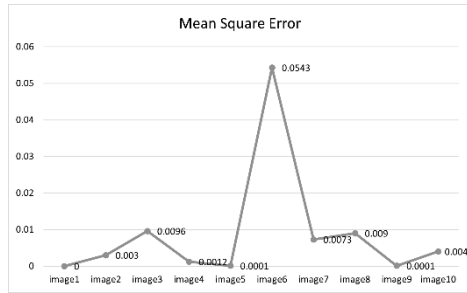


ภาพที่ 9: ตัวอย่างภาพรถยนต์ก่อนและหลังซ่อนข้อมูล

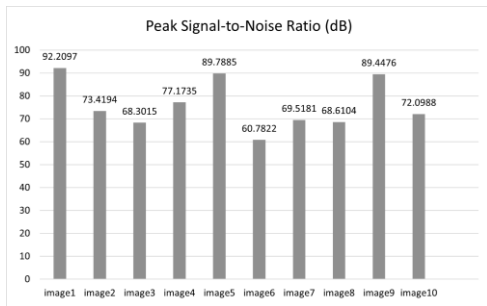
หลังจากนั้นเมื่อเรานำภาพดิจิทัลมาสกัดข้อมูล และนำข้อมูลไปถอดรหัสลับ ปรากฏว่าข้อมูลรหัสลับนั้นยังคงเหมือนเดิมดังภาพที่ 2 ภาพที่ 7 และภาพที่ 8

4.2 ผลการประเมิน

ค่าคลาดเคลื่อนกำลังสองเฉลี่ยดังภาพที่ 10 จากกราฟแสดงให้เห็นว่าภาพต้นฉบับกับภาพที่ซ่อนข้อมูลลับมีค่าความคลาดเคลื่อนไม่แตกต่างกันมาก โดยเฉลี่ยมีค่าประมาณ 0 ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุดดังภาพที่ 11 จากแผนภูมิแสดงให้เห็นว่าภาพต้นฉบับกับภาพที่ซ่อนข้อมูลลับมีค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุดอยู่ในเกณฑ์ที่สูง คือตั้งแต่ค่าประมาณ 60-92 dB และค่าดัชนีความคล้ายคลึงกันของโครงสร้างโดยการใช้สมการที่ 3-3 คำนวณภาพต้นฉบับกับภาพที่ซ่อนข้อมูลลับปรากฏว่ามีค่าเท่ากับ 1 ซึ่งแสดงว่าภาพต้นฉบับกับภาพที่ซ่อนข้อมูลลับมีความเหมือนกัน



ภาพที่ 10: ค่าคลาดเคลื่อนกำลังสองเฉลี่ย



ภาพที่ 11: ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด

5. สรุปและอภิปราย

ผู้วิจัยได้ทดลองใช้โปรแกรมเข้ารหัสลับร่วมกับโปรแกรมซ่อนข้อมูลในภาพ หลังจากนั้นทดสอบหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) และค่าดัชนีความคล้ายคลึงกันของโครงสร้าง (SSIM) ผลการทดสอบเพื่อประเมินความเหมือนของภาพ ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าประมาณ 0 ค่าอัตราส่วนสัญญาณต่อสัญญาณรบกวนสูงสุด (PSNR) มีค่าสูง และค่าดัชนีความคล้ายคลึงกันของโครงสร้าง (SSIM) เท่ากับ 1 แสดงว่าภาพดิจิทัลที่ถูกซ่อนข้อมูลไม่เกิดการเปลี่ยนแปลงหลังจากการซ่อนข้อมูล และเมื่อสกัดข้อมูลรวมทั้งถอดรหัสลับ ผลปรากฏว่าข้อมูลต้นฉบับยังเป็นข้อความเดิม จากผลการดำเนินงานวิจัยได้ข้อสรุปว่าเราสามารถประยุกต์ใช้เทคนิคการเข้ารหัสลับร่วมกับการซ่อนข้อมูลในภาพเพื่อช่วยเพิ่มความมั่นคงปลอดภัยข้อมูลได้

6. งานวิจัยในอนาคต

ในงานวิจัยนี้ยังขาดกระบวนการในการทดสอบความแข็งแกร่งของภาพ (Robustness) ซึ่งเมื่อมีการใช้งานจริงยังมีปัจจัยที่ต้องคำนึงถึง เช่น เมื่อภาพที่ถูกซ่อนข้อมูลถูก

โจมตี เช่น การบันทึกข้อมูลภาพ (Saving) หลายครั้ง การปรับค่าความสว่างของภาพ (Brightness intensity) การขยายภาพ (Scaling) การหมุน (Rotation) ฯลฯ ภาพนั้นจะยังคงสามารถถอดรหัสได้อยู่หรือไม่ ซึ่งผู้วิจัยจะได้ทำการทดสอบกระบวนการดังกล่าวต่อไป

เอกสารอ้างอิง

- [1] National Institute of Standards and Technology (NIST) Cybersecurity Framework. Retrieved 20 February 2023. From: <https://www.nist.gov/cyberframework>.
- [2] Yang Yu. "Encryption Technology for Computer Network Data Security Protection." *Hindawi Security and Communication Networks*, Volume 2022.
- [3] Shehab D.A. and Alhaddad, M.J. "Survey of Multimedia Steganalysis: Techniques, Evaluations, and Trends in Future Research." *MDPI, Symmetry*, 2022, 14, 117.
- [4] Y. Manjula and K. B. Shivakumar, "Enhanced secure image steganography using double encryption algorithms," *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2016, pp. 705-708.
- [5] M. Kumar, A. Soni, A. R. S. Shekhawat and A. Rawat, "Enhanced Digital Image and Text Data Security Using Hybrid Model of LSB Steganography and AES Cryptography Technique," *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, Coimbatore, India, 2022, pp. 1453-1457.
- [6] C. Biswas, U. D. Gupta and M. M. Haque, "An Efficient Algorithm for Confidentiality, Integrity and Authentication Using Hybrid Cryptography and Steganography," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh, 2019, pp. 1-5.
- [7] AES calculator. Retrieved 20 February 2023. From: decryption.
- [8] All keys generator. Retrieved 20 February 2023. From: <https://www.allkeysgenerator.com/>

**การออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการ
และควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชน**

**Design and Development of a Data Storage and Exchange Systems
for the Air Force Command and Control Systems Using Blockchain Technology**

ธนภุต เพ็งเคียน (Thanakrit Pengkian)¹ ประสงค์ ปรามิตพลกรัง (Prasong Praneetpolgrang)²

และพายัพ ศิรินาม (Payap Sirinam)³

หลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีป้องกันประเทศ

สำนักบัณฑิตศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

¹Intelligent.it4@gmail.com, ²prasongspu@gmail.com, ³p.sirinam@gmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชนซึ่งเป็นเทคโนโลยีการประมวลผลและจัดเก็บข้อมูลแบบกระจายศูนย์ หรือที่เรียกว่า Distributed Ledger Technology (DLT) และเป็นรูปแบบการบันทึกข้อมูลที่ใช้หลักการ Cryptography ร่วมกับกลไก Consensus โดยข้อมูลที่ถูกรับบันทึกในระบบบล็อกเชนนี้นจะสามารถทำการแก้ไขเปลี่ยนแปลงได้ยาก ช่วยเพิ่มความถูกต้องและความน่าเชื่อถือของข้อมูล งานวิจัยนี้จะกล่าวถึงกระบวนการออกแบบสัญญาอัจฉริยะ (Smart Contract) ซึ่งเป็นส่วนของการควบคุมการเข้าถึงข้อมูลระบบบัญชาการและควบคุมของกองทัพอากาศ โดยผ่านกระบวนการตรวจสอบข้อมูลเพื่อให้มั่นใจว่า ข้อมูลที่จัดเก็บอยู่ในระบบนั้นสามารถนำไปใช้ในการปฏิบัติการกิจของกองทัพอากาศได้

คำสำคัญ: บล็อกเชน สัญญาอัจฉริยะ ระบบบัญชาการและควบคุม

Abstract

This research aims to design and develop data storage and exchange systems for air force command and control systems. The Blockchain technology, also

known as Distributed Ledger Technology (DLT), is used as a decentralized data processing and storage technology. It is also applied as a data recording model that uses cryptography principles in conjunction with consensus mechanisms. The fact that data recording in Blockchain system is difficult to change helps increasing the accuracy and credibility of data. Thus, this research will discuss the Smart Contract design process as part of the access control of Air Force's command and control systems through a data verification process. This is to ensure that the data stored in the system could be used to fulfill air force missions.

Keywords: Blockchain, Smart Contract, Command and Control Systems.

1. บทนำ

จากยุทธศาสตร์กองทัพอากาศ 20 ปี พ.ศ. 2561-2580 ได้กล่าวถึง เทคโนโลยีสารสนเทศและการสื่อสารที่มีการพัฒนาอย่างรวดเร็ว รวมทั้งการเกิดขึ้นของภัยคุกคามในมิติไซเบอร์ทั้งในรูปแบบการจารกรรมข้อมูล และการโจมตีเพื่อทำลายล้าง ล้วนก่อให้เกิดผลกระทบและความเสียหายในวงกว้าง หลายประเทศมีการจัดตั้งหน่วยงานรับผิดชอบโดยตรง ดังนั้นกองทัพอากาศจำเป็นต้องพัฒนาขีด

ความสามารถด้านไซเบอร์ให้มีความพร้อมในการเผชิญกับภัยคุกคามทางไซเบอร์ รวมทั้งการพัฒนาระบบเครือข่ายให้มีความแข็งแกร่งและปลอดภัย จากที่กล่าวมาข้างต้นทำให้เห็นถึงความสำคัญจำเป็นและผลกระทบในมิติไซเบอร์ที่มีต่อความมั่นคงของกองทัพอากาศและประเทศ การเข้ารหัสข้อมูลที่เป็นมาตรฐาน อาจจะไม่เพียงพอต่อการบุกรุกในอนาคต ผู้วิจัยจึงได้ศึกษาวิธีการออกแบบและพัฒนาระบบจัดเก็บข้อมูลแบบกระจายในระบบบัญชาการและควบคุมของกองทัพอากาศ เพื่อให้ผู้บังคับบัญชามั่นใจได้ว่า ข้อมูลมีความถูกต้อง ครบถ้วน น่าเชื่อถือ และสามารถนำไปใช้ปฏิบัติการกิจต่าง ๆ ได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคโนโลยีบล็อกเชน

บล็อกเชน [2] เป็นการจัดเก็บข้อมูลแบบ Shared Database โดยทำการบันทึกข้อมูลที่บ่งบอกว่าข้อมูลที่ถูกบันทึกไปก่อนหน้านั้น ไม่สามารถที่จะเปลี่ยนแปลงหรือแก้ไขได้ ซึ่งทุกผู้ใช้งานจะให้เห็นข้อมูลชุดเดียวกันทั้งหมด

2.1.1 ประเภทของบล็อกเชน

บล็อกเชน สามารถแบ่งออกได้เป็น 3 ประเภท [2] โดยพิจารณาจากข้อกำหนด ในการเข้าร่วมเป็นสมาชิกของเครือข่ายประกอบด้วย บล็อกเชนแบบเปิดสาธารณะ (Public Blockchain) เป็นบล็อกเชนที่อนุญาตให้ทุกคนสามารถเข้าใช้งานได้อย่างอิสระ บล็อกเชนแบบปิด (Private Blockchain) เป็นบล็อกเชนที่เข้าใช้งานได้เฉพาะผู้ที่ได้รับอนุญาตเท่านั้น และ บล็อกเชนแบบเฉพาะกลุ่ม (Consortium Blockchain) เป็นบล็อกเชนที่ผสมผสานแนวคิดระหว่างบล็อกเชนแบบเปิดสาธารณะและบล็อกเชนแบบปิด ซึ่งส่วนมากเป็นการรวมตัวกันขององค์กรที่ต้องมีการแลกเปลี่ยนข้อมูลระหว่างกัน

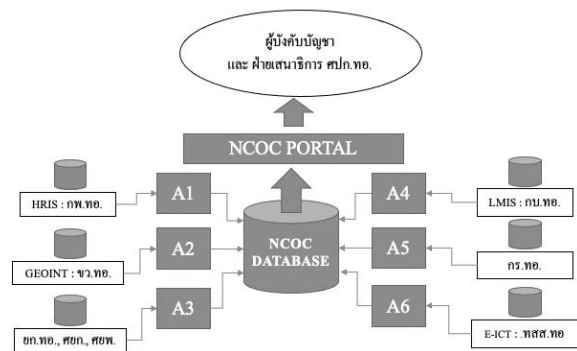
2.2 สัญญาอัจฉริยะ (Smart Contract)

สัญญาอัจฉริยะ [4] คือกระบวนการทางดิจิทัลที่กำหนดขั้นตอนการทำธุรกรรมโดยอัตโนมัติไว้ล่วงหน้า โดยไม่ต้องอาศัยตัวกลาง เช่น ธนาคาร การสร้างสัญญาอัจฉริยะที่เป็นระบบอัตโนมัติอย่างเต็มรูปแบบ โดยคู่สัญญาทั้งสอง

ฝ่ายจะมีการตกลงกันก่อนหน้านี้ถึงขั้นตอน กลไก ในการทำรายการธุรกรรมดังกล่าว

2.3 ระบบบัญชาการและควบคุม

ระบบบัญชาการและควบคุม (Command and Control Systems) คือ สิ่งอำนวยความสะดวก ยุทธภัณฑ์ การสื่อสาร ระเบียบปฏิบัติ กำลังพล การบริหารจัดการ ข้อมูล ข่าวสาร และข่าวกรอง ที่จำเป็นสำหรับผู้บังคับบัญชาเพื่อใช้ในการวางแผน อำนาจการและควบคุมการปฏิบัติต่าง ๆ ของกำลังรบตามภารกิจที่ได้รับมอบหมาย ดังภาพที่ 1



ภาพที่ 1: ระบบบัญชาการและควบคุมของกองทัพอากาศ

จากภาพที่ 1 พบว่า ระบบบัญชาการและควบคุมของกองทัพอากาศสามารถสนับสนุนและบูรณาการข้อมูลพื้นฐานทางกายภาพ และข้อมูลสถานภาพทั้ง 6 ด้าน ได้แก่ กำลังพล ข่าว ยุทธการ ส่งกำลังบำรุง กิจการพลเรือน และเทคโนโลยีสารสนเทศ จากระบบฐานข้อมูลหลักของกรมฝ่ายเสนาธิการซึ่งเป็นการรวมศูนย์ของข้อมูล

2.4 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่ผ่านมาพบว่า เทคโนโลยีบล็อกเชนสามารถนำมาใช้ประยุกต์ใช้กับงานต่าง ๆ ได้ โดยมีรายละเอียดดังนี้

สิริกันต์ ไชยกุล [3] ได้ทำการวิจัยเรื่อง การพัฒนาระบบข้อมูลการดูแลสุขภาพโดยใช้เทคโนโลยีบล็อกเชน วัตถุประสงค์เพื่อออกแบบและพัฒนาแบบจำลองระบบเพื่อรวบรวมและจัดเก็บข้อมูลการรักษาพยาบาลของผู้ป่วยลงในฐานข้อมูล และกระจายข้อมูลให้กับทุกสถานพยาบาล ผู้ป่วยสามารถตรวจสอบรายละเอียดข้อมูลทางการแพทย์ก่อนหน้าแบบเรียลไทม์ ผลการวิจัยพบว่า ข้อมูลที่จัดเก็บอยู่

ในระบบ สามารถลดขั้นตอนและเวลาสำหรับผู้ผู้ใช้ที่เคยเข้ารับการรักษาจากสถานพยาบาลอื่น ๆ

Firmansyah Ashari [8] ได้ทำการวิจัยและพัฒนา แพลตฟอร์มระดมทุน โดยใช้สัญญาอัจฉริยะและเทคโนโลยีบล็อกเชน มีวัตถุประสงค์เพื่อสร้างความโปร่งใสในการระดมทุน เพื่อให้สามารถตรวจสอบแหล่งที่มาที่ไปของเงินทุนได้ ผลการวิจัยพบว่า การนำสัญญาอัจฉริยะมาใช้สำหรับการระดมทุนสามารถเพิ่มความไว้วางใจ มีความโปร่งใส และช่วยลดขั้นตอนในการระดมทุนได้ เนื่องจากคุณสมบัติของสัญญาอัจฉริยะทำให้ทุกคนสามารถตรวจสอบข้อมูลได้แต่ไม่สามารถแก้ไขได้

3. วิธีการดำเนินงาน

งานวิจัยนี้มีวัตถุประสงค์เพื่อออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชน ซึ่งในบทความนี้จะกล่าวถึง ขั้นตอนการวิจัยและเครื่องมือที่ใช้ในการวิจัย โดยมีรายละเอียดดังนี้

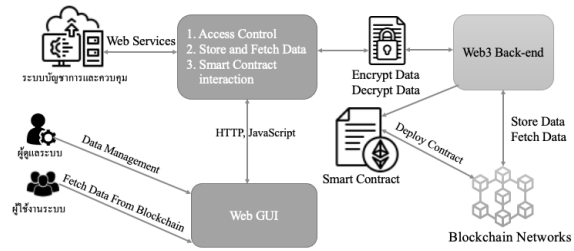
3.1 ศึกษาข้อมูลและเครื่องมือที่ใช้ในการออกแบบและพัฒนาระบบจัดเก็บข้อมูล

จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง ทำให้ผู้วิจัยเลือกใช้เครื่องมือ ดังต่อไปนี้

- 3.1.1 อีเธอเรียมบล็อกเชน (Ethereum Blockchain)
- 3.1.2 ภาษาที่ใช้ในการพัฒนา เป็นภาษา Solidity
- 3.1.3 สัญญาอัจฉริยะ (Smart Contract)

3.2 ออกแบบกระบวนการทำงานของระบบ

การออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศโดยใช้เทคโนโลยีบล็อกเชน เป็นการจัดเก็บและแลกเปลี่ยนข้อมูลต่าง ๆ ผ่านระบบเครือข่ายที่ไม่มีศูนย์กลาง แสดงให้เห็นถึงกระบวนการทำงานของระบบดังภาพที่ 2



ภาพที่ 2: แสดงกระบวนการทำงานของระบบ

จากภาพที่ 2 แสดงถึงกระบวนการทำงานของระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศ โดยที่ผู้ดูแลระบบหรือผู้ใช้งานสามารถเข้าใช้งานได้ผ่านเว็บไซต์ ซึ่งเบื้องหลังการทำงานจะมีการเชื่อมต่อกันระหว่าง ระบบหน้าบ้าน (Front-end) และ ระบบหลังบ้าน (Back-end) เพื่อใช้ในการเรียกใช้ข้อมูล และบันทึกยังบล็อกเชน โดยข้อมูลทั้งหมดที่ส่งมาจากระบบบัญชีการและควบคุมถูกออกแบบไว้ให้มีการเข้ารหัสข้อมูลก่อนส่งไปเก็บยังบล็อกเชน เพื่อความปลอดภัยและความเป็นส่วนตัวของข้อมูลในระบบ นอกจากนี้ยังมีการจัดเก็บข้อมูลแบบกระจายศูนย์เพื่อป้องกันการลักลอบแก้ไขและเพิ่มความพร้อมใช้งานของข้อมูล

3.3 ออกแบบส่วนแสดงผล (Front-end)

ระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศ มีส่วนของการแสดงผล 2 ส่วนดังนี้

3.3.1 ส่วนผู้ดูแลระบบเป็นส่วนที่มีไว้เพื่อจัดการกับข้อมูลต่าง ๆ ของระบบ เช่น กำหนดสิทธิ์ผู้ใช้งานระบบ

3.3.2 ส่วนผู้ใช้งานระบบเป็นส่วนที่ใช้ในการค้นหาและแสดงผลข้อมูล ซึ่งผู้ที่จะใช้งานในส่วนนี้จะต้องได้รับอนุญาตจากผู้ดูแลระบบก่อน

3.4 ออกแบบสัญญาอัจฉริยะ (Smart Contract)

สัญญาอัจฉริยะ (Smart Contract) เป็นสิ่งสำคัญที่ทำให้ผู้ใช้งานสามารถติดต่อแลกเปลี่ยนข้อมูลกันได้ผ่านเครือข่ายบล็อกเชน ซึ่งในการออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชน มีฟังก์ชันหลักๆ ของสัญญาอัจฉริยะ 2 ฟังก์ชันดังนี้

3.4.1 Calling เป็นส่วนของการติดต่อเพื่อเรียกใช้ข้อมูลที่บันทึกอยู่บนบล็อกเชน ฉะนั้นจึงทำให้ผู้ใช้งานสามารถรับข้อมูลได้อย่างรวดเร็ว เนื่องจากไม่ต้องรอการยืนยันข้อมูลจากโหนดอื่น ๆ

3.4.2 Sending เป็นส่วนของการติดต่อนำข้อมูลบันทึกบนบล็อกเชน ฉะนั้นจึงทำให้ใช้ระยะเวลาสั้นเนื่องจากข้อมูลที่ส่งเข้ามานั้น จะถูกกระจายไปยังโหนดต่าง ๆ เพื่อช่วยกันยืนยันข้อมูลและบันทึกลงบนบล็อกเชนต่อไป

4. ผลการดำเนินงาน

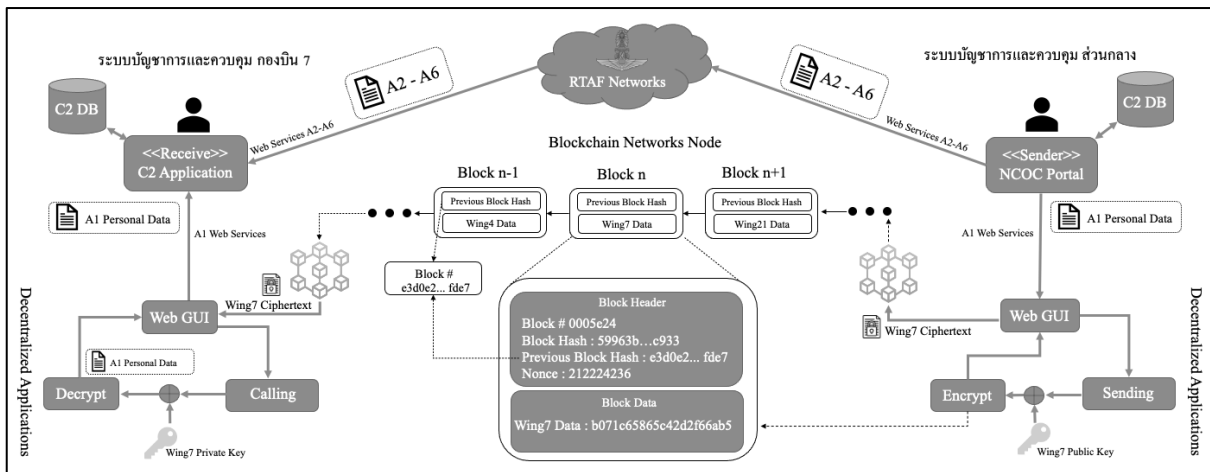
4.1 กระบวนการรับส่งข้อมูลผ่านบล็อกเชน

การออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชนให้เป็นต้นแบบของระบบของระบบบัญชาการและควบคุมของกองบิน 7 จ.สุราษฎร์ธานี

ได้คำนึงถึงความถูกต้อง โปร่งใส ความพร้อมใช้งาน การรักษาความลับของข้อมูล การกระจายอำนาจ และเป็นและเป็นมาตรฐานสากล ผู้วิจัยได้เพิ่มเทคโนโลยีการเข้ารหัสแบบอสมมาตรมาเข้ารหัสข้อมูลก่อนส่งไปจัดเก็บยังบล็อกเชน ดังภาพที่ 3

โดยใช้การเชื่อมต่อ 2 เส้นทาง เนื่องด้วยบล็อกเชนมีข้อจำกัดในการเขียนข้อมูลให้กับทุกโหนดทำได้ช้า ฉะนั้นข้อมูลที่มีความสำคัญรองลงมา อาจจะต้องใช้เครือข่ายคอมพิวเตอร์กองทัพอากาศในการส่งข้อมูล ซึ่งข้อมูลที่ส่งผ่านแอปพลิเคชันแบบกระจายอำนาจ (Decentralized Applications) จะถูกส่งไปเก็บยังทุก Node ในกองทัพอากาศ แต่ข้อมูลจะถูกเข้ารหัสแล้วคือมีข้อมูลเก็บอยู่ทุกหน่วยแต่จะสามารถเปิดข้อมูลได้จะต้องเป็นเจ้าของข้อมูลที่มีคีย์ส่วนตัว (Private Key) เท่านั้น

4.1.2 กระบวนเข้ารหัสข้อมูลภายในเครือข่ายบล็อกเชน เมื่อข้อมูลไหลเข้ามายังเครือข่ายบล็อกเชน ด้วยกลไก



ภาพที่ 3: แสดงกระบวนการรับส่งข้อมูลผ่านบล็อกเชน

จากภาพที่ 3 พบว่าระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชน มีกระบวนการรับส่งข้อมูลผ่านเครือข่ายบล็อกเชน โดยแสดงให้เห็นถึงระบบที่เกี่ยวข้องกัน เช่น ระบบเครือข่ายคอมพิวเตอร์กองทัพอากาศระบบบัญชาการและควบคุมส่วนกลาง ระบบบัญชาการและควบคุมส่วนกองบิน 7 แอปพลิเคชัน

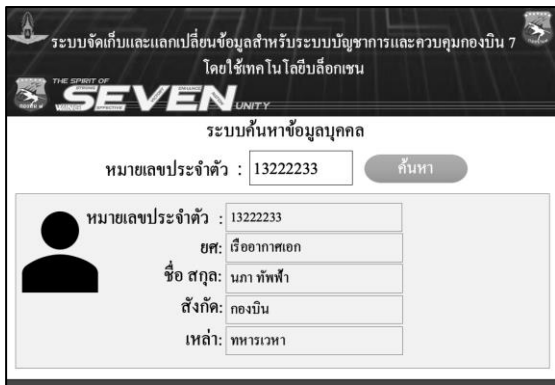
แบบกระจายอำนาจซึ่งมีกระบวนการรับส่งข้อมูลจากต้นทางถึงปลายทางสามารถสรุป ได้ดังนี้

4.1.1 กระบวนการส่งข้อมูลผ่านเครือข่ายบล็อกเชน ผู้ใช้ส่งข้อมูลจากระบบบัญชาการและควบคุมส่วนกลางให้กับระบบบัญชาการและควบคุมส่วนกองบิน 7 ของบล็อกเชนจะนำข้อมูลมาเข้ารหัสในรูปแบบ Cryptography โดยใช้อัลกอริทึม Sha256 จากนั้นกระจายไปยังทุกโหนดในเครือข่าย

4.1.3 กระบวนการรับข้อมูลผ่านเครือข่ายบล็อกเชน

เมื่อข้อมูลถูกจัดเก็บลงบล็อกเชน หน่วยปลายทางสามารถเรียกใช้ข้อมูลได้โดยผ่านทางแอปพลิเคชันแบบกระจายอำนาจ (Decentralized Applications) ซึ่งข้อมูลที่ถูกเรียกใช้จะยังไม่สามารถนำมาใช้งานได้จนกว่าจะผ่านกระบวนการถอดรหัส (Decrypt) โดยใช้คีย์ส่วนตัว (Private Key) ของเจ้าของข้อมูลเท่านั้น เมื่อถอดรหัสเรียบร้อยแล้วข้อมูลจะถูกส่งถึงผู้รับ ซึ่งถือว่าเสร็จสิ้นกระบวนการในการรับส่งข้อมูลผ่านเครือข่ายบล็อกเชน

4.1.4 การพัฒนาส่วนแสดงผลระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุมของกองทัพอากาศ โดยใช้บล็อกเชนดังภาพที่ 4 และภาพที่ 5



ภาพที่ 4: ตัวอย่างหน้าแสดงผลการเรียกใช้ข้อมูล

จากภาพที่ 4 แสดงการค้นหาข้อมูลของระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุม

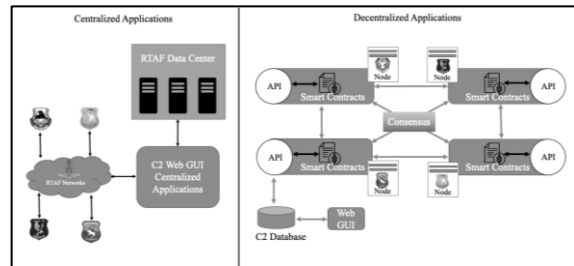


ภาพที่ 5: ตัวอย่างหน้าแสดงผลการบันทึกข้อมูล

จากภาพที่ 5 แสดงการบันทึกข้อมูลของระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชาการและควบคุมของกองทัพอากาศใช้บล็อกเชน

4.2 เปรียบเทียบการรับส่งข้อมูลแบบดั้งเดิมและระบบใหม่ที่พัฒนาขึ้นโดยใช้บล็อกเชนของกองทัพอากาศ

ระบบบัญชาการและควบคุมของกองทัพอากาศแบบดั้งเดิมมีการใช้งานในรูปแบบรวมศูนย์กล่าวคือข้อมูลทั้งหมดจะรวมศูนย์อยู่ที่ส่วนกลาง และเมื่อนำระบบกระจายศูนย์ที่ผู้วิจัยพัฒนาขึ้น มาช่วยนั้นจะทำให้การรับส่งข้อมูลมีประสิทธิภาพมากยิ่งขึ้น ดังภาพที่ 6



ภาพที่ 6: แสดงการเปรียบเทียบ การรับส่งข้อมูลระหว่าง Centralized Applications และ Decentralized Applications

จากภาพที่ 6 พบว่า เมื่อเทียบระหว่างระบบรวมศูนย์ที่เป็นระบบเดิม และระบบกระจายศูนย์ที่เป็นระบบใหม่สามารถสรุปได้ดังนี้

4.2.1 ด้านอุปกรณ์ ระบบรวมศูนย์มีการใช้เครื่องคอมพิวเตอร์แม่ข่ายจำนวนมากเพื่อใช้ในการจัดเก็บและบริหารจัดการข้อมูล ส่วนระบบกระจายศูนย์จะใช้เครื่องที่อยู่ในเครือข่ายบล็อกเชนเป็นตัวจัดเก็บข้อมูล

4.2.2 ด้านงบประมาณ ระบบรวมศูนย์จะต้องใช้งบประมาณในการจัดหาและบำรุงรักษาอุปกรณ์ประจำปี ส่วนระบบกระจายศูนย์ใช้งบประมาณเพียงแค่การบำรุงรักษาเครื่องที่ใช้งานระบบเท่านั้น

4.2.3 ด้านความเร็วในการรับส่งข้อมูล ระบบรวมศูนย์จะรับส่งข้อมูลได้ช้าหรือเร็วขึ้นอยู่กับความเสถียรของระบบเครือข่ายคอมพิวเตอร์ภายในองค์กร ส่วนระบบกระจายศูนย์สามารถรับส่งข้อมูลได้อย่างรวดเร็วเนื่องจากข้อมูลทั้งหมดถูกกระจายไปเก็บยังโหนดต่างๆ ที่อยู่ในเครือข่ายบล็อกเชน

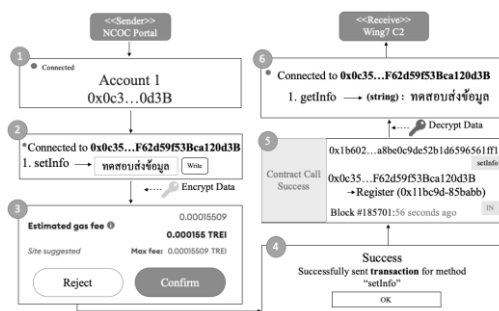
4.2.4 ด้านความปลอดภัย ระบบรวมศูนย์มีความปลอดภัยในส่วนของการบริหารจัดการระบบ เนื่องจากผู้ดูแลระบบสามารถจัดการได้ที่จุดเดียว ส่วนระบบกระจายศูนย์มีความยืดหยุ่นในเรื่องของการกำหนด

เงื่อนไขภายในสัญญาอัจฉริยะ ซึ่งจะเหมาะสำหรับผู้ดูแลระบบที่มีความชำนาญในด้านนี้

4.2.5 ด้านความปลอดภัย ระบบรวมศูนย์มีการวางมาตรการด้านการรักษาความปลอดภัยระบบสารสนเทศรวมไว้ที่ส่วนกลางเท่านั้น ส่วนระบบกระจายศูนย์มีการรักษาความปลอดภัยของข้อมูลโดยการเข้ารหัสข้อมูลในรูปแบบ Cryptography ก่อนการส่งและรับข้อมูล สามารถพิสูจน์ต้นทางและปลายทางได้ เพิ่มความน่าเชื่อถือให้กับข้อมูลมากยิ่งขึ้น

4.3 ผลการทดสอบระบบที่ผู้วิจัยได้พัฒนาขึ้น

การรับส่งข้อมูลในระบบบัญชีการและควบคุมของกองทัพอากาศโดยใช้เทคโนโลยีบล็อกเชน จะมีการรักษาความปลอดภัยของข้อมูลด้วยวิธีการเข้ารหัสแบบอสมมาตร ประกอบกับเทคโนโลยีบล็อกเชนที่มีกลไกในการป้องกันการเปลี่ยนแปลงแก้ไขข้อมูลและใช้ในการพิสูจน์ธุรกรรมที่เกิดขึ้น ด้วยการเข้ารหัสในรูปแบบ Cryptography และเชื่อมโยงข้อมูลธุรกรรมเข้าด้วยกัน ดังภาพที่ 7



ภาพที่ 7: แสดงการทดสอบรับ ส่งข้อมูลผ่านบล็อกเชน

จากภาพที่ 7 ผู้วิจัยได้ทำการเลือกใช้ Test Networks ในการทดลองรับส่งข้อมูลระหว่างระบบบัญชีการและควบคุมส่วนกลาง กับ ระบบบัญชีการและควบคุมกองบิน 7 โดยข้อมูลที่ส่งผ่านบล็อกถูกเข้ารหัสก่อนการส่ง และถอดรหัสเมื่อเรียกใช้งาน ทำให้ข้อมูลมีความปลอดภัย ถูกต้อง น่าเชื่อถือ และพร้อมใช้งานตลอดเวลา

5. สรุป

ปัญหาของการจัดเก็บและแลกเปลี่ยนข้อมูลแบบรวมศูนย์คือความพร้อมใช้งานของเครื่องแม่ข่ายที่จำเป็น

จะต้องเข้าถึงและใช้งาน ได้ตลอดเวลา การออกแบบและพัฒนาระบบจัดเก็บและแลกเปลี่ยนข้อมูลสำหรับระบบบัญชีการและควบคุมของกองทัพอากาศ โดยใช้เทคโนโลยีบล็อกเชน ที่ผู้วิจัยได้พัฒนาขึ้นนี้สามารถลดปัญหาและเพิ่มประสิทธิภาพในการจัดเก็บและแลกเปลี่ยนข้อมูลได้เป็นอย่างดี ซึ่งจัดเก็บข้อมูลแบบกระจายศูนย์ได้เพิ่มเทคโนโลยีการเข้ารหัสแบบอสมมาตร เพื่อรักษาไว้ซึ่งความลับของข้อมูล กองทัพอากาศ อีกทั้งยังช่วยลดความเสี่ยงในเรื่องของความพร้อมใช้งานของเครื่องคอมพิวเตอร์แม่ข่ายได้

เอกสารอ้างอิง

- [1] ชัยพร ทบแป และคณะ, “การพัฒนากรอบสถาปัตยกรรมการสืบเปลี่ยนเอกสารธุรกรรมทางอิเล็กทรอนิกส์ระหว่างหน่วยงานภาครัฐที่มีประสิทธิภาพและความมั่นคงปลอดภัยด้วยเทคโนโลยีบล็อกเชน,” *วารสารเทคโนโลยีสารสนเทศ*, ปีที่ 17, ฉบับที่ 1, หน้า 66-75, 2563.
- [2] มณฑชา ชยการณ์วิกรม, *การใช้เทคโนโลยีบล็อกเชนในภาครัฐ*, สำนักงานพัฒนารัฐบาลดิจิทัล(องค์การมหาชน), 2562.
- [3] สิริกานต์ ไชยกุล และคณะ, “การพัฒนาแบบจำลองการดูแลสุขภาพโดยใช้เทคโนโลยีบล็อกเชน,” *วารสารการประชุมวิชาการ ระดับชาติ ครั้งที่ 17 มหาวิทยาลัยเกษตรศาสตร์*, หน้า 1138-1146, 2563.
- [4] B. K. Mohanta et al, “An Overview of Smart Contract and Use Cases in Blockchain Technology,” *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2018.
- [5] Chirtoaca et al, “A Framework for Creating Deployable Smart Contracts for Non-Fungible Tokens on the Ethereum Blockchain,” *In Proceeding of 2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*, pp. 100-105, 2020.
- [6] D. Zheng et al, “A Traceable Blockchain-Based Access Authentication System With Privacy Preservation in VANETs,” *IEEE Access*, vol. 7, pp. 117716-117726, 2019.
- [7] E.-Y. Daraghmi et al, “MedChain: A Design of Blockchain-Based System for Medical Records Access and Permissions Management,” *IEEE Access*, vol. 7, pp. 164595-164613, 2019.
- [8] Firmansyah Ashari et al, “Smart Contract and Blockchain for Crowdfunding Platform,” *International*

Journal of Advanced Trends in Computer Science and Engineering, no. 9, 2020.

- [9] F. M. Bencic et al, "DL-Tags: DLT and Smart Tags for Decentralized, Privacy-Preserving, and Verifiable Supply Chain Management," *IEEE Access*, vol. 7, pp. 46198-46209, 2019.
- [10] I. Tonev, "Energy Trading Web Platform Based on the Ethereum Smart Contracts and Blockchain," *2020 12th Electrical Engineering Faculty Conference (BulEF)*, 2020.

การพัฒนาระบบไอโอทีสำหรับเพาะเลี้ยงถั่งเช่าสีทอง

Development of IoT System for Cultivation of *Cordyceps militaris*

สมรรถชัย จันทรัตน์ (Samatachai Jantarat)¹ เจษฎา อรุณฤกษ์ (Jedsada Arunruerk)²

และพัฒนารพี สุนันทพจน์ (Patrapee Sunantapot)³

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

¹samatachai.j@en.rmutt.ac.th, ²jedsada.a@rmutt.ac.th.ac.th, ³patrapee.s@en.rmutt.ac.th

บทคัดย่อ

บทความฉบับนี้มีวัตถุประสงค์เพื่อนำเสนอการออกแบบพัฒนาระบบไอโอที (Internet of things - IoT) สำหรับควบคุมการเพาะเลี้ยงถั่งเช่าสีทอง (*Cordyceps militaris*) ซึ่งระบบควบคุมนี้จะทำการเฝ้าติดตามและควบคุมค่าสภาพแวดล้อมแบบอัตโนมัติ เช่น ค่าอุณหภูมิ ค่าความชื้น ค่าแสง ให้เหมาะสมกับการเพาะเลี้ยงถั่งเช่าสีทองทั้งสามระยะ คือ ระยะบ่มเชื้อ ระยะกระตุ้นให้เกิดไมตาตอกและระยะเลี้ยงดอก โดยเชื้อราที่ทำการตรวจจับในงานวิจัยนี้จะตรวจสอบสองสายพันธุ์ คือ เชื้อราไตรโคเดอร์มา หรือราที่มีสีเขียวและ เชื้อราดำ สำหรับส่วนระบบควบคุม นั้นมีผู้ใช้งานในระบบจะมีอยู่สองระดับ คือ ผู้ดูแลระบบ และผู้ใช้งานทั่วไป ซึ่งติดตามค่าสภาพแวดล้อมต่าง ๆ รวมถึง ภาพถ่ายภายในแต่ละขวดที่เพาะเลี้ยงผ่านทางกล้องเว็บแคมแบบช่วงเวลาจริงและประวัติย้อนหลัง ผ่านทางโมบายล์แอปพลิเคชัน และเว็บแอปพลิเคชันหลักของระบบ ผลลัพธ์จากการทดลองระบบต้นแบบนี้มีการความแม่นยำในการตรวจจับเชื้อราที่เกิดขึ้นร้อยละ 70

คำสำคัญ: ไอโอที การเพาะเลี้ยงถั่งเช่าสีทอง การควบคุมความชื้นและอุณหภูมิ

Abstract

This paper aims to present the design and development of the Internet of things (IoT) system for controlling the cultivation of *cordyceps militaris*. It automatically monitors and controls environmental values such as temperature, humidity, and light for the three phases of *cordyceps militaris*. The fungi detected in this study examined two

species: *Trichoderma* or a green mold, and a black mold.

As for the control system, there are two levels of users in the system: administrators and general users. They can monitor various environmental values, including photos inside each bottle cultured via real-time webcam and historical history through mobile applications and the system's main web applications. The results of this prototype experiment were 70 percent accurate in detecting mold in culture bottles.

Keywords: Internet of Things, *Cordyceps militaris*, Humidity and Temperature Control.

1. บทนำ

เห็ดถั่งเช่าจัดเป็นสมุนไพรที่ได้รับความนิยมอย่างมากในการบำรุงร่างกายตั้งแต่อดีตกาล ถั่งเช่ามีหลากหลายชนิดที่นิยมและมีชื่อเสียง ได้แก่ ถั่งเช่าทิเบต (*Cordyceps sinensis*), ถั่งเช่าหิมะหรือถั่งเช่าเกาหลิ (*Isaria tenuipes*), ถั่งเช่าจ๊กจั่นหรือว่านจ๊กจั่น (*Isaria sinclairii*) และถั่งเช่าสีทอง (*Cordyceps militaris*) ซึ่งถั่งเช่าสีทองนี้เป็นเห็ดที่เกิดขึ้นจากการเพาะเลี้ยงในห้องที่ควบคุมสภาพแวดล้อม โดยมีสรรพคุณทางยาและสารออกฤทธิ์ใกล้เคียงกับถั่งเช่าทิเบต มีสรรพคุณช่วยในการไหลเวียนของเลือด ช่วยต้านอนุมูลอิสระและลดการอักเสบ จึงเกิดความต้องการจากฝั่งผู้บริโภคมากและมีราคาสูงขึ้นตามปริมาณของสารที่อยู่ในถั่งเช่า โดยมีราคาอยู่ในช่วง 18,000 – 250,000 บาทต่อกิโลกรัม [1] จากราคาและความต้องการใช้ถั่งเช่าสีทอง ทำให้มีจำนวนเกษตรกรและผู้สนใจที่จะเพาะเลี้ยงเพิ่มมากขึ้น แต่มีบางส่วนยังประสบปัญหาในการควบคุมสภาพแวดล้อม

ไม่ว่าจะเป็น อุณหภูมิ ความชื้น แสง ซึ่งมีความต่างกันในแต่ละช่วงของการเพาะเลี้ยง รวมถึงปัญหาการเกิดราชนิดต่าง ๆ ขึ้นในการเพาะเลี้ยง

ดังนั้นทางคณะผู้วิจัยจึงได้พัฒนาต้นแบบของระบบไอโอทีสำหรับควบคุมการเพาะเลี้ยงถั่งเช่าสีทองขึ้นมาเพื่อเป็นแนวทางและเครื่องมือสำหรับเกษตรกรหรือผู้สนใจที่กำลังจะเริ่มเพาะเลี้ยง โดยมีวัตถุประสงค์สองประการคือ ประการแรกเพื่อพัฒนาระบบไอโอทีสำหรับควบคุมสภาพแวดล้อมที่เหมาะสมกับการเจริญเติบโตของเห็ดถั่งเช่าและอีกประการเพื่อการติดตามปัญหาการเกิดเชื้อราจึงได้เพิ่มส่วนโมเดลสำหรับการตรวจจับการเกิดเชื้อราในขวดเพาะเลี้ยงถั่งเช่าและมีการส่งข้อมูลแจ้งเตือนไปยังผู้ดูแลเพื่อลดความเสียหายในการเพาะเลี้ยง

2. แนวคิดและงานวิจัยที่เกี่ยวข้อง

2.1 ไอโอที (Internet of Thing - IoT)

อินเทอร์เน็ตของทุกสรรพสิ่งหรือไอโอที [2] เป็นแนวคิดระบบเครือข่ายทั้งแบบมีสายและแบบไร้สาย ช่วยให้สามารถเชื่อมต่ออุปกรณ์ เครื่องมือ เช่น เซอร์โวมอเตอร์คอนโทรลเลอร์ หรือวงจรควบคุมต่าง ๆ เพื่อแลกเปลี่ยนข้อมูลต่าง ๆ กันได้ผ่านเครือข่ายอินเทอร์เน็ตซึ่งมีการใช้งานอย่างกว้างในทุกอุตสาหกรรม ไม่ว่าจะเป็นในด้านอุตสาหกรรม ธุรกิจ คริวเรือนทั้งในด้านสุขภาพอาหาร สิ่งแวดล้อม และรวมถึงในภาคการเกษตรด้วย [3] ซึ่งมีการนำเทคโนโลยีด้านไอโอทีเข้ามาตรวจวัดค่าสภาพแวดล้อมและทำการควบคุมให้อยู่ในค่าที่เหมาะสมกับการเจริญเติบโตของพืชแต่ละชนิด พร้อมทั้งอำนวยความสะดวกกับผู้เพาะปลูกให้สามารถติดตามข้อมูลต่างๆ ได้จากแอปพลิเคชันที่ให้บริการผ่านเครือข่ายอินเทอร์เน็ต

2.2 ถั่งเช่าสีทอง

ถั่งเช่าสีทองให้คุณค่าทางโภชนาการคือ กรดอะมิโน 16 ชนิด รวม 69.32 มก./ก. น้ำหนักแห้ง กรดไขมัน 6 ชนิด รวมถึงสารอดีโนซีน และคอร์โคเซปินที่ ซึ่งเป็นสารที่สำคัญ และสรรพคุณทางยาของเห็ดถั่งเช่าจะมีปริมาณมากน้อยก็ขึ้นอยู่กับปัจจัยหลายอย่าง เช่น สายพันธุ์ การเพาะเลี้ยง

อาหารที่เพาะเลี้ยง ค่าสภาพแวดล้อมต่าง ๆ ในการเพาะเลี้ยงถั่งเช่านั้นจะทำให้ได้ 2 แบบคือ การเพาะด้วยหนอนหรือดักแด้ และการเพาะด้วยอาหารที่ผสมขึ้นมาเอง เช่น จาก ดักแด้ รัชพูพืช ข้าว ข้าวโพด ข้าวฟ่าง ลูกเดือย ซึ่งต้องมีสารอาหารประเภทคาร์บอน ไนโตรเจน วิตามินและแร่ธาตุต่าง ๆ [4] ตัวอย่างสูตรอาหารที่ใช้เพาะเลี้ยงมีวัตถุดิบและปริมาณดังนี้ 1) น้ำสะอาด 1.2 ลิตร 2) ข้าวสาร 25-30 กรัม 3) มันฝรั่ง 200 กรัม 4) ยีสต์ 5 กรัม 5) เปปโตน 5 กรัม 6) กลูโคส 20 กรัม และ 7) ตัวหนอนไหมหรือดักแด้ 200 กรัม [5] ส่วนค่าสภาพแวดล้อมต่างๆ ที่ควบคุมคือค่า อุณหภูมิประมาณ 20 องศาเซลเซียส ยกเว้นระยะกระตุ้นให้เกิดปุ่มตาออกให้ประมาณ 17 องศาเซลเซียส ร้อยละของความชื้นประมาณ 60 – 70 และแสงที่ 1,000 ลักซ์ ที่ 12 ชั่วโมงต่อวัน [6]

เชื้อราที่พบได้บ่อยในการเพาะเลี้ยงถั่งเช่าสีทอง ได้แก่ เชื้อราไตรโคเดอร์มาและเชื้อราดำ เชื้อราไตรโคเดอร์มา [7] นั้นจะเป็นเชื้อราที่เป็นศัตรูต่อเชื้อราโรคพืชหลายชนิด มีลักษณะเป็นสีเขียว เจริญเติบโตได้ดีบนเศษซากพืชซากอินทรีย์วัตถุในธรรมชาติ ชอบสภาพดินที่มีความชื้นแต่ไม่แฉะ จัดเป็นปรสิต (Parasite) มีการพันรัดเส้นใยเชื้อราแล้วสร้างเอนไซม์ เข้าไปย่อยสลายผนังเส้นใยของพืชทำให้สูญเสียชีวิตลงและเชื้อราดำกลุ่มแอสเพอร์จิลลัส (Aspergillus) สร้างความเสียหายต่อผลผลิตมากและยังก่อให้เกิดสารที่เป็นอันตรายต่อมนุษย์ เช่น อะฟลาทอกซิน (Aflatoxin)

2.3 งานวิจัยที่เกี่ยวข้อง

สัญญา วรรคิต และคณะ [8] ได้พัฒนาต้นแบบของระบบควบคุมการเพาะเห็ดหลายชนิดในโรงเรือนขนาดเล็ก โดยมีการควบคุมโดยใช้ไมโครคอนโทรลเลอร์ในกลุ่มของอาร์ดูโน มาเป็นตัวหลักในการควบคุมอุณหภูมิ ความชื้นภายในโรงเรือนเพาะเห็ด มีการสื่อสารกับอุปกรณ์ภายในโรงเรือนแบบไร้สายผ่านโมดูล NRF 24L01 และส่งข้อมูลที่ตรวจวัดได้ไปเก็บไว้ยังระบบคลาวด์ของผู้ให้บริการที่ชื่อ ThingSpeak เพื่อแสดงผลแบบเวลาจริง

สมประสงค์ อินทรักษ์ และสุนันทา ศรีม่วง [9] ได้นำเสนอการพัฒนาแอปพลิเคชันสำหรับควบคุม

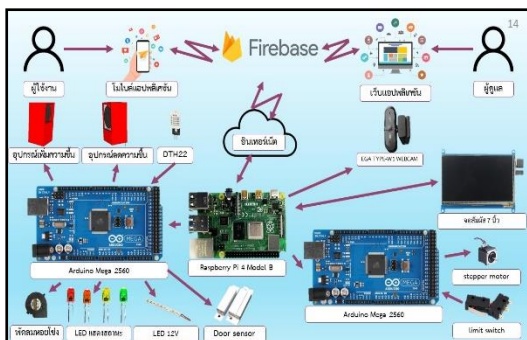
อุปกรณ์ไฟฟ้าผ่านระบบ ซึ่งเป็นการพัฒนาแอปพลิเคชันบนโทรศัพท์เคลื่อนที่เพื่อใช้ในการควบคุมเครื่องใช้ไฟฟ้าแบบระยะไกล โดยมีส่วนของไมโครคอนโทรลเลอร์ทำหน้าที่หลักในการเชื่อมต่อประสานเพื่อควบคุมการเปิดปิดอุปกรณ์ไฟฟ้าผ่านโทรศัพท์เคลื่อนที่ได้

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้ผู้วิจัยมีวัตถุประสงค์ในเพื่อพัฒนาต้นแบบระบบไอโอทีเพื่อควบคุมสภาพแวดล้อมให้เหมาะสมกับการเพาะเลี้ยงเห็ดถั่งเช่าสีทองที่อยู่ในขวดเพาะเลี้ยง และมีการตรวจจับการเกิดราเขียวหรือราดำในขวดเพาะเลี้ยง

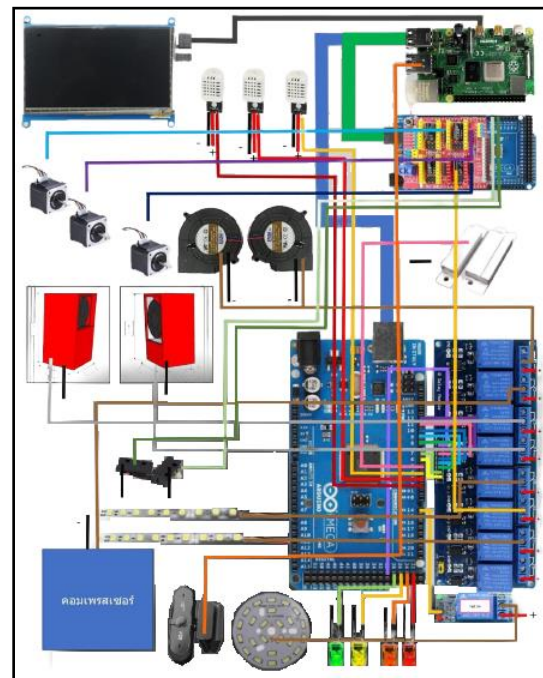
ขั้นตอนในการออกแบบและพัฒนาตัวอุปกรณ์ต่าง ๆ ของระบบผู้วิจัยได้แบ่งออกเป็น 3 ส่วน คือ ส่วนแรกเป็นส่วนตัวตู้เพาะเลี้ยง เป็นการนำตู้เย็นเก่ามาทำการตัดแปลงเป็นตู้เพาะปลูก โดยต่อเข้ากับวงจรส่วนของชุดควบคุมสภาพแวดล้อม และชุดควบคุมกล้องเว็บแคม ซึ่งเป็นส่วนที่เป็นระบบกึ่งปิดเพื่อควบคุมระดับของอุณหภูมิ ความชื้น แสง ให้เป็นไปตามค่าระดับที่ตั้งไว้ก่อนเพาะเลี้ยงในแต่ละรอบ ส่วนที่สองเป็นส่วนของเว็บแอปพลิเคชันที่เป็นซอฟต์แวร์หลักในของการควบคุมจัดการทั้งหมด และส่วนสุดท้ายเป็นซอฟต์แวร์สำหรับ โทรศัพท์เคลื่อนที่ที่ทำงานบนระบบปฏิบัติการแอนดรอยด์ (Android) รุ่น 8 ขึ้นไป เพื่อให้สะดวกในการติดตามการทำงานของกาเพาะเลี้ยง ส่วนในขั้นตอนการเพาะเลี้ยงถั่งเช่านั้นจะแบ่งเชื้อที่ใช้เพาะเลี้ยงออกเป็นสองส่วนคือส่วนที่เพาะเลี้ยงปกติที่ใช้เพาะเลี้ยงโดยการควบคุมด้วยระบบ และอีกส่วนจะเพาะเลี้ยงให้เกิดรา เพื่อใช้สำหรับเตรียมภาพไว้สำหรับสร้างโมเดลสำหรับการตรวจจับเมื่อเกิดขึ้นในขวดเพาะเลี้ยง

3.1 ภาพรวมการออกแบบระบบ



ภาพที่ 1: ภาพรวมของระบบไอโอที

จากภาพที่ 1 ซึ่งเป็นภาพรวมของระบบที่ผู้วิจัยได้ออกแบบไว้ โดยจะมี 3 ส่วนหลักคือ ส่วนผู้ใช้งาน ส่วนผู้ดูแลระบบ และส่วนของตัวตู้เพาะเลี้ยง โดยส่วนของผู้ใช้งานจะเป็นผู้ที่ทำการเพาะเลี้ยงซึ่งอาจเป็นเกษตรกรหรือผู้สนใจ โดยจะมีการใช้งานผ่านทางโมบายแอปพลิเคชัน ส่วนผู้ดูแลระบบจะเป็นผู้ที่ดูแลในส่วนของเว็บแอปพลิเคชัน และส่วนสุดท้ายส่วนของตัวตู้เพาะเลี้ยง ซึ่งเป็นส่วนหลัก โดยจะประกอบด้วยวงจร 3 ส่วน คือส่วนชุดควบคุมสภาพแวดล้อมในตัว เช่น เซ็นเซอร์ต่าง ๆ และชุดควบคุมอุณหภูมิ ความชื้น แสง ส่วนที่สองชุดควบคุมกล้องเว็บแคมใช้สำหรับควบคุมการเคลื่อนที่ของกล้องเพื่อใช้ในการถ่ายภาพขวดที่เพาะเลี้ยงทั้งหมดในตัว โดยทั้งสองส่วนจะใช้ Arduino Mega 2560 เป็นบอร์ดหลักและส่วนสุดท้ายคือส่วนที่เป็นตัวควบคุมการบันทึกภาพขวดที่เพาะเลี้ยง การติดต่อกับผู้ใช้งานผ่านหน้าที่คิดไว้ด้านหน้าของผู้ดูแลติดต่อแลกเปลี่ยนข้อมูลกับส่วนของฐานข้อมูล โดยในการทดลองนี้ใช้ Firebase เป็นหลัก และใช้ Raspberry Pi 4 Model B เป็นบอร์ดหลัก



ภาพที่ 2: การเชื่อมต่อของอุปกรณ์ตู้เพาะเลี้ยง

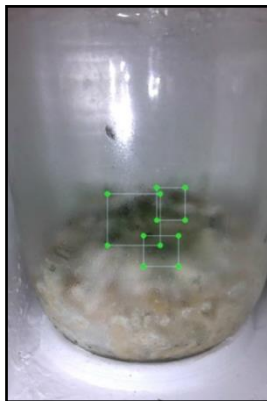


ภาพที่ 3: ด้านหน้าตู้เพาะเลี้ยง

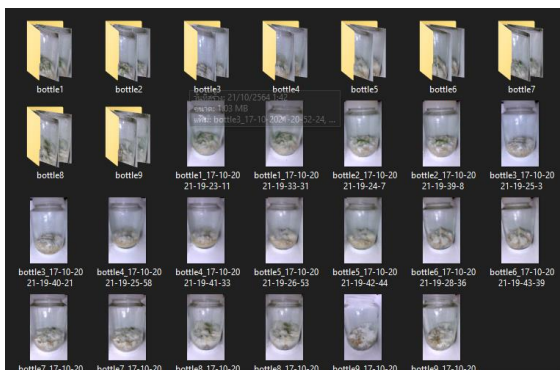
3.2 การเตรียมข้อมูล

ในเตรียมข้อมูลเพื่อนำไปสร้างโมเดลให้ตรวจจับเชื้อรา นั้นจะมีลำดับดังนี้

1) ทำการเพาะตั้งเชื้อราในขวดเพาะและใส่เชื้อราลงไป จากนั้นเมื่อเกิดเป็นราจะทำการบันทึกภาพขวดเพาะเลี้ยงเชื้อราทุกขวดดังภาพที่ 4 และ ภาพรวมทั้งหมดที่ใช้ในการเทรนโมเดลดังภาพที่ 5



ภาพที่ 4: ตัวอย่างขวดเพาะเลี้ยงที่เกิดรา



ภาพที่ 5: ชุดข้อมูลของขวดเพาะเลี้ยงที่เกิดรา

2) นำรูปที่บันทึกไปทำการระบุตำแหน่งของเชื้อรา และชนิดของเชื้อราที่เกิดขึ้นภายในขวดเพาะเลี้ยง ด้วยโปรแกรม RoboFlow [10] โดยเลือก License เป็น CC BY 4.0 และเลือก Project Type เป็น Object Detection (Bounding Box) และให้ Export ออกเป็นรูปแบบของ YOLO v5 PyTorch โดยมีจำนวนรูป 806 รูป

3) นำไฟล์ที่ได้ไปเทรนใน Google Colab โดยใช้ภาพที่อนุพัทธ์ที่มีความกว้าง 640 พิกเซล มีจำนวนข้อมูล (Batch) เท่ากับ 16 และจำนวนรอบ (Epoch) เท่ากับ 300 และเมื่อทดสอบโมเดลที่ได้จะมีความแม่นยำค่อนข้างสูงอยู่ในช่วง 0.88 – 0.95

3.3 แผนการเพาะเลี้ยง

ในการเพาะเลี้ยงตั้งเชื้อราครั้งนี้ ผู้วิจัยได้ทำการแบ่งช่วงอายุของการเพาะเลี้ยงออกเป็น 3 ระยะคือ

1) ระยะบ่มเชื้อ จะควบคุมให้อุณหภูมิอยู่ในช่วง 19 – 21 องศาเซลเซียส ความชื้นร้อยละ 75 – 80 และไม่ให้มีแสงสว่าง ใช้ระยะเวลา 21 วัน

2) ระยะกระตุ้นให้เกิดปุ่มตาดอก จะมีการควบคุมให้อุณหภูมิอยู่ในช่วง 16 – 18 องศาเซลเซียส ความชื้นร้อยละ 65 – 70 และให้มีแสงสว่าง 12 ชั่วโมงต่อวัน ใช้ระยะเวลา 28 วัน

3) ระยะเลี้ยงดอก จะมีการควบคุมให้อุณหภูมิอยู่ในช่วง 19 – 21 องศาเซลเซียส ความชื้นร้อยละ 75 – 80 และให้มีแสงสว่าง 12 ชั่วโมงต่อวัน

4. ผลการดำเนินงาน

ผลลัพธ์การออกแบบส่วนต่าง ๆ ที่ติดต่อกับผู้ใช้งาน จะแบ่งออกเป็น 3 ส่วน ดังนี้

1) ส่วนของผู้เพาะเลี้ยง

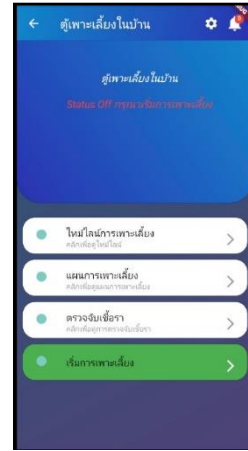
ในการเพาะเลี้ยงแต่ละรอบจะเพาะเลี้ยงจำนวน 9 ขวด โดยมีการจัดเป็น 3 แถว แถวละ 3 ขวด ดังภาพที่ 6 และหน้าจอแบบสัมผัสและหลอดแอลอีดีที่ติดต่อกับผู้ใช้งาน ดังภาพที่ 7 โดยที่เป็นส่วนที่ผู้ใช้งานสามารถเลือกเมนูเพื่อดูค่าอุณหภูมิ ความชื้น แสง โอโซน ภาพขวดเพาะเลี้ยง และมีหลอดแอลอีดีจำนวน 4 หลอดซึ่งจะกระพริบเมื่อมี

ความผิดพลาดเกิดขึ้น เช่น อุณหภูมิไม่อยู่ในช่วงที่ตั้งไว้ หรือประตูของผู้ถือเปิด

เหมือนส่วนของ โมบายล์แอปพลิเคชันแต่เพิ่มในส่วนของการจัดการบัญชีผู้ใช้ของทุกคนในระบบ ดังภาพที่ 9



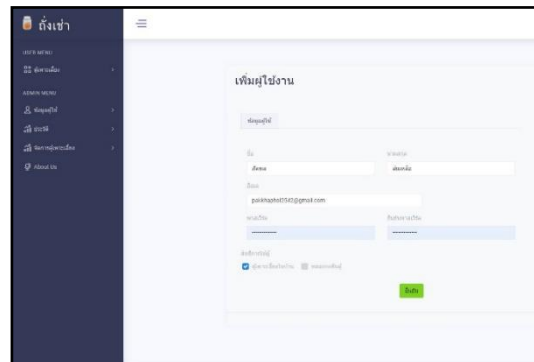
ภาพที่ 6: ชั้นวางขวดเพาะเลี้ยงในตู้



ภาพที่ 8: หน้าจอเมนูหลักของโมบายล์



ภาพที่ 7: จอสัมผัสด้านหน้าของตู้



ภาพที่ 9: เว็บแอปพลิเคชัน ส่วนเพิ่มผู้ใช้งาน

2) ส่วนการติดต่อกับผู้ใช้งานผ่าน โมบายล์แอปพลิเคชัน ซึ่งจะมีเมนูหลักอยู่ 4 รายการคือ รายการแรก โอนไลน์การเพาะเลี้ยง ใช้จัดการรอบของการเพาะเลี้ยง รายการที่สอง แผนการเพาะเลี้ยง ใช้สำหรับการปรับแต่งค่าควบคุมสภาพแวดล้อมทั้ง 3 ระยะการเพาะเลี้ยง เช่นอุณหภูมิ ความชื้น แสง ของแต่ละช่วงการเพาะเลี้ยง ซึ่งผู้ใช้งานสามารถปรับได้ตามต้องการ รายการที่สาม ตรวจสอบเชื้อรา เป็นรายการที่ให้เห็นขวดเพาะเลี้ยงใดมีราเกิดขึ้น โดยระบบจะเคลื่อนกล้องและถ่ายภาพในเวลาปัจจุบันทั้ง 9 ขวดแล้วแจ้งว่ามีราเกิดขึ้นหรือไม่ และรายการสุดท้ายเป็นรายการเริ่มการเพาะเลี้ยง เป็นการเริ่มต้นใช้ค่าต่างๆ ที่ตั้งไว้ให้เริ่มทำการควบคุม ดังภาพที่ 8 ที่แสดงภาพหน้าจอ โมบายล์แอปพลิเคชัน

ผลลัพธ์ในการทดลองเมื่อเพาะเลี้ยงมีดังนี้ ภาพที่ 10 เป็นภาพที่แสดงการเพาะเลี้ยงแสดงด้านซ้ายเป็นส่วนที่เพาะเลี้ยงและมีสีเป็นหลืองทอง และภาพด้านขวาเป็นขวดที่ตรวจจับได้ว่ามีราเกิดขึ้นในระหว่างทำการเพาะเลี้ยง

3) ส่วนของเว็บแอปพลิเคชัน เป็นส่วนงานเบื้องหลัง ที่ผู้ใช้หลักคือผู้ดูแลระบบ โดยมีความสามารถพื้นฐาน



ภาพที่ 10: ขวดที่ทำการเพาะเลี้ยงและขวดที่มีราเกิดขึ้น

สำหรับผลลัพธ์ที่ได้จากการทดลองในเรื่องการตรวจจับรา ได้นำขวดเพาะเลี้ยงที่มีเชื้อรา 5 ขวดและขวดเพาะเลี้ยงที่ไม่มีเชื้อราจำนวน 4 ขวดใส่ในตู้และสั่งให้ทำการตรวจสอบรวม 20 รอบ ได้ผลลัพธ์ดังนี้

ตารางที่ 1: ผลการตรวจจับเชื้อราในขวดเพาะเลี้ยง

รอบ	พบเชื้อรา	รอบ	พบเชื้อรา
1	5	11	6
2	6	12	6
3	5	13	7
4	5	14	6
5	5	15	5
6	5	16	5
7	5	17	5
8	5	18	5
9	6	19	5
10	5	20	5

จากผลการทดลอง พบว่าขวดที่มีเชื้อราจะตรวจพบทุกขวด แต่บางรอบจะมีความผิดพลาดโดยนับขวดที่ไม่มีเชื้อเป็นพบเชื้อด้วยคือ 6 ครั้งจาก 20 ครั้ง ดังนั้นค่าความถูกต้องจะเท่ากับ $(20 - 6)/20 = 0.7$ หรือ ร้อยละ 70

5. สรุป

งานวิจัยนี้ได้ศึกษาถึงการออกแบบและพัฒนาระบบไอโอทีสำหรับเพาะเลี้ยงถั่งเช่าสีทองโดยที่มีการเพิ่มความสามารถในการตรวจจับการเกิดราในขวดเพาะเลี้ยงโดยใช้โมเดล YOLO v5 โดยการตั้งค่าเวลาในการถ่ายภาพขวดเพาะเลี้ยงแล้วนำมาวิเคราะห์ว่ามีราเกิดขึ้นหรือไม่ โดยในการทดลองมีความแม่นยำที่ 70%

โดยในส่วนที่มีความผิดพลาดนั้นพบว่าขณะทำการทดลองนั้นพบปัญหาเรื่องการเกิดไอน้ำจับภายในขวดทดลองทำให้ขณะถ่ายภาพทำให้เกิดเงาขึ้นได้ และระยะห่างของกล้องกับตัวขวดเพาะเลี้ยงอยู่ใกล้ไปทำให้บางครั้งภาพที่ถ่ายเกิดการเบลอขึ้นได้

เอกสารอ้างอิง

- [1] วัชรินทร์ อัม ทองกลาง, “การประยุกต์ใช้ฐานเห็ดถั่งเช่าสีทองเพื่อเป็นอาหารเสริมสำหรับสุกรอนุบาล”, วิทยานิพนธ์ วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยนเรศวร, 2564.
- [2] ฉัตรพงศ์ ชูแสงนิล, อินเทอร์เน็ตแห่งสรรพสิ่ง (ออนไลน์). สืบค้น 1 ตุลาคม 2565, จาก <https://www.scimath.org/article-technology/item/9089-2018-10-18-07-59-07>
- [3] พีรพงษ์ คล้ายปาน เกษม พวงเพชร และ วีระยุทธ ประทุมมิ่ง, “ระบบการควบคุมการปลูกพืชในโรงเรือน” วิทยานิพนธ์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2558.
- [4] สุภาพร อวัญ, “สูตรอาหารเลี้ยงเชื้อและวิธีการที่เหมาะสมในการเพาะเลี้ยงเห็ดถั่งเช่าสีทอง”, วารสารวิชาการสถาบันการอาชีวศึกษาเกษตร ปีที่ 3 ฉบับที่ 2, 2562 หน้า 15 – 23.
- [5] ชัญญา ทะพิงค์แก, การเพาะเห็ดถั่งเช่าสีทอง พิมพ์ครั้งที่ 2 ห้างหุ้นส่วนจำกัด เฟรม-อัป ดีไซน์, 2559.
- [6] ยาวภา ทองอร่าม, “การศึกษาและเปรียบเทียบการเพาะเลี้ยงถั่งเช่าสีทองโดยใช้หัวเชื้อเหลวและหัวเชื้อแข็งของ Cordyceps militaris บนเมล็ดธัญพืช” วารสารวิชาการ โรงเรียนนายร้อยพระจุลจอมเกล้า ปีที่ 13 ฉบับที่ 13 2558 หน้า 54 – 64.
- [7] ศูนย์ส่งเสริมเทคโนโลยีการเกษตรด้านอารักขาพืชจังหวัดขอนแก่น เชื้อราไตรโคเดอร์มา สืบค้น 1 ตุลาคม 2565 จาก <http://www.pmc04.doae.go.th/Myweb-2011-data1/11%20Trichoderma/11%20Trichoderma.html>.
- [8] ฉัญญา ควรคิด, ก้องภพ ชามานาดย์ และ ธวัชชัย ทองเหลี่ยม, “การพัฒนาต้นแบบระบบควบคุมการเพาะเห็ดหลายชนิดภายในโรงเรือน”, วารสารงานวิจัยและพัฒนาเชิงประยุกต์ โดยสมาคม ECTI ปีที่ 1 ฉบับที่ 2 หน้า 54 – 64.
- [9] สมประสงค์ อินทรักษ์ และ สุนันทา ศรีม่วง, “การพัฒนาแอปพลิเคชันสำหรับควบคุมอุปกรณ์ไฟฟ้าผ่านระบบ”, วารสารโครงการวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ปีที่ 3 ฉบับที่ 1 2560, หน้า 57 – 62.
- [10] Roboflow <https://roboflow.com/>

การประยุกต์ใช้เทคนิคการประมวลผลภาพ และความเข้าใจภาษาธรรมชาติ ร่วมกับ ไลน์ แชทบอท ให้ความรู้ด้านการเลี้ยงกระต่าย Applying Image Processing and Natural Language Understanding with LINE Chatbot for Rabbit Caring

ปริยาวัจน์ ตมะวิโมกษ์ (Preeyarwat Tamawimok)¹ และ ณัฐวี อดุตฤกษ์ (Natavee Utakrit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹S6407021810020@email.kmutnb.ac.th, ²nattavee.u@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้เป็นการศึกษาและพัฒนาแชทบอท เพื่อช่วยเหลือ และให้คำแนะนำด้านการเลี้ยงกระต่าย โดยแชทบอทจะทำงานบนแพลตฟอร์ม แอปพลิเคชัน ไลน์ (LINE) ซึ่งเป็นสื่อสังคมออนไลน์ที่เป็นที่นิยมอย่างมากในปัจจุบัน เพื่อความสะดวกในการเข้าถึง และตอบคำถามได้อย่างรวดเร็วทันที โดยการพัฒนา แชทบอท จะใช้เทคนิคที่เรียกว่าความเข้าใจภาษาธรรมชาติ (Natural Language Understanding) เพื่อช่วยให้แชทบอท ได้เข้าใจถึงสิ่งที่ผู้ใช้งานพิมพ์ข้อความ และสามารถตีความและตอบคำถามได้อย่างถูกต้อง อีกทั้งยังได้นำ เทคนิคการประมวลผลภาพ (Image Processing) มาใช้ในการตรวจสอบสายพันธุ์กระต่ายที่ผู้ใช้งานเลี้ยง เพื่อคำนึงถึงความต้องการพิเศษของกระต่ายในแต่ละสายพันธุ์ ผลลัพธ์จากการทดสอบพบว่าค่าความถูกต้องของการตอบคำถาม อยู่ที่ 86.67 % และความแม่นยำของการตรวจสอบสายพันธุ์กระต่ายอยู่ที่ 76.67 %

คำสำคัญ: ไลน์ แชทบอท ความเข้าใจภาษาธรรมชาติ

การประมวลผลรูปภาพ

Abstract

This research was to study and create a Chatbot to help and give some advices about the rabbit care via LINE application in order to answer questions quickly. The Chatbot applied Natural Language Understanding (NLU) techniques to understand the

user's questions correctly and respond quickly. Additionally, this research also used Image Processing to identify rabbit breeds that users raise, in order to better care to the specific needs of each breed. The evaluation results showed that the accuracy of the correct answer and the rabbit breed identification are 86.67% and 76.67% respectively.

Keywords: LINE, Chatbot, Natural Language Understanding, Image processing

1. บทนำ

ปัจจุบัน การเลี้ยงสัตว์ของทั่วโลกมีแนวโน้มที่จะเพิ่มมากขึ้นทุก ๆ ปี ส่งผลให้ประชากรสัตว์เลี้ยงมีจำนวนมากขึ้นอย่างต่อเนื่อง [1] สัตว์เลี้ยงนอกจากสุนัขแล้ว ยังมีสัตว์เลี้ยงที่เราสามารถเลี้ยงได้อีกหรือที่เราเรียกกันในชื่อกลุ่มสัตว์เลี้ยง Exotic [2] หนึ่งในกลุ่มสัตว์เลี้ยง Exotic ที่คนนิยมเลี้ยงกันมานานก็คือ กระต่าย ด้วยความที่กระต่ายเป็นสัตว์หน้าตาน่ารัก ไม่ส่งเสียงรบกวนผู้เลี้ยง กระต่ายจึงเป็นสัตว์เลี้ยงยอดนิยมเนื่องจากมีความคิดว่ากระต่ายเลี้ยงในกรง ให้อาหารตามเวลาที่อยู่ได้ แต่ความจริงแล้วไม่ใช่ ผู้เลี้ยงกระต่าย ต้องทำความเข้าใจในตัวกระต่ายก่อนไม่เช่นนั้น อาจส่งผลให้มีการทอดทิ้งกระต่าย หรือทำให้กระต่ายตายโดยใช่เหตุ

ในขณะเดียวกัน เทคโนโลยีก็มีการพัฒนาหุ่นยนต์จำลองบทสนทนาของมนุษย์ หรือที่เราเรียกกันว่า แชทบอท โดยมีการใช้ เทคโนโลยี Artificial Intelligent AI หรือ

ปัญญาประดิษฐ์ใช้ในการโต้ตอบกับคู่สนทนา ซึ่งแชทบอทที่ถูกพัฒนาด้วย Natural Language Processing (NLP) โดยใช้ Machine Learning ส่งผลให้ AI Bot ได้รับความสนใจและความนิยมจากผู้ใช้เป็นอย่างมาก โดยใช้เครื่องมือที่เรียกว่า Dialogflow เนื่องจาก Dialogflow มีการใช้อัลกอริทึม Natural Language Processing หรือ การประมวลผลภาษาธรรมชาติ ช่วยให้ข้อจำกัดทางด้านโต้ตอบของแชทบอทมีน้อยลง และ Dialogflow สามารถทำให้แชทบอทมีการเรียนรู้เกี่ยวกับ Natural Language Understanding อีกด้วย

นอกจากการที่เราสามารถนำแชทบอทมาใช้ในการช่วยเหลือเรื่องการให้ความรู้ และคำแนะนำด้านการเลี้ยงกระต่ายแล้ว ทางผู้วิจัยเห็นว่าถ้าหากแชทบอทนี้สามารถจำแนกสายพันธุ์ของกระต่ายได้ผ่านการถ่ายรูปส่งผ่านใน LINE ก็จะทำให้ผู้ใช้งานแชทบอทหรือผู้เลี้ยงกระต่ายได้เข้าใจลักษณะพิเศษของกระต่ายที่เลี้ยงอยู่ จึงได้มีการนำเทคโนโลยีแชทบอท และเทคโนโลยีการประมวลผลภาพเข้ามาผนวกกับความรู้ด้านการดูแลกระต่าย บนแพลตฟอร์มโซเชียลมีเดีย ไลน์ LINE จะช่วยเหลือและให้คำแนะนำให้แก่ผู้ที่สนใจเกี่ยวกับกระต่ายเบื้องต้นได้อย่างรวดเร็ว และถูกต้องอย่างมีประสิทธิภาพ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ข้อมูลเบื้องต้นเกี่ยวกับกระต่าย

กระต่ายเป็นสัตว์เลี้ยงลูกด้วยนม มีขนาดตัวที่เล็ก ไปจนถึงขนาดใหญ่ มีขนทั่วตัว มีใบหูยาว มีร่างกายที่บอบบางเมื่อเทียบกับสัตว์อื่น ๆ ดังนั้นกระต่ายจึงเป็นสัตว์ที่ต้องการการเลี้ยงดูอย่างทะนุถนอมและต้องเอาใจใส่เป็นอย่างมาก

2.2 แชทบอท (แชทบอท)

แชทบอท คือซอฟต์แวร์ที่พัฒนาขึ้นมาเพื่อช่วยตอบกลับการสนทนาผ่านข้อความหรือเสียงแบบ อัตโนมัตินี้และรวดเร็วซึ่งสามารถใช้งานได้ทั้งบน แอปพลิเคชันไลน์ LINE / แอปพลิเคชัน Facebook / แอปพลิเคชัน Instagram รวมทั้งบนเว็บไซต์ด้วย

2.3 LINE Messaging API

Line Messaging API คือ การสื่อสารระหว่าง server ของผู้พัฒนา ไปยัง user LINE ผ่านทาง LINE Platform โดยตัว server ผู้พัฒนาจะต้องเชื่อมต่อกับ LINE Platform และเมื่อ มี user เพิ่ม account LINE เป็นเพื่อน หรือ ส่งข้อความมาหา ทาง LINE Platform จะทำการส่ง request มายัง server ที่ผู้พัฒนาลงทะเบียนผูกไว้กับ LINE account นั้นทันที วิธีนี้เรียกว่า Webhook ซึ่งมันทำให้ผู้ใช้งานรู้สึกเหมือนกับว่าได้โต้ตอบกับคนจริงๆ

2.4 การประมวลผลภาษาธรรมชาติ

การประมวลผลภาษาธรรมชาติ (Natural language processing - NLP) เป็นสาขาย่อยของภาษาศาสตร์วิทยาการคอมพิวเตอร์ วิศวกรรมสารสนเทศ และปัญญาประดิษฐ์ เนื่องจากภาษาที่มนุษย์ใช้กันนั้นมีความซับซ้อน มีวิธีการแสดงออก มากมายนับไม่ถ้วน ทั้งในด้านการสื่อสารด้วยคำพูดหรือข้อความที่เกิดขึ้นด้วยการเขียน [3] นอกจากนี้ ยังมีแขนงย่อยของวิทยาการด้าน NLP ซึ่งเรียกว่า Natural language understanding (NLU) หรือความเข้าใจในภาษาธรรมชาติ [5] ซึ่งกำลังได้รับความสนใจเพิ่มขึ้นมากในปัจจุบัน เนื่องจากสามารถตีความเจตนาที่แฝงอยู่ รวมถึงถอดความหมายของบริบทต่าง ๆ และความกำกวมของการใช้คำศัพท์ได้อีกด้วย

2.5 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) หมายถึง การนำภาพมาประมวลผลหรือคิดคำนวณด้วยคอมพิวเตอร์ เพื่อให้ได้ข้อมูลที่เรากำลังต้องการทั้งในเชิงคุณภาพและปริมาณ โดยมีขั้นตอนต่างๆ เช่น การแบ่งส่วนของวัตถุที่เราสนใจ ออกมาจากภาพ เพื่อนำภาพวัตถุที่ได้ไปวิเคราะห์หาข้อมูล หรือการปรับแต่งภาพ (Data Augmentation) [6] เช่น การหมุน แต่งภาพให้สว่างขึ้น หรือกลับด้าน จากนั้นเราสามารถนำข้อมูลที่ได้มาไปวิเคราะห์ และสร้างเป็นระบบ เช่น ระบบดูแลและตรวจสอบสภาพการจราจรบนท้องถนน ระบบตรวจสอบคุณภาพของผลิตภัณฑ์ในกระบวนการผลิตของโรงงานอุตสาหกรรม เป็นต้น [7]

2.7 งานวิจัยที่เกี่ยวข้อง

ร้อยตำรวจโทเกียรติพันธ์ศักดิ์ [8] ได้ทำการวิจัยเรื่องการประยุกต์ใช้แชทบอทสนับสนุนงานสอบสวน

กรณีศึกษาสถานีตำรวจภูธรเบตง เนื่องจากมีผู้ที่เข้ามา ร้องเรียนสอบถามปัญหาข้อกฎหมาย จำนวน 942 เรื่อง คิดเป็น 78 เรื่องต่อเดือน มีการใช้ Facebook ของสถานี ตำรวจในการร้องเรียนและสอบถามปัญหาด้านกฎหมาย ซึ่งพบปัญหาหลายอย่าง เช่น ความล่าช้าในการตอบคำถาม และเจ้าหน้าที่พนักงานทุกคนไม่มีความรู้ข้อกฎหมาย ทำให้เสียเวลาในการสอบถามผู้รู้ และส่งผลทำให้เกิดความ ล่าช้าในการให้บริการในแต่ละกรณีค่อนข้างมาก ผู้วิจัยจึง มีการพัฒนาระบบแชทบอท โดยผนวกกับการใช้ Dialogflow สำหรับใช้ใน Facebook

I-Ching Hsu and Jiun-De Yu [9] ได้ทำการวิจัยเรื่อง A medical chatbot using machine learning and natural language understanding โดยงานวิจัยนี้เป็นงานวิจัยที่สร้าง แชทบอทสำหรับให้คำแนะนำด้านอาการและโรคต่าง ๆ ด้วยภาษาจีน ซึ่งภาษาจีนเป็นภาษาที่จำเป็นต้อง ครอบรูประโยค การใช้เทคนิค Natural language process อย่างเดียวอาจจะทำให้การสื่อสารผิดพลาด ซึ่งจำเป็นต้อง ใช้เทคนิค Natural language understanding มาช่วยในการ วิเคราะห์รูประโยคเพื่อให้แชทบอทสามารถ ตอบคำถามได้ถูกต้อง

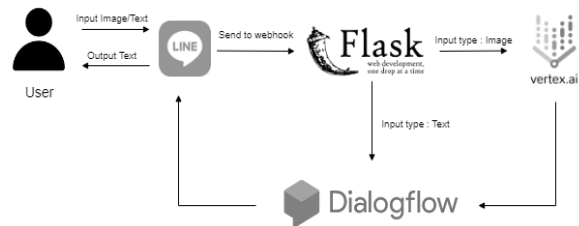
Sotir I. Sotirov และคณะ [10] ได้ทำการวิจัยเรื่อง Embedded System For Fruit Image Processing งานวิจัยนี้ จะเกี่ยวกับการประมวลผลภาพ โดยการประยุกต์ใช้งาน ของระบบในการวิเคราะห์คุณภาพของเชอร์รี่ที่เก็บรักษา ไว้เป็นเวลา 3 สัปดาห์ เพื่อช่วยในการตรวจจับว่ามีส่วน ไหนของเชอร์รี่ที่เน่าเสีย หรือมีปัญหา จากผลการวิจัย พบว่าการใช้ Image processing เพื่อวิเคราะห์การ เปลี่ยนแปลงของเชอร์รี่ในระหว่างการเก็บรักษา สามารถ ใช้ในการลักษณะเนื้อสีของผลไม้และการตรวจจับ ข้อบกพร่องบนพื้นผิวได้อย่างประสบความสำเร็จ มีโอกาสในการนำมาใช้ในการควบคุมคุณภาพของผลไม้ ในช่วงการเก็บรักษาได้เป็นอย่างดี

3. วิธีการดำเนินงานวิจัย

3.1 ศึกษาและรวบรวมข้อมูลที่เกี่ยวข้อง

เป้าหมายของงานวิจัยครั้งนี้คือการให้ความรู้ และตอบ คำถามของผู้เลี้ยงกระต่าย ฉะนั้นผู้วิจัยจึงจำเป็นต้อง ค้นคว้าข้อมูล เบื้องต้น และเบื้องลึกของกระต่ายและผู้วิจัย ต้องทำการเก็บข้อมูลคำถามที่คิดว่าผู้ใช้งานจะถามเกี่ยวกับ กระต่าย และเก็บข้อมูลตัวอย่างคำถามเหล่านั้นไว้ เพื่อที่จะ นำไปฝึกการเรียนรู้ เพื่อให้ แชทบอท สามารถตอบคำถาม ของผู้ใช้งาน ได้อย่างถูกต้อง และมีข้อผิดพลาดน้อยที่สุด

3.2 กรอบแนวคิดของระบบ



ภาพที่ 1: กรอบแนวคิดของระบบ

จากภาพที่ 1 กระบวนการการประยุกต์ใช้เทคนิค การประมวลผลภาพ และความเข้าใจภาษาธรรมชาติ ร่วมกับ LINE แชทบอท ให้ความรู้ด้านการเลี้ยงกระต่ายนั้น จะเริ่มกระบวนการ โดยเริ่มจากผู้ใช้งานได้ส่งข้อความผ่าน ทางไลน์ โดยข้อความที่ผู้ใช้งานส่งไปนั้น จะมีเพียง 2 แบบ คือแบบรูปภาพหรือ Image และแบบตัวอักษร หรือ Text ซึ่งหลังจากส่งข้อความไปหาตัว แชทบอท แล้วก็จะทำการ นำประโยคที่ผู้ใช้งานพิมพ์ ส่งไปยัง Custom Webhook (Flask) เพื่อทำการตรวจสอบว่าผู้ใช้งานส่งข้อความ ประเภทรูปภาพ หรือตัวอักษร ถ้าหากข้อความเป็นประเภท ตัวอักษร ก็จะถูกส่งต่อไปยัง Dialogflow เพื่อทำการ ประมวลผลต่อว่า ประโยคที่ผู้ใช้งานพิมพ์มานั้น แชทบอท ควรจะตอบสนองอะไรกลับไป หรือถ้าหาก ผู้ใช้งานส่งข้อความประเภทรูปภาพนั้น ไฟล์รูปภาพก็ จะถูกอัปขึ้น ไปประมวลผลภาพใน Vertex AI โดยใช้ เทคนิค Classification เพื่อบอกสายพันธุ์ของกระต่ายใน รูปภาพที่ผู้ใช้งาน ได้ส่งข้อความมา เมื่อทำการประมวลผล เป็นอันเรียบร้อยแล้ว ก็จะส่งผลลัพธ์ไปยัง Dialogflow และส่งต่อมายัง LINE ของผู้ใช้งาน

3.3 การพัฒนาระบบ

3.3.1 การเตรียมข้อมูล

การที่จะทำระบบขึ้นมานั้น จำเป็นต้องมีการเตรียม ข้อมูลก่อน ข้อมูลที่ต้องเตรียมจะมีอยู่ 2 ส่วน คือ ข้อมูลที่

เป็นตัวอักษร สำหรับการเทรนแชทบอท และข้อมูลรูปภาพ สำหรับการทำ Image processing โดยเริ่มจากการจัดกลุ่มคำเพื่อนำไปสร้างเป็น Intent โดยจะยกตัวอย่างมาเพียงบางส่วนเท่านั้น ดังตารางที่ 1

ตารางที่ 1: ตารางแสดงตัวอย่างการจัดประเภทกลุ่มคำ

กลุ่มคำ	รูปประโยค
การรับเลี้ยง	อยากเลี้ยงกระต่าย, ต้องการรับเลี้ยงกระต่าย, เริ่มเลี้ยงกระต่าย, ดูแลกระต่าย
ปลดกคอก	ขี้ปลดกคอก, ไล่ปลดกคอก, สวมปลดกคอก, ไล่สร้อยคอก
อาหาร	กินอะไรได้, กินแครอทได้ไหม, กินส้มได้ไหม, กินผักอะไรบ้าง
โรคภัย	เปลือกตาแดง, ขาเป็นแผล, หูเป็นแผล, ร้องตลอดเวลา, จมูกเปียก

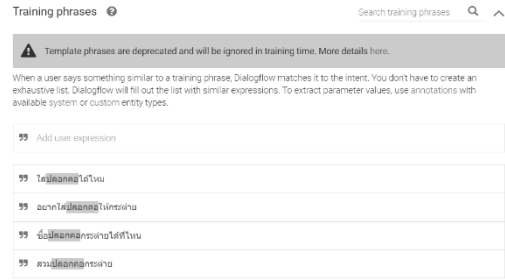
ในส่วนของ Image processing นั้น เนื่องจากข้อจำกัดด้านทุนของผู้วิจัย ทางผู้วิจัยจึงได้สร้างแบบสำรวจสายพันธุ์กระต่ายที่คนไทยส่วนใหญ่นิยมเลี้ยง ไปยังกลุ่มคนเลี้ยงกระต่ายหลายๆกลุ่ม จนเป็นผลสรุปได้ทั้งหมด 10 ลำดับสายพันธุ์กระต่ายที่คนไทยส่วนใหญ่นิยมเลี้ยง ดังตารางที่ 2

ตารางที่ 2: ตารางแสดงสายพันธุ์กระต่ายที่คนไทยนิยมเลี้ยง

ลำดับ	สายพันธุ์	ลำดับ	สายพันธุ์
1	ลอปส์	6	แองโกรา
2	เนเชอร์แลนด์ดวอป	7	เท็ดดี้แบร์
3	ไทย	8	เร็กซ์
4	คัตซ์	9	ควอร์ฟไฮโท
5	ไลออนเฮด	10	นิ ว ซี แล น ด์ ไวท์

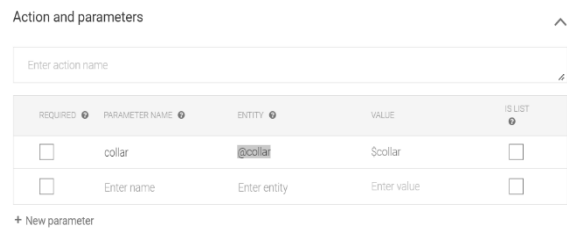
3.3.2 การฝึกแชทบอท

ในส่วนของการฝึกแชทบอท นั้น เราต้องไปที่เมนูคำสั่ง Intents และใส่ข้อมูลที่ต้องการสอน แชทบอท ในช่อง Training phrases ดังภาพที่ 2



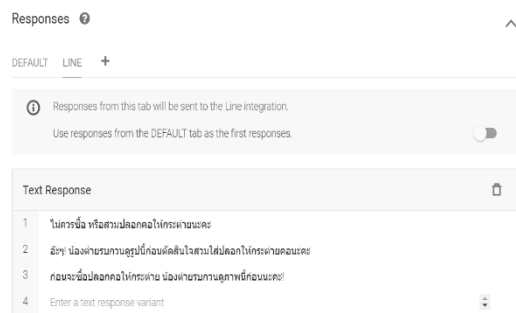
ภาพที่ 2: เมนู Training phrases ใน Dialogflow

ในขณะที่เทรนแชทบอทเราสามารถกำหนดคีย์เวิร์ดนั้น ๆ ได้ ถ้าเรากำลังเทรน แชทบอทเกี่ยวกับการไล่ปลดกคอก เราจะสร้างคีย์เวิร์ด @collar เข้ามาเพื่อกำหนดว่าคีย์เวิร์ดนี้คือปลดกคอก ถ้าผู้ใช้งานมีคำนี้อยู่ในประโยคแชทบอท ก็จะนำคำตอบที่อยู่ในหมวด @collar ที่เรากำหนดไว้มาตอบดังภาพที่ 3



ภาพที่ 3: เมนู Action parameters ใน Dialogflow

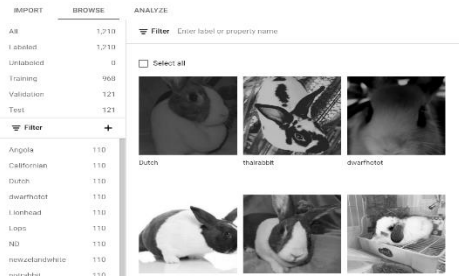
เมื่อทำการฝึกแชทบอทแล้ว ผู้วิจัยจำเป็นต้องเขียนการตอบรับของแชทบอท โดยจะเขียนการตอบรับของ Intent นั้น ๆ ยังไง ดังภาพที่ 4



ภาพที่ 4: เมนู Response ใน Dialogflow

3.3.3 การเทรนโมเดล Image processing

ในส่วนของการใช้ Image processing ผู้วิจัยได้ใช้เครื่องมือของ Google ที่ชื่อ Vertex AI ก่อนที่จะเทรนโมเดลภาพ ผู้วิจัยได้ทำการแบ่ง Label ของกระต่ายแต่ละพันธุ์แล้วอัปโหลดภาพขึ้นไปก่อนจะกดเทรนโมเดลดังภาพที่ 5

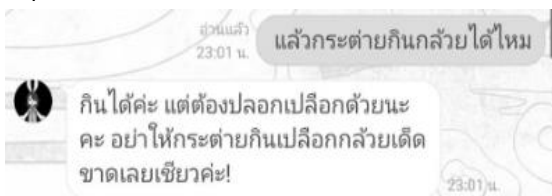


ภาพที่ 5: Vertex AI

และในส่วนสุดท้ายของการทำงานคือการ Coding เพื่อเชื่อม API ของ LINE, Dialogflow และ Google Vision API โดยตัว Framework ที่จะใช้เป็นตัวกลางคือ Flask ถ้าจะให้ระบบเริ่มทำงาน ก็สามารถพิมพ์คำสั่งเพื่อทำงานระบบได้ทันที

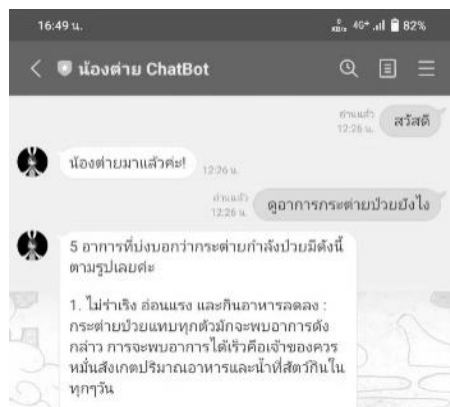
3.3.4 ผลการทดสอบ

เมื่อทำระบบเสร็จสิ้นแล้ว ก่อนเริ่มต้นการใช้งานนั้น ผู้ใช้งานจำเป็นต้อง Add LINE ของ แชนบอท ตัวนี้ก่อน จึงจะสามารถใช้งาน แชนบอท ได้ ซึ่งผู้ใช้งานสามารถถาม แชนบอท ในเรื่องที่ไม่รู้ได้ทันทีได้เลย เช่น อยากรู้ว่า กระต่ายกินผัก หรือผลไม้อะไรบ้าง หรือเราสามารถระบุถามเฉพาะชนิดผักนั้นๆก็ได้เช่นกัน ดังภาพที่ 6



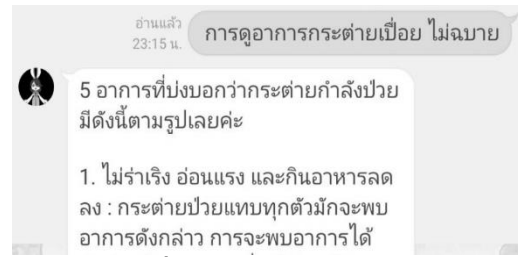
ภาพที่ 6: ตัวอย่างการถามด้านอาหารกระต่าย

นอกจากนี้ยังมีการสอบถามด้านอื่นๆอีก เช่น ถามเรื่องการใส่ปลอกคอก ถามเรื่องการที่กระต่ายไม่ยอมกินหญ้า กินแต่อาหารเม็ด หรือการดูอาการกระต่ายป่วย เป็นต้น ดังภาพที่ 7



ภาพที่ 7: ตัวอย่างการถามทั่วไป

นอกจากนี้หากผู้ใช้งานมีการเขียนผิด หรือใช้ศัพท์แสดง แชนบอทก็ยังสามารถตอบคำถามได้อีกด้วย ดังเช่นภาพที่ 8



ภาพที่ 8: ตัวอย่างการถามที่ใช้ศัพท์แสดง

ในส่วนของการทำนายสายพันธุ์กระต่าย ผู้ใช้งานกดปุ่ม “ตรวจสอบสายพันธุ์กระต่ายของคุณ” ใน Rich Menu ด้านล่าง แชนบอทจะบอกให้ส่งรูปมากระต่ายมา ไฟล์รูปกระต่ายที่ส่งไม่จำกัดขนาด หลังจากที่ได้รูปแล้ว แชนบอทจะตอบว่ากระต่ายตัวนี้ เป็นกระต่ายสายพันธุ์อะไร ดังภาพที่ 9



ภาพที่ 9: การทำนายสายพันธุ์กระต่าย

4. สรุปผลการประเมิน

การประเมินระบบจะแบ่งเป็นการประเมินความถูกต้องของระบบในด้านของการตอบคำถาม แชนบอทกับ Image processing โดยผู้เชี่ยวชาญทั้ง 3 คนได้แก่ สัตวแพทย์เฉพาะทางด้านกระต่าย 1 ท่าน และผู้มีประสบการณ์ด้านการเลี้ยงกระต่ายมากกว่า 5 ปี 2 ท่าน ซึ่งผลการประเมินมีดังนี้

ความถูกต้องของการตอบคำถามในส่วนขอ แชนบอท จากผู้เชี่ยวชาญทั้งหมด 3 คน คนละ 10 ข้อคำถาม ดังตารางที่ 3

ตารางที่ 3: ตารางแสดงความถูกต้องของการตอบคำถามในส่วน
ของแชทบอท

ผู้ทดสอบ	จำนวนทั้งหมด	จำนวนที่ถูกต้อง	จำนวนที่คลาดเคลื่อน	ร้อยละความถูกต้อง
1	10	9	1	90.00 %
2	10	9	1	90.00 %
3	10	8	2	80.00 %
ค่าเฉลี่ยรวมทั้งหมด				86.67 %

จากตารางที่ 3 ผลการทดสอบความถูกต้องของแชทบอท โดยผู้ทดสอบคนที่ 1 มีค่าความถูกต้องอยู่ที่ 90.00 % ผู้ทดสอบคนที่ 2 มีค่าความถูกต้องอยู่ที่ 90.00 % ผู้ทดสอบคนที่ 3 มีค่าความถูกต้องอยู่ที่ 80.00 % โดยค่าเฉลี่ยรวมความถูกต้องของการตอบคำถามของแชทบอทอยู่ที่ 86.67 %

ความถูกต้องของการตอบคำถามในส่วนของ Image processing จากผู้เชี่ยวชาญทั้งหมด 3 คน คนละ 10 ข้อคำถาม ดังตารางที่ 4

ตารางที่ 4: ตารางแสดงความถูกต้องของการตอบคำถามในส่วน
ของ Image Processing

ผู้ทดสอบ	จำนวนทั้งหมด	จำนวนที่ถูกต้อง	จำนวนที่คลาดเคลื่อน	ร้อยละความแม่นยำ
1	10	7	3	70.00 %
2	10	7	3	70.00 %
3	10	9	1	90.00 %
ค่าเฉลี่ยรวมทั้งหมด				76.67 %

จากตารางที่ 4 ผลการทดสอบความถูกต้อง Image Processing โดยผู้ทดสอบคนที่ 1 มีค่าความถูกต้องอยู่ที่ 70.00 % ผู้ทดสอบคนที่ 2 มีค่าความถูกต้องอยู่ที่ 70.00 % ผู้ทดสอบคนที่ 3 มีค่าความถูกต้องอยู่ที่ 90.00 % โดยค่าเฉลี่ยรวมความถูกต้องของการตอบคำถามของ Image Processing อยู่ที่ 76.67 %

5. ข้อเสนอแนะ

การประยุกต์ใช้เทคนิคการประมวลผลภาพ และความเข้าใจภาษาธรรมชาติร่วมกับ ไลน์ แชทบอท ให้ความรู้ด้านการเลี้ยงกระต่ายนั้น ในส่วนของ Image processing ยังมีค่าความถูกต้องค่อนข้างต่ำ เนื่องจากกระต่ายบนโลกมีทั้งหมด 51 สายพันธุ์ แต่ในงานวิจัยนี้ นำมาเพียง 10 สายพันธุ์เท่านั้น ส่งผลให้ตอนผู้เชี่ยวชาญทดสอบระบบ

ไม่สามารถระบุสายพันธุ์กระต่ายที่นอกเหนือจากการ Train ไปได้อย่างถูกต้องอย่างสมบูรณ์ และไม่สามารถระบุกระต่ายพันธุ์ผสมได้อย่างแม่นยำ เนื่องจากความกำกวมของลักษณะของพันธุ์กระต่ายแต่ละพันธุ์ ในส่วนของแชทบอทอาจจะยังไม่ครอบคลุมคำถามสนทนาทั้งหมด หรือคำตอบบางคำเนื่องจากแชทบอทตัวนี้ถูกสอนมาเพื่อตอบคำถามเฉพาะทางที่เกี่ยวกับกระต่ายเท่านั้น โดยมี Intent ทั้งหมดเพียง 62 Intent จึงอาจทำให้เกิดความไม่ลื่นไหลของการสนทนา

เอกสารอ้างอิง

- [1] อธิป จิตตฤกษ์. (2565). "ธุรกิจสัตว์เลี้ยง' เทรนด์อุตสาหกรรมใหญ่มาแรง โอกาสทำเงินที่เหมาะสมแก่การลงทุน". (ออนไลน์). จาก <https://www.gqthailand.com/gq-hype/article/gq-hype-vol-107>
- [2] Chonticha. (2565). "Exotic Pet สัตว์เลี้ยงสุดแปลก แตกต่างไม่เหมือนใคร แต่น่ารักสุด ๆ". (ออนไลน์). จาก <https://www.wongnai.com/articles/exotic-pet>
- [3] Iconext. (2565). "Chatbot คืออะไร? ประโยชน์และตัวอย่างการนำไปใช้ในธุรกิจ". (ออนไลน์). จาก <https://iconext.co.th/th/2022/01/27/chatbot-คืออะไร-ประโยชน์และตัว/>
- [4] Mangozero. (2562). "รู้จัก "Thai NLP" ตัวช่วยสุดเจ๋งที่ทำให้คอมพิวเตอร์เข้าใจภาษาไทยได้มากขึ้น!". (ออนไลน์). จาก <https://www.mangozero.com/what-is-thai-nlp/>
- [5] TechTarget Contributor. (2021). "Natural Language Understanding (NLU) ". (ออนไลน์). จาก <https://www.techtarget.com/searchenterpriseai/definition/natural-language-understanding-NLU>
- [6] Cem Dilmegani. (2022). "What is Data Augmentation". (ออนไลน์). จาก <https://research.aimultiple.com/data-augmentation/>
- [7] Daily Tech. (2563). "Image Processing คือ ความฉลาดของเทคโนโลยีที่ทำให้ชีวิตเราง่ายขึ้น".(ออนไลน์). จาก <https://www.dailytech.in.th/image-processing/>
- [8] เกียรติพันธ์ศักดิ์ บิลอับดุลลาห์. (2564). "การประยุกต์ใช้แชทบอท สนับสนุนงานสอบสวน กรณีศึกษาสถานีตำรวจภูธรเบตง". (สารนิพนธ์ ปริญญาวิทยาศาสตรมหาบัณฑิต). มหาวิทยาลัยสงขลานครินทร์
- [9] I-Ching Hsu and Jiun-De Yu. (2022). "A medical chatbot using machine learning and natural language understanding". (ออนไลน์). สืบค้นจาก <https://link.springer.com/article/10.1007/s11042-022-12820-4>

- [10] Sotirov S.I., Zhelyazkov S.P., “Marudova M.G., Zsivanovits G., and Tokamkov D.M. Embedded System for Fruit Image Processing”. International Scientific Conference Electronics. September 16-18, pp. 1-3.

การพัฒนาเสื้อชูชีพเฝ้าระวังผู้ประสบอุบัติเหตุทางน้ำด้วยอินเทอร์เน็ตของสรรพสิ่ง

Development of Life Jacket to Monitor Water Accident Victim with the Internet of Things

สมโชค กิตติสิทธิพงษ์ (Somchoke Kittisitthipong)¹ และมหศักดิ์ เกตุฉ่ำ (Mahasak Ketcham)²

^{1,2}ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856038@email.kmutnb.ac.th, ²mahasak.k@itd.kmutnb.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อศึกษา สร้าง และทดลองประยุกต์ใช้อินเทอร์เน็ตของสรรพสิ่ง (IoT) เพื่อใช้ในการแจ้งเตือนการตกน้ำของผู้ประสบเหตุ พร้อมมีการแสดงผลพิกัดตำแหน่งบนแผนที่แบบ Real-Time ผ่าน Protocol MQTT และ Websocket จากการทดสอบพบว่าระบบสามารถแจ้งเตือนบน Application Line เมื่อ Sensor สัมผัสน้ำภายในระยะเวลาประมาณ 1 วินาที ซึ่งเป็นเวลาไม่เกิน 4 นาทีหลังจากร่างกายเริ่มขาดออกซิเจน และเป็นเวลาดำกว่า 6 ชั่วโมงซึ่งเป็นเวลาดำที่เสื้อชูชีพสามารถลอยน้ำได้ สามารถแสดงผลพิกัดปัจจุบันของเสื้อชูชีพได้แบบ Real-Time สามารถส่งข้อมูลไปยัง MQTT Broker ได้ภายในระยะเวลาเฉลี่ย 371 มิลลิวินาที ทั้งนี้จากการทดสอบในสถานที่จริงผู้วิจัยสามารถวัดค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งสำหรับบริเวณลำคลองซึ่งเป็นพื้นที่ที่ไม่มีวัตถุกีดขวางในระยะ 10 เมตร อยู่ที่ 1.50 เมตร ค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งบริเวณสระว่ายน้ำแห่งที่ 1 เป็นพื้นที่เปิด ซึ่งในระยะ 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน กำแพงและต้นไม้สูง 4 เมตร 3 ด้าน อยู่ที่ 3.60 เมตร และค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งบริเวณสระว่ายน้ำแห่งที่ 2 ซึ่งเป็นพื้นที่เปิดที่มีรัศมี 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน อยู่ที่ 1.31 เมตร ซึ่งทั้งหมดยังเป็นระยะสายตาสีฟ้าที่สามารถมองเห็นได้ ทำให้ผู้เฝ้าระวังสามารถหยิบยื่นความช่วยเหลือได้ทันเวลา

คำสำคัญ: อินเทอร์เน็ตของสรรพสิ่ง ระบบการหาตำแหน่งทั่วโลก เสื้อชูชีพ

Abstract

This article aims to study, develop, and experiment with the Internet of Things (IoT) to notify individuals of a water accident, including a real-time display of coordinates on a map through MQTT protocol and Websocket. The test results show that the system can send a notification on the Line application within approximately 1 second of the sensor coming into contact with water, which is less than 4 minutes after the body starts to lose oxygen and less than 6 hours, which is the minimum time a life jacket can float on water. The system can display the current coordinates of the life jacket in real-time and send data to the MQTT Broker in an average of 371 milliseconds. In actual testing, the researchers measured the average deviation of the location for the canal area, which is an unobstructed area within a 10-meter radius, at 1.50 meters. The average deviation of the location for the first swimming pool, an open area with a 10 meter radius, where a 2-storey building is located on one side and a wall and trees are 4 meters high on three sides, was 3.60 meters. The average deviation of the location for the second swimming pool, which is also an open area with a 10-meter radius, where a 2-storey building is located on one side, was 1.31 meters, all of which are within the visible range, allowing rescuers to reach the site in time to help.

Keywords: Internet of Things, Global Positioning System, Life Jacket

1. บทนำ

เสื้อชูชีพแม้เป็นสิ่งที่สามารถช่วยชีวิตผู้คนที่ได้ แต่ก็ยังมีข้อจำกัดในเรื่องของระยะเวลาการพองตัว โดยข้อมูลของสำนักงานกองทุนสนับสนุนการส่งเสริมสุขภาพ ระบุว่า สำหรับเสื้อชูชีพทั่วไปจะลอยตัวได้ประมาณ 6 ชั่วโมง [1] และตามระเบียบกรมเจ้าท่าว่าด้วยหลักเกณฑ์การรับรองคุณภาพมาตรฐานอุปกรณ์ประจำเรือ (เสื้อชูชีพ) ให้แก่ภาคเอกชนหรือภาคส่วนอื่น พ.ศ. 2554 [2] กำหนดมาตรฐานระยะเวลาการลอยตัวไว้ไม่ต่ำกว่า 24 ชั่วโมง ในน้ำจืด อย่างไรก็ตามนอกจากข้อจำกัดเรื่องเวลาดังกล่าวแล้ว ผู้ประสบเหตุตกน้ำยังต้องพบปัจจัยด้านอุณหภูมิของน้ำ โดยจากการศึกษาพบว่า หากร่างกายสัมผัสน้ำที่อุณหภูมิต่ำกว่า 11.6 องศาเซลเซียสแบบทั้งตัวแม้ใส่เสื้อชูชีพ จะทำให้ร่างกายเข้าสู่สภาวะ Cold Shock และจะเสียชีวิตภายในเวลาประมาณ 30 นาที [3] ในส่วนของประเทศไทย อุณหภูมิของน้ำที่เย็นจัดอาจไม่ใช่ปัญหาแต่หากมองในแง่ของความเป็นจริงการตกน้ำโดยไม่ได้ตั้งใจ อาจทำให้ผู้ประสบเหตุหมดสติได้เช่นกัน ส่งผลให้ช่วยเหลือตัวเองไม่ได้ ทำให้ร่างกายอยู่ในท่าที่ไม่สามารถทำให้งอกและปากพื้นผิวน้ำได้ทำให้ขาดออกซิเจน เพราะฉะนั้นผู้ประสบเหตุจึงต้องได้รับการช่วยเหลือโดยเร็วที่สุดภายในระยะเวลา 4 นาที นับจากที่ร่างกายเริ่มขาดออกซิเจน [4]

จากข้อความข้างต้น จะเห็นได้ว่าเวลาเป็นปัจจัยสำคัญที่ส่งผลถึงโอกาสรอดชีวิตของผู้ประสบเหตุ ผู้วิจัยจึงมีแนวคิดที่จะนำเทคโนโลยี IoT เข้ามาใช้เป็นส่วนหนึ่งของงานวิจัยเพื่อจุดประสงค์ในการอำนวยความสะดวก หรือแม้กระทั่งการเฝ้าระวังและแจ้งเตือนภัยต่าง ๆ รอบตัว โดยสิ่งที่ผู้วิจัยมุ่งเน้น คือ การนำ GPS เข้ามาใช้ในงาน IoT เพื่อแก้ไขปัญหาในด้านต่าง ๆ ผู้วิจัยจึงมีแนวคิดในการนำ IoT เข้ามามีส่วนช่วยในการป้องกัน แจ้งเตือน และแจ้งพิกัดกรณีเกิดเหตุการณ์จมน้ำ เพื่ออำนวยความสะดวกในการติดตามตำแหน่งของผู้ประสบเหตุให้ผู้ช่วยเหลือสามารถหิบบ้นความช่วยเหลือได้ทันเวลา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบกำหนดตำแหน่งบนโลก (GPS)

GPS (Global Positioning System) คือ ระบบบอกตำแหน่งบนผิวโลกโดยอาศัยการคำนวณจากสัญญาณที่ส่งมาจาก

ดาวเทียมนำทาง คำนวณหาตำแหน่งจากจุดที่ส่งค่าพิกัด ซึ่งจะบอกเป็นค่าละติจูดกับลองจิจูด

หลักการพื้นฐานของการทำงาน GPS Tracking [5] จะใช้ตำแหน่งของดาวเทียมในอวกาศเป็นจุดอ้างอิงวัฏระยะจากดาวเทียมและใช้หลักการทางเรขาคณิตในการคำนวณหาตำแหน่งบนพื้นโลกระหว่าง ดาวเทียมกับเครื่องรับ GPS เมื่อเครื่องรับ GPS สามารถรับสัญญาณจากดาวเทียมได้ 3 ดวงขึ้นไปจะมีการคำนวณระยะทางระหว่างดาวเทียมถึงเครื่องรับ GPS โดยสูตรคำนวณทางทริโกโนเมตรี ระยะทางเท่ากับความเร็วคูณด้วยเวลา โดยดาวเทียมทั้ง 3 ดวงจะส่งสัญญาณคลื่นวิทยุที่เคลื่อนที่ด้วยความเร็วคงที่มายังเครื่องรับ GPS ที่ 186,000 ไมล์ต่อวินาที

2.2 เสื้อชูชีพ

เสื้อชูชีพทำงานบนหลักการลอยตัว [6] กล่าวคือ แรงลอยตัวที่กระทำต่อวัตถุมีค่าเท่ากับน้ำหนักของของเหลวที่ถูกแทนที่ เมื่อบุคคลอยู่ในน้ำแรงลอยตัวที่สร้างขึ้นจะเท่ากับน้ำหนักของบุคคล เสื้อชูชีพทำมาจากรูปทรงที่มีน้ำหนักเบาซึ่งมีน้ำหนักน้อยกว่าความหนาแน่นเฉลี่ยของมนุษย์เป็นผลให้เมื่อมีลมสวมเสื้อชูชีพ น้ำหนักรวมที่จะแทนที่ด้วยน้ำจะน้อยกว่ามากเมื่อเทียบกับน้ำหนักของบุคคลเพียงคนเดียวและบุคคลนั้นก็จะสามารถลอยได้ นอกจากนี้ร่างกายเมื่ออยู่ในน้ำจะสร้างแรงลอยตัวที่ดันขึ้นไปด้านบน ดังนั้นเสื้อชูชีพจึงไม่จำเป็นต้องรองรับน้ำหนักของร่างกายทั้งหมด สิ่งสำคัญคือต้องสวมเสื้อชูชีพที่พอดีตัว ปกป้องบุคคลจากการบาดเจ็บเนื่องจากการกระแทก และรับรองความปลอดภัยสูงสุดจนกว่าความช่วยเหลือจะมาถึง

2.3 เทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (IoT)

เป็นโครงสร้างพื้นฐานเครือข่ายทั่วโลกแบบไดนามิกพร้อมความสามารถในการกำหนดค่าด้วยตนเองตามโปรโตคอลการสื่อสารมาตรฐานและระหว่างกัน IoT เป็นเครือข่ายระหว่างอุปกรณ์ทางกายภาพ ยานพาหนะ อาคาร และอื่น ๆ ที่ฝังอยู่ในอุปกรณ์อิเล็กทรอนิกส์ Software Sensor Actuator และการเชื่อมต่อเครือข่ายที่ทำให้วัตถุเหล่านี้สามารถรวบรวมและแลกเปลี่ยนข้อมูลได้ โดยสามารถสั่งการ การทำงานของโปรแกรมหรือสั่งการควบคุมการใช้งานอุปกรณ์ผ่านทางเครือข่ายอินเทอร์เน็ต องค์ประกอบของสถาปัตยกรรมเทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (IoT) โดยอ้างอิงตามแบบเรียนในรายวิชา

อินเทอร์เน็ตของสรรพสิ่งสำหรับธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จะประกอบไปด้วย 4 ส่วน [7] คือ

2.3.1 Sensor Layer

ชั้นนี้เป็นส่วนสำคัญของระบบ IoT และมีารเชื่อมต่อเครือข่ายไปยังขั้นถัดไปซึ่งเป็นเกตเวย์ (Gateway) และชั้นเครือข่าย วัตถุประสงค์หลักของเซ็นเซอร์คือการรวบรวมข้อมูลจากสภาพแวดล้อมโดยรอบอุปกรณ์ตรวจจับ (Sensors) หรือสิ่งของ (Things) ของระบบฯ ซึ่งจะรอการเชื่อมต่อข้อมูลทางตรงหรือทางอ้อมกับเครือข่าย IoT หลังจากการแปลงและการประมวลผลสัญญาณจากที่ได้รับจากเซ็นเซอร์หรือสิ่งของแล้ว โดยในชั้นนี้ยังรวมถึงอุปกรณ์ที่แสดงผล (Actuators) ด้วย

2.3.2 Gateway and Network Layer

ทำหน้าที่เชื่อมต่อไปยังระบบเครือข่าย (Network) ทำให้อุปกรณ์ทั้งหมดในโครงข่าย IoT ทั้งหมดส่งข้อมูลเข้าสู่อินเทอร์เน็ตได้ ทั้งนี้ โดยทั่วไปค่าเริ่มต้นของ Gateway มักตั้งค่าให้อยู่ภายใต้ระบบเครือข่ายภายใน (Local network) ซึ่งจะสามารถการตั้งค่าเพื่อเชื่อมต่อไปยัง Internet ได้

2.3.3 Management Service Layer

ชั้นนี้ทำหน้าที่เป็นส่วนต่อประสานระหว่าง Gateway and Network Layer และ Application Layer [8] แบบสองทิศทาง กล่าวคือมีหน้าที่จัดการอุปกรณ์ จัดการข้อมูล รวมไปถึงรวบรวมข้อมูลจำนวนมากและดึงข้อมูลที่เกี่ยวข้องจากข้อมูลที่เก็บไว้ มีการวิเคราะห์ข้อมูล และควบคุมความปลอดภัย

2.3.4 Application Layer

ส่วนต่อประสานกับผู้ใช้ระบบ IoT แอปพลิเคชันนี้สามารถพัฒนาให้เข้ากับอุปกรณ์ได้หลายรูปแบบ เช่น บนอุปกรณ์มือถือ หรือ แอปพลิเคชันบน Web Browser ที่สามารถรองรับระบบปฏิบัติการได้ทุกรูปที่สามารถติดตั้ง Web Browser ได้

3. วิธีดำเนินการวิจัย

3.1 การออกแบบระบบ

ผู้วิจัยอ้างอิงตามองค์ประกอบของสถาปัตยกรรมเทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง (Internet of Things)

ซึ่งประกอบไปด้วย Sensor Layer, Gateway and Network Layer, Management Service Layer และ Application Layer

3.1.1 Sensor Layer

ผู้วิจัยเลือกใช้เสื้อชูชีพแบบโฟมอยู่ในตัวเสื้อ ซึ่งเป็นเสื้อชูชีพที่ใช้งานตามมาตรฐานทั่วไปและนิยมใช้กันมากที่สุด โดยอุปกรณ์ที่ติดตั้งในเสื้อชูชีพ มีดังนี้

1) Microcontroller ได้แก่ ESP32

2) Sensor ได้แก่ GPS Sensor Module, Rain Water Detection Sensor และ Push Switch

3) Actuator ได้แก่ สีขาว สีเหลือง สีน้ำเงิน สีเขียว และ สีแดง

3.1.2 Gateway and Network Layer

จากการหาข้อมูล พบว่าเครือข่ายแบบ LoRa และ Zigbee จำเป็นต้องใช้ Gateway เพิ่มเพื่อการออกสู่ Internet จึงไม่เหมาะสำหรับการติดตั้งบนเสื้อชูชีพ เนื่องจากจะทำให้มีอุปกรณ์จำนวนมาก หากเป็นเทคโนโลยี NB-IoT จะทำให้ไม่คล่องตัวในการเปลี่ยนสัญญาณเครือข่าย และ 3G/4G module จำเป็นต้องอาศัย MCU ในการส่งข้อมูล ทำให้หากต้องการเปลี่ยน SIM Card จำเป็นต้องถอดชิ้นส่วนบริเวณใกล้กับ MCU จึงอาจลดประสิทธิภาพในการกันน้ำ และความชื้น ดังนั้น ผู้วิจัยจึงได้เลือกใช้ Pocket Wi-Fi ที่สามารถเปลี่ยน SIM Card ได้ เพื่อเพิ่มความสะดวกในการเลือกใช้ผู้ให้บริการสัญญาณเครือข่ายที่ดีที่สุดตามแต่ละพื้นที่ อีกทั้งอุปกรณ์ดังกล่าวยังมีขนาดเล็ก จึงติดตั้งภายในช่องกันน้ำแยกจาก MCU ได้อีกด้วย

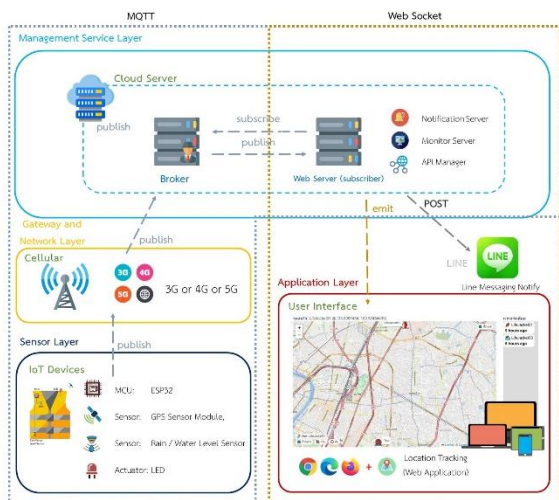
3.2.3 Management Service Layer

ใช้ภาษา JavaScript ที่ติดตั้งบน Cloud Server มี Public IP สามารถเชื่อมต่อได้ผ่านเครือข่ายอินเทอร์เน็ต และ Line Messaging API ที่จะบริหารจัดการข้อความแจ้งเตือนผ่าน Application Line

3.1.4 Application Layer

ผู้วิจัยเลือกพัฒนาส่วนติดต่อกับผู้ใช้งานด้วย Web Application โดยใช้ภาษา JavaScript เนื่องจาก Web Application สามารถเข้าถึงผู้ใช้งานได้เป็นส่วนใหญ่ ซึ่งผู้ใช้งานเกือบทุกระบบปฏิบัติการที่ติดตั้ง Web Browser จะสามารถเข้าถึงได้

3.2 การพัฒนาระบบ

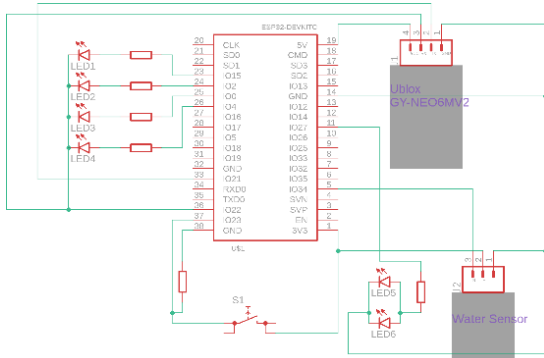


ภาพที่ 1: แสดงกรอบแนวคิดของงาน

จากการศึกษา และค้นคว้าข้อมูลเกี่ยวกับการพัฒนา IoT ผู้วิจัย ได้ออกแบบระบบฯ โดยใช้ MQTT Protocol สำหรับส่งข้อมูลจาก IoT Device ขึ้นสู่ MQTT Broker บนระบบ Cloud ผ่านระบบเครือข่าย 4G และใช้ Web Socket ในการส่งข้อมูลที่รับจาก IoT Device ไปแสดงผลยัง Client ด้วย Web Browser

3.2.1 Sensor Layer

กำหนดให้ ESP 32 ตรวจจับค่าจาก Rain / Water Level Sensor เพื่อตรวจสอบว่าอุปกรณ์มีการสัมผัสน้ำหรือไม่ ในขณะเดียวกันจะมีการหาค่าพิกัด GPS ควบคู่ไปด้วย หากมีการตรวจพบการสัมผัสน้ำ ESP32 จะนำส่งข้อมูล GPS ออกสู่อินเทอร์เน็ตไปยัง Server ที่ได้ตั้งค่าเอาไว้ และสั่งให้ LED เกิดแสงกะพริบ สำหรับการรับข้อมูลจาก Sensor จะทำงาน โดย Multicore และ Multithread เพื่อให้สามารถทำงานแบบขนาน (Parallel) กันได้ สำหรับการเชื่อมต่อวงจรจะเป็นดังภาพที่ 2



ภาพที่ 2: แสดง Schematic ของการเชื่อมต่อวงจร



ภาพที่ 3: แสดงวงจจริง



ภาพที่ 4: แสดงเสื้อชูชีพ

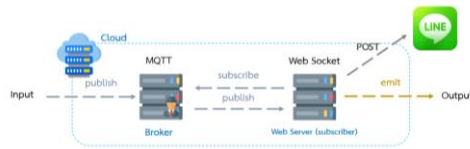
3.2.2 Gateway and Network Layer

จากการค้นหาข้อมูลพบว่า การใช้พลังงานของสัญญาณ 3G 4G และ 5G นั้น หากเลือกใช้สัญญาณใน Generation ที่ใหม่กว่า จะยิ่งส่งผลให้การใช้พลังงานในการส่งข้อมูลต่อ 1 bit น้อยกว่าแต่ขณะเดียวกันก็สามารถส่งจำนวน bit ข้อมูลได้มากขึ้นหากเทียบในหน่วยของเวลา ฉะนั้น หากพิจารณาถึงการสูญเสียพลังงานไปในหนึ่งหน่วยเวลานั้น สัญญาณใน Generation ใหม่ จะมีการใช้พลังงานที่มากกว่า ประกอบกับ ผู้วิจัยพบว่า Pocket Wi-Fi แต่ละรุ่นมีชั่วโมงการใช้งานที่แตกต่างกันขึ้นอยู่กับปริมาณแบตเตอรี่ที่ติดตั้งมากับเครื่อง ผู้วิจัยจึงเลือกใช้ Pocket Wi-Fi แบบ 4G ที่สามารถส่งข้อมูลได้เร็วกว่า 3G แต่ ใช้พลังงานต่อ 1 หน่วยเวลาน้อยกว่า 5G และหากดูปริมาณการครอบคลุมของสัญญาณแล้ว 4G ครอบคลุมพื้นที่ทางทะเลมากกว่า 5G อีกด้วย

3.2.3 Management Service Layer

3.2.3.1 MQTT Broker จะรอรับข้อมูลจาก ESP32 และส่งข้อมูลต่อไปยัง Subscriber ซึ่งสำหรับการวิจัยครั้งนี้คือ Web Server

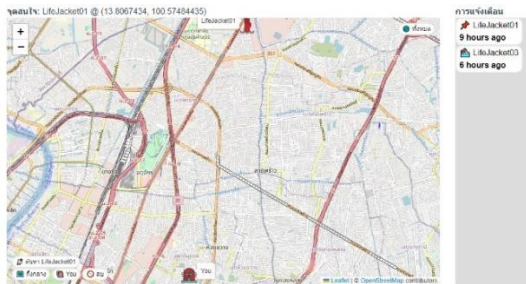
3.2.3.2 Web Server จะทำการ Subscribe ไปยัง Broker เพื่อจะทำการส่งข้อมูลพิกัดจุดบนแผนที่ไปยังส่วนของ Web Application ด้วย Web Socket ผ่าน Framework ที่ชื่อว่า Socket.io โดยจะมีการเขียนเงื่อนไขในการแจ้งเตือนไปยัง Application Line ผ่าน LINE Messaging API อีกด้วย



ภาพที่ 5: แสดงการทำงานของ Cloud Server

3.2.4 Application Layer

ผู้วิจัยได้นำ Vue.js ที่เป็นหนึ่งใน Web Framework ของ JavaScript ในการพัฒนา Web Application โดยจะมีหน้าที่แสดงแผนที่ระบุจุดเกิดเหตุเป็นหลัก โดยหน้าจอก็จะมีเพียง 1 หน้าจอหลัก แสดงผลแบบ Responsive Web Design กล่าวคือสามารถแสดงผลได้อย่างสมบูรณ์บนอุปกรณ์ แสดงผลได้หลากหลายความละเอียดโดยไม่ต้องเข้า URL อื่น



ภาพที่ 6: แสดง Web Application

4. ผลการดำเนินงาน

4.1 ทดสอบการทำงานอย่างถูกต้องของระบบ

ผลการทดสอบการใช้งานทั่วไปของเสื้อชูชีพได้แก่ ทดสอบสัญญาณไฟกะพริบ ทดสอบการแสดงผลตำแหน่งบนแผนที่ ทดสอบการแจ้งเตือน จำนวน 5 ครั้ง ได้ผลดังตารางที่ 1

ตารางที่ 1: แสดงผลการทดสอบการใช้ทั่วไป

ครั้งที่	สัญญาณไฟ	แสดงตำแหน่ง	การแจ้งเตือน
1	✓	✓	✓ (~ 1s)
2	✓	✓	✓ (~ 1s)
3	✓	✓	✓ (~ 1s)
4	✓	✓	✓ (~ 1s)
5	✓	✓	✓ (~ 1s)

4.2 ทดสอบค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่ง

ผู้วิจัยจะวัดระยะห่างระหว่างตำแหน่งจริงและตำแหน่งที่ได้จาก GPS Module ด้วยสูตร

$$a = \sin^2(\Delta\phi/2) + \cos \phi_1 \cdot \cos \phi_2 \cdot \sin^2(\Delta\phi/2)$$

$$c = 2 \cdot \text{atan2}(\sqrt{a}, \sqrt{1-a})$$

$$d = R \cdot c \tag{4-1}$$

โดยทำการบันทึกตำแหน่งจากสถานที่จริง ได้แก่

4.2.1 ลำคลอง เป็นพื้นที่เปิด ไม่มีสิ่งกีดขวางในรัศมี 10 เมตร

4.2.2 สระว่ายน้ำแห่งที่ 1 เป็นพื้นที่เปิด ซึ่งในระยะ 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน กำแพงและต้นไม้สูง 4 เมตร 3 ด้าน

4.2.3 สระว่ายน้ำแห่งที่ 2 เป็นพื้นที่เปิด ซึ่งในระยะ 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน

เก็บค่าทั้งหมด 10 ครั้ง ภายใน 1 นาที หลังจากนั้นหาค่าเฉลี่ย ได้ผลดังตารางที่ 2

ตารางที่ 2: แสดงผลการทดสอบความคลาดเคลื่อนของตำแหน่ง

ครั้งที่	คลาดเคลื่อน (คลอง)	คลาดเคลื่อน (สระน้ำ1)	คลาดเคลื่อน (สระน้ำ2)
1	1.07 เมตร	2.94 เมตร	0.76 เมตร
2	1.08 เมตร	2.95 เมตร	1.13 เมตร
3	1.22 เมตร	3.15 เมตร	1.55 เมตร
4	1.37 เมตร	3.42 เมตร	1.44 เมตร
5	1.49 เมตร	3.59 เมตร	1.37 เมตร
6	1.67 เมตร	3.74 เมตร	1.28 เมตร
7	1.73 เมตร	3.81 เมตร	1.37 เมตร
8	1.76 เมตร	3.79 เมตร	1.33 เมตร
9	1.78 เมตร	4.25 เมตร	1.12 เมตร
10	1.85 เมตร	4.27 เมตร	1.77 เมตร
เฉลี่ย	1.50 เมตร	3.60 เมตร	1.31 เมตร

4.3 ทดสอบความคลาดเคลื่อนของเวลาในการส่งข้อมูล

ผู้วิจัยจะทำการบันทึกเวลาเมื่ออุปกรณ์ IoT ได้รับค่าจาก GPS จนถึงเวลาที่ Broker ได้รับข้อมูล โดยใช้สูตร

$$\Delta t = t_1 - t_2 \tag{4-2}$$

เก็บค่าทั้งหมด 10 ครั้ง ภายใน 1 นาที หลังจากนั้นหาค่าเฉลี่ย

ตารางที่ 3: แสดงผลความคลาดเคลื่อนของเวลาในการส่งข้อมูล

ครั้งที่	คลาดเคลื่อน
1	372 มิลลิวินาที
2	357 มิลลิวินาที
3	360 มิลลิวินาที
4	365 มิลลิวินาที
5	368 มิลลิวินาที
6	363 มิลลิวินาที
7	376 มิลลิวินาที
8	384 มิลลิวินาที
9	381 มิลลิวินาที
10	384 มิลลิวินาที
เฉลี่ย	371 มิลลิวินาที

5. สรุป

งานวิจัยครั้งนี้ผู้วิจัยได้ทำการพัฒนาเสื้อชูชีพที่ผสานรวมกับเทคโนโลยี Internet of Things พร้อมทำการวัดผลประสิทธิภาพการทำงานของเสื้อชูชีพ จากการลงน้ำจริง ซึ่งระบบมีการแสดงผลที่กดบนแผนที่แบบ Real-Time มีค่าใช้จ่ายโดยประมาณ 2,909 บาท (ไม่รวมค่า Cloud Server และค่า Sim Card ที่ต้องจ่ายแบบรายเดือน)

ผลการทดสอบพบว่าระบบมีการแสดงผลข้อมูลโดยมีเวลาคลาดเคลื่อนเฉลี่ย 371 มิลลิวินาที ในแต่ละชุดข้อมูล และสามารถแจ้งเตือนผ่าน Application Line โดยมีเวลาคลาดเคลื่อนประมาณ 1 วินาที ซึ่งเป็นเวลาไม่เกิน 4 นาทีหลังจากร่างกายเริ่มขาดออกซิเจน และเป็นเวลาดำกว่า 6 ชั่วโมง ซึ่งเป็นเวลานานที่เสื้อชูชีพสามารถลอยน้ำได้ ทั้งนี้จากการวิจัย สามารถวัดค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งสำหรับบริเวณลำคลองซึ่งเป็นพื้นที่ที่ไม่มีวัตถุกีดขวางในระยะ 10 เมตรอยู่ที่ 1.50 เมตร ค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งบริเวณสระว่ายน้ำแห่งที่ 1 เป็นพื้นที่เปิด ซึ่งในระยะ 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน กำแพงและต้นไม้สูง 4 เมตร 3 ด้าน อยู่ที่ 3.60 เมตร และค่าเฉลี่ยความคลาดเคลื่อนของตำแหน่งบริเวณสระว่ายน้ำแห่งที่ 2 ซึ่งเป็นพื้นที่เปิดที่ในรัศมี 10 เมตร มีอาคารสูง 2 ชั้น 1 ด้าน อยู่ที่ 1.31 เมตร ซึ่งทั้งหมดยังเป็นระยะสายตา

ที่สามารถมองเห็นได้ ทำให้ผู้เฝ้าระวังสามารถหยิบยื่นความช่วยเหลือ ได้ทันเวลา

ข้อเสนอแนะในการวิจัยครั้งต่อไป ในกรณีที่ต้องการเพิ่มระยะเวลาการใช้งานอุปกรณ์ให้นานขึ้น สามารถเพิ่มขนาด mAh ได้ เนื่องจากแหล่งจ่ายพลังงานคือ Power Bank โดยการวิจัยในครั้งนี้ใช้ขนาด 5000 mAh ซึ่งสามารถใช้งานได้สูงสุด 16 ชั่วโมง รวมถึง อุปกรณ์ Pocket Wi-Fi ซึ่งรองรับการใช้งานสูงสุด 8 ชั่วโมง โดยหากต้องการเพิ่มระยะเวลาการใช้งานสามารถแก้ไขได้ด้วยการเปลี่ยนรุ่น เพิ่มขนาดแบตเตอรี่ หรือใช้เทคโนโลยี Narrowband IoT (NB-IoT) ทั้งนี้ การใช้งาน NB-IoT ควรดูเรื่องความเร็วการส่งข้อมูลที่ NB-IoT ทำได้ และความต้องการของระบบที่จะพัฒนาเป็นสำคัญ [9] อย่างไรก็ตามงานวิจัยครั้งนี้ใช้เครือข่ายผู้ให้บริการสัญญาณเครือข่ายภายในประเทศไทยครอบคลุมบริเวณชายฝั่ง และอาณาเขตทางทะเลของประเทศไทยเท่านั้น การวิจัยในประเทศอื่นอาจต้องพิจารณาผู้ให้บริการอีกครั้ง [10]

เอกสารอ้างอิง

- [1] สำนักงานกองทุนสนับสนุนการส่งเสริมสุขภาพ (สสส.). “สพฉ. แนะนำ วิธีเอาตัวรอดจากการจมน้ำ.” [สืบค้นวันที่ 6 กุมภาพันธ์ 2565]. จาก <https://www.thaihealth.or.th/สพฉ-แนะนำ-วิธีเอาตัวรอดจ/, 2561>
- [2] กรมเจ้าท่า. ระเบียบกรมเจ้าท่าว่าด้วยหลักเกณฑ์การถ่ายโอนภารกิจด้านการรับรองคุณภาพมาตรฐานอุปกรณ์ประจำเรือ (เสื้อชูชีพ) ให้แก่ภาคเอกชนหรือภาคส่วนอื่น พ.ศ. ๒๕๕๔. จท.(หมวด 1-03-044 เรื่องที่ 39), 2554
- [3] Buck, P.J., Roberts, C.W., & Minehane, C.K., “The Consequences of Cold Water Immersion: Impacts and Treatment,” International Journal of Aquatic Research and Education, Vol. 11, No. 4, pp. 1-24, 2019.
- [4] โรงพยาบาลสมิติเวช. “4 นาทีชีวิต เมื่อถูกจมน้ำ.” [สืบค้นวันที่ 6 กุมภาพันธ์ 2565]. จาก <https://www.samitivejhospitals.com/th/article/detail/4 นาทีชีวิตเมื่อถูกจมน้ำ, 2560>
- [5] กุลปริญ นกดี. “การยอมรับเทคโนโลยี GPS Tracking ของบริษัท ที.ที.ทรานส์ เอ็กซ์เพรส จำกัด.” การค้นคว้าอิสระบริหารธุรกิจมหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยกรุงเทพ, 2557.
- [6] นิตวารรณ ช้างทอง. “การสร้างเลิร์นนิ่งอ้อมเจดีย์วิชาฟิสิกส์ เรื่อง แรงลอยตัว สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5.”

- การค้นคว้าแบบอิสระศึกษาศาสตรมหาบัณฑิต
คณะเทคโนโลยีทางการศึกษา มหาวิทยาลัยเชียงใหม่, 2557
- [7] มหศักดิ์ เกตุกล้า. (2564). [ออนไลน์]. IoT Architecture and Technologies. (วิชาอินเทอร์เน็ตของสรรพสิ่งสำหรับธุรกิจดิจิทัล). [สืบค้นวันที่ 20 กรกฎาคม 2565].
- [8] Naveen, S., “Study of IoT: Understanding IoT Architecture, Applications, Issues and Challenges,” Paper presented at International Conference on Innovations in Computing & Networking (ICICN16), Bangalore, India, May, 2016.
- [9] Chen, M., Miao, Y., Hao, Y., & Hwang, K. Narrow Band Internet of Things. *Institute of Electrical and Electronics Engineers (IEEE)*. 5, 20557 – 20577, 2017.
- [10] Nperf. “3G / 4G / 5G coverage map, Thailand.” [cited 6 February 2023]. Available from <https://www.nperf.com/en/map/TH/>, n.d.

ปัจจัยที่มีอิทธิพลต่อพฤติกรรมการใช้งานคลับเฮาส์ เพื่อเป็นช่องทางในการสื่อสารและการแสดงออกทางความคิด

Acceptance Factors for the use of Clubhouse as the Communication Channel and the Expression of Thoughts

รัชฎ ศรีอรุณ (Rachata Sriaroon)¹ และปราโมทย์ ลีโอนาม (Pramote Luenam)²
สาขาบริหารเทคโนโลยีสารสนเทศ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์
¹rachata.sria@stu.nida.ac.th, ²pramote@nida.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ในการศึกษาปัจจัยที่มีผลต่อพฤติกรรมการใช้งานคลับเฮาส์เพื่อเป็นช่องทางในการสื่อสารและการแสดงออกทางความคิด เครื่องมือที่ใช้คือแบบสอบถามออนไลน์ รวบรวมข้อมูลด้วยการสุ่มแบบสะดวกจากกลุ่มประชากรบุคคลทั่วไปที่มีประสบการณ์การใช้งานคลับเฮาส์ หรือกำลังใช้งานคลับเฮาส์ จำนวน 408 คน ทำการเก็บข้อมูลในช่วงระหว่างเดือน เมษายนถึง พฤษภาคม 2565 และนำข้อมูลมาวิเคราะห์ด้วยโมเดลสมการโครงสร้างโดยใช้โปรแกรม SPSS และ AMOS เพื่อทดสอบสมมติฐานและหาความสัมพันธ์กับข้อมูลเชิงประจักษ์ โมเดลแสดงความเหมาะสมด้วยค่า $P\text{-value}=0.151$ $\chi^2/df = 1.073$ $RMSEA=0.013$ $CFI=0.995$ และ $TLI=0.993$ ผลการศึกษาพบว่า พฤติกรรมความตั้งใจใช้และความไว้วางใจ มีอิทธิพลทางตรงต่อพฤติกรรมการใช้งานคลับเฮาส์ การรับรู้ความปลอดภัยมีอิทธิพลทางอ้อมต่อพฤติกรรมการใช้งานคลับเฮาส์ ความรู้ความเข้าใจส่วนตัวไม่มีอิทธิพลทางอ้อมต่อพฤติกรรมการใช้งานคลับเฮาส์ ความคาดหวังในประสิทธิภาพ ความคาดหวังในความพยายาม และอิทธิพลทางสังคมไม่มีอิทธิพลทางอ้อมต่อพฤติกรรมการใช้งานคลับเฮาส์ ความพึงพอใจในการต้องการความสัมพันธ์ ไม่มีอิทธิพลทางตรงต่อพฤติกรรมการใช้งานคลับเฮาส์ การตอบกลับทันทีและความสำคัญเฉพาะตัวบุคคลไม่มีอิทธิพลทางอ้อมต่อพฤติกรรมการใช้งานคลับเฮาส์

คำสำคัญ: คลับเฮาส์ ทฤษฎีรวมของการยอมรับและการใช้เทคโนโลยี สื่อสังคมออนไลน์ การสื่อสาร ทฤษฎีสรรพภาพของสื่อ

Abstract

The purpose of this research was to study the acceptance factors for Clubhouse as one of as the communication channel and the expression of thoughts. A questionnaire was used as a tool to collect the data. The data were collected from convenience sample of 408 users who have experience of using Clubhouse. The data were analyzed through Structural Equation Model to test the hypotheses and derive model fit. The model showed a goodness-of-fit with $P\text{-value}=0.151$ $\chi^2/df=1.073$ $RMSEA=0.013$ $CFI=0.995$ and $TLI=0.993$. The results indicated that behavioral intention and trust have a significant direct effect on the use behavior of Clubhouse. Perceived privacy and perceived security have a significant indirect effect on the use behavior of Clubhouse. Performance expectancy, effort expectancy, social influence, and system and service quality do not have a significant indirect effect on the use behavior of Clubhouse. Relatedness need satisfaction does not have a significant direct effect on the use behavior of Clubhouse. However, the immediate feedback and personal focus do not have an indirect effect on the use behavior of Clubhouse.

Keywords: Clubhouse , Unified Theory of Acceptance and Use of Technology, Social media, Communication, Media Richness Theory.

1. บทนำ

ปัจจุบันการสื่อสารผ่านทางสื่อสังคมออนไลน์มีการพัฒนาอย่างรวดเร็ว โดยที่สื่อสังคมออนไลน์เป็นเทคโนโลยี

เชิงโต้ตอบที่อนุญาตให้มีการสร้าง แบ่งปัน แลกเปลี่ยน ข้อมูลทางความคิดความสนใจ รวมถึงการแสดงออก ผ่านชุมชนหรือเครือข่ายเสมือนจริง [1] คลับเฮาส์เป็นสื่อสังคมออนไลน์ ที่ได้รับความสนใจหลังจากมีการเปิดตัวไปในช่วงเดือนเมษายน พ.ศ. 2563 ปัจจุบันการใช้งานคลับเฮาส์มีการเติบโตขึ้นอย่างรวดเร็วจนมีผู้ใช้มากกว่า 10 ล้านคนทั่วโลก โดยคุณสมบัติเด่นคือผู้ใช้สามารถใช้งานสื่อสารระหว่างกันได้ ทำให้ผู้ใช้สามารถเข้าถึงและสำรวจหัวข้อที่ต้องการสนทนาได้ การแบ่งปัน การตั้งคำถาม การอภิปราย และการโต้วาที โดยยังผู้ใช้งานมีข้อมูลที่มีความน่าสนใจมากขึ้นจะทำให้มีการแนะนำให้มีการเข้าร่วมมากขึ้น [2] จากการศึกษาสำรวจของ YouGov PLC. [3] ผ่านช่องทางออนไลน์จากกลุ่มประชากรทั้งหมด 2,074 คนในประเทศไทย มีคนรู้จักคลับเฮาส์ ถึงร้อยละ 72 แต่มีผู้ใช้งานจริงอยู่เพียงร้อยละ 25 ซึ่งมีบางส่วนสนใจใช้งานแต่ไม่สามารถใช้งานได้ร้อยละ 15 เนื่องจากข้อจำกัดของแอปพลิเคชัน และมีกลุ่มที่กำลังตัดสินใจที่จะใช้งานในเร็วนี้้อยู่ถึงร้อยละ 45 และกลุ่มที่ไม่สนใจที่จะใช้งานอยู่ร้อยละ 25 สำหรับหัวข้อหลักในการที่ผู้ใช้งานเข้าใช้งานเป็นเรื่องของความรู้ทั่วไป ความบันเทิง และสังคมการเมือง ตามลำดับ ซึ่งมีอัตราส่วนใกล้เคียงกัน

จากการคาดการณ์จำนวนผู้ใช้คลับเฮาส์ในประเทศไทยจะมีมากกว่า 2 ล้านคน ซึ่งทำให้คลับเฮาส์เป็นสื่อสังคมออนไลน์ที่มีการเจริญเติบโตอย่างรวดเร็ว และสามารถสร้างตัวตนของบุคคลได้ในระยะเวลาอันรวดเร็ว ความนิยมที่เกิดขึ้นในคลับเฮาส์แสดงให้เห็นถึงความต้องการความรู้ใหม่ ทำให้รู้ว่าความรู้ในอดีตไม่สามารถใช้งานได้ในปัจจุบัน แต่นอกจากแนวทางการเริ่มใช้งานในการเปลี่ยนแปลงเหนือการคาดการณ์ อย่างไรก็ตามยังมีความกังวลสำหรับการใช้งาน [4] ทางโฆษกกองบัญชาการตำรวจสืบสวนอาชญากรรมทางเทคโนโลยีเตือนการใช้ด้วยการสนทนาผ่านเสียงได้อย่างมีอิสระ อาจจะทำให้ผู้ใช้งานคนพูดแสดงความเห็นในลักษณะพาดพิงหรือละเมิดผู้อื่นจนเกิดความเสียหาย หรือการกล่าวข้อมูลบิดเบือนจนทำให้มีโทษทางกฎหมาย อาจเข้าข่าย

ความผิดฐานหมิ่นประมาท และรวมถึงการนำข้อมูลเท็จเข้าสู่ระบบ [5]

จากข้อมูลข้างต้นทั้งแนวคิดในการคาดการณ์ผลการนำไปใช้ในอนาคต หรือข้อควรระวังในการใช้งานคลับเฮาส์ โดยที่ยังไม่มีการศึกษาอย่างแน่ชัดถึงปัจจัยในแต่ละด้านในเชิงสถิติ ที่จะสามารถบ่งบอกถึงแนวทางในการใช้งานของคลับเฮาส์ที่มีการสำรวจว่ามีการใช้งานเพิ่มสูงขึ้น แต่ยังไม่ได้ใช้งานอย่างเป็นวงกว้าง

ผู้วิจัยจึงมีแนวคิดที่จะศึกษาปัจจัยที่มีอิทธิพลต่อพฤติกรรมการใช้งานคลับเฮาส์เพื่อ เพื่อให้ทราบถึงปัจจัยที่มีอิทธิพลต่อพฤติกรรมการใช้งานคลับเฮาส์ เป็นช่องทางในการสื่อสารและการแสดงออกทางความคิด รวมถึงเป็นแนวทางในการพัฒนาระบบการทำงานของคลับเฮาส์ การสื่อสาร เพื่อทำให้เกิดผู้ใช้งานคลับเฮาส์เพิ่มมากขึ้นในอนาคต

2. งานวิจัยและทฤษฎีที่เกี่ยวข้อง

ในการศึกษาปัจจัยที่มีอิทธิพลต่อพฤติกรรมการความตั้งใจใช้งานคลับเฮาส์ เพื่อเป็นช่องทางในการสื่อสารและการแสดงออกทางความคิด ผู้วิจัยได้ศึกษา ทฤษฎีงานวิจัยที่เกี่ยวข้อง และการทบทวนวรรณกรรม ดังนี้

ทฤษฎีรวมของการยอมรับและการใช้เทคโนโลยี (Unified Theory of Acceptance and Use of Technology: UTAUT) เป็นการรวบรวมทฤษฎีที่เกี่ยวกับการศึกษาถึงพฤติกรรมมนุษย์เข้าด้วยกัน [6] โดยพฤติกรรมการใช้ (Use Behavior) ได้รับแรงขับเคลื่อนมาจากพฤติกรรมความตั้งใจ (Behavior Intention: BI) ปัจจัยที่มีอิทธิพลต่อพฤติกรรมความตั้งใจประกอบด้วยปัจจัยหลัก 3 ประการ ได้แก่ ความคาดหวังในประสิทธิภาพ (Performance Expectancy: PE) ความคาดหวังในความพยายาม (Effort Expectancy: EE) และอิทธิพลทางสังคม (Social Influence: SI)

จากการศึกษาแบบจำลองของทฤษฎีรวมของการยอมรับและการใช้เทคโนโลยี [6] พบว่าความสัมพันธ์ระหว่างพฤติกรรมความตั้งใจ (BI) ที่ได้รับอิทธิพลมาจาก 3 ปัจจัยหลัก มีอิทธิพลโดยตรงต่อพฤติกรรมการใช้งานคลับเฮาส์ (Use Behavior: USE) เพื่อเป็นช่องทางในการสื่อสาร

และการแสดงออกทางความคิด จากการศึกษาทฤษฎีดังกล่าว จึงได้ตั้งสมมติฐานดังนี้

สมมติฐานที่ 1 (H1) : ความคาดหวังในความพยายาม (EE) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)

สมมติฐานที่ 2 (H2) : ความคาดหวังในประสิทธิภาพ (PE) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)

สมมติฐานที่ 3 (H3) : อิทธิพลทางสังคม (SI) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)

สมมติฐานที่ 4 (H4) : พฤติกรรมความตั้งใจ (BI) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE)

คุณภาพของระบบและบริการ (System and Service Quality : SSQ) หมายถึง ระดับที่ผู้ใช้งานเชื่อว่าบริการโดยรวมและประสิทธิภาพของระบบนั้นตรงตามที่คาดหวัง [7], [8] จากการศึกษาโดย [9] ระบุว่าคุณภาพของระบบและบริการเป็นปัจจัยสำคัญที่เอื้อต่อการใช้งานในบริการ โขเดียลมีเดียและเครือข่าย จากการศึกษาของ [10] วิเคราะห์แรงผลักดันที่ทำให้บริการเครือข่ายสังคมออนไลน์ประสบความสำเร็จ การวิเคราะห์เปรียบเทียบการยอมรับของผู้ใช้เฟซบุ๊กและทวิตเตอร์ พบว่า คุณภาพและการบริการมีบทบาทสำคัญในการกำหนดทัศนคติของผู้ใช้

จากการศึกษาทฤษฎีดังกล่าว ทางผู้วิจัยคาดการณ์ว่าคุณภาพของระบบและบริการ (SSQ) มีอิทธิพลต่อพฤติกรรมความตั้งใจใช้คลัสเฮาส์ จึงตั้งสมมติฐานดังนี้

สมมติฐานที่ 5 (H5) : คุณภาพของระบบและบริการ (SSQ) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)

การรับรู้ความเป็นส่วนตัว (Perceived Privacy: PP) หมายถึง การที่ผู้ใช้งาน การประเมินว่า มีการรับรู้ทราบถึงการเข้าถึงข้อมูลส่วนบุคคลของตนเองจากบุคคลภายนอก หรือการรับรู้เกี่ยวกับการเข้าถึงข้อมูลส่วนบุคคลได้อย่างจำกัด [11] ซึ่งหากรับรู้ถึงความเป็นส่วนตัว จะนำไปสู่การเกิดความรู้สึกเชิงบวก

การรับรู้ความปลอดภัย (Perceived Security: PS) หมายถึง การที่ผู้ใช้งาน การประเมินการรับรู้ว่าจะระบบมีมาตรการรักษาความปลอดภัยที่น่าเชื่อถือได้ [12] โดยที่ผู้ใช้งานจะรับรู้ว่ามีบริการจะปฏิบัติตามข้อกำหนดทางด้านความปลอดภัย เช่น ความถูกต้อง ความสมบูรณ์ และการเข้ารหัสในการเข้าถึงข้อมูล จากการศึกษาของ [13] เห็นว่าการรับรู้ความปลอดภัย นำไปสู่ความรู้สึกเชิงบวก และรับรู้การควบคุมระบบหรือบริการมากขึ้น

ความไว้วางใจ (Trust: TR) หมายถึง ความตั้งใจกระทำต่อผู้อื่นบนพื้นฐานความคาดหวังเชิงบวก ภายใต้ความเสี่ยงและการพึ่งพากัน [14] โดยเป็นความคาดหวังต่อความน่าเชื่อถือ ความเชื่อมั่นจากคำพูด การกระทำ คำมั่นสัญญาจากการเขียนหรือข้อความที่ได้ชี้แจงไว้ [15] ความไว้วางใจมีความสำคัญอย่างยิ่งในเทคโนโลยีสารสนเทศ [16] โดยมีงานวิจัยมากมายที่เน้นย้ำถึงความสำคัญของความไว้วางใจในการแบ่งปันข้อมูลและการเปิดเผยข้อมูลส่วนบุคคล

[17] ได้มีการศึกษาเกี่ยวกับ ความสัมพันธ์ระหว่างการรับรู้ถึงความปลอดภัย ความเป็นส่วนตัว และความตั้งใจทางการค้าทางสังคม พบว่าการค้าในเครือข่ายสังคมออนไลน์มักได้รับความเชื่อถือเกี่ยวกับการใช้งาน และพบว่า TR, PP และ PS นั้นมีความสอดคล้องกัน

จากงานวิจัยข้างต้นทางผู้วิจัยนำเรื่องของ การรับรู้ความเป็นส่วนตัว (Perceived Privacy: PP) และการรับรู้ความปลอดภัย (Perceived Security:PS) เป็นตัวแปรในการส่งผลถึงความไว้วางใจ (Trust) ซึ่งทางผู้วิจัยคาดการณ์ว่าตัวแปรดังกล่าวมีอิทธิพลต่อพฤติกรรมความตั้งใจใช้งานคลัสเฮาส์ จึงได้ตั้งสมมติฐานดังนี้

สมมติฐานที่ 6 (H6) : การรับรู้ความเป็นส่วนตัว (PP) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR)

สมมติฐานที่ 7 (H7) : การรับรู้ความปลอดภัย (PS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR)

สมมติฐานที่ 8 (H8) : ความไว้วางใจ (TR) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE)

[18] อธิบายแนวคิดสมรรถภาพของสื่อ (MRT) เป็นกรอบงานที่ใช้อธิบายความสามารถของสื่อในการสื่อสารสำหรับข้อมูลที่ส่งผ่านเพื่อเป็นส่วนขยายของการประเมินความสมบูรณ์ โดยในทฤษฎีมีการกล่าวไว้ว่า ยิ่งการเรียนรู้ที่สามารถแสดงผ่านสื่อได้มากเท่าใด สื่อก็ยิ่งสมบูรณ์มากขึ้นเท่านั้น ความสมบูรณ์ของสื่อเป็นหน้าที่ของคุณลักษณะต่าง ๆ ดังต่อไปนี้

การตอบกลับทันที (Immediate Feedback: IF) หมายถึง ความสามารถของสื่อกลาง ตัวกลางหรือระบบที่เปิดโอกาสให้ผู้สื่อสารความสามารถในการตอบกลับได้อย่างรวดเร็วและทันที ทำให้ดีความได้อย่างรวดเร็ว [18]

ความสำคัญเฉพาะตัวบุคคล (Personal Focus: PF) หมายถึง ความสามารถในการใช้งานของสื่อกลาง ตัวกลางหรือระบบที่สามารถให้ผู้สื่อสาร นั้นแสดงถึงตัวตนของผู้สื่อสาร [18] รวมถึงสามารถปรับให้เข้ากับตัวผู้สื่อสาร เพื่อให้ตรงกับความต้องการและสถานการณ์ปัจจุบันของผู้สื่อสาร สามารถออกแบบได้เฉพาะตัว

ความพึงพอใจในความต้องการความสัมพันธ์ (Relatedness Need Satisfaction : RNS) หมายถึง การเติมเต็มความต้องการในการเชื่อมต่อทางสังคม ความสัมพันธ์ทางสังคม [19] โดยที่เป็นความพึงพอใจที่เชื่อมโยงในทัศนคติเชิงบวก [20] ซึ่งช่วยเสริมความผูกพันกับแอปพลิเคชันบนอุปกรณ์เคลื่อนที่

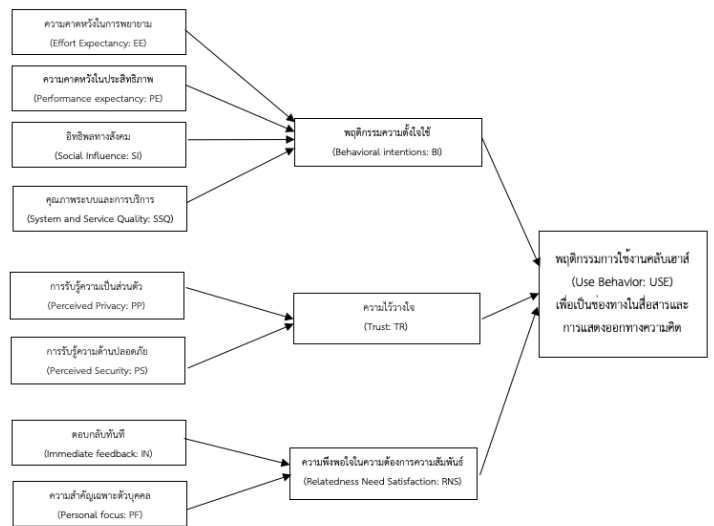
จากงานวิจัยข้างต้นทางผู้วิจัยนำเรื่องของ การตอบกลับทันที (Immediate Feedback: IF) และความสำคัญเฉพาะตัวบุคคล (Personal Focus: PF) เป็นตัวแปรในการส่งผลถึง ความพึงพอใจในความต้องการความสัมพันธ์ (RNS) จึงได้ตั้งสมมติฐานดังนี้

สมมติฐานที่ 9 (H9) : การตอบกลับทันที (IF) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านทางความพึงพอใจในความต้องการความสัมพันธ์ (RNS)

สมมติฐานที่ 10 (H10) : ความสำคัญเฉพาะตัวบุคคล (PF) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE) ผ่านทางความพึงพอใจในความต้องการความสัมพันธ์ (RNS)

สมมติฐานที่ 11 (H11) : ความพึงพอใจในความต้องการความสัมพันธ์ (RNS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัสเฮาส์ (USE)

จากการที่ผู้วิจัยได้ศึกษาแนวคิด และทฤษฎีที่เกี่ยวข้อง สามารถกำหนดกรอบแนวคิดของการวิจัย ดังแสดงในภาพที่ 1



ภาพที่ 1: กรอบแนวคิดการวิจัย

3. วิธีการดำเนินงานวิจัย

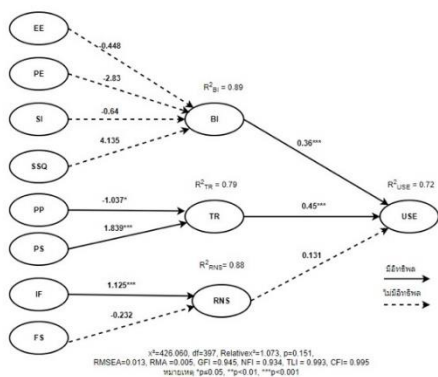
ในการศึกษางานวิจัยนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) เครื่องมือที่ใช้คือแบบสอบถามออนไลน์ รวบรวมข้อมูลด้วยการสุ่มแบบสะดวก ในช่วงเดือนเมษายนถึงพฤษภาคม 2565 จากกลุ่มตัวอย่างบุคคลทั่วไปที่มีประสบการณ์การใช้งานคลัสเฮาส์หรือกำลังใช้งานคลัสเฮาส์ เพื่อเป็นช่องทางในสื่อสารและการแสดงออกทางความคิด จำนวน 408 คน ผู้วิจัยใช้โปรแกรม G*Power ในการคำนวณขนาดของกลุ่มตัวอย่าง โดยกำหนดค่าเพาเวอร์ (1-β) เท่ากับ 0.95 ค่าอัลฟา (α) เท่ากับ 0.05 ขนาดของอิทธิพล (Effect Size) เท่ากับ 0.03 จำนวนตัวแปรทำนายเท่ากับ 12 ตัวแปร ผลที่ได้คือขนาดของกลุ่มตัวอย่างเท่ากับ 363 ตัวอย่าง และเพิ่มการเก็บข้อมูลที่ใช้สำหรับการศึกษานำร่อง เพื่อหาความเชื่อมั่นอีก 30 ตัวอย่าง รวมเป็น 393 ตัวอย่าง ซึ่งผู้วิจัยได้ปรับจำนวนขนาดตัวอย่างแบบสอบถามเพิ่มเติมเป็น 400 ตัวอย่าง

การวัดคุณภาพเครื่องมือแบบสอบถาม ทำโดยการวัดความเที่ยงตรง (Valid) พบว่าทุกข้อมีค่าผ่านเกณฑ์มาตรฐาน แล้วจึงนำมาวัดค่าความเที่ยง (Stability) ได้ค่า Cronbach's alpha coefficient เท่ากับ 0.764-0.927 ถือได้ว่าแบบสอบถามที่ใช้มีความน่าเชื่อถืออยู่ในระดับดีมาก งานวิจัยนี้ใช้การวิเคราะห์โมเดลสมการ โครงสร้าง (Structural Equation Modeling: SEM) ในการทดสอบสมมติฐานเพื่อวิเคราะห์อิทธิพลของตัวแปรอิสระ ตัวแปรตาม และตัวแปรคั่นกลาง ในการทดสอบสมมติฐานจะทำการทดสอบที่ระดับนัยสำคัญทางสถิติ 0.05

4. ผลการวิจัย

ในการศึกษาวิจัยครั้งนี้มีผู้ตอบแบบสอบถามจำนวน 451 คน โดยเป็นกลุ่มประชากรกลุ่มประชากรบุคคลทั่วไปที่มีประสบการณ์การใช้งานคลัดเฮาส์ หรือกำลังใช้งานคลัดเฮาส์ เพื่อเป็นช่องทางในสื่อสารและการแสดงออกทางความคิดที่สามารถนำมาวิเคราะห์ข้อมูลได้เป็นจำนวน 408 คน

การวิจัยใช้การวิเคราะห์รูปแบบเชิงโครงสร้าง โดยเป็นการวิเคราะห์เส้นทาง (path analysis) เพื่อหาความสัมพันธ์ทางตรงและทางอ้อมของตัวแปรอิสระว่ามีผลต่อตัวแปรตามมากน้อยเพียงใดและทดสอบสมมติฐาน โมเดลแสดงความเหมาะสมด้วยค่า P-value=0.151 $\chi^2/df = 1.073$ RMSEA = 0.013 CFI = 0.995 และ TLI = 0.993 ดังแสดงในภาพที่ 2 ผลการวิเคราะห์อิทธิพลเชิงสาเหตุภายในโมเดลแสดงในตารางที่ 1



ภาพที่ 2: สัมประสิทธิ์เส้นทางมาตรฐานของโมเดล

ตารางที่ 1: ผลการวิเคราะห์เชิงสาเหตุของ โมเดล

ตัวแปรผล	อิทธิพล	เชิงสาเหตุ										
		EE	PE	SI	SSQ	PP	PS	IF	FS	BI	TR	RNS
BI	DE	-0.448	-2.83	-0.64	4.135	-	-	-	-	-	-	-
	IE	-	-	-	-	-	-	-	-	-	-	-
TR	DE	-	-	-	-	-1.037*	1.839***	-	-	-	-	-
	IE	-	-	-	-	-	-	-	-	-	-	-
RNS	DE	-	-	-	-	-	-	1.125***	-0.232	-	-	-
	IE	-	-	-	-	-	-	-	-	-	-	-
USE	DE	-	-	-	-	-	-	-	-	0.36***	0.45***	0.131
	IE	-0.161	-1.018	-0.23	1.488	-0.466	0.827	0.147	-0.003	-	-	-

หมายเหตุ *p<0.05, **p<0.01, ***p<0.001

ผลที่ได้จากการวิเคราะห์อิทธิพลเชิงสาเหตุภายในโมเดล สามารถนำมาในตารางที่ 2 ดังนี้

ตารางที่ 2: สรุปผลการทดสอบสมมติฐาน

สมมติฐาน	ค่าอิทธิพล	ค่า P	ผลการทดสอบ
H1: ความคาดหวังในความพยายาม (EE) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)	0.161	0.210	ปฏิเสธ
H2: ความคาดหวังในประสิทธิภาพ (PE) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)	-1.018	0.003	ปฏิเสธ
H3: อิทธิพลทางสังคม (SI) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)	-0.23	0.759	ปฏิเสธ
H4: พฤติกรรมความตั้งใจ (BI) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE)	0.36	0.001	ยอมรับ
H5: คุณภาพของระบบและบริการ (SSQ) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI)	1.488	0.001	ปฏิเสธ
H6: การรับรู้ความเป็นส่วนตัว (PP) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR)	-0.466	0.132	ปฏิเสธ
H7: การรับรู้ความปลอดภัย (PS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR)	0.827	0.010	ยอมรับ
H8: ความไว้วางใจ (TR) มีอิทธิพลพฤติกรรมการใช้งานคลัดเฮาส์ (USE)	0.147	0.001	ยอมรับ
H9: การตอบกลับทันที (IF) มีอิทธิพลพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านทางความพึงพอใจในความต้องการความสัมพันธ์ (RNS)	0.003	0.431	ปฏิเสธ
H10: ความสำคัญเฉพาะตัวบุคคล (FS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE) ผ่านทางความพึงพอใจในความต้องการความสัมพันธ์ (RNS)	0.003	0.331	ปฏิเสธ
H11: ความพึงพอใจในความต้องการความสัมพันธ์ (RNS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE)	0.131	0.199	ปฏิเสธ

5. สรุป

ผลการศึกษาวิจัยที่มีผลต่อพฤติกรรมการใช้งานคลัดเฮาส์ เพื่อเป็นช่องทางในการสื่อสารและการแสดงออกทางความคิดผู้วิจัยได้ดำเนินการวิเคราะห์และอภิปรายผลการทดสอบเป็นประเด็นต่าง ๆ ได้ดังนี้

ปัจจัยพฤติกรรมความตั้งใจใช้ (BI) มีอิทธิพลเชิงบวกต่อพฤติกรรมการใช้งานคลัดเฮาส์ (USE)

ซึ่งสอดคล้องกับสมมติฐานที่ได้ตั้งไว้ และงานสอดคล้องกับงานวิจัยการศึกษาแบบจำลองของทฤษฎีรวมของการยอมรับและการใช้เทคโนโลยี [6] ซึ่งกล่าวได้ว่าเมื่อผู้ใช้งานมีระดับความตั้งใจใช้งาน ความต้องการใช้งาน และคลิกเฮาส์สามารถตอบสนองต่อการใช้งานได้จะส่งผลถึงการใช้งานคลิกเฮาส์ มีโอกาสที่ทำให้คลิกเฮาส์สามารถเพิ่มจำนวนผู้ใช้งานได้เพิ่มมากขึ้น ซึ่งในงานวิจัยได้ศึกษาตามทฤษฎี [6]

ความคาดหวังในความพยายาม (EE) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI) ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ซึ่งกล่าวได้ว่าผู้ใช้งานคลิกเฮาส์ยังไม่รับรู้ถึงความสะดวก ความง่ายของการใช้งานคลิกเฮาส์ ทั้งนี้อาจเกิดจากผู้ใช้ส่วนใหญ่อยู่ในช่วงวัยที่สามารถเรียนรู้การใช้งานได้อย่างรวดเร็วในการใช้งานแอปพลิเคชัน รวมถึงการเรียนรู้การทำงานจากระบบโดยปราศจากขั้นตอนคำแนะนำคือถึงแม้ว่าผู้ใช้ส่วนใหญ่อาจจะมีโอกาสที่สามารถเรียนรู้การใช้งานได้ แต่คลิกเฮาส์ยังควรที่จะพัฒนาระบบอย่างต่อเนื่อง เช่นรูปแบบลงทะเบียน การเข้าถึงห้องสื่อสาร การค้นหาตามความสนใจยังเป็นโอกาสกระตุ้นให้เกิดผู้ใช้งานได้มากขึ้น

ความคาดหวังในประสิทธิภาพ (PE) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI) ไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ ซึ่งกล่าวได้ว่า ผู้ใช้งานเชื่อว่าคลิกเฮาส์ยังไม่สามารถช่วยเพิ่มประสิทธิภาพ หรือช่วยทำให้บรรลุผลในการสื่อสารได้เพิ่มมากขึ้น ข้อเสนอแนะคือควรมีการพัฒนาประสิทธิภาพของระบบ และการพัฒนาอย่างต่อเนื่องโดยยึดความต้องการของผู้ใช้งาน และรับรู้ถึงปัญหาจากผู้ใช้งาน

อิทธิพลทางสังคม (SI) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI) ไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ กล่าวได้ว่าผู้ใช้งานมองว่าบุคคลอื่นไม่ได้มีความสำคัญ หรือทำให้เชื่อว่าต้องใช้งานคลิกเฮาส์ ซึ่งอาจเกิดจากการที่ผู้ใช้อยู่ในช่วงอายุที่สามารถตัดสินใจหรือปรับพฤติกรรมได้ด้วยตัวเอง

ทำให้บุคคลรอบข้างไม่ได้มีความสำคัญในการตัดสินใจที่ให้ความเชื่อว่าควรใช้หรือไม่ควรใช้งานคลิกเฮาส์

คุณภาพของระบบและบริการ (SSQ) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านพฤติกรรมความตั้งใจ (BI) ไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ กล่าวได้ว่าผู้ใช้งานเชื่อว่าการบริการโดยรวมและประสิทธิภาพของระบบไม่ตรงตามคาดหวัง คุณภาพและการบริการไม่ได้มีบทบาทสำคัญต่อพฤติกรรมการใช้งานคลิกเฮาส์ ซึ่งทางคลิกเฮาส์จะต้องเพิ่มเรื่องของคุณภาพระบบและบริการ ข้อเสนอแนะของผู้วิจัย การพัฒนาการใช้งานคลิกเฮาส์นั้นยังต้องพัฒนาเพื่อให้ครอบคลุมการใช้งานกับทุกรูปแบบ ทั้งความพึงพอใจในคุณภาพของเสียง การตัดเสียงรบกวน รวมถึงข้อจำกัดของจำนวนคนในห้อง เพื่อให้ผู้ใช้ได้รับการบริการและประสิทธิภาพอย่างเต็มที่

ความไว้วางใจ (TR) มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ซึ่งสอดคล้องกับสมมติฐานที่ได้ตั้งไว้ กล่าวได้ว่าผู้ใช้งานเมื่อมีความตั้งใจกระทำด้วยความคาดหวังเชิงบวก และคาดหวังต่อความน่าเชื่อถือจากมาตรการรักษาความปลอดภัย ข้อจำกัดด้านความปลอดภัย รับรู้ถึงการเข้าถึงข้อมูลส่วนบุคคล ส่งผลต่อพฤติกรรมการใช้งานคลิกเฮาส์ ซึ่งหากคลิกเฮาส์นั้นสามารถรักษาถึงปัจจัยดังกล่าวได้อย่างต่อเนื่องจะสามารถทำให้ผู้ใช้งานเกิดความไว้วางใจและมีโอกาสเพิ่มจำนวนผู้ใช้งานได้

การรับรู้ความเป็นส่วนตัว (PP) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR) ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ กล่าวได้ว่าผู้ใช้งานประเมินว่าการรับรู้ถึงการเข้าถึงข้อมูลส่วนบุคคล หรือการรับรู้เกี่ยวกับการเข้าถึงข้อมูลได้อย่างจำกัดส่งผลต่อความไว้วางใจ ซึ่งมีผลต่อความรู้สึกเชิงบวกทำให้เกิดความไว้วางใจ แต่ยังไม่ส่งผลถึงการที่ผู้ใช้จะต้องการใช้งานคลิกเฮาส์เพื่อแสดงออกทางความคิด

การรับรู้ความปลอดภัย (PS) มีอิทธิพลต่อพฤติกรรมการใช้งานคลิกเฮาส์ (USE) ผ่านทางความไว้วางใจ (TR) ซึ่งสอดคล้องกับสมมติฐานที่ได้ตั้งไว้ ซึ่งตรงกับทฤษฎี

และงานวิจัย [17] กล่าวว่าได้ว่าผู้ใช้งานรับรู้ว่าคุณคลับเฮาส์มีมาตรการรักษาความปลอดภัยที่น่าเชื่อถือได้ โดยที่คลับเฮาส์จะปฏิบัติตามข้อกำหนดทางด้านความปลอดภัย เช่น ความถูกต้อง ความสมบูรณ์ และการเข้ารหัสในการเข้าถึงข้อมูล ทำให้ผู้ใช้งานเกิดความรู้สึกเชิงบวกมากขึ้น ส่งผลความไว้วางใจ นำสู่พฤติกรรมการใช้งานคลับเฮาส์เพื่อการแสดงออกทางความคิด

ความพึงพอใจในความต้องการความสัมพันธ์ (RNS) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลับเฮาส์ (USE) ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ กล่าวว่าผู้ใช้งานเชื่อว่าคลับเฮาส์ไม่ได้ช่วยความต้องการในการสร้างความสัมพันธ์ทางสังคม

ความสำคัญเฉพาะตัวบุคคล (FS) ไม่มีอิทธิพลต่อพฤติกรรมการใช้งานคลับเฮาส์ (USE) ผ่านทางความพึงพอใจในความต้องการความสัมพันธ์ (RNS) ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ กล่าวว่าผู้ใช้งานยังไม่สามารถทำให้ผู้ใช้งานสื่อสาร แสดงถึงตัวตนออกมาได้อย่างเต็มที่ ไม่สามารถเชื่อมต่อกับผู้ใช้งานกับสังคมความสัมพันธ์กับสังคมได้ ทำให้ยังไม่ส่งผลถึงพฤติกรรมการใช้งานคลับเฮาส์

ข้อเสนอแนะสำหรับงานวิจัยในครั้งถัดไป การศึกษาคควรนำเอาปัจจัยด้านเงื่อนไขการอำนวยความสะดวกเข้ามาเป็นปัจจัยเพิ่มเติม รวมถึงทฤษฎีที่เกี่ยวข้องกับการสื่อสารเพื่อศึกษาปัจจัยอื่นที่อาจจะเป็นอิทธิพลต่อพฤติกรรมการใช้งาน และเนื่องจากคลับเฮาส์มีการพัฒนาระบบอย่างต่อเนื่อง ควรมีการใช้เวลาในการทำแบบสอบถามและอภิปรายผลอย่างทันถ่วงที เพื่อควบคุมตัวแปรให้สอดคล้องกับช่วงเวลาในการดำเนินการศึกษา

เอกสารอ้างอิง

[1] Kietzmann, Jan H , (2011). Social Media? Get serious! Understanding the functional building blocks of social media. 54 (3): 241–251.

[2] Touma, R. (2021). Clubhouse app: what is it and how do you get an invite to the exclusive audio app? *Theguardian.Com*. <https://www.theguardian.com/technology/2021/feb/17/>

clubhouse-app-invite-what-is-it-how-to-get-audio-chat-elon-musk.

[3] AdAddictTH, “สำรวจสถิติที่น่าสนใจ ‘คนไทย กับ Clubhouse’ จากการสำรวจข้อมูล โดย YouGov” 2564. [ออนไลน์]. <https://adaddictth.com/knowledge/Data-Clubhouse-YouGov>.

[4] เบญจวรรณ รัตนวิจิตร, “ส่องธุรกิจส่องโอกาสธุรกิจ คลับเฮาส์” 2564. [ออนไลน์]. <https://forbesthailand.com/news/ส่องโอกาสธุรกิจใน-Clubhouse.html>

[5] สถานีโทรทัศน์ตำรวจแห่งชาติเพื่อประชาชน, “บช.สอท (ตำรวจไซเบอร์) เตือนใช้แอป Clubhouse ต้องระวัง” 2564. [ออนไลน์]. <https://policetv.tv/archives/16701>.

[6] Venkatesh V., Morris M.G., Davis G.B., and Davis F.D. (2003). “User acceptance of information technology: toward a unified view.” *MIS Quarterly*, 27(3): 425-478.

[7] DeLone, W. H., & McLean E. R. (1992). Information system success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60-95

[8] DeLone, W. H., & McLean E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.

[9] Liu, S. H., Liao, H. L., & Pratt, J. A. (2009). Impact of media richness and flow on e-learning technology acceptance. *Computers & Education*, 52(3), 599-607.

[10] Sang Jib Kwon, Eunil Park & Ki Joon Kim (2014). What drives successful social networking services? A comparative analysis of user acceptance of Facebook and Twitter. *The Social Science Journal*, 51(4), 534-544.

[11] Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems*, 22(3), 295-316

[12] Yenisey, M. M., Ozok, A. A., & Salvendy, G. (2005). Perceived security determinants in e-commerce among Turkish university students. *Behaviour & Information Technology*, 24(4), 259-274

[13] Kim, D. J., Ferrin, D. L., & Rao, H. R. (2009). Trust and satisfaction, two stepping stones for successful e-commerce relationships: A longitudinal exploration. *Information systems research*, 20(2), 237-257.

- [14] Davis, J. H., Schoorman, F. D., Mayer, R. C., & Tan, H. H. (2000). The trusted general manager and business unit performance: Empirical evidence of a competitive advantage. *Strategic management journal*, 21(5), 563-576.
- [15] Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American psychologist*, 26(5), 443.
- [16] Bansal, G., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision support systems*, 49(2), 138-150.
- [17] Amro, A. M., Ariffin, A. M., & Norsiah, A. H. The Relationship between Perceived Security, Perceived Privacy and Social Commerce Intention among University Students in Saudi Arabia: The Mediating Role of Trust, Attitude and Satisfaction.
- [18] Daft, R. L., & Lengel, R. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32, 554-571.
- [19] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68.
- [20] Kim, S., & Karrila, S. J. (2013). *Microhydrodynamics: principles and selected applications*. Courier Corporation.

ปัจจัยที่มีผลต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง

Factors Affecting the Continuance Intention to Use of QR Payment Services

นภัตกนก ศรีกาญจ โนคม (Napatkanok Srikanjanodom)¹ และปราโมทย์ ลื่อนาม (Pramote Luenam)²

สาขาบริหารเทคโนโลยีสารสนเทศ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

¹napatkanok.sri@stu.nida.ac.th, ²pramote@nida.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ในการศึกษาปัจจัยที่มีผลต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยศึกษาจากวิธีการสุ่มตัวอย่างแบบสะดวกจำนวน 416 คน ที่มีการใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างน้อยเดือนละครั้ง เป็นระยะเวลา 3 เดือนขึ้นไป ไม่ว่าจะเป็นการชำระเงินผ่านแอปพลิเคชันใดก็ตาม เครื่องมือที่ใช้ในงานวิจัยคือแบบสอบถามออนไลน์ที่ผ่านการทดสอบความน่าเชื่อถือและความถูกต้อง เมื่อนำข้อมูลมาวิเคราะห์ด้วยโมเดลสมการโครงสร้าง เพื่อทดสอบสมมติฐาน และหาความสัมพันธ์กับข้อมูลเชิงประจักษ์ โมเดลแสดงความเหมาะสมด้วยค่า $P\text{-value}=0.273$, $\chi^2/df=1.074$, $RMR=0.009$, $RMSEA=0.013$, $GFI=0.971$, $CFI=0.996$, $NFI=0.949$ และ $TLI=0.994$ จากผลการศึกษาวินิจฉัยพบว่า ปัจจัยการรับรู้ประโยชน์ ความไว้วางใจทางออนไลน์ และการรับรู้ความเสี่ยงมีอิทธิพลเชิงบวกทางตรงต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง และอิทธิพลทางสังคมมีอิทธิพลเชิงบวกทางอ้อมต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยผ่านการรับรู้ประโยชน์ ส่วนความไว้วางใจทางออนไลน์มีอิทธิพลเชิงลบต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยผ่านการรับรู้ความเสี่ยงอย่างมีนัยสำคัญทางสถิติและความสามารถทางนวัตกรรมส่วนบุคคลไม่มีอิทธิพลเชิงบวกทางอ้อมต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยผ่านการรับรู้ความเสี่ยงอย่างไม่มีนัยสำคัญทางสถิติ ผลที่ได้จากการศึกษานี้สามารถเป็นแนวทางให้กับหน่วยงานภาครัฐ

ภาครัฐกิจ และสถาบันการเงิน ในการพัฒนานวัตกรรมกรให้บริการที่สอดคล้องกับความต้องการของผู้ใช้บริการ
คำสำคัญ: การชำระเงินด้วยคิวอาร์โค้ด ความตั้งใจอย่างต่อเนื่อง การชำระเงินอิเล็กทรอนิกส์

Abstract

This study aims to evaluate the factors that affect the continuance intention to use of QR payment services. The study was conducted from a convenience sampling of 416 Thais who have experience using QR payment services at least once a month for a period of no less than 3 months, regardless of the applications used. An online questionnaire is used as the research instrument whose validity and reliability have been tested. The data were analyzed through Structural Equation Model to test the hypotheses and derive model fit. The model showed a goodness-of-fit with $P\text{-value}=0.273$, $\chi^2/df=1.074$, $RMR=0.009$, $RMSEA=0.013$, $GFI=0.971$, $CFI=0.996$, $NFI=0.949$ and $TLI=0.994$. According to analysis results, perceived usefulness, online trust and perceived risk have a significant direct positive effect to the continuance intention to use of QR payment services. Social influence has a significant positive indirect effect to the continuance intention to use of QR payment services. Online trust has a negative effect to the intention to continuance use of QR payment services through perceived risk. Conversely, personal innovativeness does not have a positive indirect to the continuance

intention to use of QR payment services through perceived risk. The findings of the study are useful for government agencies, business sector and financial institution to be more innovative in their service development and responsive to emerging consumer demands.

Keywords: QR Payment, continuance intention, e-Payment.

1. บทนำ

ปัจจุบันประเทศไทยมีแนวคิดส่งเสริมเกี่ยวกับเศรษฐกิจดิจิทัลทำให้กระทรวงการคลังขับเคลื่อนโครงการการชำระเงินอิเล็กทรอนิกส์แห่งชาติตั้งแต่ปี 2558 เพื่อพัฒนาโครงสร้างพื้นฐานการชำระเงินทางอิเล็กทรอนิกส์หรือระบบ “พร้อมเพย์” เพื่อผลักดันประเทศไทยเข้าสู่เศรษฐกิจดิจิทัลสังคมไร้เงินสด ที่เอื้อให้การโอนเงินและชำระเงินทางอิเล็กทรอนิกส์ทำได้ถูกสะดวก รวดเร็ว และปลอดภัยมากขึ้น [1]

ธนาคารแห่งประเทศไทยและสมาคมธนาคารไทยได้ร่วมกำหนดรูปแบบคิวอาร์โค้ดมาตรฐานของไทย หรือ Thai QR Payment ในการชำระเงินผ่านระบบพร้อมเพย์ ทำสอดคล้องกับรูปแบบการใช้ชีวิตคนยุคใหม่ โดยเฉพาะช่วงการแพร่ระบาดของโควิด-19 สังคมไร้เงินสดช่วยเพิ่มความสะดวกในการชำระเงิน เพิ่มความปลอดภัย ตรวจสอบความถูกต้องได้ง่าย ลดการแพร่กระจายของเชื้อโรคบนธนบัตร อีกทั้งช่วยประหยัดต้นทุนค่าใช้จ่ายในการผลิตธนบัตร เหรียญ ค่าบริหารจัดการเงินสด เป็นต้น [2]

การชำระเงินอิเล็กทรอนิกส์ด้วยมาตรฐานคิวอาร์โค้ดเป็นช่องทางที่ร้านค้าและผู้บริโภคสามารถเข้าถึงการใช้บริการได้ง่าย สะดวก ถูกต้อง รวดเร็ว ไม่มีค่าธรรมเนียม ลดความซ้ำซ้อนของการมีคิวอาร์โค้ดหลายประเภท จากการสำรวจพฤติกรรมการใช้ e-Payment ของประชาชน [3] พบว่าเป็นการโอนเงินให้บุคคลอื่น 62% ซื้อสินค้าและบริการออนไลน์ 52% ข้อมูลในเดือนพฤศจิกายน 2564 ระบุว่ามีการใช้เฉลี่ย 301 รายการ/คน/ปี ธุรกิจออนไลน์เพิ่มขึ้นทุกช่องทาง โดยเฉพาะ Mobile Banking ได้รับความนิยมสูงสุด ส่วนช่องทางดั้งเดิม เช่น เช็ค เอทีเอ็ม และสาขา ยังคงมีแนวโน้มลดลงอย่างต่อเนื่อง [4]

จากข้อมูลการตอบรับการใช้บริการชำระเงินด้วยคิวอาร์โค้ดที่ดีจากประชาชน ผู้วิจัยจึงมีความสนใจศึกษาปัจจัยที่มีผลต่อความตั้งใจใช้บริการของผู้บริโภคอย่างต่อเนื่อง เพื่อนำมาใช้ประโยชน์และเป็นแนวทางในการขยายฐานผู้ใช้บริการ อีกทั้งเป็นแนวทางในการพัฒนานวัตกรรมและปรับปรุงบริการให้สอดคล้องกับความต้องการของผู้ใช้บริการและเข้าถึงประชาชนได้มากยิ่งขึ้น

2. งานวิจัยและทฤษฎีที่เกี่ยวข้อง

ความตั้งใจใช้งาน หมายถึง ความพร้อมหรือความเป็นไปได้ของบุคคลที่จะแสดงพฤติกรรมของการใช้งาน มีความพยายามเพื่อที่จะกระทำพฤติกรรมนั้น [5] ความตั้งใจเป็นตัวที่สามารถชี้ให้เห็นว่าบุคคลมีการวางแผนในการแสดงออกของพฤติกรรม [6] ดังนั้น ความตั้งใจใช้บริการอย่างต่อเนื่อง จึงหมายถึง การแสดงพฤติกรรมหรือการตัดสินใจจะใช้บริการชำระเงินด้วยคิวอาร์โค้ด โดยมีแบบแผนหรือมีความพยายามที่จะใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่องสม่ำเสมอทุกเดือน และทุกครั้งเมื่อมีโอกาสที่เหมาะสม เป็นการแสดงพฤติกรรมออกมาในลักษณะที่สอดคล้องกับความตั้งใจ อีกทั้งมีความตั้งใจเชิงพฤติกรรมในการพูดถึงประสบการณ์ดี ๆ ที่ได้รับจากการใช้บริการชำระเงินด้วยคิวอาร์โค้ด โดยการแนะนำจากบุคคลอื่นให้มาใช้บริการชำระเงินด้วยคิวอาร์โค้ด

ความไว้วางใจทางออนไลน์ (Online Trust : OT) หมายถึง ความเชื่อมั่น จงรักภักดี เต็มใจ ที่ผู้ใช้บริการมีให้กับผู้ให้บริการ ในการใช้บริการผ่านอินเทอร์เน็ต [7] อธิบายว่า ความไว้วางใจจะถูกเชื่อมโยงเข้ากับความตั้งใจซื้อ กล่าวคือ หากผู้บริโภคเกิดความไว้วางใจ และความเชื่อมั่นต่อผู้ให้บริการในระดับสูง จะส่งผลให้ผู้บริโภคเกิดความตั้งใจซื้อสินค้า บริการสูงขึ้นตามไปด้วย มีงานวิจัยของ Hankun และคณะ [8] พบว่าในสหรัฐอเมริกาความกังวลเรื่องความเป็นส่วนตัวมีผลกระทบในทางลบอย่างมีนัยสำคัญต่อทัศนคติของผู้ใช้บริการชำระเงินออนไลน์ แต่ในจีนความกังวลเรื่องความเป็นส่วนตัวมีผลกระทบเชิงบวกต่อทัศนคติของผู้ใช้บริการชำระเงินออนไลน์ อีกทั้ง [9] ได้ศึกษาปัจจัยความไว้วางใจบนพื้นฐานของการรับรู้ที่เกิด

จากผู้บริโภคต่อคุณสมบัติและลักษณะของผู้ขายออนไลน์ไว้ได้แก่ การรับรู้ถึงการรักษาความปลอดภัยคุณภาพของสารสนเทศ และการรับรู้ถึงความเป็นส่วนตัว ผู้วิจัยจึงตั้งเป็นสมมติฐานดังนี้

สมมติฐานที่ 1 (H1): ความไว้วางใจทางออนไลน์มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง

การรับรู้ความเสี่ยง (Perceived Risk: PR) หมายถึง ความสามารถ ความตระหนักรู้ ในการประเมินค่าระดับความเสี่ยงที่ผู้ใช้บริการต้องเผชิญเกี่ยวกับการใช้บริการชำระเงินด้วยคิวอาร์โค้ดที่อาจก่อให้เกิดความสูญเสียต่าง ๆ ต่อผู้ใช้บริการ ไม่ว่าจะเป็นความเสี่ยงด้านความปลอดภัย เช่น การอาชญากรรมทางไซเบอร์ ไม่จะเป็นการเปลี่ยนเส้นทางผู้ใช้ไปยังหน้าเว็บไซต์หลอกลวง การแก้ไขเลขที่บัญชีปลายทาง ความเสี่ยงด้านประสิทธิภาพของระบบ เช่น การทำงานผิดพลาด หรือความไม่แน่นอนของระบบ ความเสี่ยงด้านการสูญเสียเวลา เช่น กระบวนการรับชำระเงินใช้เวลานานมากกว่าปกติ ความเสี่ยงด้านการสูญเสียความเป็นส่วนตัวของข้อมูล งานวิจัยของ Heng และคณะ [10] กล่าวว่า ทศนคติเกี่ยวกับการรับรู้ความเสี่ยงจะส่งผลเชิงลบ และมีความสัมพันธ์กับความเชื่อมั่นและความตั้งใจกระทำ มีงานวิจัยของ Hankun และคณะ [8] แสดงให้เห็นว่าอิทธิพลของการรับรู้ความเสี่ยง การรับรู้ถึงความปลอดภัย ความไว้วางใจ การรับรู้ถึงความสะดวกในการใช้บริการของผู้ใช้มีผลต่อการชำระเงินออนไลน์ และงานวิจัยของ Nattawat [11] ยืนยันว่าการรับรู้ด้านความเสี่ยงเป็นปัจจัยที่สำคัญที่สุดที่ชักจูงให้คนไทยหันมาใช้บริการพร้อมเพย์ ดังนั้นผู้วิจัยจึงตั้งเป็นสมมติฐานดังนี้

สมมติฐานที่ 2 (H2): ความไว้วางใจทางออนไลน์มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่องผ่านการรับรู้ความเสี่ยง

สมมติฐานที่ 3 (H3): การรับรู้ความเสี่ยงมีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง

ความสามารถทางนวัตกรรมส่วนบุคคล (Personal Innovativeness : PI) หมายถึง ลักษณะบุคลิกภาพของ

ผู้ใช้บริการชำระเงินด้วยคิวอาร์โค้ดที่เต็มใจปรับเปลี่ยน และมีความตั้งใจที่จะทดลองใช้เทคโนโลยีใหม่ๆ สามารถรับความเสี่ยงและมีแนวโน้มที่จะใช้เทคโนโลยีใหม่ได้เร็วกว่าคนอื่น มีงานวิจัยของ Liébana-Cabanillas [12] พบว่าทัศนคติต่อนวัตกรรมเป็นปัจจัยกำหนดความตั้งใจในอนาคตที่จะใช้เทคโนโลยีการชำระเงินผ่านมาตรฐานคิวอาร์โค้ด งานวิจัยของรัชมัง วิริยะพันธ์ และ ปวีณา คำพุกกะ [13] พบว่าความสอดคล้องและนวัตกรรมส่วนบุคคลมีผลต่อความตั้งใจในการชำระเงินผ่านสมาร์ตโฟนในแต่ละยุคสมัย ดังนั้นผู้วิจัยจึงตั้งเป็นสมมติฐานดังนี้

สมมติฐานที่ 4 (H4): ความสามารถทางนวัตกรรมส่วนบุคคลมีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่องผ่านการรับรู้ความเสี่ยง

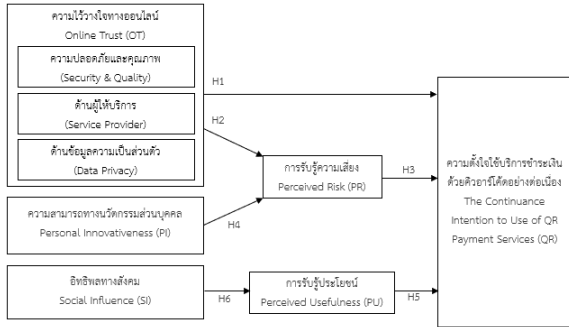
การรับรู้ประโยชน์ (Perceived Usefulness: PU) เป็นหนึ่งในปัจจัยสำคัญของแบบจำลองการยอมรับเทคโนโลยี (TAM) ที่คิดค้นโดย [14] การรับรู้ประโยชน์ หมายถึง การที่ผู้ใช้งานรับรู้ว่าการใช้บริการชำระเงินด้วยคิวอาร์โค้ดมีขั้นตอนที่ง่าย ไม่ซับซ้อน ชำระเงินได้รวดเร็ว ลดความผิดพลาด ตรวจสอบข้อมูลการชำระเงินได้ลดต้นทุนค่าใช้จ่ายพวกค่าธรรมเนียมในการชำระเงิน อีกทั้งชำระเงินได้ตลอด 24 ชั่วโมง อัจฉราวรรณ กองเพชร และคณะ [15] พบว่าการรับรู้ถึงประโยชน์และการรับรู้ถึงความง่ายในการใช้งานมีอิทธิพลต่อทัศนคติการใช้ระบบ และส่งผลต่อพฤติกรรมความตั้งใจใช้ระบบ (Behavioral Intention) และพฤติกรรมการยอมรับใช้งานจริง (Actual Usage Behavior) ผู้วิจัยจึงตั้งเป็นสมมติฐานดังนี้

สมมติฐานที่ 5 (H5): การรับรู้ประโยชน์มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง

อิทธิพลทางสังคม (Social Influence : SI) หมายถึง การที่พฤติกรรมของบุคคลหนึ่งมีผลต่อทัศนคติและพฤติกรรมของอีกบุคคลหนึ่ง และอาจทำให้เกิดการเปลี่ยนแปลงทัศนคติและพฤติกรรมให้เป็นไปในรูปแบบที่สังคมยอมรับจนส่งผลการตอบสนองเกิดเป็นแรงกระตุ้นให้เกิดความต้องการที่จะลอกเลียนแบบพฤติกรรมนั้น ๆ ทั้งทางตรงและทางอ้อม ผู้วิจัยจึงตั้งเป็นสมมติฐานดังนี้

สมมติฐานที่ 6 (H6): อิทธิพลทางสังคมมีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่องผ่านการรับรู้ประโยชน์

จากการศึกษางานวิจัยและทฤษฎีที่เกี่ยวข้อง ผู้วิจัยกำหนดเป็นกรอบแนวคิดของการวิจัย ดังแสดงในภาพที่ 1



ภาพที่ 1: แสดงกรอบแนวคิดของการวิจัย

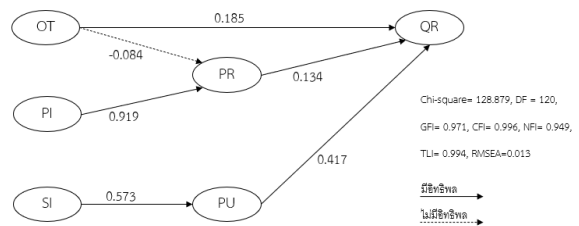
3. วิธีการดำเนินงานวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) โดยใช้แบบสอบถามเป็นเครื่องมือเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างระหว่างมิถุนายน - เดือนกรกฎาคม 2565 ประชากรคือกลุ่มประชาชนผู้ใช้ทั่วไปที่อาศัยอยู่ในประเทศไทย มีประสบการณ์การใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยใช้โปรแกรม G*Power ในการคำนวณขนาดของกลุ่มตัวอย่าง กำหนดค่าเพาเวอร์ (1-β) เท่ากับ 0.95 ค่าอัลฟา (α) เท่ากับ 0.05 จำนวนตัวแปรทำนาย 6 ตัวแปร ขนาดของอิทธิพล (Effect Size) เท่ากับ 0.03 ผลที่ได้คือ ขนาดของกลุ่มตัวอย่างเท่ากับ 363 ตัวอย่าง และเพิ่มการเก็บข้อมูลที่ใช้สำหรับการศึกษานำร่อง (Pilot study) เพื่อวัดความเที่ยงอีก 30 ตัวอย่าง รวมเป็น 393 ตัวอย่าง

ผู้วิจัยได้ทำการทดสอบสมมติฐานของงานวิจัยโดยใช้โปรแกรม SPSS และ AMOS ในการวิเคราะห์รูปแบบเชิงโครงสร้าง (Structural Equation Modeling: SEM) เพื่อวิเคราะห์หาค่าอิทธิพลทั้งทางตรงและทางอ้อม และได้มีการปรับรูปแบบของโมเดล เพื่อทดสอบความสอดคล้องของโมเดลกับกับข้อมูลเชิงประจักษ์ (Model Fit)

4. ผลการวิจัย

การวิจัยใช้การวิเคราะห์รูปแบบเชิงโครงสร้าง โดยเป็นการวิเคราะห์เส้นทาง (Path Analysis) เพื่อหาค่าอิทธิพลทั้งทางตรงและทางอ้อมของตัวแปรอิสระว่ามีผลต่อตัวแปรตามมากน้อยเพียงใด เมื่อนำแบบสอบถามมาตรวจสอบความเชื่อมั่น (Reliability) ได้ค่าสัมประสิทธิ์แอลฟาของครอนบาคแต่ละปัจจัยเป็น 0.84 การวัดความเที่ยงตรง (Validity) โดยใช้ผู้เชี่ยวชาญ 3 คนมาประเมินแล้วหาค่าเฉลี่ยได้เกินกว่า 0.66 ทุกคำถาม และได้ค่าความสอดคล้องของโมเดล (Goodness-of-Fit) ดังนี้ P-value=0.273, $\chi^2/df=1.074$, RMR=0.009, RMSEA=0.013, GFI=0.971, CFI=0.996, NFI=0.949 และ TLI= 0.994 สามารถแสดงผลสัมประสิทธิ์โมเดลการวิจัยในภาพที่ 2 และแสดงผลการวิเคราะห์เชิงสาเหตุของโมเดลในตารางที่ 1



ภาพที่ 2: สัมประสิทธิ์เส้นทางมาตรฐานของโมเดล

ตารางที่ 1: ผลการวิเคราะห์เชิงสาเหตุของโมเดล

ตัวแปรผล	อิทธิพล	ตัวแปรเชิงสาเหตุ				
		OT	PI	SI	PR	PU
PR	DE	-0.084	0.919***	-	-	-
	IE	-	-	-	-	-
PU	DE	-	-	0.573***	-	-
	IE	-	-	-	-	-
QR	DE	0.185**	-	-	0.134*	0.417***
	IE	-0.011*	0.123	0.239*	-	-

หมายเหตุ *p<0.05, **p<0.01, ***p<0.001

ผลที่ได้จากการวิเคราะห์อิทธิพลเชิงสาเหตุภายในโมเดล สามารถนำมาสรุปผลการทดสอบของสมมติฐานต่างๆ ที่ตั้งไว้มีรายละเอียดที่แสดงในตารางที่ 2 ดังนี้

ตารางที่ 2: สรุปผลการทดสอบสมมติฐาน

สมมติฐาน	ค่า อิทธิพล	ค่า P	ผลการ ทดสอบ
H1: ความไว้วางใจทางออนไลน์ (OT) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR)	0.185	0.005	ยอมรับ
H2: ความไว้วางใจทางออนไลน์ (OT) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ผ่านการรับรู้ความเสี่ยง (PR)	-0.011	0.015	ปฏิเสธ
H3: การรับรู้ความเสี่ยง (PR) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR)	0.134	0.049	ยอมรับ
H4: ความสามารถทางนวัตกรรมส่วนบุคคล (PI) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ผ่านการรับรู้ความเสี่ยง (PR)	0.123	0.064	ปฏิเสธ
H5: การรับรู้ประโยชน์ (PU) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR)	0.417	<0.001	ยอมรับ
H6: อิทธิพลทางสังคม (SI) มีอิทธิพลเชิงบวกต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ผ่านการรับรู้ประโยชน์ (PU)	0.239	0.019	ยอมรับ

5. สรุป

ผลการศึกษาปัจจัยที่มีผลต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง นำมาอภิปรายผลในประเด็นต่าง ๆ ดังนี้

ความไว้วางใจทางออนไลน์ (OT) มีอิทธิพลเชิงบวกทางตรงต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้เป็นเพราะว่าผู้ใช้บริการเกิดความเชื่อมั่นว่าบริการชำระหนี้ด้วยคิวอาร์โค้ด มีความน่าเชื่อถือ ปลอดภัย และถูกต้อง มีความไว้วางใจในการให้บริการบนสารสนเทศที่มีคุณภาพของผู้พัฒนา อีกทั้งผู้ให้บริการยังมีความซื่อสัตย์และมีกลไกการพัฒนาระบบชำระหนี้ที่น่าเชื่อถือ มีการป้องกันและไม่เปิดเผยข้อมูลความเป็นส่วนตัวของผู้บริโภค จึงทำให้เกิดความตั้งใจใช้บริการ ซึ่งสอดคล้องกับงานวิจัยของ Kim และคณะ [7] ว่าหากผู้บริโภคเกิดความไว้วางใจ และความเชื่อมั่นต่อผู้ให้บริการในระดับสูง จะส่งผลให้ผู้บริโภคเกิดความตั้งใจซื้อสินค้า และงานวิจัยของ McAllister [9] ว่าความไว้วางใจพื้นฐานเกิดจากการรับรู้ของผู้บริโภคที่มีต่อคุณสมบัติและลักษณะของผู้ขายออนไลน์ จากผลวิจัยที่ได้ ผู้วิจัยคิดว่าหากผู้พัฒนาและผู้ให้บริการสร้างคอนเทนต์และสื่อสารให้ความรู้ตามช่องทางต่าง ๆ สม่ำเสมอ

จะสร้างความเชื่อมั่น ของผู้ใช้บริการ ซึ่งจะทำให้มีผู้ใช้บริการเพิ่มมากขึ้น

การรับรู้ความเสี่ยง (PR) มีอิทธิพลเชิงบวกทางตรงต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้เป็นเพราะการที่ผู้ใช้บริการรับรู้ถึงความเสี่ยงในด้านต่าง ๆ ที่อาจเกิดทำให้ผู้ใช้บริการมีความเข้าใจและใช้งานบริการด้วยความระมัดระวังยิ่งขึ้น และมีความกล้าที่จะใช้ ซึ่งสอดคล้องกับงานวิจัยของ Hankun และคณะ [8] ที่ว่าการรับรู้ความเสี่ยงและความไว้วางใจมีผลต่อการการชำระหนี้ออนไลน์ และงานวิจัยของ Nattawat [11] ว่าการรับรู้ความเสี่ยงเป็นปัจจัยที่สำคัญที่สุดที่ชักจูงให้คนไทยหันมาใช้บริการพร้อมเพย์ จากผลวิจัยที่ได้หากผู้พัฒนาและผู้ให้บริการ รวมถึงภาครัฐที่คอยกำกับดูแล มีการให้ความรู้ วิธีการสังเกตและป้องกันความเสี่ยงที่อาจเกิดขึ้นในการใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดที่เข้าใจง่ายก็จะทำให้มีผู้ใช้บริการเพิ่มมากขึ้น

อย่างไรก็ตามพบว่า ความไว้วางใจทางออนไลน์ มีอิทธิพลเชิงลบต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่องโดยผ่านการรับรู้ความเสี่ยงอย่างมีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่สอดคล้องกับงานวิจัยของ Heng และคณะ [10] ที่ว่าการรับรู้ความเสี่ยงมีอิทธิพลเชิงลบต่อความตั้งใจใช้ ทั้งนี้อาจเพราะผู้บริโภครับรู้ถึงความเสี่ยงด้านต่าง ๆ มากมาย ทำให้ความไว้วางใจที่จะใช้บริการมีน้อยลง ยิ่งหากไม่รู้ถึงวิธีป้องกันการถูกโจรกรรม จะส่งผลให้ความสนใจในการใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดลดลง ดังนั้นผู้วิจัยแนะนำว่าควรมีการให้ข้อมูลความเสี่ยงที่เข้าใจง่าย ให้ความรู้และวิธีป้องกันตนเองจากการถูกโจรกรรม เพื่อสร้างความมั่นใจ ความไว้วางใจ ความปลอดภัยให้กับผู้ใช้บริการ

ความสามารถทางนวัตกรรมส่วนบุคคล (PI) ไม่มีอิทธิพลเชิงบวกทางอ้อมต่อความตั้งใจใช้บริการชำระหนี้ด้วยคิวอาร์โค้ดอย่างต่อเนื่อง (QR) ผ่านการรับรู้ความเสี่ยง (PR) ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ แต่จากผลวิจัยที่ได้ ความสามารถทางนวัตกรรมส่วนบุคคลมีอิทธิพลเชิงบวกทางตรงต่อการรับรู้ความเสี่ยง ทำให้การรับรู้ถึงความเสี่ยงของผู้ใช้บริการจะขึ้นอยู่กับระดับความรู้

ความสามารถทางนวัตกรรมของแต่ละบุคคล และจากที่กล่าวมาว่าการรับรู้ความเสี่ยงมีอิทธิพลเชิงบวกทางตรงต่อความตั้งใจใช้บริการอย่างมีนัยสำคัญ อย่างไรก็ตามแม้ว่าความสามารถทางนวัตกรรมส่วนบุคคลจะไม่ได้ส่งผลทางตรงต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง แต่หากผู้ใช้บริการเป็นผู้ที่มีความสามารถทางนวัตกรรม ก็จะทำให้รับรู้และระมัดระวังความเสี่ยงได้ดีขึ้นและส่งผลให้เกิดการใช้งานเพิ่มขึ้นได้ ดังนั้นผู้ให้บริการควรส่งเสริมการให้ความรู้ทางด้านนวัตกรรม เพื่อให้ผู้ใช้รับมือกับความเสี่ยง ได้ดียิ่งขึ้น และเกิดความไว้วางใจที่จะใช้บริการชำระเงินด้วยคิวอาร์โค้ด

การรับรู้ประโยชน์ (PU) มีอิทธิพลเชิงบวกทางตรงต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้เป็นเพราะการที่ผู้ใช้บริการรับรู้ถึงประโยชน์ที่จะได้รับจากการใช้บริการชำระเงินด้วยคิวอาร์โค้ด ไม่ว่าจะเป็นการใช้งานที่ง่ายขึ้นตอนที่ยังไม่ซัฟชั่น รวดเร็ว ช่วยลดความผิดพลาดในการกรอกข้อมูลบัญชีปลายทาง และสามารถตรวจสอบข้อมูลก่อนการชำระเงินได้ ทำให้ผู้บริโภคหันมาใช้บริการมากขึ้น สอดคล้องกับงานวิจัยของอัจฉราวรรณ กองเพชร และคณะ [15] ที่กล่าวไว้ว่าการรับรู้ถึงประโยชน์และการรับรู้ถึงความง่ายในการใช้งานเป็นปัจจัยหลักที่มีอิทธิพลต่อทัศนคติการใช้ระบบ

อิทธิพลทางสังคม (SI) มีอิทธิพลเชิงบวกทางอ้อมต่อความตั้งใจใช้บริการชำระเงินด้วยคิวอาร์โค้ดอย่างต่อเนื่อง โดยผ่านการรับรู้ประโยชน์ ซึ่งสอดคล้องกับสมมติฐานที่ตั้งไว้ ทั้งนี้เป็นเพราะการที่ร้านค้ามีบริการชำระเงินด้วยคิวอาร์โค้ด ในช่วงการแพร่ระบาดโควิด-19 ผู้คนส่วนใหญ่ต่างหันมาใช้กระเป๋าเงินดิจิทัลมากขึ้น เพื่อลดความเสี่ยงในการติดเชื้อ อีกทั้งทางภาครัฐได้คิดมาตรการต่างๆ เช่น โครงการคนละครึ่ง ทำให้มีผู้บริโภคหันมาใช้บริการชำระเงินด้วยคิวอาร์โค้ดมากกว่าปกติ จากคนที่ไม่เคยใช้บริการที่ต่างสมัครบริการเพื่อเข้าร่วมโครงการในช่วงที่ต้องรับมือกับสถานะการเศรษฐกิจที่ฝืดเคือง จากผลวิจัยที่ได้ ผู้วิจัยคิดว่าหากหน่วยงานภาครัฐ ภาคธุรกิจ สถาบันการเงิน ต้องการผลักดันให้ผู้ใช้บริการหันมาเปลี่ยนวิธีการชำระ

เงินเป็นรูปแบบอิเล็กทรอนิกส์ ควรมีการบอกประโยชน์ที่หากผู้ใช้บริการหันมาใช้บริการชำระเงินด้วยคิวอาร์โค้ด แทนการใช้บริการชำระเงินรูปแบบเดิม และหากหน่วยงานมีการจัดทำโครงการเพื่อสนับสนุน และมีความตั้งใจให้ใช้งาน เช่น มีโปรโมชันส่วนลด ไม่มีค่าธรรมเนียม หรืออาจจะจัดให้มีการชำระเงินด้วยคิวอาร์โค้ดเท่านั้นในงานมหกรรมซื้อขายสินค้า เพื่อกระตุ้นให้ผู้บริโภคทดลองใช้ และสัมผัสถึงประโยชน์จากการใช้

ข้อเสนอแนะสำหรับงานวิจัยในครั้งถัดไป ควรใช้การวิจัยแบบผสมผสาน เพิ่มวิธีเก็บข้อมูลด้วยการสัมภาษณ์ซึ่งจะทำให้ได้ความคิดเห็นจากผู้ใช้บริการในมุมมองที่กว้าง และรับรู้ถึงความรู้สึกของผู้ใช้บริการมากขึ้น ซึ่งจะช่วยให้เข้าใจในความสัมพันธ์ระหว่างปัจจัยชัดเจนยิ่งขึ้น

เอกสารอ้างอิง

- [1] ฝ่ายนโยบายระบบการชำระเงิน “PromptPay the Game Changer for Payments,” [ออนไลน์], BOT พระสยาม MAGAZINE (Electronic), 2563, เล่มที่ 3, หน้า 19-22. Available: https://www.bot.or.th/Thai/BOTMagazine/Pages/256303_TheKnowledge_PromptPay.aspx
- [2] <https://www.krungsri.com/th/plearn-plearn/promptpay-gateway-to-cashless-society>
- [3] ทีมพัฒนาและวิเคราะห์ข้อมูล ฝ่ายนโยบายระบบการชำระเงิน ธนาคารแห่งประเทศไทย, 2564. “Executive Summary เดือนพฤศจิกายน 2564,” [ออนไลน์]. Available: https://www.bot.or.th/Thai/PaymentSystems/Publication/payment_data_indicators/Documents/Payment%20Data%20Indicators_BOT%20Website_Nov2021_TH.pdf
- [4] Bi-monthly PAYMENTINSIGHT VOL 11/ ตุลาคม 2564. “พฤติกรรมกรรมการชำระเงินผ่าน Payment Diary,” [ออนไลน์]. Available: https://www.bot.or.th/Thai/PaymentSystems/Publication/payment_insight/Documents/Bi-monthly_report_Vol11-2021_October.pdf
- [5] Fishbein, M., and Ajzen, I. (2010). Predicting and changing behavior: the reasoned action approach. New York: Psychology Press.
- [6] Ajzen, and Fishbein. (1980). Understanding Attitudes and Predicting Social Behavior. New Jersey: Prentice Hall Inc
- [7] Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce:

- The role of trust, perceived risk, and their antecedents. *Decision support systems*, 44(2), 544-564.
- [8] Hankun, H., Yafang, L., Xuemei, H., and Jing, F.. “A comparative study of China and US users' acceptance of online payment,” in 2016 13th International Conference on Service Systems and Service Management (ICSSSM), Kunming, China. Available: IEEE Xplore (11 August 2016)
- [9] McAllister, D. J. (1995). Affect and cognitive based trust as foundation for interpersonal cooperation in organizations. *Academy of Management Journal*, 38(1), 24-59.
- [10] Heng, X., Hock-Hai, T., & Bernard, C.Y.T. (2005). Predicting the Adoption of Location-Based Services: The Role of Trust and Perceived Privacy Risk. Unpublished research report, Department of Information Systems, National University of Singapore.
- [11] Nattawat Poopattanapong, 2017. “Factors that Influence Thai People Using Promptpay,” [ออนไลน์]. Available: http://ethesisarchive.library.tu.ac.th/thesis/2017/TU_2017_5902040145_8391_9370.pdf
- [12] F. Liébana-Cabanillas. (2015). “User behaviour in QR mobile payment system: the QR Payment Acceptance Model” [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/09537325.2015.1047757>,
- [13] รังมั่ง วิริยะพันธ์, ปวีณา คำพุกกะ. (2560). อิทธิพลของความสะดวกต้อง และนวัตกรรมส่วนบุคคลที่มีต่อความตั้งใจในการชำระเงินผ่านสมาร์ตโฟน. วารสารวิชาการมหาวิทยาลัยฟาร์อีส-เทอร์น. Vol 11 No 3 (2017) : กรกฎาคม 2560 - กันยายน 2560
- [14] Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003.
- [15] อัจฉรวรรณ กองเพชร, ธัญญา เวียรร์รา, และอรพรรณ คงมาลัย. (2555). “การยอมรับและใช้งานระบบ SAP BI: กรณีศึกษาการบริหารงานจัดซื้อจัดจ้างในองค์กรไม่แสวงหากำไร”. วิทยานิพนธ์ วท.ม. (การบริหารเทคโนโลยี). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์.

ปัจจัยที่ส่งผลต่อการรับรู้ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้ บุคคลธรรมดาผ่านทางอินเทอร์เน็ต

Factory Affecting User's Perceive Benefits of Personal Income Tax Payment Through Internet

อภิสร่า อ่อนท้าว (Apisara Onthao)¹ และชนพล เจนสุทธิเวชกุล (Tanapon Jentsuttiwetchakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857182@email.kmutnb.ac.th, ²tanapon.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อการรับรู้ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต โดยนำทฤษฎีความสำเร็จของระบบสารสนเทศ (ISSM Model) มาใช้ร่วมกับทฤษฎีความกังวลในเทคโนโลยี งานวิจัยนี้เก็บข้อมูลโดยใช้แบบสอบถามออนไลน์ กลุ่มตัวอย่างคือผู้เสียภาษี จำนวนตัวอย่าง 359 คน จากผลการวิจัย พบว่าคุณภาพระบบ คุณภาพข้อมูล คุณภาพบริการ ความกังวลต่อเทคโนโลยี มีผลต่อ ความพึงพอใจของผู้ใช้งาน และนำไปสู่ การรับรู้ประโยชน์ของระบบ โดยค่า R^2 แสดงให้เห็นว่าแบบจำลองนี้สามารถอธิบายปัจจัยที่ส่งผลต่อความพึงพอใจ และการรับรู้ประโยชน์ของระบบ ได้ 71.8% และ 66.4% ตามลำดับ

คำสำคัญ: กรมสรรพากร ความพึงพอใจ ปัจจัยเทคโนโลยี

Abstract

The objective of this research is to study factors affecting perceive benefits of income tax payment through Internet. IS Success Model (ISSM) and technology anxiety were used to develop research model. An online questionnaire was used in data collection. Moreover, 359 samples were used in statistical analysis. The results indicated that system quality, data quality, service quality and technology anxiety have significantly influenced on

user's satisfaction which affecting on perceived benefits. Moreover, R^2 indicated that these factors can explain user's satisfaction and perceived benefits about 71.8% and 66.4% respectively.

Keywords: The Revenue Department, Satisfaction, IS Success Model (ISSM), Technology Anxiety.

1. บทนำ

ในปีพ.ศ 2563 มีจำนวนผู้เสียภาษีที่ยื่นแบบและชำระภาษีด้วยกระดาษทั่วประเทศจำนวน 1,265,010 คน ซึ่งกรมการบางตามประเภทภาษีเงินได้บุคคลธรรมดา จำนวนผู้เสียภาษีที่ใช้กระดาษ 501,568 คน ภาษีมูลค่าเพิ่มจำนวนผู้เสียภาษีที่ใช้กระดาษ 702,899 คน และอื่น ๆ 60,543 คน ปัจจุบันกรมสรรพากรจึงได้อำนวยความสะดวกให้ประชาชนสามารถยื่นแบบและชำระภาษีผ่านอินเทอร์เน็ต แทนการกรอกแบบและยื่นแบบฯ ด้วยกระดาษ พร้อมชำระภาษีผ่านช่องทางอิเล็กทรอนิกส์ เพื่อความสะดวกมากยิ่งขึ้น

งานวิจัยนี้จึงต้องการที่จะศึกษาปัจจัยที่ส่งผลต่อการรับรู้ประโยชน์ ความพึงพอใจ และความวิตกกังวลในการใช้ระบบ ดังกล่าว โดยนำทฤษฎีความสำเร็จของระบบสารสนเทศ (ISSM Model) มาใช้ร่วมกับทฤษฎีความกังวลในเทคโนโลยี

โดยมุ่งหวังว่าผลที่ได้จากงานวิจัยนี้จะช่วยให้เข้าใจผู้เสียภาษีที่เคยใช้เว็บไซต์การชำระหนี้ได้บุคคลธรรมดาที่มากขึ้นนั้นมีประโยชน์และเกิดปัญหาในการใช้งานอย่างไร เป็นแนวทางในการพัฒนาระบบให้ตอบโจทย์ความต้องการของผู้ใช้มากขึ้นเพื่อค้นหาปัจจัยที่ส่งผลกระทบต่อการใช้บริการเว็บไซต์การชำระหนี้ได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต จากกลุ่มตัวอย่างด้วยแบบสอบถามที่จัดทำขึ้น

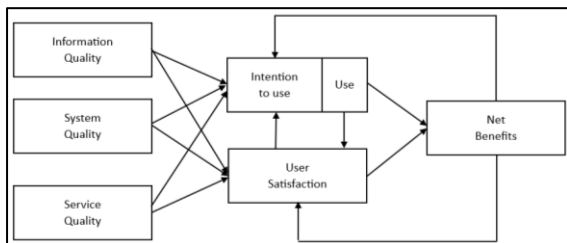
2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดเกี่ยวกับระบบยื่นแบบผ่านอินเทอร์เน็ต

การให้บริการยื่นแบบและชำระภาษี ผ่านอินเทอร์เน็ต โดยผู้เสียภาษีเป็นผู้ป้อนข้อมูลที่จะต้องกรอกในแบบแสดงรายการภาษีที่ได้รับอนุญาตให้ยื่นผ่านทางอินเทอร์เน็ตแทนการกรอกแบบแสดงรายการภาษีและยื่นแบบด้วยกระดาษ พร้อมชำระภาษีโดยใช้คำสั่งทางอิเล็กทรอนิกส์ให้ธนาคาร โอนเงินค่าภาษีให้กรมสรรพากร ซึ่งเป็นทางเลือกใหม่อีกทางหนึ่งที่ไม่บังคับ แต่ช่วยให้ผู้เสียภาษีที่สนใจ เลือกใช้บริการเพื่อความสะดวกมากยิ่งขึ้น ในขณะที่การยื่นแบบฯ และชำระภาษีที่สำนักงานสรรพากรพื้นที่สาขายังคงมีอยู่ตามปกติ [1]

2.2 ทฤษฎีโมเดลแห่งความสำเร็จระบบสารสนเทศของ DeLone และ McLean

ทฤษฎีทางด้านความสำเร็จของระบบสารสนเทศ (IS Success Model) ของ DeLone และ McLean ได้ทำการศึกษาวิจัยในบริบทของการใช้บริการเว็บไซต์การชำระภาษีเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต [3]



ภาพที่ 1 : ทฤษฎีโมเดลของ DeLone และ McLean

2.2.1 ปัจจัยด้านคุณภาพระบบ (System Quality)

การวัดคุณภาพระบบมีตัวแปรที่เกี่ยวข้อง ดังนี้ ความง่ายในการใช้งาน (Ease-of-use) ระบบต้องมีความเสถียรภาพ (Reliability) ความรวดเร็วในการตอบสนอง (Response time) ความพร้อมในการใช้งาน (Availability) ความสะดวกและใช้งานได้ทุกที่ทุกเวลา (Convenience of access) มีความสามารถในการปรับระบบ (Adaptability)[4]

2.2.2 ปัจจัยด้านคุณภาพของข้อมูล (Information Quality)

การวัดคุณภาพของข้อมูลที่ถูกส่งออกมาจากระบบ โดยผู้ใช้จะวัดจากการใช้งานข้อมูล ในการวัดคุณภาพของข้อมูลมีองค์ประกอบที่สำคัญ ได้แก่ ความสมบูรณ์และถูกต้อง (Completeness) ความง่ายในการเข้าใจ (Ease of understanding) ข้อมูลมีความเป็นส่วนบุคคล(Personalization) ความสัมพันธ์ของข้อมูล (Relevance) ข้อมูลมีความปลอดภัย (Security)

2.2.3 ปัจจัยด้านคุณภาพบริการ (Service Quality)

การวัดเกี่ยวกับความสำเร็จของระบบสารสนเทศว่า คุณภาพในการบริการอิเล็กทรอนิกส์ หมายถึงการรับรู้ของผู้ใช้บริการทางด้านคุณภาพที่ส่งผลให้การพาณิชย์อิเล็กทรอนิกส์ประสบความสำเร็จและมีผลกับการเพิ่มขึ้นและลดลงของผู้ใช้บริการ มีปัจจัยในการวัดด้านคุณภาพบริการ

2.2.4 ความตั้งใจในการใช้ (Intention to Use) หรือการใช้งาน (Use)

การวัดผลจากการใช้งานระบบ (Use) จะมีประสิทธิภาพและสัมฤทธิ์ผลต่อเมื่อผู้ใช้งานมีความสมัครใจ และสามารถวัดผลได้จาก ความถี่ในการใช้งาน เวลาในการใช้งาน จำนวนครั้งในการเข้าถึงระบบ และรูปแบบในการใช้งาน เป็นต้น ซึ่งมีความเกี่ยวข้องกับผลกระทบส่วนบุคคล (Individual Impacts) ในแง่ของประสิทธิภาพในการ ทำงาน (Job performance) และประสิทธิภาพในการตัดสินใจ (Decision making performance) ตัววัดความสำเร็จของระบบ E-Commerce

2.2.5 ความพึงพอใจของผู้ใช้งาน (User satisfaction)

ความพึงพอใจเป็นปัจจัยที่สำคัญต่อความสำเร็จของระบบสารสนเทศ ซึ่งเกิดจากการที่ผู้ใช้บริการได้รับการตอบสนองต่อความต้องการได้สำเร็จตามความคาดหวังในทฤษฎีแบบจำลองความสำเร็จของระบบสารสนเทศ

2.2.6 การรับรู้ประโยชน์ (Net Benefits)

การวัดระดับที่ผู้ใช้งานเชื่อมั่นว่าการใช้ระบบสารสนเทศจะส่งผลให้ประสิทธิภาพของงานเพิ่มขึ้น

2.3 แนวคิดและทฤษฎีที่เกี่ยวกับความพึงพอใจ

เป็นความรู้ของผู้บริโภคที่เพิ่มขึ้นในระยะเวลาและระยะยาวที่มีต่อคุณภาพการบริการต่าง ๆ ทั้งในระดับแคบที่เกี่ยวกับลักษณะบริการที่มีต่อคุณภาพการบริการ และในระดับกว้างที่เป็นมุมมองของผู้รับบริการที่ได้จากบริการทุกประเภทที่นำไปเป็นข้อสรุปรวมความพึงพอใจของผู้ใช้บริการ[2]

2.4 แนวคิดทฤษฎีความกังวลต่อเทคโนโลยี (Technology Anxiety)

เป็นแนวคิดที่ศึกษาเทคโนโลยี และศึกษาปัจจัยที่คาดว่าจะมีอิทธิพลต่อความสนใจในการใช้เทคโนโลยีใหม่ๆ โดยจากแนวความคิดจากเทคโนโลยี มีการกล่าวถึงความกังวลในการใช้งานคอมพิวเตอร์ โดยให้นิยามว่า คือ ความกลัว ความหวาดหวั่น และความหวังที่คนจะรู้สึกเมื่อพิจารณาที่จะใช้ หรือใช้เทคโนโลยีคอมพิวเตอร์จริง โดยความกังวลดังกล่าว คือความประหม่าในการใช้คอมพิวเตอร์ ความคิดเห็นในเชิงลบต่อการใช้คอมพิวเตอร์ ความพยายามที่จะลดระยะเวลาในการใช้คอมพิวเตอร์และรวมถึงการหลีกเลี่ยงการใช้คอมพิวเตอร์ด้วย[5]

2.5 การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis)

การวิเคราะห์ตัวประกอบ (Factor Analysis) หรือ การวิเคราะห์ปัจจัยเป็นเทคนิคที่จับกลุ่มหรือรวมตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มหรือปัจจัยเดียวกัน ตัวแปรที่อยู่ในปัจจัยเดียวกันจะมีความสัมพันธ์กันมาก โดย

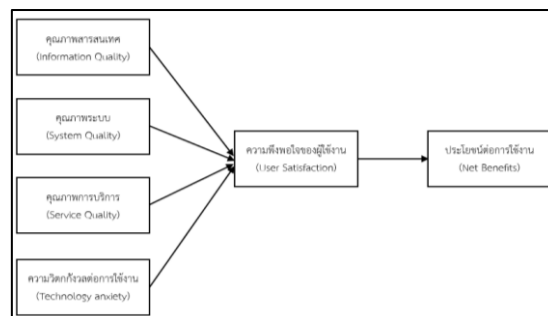
ความสัมพันธ์นั้นอาจจะเป็นในทิศทางบวก (ไปในทิศทางเดียวกัน) หรือทิศทางลบ (ไปในทางตรงกันข้าม) ส่วนตัวแปรต่างองค์ประกอบกัน จะไม่มีความสัมพันธ์

2.6 การวิเคราะห์การถดถอยพหุคูณ (Multiple Linear Regression)

การวิเคราะห์การถดถอยเป็นวิธีการทางสถิติที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variable) กับตัวแปรตาม (Dependent Variable) จะเป็นการศึกษาความสัมพันธ์เชิงเส้นตรง (Linearity) ถ้าศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหนึ่งตัวกับตัวแปรตามหนึ่งตัว เรียกว่า การวิเคราะห์ถดถอยเชิงเส้นเชิงเดียว หรือการวิเคราะห์ถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Analysis) ถ้าตัวแปรอิสระมีมากกว่าหนึ่งตัวกับตัวแปรตามหนึ่งตัว เรียกว่า การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)[6]

3. วิธีการดำเนินงาน

กรอบแนวความคิดการดำเนินงานวิจัยปัจจัยที่ส่งผลต่อการรับรู้ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต

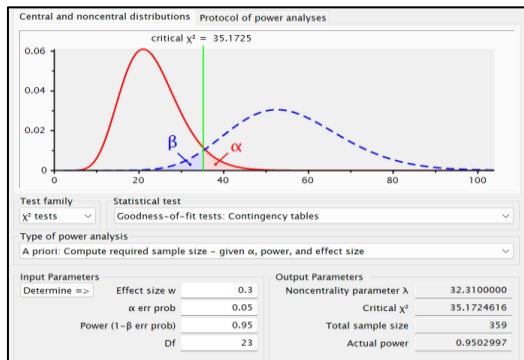


ภาพที่ 2 : กรอบแนวความคิดการดำเนินงานวิจัย

3.1 การกำหนดประชากรกลุ่มตัวอย่าง

ผู้วิจัยได้กำหนดกลุ่มตัวอย่างด้วยใช้โปรแกรม G*Power 3.1 (Erdfelder et al., 2007) โดยกำหนดค่า effect size เท่ากับ 0.3 หมายถึงมีขนาดอิทธิพลในระดับกลาง (medium) ซึ่งส่วนใหญ่ขึ้นอยู่กับ effect size เท่ากับ 0.3 ที่ระดับนัยสำคัญทางสถิติ เท่ากับ 0.05 ค่าอำนาจทดสอบ (power of test) เท่ากับ 0.08 (Hair et al.,

2010) ค่าองศาอิสระ ของตัวแปร (DF) เท่ากับ 14 ที่นำมา จากข้อคำถามจากแบบสอบถามปัจจัยที่ส่งผลต่อการรับรู้ ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้ บุคคลธรรมดาผ่านทางอินเทอร์เน็ตจากโปรแกรม G*Power ได้ขนาดกลุ่มตัวอย่างเท่ากับ 359 คน



ภาพที่ 3 : กำหนดกลุ่มตัวอย่างด้วยโปรแกรม G*Power

3.2 การเก็บรวบรวมข้อมูล

3.2.1 สร้างเครื่องมือแบบสอบถามทางผู้วิจัยได้ สร้างเครื่องมือแบบสอบถามจากการศึกษาทฤษฎีและ งานวิจัยที่เกี่ยวข้องต่าง ๆ ที่สามารถประยุกต์ใช้กับ งานวิจัยของผู้วิจัย

3.2.2 นำแบบสอบถามที่ผ่านการตรวจสอบไป สํารวจกลุ่มตัวอย่างเพื่อหาปัจจัยที่ส่งผลต่อการรับรู้ ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้ บุคคลธรรมดาผ่านทางอินเทอร์เน็ต

3.2.3 ทำการรวบรวมข้อมูลที่ได้จาก แบบสอบถามที่นำมาจากกลุ่มตัวอย่าง เพื่อนำมา วิเคราะห์หาปัจจัยที่ส่งผลต่อการรับรู้ประโยชน์ในการ ใช้ บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทาง อินเทอร์เน็ต

3.3 วิธีการวิเคราะห์ข้อมูล

ผู้วิจัยได้นำข้อมูลจำนวนรายการที่ผ่านการตรวจสอบ ความถูกต้องและความสมบูรณ์ของข้อมูลเป็นที่เรียบร้อย แล้วนำข้อมูลที่ได้นำมาทำการแยกหมวดหมู่เพื่อวิเคราะห์ โดยทำการวิเคราะห์ข้อมูลด้วยโปรแกรมประมวลผล สำเร็จรูปทางสถิติ SPSS เพื่อหาคำตอบของวัตถุประสงค์ และ สมมติฐาน โดยสถิติที่ใช้ในการวิเคราะห์ข้อมูล ดังนี้

3.3.1 นำข้อมูลทั่วไปของประชากรตัวอย่างมา วิเคราะห์ โดยใช้สถิติเชิงพรรณนา (Descriptive

Statistics) ได้แก่ วิเคราะห์ข้อมูลโดยการ ใช้การแจกแจง ความถี่ (Frequency)

3.3.2 นำข้อมูลตัวแปรมาวิเคราะห์องค์ประกอบ เชิงสํารวจ (EFA) เพื่อตรวจสอบตัวแปรและความ เหมาะสมของตัวแปร

3.3.3 นำตัวแปรแต่ละด้านมาทดสอบความ เชื่อมั่นของแบบสอบถาม เพื่อตรวจสอบความน่าเชื่อถือ ของแบบสอบถาม

3.3.4 นำข้อมูลตัวแปรมาวิเคราะห์การถดถอย พหุคูณ (Multiple Regression Analysis) เพื่อหา ความสัมพันธ์ระหว่างตัวแปร และหาความน่าเชื่อถือของ โมเดลที่ระดับนัยสําคัญทางสถิติที่ 0.05

4. ผลการดำเนินงาน

4.1 ผลการตรวจสอบความเหมาะสมของข้อมูล

ตัวแปรที่นำมาใช้ในการวิเคราะห์องค์ประกอบปัจจัย ที่ส่งผลต่อการรับรู้ประโยชน์ในการใช้บริการเว็บไซต์ การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ตนั้น มี ค่า Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) เท่ากับ 0.956 ซึ่งมีค่ามากกว่า 0.50 แสดงว่า ตัว แปรมีความสัมพันธ์กัน ในขนาดที่เหมาะสมในการนำไป วิเคราะห์เชิงสํารวจได้ และเมื่อทดสอบนัยสําคัญทาง สถิติด้วย Bartlett's Test of Sphericity พบว่า มีนัยสําคัญ ทางสถิติที่ระดับ 0.05 แสดงว่าเมทริกซ์สหสัมพันธ์ที่ได้นั้นไม่เป็นเมทริกซ์เอกลักษณะ ข้อมูลมีความเหมาะสมใน การวิเคราะห์องค์ประกอบได้

ตารางที่ 1 : KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.956
Bartlett's Test of Sphericity	Approx. Chi-Square	7939.381
	df	253
	sig.	.000

4.2 ผลค่าน้ำหนักองค์ประกอบ

Rotated Factor Matrix พบว่าค่าน้ำหนักองค์ประกอบ (Factor loading) ที่มีค่ามากกว่า 0.45 มีทั้งหมด 6 องค์ประกอบ โดยองค์ประกอบที่ 1 มีจำนวนตัวแปรมากที่สุด 5 ตัว องค์ประกอบที่ 2, 3, 4 และองค์ประกอบที่ 6 มี ตัวแปร จำนวน 3 ตัวแปร และองค์ประกอบที่ 5 มีจำนวน ตัวแปรน้อยที่สุด 2 ตัวแปร ตามตารางดังนี้

ตารางที่ 2 : Rotated Factor Matrix

	Factor					
	1	2	3	4	5	6
Benefits2	0.680					
Benefits4	0.636					
Benefits1	0.629					
Benefits3	0.619					
Benefits5	0.549					
Service2		0.818				
Service1		0.667				
Service3		0.584				
Satisfaction2			0.718			
Satisfaction3			0.637			
Satisfaction1			0.506			
System1				0.694		
System2				0.693		
System3				0.472		
Information2					0.877	
Information3					0.471	
Anxiety1						0.886
Anxiety3						0.869
Anxiety2						0.849

4.3 ผลจากการทดสอบความเชื่อมั่นของแบบสอบถาม

ทดสอบเครื่องมือจากการสำรวจกลุ่มตัวอย่างทั้งหมด 359 คน เพื่อทดสอบความเข้าใจ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน และวิเคราะห์ความเชื่อมั่นของแบบสอบถาม

ตารางที่ 3 : Reliability

	Mean	S.D	Cronbach's Alpha
Benefits	5.578	1.421	0.936
Service	5.155	1.55	0.912
Satisfaction	5.253	1.398	0.917
System	5.188	1.405	0.847
Information	5.418	1.249	0.832
Anxiety	4.134	2.744	0.896

พบว่าการวิเคราะห์ ความเชื่อมั่นของแบบสอบถาม พบว่ามีค่าสัมประสิทธิ์ครอนบาคแอลฟา (Cronbach's Alpha) อยู่ ระหว่าง 0.832 ถึง 0.936 ซึ่งอยู่ในระดับที่ยอมรับได้แสดงว่าเครื่องมือแบบสอบถามที่ใช้ในการสำรวจฉบับนี้มีความน่าเชื่อถือและผลการวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 6 ด้าน พบว่าการรับรู้ประโยชน์ในการใช้งานมีค่าเฉลี่ยมากที่สุด (Mean = 5.578, SD = 2.744)

4.4 การวิเคราะห์สมการถดถอยพหุคูณ

Model Summary									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	.848 ^a	.718	.715	.58910	.718	225.781	4	354	.000

a. Predictors: (Constant), Anxiety, System, Information, Service

ภาพที่ 4 : Model Summary

ผลลัพธ์ที่แสดงออกมาจากตาราง Model Summary จะบอกให้ทราบสมการพยากรณ์ที่ได้มีประสิทธิภาพในการพยากรณ์มากน้อยแค่ไหน โดยดูจาก R Square ซึ่งจากตารางที่ 4 นี้ R Square = 0.718 แสดงว่าสมการพยากรณ์ที่ได้ มีประสิทธิภาพเพียงร้อยละ 71.8

Coefficients ^a											
Model	Unstandardized Coefficients			Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
	B	Std. Error	Beta	Lower Bound	Upper Bound			Tolerance	VIF		
1	(Constant)	.093	.193		.431	.667		-.296	.462		
	Service	.361	.843	.380	8.467	.000	.277	.445	.395	2.531	
	System	.330	.847	.314	7.864	.000	.238	.423	.401	2.481	
	Information	.257	.847	.242	5.463	.000	.165	.350	.406	2.463	
	Anxiety	.046	.821	.062	2.210	.028	.005	.086	.895	1.005	

a. Dependent Variable: Satisfaction

ภาพที่ 5 : Coefficients

พบว่าทุกด้านมีค่าเป็น Sig. อย่างมีระดับนัยสำคัญที่ 0.05 แสดงตัวแปรที่สามารถใช้งานได้จริง และจากผลของการวิเคราะห์ของแต่ละตัวแปรได้ค่าสัมประสิทธิ์ของสมการถดถอย (Beta Coefficient) ซึ่งเป็นค่าที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ซึ่งผลการวิเคราะห์คุณภาพการบริการมีผลมากที่สุด รองลงมาคือคุณภาพของระบบและคุณภาพของข้อมูล น้อยสุดคือความวิตกกังวลการใช้

Model Summary									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	.815 ^a	.664	.663	.61740	.664	705.876	1	357	.000

a. Predictors: (Constant), Satisfaction

ภาพที่ 6 : Model Summary

ผลลัพธ์ที่แสดงออกมาจากตาราง Model Summary จะบอกให้ทราบสมการพยากรณ์ที่ได้มีประสิทธิภาพในการพยากรณ์มากน้อยแค่ไหน โดยดูจาก R Square ซึ่งจากตารางที่ 6 นี้ R Square = 0.664 แสดงว่าสมการพยากรณ์ที่ได้ มีประสิทธิภาพเพียงร้อยละ 66.4

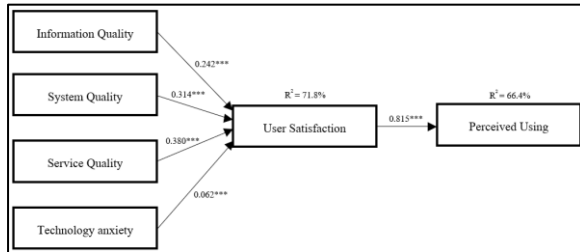
Coefficients ^a										
Model	Unstandardized Coefficients			Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics
	B	Std. Error	Beta	Lower Bound	Upper Bound			Tolerance	VIF	
1	(Constant)	1.427	.160		8.932	.000	1.113	1.741		
	Satisfaction	.791	.030	.815	26.568	.000	.732	.849	1.000	1.000

a. Dependent Variable: Benefits

ภาพที่ 7 : Coefficients

พบว่าทุกด้านมีค่าเป็น Sig. อย่างมีระดับนัยสำคัญที่ 0.05 แสดงตัวแปรที่สามารถใช้งานได้จริง และจากผล

ของการวิเคราะห์ของแต่ละตัวแปรได้ค่าสัมประสิทธิ์ของสมการถดถอย (Beta Coefficient) ซึ่งเป็นค่าที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ซึ่งผลการวิเคราะห์ความพึงพอใจของผู้ใช้งานมีระดับนัยสำคัญ



ภาพที่ 8 : โมเดลที่ส่งผลกระทบต่อประโยชน์ในการใช้งานผู้ใช้บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต

จากภาพที่ 4 โมเดลที่ส่งผลกระทบต่อประโยชน์ในการใช้งานผู้ใช้บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต สามารถพยากรณ์ได้ 71.8% (R-Squared= 0.718) และความพึงพอใจของผู้ใช้งานส่งผลกระทบต่อรับรู้ประโยชน์ต่อการใช้งาน สามารถพยากรณ์ได้ 66.4% (R-Squared=0.664) แสดงว่าเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ตส่งผลประโยชน์ต่อการใช้งานของผู้ใช้บริการ

5. สรุปผลและอภิปรายผล

งานวิจัยครั้งนี้ ผู้วิจัยได้ผลการวิเคราะห์ปัจจัยที่ส่งผลกระทบต่อรับรู้ประโยชน์ในการใช้บริการเว็บไซต์การชำระเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต ค่าความเหมาะสมของข้อมูล (KMO) ตัวแปรมีความสัมพันธ์ถึง 0.956 ซึ่งมีค่ามากกว่า 0.5 เป็นข้อมูลที่สามารถนำมาตั้งเป็นข้อคำถามได้ ผลน้ำหนักรวมขององค์ประกอบมีค่ามากกว่า 0.45 มีทั้งหมด 6 องค์ประกอบ ประกอบไปด้วยคุณภาพของระบบ คุณภาพของข้อมูล คุณภาพการบริการ ความพึงพอใจของผู้ใช้งาน ผลประโยชน์ในการใช้งาน และความวิตกกังวลการใช้ จึงนำมาทดสอบความเชื่อมั่นของแบบสอบถามกับกลุ่มตัวอย่าง 359 คน มีความเชื่อมั่นอยู่ที่ 0.832 ถึง 0.836 ซึ่งอยู่ในระดับที่น่าเชื่อถือ

และผลวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 6 ด้านพบว่า รับรู้ถึงความง่ายที่มีค่าเฉลี่ยมากที่สุด (Mean=5.578,SD=2.744) ซึ่งผลการวิเคราะห์ถดถอยพหุคูณ ทุกด้านมีค่าเป็น Sig. อย่างมีระดับนัยสำคัญที่ 0.05 แสดงตัวแปรที่สามารถใช้งานได้จริง และจากผลของการวิเคราะห์ของแต่ละตัวแปรได้ค่า สัมประสิทธิ์ของสมการถดถอย แสดงถึงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม

อย่างไรก็ตามงานวิจัยนี้ หากมีการศึกษาครั้งต่อไป จึงควรปรับปรุงกรอบการวิจัยให้กว้างขึ้น โดยเพิ่มการศึกษาปัจจัยอื่นๆที่สามารถนำมาวิเคราะห์ความพึงพอใจของผู้ใช้งานเว็บไซต์ได้ หรือปัญหาผลกระทบต่างๆ

เอกสารอ้างอิง

- [1] ว. นุตโร, “ปัจจัยที่มีผลต่อการใช้บริการยื่นแบบภาษีเงินได้บุคคลธรรมดาผ่านทางอินเทอร์เน็ต ในเขตสำนักงานสรรพากรพื้นที่สาขาท่ามะกา จังหวัดกาญจนบุรี”, ปี 2563.
- [2] ว. พิมพ์สุวรรณ, “ปัจจัยที่มีผลต่อความสนใจในการเสียภาษีเงินได้บุคคลธรรมดา ของผู้มีเงินได้ในเขตภาคตะวันออกเฉียงเหนือ”, ปี 2562, 15.
- [3] K. Benmoussa, M. Laaziri, S. Khouliji, M. L. Kerkeb, และ A. El Yamami, “Impact of System Quality, Information Quality and Service Quality on the efficiency of information system”, ใน *Proceedings of the 3rd International Conference on Smart City Applications*, Tetouan Morocco, ต.ค. 2018, น. 1-6. doi: 10.1145/3286606.3286818.
- [4] น. มนุสรพรวัฒนา, “ทัศนคติและปัจจัยความสำเร็จของระบบสารสนเทศที่มีอิทธิพล ต่อความพึงพอใจในการใช้บริการ Netflix ในแต่ละ Generation”, ปี 2562.
- [5] น. ธรรมจาร, “การศึกษาปัจจัยที่ส่งผลต่อความตั้งใจในการใช้เทคโนโลยีสุขภาพ ของผู้สูงอายุในเขตกรุงเทพมหานครและปริมณฑล”, ปี 2561.
- [6] น. ไพบุญย์, “ปัจจัยที่ส่งผลต่อความตั้งใจซื้อบริการสตรีมมิ่งเพลงดิจิทัล”, ปี 2563.

**การวิเคราะห์องค์ประกอบปัจจัยในการเลือกเรียน
คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
Factors Analysis Decision Making in Selecting Institutions
Faculty of Information Technology and Digital Innovation**

ชลลดา ชูชมชื่น (Chonlada Chuchomchuen)¹ และมณเฑียร รัตน์ศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857042@email.kmutnb.co.th, ²montean.r@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาองค์ประกอบปัจจัยในการเลือกเรียนคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล โดยได้ดำเนินการสำรวจและนำข้อมูลที่ได้ออกมาวิเคราะห์องค์ประกอบของปัจจัยผลการวิจัยพบว่า ในการวิเคราะห์องค์ประกอบเชิงสำรวจได้องค์ประกอบปัจจัยในการเลือกเรียนจำนวน 4 องค์ประกอบ ได้แก่ ด้านภาพลักษณ์ด้านสวัสดิการและการบริการ ด้านอาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์ และ ด้านปัจจัยส่วนบุคคล การวิเคราะห์องค์ประกอบเชิงยืนยันพบว่าค่า *CMIN/DF* (1.162), *GFI* (0.951), *AGFI* (0.909), *CFI* (0.995), *RMSEA* (0.023), *RMR* (0.019) และการวิเคราะห์เส้นทาง พบว่าปัจจัยด้านสวัสดิการและการบริการมีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ ปัจจัยด้านอาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์มีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ และปัจจัยด้านภาพลักษณ์มีอิทธิพลทางตรงเชิงบวกต่อปัจจัยส่วนตัว ซึ่งแบบจำลองความสัมพันธ์เชิงสาเหตุที่พัฒนาขึ้นมีความเหมาะสมซึ่งสอดคล้องกับข้อมูลเชิงประจักษ์มีความสามารถในการทำนาย (R^2) ด้านปัจจัยภาพลักษณ์ คิดเป็นร้อยละ 0.63 และด้านปัจจัยส่วนตัวคิดเป็นร้อยละ 0.54

คำสำคัญ: การวิเคราะห์เชิงสำรวจ การวิเคราะห์เชิงยืนยัน การวิเคราะห์เส้นทาง ปัจจัยการเลือกเรียน

Abstract

The purpose of the research was to inform decision-making in selecting institutions for the Faculty of Information Technology and Digital Innovation. The data obtained were analyzed based on the components. The results showed that in exploratory component analysis There were 4 factors in choosing to study, namely Image Factor, Welfare and service Factor, Instructor, Courses and Products and Personal Factors. Confirmatory component analysis showed *CMIN/DF* (1.162), *GFI* (0.951), *AGFI* (0.909), *CFI* (0.995), *RMSEA* (0.023), and *RMR* (0.019) and Path Analysis showed Welfare and Service factors had a positive direct influence on Image Factor. Instructor Courses and Products factor have a positive direct influence on Image Factor and the Image Factor had a positive direct influence on Personal Factors. The developed causal relationship model is appropriate, consistent with the empirical data, and could predict (R^2) in terms of the Image Factor. representing 0.63 percent, and Personal Factors representing 0.54 percent.

Keywords: Exploratory Factor Analysis, Confirmatory Factor Analysis, Path Analysis, Factors in Affecting Decision.

1. บทนำ

ปัจจุบันแนวโน้มทางการศึกษากำลังถูกพัฒนาขึ้นมาก ซึ่งเป็นกระแสทำให้เกิดความกระตือรือร้นพยายามหาทางศึกษาต่อในระดับที่สูงขึ้น นำไปสู่การแข่งขันในช่วงชีวิตวัยทำงาน รายได้ที่เพิ่มขึ้น ตลอดจนเป็นที่ยอมรับจากเพื่อนฝูง แล้วยังมีการแข่งขันทางการศึกษาของมหาวิทยาลัยต่าง ๆ ทั้งเรื่องภาพลักษณ์ของสถาบัน คุณภาพการศึกษา สวัสดิการของนักศึกษาบุคลากรเพื่อทำให้เกิดแรงจูงใจในการสมัครเข้าศึกษาต่อในสถาบันนั้น

ดังนั้นผู้วิจัยได้ตระหนักถึงบทบาทและหน้าที่ของคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล ในการจัดการศึกษาและให้บริการทางการศึกษาให้สามารถนำเอาความรู้ที่ได้รับไปใช้ในการพัฒนาอาชีพตามความต้องการของตลาดแรงงาน และประเทศชาติต่อไป จึงได้มองเห็นว่าปัจจัยในการเข้ารับการศึกษาต่อมีความสำคัญต่อการนำข้อมูลที่ได้จากการศึกษาวิจัยเป็นข้อมูลพื้นฐานในการวางแผนเพื่อการพัฒนาหลักสูตร ตลอดจนแนวทางการเรียนการสอนเพื่อการจัดการศึกษาและให้บริการที่มีประสิทธิภาพ เหมาะสมกับความต้องการของผู้ที่ต้องการเข้ารับการศึกษาต่อในอนาคต

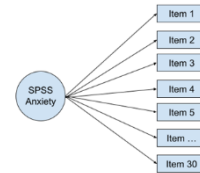
2. รูปแบบบทความ

2.1 ทฤษฎีการวิเคราะห์องค์ประกอบ

การวิเคราะห์องค์ประกอบการวิเคราะห์องค์ประกอบเป็นวิธีการทางสถิติในการลดปริมาณตัวแปรในงานวิจัยอาศัยหลักการการรวบรวมตัวแปรที่มีความสัมพันธ์กันมาสร้างเป็นองค์ประกอบใหม่ (Factor) เป็นเบื้องหลังของตัวแปรนั้นๆ ทำให้นักวิจัยที่เข้าใจโครงสร้างตัวแปรในองค์ประกอบนั้นได้ดีและจุดเด่นของปัจจัยที่วิเคราะห์ออกมาคือ ลดจำนวนความคลาดเคลื่อนในการวัดลง มีประโยชน์ที่นำปัจจัยวิเคราะห์ได้ไปทำการศึกษาต่อ หรือรวมตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มเดียวกันซึ่งความสัมพันธ์เป็นไปได้ทั้งทางบวกและทางลบ ตัวแปรภายในองค์ประกอบเดียวกันจะมีความสัมพันธ์กันสูง

2.1.1 วัตถุประสงค์ของเทคนิคการวิเคราะห์องค์ประกอบ (Factor Analysis)

2.1.1.1 เพื่อศึกษาว่าองค์ประกอบรวมที่จะสามารถอธิบายความสัมพันธ์ร่วมกันระหว่างตัวแปรต่าง ๆ โดยที่จำนวนองค์ประกอบรวมที่หาได้จะมีจำนวนน้อยกว่าจำนวนตัวแปรนั้น จึงทำให้ทราบว่ามืองค์ประกอบรวมอะไรบ้าง โมเดลนี้ เรียกว่า Exploratory Factor Analysis Model: EFA



ภาพที่ 1 : โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ

2.1.1.2 เพื่อต้องการทดสอบสมมุติฐานเกี่ยวกับโครงสร้างขององค์ประกอบว่าองค์ประกอบแต่ละองค์ประกอบด้วยตัวแปรอะไรบ้าง และตัวแปรแต่ละตัวควรมีน้ำหนักหรืออัตราความสัมพันธ์กับองค์ประกอบมากน้อยเพียงใด ตรงกับที่คาดคะเนไว้หรือไม่ หรือสรุปได้ว่าเพื่อต้องการทดสอบว่าตัวประกอบอย่างนี้ตรงกับโมเดลหรือตรงกับทฤษฎีที่มีอยู่หรือไม่ โมเดลนี้เรียกว่า Confirmatory Factor Analysis Model: CFA

2.2 การวิเคราะห์เส้นทาง (Path Analysis)

เป็นการศึกษาความสัมพันธ์ของตัวแปรในเชิงเหตุและผล เป็นวิธีที่มีพื้นฐานทางสถิติมาจากการวิเคราะห์การถดถอย (Regression analysis) โดยอาศัยแผนภาพและสมการ โครงสร้างของแผนภาพเป็นหลักในการนำมาวิเคราะห์และอธิบายความสัมพันธ์ของตัวแปรเหตุที่มีต่อตัวแปรผลทั้งในด้านขนาดและทิศทาง อีกทั้งยังสามารถอธิบายความสัมพันธ์ได้ทั้งความสัมพันธ์ทางตรงและทางอ้อม

2.2.1 เป็นการศึกษาอิทธิพลระหว่างตัวแปรต่าง ๆ เพื่อดูว่ามีอิทธิพลทางตรงและอิทธิพลทางอ้อมของตัวแปรที่สันนิษฐานว่าเป็นสาเหตุต่อตัวแปรที่เป็นผลหรือไม่

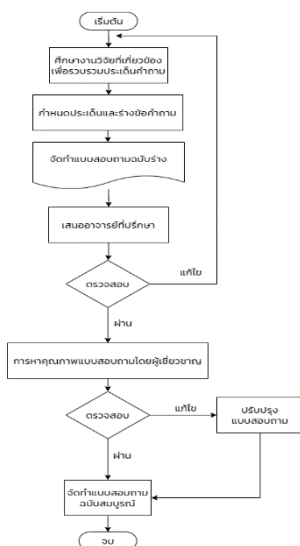
2.2.2 การวิจัยที่มุ่งแสวงหาความสัมพันธ์เชิงสาเหตุ (Causal Analysis) โมเดลเชิงสาเหตุ (Causal Relationship Model) มีดังนี้ 1) การวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์มุ่งบรรยาย (description) มุ่งอธิบาย (explanation) การพยากรณ์ (prediction) และการควบคุม (control)

ปรากฏการณ์ทางธรรมชาติ 2) โมเดลการวิจัย เป็นสิ่งที่นักวิจัยสร้างขึ้นตามทฤษฎีแทนปรากฏการณ์ที่เกิดขึ้นจริงตามธรรมชาติ

3. วิธีการดำเนินงาน

ในการวิเคราะห์หองค์ประกอบปัจจัยของการตัดสินใจในการเลือกเรียนคณะ เพื่อให้ได้ผลการวิจัยเป็นไปตามวัตถุประสงค์หลัก ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องของพร้อมศึกษาตัวแปรที่มีความสัมพันธ์ โดยคัดเลือกตัวแปรที่สำคัญด้วยการจัดระบบความสัมพันธ์ของตัวแปร จากนั้นจึงทำการตรวจสอบความกลมกลืนโมเดลกับข้อมูลเชิงประจักษ์ โดยใช้โปรแกรม SPSS และ AMOS ผู้วิจัยได้แบ่งขั้นตอนดำเนินงาน ดังนี้

3.1 การสร้างและหาคุณภาพของเครื่องมือ



ภาพที่ 2 : Flow การดำเนินงานสร้างแบบสอบถาม

3.1.1 ศึกษาข้อมูลจากทฤษฎีและงานวิจัยที่เกี่ยวข้องเพื่อกำหนดประเด็นและร่างข้อคำถามเพื่อจัดทำร่างแบบสอบถามซึ่งแบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ส่วนที่ 2 ปัจจัยที่มีผลต่อการเลือกเรียนคณะ แบบมาตราส่วนประมาณค่า (Rating Scale) ให้เลือกตอบ 5 ระดับ

3.1.2 นำแบบสอบถามเสนอต่ออาจารย์ที่ปรึกษาเพื่อตรวจสอบข้อคำถาม แล้วนำมาปรับปรุงตามคำแนะนำของอาจารย์ที่ปรึกษา

3.1.3 การหาคุณภาพแบบสอบถามโดยผู้เชี่ยวชาญด้วยการตรวจสอบความเที่ยงตรง (Validity) และความเชื่อมั่น (Reliability) ของแบบสอบถาม

3.1.4 จัดทำแบบสอบถามฉบับสมบูรณ์ เพื่อนำไปเก็บรวบรวมข้อมูล

3.1.5 นำแบบสอบถามไปเก็บข้อมูลจากนักศึกษาในคณะระดับบัณฑิตศึกษาทั้งที่กำลังศึกษาอยู่และสำเร็จการศึกษาไปแล้ว จำนวนทั้งสิ้น 314 คน

3.2 การวิเคราะห์องค์ประกอบเชิงสำรวจ

3.2.1 เตรียมเมตริกซ์สหสัมพันธ์ คำนวณหาค่าความสัมพันธ์เส้นตรงวิธีของเพียร์สันคลอเรนซ์ระหว่างตัวแปรทั้งหมดในงานวิจัยนำมาจัดกลุ่มกัน เพื่อที่จะทำการลดขนาดตัวแปรที่ไม่มีความสัมพันธ์ออกไปหรือจับกุมตัวแปรที่มีความสัมพันธ์ไปในทิศทางเดียวกันให้อยู่ในองค์ประกอบเดียวกัน

3.2.2 การสกัดองค์ประกอบ เป็นการหาจำนวนองค์ประกอบที่สามารถใช้ตัวแปรทั้งหมดทุกตัวได้ ใช้การสกัดองค์ประกอบ แบบองค์ประกอบหลัก (Principal Component Analysis)

3.2.3 การหมุนแกนองค์ประกอบ เป็นขั้นตอนที่จะดำเนินการแยกตัวแปรให้เห็นเด่นชัดว่าตัวแปรหนึ่งๆ ควรจะจัดอยู่ในกลุ่มหรือในองค์ประกอบใด

3.2.4 การตั้งชื่อองค์ประกอบ เป็นการให้ความหมายแต่ละองค์ประกอบด้วยการกำหนดชื่อ โดยพิจารณาว่าในปัจจุบันนี้ ๆ อาจตั้งชื่อตามความคล้ายคลึงกันของทุกตัวแปรที่อยู่ในองค์ประกอบ

3.3 การวิเคราะห์องค์ประกอบเชิงยืนยัน

3.3.1 การกำหนดข้อมูลเฉพาะของโมเดล ผู้วิจัยระบุโมเดลจำเพาะที่ใช้ในการยืนยันความสอดคล้องกับข้อมูลที่อยู่ในรูปแบบของความแปรปรวน-ความแปรปรวนร่วม

3.3.2 การระบุความเป็นไปได้ค่าเดียวของโมเดล เป็นการเชื่อมโยงระหว่างโมเดลสมการ โครงสร้างอ่านการกำหนดข้อมูลเฉพาะของโมเดล โดยใช้ความสมเหตุสมผลอย่างถี่ถ้วนกับโปรแกรมที่ใช้ตรวจสอบความสอดคล้อง

3.3.3 การประมาณค่าพารามิเตอร์ของโมเดล การระบุค่าความเป็นไปได้ค่าเดียวของโมเดล ใช้วิธี Maximum likelihood (ML)

3.3.4 การตรวจสอบความสอดคล้องของโมเดล

3.3.5 การปรับโมเดล เป็นขั้นตอนจะทำเมื่อมีค่าพารามิเตอร์บางค่าที่ไม่แตกต่างจากศูนย์หรือมีทิศทางค่าพารามิเตอร์ไม่ตรงกับทฤษฎีที่กำหนดไว้

3.3.6 การแปรผลการวิเคราะห์โมเดล แสดงผลการวิเคราะห์เพื่ออธิบายความสอดคล้องของโมเดล

3.4 การวิเคราะห์เส้นทาง

ศึกษาอิทธิพลระหว่างตัวแปรต่าง ๆ เพื่อความีอิทธิพลทางตรงและอิทธิพลทางอ้อมของตัวแปรที่สันนิษฐานว่าเป็นสาเหตุต่อตัวแปรที่เป็นผลหรือไม่

4. ผลการดำเนินงาน

4.1 ผลการวิเคราะห์องค์ประกอบเชิงสำรวจ

ผู้วิจัยได้นำข้อมูลไปวิเคราะห์ด้วยโปรแกรม SPSS พบว่า ค่าสถิติ Bartlett's Test of Sphericity มีค่าเท่ากับ 6119.25 โดยมีค่า Sig (Significance) น้อยกว่า 0.05 และมีค่าดัชนี KMO เท่ากับ 0.94 ซึ่งถือว่าอยู่ในเกณฑ์ที่มีความเหมาะสมในระดับดีในการนำมาวิเคราะห์องค์ประกอบ ซึ่งจากการสกัดองค์ประกอบ พบว่า จำนวนองค์ประกอบที่สามารถอธิบายความผันแปรของข้อมูลที่มีค่า Eigenvalues มากกว่า 1 มีจำนวนทั้งหมด 4 องค์ประกอบ ดังนี้

ตารางที่ 1: องค์ประกอบที่ 1 ด้านภาพลักษณ์ของมหาวิทยาลัย มีค่า Eigenvalues เท่ากับ 5.44

Component	Initial Eigenvalues		
	Total	% Of Variance	Cumulative %
IMA1	5.44	67.95	67.95
IMA2	0.86	10.79	78.74
IMA3	0.49	6.35	84.97
IMA4	0.38	4.77	89.74
IMA5	0.27	3.37	93.11
IMA6	0.23	2.84	95.95
IMA7	0.19	2.39	98.34
IMA8	0.13	1.66	100.00
Kaiser Meyer Olkin (KMO) = 0.872, Bartlett's Test of Sphericity = 1199.31, df = 15			

ตารางที่ 2: องค์ประกอบที่ 2 ด้านสวัสดิการและการบริการ Eigenvalues เท่ากับ 4.10

Component	Initial Eigenvalues		
	Total	% Of Variance	Cumulative %
WAS1	4.10	68.33	68.33
WAS2	0.60	10.05	78.38
WAS3	0.47	7.83	86.21
WAS4	0.38	6.34	92.55
WAS5	0.27	4.57	97.12
WAS6	0.17	2.88	100.00
Kaiser Meyer Olkin (KMO) = 0.906, Bartlett's Test of Sphericity = 2036.37, df = 28			

ตารางที่ 3: องค์ประกอบที่ 3 ด้านอาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์ มีค่า Eigenvalues เท่ากับ 4.53

Component	Initial Eigenvalues		
	Total	% Of Variance	Cumulative %
ICP1	4.53	56.66	56.66
ICP2	0.967	12.12	68.78
ICP3	0.68	8.55	77.33
ICP4	0.51	6.33	83.67
ICP5	0.43	5.37	89.04
ICP6	0.35	4.33	93.37
ICP7	0.31	3.89	97.26
ICP8	0.22	2.75	100.00
Kaiser Meyer Olkin (KMO) = 0.878, Bartlett's Test of Sphericity = 1289.95, df = 28			

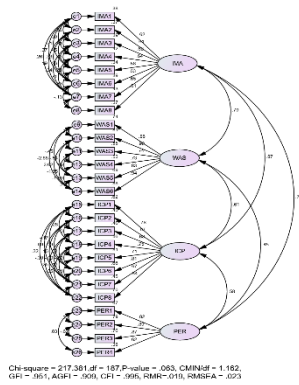
ตารางที่ 4: องค์ประกอบที่ 4 ด้านปัจจัยส่วนบุคคล มีค่า Eigenvalues เท่ากับ 2.89

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
PER1	2.89	72.12	72.12
PER2	0.55	13.68	85.80
PER3	0.37	9.28	95.09
PER4	0.20	4.91	100.00
Kaiser Meyer Olkin (KMO) = 0.782, Bartlett's Test of Sphericity = 671.74, df = 6			

จากทั้งหมด 4 องค์ประกอบ พบว่า ค่าน้ำหนักความสำคัญของตัวแปรทุกตัวมีความสัมพันธ์กัน

4.2 ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน

ผู้วิจัยนำผลที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจมาวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม SPSS AMOS เพื่อตรวจสอบความถูกต้องเหมาะสมตามทฤษฎีได้ผลการวิเคราะห์โมเดลเพื่อทดสอบความสอดคล้องกลมกลืนของโมเดล ดังนี้



ภาพที่ 3 : โมเดลองค์ประกอบเชิงยืนยันของปัจจัยที่มีผลในการเลือกเรียนคณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

องค์ประกอบที่ 1 ด้านภาพลักษณ์ มีค่า Chi-Square เท่ากับ 5.322, df เท่ากับ 4, P-value เท่ากับ 0.256, RMSEA เท่ากับ 0.033, GFI เท่ากับ 0.996, AGFI เท่ากับ 0.962, CFI เท่ากับ 0.999 และ RMR เท่ากับ 0.006

องค์ประกอบที่ 2 ด้านสวัสดิการและการบริการ มีค่า Chi-Square เท่ากับ 2.730, df เท่ากับ 5, P-value เท่ากับ 0.742, RMSEA เท่ากับ 0.000, GFI เท่ากับ 0.997, AGFI เท่ากับ 0.988, CFI เท่ากับ 1.000 และ RMR เท่ากับ 0.005

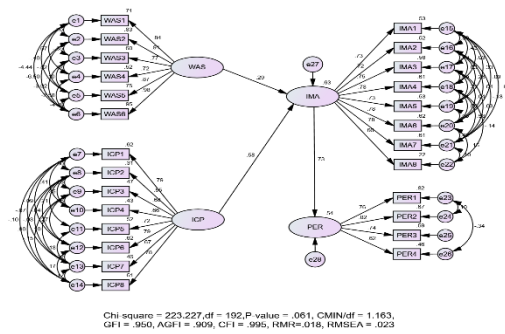
องค์ประกอบที่ 3 ด้านอาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์ มีค่า Chi-Square เท่ากับ 9.402, df เท่ากับ 6, P-value เท่ากับ 0.152, RMSEA เท่ากับ 0.043, GFI เท่ากับ 0.993, AGFI เท่ากับ 0.956, CFI เท่ากับ 0.998 และ RMR เท่ากับ 0.009

องค์ประกอบที่ 4 ด้านปัจจัยส่วนตัว เมื่อพิจารณาจากค่า Chi-Square เท่ากับ 0.092, df เท่ากับ 1, P-value เท่ากับ 0.762, RMSEA เท่ากับ 0.000, GFI เท่ากับ 1.00, AGFI เท่ากับ 0.999, CFI เท่ากับ 1.000 และ RMR เท่ากับ 0.001

พบว่า องค์ประกอบทั้ง 4 โมเดลตามทฤษฎี มีความคล้อยกับข้อมูลเชิงประจักษ์

4.3 ผลวิเคราะห์เชิงเส้น

ผู้วิจัยได้ทดสอบความเหมาะสมของแบบจำลอง (Model fit) จากนั้นทำการวิเคราะห์เส้นทาง ซึ่งพิจารณาได้จากค่าสัมประสิทธิ์เส้นทางความสัมพันธ์ โดยงานวิจัยนี้มีค่าสัมประสิทธิ์เส้นทางความสัมพันธ์ที่วัด ตามภาพที่ 4



ภาพที่ 4 : ความสัมพันธ์ระหว่างตัวแปรในแบบจำลองสมการโครงสร้างปัจจัยของการตัดสินใจในการเลือกเรียนคณะของนักศึกษาระดับบัณฑิตศึกษา

ผลการทดสอบสมมติฐานในแบบจำลองโครงสร้างพบว่า มีการยอมรับสมมติฐาน จำนวน 3 สมมติฐาน ได้แก่ สมมติฐานที่ 1 (H1): สวัสดิการและการบริการมีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ พบว่า สวัสดิการและการบริการมีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ มีค่าสัมประสิทธิ์ความสัมพันธ์เป็นบวกเท่ากับ 0.29 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

สมมติฐานที่ 2 (H2): อาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์มีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ พบว่า อาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์มีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ มีค่าสัมประสิทธิ์ความสัมพันธ์เป็นบวกเท่ากับ 0.58 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

สมมติฐานที่ 3 (H3): ภาพลักษณ์มีอิทธิพลทางตรงเชิงบวกต่อปัจจัยส่วนตัว พบว่า ภาพลักษณ์มีอิทธิพลทางตรงเชิงบวกต่อปัจจัยส่วนตัว มีค่าสัมประสิทธิ์ความสัมพันธ์เป็นบวกเท่ากับ 0.73 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

5. สรุป

จากการเก็บข้อมูลจาก นักศึกษาในคณะระดับบัณฑิตศึกษาทั้งที่กำลังศึกษาอยู่และสำเร็จการศึกษาไปแล้ว เพศหญิง จำนวน 160 คน คิดเป็นร้อยละ 51.10 เพศชาย จำนวน 141 คน คิดเป็นร้อยละ 44.90 และเพศ LGBTQ+ จำนวน 13 คน คิดเป็นร้อยละ 4.10 โดยส่วนใหญ่มีอายุระหว่าง 25-35 ปี คิดเป็นร้อยละ 64.30 ส่วนใหญ่

ประกอบอาชีพ พนักงานบริษัทเอกชน คิดเป็นร้อยละ 31.50 มีรายได้ต่อเดือน 20,001-30,000 บาท คิดเป็นร้อยละ 35.40 แบ่งเป็นนักศึกษาที่กำลังศึกษาจำนวน 152 คน คิดเห็นร้อยละ 48.4 และสำเร็จการศึกษาระดับปริญญาโท จำนวน 162 คิดเป็นร้อยละ 51.60

พบว่า ค่าน้ำหนักองค์ประกอบของตัวประกอบทั้งหมด มีค่าเป็นบวก มีขนาดตั้งแต่ 0.56-0.98 โดยปัจจัยที่มีค่าน้ำหนักองค์ประกอบสูงสุด 3 อันดับแรก ได้แก่ ปัจจัยมีบริการอินเทอร์เน็ตไร้สาย (WIFI) อย่างทั่วถึง ในมหาวิทยาลัย มีน้ำหนักองค์ประกอบเท่ากับ 0.98 ปัจจัยสิ่งแวดล้อมและบรรยากาศน่าอยู่ น่าเรียน มีอาคารปฏิบัติการที่เหมาะสมเพียงพอ มีน้ำหนักองค์ประกอบเท่ากับ 0.91 ปัจจัยมีเอกสาร ตำราหนังสือ ทรัพยากรสารสนเทศและระบบสืบค้นข้อมูลอิเล็กทรอนิกส์ที่ทันสมัยเพียงพอสำหรับนักศึกษาหาความรู้เพิ่มเติมได้ด้วยตนเอง มีน้ำหนักองค์ประกอบเท่ากับ 0.87 โดยมีค่าสถิติอธิบายความผันแปรของตัวแปรทั้ง 3 ตัวแปรอยู่ระหว่าง 0.75-0.95 จึงกล่าวได้ว่าองค์ประกอบหรือปัจจัยแฝงนี้สามารถจัดได้จากตัวแปรสังเกตได้ดังกล่าว

ปัจจัยที่มีน้ำหนักองค์ประกอบน้อยที่สุด 3 อันดับ ได้แก่ ปัจจัยมีทุนสนับสนุนการศึกษาของสถาบัน และปัจจัยอาจารย์ผู้สอนเป็นผู้มีความรู้ความเชี่ยวชาญและมีประสบการณ์การสอนที่ยาวนานเป็นที่น่าเชื่อถือ เท่ากับ 0.66 เท่ากัน ปัจจัยต้องการเพิ่มวุฒิการศึกษา เท่ากับ 0.62 และปัจจัยอาจารย์ผู้สอนส่วนใหญ่มีตำแหน่งทางวิชาการ และมีคุณวุฒิด้านวิชาการสูง มีน้ำหนักองค์ประกอบเท่ากับ 0.56 โดยมีค่าสถิติอธิบายความผันแปรของตัวแปรทั้ง 3 ตัวแปรอยู่ระหว่าง 0.22-0.46 จึงกล่าวได้ว่าองค์ประกอบหรือปัจจัยแฝงนี้สามารถจัดได้จากตัวแปรสังเกตได้

สมมติฐานที่ 1 (H1): สวัสดิการและการบริการมีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ สอดคล้องกับสมมติฐาน แสดงให้เห็นว่า สวัสดิการและการบริการมีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์

สมมติฐานที่ 2 (H2): อาจารย์ผู้สอน หลักสูตรและผลิตภัณฑ์มีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์ สอดคล้องกับสมมติฐาน แสดงให้เห็นว่า อาจารย์ผู้สอน

หลักสูตรและผลิตภัณฑ์มีอิทธิพลทางตรงเชิงบวกต่อภาพลักษณ์

สมมติฐานที่ 3 (H3): ภาพลักษณ์มีอิทธิพลทางตรงเชิงบวกต่อปัจจัยส่วนตัว สอดคล้องกับสมมติฐาน แสดงให้เห็นว่า ภาพลักษณ์มีอิทธิพลทางตรงเชิงบวกต่อปัจจัยส่วนตัว

และค่าสัมประสิทธิ์การทำนาย (R²) เพื่อตรวจสอบความตรงของแบบจำลอง พบว่า ตัวแปรภาพลักษณ์ มีค่าสัมประสิทธิ์การทำนายเท่ากับ 0.63 หรือคิดเป็นร้อยละ 63 และตัวแปรปัจจัยส่วนตัว มีค่าสัมประสิทธิ์การทำนายเท่ากับ 0.54 หรือคิดเป็นร้อยละ 54

เอกสารอ้างอิง

- [1] ธนพัฒน์ อินทวิ, “ปัจจัยที่มีผลต่อการตัดสินใจในการเลือกสถาบันศึกษาเพื่อเรียนต่อในระดับอุดมศึกษา,” *วารสารสถาบันเทคโนโลยีนวัตกรรมทางการศึกษาและการวิจัยแห่งสุวรรณภูมิ (สนว.)* ปีที่ 5 ฉบับที่ 1 มกราคม - เมษายน พ.ศ. 2564 หน้า 1-14.
- [2] มัชฌิมา รัตนลัมภ์ และ วชิราภรณ์ ทองคุ้ม, “ปัจจัยที่ส่งผลต่อการตัดสินใจเลือกศึกษาต่อระดับบัณฑิตศึกษา สาขาวิชาวิศวกรรมชีวการแพทย์ (หลักสูตรนานาชาติ) คณะวิศวกรรมศาสตร์มหาวิทยาลัยมหิดล,” *Mahidol R2R e-Journal* ปีที่ 8 ฉบับที่ 2 สิงหาคม พ.ศ. 2564 หน้า 160-172.
- [3] สุขพัฒน์ อนนท์จารย์, ภาสกร ดอกจันทร์ และ คณะ, “ปัจจัยที่ส่งผลต่อการตัดสินใจเลือกเข้าศึกษาต่อระดับปริญญาโทในมหาวิทยาลัยมหามกุฏราชวิทยาลัย วิทยาเขตศรีล้านช้าง” *วารสารวิชาการวิทยาลัยบริหารศาสตร์* ปีที่ 3 ฉบับที่ 1 มกราคม-มีนาคม พ.ศ. 2663 หน้า 49-60.
- [4] S.O. Pinaraswati and S. Saibat, “Factors Affecting Students Choosing Management Study Program (Case Study on The Faculty of Economics and Business Dr. SOETOMO UNIVERSITY, SURABAYA),” *International Journal of Economics, Business and Accounting Research (IJEBAR)*, Vol 4, No.03, pp. 242-255, 2020.
- [5] D. Rosalina, K. Yuliani, W. Purnamasari and M.R. Zati, “Factors Affecting Intention in Accounting Study Program Students Choosing the Public Accountant Profession,” *Jurnal Akuntansi Dan Bisnis, Vol. 6 No.1*, pp. 86-95, 2020.

- [6] I. Dalbudak, S. Çelik and M. Balyan, Examining the Factors Affecting the Selection of the Recreation Activity Type of University Students,” *Jurnal of Educational Psychology - Propósitos y Representaciones*, Vol. 09, May 2021.
- [7] F. Aryani and N. F. Umar, “Factors Affecting Z Generation on Selecting Majors in the University: An Indonesian Case,” *Journal of Social Studies Education Research*, Vol. 11, pp. 109-133, 2020.

ปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ

Factors Affecting the Adoption of Automatic Beverage Dispenser Technology

สุพิชญา จิตต์อารมย์ (Supichaya Jitarom)¹ และธนพล เจนสุทธิเวชกุล (Tanapon Jentsuttiwetchakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857158@email.kmutnb.ac.th, ²tanapon.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติของผู้บริโภคในประเทศไทย เก็บข้อมูลโดยใช้แบบสอบถามจากผู้บริโภคชาวไทยที่เคยใช้บริการเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ จำนวน 400 คน โดยใช้สถิติในการวิเคราะห์ข้อมูล ได้แก่ การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) และการวิเคราะห์สมการถดถอยเชิงพหุคูณ (Multiple Linear Regression) ผลการวิจัย พบว่า คุณภาพของข้อมูล คุณภาพของการบริการ ความสามารถในการควบคุมพฤติกรรม การรับรู้ด้านราคา มีผลต่อการรับรู้คุณค่า และนำไปสู่ความตั้งใจซื้อ โดยค่า R^2 แสดงให้เห็นว่าแบบจำลองนี้สามารถอธิบายปัจจัยที่ส่งผลต่อการรับรู้คุณค่า และความตั้งใจซื้อสินค้า ได้ 62.1% และ 59.4% ตามลำดับ

คำสำคัญ: ความตั้งใจซื้อ เทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ ผู้บริโภค

Abstract

The objective of this research is to study factors affecting the adoption of automatic beverage dispenser technology in Thailand. Data were Collected by an online questionnaire. 400 Thai consumers were used as sample. To test the hypotheses, Exploratory factor analysis (EFA) and multiple linear regression were used as data analysis. The results showed that the Information Quality Service Quality, Perceived Behavioral Control,

and Perceived Value affect the Perceived Value and lead to Purchase intention. The R^2 value shows that the model can explain the factors affecting perceived value. And purchase intent was 62.5% and 59.4%, respectively.

Keywords: Purchase Intention, Automatic Beverage Dispenser Technology, Consumers.

1. บทนำ

ปัจจุบันเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติเข้ามามีบทบาทในชีวิตประจำวันของมนุษย์มากขึ้น ด้วยความก้าวหน้าทางเทคโนโลยีทำให้มีการพัฒนาอุตสาหกรรมให้บริการในรูปแบบต่างๆ ที่มีความสะดวกขึ้น จะพบเห็นได้ตามสถานที่ทั่วไป ผู้บริโภคจึงสามารถเข้าถึงเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติได้สะดวกมากยิ่งขึ้น และกิจกรรมส่วนใหญ่ในการใช้งานเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ อาทิ การสะสมแต้มแลกของรางวัล การรับส่วนลดในวันเกิด สามารถตอบสนองความต้องการของผู้บริโภคได้ครบถ้วน

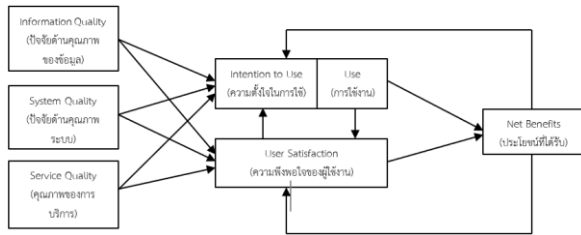
จากที่กล่าวข้างต้นข้อมูลเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ เป็นอีกทางเลือกในการเลือกซื้อเครื่องดื่มมีให้เลือกหลากหลาย และพฤติกรรมของผู้บริโภคที่เลือกใช้ที่แตกต่างกันไป ผู้วิจัยจึงมีความสนใจที่จะศึกษาปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติ โดยศึกษาจากปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีตู้กดเครื่องดื่มอัตโนมัติดังกล่าว โดยนำทฤษฎีความสำเร็จของระบบสารสนเทศ (ISS Model) มาใช้ร่วมกับทฤษฎีพฤติกรรมตามแผนและแนวคิดเกี่ยวกับการรับรู้ด้านราคามาเป็นเครื่องมือช่วยในการศึกษาและให้การศึกษ

นี้เป็นแนวทางสำหรับผู้ที่อยู่ในอุตสาหกรรมนี้สามารถนำข้อมูลที่ได้เป็นแนวทางไปส่งเสริมและพัฒนากลยุทธ์เพื่อตอบสนองความต้องการของผู้บริโภคและสร้างข้อได้เปรียบ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แบบจำลองความสำเร็จของระบบสารสนเทศ

(IS Success Model)



ภาพที่ 1: แบบจำลองความสำเร็จของระบบสารสนเทศ

(IS Success Model)

แบบจำลองความสำเร็จของระบบสารสนเทศ ของ Delone and Mclean (IS Success Model) โดย ทฤษฎีดังกล่าวมีปัจจัยหลัก 3 ปัจจัยได้แก่

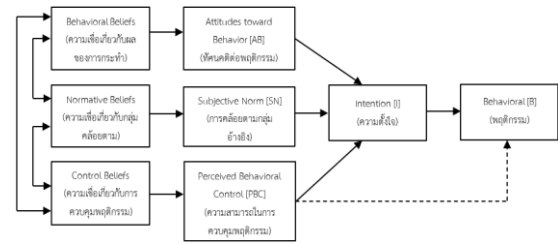
(1) ปัจจัยด้านคุณภาพของข้อมูล (Information Quality) หมายถึง ข้อมูลสารสนเทศที่สื่อสารได้ตรงกับความต้องการของผู้ใช้บริการ โดยข้อมูลที่เป็นความจริง มีความหมายรายละเอียดที่ชัดเจน ครบถ้วนสมบูรณ์ มีความถูกต้อง มาจากแหล่งข้อมูลที่เชื่อถือได้และมีความทันต่อเวลา

(2) ปัจจัยด้านคุณภาพระบบ (System Quality) หมายถึง ระบบที่มีความรวดเร็วในการให้บริการอย่างราบรื่น มีความน่าเชื่อถือของข้อมูลที่เกิดจากการประมวลผล

(3) คุณภาพของการบริการ (Service Quality) หมายถึง ความสามารถด้านการบริการตอบสนองความต้องการของผู้ใช้บริการ และการรับประกันการบริการที่มีการควบคุม คุณภาพ มาตรฐาน เพื่อสร้างความมั่นใจให้กับผู้ให้บริการ ซึ่งทั้ง 3 ปัจจัยจะเป็นปัจจัยภายนอกที่สำคัญส่งอิทธิพลให้ผู้ให้บริการรู้สึกถึงการรับรู้จนเกิดการยอมรับใช้ [1]

2.2 ทฤษฎีพฤติกรรมตามแผน (Theory of Planned Behavior: TPB)

Behavior: TPB)



ภาพที่ 2: แบบจำลองความสัมพันธ์ระหว่างปัจจัยในทฤษฎีพฤติกรรมตามแผน

ทฤษฎีพฤติกรรมตามแผน (TPB) นำเสนอโดย Ajzen หลักการของ TPB คือ การแสดงพฤติกรรมของมนุษย์ (Behavior) จะได้รับอิทธิพลโดยตรงจากความตั้งใจแสดงพฤติกรรม (Behavioral Intention, I) และสิ่งที่มีอิทธิพลต่อความตั้งใจแสดงพฤติกรรมนั้น ประกอบด้วยปัจจัยหลัก 3 ประการ ได้แก่

- (1) ทัศนคติที่มีต่อพฤติกรรมนั้น (Attitudes towards the behavior, ATT)
- (2) การคล้อยตามอิทธิพลของกลุ่มอ้างอิง (Subjective norm, SN)
- (3) การรับรู้ถึงการควบคุมพฤติกรรมของตนเองในการแสดงพฤติกรรมใดๆ (Perceived Behavioral Control, PBC) โดยการรับรู้ถึงการควบคุมพฤติกรรมนี้ ยังมีอิทธิพลโดยตรงต่อพฤติกรรมด้วยความสัมพันธ์ระหว่างปัจจัยตามทฤษฎี TPB ตามที่กล่าวข้างต้น[2]

2.3 แนวคิดเกี่ยวกับความตั้งใจซื้อ (Purchase Intention)

ความตั้งใจซื้อ หมายถึง ความตั้งใจที่จะซื้อสินค้าใดๆ ที่สามารถเข้าถึงความพึงพอใจที่ผู้บริโภคคาดหวังไว้ ซึ่งจะเกิดขึ้นหลังจากผู้บริโภคได้ประเมินตราสินค้านั้นแล้ว ซึ่งอาจก่อให้เกิดพฤติกรรมการซื้อตามมารวมถึงพฤติกรรมการกลับมาซื้อตราสินค้าเดิมในอนาคตความตั้งใจซื้อเกิดจากความรู้สึกส่วนบุคคลของผู้บริโภคที่เกิดจากการสะสมความรู้และการประเมินผ่านประสบการณ์ที่เกิดขึ้นต่อตราสินค้าและบริการจากข้อความที่ถูกส่งผ่าน ซึ่งสอดคล้องงานวิจัยที่กล่าวไว้ว่า ความตั้งใจซื้อคือ อารมณ์ที่แสดงออกถึงความต้องการในระดับเข้มข้น

ซึ่งเกิดจากการถูกกระตุ้นด้วยแรงจูงใจภายในและภายนอก (Internal and external motivation)[3]

2.4 แนวคิดเกี่ยวกับการรับรู้คุณค่า (Perceived Value)

การรับรู้คุณค่านับเป็นแนวคิดที่มีความสำคัญ เนื่องจากการรับรู้คุณค่าไม่เพียงแต่จะมีอิทธิพลต่อความตั้งใจซื้อสินค้าของผู้บริโภคเท่านั้น แต่ยังส่งผลถึงการเพิ่มขึ้นของส่วนครองตลาด (Market share) และความสามารถในการทำกำไรให้กับตราสินค้าได้อย่างมีประสิทธิภาพการรับรู้คุณค่า คือ การที่ลูกค้าประเมินอรรถประโยชน์ที่ได้รับผลิตภัณฑ์หรือการบริการบนพื้นฐานของสิ่งที่เขาได้รับรู้ โดยระบุความหมายของคุณค่านั้นไว้ว่าลูกค้าบางคนจะรับรู้ถึงคุณค่าเมื่อสินค้าหรือบริการนั้นมีราคาต่ำกว่าคนรับรู้ถึงคุณค่าเมื่อคุณภาพสอดคล้องกับราคา คุณค่าเป็นอะไรก็ตามที่ผู้บริโภคต้องการในผลิตภัณฑ์หรือบริการเป็นคุณภาพที่ได้รับจากราคาที่จ่ายไป จึงให้คำจำกัดความของการรับรู้คุณค่าของการบริการว่าเป็นการเลือกของลูกค้ำที่จะรับรู้ระหว่างสิ่งที่ลูกค้ำให้เพื่อได้รับการบริการนั้นการกับบริการที่ได้รับกลับมา[4]

2.5 แนวคิดเกี่ยวกับการรับรู้ด้านราคา (Perceived Price)

แนวคิดการรับรู้ราคา หมายถึงราคาสินค้าที่มาจาก การอนุมานของผู้บริโภคซึ่งแตกต่างจากราคาจริงของสินค้าผู้บริโภคบางคนอาจคิดว่าสินค้านี้มีราคาแพงหรือบางคนอาจคิดว่าสินค้านี้มีราคาถูก ในทำนองเดียวกันการรับรู้ราคา คือการที่ผู้บริโภคจะเปรียบเทียบราคาของสินค้าในแต่ละตราสินค้าตามความพึงพอใจของตนเองการรับรู้ราคาว่าผู้บริโภคส่วนใหญ่จะไม่จดจำราคาจริงของสินค้า แต่จะอนุมานราคาของสินค้าและจดจำความหมายที่ได้จากการอนุมานครั้งนั้น ซึ่งการรับรู้ราคาของผู้บริโภคจะส่งผลต่อการตัดสินใจซื้อสินค้า[4]

2.6 การวิเคราะห์ห้่องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis)

การวิเคราะห์ห้่องค์ประกอบเชิงสำรวจเป็นเทคนิคที่ศึกษาเกี่ยวกับโครงสร้างความสัมพันธ์ของตัวแปรอิสระ

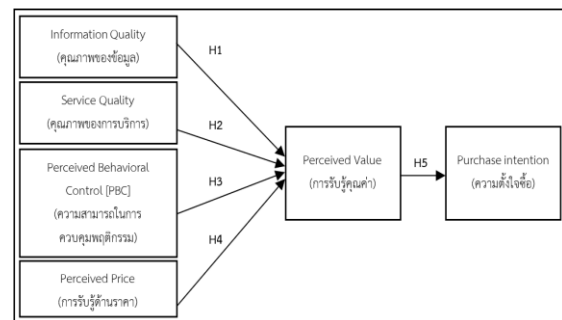
(Independent Variable) กับตัวแปรตาม (Dependent Variable) จะเป็นการศึกษาความสัมพันธ์เชิงเส้นตรง (Linearity)[5]

2.7 การวิเคราะห์ห้สมการถดถอยพหุคูณ (Multiple Regression Analysis)

การวิเคราะห์ห้การถดถอยเป็นวิธีการทางสถิติที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variable) กับตัวแปรตาม (Dependent Variable) ถ้าตัวแปรอิสระมีมากกว่าหนึ่งตัวกับตัวแปรตามหนึ่งตัว เรียกว่า การวิเคราะห์ห้การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)[6]

3. วิธีการดำเนินงาน

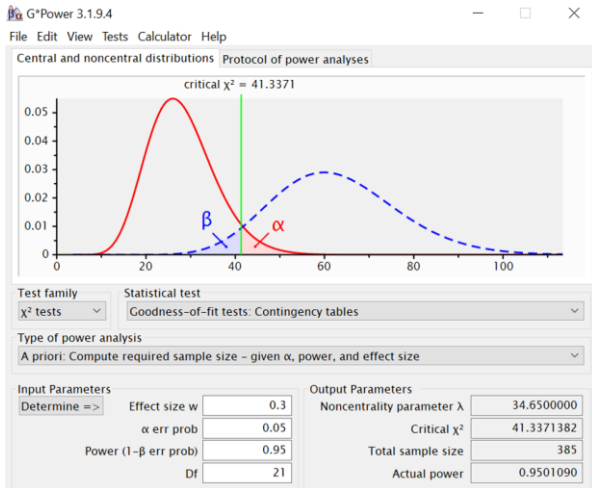
3.1 กรอบแนวคิดการวิจัย



ภาพที่ 3: กรอบแนวคิดของการวิจัยเรื่อง ปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มอัตโนมัติ

3.2 การกำหนดประชากรและกลุ่มตัวอย่าง

ผู้วิจัยได้กำหนดกลุ่มตัวอย่างด้วยใช้โปรแกรม G*Power 3.1 (Erdfelder et al., 2007) โดยกำหนดค่า effect size เท่ากับ 0.3 หมายถึงมีขนาดอิทธิพลในระดับกลาง (medium) ซึ่งส่วนใหญ่นิยมค่า effect size เท่ากับ 0.3 ที่ระดับนัยสำคัญทางสถิติเท่ากับ 0.05 ค่าอำนาจทดสอบ (power of test) เท่ากับ 0.95 (Hair et al., 2010) ค่าองศาอิสระของตัวแบบ (DF) เท่ากับ 28 ที่นำมาจากข้อคำถามจากแบบสอบถามปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มอัตโนมัติ ผลลัพธ์จากโปรแกรม G*Power ได้ขนาดกลุ่มตัวอย่างเท่ากับ 385 ซึ่งผู้วิจัยปัดเป็นตัวเลขเป็น 400 คน ดังนั้นขนาดของกลุ่มตัวอย่างที่ใช้การวิจัย คือ 400 คน



ภาพที่ 4: กำหนดขนาดกลุ่มตัวอย่างด้วยโปรแกรม G*Power

3.3 เครื่องมือที่ใช้ในการวิจัย

เก็บรวบรวมข้อมูลโดยแบบสอบถามออนไลน์ (Online Questionnaire) งานวิจัยครั้งนี้ถูกแบ่งออกเป็น 3 ตอน ได้แก่ 1. ข้อมูลพื้นฐานของผู้ตอบแบบสอบถาม 2. ชุดคำถามเกี่ยวกับพฤติกรรมการใช้เทคโนโลยีผู้คิดเครื่องดื่มอัตโนมัติ 3. คำถามวัดระดับทัศนคติของผู้ตอบแบบสอบถาม โดยใช้มาตรวัด 7 ระดับ แบบ Likert's Scale ดังความหมายต่อไปนี้ 7 หมายถึง สอดคล้องกับความคิดเห็นของท่านมากที่สุด, 6 หมายถึง สอดคล้องกับความคิดเห็นของท่านมาก, 5 หมายถึง สอดคล้องกับความคิดเห็นของท่านค่อนข้างมาก, 4 หมายถึง สอดคล้องกับความคิดเห็นของท่านปานกลาง, 3 หมายถึง สอดคล้องกับความคิดเห็นของท่านค่อนข้างน้อย, 2 หมายถึง สอดคล้องกับความคิดเห็นของท่านอย่างน้อย, 1 หมายถึง สอดคล้องกับความคิดเห็นของท่านอย่างน้อยที่สุด

3.4 การเก็บรวบรวมข้อมูล

ทางผู้วิจัยได้สร้างเครื่องมือแบบสอบถามจากการศึกษาทฤษฎี บทความและงานวิจัยที่เกี่ยวข้องต่างๆ ที่สามารถประยุกต์ใช้งานกับงานวิจัยได้ จากนั้นนำแบบสอบถามที่ผ่านการตรวจสอบไปสำรวจกับกลุ่มตัวอย่างทั้ง 400 คนผ่านแบบสอบถามออนไลน์

3.5 วิธีการวิเคราะห์ข้อมูล

1) การคัดเลือกข้อมูล (Data Selection) เพื่อให้มั่นใจว่าข้อมูลที่นำมาคำนวณมีความสมบูรณ์

2) ทำการวิเคราะห์ข้อมูลโดยใช้โปรแกรม SPSS สำหรับประมวลผลสำเร็จรูปทางสถิติ เพื่อตอบวัตถุประสงค์โดยสถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ วิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และการวิเคราะห์สมการถดถอยเชิงพหุคูณ (Multiple Linear Regression)

4. ผลการดำเนินงานวิจัย

4.1 ผลจากการวิเคราะห์ข้อมูลทั่วไป

จากการวิเคราะห์ข้อมูลทั่วไปกับกลุ่มตัวอย่าง 395 คน เนื่องจากมีผู้ไม่ประสงค์ออกความคิดเห็น 5 คน

ตารางที่ 1 : ตารางข้อมูลทั่วไป

ข้อมูลทั่วไป	จำนวน (คน)	ร้อยละ (%)
เพศ		
หญิง	248	62.7
ชาย	144	36.4
อื่นๆ	3	0.7
รวม	395	100
อายุ		
ต่ำกว่า 18 ปี	2	0.5
18 - 24 ปี	116	29.3
25 - 34 ปี	179	45.3
35 - 45 ปี	42	10.6
45 ปีขึ้นไป	56	14.1
รวม	395	100
ระดับการศึกษาสูงสุด		
ต่ำกว่าปริญญาตรี	58	14.6
ปริญญาตรี	281	71.1
ปริญญาโท	53	13.4
ปริญญาเอก	3	0.7
สูงกว่าปริญญาเอก	0	0
รวม	395	100
รายได้ต่อเดือน		
ไม่เกิน 10,000 บาท	69	17.4
10,001 - 20,000 บาท	114	28.8
20,001 - 30,000 บาท	96	24.3
30,001 - 40,000 บาท	57	14.4
40,001 - 50,000 บาท	20	5.0
50,001 บาทขึ้นไป	39	9.8
อื่นๆ	0	0.0
รวม	395	100

4.2 ผลจากการวิเคราะห์ข้อมูลทั่วไป

การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) เพื่อสำรวจหรือค้นหาตัวแปรแฝงที่ซ่อนอยู่ภายใต้ตัวแปรที่สังเกตได้หรือวัดได้ ผลการทดสอบความเหมาะสมของชุดตัวแปรที่ทำการศึกษาพบว่า สามารถจัดกลุ่มใหม่เพื่อถ่วงน้ำหนักข้อคำถามแบ่งได้ ทั้งหมด 6 ด้าน โดยในแต่ละด้านมีผลการวิเคราะห์เชิงสำรวจได้ดังนี้

ตารางที่ 2: แสดงค่า KMO และ Bartlett's Test of Sphericity

KMO and Bartlett's Test			
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.929	
Bartlett's Test of Sphericity	Approx. Chi-Square	4667.880	
	df	190	
	Sig.	0.000	

จากตารางที่ 2 จากการวิเคราะห์ข้อมูลด้วยข้อคำถามทั้งหมด 21 ตัวแปร พบว่ามีค่า KMO เท่ากับ 0.929 ซึ่งมีค่ามากกว่า 0.5 และเข้าใกล้ 1 แสดงถึงความสัมพันธ์กันของตัวแปรในขนาดที่เหมาะสม สามารถนำไปวิเคราะห์เชิงสำรวจได้ และค่าการทดสอบนัยสำคัญทางสถิติด้วย Bartlett's Test of Sphericity พบว่ามีค่านัยสำคัญทางสถิติที่ระดับ 0.05 สามารถสรุปได้ว่าข้อมูลมีความเหมาะสมในการวิเคราะห์องค์ประกอบได้

ตารางที่ 3: แสดงค่า Rotated Factor Matrix

	Rotated Factor Matrix ^a					
	Factor					
	1	2	3	4	5	6
PurchaseIntention4	0.795					
PurchaseIntention6	0.681					
PurchaseIntention3	0.648					
PurchaseIntention2	0.642					
PurchaseIntention5	0.588					
PerceivedBehavioralControl1		0.790				
PerceivedBehavioralControl3		0.745				
PerceivedBehavioralControl2		0.660				
ServiceQuality3			0.685			
ServiceQuality4			0.567			
ServiceQuality2			0.522			
PerceivedValue2				0.566		
PerceivedValue1				0.551		
PerceivedValue4				0.529		
PerceivedValue3				0.507		
PerceivedPrice2					0.820	
PerceivedPrice3					0.772	
PerceivedPrice1					0.739	
InformationQuality1						0.669
InformationQuality3						0.598

Extraction Method: Maximum Likelihood.
Rotation Method: Equamax with Kaiser Normalization.

a. Rotation converged in 37 iterations.

จากตารางที่ 3 แสดงค่าสัมประสิทธิ์ หรือที่เรียกกันว่า Factor Loadings เป็นค่าที่แสดงความสัมพันธ์ของตัวแปรกับองค์ประกอบ หรือตัวแปรทั้ง 6 ตัวโดยการหมุนแกนปัจจัยด้วยเทคนิค Maximum Likelihood และ Equamax ซึ่งได้ค่า Factor Loadings ทั้ง 6 ตัวแปรได้แก่ 1.ความตั้งใจซื้อ 2. ความสามารถในการควบคุม 3. คุณภาพของการบริการ 4. การรับรู้คุณค่า 5. การรับรู้ด้านราคา 6. ปัจจัยด้านคุณภาพของข้อมูล โดยองค์ประกอบที่ 1 มีจำนวนตัวแปรมากที่สุด 5 ตัวแปร องค์ประกอบที่ 4 มีจำนวน 4 ตัวแปร องค์ประกอบที่ 2, 3, 5 มีจำนวน 3 ตัวแปร และองค์ประกอบที่ 6 มีตัวแปรน้อยที่สุด 2 ตัวแปร

ตารางที่ 4: ผลการวิเคราะห์ค่าความเชื่อมั่น ค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐาน

ปัจจัย	ค่าเฉลี่ย	S.D	Cronbach's Alpha
PI	5.218	1.849	1.849
PBC	5.819	1.336	1.336
SVQ	5.501	1.338	1.338
PV	5.183	1.636	1.636
PP	4.786	2.387	2.387
INFQ	5.589	1.467	1.467

จากตารางที่ 4 ปัจจัย PI คือ ความตั้งใจซื้อ

ปัจจัย PBC คือ ความสามารถในการควบคุม พฤติกรรม

ปัจจัย SVQ คือ คุณภาพของการบริการ

ปัจจัย PV คือ การรับรู้คุณค่า

ปัจจัย PP คือ การรับรู้ด้านราคา

ปัจจัย INFQ คือ ปัจจัยด้านคุณภาพของข้อมูล

พบว่าผลการวิเคราะห์ความเชื่อมั่นของแบบสอบถามพบว่ามีค่าสัมประสิทธิ์ครอนบาคแอลฟา (Cronbach's alpha) อยู่ระหว่าง 1.336 ถึง 2.387 ซึ่งอยู่ในระดับที่ยอมรับได้ แสดงว่าเครื่องมือแบบสอบถามที่ใช้ในการสำรวจฉบับนี้มีความน่าเชื่อถือ

ตารางที่ 5: แสดงค่า Summary Model 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.790 ^a	0.625	0.620	0.66851	0.625	124.520	4	299	0.000

a. Predictors: (Constant), PerceivedPrice, InformationQuality, PerceivedBehavioralControl, ServiceQuality

จากตารางที่ 5 แสดงค่า Model 1 Summary จะบอกให้ทราบสมการพยากรณ์ที่ได้มีประสิทธิภาพในการพยากรณ์มากที่สุดเท่าใด โดยดูจาก R Square จะพบว่า R Square = 0.625 แสดงว่าสมการพยากรณ์ที่ได้มีประสิทธิภาพ ร้อยละ 62.5

ตารางที่ 6: ผลการวิเคราะห์การถดถอยเชิงพหุคูณ Model 1

Model	Coefficients ^a									
	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
	B	Std. Error	Beta				Lower Bound	Upper Bound	Tolerance	VIF
1 (Constant)	0.150	0.247			0.608	0.544	-0.336	0.635		
InformationQuality	0.204	0.055	0.207	3.738	0.000	0.096	0.311	0.609	2.444	
ServiceQuality	0.569	0.061	0.547	9.319	0.000	0.449	0.690	0.365	2.743	
PerceivedBehavioralControl	-0.004	0.055	-0.004	-0.068	0.946	-0.113	0.105	0.645	2.245	
PerceivedPrice	0.164	0.030	0.203	5.486	0.000	0.105	0.223	0.919	1.088	

a. Dependent Variable: PerceivedValue

จากตารางที่ 6 เมื่อพิจารณาจากค่า sig. จะสามารถสรุปได้ว่าปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีเครื่องตัดโน้มนัติ พบว่าปัจจัยด้านคุณภาพของข้อมูล, คุณภาพของการบริการ และการรับรู้ด้านราคา มีค่าเป็น Sig. เท่ากับ 0.000 ซึ่งปัจจัยทั้ง 3 ด้านนี้ มีผลต่อการรับรู้

คุณค่า ส่วนความสามารถในการควบคุมพฤติกรรม มีค่า Sig. เท่ากับ 0.946 และปัจจัยด้านนี้ไม่มีผลต่อการรับรู้คุณค่า

ตารางที่ 7: แสดงค่า Summary Model 2

Model Summary									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F	df1	df2	Sig. F Change
1	.770 ^a	0.594	0.592	0.76136	0.594	441.194	1	302	0.000

a. Predictors: (Constant), PerceivedValue

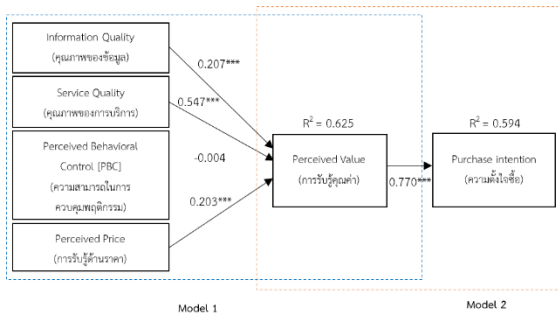
จากตารางที่ 7 แสดงค่า Model 2 Summary จะบอกให้ทราบสมการพยากรณ์ที่ได้มีประสิทธิภาพในการพยากรณ์มากน้อยแค่ไหน โดยดูจาก R Square จะพบว่า R Square = 0.594 แสดงว่าสมการพยากรณ์ที่ได้ มีประสิทธิภาพเพียงร้อยละ 59.4

ตารางที่ 8: ผลการวิเคราะห์การถดถอยเชิงพหุคูณ Model 2

Coefficients ^a										
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	0.831	0.214		3.891	0.000	0.411	1.252		
	PerceivedValue	0.847	0.040	0.770	21.005	0.000	0.768	0.927	1.000	1.000

a. Dependent Variable: PurchaseIntention

จากตารางที่ 8 เมื่อพิจารณาจากค่า sig. จะสามารถสรุปได้ว่าปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี พบว่าปัจจัยการรับรู้คุณค่ามีค่าเป็น Sig. เท่ากับ 0.000 และจากผลของการวิเคราะห์ของแต่ละตัวแปรได้ การรับรู้คุณค่าส่งผลต่อความตั้งใจซื้อ



ภาพที่ 5: โมเดลส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี

จากภาพที่ 5 คุณภาพของข้อมูล คุณภาพของการบริการ ความสามารถในการควบคุมพฤติกรรม และการรับรู้ด้านราคา พยากรณ์ได้ 62.5% (R-Squared=0.625) และการรับรู้คุณค่าส่งผลต่อความตั้งใจซื้อสามารถพยากรณ์ได้ 59.4% (R-Squared=0.594) แสดงว่าคุณภาพ

ของการบริการส่งผลมากที่สุดต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี

5. สรุปผล

งานวิจัยครั้งนี้ ผู้วิจัยได้ผลการวิเคราะห์ปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี ค่าความเหมาะสมของข้อมูล (KMO) ตัวแปร มีความสัมพันธ์ถึง 0.929 ซึ่งมีค่ามากกว่า 0.5 เป็นข้อมูลที่สามารถนำมาตั้งเป็นข้อคำถามได้ ผลน้ำหนักรองค์ประกอบมีค่ามากกว่า 0.5 มีทั้งหมด 6 องค์ประกอบประกอบไปด้วย คุณภาพของข้อมูล, คุณภาพการบริการ, ความสามารถในการควบคุมพฤติกรรม, การรับรู้ด้านราคา, การรับรู้คุณค่า และความตั้งใจซื้อ จึงนำมาทดสอบความเชื่อมั่นของแบบสอบถามกับกลุ่มตัวอย่าง 395 คน มีความเชื่อมั่นอยู่ที่ 1.336ถึง 2.387 ซึ่งอยู่ในระดับที่น่าเชื่อถือและผลวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 6 ด้าน พบว่า ความสามารถในการควบคุมพฤติกรรมที่มีค่าเฉลี่ยมากที่สุด (Mean=5.819,SD=1.336) ซึ่งผลการวิเคราะห์ถดถอยพหุคูณ ปัจจัยทุกด้านมีค่าเป็น Sig. ยกเว้นความสามารถในการควบคุมพฤติกรรมที่ไม่มีค่าเป็น sig. ปัจจัยอย่างมีระดับนัยสำคัญที่ 0.05 แสดงตัวแปรที่สามารถใช้งานได้จริง และจากผลของการวิเคราะห์ของแต่ละตัวแปรได้ค่าสัมประสิทธิ์ของสมการถดถอย (Beta Coefficient) ซึ่งเป็นค่าที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม

อย่างไรก็ตามงานวิจัยนี้เป็นเพียงส่วนหนึ่งของปัจจัยที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี ยังคงมีข้อจำกัดตามขอบเขต งานวิจัยที่กำหนด ดังนั้นหากมีการศึกษาครั้งต่อไป จึงควรปรับปรุงกรอบการวิจัยให้กว้างขึ้น จึงควรศึกษาวิจัยเพิ่มเติมเกี่ยวกับปัจจัยอื่นๆที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี เช่น ปัจจัยทางการตลาดที่ส่งผลต่อการยอมรับเทคโนโลยีผู้กดเครื่องดื่มนมดี

เอกสารอ้างอิง

[1] อธิธา มธุรสพรวัฒนา, “ทัศนคติและปัจจัยความสำเร็จของระบบสารสนเทศที่มีอิทธิพลต่อความพึง

- พอใจในการใช้บริการ Netflix ในแต่ละ Generation
”ปี 2561.
- [2] ชาญวิทย์ ตั้งสุวรรณกุล (2015) ทฤษฎีพฤติกรรมตาม
แผน (Theory of Planned Behavior : TPB)
- [3] Eagly & Chaiken (1993) Eagly, A., & Chaiken, S.
(1993). The psychology of attitudes. Orlando:
Harcourt Brace Jovanovich College Publishers.
- [4] Zeithaml, V.A.(1982). Consumer response to instore
price information environments. Journal of
Consumer Research, 8(4), 357-369.
- [5] กัลยา วานิชย์บัญชา (2551) การวิเคราะห์ข้อมูลหลาย
ตัวแปร (พิมพ์ครั้งที่ 3) กรุงเทพมหานคร บริษัทธรรม
สาร
- [6] สถิติ และการวิเคราะห์ข้อมูลในงานวิจัยเบื้องต้น.
สุทิน ชนะบุญ

ปัจจัยที่มีอิทธิพลต่อทัศนคติด้านการยอมรับความเป็นส่วนตัวของข้อมูล ผู้ใช้บริการโบายแบงก์กิ้ง

Factors Affecting Attitude of Data Privacy Acceptance of Mobile Banking Users

วัลย์ลักษณ์ เกตุ ไชยสง (Walailak Gattaisong)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856097@email.kmutnb.ac.th, ²sakchai.t@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อทัศนคติด้านการยอมรับความเป็นส่วนตัวของข้อมูล (Data Privacy) ผู้ใช้บริการโบายแบงก์กิ้ง โดยใช้แบบสอบถามเป็นเครื่องมือวิจัย และเก็บข้อมูลจากผู้เคยมีการใช้งานโบายแบงก์กิ้งสถาบันการเงินต่าง ๆ จำนวน 270 คน ซึ่งข้อมูลงานวิจัยนี้ใช้สถิติในการวิเคราะห์ข้อมูล ได้แก่ สถิติพื้นฐาน การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis :EFA) และการวิเคราะห์สมการถดถอยเชิงพหุคูณ (Multiple Linear Regression) ผลการวิจัยพบว่าปัจจัยที่ส่งผลต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูลของผู้ใช้บริการ โบายแบงก์กิ้งมากที่สุดคือการรับรู้ความง่ายในการใช้งาน รองลงมาคือการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัวของข้อมูล สำหรับปัจจัยที่ไม่ส่งผลต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูลของผู้ใช้บริการ โบายแบงก์กิ้งได้แก่ ด้านการรับรู้ชื่อเสียงของผู้ให้บริการ และด้านการรับรู้ความเสี่ยงความเป็นส่วนตัว ซึ่งผลสรุปจากงานวิจัยนี้มีค่าที่ยอมรับได้ที่ร้อยละ 56 ($R\text{-Squared} = 0.561$)

คำสำคัญ: โบายแบงก์กิ้ง ความเป็นส่วนตัวของข้อมูล

Abstract

The purpose of this research was to study factors affecting attitude of data privacy acceptance of mobile banking users. This research tool used a questionnaire to collect data from 270 people who used mobile

banking services from various financial institutions. The data of this research was analyzed with basic statistics, exploratory factor analysis (EFA), and multiple linear regression analysis. The research results revealed that the factor most affecting the attitude of data privacy acceptance of mobile banking users were the perception of ease of use. And followed by the perceived effectiveness of the data privacy policy. The factors that do not affect the acceptance of data privacy attitudes of mobile banking users were the reputation of the service provider and data privacy risk perception. The results of this research have an acceptable value of 56% ($R\text{-squared} = 0.561$)

Keywords: Mobile Banking, Data Privacy, Personal Data Protection Act.

1. บทนำ

ปัจจุบันการทำธุรกรรมกับทางธนาคาร ส่วนใหญ่จะใช้ผ่านระบบโบายแบงก์กิ้ง (Mobile Banking) จากเดิมที่เคยใช้บริการผ่านสาขาหรือตู้เอทีเอ็ม ซึ่งมีข้อจำกัดด้านเวลาและระยะทาง กลายเป็นความสะดวก รวดเร็ว ด้วยเทคโนโลยีที่เอื้ออำนวย และพฤติกรรมของผู้บริโภคที่คุ้นชินกับการใช้สมาร์ทโฟน เริ่มถือเงินสดน้อยลง

ในช่วง 2-3 ปีที่ผ่านมา มีการทำธุรกรรมผ่านระบบโบายแบงก์กิ้งของธนาคารมากขึ้น [1] ด้วยแรงสนับสนุนจากภาครัฐผ่านแผนยุทธศาสตร์การพัฒนาโครงสร้างพื้นฐานระบบการชำระเงินแบบอิเล็กทรอนิกส์แห่งชาติ

(National e-Payment Master Plan) [2] จะเป็นส่วนสำคัญที่จะผลักดันให้เกิดสังคมไร้เงินสด (Cashless Society) ในประเทศไทยเร็วขึ้น โดยมีธนาคารขับเคลื่อน

ผู้วิจัยจึงมีความสนใจที่จะศึกษาปัจจัยที่มีผลกระทบต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูลของผู้ใช้บริการ 모바일แบงก์กิ้ง ในส่วนข้อมูลส่วนบุคคลของผู้ใช้บริการที่ธนาคารต้องปฏิบัติตาม พ.ร.บ.คุ้มครองข้อมูลส่วนบุคคล (PDPA: Personal Data Protection Act) ว่าปัจจัยใดส่งผลกระทบต่อความยินยอมของผู้ใช้บริการ ทั้งยังเป็นแนวทางสำหรับผู้ที่อยู่ในอุตสาหกรรมธนาคารสามารถนำข้อมูลที่ได้เป็นแนวทางไปส่งเสริมและพัฒนากลยุทธ์เพื่อตอบสนองความต้องการของผู้ใช้บริการมากขึ้น

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 โมบายแบงก์กิ้ง (Mobile Banking)

เป็นแอปพลิเคชันของสถาบันการเงินที่ให้บริการธุรกรรมออนไลน์ผ่านโทรศัพท์เคลื่อนที่ เช่น การโอนเงินระหว่างบัญชีธนาคาร การตรวจสอบยอดบัญชี การซื้อขายตราสาร/กองทุน การตั้งระบบแจ้งเตือนอัตโนมัติ เบิกถอนโอน จ่ายบิล เป็นต้น [3]

2.2 การรับรู้ความง่ายในการใช้งาน

คือการรับรู้ความง่ายในการให้ความยินยอมข้อมูลส่วนบุคคลผ่าน โมบายแบงก์กิ้ง ว่าทำได้ง่าย สะดวก และสามารถศึกษานโยบายความเป็นส่วนตัวของธนาคารต่างๆ ได้เองผ่านระบบ โมบายแบงก์กิ้งที่มีการพัฒนาขึ้นก่อนตัดสินใจเลือก “ยอมรับ” หรือ “ปฏิเสธ” การให้ข้อมูล รวมถึงสามารถเปลี่ยนแปลงการให้ความยินยอมด้วยตัวเองได้ตลอดเวลา โดยไม่ต้องใช้ความพยายามในการเรียนรู้ หรือทำความเข้าใจระบบมากนัก [4]

2.3 ทัศนคติที่มีต่อการใช้งาน

ความเชื่อว่าการใช้เทคโนโลยี โมบายแบงก์กิ้ง สำหรับการเก็บข้อมูลส่วนบุคคลมีประโยชน์ เหมาะสม เก็บข้อมูลส่วนบุคคลได้ปลอดภัย และผู้ให้บริการมีความพึงพอใจในการเก็บข้อมูลผ่านการใช้งาน โมบายแบงก์กิ้ง หากผู้งานมีทัศนคติที่ดีการยอมรับความเป็นส่วนตัวของข้อมูล จะส่งผลกระทบต่อการใช้งานระบบ โมบายแบงก์กิ้งต่อไป [5]

2.4 การรับรู้ความเสี่ยงความเป็นส่วนตัวของข้อมูล

การรับรู้ถึงความเสี่ยงด้านความเป็นส่วนตัว หมายถึง การที่ผู้ใช้บริการมีความเข้าใจว่าข้อมูลส่วนบุคคลถูกรู้ว่าเป็นส่วนตัว ถูกเข้าถึง ติดตาม และนำไปใช้โดยไม่รู้ตัว ซึ่งการพัฒนาการของอินเทอร์เน็ตและแอปพลิเคชันบนสมาร์ตโฟน เป็นช่องทางการเข้าถึงข้อมูลส่วนบุคคลจากปัจจัยภายนอก หากไม่ได้รับการจัดการอย่างเหมาะสม ข้อมูลส่วนบุคคลอาจนำไปสู่การคุกคามที่ร้ายแรงได้ [6] การถูกละเมิดความเป็นส่วนตัว ทำให้ผู้ใช้บริการเกิดความไม่พึงพอใจ และยกเลิกใช้บริการไปในที่สุด ความเสี่ยงด้านความเป็นส่วนตัวของข้อมูลนี้ยังรวมถึงการถูกขโมยข้อมูลส่วนบุคคลเช่น หมายเลขบัตรเครดิต หมายเลขบัตรประชาชน เลขที่บัญชีธนาคาร [7]

หากทางผู้ให้บริการมีมาตรฐานรองรับ มีความชัดเจนในวัตถุประสงค์ในการขอข้อมูล สามารถตรวจสอบได้ ช่วยทำให้ผู้ใช้บริการลดความกังวลและเกิดทัศนคติที่ดีต่อการใช้งาน โมบายแบงก์กิ้งมากยิ่งขึ้น

2.5 การรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัวของข้อมูล

เป็นสิ่งที่ผู้ใช้บริการ โมบายแบงก์กิ้งเชื่อว่านโยบายความเป็นส่วนตัว (Privacy Policy) ของผู้ใช้งานที่มีการประกาศผ่านทางโมบายแบงก์กิ้งนั้นมีความถูกต้องและมีความน่าเชื่อถือ ทำให้ผู้ใช้บริการมีความมั่นใจว่าข้อมูลที่ยินยอมให้จัดเก็บมีความปลอดภัย และถูกนำไปใช้งานได้ตรงตามจุดประสงค์ รวมถึงนโยบายความเป็นส่วนตัวที่ประกาศออกมานำไปปฏิบัติให้เกิดผลได้จริง [8] รวมถึงต้องแจ้งให้เจ้าของข้อมูลส่วนบุคคลก่อนที่จะมีการเก็บข้อมูล ว่าตัวของข้อมูลส่วนบุคคลที่จะมีการจัดเก็บมีอะไรบ้าง ใช้ข้อมูลอย่างไร จัดเก็บไว้ที่ไหน [9] แจ้งสิทธิของเจ้าของข้อมูลส่วนบุคคล ซึ่งเป็นสิทธิตามกฎหมาย

การเน้นย้ำถึงความสำคัญนโยบายความเป็นส่วนตัวคือการ “ประกาศ” หากผู้ให้บริการมีการรับประกันว่าข้อมูลที่ผู้บริการให้มีความปลอดภัย และไม่ถูกนำไปใช้ในการที่ผิดวัตถุประสงค์ ผู้ใช้บริการจะเกิดความเชื่อถือในประสิทธิภาพนโยบายความเป็นส่วนตัวของข้อมูล

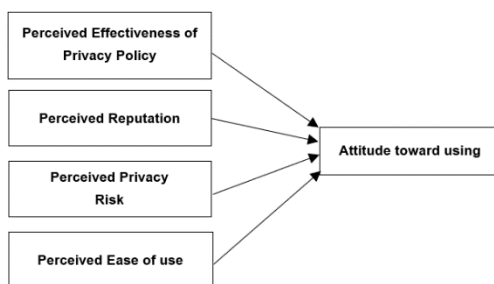
นอกจากนี้ ในกรณีที่ต้องขอเก็บรวบรวมข้อมูลส่วนบุคคลต้องให้อิสระเจ้าของข้อมูลในการเลือกให้ความยินยอม

2.6 การรับรู้ชื่อเสียง

การรับรู้ชื่อเสียงองค์กร คือการรับรู้คุณลักษณะหรือภาพลักษณ์ที่องค์กรส่งสมมาอย่าง ต่อเนื่องยาวนานตั้งแต่อดีตจนถึงปัจจุบัน[10]ทั้งในด้านบวกและด้านลบที่มีผลต่อความเชื่อมั่นของการใช้บริการ การรับรู้ถึงมาตรฐานหรือมาตรการต่างๆ ที่เกี่ยวกับการเก็บ การใช้หรือการประมวลผลข้อมูลส่วนบุคคลของผู้ใช้บริการผ่านนโยบายแบงก์กิ้ง และการที่ธนาคารฯ รับข้อกำหนดจากกฎหมายที่ว่าด้วยการแจ้งให้ผู้ให้บริการรับทราบถึงสิทธิของเจ้าของข้อมูล แสดงออกถึงความมุ่งมั่นที่จะปกป้องข้อมูลส่วนบุคคลของผู้ใช้บริการ และหลีกเลี่ยงความเสี่ยงจากผลกระทบจากการสูญเสียชื่อเสียงจากการไม่ปฏิบัติตามกฎหมายอีกด้วย

ดังนั้นในงานวิจัยที่เกี่ยวข้องของผู้วิจัยเลือกใช้ทฤษฎีการยอมรับเทคโนโลยี (Technology Acceptance Model: TAM) 2 ปัจจัย คือการรับรู้ความง่ายในการใช้งาน และทัศนคติที่มีต่อการใช้งาน และปัจจัยอื่น ๆ คือการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัวของข้อมูล, การรับรู้ความเสี่ยงความเป็นส่วนตัวของข้อมูล, การรับรู้ชื่อเสียง เพื่อหาว่าปัจจัยใดมีอิทธิพลต่อทัศนคติด้านการยอมรับความเป็นส่วนตัวของข้อมูลผู้ใช้บริการ โบนัสแบงก์กิ้ง

3. วิธีดำเนินการวิจัย



ภาพที่ 2 : กรอบแนวความคิดการดำเนินงานวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

การกำหนดขนาดของตัวอย่างได้ใช้หลักการของ Jackson จากกฎ $N:q [10]$ เมื่อ N คือ ขนาดของกลุ่มตัวอย่าง และ q คือ จำนวนพารามิเตอร์ โดยโมเดลที่ต้องการ

วิเคราะห์สมการเชิงโครงสร้าง ควรมีขนาดตัวอย่างขั้นต่ำเท่ากับ 10:1 ดังนั้นงานวิจัยนี้มีจำนวนพารามิเตอร์ที่ศึกษาจำนวน 15 พารามิเตอร์ ขนาดตัวอย่างของงานวิจัยนี้ควรมีอย่างน้อยที่สุด 150 ตัวอย่าง และเพื่อป้องกันความผิดพลาดผู้วิจัยจึงได้ทำการเพิ่มขนาดกลุ่มตัวอย่างขึ้นคือ 270 ราย

3.2 การเก็บรวบรวมข้อมูล

เก็บรวบรวมข้อมูลโดยการแจกแบบสอบถามออนไลน์ (Google Form) แบบสอบถามที่ใช้งานวิจัยครั้งนี้ถูกแบ่งออกเป็น 3 ส่วน ได้แก่ 1. ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม 2. ชุดคำถามเกี่ยวกับการรับรู้ข่าวสารและพฤติกรรมในการทำธุรกรรมกับธนาคาร 3. เป็นคำถามเกี่ยวกับปัจจัยที่มีอิทธิพลต่อทัศนคติด้านการยอมรับความเป็นส่วนตัวของข้อมูลผู้ใช้บริการ โบนัสแบงก์กิ้ง โดยใช้มาตรประเมินค่า 7 ระดับ (7 Point Likert Scales) โดยการวัดข้อมูลประเภทอันตรภาคชั้นแบ่งเป็น 7 ระดับ ได้แก่ 1 = น้อยที่สุด 2 = น้อย 3 = ค่อนข้างน้อย 4 = ปานกลาง 5 = ค่อนข้างมาก 6 = มาก และ 7 = มากที่สุด

3.3 วิธีการวิเคราะห์ข้อมูล

ผู้วิจัยได้นำข้อมูลจำนวนรายการที่ผ่านการตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูลเป็นที่เรียบร้อยแล้วนำข้อมูลที่ได้นำมาทำการแยกหมวดหมู่เพื่อวิเคราะห์โดยทำการวิเคราะห์ข้อมูลด้วย โปรแกรมประมวลผลสำเร็จรูปทางสถิติ SPSS

3.3.1) การคัดเลือกข้อมูล (Data Selection) เพื่อให้มั่นใจว่าข้อมูลที่นำมาคำนวณมีความสมบูรณ์

3.3.2) นำข้อมูลทั่วไปของประชากรตัวอย่างมาวิเคราะห์โดยการหาค่าการแจกแจงความถี่ (Frequency Distribution) ค่าร้อยละ (Percentage) ค่าเฉลี่ย (Mean) และค่าความเบี่ยงเบนมาตรฐาน (Standard Deviation)

3.3.3) ทำการวิเคราะห์ข้อมูล โดยใช้โปรแกรม SPSS สำหรับประมวลผลสำเร็จรูปทางสถิติ เพื่อตอบวัตถุประสงค์โดยสถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ วิเคราะห์องค์ประกอบเชิงสำรวจ (EFA) และการวิเคราะห์สมการถดถอยเชิงพหุคูณ (Multiple Linear Regression)

4. ผลการดำเนินการวิจัย

4.1 ผลจากการวิเคราะห์ข้อมูลทั่วไป

ตารางที่ 1: ตารางข้อมูลทั่วไป

ข้อมูลทั่วไป		จำนวน(คน)	ร้อยละ %
เพศ	หญิง	150	56
	ชาย	120	44
	รวม	270	100
อายุ	ต่ำกว่า 25 ปี	12	5
	25-40 ปี	184	68
	41-50 ปี	66	24
	51 ปีขึ้นไป	8	3
	รวม	270	100
อาชีพ	นักเรียนนักศึกษา	7	2
	รับราชการ/รัฐวิสาหกิจ	34	13
	พนักงานบริษัทเอกชน	183	68
	ธุรกิจส่วนตัว	37	14
	อื่นๆ	9	3
	รวม	270	100
ระดับการศึกษาสูงสุด	ต่ำกว่าปริญญาตรี	40	15
	ปริญญาตรี	163	60
	ปริญญาโท	63	23
	สูงกว่าปริญญาโท	4	2
	รวม	270	100

4.2 ผลการตรวจสอบความเหมาะสมของข้อมูล

ตารางที่ 2: KMO และ Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.859
Bartlett's Test of Sphericity	Approx. Chi-Square	2312.62
	df	105
	Sig.	0

จากตารางที่ 2 การวิเคราะห์ ข้อมูลด้วยข้อคำถามทั้งหมด 15 ตัวแปร พบว่ามีค่า KMO เท่ากับ 0.859 ซึ่งมีค่ามากกว่า 0.5 และเข้าใกล้ 1 แสดงถึงความสัมพันธ์กันของตัวแปรในขนาดที่เหมาะสม สามารถนำไปวิเคราะห์เชิงสำรวจได้ และค่าการทดสอบนัยสำคัญทางสถิติด้วย Bartlett's Test of Sphericity พบว่ามีค่านัยสำคัญทางสถิติที่ระดับ 0.05 สามารถสรุปได้ว่าข้อมูลมีความเหมาะสมในการวิเคราะห์องค์ประกอบ

ตารางที่ 3: Rotated Factor Matrix

Rotated Factor Matrix ²						
	Factor					
	1	2	3	4	5	6
PPR2		0.973				
PPR1		0.760				
PPR3		0.733				
PEPP2			0.755			
PEPP1			0.703			
PEPP3			0.504			
PEU2				0.886		
PEU1				0.664		
PEU3				0.501		
ATU3					0.777	

ATU4					0.534	
ATU2					0.517	
ATU1					0.515	
PR1						0.960
PR2						0.559

จากตารางที่ 3 Rotated Factor Matrix แสดงค่าสัมประสิทธิ์ หรือที่เรียกกันว่า Factor Loadings เป็นค่าที่แสดงความสัมพันธ์ของตัวแปรกับองค์ประกอบหรือพบว่าค่า Factor Loadings ทั้ง 5 ตัวแปรได้แก่ 1. ด้านการรับรู้ความเสี่ยงความเป็นส่วนตัว (Perceived Privacy Risk) 2. ด้านการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว (Perceive Effectiveness of Privacy Policy) 3. ด้านการรับรู้ความง่ายในการใช้งาน (Perceived Ease of Use) 4. ด้านทัศนคติต่อการใช้งาน (Attitude Toward Using) 5. ด้านการรับรู้ชื่อเสียง (Perceived Reputation) โดยองค์ประกอบที่ 4 มีจำนวนตัวแปรมากที่สุด 4 ตัวแปร องค์ประกอบที่ 1,2,3 มีจำนวน 3 ตัวแปร และองค์ประกอบที่ 5 มีตัวแปรน้อยที่สุด 2 ตัวแปร

4.3 ผลจากการทดสอบความเชื่อมั่นของแบบสอบถาม

ตารางที่ 4: ทดสอบเครื่องมือจากการสำรวจกลุ่มตัวอย่างทั้งหมด 270 คน เพื่อทดสอบ ค่าเฉลี่ย ส่วนเบี่ยงเบน และวิเคราะห์ความเชื่อมั่นของแบบสอบถาม

Factor	ค่าเฉลี่ย	S. D	Alpha
PPR	4.854	2.153	0.859
PEPP	4.753	1.603	0.841
PEU	5.243	1.249	0.827
ATU	4.831	1.531	0.846
PR	5.293	1.409	0.783

โดยที่

- PPR หมายถึง ด้านการรับรู้ความเสี่ยงความเป็นส่วนตัว
- PEPP หมายถึง ด้านการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว
- PEU หมายถึง ด้านการรับรู้ความง่ายในการใช้งาน
- AUT หมายถึง ด้านทัศนคติต่อการใช้งาน
- PR หมายถึง ด้านการรับรู้ชื่อเสียง

พบว่าผลการวิเคราะห์ ความเชื่อมั่นของแบบสอบถาม พบว่ามีค่าสัมประสิทธิ์ครอนบาคแอลฟา (Cronbach's alpha) อยู่ ระหว่าง 0.783 ถึง 0.88 ซึ่งอยู่ในระดับที่ยอมรับได้

แสดงว่าเครื่องมือแบบสอบถามที่ใช้ในการสำรวจฉบับนี้มีความน่าเชื่อถือ และผลวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 5 ด้าน พบว่าการรับรู้ถึงความด้านารรับรู้ชื่อเสียงมีค่าเฉลี่ยมากที่สุดและด้านการรับรู้ความเลียงความเป็นส่วนตัวมีส่วนเบี่ยงเบนมาตรฐานมากที่สุด (Mean = 5.293, SD = 2.153)

4.4 ผลการวิเคราะห์การถดถอยเชิงพหุคูณ

ตารางที่ 5: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.749 ^a	.561	.555	.69	.561	84.826	4	265	0

a Predictors: (Constant), PR, PPR, PEU, PEPP

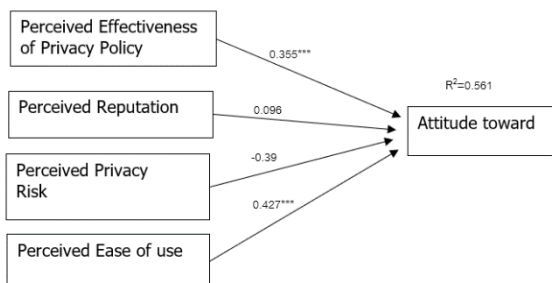
ผลลัพธ์ที่แสดงออกมาจากตาราง Model Summary จะบอกถึงสมการพยากรณ์ที่ได้มีประสิทธิภาพในการพยากรณ์มากขึ้นแค่ไหน โดยดูจาก R Square ซึ่งจากตารางที่ 5 นี้ R Square = 0.561 แสดงว่าสมการพยากรณ์ที่ได้ มีประสิทธิภาพเพียงร้อยละ 56

ตารางที่ 6: Coefficients Model

Model	Unstandardized Coefficients			Standardized Coefficients Beta	Collinearity Statistics			
	B	Std. Error	t		Sig.	Tolerance	VIF	
1	(Constant)	0.524	0.309	1.693	0.092			
	PPR	-0.024	0.026	-0.039	-0.925	0.356	0.917	1.091
	PEPP	0.331	0.053	0.355	6.297	0	0.521	1.918
	PEU	0.456	0.052	0.427	8.692	0	0.685	1.46
	PR	0.091	0.049	0.096	1.843	0.066	0.606	1.65

a Dependent Variable: ATU

จากตารางที่ 6 พบว่าด้านการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว และด้านการรับรู้ความง่ายในการใช้งาน มีค่าเป็น Sig. เท่ากับ 0.000 ซึ่งปัจจัยทั้ง 2 ด้านนี้มีผลต่อทัศนคติต่อการใช้งานส่วนด้านการรับรู้ความเลียงความเป็นส่วนตัว และด้านการรับรู้ชื่อเสียงมีค่า Sig. เท่ากับ 0.356, 0.066 ตามลำดับ ดังนั้นปัจจัยทั้ง 2 ด้านนี้ไม่มีผลต่อทัศนคติต่อการใช้งาน



ภาพที่ 3: โมเดลที่อธิบายปัจจัยที่มีผลกระทบต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล (Data Privacy) ของผู้ใช้บริการ โฆษณาเบงกี้

จากภาพที่ 3 จากปัจจัยต่างๆที่ได้นำเป็นวิเคราะห์ทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล (Data Privacy)

ของผู้ใช้บริการ โฆษณาเบงกี้สามารถพยากรณ์ได้ 56% (R-Squared= 0.561)

5. สรุปผล

งานวิจัยครั้งนี้ ผู้วิจัยได้ผลการวิเคราะห์ปัจจัยที่ส่งผลกระทบต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูลของผู้ใช้บริการ โฆษณาเบงกี้ค่าความเหมาะสมของข้อมูล (KMO) ตัวแปรมีความสัมพันธ์ถึง 0.859 ซึ่งมีค่ามากกว่า 0.5 เป็นข้อมูลที่สามารถนำมาตั้งเป็นข้อคำถามได้ ผลน้ำหนักรององค์ประกอบมีค่ามากกว่า 0.5 มีทั้งหมด 5 องค์ประกอบ มาทดสอบความเชื่อมั่นของแบบสอบถามกับกลุ่มตัวอย่าง 270 ตัวอย่าง มีความเชื่อมั่นอยู่ที่ 0.783 ถึง 0.88 ซึ่งอยู่ในระดับที่น่าเชื่อถือและผลวิเคราะห์ค่าส่วนเบี่ยงเบนมาตรฐานทั้ง 5 ด้าน พบว่าการรับรู้ถึงความด้านารรับรู้ชื่อเสียงมีค่าเฉลี่ยมากที่สุดและด้านการรับรู้ความเลียงความเป็นส่วนตัวมีส่วนเบี่ยงเบนมาตรฐานมากที่สุด (Mean = 5.293, SD = 2.153) จากการวิเคราะห์การถดถอยเชิงพหุคูณพบว่าด้านการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว และด้านการรับรู้ความง่ายในการใช้งาน มีค่าเป็น Sig. เท่ากับ 0.000 ซึ่งปัจจัยทั้ง 2 ด้านนี้มีผลต่อทัศนคติต่อการใช้งาน 56% (R-Squared= 0.561) ส่วนปัจจัยด้านการรับรู้ความเลียงความเป็นส่วนตัว และด้านการรับรู้ชื่อเสียงมีค่า Sig. เท่ากับ 0.356 และ 0.066 ตามลำดับ สามารถอภิปรายผลแต่ละด้านได้ดังต่อไปนี้

ด้านที่ 1 ปัจจัยการรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว ผลการวิเคราะห์พบว่า ผู้ใช้บริการเห็นว่าประกาศนโยบายความเป็นส่วนตัวบนโฆษณาเบงกี้ของธนาคาร เป็นวิธีที่สื่อถึงกับผู้ใช้บริการได้ง่าย และรวดเร็ว เป็นการแสดงถึงกระบวนการในการปกป้องข้อมูลส่วนบุคคลของของผู้ใช้บริการ เพื่อสร้างความมั่นใจในการใช้บริการ ทั้งนี้ในด้านดังกล่าวส่งผลกระทบต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล

ด้านที่ 2 ด้านการรับรู้ชื่อเสียง ผลการวิเคราะห์พบว่าจากการหาปัจจัยด้านภาพลักษณ์ของธนาคาร ในการทำธุรกิจอย่างยาวนานไม่ส่งผลกระทบต่อความเชื่อมั่นในการเก็บ, การใช้หรือการประมวลผลข้อมูลส่วนบุคคลของ

ผู้ใช้บริการผ่าน โฆษณาเบงกั้ง ทั้งนี้ในด้านดังกล่าวไม่ได้ส่งผลต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล เนื่องจากจากระบบโฆษณาเบงกั้ง เป็นสิ่งอำนวยความสะดวกให้กับผู้ใช้งาน ในการทำธุรกรรมอยู่แล้ว โดยผู้ใช้บริการไม่ได้ตัดสินใจใช้งานจากชื่อเสียงองค์กร และอีกสาเหตุคือ การผูกบัญชีเงินเดือนจากทางบริษัทที่กำหนดธนาคารที่ต้องเปิดบัญชีเพื่อรับเงินเดือน ดังนั้นสาเหตุการเลือกใช้บริการจึงไม่ได้ขึ้นอยู่กับทัศนคติการใช้บริการจากตัวผู้ใช้งานเอง

ด้านที่ 3 ปัจจัยการรับรู้ความเสี่ยงความเป็นส่วนตัว ผลการวิเคราะห์พบว่า ในการให้ความยินยอมในการเปิดเผยข้อมูลส่วนบุคคลเพื่อวัตถุประสงค์ต่างๆของทางธนาคาร ไม่ได้ส่งผลต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล

ด้านที่ 4 ปัจจัยการรับรู้ความง่ายในการใช้งาน ผลการวิเคราะห์พบว่า ในการยินยอมการให้ข้อมูลส่วนบุคคลผ่าน โฆษณาเบงกั้ง ว่าทำได้ง่าย สะดวก และสามารถศึกษา นโยบายความเป็นส่วนตัวของธนาคารต่างๆได้เองผ่านระบบโฆษณาเบงกั้งที่มีการพัฒนาขึ้น ก่อนตัดสินใจเลือก “ยอมรับ” ซึ่งส่งผลต่อทัศนคติการยอมรับเป็นส่วนตัวของข้อมูล

ซึ่งผลจากการวิเคราะห์การรับรู้ความเสี่ยงความเป็นส่วนตัว การรับรู้ประสิทธิภาพของนโยบายความเป็นส่วนตัว การรับรู้ความง่ายในการใช้งาน และการรับรู้ชื่อเสียง สามารถพยากรณ์ได้ 56% (R-Squared= 0.561) อย่างไรก็ตาม หากมีการศึกษาครั้งต่อไป จึงควรปรับปรุงกรอบการวิจัยให้กว้างขึ้น โดยเพิ่มการศึกษาปัจจัยอื่นๆที่สามารถนำมาวิเคราะห์ต่อยอดได้ เช่น พรบ.คุ้มครองข้อมูลส่วนบุคคล เป็นต้น

เอกสารอ้างอิง

- [1] ธนาคารแห่งประเทศไทย [สืบค้นวันที่ 1 กุมภาพันธ์ 2566]. จ 1 ก https://www.bot.or.th/App/BTWS_STAT/statistics/ReportPage.aspx?reportID=949&language=th
- [2] National e-Payment [สืบค้นวันที่ 1 กุมภาพันธ์ 2566]. จ 1 ก <http://www.epayment.go.th/home/app/overview>.
- [3] Umayplus [สืบค้นวันที่ 1 กุมภาพันธ์ 2566].จาก [https://www.umayplus.com/content/blogใช้โฆษณาเบงกั้ง อย่างไรก็ตาม ปล่อยให้ปลอดภัย/blogdetail?blogid=2022-010000007](https://www.umayplus.com/content/blogใช้โฆษณาเบงกั้ง อย่างไรก็ตาม ปล่อยให้ปลอดภัย/blogdetail?blogid=2022-010000007)
- [4] พงษ์ พิมพ์ระเบียบ, “อิทธิพลกำกับของการรับรู้ความเสี่ยงออนไลน์ด้านความเป็นส่วนตัวส่วนบุคคลในความสัมพันธ์ระหว่างปัจจัยด้านลักษณะของผู้ประกอบการค้าและความเต็มใจในการเปิดเผยข้อมูลส่วนบุคคล” *จุฬาลงกรณ์ธุรกิจปริทัศน์* ปีที่ 43 ฉบับที่ 170 ต.ค. -ธ.ค. 2564 หน้า 38-59.
- [5] สิทธิชัย ญาษาแก้ว, “ปัจจัยการยอมรับเทคโนโลยีที่มีผลต่อการตัดสินใจใช้บริการ Grab ของผู้บริโภคในกรุงเทพมหานคร” (ค้นคว้าอิสระปริญญาบริหารธุรกิจมหาบัณฑิต สาขาวิชาบริหารธุรกิจบัณฑิตวิทยาลัย).มหาวิทยาลัยกรุงเทพ.2561
- [6] Joohee, L., Songmi, K., Wonjoon,K. “The Effects of Consumers' Perceived Privacy Control on Perceived Privacy Risk in Location-Based Services.” *International Journal of Contents* [serial online] 2017. Vol13 No.1:22-30
- [7] Bianchi, C., & Andrews, L. “Risk, trust, and consumer online purchasing behavior: A Chilean perspective”. *International Marketing Review*. 3(2012) : 253-276
- [8] สมาคมธนาคารไทย, แนวปฏิบัติการคุ้มครองข้อมูลส่วนบุคคลภาคธุรกิจธนาคาร. เผยแพร่ ณ วันที่ 28 เมษายน 2564
- [9] ศูนย์วิจัยกฎหมายและการพัฒนา, คณะนิติศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. Thailand Data Protection Guidelines 3.2 แนวปฏิบัติเกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคล. พิมพ์ครั้งที่ 1. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2564

- [10] ศุภมิตร ลือรุ่งโรจน์วุฒิ. ศึกษาปัจจัยด้านองค์กรที่มีผลต่อความพึงพอใจในการเลือกทำงานของพนักงานระดับหัวหน้างานในเขตพื้นที่นิคมอุตสาหกรรมแหลมฉบัง. (การค้นคว้าอิสระ สาขาวิชาบริหารธุรกิจ). มหาวิทยาลัยบูรพา.2559
- [11] กนกวรรณ บุญวันต์, “ปัจจัยการตลาดออนไลน์ที่มีผลต่อทัศนคติในการใช้ Café & Working Space”. (การค้นคว้าอิสระ สาขาวิชาระบบสารสนเทศเพื่อการจัดการ).มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.2563

การพัฒนาเครื่องมืออัตโนมัติสำหรับการสร้างแบบจำลองความสัมพันธ์ของส่วนต่อประสานโปรแกรมประยุกต์เว็บเซอร์วิสแบบเรสต์ฟูล

Development of an Automatic Tools for Modeling the Relationships of RESTful API Web Service

วิภาดา กลิ่งเทศ (Wipada Klungtes)¹ และเนืองวงศ์ ทวยเจริญ (Nuengwong Tuaycharoen)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹6470275321@student.chula.ac.th, ²nuengwong.t@chula.ac.th

บทคัดย่อ

ในอุตสาหกรรมซอฟต์แวร์ นิยมนำซอฟต์แวร์กลับมาใช้ใหม่เป็นจำนวนมาก เนื่องจากการลดต้นทุนในการพัฒนาซอฟต์แวร์ เอกสารต่าง ๆ ในการพัฒนาระบบจึงมีความสำคัญในการอ้างอิง งานวิจัยนี้ให้ความสนใจกับแผนภาพยูเอ็มแอลที่เป็นส่วนหนึ่งของเอกสารส่วนต่อประสานโปรแกรมประยุกต์ จึงได้นำเสนอการออกแบบและพัฒนาเครื่องมืออัตโนมัติสำหรับการสร้างแบบจำลองความสัมพันธ์ของส่วนต่อประสานโปรแกรมประยุกต์เว็บเซอร์วิสแบบเรสต์ฟูล โดยใช้เครื่องมือเสริม PlantUML ดังนั้น เมื่อนำเครื่องมือที่พัฒนาขึ้นไปประยุกต์ใช้ จะช่วยให้ปรับปรุงเอกสารส่วนต่อประสานโปรแกรมประยุกต์ได้ง่ายยิ่งขึ้น และยังช่วยให้เอกสารตรงกับรหัสต้นฉบับ จากการประยุกต์ใช้เครื่องมือกับโครงการ ทำให้ได้ผลลัพธ์ของความถูกต้องเป็น 100%

คำสำคัญ: แผนภาพยูเอ็มแอล เรสต์ฟูล แผนภาพคลาส โครงสร้างและพฤติกรรมของระบบ เครื่องมืออัตโนมัติ

Abstract

In the software industry, the software has often been made reusable because it can reduce the cost of the new project. So, documentation in developing software is important for later references. This paper focuses on the UML diagram that is a part of API documentation. This paper proposes to design and develop an automatic tool for creating a diagram

representing RESTful API using PlantUML, an open-source tool for creating a diagram. When we apply this developed tool, we can update the API documentation more easily. Therefore, the API documentation will align with the API's source code. After we validate the diagram from our tool, the correctness of the tool is 100%.

Keywords: UML Diagram, RESTful, Class Diagram, Structural and Behavioral, Automatic Tools.

1. บทนำ

ในปัจจุบันการนำซอฟต์แวร์กลับมาใช้ใหม่เป็นที่นิยมนอย่างมาก เนื่องจากการนำซอฟต์แวร์ที่มีคุณภาพมาพัฒนาให้มีคุณภาพมากขึ้นกว่าเดิม และการลดต้นทุนในการพัฒนาซอฟต์แวร์ ดังนั้น เอกสารของระบบเดิมจึงมีความสำคัญเพื่อใช้ในการอ้างอิงซอฟต์แวร์ใหม่ในอนาคต เช่น เอกสารความต้องการของระบบ เอกสารแสดงส่วนต่อประสานโปรแกรมประยุกต์ (API: Application Program Interface) เอกสารรายละเอียดของฐานข้อมูลของระบบ (Data Dictionary) เป็นต้น

งานวิจัยนี้ให้ความสำคัญกับเอกสารแสดงส่วนต่อประสานโปรแกรมประยุกต์ เป็นเอกสารอธิบายรายละเอียดของส่วนต่อประสานโปรแกรมประยุกต์ที่อยู่ในระบบ ซึ่งอธิบายถึงการรับส่งข้อมูลในรูปแบบ GET (Get), POST (Post), PUT (Put), DELETE (Delete) และระบุถึงพารามิเตอร์ที่เกี่ยวข้อง [8], [10], [11] โดยส่วนใหญ่เอกสารนี้ถูกทำขึ้นจากการออกแบบโดยผู้พัฒนา

ตั้งแต่ขั้นตอนเริ่มต้น และระหว่างการพัฒนาอาจนำเครื่องมือเสริมโอเพ่นเอพีไอ (OpenAPI) เข้ามาช่วยในการสื่อสารกันระหว่างทีมผู้พัฒนา เมื่อระบบอยู่ในขั้นตอนการพัฒนาหรือพัฒนาเสร็จสิ้นแล้ว หากส่วนต่อประสานโปรแกรมประยุกต์มีการแก้ไข ผู้พัฒนาระบบส่วนใหญ่จะละเลยเอกสารนี้ และนอกจากนี้ผู้ที่จำเป็นต้องใช้เอกสารอาจไม่เพียงแต่เป็นผู้พัฒนาระบบเท่านั้น แต่อาจเป็นผู้วิเคราะห์ระบบ หรือผู้ที่ต้องการทำความเข้าใจในการรับส่งข้อมูลของระบบ หากเอกสารนี้ถูกจัดทำขึ้นในรูปแบบของแผนภาพ จะช่วยให้ผู้อ่านเอกสารเข้าใจได้ง่ายขึ้น อีกทั้งยังช่วยในการสื่อสารกันของนักพัฒนาระบบ

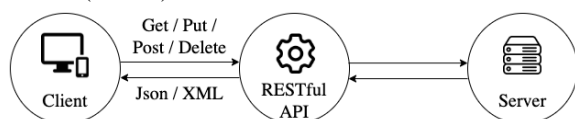
จากปัญหาดังกล่าวข้างต้น ทางผู้วิจัยมีแนวคิดในการออกแบบ และพัฒนาเครื่องมืออัตโนมัติสำหรับการสร้าง และปรับปรุงแผนภาพที่ใช้ในการรับส่งข้อมูล ซึ่งใช้เป็นแผนภาพยูเอ็มแอล (UML: Unified Modeling Language) ที่เรียกว่า ยูเอ็มแอลโพรไฟล์ (UML Profile) [12] โดยประกอบด้วย เส้นทางการรับส่งข้อมูล พารามิเตอร์สำหรับเส้นทางนั้น ๆ เมื่อมีการนำเครื่องมือที่พัฒนาขึ้นอ่านโค้ดของระบบ แผนภาพยูเอ็มแอลจะถูกสร้างขึ้นโดยอัตโนมัติ ดังนั้น เมื่อมีการปรับปรุงรหัสต้นฉบับที่มีผลกระทบต่อการรับส่งข้อมูล แผนภาพนี้จะถูกปรับปรุงโดยอัตโนมัติ และสามารถนำไปใช้ในเอกสารได้ทันที เพื่อป้องกันการละเลยการปรับปรุงเอกสาร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้มีทฤษฎีที่เกี่ยวข้อง ดังนี้

2.1 ส่วนต่อประสานโปรแกรมประยุกต์เว็บเซอร์วิซแบบเรสท์ฟูล (RESTful API Web Services)

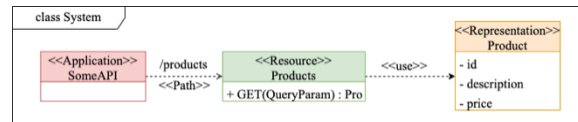
เป็นเว็บเซอร์วิซที่ใช้สถาปัตยกรรมแบบเรสท์ โดยอนุญาตให้มีผู้ร้องขอการเข้าถึงทรัพยากร เรียกว่า เครื่องลูกข่าย (Client) และมีผู้ให้บริการทรัพยากร เรียกว่า เครื่องแม่ข่าย (Server)



ภาพที่ 1: กระบวนการทำงานของเรสท์ฟูลเว็บเซอร์วิซ

จากภาพที่ 1 กระบวนการทำงานของเว็บเซอร์วิซแบบเรสท์ฟูล คือ เครื่องลูกข่ายเป็นผู้ร้องขอข้อมูลจากเครื่องแม่ข่ายผ่าน Http หรือ Https ในรูปแบบของเมธอดเก็ท โพสต์ พุด ดีลิต ผ่านเว็บเซอร์วิซแบบเรสท์ฟูล และเครื่องแม่ข่ายจะส่งผลตอบสนองกลับในรูปแบบของ Json หรือ XML ผ่านเว็บเซอร์วิซแบบเรสท์ฟูล

2.2 แผนภาพยูเอ็มแอลโพรไฟล์แสดงส่วนต่อประสานโปรแกรมประยุกต์เว็บเซอร์วิซแบบเรสท์ฟูล



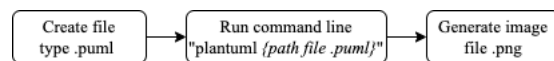
ภาพที่ 2: ตัวอย่างแผนภาพส่วนต่อประสาน โปรแกรมประยุกต์เว็บเซอร์วิซแบบเรสท์ฟูล

จากภาพที่ 2 แผนภาพส่วนต่อประสาน โปรแกรมประยุกต์ [9][13] มีส่วนประกอบ 3 ส่วน ดังนี้

1. <<Application>> เป็นคลาสของตัวกลางในการรับส่งข้อมูล
2. <<Resource>> เป็นคลาสของแอ็กชันต่าง ๆ โดยมีการดำเนินงานอยู่ภายใน และมีเส้นทางการขอข้อมูลจาก <<Application>> ไปยัง <<Resource>> โดยที่เส้นเชื่อมระหว่างคลาสมี่ <<Path>> กำกับ ที่ระบุชื่อของเส้นทางการขอข้อมูล
3. <<Representation>> เป็นคลาสของทรัพยากรของข้อมูลโดยมีคุณลักษณะอยู่ภายใน ที่เส้นเชื่อมระหว่างคลาสมี่ <<use>> กำกับเพื่อบอกถึงการเรียกใช้จาก <<Resource>>

2.3 PlantUML

PlantUML [6] เป็นเครื่องมือเสริม ที่ช่วยในการสร้างแผนภาพต่าง ๆ ด้วยภาษาเฉพาะของ PlantUML โดยงานวิจัยนี้นำมาพัฒนาร่วมกันกับภาษาซีชาร์ปดิ้อทเน็ตเวอร์ชัน 3.1 [7] มีกระบวนการทำงานดังภาพที่ 3



ภาพที่ 3: กระบวนการทำงานของเครื่องมือ PlantUML

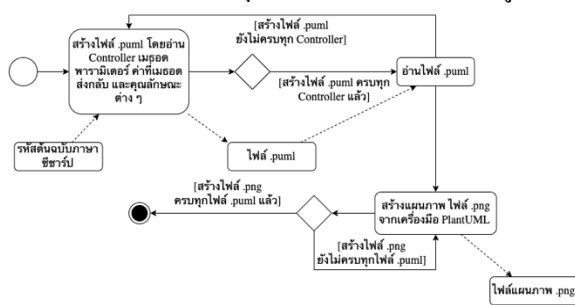
2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ Kristian Nybom และคณะ [1] ได้ทำการวิเคราะห์ถึงความสำคัญของเอกสารแสดงส่วนต่อประสานโปรแกรมประยุกต์ ที่เป็นเอกสารสำหรับการ

นำเสนอภาพรวมของระบบ โดยอธิบายความซับซ้อน และโครงสร้างของระบบในรูปแบบภาษาที่เข้าใจง่าย รวมถึงการวิเคราะห์เกี่ยวกับการบำรุงรักษาเอกสารแสดงส่วนต่อประสาน โปรแกรมประยุกต์ด้วย งานวิจัยของ Digvijaysinh M.Rathod และคณะ [2] และงานวิจัยของ Ivan Porres และ Irum Rauf [4] ได้ทำการวิเคราะห์โครงสร้าง และความสัมพันธ์ของทรัพยากร โดยใช้แบบจำลองแผนภาพยูเอ็มแอล ของแอปพลิเคชันการจองห้องของโรงแรม งานวิจัยของ Hamza Ed-douibi และคณะ [3] ได้นำเสนอเครื่องมือสำหรับการทำ OpenAPI โดยใช้เครื่องมือชื่อ WAPIml เป็นเครื่องมือที่แก้ไขไฟล์ JSON เมื่อมีการแก้ไขสคิมา

3. วิธีการดำเนินงาน

งานวิจัยนี้ได้ออกแบบ และพัฒนาเครื่องมืออัตโนมัติ สำหรับสร้างแบบจำลองความสัมพันธ์ของส่วนต่อประสาน โปรแกรมประยุกต์เว็บเซอร์วิซแบบเรสท์ฟูล



ภาพที่ 4: กระบวนการทำงานของเครื่องมือที่พัฒนาขึ้น

จากภาพที่ 4 การออกแบบกระบวนการทำงานของเครื่องมือ มีกระบวนการดังนี้

1. นำเครื่องมือที่ได้พัฒนาขึ้นไปประยุกต์ใช้กับโครงการที่ต้องการสร้างแผนภาพ และทำการรันโครงการ
2. เครื่องมือที่ได้พัฒนาขึ้น จะทำการสร้างไฟล์ .puml โดยการอ่านคอนโทรลเลอร์ ที่ภายในประกอบด้วยเมธอดต่าง ๆ เช่น พารามิเตอร์ ค่าที่ส่งกลับ และอ่านคุณลักษณะภายใน ซึ่งเครื่องมือจะทำการสร้างไฟล์ .puml จนครบทุกคอนโทรลเลอร์ในโครงการ
3. เครื่องมือที่ได้พัฒนาขึ้น จะทำการสร้างไฟล์ .png โดยรันคำสั่ง “plantuml {path file}” เพื่อสร้างไฟล์แผนภาพตามไฟล์ .puml จากข้อ 2 ให้อัตโนมัติ โดยใช้เครื่องมือเสริม PlantUML

```

9 [Route("[controller]")]
10 public class AccountController : Controller
11 {
12     [HttpPost("register")]
13     public async Task<ActionResult<UserDto>> Register(RegisterDto registerDto)
14     {
15         return new UserDto
16         {
17             DisplayName = "Display Name",
18             Email = "email@gmail.com",
19             Token = "1234"
20         };
21     }
    
```

ภาพที่ 5: ตัวอย่างรหัสต้นฉบับ

เครื่องมือจะเริ่มดำเนินการทำงานจากการอ่านรหัสต้นฉบับ และทำการค้นหาคอนโทรลเลอร์ ดังภาพที่ 5 เป็นตัวอย่างรหัสต้นฉบับใน 1 คอนโทรลเลอร์ โดยในที่นี้จะสนใจที่ AccountController

จากภาพที่ 5 เครื่องมือจะค้นหาองค์ประกอบต่าง ๆ จากรหัสต้นฉบับ และนำมาเขียนให้อยู่ในรูปแบบของเครื่องมือ PlantUML โดยให้อยู่ในไฟล์ .puml ได้ดังตัวอย่างต่อไปนี้

1. เครื่องมืออ่านเจอ AccountController จากการประกาศในบรรทัดที่ 10 จะถูกนำไปแปลงเป็น <<Application>> AccountController
2. เครื่องมืออ่านเมธอดภายใน AccountController ในบรรทัดที่ 13 ซึ่งเจอเมธอดชื่อ Register จะถูกนำไปแปลงเป็น <<Resource>> Register ประเภทโพสต์จากการประกาศในบรรทัดที่ 12 และทำให้เกิดเส้นความสัมพันธ์ โดยกำหนดให้เป็นลูกศรเส้นประ ซึ่งระบุเส้นทางการสื่อสารของข้อมูล และกำกับว่า <<PATH>>
3. ดั่งบรรทัดที่ 13 เครื่องมืออ่านเจอพารามิเตอร์ที่รับเข้าเป็น RegisterDto จะถูกนำไปแปลงเป็น <<Representation>> RegisterDto และเจอค่าพารามิเตอร์ที่เมธอดส่งออกเป็น UserDto จะถูกนำไปแปลงเป็น <<Representation>> UserDto และทำให้เกิดเส้นความสัมพันธ์โดยกำหนดให้เป็นลูกศรเส้นประ โดยกำกับ <<use>> เพื่อบ่งบอกถึงการกล่าวถึงจาก <<Resource>>

```

public class RegisterDto
{
    [Required]
    0 references
    public string DisplayName { get; set; }

    [Required]
    [EmailAddress]
    0 references
    public string Email { get; set; }

    [Required]
    0 references
    public string Password { get; set; }
}
    
```

ภาพที่ 6: คุณลักษณะต่างๆ ของ RegisterDto

4. จากข้อ 3 เมื่อพารามิเตอร์ที่รับเข้า หรือส่งออก เป็นพารามิเตอร์แบบมีคุณลักษณะภายใน เครื่องมือจะทำการอ่านคุณลักษณะภายใน เช่น ค่าพารามิเตอร์ที่รับเข้า เป็น RegisterDto เครื่องมือจะทำการอ่านคุณลักษณะภายใน จากระหัสต้นฉบับดังกล่าวที่ 6 และนำไปเขียนภายใต้ คลาส RegisterDto เป็นต้น

```

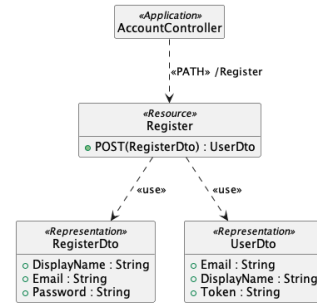
@startuml
class AccountController <<Application>>
{
}
class Register <<Resource>>
{
+ POST(RegisterDto) : UserDto
}
class RegisterDto <<Representation>>
{
+ DisplayName : String
+ Email : String
+ Password : String
}
class UserDto <<Representation>>
{
+ Email : String
+ DisplayName : String
+ Token : String
}
AccountController ..> Register : <<PATH>> /Register
Register ..> UserDto : <<use>>
Register ..> RegisterDto : <<use>>
@enduml
    
```

ภาพที่ 7: ตัวอย่างโค้ดในไฟล์ .puml

ตารางที่ 1: ตารางเปรียบเทียบองค์ประกอบต่างๆ

องค์ประกอบ	สัญลักษณ์ในแผนภาพ	ตัวอย่างโค้ดที่อยู่ในไฟล์ .puml
ชื่อของคอนโทรลเลอร์: Account Controller	«Application» AccountController	หมายเลข 1 จากภาพที่ 7
ชื่อของ Path: /register	«PATH» /register	หมายเลข 5 จากภาพที่ 7
ชื่อ และชนิดของเมธอด: Register ชนิด POST	«Resource» Register + POST(RegisterDto) : UserDto	หมายเลข 2 จากภาพที่ 7
พารามิเตอร์ที่เมธอดรับ: RegisterDto	«Representation» RegisterDto + DisplayName : String + Email : String + Password : String	หมายเลข 3 จากภาพที่ 7
พารามิเตอร์ที่เมธอดส่งออก: UserDto	«Representation» UserDto + DisplayName : String + Email : String + Token : String	หมายเลข 4 จากภาพที่ 7

จากตารางที่ 1 ได้เปรียบเทียบองค์ประกอบที่สำคัญกับรหัสต้นฉบับในภาพที่ 5 และทำการแปลงรหัสต้นฉบับให้อยู่ในรูปแบบของสัญลักษณ์บนแผนภาพ เมื่อเครื่องมืออ่านโค้ดตามลำดับที่กำหนด จึงเกิดเป็นโค้ดดังภาพที่ 7 จากนั้นจึงใช้คำสั่ง “plantuml {path .puml file}” [14] เพื่อสร้างแผนภาพแบบจำลองความสัมพันธ์ของส่วนต่อประสานโปรแกรมประยุกต์แบบเรสต์ฟูลเป็นไฟล์ภาพนามสกุล .png ดังภาพที่ 8

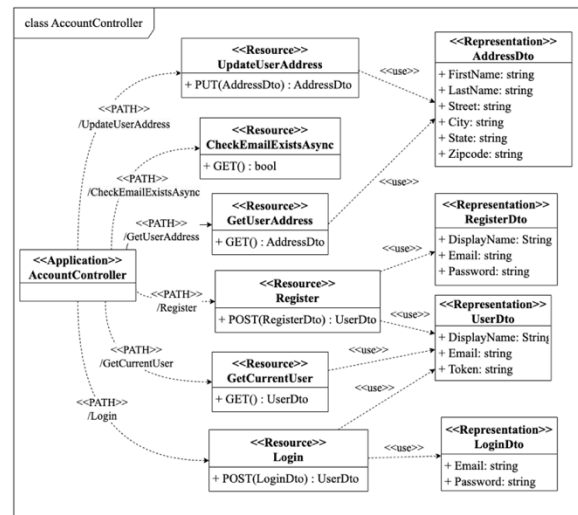


ภาพที่ 8: ตัวอย่างแผนภาพที่ทดลองร่างได้

การทดสอบประสิทธิภาพของเครื่องมือที่ได้พัฒนาขึ้น มีกระบวนการทดสอบ ดังนี้

1. นำรหัสต้นฉบับมาสร้างแผนภาพโดยผู้ทดสอบ

เป็นผู้วาด ดังภาพที่ 9



ภาพที่ 9: ตัวอย่างแผนภาพที่วาดโดยผู้ทดสอบ

2. นำแผนภาพที่ได้จากการสร้างโดยเครื่องมือ ดังภาพที่ 10 มาเปรียบเทียบกับแผนภาพที่ผู้ทดสอบเป็นผู้วาด ดังภาพที่ 9

ผู้ทดสอบจะนำรหัสต้นฉบับจาก 3 โครงการ มาสร้างแผนภาพโครงการละ 3 คอนโทรลเลอร์ ผลการทดสอบที่

ได้รับ คือ แผนภาพที่เกิดจากการวาดโดยผู้ทดสอบ และแผนภาพที่เกิดจากเครื่องมือจะต้องได้แผนภาพที่ตรงกันทั้งหมด โดยมีการกำหนดเงื่อนไขการตรวจสอบ ดังนี้

- จำนวนคลาสของ <<Application>>, <<Resource>>, <<Representation>> ของแผนภาพ จะต้องมีความเท่ากัน และข้อมูลตรงกัน

- ภายในคลาสของ <<Resource>> ต้องประกอบด้วยวิธีการดำเนินงานเท่านั้น และภายในคลาสของ <<Representation>> ต้องประกอบด้วยคุณลักษณะเท่านั้น

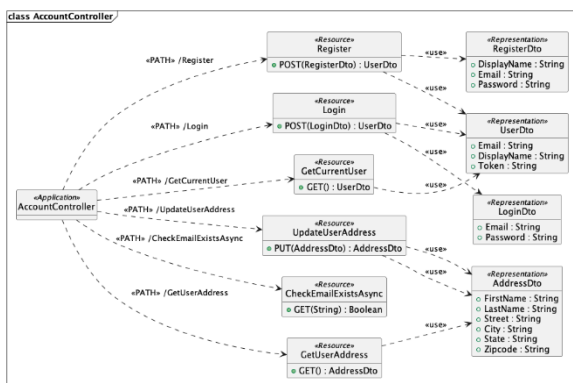
- เมื่อเกิดความสัมพันธ์ระหว่าง <<Application>> กับ <<Resource>> ให้มีเส้นกำกับเป็น <<Path>> ที่ระบุเส้นทางการสื่อสารข้อมูล

- เมื่อเกิดความสัมพันธ์ระหว่าง <<Resource>> กับ <<Representation>> ให้มีเส้นกำกับเป็น <<use>>

ซึ่งการเปรียบเทียบแผนภาพเพื่อตรวจสอบนี้ จะให้ความสำคัญกับความถูกต้อง และความครบถ้วนของข้อมูลเท่านั้น และจะมองข้ามรูปแบบการจัดวางของแผนภาพ เช่น ตำแหน่ง ช่องว่าง สี และลักษณะอักษร เป็นต้น

4. ผลการดำเนินงาน

เมื่อนำเครื่องมือที่พัฒนาขึ้นไปประยุกต์ใช้ เครื่องมือที่พัฒนาขึ้นจะทำการอ่านคอนโทรลเลอร์ทั้งหมดของโครงการนั้น ๆ และทำการสร้างแผนภาพโดยอัตโนมัติ ดังภาพที่ 10 เป็นตัวอย่างของแผนภาพแบบจำลองความสัมพันธ์ของ AccountController



ภาพที่ 10: แผนภาพที่ได้จากการพัฒนาเครื่องมือ

ตารางที่ 2: ผลลัพธ์ของ <<Application>>

<<Application>>			
ชื่อโครงการ	ผู้ทดสอบวาด	เครื่องมือสร้าง	ผลลัพธ์
Dating [15]	7	7	100%
eCommerce [5]	7	7	100%
Routine [16]	4	4	100%
ผลรวม:			100%

ตารางที่ 3: ผลลัพธ์ของ <<Resource>>

<<Resource>>			
ชื่อของ Controller	ผู้ทดสอบวาด	เครื่องมือสร้าง	ผลลัพธ์
โครงการ: Dating			
LikesController	2	2	100%
AccountController	5	5	100%
UsersController	6	6	100%
โครงการ: eCommerce			
OrdersController	1	1	100%
AccountController	6	6	100%
ProductsController	4	4	100%
โครงการ: Routine			
RootController	1	1	100%
EmployeeController	5	5	100%
CompaniesController	5	5	100%
ผลรวม:			100%

ตารางที่ 4: ผลลัพธ์ของ <<Representation>>

<<Representation>>			
ชื่อของ Controller	ผู้ทดสอบวาด	เครื่องมือสร้าง	ผลลัพธ์
โครงการ: Dating			
LikesController	2	2	100%
AccountController	6	6	100%
UsersController	5	5	100%
โครงการ: eCommerce			
OrdersController	2	2	100%
AccountController	4	4	100%
ProductsController	4	4	100%

โครงการ: Routine			
RootController	1	1	100%
EmployeeController	3	3	100%
CompaniesController	3	3	100%
ผลรวม:			100%

จากการตรวจสอบความถูกต้องของการสร้างแผนภาพ ได้ผลการทดสอบ ดังตารางที่ 2-4 จำนวนของ <<Application>>, <<Resource>>, <<Representation>> ของแผนภาพที่เกิดจากการวาดโดยผู้ทดสอบ และแผนภาพที่เกิดจากเครื่องมือ ตรงกัน 100% โดยไม่นับจำนวนของ <<Resource>> ที่เป็นฟังก์ชันภายในหรือที่ไม่ใช่เมธอด และไม่นับจำนวนของ <<Representation>> ที่ประกาศอยู่ภายนอกโครงการ และเมื่อเกิดความสัมพันธ์ระหว่าง <<Application>> กับ <<Resource>> สามารถแสดงเส้นกำกับที่มีการระบุ <<Path>> /{ชื่อของ path} ได้และเมื่อเกิดความสัมพันธ์ระหว่าง <<Resource>> กับ <<Representation>> สามารถแสดงเส้นกำกับที่มีการระบุ <<use>> ได้ถูกต้อง และครบถ้วนทั้งหมด โดยคิดเป็น $(100\% + 100\% + 100\%) / 3 = 100\%$

5. ผลการดำเนินงาน

ในงานวิจัยนี้ได้ทำการออกแบบและพัฒนาเครื่องมือสำหรับการสร้างแผนภาพจำลองความสัมพันธ์ของส่วนต่อประสานโปรแกรมประยุกต์เว็บเซอร์วิซแบบเรสต์ฟูล เพื่อช่วยในการปรับปรุงเอกสารแสดงส่วนต่อประสานโปรแกรมประยุกต์ ที่สามารถนำแผนภาพนี้ ไปเป็นส่วนหนึ่งของเอกสารแสดงส่วนต่อประสานโปรแกรมประยุกต์ เพื่อช่วยให้ผู้อ่านสามารถเข้าใจได้ง่ายยิ่งขึ้น เมื่อนำเครื่องมือที่พัฒนาขึ้น ไปประยุกต์ใช้กับโครงการต่าง ๆ และทำการทดสอบโดยการสุ่มจากโครงการต่าง ๆ จำนวน 3 โครงการ โครงการละ 3 คอนโทรลเลอร์ ทำให้ได้ผลลัพธ์ของความถูกต้อง และครบถ้วน เป็น 100% แต่มีข้อกำหนดของเครื่องมือ ในส่วนของเมธอด หากเป็นฟังก์ชันภายใน หรือไม่ใช่เมธอดแต่ถูกประกาศไว้ในรหัสต้นฉบับ หรือเป็นส่วนของคุณลักษณะที่มีการประกาศอยู่ภายนอกโครงการ จะไม่ถูกนำไปสร้างแผนภาพ

เอกสารอ้างอิง

- [1] Kristian Nybom, Adnan Ashraf, Ivan Porres. "A Systematic Mapping Study on API Documentation Generation Approaches", 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). 2018
- [2] Digvijaysinh M.Rathod, Satyen M.Parikh, B. V. Buddhadev. "Structural and Behavioral modeling of RESTful web service interface using UML", International Conference on Intelligent Systems and Signal Processing (ISSP). 2013
- [3] Hamza Ed-douibi, Javier Luis Canovas Izquierdo, Francis Bordeleau, Jordi Cabot. "WAPIML: Towards a Modeling Infrastructure for Web APIs", ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C). 2019
- [4] Ivan Porres, IrumRauf. "Modeling Behavioral RESTful Web Service Interfaces in UML", ACM Symposium on Applied Computing. 2011.
- [5] Chiantine P. Manigos. "E-CommerceApp" [Online]. 2021. Available: <https://github.com/xchanmolx/E-CommerceApp>
- [6] PlantUML Version 1.2021.2. [Online]. Available: <https://plantuml.com/guide>
- [7] Hamza Ed-douibi. "OpenAPItoUML: a Tool to Generate UML Models from OpenAPI Definitions", The International Conference on Web Engineering (ICWE). 2018
- [8] Mahesh Haldar. "RESTful API Designing guidelines – The best practices" [Online]. 2017. Available: <https://medium.com/hackernoon/restful-api-designing-guidelines-thebest-practices-60e1d954e7c9>
- [9] Create Spec for REST API in Enterprise Architect [Online]. 2019. Available: <https://newbedev.com/create-spec-for-rest-api-in-enterprise-architect>
- [10] The Hungry Brain, "REST API Architectural Constraints" [Online]. 2022. Available: <https://www.geeksforgeeks.org/rest-api-architectural-constraints/>
- [11] Archit Patel, "Best Practices to Follow for REST API Development" [Online]. 2020. Available: <https://www.mindinventory.com/blog/best-practices-rest-api-development/>

- [12] Wathichai and Twittie Senivongse, “A Model-Driven Development of Web-Based Applications Using AngularJS Framework”, IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS). 2016
- [13] Microsoft, “Guide to UML Diagramming and Database Modeling” [Online]. 2022. Available: <https://www.microsoft.com/th-th/microsoft-365/business-insightsideas/resources/guide-to-uml-diagramming-and-database-modeling>
- [14] Ubuntu Manpage, “Plantuml - Generate UML Diagrams from a Text Description” [Online]. 2019. Available: <https://manpages.ubuntu.com/manpages/impish/man1/plantuml.1.html#name>
- [15] Anthony Matute. “DatingApp” [Online]. 2020. Available: <https://github.com/tonymatute/DatingApp>
- [16] Surbowl. “Routine” [Online]. 2020. Available: <https://github.com/Surbowl/ASP.NET-Core-RESTful-Note>

การรายงานผลการทดสอบอัตโนมัติโดยใช้บีดีดีด้วยภาษาเชอร์กิน

Test Automation Report using BDD with Gherkin

จารุ โสพรรณพนิชกุล (Jare Sophanpanitchakul)¹ และธาราทิพย์ สุวรรณศาสตร์ (Taratip Suwannasart)²

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹6272014021@student.chula.ac.th, ²taratip.s@chula.ac.th

บทคัดย่อ

การทดสอบซอฟต์แวร์เป็นกระบวนการที่สำคัญในการพัฒนาซอฟต์แวร์ เพื่อให้ได้ซอฟต์แวร์ที่มีคุณภาพปราศจากข้อผิดพลาด ตรงตามความต้องการของลูกค้า โดยปัจจุบันการทดสอบซอฟต์แวร์แบบอัตโนมัติได้เข้ามามีบทบาทอย่างมีนัยสำคัญในกระบวนการทดสอบซอฟต์แวร์ เนื่องจากช่วยให้นักทดสอบซอฟต์แวร์สามารถทำงานเสร็จได้เร็วขึ้น โดยที่ยังคงคุณภาพของงานที่เท่าเทียมหรือดีขึ้นกว่าการทดสอบด้วยมือ ซึ่งการเขียนสคริปต์เพื่อทำการทดสอบแบบอัตโนมัติต้องใช้บุคลากรที่มีความรู้ความเข้าใจในการเขียนโค้ด อีกทั้งผลลัพธ์จากการทดสอบจะอยู่ในรูปแบบของภาษาโปรแกรม ซึ่งทำให้ยากต่อการทำความเข้าใจจากผู้มีส่วนได้ส่วนเสียจากหน่วยธุรกิจ ที่ไม่ได้มีความรู้ความเข้าใจเกี่ยวกับภาษาของโปรแกรมมาก่อนในการตรวจสอบและทวนสอบความถูกต้องของงานก่อนรับมอบ งานวิจัยนี้จึงเป็นการช่วยนักทดสอบซอฟต์แวร์ในการแปลงเอกสารการทดสอบที่อยู่ในรูปแบบ บีดีดี ด้วยภาษาเชอร์กินให้เป็นไฟล์สคริปต์เพื่อเป็นข้อมูลนำเข้าสำหรับการทดสอบแบบอัตโนมัติบนโปรแกรมวิซวลสตูดิโอ และใช้ไลบรารีชื่อเอ็กซ์เท็นด์รีพอร์ตในการสร้างผลลัพธ์จากการทดสอบแบบอัตโนมัติที่ระบุขั้นตอนการทดสอบในรูปแบบของภาษาเชอร์กินที่สามารถเข้าใจได้ และถูกบันทึกอยู่ในรูปแบบของไฟล์นามสกุลเอชทีเอ็มแอล

คำสำคัญ: การทดสอบแบบอัตโนมัติ บีดีดี ภาษาเชอร์กิน

Abstract

Software testing is an important process in software development to produce good software quality without bugs and meet customer needs.

At present, automated software testing is an essential part of the software testing process because it can help software testers speed up their work while maintaining the same or better quality than manual testing. However, scripting has required a person who has coding skills. Also, the outcome of testing is generally in a coding format which is complicated for stakeholders from the business units who have no experience in programming to verify and validate the testing result. This research aims to help a tester convert a test case document which is in Gherkin language format to a script file, then import it into Visual studio to perform automated testing, using Extent report library to generate a testing report in Gherkin language which is understandable and recorded in HTML file format.

Keywords: Automation Testing, BDD, Gherkin Language.

1. บทนำ

การทดสอบซอฟต์แวร์เป็นกระบวนการที่สำคัญในการพัฒนาซอฟต์แวร์ เพื่อให้ได้ซอฟต์แวร์ที่มีคุณภาพปราศจากข้อผิดพลาด ตรงตามความต้องการของลูกค้า อีกทั้งยังสร้างความน่าเชื่อถือให้กับลูกค้าและองค์กรอีกด้วย

กระบวนการทดสอบซอฟต์แวร์มีทั้งแบบการทดสอบด้วยมือ (Manual Testing) ซึ่งเป็นการทำการทดสอบโดยนักทดสอบซอฟต์แวร์ และการทดสอบแบบอัตโนมัติ (Automation Testing) ซึ่งเป็นการทำการทดสอบโดยใช้โปรแกรมหรือสคริปต์ที่ถูกเขียนขึ้นมาโดยเฉพาะสำหรับงานนั้น ๆ ในการทดสอบซอฟต์แวร์เพื่อให้ได้ผลลัพธ์ตามที่ต้องการ

ในปัจจุบันการทดสอบซอฟต์แวร์แบบอัตโนมัติได้เข้ามามีบทบาทอย่างมีนัยสำคัญในกระบวนการการทดสอบซอฟต์แวร์ เนื่องจากช่วยให้นักทดสอบซอฟต์แวร์สามารถทำงานเสร็จได้เร็วขึ้น โดยที่ยังคงคุณภาพของงานที่เท่าเทียมหรือดีขึ้นกว่าการทดสอบด้วยมือ อีกทั้งยังช่วยเพิ่มความความสามารถในการทำการทดสอบซอฟต์แวร์ที่มีการเปลี่ยนแปลง (Regression test) ได้มีประสิทธิภาพมากยิ่งขึ้น ทั้งด้านบุคลากร เวลา ค่าใช้จ่าย และความน่าเชื่อถือ ซึ่งช่วยให้องค์กรมีความคล่องตัวต่อความเปลี่ยนแปลง รวมถึงช่วยให้องค์กรสามารถบริหารทรัพยากรได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

การทดสอบทดสอบซอฟต์แวร์แบบอัตโนมัติจะต้องใช้บุคลากรที่มีความรู้ความเข้าใจในการเขียนโค้ดเพื่อสังเกตุส่วนประกอบต่าง ๆ บนเว็บไซต์หรือบนแอปพลิเคชัน เช่น ปุ่มกด กล้องข้อความและอื่น ๆ ให้เป็นไปตามขั้นตอนในกรณีทดสอบ ซึ่งผลลัพธ์ที่ได้นั้นจะอยู่ในรูปแบบของภาษาโปรแกรมทำให้ยากต่อการทำความเข้าใจจากผู้มีส่วนได้ส่วนเสียจากหน่วยธุรกิจ (Business Unit) ที่ไม่ได้มีความรู้ความเข้าใจเกี่ยวกับภาษาของโปรแกรมมาก่อนในการตรวจสอบและทวนสอบความถูกต้องของงานก่อนรับมอบ

จากปัญหาดังกล่าวจึงมีความคิดที่จะนำเสนอการแสดงผลการทดสอบซอฟต์แวร์แบบอัตโนมัติที่อยู่ในรูปแบบ บีดีดี (BDD: Behavioral Driven Development) ด้วยภาษา เมอร์กิน (Gherkin Language) ที่มีความคล้ายคลึงกับภาษาที่มนุษย์ใช้ในการสื่อสารกันมากที่สุด โดยเริ่มจากการแปลงไฟล์กรณีทดสอบ (Test Case) ที่ถูกเขียนจากนักทดสอบซอฟต์แวร์ให้กลายเป็นสคริปต์ที่อยู่ในรูปแบบของภาษาเมอร์กิน เพื่อเป็นข้อมูลนำเข้าให้กับซอร์สโค้ดที่ใช้ในการควบคุมส่วนประกอบของเว็บไซต์จนกระทั่งได้รายงานผลลัพธ์ของการทดสอบแบบอัตโนมัติที่อยู่ในรูปแบบไฟล์เอชทีเอ็มแอล (HTML)

ผลลัพธ์ที่ได้จากงานวิจัยนี้ จะช่วยให้ผู้มีส่วนได้ส่วนเสียที่ทำหน้าที่ในการตรวจสอบผลลัพธ์ของการทดสอบซอฟต์แวร์แบบอัตโนมัติสามารถทำความเข้าใจในสิ่งที่ถูกทดสอบได้ง่าย และรวดเร็วยิ่งขึ้น อีกทั้งยังสามารถนำไปตรวจสอบความครบถ้วน ถูกต้อง สมบูรณ์ โดยอ้างอิงจาก

เอกสารความต้องการก่อนที่จะดำเนินการอนุมัติ และรับมอบผลิตภัณฑ์ในขั้นตอนถัดไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทดสอบแบบอัตโนมัติ [1]

การทดสอบแบบอัตโนมัติเป็นวิธีหนึ่งในการทดสอบซอฟต์แวร์จากเครื่องมืออัตโนมัติ เพื่อควบคุมการดำเนินการทดสอบแทนผู้ทดสอบที่เป็นมนุษย์ จากนั้นจึงทำการเปรียบเทียบผลลัพธ์ที่ได้จากทดสอบจริง เทียบกับผลลัพธ์ที่คาดหวัง อีกทั้งการทดสอบแบบอัตโนมัติช่วยทำให้ประสิทธิภาพในการทดสอบดีขึ้นทั้งในเรื่องของบุคลากร เวลา ค่าใช้จ่าย และความน่าเชื่อถือ

2.2 กรอบการทำงานการทดสอบแบบอัตโนมัติ (Test Automation Framework) [2]

กรอบการทำงานการทดสอบแบบอัตโนมัติคือโครงสร้างที่รวบรวมโปรแกรม คอมไพเลอร์ เครื่องมือ และสิ่งอื่นๆ เข้าด้วยกัน ซึ่งเป็นการเตรียมสภาพแวดล้อมสำหรับการเรียกใช้สคริปต์สำหรับการทดสอบแบบอัตโนมัติ เพื่อให้ง่ายต่อการพัฒนาและการบำรุงรักษา สคริปต์สำหรับทดสอบ รวมถึงส่งเสริมการนำกลับมาใช้ใหม่อีกด้วย

2.3 การพัฒนาที่ขับเคลื่อนด้วยพฤติกรรม – บีดีดี [3]

การพัฒนาที่ขับเคลื่อนด้วยพฤติกรรม - บีดีดี เป็นวิธีการทดสอบซอฟต์แวร์ที่ต่อยอดมาจากการพัฒนาที่ขับเคลื่อนด้วยการทดสอบ – ทีดีดี โดยการทดสอบแบบบีดีดีจะขึ้นอยู่กับพฤติกรรมของระบบเป็นหลัก โดยมีการกำหนดวิธีการในการพัฒนาตามคุณลักษณะของพฤติกรรม เช่นการเขียนบีดีดีด้วยภาษาเมอร์กิน ซึ่งอยู่ในรูปแบบ Given-When-Then ดังตัวอย่างดังต่อไปนี้

Given the user has entered valid login credentials

When a user clicks on the login button

Then display the successful validation message

จากตัวอย่างดังกล่าวจะเห็นได้ว่าการพัฒนาที่ขับเคลื่อนด้วยพฤติกรรม – บีดีดี จะใช้ภาษาอังกฤษที่ใช้ในการสื่อสารที่เข้าใจได้ง่าย ซึ่งช่วยให้ทุกคนในทีมเข้าใจ

พฤติกรรมการทำงานของระบบได้อย่างรวดเร็วผ่านรูปแบบของภาษา

2.5 ภาษาเมอร์กิน [4]

ภาษาเมอร์กินเป็นภาษาที่ใช้ในการกำหนดกรณีทดสอบ โดยถูกออกแบบให้เป็นภาษาที่มนุษย์สามารถทำความเข้าใจได้ และถูกใช้เพื่ออธิบายพฤติกรรมที่เกี่ยวข้องกับการทำงานของซอฟต์แวร์ โดยสามารถใช้ร่วมกับภาษาโปรแกรมได้หลายภาษา เช่น ภาษาซีชาร์ป ภาษาจาวา ภาษาไพทอน ดังแสดงในภาพที่ 1

```
Scenario Outline: A user withdraws money from an ATM
  Given <Name> has a valid Credit or Debit card
  And their account balance is <OriginalBalance>
  When they insert their card
  And withdraw <WithdrawalAmount>
  Then the ATM should return <WithdrawalAmount>
  And their account balance is <NewBalance>
```

Examples:				
Name	OriginalBalance	WithdrawalAmount	NewBalance	
Eric	100	45	55	
Gaurav	100	40	60	
Ed	1000	200	800	

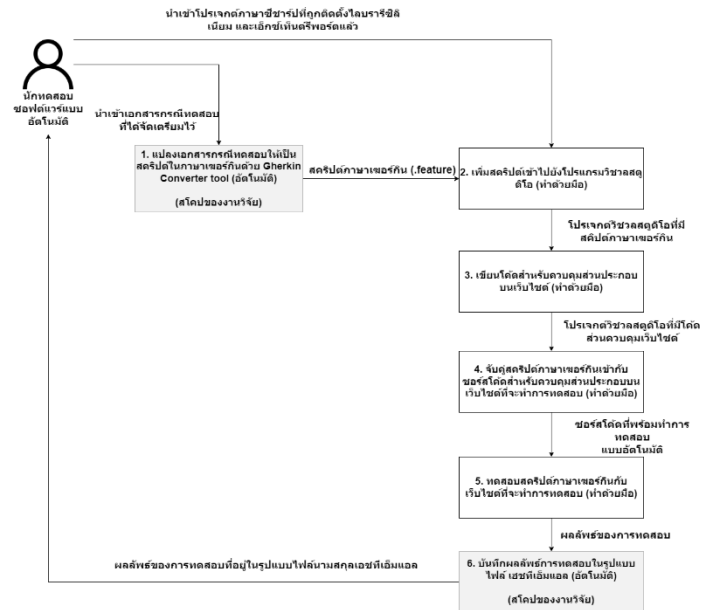
ภาพที่ 1: ตัวอย่างภาษาเมอร์กินที่ถูกนำไปใช้กับโค้ดภาษาซีชาร์ป

2.6 ซิลิเนียม ไลบรารี (Selenium Library) [5]

ซิลิเนียม ไลบรารี เป็นไลบรารีสำหรับการควบคุมส่วนประกอบ (Component) บนเว็บไซต์แบบอัตโนมัติ อีกทั้งยังสามารถติดตั้งและใช้งานร่วมกับภาษาโปรแกรมที่หลากหลา รวมถึงมีการเรียกใช้คำสั่งที่เป็นลึขั้วีร์ดที่เข้าใจได้ง่าย ทำให้สามารถพัฒนาสคริปต์ทดสอบแบบอัตโนมัติได้อย่างรวดเร็วและมีประสิทธิภาพ

3. วิธีการดำเนินงาน

งานวิจัยนี้ได้นำเสนอกระบวนการในการแปลงเอกสารกรณีทดสอบให้อยู่ในรูปแบบของไฟล์สคริปต์ภาษาเมอร์กิน เพื่อใช้เป็นข้อมูลนำเข้าในโปรแกรมวิซวลสตูดิโอสำหรับสร้างรายงานผลลัพธ์ของการทดสอบแบบอัตโนมัติที่อยู่ในรูปแบบภาษาเมอร์กินและถูกบันทึกอยู่ในไฟล์นามสกุลเอชทีเอ็มแอล ดังแสดงในภาพที่ 2 และมีขั้นตอนดังนี้



ภาพที่ 2: ภาพรวมของงานวิจัย

1) แปลงเอกสารกรณีทดสอบให้เป็นสคริปต์ในภาษาเมอร์กินด้วย Gherkin Converter tool

ในขั้นตอนนี้ นักทดสอบซอฟต์แวร์แบบอัตโนมัติจะทำการนำเข้าเอกสารกรณีทดสอบที่อยู่ในรูปแบบของไฟล์เอ็กซ์เซลซึ่งมีรูปแบบตามที่กำหนด เพื่อแปลงเนื้อหาในเอกสารกรณีทดสอบให้อยู่ในรูปแบบของไฟล์สคริปต์ภาษาเมอร์กิน และเพื่อใช้เป็นข้อมูลนำเข้าให้กับโปรแกรมวิซวลสตูดิโอในขั้นตอนถัดไป ดังแสดงในภาพที่ 3

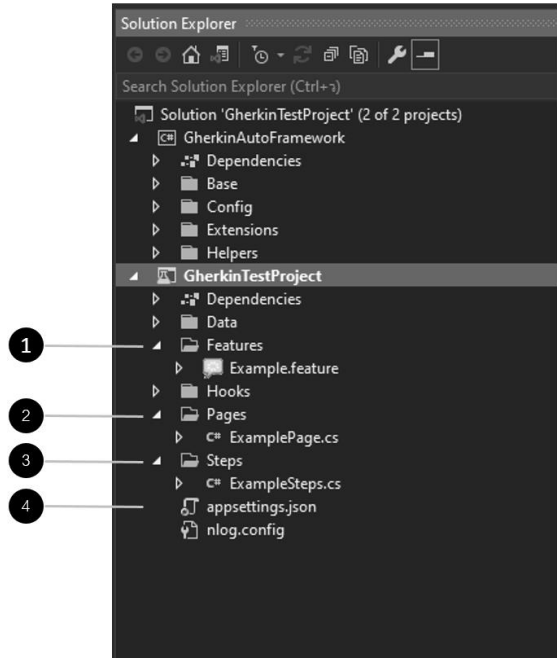


ภาพที่ 3: กระบวนการในการแปลงไฟล์เอกสารกรณีทดสอบให้เป็นสคริปต์ภาษาเมอร์กิน

2) นำเข้าไฟล์สคริปต์ภาษาเมอร์กินมายังโปรแกรมวิซวลสตูดิโอ 2019 [6]

โปรแกรมวิซวลสตูดิโอ เป็นโปรแกรมสำหรับเขียนซอร์สโค้ดบนเฟรมเวิร์คคอตเน็ต (.Net) ซึ่งถูกพัฒนาโดยบริษัทไมโครซอฟต์ โดยผู้วิจัยได้ทำการสร้างโปรเจกต์ภาษาซีชาร์ปด้วยวิซวลสตูดิโอ 2019 สำหรับทำการทดสอบแบบอัตโนมัติ พร้อมทั้งติดตั้งไลบรารีที่จำเป็นเช่นซิลิเนียมไลบรารีสำหรับควบคุมส่วนประกอบบนเว็บไซต์และเอ็กซ์ทีเอ็นดีทีพอร์ตสำหรับสร้างรายงานผลลัพธ์การ

ทดสอบที่อยู่ในรูปแบบไฟล์เอชทีเอ็มแอล โดยไฟล์สคริปต์ภาษาเจอริน (.feature) ที่ได้จากการแปลงจะถูกจัดเก็บอยู่ภายใต้โฟลเดอร์ Feature ดังแสดงในภาพที่ 4 (หมายเลข 1)



ภาพที่ 4: โครงสร้างโปรเจกต์ภาษาซีชาร์ปสำหรับการทดสอบแบบอัตโนมัติ

3) เขียนโค้ดเพื่อควบคุมส่วนประกอบบนเว็บไซต์ที่จะทำการทดสอบ

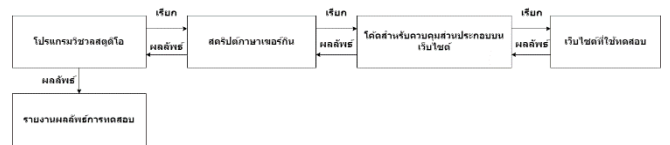
ในส่วนนี้จะเป็นการใช้ความสามารถของซีลีเนียมไลบรารีในการควบคุมส่วนประกอบต่าง ๆ ที่อยู่บนเว็บไซต์ที่จะทำการทดสอบ เช่น สั่งให้กดปุ่ม เติมข้อความ ในกล่องข้อความ ตรวจสอบข้อความที่ปรากฏ เป็นต้น โดยจะต้องอาศัยนักพัฒนาหรือทดสอบซอฟต์แวร์แบบอัตโนมัติมีความรู้ความเข้าใจในการเขียนโค้ด โดยโค้ดที่ถูกเขียนจะอยู่ในรูปแบบคลาสและเมธอด และจะต้องเขียนให้ครอบคลุม ครอบคลุมตามเงื่อนไขที่ต้องการจะทดสอบ โดยโค้ดคำสั่งต่าง ๆ จะถูกบันทึกเป็นไฟล์นามสกุล .cs และถูกเก็บไว้ภายใต้โฟลเดอร์ชื่อ Pages ดังแสดงภาพที่ 5 (หมายเลข 2)

4) จับคู่สคริปต์ภาษาเจอรินเข้ากับซอร์สโค้ดสำหรับควบคุมส่วนประกอบบนเว็บไซต์ที่จะทำการทดสอบ

ในขั้นตอนนี้จะเป็นการจับคู่สคริปต์ภาษาเจอริน (.feature) ที่อยู่ภายใต้โฟลเดอร์ Features เข้ากับซอร์สโค้ดสำหรับควบคุมส่วนประกอบบนเว็บไซต์ (.cs) ที่อยู่ภายใต้โฟลเดอร์ Pages โดยจะต้องทำการสร้างไฟล์ภาษาซีชาร์ปนามสกุล .cs ไว้ภายใต้โฟลเดอร์ Steps ดังแสดงในภาพ 5 (หมายเลข 3)

5) ทดสอบสคริปต์ภาษาเจอรินกับเว็บไซต์ที่จะทำการทดสอบ

ในขั้นตอนนี้จะเป็นการทดสอบสคริปต์ภาษาเจอรินร่วมกับซอร์สโค้ดที่ได้เตรียมไว้ผ่านโปรแกรมวิชวลสตูดิโอ 2019 เพื่อทำการทดสอบเว็บไซต์ที่ได้เตรียมไว้ โดยมีขั้นตอนการทำงาน ดังแสดงในภาพที่ 5



ภาพที่ 5: ขั้นตอนการทำงานของชุดคำสั่งภายในโปรแกรมวิชวลสตูดิโอ

6) บันทึกรายงานผลลัพธ์การทดสอบในรูปแบบไฟล์เอชทีเอ็มแอล

การที่ได้มาซึ่งรายงานที่อยู่ในรูปแบบของเอชทีเอ็มแอลนั้น จำเป็นต้องทำการติดตั้งไลบรารีเอ็กทีเอ็นตริฟอร์ต [7] เข้ากับโปรแกรมวิชวลสตูดิโอก่อน จากนั้นจึงทำการเขียนโค้ดภาษาซีชาร์ปเพื่อเรียกใช้งานไลบรารีดังกล่าว โดยหลังจากทำการรันสคริปต์ภาษาเจอรินด้วยวิชวลสตูดิโอแล้ว วิชวลสตูดิโอจะทำการอ่านสคริปต์ภาษาเจอรินเพื่อไปสั่งการเว็บไซต์ที่จะทำการทดสอบ และนำส่งผลลัพธ์ของการทดสอบไปให้เอ็กทีเอ็นตริฟอร์ตเพื่อทำการสร้างรายงานการทดสอบที่ถูกบันทึกอยู่ในรูปแบบของไฟล์เอชทีเอ็มแอล ตามไคเรกทอรีที่ถูกระบุไว้ในไฟล์ appsettings.json ดังแสดงในภาพที่ 4 (หมายเลข 4)

4. ผลการดำเนินงาน

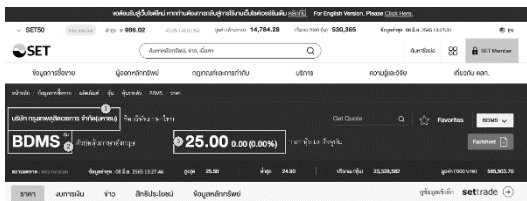
4.1 การทดสอบแบบอัตโนมัติบนเว็บไซต์ ด้วยวิชวลสตูดิโอ ร่วมกับซีลีเนียมไลบรารี และเอ็กทีเอ็นตริฟอร์ต

www.set.or.th [8] เป็นเว็บไซต์ตลาดหลักทรัพย์แห่งประเทศไทยซึ่งได้รวบรวมข้อมูลเกี่ยวกับบริษัทจดทะเบียน

ทั้งหมด เช่น ข้อมูลบริษัท งบการเงิน ข่าวสาร ราคาหุ้น เป็นต้น

ผู้วิจัยต้องการทดสอบฟังก์ชัน Search บนเว็บไซต์ดังกล่าว เพื่อทำการตรวจสอบชื่อบริษัทภาษาไทย ด้วยย่อหุ้นภาษาอังกฤษ และราคาหุ้นปัจจุบัน ว่าถูกต้องหรือไม่ โดยมีขั้นตอนการทดสอบดังนี้

- 1) ทำการเข้าสู่เว็บไซต์ www.set.or.th
- 2) นำเอาชื่อย่อของบริษัทจดทะเบียนไปค้นหาในกล่องค้นหาที่หน้าโฮมเพจ
- 3) ทำการตรวจสอบ ชื่อบริษัทภาษาไทย ด้วยย่อหุ้นภาษาอังกฤษ และราคาหุ้น

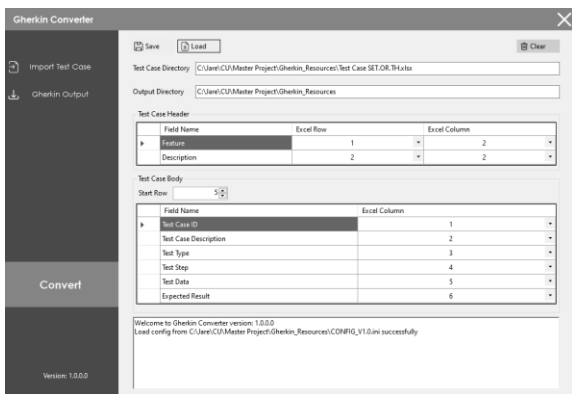


ภาพที่ 6: ข้อมูลบนเว็บไซต์ที่จะทำการตรวจสอบ

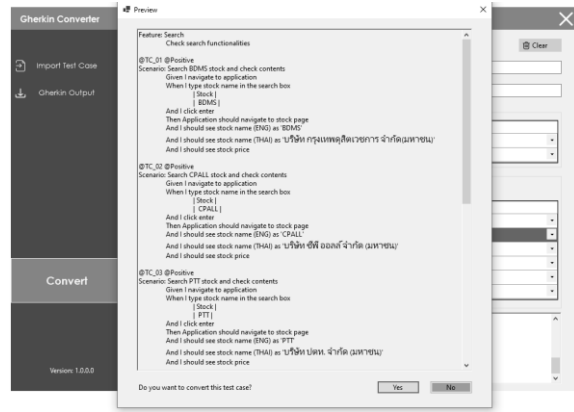
การทดสอบจะเริ่มจากการแปลงเอกสารกรณีสอบที่อยู่ในรูปแบบเอ็กซ์เซลให้เป็นไฟล์ภาษาเมอร์กินด้วย Gherkin Converter ดังแสดงในภาพที่ 7, 8 และ 9

Scenario ID	Scenario Name	Priority	Tags	Created By	Created Date	Updated Date	Version
TC_01	Search BDMC stock and check contents	Pass		SAK,BAK			
TC_02	Search CPALL stock and check contents	Pass		SAK,BAK			
TC_03	Search PTT stock and check contents	Pass		SAK,BAK			

ภาพที่ 7: เอกสารกรณีสอบในรูปแบบไฟล์เอ็กซ์เซล



ภาพที่ 8: การแปลงไฟล์กรณีสอบด้วย Gherkin Converter



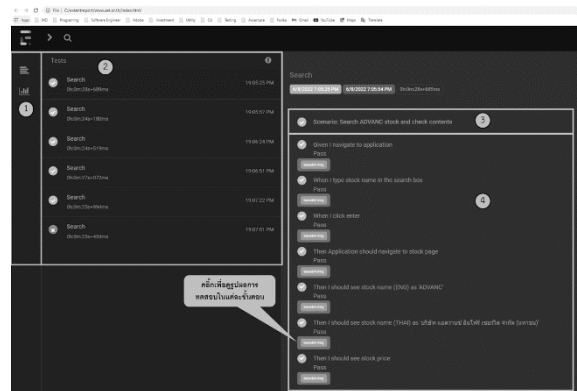
ภาพที่ 9: รูปแบบภาษาเมอร์กินที่ได้จากการแปลงไฟล์กรณีสอบ

จนกระทั่งได้ไฟล์สคริปภาษาเมอร์กินที่พร้อมนำเข้าสู่โปรแกรมวิชวลสตูดิโอ 2019 เพื่อนำไปทดสอบแบบอัตโนมัติ ดังแสดงในภาพที่ 10

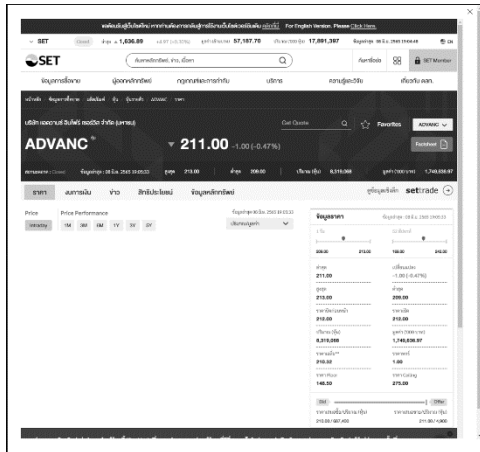
Name	Date modified	Type	Size
Gherkin_20223708043722	6/8/2022 4:37 PM	Gherkin Specificat...	3 KB

ภาพที่ 10: ไฟล์สคริปต์ภาษาเมอร์กินที่อยู่ในรูปแบบ .feature

ผลลัพธ์ที่ได้จากการทดสอบด้วยวิชวลสตูดิโอ 2019 ที่ติดตั้งไลบรารีเอ็กซ์เท็นชันดีพรีพอร์เตอร์จะอยู่ในรูปแบบของไฟล์เอชทีเอ็มแอล ที่สามารถเปิดได้ด้วยเว็บเบราว์เซอร์ และอยู่ในรูปแบบภาษาเมอร์กินที่ผู้มีส่วนได้ส่วนเสียสามารถที่จะอ่านและทำความเข้าใจได้ง่ายเนื่องจากมีความใกล้เคียงกับภาษาที่มนุษย์ใช้ในการสื่อสาร ดังแสดงในภาพที่ 11 และ 12



ภาพที่ 11: ผลลัพธ์การทดสอบในรูปแบบเอชทีเอ็มแอล



ภาพที่ 12: ภาพผลลัพธ์การทดสอบในแต่ละขั้นตอน

โดยไฟล์ผลลัพธ์ที่ได้จากการทดสอบมีส่วนประกอบที่สำคัญ ดังนี้

1. แถบเมนูเพื่อเลือกระหว่างการแสดงผลแบบมีรายละเอียด หรือแบบสรุป
2. ฟิวเจอร์สำหรับสิ่งที่ถูกทดสอบ
3. Scenario สำหรับการทดสอบ
4. ขั้นตอนการทดสอบที่อยู่ในรูปแบบภาษาเมอร์กิน พร้อมทั้งผลลัพธ์ของการทดสอบที่ถูกบันทึกเป็นไฟล์รูปภาพ

5. สรุป

งานวิจัยนี้ได้นำเสนอการออกแบบและแปลงไฟล์เอกสารกรณีทดสอบให้เป็นสคริปต์ภาษาเมอร์กินเพื่อใช้เป็นข้อมูลนำเข้าให้กับโปรแกรมสำหรับทดสอบแบบอัตโนมัติด้วยโปรแกรมวิชวลสตูดิโอ 2019 จากนั้นจึงใช้ความสามารถของไลบรารีเอ็กซ์เท็นดิฟอรัทในการสร้างรายงานผลลัพธ์การทดสอบที่มีรายละเอียดขั้นตอนการทดสอบอยู่ในรูปแบบของภาษาเมอร์กินที่สามารถเข้าใจได้ง่าย และถูกบันทึกอยู่ในไฟล์นามสกุลเอชทีเอ็มแอล

ซึ่งจากการทดสอบเว็บไซต์ทั้ง 3 แบบ พบว่าเครื่องมือสามารถแปลงไฟล์เอกสารกรณีทดสอบได้ถูกต้องครบถ้วน สมบูรณ์ และสามารถนำเข้าไปในโปรแกรมวิชวลสตูดิโอได้ทันทีโดยไม่มีการแสดงข้อผิดพลาด อีกทั้งยังสามารถส่งข้อมูลการทดสอบเข้าไปยังเมธอดภาษาซีชาร์ปเพื่อทำการป้อนให้กับเว็บไซต์ทดสอบได้อย่างถูกต้อง รวมถึง

สามารถสร้างผลลัพธ์การทดสอบที่อยู่ในรูปแบบของภาษาเมอร์กินพร้อมทั้งรูปประกอบในรูปแบบของไฟล์เอชทีเอ็มแอลที่สามารถเปิดบนเว็บเบราว์เซอร์ได้ทันที

6. งานในอนาคต

1. การวิเคราะห์โครงสร้างของเว็บไซต์แบบอัตโนมัติ การวิเคราะห์โครงสร้างของเว็บไซต์เป็นสิ่งที่จำเป็นต่อการทดสอบแบบอัตโนมัติ เพื่อที่จะได้ทราบถึงส่วนประกอบต่างๆ บนหน้าเว็บไซต์ รวมถึงสามารถเขียนโค้ดเพื่อควบคุมและสั่งงานส่วนประกอบเหล่านั้นได้อย่างถูกต้องและแม่นยำ

2. การจับคู่โค้ดภาษาเมอร์กินเข้ากับโค้ดสำหรับควบคุมเว็บไซต์แบบอัตโนมัติ

การจับคู่โค้ดภาษาเมอร์กินเข้ากับโค้ดสำหรับควบคุมเว็บไซต์แบบอัตโนมัติสามารถช่วยนักทดสอบในการลดระยะเวลา รวมถึงข้อผิดพลาดในการพัฒนาทดสอบ เพื่อให้ได้ผลลัพธ์การทดสอบที่ต้องการ ครบถ้วน สมบูรณ์ตามที่ระบุไว้ในเอกสารการทดสอบ

เอกสารอ้างอิง

- [1] Bluerotor, "Automation testing" Available on: <https://www.techterrotor.com/2019/08/automation-testing.html>
- [2] N. Palani, "Software Automation Testing Secrets Revealed, revised edition 2016"
- [3] Jash Unadkat, "BrowserStack, App & Brower Testing Made Easy" Available on: <https://www.browsersstack.com/guide/tdd-vs-bdd-vs-atdd>
- [4] Rose, "The Cucumber for Java Book", published February 2015
- [5] "Selenium Library" Available on: <https://www.selenium.dev/>
- [6] "Visual Studio 2019 Platform Targeting and Compatibility" Available on: <https://learn.microsoft.com/en-us/visualstudio/releases/2019/compatibility>
- [7] "Extent Reports" Available on: <https://www.extentreports.com/>
- [8] "SET" Available on: <https://www.set.or.th/th/home>

การพัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคล Development of Personalized Programming Tutoring Systems

ทศพล บ้านคลองสี่ (Todsapon Banklongsi)¹ และจรูญ แสนราช (Charun Sanrach)²

¹ภาควิชาวิศวกรรมคอมพิวเตอร์และหุ่นยนต์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยกรุงเทพ

²ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระเจ้าเกด้าพระนครเหนือ

¹todsapon.b@bu.ac.th, ²charun.s@fte.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันการเรียนรู้การเขียนโปรแกรมนั้นมีหลายภาษา ซึ่งแต่ละภาษามีความนิยมที่ต่างกัน และภาษาซีเป็นหนึ่งในภาษาพื้นฐานที่เหมาะสมแก่การเรียนรู้ถึงโครงสร้างทางภาษาต่าง ๆ ที่สามารถนำไปปรับใช้กับภาษาอื่นได้ งานวิจัยนี้จึงได้นำเสนอการพัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคลในรูปแบบเว็บแอปพลิเคชัน มีวัตถุประสงค์เพื่อใช้เป็นเครื่องมือในการจัดการเรียนการสอนให้ผู้เรียนสามารถเรียนรู้และประเมินความสามารถในการเขียนโปรแกรมได้ด้วยตนเอง ระบบประกอบด้วย การลงทะเบียนและลงชื่อเข้าใช้งาน การจัดการคลังความรู้ การแนะนำการเรียนส่วนบุคคล การประเมินทักษะการเขียนโปรแกรม 4 ทักษะ ได้แก่ ความรู้พื้นฐาน การติดตาม การอธิบาย และการเขียนโปรแกรม การทำโครงงาน และการสรุปผลของผู้เรียน ผลลัพธ์ที่ได้สามารถช่วยกระตุ้นให้ผู้เรียนมีความสนใจในการเขียนโปรแกรม เข้าใจหลักการเขียนโปรแกรม การพัฒนาทักษะความสามารถด้านการเขียนโปรแกรมได้ดีขึ้น และเป็นต้นแบบเครื่องมือในการพัฒนาการสอนการเขียนโปรแกรมได้

คำสำคัญ: การเขียนโปรแกรมคอมพิวเตอร์ ระบบการสอน การเขียนโปรแกรม การแนะนำแบบตั้งกฎ ทักษะการเขียนโปรแกรม

Abstract

Coding classes are now accessible in several languages. Each language has a particular popularity that C language is one of the foundational languages

that can be used to learn different language patterns that can be applied to other languages. This research presents the development of personalized programming tutoring systems. Its purpose is to be used as a tool for teaching and learning so that learners may develop and evaluate their own programming abilities. The system includes registration and sign-in, knowledge base management, individualized learning coaching, and assessments of four programming skills: basic knowledge, tracking, explaining, and writing, projects, and summarizing the outcomes of the learners. The findings may be used to inspire learners to pursue a career in programming, grasp programming concepts, improve programming abilities, and serve as a model for developing teaching programming tools.

Keywords: Computer Programming, Programming Tutoring System, Rule-based Recommendation, Programming Skills

1. บทนำ

การเรียนรู้การเขียนโปรแกรมเป็นสิ่งที่สำคัญมากในยุคปัจจุบัน อีกทั้งจากสถานการณ์การแพร่ระบาดของเชื้อโควิด-19 ที่ผ่านมา ทำให้สถาบันการศึกษาต่าง ๆ ได้ปรับการเรียนการสอนมาอยู่ในแพลตฟอร์มออนไลน์มากขึ้น ซึ่งปัญหาที่พบส่วนใหญ่ผู้เรียนแต่ละคนมีทักษะการเรียนรู้ที่แตกต่างกัน จึงทำให้ความสามารถในการเรียนรู้การเขียนโปรแกรมของผู้เรียนไม่ดีเท่าที่ควร ถ้ามีระบบที่ให้คำแนะนำและจัดเส้นทางเรียนที่ตอบโจทย์ลักษณะการ

เรียนของผู้เรียนนั้นได้ จะช่วยเพิ่มความสามารถในการเรียนรู้การเขียนโปรแกรมได้ดียิ่งขึ้น งานวิจัยนี้จึงได้พัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคลในรูปแบบเว็บแอปพลิเคชัน โดยใช้ภาษาซี [1] เป็นภาษาสำหรับการสอนการเขียนโปรแกรม การแนะนำแบบกฎ [2] ช่วยแนะนำเส้นทางการเรียนของแต่ละบุคคล นำแบบจำลองเอดีอาร์ไอ [3] มาช่วยในการออกแบบแบบทดสอบและการทำโครงการ เพื่อประเมินทักษะการเขียนโปรแกรม 4 ทักษะ ได้แก่ การทำความเข้าใจ การติดตาม การอธิบาย และการเขียนโปรแกรม พร้อมทั้งประเมินผลทักษะการเขียนโปรแกรมของผู้เรียน ซึ่งสามารถช่วยกระตุ้นให้ผู้เรียนมีความสนใจในการเขียนโปรแกรม เข้าใจหลักการเขียนโปรแกรม และพัฒนาทักษะความสามารถด้านการเขียนโปรแกรมส่วนบุคคลได้

2. ทฤษฎีที่เกี่ยวข้อง

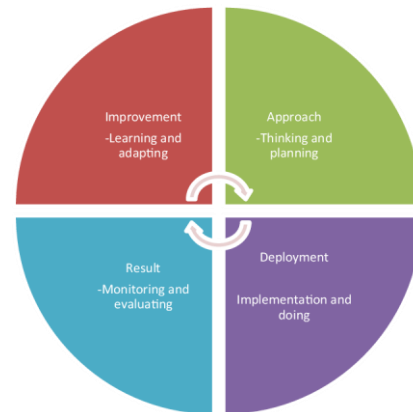
2.1 เนื้อหาการเรียนเขียนโปรแกรม

เนื้อหาการเรียนการเขียนโปรแกรมคอมพิวเตอร์ประกอบด้วย 11 หัวข้อ ได้แก่ การแนะนำภาษาซี (Introduction) ชนิดข้อมูล (Data Types) ตัวดำเนินการ (Operators) อินพุต/เอาต์พุต (Input/Output) คำสั่งเลือกเงื่อนไข (Selection Statements) คำสั่งวนซ้ำ (Looping Statements) อาร์เรย์ (Array) สตริง (String) พอยน์เตอร์ (Pointer) ฟังก์ชัน (Function) และ โครงสร้าง (Structure) ซึ่งใช้ในการเรียนการสอนวิชา CS122 ในหมวดวิชาเฉพาะด้านบังคับเรียน หลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์และหุ่นยนต์ หลักสูตรปรับปรุง พ.ศ. 2561 มหาวิทยาลัยกรุงเทพ

2.2 การออกแบบแบบทดสอบการเรียนเขียนโปรแกรม

การออกแบบแบบทดสอบการเรียนเขียนโปรแกรมด้วยแบบจำลอง ADRI [2] ซึ่งเป็นแบบจำลองที่ประกอบด้วย 4 ขั้นตอน ได้แก่ วิธีการ (Approach) เป็นขั้นวิธึสร้างกลยุทธ์การแก้ไขปัญหา ซึ่งเป็นส่วนของอัลกอริทึมและผังงาน การปรับใช้ (Deploy) เป็นขั้นการปรับใช้ความรู้ด้วยการเขียนโปรแกรม ผลลัพธ์ (Result) เป็นขั้นการแสดงผลลัพธ์ที่ได้จากการเขียนโปรแกรม และการปรับปรุง

(Improvement) เป็นขั้นการปรับปรุงการเรียนรู้และการแนะนำ ดังภาพที่ 1 และ 2



ภาพที่ 1: แบบจำลอง ADRI [2]

Write a C++ program that prompts the user to enter the elapsed time for an event in hours, minutes, and seconds. The program then calculates and outputs the elapsed time in seconds.

Step 1: Approach – Problem solving strategies

Pseudo-code

1. Start
2. Read Second, Minutes, Hour
3. Total_time_second = Second + Minutes*60+ Hour * 3600
4. Print Total_time_second
5. Stop

A: Approach

Flowchart

```

graph TD
    Start([Start]) --> Read([Read Second, Minutes, Hour])
    Read --> Calc[Total_time_second = Second + Minutes * 60 + Hour * 3600]
    Calc --> Print[/Print Total_time_second/]
    Print --> Stop([Stop])
                    
```

Step 2: Programming Knowledge

```

#include <iostream>
using namespace std;
int main()
{
    int second, hour, minute, total_time_second;
    cout<<"enter the time in second:";
    cin>>second;
    cout<<"\n";
    cout<<"enter the time in minutes:";
    cin>>minute;
    cout<<"\n";
    cout<<"enter the time in hours:";
    cin>>hour;
    cout<<"\n";
    total_time_second = second + minute*60 + hour*3600;
    cout<<"The total time in second = "<<total_time_second<<"\n";
    return 0;
}
                    
```

D: Deployment

Step 3: Result

Expected output: enter the time in second: 65
 enter the time in minutes: 20
 enter the time in hours: 2
 The total time in second: 8465

Process: total_time_second = 65 + 20*60 + 2*3600

Achieved output: Achieved: Not Achieved:

Goals: Achieved: Not Achieved:

Syntax Error: Incompatible data types / Initialization

Example:

```

7 int num = "Hello";
                    
```

Reason: integer value is expected in variable num on line 7. The string value (Hello) is assigned to integer variable (num).

R: Result

ภาพที่ 2: บททดสอบการเขียนโปรแกรมด้วย ADRI [2]

2.3 ทักษะการเขียนโปรแกรม

ทักษะการเขียนโปรแกรมโดยใช้ทฤษฎีการวัดแบบจำลองราสซ์ [3] ประกอบด้วย 4 ระดับ ดังตารางที่ 1

ตารางที่ 1: ระดับทักษะการเรียนเขียนโปรแกรม

ระดับทักษะ	คำอธิบาย
1: Knowledge	ตอบคำถามโดยใช้ความรู้พื้นฐานได้
2: Tracing	ติดตามผลลัพธ์/ค่าของตัวแปรของโค้ดได้
3: Explaining	อธิบายการทำงานของโปรแกรมได้
4: Writing	เขียนโปรแกรมตามโจทย์ที่กำหนดได้

2.4 การเก็บคะแนน

การประเมินทักษะการเขียนโปรแกรมโดยใช้เกณฑ์การให้คะแนนแบบรูบริก (Rubric) [4] ซึ่งเป็นเกณฑ์การให้คะแนนที่กำหนดคุณภาพคำตอบของแต่ละระดับคะแนนไว้อย่างชัดเจนเพื่อช่วยให้การตรวจคำตอบ โดยใช้เกณฑ์การให้คะแนนแบบภาพรวม (Holistic Score) เป็นการกำหนดระดับคะแนนแล้วอธิบายหลักเกณฑ์ของคะแนนในแต่ละระดับแบบรวม ๆ ให้ลดหลั่นตามลำดับคุณภาพของคำตอบ ตัวอย่างดังตารางที่ 2

ตารางที่ 2: ตัวอย่างเกณฑ์การให้คะแนนแบบภาพรวม

คะแนน	เกณฑ์การให้คะแนน
4	วิเคราะห์ได้ว่าโจทย์กำหนดอะไรให้บ้าง แสดงวิธีหาคำตอบชัดเจน คำตอบถูก
3	วิเคราะห์ได้ว่าโจทย์กำหนดอะไรให้บ้าง แสดงวิธีหาคำตอบชัดเจน แต่คำตอบผิด
2	วิเคราะห์ได้ว่าโจทย์กำหนดอะไรให้บ้าง แสดงวิธีหาคำตอบแต่ไม่ชัดเจน
1	วิเคราะห์ได้ว่าโจทย์กำหนดอะไรให้บ้าง ไม่แสดงวิธีหาคำตอบ
0	วิเคราะห์ไม่ได้ว่าโจทย์กำหนดอะไรบ้าง และไม่แสดงวิธีหาคำตอบ

2.5 การแนะนำแบบตั้งกฎ

การแนะนำแบบตั้งกฎ (Rule-based Recommendation) [2] เป็นวิธีการเขียนกฎหรือเงื่อนไข เพื่อแยกแยะผลลัพธ์ของสิ่งต่าง ๆ ซึ่งนิยมนำมาใช้ในการพัฒนาเป็นระบบผู้เชี่ยวชาญ โดยใช้รูปแบบคำสั่ง IF (การแสดงผลเงื่อนไข) ... THEN (การแสดงผลสรุป) เช่น IF ไฟฟ้ามอดนเป็นสีเขียว THEN ให้ข้ามถนนได้

3. วิธีการดำเนินงานวิจัย

วิธีการดำเนินงานวิจัยใช้รูปแบบ ADDIE Model [5] ดังนี้

3.1 การวิเคราะห์ (Analysis)

ผู้วิจัยได้ทำการค้นคว้าทฤษฎีและงานวิจัยที่เกี่ยวข้อง การพัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคล แล้วพบว่าการนำแบบจำลองเอ็ดอาร์ไอมาประยุกต์ใช้ในการออกแบบการสอนการเขียนโปรแกรมสามารถช่วย

ส่งเสริมและพัฒนาทักษะการเขียนโปรแกรมของผู้เรียนได้ [3] จึงนำมาใช้ในการพัฒนาระบบของงานวิจัยนี้ โดยนำเนื้อหาการเขียนโปรแกรมภาษาซี การออกแบบบททดสอบด้วยแบบจำลอง ADRI การวัดทักษะการเขียนโปรแกรมด้วยแบบจำลองราล์ช การประเมินผลคะแนนแบบรูบริก และการแนะนำแบบตั้งกฎ มาใช้ในการพัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคลในรูปแบบเว็บแอปพลิเคชันเพื่อช่วยสนับสนุนการเรียนการสอนการเขียนโปรแกรม

3.2 การออกแบบ (Design)

การออกแบบระบบแบ่งเป็น 2 ส่วน ดังนี้

3.2.1 ออกแบบการทำงานของระบบ

ประกอบด้วย 7 ส่วน ดังภาพที่ 3

1) การเก็บข้อมูลลงทะเบียนและลงชื่อเข้าใช้ระบบ คือ ส่วนแรกที่ใช้ใช้งานทั้งผู้เรียนและผู้สอนต้องทำการสมัครและลงชื่อเข้าใช้งาน ถ้าผู้ใช้งานเป็นผู้เรียนใหม่ระบบจะทำการส่งไปยังหน้าแบบทดสอบก่อนเรียน ถ้าเป็นผู้สอนระบบจะทำการส่งไปหน้าส่วนของฝั่งผู้สอน

2) การจัดการคลังความรู้ของการเขียนโปรแกรม เป็นส่วนที่ทำการจัดการเนื้อหาการเรียนของผู้ใช้งานฝั่งผู้เรียน เมื่อผู้เรียนเข้าสู่เนื้อหาจะทำการตรวจสอบการเรียนของผู้เรียน ถ้าตรวจสอบว่าเรียนครบทุกเนื้อหา ระบบจะทำการส่งไปยังแบบทดสอบหลังเรียน ถ้ายังไม่ครบจะส่งไปยังแบบทดสอบภายในเนื้อหา

3) การแนะนำการเรียนส่วนบุคคล เป็นส่วนการแนะนำ 2 เส้นทาง ได้แก่ แนะนำแบบหาเส้นทางที่เร็วที่สุด โดยแนะนำโครงการที่ใช้จำนวนเนื้อหาที่ต้องเรียนน้อยที่สุดก่อน ถ้ามีจำนวนเนื้อหาเท่ากันจะพิจารณาโดยเปรียบเทียบว่าต้องการเนื้อหาใดก่อนตามลำดับความสำคัญ และแนะนำจากการเลือกโครงการ โดยผู้เรียนทำการเลือกโครงการที่สนใจ ระบบจะทำการแนะนำเนื้อหาที่เกี่ยวข้องกับโครงการนั้น ผลลัพธ์ของระบบแนะนำจะนำไปสู่การเรียนเนื้อหาและบททดสอบต่าง ๆ ที่ช่วยทำให้การทำโครงการการเขียนโปรแกรมต่าง ๆ สำเร็จ

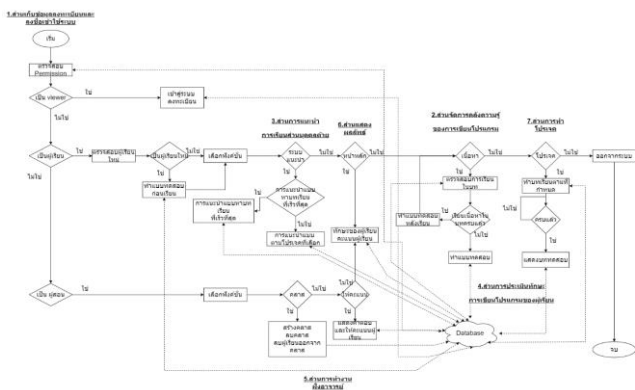
4) การประเมินทักษะการเขียนโปรแกรมของผู้เรียน เป็นส่วนการให้คะแนนในบททดสอบต่าง ๆ โดยเริ่มการ

ทำงานฝั่งผู้เรียน เมื่อผู้เรียนทำบททดสอบแล้ว ระบบจะทำการตรวจและให้คะแนนในส่วนของทักษะ Knowledge, Tracing และ Explaining ส่วนทักษะ Writing ระบบจะทำการส่งคำตอบของผู้ให้ผู้สอนเป็นผู้ตรวจและให้คะแนนตามเกณฑ์บริบทที่ผู้สอนกำหนดขึ้น

5) การทำงานฝั่งผู้สอน แบ่งเป็น 2 ส่วน คือ ส่วนการจัดการห้องเรียน ผู้สอนสามารถสร้างห้องเรียน ตั้งชื่อห้องเรียน และลบห้องเรียน ได้โดยใช้รหัสโทเค็นที่ได้จากการสร้างห้องเรียน สามารถลบผู้เรียนออกจากห้องเรียนได้ด้วยการใส่อีเมลของผู้เรียนที่ต้องการลบออกจากห้องเรียน ในกรณีของผู้เรียนสามารถเข้าห้องเรียนได้ด้วยการใส่โทเค็นที่ผู้สอนได้จากการสร้างห้องเรียน และส่วนการให้คะแนนข้อเขียนทักษะ Writing ของผู้เรียน โดยผู้สอนทำการประเมินและให้คะแนน ประกอบด้วย การให้คะแนนส่วนบททดสอบทักษะ Writing ของเนื้อหาต่างๆ และส่วนโครงงาน เมื่อผู้สอนตรวจแล้วระบบจะทำการบันทึกคะแนนในฐานข้อมูลและแสดงผลคะแนนให้กับผู้เรียน

6) การแสดงผลลัพธ์ เป็นส่วนแสดงผลลัพธ์ของทักษะต่าง ๆ ของผู้เรียน ถ้าเป็นผู้เรียนจะแสดงทักษะส่วนบุคคลของผู้เรียนคนนั้น ๆ ถ้าเป็นผู้สอนระบบจะแสดงผลสรุปของทักษะที่อธิบายความสามารถของผู้เรียนทั้งหมดในห้องเรียนว่าเก่งหรือด้อยในทักษะใด ทำให้ผู้สอนรู้ถึงสถานการณ์และนำไปปรับวิธีการสอนได้

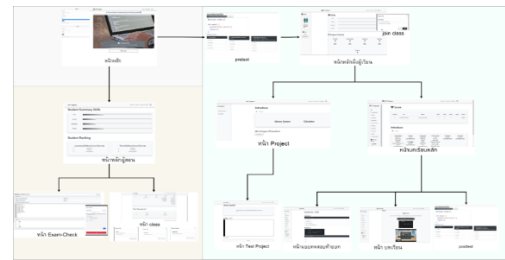
7) การทำโครงงาน เป็นส่วนเข้าทำโครงงานของผู้เรียน โดยระบบทำการตรวจสอบว่าผู้เรียนได้เรียนเนื้อหาและผ่านการทำบททดสอบที่เกี่ยวข้องกับโครงงานนั้น ๆ ครบหรือไม่ ถ้าครบตามเงื่อนไข ระบบจะทำการปลดล็อกโครงงานนั้น แล้วผู้เรียนสามารถเข้าไปทำโครงงานนี้ได้



ภาพที่ 3: ผังงานการทำงานของระบบ

3.2.2 ออกแบบส่วนต่อประสานผู้ใช้ของระบบ

ออกแบบส่วนต่อประสานผู้ใช้ของระบบ ดังภาพที่ 4



ภาพที่ 4: การออกแบบส่วนต่อประสานผู้ใช้ของระบบ

3.3 การพัฒนา (Development)

ผู้วิจัยพัฒนาเขียนโปรแกรมแบบ Full Stack เพื่อพัฒนาระบบในรูปแบบเว็บแอปพลิเคชัน ทั้งในส่วน Front-end ที่ใช้โปรแกรม Figma ในการออกแบบหน้าตาของเว็บแอปพลิเคชัน ใช้ภาษา HTML, CSS, JavaScript ในการพัฒนาการทำงานของหน้าเว็บแอปพลิเคชันฝั่ง Client และส่วน Back-end ที่ใช้ภาษา NodeJS, Express และ SQL ในการพัฒนาการทำงานของฝั่ง Server ที่เชื่อมโยงข้อมูลระหว่างฐานข้อมูลกับการประมวลผลฝั่ง Back-end และ Front-end ให้สามารถทำงานร่วมกันได้

3.4 การนำไปใช้ (Implementation)

ผู้วิจัยได้พัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคลโดยสร้างเป็นเว็บแอปพลิเคชันที่สามารถเปิดผ่านเว็บเบราว์เซอร์ใช้งานได้ทั้งเครื่องคอมพิวเตอร์และสมาร์ทโฟน แล้วนำไปใช้จริงเพื่อทดสอบความถูกต้องและเหมาะสมกับกลุ่มตัวอย่างที่เป็นผู้เชี่ยวชาญด้านการเขียนโปรแกรม จำนวน 26 คน เพื่อให้สามารถนำผลลัพธ์ที่ได้มาใช้ประโยชน์ในการพัฒนาระบบก่อนนำไปใช้สอนกับผู้เรียนจริงในอนาคต

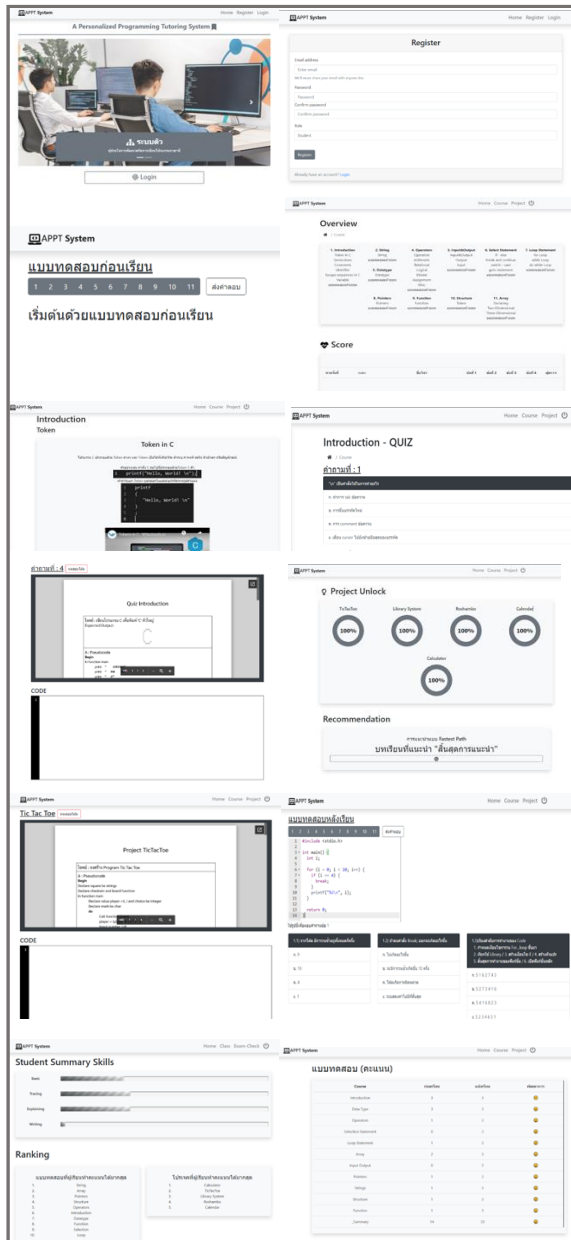
3.5 การประเมิน (Evaluation)

ระบบการสอนการเขียนโปรแกรมส่วนบุคคลนี้ได้ปรับแต่งเนื้อหาบทเรียน การทดสอบ การวัดและประเมินผล ให้เหมาะสมกับผู้เรียนตามผลลัพธ์ที่ได้ของงานวิจัยที่เกี่ยวข้องต่างๆ และตามคำแนะนำของผู้สอนในรายวิชาการเขียนโปรแกรมเดิม ซึ่งได้ทำการประเมินคุณภาพเบื้องต้นของเครื่องมือดังกล่าว ก่อนนำไปใช้ทดสอบจริงกับกลุ่มตัวอย่างเพื่อประเมินความเหมาะสมกับผู้เชี่ยวชาญด้านการเขียนโปรแกรม

4. ผลการทดลอง

4.1 ผลการทดลองการทำงานของระบบ

ผลการทดสอบระบบการเขียนโปรแกรมส่วนบุคคลสามารถทำงานได้อย่างถูกต้อง ประกอบด้วย การเก็บข้อมูลลงทะเบียนและลงชื่อเข้าใช้ระบบ การจัดการคลังความรู้ของการเขียนโปรแกรม การแนะนำการเรียนส่วนบุคคล การประเมินทักษะการเขียนโปรแกรมของผู้เรียน การทำงานฝั่งผู้สอน การทำโครงการ และการแสดงผลลัพธ์ ดังภาพที่ 5



ภาพที่ 5: ผลการทดลองการทำงานของระบบ

4.2 ผลการประเมิน

ประเมินจากกลุ่มตัวอย่างผู้เชี่ยวชาญด้านการเขียนโปรแกรม จำนวน 26 คน โดยกำหนดค่าเฉลี่ยความพึง

พอใจเป็นมาตราส่วนคะแนน 5 ระดับ โดยค่าเฉลี่ยตามเกณฑ์ดังนี้

ค่าเฉลี่ย 4.51 – 5.00 คือ มีความเหมาะสมมากที่สุด

ค่าเฉลี่ย 3.51 – 4.50 คือ มีความเหมาะสมมาก

ค่าเฉลี่ย 2.51 – 3.50 คือ มีความเหมาะสมปานกลาง

ค่าเฉลี่ย 1.51 – 2.50 คือ มีความเหมาะสมน้อย

ค่าเฉลี่ย 1.01 – 1.50 คือ มีความเหมาะสมน้อยที่สุด

และได้ผลการประเมินความพึงพอใจ ดังตารางที่ 3

ตารางที่ 3: ความพึงพอใจของระบบการเขียนโปรแกรมส่วนบุคคล

หัวข้อการประเมิน	ค่าเฉลี่ย
1. เนื้อหา	
1.1 ข้อมูลมีความถูกต้องครบถ้วน	4.5
1.2 สามารถนำเนื้อหาไปใช้งานได้จริง	4.4
2. การแนะนำ	
2.1 ระบบ Rule-based แนะนำเนื้อหาที่ต้องการได้จริง	4.5
3. รูปแบบการเข้าใช้งาน	
3.1 การจัดวางองค์ประกอบของเว็บแอปพลิเคชัน	3.8
3.2 ขนาดและสีของตัวอักษรที่แสดง	3.6
3.3 การออกแบบกราฟิก	3.6
4. ความพึงพอใจต่อการใช้งาน	
4.1 บทเรียน	4.4
4.2 ห้องเรียน	4.1
4.3 บททดสอบ	4.2
4.4 ระบบ Rule-based การแนะนำ	4.1
ค่าเฉลี่ยรวม	4.12
ส่วนเบี่ยงเบนมาตรฐาน	0.33

5. สรุป

งานวิจัยนี้ได้ทำการพัฒนาระบบการเขียนโปรแกรมส่วนบุคคล เพื่อนำไปใช้เป็นเครื่องมือช่วยในการเรียนการสอนและประเมินทักษะการเขียนโปรแกรมของผู้เรียน ซึ่งในงานวิจัยนี้ได้สร้างต้นแบบระบบการเขียนโปรแกรมส่วนบุคคลและประเมินกับผู้เชี่ยวชาญด้านการเขียน

โปรแกรมก่อนนำไปใช้กับผู้เรียน ผลการตอบรับได้คะแนนเฉลี่ยอยู่ที่ 4.12 คะแนน จากคะแนนเต็ม 5 คะแนน ซึ่งอยู่ในเกณฑ์ระดับที่มีความเหมาะสมมาก ซึ่งถือได้ว่างานวิจัยนี้สามารถนำไปใช้เป็นตัวแบบในการพัฒนาเครื่องมือสำหรับใช้สอนและช่วยส่งเสริมการเขียนโปรแกรมให้กับผู้เรียนได้เรียนรู้เข้าใจ และพัฒนาทักษะด้านการเขียนโปรแกรมของตนเองได้

6. กิตติกรรมประกาศ

ขอขอบคุณนักศึกษา และผู้ร่วมทดสอบงานวิจัย คณาจารย์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยกรุงเทพ ที่ให้ความช่วยเหลือและสนับสนุนการพัฒนาระบบการสอนการเขียนโปรแกรมส่วนบุคคลมาเป็นอย่างดี

เอกสารอ้างอิง

- [1] M Greg, and M. Dean, *C programming: absolute beginner's guide*, Pearson Education, 2013.
- [2] J. Tarus, Z. Niu, G. Mustafa “Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning,” *Artificial intelligence review*. pp. 21-48, 2018.
- [3] S. Malik, & J. Coldwell-Neilson, “A model for teaching an introductory programming course using ADRI,” *Education and Information Technologies*, Vol. 22, pp. 1089-1120, 2017.
- [4] A. Mustapha, et al. “Generic Assessment Rubrics for Computer Programming Courses,” *Turkish Online Journal of Educational Technology*, Vol.15 No.1, pp.53-68, 2016.
- [5] T. Trust and E. Pektas, “Using the ADDIE Model and Universal Design for Learning Principles to Develop an Open Online Course for Teacher Professional Development,” *Journal of Digital Learning in Teacher Education*, Vol. 34, No. 4, pp. 219-233, 2018.

การตรวจสอบพื้นที่จอดรถผ่านแอปพลิเคชันไลน์ The Checking the Parking with LINE Application

สิทธิชอบ รัตรอดกิจ (Sittichop Ratrojki) ¹ พัชรพงษ์ จรัสโรจน์โกภา (Pattarapong Jarrusrojpo) ²
ธนพนธ์ โยชนา (Thanaphon Yochana) ³ อรวรรณ ชุมหปราม (Orawan Chunhapran) ⁴ และรัฐ บุรีรัตน์ (Rath Burirat) ⁵
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก วิทยาเขตจรัลพนธ์
¹sittichop.rat@rmutto.ac.th, ²pattarapong.jar@rmutto.ac.th, ³thanaphon.yox@rmutto.ac.th,
⁴orawan_ch@rmutto.ac.th, ⁵rath_bu@rmutto.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบพื้นที่จอดรถผ่านแอปพลิเคชันไลน์ โดยนำแนวคิดเทคโนโลยีอินเทอร์เน็ตทุกสรรพสิ่งมาประยุกต์ใช้เพื่อให้ผู้ใช้รถสามารถค้นหาที่จอดรถได้ พัฒนาโปรแกรมด้วยภาษาซีในการควบคุมตรวจสอบพื้นที่ว่างผ่านเครือข่ายอินเทอร์เน็ตและแสดงสถานะพื้นที่ว่างผ่านแอปพลิเคชันบนสมาร์ตโฟน ในส่วนของการตรวจสอบพื้นที่ว่างนั้นได้นำอุปกรณ์บอร์ดอาร์ดูโนเอสพี 32 ควบคุมการทำงานของเซ็นเซอร์อินฟราเรด เพื่อตรวจสอบระยะห่างระหว่างรถยนต์กับอุปกรณ์ตรวจสอบพื้นที่จอดรถ และออกแบบโมเดลพื้นที่จอดรถด้วยแผ่นโพลีโพรพิลีน ผลการทดลองแสดงให้เห็นว่าเมื่อไม่มีรถยนต์จอดในช่องจอดรถยนต์แต่ละช่องจะส่งค่าข้อมูลเป็น 0 และส่งค่าเป็น 1 เมื่อตรวจพบว่ามีรถยนต์เข้ามาจอดในพื้นที่จอดรถ ซึ่งระบบสามารถแจ้งข้อมูลให้ผู้ใช้ตัดสินใจในการหาพื้นที่จอดรถในอาคารจอดรถต่าง ๆ ได้จริง

คำสำคัญ: อินเทอร์เน็ตทุกสรรพสิ่ง แอปพลิเคชัน อาร์ดูโน เซ็นเซอร์

Abstract

The purpose of this research was to examine the vacant parking spaces via the LINE application. The concept of Internet of Things technology is applied to allow car users to find parking spaces. With the development of a program in C to control and check available space through the Internet network

and display the status of free space through an application on a smartphone, the Arduino ESP32 board device is used to check the availability of free space, control the operation of the infrared sensor to measure the distance between the car and the parking monitoring device, and create a polypropylene sheet model of a parking space. The experimental results show that when no car is parked in each parking space, the system sends a value of 0, and the system sends a value of 1 when a car is detected in the parking area. This system can inform the user and help them decide on finding parking spaces in different buildings.

Keywords: Internet of Things, Application, Arduino, Sensor

1. บทนำ

ปัจจุบันความก้าวหน้าทางเทคโนโลยีนั้นมีการพัฒนาแบบก้าวกระโดด นวัตกรรมและสิ่งประดิษฐ์ใหม่ๆ ที่เกิดขึ้นเกือบทั้งหมดล้วนสามารถเชื่อมต่อกันผ่านอินเทอร์เน็ตได้โดยจะเรียกอุปกรณ์หรือสิ่งเหล่านี้ว่า อินเทอร์เน็ตเพื่อสรรพสิ่ง ไม่ว่าจะเป็นเซ็นเซอร์ตรวจวัดต่างๆ ที่สามารถตรวจวัดสภาพแวดล้อม การจราจรหรือแม้กระทั่งการใช้พลังงานไฟฟ้าของแต่ละพื้นที่ การเพิ่มขึ้นของ IoT มีผลต่อวิถีชีวิตของคนไทย ทั้งการใช้ชีวิตที่ดีขึ้นและช่องทางการสร้างมูลค่าเพิ่มให้กับ IoT ที่มากขึ้น

อินเทอร์เน็ตของสรรพสิ่ง (Internet of Things: IoT) ตามนิยามของสหภาพโทรคมนาคมระหว่างประเทศ (International Telecommunication Union: ITU) หมายถึง โครงข่ายสื่อสารที่มีการเชื่อมต่ออุปกรณ์สื่อสาร เครื่องใช้ไฟฟ้า ยานพาหนะ ฯลฯ โดยอาศัยการส่งระบบ อิเล็กทรอนิกส์ ซอฟต์แวร์ อุปกรณ์เซ็นเซอร์ และส่วน เชื่อมต่อโครงข่าย ที่จะช่วยให้อุปกรณ์และวัตถุดังกล่าว สามารถเก็บหรือแลกเปลี่ยนข้อมูลระหว่างกันได้ หาก จินตนาการถึงอนาคตอันใกล้ ผู้คนจำนวนมากอาจจะมี อุปกรณ์ที่สามารถบันทึกข้อมูลการเดินทาง เช่น ระยะทาง ช่วงเวลา การจราจร สถานที่จอดรถ เพื่อให้ผู้ใช้งาน สามารถวิเคราะห์ประสิทธิภาพในการเดินทางของตน

ตัวอย่างของสถานการณ์ดังกล่าวเป็นสิ่งที่ใกล้ตัวเข้ามา ทุกที อย่างไรก็ตาม ในปัจจุบันเทคโนโลยี อินเทอร์เน็ต ของสรรพสิ่ง นั้นมีแนวโน้มของการเจริญเติบโตที่สูงมาก การเจริญเติบโตดังกล่าวจะเปิดโอกาสให้มีการสร้าง นวัตกรรมและการบริการรูปแบบใหม่ รวมถึงช่วยให้ มนุษย์สามารถเข้าถึงข้อมูลอย่างหลากหลาย ในระดับที่ ไม่เคยเป็นมาก่อนซึ่งจะช่วยสร้างเครื่องมืออำนวยความสะดวกให้กับมนุษย์และสร้างมูลค่าทางเศรษฐกิจอย่าง มหาศาลในอนาคต ซึ่งในปัจจุบันผู้คนมีการใช้รถยนต์ ในการเดินทางมากขึ้น [1] เพราะสะดวกสบายในการเดินทาง โดยเฉพาะการเดินทางไปยังสถานที่ท่องเที่ยวหรือตามงาน ต่างๆ ที่มีการจัดขึ้น ตามเทศกาลต่าง ๆ แต่ในการไปยัง สถานที่นั้น อันดับแรกของการเดินทางไปยังยังสถานที่ ปลายทาง คือ การหาที่จอดรถ เมื่อคนเยอะและสถานที่ จอดรถที่มีจำกัด ทำให้ต้องใช้ เวลาในการหาที่จอดรถ พอสมควร อีกทั้งยังทำให้การจราจรติดขัดเป็นอย่างมาก [2] โดยที่จอดรถนั้น ส่วนใหญ่ยังใช้ระบบเดิมอยู่คือมี เจ้าหน้าที่รักษาความปลอดภัยดูแลเรื่องการจัดระเบียบใน การจอดรถ ทำให้รถที่เข้ามาจอดไม่เป็นระเบียบการจราจร ติดขัด หาที่จอดยาก เนื่องจากไม่สามารถรู้ได้ว่าพื้นที่ใน การจอดรถตรงไหนยังว่างและสามารถเข้าจอดได้ ซึ่งใน ปัจจุบันงานทุกระบบมีการนำเครื่องมือทางคอมพิวเตอร์ เข้ามาใช้งานในการบริหารงานมากขึ้น เพื่อความทันสมัย และลดแรงงาน ดังนั้นการนำเครื่องมือทางคอมพิวเตอร์ มาช่วยในการบริหารจัดการ จึงมีบทบาทในการทำงาน

เพื่อเพิ่มประสิทธิภาพมากขึ้น ดังนั้นงานวิจัยนี้จึงเสนอ ระบบตรวจสอบพื้นที่จอดรถผ่านแอปพลิเคชัน ไลน์ โดยการนำเทคโนโลยีอาร์ดูโน (Arduino) มาประยุกต์ [3] ใช้เพื่อให้ผู้ใช้งานสามารถตรวจสอบพื้นที่จอดรถผ่าน ทางแอปพลิเคชันไลน์

2. วรรณกรรมที่เกี่ยวข้อง

A. Joshi, A. T. Hariram, K. M. Vishal Somaiya and M. Hussain [2] ศึกษาการค้นหาค้นหาพื้นที่ว่างและจัดการกับ ยานพาหนะที่เข้าและออกโดยใช้เซ็นเซอร์อินฟราเรดและ บอร์ดอาร์ดูโนโอโน เพื่อตรวจจับสภาพแวดล้อมบางส่วน เซ็นเซอร์อินฟราเรดสามารถวัดความอบอุ่นของวัตถุและ ตรวจจับการเคลื่อนไหวได้โดยระบบนี้ใช้เพื่อค้นหา ตำแหน่งที่จอดรถที่ว่างอยู่ และนำทางคนขับไปยังพื้นที่ ที่ต้องการ นอกจากนี้ไฟไดโอดเปล่งแสง (Light Emitting Diode: LED) จะถูกวางไว้ตามทางเดินรถ เพื่อใช้ในการ ตรวจสอบ ว่าเซ็นเซอร์ทำงานอย่างปลอดภัย โดยใช้แบตเตอรี่ 12 โวลต์ เพื่อจ่ายไฟให้อุปกรณ์ทั้งหมด

S. Nandyal, S. Sultana and S. Anjum [4] โครงการนี้ ประกอบด้วยกลยุทธ์การจัดกลุ่มที่ขับเคลื่อนด้วยเวลา ซึ่งดูแลปัญหาการจอดรถโดยใช้เทคนิคการเปิดงานการ ค้นหาจุดจอดรถ ที่ช่วยให้ผู้ขับขี่สามารถระบุตำแหน่งที่ ว่างได้ โดยใช้เซ็นเซอร์ตรวจจับคลื่นอัลตราโซนิกเชื่อมต่อกับ อาร์ดูโนโอโน เพื่อระบุพื้นที่ว่างด้วยการวัดระยะทาง ในการค้นหาช่องว่างในที่จอดรถเพื่อจอดรถและช่วยให้ผู้ ขับขี่ค้นหาช่องว่างได้ง่ายด้วยการแสดงไฟไดโอดเปล่งแสง (Light Emitting Diode: LED) และช่วยลดระยะเวลาในการ ค้นหาพื้นที่จอดรถ

D. Santoso, M. S. Hakim, M. N. Rachfian, M. L. Hidayat [5] เป็นงานออกแบบระบบลานจอดรถอัจฉริยะที่ ช่วยให้ผู้ใช้สามารถทราบข้อมูลเกี่ยวกับพื้นที่ในการจอด รถโดยได้นำอุปกรณ์มาใช้คือบอร์ดอาร์ดูโนเอสพี 8266 และเซ็นเซอร์อัลตราโซนิกและได้ออกแบบหลักการ ทำงานของแอปพลิเคชันคือให้ผู้ใช้ล็อกอินเข้าสู่ระบบ จากนั้นจะแสดงหน้าต่างเมนูขึ้นมาให้เลือก Search Mall Location จากนั้นเซ็นเซอร์จะทำการตรวจจับว่ามีรถยนต์

จอดอยู่หรือไม่ในกรณีที่มีรถจอดอยู่ให้ขึ้นแสดงว่า “DISPLAY NOTIFICATION THERE IS NO AVAILABLE PARKING SPACE” ในกรณีที่ไม่มีรถจอดอยู่ให้แสดงว่า “DISPLAY AVAILABLE PARKING SPACE BASED FLOOR THET IN”

H. Tanti, P. Kasodariya, S. Patel, D. H. Rangrej [6] ศึกษาการใช้อินเทอร์เน็ตทุกสรรพสิ่ง ตามระบบการจัดการที่จอดรถ โดยใช้เซ็นเซอร์อินฟราเรดและโหนดเอ็มซียู 8266 เพื่อเชื่อมต่อโครงสร้างพื้นฐานพื้นที่จอดรถจริงกับเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งให้บริการการจัดการอัจฉริยะบนคลาวด์ เพื่อนำแนวคิดนี้ไปใช้จะต้องมีการพัฒนาแอปพลิเคชันบนมือถือ แอปพลิเคชันมือถือนี้จะอนุญาตให้ผู้ใช้ตรวจสอบพื้นที่ว่างของที่จอดรถและจองที่จอดรถ ที่จอดรถแต่ละแห่งจะติดตั้งระบบควบคุมที่ช่วยให้สามารถตรวจสอบจำนวนสถานที่จอดรถว่างและมีผู้ครอบครองและแจ้งให้ผู้ใช้ทราบเกี่ยวกับสถานะที่จอดรถ และแอปพลิเคชันจะแสดงการชำระค่าบริการที่จอดรถ ตามระยะเวลาการจอดรถ

ณัทกิตต์ ชาวชน และคณะ [7] ศึกษาการสร้างแบบจำลองลานจอดรถและพัฒนาาระบบเพื่อตรวจสอบช่องจอดว่างผ่านเว็บแอปพลิเคชัน พร้อมทั้งสามารถนำทางไปยังช่องจอดว่างที่ใกล้ที่สุด และสามารถแสดงสถานะไฟช่องจอดรถได้ งานวิจัยชิ้นนี้ได้นำเทคนิคการจัดเรียงจากน้อยไปมากมาประยุกต์ใช้ในการจัดเรียงช่องจอดรถที่ใกล้ที่สุดเพื่อแจ้ง ไปยังเว็บแอปพลิเคชัน ประกอบกับการใช้เซ็นเซอร์ (Sensor) มาช่วยในการตรวจสอบสถานะของช่องจอดแต่ละช่องจอด โดยใช้บอร์ดไมโครคอนโทรลเลอร์ (Board Microcontroller) โหนดอีเอสพี 8266 (Esp 8266) ภาษาซีซีพีลัสพลัส เซ็นเซอร์อัลตราโซนิก (Ultrasonic Sensors), ไฟแอลอีดี แสดงสถานะ และฐานข้อมูล Firebase มาเป็นเครื่องมือในการพัฒนาผลจากการพัฒนาระบบพบว่าสามารถตรวจสอบช่องจอดว่างได้พร้อมกับสามารถนำทางผู้ใช้งานไปยังช่องจอดที่ว่างที่ใกล้ที่สุด และสามารถแสดงไฟสถานะช่องจอดรถได้อย่างมีประสิทธิภาพ

เกรียงไกร สว่างวงศ์ และคณะ [8] ได้พัฒนาระบบจองที่จอดรถอัจฉริยะ โดยใช้เทคโนโลยี อินเทอร์เน็ตของทุก

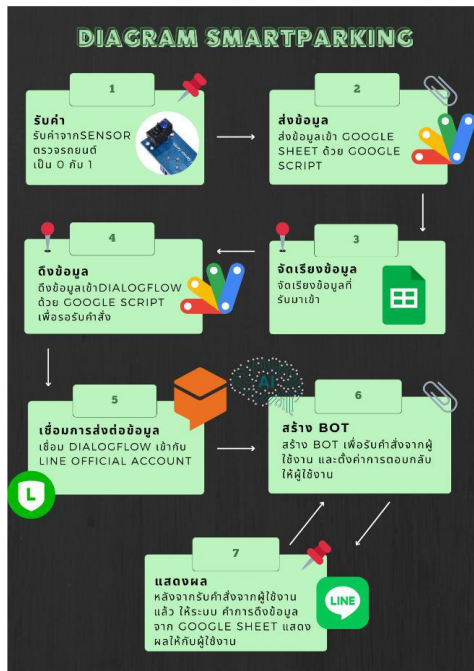
สรรพสิ่ง มาใช้ในการแก้ไขปัญหาในการหาพื้นที่สำหรับจอดรถ โดยผู้วิจัยได้พัฒนาแบบจำลองลานจอดรถ อาคาร 8 คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพระนคร มีการนำอุปกรณ์บอร์ดอาร์ดูโน ยูโน อาร์สาม (Arduino Uno R3) บอร์ดโหนดเอ็มซียู อีเอสพี 8266 และเซ็นเซอร์วัดระยะทางอัลตราโซนิก มาใช้ในการจับวัตถุด้านหน้าเพื่อแสดงสถานะของที่จอดรถ โดยผู้ใช้สามารถจองที่จอดรถได้ผ่านแอปพลิเคชันสำหรับจองที่จอดรถ

มนิโรส บริรักษ์อรารินท และคณะ [9] ได้วิจัยเกี่ยวกับอินเทอร์เน็ตทุกสรรพสิ่ง กรณีศึกษา ระบบลานจอดรถอัจฉริยะ เพื่อให้ผู้ใช้รถสามารถค้นหาที่จอดรถได้สะดวก ซึ่งมีการนำอาร์เอฟไอดี และเทคโนโลยีรหัสแท่งมาใช้ในการระบุตำแหน่งที่จอดรถ และพัฒนาระบบในลักษณะของเว็บแอปพลิเคชัน และโมบายแอปพลิเคชันโดยอาศัยโปรแกรมภาษาพีเอชพี และระบบจัดการฐานข้อมูลมายเอสคิวแอล

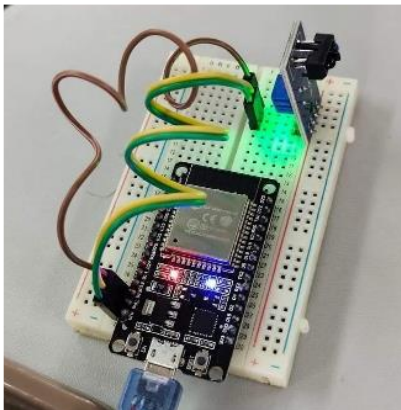
กษัตริธ อัสวพงศ์วานิช [13] ได้พัฒนาไลน์บอทของร้านอาหาร: กรณีศึกษาร้านกาแฟ ชรรรมาชาติ รีสอร์ท นครพนม และศึกษาความพึงพอใจผู้ใช้ระบบไลน์บอทในการสั่งอาหาร เพื่อให้ระบบไลน์บอทเข้ามาช่วยในการสั่งอาหารและตอบคำถามของลูกค้า พบผลการสอบถามความพึงพอใจของผู้ใช้งานระบบไลน์บอทด้านการนำมาใช้ประโยชน์ มีค่าเฉลี่ยรวม 4.88 โดยรวมอยู่ในระดับมากที่สุด

ซึ่งเป็นที่มาในการเลือกระบบไลน์บอทมาใช้ร่วมกับเทคโนโลยีอินเทอร์เน็ตทุกสรรพสิ่ง (IoT) เพื่อให้ผู้ใช้รถสามารถค้นหาที่จอดรถได้เพื่อตรวจสอบพื้นที่จอดรถผ่านแอปพลิเคชันไลน์ และช่วยในการควบคุมตรวจสอบพื้นที่ว่างผ่านเครือข่ายอินเทอร์เน็ต และแสดงสถานะพื้นที่ว่างผ่านแอปพลิเคชันบนสมาร์ตโฟน

3. วิธีการดำเนินการวิจัย



ภาพที่ 1: ไคอะแกรมการดำเนินการวิจัย



ภาพที่ 2: การเชื่อมต่อบอร์ด

3.1 การออกแบบฮาร์ดแวร์และซอฟต์แวร์

ออกแบบอุปกรณ์ตรวจสอบพื้นที่จอดรถโดยการใช้บอร์ดอาร์ดูโนโอเอสพี 32 (Arduino Esp32) ในการควบคุมการทำงานของเซ็นเซอร์อินฟราเรด (Infrared Sensor) เพื่อใช้ตรวจสอบระยะห่างระหว่างรถยนต์กับอุปกรณ์ตรวจที่จอดรถและออกแบบโมเดลพื้นที่จอดรถโดยใช้แผ่นโพลีโพรพิลีน (Polypropylene) ในการประกอบเป็นตัวเคสของอุปกรณ์ หลังจากนั้นจึงทำการติดตั้งบอร์ด อาร์ดู

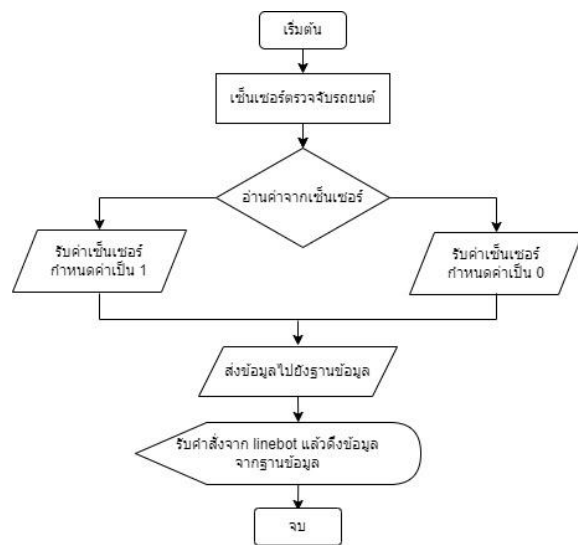
โนโอเอสพี 32 และเซ็นเซอร์อินฟราเรดที่เชื่อมต่อกันแล้ว ดังที่แสดงในภาพที่ 2 ลงในตัวอย่างอุปกรณ์



ภาพที่ 3: โมเดล

ออกแบบโปรแกรมด้วยการใช้โปรแกรมอาร์ดูโนโอไอดีอี (Arduino IDE) โดยการใช้ภาษาซี

3.2 การทำงานของระบบ



ภาพที่ 4: แสดงขั้นตอนการทำงานของอุปกรณ์

เริ่มต้นอ่านค่าเซ็นเซอร์ตรวจสอบสถานะของเซ็นเซอร์ จากนั้นส่งค่าข้อมูลไปเก็บไว้ในกูเกิลชีต (Google Sheets) โดยกำหนดค่าให้มีรถจอดเท่ากับ 1 และไม่มีรถจอดเท่ากับ 0 เมื่อระบบได้รับข้อมูลจะใช้กูเกิลสคริปต์และนำข้อมูลจากกูเกิลชีตเข้าสู่แอปพลิเคชันไลน์ เพื่อรอรับคำสั่งการใช้งานเมื่อผู้ใช้งานระบบต้องการตรวจสอบพื้นที่จอดรถให้ผู้ใช้งานทำการเพิ่มเพื่อน Smart Parking ดังภาพที่ 5 หลังจากเพิ่มเพื่อนแล้วให้ผู้ใช้งานพิมพ์เข้าไปในแอปพลิเคชันไลน์

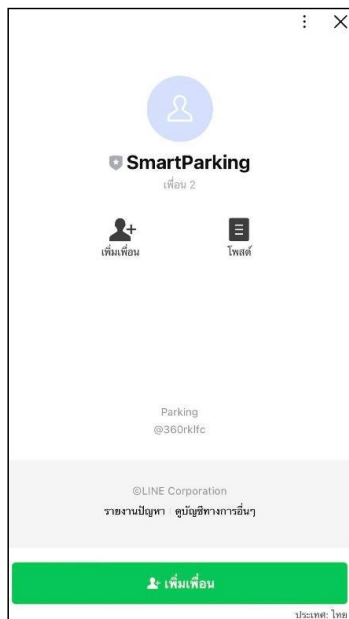
ว่า “ตรวจสอบที่จอดรถ” จากนั้นระบบจึงแสดงข้อมูลพื้นที่จอดรถ



ภาพที่ 5: สแกนคิวอาร์โค้ด Smart Parking เพิ่มเพื่อน



ภาพที่ 7: แสดงข้อมูลพื้นที่จอดรถ



ภาพที่ 6: แสดงไลน์ Smart Parking

3.3 การประเมินประสิทธิภาพของอุปกรณ์

การประเมินประสิทธิภาพของอุปกรณ์ตรวจสอบพื้นที่จอดรถโดยผลการทดสอบจะแบ่งออกเป็น 2 ส่วน คือ 1. การทดสอบอุปกรณ์ตรวจสอบที่จอดรถผ่านโมเดลลานจอดรถพร้อมกับเซ็นเซอร์อินฟราเรด (Infrared Sensor) ที่ใช้ในการตรวจจับรถยนต์ในช่องจอดรถยนต์ 2. ทดสอบการทำงานของแอปพลิเคชัน

ผลการทดสอบการรับส่งข้อมูลระหว่างฐานข้อมูลและอุปกรณ์โดยให้เซ็นเซอร์ทำการตรวจจับ กรณีที่พบว่ามียอดรถจอดอยู่ให้ทำการส่งข้อมูลเป็น 1 กรณีที่ไม่พบว่ามีรถยนต์จอดอยู่ให้ทำการส่งข้อมูลเป็น 0

4. ผลการดำเนินงาน

จากการตรวจสอบพื้นที่จอดรถผ่านแอปพลิเคชันไลน์พบว่า ผลการทดสอบในการใช้งานอุปกรณ์ตรวจสอบพื้นที่จอดรถและการทำงานของแอปพลิเคชันบนโทรศัพท์มือถือพบว่าผลการทดสอบครั้งที่ 1 และ 2 แอปพลิเคชันสามารถทำงานได้เป็นปกติแต่อุปกรณ์ไม่สามารถส่งค่าข้อมูลไปยังฐานข้อมูลได้ พบมีความผิดพลาด คิดเป็นร้อยละ 20 ของการทดลอง 10 ครั้ง และผลการทดสอบครั้งที่

ที่ 3 – 10 พบมีความถูกต้อง คิดเป็นร้อยละ 80 ของการทดลอง 10 ครั้งทั้งตัวแอปพลิเคชันและอุปกรณ์ แสดงให้เห็นว่าสามารถนำไปประยุกต์ใช้กับการใช้งานจริงได้

ตารางที่ 1: การทดสอบการใช้งานแอปพลิเคชันและอุปกรณ์

ครั้งที่	การทำงานของอุปกรณ์		การทำงานของแอปพลิเคชัน	
	ปกติ	ไม่ปกติ	ปกติ	ไม่ปกติ
1		✓		✓
2		✓		✓
3	✓		✓	
4	✓		✓	
5	✓		✓	
6	✓		✓	
7	✓		✓	
8	✓		✓	
9	✓		✓	
10	✓		✓	

5. สรุปผลการวิจัย

ผลการทดสอบประสิทธิภาพของอุปกรณ์ตรวจจับรถยนต์และการแสดงผลผ่านแอปพลิเคชัน พบว่าการทำงานของอุปกรณ์และแอปพลิเคชันเป็นไปได้อย่างมีประสิทธิภาพ คิดเป็นร้อยละ 80 ของการทดสอบ 10 ครั้ง โดยวิธีการที่จะทำการตรวจสอบพื้นที่จอดรถ เมื่อมีรถยนต์เข้ามาจอดค่าสถานะจะเปลี่ยนเป็น 1 แต่ถ้าหากไม่มีรถเข้ามาจอดหรือรถที่จอดอยู่ได้ขับออกไปแล้วค่าสถานะจะกลับเป็น 0 โดยทำการทดสอบทั้งหมดทั้งสิ้นจำนวน 10 รอบ แม้จะพบว่าการทำงานในช่วงที่ 1 – 2 จะมีปัญหาทางด้านการทำงานของอุปกรณ์อยู่บ้าง แต่เมื่อแก้ไขแล้ว พบว่าอุปกรณ์สามารถทำงานได้อย่างเป็นปกติ ผลการทดสอบของอุปกรณ์ตรวจสอบพื้นที่จอดรถ การทดลองนี้เป็นการทดสอบการทำงานของเซ็นเซอร์และบอร์ดอาร์ดูโนโอเอสพี 32

เอกสารอ้างอิง

[1] กรมการขนส่งทางบก. *กลุ่มสถิติการขนส่ง กองแผนงาน กรมการขนส่งทางบก*. Available at: <https://web.dlt.go.th/statistics/> (Accessed: 23 ตุลาคม 2564).

[2] A. Joshi, A. T. Hariram, K. M. Vishall Somaiya and M. Hussain. (2020). Smart Car Parking System. *International Journal of Engineering Research & Technology (IJERT)*, 9 (9).

[3] โชคชัย แจวจารณ์. (2563). *การประยุกต์ใช้ Arduino สำหรับปฏิบัติการรายคาบ*. [วิทยานิพนธ์คณะวิทยาศาสตร์]. มหาวิทยาลัยอุบลราชธานี.

[4] Nandyal, S., Sultana, S. and Anjum, S. (2017). Smart car parking system using Arduino Uno. *International Journal of Computer Applications*, 169(1), pp. 13–18. Available at: <https://doi.org/10.5120/ijca2017914425>.

[5] Santoso, D. et al. (2022). Mobile application for Smart Parking Systems. *Algorithms for Intelligent Systems*, pp. 485–497. Available at: https://doi.org/10.1007/978-981-19-3311-0_41.

[6] H. Tanti, P. Kasodariya, S. Patel, D. H. Rangrej. Smart Parking System based on IOT. *International Journal of Engineering Research & Technology (IJERT)*, 9 (5)

[7] นันทศักดิ์ ชาวชน และคณะ. (2563). แบบจำลองระบบการจัดการลานจอดรถผ่านเว็บแอปพลิเคชัน. *การประชุมวิชาการระดับชาติมหาวิทยาลัยราชภัฏภูเก็ต ครั้งที่ 13*. ประเทศไทย.

[8] เกริญไกร สว่างวงศ์ และคณะ. (2563). การพัฒนาระบบจองที่จอดรถอัจฉริยะโดยใช้เทคโนโลยีอินเทอร์เน็ตของทุกสรรพสิ่ง. *วารสารวิชาการวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสงขลา*, 2 (1), 57-68.

[9] มโมรส บริรักษ์อรวิวิท และคณะ. (2019). อินเทอร์เน็ตทุกสรรพสิ่ง ภูมิศึกษา: ระบบลานจอดรถอัจฉริยะ. *The 18th National Conference on Computing and Information Technology*, 36-41.

[10] W. Alsafery, B. Alturki, S. Reiff-Marganec and K. Jambi, "Smart Car Parking System Solution for the Internet of Things in Smart Cities," *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, 2018, pp. 1-5, doi: 10.1109/CAIS.2018.8442004.

[11] ปานวิทย์ ชูวะนุด. (2560). *สำนักงานรัฐบาลอิเล็กทรอนิกส์ องค์การมหาชน หลักสูตรผู้บริหารรัฐบาลอิเล็กทรอนิกส์รุ่นที่ 7*. คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

[12] กัลยา ธนาสินธุ์ และคณะ. (2563). ระบบตรวจสอบที่ว่างของช่องจอดรถยนต์ในโมเดลลานจอดผ่านแอปพลิเคชันระบบแอนดรอยด์บนสมาร์ทโฟน. *Advanced Science Journal*, 20 (2), 35-50.

[13] กษิทธิร อัสวพงศ์วานิช. (2565). การพัฒนาไลน์บอทของร้านอาหาร: กรณีศึกษา ร้านกาแฟ & ธรรมชาติ รีสอร์ท @ นครพนม. *วารสารวิชาการจัดการปริทัศน์*, 24 (3), 149-158.

แอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์โดยใช้บริบทของแรงงานจังหวัดกาญจนบุรี

Application for Training Artificial Intelligence Skills in the Context of Employees in Kanchanaburi Province

สุทัศน์ กำมณี (Sutat Gammanee)¹ ณรงค์เดช รัตนานนท์เสถียร (Narongdech Rattananonsathien)²
และขวัญนรี กล้าปราบโจร (Kwannaree Klaprabchone)³

¹สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

²สาขาวิชาคอมพิวเตอร์ศึกษา คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

³สาขาวิชาบัญชี คณะวิทยาการจัดการ มหาวิทยาลัยราชภัฏกาญจนบุรี

¹sutat@kru.ac.th, ²narongdech@kru.ac.th, ³kwannaree2992@kru.ac.th

บทคัดย่อ

การวิจัยมีวัตถุประสงค์ 1) เพื่อพัฒนาแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์สำหรับการทำงาน 2) เพื่อวัดประสิทธิภาพบทเรียนแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์ 3) เพื่อวัดความพึงพอใจของผู้ใช้งานแอปพลิเคชันสำหรับฝึกทักษะปัญญาประดิษฐ์ เครื่องมือเพื่อสำรวจบริบทของปัญญาประดิษฐ์ในจังหวัดกาญจนบุรี ได้แก่ แบบสอบถาม กลุ่มตัวอย่างจำนวน 369 คน และแบบสัมภาษณ์ จำนวน 30 คน ขั้นตอนต่อไปจึงนำบริบทมาออกแบบบทเรียนสร้างแอปพลิเคชันแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์ แล้วจึงทดสอบแอปพลิเคชันกับผู้เรียนจำนวน 228 คน เพื่อประเมินประสิทธิภาพของบทเรียนและประเมินความพึงพอใจ ผลการทดสอบประสิทธิภาพบทเรียนพบว่ามีความแตกต่างของคะแนนระหว่างก่อนเรียนและหลังเรียนเท่ากับ 51.19 คะแนน ผลประเมินความพึงพอใจมีค่าเฉลี่ย 6 ด้านเท่ากับ 4.47

คำสำคัญ: แอปพลิเคชัน ปัญญาประดิษฐ์ แรงงาน

Abstract

This research aims to 1) to develop an artificial intelligence training application for work, 2) to measure the effectiveness of an artificial intelligence training application lesson, and 3) to measure app user satisfaction. The tools to explore the context of artificial intelligence in Kanchanaburi province are the

questionnaire, a sample of 369 people and an interview of 30 people. Then, apply the context to develop a lesson plan for an application that will train artificial intelligence. Afterwards, examine the application with 228 learners to assess the effectiveness of the lessons and assess their satisfaction. The results of the lesson efficiency examination present that there is a difference in scores between pretest and posttest is 51.19 points. The satisfaction assessment result has an average of 6 parts is 4.47.

Keywords: Application, Artificial Intelligence, Employee.

1. บทนำ

ในปัจจุบันปัญญาประดิษฐ์เข้ามามีบทบาทในการทำงานทั้งในภาครัฐบาลและเอกชน รวมถึงในกลุ่มอุตสาหกรรม [1], [2], [3] เพราะปัญญาประดิษฐ์สามารถทำงานแทนมนุษย์ได้หลายรูปแบบ สามารถวิเคราะห์ การให้เหตุผล เพื่อนำไปสู่การหาคำตอบในปัญหาที่ซับซ้อนเกินกว่าที่มนุษย์จะสามารถคิดได้ นอกจากนี้แล้วยังสามารถทำงานต่อเนื่องในรูปแบบเดิมและได้ผลลัพธ์ในมาตรฐานเดิมทุกครั้ง ทำงานได้ตลอดเวลา ส่งผลให้สามารถลดต้นทุน ลดเวลา และเพิ่มกำไรได้

บริษัทภาคเอกชนหรือภาคอุตสาหกรรมในปัจจุบันที่นำปัญญาประดิษฐ์มาประยุกต์ใช้ในงานภายใน

จะทำให้เกิดรายได้เพิ่มมากขึ้นอย่างมีนัยสำคัญ ด้วยเหตุนี้ทำให้บริษัทมีความเชื่อถือในการใช้ปัญญาประดิษฐ์เพื่อช่วยเหลือในงานของตน ซึ่งนำไปสู่การจ้างแรงงานที่มีความรู้ในทักษะปัญญาประดิษฐ์เพิ่มมากขึ้น หากแรงงานมีความรู้และสามารถประยุกต์ใช้ปัญญาประดิษฐ์ได้อย่างมีประสิทธิภาพก็จะทำให้ผลประโยชน์ของบริษัทติดตามไปเช่นกัน

แผนปฏิบัติการด้านการพัฒนาอุตสาหกรรมจังหวัดกาญจนบุรี ระยะเวลา 5 ปีตั้งแต่ปี พ.ศ. 2566-2570 กำหนดแนวทางการพัฒนาในการนำนวัตกรรมมาประยุกต์ใช้ในงาน สนับสนุนให้มีการปรับตัวสู่ธุรกิจรูปแบบใหม่ พัฒนาผู้ประกอบการให้มีทักษะวิเคราะห์และมององค์ความรู้ด้านการจัดการฐานข้อมูล [4] แสดงให้เห็นถึงความสำคัญของการเติบโตและการปรับตัวภาคอุตสาหกรรมในการทำให้ผู้ที่เกี่ยวข้องจำเป็นต้องมีทักษะความเข้าใจในเรื่องของปัญญาประดิษฐ์

จากความสำคัญที่กล่าวมานักวิจัยจึงมีแนวคิดในการสร้างและพัฒนาแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์เพื่อสร้างบุคลากรที่มีความรู้ด้านปัญญาประดิษฐ์ให้เพิ่มมากขึ้นทันต่อความต้องการตลาดแรงงานที่มีแนวโน้มการใช้ปัญญาประดิษฐ์เพิ่มขึ้นในอนาคต โดยศึกษาจากบริบทแรงงานในจังหวัดกาญจนบุรีเป็นพื้นที่ในการศึกษาและนำมาออกแบบบทเรียน เพื่อให้แรงงานเมื่อเรียนแล้วจะมีทักษะด้านปัญญาประดิษฐ์เพิ่มมากขึ้นจนนำไปสู่การประยุกต์ใช้ในการทำงานเพื่อผลประโยชน์ต่อผู้ประกอบการและองค์กร

2. วัตถุประสงค์

2.1 เพื่อพัฒนาแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์สำหรับการทำงาน

2.2 เพื่อวัดประสิทธิภาพบทเรียนแอปพลิเคชันการฝึกทักษะปัญญาประดิษฐ์

2.3 เพื่อวัดความพึงพอใจของผู้ใช้งานแอปพลิเคชันสำหรับฝึกทักษะปัญญาประดิษฐ์

3. แนวคิดและทฤษฎีที่เกี่ยวข้อง

การวิจัยครั้งนี้เป็นการพัฒนาแอปพลิเคชันเพื่อฝึกทักษะด้านปัญญาประดิษฐ์ซึ่งองค์ความรู้ด้านปัญญาประดิษฐ์มีวิวัฒนาการตั้งแต่ในอดีตต่อเนื่องมาถึงยุคปัจจุบัน

ในยุคแรกของการกำเนิดเครื่องคอมพิวเตอร์ก็เพื่อคำนวณและประมวลผลในระบบงานที่มนุษย์ต้องการโดยมนุษย์จะเป็นผู้เขียนโปรแกรมคำสั่ง คอมพิวเตอร์มีหน้าที่เพียงประมวลผลตามคำสั่งที่กำหนดได้จากข้อมูลที่มนุษย์ป้อนเข้า คอมพิวเตอร์ยังไม่สามารถตัดสินใจหรือเสนอแนวทางให้กับมนุษย์ได้ เมื่อคอมพิวเตอร์มีการพัฒนาด้านประสิทธิภาพและความเร็วมากขึ้น มนุษย์จึงพัฒนาให้คอมพิวเตอร์สามารถวิเคราะห์และให้เหตุผลเลียนแบบมนุษย์จนสามารถเสนอแนวทางการตัดสินใจให้กับมนุษย์บนพื้นฐานข้อมูลที่เรียนรู้ กระบวนการให้เหตุผลและการคิดดังกล่าวสามารถทดแทนหรือบางอย่างอาจจะดีกว่าการคิดโดยมนุษย์ซึ่งเรียกกว่า “ปัญญาประดิษฐ์”

ปัญญาประดิษฐ์แบ่งออกเป็นประเภทดังต่อไปนี้

1. ระบบผู้เชี่ยวชาญ (Expert System) [5] เป็นระบบที่นิยมใช้กันในภาคเอกชนที่มีการแข่งขันด้านธุรกิจสูงที่จำเป็นต้องมีการตัดสินใจในสถานการณ์รูปแบบต่าง ๆ จึงต้องอาศัยความสามารถในการแก้ปัญหาที่มีโครงสร้างซับซ้อนเพื่อเสนอแนวทางการแก้ปัญหา รวมถึงการทำนายเหตุการณ์ล่วงหน้าเพื่อความสะดวกเปรียบทางธุรกิจ

2. การให้เหตุผลเชิงกรณีศึกษา (Case-based Reasoning) [6] เป็นการให้เหตุผลโดยการนำการเทียบเคียงการให้เหตุผลของมนุษย์เพื่อเป็นแนวทางในการพัฒนา โดยจะต้องมีข้อมูลการให้เหตุผลเดิมเพื่อใช้ในการสร้างแบบจำลอง ก่อนนำแบบจำลองไปเทียบเคียงเพื่อแก้ปัญหาใหม่ ปัญญาประดิษฐ์ประเภทนี้มีใช้ในงานอุตสาหกรรมสำหรับการลดข้อผิดพลาดจากกระบวนการทำงาน

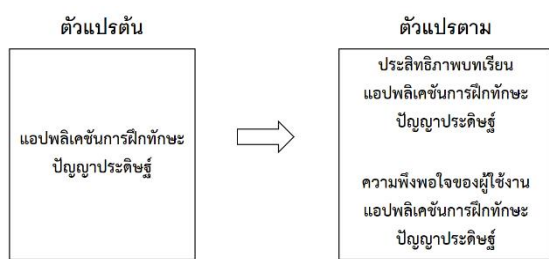
3. การเรียนรู้เครื่องจักร (Machine Learning) [7] เป็นการนำมาประยุกต์ใช้ด้านการเรียนรู้ โดยมีลักษณะการเรียนรู้จากสถานการณ์จากองค์ความรู้เดิมที่เกิดขึ้นหรือสิ่งที่ไม่เคยเกิดขึ้นจนสามารถนำมาใช้ในการพยากรณ์สิ่งที่

จะเกิดขึ้นล่วงหน้า และสามารถจัดการปัญหาทั้งที่มีรูปแบบและไม่มีรูปแบบได้

4. ระบบโครงข่ายประสาทเทียม (Neural Network) [8] เป็นแนวคิดในการจำลองโครงข่ายประสาทของมนุษย์ มาใช้ในการประมวลผล ซึ่งสามารถทำงานได้อย่างรวดเร็ว มีความสามารถในการคิดวิเคราะห์ปัญหาที่ซับซ้อน ระบบโครงข่ายประสาทประกอบด้วยโหนดและเชื่อมโยงกัน เป็นเครือข่าย

แนวคิดเรื่องปัญญาประดิษฐ์ยังมีการพัฒนาและต่อยอด แลกเปลี่ยนความรู้อย่างต่อเนื่อง ทำให้ปัจจุบันยังไม่มีกรอบทฤษฎีหรือกระบวนการที่แน่ชัดที่เป็นแนวทางที่ชัดเจน มีบางประเด็นที่เกิดขึ้นและนักวิจัยยังไม่สามารถหาคำตอบได้ เช่น จริยธรรมของปัญญาประดิษฐ์ การทำงานได้ทดแทนคน การมีพฤติกรรมเลียนแบบมนุษย์ หรือการสร้างความปลอดภัยขึ้นมาด้วยตนเอง แม้หลายประเด็นจะถูกตั้งขึ้นและอาจยังไม่ได้รับคำตอบในเร็ววันนี้ แต่การพัฒนาด้านปัญญาประดิษฐ์เพื่อใช้งานก็ยังมีพัฒนาต่อไป เราจะได้พบความรู้ในปัญญาประดิษฐ์ที่ฉลาดมากขึ้นต่อไปในอนาคต สามารถทำงานได้หลากหลายจนปัญญาประดิษฐ์เข้ามาอยู่ในทุกสิ่งทุกอย่างรอบตัวเรา

4. วิธีการดำเนินการวิจัย



ภาพที่ 1: กรอบแนวคิดการวิจัย

การสร้างแอปพลิเคชันปัญญาประดิษฐ์สำหรับ “การทำงาน มีขั้นตอนดังนี้ 1.การวิเคราะห์ความต้องการด้านทักษะการอบรม ในทักษะปัญญาประดิษฐ์ 2. การออกแบบบทเรียน โดยนำผลลัพธ์ที่ได้จากการวิเคราะห์ 3. เขียนผังงาน 3. สร้างและพัฒนาแอปพลิเคชัน 4. การทดสอบโดยการนำแอปพลิเคชันไปทดลองใช้กับกลุ่ม

ตัวอย่าง จำนวน 228 คน 5. การประเมินประสิทธิภาพและความพึงพอใจ

การวิจัยจะใช้ผลการสำรวจในเชิงปริมาณในส่วนของรูปแบบบทเรียน และใช้การสัมภาษณ์เชิงคุณภาพเพื่อให้ได้มาซึ่งความต้องการใน โครงสร้างและรูปแบบของแอปพลิเคชัน

4.1 ประชากรและกลุ่มตัวอย่าง

4.1.1 ประชากรสำหรับสำรวจความต้องการทักษะการอบรมในทักษะปัญญาประดิษฐ์ ได้แก่ นักศึกษาผู้ประกอบการ แรงงาน ในตลาดแรงงานในจังหวัดกาญจนบุรี โดยเลือกกลุ่มตัวอย่างมาจำนวน 369 ราย

4.1.2 กลุ่มตัวอย่างในการวิจัยเพื่อเก็บรวบรวมข้อมูลเชิงคุณภาพโดยใช้เทคนิค การสัมภาษณ์และการประชุมเชิงปฏิบัติการ ผู้วิจัยใช้การสุ่มตัวอย่างแบบเจาะจง (Purposive Simple) จำนวน 30 ราย

4.2 เครื่องมือในการวิจัย

4.2.1 แบบสอบถามเพื่อเก็บรวบรวมข้อมูลเกี่ยวกับบริบทและระดับความคิดเห็นเกี่ยวกับทักษะและความสามารถในการใช้ปัญญาประดิษฐ์ [9] ซึ่งได้ข้อมาเฉพาะผลที่เกี่ยวข้องกับการพัฒนาแอปพลิเคชันดังต่อไปนี้

ตารางที่ 1: รูปแบบปัญญาประดิษฐ์ในการทำงาน

รูปแบบการใช้ปัญญาประดิษฐ์ในการทำงาน	ร้อยละ
Cognitive Science	30.49
Robotics	12.56
Hybrid AI Systems	23.17
Natural Interface	20.18
ไม่มีการใช้ปัญญาประดิษฐ์ในการทำงาน	13.60
รวม	100.00

จากตารางที่ 1 แสดงบริบทรูปแบบการใช้ปัญญาประดิษฐ์ในจังหวัดกาญจนบุรี พบว่าด้าน Cognitive Science สูงที่สุด ร้อยละ 30.49 รองลงมาได้แก่ ด้าน Hybrid AI Systems

4.2.2 แบบสัมภาษณ์เพื่อเก็บรวบรวมข้อมูลเกี่ยวกับแนวทางของทางการเรียนรู้ปัญญาประดิษฐ์สำหรับการเพิ่มขีด

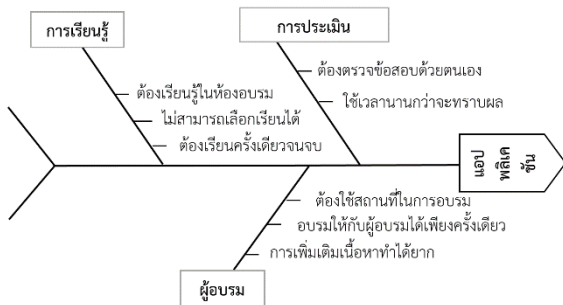
ความสามารถของงานและความสามารถในการทำงานในจังหวัดกาญจนบุรี ซึ่งได้ยกมาเฉพาะผล [9] ที่เกี่ยวข้องกับพัฒนาแอปพลิเคชันดังต่อไปนี้

“การใช้ข้อมูลปัญญาประดิษฐ์ในองค์กรส่วนใหญ่ใช้ในการวิเคราะห์ข้อมูล เช่น การวิเคราะห์ข้อมูลลูกค้า วิเคราะห์ข้อมูลความต้องการในองค์กร”

“การใช้ปัญญาประดิษฐ์สำหรับการบริหาร ปัจจุบันยังไม่มีปัญหาในการใช้ปัญญาประดิษฐ์เข้ามาใช้ในการบริหารงานหรืองานด้านอื่น ในงานเกี่ยวกับการวิเคราะห์ข้อมูลหรืองานที่ต้องตัดสินใจก็ใช้ปัญญาประดิษฐ์ในการทำข้อมูลที่เป็นแก่นมนุษย์ก็ต้องเป็นผู้ที่ตัดสินใจจากข้อมูลเหล่านั้นเอง ไม่ได้อนุญาตให้ปัญญาประดิษฐ์มีอำนาจในการตัดสินใจเอง ในส่วนของงานอื่นมีปัญหาบ้าง”

4.2.3 แบบสอบถามความพึงพอใจในการใช้ แอปพลิเคชันปัญญาประดิษฐ์สำหรับการทำงาน

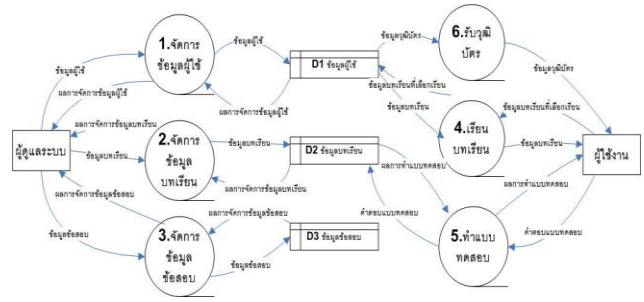
4.2.4 ผังก้างปลา เพื่อแสดงความสัมพันธ์ของปัญหาในการเรียนรู้และด้านการพัฒนาทักษะปัญญาประดิษฐ์



ภาพที่ 2: แพนผังก้างปลา

จากแผนผังก้างปลาแสดงถึงปัญหาในการพัฒนาทักษะด้านปัญญาประดิษฐ์ ประเด็นหลักคือไม่สะดวกในการเรียนรู้ในห้องอบรม ไม่สามารถใช้เวลาเรียนครั้งเดียวจบ และต้องการที่จะเรียนรู้ได้หลายครั้งเพราะเป็นหัวข้อที่เข้าใจได้ยาก

4.2.5 แผนภาพกระแสข้อมูล เพื่ออธิบายการเดินทางขอข้อมูลในการออกแบบแอปพลิเคชันปัญญาประดิษฐ์สำหรับการทำงาน มีผลการออกแบบดังภาพที่ 3



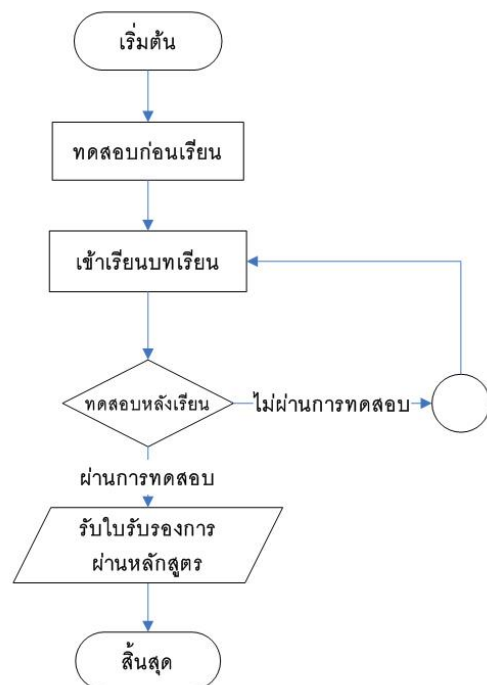
ภาพที่ 3: แพนภาพกระแสข้อมูล

จากภาพที่ 3 แสดงแผนภาพกระแสข้อมูลโดยมีกระบวนการทำงานจำนวน 6 กระบวนการ มีผู้ใช้งานจำนวน 2 กลุ่ม ได้แก่ ผู้ดูแลระบบ ผู้ใช้งาน และฐานข้อมูลเพื่อดำเนินการเก็บข้อมูลจำนวน 3 ฐานข้อมูล

5. ผลการวิจัย

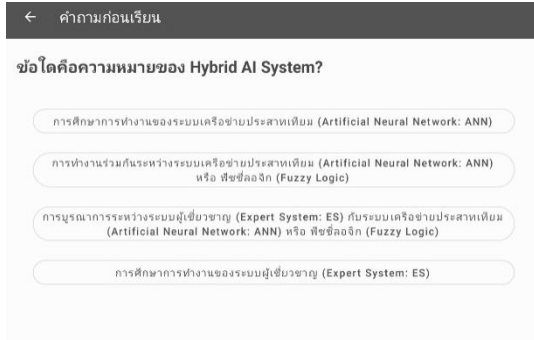
5.1 ขั้นตอนการทำงานแอปพลิเคชัน

จากวิธีการวิจัยได้นำมาสรุปและออกแบบเป็นขั้นตอนแอปพลิเคชันปัญญาประดิษฐ์สำหรับการทำงาน โดยมีขั้นตอนดังภาพที่ 4



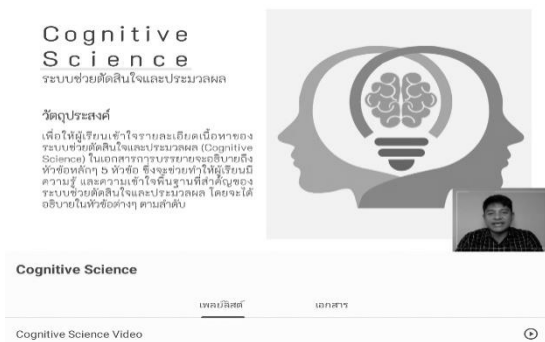
ภาพที่ 4: ผังงานลำดับการทำงานแอปพลิเคชัน

จากภาพที่ 4 แสดงผังงานเริ่มจากทำแบบทดสอบก่อนเรียนแล้วจึงเข้าเรียนบทเรียน เมื่อเรียนเสร็จจึงทำแบบทดสอบหลังเรียนต้องผ่านการทดสอบทุกบทเรียนจึงจะได้รับใบรับรองการผ่านหลักสูตร



ภาพที่ 5: การทดสอบก่อนเรียน

จากภาพที่ 5 เป็นการทำงานของแอปพลิเคชันในการทดสอบก่อนเรียน แล้วจึงเข้าสู่การเรียนจากบทเรียนดังภาพที่ 6



ภาพที่ 6: การเรียนในบทเรียน

เมื่อเรียนเสร็จแล้วจะสามารถทำแบบทดสอบหลังเรียน เมื่อทำแบบทดสอบหลังเรียนผ่านทุกบทแล้วจะได้รับใบรับรองการผ่านหลักสูตร ดังภาพที่ 7



ภาพที่ 7: การรับใบรับรองการผ่านหลักสูตร

5.2 การออกแบบบทเรียน

บทเรียนในแอปพลิเคชันปัญญาประดิษฐ์สำหรับการทำางานผู้วิจัยได้ออกแบบตามผลจากแบบสอบถามและแบบสัมภาษณ์มีจำนวน 8 บท ดังนี้

บทที่ 1 ปัญญาประดิษฐ์พื้นฐาน

บทที่ 2 Cognitive Science

บทที่ 3 Hybrid AI System

บทที่ 4 การประมวลผลภาษาธรรมชาติ

บทที่ 5 ปัญญาประดิษฐ์สำหรับงานอาชีพ

บทที่ 6 เครื่องมือการวิเคราะห์ปัญญาประดิษฐ์

(WEKA)

บทที่ 7 เครื่องมือการวิเคราะห์ปัญญาประดิษฐ์

(RAPID MINER)

บทที่ 8 กฎหมายดิจิทัลกับปัญญาประดิษฐ์

5.3 ผลทดสอบประสิทธิภาพบทเรียน

การทดสอบประสิทธิภาพบทเรียนจะทดสอบกับกลุ่มตัวอย่าง จำนวน 228 คน โดยการเปรียบเทียบคะแนนจากแบบทดสอบก่อนเรียนและแบบทดสอบหลังเรียน ได้ผลดังตารางที่ 2

ตารางที่ 2: ประสิทธิภาพของบทเรียน

เนื้อหา	เฉลี่ยคะแนน	เฉลี่ยคะแนน	ผลต่างคะแนน
	ก่อนเรียน (100 คะแนน)	หลังเรียน (100 คะแนน)	
ปัญญาประดิษฐ์พื้นฐาน	33.39	89.82	56.43
Cognitive Science	39.12	90.79	51.67
Hybrid AI System	39.39	89.74	50.35
การประมวลผลภาษาธรรมชาติ	41.23	90.18	48.95
ปัญญาประดิษฐ์สำหรับงานอาชีพ	39.21	90.53	51.32
เครื่องมือ การวิเคราะห์ปัญญาประดิษฐ์ (WEKA)	39.82	90.26	50.44
เครื่องมือ การวิเคราะห์ปัญญาประดิษฐ์ (RAPID MINER)	40.79	90.44	49.65
กฎหมายดิจิทัลกับปัญญาประดิษฐ์	39.04	89.74	50.7
ค่าเฉลี่ย	39.00	90.19	51.19

จากตารางที่ 2 พบว่าทุกบทเรียนมีค่าเฉลี่ยคะแนนหลังเรียนเปรียบเทียบกับก่อนเรียนดีขึ้นทุกบทเรียน และผลต่างค่าเฉลี่ยรวมของทุกบทเรียนระหว่างก่อนเรียนและหลังเรียนมีค่าเท่ากับ 51.19

หลังจากนั้นจึงทดสอบว่าบทเรียนมีประสิทธิภาพทำให้มีผลสัมฤทธิ์ทางการเรียนหลังเรียนสูงกว่าก่อนเรียนหรือไม่ที่ระดับนัยสำคัญ .05

ตั้งสมมติฐานทางสถิติ

$$H_0 : \mu_{post} = \mu_{pre}$$

$$H_1 : \mu_{post} > \mu_{pre}$$

กำหนดระดับนัยสำคัญ $\alpha = .05$

ค่าวิกฤตของ t ที่ $df = 7$ และ $\alpha = .05$ จะได้ค่าวิกฤตของ t = 1.89

เนื่องจาก t คำนวณ = 63.32 > t จากตาราง เท่ากับ 1.89 ดังนั้น จึงปฏิเสธ H_0 ยอมรับ H_1 นั่นคือหลังเรียนมีผลสัมฤทธิ์สูงกว่าก่อนเรียนได้อย่างมีนัยสำคัญ ที่ระดับ .05

การหาความเชื่อมั่นของแบบทดสอบใช้การวัดความมั่นคงในการตอบแบบสอบถาม โดยใช้สมการดังนี้

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}} \quad (1)$$

โดย r คือ ความเชื่อมั่นของแบบทดสอบ

X คือ คะแนนจากการสอบครั้งแรก

Y คือ คะแนนจากการสอบครั้งหลัง

N คือ จำนวนผู้สอบเท่ากับ 228 คน

ซึ่งได้ค่าดังตารางที่ 3

ตารางที่ 3: ตารางแสดงค่าการสอบเพื่อหาความเชื่อมั่น

$\sum X$	$\sum X^2$	$\sum Y$	$\sum Y^2$	$\sum XY$
581	43981	721.5	65071.3662	52416.34

เมื่อคำนวณแล้วได้ค่าความเชื่อมั่นเท่ากับ 0.95 ถือว่ามีความเชื่อมั่นที่สูงมาก แสดงว่าแบบทดสอบนี้เชื่อถือได้

5.4 ผลการประเมินความพึงพอใจ

ผลการประเมินความพึงพอใจแอปพลิเคชันปัญญาประดิษฐ์สำหรับการทำงานมีผลดังตารางที่ 3

ตารางที่ 4: ผลประเมินความพึงพอใจ

หัวข้อการประเมิน	ค่าเฉลี่ย	ค่าเฉลี่ยความแปรปรวน	ระดับ
ด้านการใช้งานระบบ	4.49	0.64	ดี
ด้านความถูกต้องการทำงานระบบ	4.50	0.59	ดีมาก
ด้านความง่ายต่อการใช้งาน	4.50	0.58	ดีมาก
ด้านความเหมาะสมของหลักสูตร	4.43	0.60	ดี
ด้านผู้บรรยาย	4.50	0.58	ดีมาก
ด้านสื่อการสอน	4.41	0.61	ดี
ค่าเฉลี่ย	4.47	0.60	ดี

จากการประเมินความพึงพอใจพบว่าด้านที่มีระดับความพึงพอใจสูงที่สุด ได้แก่ ด้านความถูกต้องการทำงาน ของระบบ ด้านความง่ายต่อการใช้งาน ด้านผู้บรรยาย เท่ากันที่ระดับดีมาก คะแนน 4.50 และด้านน้อยที่สุด ได้แก่ ด้านสื่อการสอน 4.41

6. สรุปและอภิปรายผล

จากผลการวิจัยในรูปแบบสอบถามและแบบสัมภาษณ์เพื่อให้เข้าใจบริบทของการใช้ปัญญาประดิษฐ์ในแรงงานในจังหวัดกาญจนบุรี พบว่า รูปแบบปัญญาประดิษฐ์ที่มีการใช้มากที่สุด ได้แก่ รูปแบบ Cognitive Science และมีความต้องการให้หลักสูตรเป็นลักษณะออนไลน์จึงออกแบบเป็นลักษณะแอปพลิเคชันเพื่อสะดวกในการเรียนที่ไม่ขึ้นอยู่กับ สถานที่ เวลา และสามารถเรียนซ้ำจนกว่าจะเข้าใจได้ แอปพลิเคชันประกอบด้วยบทเรียนจำนวน 8 บทเรียน

จากการประเมินประสิทธิภาพและการประเมินความพึงพอใจ พบว่ามีค่าเฉลี่ยทุกบทเรียนดีขึ้น 51.19 คะแนน จาก 100 คะแนน ผลประเมินความพึงพอใจพบว่ามีความพึงพอใจในภาพรวมระดับดีได้คะแนนทุกด้าน 4.47 ซึ่งผู้ใช้ได้ให้ข้อเสนอแนะว่าระดับความยากง่ายแต่ละหัวข้อแตกต่างกัน และอยากพัฒนาลงบนระบบปฏิบัติการ ไอโอเอส แอปพลิเคชันมีคุณภาพดี และเลือกได้หลายภาษา เหมาะสมสำหรับนำไปใช้ในการทำงาน

จากการสัมภาษณ์กลุ่มตัวอย่างพบว่ามี การประยุกต์ใช้ปัญญาประดิษฐ์ในปริมาณน้อยแต่มีความเห็น ว่าปัญญาประดิษฐ์จะมีความสำคัญมากขึ้นในอนาคต การเรียนรู้มีความจำเป็นต่อแรงงานในระดับการประยุกต์ใช้ งาน การพัฒนาปัญญาประดิษฐ์จะยังคงอยู่ในกลุ่มของ ผู้เชี่ยวชาญด้านคอมพิวเตอร์ บริบทการเรียนรู้เหมาะสมที่จะ นำไปออกแบบหลักสูตรระยะสั้นสำหรับคนที่สนใจ

จากผลการวิจัยพบว่าประสิทธิภาพพบที่เรียนที่ ออกแบบมีประสิทธิภาพและมีความน่าสนใจตรงตาม ความต้องการตามบริบทการพัฒนาทักษะด้าน ปัญญาประดิษฐ์ในจังหวัดกาญจนบุรีเนื่องจากมีการ ออกแบบบทเรียนที่มีการวิเคราะห์จากบริบทการทำงานใน ปัจจุบัน ซึ่งจากผลการเรียนสะท้อนให้เห็นว่าแรงงานมี ความตระหนักและเข้าใจถึงความสำคัญถึงสิ่งที่จะต้อง ปรับตัวในการทำงานกับปัญญาประดิษฐ์ในอนาคตเป็น ตามผลการวิจัย [10][11] ที่กล่าวถึง ในอนาคต ปัญญาประดิษฐ์จะมีส่วนในการออกแบบความคิดต่อการ พัฒนาทรัพยากรบุคคล หากจุดสมดุลเพื่ออยู่ร่วมกันในการ ทำงานระหว่างคนและปัญญาประดิษฐ์

เอกสารอ้างอิง

- [1] B. Chaib-draa, "Industrial applications of distributed AI," *Commun. ACM*, vol. 38, no. 11, pp. 49–53, Nov. 1995, doi: 10.1145/219717.219761.
- [2] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook," *IEEE Access*, vol. 8, pp. 220121–220139, 2020, doi: 10.1109/ACCESS.2020.3042874.
- [3] S. Wu and T. W. S. Chow, "Induction machine fault detection using SOM-based RBF neural networks" *IEEE Trans. on Industrial Electronics*, vol. 51, no. 1, pp. 183-194, 2004.
- [4] สำนักงานอุตสาหกรรมจังหวัดกาญจนบุรี, (2565, กุมภาพันธ์ 10), "แผนปฏิบัติการด้านการพัฒนาอุตสาหกรรม (พ.ศ.2566-2570), (ออนไลน์), จาก : <https://kanchanaburi.industry.go.th/>
- [5] B. G. Buchanan and R. G. Smith, "Fundamentals of Expert Systems," *Annual Review of Computer Science*, vol. 3, no. 1, pp. 23–58, 1988, doi: 10.1146/annurev.cs.03.060188.000323.
- [6] J. L. Kolodner, "An introduction to case-based reasoning," *Artif Intell Rev*, vol. 6, no. 1, pp. 3–34, Mar. 1992, doi: 10.1007/BF00155578.
- [7] G. Carleo et al., "Machine learning and the physical sciences," *Rev. Mod. Phys.*, vol. 91, no. 4, p. 045002, Dec. 2019, doi: 10.1103/RevModPhys.91.045002.
- [8] G. Carleo et al., "Machine learning and the physical sciences," *Rev. Mod. Phys.*, vol. 91, no. 4, p. 045002, Dec. 2019, doi: 10.1103/RevModPhys.91.045002.
- [9] สุทัศน์ ก้ามณี. (2565), การพัฒนาทักษะด้านการใช้งาน ปัญญาประดิษฐ์เพื่อเพิ่มมูลค่าให้กับงานและความสามารถในการทำงาน, *กาญจนบุรี:มหาวิทยาลัยราชภัฏกาญจนบุรี*
- [10] วัชรพล ขวณิชชย์. "มนุษย์ ปัญญาประดิษฐ์และองค์กร: วัฒนธรรมและสมดุลของการอยู่ร่วมกัน". *วารสารวิชาการมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏธนบุรี*. ฉบับที่ 3 ปีที่ 5 หน้า 134-146, 2565
- [11] ทศพล มะหะหมัด และมนัส สุทธิการ. "ปัญญาประดิษฐ์ การจัดการทรัพยากรบุคคลสมัยใหม่ในองค์กร". *RMUTT Global Business and Economics Review*. ฉบับที่ 1 ปีที่ 15 หน้า 75-89, 2563

การออกแบบประสบการณ์ผู้ใช้แอปพลิเคชันสมาร์ททีวีสำหรับผู้สูงอายุในประเทศไทย

User Experience Design of Smart TV Applications for the Elderly in Thailand

สิทธิโชค สาคร (Sittichoke Sakon)¹ และวารพร จิระพัชรูทอง (Waraporn Jirapanthong)²

สาขาวิชาวิศวกรรมเว็บและการพัฒนาแอปพลิเคชันบนอุปกรณ์พกพา

หลักสูตรวิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยธุรกิจบัณฑิตย์

¹sittichokesakon@gmail.com, ²waraporn.jir@dpu.ac.th

บทคัดย่อ

จุดประสงค์ของการศึกษานี้คือเพื่อศึกษาแอปพลิเคชันประเภทสมาร์ททีวี ให้เหมาะกับกลุ่มผู้ใช้สูงอายุในประเทศไทย ได้ทำการศึกษาพฤติกรรมการใช้ในปัจจุบัน สภาพปัญหาในการใช้งานแอปพลิเคชันสมาร์ททีวีของกลุ่มผู้ใช้สูงอายุ ตลอดจนศึกษาถึงแรงจูงใจที่มีผลต่อมุมมองการใช้งานแอปพลิเคชันสมาร์ททีวีของผู้สูงอายุ โดยใช้วิธีการเรียนเชิญผู้เข้ารับการทดสอบซึ่งเป็นผู้สูงอายุที่อยู่ในอายุระหว่าง 50 - 70 ปี จำนวน 50 คน ซึ่งมีที่พักอาศัยในกรุงเทพมหานครด้วยวิธีการเก็บข้อมูล โดยการสร้างแบบสอบถามเพื่อขอทราบข้อมูลของผู้เข้ารับการทดสอบทั้งนี้การเก็บข้อมูลผู้เข้ารับการทดสอบภายใต้กรอบสมมติฐานแบบจำลองการยอมรับเทคโนโลยี (TAM) ผลการวิจัยพบว่าผู้เข้าทดสอบทั้งหมดมีประสบการณ์ในการใช้งานแอปพลิเคชันประเภทสื่อสังคมในชีวิตประจำวัน แต่เมื่อต้องใช้แอปพลิเคชันสมาร์ททีวี ข้อมูลพบว่าผู้ใช้สูงอายุยังจำเป็นต้องมีผู้ใกล้ชิดช่วยสอนการใช้งาน โดยปัญหาส่วนใหญ่จะเป็นรูปแบบปัญหาการใช้งานด้านฟังก์ชันที่ยากเกินไปในการใช้งาน อาทิ เมนูการแบ่งโหมดหมู่รายการตลอดจนการใช้งานด้านเมนูคำสั่ง และยังพบข้อเสนอและเกี่ยวกับความต้องการด้านปฏิสัมพันธ์ในกลุ่มผู้เข้ารับการทดสอบ อาทิ ความต้องการอยากให้ แอปพลิเคชันสมาร์ททีวีเป็นสื่อสังคมยุคใหม่ที่ช่วยประสานผู้ใช้เข้าหากันง่ายขึ้นในรูปแบบกลุ่มสังคมระหว่างผู้ใช้โดยไม่ต้องพึ่งแพลตฟอร์ม ตลอดจนต้องการให้แอปพลิเคชันเข้ามาเป็นส่วนหนึ่งในการอำนวยความสะดวกในการใช้ชีวิตและส่งเสริมกิจกรรมภายในที่พักอาศัย (TAM)

คำสำคัญ: สมาร์ททีวี สตรีมมิง

Abstract

The purpose of this study is to study smart TV applications. To suit the elderly users in Thailand by studying current usage behavior problems in using smart TV applications of elderly users. Also, study the influence that affects usage attitudes smart TV apps of elderly. The method of study is to invite 50 elderly test takers between the ages of 50 - 70 who have residences in Bangkok by collecting data. By creating a questionnaire to request information on all these tests and collecting data for internal test-takers under the framework of the technology acceptance model (TAM). The results show that all test testers have experience in using social media-type apps on a daily basis. But when using the Business Smart TV app still needs a close person to help teach usage Most problems are forms of function usage problems that are too difficult. Artem menu, group mode division, and command menu usage. And also found proposals about the need for relativity among the Ati Testers Needs The Privatization TV app is modern social media that helps to communicate with users more easily in the form of social groups between users without having to rely on the platform. As well as wanting the application to be a part of facilitating life and promoting residential activities.

Keywords: Smart TV, Streaming

1. บทนำ

ข้อมูลของสหประชาชาติ ซึ่งให้เห็นเป็นครั้งแรกว่าโลกมีผู้สูงอายุมากกว่าคนในวัยเด็ก ซึ่งปลายปี 2018 ตัวเลขคาดการณ์ว่า ตัวเลขผู้สูงอายุจะมากขึ้นในปี 2022 ราว 705 ล้านคน ซึ่งสูงกว่าเด็กอายุ 5 ขวบรวม 680 ล้านคน และภายในปี 2050 สัดส่วนคนอายุ 65 ปี ต่อสัดส่วนคนอายุ 0-4 ปี เทียบเท่ากับ 2 ต่อ 1 ของประชากร [1] และมีแนวโน้มว่าผู้ที่กำลังเข้าสู่วัยสูงอายุ มีอัตราการใช้เทคโนโลยีเพิ่มมากขึ้นอย่างต่อเนื่องเช่นกัน เนื่องจากมีปัจจัยที่ส่งเสริมทั้งระดับการศึกษาและมีฐานะทางเศรษฐกิจที่ดีขึ้น อีกทั้งยังดำรงชีวิตอยู่ในสภาพแวดล้อมที่อุดมไปด้วยเทคโนโลยีและการเปลี่ยนแปลงไปของโลกในยุคการพัฒนาทางเทคโนโลยีที่ไม่หยุดนิ่ง ด้วยการขับเคลื่อนของเทคโนโลยีได้เน้นมนุษย์เป็นศูนย์กลาง เพื่อให้สอดคล้องกับวิถีชีวิตของมนุษย์ให้มากยิ่งขึ้น ด้วยเหตุนี้ส่งผลให้ผู้สูงอายุมีโอกาสและทางเลือกสามารถเข้าถึงเทคโนโลยีได้ง่ายขึ้นจากอดีตที่ผ่านมา [2] ซึ่งพิจารณาจากการเพิ่มปริมาณของจำนวนประชากรผู้สูงอายุและผู้ที่กำลังเข้าสู่วัยสูงอายุ ทำให้เห็นความสำคัญของถึงเทคโนโลยีผู้สูงอายุที่กำลังมีบทบาทเพิ่มขึ้นในอนาคต

โดยเฉพาะอย่างยิ่ง เทคโนโลยีโทรทัศน์เป็นหนึ่งในเครื่องมือหลักสำหรับความบันเทิงภายในบ้าน โดยอยู่คู่สังคมของมนุษย์มากกว่า 50 ปี มีอิทธิพลอย่างมากต่อสภาพแวดล้อมภายในบ้านตลอดจนถึงสถานประกอบการต่างๆ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 สมาร์ททีวี

โทรทัศน์มีการเปลี่ยนแปลงอย่างมีนัยสำคัญ โดยเฉพาะรูปแบบรวมทั้งการโต้ตอบกับผู้ใช้หรือสื่อสารผ่านอุปกรณ์ เช่น เครื่องมือรับสัญญาณเคเบิล หรือดาวเทียม อีกทั้งยังมีตัวควบคุมรีโมทคอนโทรลชนิดต่าง ๆ ระบบรู้จำท่าทางที่ถูกติดตั้งและรวมเข้ากับโทรทัศน์ ยิ่งไปกว่านั้น การสืบทอดของนวัตกรรมทางเทคโนโลยีบางอย่างทำให้โทรทัศน์มีรูปลักษณ์ใหม่และความสามารถใหม่ เช่น ภาพสี่เหลี่ยมหน้าจอเลเซอร์ ความหนาของเคสน้อยลง การเชื่อมต่อบรอดแบนด์ LAN หรือ Wi-Fi การแชร์สื่อ

การเชื่อมต่อสมาร์ทโฟน ดังนั้นโทรทัศน์หรือโทรทัศน์ในยุคนี้คือสมาร์ททีวี จึงสามารถเข้าใช้ App Store, เกม, การเล่นเกมวิดีโอ, โซเชียลเน็ตเวิร์ก, เว็บเบราว์เซอร์ ตลอดจนการสตรีมภาพยนตร์ อย่างไรก็ตาม แอปทีวีที่มีให้บริการในปัจจุบันไม่ได้แตกต่างจากแอปในสมาร์ทโฟนมากนัก

ปัจจุบันเป็นเรื่องปกติที่ผู้ให้บริการในอุตสาหกรรมการสตรีมวิดีโอ (Streaming) หรือ Over-the-top (OTT) จะให้บริการผ่านอินเทอร์เน็ต แต่ไม่จำเป็นต้องเป็นเจ้าของเครือข่ายอินเทอร์เน็ตด้วยตนเอง เมื่อรวมคำว่า Over-the-top (OTT) กับคำว่า TV จะหมายถึงผู้ให้บริการวิดีโอผ่านอินเทอร์เน็ต เช่น YouTube, Line TV และ Netflix จากการประมาณการโดย Netflix ผู้ให้บริการวิดีโอสตรีมมิงในประเทศไทย จำนวนผู้ใช้ที่ลงทะเบียนในประเทศไทยในปี 2561 อยู่ที่ 287,194 ราย โดย 244,074 รายจ่ายค่าสมัครรายเดือน และ 43,120 รายใช้บริการทดลองใช้ฟรี ปัจจัยหลักที่มีอิทธิพลต่อการตัดสินใจสมัครเป็นสมาชิกของแอปพลิเคชันวิดีโอสตรีมมิงในกรุงเทพฯ ได้แก่ ค่าบริการ ความหลากหลายและคุณภาพของเนื้อหา และความสะดวกในการเข้าถึง [3]

ด้วยข้อจำกัดของการจัดจาและการเชื่อมต่อของผู้ใช้สูงอายุที่เป็นอุปสรรคสำคัญต่อการเริ่มต้นเรียนรู้ การสร้างความเข้าใจ และการปฏิสัมพันธ์กับเทคโนโลยีใหม่ซึ่งมีรูปแบบและลักษณะการทำงานที่เปลี่ยนแปลงจากเทคโนโลยีในอดีตที่ผ่านมาค่อนข้างมาก [4]

2.2 แอปสตรีมมิงวิดีโอ

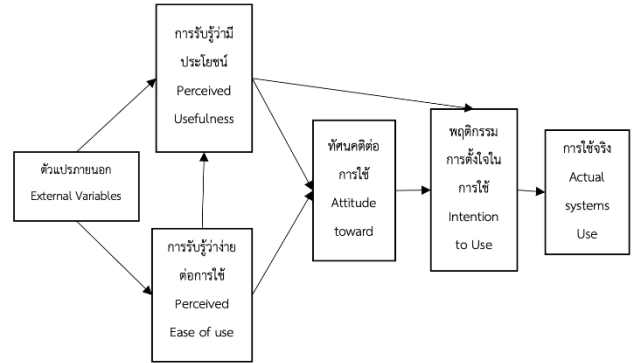
อิทธิพลต่อการยอมรับการใช้งานแอปสตรีมมิงวิดีโอ (Streaming) ในเขตกรุงเทพฯ และปริมณฑล ปัจจัยหนึ่งคือความสะดวกในการสตรีม เนื่องจากผู้ใช้สามารถเข้าถึงเนื้อหาโดยไม่จำเป็นต้องดาวน์โหลดไฟล์ขนาดใหญ่ซึ่งอาจใช้เวลาานาน อีกปัจจัยหนึ่งคือประสิทธิภาพของอินเทอร์เน็ตและหน่วยความจำของคอมพิวเตอร์หรืออุปกรณ์ซึ่งต้องเพียงพอต่อการรองรับการสตรีม นอกจากนี้ ความพร้อมใช้งานและความสามารถในการเข้าถึงเนื้อหาในแอปอาจส่งผลต่อการนำไปใช้ ตลอดจนอิทธิพลทางสังคมของเพื่อนและองค์กรที่ใช้แอป คุณลักษณะส่วนบุคคล เช่น อายุ เพศ และความสามารถทางเทคโนโลยี ปัจจัยเหล่านี้ล้วนส่งผลต่อการรับรู้ของแต่ละคนเกี่ยวกับ

ประโยชน์และความสะดวกในการใช้งานของแอปสตรีมวิดีโอ ซึ่งทำที่สุดจะนำไปสู่การนำไปใช้ โดยผู้ให้บริการสตรีมวิดีโออย่าง Netflix มีทัศนคติและปัจจัยที่นำไปสู่ความสำเร็จของระบบข้อมูลในด้านความพึงพอใจของลูกค้าต่อการใช้บริการ Netflix นั้นแตกต่างกันไปในแต่ละช่วงอายุคน การศึกษาพบว่ามีปัจจัยหลัก 3 ปัจจัยที่ส่งผลต่อความพึงพอใจของลูกค้าที่มีต่อ Netflix ได้แก่ ประโยชน์สุทธิ คุณภาพของข้อมูล และคุณภาพของระบบ ในทางตรงกันข้าม คุณภาพการบริการและการนำเสนอไม่ได้ส่งผลกระทบต่อความพึงพอใจของลูกค้าที่มีต่อ Netflix อย่างมีนัยสำคัญ ในแง่ของลักษณะทางประชากร การศึกษาพบว่าเพศและอายุไม่ส่งผลกระทบต่อความพึงพอใจของลูกค้าที่มีต่อ Netflix อย่างมีนัยสำคัญ ในขณะที่ระดับการศึกษามีผลกระทบ โดยเฉพาะอย่างยิ่ง ระดับการศึกษาที่สูงขึ้นมีความสัมพันธ์กับระดับความพึงพอใจต่อบริการที่สูงขึ้น นอกจากนี้ การศึกษาายังพบว่าความถี่ในการใช้งานเป็นตัวทำนายความพึงพอใจของลูกค้าที่มีต่อ Netflix อย่างมีนัยสำคัญ โดยการใช้งานที่สูงขึ้นจะนำไปสู่ระดับความพึงพอใจที่สูงขึ้น [5]

2.3 พฤติกรรมการยอมรับเทคโนโลยี

ทฤษฎีพฤติกรรมการยอมรับเทคโนโลยีสารสนเทศ ทฤษฎีการกระทำตามหลักเหตุและผล (The theory of reasoned action: TRA) [6] ซึ่งเป็น ทฤษฎีทางจิตวิทยา สังคม (Social psychology) ที่ถูกนำมาใช้ในการศึกษา พฤติกรรมของมนุษย์มากที่สุด โดยได้อธิบายถึงความสัมพันธ์ ระหว่างความเชื่อและทัศนคติที่มีต่อพฤติกรรมว่า การเปลี่ยนแปลงพฤติกรรมเป็นผลมาจากการเปลี่ยนแปลง ความเชื่อบุคคลจะแสดงพฤติกรรมเพราะคิดว่าเป็นสิ่งที่ถูกต้องเหมาะสม เนื่องจากจะต้องพิจารณาเหตุและผล ก่อนการกระทำเสมอหลักการของ TRA คือ ปัจจัยที่เป็นตัวกำหนดพฤติกรรมของบุคคลนั้นคือ ความตั้งใจ แสดง พฤติกรรม (Behavioural intention) ซึ่งความตั้งใจแสดงพฤติกรรมจะได้รับอิทธิพลจาก ทัศนคติ ที่มีต่อพฤติกรรม (Attitudes towards the behaviour) คือปัจจัยที่เกิดขึ้นภายในตัวบุคคลประกอบด้วย ความเชื่อ ความรู้สึก ทั้งทางบวกและทางลบเกี่ยวกับพฤติกรรม โดยเชื่อว่าถ้าบุคคลมีความเชื่อว่าให้ผลทางบวกก็จะมีทัศนคติ ที่ดีต่อสิ่ง

นั้นด้วยในทางตรงข้ามกันถ้าบุคคลมีความเชื่อว่าให้ผลทางลบก็จะมีทัศนคติที่ไม่ดีต่อสิ่งนั้นเช่นกัน ซึ่งส่งผลต่อพฤติกรรมการยอมรับการใช้งานในรูปแบบ



ภาพที่ 1: รูปแบบการยอมรับเทคโนโลยี

การจำลองการยอมรับเทคโนโลยี (TAM) [7] การยอมรับเทคโนโลยีของแต่ละบุคคลได้รับอิทธิพลจากตัวแปรภายนอกที่กำหนดการรับรู้ของแต่ละบุคคล ตัวแปรเหล่านี้อาจรวมถึงความรู้ ความเชื่อ ประสบการณ์ หรือพฤติกรรมทางสังคม ซึ่งนำไปสู่กระบวนการยอมรับเทคโนโลยี

3. วิธีการดำเนินงาน

ด้วยกลุ่มตัวอย่างที่ใช้ในการศึกษานี้คือ ผู้สูงอายุในประเทศไทยทั้งชายและหญิงที่มีช่วงอายุ 60 ปีขึ้นไป เป็นกลุ่มผู้สูงอายุในกรุงเทพมหานครหรือ มีที่พักอาศัยอยู่ในกรุงเทพมหานคร ในปัจจุบัน จำนวน 50 คน โดยใช้วิธีการเรียนเชิญ โดยการเลือกตัวอย่างแบบเฉพาะเจาะจงจากผู้สูงอายุที่มีความสามารถในการใช้เทคโนโลยีแต่ละระดับ อาทิ ผู้ใช้งานสมาร์ทโฟนในการใช้แอปพลิเคชันทางโซเชียลมีเดีย อาทิ Facebook, Line และ Youtube ซึ่งสมาร์ทโฟนเป็นอุปกรณ์แรกๆ ในยุคนี้ที่สามารถเข้าถึงกลุ่มผู้สูงอายุได้ และเป็นผู้ที่เคยใช้ทีวีดิจิทัล สมาร์ททีวี ตลอดจนผู้ที่ยังไม่เคยใช้สมาร์ททีวีเลย เพราะผู้วิจัยสนใจที่จะเข้าใจว่าการใช้เทคโนโลยีส่งผลกระทบต่อชีวิตของผู้สูงอายุอย่างไร

การเก็บรวบรวมข้อมูลโดยพื้นที่เป้าหมายในการเก็บข้อมูลในงานวิจัยครั้งนี้อยู่ในกรุงเทพมหานคร โดยจะเน้นไปที่สมาคมหรือชมรมผู้สูงอายุในกรุงเทพมหานคร ตลอดจนผู้สูงอายุที่สมัครใจเข้ารับการทดสอบ และผู้สูงอายุที่ถูกรับเชิญมาทดสอบ โดยใช้การวิเคราะห์ข้อมูลแบบ (Chris Brooks, 2020) [8] โดยใช้ค่าสถิติ confidence interval หรือช่วงความเชื่อมั่น ว่าค่ากลุ่มประชากรจะมีค่าอยู่ในช่วงความเชื่อมั่น

$$\bar{x} \pm \frac{t_{n-1}^{0.025} S}{\sqrt{n}} = 95\%$$

ข้อมูลที่น่ามาวิเคราะห์จะอาศัยโปรแกรมสำเร็จรูป IBM SPSS Statistics มาช่วยในการประมวลผลได้ผลการวิเคราะห์ดังนี้

4. วิธีการดำเนินงาน

4.1 ผลการเก็บข้อมูล

ทดสอบโดยใช้แบบสอบถามในการเก็บค่าความเชื่อมั่นจากผู้เข้ารับการทดสอบ โดยทำการเก็บข้อมูลทั่วไป ข้อมูลด้านปัญหาการใช้งาน ข้อมูลการรับรู้ประโยชน์ในการใช้งาน ตลอดจนข้อมูลการรับรู้ประโยชน์ในการใช้แอปพลิเคชัน

ตารางที่ 1: คุณลักษณะส่วนบุคคลของผู้ตอบแบบสอบถาม จากจำนวนของผู้สูงอายุผู้เข้ารับการทดสอบ 50 คน มีที่พักอาศัยในกรุงเทพมหานคร

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
ชาย	26	52.0
หญิง	24	48.0

จากตารางที่ 1 พบว่าผู้สูงอายุที่มากที่สุดเป็นเพศชาย คิดเป็นร้อยละ 52.0 ผู้สูงอายุที่เป็นเพศหญิงคิดเป็นร้อยละ 48.0 โดยส่วนใหญ่ของอายุที่เข้ารับการทดสอบอยู่ในช่วงอายุที่ 54 ปี คิดเป็นร้อยละ 12.0

ตารางที่ 2: คุณลักษณะส่วนบุคคลของผู้ตอบแบบสอบถาม ในกลุ่มผู้สูงอายุจำแนกตามระดับการศึกษาของผู้ตอบแบบสอบถาม

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
ประถมศึกษา หรือต่ำกว่า	9	18.0
มัธยมศึกษาปีที่ 3	12	24.0
มัธยมศึกษาปีที่ 6 / ปวช	10	20.0
ปวส./ปวท./อนุปริญญา	11	22.0
ปริญญาตรี	8	16.0

จากตารางที่ 2 พบว่าผู้สูงที่เข้ารับการสำรวจมีระดับการศึกษาอยู่ที่ระดับ มัธยมศึกษาปีที่ 3 มากที่สุดคิดเป็นร้อยละ 24.0 และรองลงมาจะเป็นระดับ ปวส./ปวท./อนุปริญญา คิดเป็นร้อยละ 22.0

ตารางที่ 3: สถานะในการพักอาศัยของผู้สูงอายุจำแนกตามที่พักอาศัยของผู้สูงอายุ

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
อยู่ตามลำพัง	2	4.0
บิดา/มารดา/ญาติผู้ใหญ่	2	4.0
พี่น้อง	5	10.0
คู่สมรส	10	20.0
บุตร	24	48.0
บุตรเขย/บุตรสะใภ้	6	12.0
หลาน	1	2.0

จากตารางที่ 3 พบว่าผู้สูงอายุที่มีสถานะการพักอาศัย โดยอยู่พักอาศัยกับบุตรมากที่สุดคิดเป็นร้อยละ 48.0 และรองลงมาเป็นการพักอาศัยกับคู่สมรส คิดเป็นร้อยละ 20.0

ตารางที่ 4: สถานะของประสบการณ์การได้เข้าใช้งานแอปพลิเคชัน
สมาร์ตทีวี

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
เคย	36	68.0
ไม่เคย	16	32.0

จากตารางที่ 4 พบว่าผู้สูงอายุกลุ่มตัวอย่างส่วนใหญ่
เคยใช้งานแอปพลิเคชันสมาร์ตทีวีโดยตรงคิดเป็นร้อยละ 68.0

ตารางที่ 5: พฤติกรรมการใช้งานแอปพลิเคชันสมาร์ตทีวี
แสดงระยะเวลาการใช้งานแอปพลิเคชันสมาร์ตทีวี
ระยะเวลาการใช้งานอินเทอร์เน็ตและการเรียนรู้
ผู้ใช้งาน

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
ท่านใช้สมาร์ตทีวีเป็นระยะเวลาานเท่าใด		
2 ชั่วโมงต่อวัน	10	20.0
6 ชั่วโมงต่อวัน	22	44.0
8 ชั่วโมงต่อวัน	16	32.0
11 ชั่วโมงต่อวัน	2	4.0
ท่านใช้อินเทอร์เน็ตเป็นระยะเวลาานเท่าใด		
2 ชั่วโมงต่อวัน	10	20.0
6 ชั่วโมงต่อวัน	13	26.0
8 ชั่วโมงต่อวัน	25	50.0
11 ชั่วโมงต่อวัน	2	4.0
ท่านใช้วิธีใดในการเรียนรู้การใช้งาน		
ให้ผู้ใกล้ชิดช่วยสอน	33	66.0
ค้นหาวิธีใช้จากเว็บไซต์	8	16.0
ทดลองด้วยตัวเอง	9	18.0

จากตารางที่ 5 พบว่ากลุ่มผู้สูงอายุที่ใช้งานแอปพลิเคชัน
สมาร์ตทีวีมากที่สุดจำนวน 6 ชั่วโมงต่อวันคิดเป็นร้อยละ
44.0 เมื่อนำมาเปรียบเทียบกับระยะเวลาการใช้งาน
อินเทอร์เน็ตไม่ว่าจะใช้งานผ่านอุปกรณ์ต่าง ๆ ผลปรากฏ
ผู้สูงอายุใช้งานอินเทอร์เน็ตมากที่สุด 8 ชั่วโมงต่อวัน โดย

คิดเป็นร้อยละ 50.0 และเมื่อดูผลสำรวจการเรียนรู้การใช้งาน
แอปพลิเคชันสมาร์ตทีวีพบว่า ผู้สูงอายุส่วนใหญ่ให้ผู้
ใกล้ชิดช่วยสอนการใช้งานมากที่สุด โดยคิดเป็นร้อยละ 66.0

ตารางที่ 6: ปัญหาในการใช้งานแอปพลิเคชันสมาร์ตทีวี โดยจะทำ
การเก็บค่าความเชื่อมั่นจากผู้เข้ารับการทดสอบ

ปัญหาในการใช้งาน	\bar{X}	S.D.	ช่วงความ เชื่อมั่น
สามารถค้นหา รายการที่สนใจได้	3.00	.639	2.8185 - 3.1815
สามารถแก้ไขหน้า โปรไฟล์ได้ง่าย	3.16	.710	2.9583 - 3.3617
ท่านตัดสินใจ รับชมรายการจัด อันดับ ที่แอปพลิเคชัน แนะนำเป็นหลัก	3.08	.829	2.8445 - 3.3155
เลือกช่วงเวลา การรับชมได้	3.16	.710	2.9583 - 3.3617
สามารถที่จะเรียนรู้ เมื่อมีฟังก์ชัน ใหม่เข้ามา	3.18	.691	2.9837 - 3.3763
มองเห็นเนื้อหา ของ รายละเอียดที่ สามารถเข้าใจได้ ง่าย	3.10	.647	2.9162 - 3.2838
เมนูที่ใช้งานยาก เกินไป	3.64	.851	3.3983 - 3.8817
ขนาดตัวอักษรไม่ ชัดเจน	3.72	.730	3.5126 - 3.9274
การแบ่งหมวดหมู่ การรับชมที่ ซับซ้อน	3.86	.756	3.6453 - 4.0747

จากตารางที่ 6 พบว่ากลุ่มผู้สูงอายุที่ใช้งานแอปพลิเคชัน
สมาร์ตทีวีพบปัญหาในการใช้งาน โดยเรียงลำดับจาก
ควรปรับปรุงจากมากที่สุดไปน้อยที่สุดเริ่มจาก การแบ่ง

โหมคหมู่การรับชมที่ซับซ้อน โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.51 - 3.92 เมนูที่ใช้งานยากเกินไป โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.39 - 3.88 การแบ่งโหมคหมู่การรับชมที่ซับซ้อน โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.64 - 4.07

ตารางที่ 7: ปัญหาในการใช้งานแอปพลิเคชันสมาร์ตทีวี โดยการเก็บค่าความพึงพอใจในการใช้งานในส่วนของการแก้ไขปัญหา

คุณลักษณะส่วนบุคคล	จำนวน	ร้อยละ
เมื่อเกิดปัญหาท่านมักแก้ไขด้วยตนเองทุกครั้งหรือไม่		
แก้ด้วยตนเองทุกครั้ง	17	34.0
แก้ด้วยคนอื่นบางครั้ง	13	26.0
แก้ไขด้วยคนอื่นทุกครั้ง	20	40.0

จากตารางที่ 7 พบว่ากลุ่มผู้สูงอายุที่ใช้ที่ใช้งานแอปพลิเคชันสมาร์ตทีวีเมื่อพบปัญหาขณะใช้งานผู้เข้ารับการทดสอบส่วนมากคิดเป็นร้อยละ 40.0 จะใช้วิธีการแก้ไขโดยการขอความช่วยเหลือหรือกับคนอื่นหรือผู้ใกล้ชิด

ตารางที่ 8: ปัจจัยด้านการยอมรับการใช้งานแอปพลิเคชันสมาร์ตทีวี โดยทำการเก็บค่าความเชื่อมั่น

การรับรู้ประโยชน์ของแอปพลิเคชัน	\bar{X}	S.D.	ช่วงความเชื่อมั่น
ได้รับประโยชน์จากการดูที่แนะนำ	3.00	.553	2.8429 - 3.1571
ได้ประโยชน์จากการค้นหารายการที่สนใจ	3.16	.580	2.9952 - 3.3248
ได้ประโยชน์จากการแก้ไขหน้าโปรไฟล์	3.08	.528	2.9300 - 3.2300
ได้ประโยชน์จากการค้นหารายการที่สนใจ	3.16	.493	3.0200 - 3.3000

ได้ประโยชน์จากการตัดสินใจรับชมรายการจากการจัดอันดับ	3.18	.769	2.9616 - 3.3984
ได้ประโยชน์การใช้งานการเลือกเวลาที่รับชมได้	3.10	.678	2.9074 - 3.2926

จากตารางที่ 8 พบว่ากลุ่มผู้สูงอายุที่เข้าทำการทดสอบแอปพลิเคชันด้านการรับรู้ประโยชน์ของแอปพลิเคชัน โดยเชื่อมั่นเกี่ยวกับการรับรู้ประโยชน์มากที่สุดคือการได้ประโยชน์จากการค้นหารายการที่ท่านสนใจ โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.02 - 3.30

ตารางที่ 9: ปัจจัยด้านการยอมรับการใช้งานแอปพลิเคชันสมาร์ตทีวี โดยทำการเก็บค่าความเชื่อมั่น

การรับรู้ประโยชน์ของแอปพลิเคชัน	\bar{X}	S.D.	ช่วงความเชื่อมั่น
ขนาดความพอดีของปุ่มและรูปภาพ	3.14	.639	2.9585 - 3.3215
ความสว่างของหน้าจอ	3.08	.528	2.9300 - 3.2300
รู้สึกสับสนสัญลักษณ์ภาพและคำสั่ง	3.74	.664	3.5514 - 3.9286
เมนูคำสั่งมีมากเกินไป	3.72	.858	3.4763 - 3.9637
เมนูตัวเลือกใช้งานยาก	3.84	.710	3.6383 - 4.0417
ไม่สามารถใช้งานได้แม้ว่าผู้ใช้เคยมีประสบการณ์ใช้แอปพลิเคชันอื่นมาก่อน	3.82	.748	3.6075 - 4.0325

จากตารางที่ 9 พบว่ากลุ่มผู้สูงอายุที่เข้าทำการทดสอบแอปพลิเคชันด้านการรับรู้ประโยชน์ของ

แอปพลิเคชัน โดยเชื่อมั่นเกี่ยวกับการรับรู้ประโยชน์มากที่สุดคือ พีเจอร์และเมนูตัวเลือกใช้งานยากโดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.63 - 4.04 รองลงมาเป็น ไม่สามารถใช้งานได้แม้ว่าผู้ใช้เคยมีประสบการณ์ใช้แอปพลิเคชันอื่นมาก่อน โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.60 - 4.03 มีความรู้สึกสับสนในสัญลักษณ์ภาพและคำสั่ง โดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.55 - 3.92 และ เมนูคำสั่งมีจำนวนมากเกินไปโดยค่าช่วงความเชื่อมั่นอยู่ที่ 3.47 - 3.96 ซึ่งจากการทดสอบของกลุ่มตัวอย่างสรุปว่าตัวเลือกที่ผู้ทดสอบได้ทดสอบมาโดยให้คะแนนค่อนข้างมากในเรื่องของข้อเสียในแอปพลิเคชัน

ตารางที่ 10: ข้อเสนอแนะ โดยภาพรวมของผู้สูงอายุเกี่ยวกับแนวทางการออกแบบแอปพลิเคชันสมาร์ตทีวี

ลำดับ	ประเด็นข้อเสนอแนะ
1	ต้องการให้แอปพลิเคชันสมาร์ตทีวีเป็นส่วนหนึ่งของอุปกรณ์ที่ให้ความสะดวกภายในบ้าน ได้อย่างมีประสิทธิภาพ
2	ต้องการแชร์รายการให้เพื่อนดูได้หรือ มีกลุ่มที่สามารถพูดคุยแลกเปลี่ยนเป็นสังคมออนไลน์
3	การใช้สัญลักษณ์หรือรูปแบบควรเข้าใจง่ายและมีข้อความที่ชัดเจน
4	ต้องการให้มีการสั่งใช้งานด้วยคำสั่งเสียงได้
5	ต้องการให้ปุ่มสามารถมองเห็นได้ง่ายหรือใช้ได้ง่ายขึ้น
6	ต้องการรูปแบบการใช้งานที่เรียบง่ายไม่ซับซ้อน
7	อยากให้มีช่องทางลัดที่เข้าถึงรายการที่รับชมอยู่บ่อย ๆ

จากตารางที่ 10 ข้อเสนอแนะพบว่า ผู้ทำแบบสอบถามทั้ง 7 คน ต้องการเน้น ที่คุณสมบัติหลักที่จำเป็นในแอปพลิเคชันสมาร์ตทีวีให้สอดคล้องความสะดวกสบายในบ้านที่ได้อย่างมีประสิทธิภาพ พร้อมด้วยสัญลักษณ์หรือรูปแบบที่ใช้งานง่ายรวมถึงการใช้ข้อความที่ชัดเจนแล้ว ผู้ใช้หรือผู้ทดสอบแบบสอบถามยังต้องการความสามารถในการแชร์รายการกับเพื่อนหรือชุมชนออนไลน์ ตลอดจนต้องการให้สามารถใช้คำสั่งเสียงในการสั่งการใช้งานได้

และมีปุ่มที่มองเห็นได้และใช้งานง่าย นอกจากนี้ ผู้ใช้ยังชอบรูปแบบการใช้งานที่ง่ายและไม่ซับซ้อน พร้อมด้วยต้องการวิธีที่สะดวกในการเข้าถึงรายการที่รับชมได้ง่ายกว่า

5. สรุป

5.1 ผลการศึกษา

การเก็บข้อมูลด้านทัศนคติเกี่ยวกับการใช้เทคโนโลยีการรับรู้ถึงการควบคุม พบว่ากลุ่มตัวอย่างผู้สูงอายุมีความสามารถในการใช้งานได้ในระดับหนึ่งเมื่อดูจากค่าร้อยละของความสามารถในการใช้งานจากการเลือกช่วงเวลาการรับชมโดยจะได้ค่าร้อยละอยู่ที่ 2.95 - 3.36 เมื่อนำมาเปรียบเทียบกับความสามารถในการใช้งานการเลือกช่วงเวลา การรับชมของผู้สูงอายุโดยคิดเป็นร้อยละ 2.99 - 3.32 โดยมีค่าเฉลี่ยใกล้เคียงกันซึ่งสามารถสันนิษฐานตามค่าช่วงความเชื่อมั่นของกลุ่มประชากรได้ว่าแอปพลิเคชันสมาร์ตทีวีในปัจจุบันผู้สูงอายุสามารถใช้งานได้ในระดับที่การใช้งานค่อนข้างยาก

การเก็บข้อมูลการสอดคล้องกับการใช้งาน พบว่ากลุ่มตัวอย่างผู้สูงอายุมีความสับสนในสัญลักษณ์ภาพและคำสั่ง โดยค่าร้อยละของช่วงความเชื่อมั่นอยู่ที่ 3.55 - 3.92 เมื่อนำมาเปรียบเทียบกับด้านขนาดความพอดีของปุ่มและขนาดของรูปภาพหรือสัญลักษณ์ [11] พบว่า ค่าร้อยละของช่วงความเชื่อมั่นอยู่ที่ 2.95 - 3.32 โดยค่าเฉลี่ยค่อนข้างน้อยจึงสันนิษฐานตามค่าช่วงความเชื่อมั่นของกลุ่มผู้ใช้งานผู้สูงอายุที่มีต่อความสอดคล้องกับการใช้งานแอปพลิเคชันสมาร์ตทีวีมีความยากในระดับที่สูง

การเก็บข้อมูลด้านปัจจัยด้านสิ่งแวดล้อม ว่ามีผลต่อการใช้งานแอปพลิเคชันสมาร์ตทีวีในกลุ่มผู้สูงอายุพบว่า ในสถานการณ์ที่เกิดปัญหาด้านการใช้งานพบว่าผู้สูงอายุจะขอความช่วยเหลือจากคนใกล้ชิด โดยคิดเป็นร้อยละ 40.0 ซึ่งสันนิษฐานได้ว่าสิ่งแวดล้อมและผู้ใช้ใกล้ชิดมีความสำคัญในการใช้งานแอปพลิเคชันสมาร์ตทีวีของผู้สูงอายุ

การเก็บข้อมูลด้านปัจจัยด้านการรับรู้ประโยชน์ในการใช้งานแอปพลิเคชันสมาร์ตทีวีในกลุ่มผู้สูงอายุพบว่า ประโยชน์การใช้งานการเลือกเวลาที่รับชมมีค่าเฉลี่ยช่วยความเชื่อมั่นอยู่ที่ 2.90 - 3.29 ประโยชน์จากการตัดสินใจ

รับชมรายการจากการจัดอันดับ ค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 2.96 - 3.39 ประโยชน์จากการค้นหารายการที่สนใจ ค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.02 - 3.30 ประโยชน์จากการแก้ไขหน้า โปรไฟล์ของท่านค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 2.93 - 3.23 ประโยชน์จากการค้นหารายการที่ท่านสนใจค่าเฉลี่ยช่วงความเชื่อมั่น อยู่ที่ 2.99 - 3.32 และประโยชน์จากการดูแอปพลิเคชันที่แนะนำค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 2.84 - 3.15 เมื่อทำการพิจารณาค่าเฉลี่ยเปรียบเทียบทั้งหมดพบว่า การใช้งานสมาร์ตทีวีในกลุ่มผู้สูงอายุด้านปัจจัยรวมถึงด้านการรับรู้ประโยชน์นั้น ผลค่าความเชื่อมั่นค่อนข้างน้อยซึ่งสันนิษฐานฟังก์ชันที่อยู่ในสมาร์ตทีวีไม่ค่อยตอบโจทย์ต่อการใช้งานในกลุ่มผู้สูงอายุ

นอกจากนี้ยังพบอีกว่าด้านความยากในการใช้งานของผู้สูงอายุที่มีต่อแอปพลิเคชันสมาร์ตทีวี ในส่วนของเมนูที่ใช้งานยากเกินไป โดยพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.39 - 3.88 ขนาดของตัวอักษรไม่ชัดเจนพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.51 - 3.92 ประเภทการแบ่งหมวดหมู่การรับชมที่ซับซ้อนมากเกินไปพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.64 - 4.07 เมนูคำสั่งมีจำนวนมากเกินไปพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.47 - 3.96 ฟีเจอร์และเมนูตัวลือกใช้งานยากเกินไปพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.63 - 4.04 และการใช้งานไม่สามารถใช้งานได้แม้ว่าผู้ใช้เคยมีประสบการณ์ใช้แอปพลิเคชันอื่นมาก่อนพบว่าค่าเฉลี่ยช่วงความเชื่อมั่นอยู่ที่ 3.60 - 4.03 ซึ่งผลค่าเฉลี่ยความเชื่อมั่นในเรื่องของการใช้งาน โดยรวมค่อนข้างเป็นปัญหาแก่ผู้ใช้งานแอปพลิเคชันสมาร์ตทีวี เนื่องจากต้องใช้เวลาพยายามในการเข้าใจและเรียนรู้ในตัวแอปพลิเคชันสมาร์ตทีวีอย่างมากจึงจะสามารถเริ่มใช้งานตัวแอปพลิเคชันสมาร์ตทีวีได้ ซึ่งด้วยเหตุนี้เองจึงอาจจะเป็นสาเหตุหนึ่งที่ทำให้ผู้ใช้ล้มเลิกความตั้งใจในการใช้งานแอปพลิเคชันไปโดยสิ้นเชิง

5.2 ข้อเสนอแนะ

ในการเก็บข้อมูลครั้งต่อไปเพื่อนำมาสร้างแบบทดสอบเสมือนจริงในการทดสอบครั้งต่อไปโดยอิงจากฐานข้อมูลการเก็บแบบสำรวจที่กล่าวมาข้างต้น โดยผู้วิจัยจะทำการทดสอบภายใต้กรอบสมมติฐานแบบจำลอง

การยอมรับเทคโนโลยี (TAM) ผ่านรูปแบบการเก็บแบบสอบถามพร้อมกับการทดสอบด้วยการทดลองให้ผู้ทดสอบลองใช้งานแอปพลิเคชันสมาร์ตทีวี ในรูปแบบแอปพลิเคชันเสมือนจริงและทำการทดสอบเก็บค่ากับกลุ่มตัวอย่างผู้สูงอายุเพื่อเป็นแนวทางในการออกแบบแอปพลิเคชันสมาร์ตทีวี ให้เหมาะกับกลุ่มผู้สูงอายุในประเทศไทยต่อไป

เอกสารอ้างอิง

- [1] BBC NEWS Thai Apr. 2019 [:https://www.bbc.com/thai/international-47801600](https://www.bbc.com/thai/international-47801600)
- [2] รุจา รอดเข็มและ สุภารัตน์ ไชยประสิทธิ์ “สังคมสูงวัย เทคโนโลยีผู้สูงอายุ มหาวิทยาลัยอีสเทิร์นเอเซีย พ.ศ. 2562”
- [3] พิมพ์ชนก พึ่งสิริรัตน์ อินทกะ พิริยะกุล การวิเคราะห์พฤติกรรมผู้บริโภคตลาดสตรีมมิงด้วยวิธีการวิเคราะห์แบบคอมจอยท์ พ.ศ. 2564
- [4] J. M. Beer and L. Takayama, “Mobile remote presence systems for older adults: acceptance, benefits, and concerns”, *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Lausanne, Switzerland, 2011
- [5] ถิรดา มธุรสพรวัฒนา “ทัศนคติและปัจจัยความสำเร็จของระบบสารสนเทศที่มีอิทธิพลต่อความพึงพอใจในการใช้บริการ Netflix” พ.ศ. 2561
- [6] Thomas J Madden Pamela Scholder Ellen Ice Ajzen A Comparison of the theory of Planned Behaviour and the Theory of Reasoned Action 1992
- [7] Fred D. Davis, Jr. B.S., Industrial Engineering Wayne State University A TECHNOLOGY ACCEPTANCE MODEL FOR EMPIRICALLY TESTING NEW END-USER INFORMATION SYSTEMS: THEORY AND RESULTS 1980
- [8] Brooks, C. Introductory econometrics for finance. Cambridge University Press. 2020

การทดสอบส่วนต่อประสานผู้ใช้งานอัตโนมัติของเว็บแอปพลิเคชันโดยใช้ ไซเปรสเฟรมเวิร์ก: กรณีศึกษาระบบเว็บการประเมินออนไลน์ Automatic User Interface Testing for Web Application with Cypress Framework: A Case Study of an Online Assessment Web Application

รัชพันธ์ ศรีสมุท (Tatchanon Srisamudh)¹ อติวงศ์ สุชาโต (Atiwong Suchato)² ทิตยา หวานวารี (Dittaya Wanvarie)³

นฤมล ประทานวิช (Naruemon Pratanwanich)⁴ และเนื้องวงศ์ ทวยเจริญ (Nuengwong Tuaycharoen)⁵

^{1, 2, 5}ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

^{3, 4}ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

¹6470194121@student.chula.ac.th, ²atiwong.s@chula.ac.th, ³Dittaya.W@chula.ac.th,

⁴naruemon.p@chula.ac.th, ⁵nuengwong.t@chula.ac.th

บทคัดย่อ

ปัจจุบันในการทดสอบเว็บแอปพลิเคชัน หากใช้มนุษย์จะต้องใช้ระยะเวลาในการทดสอบนานกว่าการนำเฟรมเวิร์กสำหรับการทำการทดสอบแบบอัตโนมัติ ซึ่งเฟรมเวิร์กที่กำลังได้รับความนิยมได้แก่ไซเปรสเฟรมเวิร์ก (Cypress Framework) เนื่องจากรองรับภาษาที่ง่ายต่อการใช้งาน และสามารถดำเนินการทดสอบโดยใช้เวลาไม่มากเพื่อให้ซอฟต์แวร์สามารถถูกส่งมอบให้ผู้ใช้งานได้อย่างรวดเร็ว งานวิจัยนี้ได้จัดทำเพื่อเป็นกรณีศึกษาในการนำไซเปรสเฟรมเวิร์กมาใช้ในการทดสอบระบบการประเมินออนไลน์อัตโนมัติ สำหรับการทดสอบส่วนต่อประสานผู้ใช้งานของเว็บแอปพลิเคชันที่ยังมีการเปลี่ยนแปลงอย่างต่อเนื่องจากการพัฒนา ให้สามารถทำการทดสอบเพื่อค้นพบข้อผิดพลาดได้อย่างรวดเร็ว โดยมีรายงานการรับประกันซอฟต์แวร์ได้ว่ามีคุณภาพ และสามารถลดระยะเวลาหลายชั่วโมงในการทดสอบด้วยมือได้ โดยใช้เวลาในการทดสอบทั้งระบบเพียง 3 นาที 55 วินาที

คำสำคัญ: การทดสอบซอฟต์แวร์ ไซเปรสเฟรมเวิร์ก

เฟรมเวิร์กการทดสอบอัตโนมัติ

การทดสอบอัตโนมัติ เว็บแอปพลิเคชัน

Abstract

Currently, testing web applications manually requires significantly more time than using a test automation framework. The Cypress Framework has gained popularity due to its user-friendly language support and ability to conduct efficient tests. This research serves as a case study for using the Cypress Framework to test an online evaluation system's user interface. The system undergoes continuous development, making it crucial to quickly identify and resolve errors. The software quality assurance report confirms the effectiveness of the Cypress Framework in reducing manual testing time by more than hours while ensuring software quality with a time duration is 3 minutes and 55 seconds.

Keywords: Software Testing, Cypress Framework

Test automation framework,

Test Automation, Web Application

1. บทนำ

ในปัจจุบันเว็บแอปพลิเคชันเป็นที่นิยมมากขึ้นเนื่องจากเหมาะกับองค์กรทุกขนาดและสามารถใช้งานได้ง่ายเพียงมีเว็บเบราว์เซอร์อยู่บนอุปกรณ์ การทดสอบความถูกต้องของการทำงานของเว็บแอปพลิเคชันจึงเป็น

สิ่งจำเป็น เพื่อตรวจสอบความถูกต้องของเว็บแอปพลิเคชัน ก่อนส่งมอบให้ผู้ใช้งาน ซึ่งในระหว่างการทดสอบ เว็บแอปพลิเคชันอาจมีการเปลี่ยนแปลงจากการพัฒนา ระบบตลอดเวลา เนื่องจากการทดสอบเป็นการทดสอบในขณะที่ทีมพัฒนา ยังคงพัฒนาเว็บแอปพลิเคชันอยู่ด้วย เพื่อให้พบข้อผิดพลาดได้อย่างรวดเร็ว และยังเป็นการลด ต้นทุนการพัฒนาซอฟต์แวร์ จึงจำเป็นต้องมีการทำการ ทดสอบอย่างสม่ำเสมอ โดยในสมัยก่อนมักใช้มนุษย์ใน การทดสอบซอฟต์แวร์ ซึ่งใช้ทรัพยากรและเวลามาก และ อาจเกิดความผิดพลาดจากมนุษย์ได้ ซึ่งส่งผลให้ซอฟต์แวร์ ถูกส่งออกไปให้ผู้ใช้งานได้ช้า ทำให้เกิดปัญหาต้นทุนการ พัฒนาที่สูงขึ้นตามมา ในปัจจุบันจึงนิยมนำการทดสอบ แบบอัตโนมัติด้วยเฟรมเวิร์กต่าง ๆ [5] เข้ามาใช้กันมากขึ้น เพื่อเพิ่มประสิทธิภาพของซอฟต์แวร์ ลดระยะเวลา และทรัพยากรในการทดสอบ

ในปัจจุบัน มีสถานการณ์โรคระบาดทำให้มีการเรียน การสอนในรูปแบบออนไลน์ รวมถึงการสอบออนไลน์ ทำให้ระบบการประเมินผลการเรียนออนไลน์มีความ จำเป็นอย่างยิ่ง เพื่อตอบสนองความต้องการในการ จัดการเรียนการสอนออนไลน์ ให้กับนิสิตจุฬาลงกรณ์ มหาวิทยาลัย จึงจำเป็นต้องพัฒนา และอัปเดตซอฟต์แวร์ ให้ทันต่อการเปลี่ยนแปลงของการจัดการศึกษา และ เทคโนโลยีที่เปลี่ยนแปลงอย่างรวดเร็ว การทดสอบซอฟต์แวร์ ด้วยมนุษย์ทำให้ส่งมอบซอฟต์แวร์ให้ผู้ใช้งานได้ช้า จึงจำเป็นต้องใช้การทดสอบแบบอัตโนมัติ เพื่อลดปัญหา ดังกล่าว โดยการใช้ไซเปรสเฟรมเวิร์กในการเขียน โปรแกรมเพื่อทดสอบส่วนต่อประสานผู้ใช้งานของระบบ การประเมินออนไลน์ให้เป็นอัตโนมัติ และสามารถออกรายงานผลการทดสอบเพื่อใช้ในการประกันคุณภาพของระบบได้ งานวิจัยนี้อธิบายแนวทางการใช้ไซเปรสเฟรม เวิร์กในการทดสอบระบบแบบอัตโนมัติและวัด ประสิทธิภาพการทำงานของเฟรมเวิร์กดังกล่าวโดยใช้ ระบบเว็บประเมินผลการเรียนออนไลน์เป็นกรณีศึกษา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ไซเปรสเฟรมเวิร์ก (Cypress Framework)

เป็นเครื่องมือที่มีความน่าเชื่อถือสำหรับทำการทดสอบ บนเว็บเบราว์เซอร์โดยเฉพาะการทดสอบส่วนต่อประสาน ผู้ใช้ของเว็บแอปพลิเคชัน และเป็น Open Source ทำให้ คนที่สนใจสามารถใช้งานได้ฟรี และยังมีคุณสมบัติที่ สำคัญหลายอย่าง เช่น ความสามารถประเมินได้ว่าเพราะอะไร การทดสอบจึงล้มเหลว [1] และยังสามารถทำงานได้อย่าง รวดเร็ว สามารถถ่ายภาพหน้าจอแบบวิดีโอ และภาพนิ่งได้ ในขณะที่ทำการทดสอบไปพร้อม ๆ กัน รวมถึงมีขั้นตอน การเขียนโปรแกรมที่ไม่ยากเมื่อเทียบกับเฟรมเวิร์กอื่น ๆ [5]

2.2 การทดสอบการถดถอย (Regression Test)

ทดสอบการถดถอย เป็นแนวปฏิบัติในการทดสอบ ซอฟต์แวร์อย่างหนึ่งซึ่งช่วยให้มั่นใจว่าระบบยังคงทำงาน ตามที่คาดไว้หลังจากที่มีการปรับปรุงเปลี่ยนแปลงส่วนใด ส่วนหนึ่งภายในระบบ ซึ่งการทำการทดสอบการถดถอย ต้องมีการจัดลำดับความสำคัญ บนพื้นฐานของความเสี่ยง [6] และมีการทดสอบอย่างต่อเนื่องด้วย โดยทั่วไประบบ ต้องผ่านการทดสอบหลายครั้งและการทดสอบการถดถอย เป็นขั้นตอนสุดท้ายเนื่องจากการตรวจสอบการทำงานของระบบโดยรวมก่อนส่งมอบให้ผู้ใช้งานต่อไป

2.3 End to End Test (E2E)

เป็นแนวปฏิบัติในการทดสอบซอฟต์แวร์อย่างหนึ่ง โดยทำการทดสอบตั้งแต่ต้นจนจบเพื่อให้แน่ใจว่าขั้นตอน การทำงานของแอปพลิเคชันทำงานตามที่คาดไว้ โดยเป็น การทดสอบจากประสบการณ์ของผู้ใช้งานโดยการจำลอง สถานการณ์ผู้ใช้จริง ซึ่งมีการตรวจสอบระบบภายใต้การ ทดสอบ และความสมบูรณ์ของข้อมูล [2] ในปัจจุบันระบบ ซอฟต์แวร์มีความซับซ้อนและเชื่อมโยงกับระบบย่อย จำนวนมากจึงมีความเสี่ยงที่ ถ้าหากระบบย่อยใดล้มเหลว ระบบซอฟต์แวร์ทั้งหมดอาจหยุดทำงานได้จึงทำให้ระบบ ซอฟต์แวร์สามารถหลีกเลี่ยงได้ด้วยการทดสอบแบบ End to End

2.4 Test Condition Matrix

เป็นเมทริกซ์เงื่อนไขการทดสอบที่ทำให้เห็นมุมมอง ระดับสูงของข้อมูลการทดสอบที่ใช้ในแต่ละกรณีทดสอบ ของแต่ละสถานการณ์โดยเมทริกซ์เงื่อนไขการทดสอบมี จุดประสงค์เพื่ออำนวยความสะดวกในการสร้างกรณี ทดสอบที่ไม่ซ้ำกันสำหรับแต่ละสถานการณ์โดยใช้การ เรียงสับเปลี่ยน และข้อมูลการทดสอบที่แตกต่างกัน ทำให้

สามารถตรวจสอบให้แน่ใจว่าเงื่อนไขการทดสอบที่จำเป็นทั้งหมดสำหรับกรณีการใช้งานนั้นครอบคลุมโดยกรณีทดสอบ [3]

2.5 Cucumber Reports

ปลั๊กอินที่ใช้เพื่อสร้างรายงานที่มีข้อมูลเกี่ยวกับผลของการทดสอบที่ผ่าน หรือล้มเหลวหลังจากที่มีการทำการทดสอบแบบอัตโนมัติ ซึ่งเป็นปลั๊กอินที่ถูกริใช้งานแพร่หลายในด้านของการสร้างรายงานเพื่อใช้ในการยืนยันความถูกต้องการทำงานของซอฟต์แวร์หลังจากการทดสอบแบบอัตโนมัติ และยังมีกลุ่มสังคมออนไลน์ที่คอยให้ความช่วยเหลือเกี่ยวกับข้อมูลของการใช้งาน [4]

2.6 Gherkin Style

การนำ Gherkin Syntax [10] มาใช้ในการเขียนโปรแกรมช่วยให้ผู้อื่นสามารถอ่านและเข้าใจโปรแกรมได้ง่ายโดยไม่จำเป็นต้องมีความรู้เกี่ยวกับภาษาที่ใช้ในการเขียนโปรแกรมมากนัก โดยมี Syntax หลักคือ Given When และ Then [10] ซึ่งทำให้มีความเป็นภาษาธรรมชาติมากยิ่งขึ้น โดยสามารถทำงานร่วมกันกับไซเปรสเฟรมเวิร์กได้ผ่านไฟล์ Feature

2.7 กระบวนการทดสอบ (Test Process)

ISO/IEC/IEEE 29119-2 [7] มีการกำหนดกระบวนการทดสอบซอฟต์แวร์โดยแบ่งออกเป็น 3 ขั้นตอน ได้แก่

1) การวางแผนการทดสอบ ต้องมีการวางแผนไว้แล้วว่ามีการทดสอบแบบใดบ้าง

2) การเฝ้าดูและควบคุมการทดสอบ ต้องมีการควบคุมการทดสอบอยู่เสมอว่าผลของการทดสอบในแต่ละครั้งนั้นได้ผลลัพธ์ที่ต่างไปจากตอนที่ได้ทำการวางแผนการทดสอบหรือไม่

3) การเสร็จสิ้นการทดสอบ ต้องมีการจัดทำรายงานผลการทดสอบเพื่อสรุปผลว่าการทดสอบนั้นได้เสร็จสิ้นแล้ว ซึ่งเราสามารถนำมาปรับใช้เข้ากับกระบวนการทดสอบในครั้งนี้ด้วย เพื่อให้ได้ซอฟต์แวร์มีกระบวนการทดสอบที่เป็นไปตามมาตรฐาน [8]

2.8 myCourseVille

ระบบ myCourseVille เป็นระบบการจัดการการเรียนการสอน (Learning Management System: LMS) แบบออนไลน์ผ่านอินเทอร์เน็ต ซึ่งระบบนี้ถูกสนับสนุนโดย

ศูนย์นวัตกรรมการเรียนรู้พาลงกรณ์มหาวิทยาลัย เพื่อช่วยอำนวยความสะดวกในเรื่องของการเรียนการสอนให้สอดคล้องกับทิศทางการพัฒนาของสังคมในปัจจุบันที่มีแนวโน้มไปสู่สังคมเครือข่ายออนไลน์เพิ่มมากขึ้น [9]

2.9 myCourseVille Assessment

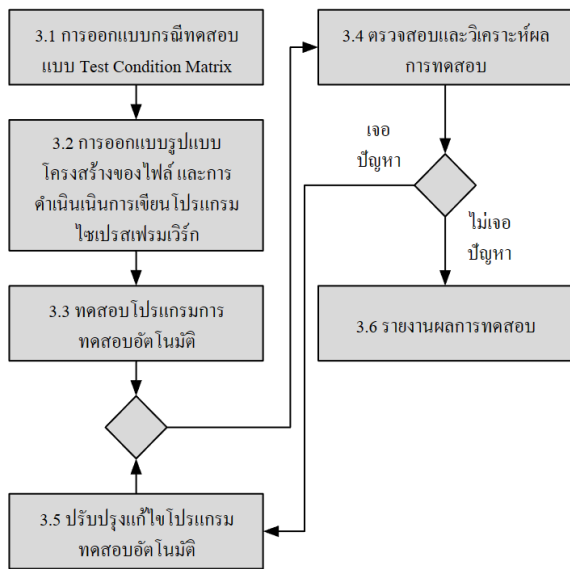
ระบบ myCourseVille Assessment เป็นแพลตฟอร์มการประเมินการเรียนรู้นบนระบบ myCourseVille เนื่องจากวิกฤติโรค COVID-19 ที่ผ่านมามีการเรียนการสอนต้องดำเนินการแบบออนไลน์เท่านั้น และมีการวัดประเมินผู้เรียน ตามวัตถุประสงค์การเรียนรู้ของแต่ละรายวิชา ซึ่งมีผู้เรียนจำนวนมาก ทำให้ต้องการระบบที่สามารถรองรับการใช้งานจำนวนมาก ๆ ได้ จึงทำให้เกิดระบบ myCourseVille Assessment มาช่วยรองรับความต้องการในด้านต่าง ๆ ดังกล่าว ทั้งในส่วนของผู้สอนและส่วนของผู้เรียนที่เข้ารับการประเมิน โดยระบบในฝั่งผู้สอนมีความสามารถในการออกข้อสอบและทำการตรวจข้อสอบอัตโนมัติได้ (Auto Grade) ส่วนในฝั่งผู้เรียนสามารถเข้าไปทำข้อสอบได้ เนื่องจากเป็นระบบที่มีขนาดใหญ่ และมีผู้ใช้งานที่หลากหลาย อีกทั้งยังมีเวลาในการทดสอบที่จำกัด และยังเป็น การทดสอบไปพร้อม ๆ กับทีมที่ยังคงทำการพัฒนาซอฟต์แวร์อยู่อย่างต่อเนื่อง จึงจำเป็นต้องมีกระบวนการทดสอบที่มีมาตรฐาน และมีการออกแบบกรณีทดสอบที่ครอบคลุม อีกทั้งยังต้องมีการทดสอบการถดถอยอย่างสม่ำเสมอ ซึ่งในงานวิจัยนี้จึงได้เลือกใช้การทดสอบแบบอัตโนมัติโดยเลือกใช้ไซเปรสเฟรมเวิร์ก มาใช้ในการทำการทดสอบ

2.10 เฟรมเวิร์ก playwright (Playwright Framework)

เป็นเครื่องมือการทดสอบ End to End ที่สามารถรองรับได้หลายภาษา เช่น JavaScript, TypeScript, Python และ Java โดยสามารถทำการทดสอบแบบหลายเซสชันของเบราว์เซอร์ หรือการทดสอบแบบหลาย ๆ หน้าต่างได้ มีเบราว์เซอร์ที่สามารถเลือกมาใช้ในการทำการทดสอบได้หลากหลาย เช่น Chromium, Firefox, Webkit [11] อีกทั้งยังมีการแก้ปัญหา ที่อาจเกิดการทดสอบที่ผิดพลาด และไม่น่าเชื่อถือ มาให้เรียบร้อยแล้ว เช่น การมี Auto Wait มาให้ โดยที่ไม่ต้องเขียนโปรแกรมรอให้เจอให้เจอองค์ประกอบก่อน แล้วจึงทำการกดปุ่มได้

3. ขั้นตอนวิธีการดำเนินการวิจัย

กระบวนการทดสอบระบบการประเมินออนไลน์ อัตโนมัตินี้ด้วยไซเปรสเฟรมเวิร์ก สำหรับการทดสอบเว็บ แอปพลิเคชันของงานวิจัยนี้ประกอบไปด้วย 6 ขั้นตอน ดัง ภาพที่ 1



ภาพที่ 1: แผนภาพขั้นตอนการดำเนินการวิจัย

3.1 การออกแบบกรณีทดสอบแบบ Test Condition Matrix

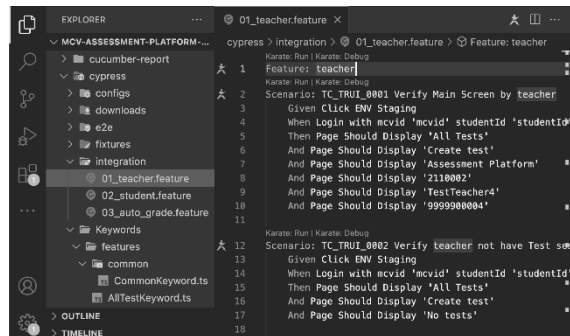
ขั้นตอนนี้เริ่มจากการกำหนดรูปแบบของรหัสลำดับกรณีทดสอบ เพื่อใช้ในการตามรอยกรณีทดสอบย้อนหลัง และใช้ในการระบุกรณีทดสอบในขั้นตอนการสร้างการทดสอบแบบอัตโนมัติ และการรายงานผลการทดสอบ

จากนั้นกำหนดกรณีทดสอบในไฟล์ Excel โดยที่คอลัมน์ที่ 1 คือรหัสของกรณีทดสอบ คอลัมน์ที่ 2 คือชื่อของกรณีทดสอบ คอลัมน์ที่ 3 คือผลลัพธ์ที่คาดหวังว่าจะได้รับจากกรณีทดสอบ คอลัมน์ที่ 4 เป็นต้นไปจะเป็นรายละเอียดของแต่ละหน้าที่ต้องการตรวจสอบความถูกต้อง เช่น ชื่อหน้าเพจส่วนต่อประสานที่เกี่ยวข้อง ที่ต้องการทำการทดสอบ ชื่อชนิดขององค์ประกอบ และชื่อองค์ประกอบ ใช้ตรวจสอบว่าสถานการณ์ของกรณีทดสอบนั้น ๆ มีข้อมูล ปุ่ม หรือ การแจ้งเตือนเป็นไปตามที่กำหนดไว้หรือไม่ คอลัมน์สุดท้ายคือผลลัพธ์ของการทดสอบ ถ้าผลลัพธ์ของการทดสอบมีความสอดคล้องกับ

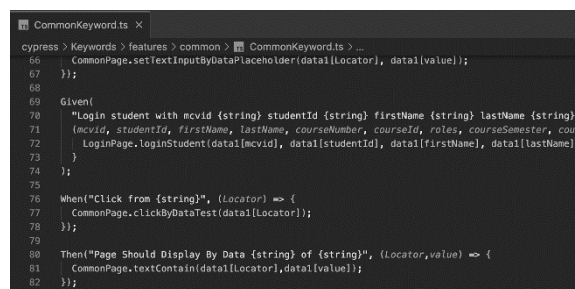
ผลลัพธ์ที่คาดหวังไว้ทั้งหมด จะสามารถระบุผลการทดสอบได้ว่าผ่าน แต่ถ้าต่างกันจะระบุว่าไม่ผ่าน

3.2 การออกแบบรูปแบบโครงสร้างของไฟล์ และการดำเนินการเขียนโปรแกรมไซเปรสเฟรมเวิร์ก

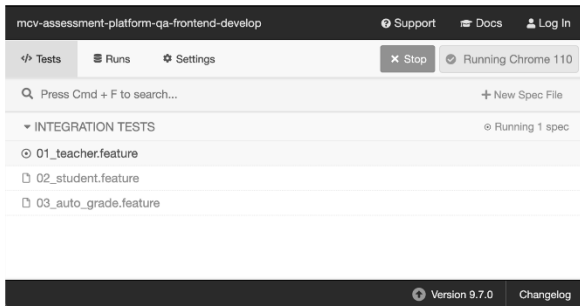
ในขั้นตอนนี้ทำการกำหนดรูปแบบโครงสร้างของไฟล์ เพื่อให้สามารถเขียนโปรแกรมในการทดสอบได้ง่าย ต่อการปรับปรุงแก้ไขเท่านั้น ขั้นตอนการดำเนินการเขียนโปรแกรมไซเปรสเฟรมเวิร์ก เริ่มจากการกำหนด Environment ที่ทำการทดสอบภายในไฟล์ envConfig.json และเริ่มทำการเขียนโค้ดส่วนที่มีลักษณะการทำงานที่สามารถเรียกใช้ได้อบย ๆ และใช้ได้กับทุก ๆ หน้าส่วนต่อประสานที่ต้องการทดสอบ จากนั้นทำการสร้างตัวแปร กำหนดให้กับข้อความ หรือปุ่มกดต่าง ๆ ที่ต้องการตรวจสอบไว้ที่ pagesName.ts โดยมีการตั้งชื่อเปลี่ยนไปตามชื่อของหน้าส่วนต่อประสานนั้น ๆ จากนั้นทำการเขียนโค้ดคำสั่งต่าง ๆ ที่ต้องการทำ หรือตรวจสอบบนหน้าส่วนต่อประสานนั้น ๆ ภายใน featuresName.ts โดยมีการตั้งชื่อเปลี่ยนไปตามชื่อของหน้านั้น ๆ และนำโค้ดที่เขียนจาก featuresName.ts ดังภาพที่ 3 และ 4 มาประกาศใน scenarioName.feature ดังภาพที่ 2 เพื่อเรียงลำดับขั้นตอนตามที่ออกแบบไว้ในกรณีทดสอบ



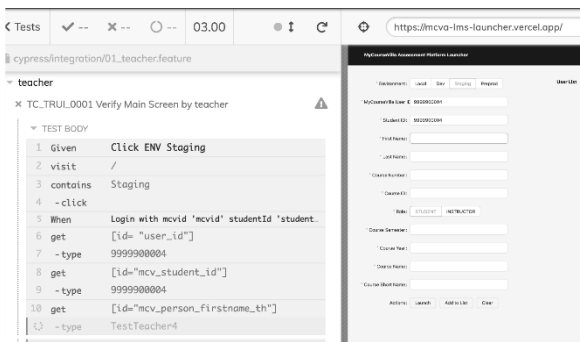
ภาพที่ 2: ตัวอย่าง scenarioName.feature แบบใช้ Gherkin Style



ภาพที่ 3: ตัวอย่าง featuresName.ts



ภาพที่ 4: ตัวอย่าง Cypress GUI App



ภาพที่ 5: ตัวอย่างขณะทำการทดสอบด้วยไซเปรสเฟรมเวิร์ก

3.3 ทดสอบโปรแกรมการทดสอบอัตโนมัติ

ดำเนินการโดยใช้บรรทัดคำสั่ง (Command line) ในการระบุ Environment ที่ต้องการทดสอบ และต่อด้วยคำสั่งการเปิดใช้งานไซเปรสเฟรมเวิร์ก ในการทำการทดสอบ จากนั้นไซเปรสเฟรมเวิร์กจะทำการเปิดหน้าต่างเว็บเบราว์เซอร์ขึ้นมาใหม่ และแสดง Cypress GUI App จากนั้นสามารถเลือกกรณีทดสอบที่ต้องการทดสอบดังภาพที่ 4 พร้อมทั้ง สามารถดูการทำงานของเว็บเบราว์เซอร์ขณะที่เรากำลังทำการทดสอบ ได้ดังภาพที่ 5

3.4 ตรวจสอบและวิเคราะห์ผลการทดสอบ

ขั้นตอนนี้ได้ผลการทดสอบออกมาในรูปแบบไฟล์เอชทีเอ็มแอล ซึ่งสามารถดูขั้นตอนการทดสอบได้อย่างละเอียด เพื่อทำการตรวจสอบและวิเคราะห์ผลการทดสอบได้อย่างถูกต้องว่าโปรแกรมการทดสอบแบบอัตโนมัติด้วยไซเปรสเฟรมเวิร์กที่ได้เขียน และส่วนต่อประสานผู้ใช้งานของเว็บแอปพลิเคชันระบบการประเมินออนไลน์นั้นสามารถการทดสอบทำงานได้ถูกต้องหรือไม่

3.5 ปรับปรุงแก้ไขโปรแกรมทดสอบอัตโนมัติ

เป็นขั้นตอนในการปรับปรุงแก้ไขโปรแกรมทดสอบอัตโนมัติ ที่ต่อเนื่องจากการตรวจสอบ และวิเคราะห์ผลการทดสอบ เพื่อให้มีการทดสอบที่ถูกต้องตามที่คาดหวังไว้ และหลังจากนั้นทำการทดสอบใหม่อีกครั้ง

3.6 รายงานผลการทดสอบ

เป็นขั้นตอนในการรายงานผลการทดสอบ ซึ่งรายงานถูกสร้างมาจาก ปลั๊กอิน Cucumber Reports ให้อยู่ในรูปแบบไฟล์ HTML ซึ่งสามารถเปิดด้วยเว็บเบราว์เซอร์ได้ ภายในรายงานแสดงจำนวนกรณีทดสอบของสถานการณ์ทั้งหมด แสดงจำนวนกรณีทดสอบที่ทำการทดสอบผ่านและไม่ผ่าน มีการแสดงเวลาที่ใช้ในการทดสอบทั้งหมด อีกทั้งยังมีการแสดงรายละเอียดระบบปฏิบัติการ และอุปกรณ์ที่ใช้ในการทดสอบ ซึ่งรายงานนี้ สามารถใช้เป็นการรับประกันซอฟต์แวร์ได้ว่ามีคุณภาพก่อนส่งมอบให้กับผู้ใช้ได้ใช้งานจริง

4. ผลการดำเนินการวิจัย

4.1 เอกสารกรณีทดสอบของระบบการประเมินออนไลน์อัตโนมัติในรูปแบบของ Test Condition Matrix

เอกสารที่ใช้ในการออกแบบกรณีทดสอบแบบ Test Condition Matrix โดยที่สถานการณ์ของกรณีทดสอบที่อยู่ในสถานการณ์เดียวกันจะถูกรวมกลุ่มไว้อยู่ด้วยกัน โดยมีลักษณะดังภาพที่ 6

Test Case ID	Test Case Name	Expected Result	Main		Test Session		Result							
			Message	button	Message	button								
			All Tests	No test	Create test	Tap create new	Add instructions	Add description	Name	Student ID	Department	Under part		
TC_TRUI_0002	Verify teacher not have Test session	The User sees No test text The User sees Create test button		x	x									PASS
TC_TRUI_0003	Verify teacher when Create test by Add essay question (Answer Type: text)	The User Click Create test button The User sees Add instructions text The User input Add instructions The User tick Signature agreement The User sees Add description text The User input Add description The User click Add essay question The User sees Question text The User input Question The User tick Answer Type(Text) The User sees and input Points The User click Preview The User sees and Preview text The User sees Answer type: Text The User click Create			x		x	x	x	x	x	x		PASS
TC_TRUI_0004	Verify teacher when	The User Click Create test button			x		x	x	x	x				

ภาพที่ 6: ตัวอย่างของเอกสาร Test Condition Matrix

4.2 รายงานผลลัพธ์ของการทดสอบแบบอัตโนมัติ

ด้วยไซเปรสเฟรมเวิร์ก

รายงานผลการทดสอบจากไซเปรสเฟรมเวิร์กถูกสร้างจาก Cucumber Reports อยู่ในรูปแบบไฟล์ HTML โดยมีการแสดงเวลาที่ใช้ในการดำเนินการทดสอบ ทั้งที่แบ่งไปแต่ละสถานการณ์ และของสถานการณ์ทั้งหมด จากการดำเนินการทดสอบระบบการประเมินออนไลน์ทั้งฟังก์ชันในฝั่งผู้สอน และฝั่งผู้เรียน โดยการทดสอบอัตโนมัติใช้เวลาทั้งหมด 3 นาที 55 วินาที แสดงดังตารางที่ 1 จาก 16 กรณีทดสอบที่ได้ทำการออกแบบไว้ในรูปแบบของ Test Condition Matrix โดยได้ผลลัพธ์จากการทดสอบว่าผ่านทั้งหมด 100 เปอร์เซ็นต์

นอกจากนี้ ผู้วิจัยได้ทดสอบเปรียบเทียบ ระยะเวลาในการทดสอบกับ เฟลโรทเฟรมเวิร์ก ซึ่งผลที่ได้ไม่มีความแตกต่างในเรื่องของระยะเวลาที่ใช้ อย่างมีนัยสำคัญ โดยต่างกัน 1 ถึง 2 วินาที เนื่องจากพบว่าระยะเวลาที่ใช้ทดสอบขึ้นกับความเร็วในการโหลดหน้าต่าง เพราะถ้าหากมีการโหลดหน้าต่างใหม่ซ้ำ ส่งผลให้การทดสอบต้องรอกดปุ่มหรือตรวจสอบค่าเพื่อดำเนินการทดสอบต่อไปจึงส่งผลต่อระยะเวลาในการดำเนินการทดสอบ รวมถึงจำนวนขั้นตอนที่มีจำนวนมากทำให้แต่ละกรณีการทดสอบใช้เวลานานขึ้น สามารถแก้ไขได้โดยการทำการทดสอบโดยเขียนโปรแกรมให้มีการทำงานอยู่เบื้องหลังในลักษณะที่เป็นโปรแกรมส่วนกลางทำงานก่อนที่การทดสอบต้องมาเรียกใช้ เช่น การเข้าสู่ระบบ เพื่อลดระยะเวลาในการทดสอบระบบ และยังเป็น การเพิ่มประสิทธิภาพในด้านความเร็วของเฟรมเวิร์กที่ใช้ในการดำเนินการทดสอบนั้น ๆ ด้วย

ตารางที่ 1: ระยะเวลาที่ใช้ในการทดสอบด้วยไซเปรสเฟรมเวิร์ก

Test Feature	Time Duration
Teacher	2 minutes 12 seconds
Student	1 minutes 15 seconds
Auto Grade	0 minutes 28 seconds
Total	3 minutes 55 seconds

5. สรุป

งานวิจัยนี้ได้ทำการทดสอบระบบการประเมินออนไลน์แบบอัตโนมัติ ด้วยไซเปรสเฟรมเวิร์กสำหรับการทดสอบ ส่วนต่อประสานผู้ใช้งานของเว็บแอปพลิเคชัน เพื่อให้เป็นกรณีศึกษาในการช่วยลดระยะเวลาในการทำการทดสอบแบบถดถอย ซึ่งทำให้สามารถส่งมอบเว็บแอปพลิเคชันของระบบการประเมินออนไลน์ให้ผู้ใช้งานได้ใช้งานได้อย่างรวดเร็ว และยังช่วยลดความผิดพลาดของการทำการทดสอบที่เกิดจากมนุษย์ในระหว่างทำการทดสอบ ซึ่งได้ผลลัพธ์ของการทดสอบออกมาเป็นรายงานเพื่อใช้ในการประกันคุณภาพของเว็บแอปพลิเคชัน ของระบบการประเมินออนไลน์ได้ จากการทดสอบพบว่าระยะเวลาในการดำเนินการทดสอบทั้งระบบใช้เวลาเพียง 3 นาที 55 วินาที ซึ่งเวลาขึ้นอยู่กับการออกแบบกรณีทดสอบ และจำนวนกรณีทดสอบในรอบนั้น ๆ และจากการทำการทดสอบที่ได้พบว่าระยะเวลาที่ใช้ทดสอบนั้นยังขึ้นกับความเร็วในการโหลดหน้าต่างนั้น หากพบว่ามี การโหลดหน้าต่างที่ใช้เวลานาน จะส่งผลกระทบต่อเวลาที่ใช้ในการดำเนินการทดสอบ

เอกสารอ้างอิง

- [1] “Why Cypress?” [online]. Available: <https://docs.cypress.io/guides/overview/why-cypress#Our-mission>. [Accessed: 25-Oct-2022].
- [2] “What is END-To-END Testing? E2E Example” [online]. Available: <https://www.guru99.com/end-to-end-testing.html>. [Accessed: 11-Feb-2023].
- [3] “Updating the test case document: Test conditions matrix (TestCondMatrix)” [online]. Available: https://help.hcltechsw.com/commerce/7.0.0/com.ibm.commerce.starterstores.doc/tasks/tsm_TestCondMatrix.html. [Accessed: 25-Oct-2022].
- [4] “Reporting” [online]. Available: <https://cucumber.io/docs/cucumber/reporting/?lang=java>. [Accessed: 25-Oct-2022].
- [5] Elis Pelivani, Adrian Besimi, Betim Cico. “A Comparative Study of UI Testing Framework” 11th Mediterranean Conference on Embedded Computing (MECO), 2022

- [6] Nikhil Govil, Ashish Sharma. “A Game Plan to Build Optimized Regression Testing in Agile Methodologies Using Test Prioritization” 5th International Conference on Information Systems and Computer Networks (ISCON), 2021
- [7] Yi Zhao, Yun Hu, Jiayu Gong. “Research on International Standardization of Software Quality and Software Testing” IEEE/ACIS 20th International Fall Conference on Computer and Information Science (ICIS Fall), 2021
- [8] ISO/IEC/IEEE. Software and systems engineering—Software testing—Part 2: Test processes: ISO/IEC/IEEE 29119-2:2013[S]. Switzerland:ISO,2013.
- [9] “myCourseVille” [online]. Available: https://lic.chula.ac.th/?page_id=4900. [Accessed: 25-Oct-2022].
- [10] “Gherkin Reference” [online]. Available: <https://cucumber.io/docs/gherkin/reference/>. [Accessed: 24-Feb-2023].
- [11] “Playwright” [online]. Available: <https://playwright.dev/>. [Accessed: 25-Feb-2023].

การพัฒนาฟังก์ชันมาตรฐานรักษาความปลอดภัยของข้อมูลส่วนบุคคล ในระบบการวางแผนทรัพยากรองค์กร

The Development of Personal Data Protection Act Function in Enterprise Resource Planning System

ชไมพร ปิสิริ (Chamaiporn Peesiri)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857034@email.kmutnb.ac.th, ²sakchai.t@itd.kmutnb.ac.th

บทคัดย่อ

การวิจัยนี้ได้มีการพัฒนาฟังก์ชันมาตรฐานรักษาความปลอดภัยของข้อมูลส่วนบุคคลในระบบการวางแผนทรัพยากรองค์กร โดยทำการเก็บรวบรวมข้อมูลความต้องการจากผู้ใช้งานผ่านการสัมภาษณ์เชิงลึก เพื่อนำข้อมูลมาทำการวิเคราะห์และออกแบบฟังก์ชันมาตรฐานของระบบให้สอดคล้องกับมาตรฐานรักษาความปลอดภัยของข้อมูลส่วนบุคคล ซึ่งระบบที่พัฒนาขึ้นจะต้องตอบสนองความต้องการของผู้ใช้งาน มีความถูกต้องและเหมาะสม สามารถรองรับมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคลได้อย่างมีประสิทธิภาพ ตรงตามเป้าหมายของธุรกิจ ผลการประเมินความเหมาะสมของระบบที่มีการพัฒนาขึ้น ได้มีการจำลองสถานการณ์ธุรกิจและแบ่งแยกหน้าที่งานให้กับผู้เชี่ยวชาญจำนวน 5 ท่าน สำหรับการทำการทดสอบและประเมินความเหมาะสมของระบบที่ได้พัฒนาขึ้น ซึ่งผลการประเมินความเหมาะสมของระบบที่ได้พัฒนา มีค่าเฉลี่ยอยู่ที่ 4.78 และมีค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.42 สามารถสรุปได้ว่าความเหมาะสมของระบบที่ได้พัฒนาค่ามากกว่าระดับ 4.0 และอยู่ในระดับดีและผลการประเมินความพึงพอใจของผู้ใช้งานระบบ โดยได้รับความร่วมมือจากผู้ใช้งานจำนวน 6 ท่าน ในการประเมินความพึงพอใจของระบบที่ได้พัฒนาขึ้นเปรียบเทียบกับความต้องการของผู้ใช้งาน ซึ่งผลการประเมินความพึงพอใจของระบบที่ได้พัฒนา มีค่าเฉลี่ยอยู่ที่ 4.20 และมีค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่

0.88 สามารถสรุปได้ว่าความพึงพอใจในระบบที่ได้พัฒนามีค่ามากกว่าระดับ 4.0 ซึ่งอยู่ในระดับดี

คำสำคัญ: ระบบการวางแผนทรัพยากรองค์กร การรักษาความปลอดภัยของข้อมูลส่วนบุคคล เอสเอพี

Abstract

This research has developed the personal data security standard function in the enterprise resource planning system. By collecting information from users through in-depth interviews. To bring the information to analyze and design the standard functions of the system in accordance with the security measures of personal information. The developed system must meet the needs of users. It is accurate and appropriate and can effectively support the security measures of personal information. meet business goals. The results of the evaluation of the suitability of the developed system. Business situations were simulated, and tasks were assigned to 5 experts for testing and evaluating the suitability of the developed system. The results of the evaluation of the suitability of the developed system the average value was 4.78 and the standard deviation was 0.42. It can be concluded that the suitability of the developed system was greater than the level of 4.0 and was at a good level and the result of assessing the satisfaction of the users of the system. By getting cooperation from 6 users in

evaluating the satisfaction of the developed system compared to the needs of users. The satisfaction assessment results of the developed system the mean is 4.20 and the standard deviation is 0.88. It can be concluded that the satisfaction with the developed system is greater than the 4.0 level, which is a good level. Then the 4.0 level, which is a good level.

Keywords: Enterprise Resource Planning System, Personal Data Protection Act, SAP.

1. บทนำ

ปัจจุบันระบบการวางแผนทรัพยากรองค์กรจึงได้มีบทบาทสำคัญต่อการดำเนินงานในองค์กรเป็นอย่างมาก เนื่องจากเป็นระบบที่ทำให้ผู้ใช้งานสามารถทำงานได้รวดเร็ว ลดความผิดพลาดในการทำงานและจัดเก็บข้อมูลอยู่ในฐานข้อมูลเดียวกัน ทำให้สามารถดึงข้อมูลไปใช้งานเป็นแบบทันทีและมีความน่าเชื่อถือมากยิ่งขึ้น โดยองค์กรจะต้องมีการจัดเก็บ การบันทึกข้อมูลต่างๆ เช่น ข้อมูลของลูกค้า ประวัติการซื้อขาย ซึ่งข้อมูลดังกล่าว ที่เป็นข้อมูลส่วนบุคคลจะต้องได้รับการยินยอมจากเจ้าของข้อมูลและห้ามเปิดเผยแก่บุคคลภายนอกหรือสาธารณะ โดยเด็ดขาด เพื่อเป็นการรักษาความปลอดภัยของข้อมูลส่วนบุคคลที่จัดเก็บในฐานข้อมูลขององค์กร เช่น ชื่อ - นามสกุล เลขที่บัญชีธนาคาร เป็นต้น จากประเด็นดังกล่าวผู้วิจัยจึงได้มีความคิดในการพัฒนาระบบการวางแผนทรัพยากรองค์กร โดยได้มีการพัฒนาผ่านฟังก์ชันมาตรฐานของระบบ SAP Business One Version 10.0 ตามข้อมูลที่ได้จากการรวบรวมความต้องการผ่านการสัมภาษณ์เชิงลึกจากผู้เชี่ยวชาญและผู้ใช้งานระบบเกี่ยวกับการจัดเก็บ แก้ไข หรือความต้องการลบข้อมูลของเจ้าของข้อมูล ขั้นตอนการทำงานเพื่อนำมาข้อมูลมาทำการวิเคราะห์และออกแบบระบบผ่านฟังก์ชันมาตรฐานของระบบ มุ่งเน้นการรักษาความปลอดภัยของข้อมูล ตอบสนองความต้องการของผู้ใช้งานและเพื่อให้ระบบมีประสิทธิภาพในการทำงาน

สามารถรองรับมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคล

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การพัฒนาระบบสารสนเทศ

วิธีการพัฒนาระบบสารสนเทศ คือ ขั้นตอนกิจกรรมและปัจจัยที่มีผลต่อการพัฒนาระบบ โดยมุ่งเน้นการศึกษา“วงจรการพัฒนาระบบ”เพื่อนำมาใช้ในการวางแผนขั้นตอนการวิเคราะห์และออกแบบระบบสารสนเทศให้ปลอดภัยและมีประสิทธิภาพที่สุด [1]

2.2 PDAPA

กฎหมาย PDPA ว่าด้วยการคุ้มครองข้อมูลส่วนบุคคลขึ้นเพื่อกำหนดหลักเกณฑ์ กลไก หรือมาตรการกำกับดูแลเกี่ยวกับการให้ความคุ้มครองข้อมูลส่วนบุคคลที่รวมถึงการเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคลขึ้น โดยข้อมูลส่วนบุคคล คือ ข้อมูลเกี่ยวกับบุคคลซึ่งทำให้สามารถระบุตัวบุคคลนั้นได้ ไม่ว่าทางตรงหรือทางอ้อม [2], [3] สามารถจำแนกได้ 2 ประเภท คือ ข้อมูลส่วนบุคคลทั่วไป ได้แก่ ชื่อ-นามสกุล ที่อยู่ และข้อมูลส่วนบุคคลที่มีความละเอียดอ่อน เช่น ข้อมูลทางการแพทย์หรือสุขภาพ ประวัติอาชญากรรม ซึ่งข้อมูลอ่อนไหวถือเป็นข้อมูลสูงเสี่ยงต่อการถูกใช้ในการเลือกปฏิบัติอย่างไม่เป็นธรรม ดังนั้น เพื่อลดปัญหาดังกล่าวจึงควรมีการจำกัดสิทธิการมองเห็นข้อมูลเฉพาะผู้ที่เกี่ยวข้องหรือได้รับอนุญาตเท่านั้น เช่น การป้องกันการมองเห็นข้อมูลโดยแสดงเป็นสัญลักษณ์หรือการเข้ารหัสแทนข้อมูลจริงในระบบต่างๆ เพื่อเป็นการรักษาความลับและความปลอดภัยของข้อมูลส่วนบุคคล

2.3 SAP Business One

โปรแกรม SAP Business One คือ ซอฟต์แวร์ ERP สำหรับธุรกิจขนาดเล็กและขนาดกลาง โดย SAP เป็นผู้พัฒนาซอฟต์แวร์ ERP ที่มีชื่อเสียงระดับโลก และยังเป็น ERP Software ที่เชื่อมโยงทุกการทำงานของหน่วยงานต่าง ๆ ภายในองค์กรเข้าด้วยกันอย่างมีประสิทธิภาพและยืดหยุ่นสูง ครอบคลุมทุก Business Process และมีการพัฒนาให้สามารถรองรับการรักษา

ความปลอดภัยของข้อมูลส่วนบุคคลตามกฎหมายของสหภาพยุโรปในการจัดเก็บ แก้ไข มองเห็นข้อมูลของผู้ใช้งานระบบ [4]

2.4 GDPR

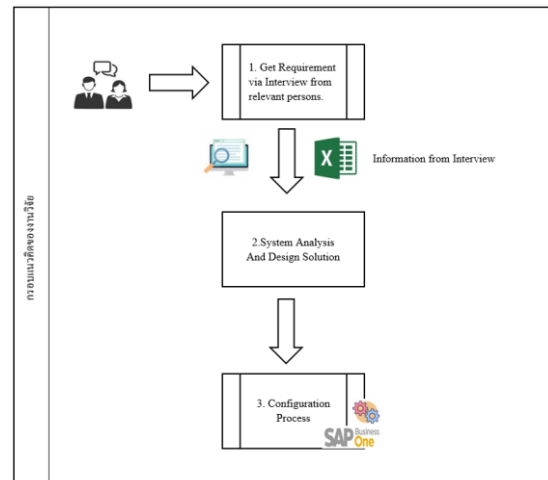
กฎหมายคุ้มครองข้อมูลส่วนบุคคล หรือ General Data Protection Regulation : GDPR คือ บริษัทธุรกิจที่จัดเก็บและจัดการข้อมูลส่วนบุคคลของพลเมืองสหภาพยุโรปจะต้องเพิ่มมาตรการปกป้องข้อมูลต่างๆ นอกจากนี้ในกฎหมายกำหนดไว้ว่า ข้อมูลเหล่านี้จะไม่สามารถนำไปใช้ในเชิงพาณิชย์ได้ หากไม่ได้รับความยินยอมจากเจ้าของข้อมูล การละเมิดกฎระเบียบนี้อาจถูกปรับเป็นจำนวนเงินค่อนข้างสูง และกฎหมายนี้จะมีผลบังคับใช้ปกป้องข้อมูลพลเมืองสหภาพยุโรปไม่ว่าจะอยู่ที่ใด [5]

3. วิธีการดำเนินการวิจัย

3.1 กรอบแนวคิดของงานวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงประยุกต์โดยใช้ระบบการวางแผนทรัพยากรองค์กร คือ โปรแกรม SAP Business One มาใช้สำหรับการศึกษาและพัฒนาระบบให้สามารถรองรับการรักษาความปลอดภัยของข้อมูลตามมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคลด้วยการศึกษาถึงหลักการ ทฤษฎี แนวทางและข้อควรปฏิบัติในการจัดเก็บข้อมูล การกำหนดสิทธิตามหน้าที่ รวมถึงไปถึงศึกษาถึงปัญหาต่างๆ เกี่ยวกับการจัดเก็บ แก้ไข หรือลบข้อมูล การวิเคราะห์ถึงวิธีการตรวจสอบที่ควรจะเป็น ซึ่งข้อมูลต่าง ๆ เป็นความคิดเห็นและความต้องการ (Requirement) จากกลุ่มผู้ใช้งานและผู้เชี่ยวชาญผ่านการสัมภาษณ์เชิงลึกด้านการใช้งาน ความเหมาะสม การรักษาความปลอดภัยของระบบเดิม ความต้องการพัฒนาจากระบบใหม่ขึ้น และได้นำข้อมูลทั้งหมดที่ได้มาทำการวิเคราะห์และออกแบบระบบ โดยการพัฒนาระบบการวางแผนทรัพยากรองค์กรให้ตรงต่อความต้องการและได้โซลูชันสำหรับการแก้ไขปัญหาดังกล่าว เพื่อเพิ่มประสิทธิภาพและความเหมาะสมในการทำงานของระบบและความพึงพอใจแก่ผู้ใช้งาน โดยสามารถสรุปภาพรวมของการพัฒนาระบบการวางแผนทรัพยากร

องค์กร ได้ดังภาพที่ 1 กรอบแนวคิดของงานวิจัย



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

3.2 ศึกษาข้อมูลและเครื่องมือในการพัฒนาระบบ

การศึกษาข้อมูลและเครื่องมือในการพัฒนาระบบ ได้แก่ ศึกษาเทคนิค ขั้นตอน กระบวนการทำงานทางธุรกิจ การวิเคราะห์และออกแบบ เครื่องมือที่ใช้ในการพัฒนา การจัดเก็บข้อมูล ฟังก์ชันการทำงานในระบบ การวางแผนทรัพยากรองค์กรที่เกี่ยวข้อง โดยได้มีศึกษาจากข้อมูลที่เกี่ยวข้อง ซึ่งประกอบด้วย

3.2.1 มาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคล โดยทำการศึกษาถึงข้อกำหนด กฎระเบียบ รวมไปถึงข้อบังคับ กระบวนการ บุคคลที่เกี่ยวข้อง

3.2.2 คู่มือการใช้งานและการตั้งค่าระบบ SAP Business One Version 10.0 ที่พัฒนามาเพื่อรองรับกับมาตรการของกฎหมายการรักษาความปลอดภัยของข้อมูลส่วนบุคคลสำหรับประเทศทางสหภาพยุโรป ดังนั้นผู้วิจัยจึงทำการศึกษาข้อมูลการตั้งค่า การใช้งาน การป้องกันหรือการลบข้อมูล การพิมพ์รายงานสถานะ จัดเก็บข้อมูล ตามมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคลประเทศไทย

3.2.3 การใช้งานโปรแกรม Microsoft SQL Server Studio และการเขียนภาษา SQL Language สำหรับทดสอบการตรวจสอบในการเข้าถึงฐานข้อมูลด้วยใช้คำสั่ง Query สำหรับเรียกข้อมูลที่มีการป้องกันไว้

3.3 รวบรวมข้อมูลที่ต้องการเกี่ยวกับการพัฒนาฟังก์ชันมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคล

การเก็บรวบรวมข้อมูลต่างๆ หลักการและการพัฒนาระบบ รวมถึงปัญหาสถานการณ์ที่เคยมองเกี่ยวกับความเสี่ยง ช่องโหว่ ความต้องการของผู้ใช้งานและผู้เชี่ยวชาญผ่านการสัมภาษณ์เชิงลึก ให้เป็นไปมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคลในการจัดเก็บแก้ไข หรือการลบและการตรวจสอบการเปลี่ยนแปลงการจำกัดสิทธิการเข้าถึงข้อมูลตามหน้าที่งานที่ได้รับมอบหมายเท่านั้น ขณะเดียวกันระบบที่พัฒนาขึ้นจะต้องมีความเหมาะสมกับการทำงานของผู้ใช้งานแต่ละท่านแยกตามหน้าที่งาน โดยสามารถสรุปได้ดังตารางที่ 1 สรุปความต้องการจากการสัมภาษณ์เชิงลึก

ตารางที่ 1: สรุปความต้องการจากการสัมภาษณ์เชิงลึก

ลำดับ	ปัญหาที่พบในระบบเดิม	ความต้องการระบบที่พัฒนาขึ้น
1	ระบบเดิมไม่ได้จำกัดการมองเห็นข้อมูลแยกตามหน้าที่งาน โดยจะสามารถมองเห็นข้อมูลผู้ประกอบการและออกคำสั่งชำระทั้งหมด ไม่มีการจำกัดสิทธิ์แยกกันระหว่างกลุ่มลูกค้าให้เห็นเฉพาะผู้ขายและกลุ่มเจ้าหน้าที่เห็นเฉพาะฝ่ายบัญชี	ต้องการให้ระบบสามารถจำกัดสิทธิการมองเห็นที่ชัดเจน โดยแบ่งแยกตามหน้าที่งาน 1. พนักงานขาย ให้มองเห็นเฉพาะข้อมูลลูกค้าเท่านั้น 2. พนักงานจัดซื้อ ให้แสดงเฉพาะข้อมูลเจ้าหน้าที่เท่านั้น
2	ระบบเดิมไม่ได้กำหนดแบ่งแยกหน้าที่งานตามฟังก์ชันในระบบตามหน้าที่งานที่ได้รับมอบหมาย เนื่องจากเพื่อให้สอดคล้องต่อการใช้งานของผู้ใช้งาน	ระบบที่พัฒนาขึ้นจะต้องสามารถกำหนดสิทธิการมองเห็นแยกตามหน้าที่งานที่ได้รับมอบหมายและเหมาะสมกับการทำงานของผู้ใช้งาน มีความสอดคล้องกับการทำงาน
3	ระบบเดิมยังไม่มีกรนำมาตรวจรักษาความปลอดภัยมาใช้ภายในองค์กร เช่น การ บล็อกไม่ให้มีข้อมูลไปทำรายการหรือการพิมพ์รายงานให้กับเจ้าของข้อมูล	ต้องการให้ระบบมีฟังก์ชันการรักษาความปลอดภัยของข้อมูลส่วนบุคคล ในการจำกัดสิทธิการบล็อคข้อมูลหรือการลบข้อมูลออกจากระบบตามความต้องการของเจ้าของข้อมูล อีกทั้งต้องการรายงานสำหรับแสดงสถานะข้อมูลส่วนบุคคลปฏิบัติการทำการในระบบ เพื่อพิมพ์รายงานส่งมอบให้แก่เจ้าของข้อมูล
4	ระบบเดิมไม่สามารถตรวจสอบประวัติการแก้ไข เปลี่ยนแปลงข้อมูลต่างๆ ได้	ต้องการให้ระบบสามารถตรวจสอบประวัติการเปลี่ยนแปลง แก้ไข ทั้งประวัติการซื้อขายและข้อมูลผู้ประกอบการในระบบ
5	ระบบเดิมไม่มีการตั้งค่าป้องกัน ล็อกหน้าจอเมื่อผู้ใช้งานไม่ได้ใช้งานระบบ และกำหนดการเปลี่ยนรหัสผ่านอัตโนมัติ	ต้องการให้ระบบมีการป้องกันด้านกรแจ้งเตือนเปลี่ยนรหัสผ่านอัตโนมัติและล็อกหน้าจอเมื่อผู้ใช้งานไม่ได้ใช้งานมากกว่าระบบ 30 นาที

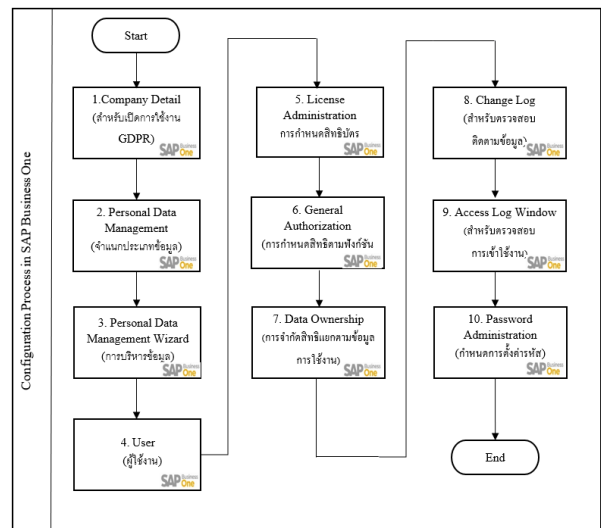
3.4 การวิเคราะห์และออกแบบระบบ

การศึกษาและพัฒนาระบบผ่านโปรแกรม SAP Business One ที่ได้มีการพัฒนาฟังก์ชัน GDPR สำหรับรองรับการรักษาความปลอดภัยข้อมูลส่วนบุคคลสหภาพยุโรป เพื่อนำมาพัฒนาต่อยอดให้รองรับมาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคลของประเทศไทย โดยการใช้แผนภาพกระแส กระบวนการทางธุรกิจและการพัฒนาระบบตามขั้นตอน SDCL รวมไปถึงพัฒนาผ่านฟังก์ชันการทำงาน การตั้งค่าและการกำหนดสิทธิ

ข้อจำกัดของระบบ ให้เหมาะสมกับหน้าที่งานที่ได้รับมอบหมายตามความต้องการของผู้ใช้งาน ผ่านการจัดเก็บข้อมูลระดับฐานข้อมูล ที่ได้มีการจัดเก็บบนโปรแกรม My SQL ซึ่งข้อมูลดังกล่าวจะต้องสามารถจำกัดสิทธิการมองเห็นได้ทั้งหมดมองผู้ใช้งานจากระบบและผู้มีความรู้ด้านภาษา SQL ที่เรียกข้อมูลจากฐานข้อมูล ให้แสดงผลเป็น Encode ในการแสดงผลข้อมูลอ่อนไหวที่ต่างๆ

3.5 การพัฒนาระบบ

การพัฒนาระบบการวางแผนทรัพยากรองค์กร โดยพัฒนาจากฟังก์ชันมาตรฐานของระบบ ในการจัดเก็บ แก้ไข หรือความต้องการลบ หรือการตรวจสอบติดตามการเปลี่ยนแปลงของข้อมูล การจำกัดสิทธิการเข้าถึงของผู้ใช้งานแต่ละบุคคลแยกตามหน้าที่งาน โดยสามารถสรุปฟังก์ชันการตั้งค่าที่เกี่ยวข้องได้ ดังภาพที่ 2



ภาพที่ 2: ขั้นตอนการตั้งค่าพื้นฐาน (Configuration)

3.6 การทดสอบและประเมินผลของระบบหลังพัฒนา

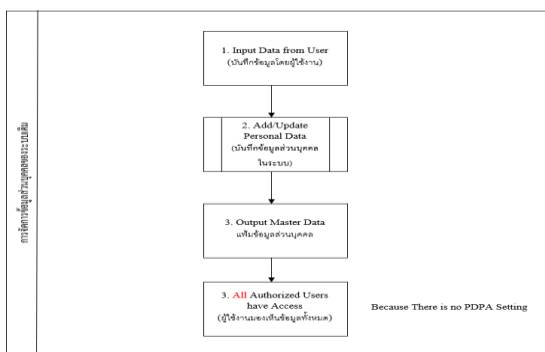
การทดสอบได้มีการจำลองสถานการณ์จริงขึ้นมา โดยผู้เชี่ยวชาญและผู้ใช้งานแต่ละท่านจะได้รับบทบาทหน้าที่ตามหน้าที่งานตามความเหมาะสม เพื่อทำการทดสอบระบบตามหน้าที่งานที่ได้รับมอบหมายภายในระยะเวลา 30 นาทีและหลังจากทดสอบระบบเรียบร้อยแล้ว ผู้เชี่ยวชาญและผู้ใช้งานจะดำเนินการตอบแบบสอบถามออนไลน์ เพื่อการประเมินความเหมาะสมและความพึงพอใจของระบบที่ได้พัฒนาขึ้น โดยเป็นการประเมินเปรียบเทียบกับความต้องการเพิ่มจากระบบเดิมดังตารางที่ 1 สรุปความต้องการจากการสัมภาษณ์เชิงลึก

3.7 สรุปผลการวิจัย

หลังจากที่ได้ดำเนินการทดสอบแล้ว ผู้เชี่ยวชาญ และผู้ใช้งานจะทำการประเมินระบบที่ได้พัฒนา ซึ่งผู้วิจัย จะทำการสรุปผลลัพธ์ของงานวิจัย โดยเป็นการประเมิน ด้านความเหมาะสมและความพึงพอใจของระบบที่ได้มีการวิเคราะห์และออกแบบผ่านฟังก์ชันมาตรฐานของระบบการวางแผนทรัพยากรองค์กรให้สามารถรองรับ มาตรการรักษาความปลอดภัยของข้อมูลส่วนบุคคล ในการจัดเก็บ แก้ไข การลบ การตรวจสอบประวัติการเปลี่ยนแปลงข้อมูล เพื่อให้ได้ระบบการวางแผน ทรัพยากรองค์กรที่มีความเหมาะสมตรงต่อความต้องการ ของผู้ใช้งานและสามารถรักษาความปลอดภัยของข้อมูล ส่วนบุคคลได้อย่างมีประสิทธิภาพ

4. ผลการดำเนินงาน

จากการวิจัยการพัฒนาฟังก์ชันมาตรการรักษา ความปลอดภัยของข้อมูลส่วนบุคคลในระบบการ วางแผนทรัพยากรองค์กร สามารถเปรียบเทียบความ แตกต่างของการจัดเก็บและการเข้าถึงข้อมูลส่วนบุคคล ได้ดังภาพที่ 3 การจัดการข้อมูลส่วนบุคคลในระบบเดิม และภาพที่ 4 การจัดการข้อมูลส่วนบุคคลในระบบที่ พัฒนาขึ้น โดยมีรายละเอียด ดังนี้

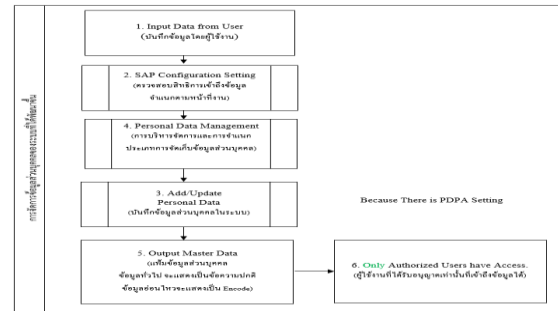


ภาพที่ 3: การจัดการข้อมูลในระบบเดิม

จากภาพที่ 3 การจัดการข้อมูลส่วนบุคคลในระบบเดิม พบว่าเมื่อผู้ใช้งานทำการบันทึกข้อมูลเข้าสู่ระบบ ระบบจะทำการบันทึก/แก้ไขข้อมูลส่วนบุคคลทันที โดยผู้ใช้งานทุกคนจะสามารถเข้าไปเพิ่ม แก้ไข ลบ

ข้อมูลหรือสามารถนำข้อมูลส่วนบุคคลไปทำธุรกรรมอื่น ได้นอกเหนือหน้าที่งานที่รับผิดชอบ

ผู้วิจัยจึงได้ทำการวิเคราะห์และออกแบบระบบที่ พัฒนาขึ้นสามารถบริหารจัดการข้อมูลส่วนบุคคลให้ เหมาะสมและรองรับมาตรการรักษาความปลอดภัยของ ข้อมูลส่วนบุคคล ดังภาพที่ 4



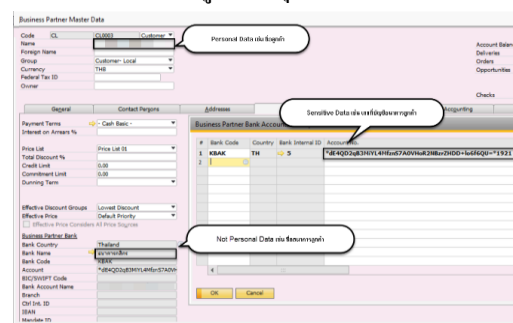
ภาพที่ 4: การจัดการข้อมูลในระบบที่พัฒนาขึ้น

จากภาพที่ 4 การจัดการข้อมูลส่วนบุคคลในระบบ ที่พัฒนาขึ้น พบว่าเมื่อผู้ใช้งานได้มีการบันทึก/แก้ไขข้อมูล ส่วนบุคคลเข้าสู่ระบบ ระบบจะทำการตรวจสอบสิทธิ การเข้าถึงข้อมูลส่วนบุคคลแยกตามหน้าที่งานที่ได้ที่ ได้ มีการกำหนดไว้ จากนั้นระบบจะทำการตรวจสอบการ บริหารจัดการข้อมูลส่วนบุคคล ตามประเภทการจัดเก็บ และการแสดงผลให้กับผู้ใช้งานแต่ละราย จึงทำการ บันทึกและแก้ไขข้อมูลดังกล่าว โดยระบบจะแสดงผล ให้เห็นเฉพาะผู้ใช้งานที่ได้รับอนุญาตเท่านั้น สามารถสรุป การแสดงผลส่วนบุคคลแต่ละประเภท ได้ดังตารางที่ 2

ตารางที่ 2: การแสดงประเภทข้อมูลส่วนบุคคลในระบบ

ประเภทข้อมูล	การแสดงผล
ข้อมูลส่วนบุคคล (Personal Data)	ข้อความ
ข้อมูลอ่อนไหว (Sensitive Data)	Encode
ข้อมูลทั่วไป (Not Personal)	ข้อความ

ตัวอย่างการแสดงผลข้อมูลส่วนบุคคลในระบบที่พัฒนาขึ้น



ภาพที่ 4: ตัวอย่างการแสดงผลข้อมูลในระบบที่พัฒนาขึ้น

จากการสรุปผลแบบประเมินการพัฒนาฟังก์ชัน มาตราการรักษาความปลอดภัยของข้อมูลส่วนบุคคลในระบบการวางแผนทรัพยากรองค์กรหลังจากสามารถแยกผลการประเมินออกเป็น 2 ส่วน ดังนี้

ตารางที่ 3 ผลการประเมินความเหมาะสมของระบบ

รายการประเมิน	ผลการประเมิน		
	X	S.D.	ระดับ
1. ความเหมาะสมในการกำหนด การตั้งค่า การจัดเก็บข้อมูลส่วนบุคคล	4.8	0.45	ดี
2. ความเหมาะสมในการตั้งค่าระบบ สำหรับจำกัดสิทธิการใช้งาน	4.4	0.55	ดี
3. ความเหมาะสมในการมองเห็น การใช้งานข้อมูลส่วนบุคคลและในการทำงาน	4.8	0.45	ดี
4. ความเหมาะสมด้านการจำกัดสิทธิ การเพิ่ม แก้ไข หรือการลบข้อมูลส่วนบุคคล	5	0	ดี
5. ความเหมาะสมของข้อมูลที่แสดงในรายงานข้อมูลส่วนบุคคล (Personal Data Report)	4.6	0.55	ดี
6. ความเหมาะสมของข้อมูลส่วนบุคคลในระบบ ในการจัดการปิดกั้น เปลี่ยนสถานะตามความต้องการของเจ้าของข้อมูล	4.8	0.45	ดี
7. ความเหมาะสมในการตรวจสอบ การบันทึก แก้ไข ปรับปรุงข้อมูลในระบบ Change Log	5	0	ดี
8. ความปลอดภัยของระบบโดยรวม เช่น การกำหนด Auto Lock Screen	4.8	0.45	ดี
9. ความสะดวก ความง่ายในการใช้งานระบบ (User Friendly)	4.8	0.45	ดี
เฉลี่ย	4.78	0.42	ดี

ส่วนที่ 1 การประเมินด้านความเหมาะสมของระบบที่ได้พัฒนาขึ้น ได้มีการกำหนดบทบาทหน้าที่งานให้กับทางผู้เชี่ยวชาญ ทดสอบระบบภายในเวลา 30 นาที โดยสามารถสรุปการประเมินได้ดังตารางที่ 3

ส่วนที่ 2 การประเมินด้านความพึงพอใจของระบบที่ได้พัฒนาขึ้น ได้มีการกำหนดบทบาทหน้าที่งานให้กับทางผู้ใช้งานระบบ ภายในระยะเวลา 30 นาที โดยผู้ใช้งานจะทำการประเมินเปรียบเทียบระหว่างความต้องการเพิ่มเติมจากระบบเดิมที่ใช้งานและระบบที่ได้พัฒนาขึ้นว่าตรงต่อความต้องการของผู้ใช้งานหรือไม่ นาที โดยสามารถสรุปการประเมินได้ดังตารางที่ 4

ตารางที่ 4: ผลการประเมินความพึงพอใจของระบบ

รายการประเมิน	ผลการประเมิน		
	X	S.D.	ระดับ
1. ความพึงพอใจในการกำหนด การตั้งค่า การจัดเก็บข้อมูลส่วนบุคคล	4.17	0.98	ดี
2. ความพึงพอใจในการตั้งค่าระบบ สำหรับจำกัดสิทธิการใช้งาน	3.83	0.75	ดี
3. ความพึงพอใจในการมองเห็น การใช้งานข้อมูลส่วนบุคคลและในการทำงาน โดยเฉพาะอย่างยิ่งกับหน้าที่งานที่	4.33	0.82	ดี
4. ความพึงพอใจด้านการจำกัดสิทธิ การเพิ่ม แก้ไข หรือการลบข้อมูลส่วนบุคคล	4.17	0.98	ดี
5. ความพึงพอใจของข้อมูลที่แสดงในรายงานข้อมูลส่วนบุคคล	4.33	1.03	ดี
6. ความพึงพอใจของข้อมูลส่วนบุคคลในระบบ ในการจัดการปิดกั้น เปลี่ยนสถานะตามความต้องการของเจ้าของข้อมูล	4.33	0.82	ดี
7. ความพึงพอใจในการตรวจสอบ การบันทึก แก้ไข ปรับปรุงข้อมูลในระบบ	4.00	0.89	ดี
8. ความปลอดภัยของระบบโดยรวม เช่น การกำหนด Auto Lock Screen	4.33	1.03	ดี
9. ความสะดวก ความง่ายในการใช้งานระบบ (User Friendly)	4.33	1.03	ดี
เฉลี่ย	4.20	0.88	ดี

5. สรุป

จากการพัฒนาฟังก์ชันมาตราการรักษาความปลอดภัยของข้อมูลส่วนบุคคลในระบบการวางแผนทรัพยากรองค์กร สามารถสรุปผลการประเมินของระบบที่ได้พัฒนาขึ้นออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 ผลการประเมินความเหมาะสมของระบบที่ได้พัฒนาขึ้นที่ได้รับการประเมินจากผู้เชี่ยวชาญจำนวน 5 ท่าน ในการทดสอบระบบตามหน้าที่งานที่ได้มอบหมายตามความเหมาะสมและความถนัดของผู้เชี่ยวชาญ ภายในระยะเวลา 30 นาที พบว่ามีค่าเฉลี่ยอยู่ที่ 4.78 มีค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.42 สามารถสรุปได้ว่าความเหมาะสมของระบบที่ได้พัฒนาค่ามากกว่าระดับ 4.0 และอยู่ในระดับดี

ส่วนที่ 2 ผลการประเมินความพึงพอใจของระบบที่ได้พัฒนาขึ้นที่ได้รับการประเมินจากผู้ใช้งานระบบจำนวน 6 ท่าน โดยเป็นการประเมินเปรียบเทียบกันระหว่างระบบเดิมและระบบที่ได้พัฒนาขึ้นว่าตรงต่อความต้องการของผู้ใช้งานหรือไม่ พบว่ามีค่าเฉลี่ยอยู่ที่ 4.20 มีค่าส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.88 สามารถสรุปได้ว่าระบบที่ได้พัฒนาขึ้นได้รับความพึงพอใจจากผู้ใช้งานมีค่ามากกว่าระดับ 4.0 และอยู่ในระดับดี

นอกจากนี้ผู้วิจัยได้มีการทดสอบเพิ่มเติมโดยการใช้ความรู้ด้านการใช้ภาษา SQL ในการเรียกดูข้อมูลส่วนบุคคลจากฐานข้อมูลผ่าน โปรแกรม Microsoft SQL Server Management พบว่าระบบที่ได้พัฒนาที่การป้องกันการมองเห็นข้อมูลส่วนบุคคลทั้งการใช้งานในระบบ SAP Business One ของผู้ใช้งานหรือการป้องกันผู้ไม่มีสิทธิ์ที่มีความรู้ด้านการใช้ภาษาคอมพิวเตอร์เพื่อเข้ามาเรียกดูข้อมูลส่วนบุคคลในระบบที่ได้พัฒนาขึ้นนั่นเอง

เอกสารอ้างอิง

- [1] มหาวิทยาลัยราชภัฏราชนครินทร์, การพัฒนาระบบเทคโนโลยีสารสนเทศ. Available Online at <https://sites.google.com/site/ispmmzizijex/kar-phathna-rabb-sarsnthes>, accessed on 30 October 2022.
- [2] บริษัท เคต้า วิว จำกัด, Data Privacy VS Data Security ความต่างในความเหมือนกับการคุ้มครองข้อมูลตาม

- น ฎ ฬ ม ฎ ฎ PDPA. Available Online at <https://pdpacore.com/th/blogs/the-difference-between-data-privacy-and-data-security-in-pdpa-perspective>, accessed on 30 October 2022.
- [3] Easy Company Group Company Limited, PDPA ฎ ฎ ? Available Online at <https://easypdpa.com/article/easypdpa-summary-what-is-pdpa>, accessed on 17 July 2022.
- [4] Aware Company Limited, SAP Business One?. Available Online at <https://www.aware.co.th/sap-b1-for-business/>, accessed on 17 July 2022.
- [5] น ฎ ม น ฎ ฎ ฎ ฎ , GRPR ฎ ฎ ? Available Online at <https://www.dft.go.th/thth/DetailHotNews/ArticleId/10986/10986GDPR>, accessed on 22 March 2023.

เทคนิคการจำแนกประเภทข้อความสำหรับการรับแจ้งปัญหาอัตโนมัติ

Text Classification for Automatic Problem Reporting

ภูวนาด ภูพันธ์กุล (Puwanart Poopankul)¹ และพยุ่ง มีสังข์ (Phayung Meesad)²
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021856062@emmail.kmutnb.ac.th, ²phayung.m@itd.kmutnb.ac.th

บทคัดย่อ

ปัญหาการใช้งานสามารถเกิดขึ้นได้ตลอดเวลา โดยปัญหาที่ได้รับแจ้งเข้ามามีหลากหลายคำถาม ทำให้ยากต่อการจำแนกประเภทของปัญหาที่ถามเข้ามา งานวิจัยนี้มีวัตถุประสงค์เพื่อจำแนกประเภทข้อความสำหรับการรับแจ้งปัญหาอัตโนมัติ โดยแบ่งงานวิจัยออกเป็น 2 ส่วน ดังนี้ ส่วนที่ 1 ส่วนของการเปรียบเทียบแบบจำลองเพื่อทำการจำแนกประเภทข้อความจำนวน 4 แบบจำลอง ได้แก่ LSTM, Bi-LSTM, GRU และ Bi-GRU ข้อมูลที่ใช้เป็นข้อมูลคำถามและคำตอบ ที่ผ่านกระบวนการเตรียมข้อมูลให้พร้อมก่อนสร้างแบบจำลองจำนวน 6,648 ข้อความ แบ่งเป็น 168 ประเภทคำถาม จากนั้นนำข้อมูลไปสร้างแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพในการจำแนก พบว่าแบบจำลอง Bi-LSTM ให้ประสิทธิภาพสูงที่สุด โดยมีถูกต้องเฉลี่ยเท่ากับร้อยละ 91.16 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.940 ค่าระลอกเฉลี่ยเท่ากับ 0.912 และค่าถ่วงดุลเฉลี่ยเท่ากับ 0.926 ส่วนที่ 2 พัฒนาระบบสำหรับทดสอบแบบจำลอง โดยนำแบบจำลองที่มีประสิทธิภาพสูงที่สุดจากการเปรียบเทียบไปทำการพัฒนาระบบสำหรับแสดงผลการจำแนกประเภทคำถามจากผู้ใช้

คำสำคัญ: เทคนิคการจำแนกประเภทข้อความ การรับแจ้งปัญหาอัตโนมัติ การเรียนรู้เชิงลึก

Abstract

Problems can occur all the time, and the problem is various, making it difficult to classify the type of problem. The purpose of this study was to propose a text classification technique for automatic problem

reporting, divided into two parts. The first part compares four models for text classification: LSTM, Bi-LSTM, GRU, and Bi-GRU. The data used consists of 6,648 preprocessed questions and answers, categorized into 168 question categories, and to create and test the models to compare classification performance. The results showed that the Bi-LSTM model achieved the highest accuracy at 91.16%, precision at 0.940, recall at 0.912, and F1-score at 0.926. The second part involved the development of a system to test the model, where the best-performing model from the comparison was used to systems developed for displaying the classification results of user problems.

Keywords: Text Classification, Automatic Problem Reporting, Deep learning

1. บทนำ

ในปัจจุบันเทคโนโลยีเพื่อการสื่อสารมีความก้าวหน้ามากขึ้น โดยส่งผลกระทบต่อการใช้งานในแต่ละองค์กร เราจะเห็นว่าหลาย ๆ องค์กรมีความจำเป็นที่จะต้องปรับตัวเพื่อความอยู่รอดและความเจริญก้าวหน้าในอนาคต พร้อมทั้งมีการปรับเปลี่ยนโครงสร้างพื้นฐานองค์กร โดยในบางองค์กรได้มีการนำเทคโนโลยีสารสนเทศเข้ามาช่วยลดกระบวนการขั้นตอนการทำงานช่วยในการสื่อสารประสานงานระหว่างหน่วยงานได้อย่างรวดเร็ว และมีประสิทธิภาพที่ดีมากยิ่งขึ้น อีกทั้งบุคลากรในองค์กรต้องมีความรู้ความสามารถในการใช้งานเทคโนโลยี สามารถปรับเปลี่ยนพฤติกรรมในการใช้งานเทคโนโลยีได้เป็น

อย่างดี รวมไปถึงแนวโน้มที่มีการใช้งานเทคโนโลยีเพิ่มขึ้นอย่างรวดเร็ว ดังนั้นในส่วนของกาให้บริการการสื่อสาร การรับมือการแก้ไขปัญหาที่เกี่ยวข้องกับเทคโนโลยี จำเป็นจะต้องมีการเตรียมการบริการให้สามารถรองรับปัญหาที่เกิดขึ้นและสอดคล้องไปพร้อมกับแนวทางการปรับเปลี่ยนโครงสร้างพื้นฐานองค์กร

เนื่องจากการใช้งานระบบสารสนเทศจากบุคลากรในองค์กรจำนวนมาก ๆ นั้นย่อมมีปัญหาการใช้งานที่สามารถเกิดขึ้นได้ตลอดเวลา โดยปัญหาที่ได้รับแจ้งเข้ามา มีหลากหลายคำถามอย่างเช่น การเข้าใช้งานระบบไม่ได้ คอมพิวเตอร์เปิดใช้งานไม่ได้ ความพร้อมใช้งานระบบระบบช้า เป็นต้น ทำให้ยากต่อการจำแนกประเภทของปัญหาที่ถามเข้ามา

จากปัญหาดังกล่าวข้างต้นผู้วิจัยจึงนำเทคนิควิธีการจำแนกประเภทข้อความมาใช้สำหรับการจำแนกประเภทคำถามสำหรับการรับแจ้งปัญหาอัตโนมัติ โดยเลือกแบบจำลอง LSTM, Bi-LSTM, GRU และ Bi-GRU มาเปรียบเทียบประสิทธิภาพในการจำแนกประเภทข้อความและนำแบบจำลองที่เหมาะสมมาพัฒนาระบบเพื่อทดสอบการแสดงผลการจำแนกประเภทคำถามสำหรับการรับแจ้งปัญหาอัตโนมัติต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาษาธรรมชาติ (Natural

Language Processing: NLP)

NLP เป็นเทคนิคทางคอมพิวเตอร์ที่มีพื้นฐานทฤษฎีในการวิเคราะห์และเข้าใจภาษาธรรมชาติ (natural language) ในระดับภาษาศาสตร์ที่หลากหลาย หรือมากกว่านั้น เพื่อให้บรรลุการประมวลผลภาษาอย่างมีประสิทธิภาพและใกล้เคียงกับมนุษย์ สามารถนำไปประยุกต์ใช้ในงานหรือแอปพลิเคชันต่าง ๆ [1]

2.2 การเรียนรู้เชิงลึก (Deep Learning)

Deep Learning เป็นส่วนหนึ่งของการเรียนรู้ของเครื่องจักร (Machine Learning) มีพื้นฐานมาจากโครงข่ายประสาทเทียม (Artificial Neural Network : ANN) เป็นวิธีที่สร้างขึ้นเพื่อทำให้เครื่องจักรสามารถเรียนรู้ได้โดยใช้

ต้นแบบมาจากระบบประสาทของมนุษย์ โดยต้องใส่ข้อมูลเข้าไปในชั้นรับข้อมูล (Input Layer) จากนั้นเครื่องจักรจะนำข้อมูลไปประมวลผลในชั้นซ่อน (Hidden Layer) แล้วจะนำเสนอข้อมูลผลลัพธ์ในชั้นแสดงผล (Output Layer) แสดงดังภาพที่ 2-1 โครงข่ายประสาทเทียมได้มีการพัฒนาอย่างต่อเนื่องซึ่งวิธีที่นิยมใช้ในปัจจุบันคือ โครงข่ายประสาทคอนโวลูชัน (Convolution Neural Network: CNN), LSTM และ GRU [2]

2.3 Long Short-Term Memory: LSTM

เป็นโครงข่ายประเภท RNN รูปแบบหนึ่งที่ถูกพัฒนาขึ้นมาให้มีความเสถียรและมีประสิทธิภาพมากขึ้น โดยมีหลักการทำงานคือ สามารถเก็บ ‘สถานะ’ หรือข้อมูลของแต่ละโหนดเอาไว้เพื่อที่เวลาย้อนกลับไปได้จะดูทราบถึงที่มาของข้อมูลค่าดังกล่าวว่าเดิมเป็นค่าอะไร และจุดเด่นของแบบจำลอง LSTM คือฟังก์ชันพิเศษที่มีหน้าที่เสมือนประตู (Gate) ที่คอยควบคุมข้อมูลที่จะเข้าไปในแต่ละโหนด ซึ่งประกอบด้วย Forget gate layer, Input gate layer และ Output gate layer [3]

2.3 Bidirectional LSTM: Bi-LSTM

เป็นอีกรูปแบบหนึ่งของ LSTM มีหลักการทำงานเหมือนกับ LSTM เพียงแต่มีการเพิ่มหนึ่งฟังก์ชันในการทำงานคือการป้อนข้อมูลแบบย้อนกลับเข้าไปด้วย นั้นหมายความว่าขั้นตอนในการทดสอบแบบ Bi-LSTM จะช้ากว่า LSTM [4]

2.4 Gated Recurrent Unit: GRU

เป็นรูปแบบหนึ่งของโครงข่ายระบบประสาทแบบย้อนกลับ (RNN) มีจุดประสงค์เพื่อแก้ปัญหาการลดลงของเกรเดียนต์ (Vanishing Gradient) หรือก็เพื่อแก้ปัญหาการมีหน่วยความจำระยะสั้นของ RNN เช่นเดียวกับ LSTM ซึ่งการทำงานภายในของ GRU จะมีความคล้ายคลึงกับ LSTM เช่นกัน แต่จะลดความซับซ้อนของประตู (Gate) ลงเหลือเพียงแค่ 2 ประตู เพื่อให้เกิดความรวดเร็วในการคำนวณ ได้แก่ Update Gate เป็นประตูนำเข้าข้อมูลเพื่อนำไปคำนวณในการปรับสถานะของค่าเซลล์และ Reset Gate เป็นประตูที่ใช้ในการตัดสินใจว่าจะปรับสถานะเซลล์เมื่อมีข้อมูลใหม่เข้ามาหรือไม่ [5]

2.5 Gated Recurrent Unit: GRU

Bi-GRU คล้ายกับ Bi-LSTM สามารถใช้ข้อมูลไปข้างหน้าและย้อนกลับได้พร้อมกัน เนื่องจากหน่วย GRU นั้นง่ายกว่าหน่วย LSTM รุ่น Bi-GRU จึงง่ายกว่ารุ่น Bi-LSTM [7]

2.6 งานวิจัยที่เกี่ยวข้อง

Mahara, et al. [6] ได้นำเสนอเกี่ยวกับการตรวจจับข่าวปลอม ในงานวิจัยนี้ได้ทำการเปรียบเทียบแบบจำลองระหว่าง LSTM และ Bi-LSTM เพื่อทำการตรวจจับข่าวปลอม โดยชุดข้อมูลที่ใช้ในการทดลอง คือชุดข้อมูลสาธารณะประเภทข่าวสารที่มีอยู่บน Kaggle โดยแบ่งข้อมูลออกเป็นข้อมูล Training set เท่ากับร้อยละ 70 และข้อมูล Test set เท่ากับร้อยละ 30 สำหรับสร้างแบบจำลองด้วย ผลการวิจัยพบว่าแบบจำลอง Bi-LSTM มีความแม่นยำสูงสุดถึง 94% และ 93% เมื่อใช้ค่าพารามิเตอร์ Dropout เท่ากับ 0.1 และ 0.2 ตามลำดับ

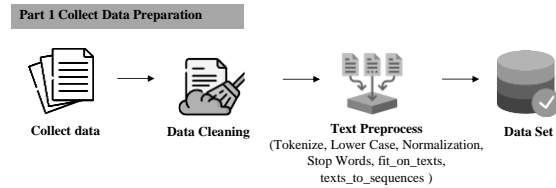
Yang, et al. [7] ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่าง LSTM และ GRU โดยใช้ข้อมูลจาก Yelp Open Source เป็นชุดข้อมูลเกี่ยวกับการแสดงความคิดเห็นเกี่ยวกับร้านอาหาร ห้างสรรพสินค้า โรงแรม และการท่องเที่ยวจากทั่วโลก ผลการเปรียบเทียบพบว่าแบบจำลอง GRU ใช้เวลาในการฝึกสอนน้อยกว่าร้อยละ 29.29 และให้ประสิทธิภาพสูงกว่าแบบจำลอง LSTM โดยมีค่าความถูกต้องสูงกว่า ร้อยละ 23.45 ค่าระลอกสูงกว่าร้อยละ 27.69 ค่าการถ่วงดุลสูงกว่าร้อยละ 26.95

3. วิธีดำเนินการวิจัย

การพัฒนาเทคนิคการจำแนกประเภทข้อความสำหรับระบบปรับแจ้งปัญหาอัตโนมัติ มีการดำเนินการงาน 5 ขั้นตอนดังต่อไปนี้

3.1 การรวบรวมข้อมูลและเตรียมข้อมูล

ผู้วิจัยได้ทำการรวบรวมข้อมูลและเตรียมข้อมูลก่อนการประมวลผล ซึ่งกระบวนการทำงานแสดงดังภาพ 1

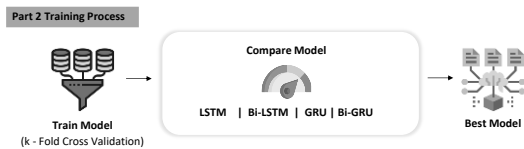


ภาพที่ 1 : การรวบรวมข้อมูลและเตรียมข้อมูล

จากภาพที่ 1 ผู้วิจัยได้เก็บรวบรวมคำถามที่ผู้ใช้งานมีการแจ้งปัญหาด้านต่างๆ ผ่านช่องทางไลน์แอปพลิเคชัน และช่องทางโทรศัพท์ ที่ได้รวบรวมเก็บไว้ ของศูนย์เทคโนโลยีสารสนเทศ กลุ่มถ่ายทอดเทคโนโลยีสารสนเทศ สำนักงานการตรวจเงินแผ่นดิน จากนั้นนำข้อมูลที่ได้ไปผ่านกระบวนการทำความสะอาดข้อมูล (Data Cleaning) เพื่อนำข้อมูลที่ไม่เกี่ยวข้องออก ซึ่งได้ข้อมูลผ่านการบวนการทำความสะอาดแล้วได้ข้อมูลชุดคำถามและคำตอบจำนวน 6,648 และแบ่งประเภทข้อมูลออกเป็น 168 ประเภท มาเตรียมความพร้อมโดยผ่านกระบวนการเตรียมข้อมูลเพื่อการวิเคราะห์ข้อความ (Text Preprocessing) ได้แก่ การกำจัดคำหยุด กระบวนการตัดคำ ผู้วิจัยได้ทำการเลือกใช้เทคนิค Attacut ในการตัดคำเนื่องจากมีประสิทธิภาพในการตัดคำภาษาไทยที่ดีและใช้เวลาในการประมวลผลน้อย จากนั้นใช้เครื่องมือในไลบรารี Keras ได้แก่ fit_on_texts ในการแปลงข้อมูลหลายๆประโยคจากชุดข้อมูลคำถาม ให้เป็นคำ ๆ และกำหนด Index ให้แต่ละคำ จากนั้นใช้ texts_to_sequences แปลงประโยคให้เป็น Sequence (Array of Integer) จากนั้นใช้ pad_sequences ทำทุกประโยคมีความยาวเท่ากันด้วยการเติม 0 ท้ายประโยค และใช้เครื่องมือในไลบรารี pandas คือ get_dummies ในการทำ One-Hot encoding เพื่อเปลี่ยนประเภทคำถามเป็นตัวเลข เพื่อให้สามารถนำข้อมูลไปประมวลผลในการสร้างแบบจำลองได้

3.2 ขั้นตอนการสร้างแบบจำลอง และการเปรียบเทียบประสิทธิภาพ

ผู้วิจัยได้ทำการสร้างแบบจำลอง และการเปรียบเทียบประสิทธิภาพ ซึ่งกระบวนการทำงานแสดงดังภาพ 2



ภาพที่ 2: กระบวนการสร้างแบบจำลอง และการเปรียบเทียบประสิทธิภาพ

3.2.1 ขั้นตอนการสร้างแบบจำลอง

จากภาพที่ 2 ผู้วิจัยได้ใช้วิธีการตรวจสอบไขว้ (k – Fold Cross Validation) กำหนดค่า K เท่ากับ 10 และแบ่งข้อมูลออกเป็น ส่วน ๆ ให้เท่า ๆ กันแบบสุ่มจำนวน K ส่วน สำหรับใช้ในการสร้างแบบจำลองและทดสอบ ทำการเปรียบเทียบด้วยการปรับพารามิเตอร์ จำนวน 1 พารามิเตอร์ คือ ค่าพารามิเตอร์ Hidden size เท่ากับ 32, 64, 128, 256 เพื่อหาประสิทธิภาพสูงที่สุดในแต่ละแบบจำลอง ซึ่งผู้วิจัยได้กำหนดพารามิเตอร์อื่น ๆ ที่ใช้ในแบบจำลอง ได้แก่ ค่าพารามิเตอร์ Word Embedding Dimension เท่ากับ 392 (จำนวนคำที่เป็นเอกลักษณ์ของชุดข้อมูล) ค่าพารามิเตอร์ shape เท่ากับ 20 (จำนวนความยาวสูงสุดของประโยคคำถาม) ค่าพารามิเตอร์ Dropout เท่ากับ 0.1 ค่าพารามิเตอร์ recurrent dropout เท่ากับ 0.2 ค่าพารามิเตอร์ First Dense Layer เท่ากับ 168 (จำนวนประเภทคำถาม) (Active Learning เท่ากับ SoftMax) ค่าพารามิเตอร์ Epochs เท่ากับ 100 ค่าพารามิเตอร์ Batch size เท่ากับ 100 ค่าพารามิเตอร์ Optimizer เท่ากับ ADAM ค่าพารามิเตอร์ Loss Function เท่ากับ Categorical Cross entropy Loss จากนั้นทำการทดสอบประสิทธิภาพของแบบจำลอง

3.2.2 การทดสอบประสิทธิภาพ

จากภาพที่ 2 ผู้วิจัยได้ใช้การประเมินประสิทธิภาพของตัวแบบจำลองสามารถพิจารณาจากความถูกต้องจากตาราง Confusion Matrix [8] โดยมีรายละเอียดดังนี้

ตารางที่ 1: ตาราง Confusion Matrix

Actual Values	Predicted Values	
	Positive (Yes)	Negative (No)
Positive (Yes)	True Positive (TP)	False Negative (FN)
Negative (No)	False Positive (FP)	True Negative (TN)

ตารางที่ 1 สามารถอธิบายได้ดังนี้

True Positive คือ สิ่งที่แบบจำลองทำนายว่า “จริง” และมีค่าเป็น “จริง”

False Negative คือ สิ่งที่ แบบจำลองทำนายว่า “จริง” และมีค่าเป็น “ไม่จริง”

True Negative คือ สิ่งที่แบบจำลองทำนายว่า “ไม่จริง” และมีค่าเป็น “ไม่จริง”

False Positive คือ สิ่งที่แบบจำลองทำนายว่า “ไม่จริง” และมีค่าเป็น “จริง”

- ค่าความถูกต้อง (Accuracy) หมายถึง ค่าจำนวนข้อมูลที่ทำนายถูกในคลาสนั้นหารด้วยจำนวนข้อมูลทั้งหมดในคลาสนั้น ดังสมการที่ 1

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{1}$$

- ความแม่นยำ (Precision) หมายถึง ค่าจำนวนข้อมูลที่ทำนายถูกต้องคำนวณได้จากจำนวนข้อมูลที่ทำนายถูกในคลาสนั้นหารด้วยจำนวนข้อมูลทั้งหมดที่การทำนายให้ผลลัพธ์เดียวกันในคลาสนั้น ดังสมการที่ 2

$$Precision = \frac{TP}{(TP+FP)} \tag{2}$$

- ค่าระลึก (Recall) หมายถึง ค่าการทำนายด้วยโมเดลที่ตรงกับความเป็นจริง โดยมีค่าเท่ากับค่า True Positive Rate หารด้วยความจริงทั้งหมด ดังสมการที่ 3

$$Recall = \frac{TP}{(TP+FN)} \tag{3}$$

- ค่าความถ่วงดุล (F-measure) หมายถึง ค่าการวัดประสิทธิภาพโดยรวมที่เกิดจากการ เปรียบเทียบระหว่างค่า Precision และ Recall ของแต่ละคลาสเป้าหมาย ดังสมการที่ 4

$$F - measure = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \tag{4}$$

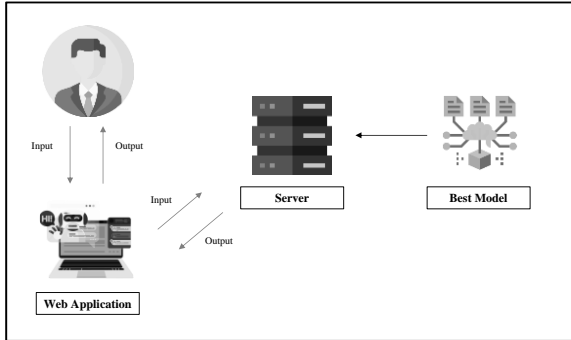
3.2.3 การเปรียบเทียบประสิทธิภาพ

จากภาพที่ 2 ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพแบบจำลองในการจำแนกประเภทข้อความสำหรับระบบรับแจ้งปัญหาอัตโนมัติของทั้ง 4 แบบจำลอง ได้แก่ LSTM, Bi-LSTM, GRU และ Bi-GRU โดยเปรียบเทียบค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าระลึก (Recall) ค่าการถ่วงดุล (F-Measure) เพื่อหาแบบจำลองที่ดีที่สุด

3.3 การออกแบบและพัฒนาระบบสำหรับทดสอบ

แบบจำลอง

ขั้นตอนการออกแบบและพัฒนาระบบรับแจ้งปัญหาอัตโนมัติเพื่อใช้สำหรับทดสอบแบบจำลองที่ดีที่สุด หลังจากการเปรียบเทียบประสิทธิภาพ มีการออกแบบและพัฒนาระบบรับแจ้งปัญหาอัตโนมัติ แสดงดังภาพที่ 3



ภาพที่ 3 : การออกแบบและพัฒนาระบบรับแจ้งปัญหาอัตโนมัติ

จากภาพที่ 3 การพัฒนาระบบแสดงผลการจำแนกประเภทข้อความสำหรับการรับแจ้งปัญหาอัตโนมัติสามารถแบ่งการดำเนินงานออกเป็น 2 ส่วน ซึ่งมีรายละเอียดดังนี้

ส่วนที่ 1 : การสร้างส่วน Font-end สำหรับติดต่อผู้ใช้งาน ในรูปแบบของ Web application ใช้ภาษา HTML ในการจัดการ โครงสร้างและรูปร่างของระบบ และตกแต่งระบบด้วย ภาษา CSS และเขียนการทำงานของปุ่มส่งข้อมูลด้วย JavaScript ในการส่งข้อมูลที่ Input เข้ามาไปยังเส้นทางที่กำหนด

ส่วนที่ 2 : การสร้างส่วน Back-end สำหรับการประมวลผลและการติดต่อกันระหว่าง Web และ Server ใช้ Flask ในการรับข้อมูลที่ request เส้นทางโดยการรับข้อมูลมาในรูปแบบ JSON นำข้อมูลที่ส่งเข้ามาไปที่แบบจำลองที่มีประสิทธิภาพที่ดีที่สุดในขั้นตอนที่ 3.3 เพื่อทำการประมวลผลในการจำแนกคำถามที่รับเข้ามา และทำการตอบกลับจากผลการจำแนกที่ได้ส่งไปยังส่วนของ User Interface โดยภาษาที่ใช้ในการพัฒนาคือ Python

3.4 ทดสอบระบบ

สำหรับขั้นตอนนี้เป็นขั้นตอนของการทดสอบแบบจำลอง โดยจำลองการใช้งานผ่านระบบรับแจ้งปัญหาอัตโนมัติ ในรูปแบบ Web application เพื่อดูว่าแบบจำลองที่ได้พัฒนามีการตอบกลับคำถามได้หรือไม่

4. ผลการดำเนินงานวิจัย

4.1 ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

การสร้างแบบจำลองสำหรับการจำแนกข้อความสำหรับระบบรับแจ้งปัญหาอัตโนมัติ มีผลการเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่าง 4 แบบจำลอง แสดงดังตารางที่ 2

ตารางที่ 2 : ผลการเปรียบเทียบอัลกอริทึม

Model	Hidden Layer	Accuracy	Precision	Recall	F-Measure
LSTM	256	91.16	0.934	0.912	0.923
Bi-LSTM	64	91.16	0.940	0.912	0.926
GRU	128	91.16	0.933	0.912	0.922
Bi-GRU	256	91.16	0.939	0.912	0.925

จากตารางที่ 2 สรุปได้ว่า Bi-LSTM ที่ปรับค่าพารามิเตอร์ Hidden size เท่ากับ 64 ได้ประสิทธิภาพสูงที่สุด โดยมีค่าความถูกต้องเฉลี่ยเท่ากับร้อยละ 91.16 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.940 ค่าระลึกลเฉลี่ยเท่ากับ 0.912 และค่าถ่วงดุลเฉลี่ยเท่ากับ 0.926 ในส่วนของ Bi-GRU ที่ปรับค่าพารามิเตอร์ Hidden size เท่ากับ 256 ได้ประสิทธิภาพลำดับที่สอง โดยมีความถูกต้องเฉลี่ยเท่ากับร้อยละ 91.16 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.939 ค่าระลึกลเฉลี่ยเท่ากับ 0.912 และค่าถ่วงดุลเฉลี่ย เท่ากับ 0.925 ส่วน LSTM ที่ปรับค่าพารามิเตอร์ Hidden size เท่ากับ 256 ได้ประสิทธิภาพลำดับที่สาม โดยมีความถูกต้องเฉลี่ยเท่ากับร้อยละ 91.16 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.934 ค่าระลึกลเฉลี่ยเท่ากับ 0.912 และค่าถ่วงดุลเฉลี่ยเท่ากับ 0.923 และ GRU ที่ปรับค่าพารามิเตอร์ Hidden size เท่ากับ 128 ได้ประสิทธิภาพน้อยที่สุด โดยมีค่าความถูกต้องเท่ากับร้อยละ 91.16 ค่าความแม่นยำเท่ากับ 0.933 ค่าระลึกลเท่ากับ 0.912 และค่าถ่วงดุลเท่ากับ 0.922

4.2 ผลการพัฒนาสำหรับทดสอบแบบจำลอง

ผลลัพธ์จากการพัฒนาระบบแสดงผลการจำแนกประเภทข้อความสำหรับการรับแจ้งปัญหาอัตโนมัติ โดยระบบจะรับข้อมูลปัญหาที่ส่งเข้ามา เพื่อส่งไปประมวลผล

(Backend) ด้วย Flask framework ที่ใช้ในควบคุมการเชื่อมต่อระหว่างส่วนติดต่อผู้ใช้งานกับส่วนของ Server ของระบบโดยมีการรับ-ส่งข้อมูลในรูปแบบ JSON จากนั้นทำการส่งข้อความตอบกลับ ซึ่งในขั้นตอนนี้ Web UI ของระบบจะทำการถอดรหัสข้อมูลทศนทนา และทำการประมวลเพื่อตอบกลับคำถามผู้ใช้งานดังภาพที่ 4



ภาพที่ 4: ผลลัพธ์จากการพัฒนาระบบรับแจ้งปัญหาอัตโนมัติ

5. การอภิปรายผลการวิจัย

ผลการวิจัยพบว่าแบบจำลอง Bi-LSTM ที่กำหนดค่าพารามิเตอร์ Hidden size เท่ากับ 64 จะให้ประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเฉลี่ยเท่ากับร้อยละ 91.16 ค่าความแม่นยำเฉลี่ยเท่ากับ 0.940 ค่าระลึกเฉลี่ยเท่ากับ 0.912 และค่าถ่วงดุลเฉลี่ยเท่ากับ 0.926 โดย Bi-LSTM นั้นเป็นการปรับปรุงมาจาก LSTM เพียงแต่มีการป้อนข้อมูลแบบย้อนกลับเข้าไปด้วย ทำให้สามารถจดจำสถานะจากทั้งสองทิศทางได้ Bi-LSTM จึงมีจำนวนพารามิเตอร์ที่เยอะกว่า ทำให้มีค่าความแม่นยำที่มากกว่า ขณะเดียวกันก็มีพารามิเตอร์ที่เยอะกว่าแบบจำลองประเภทอื่น ๆ เช่นกัน ซึ่งสอดคล้องกับงานวิจัยของ Mahara, et al. [9] ที่ทำการศึกษาการตรวจจับข่าวปลอมด้วยเทคนิค Deep learning และ Zhang and Guo [10] ที่ทำการศึกษาการวิเคราะห์ความรู้สึกข้อความภาษาจีน พบว่า แบบจำลอง Bi-LSTM ให้ประสิทธิภาพในการจำแนกสูงที่สุดเช่นกัน

เอกสารอ้างอิง

[1] Liddy, E.D. (2001). Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc.

[2] ธนดล สิงขรอาสน์. (2021). การเรียนรู้เชิงลึกสำหรับการตรวจจับและรู้จากคำบรรยายในวิดีโอ. ปรียญานพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.

[3] วิศรุต แก้วมหา และวริศ ปัญญาฉัตรพร, “การคาดการณ์ผลตอบแทนในอนาคตของตราสารทุนหุ้นสามัญโดยการใช้ระบบคอมพิวเตอร์เรียนรู้ได้ด้วยตนเอง” วารสารนวัตกรรมธุรกิจ การจัดการ และสังคมศาสตร์ ฉบับที่ 2 ก.ย.-ธ.ค. 2564 หน้า 108-123.

[4] ปิยศรีณย์ ลิขิตาภรณ์, สมพร ปั่นโกษา และ บำรุง พ่วงเกิด, “การพยากรณ์ดัชนีราคาตลาดหุ้นโดยใช้ Wavelet Transform, LSTM, และ Bi-LSTM” การประชุมนำเสนอผลงานวิจัยบัณฑิตศึกษาระดับชาติ ครั้งที่ 16 2566 หน้า 153-164.

[5] รัชกษณา ภูสีเขียว และสุรณพรีร์ ภูมิวุฒิสาร, “การพยากรณ์ปริมาณน้ำฝนระยะสั้นในบริเวณพื้นที่สนามบินสุวรรณภูมิด้วยโครงข่ายระบบประสาทแบบย้อนกลับ” Journal of Information Science and Technology Vol.2 No.1: 13–26.

[6] G. S. Mahara and S. Gangele, “Fake news detection: A RNN-LSTM, Bi-LSTM based deep learning approach.” in IEEE 1st International Conference on Data, Decision and Systems (ICDDS). Bangalore, India, Dec. 2022, pp. 01–06.

[7] S. Yang, X. Yu, and Y. Zhou, “LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example,” in 2020 International Workshop on Electronic Communication and Artificial Intelligence. (IWECIAI), Shanghai, China, Jun. 2020, pp. 98–101.

[8] Powers, David M W. (2011). “Evaluation: From Precision, Recall and F-Measure to ROC, Informed Ness, Markedness & Correlation.” Journal of Machine Learning Technologies. Vol.2 No.1 : 37–63.

[9] G. S. Mahara and S. Gangele. (2022). “Fake news detection: A RNN-LSTM, Bi-LSTM based deep learning approach.” in IEEE 1st International Conference on Data, Decision and Systems (ICDDS). Bangalore, India, Dec. 2022, pp. 01–06. Doi: 10.1109/ICDDS56399.2022.10037403

[10] T. Zhang and J. Guo. (2021). “Research on Chinese Sentiment Analysis Based on Bi-LSTM Networks,” in IEEE/ACIS 20th International Fall Conference on Computer and Information Science (ICIS Fall). Xi’an, China, Oct. 2021, pp. 198–203. Doi: 10.1109/ICISFall51598.2021.9627460.

การจำแนกรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมืองข้อมูล

Classification of Packet Pattern for Network Intrusion

by Using Data Mining

นิชาภา แซ่เหวียง (Nichapha Saewiang)¹ และมณเฑียร รัตนศิริวงษ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยมหาวิทาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857107@email.kmutnb.ac.th, ²montean.r@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึมและค้นหาอัลกอริทึมที่ดีที่สุดของรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่าย โดยใช้ข้อมูลจากเว็บไซต์ NSL-KDD จำนวน 50,000 เรคคอร์ด แบ่งเป็น 42 แอตทริบิวต์ ซึ่งทั้งหมดได้ผ่านกระบวนการทำความสะอาดข้อมูลตัดข้อมูลซ้ำซ้อนออกและจัดรูปแบบข้อมูลให้ถูกต้อง จากนั้นเข้าสู่กระบวนการเรียนรู้และทดสอบ โดยใช้หลักการ 10-Fold Cross Validation ในการเปรียบเทียบประสิทธิภาพของ 4 อัลกอริทึม ได้แก่ อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier อัลกอริทึม Bayesian Network อัลกอริทึม SPAARC และอัลกอริทึม ForEx++ ผลการวิจัยพบว่า อัลกอริทึม SPAARC มีประสิทธิภาพในการจำแนกที่ดีที่สุด โดยมีค่าความถูกต้อง (Accuracy) เท่ากับ 99.58% ค่าความแม่นยำ (Precision) เท่ากับ 71.66% ค่าระลึก (Recall) เท่ากับ 70.98% และค่าความถ่วงดุล (F-measure) เท่ากับ 71.28%

คำสำคัญ: เทคนิคการจำแนกข้อมูล การทำเหมืองข้อมูล การโจมตีระบบเครือข่าย

Abstract

This independent study aims to compare the performance of various algorithms and determine the most effective one for attacking packet patterns. Data from the NSL-KDD website consisting of 50,000 records with 42 attributes that have been processed will be

used. The data will be cleaned, redundant data will be removed, and it will be formatted correctly. The learning and testing process will use the 10-Fold Cross Validation principle to compare the performance of four algorithms: the Multi-objective Evolutionary Fuzzy Classifier Algorithm, Bayesian Network Algorithm, SPAARC Algorithm, and ForEx++ Algorithm. The results indicate that the SPAARC algorithm has the best classification performance with an accuracy of 99.58%, precision of 71.66%, recall of 70.98%, and F-measure of 71.28%

Keywords: Classification Techniques, Data Mining, Network Intrusion

1. บทนำ

ปัจจุบันการใช้งานอินเทอร์เน็ตในช่วงวิกฤตสถานการณ์แพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา (COVID-19) ทำให้ต้องปรับเปลี่ยนพฤติกรรมจากการทำงานที่องค์กรมาเป็นการทำงานจากที่ไหนก็ได้ (Work From Anywhere) ก่อให้เกิดความเสี่ยงช่องโหว่ในเรื่องของความปลอดภัยที่อาจนำไปสู่ภัยคุกคามทางอินเทอร์เน็ต จนเกิดความเสียหายเรื่องข้อมูลสำคัญขององค์กรและข้อมูลส่วนบุคคลรั่วไหลออกไปสู่สาธารณชน

ในปี 2565 สถิติภัยคุกคามจาก ThaiCERT [1] พบว่าประเภทภัยคุกคามที่ถูกโจมตีมากที่สุดอันดับที่ 1 ได้แก่ โปรแกรมไม่พึงประสงค์ (Malicious Code) จำนวน 1,187 ครั้งต่อปี อันดับที่ 2 จุดอ่อน หรือช่องโหว่อันตราย

(Vulnerability) จำนวน 637 ครั้งต่อปี และอันดับที่ 3 ภัยการบุกรุก หรือเจาะเข้าระบบ (Intrusion Attempts) จำนวน 207 ครั้งต่อปี ในการเกิดเหตุการณ์ภัยคุกคามทางไซเบอร์แต่ละครั้ง ทำให้ได้รับความเสียหายทางธุรกิจ ซึ่งการการศึกษางานวิจัยที่เกี่ยวกับการจำแนกข้อมูลโดยใช้ อัลกอริทึม ได้แก่ อัลกอริทึม Bayesian Network อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier อัลกอริทึม SPAARC และ อัลกอริทึม ForEx++ ซึ่งเป็นอัลกอริทึมที่สามารถจำแนกข้อมูลได้อย่างมีประสิทธิภาพ แต่ยังไม่เคยมีงานวิจัยไหนที่นำ 4 อัลกอริทึมนี้มาเปรียบเทียบประสิทธิภาพของแต่ละ อัลกอริทึม และค้นหาอัลกอริทึมที่ดีที่สุดของรูปแบบแพ็คเกจที่ใช้โจมตีระบบเครือข่าย

ดังนั้นในงานวิจัยนี้เครื่องมือที่ใช้คือ โปรแกรมวิเคราะห์ข้อมูล WEKA และการนำเสนอข้อมูลด้วยโปรแกรม Power BI โดยใช้ข้อมูลจากเว็บไซด์ NSL-KDD Dataset จำนวน 50,000 เรคคอร์ด แบ่งเป็น 42 แอตทริบิวต์ ซึ่งทั้งหมดได้ผ่านกระบวนการทำความสะอาดข้อมูลแล้ว จากนั้นเข้าสู่กระบวนการเรียนรู้และทดสอบ โดยใช้หลักการของ 10-Fold Cross-Validation ทำการหาประสิทธิภาพของอัลกอริทึม Bayesian Network อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier อัลกอริทึม SPAARC และอัลกอริทึม ForEx++ เพื่อเปรียบเทียบประสิทธิภาพของแต่ละ อัลกอริทึม และค้นหาอัลกอริทึมที่ดีที่สุดของรูปแบบแพ็คเกจที่ใช้โจมตีระบบเครือข่ายมาเป็นต้นแบบในการวิเคราะห์ใช้ต่อยอดสำหรับนักพัฒนาระบบต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งออกเป็น 2 ส่วน ดังนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การทำเหมืองข้อมูล (Data Mining)

กระบวนการทำงานที่ผ่านการกลั่นกรองข้อมูลจากชุดฐานข้อมูลขนาดใหญ่ เพื่อให้ได้สารสนเทศที่น่ามาใช้งาน ค้นหา รูปแบบ และความสัมพันธ์ที่อยู่ในชุดข้อมูล

การทำเหมืองข้อมูล ในปัจจุบันได้ถูกนำไปประยุกต์ใช้กับงานหลากหลายประเภท เช่น การใช้ข้อมูลเพื่อช่วยในการตัดสินใจของผู้บริหาร ซึ่งเป็นส่วนสำคัญที่ช่วยในการตัดสินใจทำธุรกิจ โดยการทำเหมืองข้อมูลเป็นขั้นตอนสำคัญในการค้นหาข้อมูลด้วยความรู้ในฐานข้อมูล (Knowledge Discovery in Database) หรือย่อ KDD [2]

2.1.2 เทคนิคการจำแนกประเภทข้อมูล(Classification)

เป็นการสร้างโมเดลจำแนกประเภทข้อมูลจากการหาความสัมพันธ์ของข้อมูลในฐานข้อมูลขนาดใหญ่ โดยข้อมูลจะแบ่งเป็น 2 ส่วน ได้แก่ ข้อมูลชุดเรียนรู้ (Training Set) นำมาแยกประเภทผลลัพธ์ด้วยการติดป้ายกำกับ (Label/Class) เป็นผลเฉลย จากนั้นนำข้อมูลที่ติดป้ายแล้วไปใช้ในการฝึกฝนผ่านอัลกอริทึมสำหรับสร้างโมเดลที่ใช้ในการทำนายผลลัพธ์ โดยนำข้อมูลใหม่หรือข้อมูลชุดทดสอบ (Test Set) ไปใช้ทดสอบประสิทธิภาพของโมเดล ประเมินความถูกต้องของโมเดลจำแนกประเภทข้อมูล [3]

2.1.3 อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier

เป็นหนึ่งในอัลกอริทึมของโปรแกรม WEKA วิธีการนี้สามารถระบุวิธีการที่ไม่ถูกจำกัดตามกฎการแยกประเภทของฟัซซี่ แก้ไขปัญหาปรับความเหมาะสมในเรื่องของข้อจำกัดหลายวัตถุประสงค์ซึ่งสามารถเป็นไปได้ โดยกำหนดตัวอย่างของปัญหา แสดงได้ดังสมการที่)2-1) [4]

$$\begin{aligned}
 &Max./Min. && FD(\Gamma) \\
 &Min. && NR(\Gamma) \\
 &subject\ to : && NR(\Gamma) \geq \omega \\
 & && NR(\Gamma) \leq M_{max} \\
 & && NR(\Gamma) \leq L_{max} \\
 & && S(\Gamma) \leq g_s
 \end{aligned} \tag{2-1}$$

โดยที่ Γ คือ การจำแนกประเภทข้อมูลของฟัซซี่

$NR(\Gamma)$ คือ กฎฟัซซี่

2.1.4 อัลกอริทึม Bayesian Network

การเรียนรู้แบบเบย์ที่ใช้องค์ความรู้ที่เกี่ยวกับทฤษฎีความน่าจะเป็น สถิติหรือการอนุมานมาใช้กับทฤษฎีและแบบจำลองของกราฟความน่าจะเป็นหรือที่ว่า Bayes Net

ซึ่งเป็นวิธีการเรียนรู้ที่ลดข้อจำกัดของการเรียนรู้แบบง่ายในสมมติฐานความไม่ขึ้นต่อกันระหว่างคุณสมบัติ แสดงได้ดังสมการที่)2-2) [5]

$$P(X|Y, Z) = P(X|Z) \quad)2-2)$$

โดยที่ P(X) คือ ค่าความน่าจะเป็นเหตุการณ์ X

P(Y) คือ ค่าความน่าจะเป็นเหตุการณ์ Y

P(Z) คือ ค่าความน่าจะเป็นเหตุการณ์ Z

P(X|Z) คือ ความน่าจะเป็นแบบมีเงื่อนไขที่เหตุการณ์ X เกิดขึ้น เมื่อเกิดเหตุการณ์ Z แล้ว

P(X|Y, Z) คือ ความน่าจะเป็นแบบมีเงื่อนไขที่เหตุการณ์ X เกิดขึ้น เมื่อเกิดเหตุการณ์ Y, Z แล้ว

2.1.5 อัลกอริทึม Split-Point and Attribute-Reduction Classifier (SPAARC)

ถูกพัฒนามาจากอัลกอริทึมต้นไม้ตัดสินใจ (Decision Trees) ซึ่งเป็นวิธีการหาค่าข้อมูล และค้นหาความรู้ที่ได้รับ ความนิยม โดยสามารถดึงข้อมูลที่ซ่อนอยู่ออกจากชุดข้อมูลที่ประกอบด้วยแอตทริบิวต์ทั้งแบบระบุชื่อ และแบบตัวเลข ซึ่งต้องทดสอบในทุกแอตทริบิวต์ที่โหนดต้นไม้ทุกโหนดอาจทำให้เสียเวลาในการรอกอย วิธี SPAARC เสนอการเร่งกระบวนการเหนี่ยวนำแผนผังต้นไม้ตัดสินใจ ซึ่งประกอบด้วย 2 องค์ประกอบ ได้แก่ การสุ่มตัวอย่างจุดแยกของโหนดต้นไม้แอตทริบิวต์ที่เป็นตัวเลข และการปรับพื้นที่การเลือกแอตทริบิวต์ของโหนดแบบไดนามิก ด้วยวิธีการเหล่านี้สามารถนำไปใช้กับอัลกอริทึมแผนผังการตัดสินใจได้เกือบทุกชนิด [6]

2.1.6 อัลกอริทึม ForEx++

ถูกพัฒนามาจากอัลกอริทึมต้นไม้ตัดสินใจเป็นที่นิยมใช้ในปัญหาสำหรับการทำนายและการค้นพบกฎการจำแนกประเภทลอลจิก โดยอัลกอริทึมป่าสุ่ม (Random Forest) ที่ถูกพัฒนามาจากต้นไม้ตัดสินใจนั้น มักถูกสร้างขึ้นเพื่อให้ได้ประสิทธิภาพการทำนายที่ดีขึ้น ซึ่งมีศักยภาพสูงในการค้นพบองค์ความรู้ อย่างไรก็ตาม อัลกอริทึมป่าสุ่มมักจะสร้างกฎจำนวนมากเกินความจำเป็น จึงได้นำเสนออัลกอริทึมใหม่ชื่อว่า ForEx++ สำหรับการวิเคราะห์ในชุดข้อมูลอิสระแยกกฎที่ให้ผลลัพธ์ใน

การวิเคราะห์ข้อมูลที่ค่อนข้างแม่นยำ ครอบคลุม รัศมมากกว่าอัลกอริทึมป่าสุ่ม แสดงได้ดังสมการที่)2-3) [7]

$$R_{ForEx++} = U_{C_k \in C, \forall C_k} R_{Avg, C_k}^{Acc, Cov, Len} \quad (2-3)$$

โดยที่ $R_{ForEx++}$ คือ กฎที่ถูกสร้างขึ้นมาโดยเรียกขอฟอร์เอ็กซ์พลัสพลัส

$R_{Avg, C_k}^{Acc, Cov, Len}$ คือ การอินเตอร์เซกชันกันของ

ค่าความแม่นยำ (Accuracy) ความครอบคลุม (Coverage) ความยาว (Length) ค่าเฉลี่ย (Average) ค่าคลาส (Class) ตามลำดับ

2.2 งานวิจัยที่เกี่ยวข้อง

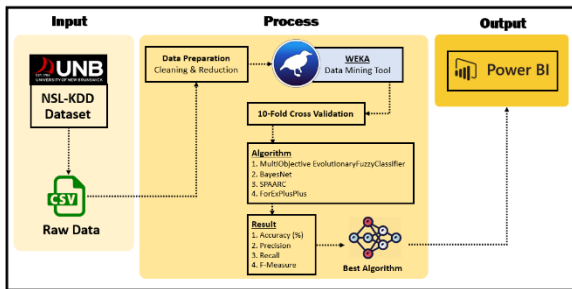
Qasim and Alsadi [8] ได้ทำการศึกษางานวิจัยเรื่องการตรวจจับการโจมตีของภัยคุกคามมัลแวร์เวิร์ม (Worm) ที่ผู้โจมตีเครือข่ายใช้ในการบุกรุกเครือข่าย ในบทความนี้ได้นำเสนอระบบตรวจจับที่ตรวจจับมัลแวร์เวิร์ม โดยใช้ชุดข้อมูลจากเว็บไซต์ของแคสเปอร์สกี (Kaspersky) และแมคอาฟี (McAfee) ตัวระบบได้จำลองสภาพแวดล้อมแบบไฮสปีด ประกอบด้วยอัลกอริทึมไฮบริด 2 แบบ ได้แก่ อัลกอริทึม Hybrid Naive Bayes & K-Mean และอัลกอริทึม Naive Bayes ผลการวิจัยพบว่า อัลกอริทึม NBKM มีความแม่นยำเท่ากับ 88% และอัลกอริทึม NB Classifier มีความแม่นยำเท่ากับ 81% ของอัตรา การตรวจจับ และทั้งสองอัลกอริทึมมีอัตราการแจ้งเตือนที่ผิดพลาด ซึ่งคิดได้เท่ากับ 21% ของข้อมูลในประสิทธิภาพของทั้งสองรุ่น

Purnachary and et al. [9] ได้ทำการศึกษาข้อมูลเพื่อวิเคราะห์ประสิทธิภาพของวิธีการจำแนกประเภทอัลกอริทึมตระกูล Bayes คือ อัลกอริทึม Bayesian Network และอัลกอริทึม Naive Bayes ใช้โปรแกรมวิเคราะห์ข้อมูล WEKA โดยใช้ชุดข้อมูลจาก Bank Marketing ที่ดึงมาจากเว็บไซต์ UCI Repository เทียบตัวแยกประเภทอัลกอริทึม Bayes 2 ชนิด คือ อัลกอริทึม Bayesian Network และอัลกอริทึม Naive Bayes ถูกนำมาใช้เพื่อเปรียบเทียบความถูกต้องของการจัดประเภทได้อย่างถูกต้อง อินสแตนซ์ที่เกี่ยวข้องกับชุดข้อมูล การตลาดของธนาคารพบว่า อัลกอริทึม Bayesian Network มีค่าความแม่นยำเท่ากับ 81.49% ซึ่งมากกว่า

อัลกอริทึม Naive Bayes ที่ค่าความแม่นยำเท่ากับ 78%
สรุปความแม่นยำอัลกอริทึม Bayesian Network ดีที่สุด

3. วิธีการดำเนินการวิจัย

จากรูปแสดงภาพกรอบแนวความคิดการดำเนินงานวิจัยในการจำแนกรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมืองข้อมูล แสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวความคิดการดำเนินงานวิจัย

จากภาพที่ 1 แสดงกรอบแนวคิดในการทำวิจัย โดยได้แบ่งการทำงานออกเป็น 3 ส่วน ได้แก่ 1. การรวบรวมข้อมูล 2. การเตรียมข้อมูล (Data Preparation) โดยใช้ข้อมูลจากเว็บไซต์ NSL-KDD จำนวน 50,000 เรคคอร์ด แบ่งเป็น 42 แอตทริบิวต์ ซึ่งทั้งหมดได้ผ่านกระบวนการทำความสะอาดข้อมูลแล้ว จากนั้นเข้าสู่กระบวนการเรียนรู้และทดสอบ โดยใช้หลักการ 10-Fold Cross Validation โดยใช้เครื่องมือวิเคราะห์เหมืองข้อมูล WEKA ในการเปรียบเทียบประสิทธิภาพของ 4 อัลกอริทึมต่าง ๆ คือ อัลกอริทึม Bayesian Network อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier อัลกอริทึม SPAARC และอัลกอริทึม ForEx++ 3. ผลลัพธ์เป็นการเปรียบเทียบประสิทธิภาพของแต่ละอัลกอริทึมค้นหาอัลกอริทึมที่ดีที่สุดของรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่าย (Best Algorithm) และนำเสนอผ่าน โปรแกรมธุรกิจอัจฉริยะ

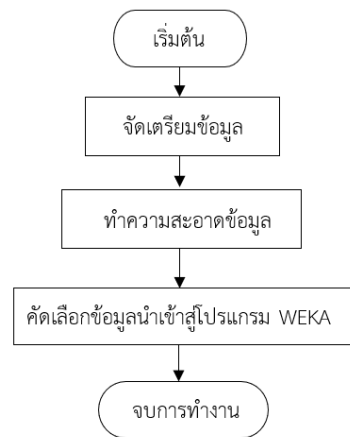
3.1 การรวบรวมข้อมูล

ผู้วิจัยได้เลือกใช้ชุดข้อมูล NSL-KDD ที่เป็นชุดข้อมูลที่ใช้สำหรับวิเคราะห์รูปแบบของการโจมตี และพัฒนามาจากชุดข้อมูล KDD-Cup99 ปรับปรุงจากข้อผิดพลาดหลายด้าน เช่น ข้อมูลที่ซ้ำ ๆ ทำให้เกิดการสับสนของการประมวลผลข้อมูล และตัวข้อมูลมีมากเกินไปจนจำเป็น โดย

ชุดข้อมูล NSL-KDD เป็นที่ยอมรับระดับสากลในหลายงานวิจัยด้วยผลลัพธ์ที่แม่นยำมากกว่า ทางผู้วิจัยจึงเลือกนำชุดข้อมูลมาใช้สำหรับการทำวิจัยในครั้งนี้

3.2 การเตรียมข้อมูล (Data Preparation)

3.2.1 กระบวนการที่ผู้วิจัยต้องกระทำกับข้อมูลดิบ (Raw Data) ที่ได้รับมาปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสมก่อนที่จะนำไปโหลดเข้าฐานข้อมูล หรือนำไปวิเคราะห์ต่อไป การจัดเตรียมข้อมูลเป็นขั้นตอนขั้นต้นที่สำคัญอย่างมากซึ่งประกอบด้วย 3 ขั้นตอน ได้แก่ การทำความสะอาดข้อมูล การลดขนาดข้อมูล และการคัดเลือกข้อมูลเพื่อนำเข้าสู่กระบวนการจัดทำเหมืองข้อมูลด้วยโปรแกรม WEKA แสดงดังภาพที่ 2



ภาพที่ 2: แสดงขั้นตอนการจัดเตรียมข้อมูลและการคัดเลือกข้อมูล

3.2.2 ข้อมูลที่นำมาใช้ในงานวิจัยครั้งนี้ เป็นชุดข้อมูล โดยได้ดาวน์โหลดมาจากทางเว็บไซต์ University of New Brunswick ข้อมูลที่ใช้จากฐานข้อมูล NSL-KDD เกี่ยวกับการโจมตีบุกรุกเครือข่ายที่เป็นเวอร์ชันอัปเดตข้อมูลมาจาก KDD Cup'99 โดยจะขจัดจำนวนข้อมูลที่ซ้ำซ้อนลดลง และเพิ่มจำนวนประเภทการบุกรุกต่าง ๆ หลากหลายเทคนิคมากขึ้น และผู้วิจัยได้คัดเลือกข้อมูลจำนวน 50,000 เรคคอร์ด และ 42 แอตทริบิวต์ ส่วนคลาสผลลัพธ์ประกอบด้วยจำนวน 22 คลาส แสดงดังภาพที่ 3

duration	protocol	service	flag	src_bytes	dst_bytes	land	wrong_fr	urgent	hot	num_failed	logged_in
0	tcp	ftp_data	SF	491	0	0	0	0	0	0	0
0	udp	other	SF	146	0	0	0	0	0	0	0
0	tcp	private	SF	0	0	0	0	0	0	0	0
0	tcp	http	SF	232	8153	0	0	0	0	0	1
0	tcp	http	SF	199	420	0	0	0	0	0	1
0	tcp	private	REJ	0	0	0	0	0	0	0	0
0	tcp	private	SF	0	0	0	0	0	0	0	0
0	tcp	private	SF	0	0	0	0	0	0	0	0
0	tcp	remote_js	SF	0	0	0	0	0	0	0	0
0	tcp	private	SF	0	0	0	0	0	0	0	0
0	tcp	private	REJ	0	0	0	0	0	0	0	0
0	tcp	private	SF	0	0	0	0	0	0	0	0
0	tcp	http	SF	287	2251	0	0	0	0	0	1
0	tcp	ftp_data	SF	334	0	0	0	0	0	0	1
0	tcp	name	SF	0	0	0	0	0	0	0	0

ภาพที่ 3: ตัวอย่างชุดข้อมูลที่เข้าสู่กระบวนการจัดทำเหมืองข้อมูล

3.2.3 ข้อมูลที่นำเข้า (Input Data) เป็นข้อมูลจากชุดฐานข้อมูล NSL-KDD ที่เลือกมาเพื่อนำไปสร้างแบบจำลองในโปรแกรมวิเคราะห์ข้อมูล WEKA โดยข้อมูลดังกล่าวจะถูกนำไปสร้างแบบจำลองที่มีคุณลักษณะเกี่ยวกับชื่อข้อมูล และชนิดข้อมูลจำนวน 42 แอตทริบิวต์ แสดงดังตารางที่ 1 และข้อมูลคลาสจำนวน 22 คลาส แสดงดังตารางที่ 2

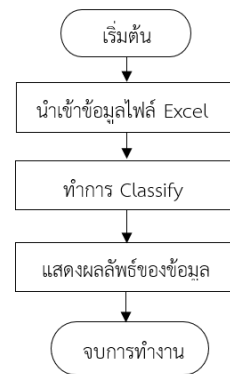
ตารางที่ 1: แอตทริบิวต์ข้อมูล NSL-KDD

ลำดับ	ชื่อข้อมูล	ชนิดข้อมูล
1	Duration	Continuous
2	Protocol_type	Symbolic
3	Service	Symbolic
4	Flag	Symbolic
5	Src_bytes	Continuous
6	Dst_bytes	Continuous
7	Land	Symbolic
8	Wrong_fragment	Continuous
9	Urgent	Continuous
10	Hot	Continuous
11	Num_failed_logins	Continuous
12	Logged_in	Symbolic
13	Num_compromised	Continuous
14	Root_shell	Continuous
15	Su_attempted	Continuous
16	Num_root	Continuous
17	Num_file_creations	Continuous
18	Num_shells	Continuous
19	Num_access_files	Continuous
20	Num_outbound_cmds	Continuous
21	Is_host_login	Symbolic
22	Is_guest_login	Symbolic
23	Count	Continuous
24	Srv_count	Continuous
25	Serror_rate	Continuous
26	Srv_serror_rate	Continuous
27	Error_rate	Continuous
28	Srv_rerror_rate	Continuous
29	Same_srv_rate	Continuous
30	Diff_srv_rate	Continuous
31	Srv_diff_host_rate	Continuous
32	Dst_host_count	Continuous
33	Dst_host_srv_count	Continuous
34	Dst_host_same_srv_rate	Continuous
35	Dst_host_same_src_port_rate	Continuous
36	Dst_host_diff_srv_rate	Continuous
37	Dst_host_serror_rate	Continuous
38	Dst_host_srv_diff_host_rate	Continuous
39	Dst_host_rerror_rate	Continuous
40	Dst_host_srv_rerror_rate	Continuous
41	Dst_host_srv_serror_rate	Continuous
42	Labels	No

ตารางที่ 2: ข้อมูลคลาส (Class)

ลำดับ	ชื่อคลาส	ลำดับ	ชื่อคลาส
1	Normal	12	Guess password
2	Neptune	13	FTP write
3	Warez client	14	Multi hop
4	IP sweep	15	Root kit
5	Port sweep	16	Buffer overflow
6	Teardrop	17	Imap
7	Nmap	18	Warez master
8	Satan	19	PHF
9	Smurf	20	Land
10	Pod	21	Load module
11	Back	22	Spy

3.2.4 การสร้างแบบจำลองโดยใช้โปรแกรมวิเคราะห์เหมือนข้อมูล WEKA ผ่านกระบวนการเรียนรู้และทดสอบโดยใช้หลักการ 10-Fold Cross Validation เป็นวิธีที่นิยมใช้ในการทดสอบประสิทธิภาพของอัลกอริทึม และแบบจำลองทั้ง 4 อัลกอริทึม แสดงดังภาพที่ 4



ภาพที่ 4: กระบวนการทดสอบข้อมูลในโปรแกรมWEKA

4. ผลการดำเนินงาน

ผลลัพธ์ของการเปรียบเทียบประสิทธิภาพของแบบจำลองการจำแนกรูปแบบแพ็กเก็ตที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมือนข้อมูลที่ใช้ทั้ง 4 อัลกอริทึม ผลการวิจัยพบว่า SPAARC มีประสิทธิภาพในการจำแนกที่มีประสิทธิภาพสูงที่สุด โดยแสดงดังตารางที่ 3 และแสดงดังตารางที่ 4

ตารางที่ 3: ผลการเปรียบเทียบค่าความถูกต้อง (Accuracy)

Algorithms	Accuracy (%)
BayesNet	97.09
ForEx++	99.53
SPAARC	99.58

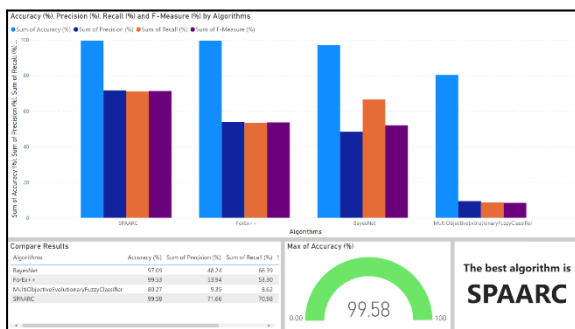
Multi-objective Evolutionary Fuzzy Classifier	80.27
---	-------

ตารางที่ 4: ผลการเปรียบเทียบค่าทางสถิติทั้ง 4 อัลกอริทึม

Algorithms	Precision (%)	Recall (%)	F-Measure (%)
BayesNet	48.24	66.39	51.95
ForEx++	53.94	53.30	53.60
SPAARC	71.66	70.98	71.28
Multi-objective Evolutionary Fuzzy Classifier	9.35	8.62	8.38

4.1 การแสดงผลผ่านโปรแกรม Power BI

จากการเปรียบเทียบผลลัพธ์ที่ได้การจำแนกรูปแบบแพ็คเกจที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมืองข้อมูลนำเสนอผ่านโปรแกรม Power BI แสดงดังภาพที่ 5



ภาพที่ 5: ผลการเปรียบเทียบแบบจำลองทั้ง 4 แบบ

5. สรุปผล และอภิปรายผล

งานวิจัยครั้งนี้ผู้วิจัยได้นำเสนอการเปรียบเทียบประสิทธิภาพของแบบจำลองการจำแนกรูปแบบแพ็คเกจที่ใช้โจมตีระบบเครือข่ายด้วยเทคนิคเหมืองข้อมูลที่ใช้ อัลกอริทึม Bayesian Network อัลกอริทึม Multi-objective Evolutionary Fuzzy Classifier อัลกอริทึม SPAARC และอัลกอริทึม ForEx++ โดยใช้ข้อมูลจากเว็บไซต์ NSL-KDD Dataset จำนวน 50,000 เรคคอร์ด แบ่งเป็น 42 แอดทริบิวต์ การแสดงผลลัพธ์ของคลาสผลลัพธ์จำนวน 22 คลาส เป็นชุดข้อมูลที่ใช้ในการเรียนรู้ และทดสอบด้วยวิธีการ 10-Fold Cross Validation ผลการวิจัยพบว่าอัลกอริทึม SPAARC มีประสิทธิภาพดี ที่สุดโดยมีค่าความถูกต้อง (Accuracy) เท่ากับ 99.58% ค่าความแม่นยำ (Precision) เท่ากับ 71.66% ค่าระลึ

(Recall) เท่ากับ 70.98% และค่าความถ่วงดุล (F-measure) เท่ากับ 71.28%

เอกสารอ้างอิง

- [1] ETDA, (2023), *Threat Statistics for the year 2022*. Retrieved date 25 February 2023, from <https://www.etcha.or.th/Our-Service/thaicert>.
- [2] ประเทือง วงษ์ทอง, วาฤทธิ์ กันแก้ว, อภิชนญาขวัญแก้ว, และ ยุวดี โฉมแดง. “การประยุกต์เทคนิคเหมืองข้อมูลค้นหาลักษณะนิสัยของผู้ประกอบอาชญากรรมด้านคอมพิวเตอร์” วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ, 2563.
- [3] อนันต์ชัย ชูติภาสเจริญ และจรัญ แสนราช. (2561). “การเปรียบเทียบประสิทธิภาพของอัลกอริทึมและการคัดเลือกคุณลักษณะที่เหมาะสมเพื่อการพยากรณ์โอกาสความสำเร็จในการโอนเงินข้ามประเทศของบุคคลทั่วไป” วารสารวิจัยมหาวิทยาลัยขอนแก่น. ปีที่ 6 ฉบับที่ 3 (กันยายน-ธันวาคม): 105-113.
- [4] F. Jimenez, C. Martinez, E. Marzano, J. T. Palma, G. Sanchez, and G. Sciavicco, (2019). “Multiobjective Evolutionary Feature Selection for Fuzzy Classification.” *IEEE Transactions on Fuzzy Systems*, 27(5), 1085–1099. doi.org/10.1109/TFUZZ.2019.2892363
- [5] J. Kuipers, P. Suter, G. Moffa S, (2022). “Efficient Sampling and Structure Learning of Bayesian Networks.” *Journal of Computational and Graphical Statistics*, 31(3), 639-650.
- [6] D. Yates, M. Z. Islam, and J. Gao, (2019). “SPAARC: A Fast Decision Tree Algorithm.” *Data Mining*, Vol.996, pp. 43–55, 2019.
- [7] M. N Adnan, and M. Z. Islam, (2017). “ForEx++: A New Framework for Knowledge Discovery from Decision Forests.” *Australasian Journal of Information Systems*, Vol.2, pp. 1-20, 2017. doi.org/10.3127/ajis.v21i0.1539
- [8] Osama Mobammed Qasim, Karim Hashim Al-Saedi. (2022). “Detection System for Detecting Worms using Hybrid Algorithm of Naïve Bayesian classifier and K-Means.” 2019 2nd International Conference on Engineering Technology and its Applications (IICETA). pp. 1–6, 2022.

- [9] M. Purnachary., et al. (2018). “Performance Analysis of Bayes Classification Algorithms in WEKA Tool using Bank Marketing Dataset” International Journal of Engineering Research in Computer Science and Engineering. Vol 5, Issue 2, February 2018

การจำแนกประเภทโรคติดต่อจากอาการตามมาตรฐาน ICD-10 ด้วยเทคนิคการเรียนรู้เชิงลึก Classifying Infectious Disease Symptoms based on ICD-10 Standards using Deep Learning

ดวงฤดี เพชรพิรุณ (Duangrudee Phetphirun)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856011@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อจำแนกข้อความประเภทโรคติดต่อจากอาการตามมาตรฐาน ICD-10 โดยทำการเปรียบเทียบประสิทธิภาพของเทคนิคการเรียนรู้เชิงลึก 2 อัลกอริทึม ได้แก่ อัลกอริทึมโครงข่ายประสาทเทียมแบบสั่งวัตนาการ และอัลกอริทึมหน่วยความจำระยะสั้นแบบยาว โดยใช้ชุดข้อมูลอาการของโรคติดต่อที่อยู่ในรูปแบบภาษาไทย ทั้งหมด 5,000 ข้อความ ผ่านกระบวนการเตรียมข้อมูล และแบ่งชุดข้อมูลออกเป็น 3 ส่วน ได้แก่ ชุดข้อมูลสอน ชุดข้อมูลตรวจสอบ และชุดข้อมูลทดสอบ จากนั้นทำการกำหนดและปรับค่าพารามิเตอร์ของแบบจำลอง เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยพิจารณาจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล ผลลัพธ์แสดงให้เห็นว่า อัลกอริทึมหน่วยความจำระยะสั้นแบบยาวเป็นเทคนิคที่ดีที่สุด โดยค่าความถูกต้องร้อยละ 96.76 ค่าความแม่นยำ 0.97 ค่าระลอก 0.97 และค่าการถ่วงดุล 0.97 ตามลำดับ

คำสำคัญ: การเรียนรู้เชิงลึก โรคติดต่อ การจำแนก อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว อัลกอริทึมโครงข่ายประสาทเทียมแบบสั่งวัตนาการ

Abstract

This paper presents the classification of infectious disease symptoms based on ICD-10 Standards aiming

to compare two deep learning, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) using 5,000 messages of infectious disease symptoms in Thai. The collected data was preprocessed using text filtering before entering the text classification model. The dataset was divided into three parts, training, validation, and testing. The algorithm's parameters were defined and adjusted to compare performance in terms of precision, recall, accuracy, and f-measure. The comparative results showed that LSTM best performs with an accuracy 96.76%, precision 0.97, recall 0.97, and f-measure 0.97.

Keywords: Deep learning, Contagious Disease, Classification, Convolutional Neural Network, Long Short-Term Memory.

1. บทนำ

การวินิจฉัยโรค เป็นกระบวนการหนึ่งที่แพทย์ต้องทำ เพื่อให้ทราบถึงสาเหตุของโรค ในปัจจุบันการวินิจฉัยโรคที่คลาดเคลื่อนนับเป็นปัญหาสำคัญในวงการแพทย์ สาเหตุจากหลายปัจจัย เช่น ข้อจำกัดของความรู้และความสามารถในการเก็บรวบรวมข้อมูลที่ไม่เพียงพอต่อการวินิจฉัย เป็นต้น ในการระบุโรค จำเป็นต้องอาศัยข้อมูลของผู้ป่วย ได้แก่ ประวัติการเจ็บป่วย อาการแสดงจากข้อความบอกอาการสำคัญ และต้องอาศัยปัจจัยที่สำคัญ ได้แก่ บุคลากร และเวลา หากมีการนำเทคโนโลยีเข้ามาช่วยในการประมวลผล

จะสามารถลดจำนวนปัจจัยดังกล่าวเพื่อให้เกิดประโยชน์ และได้ผลลัพธ์ที่ถูกต้องได้ แต่ข้อความบอกอาการสำคัญมักถูกเก็บอยู่ในรูปแบบภาษาธรรมชาติที่มีความแตกต่างไปตามลักษณะโครงสร้างของภาษา จำเป็นต้องมีการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก มาช่วยในการจำแนกประเภทโรคจากอาการเจ็บป่วย ซึ่งเป็นอีกหนึ่งวิธีที่ทำให้ได้มาซึ่งผลลัพธ์ที่น่าเชื่อถือ และยังมีผู้วิจัยนำเสนอการวิเคราะห์ข้อมูลโดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยอัลกอริทึมต่าง ๆ เพื่อวิเคราะห์ข้อมูล ค้นหารูปแบบโครงสร้าง และความสัมพันธ์ต่าง ๆ ที่ซ่อนอยู่ในชุดข้อความ เช่น การนำอัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network) มาใช้ในการจัดประเภทข้อความเพื่อวิเคราะห์ความรู้สึก [1] ใช้อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory) สำหรับวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์ [2] การพัฒนาแบบจำลองการวิเคราะห์ความรู้สึกบนสื่อสังคมออนไลน์ไทยโดยใช้เทคนิค Long Short-Term Memory [3] การจัดประเภทเอกสารทางการแพทย์โดย Convolutional Neural Network [4] และอัลกอริทึม Long Short-Term Memory มาใช้สำหรับจำแนกชุดข้อมูลบทความข่าว [5] เป็นต้น หากสามารถนำเทคนิคดังกล่าวมาประยุกต์ใช้ในการจำแนกประเภทโรคจากอาการเจ็บป่วยที่เกิดขึ้นได้ จะสามารถเพิ่มความรวดเร็ว แม่นยำ และความถูกต้องในการวินิจฉัยโรคของแพทย์ได้ดียิ่งขึ้น

ดังนั้น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อใช้อัลกอริทึมการเรียนรู้เชิงลึกในการจำแนกประเภทโรคคิดต่อจากอาการตามมาตรฐาน ICD-10 และทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมโดยพิจารณาค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 อัลกอริทึมการเรียนรู้เชิงลึก

สำหรับงานวิจัยนี้ได้มีการนำอัลกอริทึมการเรียนรู้เชิงลึก มาใช้ในการสร้างแบบจำลองจำนวน 2 อัลกอริทึม ได้แก่

2.1.1 อัลกอริทึม Convolutional Neural Network เป็นอัลกอริทึมที่มีหลายเลเยอร์ และมีโครงสร้างเฉพาะตัว โดยปกติแล้วจะนำไปใช้กับข้อมูลประเภทรูปภาพ เมื่อนำอัลกอริทึมมาใช้กับข้อมูลที่เป็นข้อความแทนข้อมูลที่เป็นรูปภาพนั้น จำเป็นต้องใช้การดำเนินการระหว่างข้อมูล 1 มิติ โดยเริ่มจากการรับ Input กำหนดค่าในเคอร์เนล หรือตัวกรอง เพื่อตรวจสอบ และทำการแยกองค์ประกอบต่อมาทำการกรองค่า เพื่อลดขนาดหรือมิติ หรือที่เรียกว่า Pooling โดยจะใช้ Max Pooling กรองเพื่อหาค่าที่สูงที่สุดในบริเวณที่ตัวกรองทาอยู่มาเป็นผลลัพธ์ของการทำนาย [6]

2.1.2 อัลกอริทึม Long Short-Term Memory เป็นอัลกอริทึมที่มีลักษณะการทำงานภายในโดยการเพิ่ม Cell State เพื่อเพิ่มความจำข้อมูลลำดับก่อนหน้าได้ดี โดยอัลกอริทึม Long Short Term Memory สามารถช่วยในการแก้ปัญหาของข้อมูลสำคัญที่หายไปบางส่วนได้ โดยจะทำหน้าที่เหมือนกับ Memory ในการบันทึกข้อมูลในแต่ละลำดับ และมีการควบคุม Gate หลักที่เป็นบทบาทสำคัญ 3 Gate โดย Input gate เป็นตัวตัดสินใจว่าจะเขียนข้อมูลลง Cell State หรือไม่โดย Input Gate จะมี Sigmoid Layer ที่ให้ค่าออกมาระหว่าง 0 กับ 1 Forget Gate เป็นส่วนที่ตัดสินใจว่าข้อมูลใดที่จะเก็บไว้และข้อมูลใดที่จะทิ้ง ส่วน Output Gate เป็นส่วนของการประมวลผลผลลัพธ์ [7]

2.2 งานวิจัยที่เกี่ยวข้อง

Muh Amin Nurrohmat and Azhari SN [8] ใช้วิธี Long Short-Term Memory วิเคราะห์ความรู้สึกของบทวิจารณ์นวนิยายภาษาชาวอินโดนีเซีย ในกระบวนการทดสอบ จะเปรียบเทียบวิธี Long Short-Term Memory กับวิธีนาอิวเบย์ (Naïve Bayes(ผลลัพธ์ที่ได้พบว่า วิธี Long Short-Term Memory มีผลความถูกต้องที่ดีกว่าวิธี Naïve Bayes โดยมีค่าความถูกต้องร้อยละ 72.85 เมื่อเทียบกับผลลัพธ์ของความถูกต้องของวิธีการ Naïve Bayes มีค่าความถูกต้องร้อยละ 67.88

Dr. G. S. N. Murthy at al. [9] เสนอวิธีการจัดประเภทความคิดเห็นเชิงบวกและเชิงลบเกี่ยวกับทัศนคติของผู้คนที่ม่ต่อผลิตภัณฑ์ การเมือง กิจกรรมทางสังคม โดยใช้ อัลกอริทึม Long Short-Term Memory ผลลัพธ์ที่ได้ให้ค่าความถูกต้องร้อยละ 85

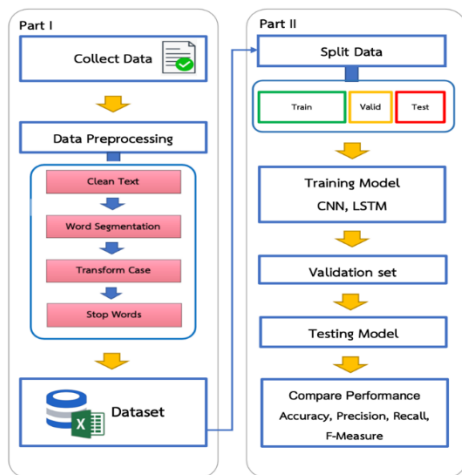
Hamed Jelodar at al. [10] ทดลองเกี่ยวกับ การจำแนกประเภทของความคิดเห็นเกี่ยวกับ COVID-19 โดยใช้ อัลกอริทึม Long Short-Term Memory และอัลกอริทึมการเรียนรู้ของเครื่องอื่น ๆ ผลลัพธ์พบว่า อัลกอริทึม Long Short-Term Memory ให้ค่าความถูกต้องร้อยละ 81.15

Pisit B. at al. [11] ศึกษาการวิเคราะห์ความรู้สึกทางอารมณ์ในบทความแนะนำสินค้าออนไลน์ โดยใช้เทคนิคการตัดคำภาษาไทยชื่อ Attacut สามารถนำไปใช้ในการวิเคราะห์ตัดคำภาษาไทยได้อย่างแม่นยำ และรวดเร็ว และเพิ่มประสิทธิภาพในการสร้างชุดข้อมูลสอน และชุดข้อมูลทดสอบที่ในการสร้างแบบจำลอง ให้มีความแม่นยำยิ่งขึ้น

Agung Triayudi [12] พัฒนาแบบจำลองสำหรับประมวลผลภาษาธรรมชาติโดยใช้ Convolutional Neural Network เพื่อจำแนกข้อมูลเกี่ยวกับแผนฉุกเฉิน ผลลัพธ์พบว่า Convolutional Neural Network เหมาะสมที่สุดในการจำแนกข้อมูลนี้ โดยมีค่าความถูกต้องถึงร้อยละ 98

3. วิธีการดำเนินการวิจัย

ผู้วิจัยศึกษาความรู้ และทฤษฎีที่เกี่ยวข้องกับงานวิจัยรวมทั้งศึกษาเครื่องมือที่นำมาใช้ในการสร้างแบบจำลอง โดยมีกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากภาพที่ 1 กรอบแนวคิดของงานวิจัยนี้ ผู้วิจัยได้แบ่งการทำงานเป็น 2 ส่วน ได้แก่ ขั้นตอนการรวบรวมและเตรียมข้อมูล และการสร้างแบบจำลองและการประเมินผล

3.1 การรวบรวมและเตรียมข้อมูล

ผู้วิจัยทำการรวบรวมข้อมูลที่เป็นอาการของโรคติดต่อจากเว็บไซต์หน่วยงานสาธารณสุข และเว็บไซต์สาธารณะจำนวน 5,000 ชุด แบ่งเป็นข้อมูลอาการของโรค 10 โรค โรคละ 500 ชุด ตัวอย่างข้อมูลแสดงดังตารางที่ 1

ตารางที่ 1: ตัวอย่างข้อมูลอาการของโรคติดต่อ

Symptom	Diseases
อักเสบที่ท่อปัสสาวะ (Urethra) & ปากมดลูก (Cervix)	หนองในเทียม
#ไข้สูงลอย#เบื่ออาหาร#อาเจียน#ปวดหัว#หน้าแดงตัวแดง#มีผื่น	ไข้เลือดออก
ผื่นขึ้นตามตัว : เป็นผื่นแดง -- จากนั้นจะเป็นตุ่มใสก่อนหน้านั้นเกิดอาการเป็น ไข้อ่อนเพลีย	อีสุกอีใส
เบื่ออาหารคลื่นไส้ปวดศีรษะรุนแรง@ไข้สูงมากอ่อนเพลียอย่างเฉียบพลัน	ไข้หวัดใหญ่
ไข้สูง อ่อนเพลีย ไอเยอะมาก และเจ็บคอ การรับรสผิดปกติออกซิเจน(Oxygen)ในเลือดต่ำ	โควิด 19

จากตารางที่ 1 ข้อมูลอาการโรคติดต่อที่รวบรวมยังมีเครื่องหมายวรรคตอน สัญลักษณ์ และคำศัพท์ภาษาอังกฤษที่มีตัวพิมพ์ใหญ่และตัวพิมพ์เล็กผสมกัน จึงต้องนำมาผ่านกระบวนการเตรียมข้อมูล (Data Preprocessing) ได้แก่ การทำความสะอาดข้อมูล การตัดคำภาษาไทย การแปลงตัวพิมพ์ใหญ่เป็นตัวพิมพ์เล็ก การลบคำฟุ่มเฟือยภาษาไทย และการกำหนดป้ายกำกับชื่อโรคด้วยรหัส ICD-10 แสดงดังตารางที่ 2 ถึง 6

ตารางที่ 2: การทำความสะอาดข้อมูล

Symptom	Diseases
อักเสบที่ท่อปัสสาวะ Urethraปากมดลูก Cervix	หนองในเทียม
ไข้สูงลอยเบื่ออาหารอาเจียนปวดหัวหน้าแดงตัวแดงมีผื่น	ไข้เลือดออก
ผื่นขึ้นตามตัวเป็นผื่นแดงจากนั้นจะเป็นตุ่มใส	อีสุกอีใส
เบื่ออาหารคลื่นไส้ปวดศีรษะรุนแรงไข้สูงมากอ่อนเพลียอย่างเฉียบพลัน	ไข้หวัดใหญ่
ไข้สูง อ่อนเพลีย ไอเยอะมาก และเจ็บคอ การรับรสผิดปกติออกซิเจนOxygenในเลือดต่ำ	โควิด 19

จากตารางที่ 2 แสดงข้อมูลที่ผ่านมากระบวนการทำความเข้าใจ
 สะอาดข้อมูลโดยการเอาเครื่องหมายวรรคตอน และ
 สัญลักษณ์ต่าง ๆ ออก

ตารางที่ 3: การตัดคำภาษาไทย

Symptom	Diseases
['อึกเสบ', 'ที่', 'ท่อ', 'ปัสสาวะ', 'Urethra', 'ปาก', 'มดลูก', 'Cervix']	หนองในเทียม
['ไอ้', 'สูง', 'ลอย', 'เบือ', 'อาหาร', 'อาเจียน', 'ปวด', 'หัว', 'หน้า', 'แดง', 'ตัว', 'แดง', 'มี', 'ผื่น']	ไข้เลือดออก
['ผื่น', 'ขึ้น', 'ตาม', 'ตัว', 'เป็น', 'ผื่น', 'แดง', 'จาก', 'นั้น', 'จะ', 'เป็น', 'ตุ่ม', 'ใส']	อีสุกอีใส
['เบือ', 'อาหาร', 'คลื่นไส้', 'ปวดศรีษะ', 'รุนแรง', 'ไอ้', 'สูง', 'มาก', 'อ่อนเพลีย', 'อย่าง', 'เจ็บปวกัน']	ไข้หวัดใหญ่
['ไอ้', 'สูง', 'อ่อนเพลีย', 'ไอ', 'เยอะ', 'มาก', 'และ', 'เจ็บ', 'คอ', 'การ', 'รับ', 'รส', 'ผิด', 'ปกติ', 'ออกซิเจน', 'Oxygen', 'ใน', 'เลือด', 'ต่ำ']	โควิด 19

จากตารางที่ 3 เป็นการแสดงข้อมูลที่ผ่านมากระบวนการ
 ตัดคำภาษาไทย โดยใช้อัลกอริทึม Attacut เป็นอัลกอริทึม
 ที่สามารถตัดคำภาษาไทยได้อย่างถูกต้องแม่นยำ และ
 รวดเร็ว

ตารางที่ 4: การแปลงตัวพิมพ์ใหญ่เป็นเป็นตัวพิมพ์เล็ก

Symptom	Diseases
['อึกเสบ', 'ที่', 'ท่อ', 'ปัสสาวะ', 'urethra', 'ปาก', 'มดลูก', 'cervix']	หนองในเทียม
['ไอ้', 'สูง', 'ลอย', 'เบือ', 'อาหาร', 'อาเจียน', 'ปวด', 'หัว', 'หน้า', 'แดง', 'ตัว', 'แดง', 'มี', 'ผื่น']	ไข้เลือดออก
['ผื่น', 'ขึ้น', 'ตาม', 'ตัว', 'เป็น', 'ผื่น', 'แดง', 'จาก', 'นั้น', 'จะ', 'เป็น', 'ตุ่ม', 'ใส']	อีสุกอีใส
['เบือ', 'อาหาร', 'คลื่นไส้', 'ปวดศรีษะ', 'รุนแรง', 'ไอ้', 'สูง', 'มาก', 'อ่อนเพลีย', 'อย่าง', 'เจ็บปวกัน']	ไข้หวัดใหญ่
['ไอ้', 'สูง', 'อ่อนเพลีย', 'ไอ', 'เยอะ', 'มาก', 'และ', 'เจ็บ', 'คอ', 'การ', 'รับ', 'รส', 'ผิด', 'ปกติ', 'Oxygen', 'ใน', 'เลือด', 'ต่ำ']	โควิด 19

จากตารางที่ 4 เป็นการแปลงตัวอักษรภาษาอังกฤษ
 ตัวพิมพ์ใหญ่ให้กลายเป็นตัวอักษรพิมพ์เล็ก เพื่อให้ง่าย
 ต่อการนำไปใช้

ตารางที่ 5: การลบคำฟุ่มเฟือยภาษาไทย

Symptom	Diseases
['อึกเสบ', 'ท่อ', 'ปัสสาวะ', 'urethra', 'มดลูก', 'cervix']	หนองในเทียม
['ไอ้', 'ลอย', 'เบืออาหาร', 'อาเจียน', 'ปวดหัว', 'หน้าแดง', 'ตัว', 'แดง', 'ผื่น', 'จุด', 'เลือด', 'ผิวหนัง']	ไข้เลือดออก
['ผื่น', 'ตามตัว', 'ผื่น', 'แดง', 'ตุ่ม', 'ใส']	อีสุกอีใส
['เบืออาหาร', 'คลื่นไส้', 'ปวด', 'ศรีษะ', 'รุนแรง', 'ไอ้', 'อ่อนเพลีย', 'เจ็บปวกัน']	ไข้หวัดใหญ่
['ไอ้', 'อ่อนเพลีย', 'ไอ', 'เจ็บ', 'คอ', 'ระคาย', 'คอ', 'รับ', 'รส', 'ผิด', 'ปกติ', 'oxygen', 'เลือด', 'ต่ำ']	โควิด 19

จากตารางที่ 5 เป็นการลบคำฟุ่มเฟือยซึ่งไม่มีความสำคัญ
 ในประโยคออก โดยที่ข้อความยังคงสื่อความหมายเดิม

ตารางที่ 6: การกำหนดป้ายกำกับชื่อโรคด้วยรหัส ICD-10

Symptom	Label
['อึกเสบ', 'ท่อ', 'ปัสสาวะ', 'urethra', 'มดลูก', 'cervix']	A56
['ไอ้', 'ลอย', 'เบืออาหาร', 'อาเจียน', 'ปวดหัว', 'หน้าแดง', 'ตัว', 'แดง', 'ผื่น', 'จุด', 'เลือด', 'ผิวหนัง']	A90
['ผื่น', 'ตามตัว', 'ผื่น', 'แดง', 'ตุ่ม', 'ใส']	B01
['เบืออาหาร', 'คลื่นไส้', 'ปวด', 'ศรีษะ', 'รุนแรง', 'ไอ้', 'อ่อนเพลีย', 'เจ็บปวกัน']	J11
['ไอ้', 'อ่อนเพลีย', 'ไอ', 'เจ็บ', 'คอ', 'ระคาย', 'คอ', 'รับ', 'รส', 'ผิด', 'ปกติ', 'oxygen', 'เลือด', 'ต่ำ']	U07.1

จากตารางที่ 6 เป็นการกำหนดป้ายกำกับชุดข้อความ
 ด้วยรหัส ICD-10 ได้แก่ "A56", "A90", "B01", "J11"
 และ "U07.1" เพื่อเตรียมพร้อมสำหรับการสร้างโมเดล

3.2 การสร้างแบบจำลองการประเมินผล

ผู้วิจัยใช้แบบจำลองการเรียนรู้เชิงลึก สำหรับจำแนก
 อาการของโรคติดต่อ ได้แก่ อัลกอริทึม Convolutional
 Neural Network และ อัลกอริทึม Long Short-Term
 Memory เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลอง
 หลังจากผ่านกระบวนการเตรียมข้อมูลแล้ว แบ่งชุดข้อมูล
 ออกเป็น 3 ส่วน ได้แก่ ชุดข้อมูลสอน ชุดข้อมูลตรวจสอบ
 และชุดข้อมูลทดสอบ แล้วทำการปรับค่าพารามิเตอร์
 ของแต่ละอัลกอริทึมเพื่อหาค่าที่ดีที่สุด จากนั้นทำการ
 ประเมินประสิทธิภาพของอัลกอริทึม โดยพิจารณาจาก
 ค่าความถูกต้อง ค่าความแม่นยำ ค่าระลึก และค่าการอ้างดุล

เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองที่ดีที่สุด ซึ่งค่าพารามิเตอร์ที่นำมาพิจารณาแสดงดังตารางที่ 7 และ 8

ตารางที่ 7: พารามิเตอร์ของ Convolutional Neural Networks

Parameter	Value
Word Embedding	100, 200, 300
Number of Kernel	8, 16, 32, 64, 128, 256
Kernel Size	2, 3, 4, 5

จากตารางที่ 7 ผู้วิจัยได้ทำการปรับค่าพารามิเตอร์จำนวน 72 แบบ เพื่อทดสอบประสิทธิภาพ แบบจำลอง ได้แก่ Word Embedding มีค่าเท่ากับ 100, 200, 300 Number of Kernel มีค่าเท่ากับ 8, 16, 32, 64, 128, 256 และ Kernel Size มีค่าเท่ากับ 2, 3, 4, 5

ตารางที่ 8: พารามิเตอร์ของอัลกอริทึม Long Short-Term Memory

Parameter	Value
Word Embedding	100, 200, 300
LSTM_Out	16, 32, 64, 128, 256
Recurrent Dropout	0.0, 0.1, 0.2, 0.3, 0.4, 0.5

จากตารางที่ 8 ผู้วิจัยปรับค่าพารามิเตอร์จำนวน 90 แบบ เพื่อทดสอบประสิทธิภาพแบบจำลอง ได้แก่ Word Embedding เท่ากับ 100, 200, 300 LSTM_out เท่ากับ 16, 32, 64, 128, 256 Recurrent Dropout เท่ากับ 0.0, 0.1, 0.2, 0.3, 0.4, 0.5 ค่าพารามิเตอร์อื่นถูกตั้งค่าเป็นค่าคงที่ ตัวอย่างของพารามิเตอร์อื่น แสดงดังตารางที่ 9

ตารางที่ 9: ผู้วิจัยได้กำหนดค่าพารามิเตอร์อื่น ๆ

Parameter	Value
Batch size	100
Dropout	0.1
Validation Split	0.1
Learning Rate	0.00015
Epochs	100
Dense Layer Output parameter	10, SoftMax

จากตารางที่ 9 ผู้วิจัยกำหนดค่าพารามิเตอร์อื่น ได้แก่ Batch size เท่ากับ 100 Dropout เท่ากับ 0.1 Validation Split เท่ากับ 0.1 Learning Rate เท่ากับ 0.00015 Dense Layer Output เท่ากับ 10 (Active Learning = SoftMax) และ Epochs เท่ากับ 100

4. ผลการดำเนินงาน

ผลการทดสอบประสิทธิภาพของอัลกอริทึม แสดงดังตารางที่ 10 ถึง 11

ตารางที่ 10: ผลลัพธ์ที่ดีที่สุดของ Convolutional Neural Network

Parameter			Value			
Emb_dim	Num_kernel	Kernel_size	A	P	R	F
100	128	2	94.00	0.94	0.94	0.94
100	128	3	90.47	0.91	0.90	0.90
100	128	4	95.63	0.96	0.96	0.96
100	128	5	92.00	0.93	0.92	0.92
200	128	2	91.00	0.92	0.91	0.91
200	128	3	93.87	0.94	0.94	0.94
200	128	4	92.53	0.93	0.93	0.93
200	128	5	95.68	0.96	0.96	0.96
300	128	2	93.40	0.94	0.94	0.93
300	128	3	93.20	0.93	0.93	0.93
300	128	4	94.80	0.95	0.95	0.95
300	128	5	93.33	0.94	0.93	0.93

จากตารางที่ 10 ค่าพารามิเตอร์ที่ทำให้อัลกอริทึม Convolutional Neural Network มีประสิทธิภาพดีที่สุด ได้แก่ Word Embedding เท่ากับ 200 Number of Kernel เท่ากับ 128 และ Kernel Size เท่ากับ 5 ผลลัพธ์ที่ได้ ค่าความถูกต้องร้อยละ 95.68 ค่าความแม่นยำ 0.96 ค่าระลึก 0.96 และค่าการถ่วงดุล 0.96 ตามลำดับ

ตารางที่ 11: ผลลัพธ์ที่ดีที่สุดของ Long Short-Term Memory

Parameter			Value			
Emb_dim	LSTM_Out	Rec_dropout	A	P	R	F
200	256	0.0	96.33	0.96	0.96	0.96
200	256	0.1	95.60	0.95	0.96	0.96
200	256	0.2	96.33	0.96	0.96	0.96
200	256	0.3	96.71	0.97	0.97	0.97
200	256	0.4	96.33	0.96	0.96	0.96
200	256	0.5	95.13	0.95	0.95	0.95
300	256	0.0	95.87	0.96	0.96	0.96
300	256	0.1	96.60	0.97	0.97	0.96
300	256	0.2	96.20	0.97	0.97	0.97
300	256	0.3	96.33	0.96	0.96	0.96
300	256	0.4	96.76	0.97	0.97	0.97
300	256	0.5	96.56	0.96	0.97	0.97

จากตารางที่ 11 ค่าพารามิเตอร์ที่ทำให้อัลกอริทึม Long Short-Term Memory มีประสิทธิภาพดีที่สุด ได้แก่ Word Embedding เท่ากับ 300 LSTM_Out เท่ากับ 256 และ Recurrent Dropout เท่ากับ 0.4 ผลลัพธ์ที่ได้คือ ค่าความถูกต้องร้อยละ 96.76 ค่าความแม่นยำ 0.97 ค่าระลอก 0.97 และค่าการถ่วงดุล 0.97 ตามลำดับ

5. สรุป

งานวิจัยฉบับนี้นำเสนอการจำแนกข้อความประเภทโรคติดต่อจากอาการตามมาตรฐาน ICD-10 เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมการเรียนรู้เชิงลึก 2 อัลกอริทึม ได้แก่ อัลกอริทึม Convolutional Neural Network และอัลกอริทึม Long Short-Term Memory ชุดข้อมูลที่ใช้เป็นข้อมูลอาการของโรคติดต่ออยู่ในรูปแบบภาษาไทย จากเว็บไซต์หน่วยงานสาธารณสุข เว็บไซต์โรงพยาบาล และเว็บไซต์สาธารณะ ทั้งหมด 5,000 ข้อความ มาผ่านกระบวนการเตรียมข้อมูล ได้แก่ การทำความสะอาดข้อมูล การตัดคำภาษาไทย การแปลงตัวพิมพ์ใหญ่เป็นเป็นตัวพิมพ์เล็ก การกำจัดคำหยุด และการกำหนดป้ายกำกับชุดข้อความด้วยรหัส ICD-10 จากนั้นแบ่งชุดข้อมูลเป็น 3 ส่วน ได้แก่ ชุดข้อมูลสอน ร้อยละ 70 ชุดข้อมูลตรวจสอบ ร้อยละ 10 และชุดข้อมูลทดสอบ ร้อยละ 20 ผลการทดสอบประสิทธิภาพพบว่า อัลกอริทึม Long Short-Term Memory มีประสิทธิภาพในการจำแนกข้อความประเภทโรคติดต่อจากอาการตามมาตรฐาน ICD-10 ดีที่สุด โดยให้ค่าความถูกต้อง ร้อยละ 96.76 ค่าความแม่นยำ 0.97 ค่าระลอก 0.97 และค่าการถ่วงดุล 0.97 ตามลำดับ

เอกสารอ้างอิง

- [1] Kim, H., and Jeong, Y.S., "Sentiment Classification Using Convolutional Neural Networks." *In applied sciences*. Vol.9, No.11, pp. 1-14 (2019).
- [2] Bowornlertsutee, P. and Paireekreng, W., "The Model of Sentiment Analysis for Classifying the Online Shopping Reviews." *In Journal of Engineering and Digital Technology*, Vol.10, No.1, pp.71-79, (2022).
- [3] Jitboonyapinit, C. et al., "Development of Sentiment Analysis Model Based on Thai social media Using Deep Learning Techniques." *In Huachiew Chalermprakiet Science and Technology Journal*. Vol. 8, No.2, pp.8-18, (2022).
- [4] Hazha, S.Y., and Adnan, M.A., "Medical Text Classification Based on Convolutional Neural Network: A Review." *In International Journal of Science and Business*. Vol. 5, No. 3, pp.27-41, (2021).
- [5] Winda, K.S. et al., "Text Classification Using Long Short-Term Memory with GloVe Features." *In Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*. Vol. 5, No.1, pp. 85-100, (2019).
- [6] Amit, K.S. et al., "Sentimental Short Sentences Classification by Using CNN Deep Learning Model with Fine Tuned Word2Vec." *In International Conference on Computational Intelligence and Data Science*. Vol.167, pp. 1139-1147, (2020).
- [7] Amol, C. A., "Improved Text Classification using Long Short-Term Memory and Word Embedding Technique." *In International Journal of Hybrid Information Technology*. Vol.13, No.1, pp.19-32, (2020).
- [8] Muh, A. N. and Azhari, S., "Sentiment Analysis of Novel Review Using Long Short-Term Memory Method." *In Indonesian Journal of Computing and Cybernetics Systems*. Vol.13, No.3, pp. 209-218, (2019).
- [9] Dr. G. S. N. Murthy et al., "Text-based Sentiment Analysis using LSTM." *In International Journal of Engineering Research & Technology*. Vol. 9, No.5, pp. 299-303, (2020).
- [10] Hamed, J. et al., "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach." *In IEEE Journal of Biomedical and Health Informatics*. pp.1-12, (2020).
- [11] Pisit B. and Worapat P. "Sentiment Analysis Techniques of Online Product Reviews." *RMUTSV Research Journal*. Vol.14 No.3, pp. 755-769, (2021).
- [12] Agung, T., "Convolutional Neural Network for Text Classification on Twitter." *In Journal of Software Engineering & Intelligent Systems*. Vol.4, No.3, pp.123-131, (2019).

การเปรียบเทียบเทคนิคการเรียนรู้เชิงลึก สำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย Deep Learning Techniques Comparison for Analyzing Opinions of Thailand Post Service User

วรรณกร พึงอุทัยศรี (Wannakorn Puenguthaisri)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856089@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

บทความวิจัยฉบับนี้นำเสนอการเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย จำนวน 2 อัลกอริทึม ได้แก่ อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว และอัลกอริทึมหน่วยเวียนกลับแบบมีประตู โดยได้ทำการรวบรวมความคิดเห็นของผู้ใช้บริการที่มีต่อบริษัท ไปรษณีย์ไทย จำกัด จากผู้ใช้งานทวิตเตอร์ จำนวน 2,000 ข้อความ โดยแบ่งข้อความออกเป็นความคิดเห็นเชิงบวก จำนวน 1,000 ข้อความ และความคิดเห็นเชิงลบ จำนวน 1,000 ข้อความ นำข้อความเหล่านั้นไปผ่านกระบวนการจัดเตรียมข้อมูล ได้แก่ การลบสัญลักษณ์ การกำจัดคำซ้ำ การตัดคำ หลังจากนั้นนำข้อความที่ได้ไปสร้างเวกเตอร์และแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลสอน ร้อยละ 70 และชุดข้อมูลทดสอบ ร้อยละ 30 เพื่อนำเข้าแบบจำลอง ทำการประเมินประสิทธิภาพของอัลกอริทึม ด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก ค่าการถ่วงดุล และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน พบว่า อัลกอริทึมหน่วยความจำระยะสั้นแบบยาวให้ประสิทธิภาพสูงที่สุด โดยมีค่าความถูกต้องร้อยละ 94.53 ค่าความแม่นยำ 0.95 ค่าระลอก 0.95 ค่าการถ่วงดุล 0.95 และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน 0.95 ตามลำดับ

คำสำคัญ: อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว
อัลกอริทึมหน่วยเวียนกลับแบบมีประตู
บริษัท ไปรษณีย์ไทย จำกัด ทวิตเตอร์

Abstract

This research article compares between two deep learning algorithms for analyzing customer opinions of Thailand Post Co., Ltd. The Long Short-Term Memory Algorithm and Gated Recurrent Unit Algorithm were compared. The study collected 2,000 comments from Twitter users regarding their opinions towards Thailand Post Co., Ltd. The comments were divided into 1,000 positive comments and 1,000 negative comments, and then underwent a data preparation process that included removing symbols, eliminating duplicate words, and cutting words. The obtained comments were used to generate vectors and then divided into two sets a training dataset (70%) and a test dataset (30%). The purpose was to import the model and compare algorithm performance. To evaluate algorithm efficiency, accuracy, precision, recall, F-measure, and ROC values were used. Long Short-Term Memory Algorithm was the most efficient, with an accuracy of 94.53%, precision of 0.95%, recall of 0.95%, F-measure of 0.95%, and ROC of 0.95%.

Keywords: Long Short-Term Memory Algorithm,
Gated Recurrent Unit Algorithm,
Thailand Post, Twitter.

1. บทนำ

ปี พ.ศ. 2563 ประเทศไทยเข้าสู่การแพร่ระบาดของโรคไวรัสโควิด 19 ทำให้ประชาชนเปลี่ยนแปลงรูปแบบ

การใช้ชีวิตประจำวัน โดยการทำงานจากที่บ้าน (Work From Home) ซึ่งแตกต่างจากการใช้งานรูปแบบเดิมที่ประชาชนจะต้องมาใช้บริการ ณ ที่ทำการไปรษณีย์ การทำงานจากที่บ้านทำให้ช่องทางการค้าขายในรูปแบบออนไลน์มีอัตราขยายตัวที่สูงขึ้นเป็นอย่างมาก ส่งผลต่อการให้บริการไปรษณีย์มีผู้ใช้งานเพิ่มขึ้นตามลำดับ

ปัจจุบัน ช่องทางการสื่อสารในรูปแบบออนไลน์ (Social Media) โดยเฉพาะอย่างยิ่งทวิตเตอร์ (Twitter) เป็นสื่อสังคมออนไลน์ที่ได้รับความนิยมจากผู้ใช้งานอย่างแพร่หลาย การร้องเรียนผ่านช่องทางนี้สามารถแพร่กระจายข่าวสารได้อย่างรวดเร็ว เพื่อการวิเคราะห์ถึงความคิดเห็นของผู้ใช้บริการไปรษณีย์ว่ามีความเห็นเกี่ยวกับให้บริการไปรษณีย์อยู่ในเชิงบวกหรือเชิงลบ [1] ผู้วิจัยได้นำเทคนิคการทำเหมืองข้อความ มาประยุกต์ใช้ในการวิเคราะห์ความคิดเห็นในสื่อสังคมออนไลน์ นอกเหนือจากรูปแบบการปรับปรุงคุณภาพบริการในการสอบถามความคิดเห็นด้วยวิธีแบบสอบถาม [2] โดยใช้เทคนิคการเรียนรู้เชิงลึก ในการใช้เทคนิคการเรียนรู้เชิงลึกมีอัลกอริทึมสำหรับใช้ในการวิเคราะห์หลายอัลกอริทึม โดยได้มีผู้วิจัยนำเสนออัลกอริทึมเหมืองข้อความมาช่วยในการวิเคราะห์ข้อมูลความเห็นในเรื่องต่าง ๆ เช่น การจำแนกการวิเคราะห์ความคิดเห็นต่อสินค้าและบริการออนไลน์ภาษาไทย โดยใช้อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory:LSTM [3] และอัลกอริทึมหน่วยเวียนกลับแบบมีประตู (Gated Recurrent Unit:GRU) [4] รวมถึงเทคนิคการวิเคราะห์เหมืองข้อความ ซึ่งการประมวลผลภาษาธรรมชาติ โดยใช้วิธีการตัดคำมาประยุกต์ใช้เพื่อวิเคราะห์ความคิดเห็นที่เป็นภาษาไทยของผู้ใช้งานบนสื่อสังคมออนไลน์ โดยดึงข้อมูลผ่านส่วนต่อประสานโปรแกรมประยุกต์ (Application Programming Interface :API) ของทวิตเตอร์ และไลบรารีของ Python ที่ชื่อว่า “Tweepy API” [5] เป็นต้น

ผู้วิจัยได้สังเกตเห็นถึงปัญหาที่เกิดขึ้น และผลกระทบในเชิงลบที่อาจจะส่งผลต่อภาพลักษณ์ขององค์กร จึงศึกษาการเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกสำหรับ

การวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย ด้วยเทคนิคการประมวลผลภาษาธรรมชาติ เช่น การตัดคำ การสกัดคำของประเภทคำ และรูปประโยค เป็นต้น [6] นำมาสร้างแบบจำลอง โดยใช้อัลกอริทึม Long Short-Term Memory และอัลกอริทึม Gated Recurrent Unit ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยพิจารณาจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าระลึก (Recall) ค่าการถ่วงดุล (F-measure) และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน (ROC) เพื่อจำแนกความหมายของข้อความความคิดเห็นในเชิงบวก และเชิงลบ พร้อมนำความคิดเห็นเหล่านั้น มาปรับปรุงประสิทธิภาพการให้บริการ เพื่อให้เกิดภาพลักษณ์ที่ดีกับองค์กรต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งออกเป็น 2 หัวข้อ ดังนี้

2.1 อัลกอริทึมสำหรับการวิเคราะห์ความคิดเห็น

2.1.1 อัลกอริทึมอัลกอริทึมหน่วยความจำระยะสั้นแบบยาว เป็นอัลกอริทึมที่มีการพัฒนามาจากอัลกอริทึมโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN) เพื่อแก้ปัญหาเกี่ยวกับข้อมูลที่มีขนาดยาว และลดการสูญหายของเกรเดียน ในการแก้ปัญหาดังกล่าว จึงมีการออกแบบ Cell ใหม่ โดยการนำหน่วยความจำที่สามารถอ่าน เขียน และสามารถลบได้ คือ Forget Gate, Input Gate และ Output Gate เพื่อควบคุมการไหลของข้อมูล จึงสามารถเรียนรู้จากข้อมูลเป็นลำดับหากข้อมูลนั้นไม่จำเป็นก็จะเก็บหรือลบทิ้ง [7]

2.1.2 อัลกอริทึมหน่วยเวียนกลับแบบมีประตู เป็นอัลกอริทึมที่มีการพัฒนามาจากอัลกอริทึม Recurrent Neural Network เพื่อแก้ปัญหาเกี่ยวกับข้อมูลที่มีขนาดยาว และลดการสูญหายของเกรเดียน โดยทำการลดจำนวนหน่วยความจำจาก Forget Gate, Input Gate และ Output Gate เหลือเพียง Input Gate กับ Reset Gate จึงทำให้การทำงานมีความรวดเร็วกว่า Long Short-Term Memory [8]

2.2 งานวิจัยที่เกี่ยวข้อง

พิศิษฐ์ และวรภัทร [9] ได้ทำการศึกษาตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์โดยใช้เทคนิค Long Short-Term Memory เทคนิคการเคลื่อนลงตามความชัน (Stochastic Gradient Descent) เทคนิคการวิเคราะห์การถดถอยโลจิสติก (Logistic Regression) และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ใช้ข้อมูลตัวอย่างการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์ภาษาไทยจำนวน 12,900 ข้อมูล จากการทดลองสร้างตัวแบบการวิเคราะห์ระดับความรู้สึกทางอารมณ์ 3 ระดับ คือ ความรู้สึกเชิงบวก (Positive) เป็นกลาง (Neutral) และเชิงลบ (Negative) ผลลัพธ์ที่ได้ Long Short-Term Memory ให้ค่าความถูกต้องสูงที่สุดเท่ากับร้อยละ 81.27 ในการทำนายผลด้วยการเรียนรู้เชิงลึก จึงเหมาะสมแก่การนำไปใช้สร้างตัวแบบในการวิเคราะห์ระดับความรู้สึกกับข้อความที่เป็นภาษาไทย

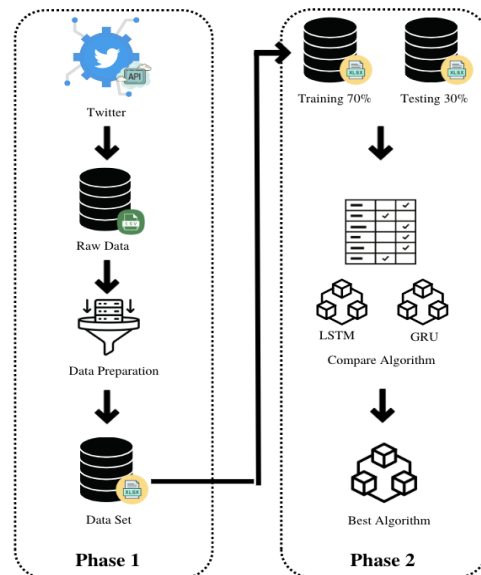
พยูง มีสัง และคณะ [10] ได้ทำการศึกษาเทคนิคทางปัญญาประดิษฐ์สำหรับการตรวจข่าวปลอมภาษาไทยโดยใช้วิธีการวิธีการค้นคืนสารสนเทศด้วยเว็บครอเลอร์ทำการเปรียบเทียบประสิทธิภาพการจำแนก ผลการวิจัยพบว่าอัลกอริทึม Long Short-Term Memory ให้ประสิทธิภาพสูงที่สุด ค่าความถูกต้องร้อยละ 93.81

Fangqin Zeng at al. [11] ได้พัฒนาแบบจำลองคำแนะนำส่วนบุคคลสำหรับใช้ใน Social Media ด้วยอัลกอริทึม Gated Recurrent Unit โดยใช้ข้อมูลของ Amazon และ MovieLens ผลการทดลองพบว่าอัลกอริทึม Gated Recurrent Unit มีความสามารถในการแนะนำเฉพาะบุคคลได้ดี

Ning Li et al. [12] ได้วิเคราะห์การรับส่งข้อมูลเครือข่ายดาวเทียม ตามการเรียนรู้ของการย้ายข้อมูลด้วยอัลกอริทึม Gated Recurrent Unit ผลการทดลองพบว่าอัลกอริทึม Gated Recurrent Unit มีความแม่นยำในการคาดคะเนกราฟิกและสามารถลดทรัพยากรการจัดเก็บการประมวลผลผ่านดาวเทียมได้อย่างมีประสิทธิภาพ

3. วิธีการดำเนินการวิจัย

การเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย ผู้วิจัยมีกรอบแนวคิดแสดง ดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากภาพที่ 1 กรอบแนวคิดการวิจัย ประกอบด้วยขั้นตอนการดำเนินงานวิจัยจำนวน 2 ขั้นตอน ได้แก่ การเก็บรวบรวมและจัดเตรียมข้อมูล และการสร้างแบบจำลองและการเปรียบเทียบประสิทธิภาพ โดยมีขั้นตอนการดำเนินการ ดังนี้

3.1 การเก็บรวบรวมข้อมูลและจัดเตรียมข้อมูล

การรวบรวมชุดข้อมูลสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย ผู้วิจัยได้รวบรวมความคิดเห็นจากทวิตเตอร์ ทำการดึงข้อมูลผ่านส่วนต่อประสานโปรแกรมประยุกต์ด้วย “Tweepy API” ในการดึงชุดข้อมูล และทำการคัดเลือกความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย จำนวน 2,000 ข้อความ แบ่งออกเป็นความคิดเห็นเชิงบวก จำนวน 1,000 ข้อความ และความคิดเห็นเชิงลบ จำนวน 1,000 ข้อความ โดยมีตัวอย่างข้อมูล แสดงดังตารางที่ 1

ตารางที่ 1: ตัวอย่างข้อมูล

ข้อความ	ความคิดเห็น
เมื่อไปร #ไปรษณีย์ไทย มีเลขให้เช็คค่าของถึงไหนแล้วและบ่อยครั้ง ไม่เคยอัปเดตนอกจาก	Negative

ตอนรับฝาก มา้อพเอาตอนของถึงแล้ว แบบอิ หยังวะมาากาากาากาากาากาาก	
ไปรษณีย์ไทยทำจั้นร้องให้อีกละ ☹️ ☹️ ☹️ ☹️ ☹️	Negative
ส่งค่วนที่ไม่ค่วน ค่าส่งแพงกว่าปกติแต่ส่งช้า มาากาากาากาากาาก ตอนนี้อยังไม่ได้ของเลย จ้า ไม่น่าใช้บริการเลย	Negative
📦 รับของถึงที่ "ฟรี" ต้องไปรษณีย์ไทย!	Positive
น่าชื่นชม 🤝 🤝 🤝 หนุ่มไปรษณีย์ไทย	Positive
ไปรษณีย์ก็มาส่งเร็วมาากาาก	Positive

จากนั้นนำข้อความมาผ่านกระบวนการจัดเตรียมข้อมูล ประกอบด้วย 5 ขั้นตอน ดังนี้

3.1.1 การลบสัญลักษณ์ ได้แก่ การจัดการกับข้อมูล โดยการลบสัญลักษณ์ อักขระพิเศษ รวมถึงการลบโมจิต่าง ๆ ซึ่งสัญลักษณ์ดังกล่าวไม่มีความจำเป็นในการจำแนกความคิดเห็น

3.1.2 การกำจัดคำซ้ำ ได้แก่ การจัดการกับข้อมูล โดยการลบตัวอักษรและสระที่มีความซ้ำออก เช่น “สระนะ” เมื่อกำจัดคำซ้ำออกจะได้เป็น “สระน้ำ” เป็นต้น

3.1.3 การตัดคำ ได้แก่ การตัดคำออกเป็น ส่วน ๆ โดยใช้ NLP PyThai ใช้ Library Attacut เพื่อใช้ตัดคำภาษาไทย ใช้สัญลักษณ์ “|” แทนการแบ่งคำแต่ละคำ เช่น คำว่า “ไปรษณีย์ส่งเร็วทันใจจริง” เป็น “ไปรษณีย์|ส่ง|เร็ว|ทันใจจริง เป็นต้น

3.1.4 การแทนค่าป้ายกำกับ ได้แก่ การแปลงค่าป้ายกำกับจาก Negative เป็น 0 และการแปลงค่าป้ายกำกับจาก Positive เป็น 1

3.1.5 การแบ่งข้อมูล ได้แก่ การนำข้อความความคิดเห็นที่ผ่านกระบวนการจัดเตรียมข้อมูลเรียบร้อยแล้ว แบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลสอน (Training Dataset) ร้อยละ 70 จำนวน 1,400 ข้อความ และชุดข้อมูลทดสอบ (Test Dataset) ร้อยละ 30 จำนวน 600 ข้อความ เพื่อนำเข้าสู่การสร้างแบบจำลอง

ตารางที่ 2: ตัวอย่างข้อความหลังผ่านกระบวนการจัดเตรียมข้อมูล

ข้อความ	ความคิดเห็น
เมื่อไปรษณีย์ไทยมีเลขให้เช็คของถึงไ ไหนแล้วและบ่อยครั้งไม่เคยอัปเดตนอก จากตอนรับฝากมา้อพเอาตอนของถึงแล้ว แบบอิหยังวะมาก	0
ไปรษณีย์ไทยทำจั้นร้องให้อีกละ	0
ส่งค่วนที่ไม่ค่วนค่าส่งแพงกว่าปกติแต่ส่ง ช้ามากตอนนี้อยังไม่ได้ของเลยจ้าไม่น่าใช้ บริการเลย	0
รับของถึงที่ฟรีต้องไปรษณีย์ไทย	1
น่าชื่นชมหนุ่มไปรษณีย์ไทย	1
ไปรษณีย์ก็มาส่งเร็วมาก	1

3.2 การสร้างแบบจำลองและการเปรียบเทียบประสิทธิภาพ

อัลกอริทึมที่ใช้ในการสร้างแบบจำลอง ประกอบด้วย 2 อัลกอริทึม ได้แก่ Long Short-Term Memory และ Gated Recurrent Unit โดยดำเนินงาน ดังนี้

กำหนดค่าพารามิเตอร์ที่ใช้สำหรับการทดสอบ อัลกอริทึมและทำการปรับค่าพารามิเตอร์ตามที่กำหนดไว้ในแต่ละอัลกอริทึม แสดงดังตารางที่ 4 ถึง 6

ตารางที่ 4: ค่าพารามิเตอร์คงที่

Parameter	Value
dropout	0.15
learning rate	0.0001
optimize	adam
activation	sigmoid
epoch	100

ตารางที่ 5: ค่าพารามิเตอร์ที่ทำการปรับค่า

Parameter	Value
embed_dim	8,16,32,64,128
units_out	8,16,32,64,128
batch_size	8,16,32,64,128

ผู้วิจัยปรับค่าพารามิเตอร์จำนวน 125 แบบ เพื่อหาค่าพารามิเตอร์ที่ดีที่สุดในแต่ละอัลกอริทึม จากนั้นนำค่าที่ดีที่สุดนำไปปรับใช้กับค่าพารามิเตอร์ epoch

ตารางที่ 6: ค่าพารามิเตอร์ epoch

Parameter	Value
epoch	50,75,100,125,150

เปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยพิจารณา ค่าความถูกต้อง ค่าความแม่นยำ ค่าระลึก ค่าการถ่วงดุล และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน

4. ผลการดำเนินงาน

ผลการดำเนินงานแสดงดังตารางที่ 7 ถึง 10 โดยมีรายละเอียด ดังนี้

ตารางที่ 7: ผลการปรับค่าพารามิเตอร์ของอัลกอริทึม Long Short-Term Memory

embed_dim	units_out	batch_size	A	P	R	F	ROC
8	16	64	94.53	0.95	0.95	0.95	0.95
8	32	128	92.70	0.93	0.93	0.93	0.93
16	8	32	94.47	0.94	0.94	0.94	0.94
16	32	8	92.30	0.92	0.92	0.92	0.92
32	16	64	93.80	0.94	0.94	0.94	0.94
32	128	64	92.03	0.92	0.92	0.92	0.92
64	8	128	94.07	0.94	0.94	0.94	0.94
64	16	64	94.03	0.94	0.94	0.94	0.94
128	8	128	93.36	0.93	0.93	0.93	0.93
128	128	16	93.20	0.93	0.93	0.93	0.93

** A = Accuracy, P = Precision, R = Recall, F = F-Measure, ROC = ROC

จากตารางที่ 7 ค่าพารามิเตอร์อัลกอริทึม Long Short-Term Memory ที่ให้ประสิทธิภาพดีที่สุด คือ ค่า embed_dim เท่ากับ 8 ค่า unit_out เท่ากับ 16 และค่า batch_size เท่ากับ 64 โดยให้ค่าความถูกต้องเท่ากับร้อยละ 94.53

ตารางที่ 8: ผลการปรับค่าพารามิเตอร์ของอัลกอริทึม Gated Recurrent Unit

embed_dim	units_out	batch_size	A	P	R	F	ROC
8	16	64	93.86	0.94	0.94	0.94	0.94
8	32	128	93.47	0.93	0.93	0.93	0.93
16	8	32	92.77	0.93	0.93	0.93	0.93
16	32	8	91.63	0.92	0.92	0.92	0.92
32	16	64	92.40	0.92	0.92	0.92	0.92
32	128	64	93.03	0.93	0.93	0.93	0.93
64	8	128	92.10	0.92	0.92	0.92	0.92
64	16	64	92.77	0.93	0.93	0.93	0.93
128	8	128	90.17	0.90	0.90	0.90	0.90
128	128	16	93.23	0.93	0.93	0.93	0.93

จากตารางที่ 8 ค่าพารามิเตอร์อัลกอริทึม Gated Recurrent Unit ที่ให้ประสิทธิภาพดีที่สุด คือ ค่า embed_dim เท่ากับ 8 ค่า unit_out เท่ากับ 16 และค่า batch_size เท่ากับ 64 โดยให้ค่าความถูกต้องเท่ากับร้อยละ 93.86

หลังจากได้ค่าพารามิเตอร์ที่ดีที่สุดของอัลกอริทึม Long Short-Term Memory และอัลกอริทึม Gated Recurrent Unit นำมาปรับค่าพารามิเตอร์ epoch มีรายละเอียดดังตารางที่ 9 ถึง 10

ตารางที่ 9: ผลการปรับค่าพารามิเตอร์ epoch ของอัลกอริทึม Long Short-Term Memory

epoch	Accuracy	Precision	Recall	F-Measure	ROC
50	90.93	0.91	0.91	0.91	0.91
75	93.50	0.94	0.94	0.94	0.94
100	94.53	0.95	0.95	0.95	0.95
125	92.90	0.93	0.93	0.93	0.93
150	93.30	0.93	0.93	0.93	0.93

ตารางที่ 10: ผลการปรับค่าพารามิเตอร์ epoch ของอัลกอริทึม Gated Recurrent Unit

epoch	Accuracy	Precision	Recall	F-Measure	ROC
50	91.90	0.92	0.92	0.92	0.92
75	93.80	0.94	0.94	0.94	0.94
100	93.86	0.94	0.94	0.94	0.94
125	93.70	0.94	0.94	0.94	0.94
150	93.63	0.94	0.94	0.94	0.94

จากตารางที่ 9 ถึง 10 พบว่าค่าพารามิเตอร์ อัลกอริทึม Long Short-Term Memory และ อัลกอริทึม Gated Recurrent Unit ให้ประสิทธิภาพดีที่สุด คือ ค่า epoch เท่ากับ 100 โดยอัลกอริทึม Long Short-Term Memory ให้ค่าความถูกต้องร้อยละ 94.53 และอัลกอริทึม Gated Recurrent Unit ให้ค่าความถูกต้องร้อยละ 93.86

หลังจากได้ค่าพารามิเตอร์ที่ดีที่สุดของอัลกอริทึม Long Short-Term Memory และอัลกอริทึม Gated Recurrent Unit เรียบร้อยแล้ว นำมาเปรียบเทียบประสิทธิภาพรายละเอียด ดังตารางที่ 11

ตารางที่ 11: ตารางเปรียบเทียบประสิทธิภาพ

ประสิทธิภาพ	อัลกอริทึม	
	LSTM	GRU
Accuracy	94.53	93.86
Precision	0.95	0.94
Recall	0.95	0.94
F-measure	0.95	0.94
ROC	0.95	0.94

จากตารางที่ 11 พบว่าอัลกอริทึม Long Short-Term Memory ให้ประสิทธิภาพดีที่สุด โดยให้ค่าความถูกต้อง

ร้อยละ 94.53 ค่าความแม่นยำ 0.95 ค่าระลอก 0.95 ค่าการถ่วงดุล 0.95 และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน 0.95 ตามลำดับ

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการเปรียบเทียบเทคนิคการเรียนรู้เชิงลึกสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการไปรษณีย์ไทย จำนวน 2,000 ข้อความ แบ่งออกเป็นความคิดเห็นเชิงบวก จำนวน 1,000 ข้อความ และความคิดเห็นเชิงลบ จำนวน 1,000 ข้อความ โดยการนำข้อความเหล่านั้นไปผ่านกระบวนการจัดเตรียมข้อมูล เช่น การลบสัญลักษณ์ การกำจัดคำซ้ำ การตัดคำ เป็นต้น หลังจากนั้นนำข้อความที่ได้ไปสร้างเวกเตอร์เพื่อนำเข้าแบบจำลอง โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลสอนร้อยละ 70 และชุดข้อมูลทดสอบร้อยละ 30 ทำการสร้างแบบจำลองและประเมินประสิทธิภาพของอัลกอริทึมจำนวน 2 อัลกอริทึม ได้แก่ อัลกอริทึม Long Short-Term Memory และอัลกอริทึม Gated Recurrent Unit ประเมินประสิทธิภาพของอัลกอริทึม โดยพิจารณาจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก ค่าการถ่วงดุล และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน ทั้งนี้ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึมพบว่า อัลกอริทึม Long Short-Term Memory ให้ประสิทธิภาพการจำแนกความคิดเห็นของผู้ใช้งานทวิตเตอร์ที่มีต่อการใช้บริการไปรษณีย์ไทยสูงที่สุด ซึ่งสอดคล้องกับงานวิจัยของ Jenq-Haur Wang et al. [6] โดยมีค่าความถูกต้องร้อยละ 94.53 ค่าความแม่นยำ 0.95 ค่าระลอก 0.95 ค่าการถ่วงดุล 0.95 และค่าการวิเคราะห์เส้นโค้งคุณลักษณะการทำงาน 0.95 ตามลำดับ

เอกสารอ้างอิง

- [1] Abdur et al., "Twitter Sentiment Analysis: A Case Study for Apparel Brands." *Journal of Physics: Conference Series*, pp.1-6.
- [2] สิริยากร ปานรอด และรัชชพงษ์ ชัชวาล, "การรับรู้คุณภาพการให้บริการบริษัทไปรษณีย์ไทย จำกัด กรณีศึกษา เขตพื้นที่การให้บริการของทำการไปรษณีย์สงขลา." *การประชุมวิชาการด้านมนุษยศาสตร์และสังคมศาสตร์ระดับชาติ ครั้งที่ 1* หน้า 231-239.
- [3] ดำรงพล สุวรรณชาติ และวัชรวิวัฒน์ จิตต์สกุล, "การจำแนกความคิดเห็นเกี่ยวกับรถยนต์ไฟฟ้าด้วยเทคนิคเหมืองข้อความ." *The Eighteenth National Conference on Computing and Information Technology* หน้า 235-240.
- [4] K.E. ArunKumar et al., "Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cell." *Chaos, Solitons and Fractals*, pp.1-12.
- [5] วิภาดา ศิลาราช พิทยารัตน์ พ่อมชฌู และธนพล ตั้งชูพงศ์, "การวิเคราะห์ความรู้สึกของข้อความ กรณีศึกษาชุดข้อมูลของทวิตเตอร์ เรื่อง การศึกษาไทย." *The Eighteenth National Conference on Computing and Information Technology* หน้า 184-189.
- [6] Jenq-Haur Wang et al., "An LSTM Approach to short Text Sentiment Classification with Word Embeddings." *The 2018 Conference on Computational Linguistics and Speech Processing ROCLING 2018*, pp. 214-223.
- [7] ชาลิสา จิตบุญญาพิณี และคณะ, "การพัฒนาแบบจำลองการวิเคราะห์ความรู้สึกบนสื่อสังคมออนไลน์ไทยโดยใช้เทคนิคการเรียนรู้เชิงลึก" *วารสารวิทยาการและเทคโนโลยีสารสนเทศ หัวเฉียวเฉลิมพระเกียรติ ปีที่ 8 ฉบับที่ 2 พ.ศ. 2565* หน้า 8-18.
- [8] รัชชกานา ภูสีเขียว และสุรณพีร์ ภูมิวุฒิสาร, "การพยากรณ์ปริมาณน้ำฝนระยะสั้นในบริเวณพื้นที่สนามบินสุวรรณภูมิด้วยโครงข่ายระบบประสาทแบบย้อนกลับ" *Journal of Information Science and Technology*. vol.12, pp. 13-26, 2022.
- [9] พิเศษฐ์ บวรเลิศสุธี และวราภรณ์ ไพรีเกรง, "ตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์" *Journal of Engineering and Digital Technology*. vol.10, หน้า 71-89.
- [10] พยุง มีสัง และคณะ, "เทคนิคทางปัญญาประดิษฐ์สำหรับการตรวจข่าวปลอมภาษาไทย" *The Journal of Applied Science*. vol.21, หน้า 1-19.
- [11] Fangqin Zeng, Rong Tang, and Yibai Wang. "User Personalized Recommendation Algorithm Based on GRU Network Model in Social Network." *Hindawi Mobile Information Systems*, pp. 1-8.
- [12] Ning Li et al. "Research on GRU Neural Network Satellite Traffic Prediction Based on Transfer Learning.", *Wireless Personal Communications*, pp. 817-826.

การเปรียบเทียบอัลกอริทึมสำหรับการจดจำใบหน้า

Comparison of Algorithms for Face Recognition

ณัฐพงศ์ จันทรศร (Nattapong Jantarasorn)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857191@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาการจดจำใบหน้า โดยเปรียบเทียบอัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ ข้อมูลที่ใช้เป็นชุดข้อมูลใบหน้าทั้งหมด 200 รูปภาพ จากบุคคลทั้งหมด 10 คน เพื่อรวบรวมข้อมูลใบหน้า 3 ลักษณะ ได้แก่ ใบหน้าตรง ใบหน้าด้านข้าง และใบหน้าสวมหน้ากากอนามัย ทำการประเมินประสิทธิภาพของอัลกอริทึมแต่ละอัลกอริทึม โดยพิจารณาจากความถูกต้อง ทั้งนี้ผลการเปรียบเทียบของอัลกอริทึม พบว่า อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ ให้ประสิทธิภาพดีที่สุด สำหรับใบหน้าตรงให้ค่าความถูกต้องร้อยละ เกือบ 91.24 สำหรับใบหน้าที่ด้านข้างให้ค่าความถูกต้องร้อยละ เกือบ 87.66 และสำหรับใบหน้าสวมหน้ากากอนามัยให้ค่าความถูกต้องร้อยละ เกือบ 82.14 ตามลำดับ ซึ่งใช้เวลาทุก ๆ 3 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง ซึ่งสามารถให้ค่าความถูกต้องดีกว่าอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ตามลำดับ

คำสำคัญ: การจดจำใบหน้า อัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ

Abstract

The research objective is to develop facing recognition by compare three algorithms, which are convolutional neural networks (CNN), Support Vector Machine (SVM) and Histogram of Oriented Gradients (HOG). A face data set from a set of 200 images of 10 people is required to collect face information from various characteristics, such as a straight face, a sideways face, and wearing mask face. For evaluating the performance, the result of individual algorithms was tested with an accuracy value. After comparison of result algorithms. CNN was efficient to face recognition with an accuracy for straight face the result showed that the average percentage accuracy was 91.24%. For sideways face the result showed that the average percentage accuracy was 87.66%. For wearing mask faces the result showed that the average percentage accuracy was 82.14%, which took every 3 seconds for each face recognition. CNN's algorithm has better the average accuracy than SVM and HOG algorithms respectively.

Keyword: Face Recognition, Histogram of Oriented Gradients, Support Vector Machine, Convolutional Neural Networks.

1. บทนำ

ปัจจุบันเทคโนโลยีการประมวลผลภาพมีความสำคัญในชีวิตประจำวันมากขึ้นเรื่อย ๆ โดยเฉพาะเทคโนโลยีจดจำใบหน้า ซึ่งใช้ในการวิเคราะห์ลักษณะใบหน้าจากรูปภาพ หรือกล้องวิดีโอ เช่น ท่าทาง การจัดแสง

และความละเอียดต่ำ เป็นต้น เพื่อระบุตัวบุคคล เนื่องจากเทคโนโลยีดังกล่าวได้รับความนิยม และสามารถนำไปประยุกต์ใช้ได้หลากหลาย แต่การใช้เทคโนโลยีการเรียนรู้จดจำใบหน้าจำเป็นต้องใช้คุณสมบัติใบหน้า หรือส่วนประกอบทั้งหมดของใบหน้า ซึ่งไม่สามารถให้ความถูกต้องเพียงพออาจเกิดจากหลายปัจจัย ได้แก่ การใส่ผ้าปิดหน้า หรือการหลบหน้า เพื่ออำพรางตัวเองทำให้เทคโนโลยีจดจำใบหน้าเป็นเรื่องที่ค่อนข้างยาก

มีการนำอัลกอริทึมจำนวนมากมาใช้ในการเรียนรู้จดจำใบหน้าในด้านต่างๆ เช่น การตรวจสอบเพศและอายุ โดยใช้อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ [1] การพัฒนาการจดจำใบหน้าด้วยเทคนิค Eigen Face, อัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน และ Principal Component Analysis [2] การพัฒนาระบบการเข้าร่วมอัตโนมัติโดยใช้อัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ [3] เป็นต้น

งานวิจัยนี้จึงมีวัตถุประสงค์ในการเปรียบเทียบอัลกอริทึม สำหรับการจดจำใบหน้า โดยใช้อัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ สำหรับการเปรียบเทียบประสิทธิภาพ โดยพิจารณาค่าความถูกต้อง เพื่อพัฒนาการเรียนรู้จดจำใบหน้าให้มีประสิทธิภาพยิ่งขึ้น สำหรับองค์กรภาครัฐ และเอกชน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 4 หัวข้อหลัก ดังนี้

2.1 โครงข่ายประสาทเทียมแบบสังวัตนาการ

เป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่ม bio-inspired และยังเป็นการเรียนรู้เชิงลึกประเภทหนึ่ง ซึ่งเหมาะกับการวิเคราะห์ภาพและวิดีโอ โดยมีพื้นฐานมาจากแนวคิดที่เกี่ยวข้องกับการใช้ตัวกรองที่เรียนรู้พื้นที่ย่อยๆ มาผสมกันกับรูปภาพอินพุต หรือแผนผังคุณลักษณะ ซึ่งออกแบบมาเพื่อแยกคุณลักษณะเฉพาะ เช่น ขอบหรือพื้นผิว จากข้อมูลที่ป้อนเข้าไป ซึ่งโครงข่ายประสาทเทียมแบบสังวัตนาการ

ประกอบด้วยชุดของเลเยอร์ ได้แก่ Convolutional Layers, Activation Layers, and Pooling Layers เป็นต้น ซึ่งใช้เพื่อจับคู่คุณสมบัติที่สกัดโดยเลเยอร์กับเอาต์พุตสุดท้ายของเครือข่ายเหล่านั้น [4]

2.2 ซัพพอร์ตเวกเตอร์แมชชีน

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาข้อมูลใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ ซึ่งประกอบด้วย Maximum Margin เรียกข้อมูลที่อยู่บน Margin เหล่านี้ว่า Support Vector [5]

2.3 เทคนิค ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์

เป็นการมองเห็นด้วยคอมพิวเตอร์และเทคนิคประมวลผลภาพที่ใช้สำหรับการตรวจจับใบหน้า ใช้เพื่อระบุใบหน้าบุคคล โดยการแบ่งภาพออกเป็นเซลล์สี่เหลี่ยมเล็ก ๆ แล้วคำนวณฮิสโตแกรมไล่ระดับสีแต่ละเซลล์สำหรับแสดงข้อมูลขอบเฉพาะที่ และการวางแนวของภาพและนำมาต่อกันเพื่อสร้างการแสดงผลโดยรวม [6]

2.4 งานวิจัยที่เกี่ยวข้อง

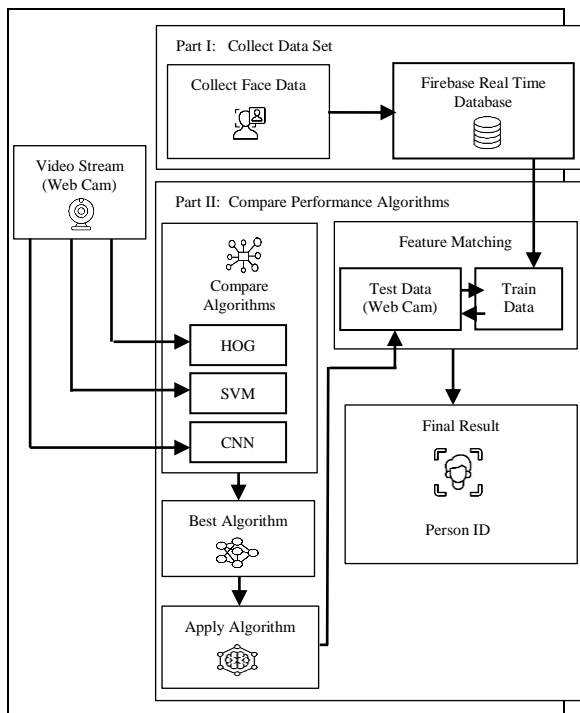
รามสวรรค์ คิวรักษา และคณะ [7] ได้ทำการตรวจจับหน้ากากบนใบหน้า โดยใช้อัลกอริทึม โครงข่ายประสาทเทียมแบบสังวัตนาการ เนื่องจากในปลายปี 2562 พบการระบาดของโรคไวรัสโคโรนา 2019 ซึ่งยังคงเป็นสาเหตุให้เกิดปัญหาเกี่ยวกับหลายด้านชีวิตและธุรกิจต่าง ๆ มากมาย จึงทำการใช้อัลกอริทึมกับข้อมูลใบหน้าที่รูปภาพมีทั้งแบบก้มหน้า มีลวดลาย และสิ่งปกปิดอัตลักษณ์บนใบหน้าทั้ง 3,835 ภาพ และใช้เครื่องมือการประเมินผลลัพธ์ของการทำนาย สามารถให้ผลลัพธ์ของค่าความถูกต้องมากที่สุดอยู่ที่ร้อยละ 97

G. Huilin and et al [8] พัฒนาระบบแอปพลิเคชันสำหรับจดจำใบหน้าแบบเรียลไทม์ โดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนทำงานร่วมกับโมเดล FaceNet โดยเทรน โมเดลกับข้อมูลบุคคลทั้งหมด 5 คน จากการทดลองพบว่า ใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนทำงานร่วมกับโมเดล FaceNet สามารถให้ค่าความถูกต้องเท่ากับร้อยละ 95

G. Mohammed and I. Melhum [9] ได้ทำการพัฒนาการตรวจจับใบหน้ามนุษย์โดยใช้การดำเนินการแยกคุณสมบัติของ อัลกอริทึมเทคนิค ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ที่ทำงานร่วมกับ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน โดยใช้ข้อมูลจากฐานข้อมูลใบหน้า MIT CBCL ทั้งหมด 7,000รูปภาพ จากการทดลองพบว่า สามารถให้ค่าความถูกต้อง ร้อยละ 0.95 และค่า Sensitivity เท่ากับ 0.75

3. วิธีการดำเนินการวิจัย

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ ผู้วิจัยมีกรอบแนวความคิด แสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวความคิด

จากภาพที่ 1 กรอบแนวความคิด สามารถแบ่งขั้นตอนการดำเนินการวิจัยเป็น 2 ส่วน ดังนี้

3.1 การรวบรวมและการจัดการข้อมูลใบหน้า

3.1.1 ภาพตัวอย่างการเตรียมชุดข้อมูลการวิจัย เป็นการจัดเตรียมชุดข้อมูลใบหน้าทั้งหมด 200 รูปภาพ จากบุคคลทั้งหมด 10 คน เพื่อเก็บรวบรวมข้อมูลใบหน้า 3 ลักษณะ ได้แก่ ใบหน้าตรง ใบหน้าด้านข้าง และใบหน้าสวมหน้ากากอนามัย แสดงตัวอย่างรูปภาพดังตารางที่ 1

ตารางที่ 1: ตัวอย่างรูปภาพใบหน้า การเตรียมชุดข้อมูลการวิจัย

Person	Straight Face	Sideways Face	Wearing Mask Faces
1			
2			
3			

3.1.2 ออกแบบส่วนการเก็บรวบรวมข้อมูลใบหน้า และจัดเก็บบนฐานข้อมูลเรียลไทม์ของ Firebase

ทำการรวบรวมข้อมูลเกี่ยวกับข้อมูลรูปภาพ ได้แก่ ใบหน้าตรง ใบหน้าด้านข้าง และใบหน้าสวมหน้ากากอนามัย หลังจากได้ข้อมูลรูปภาพมาแล้วนำมาคัดเลือกประเภทภาพที่ไม่สามารถทำการวิเคราะห์ได้ ได้แก่ ภาพปิดหน้าแบบมิดชิด และภาพที่ถ่ายไม่ให้เห็นใบหน้าแบบชัดเจน (ภาพกลางคืน) เป็นต้น เมื่อเลือกภาพที่สามารถนำไปใช้วิเคราะห์ได้แล้ว จะทำการจัดเก็บข้อมูลรูปภาพเหล่านั้นไว้ในฐานข้อมูลเรียลไทม์ของ Firebase

3.2 การคัดเลือกและเปรียบเทียบประสิทธิภาพของอัลกอริทึม

การใช้อัลกอริทึม 3 อัลกอริทึม เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมแต่ละอัลกอริทึม โดยผู้วิจัยทำการค้นหาตำแหน่งของใบหน้า (Face detection) ของแต่ละบุคคลจะทำให้ได้ค่า bounding box หลังจากนั้นจะนำค่า bounding box ไปทำกระบวนการการจดจำใบหน้า เพื่อเปรียบเทียบข้อมูลกับชุดข้อมูลใบหน้าที่ถูกจัดเก็บไว้ใน

ฐานข้อมูลเรียลไทม์ของ Firebase จำนวนทั้งหมด 200 รูปภาพ จากบุคคลทั้งหมด 10 คน โดยพิจารณาค่าความถูกต้อง เพื่อวัดความสามารถในการรู้จำใบหน้าของบุคคลคนนั้น และเพื่อหาอัลกอริทึมที่ดีที่สุด

4. ผลการดำเนินงาน

ผลการทดสอบประสิทธิภาพของอัลกอริทึม โดยพิจารณาค่าความถูกต้อง สำหรับ ใบหน้าตรง ใบหน้าด้านข้าง และใบหน้าสวมหน้ากากอนามัย แสดงดังตารางที่ 3 ถึง 11 โดยมีรายละเอียดดังนี้

4.1 การเปรียบเทียบใบหน้าที่กับฐานข้อมูล

ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึม ได้ผลลัพธ์ แสดงดังตารางที่ 2

ตารางที่ 2: การเปรียบเทียบใบหน้าที่กับฐานข้อมูล

Person	Straight Face	Sideways Face	Wearing Mask Faces
1			

4.2 ผลการทดสอบ อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ

ตารางที่ 3: Straight Face (CNN)

Person	bounding box	Confidence	Accuracy
1	155,162,229,85	0.25	88.79
2	122,163,184,100	0.25	80.75
3	76,116,166,26	0.25	90.15
4	113,121,189,46	0.25	91.24
5	130,158,182,106	0.25	86.75
6	123,137,196,63	0.25	90.65
7	129,142,191,80	0.25	89.65
8	108,156,170,93	0.25	84.36
9	122,163,184,100	0.25	86.98
10	176,147,228,95	0.25	90.64

จากตารางที่ 1 ค่าพารามิเตอร์ CNN ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [113,121,189,46] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 91.24 สำหรับใบหน้าที่ตรง ซึ่งใช้เวลาทุก ๆ 3 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 4: Sideways Face (CNN)

Person	bounding box	Confidence	Accuracy
1	136,285,220,115	0.25	87.66
2	76,175,106,96	0.25	80.03
3	92,270,199,163	0.25	80.88
4	82,225,211,96	0.25	84.88
5	118,127,161,84	0.25	81.66
6	97,270,199,163	0.25	84.47
7	121,171,196,96	0.25	86.97
8	113,154,188,79	0.25	81.57
9	105,162,179,83	0.25	81.91
10	147,118,193,96	0.25	86.78

จากตารางที่ 1 ค่าพารามิเตอร์ CNN ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [136,285,220,115] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 87.66 สำหรับใบหน้าที่ด้านข้าง ซึ่งใช้เวลาทุก ๆ 3 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 5: Wearing Mask Faces (CNN)

Person	bounding box	Confidence	Accuracy
1	138,162,213,88	0.25	80.23
2	96,129,171,55	0.25	80.11
3	130,171,204,96	0.25	80.45
4	86,155,170,66	0.25	81.77
5	141,258,183,154	0.25	80.03
6	116,155,206,66	0.25	80.87
7	121,171,196,96	0.25	79.26
8	115,163,177,180	0.25	80.48
9	113,162,180,88	0.25	82.14
10	139,158,183,106	0.25	80.23

จากตารางที่ 1 ค่าพารามิเตอร์ CNN ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [113,162,180,88] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 82.14 สำหรับใบหน้าที่สวมหน้ากากอนามัย ซึ่งใช้เวลาทุก ๆ 3 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

4.3. ผลการทดสอบ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

ตารางที่ 6: Straight Face (SVM)

Person	bounding box	Confidence	Accuracy
1	155,196,129,121	0.25	76.44

Person	bounding box	Confidence	Accuracy
2	211,196,140,67	0.25	75.69
3	76,106,166,16	0.25	81.87
4	84,145,254,56	0.25	80.14
5	141,158,193,106	0.25	79.87
6	113,137,188,63	0.25	81.02
7	129,147,191,89	0.25	75.66
8	100,148,170,93	0.25	82.03
9	122,156,104,93	0.25	75.23
10	178,147,228,95	0.25	83.78

จากตารางที่ 1 ค่าพารามิเตอร์ SVM ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [178,147,228,95] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 83.78 สำหรับใบหน้าตรง ซึ่งใช้เวลาทุก ๆ 0.5 วินาทีสำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 7: Sideways Face (SVM)

Person	bounding box	Confidence	Accuracy
1	136,205,226,216	0.25	72.63
2	108,128,170,66	0.25	72.66
3	66,116,156,20	0.25	79.88
4	92,175,199,67	0.25	72.68
5	147,197,199,135	0.25	75.12
6	121,137,196,63	0.25	79.64
7	138,154,213,79	0.25	73.46
8	96,165,186,76	0.25	79.65
9	130,171,204,96	0.25	73.89
10	164,170,216,110	0.25	80.92

จากตารางที่ 1 ค่าพารามิเตอร์ SVM ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [164,170,216,110] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 80.92 สำหรับใบหน้าด้านข้าง ซึ่งใช้เวลาทุก ๆ 0.5 วินาทีสำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 8: Wearing Mask Faces (SVM)

Person	bounding box	Confidence	Accuracy
1	146,167,221,86	0.25	72.16
2	55,162,130,88	0.25	66.89
3	153,164,205,112	0.25	72.15
4	86,145,176,56	0.25	73.65
5	124,147,176,101	0.25	71.49
6	96,167,225,38	0.25	77.96

Person	bounding box	Confidence	Accuracy
7	121,167,196,83	0.25	71.84
8	115,156,177,93	0.25	77.48
9	121,179,196,104	0.25	80.11
10	141,158,193,106	0.25	72.03

จากตารางที่ 1 ค่าพารามิเตอร์ SVM ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [121,179,196,104] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 80.11 สำหรับใบหน้าสวมหน้ากากอนามัย ซึ่งใช้เวลาทุก ๆ 0.5 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

4.4 ผลการทดสอบด้วย อัลกอริทึมเทคนิคฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์

ตารางที่ 9: Straight Face (HOG)

Person	bounding box	Confidence	Accuracy
1	138,196,213,121	0.25	75.93
2	76,145,166,56	0.25	74.91
3	66,116,156,20	0.25	80.72
4	82,253,211,124	0.25	78.18
5	141,181,193,129	0.25	78.67
6	105,137,179,63	0.25	80.78
7	121,146,191,96	0.25	74.82
8	105,154,179,79	0.25	81.93
9	135,163,193,100	0.25	74.98
10	176,163,232,100	0.25	83.52

จากตารางที่ 1 ค่าพารามิเตอร์ HOG ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [176,163,232,100] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 83.52 สำหรับใบหน้าตรง ซึ่งใช้เวลาทุก ๆ 0.5 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 10: Sideways Face (HOG)

Person	bounding box	Confidence	Accuracy
1	138,204,213,129	0.25	72.56
2	76,145,166,46	0.25	71.86
3	76,116,156,26	0.25	79.23
4	92,235,199,127	0.25	73.89
5	141,170,193,118	0.25	74.49
6	113,113,108,38	0.25	78.92
7	130,154,213,79	0.25	72.56
8	96,165,196,76	0.25	78.65
9	115,176,179,104	0.25	72.91
10	164,170,216,118	0.25	80.87

จากตารางที่ 1 ค่าพารามิเตอร์ HOG ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [164,170,216,118] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 80.87 สำหรับใบหน้าด้านข้าง ซึ่งใช้เวลาทุก ๆ 0.5 วินาทีสำหรับการจดจำใบหน้าแต่ละครั้ง

ตารางที่ 11: Wearing Mask Faces (HOG)

Person	bounding box	Confidence	Accuracy
1	146,171,221,96	0.25	71.83
2	55,162,139,88	0.25	64.71
3	129,163,191,108	0.25	70.68
4	86,145,176,56	0.25	73.01
5	124,147,176,95	0.25	70.88
6	98,242,204,56	0.25	76.61
7	121,171,196,96	0.25	71.74
8	115,163,177,100	0.25	75.93
9	135,170,198,207	0.25	75.59
10	141,152,193,101	0.25	71.96

จากตารางที่ 1 ค่าพารามิเตอร์ HOG ที่ให้ประสิทธิภาพดีที่สุด คือ bounding box = [98,242,204,56] และค่า Confidence = 0.25 โดยให้ค่าความถูกต้องเฉลี่ยร้อยละ 76.61 สำหรับใบหน้าสวมหน้ากากอนามัย ซึ่งใช้เวลาทุก ๆ 0.5 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการเปรียบเทียบอัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเทคนิคฮิสโตรแกรมของทิศทางตามค่าเกรเดียนท์ ผลการทดสอบประสิทธิภาพ พบว่าอัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการให้ประสิทธิภาพดีที่สุด สำหรับใบหน้าตรงให้ค่าความถูกต้องร้อยละเฉลี่ย เท่ากับ 91.24 สำหรับใบหน้าด้านข้างให้ค่าความถูกต้องร้อยละเฉลี่ยเท่ากับ 87.66 และสำหรับใบหน้าสวมหน้ากากอนามัยให้ค่าความถูกต้องร้อยละเฉลี่ย เท่ากับ 82.14 ตามลำดับ ซึ่งใช้เวลาทุก ๆ 3 วินาที สำหรับการจดจำใบหน้าแต่ละครั้ง ซึ่งอัลกอริทึมโครงข่ายประสาทเทียมแบบสังวัตนาการ สามารถให้ค่าความถูกต้องดีกว่าอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึม

เทคนิคฮิสโตรแกรมของทิศทางตามค่าเกรเดียนท์ตามลำดับ

เอกสารอ้างอิง

- [1] G. Huilin and et al, "Robust face recognition based on multi-task Convolutional Neural network" Electronic Information University of Science and Technology, pp. 195-209, 2021.
- [2] L. Madan and et al, "Study of face recognition techniques: A survey International Journal of Advanced Computer Science and Applications, vol. 9, 2018.
- [3] J. Automatic and et al, "Attendance System Based on Face Recognition Using HOG Features and Cosine Distance and et al., "Automatic Attendance System Based on Face Recognition Using HOG Features and Cosine Distance", International Conference on Future Data and Security Engineering, pp 130–148.
- [4] F. Yuchun and et al, "Multi-Stage CNN Architecture for Face Mask Detection iPing Data Labs LLP", Mumbai, 2020.
- [5] P. Jaruwit, "What is Support Vector Machine (SVM) [online]. 2019, source: <https://www.glurgeek.com/education/support-vector-machine/>, [15 February 2566].
- [6] T. Mrinal, "HOG (Histogram of Oriented Gradients): An Overview [online]. 2021, source: <https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f/>, [15 February 2566].
- [7] รามสวร์ คิวรักษา และคณะ, "การตรวจจับการสวมหน้ากากอนามัยโดยประยุกต์ใช้โครงข่ายประสาท เทียมแบบเชิงลึกและโครงข่ายประสาทเทียมแบบคอนโวลูชัน", เทคโนโลยีอุตสาหกรรมมหาวิทยาลัยราชภัฏกำแพงเพชร 2564.
- [8] G. Suguna and et al, "Face Recognition System for Real Time Applications using SVM Combined with FaceNet and MTCNN" International Journal of Electrical Engineering and Technology, pp.328-335, 2021.
- [9] G. Mohammed and I. Melhum, "Implementation of HOG Feature Extraction with Tuned Parameters for Human Face Detection" International Journal of Machine Learning and Computing, Vol. 10, No. 5, pp. 195-209, 2020.

การพยากรณ์ลูกหนี้ที่ค้างชำระหนี้ของสินเชื่อเพื่อที่อยู่อาศัยด้วยเทคนิคเหมืองข้อมูล

Predictive of Outstanding Debts for Housing Loan by Data Mining

กมลชนก รุ่งเจริญ (Kamonchanok Rungjaroen)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857247@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการสร้างแบบจำลองการพยากรณ์ลูกหนี้ที่ค้างชำระหนี้ของสินเชื่อเพื่อที่อยู่อาศัย โดยเปรียบเทียบอัลกอริทึม จำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน อัลกอริทึมเพื่อนบ้านใกล้ที่สุด และอัลกอริทึมเอ็กซ์ตรีมกราดิเอนบูตติ้ง โดยใช้ข้อมูลพอร์ตสินเชื่อเพื่อที่อยู่อาศัย ตั้งแต่เดือนธันวาคม 2561 - เดือนธันวาคม 2562 จำนวน 28,418 รายการ ประกอบด้วยคุณลักษณะ จำนวน 12 คุณลักษณะ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลองจำนวน 19,896 รายการ และข้อมูลสำหรับทดสอบแบบจำลองจำนวน 8,522 รายการ ทำการเปรียบเทียบประสิทธิภาพด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล พบว่า อัลกอริทึมเอ็กซ์ตรีมกราดิเอนบูตติ้ง ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 86.84 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลอกเท่ากับ 0.87 และค่าการถ่วงดุลเท่ากับ 0.87

คำสำคัญ: อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

อัลกอริทึมเพื่อนบ้านใกล้ที่สุด

อัลกอริทึมเอ็กซ์ตรีมกราดิเอนบูตติ้ง

สินเชื่อเพื่อที่อยู่อาศัย

Abstract

This research presents a forecasting model for outstanding debts for housing loans that compares three machine learning algorithms: Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost). The housing loan portfolio dataset from December 2018 to

December 2019 consists of 28,418 records and 12 attributes. The dataset is split into two parts: training data of 19,896 records and testing data of 8,522 records. This research has measured performance models with accuracy, precision, recall, and f-measure. The result showed that the Extreme Gradient Boosting (XGBoost) algorithm had the highest accuracy with an accuracy 86.84%, precision 0.87, recall 0.87, and f-measure 0.87.

Keywords: Support Vector Machines, K-Nearest Neighbors, Extreme Gradient Boosting, Housing Loans.

1. บทนำ

เนื่องจากที่อยู่อาศัยเป็นปัจจัยสำคัญ และมีความจำเป็นสำหรับการดำรงชีวิตของมนุษย์ และเป็นสิ่งที่มีอิทธิพลต่อชีวิตความเป็นอยู่ สถาบันการเงินจึงมีการให้สินเชื่อเพื่อที่อยู่อาศัยกับประชาชนที่ต้องการมีที่อยู่อาศัยให้เข้าถึงแหล่งเงินทุนอย่างทั่วถึง ในปี 2563 จนถึงปัจจุบัน การแพร่ระบาดของ COVID-19 ได้สร้างผลกระทบเชิงลบทางเศรษฐกิจต่อประชาชน และผู้ประกอบการในหลายภาคส่วน จึงส่งผลให้สถาบันการเงินต้องเผชิญกับปัจจัยเสี่ยงต่าง ๆ ที่เกิดจากผลกระทบดังกล่าว อาทิเช่น ความสามารถในการชำระหนี้ของลูกหนี้โดยรวมลดลง ไม่สามารถกลับมาชำระหนี้ได้ตามปกติ ขาดสภาพคล่อง ปัญหาการว่างงาน รายได้ลดลง และภาวะหนี้ครัวเรือนอยู่ในระดับที่สูง ประกอบกับอ้างอิงข้อมูลจำนวนหนี้ที่ไม่ก่อให้เกิดรายได้ (Non-Performing Loans: NPLs) และอัตรา NPLs ของพอร์ตสินเชื่อเพื่อที่อยู่อาศัย ย้อนหลัง

5 ปี ตั้งแต่ปี 2560 - 2564 จากธนาคารแห่งประเทศไทย [1] พบว่า ตั้งแต่ปี 2563 มีแนวโน้มจำนวน NPLs เพิ่มขึ้นอย่างต่อเนื่อง ซึ่งถ้าหากสถาบันการเงินมีปริมาณ NPLs ที่สูง สถาบันการเงินจะต้องมีการกันเงินสำรองตามหลักเกณฑ์ของธนาคารแห่งประเทศไทย และสูญเสียรายได้ดอกเบี้ยตามเกณฑ์คงค้าง ซึ่งจะส่งผลกระทบต่อสภาพคล่องในการดำเนินงานของสถาบันการเงิน

ในการพยากรณ์ลูกหนี้ที่ค้างชำระหนี้หรือการจัดกลุ่มลูกหนี้เพื่อใช้ในการพิจารณาอนุมัติสินเชื่อ ได้มีผู้วิจัยนำเสนออัลกอริทึมต่าง ๆ มาช่วยในการวิเคราะห์ข้อมูลเหล่านี้ เช่น การใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนในการเปรียบเทียบอัลกอริทึมเหมือนข้อมูลสำหรับพิจารณาการให้สินเชื่อสำหรับธนาคาร [2] การใช้อัลกอริทึมเพื่อนบ้านใกล้ที่สุดเพื่อตรวจจับการฉ้อโกงเงินอิเล็กทรอนิกส์ [3] การใช้อัลกอริทึมเอ็กซ์ตรีมกราดิเอนท์บูตติ้งในการคาดการณ์การผิดนัดชำระหนี้ของธนาคาร [4] การใช้อัลกอริทึมโครงข่ายประสาทเทียมในการการพัฒนาแบบจำลองการพิจารณาให้คะแนนสินเชื่อ [5] และ การใช้อัลกอริทึม Deep Neural Networks ในการเปรียบเทียบอัลกอริทึมเหมือนข้อมูลสำหรับการทำนายการผิดนัดชำระหนี้ที่อยู่อาศัย [6] เป็นต้น

จากปัญหา และงานวิจัยที่กล่าวมาข้างต้น ผู้วิจัยจึงมีแนวคิดในการนำเหมือนข้อมูลมาช่วยในการประมวลผลเกี่ยวกับการพยากรณ์ลูกหนี้ที่ค้างชำระหนี้ของสินเชื่อที่อยู่อาศัย เพื่อวิเคราะห์คุณภาพของกลุ่มลูกหนี้พอร์ตสินเชื่อที่อยู่อาศัย ซึ่งสามารถใช้เป็นข้อมูลสนับสนุนสำหรับหน่วยงานที่เกี่ยวข้องนำผลการวิจัยไปวางแผนวิเคราะห์ หรือกำหนดมาตรการต่าง ๆ เพื่อลดโอกาสการผิดนัดชำระหนี้ของลูกหนี้ต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 อัลกอริทึมการพยากรณ์

2.1.1 อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) ใช้การสร้างสมการเส้นตรงเพื่อแบ่งข้อมูล 2 กลุ่มออกจากกัน โดย SVM สร้างเส้นแบ่งตรงกึ่งกลางระหว่างกลุ่มให้มีระยะห่างระหว่างขอบเขตของทั้งสองกลุ่มให้มากที่สุด SVM อาศัยหลักการแปลงข้อมูลพื้นที่ข้อมูลนำเข้า (Input Space) ให้เป็นพื้นที่คุณลักษณะ (Feature Space) โดยผ่านฟังก์ชันวัดความคล้ายที่เรียกว่า Kernel Function [7]

2.1.2 อัลกอริทึมเพื่อนบ้านใกล้ที่สุด (K- Nearest Neighbor: KNN) เป็นการแบ่งกลุ่มข้อมูล และทำการวัดระยะห่างระหว่างข้อมูลที่ต้องการทำนายกับข้อมูลที่อยู่ใกล้เคียงเป็นจำนวน K ตัว จะเรียงลำดับผลลัพธ์ที่ได้จากการคำนวณค่าระยะห่างจากน้อยไปมาก หลังจากนั้นดูจากค่า K ว่ากำหนดเป็นเท่าไร แล้วนำลำดับที่เรียงได้มาเทียบกับคลาสข้อมูลแล้วนำมาตอบ โดย K เป็นจำนวนของกลุ่มที่ต้องการเลือกเป็นกลุ่มเพื่อนบ้าน ควรกำหนดเป็นจำนวนคี่ [8]

2.1.3 อัลกอริทึมเอ็กซ์ตรีม กราดิเอนท์บูตติ้ง (Extreme Gradient Boosting: XGBoost) เป็นวิธีการ Boosting โดยการทำงานที่นำต้นไม้มากมาย ต้นไม้ทำงานต่อเนื่องกันโดยมีการเพิ่มน้ำหนักให้ข้อมูลที่ถูกทำนายผิด และลดน้ำหนักให้ข้อมูลที่มีการทำนายถูก เพื่อให้การเรียนรู้ครั้งต่อไปมีการเรียนรู้ความผิดพลาดจากการเรียนรู้รอบก่อนหน้า และมีความสามารถในการลดปัญหา Overfitting ของแบบจำลอง และสามารถจัดการกับค่าว่างได้อัตโนมัติ [9]

2.2 งานวิจัยที่เกี่ยวข้อง

Koçoğlu and Ersöz [10] ได้นำเสนองานวิจัยเกี่ยวกับแอปพลิเคชันการทำเหมือนข้อมูล เพื่อเพิ่มประสิทธิภาพการตัดสินใจข้อมูลการเงินที่มีความเสี่ยง โดยใช้ชุดข้อมูลเครดิตจากเว็บไซต์ LendingClub.com จำนวน 9,578 รายการ ตั้งแต่ช่วงเดือนพฤษภาคม 2550 ถึงกุมภาพันธ์ 2553 แบ่งข้อมูลออกเป็นชุดเรียนรู้ร้อยละ 80 และชุดทดสอบร้อยละ 20 ซึ่งตัวแปรตามและตัวแปรอิสระที่เกี่ยวข้องกับชุดข้อมูลนี้ถูกกำหนดตามตัวแปรที่ใช้บ่อย

ในงานวิจัยที่เกี่ยวข้องและความคิดเห็นของผู้เชี่ยวชาญ รวมถึงข้อมูลความสัมพันธ์หรือประวัติของลูกค้าที่เคยมีกับธนาคารในอดีต เพื่อประเมินความน่าจะเป็นของการไม่ชำระคืนของสินเชื่อบุคคล โดยทำการเปรียบเทียบประสิทธิภาพด้วย 2 อัลกอริทึม คือ อัลกอริทึมการวิเคราะห์ถดถอยโลจิสติก และอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ผลการวิจัยพบว่า อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ให้ประสิทธิภาพดีที่สุด ค่าความถูกต้องร้อยละ 83.72 และค่าพื้นที่ใต้กราฟร้อยละ 58.10

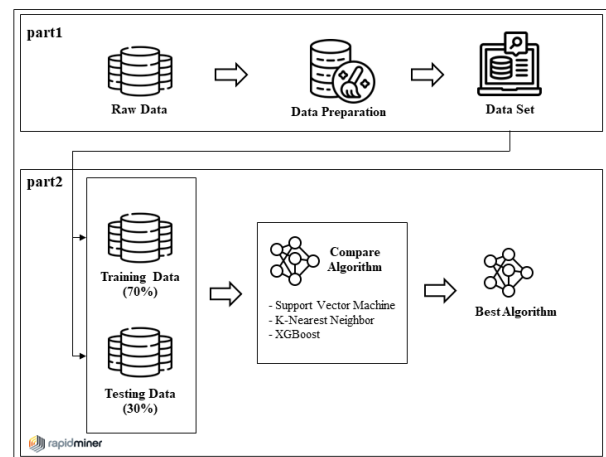
ว่าทิพย์ จักรชัย และเพชร [8] งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความรู้สึกเกี่ยวกับความคิดเห็นต่อโรงแรมในประเทศไทย งานวิจัยได้ประยุกต์ใช้เทคนิคการวิเคราะห์ความรู้สึกในรูปแบบเหมือนความคิดเห็นต่อการเข้ารับบริการ งานวิจัยนี้มีการรวบรวมความคิดเห็นจากเว็บไซต์ APT TUBE จำนวน 10,000 ประโยค แล้วจึงทำการวิเคราะห์ความคิดเห็น ด้วยเทคนิค K-means โดยแบ่งออกเป็น 5 กลุ่ม ได้แก่ การเข้าถึง กิจกรรมและความบันเทิง อาหารและเครื่องดื่ม พนักงานผู้ให้บริการ และสถานที่ จากนั้นจึงนำข้อความแต่ละกลุ่มเข้าสู่การแยกประเภทด้วยอัลกอริทึมเพื่อนบ้านใกล้ที่สุด และวัดประสิทธิภาพด้วยค่าความแม่นยำ ค่าความระลึกลับ ค่าความถูกต้อง ค่า F-Measure ร่วมกับ 10 Fold Cross Validation โดยเปรียบเทียบกับเทคนิคอัลกอริทึมต้นไม้การตัดสินใจ อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และอัลกอริทึมเพื่อนบ้านใกล้ที่สุด โดยเทคนิคที่ให้ค่าความถูกต้องมากที่สุด คือ เทคนิค K-means ร่วมกับอัลกอริทึมเพื่อนบ้านใกล้ที่สุด ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 94.8

Wang and et al. [11] ได้นำเสนองานวิจัยเกี่ยวกับการประเมินความเสี่ยงจากการผิดนัดชำระหนี้ในประเทศจีนด้วยอัลกอริทึมเอ็กริมกราเดียนบูตติ่ง เนื่องจากตลาดตราสารหนี้เป็นส่วนสำคัญของตลาดทุนของประเทศจีน แต่การผิดสัญญาในตลาดตราสารหนี้มักเกิดขึ้นบ่อยครั้งในช่วงไม่กี่ปีที่ผ่านมา จึงทำให้ต้องมีการประเมินความเสี่ยงจากการผิดนัดชำระหนี้ ด้วยตัวชี้วัดระดับเศรษฐกิจมหภาคและระดับจุลภาค รวมถึงวิเคราะห์องค์ประกอบหลักเพื่อแยกองค์ประกอบหลัก 10 อย่าง

นอกจากตัวชี้วัดเหล่านี้ และทำการทดสอบประสิทธิภาพด้วยอัลกอริทึมเอ็กริมกราเดียนบูตติ่ง โดยใช้เทคนิค 5-fold Cross-Validation เพื่อเพิ่มประสิทธิภาพสำหรับการประเมินความเสี่ยงการผิดนัดชำระหนี้ของตราสารหนี้ ผลการวิจัยพบว่า อัลกอริทึมเอ็กริมกราเดียนบูตติ่งทดสอบด้วยวิธี 5-fold CV ให้ประสิทธิภาพสูงที่สุด ค่าความถูกต้อง ร้อยละ 96.81 ค่าพื้นที่ใต้กราฟ ร้อยละ 97.99 ค่าความแม่นยำร้อยละ 95.45 ค่าระลึกร้อยละ 60 และค่า F1-score ร้อยละ 73.68

3. วิธีการดำเนินการวิจัย

การดำเนินงานวิจัยเพื่อการพยากรณ์ลูกหนี้ที่ค้างชำระหนี้ของสินเชื่อเพื่อที่อยู่อาศัย ผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวความคิด

จากภาพที่ 1 กรอบแนวความคิด สามารถแบ่งขั้นตอนการดำเนินการวิจัยเป็น 2 ส่วน ดังนี้

3.1 ขั้นตอนการรวบรวม และเตรียมข้อมูล

ข้อมูลที่ใช้วิจัยเป็นข้อมูลพอร์ตสินเชื่อเพื่อที่อยู่อาศัย ตั้งแต่เดือนธันวาคม 2561 - เดือนธันวาคม 2562 จำนวน 28,418 รายการ ประกอบด้วยคุณลักษณะจำนวน 12 คุณลักษณะ ได้แก่ เลขที่บัญชี อายุ อาชีพ เงินกู้ ป้อนุมัติ ระยะเวลาที่อนุมัติสัญญา รายได้ การปรับปรุงโครงสร้างหนี้ (Troubled Debt Restructuring: TDR) จำนวนครั้งการทำ TDR ประเภทหลักประกัน อัตราส่วน

เงินให้สินเชื่อก่อมูลค่าหลักประกัน (LTV Ratio) และสถานะ Performance 12 เดือน แสดงดังตารางที่ 1

ตารางที่ 1: คุณลักษณะของข้อมูล

ลำดับ	คุณลักษณะ
1	Account_no
2	Age
3	Cust_occ
4	Principal amount
5	Signdate
6	TimeContract
7	Income
8	Tdr_flag
9	Tdr_seq
10	Guarantee
11	Ltv_range
12	Performance 12m_status

หลังจากการรวบรวมข้อมูลข้างต้น ขั้นตอนต่อไป คือการเตรียมข้อมูล โดยการทำความสะอาดข้อมูล ตรวจสอบความครบถ้วนของข้อมูล แทนค่าข้อมูลอาชีพที่หายไปด้วยอาชีพว่างงาน และแปลงข้อมูลอายุ วงเงินกู้ และรายได้ให้อยู่ในรูปแบบช่วงข้อมูล และกำหนดเลขที่บัญชีเป็น ID และสถานะของลูกค้า Performance 12 เดือนเป็นค่าเป้าหมาย จากนั้นเตรียมเข้าสู่ขั้นตอนการคัดเลือกและทดสอบประสิทธิภาพอัลกอริทึมเพื่อสร้างแบบจำลองต่อไป

3.2 ขั้นตอนการคัดเลือก และทดสอบประสิทธิภาพอัลกอริทึม

นำข้อมูลที่ได้จากการเตรียมข้อมูลมาทำการแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลอง ร้อยละ 70 จำนวน 19,896 รายการ และข้อมูลสำหรับทดสอบแบบจำลอง ร้อยละ 30 จำนวน 8,522 รายการ โดยทำการเปรียบเทียบอัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึม SVM อัลกอริทึม KNN และอัลกอริทึม XGBoost ซึ่งแต่ละอัลกอริทึมกำหนดพารามิเตอร์ที่สนใจ และทำการปรับค่าพารามิเตอร์ตามที่กำหนดไว้ และทำการวัดประสิทธิภาพของอัลกอริทึมจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล และทำซ้ำแต่ละรูปแบบจำนวน 5 ครั้ง เพื่อคัดเลือกหาอัลกอริทึมที่ดีที่สุด โดยค่าพารามิเตอร์สำหรับการทดสอบอัลกอริทึม แสดงดังตารางที่ 2

ตารางที่ 2: ค่าพารามิเตอร์สำหรับการทดสอบอัลกอริทึม

SVM	c = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 kernel = linear, poly, rbf, simoid gamma = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
KNN	k = 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41 measure types = mixed measure mixed measure = mixed euclidean distance
XGBoost	round = 25, 50, 100, 200, 300, 500, 1000 max depth = 5, 6, 7, 8, 9, 10 learning rate = 0.1, 0.2, 0.3, 0.4

4. ผลการดำเนินงาน

ผลการทดสอบประสิทธิภาพอัลกอริทึม แสดงดังตารางที่ 3 ถึง 5 โดยมีรายละเอียดดังนี้

ตารางที่ 3: ผลการทดสอบประสิทธิภาพอัลกอริทึม SVM

Parameters			Performance			
c	kernel	gamma	Acc	P	R	F
0	linear	0	47.83	0.33	0.48	0.39
0	poly	0	47.68	0.32	0.48	0.39
0	rbf	0	79.53	0.81	0.80	0.80
0	simoid	0	49.98	0.25	0.50	0.33
0	rbf	0.1	86.60	0.87	0.87	0.87
0	rbf	0.2	86.59	0.87	0.87	0.87
0	rbf	0.3	86.42	0.86	0.86	0.86
0	rbf	0.4	86.27	0.86	0.86	0.86
0	rbf	0.5	86.18	0.86	0.86	0.86
0	rbf	0.6	86.04	0.86	0.86	0.86
0	rbf	0.7	85.79	0.86	0.86	0.86
0	rbf	0.8	85.40	0.85	0.85	0.85
0	rbf	0.9	84.89	0.85	0.85	0.85
0	rbf	1	84.47	0.85	0.84	0.85
0.1	rbf	0.1	86.36	0.86	0.86	0.86
0.2	rbf	0.1	86.44	0.86	0.86	0.86
0.3	rbf	0.1	86.51	0.87	0.87	0.87
0.4	rbf	0.1	86.61	0.87	0.87	0.87
0.5	rbf	0.1	86.63	0.87	0.87	0.87
0.6	rbf	0.1	86.66	0.87	0.87	0.87
0.7	rbf	0.1	86.62	0.87	0.87	0.87
0.8	rbf	0.1	86.61	0.87	0.87	0.87
0.9	rbf	0.1	86.60	0.87	0.87	0.87
1	rbf	0.1	86.60	0.87	0.87	0.87

** Acc = Accuracy, P = Precision, R = Recall และ F = F-measure

จากตารางที่ 3 ค่าพารามิเตอร์ของ SVM ที่ให้ประสิทธิภาพดีที่สุด คือ c เท่ากับ 0.6 kernel type เท่ากับ rbf และ gamma เท่ากับ 0.1 โดยให้ค่าความถูกต้องเท่ากับ ร้อยละ 86.66 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลอกเท่ากับ 0.87 และค่าการถ่วงดุลเท่ากับ 0.87 ตามลำดับ

เท่ากับ mixed measure และ mixed euclidean distance โดยให้ค่าความถูกต้องเท่ากับ ร้อยละ 85.05 ค่าความแม่นยำเท่ากับ 0.85 ค่าระลอกเท่ากับ 0.85 และค่าการถ่วงดุลเท่ากับ 0.85 ตามลำดับ

ตารางที่ 4: ผลการทดสอบประสิทธิภาพอัลกอริทึม KNN

Parameters			Performance			
k	measure types	mixed measure	Acc	P	R	F
3	mixed measure	mixed euclidean distance	77.76	0.80	0.78	0.79
5	mixed measure	mixed euclidean distance	78.63	0.81	0.79	0.80
7	mixed measure	mixed euclidean distance	80.60	0.82	0.81	0.81
9	mixed measure	mixed euclidean distance	81.85	0.83	0.82	0.82
11	mixed measure	mixed euclidean distance	82.93	0.84	0.83	0.83
13	mixed measure	mixed euclidean distance	83.68	0.84	0.84	0.84
15	mixed measure	mixed euclidean distance	84.10	0.85	0.84	0.84
17	mixed measure	mixed euclidean distance	84.43	0.85	0.84	0.85
19	mixed measure	mixed euclidean distance	84.63	0.85	0.85	0.85
21	mixed measure	mixed euclidean distance	84.81	0.85	0.85	0.85
23	mixed measure	mixed euclidean distance	84.93	0.85	0.85	0.85
25	mixed measure	mixed euclidean distance	85.05	0.85	0.85	0.85

จากตารางที่ 4 ค่าพารามิเตอร์ของ KNN ที่ให้ประสิทธิภาพดีที่สุด คือ k เท่ากับ 25 measure types

ตารางที่ 5: ผลการทดสอบประสิทธิภาพอัลกอริทึม XGBoost

Parameters			Performance			
round	max depth	learning rate	Acc	P	R	F
25	6	0.1	86.73	0.87	0.87	0.87
25	6	0.2	86.65	0.87	0.87	0.87
25	6	0.3	86.70	0.87	0.87	0.87
25	7	0.1	86.84	0.87	0.87	0.87
25	7	0.2	86.68	0.87	0.87	0.87
25	7	0.3	86.51	0.87	0.87	0.87
25	8	0.1	86.59	0.87	0.87	0.87
25	8	0.2	86.44	0.86	0.86	0.86
25	8	0.3	86.25	0.86	0.86	0.86
50	6	0.1	86.73	0.87	0.87	0.87
50	6	0.2	86.50	0.87	0.87	0.87
50	6	0.3	86.34	0.86	0.86	0.86
50	7	0.1	86.65	0.87	0.87	0.87
50	7	0.2	86.61	0.87	0.87	0.87
50	7	0.3	86.50	0.87	0.86	0.87
50	8	0.1	86.49	0.87	0.86	0.87
50	8	0.2	86.22	0.86	0.86	0.86
50	8	0.3	86.23	0.86	0.86	0.86

จากตารางที่ 5 ค่าพารามิเตอร์ของ XGBoost ที่ให้ประสิทธิภาพดีที่สุด คือ round เท่ากับ 25 max depth เท่ากับ 7 และ learning rate เท่ากับ 0.1 โดยให้ค่าความถูกต้องเท่ากับร้อยละ 86.84 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลอกเท่ากับ 0.87 และค่าการถ่วงดุลเท่ากับ 0.87 ตามลำดับ

จากผลการทดสอบประสิทธิภาพอัลกอริทึมทั้ง 3 อัลกอริทึม เมื่อทำการทำซ้ำทั้งหมด 5 ครั้ง แต่ละอัลกอริทึมได้ค่าเท่าเดิมทุกครั้งแสดงดังตารางที่ 6

ตารางที่ 6: ผลการเปรียบเทียบอัลกอริทึม

Algorithm	Performance			
	Acc	P	R	F
SVM	86.66	0.87	0.87	0.87
KNN	85.05	0.85	0.85	0.85
XGBoost	86.84	0.87	0.87	0.87

จากตารางที่ 6 สรุปได้ว่า อัลกอริทึม XGBoost ให้ประสิทธิภาพในการพยากรณ์ที่ดีที่สุด โดยให้ค่าความถูกต้องเท่ากับร้อยละ 86.84 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลึกเท่ากับ 0.87 และค่าการถ่วงดุลเท่ากับ 0.87 ตามลำดับ

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยนำเสนอการสร้างแบบจำลองการพยากรณ์ลูกหนี้ที่ค้างชำระหนี้ของสินเชื่อเพื่อที่อยู่อาศัย จากข้อมูลพอร์ตสินเชื่อเพื่อที่อยู่อาศัย ตั้งแต่เดือนธันวาคม 2561 - เดือนธันวาคม 2562 จำนวน 28,418 รายการ ประกอบด้วยคุณลักษณะ จำนวน 12 คุณลักษณะ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลองจำนวน 19,896 รายการ และข้อมูลสำหรับทดสอบแบบจำลองจำนวน 8,522 รายการ ทำการสร้างแบบจำลอง และทดสอบประสิทธิภาพอัลกอริทึม จำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึม SVM อัลกอริทึม KNN และอัลกอริทึม XGBoost ผลการวิจัยพบว่า อัลกอริทึม XGBoost ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 86.84 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลึกเท่ากับ 0.87 และค่าการถ่วงดุลเท่ากับ 0.87 ตามลำดับ

เอกสารอ้างอิง

- [1] Bank of Thailand (2022). [online]. Outstanding Gross NPLs Classified by Business (ISIC Rev.4). [cited 22 July 2022]. Available from: URL: <https://www.bot.or.th/Thai/Statistics/FinancialInstitutions/StatNPLsOutstanding/Pages/default.aspx>
- [2] ขวัญณา พิมพ์ชารีย์, การใช้เหมืองข้อมูลสำหรับพิจารณาการให้สินเชื่อสำหรับธนาคาร. วิทยานิพนธ์วิทยาศาสตร์

มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม, 2564.

- [3] B. Karunachandra and et al, "On the benefits of machine learning classification in cashback fraud detection" *Procedia Computer Science*, vol. 216, pp. 364-369, 2023.
- [4] R. Odegua, "Predicting bank loan default with extreme gradient boosting" *Computer Science*, 2020.
- [5] ชลลดา ม่วงธัญ สุรศักดิ์ มั่งสิงห์ และนิเวศ จิระวิชิตชัย. "การพัฒนาแบบจำลองการพิจารณาให้คะแนนสินเชื่อโดยใช้เทคนิคเหมืองข้อมูล" *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหัวเว่ยเฉลิมพระเกียรติ* ปีที่ 7 ฉบับที่ 1 มกราคม - มิถุนายน พ.ศ. 2564 หน้า 82-93.
- [6] A. Matre, *Machine Learning in Default Prediction : The Incremental Power of Machine Learning Techniques in Mortgage Default Prediction*. Master's Thesis Economics and Business Administration Major in Finance, 2019.
- [7] เทอดศักดิ์ เงินมูล และคณะ, "การคัดแยกความสูงสโตรบอร์ด้วยซอฟต์แวร์เวกเตอร์แมชชีน" *วารสารวิศวกรรมสารมหาวิทยาลัยยเรศวร* ปีที่ 12 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2560 หน้า 65-62.
- [8] วาทีศย์ คำพรมมา จักรชัย โสอินทร์ และ เพชร อิ่มทองคำ, "แบบจำลองการวิเคราะห์ความรู้สึกแบบผสมสำหรับความคิดเห็นต่อโรงแรมในประเทศไทยโดยใช้ K-means และ K-NN" *วลัยลักษณ์วิจัย* ฉบับที่ 7 มิถุนายน พ.ศ. 2562.
- [9] กาญจนมาศ เปลี่ยนสกุล, การจัดกลุ่มลูกค้าบริษัทยานยนต์ด้วยข้อมูลประชากรโดยใช้เทคนิคการเรียนรู้ของเครื่อง. การค้นคว้าอิสระวิทยาศาสตร์มหาบัณฑิต สาขาวิชาวิทยาการข้อมูล มหาวิทยาลัยศรีนครินทรวิโรฒ, 2564.
- [10] E. Koçoğlu and F. Ersöz "Data Mining Application for Financial Decision Optimization at Risk" *International Journal of 3D Printing Technologies and Digital Industry*, vol. 5, no. 2, pp. 195-209, 2021.
- [11] J. Wang and et al, "Credit Debt Default Risk Assessment Based on the XGBoost Algorithm : An Empirical Study from China" *Wireless Communications and Mobile Computing*. vol. 2022, 2022.

การเปรียบเทียบอัลกอริทึมสำหรับการพยากรณ์มูลค่าการส่งออกของวิสาหกิจ ขนาดกลางและขนาดย่อมของไทย

A Comparison of Algorithms for Export Value Forecast of Thai Small and Medium Enterprises

อภิศักดิ์ สุโน (Apisak Suno)¹ และวัชรวิธรรม จิตต์สกุล (Watchareewan Jitsakul)²
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021857174@kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบอัลกอริทึมสำหรับการพยากรณ์มูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมของไทย โดยใช้ข้อมูลย้อนหลังจากสำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม (สสว.) ที่มีลักษณะเป็นรายเดือน ตั้งแต่เดือนมกราคม พ.ศ. 2556 ถึงเดือนธันวาคม พ.ศ. 2564 รวม 9 ปี จำนวนทั้งสิ้น 108 เดือน ทำการเปรียบเทียบประสิทธิภาพในการพยากรณ์จำนวน 3 อัลกอริทึม ได้แก่ วิธีอาร์มา วิธีโครงข่ายประสาทเทียม และวิธีผสมผสาน วัดประสิทธิภาพโดยใช้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ผลการวิจัยพบว่า วิธีอาร์มาให้ประสิทธิภาพในการพยากรณ์แม่นยำที่สุด โดยให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำที่สุดเท่ากับร้อยละ 4.4 รองลงมา คือ วิธีโครงข่ายประสาทเทียมร้อยละ 8.6 และวิธีผสมผสานร้อยละ 9.6 ตามลำดับ

คำสำคัญ: การพยากรณ์ อาร์มา โครงข่ายประสาทเทียม
แบบจำลองผสมผสาน

Abstract

The objective of this research is to compare different algorithms for forecasting the export value of Thai small and medium enterprises. Using monthly historical data from the Office of Small and Medium Enterprises Promotion (OSMEP) from January 2013 to December 2021, including 9 years by 108 records.

Then, comparing the forecasting efficiency of the model with 3 algorithms: Autoregressive Integrated Moving Average (ARIMA) method, the Artificial neural network (ANN) method, and the hybrid method (ARIMA-ANN) by using percent mean absolute error (MAPE). The results showed that the ARIMA method was the most accurate in forecasting with the lowest MAPE at 4.4%, followed by the neural network method at 8.6% and the hybrid method at 9.6%, respectively.

Keywords: Forecasting, ARIMA, Artificial Neural Network, Hybrid Model

1. บทนำ

การส่งออกสินค้าของประเทศไทยนั้น นับได้ว่ามีส่วนสำคัญในการช่วยขับเคลื่อนเศรษฐกิจไทยให้เจริญเติบโตไปข้างหน้าเนื่องจากสามารถสร้างรายได้ให้กับประเทศมหาศาล รวมถึงยังมีส่วนสำคัญที่ทำให้โครงสร้างการผลิตของประเทศมีการพัฒนาอย่างรวดเร็ว ซึ่งภาคการผลิตที่สำคัญของประเทศไทยนั้น ส่วนหนึ่งคือ บริษัทต่าง ๆ ที่เป็นวิสาหกิจขนาดกลางและขนาดย่อม (SME) ซึ่งประเทศไทยได้มีการส่งออกสินค้าจำนวนมากหลากหลายประเภทไปยังประเทศต่าง ๆ ทั่วโลก สร้างรายได้เข้าประเทศปีละหลายล้านล้านบาท แต่เนื่องจากปัญหาการแพร่ระบาดของโรคไวรัสโคโรนา 2019 (COVID-19) ได้ส่งผลกระทบต่อ

การส่งออกสินค้าของประเทศไทย จะเห็นได้ว่าในช่วงปี พ.ศ.2561 ถึงปี พ.ศ.2562 ซึ่งเป็นช่วงก่อนเกิดการระบาดของโรคไวรัสโคโรนา 2019 ประเทศไทยมีมูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมเป็นมูลค่ากว่า 8.108 ล้านล้านบาท และ 7.628 ล้านล้านบาท ตามลำดับ และในช่วงปี 2563-2564 ซึ่งเป็นปีที่เกิดการแพร่ระบาดของโรคไวรัสโคโรนา 2019 ทั่วโลกทำให้ยอดมูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมของประเทศไทยลดลงเหลือเพียง 7.183 ล้านล้านบาท และ 6.202 ล้านล้านบาท ตามลำดับ แม้ว่าปัจจุบันสถานการณ์การแพร่ระบาดของโรคไวรัสโคโรนา 2019 จะค่อย ๆ ดีขึ้น เศรษฐกิจโลกค่อย ๆ ฟื้นตัว แต่แนวโน้มการส่งออกของไทยยังไม่สามารถคาดการณ์มูลค่าการส่งออกสินค้าได้ การพยากรณ์ก็เป็นอีกวิธีการหนึ่งที่จะเข้ามาช่วยในการตัดสินใจและวางแผนการผลิตเพื่อการส่งออกที่จะเกิดขึ้นในอนาคตได้

ซึ่งจากการศึกษางานวิจัยต่าง ๆ ที่ผ่านมามีผู้ที่ทำการพยากรณ์โดยใช้วิธีการวิเคราะห์หรืออัลกอริทึมต่าง ๆ เพื่อใช้ในการพยากรณ์เหตุการณ์ล่วงหน้าเป็นจำนวนมาก เช่น การเปรียบเทียบวิธีการพยากรณ์จำนวนบัญชีผู้ใช้บริการธนาคารทางอินเทอร์เน็ตด้วย วิธีบ็อกซ์-เจนกินส์ และวิธีของโฮลต์ [1] การพยากรณ์ความเร็วลมเพื่อใช้ในการคาดการณ์ปริมาณการผลิตไฟฟ้าจากพลังงานลม และลดค่าใช้จ่ายในการบริหารจัดการผลิตไฟฟ้าจากพลังงานลม โดยใช้วิธีการพยากรณ์ 3 วิธี ได้แก่ วิธีโครงข่ายประสาทเทียม วิธีอาร์มีมา และวิธีแบบผสมผสาน [2] การพยากรณ์ข้อมูลการไหลเวียนของสกุลเงินของธนาคารอินโดนีเซียโดยเปรียบเทียบระหว่างวิธีผสมผสาน ARIMAX-ANN กับวิธี ARIMAX แบบดั้งเดิม [3] เป็นต้น

จากข้อมูลข้างต้น ผู้วิจัยจึงมีแนวคิดในการทำวิจัยเพื่อหาตัวต้นแบบสำหรับการพยากรณ์ยอดมูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมของประเทศไทย โดยใช้เทคนิควิธีการพยากรณ์ทั้งหมด 3 เทคนิค ได้แก่ วิธีอาร์มีมา (Autoregressive Integrated Moving Average: ARIMA) วิธีโครงข่ายประสาทเทียม

(Artificial Neural Network: ANN) และวิธีผสมผสาน (ARIMA-ANN) เพื่อเปรียบเทียบประสิทธิภาพในการพยากรณ์ของแต่ละเทคนิค และหาเทคนิคที่มีความเหมาะสมและแม่นยำมากที่สุด จากนั้นจึงนำไปใช้สร้างแบบจำลองในการพยากรณ์มูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมของไทยในอนาคต เพื่อช่วยให้วิสาหกิจขนาดกลางและขนาดย่อมของไทยสามารถนำข้อมูลไปใช้ในการตัดสินใจในการวางแผนการผลิตสินค้า ลดต้นทุนที่อาจเกิดจากการผลิตสินค้าเกินความต้องการของตลาด และพร้อมรับมือกับวิกฤติต่าง ๆ ที่จะเกิดขึ้นในอนาคต เพื่อให้ธุรกิจมีความเติบโต เพิ่มโอกาสในการทำกำไร และลดโอกาสการสูญเสียรายได้ให้ได้อย่างมากที่สุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งออกเป็น 2 ส่วน ดังนี้

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 วิธีอาร์มีมา (Autoregressive Integrated Moving Average: ARIMA)

วิธีอาร์มีมา เป็นหนึ่งในวิธีที่เป็นที่รู้จักและใช้กันอย่างแพร่หลายในการพยากรณ์ข้อมูลที่มีลักษณะเป็นอนุกรมเวลา [4]- [7] โดยเป็นวิธีพยากรณ์ที่พิจารณาข้อมูลที่อยู่ ณ ลักษณะของอนุกรมเวลา เพื่อหาความสัมพันธ์ของข้อมูลในแต่ละช่วงเวลา [8] จากนั้นจึงนำมาวิเคราะห์เพื่อพยากรณ์ข้อมูลในอนาคต โดยตัวแบบจะเขียนอยู่ในรูปแบบ ARIMA (p, d, q) แต่หากข้อมูลที่พิจารณาามีลักษณะที่มีแนวโน้มเปลี่ยนแปลงเป็นฤดูกาล ตัวแบบของการพยากรณ์จะอยู่ในรูปของ Seasonal Autoregressive Integrated Moving Average (SARIMA) ซึ่งสามารถเขียนให้อยู่ในรูปแบบ SARIMA(p, d, q) (P, D, Q) แสดงดังสมการที่ 1

$$\phi_p(B)\phi_p(B^s)(1-B)^d(1-B^s)^pY_t = \theta_q(B)\theta_q(B^s)\varepsilon_t \quad (1)$$

โดยที่ $\phi_p(B)$ คือ สหสัมพันธ์ในตัวเองลำดับที่ p กรณีไม่มีฤดูกาล

$\phi_p(B^s)$ คือ สหสัมพันธ์ในตัวเองลำดับที่ P กรณี
มีฤดูกาล

$\theta_q(B)$ คือ ค่าเฉลี่ยเคลื่อนที่อันดับที่ q กรณีไม่มี
ฤดูกาล

$\theta_q(B^s)$ คือ ค่าเฉลี่ยเคลื่อนที่อันดับที่ Q กรณีมี
ฤดูกาล

Y_t คือ ข้อมูลอนุกรมเวลา ณ เวลา t

ε_t คือ ค่าความคลาดเคลื่อน ณ เวลา t

d คือ จำนวนครั้งของการหาผลต่าง

D คือ จำนวนครั้งของการหาผลต่างฤดูกาล

โดยวิธี ARIMA จะเหมาะกับข้อมูลที่มีลักษณะเป็น
เชิงเส้น และมีความแม่นยำสูงต่อการพยากรณ์ในระยะสั้น
แต่จะไม่เหมาะกับข้อมูลที่มีการเปลี่ยนแปลงแนวโน้ม
อย่างกระทันหัน เนื่องจากจะทำให้ค่าความคลาดเคลื่อนใน
การพยากรณ์สูง

2.1.2 วิธีโครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

วิธีโครงข่ายประสาทเทียม เป็นการศึกษากระบวนการ
ทำงานของเซลล์ประสาทภายในสมองที่จะวิเคราะห์ และ
ประมวลผลข้อมูลโดยส่งข้อมูลผ่านระบบโครงข่ายของ
เซลล์ประสาท และสามารถทำงานหลายอย่างได้ในเวลา
เดียวกันเพื่อให้ได้มาซึ่งผลลัพธ์ที่ต้องการ [9]
วิธีโครงข่ายประสาทเทียมนั้นเป็นวิธีการสร้างแบบจำลอง
แบบไม่เป็นเชิงเส้น [10] และจำเป็นต้องมีความรู้ก่อนหน้า
เกี่ยวกับความสัมพันธ์ระหว่างข้อมูลขาเข้าและขาออก ซึ่ง
วิธีโครงข่ายประสาทเทียมจะให้ผลลัพธ์ที่ดีสำหรับการ
การจัดจำรูปแบบ และการพยากรณ์ปัญหาต่าง ๆ [2]

โดยวิธีโครงข่ายประสาทเทียมที่ใช้งานกันแพร่หลาย
และประสบความสำเร็จมากที่สุด คือ โครงข่ายประสาท
เทียมแบบป้อนไปข้างหน้า (Feed-Forward Neural
Network) [3] ซึ่งโครงข่ายประสาทเทียมจะประกอบด้วย
เซลล์ประสาทแต่ละชั้น ได้แก่ ชั้นข้อมูลขาเข้า (Input
Layer) ชั้นซ่อน (Hidden Layer) และชั้นข้อมูลขาออก
(Output Layer) โดยข้อมูลที่ประมวลผลในโครงข่ายจะ
ถูกส่งไปในทิศทางเดียวจากชั้นข้อมูลขาเข้าไปยังชั้นซ่อน
จนถึงชั้นข้อมูลเอาต์พุต โดยไม่มีการย้อนกลับของข้อมูล

และแต่ละจุดต่อ (Nodes) ในชั้นเดียวกัน จะไม่มี
การเชื่อมต่อกัน

ข้อดีของวิธีโครงข่ายประสาทเทียม คือ มีความแม่นยำ
ในการพยากรณ์ค่อนข้างสูงแม้ข้อมูลจะมีการเปลี่ยนแปลง
แนวโน้มอย่างกระทันหัน แต่ข้อเสีย คือ ใช้ระยะเวลาใน
การประมวลผลค่อนข้างนาน

2.1.3 วิธีวัดค่าความแม่นยำของการพยากรณ์

งานวิจัยนี้ได้คัดเลือกความเหมาะสมในการสร้าง
แบบจำลองการพยากรณ์มูลค่าการส่งออกของวิสาหกิจ
ขนาดกลางและขนาดย่อมของไทย โดยเปรียบเทียบ
ความแม่นยำของวิธีพยากรณ์ด้วยค่าร้อยละความ
คลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ซึ่งสามารถแสดงได้
ดังสมการที่ 2

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t} \times 100 \quad (2)$$

โดยที่ y_t คือ ค่าข้อมูลจริง ณ เวลา t

\hat{y}_t คือ ค่าข้อมูลที่พยากรณ์ได้ ณ เวลา t

n คือ จำนวนข้อมูลทั้งหมด

เกณฑ์ในการเลือกตัวแบบจำลองจะพิจารณาจากค่า
MAPE ของการพยากรณ์แต่ละวิธีที่มีค่าน้อยที่สุด

2.2 งานวิจัยที่เกี่ยวข้อง

กัญญา สมฤดี และฉาไลสา [1] ได้ศึกษาเปรียบเทียบ
วิธีการพยากรณ์จำนวนบัญชีผู้ใช้บริการธนาคารทาง
อินเทอร์เน็ตด้วย วิธีบ็อกซ์-เจนกินส์ และวิธีการปรับ
ให้เรียบแบบโฮลต์ ใช้ข้อมูลจากเว็บไซต์ของธนาคาร
แห่งประเทศไทย ตั้งแต่มีนาคม พ.ศ.2553 ถึงเมษายน
พ.ศ.2561 โดยแบ่งข้อมูลชุดที่ 1 ตั้งแต่มีนาคม พ.ศ.2553
ถึงเมษายน พ.ศ.2560 เพื่อนำไปสร้างตัวแบบการพยากรณ์
และชุดที่ 2 ตั้งแต่พฤษภาคม พ.ศ.2560 ถึงเมษายน พ.ศ.
2561 ใช้เพื่อเปรียบเทียบความแม่นยำในการพยากรณ์ โดย
วัดประสิทธิภาพของการพยากรณ์จากค่า MAPE ผล
การศึกษาพบว่า วิธีบ็อกซ์-เจนกินส์ เหมาะสมมากที่สุด
เนื่องจากมีค่า MAPE เท่ากับ 0.81

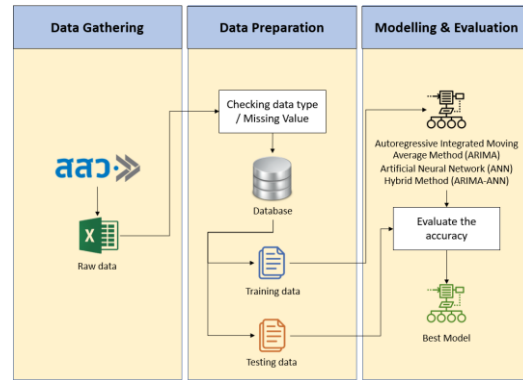
Nair, Vanitha and Jisma [2] ได้ศึกษาการพยากรณ์
ความเร็วลมเพื่อใช้ในการคาดการณ์ปริมาณการผลิตไฟฟ้า

จากพลังงานลม และลดค่าใช้จ่ายในการบริหารจัดการผลิตไฟฟ้าจากพลังงานลม โดยใช้วิธีการพยากรณ์ 3 วิธี ได้แก่ วิธี ANN วิธี ARIMA และวิธีแบบผสมผสาน (ARIMA-ANN) ใช้ข้อมูลความเร็วลมรายชั่วโมงจาก 3 สถานีในรัฐ Tamilnadu ของประเทศอินเดีย ได้แก่ Dharapuram ตั้งแต่ มกราคม ปี 2008 ถึง ธันวาคม ปี 2010 kayathar ตั้งแต่กันยายน ปี 2013 ถึงสิงหาคม ปี 2016 และจาก Nallur ตั้งแต่ ธันวาคม ปี 2013 ถึงมีนาคม ปี 2016 โดยแบ่งข้อมูล 2 ปีแรกของแต่ละชุดข้อมูลสำหรับสร้างแบบจำลอง และข้อมูลปีที่ 3 สำหรับทดสอบประสิทธิภาพของการพยากรณ์ ซึ่งวัดจากค่า MAPE ค่า MSE และค่า MAE ผลการวิจัยพบว่า วิธี ARIMA-ANN มีประสิทธิภาพดีที่สุด ให้ค่า MAE ค่า MSE และค่า MAPE ต่ำสุดที่ร้อยละ 0.32, 0.14 และ 12.4 ตามลำดับ

Prayoga, Suhartono and Rahayu [3] ได้ศึกษาการพยากรณ์ข้อมูลการไหลเวียนของสกุลเงินของธนาคารอินโดนีเซีย โดยเปรียบเทียบระหว่างวิธีผสมผสาน ARIMAX-ANN กับวิธี ARIMAX แบบดั้งเดิม โดยใช้ข้อมูลการหมุนเวียนของสกุลเงินในอินโดนีเซียจากธนาคารอินโดนีเซียตั้งแต่ ปี 2003 ถึงปี 2014 จากสำนักงานสถิติแห่งชาติของมาเลเซีย แบ่งข้อมูลเป็น 2 ชุด ได้แก่ ชุดที่ 1 ตั้งแต่ปี 2003 ถึง ปี 2013 ใช้สำหรับสร้างตัวแบบการพยากรณ์ และชุดที่ 2 คือ ข้อมูลของปี 2014 สำหรับวัดค่าความแม่นยำ ซึ่งวัดจากค่า RMSE ผลการวิจัยพบว่า วิธีผสมผสาน ARIMAX-ANN มีความแม่นยำมากที่สุด โดยให้ค่า RMSE เท่ากับ 12182.7

3. วิธีการดำเนินการวิจัย

ในการดำเนินการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดขั้นตอนในการดำเนินการไว้เป็นขั้นตอน แสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากภาพที่ 1 เป็นกรอบแนวความคิดในการทำวิจัยในครั้งนี้ โดยจะแบ่งการทำงานออกเป็น 3 ส่วน ได้แก่ การรวบรวมข้อมูล (Data Gathering) การจัดเตรียมข้อมูล (Data Preparation) การสร้างแบบจำลอง และเปรียบเทียบประสิทธิภาพการพยากรณ์แต่ละเทคนิค (Modelling & Evaluation)

3.1 การรวบรวมข้อมูล

ทำการรวบรวมข้อมูลจากเว็บไซต์ของสำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม (สสว.) โดยจะได้เป็นข้อมูลในรูปแบบไฟล์ Excel และมีลักษณะของข้อมูลเป็นรายเดือน ตั้งแต่เดือนมกราคม พ.ศ. 2556 ถึงเดือนธันวาคม พ.ศ. 2564

3.2 การจัดเตรียมข้อมูล

3.2.1 การตรวจสอบความถูกต้องครบถ้วนของข้อมูล

ทำการตรวจสอบความถูกต้องครบถ้วนของข้อมูลโดยการแทนที่ข้อมูลที่มีความสูญหายด้วยค่าเฉลี่ยของข้อมูลทั้งหมด จากนั้นนำข้อมูลทั้งหมดของแต่ละเดือนมารวมกันจนได้เป็นข้อมูลมูลค่าการส่งออกรวมของเดือนนั้น ๆ จำนวนทั้งสิ้น 108 เดือน

3.2.2 การแบ่งชุดข้อมูล

ทำการแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลสำหรับการเรียนรู้ (Training Data) จำนวน 96 เดือน ตั้งแต่เดือนมกราคม พ.ศ.2556 ถึงเดือนธันวาคม พ.ศ.2563 และชุดข้อมูลสำหรับการทดสอบ (Testing Data) จำนวน 12 เดือน ตั้งแต่เดือนมกราคม ถึงเดือนธันวาคม พ.ศ.2564

3.3 การสร้างแบบจำลอง และเปรียบเทียบประสิทธิภาพการพยากรณ์แต่ละเทคนิค

นำชุดข้อมูลสำหรับการเรียนรู้ (Training Data) มาทำการสร้างแบบจำลองด้วยเทคนิควิธีการพยากรณ์ ข้อมูลทั้งหมด 3 เทคนิค ได้แก่ วิธีออริมา (Autoregressive Integrated Moving Average: ARIMA) และวิธีโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) และวิธีผสมผสาน (ARIMA-ANN) จากนั้นทดสอบแบบจำลองด้วยชุดข้อมูลสำหรับการทดสอบ (Testing Data) เพื่อวัดประสิทธิภาพของแบบจำลองของแต่ละเทคนิค โดยพิจารณาจากค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ที่ต่ำที่สุด ซึ่งค่าพารามิเตอร์ของแต่ละอัลกอริทึมที่ใช้ทำการดำเนินการวิจัยมีรายละเอียดแสดงดังตารางที่ 1

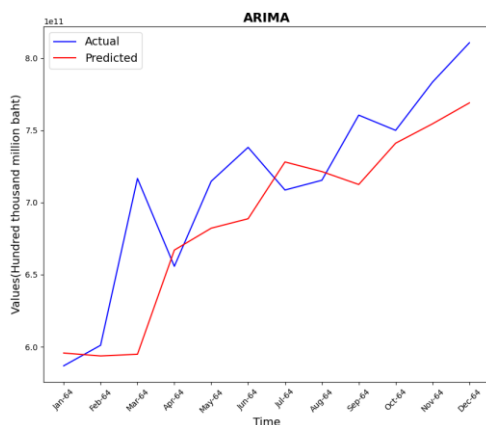
ตารางที่ 1: ค่าพารามิเตอร์ของแต่ละอัลกอริทึม

Algorithms	Parameters
ARIMA	p = 0-9, d = 0-1, q = 0-9
ANN	input = 3-96 dense = 3-96 epoch = 100-1000 batch size = 6-32
ARIMA-ANN	นำค่าพารามิเตอร์ที่ดีที่สุดของแบบจำลอง ARIMA และ ANN มาใช้

4. ผลการดำเนินงาน

4.1 ผลการพยากรณ์ด้วยแบบจำลอง ARIMA

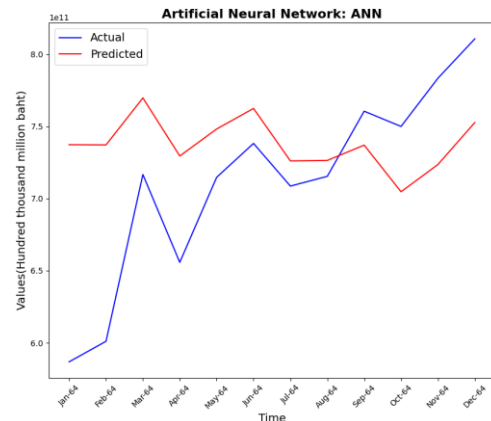
จากการทดลองด้วยอัลกอริทึม ARIMA จำนวน 200 รูปแบบ พบว่าแบบจำลองจะมีประสิทธิภาพสูงสุดเมื่อกำหนดให้พารามิเตอร์ p = 1, q = 1, d = 0 สามารถเขียนสัญลักษณ์ได้ว่า ARIMA(1, 1, 0) มีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ร้อยละ 4.4 ผลการพยากรณ์แสดงดังภาพที่ 2



ภาพที่ 2: ผลการพยากรณ์ด้วยแบบจำลอง ARIMA

4.2 ผลการพยากรณ์ด้วยแบบจำลอง ANN

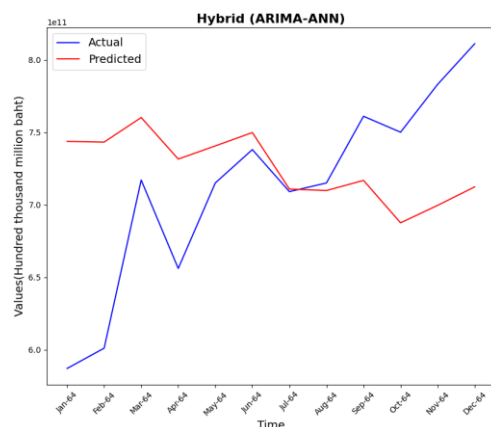
จากการทดลองด้วยอัลกอริทึม ANN จำนวน 1452 รูปแบบ พบว่าแบบจำลองจะมีประสิทธิภาพสูงสุดเมื่อกำหนดให้พารามิเตอร์ input = 12, Dense = 60, epoch = 1000, batch size = 12 มีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ร้อยละ 8.6 ผลการพยากรณ์แสดงดังภาพที่ 3



ภาพที่ 3: ผลการพยากรณ์ด้วยแบบจำลอง ANN

4.3 ผลการพยากรณ์ด้วยแบบจำลองแบบผสมผสาน (ARIMA-ANN)

จากการทดลองด้วยอัลกอริทึมแบบผสมผสานด้วยการนำค่าที่ดีที่สุดของทั้ง 2 แบบจำลองมาทดสอบ พบว่ามีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ร้อยละ 9.3 ผลการพยากรณ์แสดงดังภาพที่ 4



ภาพที่ 4: ผลการพยากรณ์ด้วยแบบจำลองแบบผสมผสาน

4.4 ผลการเปรียบเทียบประสิทธิภาพการพยากรณ์และคัดเลือกอัลกอริทึมที่เหมาะสม

จากการทดลองพบว่าอัลกอริทึม ARIMA มีประสิทธิภาพในการพยากรณ์แม่นยำที่สุด โดยมีค่าร้อยละความคลาดเคลื่อน

สัมบูรณ์เฉลี่ย (MAPE) ที่ร้อยละ 4.4

ตารางที่ 2: ผลการเปรียบเทียบอัลกอริทึมที่เหมาะสม

Algorithms	MAPE (%)
ARIMA	4.4
ANN	8.6
ARIMA-ANN	9.3

5. สรุปผล และอภิปรายผล

งานวิจัยนี้ ผู้วิจัยได้ทำการเปรียบเทียบอัลกอริทึมสำหรับการพยากรณ์มูลค่าการส่งออกของวิสาหกิจขนาดกลางและขนาดย่อมของไทย โดยใช้แบบจำลอง 3 อัลกอริทึม ได้แก่ ARIMA, ANN และ ARIMA-ANN โดยแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลสำหรับการเรียนรู้ (Training Data) จำนวน 96 เดือน ตั้งแต่เดือนมกราคม พ.ศ. 2556 ถึงเดือนธันวาคม พ.ศ. 2563 และชุดข้อมูลสำหรับการทดสอบ (Testing Data) จำนวน 12 เดือน ตั้งแต่เดือนมกราคม ถึงเดือนธันวาคม พ.ศ. 2564 ผลการวิจัยพบว่า อัลกอริทึม ARIMA มีประสิทธิภาพในการพยากรณ์แม่นยำที่สุด โดยให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ที่ร้อยละ 4.4 มีความเหมาะสมในการนำไปพัฒนาระบบเพื่อใช้ในการพยากรณ์ต่อไป

อย่างไรก็ตาม งานวิจัยนี้ได้ทำการศึกษาเปรียบเทียบอัลกอริทึมสำหรับการพยากรณ์เพียง 3 อัลกอริทึมเท่านั้น และพยากรณ์โดยใช้ข้อมูลย้อนหลังจากเว็บไซต์ของสำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม (สสว.) ตั้งแต่ปี พ.ศ. 2556 ถึงปี พ.ศ. 2564 เท่านั้น ดังนั้น หากมีการศึกษาอัลกอริทึมอื่นที่มีความเหมาะสม และมีจำนวนข้อมูลย้อนหลังมากขึ้น รวมทั้งการนำปัจจัยอื่น ๆ มีพิจารณาร่วมด้วย เช่น ปัจจัยทางด้านเศรษฐกิจ อัตราเงินเฟ้อ เป็นต้น อาจทำให้ผลการพยากรณ์มีความแม่นยำและใกล้เคียงกับข้อมูลจริงมากขึ้น

เอกสารอ้างอิง

- [1] กัญญา บวรโชคชัย, สมฤดี พงษ์เสนา และศาลีศา เนียมมณี, “การเปรียบเทียบวิธีการพยากรณ์จำนวนบัญชีผู้ใช้บริการธนาคารทางอินเทอร์เน็ตด้วย วิธีของบ็อกซ์-เจนกินส์ และ

วิธีของโฮลด์,” วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม ปีที่ 38 ฉบับที่ 4 กรกฎาคม – สิงหาคม พ.ศ.2562 หน้า 438-443.

- [2] Nair, K. R., Vanitha, V. and Jisma, M., “Forecasting of Wind Speed using ANN, ARIMA and Hybrid Models,” International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), pp. 170-175, 2017.
- [3] Prayoga, I. G. S. A., Suhartono and Rahayu, S. P., “Forecasting Currency Circulation Data of Bank Indonesia by Using Hybrid ARIMAX-ANN Model,” AIP Conference Proceedings, Vol. 1842, pp. 1-8, 2017.
- [4] Papastefanopoulos, V., Linardatos, P. and Kotsiantis, S., “COVID-19: A Comparison of Time Series Methods to Forecast Percentage of Active Cases per Population,” Applied Sciences Journal, Vol.10 No. 3880, pp. 1-15, 2020.
- [5] Lekshmi, M. and Subramanya, K. N. A., “Short-Term Load Forecasting of 400kV Grid Substation Using R-Tool and Study of Influence of Ambient Temperature on the Forecasted Load.” Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). pp. 1-5, 2019.
- [6] Karthika, S., Margaret, V. and Balaraman, K., “Hybrid Short Term Load Forecasting using ARIMA-SVM,” International Conference on Innovations in Power and Advanced Computing Technologies pp. 1-7, 2017.
- [7] Yonar, H., Yonar, A. and Tekindal, M. A., “Modeling and Forecasting for the Number of Cases of the COVID-19 Pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods,” Eurasian Journal of Medicine and Oncology, Vol.4 No.2, pp. 1-9, 2020.
- [8] อรุชา จันทรภา และนันทชัย กานดำนันตะ, “การพยากรณ์มูลค่าการส่งออกรถยนต์อุปกรณ์และส่วนประกอบของประเทศไทย ด้วยวิธีอนุกรมเวลา,” วารสารไทยการวิจัยดำเนินงาน ปีที่ 9 ฉบับที่ 2 กรกฎาคม – ธันวาคม พ.ศ.2564 หน้า 12-20.
- [9] วิรศักดิ์ ฟองเงิน วรปภา อรัราชญ์ และเผด็จ พรหมสาขา ณ สกลนคร, “การพยากรณ์ปริมาณน้ำในเขื่อน โดยใช้เทคนิคเหมืองข้อมูล,” วารสารวิชาการการจัดการเทคโนโลยี

- สารสนเทศและนวัตกรรม มหาวิทยาลัยราชภัฏมหาสารคาม
ปีที่ 4 ฉบับที่ 1 มกราคม – มิถุนายน พ.ศ.2560 หน้า 27-33.
- [10] Aquino, R. L., Alcantara, N. L. M. T. and Addawe, R. C., “A Hybrid ARIMA and Neural Network Model Applied to Forecast Catch Volumes of Selar Crumenophthalmus,” AIP Conference Proceedings, Vol. 1905, pp. 1-6, 2017.

การตรวจสอบพฤติกรรมการเบิกสิทธิประโยชน์กรณีทันตกรรม
ของสำนักงานประกันสังคมด้วยเทคนิคเหมืองข้อมูล
**Examination of The Social Security Office's Dental Benefits
Disbursement Behavior by Data Mining Techniques**

ชวรัส เกรอต (Chawarus Kerot)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857051@kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้นำเสนอการสร้างโมเดลพยากรณ์รายการการเบิกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคมที่ผิดปกติ โดยเป็นการเปรียบเทียบอัลกอริทึมจำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึมป่าสุ่ม อัลกอริทึมเพื่อนบ้านใกล้ที่สุด และอัลกอริทึมโครงข่ายประสาทเทียม โดยข้อมูลที่ใช้ในการทดสอบเป็นข้อมูลรายการการเบิกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคม ตั้งแต่วันที่ 1 สิงหาคม 2560 ถึง 31 ธันวาคม 2565 จำนวน 207,240 รายการ ประกอบด้วยคุณลักษณะ จำนวน 22 คุณลักษณะ นำคุณลักษณะทั้งหมดมาคัดเลือกคุณลักษณะที่จะนำมาใช้งาน โดยได้ทั้งสิ้น 11 คุณลักษณะ ทำการคัดเลือกข้อมูลโดยเลือกเฉพาะข้อมูลที่ครบถ้วนให้เหลือเพียง 204,620 รายการ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลองจำนวน 21,006 รายการ และข้อมูลสำหรับทดสอบจำนวน 183,614 รายการ ทำการเปรียบเทียบประสิทธิภาพด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล พบว่าอัลกอริทึมป่าสุ่มให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 97.29 ค่าความแม่นยำเท่ากับ 0.88 ค่าระลอกเท่ากับ 0.90 และค่าการถ่วงดุลเท่ากับ 0.89

คำสำคัญ: อัลกอริทึมป่าสุ่ม

อัลกอริทึมเพื่อนบ้านใกล้ที่สุด

อัลกอริทึมโครงข่ายประสาทเทียม

สิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคม

Abstract

The objective of this research is the development of predictive models to detect abnormalities in the Social Security Office's list of dental benefits reimbursements. The three algorithms are compared composed of Random Forest (RF), K-Nearest Neighbor (KNN), and Neural Network (NN). The data during 1 August 2017 to 31 December 2022 from the list of bringing the social security office's dental benefits used database. All of data is 207,240 records and 22 attributes. In addition, all features are selected to be applied only 11 attributes and just 204,620 records of complete data were chosen for data selection. The dataset is split into 2 parts: training data of 21,006 records and testing data of 183,614 records, comparing performance with accuracy, precision, recall, and f-measure. The Random Forest algorithms showed the highest accuracy, the results accuracy 97.29%, precision 0.88, recall 0.90, and f-measure 0.89.

Keywords: Random Forest, K-Nearest Neighbor, Neural Network, Social Security Office's Dental Benefits.

1. บทนำ

สำนักงานประกันสังคมเป็นหน่วยงานที่มีภารกิจในการบริหารการประกันสังคม มีภารกิจหลักในการคุ้มครองดูแลผู้ประกันตนและลูกจ้างให้ได้รับสิทธิประโยชน์ตามกฎหมายกำหนด ซึ่งการผลักดันการใช้บริการผ่านระบบอิเล็กทรอนิกส์ ของสำนักงานประกันสังคม สอดคล้องกับแผนพัฒนาารัฐบาลดิจิทัลของประเทศไทย ในปัจจุบันการขอรับสิทธิประโยชน์กรณีทันตกรรมของผู้ประกันตน ได้มีการปรับปรุงแบบการขอรับสิทธิประโยชน์กรณีทันตกรรมของผู้ประกันตน โดยไม่ต้องสำรองจ่าย ตั้งแต่ปี 2560 [1] ครอบคลุมกรณีถอนฟัน อุดฟัน ขูดหินปูน และผ่าตัดฟันคุด มีสิทธิได้รับค่าบริการทางการแพทย์เท่าที่จ่ายจริงไม่เกิน 900 บาท/คน/ปี เมื่อเข้ารับบริการผ่านสถานพยาบาลในความตกลงกับสำนักงานประกันสังคม จากข้อมูลการเบิกจ่ายเงินการใช้สิทธิประโยชน์กรณีทันตกรรม ปี 2560 – 2564 พบว่าปริมาณการเบิกจ่ายเงินโดยสถานพยาบาลในความตกลงที่เพิ่มขึ้นอย่างต่อเนื่อง โดยสำนักงานประกันสังคมได้มีการสุ่มตรวจสอบข้อมูลการอนุมัติส่งจ่ายกรณีทันตกรรม พบว่าเกิดพฤติกรรมการเบิกสิทธิประโยชน์กรณีทันตกรรมที่ไม่สมเหตุสมผล เช่น ปริมาณการเบิกจ่ายเงินที่สูงผิดปกติ เป็นต้น โดยวิธีการสุ่มตรวจสอบข้อมูลการเบิกจ่ายในปัจจุบันทำได้ยาก และใช้เวลาดำเนินการจำนวนมาก ส่งผลต่อการบริหารจัดการเพื่อให้ผู้ประกันตนได้รับสิทธิประโยชน์สูงสุด

ผู้วิจัยนำเสนอการวิเคราะห์ข้อมูลด้วยการนำเทคนิคการทำเหมืองข้อมูล (Data Mining) มาใช้ในการตรวจสอบวิเคราะห์ข้อมูล เช่น การใช้อัลกอริทึมป่าสุ่มในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล โดยใช้ Rapid Miner [2] การใช้อัลกอริทึมเพื่อนบ้านใกล้ที่สุดในการเปรียบเทียบประสิทธิภาพของอัลกอริทึมและการคัดเลือกคุณลักษณะที่เหมาะสมเพื่อการพยากรณ์โอกาสความสำเร็จในการโอนเงินข้ามประเทศของบุคคลทั่วไป [3] และอัลกอริทึมโครงข่ายประสาทเทียม ในการจำแนกข้อบกพร่องของระบบบริหารจัดการลูกค้า [4] เป็นต้น

จากปัญหา และงานวิจัยที่กล่าวถึงข้างต้น ผู้วิจัยจึงมีแนวคิดในการนำอัลกอริทึมเหมืองข้อมูล เพื่อวิเคราะห์พฤติกรรมการเบิกสิทธิประโยชน์กรณีทันตกรรมที่ไม่สมเหตุสมผล เพื่อช่วยให้สำนักงานประกันสังคมเพิ่มประสิทธิภาพในการตรวจสอบการเบิกจ่าย รวมถึงเป็นข้อมูลประกอบการปรับปรุงพัฒนาระบบสารสนเทศต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎี และงานวิจัยที่เกี่ยวข้อง แบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 อัลกอริทึมการพยากรณ์

2.1.1 อัลกอริทึมป่าสุ่ม (Random Forest: RF)

การทำงานจะนำต้นไม้ตัดสินใจมารวมกันมีลักษณะคล้ายป่า ต้นไม้ตัดสินใจแต่ละต้นจะจัดกลุ่มและจำแนกข้อมูลซึ่งป่าสุ่มจะนำผลลัพธ์ทั้งหมดที่ได้จากต้นไม้ตัดสินใจแต่ละต้นมาตรวจสอบหาค่าฐานนิยม (Mode) เพื่อโหวตผลลัพธ์ที่มีจำนวนมากที่สุดให้เป็นผลลัพธ์สุดท้าย [5]

2.1.2 อัลกอริทึมเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor: KNN)

โดยจะมีการกำหนดค่า k หรือจำนวนของเพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbors) จากนั้นจะคำนวณระยะห่างระหว่าง Data Point กับ Neighbors จำนวน k ตัว และนำระยะห่างที่ได้นั้นมาจัดลำดับ เพื่อหาว่า Data Point อยู่ใกล้กับ Neighbors ใดมากที่สุด จากนั้นจึงทำการโหวตโดยยึดตามเสียงข้างมาก (Majority Vote) เพื่อทำนายว่า Data Point นั้นจัดอยู่ในกลุ่มใด [6]

2.1.3 อัลกอริทึมโครงข่ายประสาทเทียม (Neural Network: NN)

รูปแบบของโครงข่ายประสาทเทียมพื้นฐานจะประกอบด้วยเซลล์ประสาท (Neuron) โดยที่อินพุตต่าง ๆ จะถูกส่งผ่านกลุ่มใยประสาทนำเข้า (Dendrites) เข้าสู่เซลล์ประสาท เพื่อทำการประมวลผลและส่งผลลัพธ์ออกมาทางแกนประสาท (Axon) โดยการประมวลผลจะใช้ผลรวมของผลคูณ (Sum of Product) ของข้อมูลอินพุตและค่าน้ำหนัก (Weight) ที่ใยประสาทนำเข้ารูปแบบของโครงข่ายประสาทเทียมพื้นฐานจะประกอบด้วยเซลล์ประสาท (Neuron) โดยที่อินพุตต่าง ๆ จะถูกส่งผ่านกลุ่มใยประสาทนำเข้า (Dendrites) เข้าสู่

เซลล์ประสาท เพื่อทำการประมวลผลและส่งผลลัพธ์ออกมาทางแกนประสาท (Axon) โดยการประมวลผลจะใช้ผลรวมของผลคูณ (Sum of Product) ของข้อมูลอินพุตและค่าน้ำหนัก (Weight) ที่โยประสาทนำเข้า [7]

2.2 งานวิจัยที่เกี่ยวข้อง

ภูมิพัฒน์ และรัชนา [8] ได้ทำการศึกษากการสร้างแบบจำลองการทำนายแบบอากาศยานอัตโนมัติ โดยผู้วิจัยได้สร้างแบบจำลองทำนายแบบอากาศยานอัตโนมัติไม่พึ่งพามนุษย์ โดยใช้ข้อมูลจากระบบวิเคราะห์ความเคลื่อนไหวในทางอากาศประเทศรอบบ้านของกรมข่าวทหารอากาศเป็นข้อมูลในการวิจัยเพื่อค้นหาแบบจำลองที่เหมาะสม ด้วยเทคนิคการจำแนกข้อมูล (Classification) ได้แก่ Naïve Bayes, Deep Learning, Decision Tree และ Random Forest ผลการวิจัยพบว่าแบบจำลอง Random Forest มีประสิทธิภาพสูงสุด มีค่าความถูกต้องร้อยละ 90.2 ค่าความแม่นยำร้อยละ 90.47 และค่าความระลึกร้อยละ 90.24

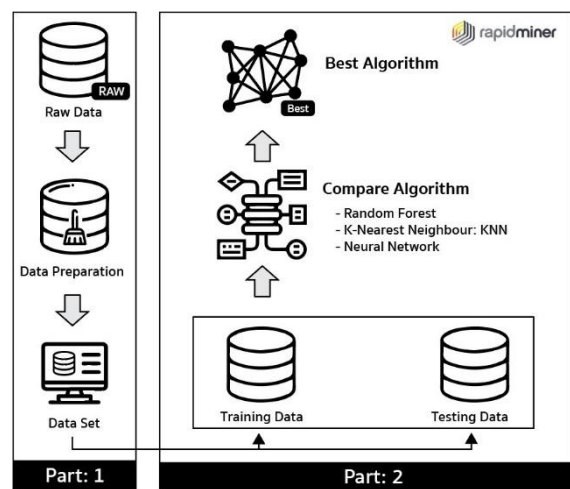
Theerthagiri and et al. [9] ได้ทำการศึกษากการทำนายความเป็นไปได้ของการรักษา COVID-19 โดยใช้อัลกอริทึมการจำแนกประเภท K- Nearest Neighbor โดยงานวิจัยนี้ได้ศึกษาอัลกอริทึมเพื่อทำนายผู้ป่วยที่หายจากโรคโควิด-19 และผู้เสียชีวิต จำแนกด้วย 5 อัลกอริทึม ได้แก่ อัลกอริทึม Logistic Regression, K-Nearest Neighbor, Decision Tree, Support Vector Machine, Multi-Layer Perceptron ใช้ชุดข้อมูลจำนวน 100,284 รายการ ผลการวิจัยพบว่า อัลกอริทึม K-Nearest Neighbor ให้ประสิทธิภาพสูงสุด ค่าความถูกต้องร้อยละ 80.36 ค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ ร้อยละ 19.63 ค่าความคลาดเคลื่อนเฉลี่ย ร้อยละ 44.31 และค่าสัมประสิทธิ์แคปปาของโคเซน ร้อยละ 41.09

นงเยาว์ [10] ได้ทำการศึกษากการเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายความเสี่ยงโรคหัวใจและหลอดเลือดโดยใช้อัลกอริทึมเหมืองข้อมูลโดยใช้อัลกอริทึมเหมืองข้อมูล ได้แก่ Neural Network, Random Forest, K-Nearest Neighbor, Naïve Bayes และ Decision Tree เปรียบเทียบประสิทธิภาพของแบบจำลองด้วยวิธี 10-Fold Cross Validation โดยใช้ชุดข้อมูลเป็นคนที่ผ่านการคัดกรองผู้ป่วยโรคหัวใจและหลอดเลือดที่

รวบรวมข้อมูลมาจากสำนักงานสาธารณสุขจังหวัดสระบุรี ระหว่างปี พ.ศ. 2561-2562 จากโรงพยาบาล ในจังหวัดสระบุรี 12 แห่ง และโรงพยาบาลส่งเสริมสุขภาพตำบล 126 แห่ง จำนวน 31,929 คน ผลการวิจัยพบว่าแบบจำลองที่มีประสิทธิภาพการทำนายที่ดีที่สุดคือ Neural Network มีค่าความถูกต้องร้อยละ 99.29

3. วิธีการดำเนินการวิจัย

ผู้วิจัยมีกรอบแนวคิดวิธีการวิจัย แสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิด

จากภาพที่ 1 กรอบแนวคิด สามารถแบ่งขั้นตอนการดำเนินการวิจัยเป็น 2 ส่วน ดังนี้

3.1 ขั้นตอนการรวบรวม และเตรียมข้อมูล

ข้อมูลที่ใช้วิจัยเป็นข้อมูลรายการการเบิกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคม ตั้งแต่วันที่ 1 สิงหาคม 2560 ถึง 31 ธันวาคม 2565 จำนวน 207,240 รายการ ประกอบด้วยคุณลักษณะ จำนวน 22 คุณลักษณะ นำคุณลักษณะทั้งหมดมาคัดเลือกคุณลักษณะที่จะนำมาใช้งาน ทำความสะอาดข้อมูล โดยได้คัดเลือกข้อมูลให้เหลือเพียง 204,620 รายการ ประกอบด้วย 11 คุณลักษณะ ได้แก่ เลขที่รับแจ้ง วันที่รับแจ้ง เลขผู้ใช้สิทธิ์ รหัส 7 หลัก รหัสสำนักงานประกันสังคม จำนวนเงินรวม รหัสการรักษา จำนวนเงินเบิก จำนวนเงินอนุมัติจ่าย สถานะแสดงจำนวนรายการต่อเดือน และสถานะแสดงจำนวนเงินอนุมัติจ่ายต่อเดือน

ตารางที่ 1: คุณลักษณะของข้อมูล

ลำดับ	คุณลักษณะ
1	PETITION_NO
2	PETITION_DATE
3	CITIZEN_ID
4	HCODE7
ลำดับ	คุณลักษณะ
5	SSO_BRANCH_CODE
6	TOTAL_AMOUNT
7	TREATMENT_CODE
8	CLAIM_AMOUNT
9	CLAIM_APPROVE_AMOUNT
10	COUNT_MONTH
11	CLAIM_APPROVE_MONTH

3.2 ขั้นตอนการคัดเลือก และทดสอบประสิทธิภาพ อัลกอริทึม

นำข้อมูลที่ได้จากขั้นตอนการรวบรวม และเตรียมข้อมูล แบ่งออกเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลอง จำนวน 21,006 รายการ และข้อมูลสำหรับทดสอบแบบจำลองจำนวน 183,614 รายการ โดยทำการเปรียบเทียบอัลกอริทึม 3 อัลกอริทึม ได้แก่ อัลกอริทึม RF อัลกอริทึม KNN และอัลกอริทึม NN ซึ่งได้มีการกำหนดพารามิเตอร์ที่สนใจในแต่ละอัลกอริทึม และทำการปรับค่าพารามิเตอร์ตามที่กำหนดไว้ จากนั้นทำการวัดประสิทธิภาพของอัลกอริทึมจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล โดยได้ทำซ้ำในแต่ละรูปแบบ เพื่อคัดเลือกหาอัลกอริทึมที่ให้ค่าความถูกต้องสูงที่สุด โดยค่าพารามิเตอร์สำหรับการทดสอบอัลกอริทึม แสดงดังตารางที่ 2

ตารางที่ 2: ค่าพารามิเตอร์สำหรับการทดสอบอัลกอริทึม

RF	number of trees = 50, 60, 70, 80, 90, 100 criterion = information_gain, gini_index voting strategy = confidence vote, majority vote
KNN	k = 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41

	measure types = mixed measure mixed measure = mixed euclidean distance
NN	training cycle = 200, 300, 400, 500 learning rate = 0.01, 0.1, 0.2 momentum = 0.9, 1

4. ผลการดำเนินงาน

ผลการทดสอบประสิทธิภาพอัลกอริทึม แสดงดังตารางที่ 3 ถึง 5 โดยมีรายละเอียดดังนี้

ตารางที่ 3: ผลการทดสอบประสิทธิภาพอัลกอริทึม RF

Parameters			Performance			
number of trees	criterion	voting strategy	Accuracy	Precision	Recall	F-Measure
50	Information_gain	confidence vote	97.13	0.86	0.90	0.88
60	Information_gain	confidence vote	97.18	0.86	0.90	0.88
70	Information_gain	confidence vote	97.17	0.86	0.90	0.88
80	Information_gain	confidence vote	97.18	0.86	0.90	0.88
90	Information_gain	confidence vote	97.19	0.86	0.90	0.88
100	Information_gain	confidence vote	97.18	0.86	0.90	0.88
50	gini_index	confidence vote	97.17	0.87	0.90	0.88
60	gini_index	confidence vote	97.20	0.87	0.90	0.88
70	gini_index	confidence vote	97.19	0.87	0.90	0.88
80	gini_index	confidence vote	97.21	0.87	0.90	0.89
90	gini_index	confidence vote	97.20	0.87	0.90	0.88
100	gini_index	confidence vote	97.19	0.87	0.90	0.88

50	Information_gain	majority vote	97.26	0.88	0.89	0.89
60	Information_gain	majority vote	97.28	0.89	0.90	0.89
<u>70</u>	<u>Information_gain</u>	<u>majority vote</u>	<u>97.29</u>	<u>0.88</u>	<u>0.90</u>	<u>0.89</u>
80	Information_gain	majority vote	97.26	0.88	0.89	0.89

จากตารางที่ 3 ค่าพารามิเตอร์ number of trees เท่ากับ 70 criterion เท่ากับ Information_gain และ voting strategy เท่ากับ majority vote ให้ค่าความถูกต้องร้อยละ 97.29 ค่าความแม่นยำเท่ากับ 0.88 ค่าระลึกเท่ากับ 0.90 และค่าการถ่วงดุลเท่ากับ 0.89 ตามลำดับ

ตารางที่ 4: ผลการทดสอบประสิทธิภาพอัลกอริทึม KNN

Parameters			Performance			
k	measure types	mixed measure	Accuracy	Precision	Recall	F-Measure
3	mixed measure	mixed euclidean distance	94.34	0.77	0.77	0.77
5	mixed measure	mixed euclidean distance	94.64	0.79	0.75	0.77
7	mixed measure	mixed euclidean distance	94.46	0.79	0.73	0.76
9	mixed measure	mixed euclidean distance	94.55	0.80	0.71	0.75
11	mixed measure	mixed euclidean distance	94.61	0.81	0.69	0.75
13	mixed measure	mixed euclidean distance	94.67	0.82	0.68	0.75
<u>15</u>	<u>mixed measure</u>	<u>mixed euclidean distance</u>	<u>94.69</u>	<u>0.83</u>	<u>0.68</u>	<u>0.74</u>
17	mixed measure	mixed euclidean distance	94.58	0.82	0.67	0.74
19	mixed measure	mixed euclidean distance	94.60	0.83	0.66	0.74
21	mixed measure	mixed euclidean distance	94.66	0.84	0.66	0.74
23	mixed measure	mixed euclidean distance	94.68	0.84	0.66	0.74
25	mixed measure	mixed euclidean distance	94.67	0.84	0.66	0.74
27	mixed measure	mixed euclidean distance	94.68	0.84	0.66	0.74

29	mixed measure	mixed euclidean distance	94.69	0.84	0.66	0.74
31	mixed measure	mixed euclidean distance	94.65	0.84	0.65	0.74
33	mixed measure	mixed euclidean distance	94.65	0.84	0.65	0.74
35	mixed measure	mixed euclidean distance	94.63	0.84	0.65	0.73

จากตารางที่ 4 ค่าพารามิเตอร์ k เท่ากับ 15 measure types เท่ากับ mixed measure และ mixed euclidean distance ให้ค่าความถูกต้องร้อยละ 94.69 ค่าความแม่นยำเท่ากับ 0.83 ค่าระลึกเท่ากับ 0.68 และค่าการถ่วงดุลเท่ากับ 0.74 ตามลำดับ

ตารางที่ 5: ผลการทดสอบประสิทธิภาพอัลกอริทึม NN

Parameters			Performance			
training cycle	learning rate	momentum	Accuracy	Precision	Recall	F-Measure
200	0.01	0.9	96.13	0.88	0.79	0.83
300	0.01	0.9	96.19	0.89	0.78	0.83
400	0.01	0.9	96.00	0.87	0.80	0.83
<u>500</u>	<u>0.01</u>	<u>0.9</u>	<u>96.27</u>	<u>0.87</u>	<u>0.82</u>	<u>0.84</u>
200	0.1	0.9	95.92	0.86	0.82	0.84
300	0.1	0.9	96.08	0.87	0.81	0.84
400	0.1	0.9	95.98	0.86	0.82	0.84
500	0.1	0.9	95.82	0.85	0.81	0.83
200	0.01	1	93.30	0.47	0.50	0.48
300	0.01	1	93.30	0.47	0.50	0.48
400	0.01	1	93.30	0.47	0.50	0.48
500	0.01	1	93.30	0.47	0.50	0.48
200	0.2	1	93.30	0.47	0.50	0.48
300	0.2	1	93.30	0.47	0.50	0.48
400	0.2	1	93.30	0.47	0.50	0.48
500	0.2	1	93.30	0.47	0.50	0.48

จากตารางที่ 5 ค่าพารามิเตอร์ training cycle เท่ากับ 500 learning rate เท่ากับ 0.01 และ momentum เท่ากับ 0.9 ให้ค่าความถูกต้องร้อยละ 96.27 ค่าความแม่นยำเท่ากับ 0.87 ค่าระลึกเท่ากับ 0.82 และค่าการถ่วงดุลเท่ากับ 0.84 ตามลำดับ

จากผลการทดสอบประสิทธิภาพอัลกอริทึมทั้ง 3 อัลกอริทึม เมื่อทำการทำซ้ำทั้งหมด 5 ครั้ง ด้วยพารามิเตอร์ ชุดเดิมแต่ละอัลกอริทึมได้ค่าเท่าเดิมทุกครั้งแสดงดังตารางที่ 6

ตารางที่ 6: ผลการเปรียบเทียบอัลกอริทึม

Algorithm	Performance			
	ACC	P	R	F
RF	97.29	0.88	0.90	0.89
KNN	94.69	0.83	0.68	0.74
NN	96.27	0.87	0.82	0.84

จากตารางที่ 6 สรุปได้ว่า อัลกอริทึม Random Forest ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 97.29 ค่าความแม่นยำเท่ากับ 0.88 ค่าระลอกเท่ากับ 0.90 และค่าการถ่วงดุลเท่ากับ 0.89 ตามลำดับ

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการสร้างโมเดลพยากรณ์รายการการเบกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคมที่ผิดปกติ เป็นข้อมูลรายการการเบกสิทธิประโยชน์กรณีทันตกรรมของสำนักงานประกันสังคม ตั้งแต่วันที่ 1 สิงหาคม 2560 ถึง 31 ธันวาคม 2565 จำนวน 204,620 รายการ ประกอบด้วยคุณลักษณะจำนวน 11 คุณลักษณะ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลองจำนวน 21,006 รายการ และข้อมูลสำหรับทดสอบแบบจำลองจำนวน 183,614 รายการ และทดสอบประสิทธิภาพของอัลกอริทึม จำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึม RF อัลกอริทึม KNN และอัลกอริทึม NN พบว่า อัลกอริทึม RF ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 97.29 ค่าความแม่นยำเท่ากับ 0.88 ค่าระลอกเท่ากับ 0.90 และค่าการถ่วงดุลเท่ากับ 0.89

เอกสารอ้างอิง

[1] สำนักงานประกันสังคม แจ้งผู้ประกันตนสามารถเข้ารับบริการทันตกรรม แบบไม่ต้องสำรองจ่าย. (2560). [ออนไลน์]. [สืบค้นวันที่ 1 สิงหาคม 2565]. จาก <https://thainews.prd.go.th/news/detail/TNSOC6011100010087>

[2] จิราภรณ์ เจริญยิ่ง, การพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยเทคนิคเหมืองข้อมูล โดยใช้ Rapid Miner. สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2563.

[3] อนัตต์ชัย ชูติภาสเจริญ, “การเปรียบเทียบประสิทธิภาพของอัลกอริทึมและการคัดเลือกคุณลักษณะที่เหมาะสมเพื่อการพยากรณ์โอกาสความสำเร็จในการโอนเงินข้ามประเทศของบุคคลทั่วไป”. *KKU Research Journal of Humanities and Social Sciences (Graduate Studies)*. Vol. 6 NO. 3: September - December pp. 105 – 113, 2018.

[4] นวรัตน์ วรณศรี, การศึกษาการประยุกต์ใช้อัลกอริทึมเหมืองข้อความในการจำแนกข้อบกพร่องของระบบบริหารจัดการลูกค้า. การค้นคว้าอิสระวิทยาศาสตรมหาบัณฑิต สาขาบริหารสารสนเทศเพื่อการจัดการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2563.

[5] กฤษณา อินทรรัตน์, “การจำแนกการใช้ประโยชน์ที่ดินในจังหวัดนครนายก ด้วยอัลกอริทึมการเรียนรู้เครื่องและภาพถ่ายจากดาวเทียม Sentinel-2”. *วารสารวิทยาศาสตร์บูรพา*. ปีที่ 27 (ฉบับที่ 2) พฤษภาคม - สิงหาคม 2565 หน้า 1153-1171.

[6] ขวัญตา ศิลป์ไพฑูรย์พานิช, การใช้เทคโนโลยีการเรียนรู้ของเครื่อง เพื่อทำนายผลการเรียนของนักเรียน. สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2562.

[7] กริช สมกันธา และคณะ, “การพัฒนาระบบประเมินบุคลากรโดยใช้โครงข่ายประสาทเทียม”. *วารสารเทคโนโลยีสารสนเทศ Information Technology Journal*. ปีที่ 9 ฉบับที่ 1 มกราคม - มิถุนายน 2556 หน้า 58 – 66.

[8] ภูมิพัฒน์ ดวงกลาง และรจนา เครือแก้ว, “แบบจำลองการทำนายอากาศยานจากข้อมูลเป้าหมายไม่ทราบฝ่ายอัตโนมัติ”. *วารสารวิทยาศาสตร์และเทคโนโลยีนายเรืออากาศ*. ปีที่ 15 มกราคม - ธันวาคม 2563 หน้า 1-8.

[9] Theerthagiri, P., Jacob, I. J., Ruby, A. U., & Vamsidhar, Y. (2020). Prediction of COVID-19 Possibilities using KNN Classification Algorithm. available at Research Square [<https://doi.org/10.21203/rs.3.rs-70985/v2>].

[10] นงเยาว์ ในอรุณ, “การเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายความเสี่ยงโรคหัวใจและหลอดเลือดโดยใช้อัลกอริทึมเหมืองข้อมูล”. *J Sci Technol MSU*. Vol 40. No 2, March-April pp. 137-147, 2021.

การเปรียบเทียบอัลกอริทึมสำหรับการจำแนกประเภทหลักทรัพย์ ในตลาดหลักทรัพย์แห่งประเทศไทย A Comparison of Algorithms for Classifying Securities in the Stock Exchange of Thailand

ฟ้าทิพย์ หวานหอม (Fathip Wanhom)¹ และธนพล เจนสุทธิเวชกุล (Tanapon Jentsuttiwetchakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹1s6407021857115@email.kmutnb.ac.th, ²tanapon.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพอัลกอริทึมการจำแนกประเภทข้อมูลที่เหมาะสมสำหรับการพยากรณ์หลักทรัพย์ด้วยเทคนิคเหมืองข้อมูลตามความผันผวนและผลตอบแทน โดยใช้ข้อมูลอัตราส่วนทางการเงินผ่านการวิเคราะห์ทั้งการเงินเฉพาะหุ้นสามัญที่อยู่ในตลาด SET จำนวน 602 รายการ ระหว่างปี พ.ศ. 2556-2559 มาทำการทดสอบ ผ่านกระบวนการทำความสะอาดข้อมูล ซึ่งได้ข้อมูลทั้งหมด 263 รายการ แต่เนื่องจากข้อมูลมีความไม่สมดุลจึงมีการนำเทคนิค Synthetic Minority Over-sampling Technique (SMOTE) เข้ามาช่วยในการแก้ปัญหาดังกล่าว จากนั้นทำการเปรียบเทียบประสิทธิภาพ 4 อัลกอริทึม ได้แก่ Decision Tree, Random Forest, AdaBoost, XGBoost ผลการวิจัยพบว่า อัลกอริทึม XGBoost มีประสิทธิภาพในการจำแนกสูงที่สุด โดยมีค่าความถูกต้อง (Accuracy) เท่ากับร้อยละ 70.22 ค่าความแม่นยำ (Precision) เท่ากับร้อยละ 70.70 และค่าระลึก (Recall) เท่ากับร้อยละ 69.40

คำสำคัญ: การพยากรณ์หลักทรัพย์ แบ่งกลุ่มหลักทรัพย์
ผลตอบแทน ความผันผวน

Abstract

The objective of this research is to compare different algorithms for classifying groups of securities based on their volatilities and returns. To

predicting the groups, this study used numerous financial ratios as input data. Financial data of 602 common stocks in the SET market during the year 2013 – 2016 were used for data analysis. However, after cleaning process, 263 common stocks were left for later analysis. The Synthetic Minority Over-sampling Technique (SMOTE) was used to solve unbalance of the selected data. Then, four algorithms were compared: Decision Tree, Random Forest, AdaBoost, XGBoost. The results showed that XGBoost algorithm has the highest classification efficiency. It has accuracy at 70.22 %, precision at 70.70 %, and recall at 69.40 %.

Keywords: Securities Forecast, Securities Segmentation, Yield, Volatility

1. บทนำ

การลงทุนในตลาดหลักทรัพย์แห่งประเทศไทยนั้นมีการปรับตัวขึ้นลงและผันผวนตลอดเวลา แสดงถึงความเสี่ยงของการลงทุน โดยเป็นผลมาจากปัจจัยต่างๆ ที่มากระทบต่อราคาของหลักทรัพย์ นำไปสู่ความเสี่ยงต่อเงินลงทุนและผลตอบแทนที่ได้รับ [1]

แม้จะมีงานวิจัยเกิดขึ้นมากมายพยายามพัฒนาวิธีการพยากรณ์เพื่อมองหาโอกาสที่จะได้รับผลตอบแทนที่มากกว่าตลาด โดยอาศัยเทคนิคการทำเหมืองข้อมูล ไม่ว่าจะเป็น Random Forest [2] [3], Decision Tree [4],

XGBoost [5], AdaBoost [6] อย่างไรก็ตามบทความและงานวิจัยส่วนใหญ่บ่งบอกแนวโน้มราคาหลักทรัพย์ หรือผลตอบแทนอย่างใดอย่างหนึ่ง มีงานวิจัยไม่มากที่บ่งบอกแนวโน้มผลตอบแทนเมื่อเทียบกับตลาดและความผันผวนของหลักทรัพย์ไปพร้อมกัน

ดังนั้นในงานวิจัยนี้ ผู้วิจัยจึงมีแนวคิดในการพยากรณ์หลักทรัพย์ตามความผันผวนและผลตอบแทน โดยใช้อัตราส่วนทางการเงินผ่านการวิเคราะห์งบการเงินเฉพาะหุ้นสามัญที่อยู่ในตลาด SET ของตลาดหลักทรัพย์แห่งประเทศไทย จำนวน 602 รายการ ระหว่างปี พ.ศ. 2556-2559 เพื่อมองหาหลักทรัพย์ที่มีโอกาสได้ผลตอบแทนมากกว่าตลาดในอีก 5 ปีข้างหน้า โดยนำเทคนิคเหมือนข้อมูลด้วยวิธีการจำแนกกลุ่ม ประกอบด้วย Decision Tree, Random Forest, AdaBoost และ XGBoost มาทำการเปรียบเทียบประสิทธิภาพ และเลือกวิธีที่เหมาะสมมากที่สุด มาเป็นแบบต้นแบบในการพยากรณ์หลักทรัพย์ตามความผันผวนและผลตอบแทน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดการวิเคราะห์อัตราส่วนทางการเงิน

อัตราส่วนทางการเงินเป็นเครื่องมือในการวิเคราะห์ข้อมูลในงบการเงิน เพื่อค้นหาความสัมพันธ์จากรายการหนึ่งไปยังรายการหนึ่ง โดยอาศัยข้อมูลในอดีตมาทำการเปรียบเทียบในรูปอัตราร้อยละกับตัวธุรกิจในอดีต [7] โดยในงานวิจัยนี้ได้นำอัตราส่วนทางการเงิน 10 อัตราส่วนมาใช้ในการวิเคราะห์ ประกอบด้วย 1) อัตราส่วนทุนหมุนเวียนเร็ว 2) อัตราหมุนเวียนสินทรัพย์ 3) อัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น 4) อัตราผลตอบแทนต่อส่วนของผู้ถือหุ้น 5) อัตราผลตอบแทนจากสินทรัพย์ 6) อัตรากำไรสุทธิ 7) มูลค่าหุ้นตามบัญชีต่อหุ้น 8) กำไรสุทธิต่อหุ้น 9) อัตราผลตอบแทนจากเงินปันผล 10) ราคาต่อหุ้นเทียบกับมูลค่าทางบัญชีต่อหุ้น

2.2 การทำเหมืองข้อมูล (Data Mining)

กระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลนั้น โดยอาศัยหลักสถิติ และการเรียนรู้ของเครื่อง จากนั้นนำผลการ

วิเคราะห์ที่ได้ไปใช้ประโยชน์เพื่อการตัดสินใจ โดยการเรียนรู้ของเครื่องมีชนิดการเรียนรู้ 2 แบบ คือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นการเรียนรู้ที่มีค่าเฉลยให้กับข้อมูล และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เป็นการเรียนรู้ที่ไม่มีค่าเฉลย ทั้งนี้กระบวนการทำเหมืองข้อมูลอาศัยอัลกอริทึมของการเรียนรู้ของเครื่องเข้ามาช่วย [8]

2.3 อัลกอริทึมการเรียนรู้ของเครื่องในการพยากรณ์

2.3.1 ต้นไม้ตัดสินใจ (Decision Tree) เป็นการนำข้อมูลมาสร้างแบบจำลองเพื่อจำแนกประเภทต่างๆ มีลักษณะโครงสร้างเป็นต้นไม้ โดยการจำแนกดังกล่าวจำเป็นต้องมีการเรียนรู้แบบมีการสอนเพื่อจำแนกโครงสร้างการตัดสินใจหรือการทำนายไว้ล่วงหน้าเพื่อใช้ในการทำนายหรือจำแนกข้อมูลที่เข้ามาใหม่ [9]

2.3.2 ป่าสุ่ม (Random Forest) เป็นแบบจำลองจำแนกประเภท Ensemble Learning หลักการคือข้อมูลชุดเดียวกันจะมีการเรียนรู้ของแบบจำลองเหมือนกันหลายครั้ง ซึ่งแต่ละครั้งของการเรียนรู้จะเลือกส่วนข้อมูลที่เรียนรู้ไม่เหมือนกัน แล้วนำการตัดสินใจของแต่ละแบบจำลองที่ได้มาโหวตว่าคลาสไหนถูกเลือกมากที่สุด [10]

2.3.3 เอดาบูท (AdaBoost) เป็นเทคนิคแบบรวมกลุ่มเริ่มด้วยการสร้างชุดข้อมูลสอนโดยกำหนดน้ำหนักให้กับข้อมูลสอนแต่ละตัว แล้วนำไปสร้างแบบจำลอง โดยแต่ละรอบการสร้างแบบจำลองใหม่ ค่าน้ำหนักจะถูกเปลี่ยนไปตามความผิดพลาดของผลลัพธ์ที่แบบจำลองในรอบนั้นๆ กระทำต่อข้อมูล ถ้าแบบจำลองตอบถูกสำหรับข้อมูลสอนตัวใด ข้อมูลสอนตัวนั้นจะถูกลดค่าน้ำหนัก และถ้าแบบจำลองตอบผิด ข้อมูลสอนตัวนั้นจะถูกเพิ่มค่าน้ำหนัก เทคนิคนี้นอกจากจะช่วยเพิ่มประสิทธิภาพของแบบจำลองแล้วยังใช้ได้กับข้อมูลที่ไม่วางใจ [11]

2.3.4 เอ็กซ์ตรีมเกรเดียนท์บูท (XGBoost) เป็นแบบจำลองที่มีการเรียนรู้โดยเอาต้นไม้ตัดสินใจมาฝึกต่อกัน โดยจะหยุดเรียนรู้เมื่อไม่เหลือค่าความผิดพลาดจากต้นไม้ตัดสินใจก่อนหน้านี้ให้เรียนรู้ ด้วยเหตุนี้จึงทำให้แบบจำลองนี้มีความแม่นยำในการทำนายเพิ่มขึ้นเรื่อยๆ โดยพัฒนามาจาก Gradient Boosting ทำให้การเรียนรู้เร็ว

ขึ้น และใช้ทรัพยากรของเครื่องน้อยลง สามารถทำงานกับข้อมูลจำนวนมากได้ เป็นอัลกอริทึมที่นิยมใช้และทำให้ได้ประสิทธิภาพสูงในปัจจุบัน [12]

2.4 เทคนิคการจำแนกประเภทข้อมูล (Classification)

เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ โดยนำข้อมูลส่วนหนึ่งมาสอนให้ระบบเรียนรู้ (Training Data) เพื่อจำแนกข้อมูลออกเป็นกลุ่มตามกำหนดไว้ ผลลัพธ์ที่ได้จากการเรียนรู้ คือ โมเดลจำแนกประเภทข้อมูล โดยนำข้อมูลที่เหลือมาเป็นข้อมูลทดสอบ (Testing Data) ซึ่งจะถูกนำมาเปรียบเทียบกับกลุ่มข้อมูลที่ได้จากโมเดล เพื่อทดสอบความถูกต้อง และปรับปรุงโมเดลจนกว่าจะได้ค่าความถูกต้องในระดับที่น่าพอใจ หลังจากนั้นนำข้อมูลผ่านโมเดลเพื่อให้สามารถทำนายกลุ่มของข้อมูลได้ [13]

2.5 ข้อมูลไม่สมดุล (Imbalanced Data)

การที่ข้อมูลกลุ่มหนึ่งมีจำนวนข้อมูลมากกว่าอีกกลุ่มหนึ่งเป็นจำนวนมาก ข้อมูลถูกแบ่งเป็นข้อมูลส่วนมาก (Majority Class) และข้อมูลส่วนน้อย (Minority Class) ดังนั้น เมื่อข้อมูลมีความไม่สมดุลเกิดขึ้น อัลกอริทึมจะเอนเอียงไปทางกลุ่มข้อมูลส่วนมาก ทำให้การจำแนกกลุ่มข้อมูลส่วนน้อยมีความผิดพลาด [14] พิจารณาจากอัตราส่วนความไม่สมดุล จากสมการที่ 1

$$\text{Imbalance Ratio (IR)} = \frac{n_{\text{majority}}}{n_{\text{minority}}} \quad (1)$$

โดยที่ n_{majority} = จำนวนข้อมูลในกลุ่มส่วนมาก

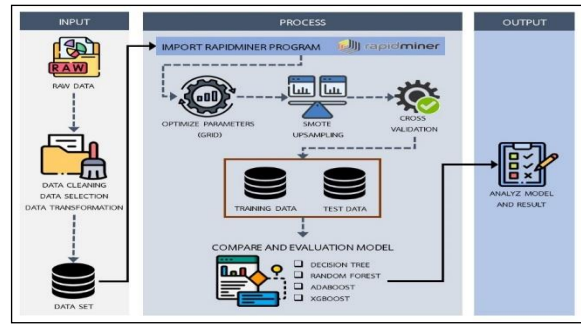
n_{minority} = จำนวนข้อมูลในกลุ่มส่วนน้อย

2.6 เทคนิคการเพิ่มความสมดุลข้อมูลด้วยวิธี SMOTE

การสุ่มเพิ่มตัวอย่างกลุ่มส่วนน้อย (Synthetic Minority Over-sampling Technique) เป็นเทคนิคที่ใช้ในการแก้ปัญหาการจำแนกข้อมูลที่ไม่สมดุล โดยทำการเพิ่มจำนวนข้อมูลกลุ่มน้อยให้เพิ่มขึ้น ทำให้การกระจายของกลุ่มข้อมูลมีความสมดุลมากขึ้น [15]

3. วิธีการดำเนินงาน

กรอบแนวความคิดการดำเนินงานวิจัยในการพยากรณ์หลักทรัพย์ด้วยเทคนิคเหมืองข้อมูล แสดงดังภาพที่ 1



ภาพที่ 1 : กรอบแนวความคิดการดำเนินงานวิจัย

3.1 การรวบรวมข้อมูล

ผู้วิจัยได้ทำการรวบรวมอัตราส่วนทางการเงินจากตลาดหลักทรัพย์แห่งประเทศไทยย้อนหลัง 4 ปี ตั้งแต่ปี 2556-2559 จากเว็บไซต์ <https://www.jitta.com/> จำนวน 602 รายการ เฉพาะหุ้นสามัญที่อยู่ในตลาด SET และทำการรวบรวมราคาปิดของ SET และหลักทรัพย์จากเว็บไซต์ www.finance.yahoo.com โดยเป็นข้อมูลรายวันตั้งแต่วันที่ 4 มกราคม 2560 ถึง วันที่ 2 ธันวาคม 2565 เพื่อนำมาหาอัตราผลตอบแทน และค่าเบต้าย้อนหลัง 5 ปี จากนั้นทำการกำหนดคลาสคำตอบให้แก่หลักทรัพย์ โดยมีเกณฑ์ในการกำหนด ดังนี้

- กลุ่ม 1 คือ ผลตอบแทนมากกว่าตลาด ความผันผวนต่ำ
- กลุ่ม 2 คือ ผลตอบแทนมากกว่าตลาด ความผันผวนสูง
- กลุ่ม 3 คือ ผลตอบแทนน้อยกว่าตลาด

3.2 การเตรียมข้อมูล

3.2.1 การทำความสะอาดข้อมูลโดยการตรวจสอบและแก้ไขข้อมูลที่ไม่สมบูรณ์ เช่น ข้อมูลไม่ครบถ้วน ข้อมูลที่มีค่าผิดปกติไปจากค่าอื่น ๆ ซึ่งการทำความสะอาดข้อมูลทำให้ข้อมูลจาก 602 รายการ เหลือเพียง 263 รายการ

3.2.2 การคัดเลือกข้อมูล ควรเลือกข้อมูลตัวแปรอิสระให้มีความสอดคล้องกับตัวแปรตาม พิจารณาจากอัตราส่วนทางการเงินที่ส่งผล จากงานวิจัยที่เกี่ยวข้องโดยสามารถคัดเลือกได้ 12 แอทริบิวต์ แสดงดังตารางที่ 1 กำหนด GROUP เป็นค่าเป้าหมาย และ STOCK เป็น id

ตารางที่ 1 : ปัจจัยนำเข้าที่ใช้ในการสร้างแบบจำลอง

No	Attribute	Description
1	STOCK	ชื่อหลักทรัพย์
2	QR	อัตราส่วนทุนหมุนเวียนเร็ว
3	TAT	อัตรามูลงเงินของสินทรัพย์รวม
4	DER	อัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น

ตารางที่ 1 : ปัจจัยนำเข้าที่ใช้ในการสร้างแบบจำลอง (ต่อ)

No	Attribute	Description
5	ROE	อัตราผลตอบแทนต่อส่วนของผู้ถือหุ้น
6	ROA	อัตราผลตอบแทนจากสินทรัพย์
7	NPM	อัตรากำไรสุทธิ
8	BVPS	มูลค่าหุ้นตามบัญชีต่อหุ้น
9	EPS	กำไรสุทธิต่อหุ้น
10	DivY	อัตราผลตอบแทนจากเงินปันผล
11	PBV	ราคาต่อหุ้นเทียบกับมูลค่าทางบัญชีต่อหุ้น
12	GROUP	ผลการดำเนินงานหลักทรัพย์

3.2.3 การแปลงรูปแบบข้อมูล เป็นการปรับค่าข้อมูลให้อัลกอริทึมสามารถเข้าใจได้ เช่น การแปลงค่ากลุ่มต่างๆ เป็นตัวเลข

3.3 การเพิ่มความสมดุลข้อมูลด้วยวิธี SMOTE

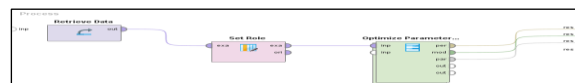
งานวิจัยนี้มีจำนวนกลุ่ม 1 และกลุ่ม 3 มากกว่าข้อมูลกลุ่ม 2 เมื่อคิดอัตราส่วนความไม่สมดุลค่าเท่ากับ 9.96 แสดงถึงความไม่สมดุลของข้อมูล แสดงดังตารางที่ 2 ดังนั้น ผู้วิจัยจึงนำเทคนิค SMOTE ซึ่งใช้กันอย่างแพร่หลายและให้ผลลัพธ์ที่ดี [16] มาช่วยเพิ่มความสมดุลข้อมูล

ตารางที่ 2 : ชุดข้อมูลงานวิจัยและอัตราส่วนความไม่สมดุล

จำนวนข้อมูล	Attribute	กลุ่ม 1	กลุ่ม 2	กลุ่ม 3	Imbalance Ratio
263	42	124	24	115	9.96

3.4 การพัฒนาแบบจำลอง

สำหรับงานวิจัยนี้ได้ทำการทดสอบ 4 อัลกอริทึม ได้แก่ Decision Tree, Random Forest, AdaBoost และ XGBoost โดยนำเครื่องมือ Optimize Parameters (Grid) เพื่อช่วยหาค่าพารามิเตอร์ที่มีประสิทธิภาพและเหมาะสมที่สุด เครื่องมือ SMOTE Upsampling เพื่อช่วยสร้างความสมดุลให้กับข้อมูล และเครื่องมือ Cross Validation เพื่อใช้ในการทดสอบประสิทธิภาพของแบบจำลอง ดังภาพที่ 2 ถึง 7



ภาพที่ 2 : การปรับพารามิเตอร์ด้วย Optimize Parameters (Grid)



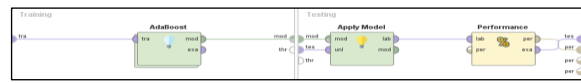
ภาพที่ 3 : การปรับพารามิเตอร์ด้วย SMOTE Upsampling และ Cross Validation



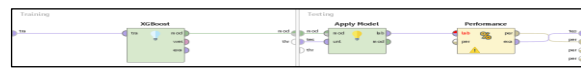
ภาพที่ 4 : การหาค่าและทดสอบพารามิเตอร์ด้วย Decision Tree



ภาพที่ 5 : การหาค่าและทดสอบพารามิเตอร์ด้วย Random Forest



ภาพที่ 6 : การหาค่าและทดสอบพารามิเตอร์ด้วย AdaBoost



ภาพที่ 7 : การหาค่าและทดสอบพารามิเตอร์ด้วย XGBoost

3.5 การวัดประสิทธิภาพและเปรียบเทียบแบบจำลอง

ในงานวิจัยนี้ทำการวัดประสิทธิภาพโดยทำการวัดค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าระลึก (Recall)

4. ผลการดำเนินงาน

4.1 ผลการปรับความไม่สมดุลของข้อมูล

จากตารางที่ 3 ข้อมูลกลุ่ม 2 ซึ่งเป็นข้อมูลกลุ่มส่วนน้อย ก่อนปรับ SMOTE มีจำนวน 24 รายการ หลังปรับ SMOTE เพิ่มจำนวนเป็น 124 เท่ากับข้อมูลกลุ่ม 1 ซึ่งเป็นกลุ่มข้อมูลส่วนมากที่มีจำนวนมากที่สุด ส่วนกลุ่ม 3 มีจำนวนเท่าเดิม

ตารางที่ 3 : การจัดการสมดุลข้อมูลก่อน-หลังด้วยเทคนิค SMOTE

กลุ่ม	ข้อมูลก่อนปรับ SMOTE	ข้อมูลหลังปรับ SMOTE
1	124	124
2	24	124
3	115	115
ผลรวม	263	363

4.2 ผลการพัฒนาแบบจำลอง

4.2.1 ผลการพัฒนาแบบจำลอง Decision Tree

จากการทดสอบพบว่าอัลกอริทึม Decision Tree มีประสิทธิภาพดีที่สุด เมื่อกำหนดค่าพารามิเตอร์ Maximal depth = 10, Confidence = 0.01, Criterion = Gini index ได้ค่าความถูกต้องร้อยละ 60.57 แสดงดังตารางที่ 4

ตารางที่ 4 : ผลการวัดประสิทธิภาพของอัลกอริทึม Decision Tree

Maximal depth	Confidence	Criterion	Accuracy
10	0.01	Gini index	59.24
10	0.01	Gini index	60.57
20	0.01	Gini index	59.01

4.2.2 ผลการพัฒนาแบบจำลอง Random Forest

จากการทดสอบพบว่าอัลกอริทึม Random Forest มีประสิทธิภาพดีที่สุด เมื่อกำหนดค่าพารามิเตอร์ Maximal

depth = 10, Number of trees = 50, Confidence = 0.01, Criterion = Gini index ได้ค่าความถูกต้องร้อยละ 65.88 แสดงดังตารางที่ 5

ตารางที่ 5 : ผลการวัดประสิทธิภาพของอัลกอริทึม Random Forest

Maximal depth	Number of trees	Confidence	Criterion	Accuracy
<u>10</u>	<u>50</u>	<u>0.01</u>	<u>Gini index</u>	<u>65.88</u>
20	50	0.01	Gini index	65.60
20	40	0.01	Information gain	65.58

4.2.3 ผลการพัฒนาแบบจำลอง AdaBoost

จากการทดสอบพบว่าอัลกอริทึม AdaBoost มีประสิทธิภาพดีที่สุด เมื่อกำหนดค่าพารามิเตอร์ Maximal depth = 20, Confidence = 0.01, Criterion = Gini index ได้ค่าความถูกต้องร้อยละ 68.85 แสดงดังตารางที่ 6

ตารางที่ 6 : ผลการวัดประสิทธิภาพของอัลกอริทึม AdaBoost

Maximal depth	Confidence	Criterion	Accuracy
10	0.01	Gini index	67.81
<u>20</u>	<u>0.01</u>	<u>Gini index</u>	<u>68.85</u>
30	0.01	Gini index	68.57

4.2.4 ผลการพัฒนาแบบจำลอง XGBoost

จากการทดสอบพบว่าอัลกอริทึม XGBoost มีประสิทธิภาพดีที่สุด เมื่อกำหนดค่าพารามิเตอร์ Maximal depth = 20, Booster = tree booster, Learning rate = 0.6 ได้ค่าความถูกต้องร้อยละ 70.22 แสดงดังตารางที่ 7

ตารางที่ 7 : ผลการวัดประสิทธิภาพของอัลกอริทึม XGBoost

Maximal depth	Booster	Learning rate	Accuracy
10	tree booster	0.8	68.04
<u>20</u>	<u>tree booster</u>	<u>0.6</u>	<u>70.22</u>
20	tree booster	1	68.60

4.3 ผลการเปรียบเทียบและคัดเลือกอัลกอริทึมที่เหมาะสม

จากการวิจัยพบว่า อัลกอริทึม XGBoost มีประสิทธิภาพมากที่สุด โดยได้ค่าความถูกต้องร้อยละ 70.22 ค่าความแม่นยำร้อยละ 70.70 และค่าระลึกร้อยละ 69.40 แสดงดังตารางที่ 8

ตารางที่ 8 : ผลการเปรียบเทียบอัลกอริทึมที่เหมาะสม

Algorithms	Accuracy	Precision	Recall
Decision Tree	60.57	59.80	60.22
Random Forest	65.88	64.63	65.47
AdaBoost	68.85	68.21	68.30
<u>XGBoost</u>	<u>70.22</u>	<u>70.70</u>	<u>69.40</u>

4.4 ผลการเปรียบเทียบค่าน้ำหนักปัจจัยแบบจำลอง

จากการวิจัยครั้งนี้ พบว่าจากปัจจัยนำเข้าทั้งหมด 10 ปัจจัย อัลกอริทึม Decision Tree, Random Forest และ XGBoost ทั้ง 10 ปัจจัยส่งผลต่อแบบจำลอง ในขณะที่อัลกอริทึม AdaBoost มีเพียง 4 ปัจจัยเท่านั้นที่ส่งผล ได้แก่ ROA, BVPS, EPS และ PBV

5. สรุปผล และอภิปรายผล

งานวิจัยครั้งนี้ ผู้วิจัยได้นำเสนอการเปรียบเทียบประสิทธิภาพสำหรับการพยากรณ์หลักทรัพย์ตามความผันผวนและผลตอบแทนด้วยการจำแนกหลักทรัพย์ออกเป็น 3 คลาส จำนวน 263 รายการ โดยใช้แบบจำลอง 4 อัลกอริทึม ได้แก่ Decision Tree, Random Forest, AdaBoost และ XGBoost ร่วมกับเทคนิค SMOTE เพื่อปรับแก้ความไม่สมดุลของข้อมูล ผลการทดสอบพบว่า ทั้ง 4 อัลกอริทึมที่มีการใช้ร่วมกับเทคนิค SMOTE มีประสิทธิภาพมากกว่า สอดคล้องกับงานวิจัย [16] ที่มีการใช้เทคนิค SMOTE ร่วมกับอัลกอริทึม k-NN โดยที่มีค่า Imbalance Ratio อยู่ใน ช่วง 2.83-5.20 และมีการจำแนกเป็น 3 คลาส เช่นเดียวกับงานวิจัยนี้ โดยได้ค่าความถูกต้อง 63.40-70.00 ในขณะที่งานวิจัย [17] ใช้อัลกอริทึม k-NN เช่นเดียวกัน จำแนก 2 คลาส มีค่า Imbalance Ratio เท่ากับ 8.71 แต่ไม่ได้ใช้เทคนิค SMOTE เข้ามาช่วย ได้ค่าความถูกต้องเพียง 58.82 เท่านั้น ดังนั้นสำหรับงานวิจัยนี้ที่จำแนกเป็น 3 คลาส และค่า Imbalance Ratio เท่ากับ 9.96 ซึ่งสูงกว่างานวิจัยทั้งสอง พบว่าอัลกอริทึม XGBoost ทำให้ได้แบบจำลองที่มีประสิทธิภาพมากที่สุด โดยมีค่าความถูกต้อง 70.22 และทั้ง 10 อัตราส่วนทางการเงินซึ่งเป็นปัจจัยนำเข้าส่งผลต่อแบบจำลองทั้งหมด ดังนั้นจึงสามารถนำข้อมูลอัตราส่วนทางการเงินดังกล่าว ซึ่งส่งผลให้ได้รับผลตอบแทนมากกว่าตลาดไปช่วยสนับสนุน และเป็นแนวทางในการตัดสินใจลงทุนหุ้นสามัญในตลาด SET ของนักลงทุนต่อไป

อย่างไรก็ตามงานวิจัยนี้ยังคงมีข้อจำกัดตามขอบเขตงานวิจัยที่กำหนด ดังนั้น จำเป็นต้องค้นหาอัลกอริทึม, เทคนิคการจัดการข้อมูลที่ไม่สมดุล ตลอดจนปัจจัย

ทางด้านเศรษฐกิจอื่น ๆ นอกเหนือจากการวิเคราะห์บริษัทผ่านอัตราส่วนทางการเงินเพียงอย่างเดียว เช่น ราคา น้ำมันดิบ คำนีราคาผู้บริโภค อัตราแลกเปลี่ยนเงินบาทต่อดอลลาร์สหรัฐ เป็นต้น โดยเป็นการวิเคราะห์ทางเศรษฐกิจซึ่งเป็นปัจจัยภายนอกบริษัทที่อาจส่งผลกระทบต่อราคาและผลตอบแทนของหลักทรัพย์เพื่อให้ได้แบบจำลองมีประสิทธิภาพมากยิ่งขึ้น

เอกสารอ้างอิง

- [1] ศุภกาญจน์ พุ่มจันทร์, “การวิเคราะห์ความผันผวนของอัตราผลตอบแทนของตลาดหลักทรัพย์ในประเทศไทย สหรัฐอเมริกา อังกฤษ และญี่ปุ่น”, *วารสารวิจัย มข. สาขามนุษยศาสตร์และสังคมศาสตร์ (ฉบับบัณฑิตศึกษา)* ปีที่ 8 ฉบับที่ 2 หน้า 88-97.
- [2] ปฎิวัติ ช่างบุญ และธนา สมพรเสริม, “การคาดการณ์ทิศทาง การเคลื่อนไหวของดัชนีตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้ตัวชี้วัดทางเทคนิคร่วมกับเทคนิคการเรียนรู้ของเครื่อง”, *การประชุมทางวิชาการมหาวิทยาลัยเกษตรศาสตร์ ครั้งที่ 58* หน้า 275-282.
- [3] I. Kumar, K. Dogra, C. Utreja, and P. Yadav, “A Comparative Study of Supervised Machine Learning Algorithms for Stock Market Trend Prediction”, in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, IEEE, Apr, 2018.
- [4] J. Xianya, H. Mo, and L. Haifeng, “Stock Classification Prediction Based on Spark”, *Procedia Computer Science*, vol. 162, pp. 243-250, 2019.
- [5] S. Zhao, “Nepal Stock Market Movement Prediction with Machine Learning”, in *2021 the 5th International Conference on Information System and Data Mining*, Silicon Valley CA USA, May, 2021.
- [6] M.Ghahramani and F. Aiolli, “Price direction prediction in financial markets, using Random Forest and Adaboost”, in *ESANN 2022 proceedings*, Bruges (Belgium) and online event, 2022, pp.139-144.
- [7] กัญญารัตน์ กิตติสนธิรักษ์, “อัตราส่วนทางการเงินและการเปลี่ยนแปลงมูลค่าของกิจการตามราคาตลาดในตลาดหลักทรัพย์แห่งประเทศไทย กลุ่ม SET100”, *มหาวิทยาลัยธุรกิจบัณฑิต*, 2562.
- [8] สุภาภณ จันทร์สกุล, “เทคนิคเหมืองข้อมูลในการวิเคราะห์ข้อมูลทางการแพทย์”, *วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเชีย*, ปีที่ 12 ฉบับที่ 2 หน้า 83-96.
- [9] ยศสยา แสงหิรัญ, “ปัจจัยที่มีผลต่อการผันแปรของหุ้นกลุ่มเทคนิคการทำเหมืองข้อมูล”, *การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา*, บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต, ปีที่ 2561 หน้า 2582-2590.
- [10] ณรงค์ศักดิ์ โควีวัลย์แสง และอัครนันท์ พงศธรวิวัฒน์, “ตัวแบบการเรียนรู้จำแนกประเภทซัพพลายเออร์แบบมีผู้สอนสำหรับปัญหาการประเมินประสิทธิภาพของซัพพลายเออร์ในระบบ SAP ERP”, *วารสารไทยการวิจัยดำเนินงาน*, ปีที่ 9, ฉบับที่ 1 หน้า 106-119.
- [11] กาญจน์ ณ ศรีระ, กิตติศักดิ์ เกิดประสพ และนิตยา เกิดประสพ, “การเปรียบเทียบเทคนิคการสุ่มตัวอย่างเพื่อการจำแนกข้อมูลที่ไม่สมดุล”, [สืบค้นวันที่ 27 ธันวาคม 2565]. จาก <https://ph01.tcithaijo.org/index.php/jait/article/download/90569/103583/370868>.
- [12] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System”, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA, Aug. 2016, pp.785-794.
- [13] สติมา มณฑาสวรรณ, “การพัฒนาาระบบค้นหาหารหัส ICD-10 สำหรับงานเวชระเบียน”, *ศิลปการ*, 2557.
- [14] กิระชาติ สุขสุทธิ, “การจำแนกข้อมูลไม่สมดุลโดยใช้การปรับปรุงข้อมูลร่วมกับการหาค่าพารามิเตอร์ที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรมที่มีการเริ่มต้นใหม่”, *เทคโนโลยีสุรนารี*, 2559.
- [15] พุทธิพร ธนธรรมเมธี และเยาวเรศ ศิริสถิตย์กุล, “เทคนิคการจำแนกข้อมูลที่พัฒนาสำหรับชุดข้อมูลที่ไม่สมดุลของภาวะข้อเข่าเสื่อมในผู้สูงอายุ”, *วารสารวิทยาศาสตร์และเทคโนโลยี*, ปีที่ 2562 ฉบับที่ 6 หน้า 1164-1178.
- [16] W. Chaipanha and P. Kaewwichian, “Smote vs. Random Undersampling for Imbalanced Data - Car Ownership Demand Model”, *Komunikácie*, vol. 24, no. 3, p. D105-D115, Jul. 2022.
- [17] R. S. Latha et al., “Stock Movement Prediction using KNN Machine Learning Algorithm”, in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, Jan. 2022, pp. 1-5.

แบบจำลองเพื่อการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐ โดยใช้เทคนิคเหมืองข้อมูล

Forecasting Model for The Number of Patients in The Public Hospital by using Data Mining Techniques

ณัฐชยาน์ แสงนาค (Nuttaya Sangnark)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857077@email.kmutnb.ac.th, ²sakchai.t@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้ทำการพัฒนาแบบจำลองเพื่อการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐ โดยใช้เทคนิคเหมืองข้อมูล ด้วยโปรแกรม Minitab 17 และใช้ชุดข้อมูลของจำนวนผู้ป่วยแบ่งแยกแผนกต่าง ๆ ดังนี้ 1) แผนกอายุรศาสตร์ 2) แผนกศัลยศาสตร์ และ 3) แผนกกุมารเวชศาสตร์ มีวัตถุประสงค์เพื่อ 1) เพื่อเปรียบเทียบค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยของการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐของโมเดลการพยากรณ์ 2) เพื่อหาแบบจำลองการพยากรณ์ที่เหมาะสมในการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐ งานวิจัยนี้ใช้เทคนิคการพยากรณ์ด้วยวิธีอนุกรมเวลา ประกอบด้วย 1) วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบสองครั้ง 2) วิธีการวิเคราะห์แนวโน้ม และ 3) วิธีแยกองค์ประกอบ เปรียบเทียบแบบจำลองด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ต่ำที่สุด ผลการวิจัยพบว่า แบบจำลองแบบแยกองค์ประกอบ ชนิด Multiplicative Model ของข้อมูลผู้ป่วยแผนกอายุรศาสตร์ แผนกศัลยศาสตร์ และแผนกกุมารเวชศาสตร์ มีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เท่ากับ 3.77, 14.35 และ 15.54

คำสำคัญ: การพยากรณ์ เหมืองข้อมูล วิธีปรับเรียบเอ็กซ์โพเนนเชียลแบบสองครั้ง วิเคราะห์แนวโน้ม วิธีแยกองค์ประกอบ

Abstract

This research was to develop a model for forecasting model for the number of patients in the public hospital by using data mining techniques and Minitab 17. The data set of the number of patients was divided into the following departments 1) Medicine department 2) Surgery department and 3) Pediatrics department. The objectives of this research are 1) compare the average absolute error percentage of the forecast of the number of patients in public hospitals of the forecasting model (2) find a suitable forecasting model to forecast the number of patients in public hospital. This research used Time-series forecasting consists of 1) Double exponential smoothing, 2) Trend Analysis and 3) Decomposition. The models were compared with the lowest mean absolute percentage error (MAPE), mean absolute deviation (MAD) and mean squared error (MSE). The final results show that the best forecasting model was Decomposition Multiplicative method due to the mean absolute error percentage (MAPE) for Medicine department, Surgery department, and Pediatrics department were 3.77, 14.35 and 15.54

Keyword: Forecasting, Data Mining, Double Exponential Smoothing, Trend Analysis, Decomposition.

1. บทนำ

การวางแผนอัตรากำลังบุคลากรทางการแพทย์เป็นเรื่องที่ทุกประเทศให้ความสนใจและศึกษาอย่างจริงจัง อย่างน้อยที่สุดนั้นแต่ละประเทศต้องการทราบถึงจำนวนความต้องการอัตรากำลังบุคลากรทางการแพทย์ทั้งในปัจจุบันและอนาคต เพื่อที่จะได้วางแผนการผลิตให้ได้เพียงพอกับความต้องการของประชาชนในประเทศ หรือไม่ให้เกิดเกินกว่าที่ประชาชนต้องการ ดังจะเห็นได้ว่าถ้ามีจำนวนบุคลากรทางการแพทย์หรือบุคลากรสาธารณสุขบางสาขาไม่เพียงพอ ก็จะทำให้บุคลากรสาธารณสุขในสาขานั้น ๆ ต้องทำงานมากกว่าที่ควรจะเป็นหรือในบางพื้นที่อาจจะไม่มีการให้บริการสาธารณสุขบางประเภท ทั้ง ๆ ที่ประชาชนมีความจำเป็นต้องได้รับการบริการนั้น ๆ เป็นต้น และในทางตรงกันข้าม ในกรณีที่ผลิตบุคลากรทางการแพทย์เกินก็จะทำให้มีจำนวนบุคลากรทางการแพทย์ไม่มีงานทำหรือได้ทำงานต่ำกว่าวุฒิการศึกษา การที่มีกำลังคนไม่เพียงพอหรือมีกำลังคนเกินกว่าความต้องการ หรือความจำเป็น ทำให้เกิดผลเสียต่อระบบการให้บริการด้านสุขภาพลดประสิทธิภาพการผลิต และส่งผลกระทบต่อขวัญและกำลังใจในการทำงานอีกด้วย

การวางแผนและการจัดการที่ดีจึงเป็นสิ่งจำเป็นที่จะช่วยให้การจัดการดูแลผู้ป่วยรวมไปถึงการบริหารการจัดการในด้านต่าง ๆ ภายในโรงพยาบาลมีประสิทธิภาพมากยิ่งขึ้น ผู้ป่วยที่มาใช้บริการเกิดความพึงพอใจ ลดอัตราการเสียชีวิตจากการไม่ได้รับการรักษาที่ทันถ่วงที หรือบุคลากรทางการแพทย์ทุกส่วนงานมีจำนวนไม่เพียงพอต่อการให้บริการ ดังนั้น เครื่องมือหรือวิธีที่จะทำให้คาดการณ์เหตุการณ์ในอนาคตได้คงหนีไม่พ้นการพยากรณ์ ในปัจจุบันการพยากรณ์มีบทบาทและมีความสำคัญเป็นอย่างมากในการวางแผนและตัดสินใจเกี่ยวกับการดำเนินงานด้านต่าง ๆ ทั้งในระยะสั้นและระยะยาว เพื่อนำข้อมูลที่ได้จากการพยากรณ์มาวางแผนควบคุม และคาดคะเนเหตุการณ์ที่จะเกิดขึ้นในอนาคต [1] วิธีการพยากรณ์สามารถแบ่งได้ 2 ประเภท คือ 1. การพยากรณ์เชิงคุณภาพ เป็นการพยากรณ์ที่ใช้ผู้ที่มีประสบการณ์ ความรู้ ความสามารถ เป็นผู้พยากรณ์

โดยไม่ใช้ตัวแบบทางคณิตศาสตร์ จึงตรวจสอบความแม่นยำของการพยากรณ์ได้ยาก 2. การพยากรณ์เชิงปริมาณ เป็นการพยากรณ์ที่ใช้ข้อมูลเชิงปริมาณ (ตัวเลข) ในอดีตเพื่อนำมาพยากรณ์ค่าในอนาคต โดยสร้างตัวแบบทางคณิตศาสตร์ ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีความสนใจที่จะศึกษาแนวทางหรือเทคนิคการพยากรณ์จำนวนผู้ป่วยที่เข้ามาใช้บริการในโรงพยาบาลรัฐ โดยใช้เทคนิคเหมืองข้อมูล และแบ่งแยกแผนกจากแผนกตัวอย่างที่ผู้วิจัยนำมาวิเคราะห์ในงานวิจัยฉบับนี้ ดังนี้ 1. แผนกอายุรศาสตร์ (Medicine) 2. แผนกศัลยศาสตร์ (Surgery) 3. แผนกกุมารเวชศาสตร์ (Pediatrics) ซึ่งแผนกที่ผู้วิจัยนำมาเป็นข้อมูลตัวอย่างในการวิเคราะห์ เป็นแผนกที่มีผู้เข้ารับบริการมากเป็นอันดับต้น ๆ ในโรงพยาบาลรัฐที่ผู้วิจัยทำการศึกษาอยู่จากการศึกษาพบว่า [2] ในการเลือกตัวแบบหรือวิธีการพยากรณ์ ลำดับแรกจำเป็นต้องศึกษารูปแบบของชุดข้อมูลอย่างละเอียดก่อน เช่น ทำการตรวจสอบว่าชุดข้อมูลมีรูปแบบอย่างไร รูปแบบองค์ประกอบ แนวโน้ม วัฏจักรฤดูกาล หรือว่ามีเพียงตัวแปรคู่เพียงอย่างเดียว ซึ่งวิธีการที่จะทำให้ทราบถึงองค์ประกอบของข้อมูลเหล่านี้ สามารถทำได้จากการวาดกราฟและการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ เมื่อทราบรูปแบบของชุดข้อมูลแล้ว จึงนำไปเลือกตัวแบบหรือวิธีการพยากรณ์

ผู้วิจัยจึงมีความสนใจนำแนวคิดและเทคนิคดังกล่าวมาประยุกต์ใช้ในการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐ โดยใช้เทคนิคเหมืองข้อมูล ซึ่งเป็นส่วนช่วยในการสนับสนุนการวางแผนการจ้างบุคลากรทางการแพทย์ได้อย่างเหมาะสม และเพื่อตอบสนองในการดูแลผู้ป่วย และพัฒนาให้การดูแลรักษาผู้ป่วยได้อย่างเหมาะสมมากยิ่งขึ้น

2. วรรณกรรมที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล

การค้นหาข้อมูลที่มีประโยชน์จากแหล่งข้อมูลที่มีเป็นจำนวนมากมายมหาศาล เพื่อดึงข้อมูลที่มีประโยชน์มาทำการวิเคราะห์ค้นหารูปแบบหรือความสัมพันธ์ที่เกิดในฐานข้อมูล และจัดทำเป็นสารสนเทศเพื่อใช้ในการวางแผนบริหารจัดการธุรกิจ โดยการแยกข้อมูลที่มีประโยชน์

ออกมาใช้งาน เปรียบเทียบคล้ายกับการทำเหมืองแร่ ที่จะต้องทำการแยกเศษหินดินทรายที่ไม่มีค่าและมีปริมาณ มากออกจากแร่ที่มีมูลค่ามากและมักจะมีปริมาณน้อย [3]

2.2 การพยากรณ์อนุกรมเวลา

เป็นการพยากรณ์เชิงปริมาณจะต้องอาศัยข้อมูลในอดีต เพื่อช่วยในการพยากรณ์สิ่งที่จะเกิดขึ้นในอนาคต ซึ่งการ พยากรณ์แบบอนุกรมเวลา เป็นเทคนิคที่ยอมรับโดยทั่วไป ว่ามีความถูกต้องมากที่สุด สามารถนำข้อมูลที่มีค่า หลากหลายจำนวนมากมาคำนวณหาค่าพยากรณ์ได้ ลักษณะของอนุกรมเวลาจะต้องใช้ข้อมูลความต้องการ สินค้าหรือข้อมูลทางสถิติในอดีต ที่ทำการเก็บรวบรวม ข้อมูลในช่วงระยะเวลาหนึ่ง โดยช่วงระยะเวลาในการเก็บ ข้อมูล เช่น การเก็บข้อมูลรายเดือน รายไตรมาส เป็น ระยะเวลา 3 ปี หรือ 5 ปี เป็นต้น [4]

2.2.1 วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียลซ้ำสอง ครั้ง (Double Exponential Smoothing Method) ดังสมการ ที่ 2-1

$$\hat{Y}_{t+1} = L_t + pT_t \quad (2-1)$$

ค่าปรับเรียบดังสมการที่ 2-2

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (2-2)$$

สมการของการประมาณค่าแนวโน้มดังสมการที่ 2-3

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2-3)$$

เมื่อ \hat{Y}_{t+1} คือ ค่าพยากรณ์ล่วงหน้า p งวด

L_t คือ ค่าปรับเรียบตัวใหม่ ณ เวลา t

α คือ ค่าคงที่ของการปรับเรียบ ($0 < \alpha < 1$)

Y_t คือ ค่าข้อมูลจริง ณ ช่วงเวลา t

β คือ ค่าคงที่ปรับเรียบสำหรับตัวประมาณ แนวโน้ม ($0 < \beta < 1$)

T_t คือ ตัวประมาณแนวโน้ม ณ ช่วงเวลา t

p คือ งวดเวลาที่ต้องการพยากรณ์ล่วงหน้า

2.2.2 วิธีการวิเคราะห์แนวโน้ม (Trend Analysis) ดัง สมการที่ 2-4 สมการที่ 2-5 และสมการที่ 2-6

$$Y = a + bx \quad (2-4)$$

เมื่อ Y คือ ค่าพยากรณ์ ณ เวลา t

a คือ ค่าตัดแกนที่ช่วงเวลา 0

b คือ ค่าความชันของเส้น

x คือ ตัวแปรอิสระ

โดยที่ $a = \bar{Y} - b\bar{X}$ (2-5)

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2-6)$$

2.2.3 วิธีการแยกองค์ประกอบ (Decomposition Method) สามารถแบ่งได้เป็น 2 รูปแบบ คือ การแยก องค์ประกอบแบบการคูณ (Multiplicative Decomposition) และ การแยกองค์ประกอบแบบการบวก (Additive Decomposition) ดังสมการที่ 2-7 และ สมการที่ 2-8

$$Y_t = T_t \times C_t \times S_t \times I_t \quad (2-7)$$

$$Y_t = T_t + C_t + S_t + I_t \quad (2-8)$$

การประมาณแนวโน้มแบบเส้นตรงดังสมการที่ 2-9

$$T_t = b_0 + b_1 t \quad (2-9)$$

การประมาณวัฏจักรดังสมการที่ 2-10

$$C_t \times I_t = \frac{Y_t}{T_t \times S_t} \quad (2-10)$$

การประมาณฤดูกาลดังสมการที่ 2-11

$$\frac{Y_t}{S_t} = T_t \times I_t \quad (2-11)$$

การประมาณรูปแบบไม่ปกติดังสมการที่ 2-12

$$I_t = \frac{C_t \times I_t}{C_t} \quad (2-12)$$

เมื่อ Y_t คือ ค่าพยากรณ์ที่เวลา t

T_t คือ ค่าการประมาณแนวโน้ม

C_t คือ ค่าการประมาณของวัฏจักร

S_t คือ ค่าการประมาณของฤดูกาล

I_t คือ ค่าการประมาณของรูปแบบไม่ปกติ

2.3 การวัดความถูกต้องของการพยากรณ์

การพยากรณ์โดยใช้รูปแบบวิธีการต่าง ๆ สามารถ เปรียบเทียบค่าที่พยากรณ์ได้กับค่าจริงที่เกิดขึ้น โดยสามารถหาความคลาดเคลื่อนที่เกิดจากการพยากรณ์ ได้จาก ค่าความคลาดเคลื่อนของการพยากรณ์ = ค่าที่ เกิดขึ้นจริง - ค่าที่พยากรณ์ [5]

ความคลาดเคลื่อนของการพยากรณ์สามารถวัดได้จาก ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อน (Mean Absolute Percent Error - MAPE) ดังสมการที่ 2-13

$$MAPE = \left[\frac{\sum [\frac{\text{ค่าที่เกิดขึ้นจริง} - \text{ค่าที่พยากรณ์}}{\text{ค่าที่เกิดขึ้นจริง}}]}{n} \right] \times 100 \quad (2-13)$$

และกำหนดเกณฑ์ที่ใช้สำหรับการพยากรณ์ ดังนี้ [6]

ค่า MAPE น้อยกว่า 10% ค่อนข้างแม่นยำ

ค่า MAPE อยู่ระหว่าง 10%-20% ดี

ค่า MAPE อยู่ระหว่าง 20%-50% พอใช้

ค่า MAPE มากกว่า 50% ไม่แม่นยำ

เช่น ถ้า MAPE = 5% แสดงว่าเทคนิคที่เลือกมีความ

ผิดพลาดในการพยากรณ์เฉลี่ย ร้อยละ 5

3. วิธีการดำเนินงานวิจัย

3.1 กรอบแนวคิดงานวิจัย

ผู้วิจัยใช้ข้อมูลจำนวนผู้ป่วยที่เข้ามาใช้บริการในโรงพยาบาลรัฐ ตั้งแต่ ปี พ.ศ. 2557-2564 (ชุดข้อมูลฝึกสอน) นำมาทำการพยากรณ์ และข้อมูลปี พ.ศ. 2565 เป็น (ชุดข้อมูลตรวจสอบ) โดยผู้วิจัยได้คัดเลือกตัวอย่างจำนวนผู้ป่วยแยกตามแผนกต่าง ๆ ได้แก่ 1.แผนกอายุรศาสตร์ (Medicine) 2.แผนกศัลยกรรม (Surgery) 3. แผนกกุมารเวชศาสตร์ (Pediatrics) เพื่อเป็นส่วนช่วยในการวางแผนการจัดหาอัตรากำลังบุคลากรทางการแพทย์ที่เหมาะสมและวางแผนงบประมาณในปีถัดไป

3.2 การรวบรวมข้อมูล

ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลจำนวนผู้ป่วยในโรงพยาบาลรัฐ จำนวน 108 เดือน ตั้งแต่ มกราคม 2557- ธันวาคม 2565 แบ่งออกเป็น 3 แผนก ดังนี้ 1.แผนกอายุรศาสตร์ (Medicine) 2.แผนกศัลยกรรม (Surgery) 3. แผนกกุมารเวชศาสตร์ (Pediatrics) โดยทั้งนี้จะแบ่งข้อมูลเป็น 2 ส่วน คือ 1.ชุดข้อมูลสำหรับการสร้างแบบจำลองการพยากรณ์ ปี พ.ศ. 2557-2564 (ชุดข้อมูลฝึกสอน) 2.ชุดข้อมูลสำหรับการเปรียบเทียบค่าความถูกต้องของการพยากรณ์ ปี พ.ศ. 2565 (ชุดข้อมูลตรวจสอบ)

ตารางที่ 1: ข้อมูลจำนวนผู้ป่วยที่เข้ารับการรักษาแผนก

อายุรศาสตร์ (Medicine) ระหว่างปี พ.ศ. 2557-2565

	แผนกอายุรศาสตร์ (Medicine)								
	2557	2558	2559	2560	2561	2562	2563	2564	2565
ม.ค.	61,148	61,052	61,485	66,214	66,552	68,880	69,891	52,033	59,975
ก.พ.	56,588	59,319	60,259	59,372	62,487	61,200	66,024	57,026	57,490
มี.ค.	63,825	65,032	70,570	71,839	66,430	68,277	56,568	71,239	64,246
เม.ย.	55,548	57,511	55,719	56,162	55,998	57,875	38,577	45,714	54,392
	แผนกอายุรศาสตร์ (Medicine)								

	2557	2558	2559	2560	2561	2562	2563	2564	2565
พ.ค.	62,546	62,742	63,129	70,921	70,083	64,151	53,906	49,391	65,735
มิ.ย.	60,429	58,070	63,087	65,552	66,394	65,237	61,535	59,897	64,052
ก.ค.	61,255	63,563	58,687	64,688	64,000	65,073	60,217	49,506	55,978
ส.ค.	60,332	61,466	69,613	68,931	68,395	68,451	67,601	51,198	68,610
ก.ย.	62,910	63,709	65,644	67,373	64,508	65,039	61,657	53,641	65,155
ต.ล.	63,584	63,297	62,867	62,285	65,043	68,674	63,748	55,115	63,714
พ.ย.	61,193	64,066	66,483	71,247	66,493	67,591	65,868	63,647	66,460
ธ.ล.	61,829	59,641	61,053	61,920	61,401	58,889	63,963	59,983	62,752
รวม	731,187	739,468	758,596	786,504	777,784	779,337	729,555	668,390	748,559

ตารางที่ 2: ข้อมูลจำนวนผู้ป่วยที่เข้ารับการรักษาแผนกศัลยกรรม (Surgery) ระหว่างปี พ.ศ. 2557-2565

	แผนกศัลยกรรม (Surgery)								
	2557	2558	2559	2560	2561	2562	2563	2564	2565
ม.ค.	34,275	32,033	31,701	33,987	34,071	34,576	35,355	18,819	28,987
ก.พ.	31,784	31,601	31,426	32,521	32,036	31,010	32,038	25,608	27,857
มี.ค.	34,755	33,428	36,059	37,300	34,298	35,043	25,664	35,039	30,136
เม.ย.	30,869	29,280	29,482	30,276	28,431	29,793	6,612	14,394	25,033
พ.ค.	34,238	32,115	30,814	35,989	35,149	33,496	14,845	13,148	28,740
มิ.ย.	34,933	33,395	33,960	35,553	34,387	33,042	30,022	23,846	30,760
ก.ค.	34,165	34,087	30,923	34,487	31,636	33,080	32,855	17,996	24,823
ส.ค.	32,821	31,432	34,554	36,653	35,334	35,348	34,707	17,278	30,630
ก.ย.	34,984	34,254	33,047	35,755	32,892	32,901	33,130	21,902	30,089
ต.ล.	35,512	34,106	31,858	32,268	34,983	34,831	32,685	25,202	29,522
พ.ย.	33,333	33,150	34,741	38,864	34,049	35,062	32,500	32,069	31,023
ธ.ล.	31,610	30,500	31,494	33,220	30,219	29,633	32,322	30,600	29,504
รวม	403,279	389,381	390,059	416,873	397,485	397,815	342,735	275,901	347,104

ตารางที่ 3: ข้อมูลจำนวนผู้ป่วยที่เข้ารับการรักษาแผนก

กุมารเวชศาสตร์ (Pediatrics) ระหว่างปี พ.ศ. 2557-2565

	กุมารเวชศาสตร์ (Pediatrics)								
	2557	2558	2559	2560	2561	2562	2563	2564	2565
ม.ค.	14,256	13,744	13,329	12,446	13,372	13,410	13,339	6,794	9,523
ก.พ.	13,269	12,831	13,391	11,705	12,388	11,512	11,495	8,089	8,292
มี.ค.	14,706	13,831	14,580	13,500	13,164	12,980	9,290	11,311	9,952
เม.ย.	12,978	12,241	11,463	11,045	11,159	11,745	3,434	6,002	7,996
พ.ค.	13,191	12,737	12,518	13,146	13,910	12,653	7,298	6,821	9,991
มิ.ย.	13,881	13,565	14,002	14,027	13,989	13,043	10,642	9,735	10,802
ก.ค.	14,094	14,050	12,120	12,917	13,022	13,363	10,418	7,565	8,453
ส.ค.	13,767	13,886	14,921	14,594	14,460	13,889	11,069	8,563	10,384
ก.ย.	14,642	13,790	14,083	13,542	12,765	13,477	10,802	8,901	10,315
ต.ล.	15,241	14,564	13,456	12,477	13,143	14,339	11,468	9,327	10,447
พ.ย.	13,365	14,020	13,593	13,809	12,930	12,774	11,903	10,247	10,344
ธ.ล.	13,203	12,870	12,512	11,880	11,546	10,904	11,027	10,199	10,097
รวม	166,593	162,129	159,968	155,088	155,848	154,089	122,185	103,554	116,596

3.3 การวิเคราะห์ข้อมูลเบื้องต้น

การวิเคราะห์ข้อมูลเบื้องต้น ผู้วิจัยทำการพิจารณาแนวโน้มและอิทธิพลของฤดูกาลของชุดข้อมูล เพื่อเลือกแบบจำลองที่มีความเหมาะสมกับลักษณะของชุดข้อมูล จากนั้นทำการหาและตัดค่าผิดปกติ (Outliers) ออกจากชุดข้อมูล และแทนที่ด้วยค่าเฉลี่ยจำนวนผู้ป่วยที่เข้ารับบริการเฉลี่ย 6 เดือนของปีในแต่ละโรคนั้น ๆ เพื่อลดความผิดพลาดในการประมาณค่าพารามิเตอร์ของแบบจำลอง โดยการวิเคราะห์ข้อมูล ผู้วิจัยทำการวิเคราะห์โดยการสร้าง Time Series Plot ผ่านโปรแกรม Minitab

3.4 การสร้างแบบจำลองและการพยากรณ์

สร้างแบบจำลองด้วยเทคนิคการวิเคราะห์อนุกรมเวลา (Time Series Analysis) โดยใช้ข้อมูลจำนวนผู้ป่วยแผนกอายุรศาสตร์ (Medicine), แผนกศัลยศาสตร์ (Surgery) และแผนกกุมารเวชศาสตร์ (Pediatrics) ตั้งแต่ปี พ.ศ. 2557-2564 จำนวน 96 เดือน (ชุดข้อมูลฝึกสอน) และใช้เครื่องมือช่วยวิเคราะห์ค่าทางสถิติด้วยโปรแกรม Minitab โดยใช้เทคนิคในการพยากรณ์ 3 วิธี คือ 1.วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง 2.วิธีการวิเคราะห์แนวโน้ม และ 3.วิธีการแยกองค์ประกอบ

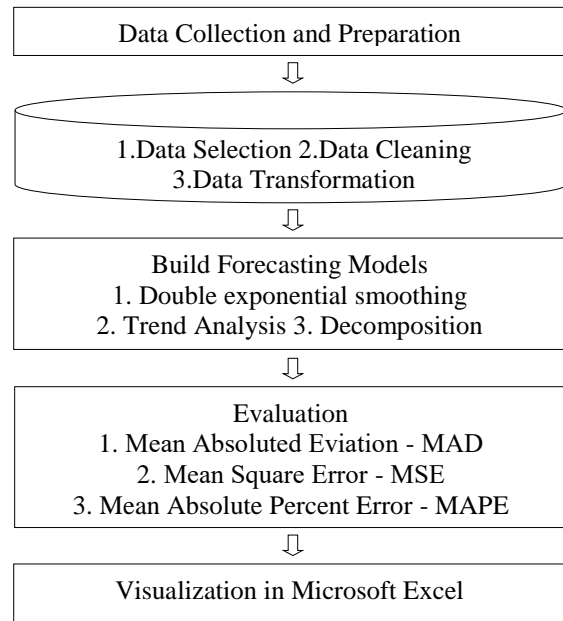
3.5 การเลือกแบบจำลองการพยากรณ์

ทำการเลือกแบบจำลองการพยากรณ์ที่เหมาะสมกับชุดข้อมูล โดยการเปรียบเทียบจากค่าความคลาดเคลื่อน (Forecast Error) ของข้อมูล ด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ต่ำที่สุด หากตัวแบบใดให้ค่าต่ำที่สุด ให้ตัวแบบพยากรณ์นั้นเป็นตัวแบบที่เหมาะสมกับชุดข้อมูล เพื่อนำค่าพยากรณ์ไปทวนสอบความแม่นยำกับชุดข้อมูลตรวจสอบ

3.6 การตรวจสอบความแม่นยำของแบบจำลอง

นำแบบจำลองมาประเมินความสามารถในการนำไปใช้ในการพยากรณ์ โดยใช้ข้อมูลปริมาณจำนวนผู้ป่วยทั้ง 3 แผนก จำนวน 12 เดือน (ชุดข้อมูลตรวจสอบ) มาหาค่า Forecast Error จากชุดข้อมูลตรวจสอบและทำการ

พิจารณาค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percent Error: MAPE)



ภาพที่ 1: กรอบแนวคิดงานวิจัย

4. ผลการดำเนินงาน

4.1 ผลการสร้างและเลือกแบบจำลองการพยากรณ์

จากผลการสร้างแบบจำลองทั้ง 3 วิธี สามารถสรุปค่าความคลาดเคลื่อนของแบบจำลองการพยากรณ์ได้ดังตารางที่ 4

ตารางที่ 4: ผลค่าความคลาดเคลื่อน MAPE MAD และ MSD ของแบบจำลองการพยากรณ์

แบบจำลอง	ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)	ค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD)	ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSD)
1.ผู้ป่วยแผนกอายุรศาสตร์ (Medicine)			
Double Exponential Smoothing	6	3972	27067248
Trend Analysis	5	3274	16745676
Decomposition Multiplicative	5	2770	14888330
Decomposition Additive	5	2787	14985867
2.ผู้ป่วยแผนกศัลยศาสตร์ (Surgery)			
Double Exponential Smoothing	6	1843	4854513
Trend Analysis	5	1576	3748495
Decomposition Multiplicative	3	958	1718858

Decomposition Additive	3	961	1725204
แบบจำลอง	ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)	ค่าความผิดพลาดสัมบูรณ์เฉลี่ย (MAD)	ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSD)
3.ผู้ป่วยแผนกกุมารเวชศาสตร์ (Pediatrics)			
Double Exponential Smoothing	9	1051	1683604
Trend Analysis	7	837	1079834
Decomposition Multiplicative	6	705	843032
Decomposition Additive	6	706	853230

4.2 ผลการตรวจสอบความแม่นยำของแบบจำลอง

จากผลการคัดเลือกแบบจำลองพยากรณ์ แบบจำลองที่มีความเหมาะสมในการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐ คือ แบบจำลองวิธีแยกองค์ประกอบ (Decomposition) ชนิด Multiplicative Model ของชุดข้อมูลผู้ป่วยแผนกอายุรศาสตร์ (Medicine), ผู้ป่วยแผนกศัลยศาสตร์ (Surgery) และผู้ป่วยแผนกกุมารเวชศาสตร์ (Pediatrics) ซึ่งผู้วิจัยนำแบบจำลองดังกล่าวมาตรวจสอบความแม่นยำด้วยค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เท่ากับ 3.77, 14.35 และ 15.54 ตามลำดับ ซึ่งมีค่าต่ำกว่าเกณฑ์ 20% จึงถือว่าอยู่ในเกณฑ์ดี

5. สรุป

การจัดทำวิจัยในครั้งนี้ผู้วิจัยได้ทำการศึกษาเพื่อหาแบบจำลองการพยากรณ์ที่เหมาะสมในการพยากรณ์จำนวนผู้ป่วยในโรงพยาบาลรัฐโดยใช้เทคนิคเหมืองข้อมูลและใช้ข้อมูลที่มีการเก็บรวบรวมจากแหล่งข้อมูลโรงพยาบาลรัฐที่เป็นข้อมูลจริง แยกแผนกดังนี้ 1.แผนกอายุรศาสตร์ (Medicine) 2.แผนกศัลยศาสตร์ (Surgery) 3. แผนกกุมารเวชศาสตร์ (Pediatrics) พบว่า แบบจำลองวิธีแยกองค์ประกอบ (Decomposition) ชนิด Multiplicative Model ของทุกชุดข้อมูล ให้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เท่ากับ 3.77, 14.35 และ 15.54 ตามลำดับ ซึ่งมีค่าต่ำกว่าเกณฑ์ 20% จึงถือว่าอยู่ในเกณฑ์ดี อย่างไรก็ตามงานวิจัยฉบับนี้พบว่ามีค่าผิดปกติ (Outliers) ของข้อมูลจำนวนผู้ป่วยที่เข้ามาใช้บริการภายใน

โรงพยาบาลรัฐมีมากในช่วงปี พ.ศ. 2562-2565 สืบเนื่องมาจากสาเหตุจากการแพร่ระบาดของเชื้อไวรัส COVID-19 ที่รัฐบาลมีมาตรการเฝ้าระวัง อยู่บ้าน หยุดเชื้อเพื่อชาติ โดยให้ออกจากที่อยู่อาศัยเท่าที่จำเป็น นั่นอาจเป็นสาเหตุหนึ่งที่ทำให้ผู้ป่วยในช่วงนั้นมีจำนวนลดลงอย่างผิดปกติ ซึ่งส่งผลกระทบต่อพยากรณ์ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา เนื่องจากอาศัยค่าในอดีตมาพยากรณ์ค่าในอนาคตเพียงปัจจัยเดียว ในการศึกษาการพยากรณ์ในครั้งต่อไป อาจต้องนำปัจจัยอื่น ๆ เข้ามาพิจารณาพร้อมด้วยเพื่อความแม่นยำของการพยากรณ์มากยิ่งขึ้น

เอกสารอ้างอิง

- [1] ยูพากรณ์ อารีพงษ์. “เทคนิคการพยากรณ์ (Forecasting Techniques).” *วารสารศิลปศาสตรและวิทยาศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์*. ปีที่ 16 ฉบับที่ 49 พฤษภาคม-สิงหาคม พ.ศ. 2545 หน้า 58-67.
- [2] คณะบริหารศาสตร์ มหาวิทยาลัยอุบลราชธานี. [สืบค้นเมื่อวันที่ 11 มิถุนายน 2565]. “การเลือกใช้เทคนิคการพยากรณ์แบบอนุกรมเวลา.” จาก <http://www.bus.ubu.ac.th/keepfile/292233358907%20บทที่%202%20ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.pdf>
- [3] กองโลจิสติกส์. [สืบค้นเมื่อวันที่ 23 กรกฎาคม 2565]. “การการทำเมืองข้อมูล.” จาก <https://doi.dip.go.th/th/category/2019-02-08-08-57-30/2019-03-15-08-49-57>.
- [4] ปุณชญา กันธิยะ. “แบบจำลองอนุกรมเวลาสำหรับการพยากรณ์จำนวนผู้ป่วยในสถานพยาบาล.” [การค้นคว้าอิสระ]. กรุงเทพมหานคร : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2562.
- [5] อติศักดิ์ ทูลธรรม. นดา เตชะบุญญาส. กวิน พินสาราย. “การพยากรณ์ความต้องการวัตถุดิบในการผลิตขนมปังบิสกิต.” *วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธนบุรี* ปีที่ 5 ฉบับที่ 1 มกราคม-มิถุนายน พ.ศ. 2564 หน้า 30-38.
- [6] สุพรรณิ อึ้งปัญสัตตวงศ์. “เทคนิคการพยากรณ์เชิงสถิติ” *ขอนแก่น : มหาวิทยาลัยขอนแก่น, 2555.*

การพัฒนาระบบแนะนำสถานที่ท่องเที่ยว ด้วยเทคโนโลยีการเรียนรู้ของเครื่อง

Development of a Tourist Attraction Recommendation System with Machine Learning Technology

สุภัทสรุ กระแสสัตย์ (Supatsara Krasaesat)¹ และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857166@email.kmutnb.ac.th, ²montean.r@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์ในการสร้างแบบจำลองการระบบแนะนำสถานที่ท่องเที่ยว ภาคเหนือ ประเทศไทย โดยใช้ อัลกอริทึมเพื่อนบ้านใกล้ที่สุด และเปรียบเทียบสูตรการหาระยะทางทั้งหมด 5 สูตร ได้แก่ Euclidean Distance, Camberra Distance, Chebyshev Distance, Dynamic Time Warping Distance และ Manhattan Distance โดยใช้ข้อมูลจากเว็บไซต์ท่องเที่ยวต่าง ๆ บนอินเทอร์เน็ต ประกอบด้วยคุณลักษณะจำนวน 4 คุณลักษณะ ทั้งหมด 500 รายการ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลองจำนวน 350 รายการ และข้อมูลสำหรับทดสอบแบบจำลองจำนวน 150 รายการ โดยนำข้อมูลมาทำการทดสอบเปรียบเทียบประสิทธิภาพสำหรับการจำแนกสถานที่ท่องเที่ยว ผลการวิจัยพบว่า สูตรการหาระยะทาง Manhattan Distance ของอัลกอริทึมเพื่อนบ้านใกล้ที่สุดนั้นมีประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องร้อยละ 85.57

คำสำคัญ: สถานที่ท่องเที่ยว อัลกอริทึมเพื่อนบ้านใกล้ที่สุด การจำแนกข้อมูล

ABSTRACT

The objective of this research is to recommend a tourist attraction in the northern region of Thailand using the nearest neighbor algorithm, and compared five formulas for distance determination, namely Euclidean Distance, Camberra Distance, Chebyshev Distance, Dynamic Time Warping Distance and Manhattan Distance by using

information from travel websites on the Internet.

It consists of 4 attributes, 500 items in total. Then, the data is divided into 2 parts: 350 items of model building data and 150 items of model testing data. The results showed that The Manhattan Distance formula of the nearest neighbor algorithm is most efficient. with an accuracy of 85.57 percent

Keywords: Tourist attractions, K-Nearest Neighbors, Classification.

1. บทนำ

ปัจจุบันประชากรทั่วโลกให้ความสนใจในการเดินทางมาท่องเที่ยวในประเทศไทยเพิ่มสูงขึ้นทุกปี ทำให้เกิดธุรกิจเกี่ยวกับการท่องเที่ยวขึ้นมาอย่างมหาศาล พร้อมกับการเจริญเติบโตของธุรกิจ มีมูลค่าตลาดที่สูงขึ้น ทั้งสถานที่ท่องเที่ยวที่ได้รับความนิยมอยู่แล้วและมีหลากหลายสถานที่ท่องเที่ยวที่เพิ่งจะถูกค้นพบและเพิ่งจะได้รับความนิยมทำให้มีตัวเลือกในการตัดสินใจเพิ่มมากขึ้น โดยภาคเหนือของประเทศไทยนั้นถือว่าได้รับความนิยมเป็นอย่างมากในด้านการท่องเที่ยวด้วยสถานที่ท่องเที่ยวที่สวยงาม หลากหลาย และแปลกตาทำให้ภาคเหนือเป็นภาคที่มีนักท่องเที่ยวเพิ่มขึ้นในทุก ๆ ปี โดยในการไปท่องเที่ยวในแต่ละครั้งผู้ใช้งานจะต้องทำการศึกษาข้อมูลเกี่ยวกับสถานที่และจังหวัดนั้น ๆ อย่างมาก เช่น สภาพอากาศ สิ่งที่น่าสนใจ หรือแม้กระทั่งงบประมาณที่ต้องใช้ ทำให้เกิดความยากในการตัดสินใจเลือกสถานที่

ท่องเที่ยวที่ต้องการจะไปได้ ประกอบกับงบประมาณที่จำกัด ช่วงเวลาที่จะไป และข้อจำกัดต่าง ๆ

จากปัญหาดังกล่าว ผู้จัดทำจึงได้เลือกที่จะพัฒนาระบบแนะนำสถานที่ท่องเที่ยวภายในภาคเหนือ ประเทศไทย เพื่อช่วยในการอำนวยความสะดวกและจัดหาสถานที่ที่ตรงกับข้อจำกัดต่าง ๆ ของผู้ใช้ให้มากที่สุด อีกทั้งผู้ใช้อย่างสามารถใช้งานระบบได้ทุกที่ ทุกเวลา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบสนับสนุนการตัดสินใจเพื่อการเดินทางท่องเที่ยวในจังหวัดอุตรดิตถ์ผ่านระบบดาวเทียมบอกพิกัด [1]

เกรียงศักดิ์, ปราโมทย์, ไมตรี และ วชิระ ได้ทำการนำ google map api ซึ่งเป็น api ที่ถูกจัดทำขึ้น โดยบริษัท google โดย google map api นั้นคือ api ที่ช่วยในการทำงานที่ซับซ้อนซึ่งเกี่ยวข้องกับแผนที่ให้ทำงานได้ง่ายขึ้นโดยสามารถเชื่อมต่อกับแอปพลิเคชัน และเว็บไซต์โดยส่วนใหญ่จะมีการใช้งานในการค้นหาร้านอาหาร สถานที่ท่องเที่ยว และเส้นทางในการเดินทาง แต่ยังมีบางสถานที่ซึ่งยังขาดหายไป ทางผู้วิจัยจึงมีแนวคิดในการแก้ปัญหาโดยการนำระบบสนับสนุนการตัดสินใจมาช่วยในการแนะนำสถานที่ และเส้นทาง โดยการรวบรวมข้อมูลแหล่งท่องเที่ยวภายในจังหวัดอุตรดิตถ์ 11 แห่ง และข้อมูลสถานที่ใกล้เคียง ซึ่งข้อมูลจะประกอบไปด้วยตำแหน่งละติจูด ลองจิจูด ประเภทสถานที่ ประเภทที่พัก ประเภทอาหาร และประเภทของฝาก จากนั้นได้นำข้อมูลมาคำนวณด้วยอัลกอริทึม K-NN การสรุปผลมีการสอบถามความพึงพอใจจากด้านต่าง ๆ 3 ด้าน 1) ด้านการทำงานได้ตรงตามความต้องการ ระบบได้รับความพึงพอใจในการนำเสนออยู่ในพอใจระดับมาก 2) ด้านการทำงานของระบบ เช่นความถูกต้องของข้อมูล ระบบได้รับความพึงพอใจในการนำเสนออยู่ในพอใจระดับมาก 3) ด้านการใช้งานของระบบ เช่นการใช้งานง่าย ระบบได้รับความพึงพอใจในการนำเสนออยู่ในพอใจระดับมาก

2.2 ระบบสนับสนุนการตัดสินใจในการวิเคราะห์ความเสี่ยงของการเกิดโรคเรื้อรัง กรณีของโรคเบาหวานและความดันโลหิตสูง [2]

ฉิชา ได้ทำการรวบรวมข้อมูลจาก 2 แหล่งข้อมูล ได้แก่ 1) การสัมภาษณ์ผู้ที่เกี่ยวข้องกับการคัดกรองผู้ป่วย 2) การศึกษาจากเอกสารและฐานข้อมูลที่เกี่ยวข้อง ได้แก่ เอกสารที่เกี่ยวข้องกับการคัดกรองความเสี่ยงของการเกิดโรคเรื้อรังกรณีของโรคเบาหวาน และความดันโลหิตสูง โดยใช้เทคนิคต้นไม้ตัดสินใจ และ KNN โดยจะแบ่งผู้ป่วยออกเป็น 7 กลุ่มตามหลักเกณฑ์ของปีงปองจรรยาชีวิต 7 สี โดยมีการใช้เทคนิค 10 – Fold Cross Validation ในการตรวจสอบประสิทธิภาพของโมเดลทั้ง 2 และมีข้อมูลในการทดสอบเพื่อนำไปสร้างเป็นโมเดลจำนวน 2,000 รายการ จากการวิจัยพบว่าเทคนิคต้นไม้ตัดสินใจให้ค่าความถูกต้องที่ดีกว่าคือร้อยละ 96.45 และได้ทำการจัดกลุ่มตามเกณฑ์ปีงปองจรรยาชีวิต 7 สี ดังนี้ 1) กลุ่มปกติ (สีเขียว) 2) กลุ่มเสี่ยง (สีเหลืองอ่อน) 3) กลุ่มป่วยระดับที่ 0 (สีเขียวเข้ม) 4) กลุ่มป่วยระดับที่ 1 (สีเหลือง) 5) กลุ่มป่วยระดับที่ 2 (สีส้ม) 6) กลุ่มป่วยระดับที่ 3 (สีแดง) 7) กลุ่มป่วยรุนแรง (สีดำ)

2.3 Machine Learning [3]

Machine Learning คือ ส่วนการเรียนรู้ของเครื่อง ถูกใช้งานเสมือนเป็นสมองของ AI (Artificial Intelligence) เราอาจพูดได้ว่า AI ใช้ Machine Learning ในการสร้างความฉลาด มักจะใช้เรียกโมเดลที่เกิดจากการเรียนรู้ของปัญญาประดิษฐ์ ไม่ได้เกิดจากการเขียนโดยใช้นมนุษย์ มนุษย์มีหน้าที่เขียนโปรแกรมให้ AI (เครื่อง) เรียนรู้จากข้อมูลเท่านั้น ที่เหลือเครื่องจัดการเอง

Machine Learning เรียนรู้จากสิ่งที่เราส่งเข้าไปกระตุ้น แล้วจดจำเอาไว้เป็นมันสมอง ส่งผลลัพธ์ออกมาเป็นตัวเลข หรือ code ที่ส่งต่อไปแสดงผล หรือให้เจ้าตัว AI นำไปแสดงการกระทำ Machine Learning เองสามารถเอาไปใช้งานได้หลายรูปแบบ ต้องอาศัยกลไกที่เป็นโปรแกรม หรือเรียกว่า Algorithm ที่มีหลากหลายแบบ โดยมี Data Scientist เป็นผู้ออกแบบ หนึ่งใน Algorithm ที่ได้รับความนิยมสูง คือ Deep Learning ซึ่งถูกออกแบบมาให้ใช้งานได้ง่าย และประยุกต์ใช้ได้หลายลักษณะงาน อย่างไรก็ตาม ในการทำงานจริง Data Scientist จำเป็นต้องออกแบบตัวแปรต่างๆ ทั้งในตัว Deep Learning เอง และต้องหา Algorithm อื่นๆ มาเป็นคู่แข่งเปรียบเทียบ เพื่อมองหา Algorithm ที่เหมาะสมที่สุดในการใช้งานจริง

2.4 K-Nearest Neighbors (K-NN) [4]

ขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors Algorithm) เป็นวิธีที่ใช้ในการจัดแบ่งคลาส โดยเทคนิคนี้จะตัดสินใจว่า คลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ๆ ได้บ้าง โดยการตรวจสอบจำนวนบางจำนวน ในขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด ของกรณีหรือเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวม (Count Up) ของจำนวนเงื่อนไขหรือกรณีต่างๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไขใหม่ๆ ให้คลาสที่เหมือนกันกับคลาสที่ใกล้เคียงกันมากที่สุด K- Nearest Neighbor เป็นอัลกอริทึมที่ใช้ในการจัดกลุ่มข้อมูล (Classification) ซึ่งเป็นอัลกอริทึมที่อยู่ในกลุ่มของ Supervised learning ซึ่งในงานวิจัยนี้ผู้วิจัยได้ใช้สูตรในการหาระยะทางของอัลกอริทึม K-Nearest Neighbors (K-NN) ทั้งหมด 5 สูตร ดังนี้

2.4.1. Euclidean Distance คือ สูตรการหาระยะทางระหว่างจุดในเส้นตรง วิธีระยะทางนี้ใช้ทฤษฎีบทพีทาโกรัส และเป็นการคำนวณระยะทางที่ใช้บ่อยที่สุดในกระบวนการแมชชีนเลิร์นนิง สูตรการหาระยะทางแบบ Euclidean เป็นผลต่างของผลต่างรากที่สองของเวกเตอร์สองตัว

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

2.4.2. Camberra Distance คือ สูตรที่ใช้เพื่อรับระยะทางจากคู่ของจุดที่ข้อมูลอยู่ในรูปของข้อมูลต้นฉบับและอยู่ในปริภูมิเวกเตอร์ สูตรการหาระยะทางแบบ Camberra จะให้ค่าเอาต์พุตสองค่า ได้แก่ TRUE และ FALSE

$$d_{ij} = \sum_{k=1}^n \frac{|x_{ik} - x_{jk}|}{|x_{ik}| + |x_{jk}|}$$

2.4.3 Chebyshev distance คือ สูตรการหาระยะทางแบบตารางหมากรุก ซึ่งอธิบายได้ว่าเป็นการใช้ค่าสัมบูรณ์ ของความแตกต่างระหว่างค่าในแนวพิกัดในมิติที่มีความแตกต่างกันมากที่สุดมาเป็นคำตอบ เช่นเดียวกับการเดินของตัวกิงบนหมากรุกฝรั่ง แสดงดังสมการที่ 3

$$d(x, V) = \max(x, -J,)$$

2.4.4 Dynamic Time Warping Distance คือ ขั้นตอนวิธีสำหรับการเปรียบเทียบความคล้ายของลำดับที่มีความแตกต่างกันในด้านเวลาหรือความเร็ว เช่น รูปแบบการเดินของคนๆหนึ่งจะถูกนับว่ามีความคล้าย ไม่ว่าคนๆนั้นจะเดินอย่างรวดเร็ว เดินอย่างเชื่องช้า หรือแม้แต่เดินด้วยความเร่ง เมื่อพิจารณาจากผู้สังเกตเดียวกัน ซึ่งไดนามิกไทม์วอร์ปิงสามารถนำไปประยุกต์ใช้กับวิดีโอ เสียง และภาพ รวมไปถึงข้อมูลต่างๆที่สามารถแปลงให้อยู่ในรูปของข้อมูลเชิงเส้นได้

2.4.5 Manhattan Distance คือ สูตรการคำนวณแบบความห่างกันของสิ่งปลูกสร้างว่าห่างกันกี่บล็อกในเมืองอย่างแมนฮัตตัน ซึ่งเป็นเขตพื้นที่ในเมืองที่พื้นที่ถูกแบ่งออกเป็นบล็อกสี่เหลี่ยมผืนผ้าขนาดเท่าๆกัน เรียงต่อกันไปเรื่อยๆทั่วทั้งเมือง เช่น ถ้าสถานที่แห่งหนึ่งอยู่ห่างจากที่ ๆ เราสนใจไปทางตะวันออก 3 บล็อกและต้องขึ้นไปทางเหนืออีก 4 บล็อก ก็จะได้ว่าสถานที่แห่งนั้นอยู่ออกจากที่แห่งนี้ออกไปเป็นระยะทาง= 3+4=7 บล็อก

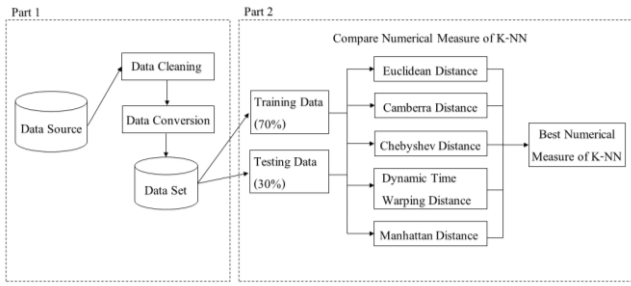
$$d(x, y) = \sum_{i=1}^N |x_i - y_i|$$

2.5 Rapid Miner [5]

Rapidminer คือซอฟต์แวร์ Data Science ใช้สำหรับการเตรียมข้อมูล การเรียนรู้เครื่อง การเรียนรู้ลึก การทำเหมืองข้อความ และการวิเคราะห์การทำนาย (Predictive analysis) เป็นซอฟต์แวร์ที่ช่วยในการจัดส่งข้อมูล และลดข้อผิดพลาดจนแทบจะไม่จำเป็นต้องเขียนโค้ดเพิ่ม แต่ที่ทำให้เป็นเครื่องมือที่เหล่า Data Scientist นิยมเลือกใช้เป็นเพราะว่าตัว Rapidminer มีขั้นตอนพร้อมสำหรับการทำ Data mining (ขุดข้อมูล) และ Machine learning ซึ่งรวมไปถึงการโหลดและการแปลงข้อมูล (ETL) การประมวลผลล่วงหน้าและการวาดภาพจากข้อมูล การวิเคราะห์เชิงพยากรณ์และการสร้างแบบจำลองทางสถิติ การประเมินผลและการปรับใช้ต่างๆ ส่วนเป็นสิ่งที่เหล่า Data Scientist จำเป็นต้องทำในการเข้าใจข้อมูลมากขึ้น

3. วิธีการดำเนินงาน

การดำเนินงานวิจัยเพื่อแนะนำสถานที่ท่องเที่ยว ภาคเหนือประเทศไทย ผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 2



ภาพที่ 2: กรอบแนวคิด

จากภาพที่ 2 กรอบแนวความคิด สามารถแบ่งขั้นตอนการดำเนินการวิจัยเป็น 2 ส่วน ดังนี้

3.1 ขั้นตอนการรวบรวม และเตรียมข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยได้ทำการศึกษา และรวบรวมข้อมูลจากเว็บท่องเที่ยวต่างๆบน www.google.com โดยนำข้อมูลมาวิเคราะห์ถึงที่มาที่ไปและความน่าเชื่อถือ เปรียบเทียบข้อมูลของสถานที่ท่องเที่ยวจากหลายๆแหล่ง และนำข้อมูลที่ได้มาแปลงให้เหมาะสมกับอัลกอริทึมที่จะใช้ ประกอบด้วยคุณลักษณะจำนวน 4 คุณลักษณะ แสดงดังตารางที่ 1

ตารางที่ 1: คุณลักษณะของข้อมูล

ลำดับ	คุณลักษณะ
1	Weather
2	Activity
3	Rank
4	Accommodation

ผู้วิจัยได้ทำการคัดเลือกคุณลักษณะของข้อมูลที่จะใช้ในการวิเคราะห์ทั้งหมด 4 คุณลักษณะ ได้แก่ สภาพอากาศ (Weather), กิจกรรมของการท่องเที่ยว (Activity), ความนิยมของสถานที่ท่องเที่ยว (Rank) และ ราคาที่พักต่อคืน โดยประมาณ (Accommodation) หลังจากได้ข้อมูลของสถานที่ท่องเที่ยวต่างๆแล้ว ก็จะนำข้อมูลไปแปลงจากตัวอักษรเป็นตัวเลข เนื่องจากอัลกอริทึมที่ใช้จำเป็นต้องใช้ชุดข้อมูลที่เป็นตัวเลขในการคำนวณหาระยะทางที่ใกล้ที่สุด จากนั้นเตรียมเข้าสู่ขั้นตอนการคัดเลือก และทดสอบประสิทธิภาพอัลกอริทึมเพื่อสร้างแบบจำลองต่อไป

3.2 ขั้นตอนการคัดเลือก และทดสอบประสิทธิภาพอัลกอริทึม

นำข้อมูลที่ได้จากการเตรียมข้อมูลมาทำการแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูลสำหรับสร้างแบบจำลอง ร้อยละ 70 จำนวน 350 รายการ และข้อมูลสำหรับทดสอบแบบจำลอง ร้อยละ 30 จำนวน 150 รายการ ซึ่งอัลกอริทึมที่จะใช้คือ K-Nearest Neighbors และเปรียบเทียบกับสูตรของอัลกอริทึมทั้งหมด 5 สูตร เพื่อคัดเลือกสูตรที่ดีที่สุด ซึ่งแต่ละสูตรจะกำหนดพารามิเตอร์ที่เหมือนกัน และทำการวัดประสิทธิภาพของอัลกอริทึมจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าระลึก (Recall) โดยพารามิเตอร์สำหรับการทดสอบอัลกอริทึมคือ K-Nearest Neighbors ได้แก่ K, Measure Types, Numerical Measure

4. ผลการดำเนินงาน

ผลการทดสอบประสิทธิภาพอัลกอริทึม แสดงดังตารางที่ 2 โดยมีรายละเอียดดังนี้

ตารางที่ 3: ผลการทดสอบประสิทธิภาพอัลกอริทึม K-NN

สูตรการหา ระยะทาง	ค่าความถูกต้อง (Accuracy)				
	K=3	K=5	K=7	K=9	K=11
Euclidean Distance	78.34%	65.91%	59.60%	57.57%	52.38%
Canberra Distance	6.08%	8.35%	9.03%	7.67%	7.40%
Chebyshev Distance	65.24%	57.81%	53.28%	50.81%	46.74%
Dynamic Time Warping Distance	77.66%	66.82%	62.09%	56.45%	53.06%
Manhattan Distance	85.57%	79.03%	72.03%	68.65%	61.42%

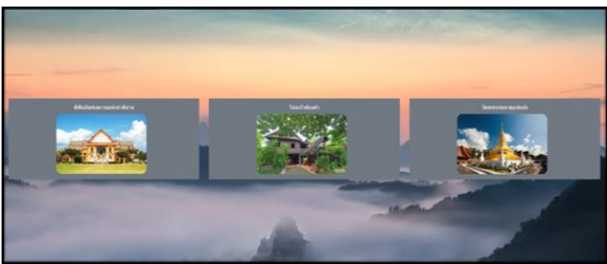
จากตารางที่ 3 ค่าพารามิเตอร์ k เท่ากับ 3 Measure Types เท่ากับ Numerical Measure และ Numerical Measure เท่ากับ Manhattan Distance ให้ประสิทธิภาพดีที่สุด โดยให้ค่าความถูกต้องเท่ากับร้อยละ 85.57 โดย Manhattan Distance

เมื่อได้ Model ในการจำแนกหาสถานที่ท่องเที่ยวแล้ว ผู้วิจัยได้ทำการพัฒนาเว็บไซต์เพื่อให้ผู้ใช้งานเข้าไปกรอกข้อมูลการ

ท่องเที่ยวที่สนใจ เพื่อให้ระบบแนะนำสถานที่ท่องเที่ยวที่
เหมาะสมกับผู้ใช้งานมากที่สุด ตามภาพที่ 3-5



ภาพที่ 3: ตัวอย่างหน้าจอกรอกความสนใจในการท่องเที่ยวของผู้ใช้งาน



ภาพที่ 4: หน้าจอแสดงสถานที่ท่องเที่ยวที่ระบบแนะนำ 3 อันดับแรก



ภาพที่ 5: ตัวอย่างหน้าจอแสดงสถานที่ท่องเที่ยว

5. อภิปรายผลและข้อเสนอแนะ

5.1 อภิปรายผล

งานวิจัยนี้นำเสนอการสร้างแบบจำลองการระบบแนะนำ
สถานที่ท่องเที่ยว ภาคเหนือ ประเทศไทย โดยใช้อัลกอริทึม
เพื่อนบ้าน ใกล้ที่สุด และเปรียบเทียบสูตรการหาระยะทาง
ทั้งหมด 5 สูตร ได้แก่ Euclidean Distance, Canberra Distance,
Chebyshev Distance, Dynamic Time Warping Distance และ
Manhattan Distance โดยใช้ข้อมูลจากเว็บท่องเที่ยวต่างๆบน
อินเทอร์เน็ต ประกอบด้วยคุณลักษณะ จำนวน 4 คุณลักษณะ
ทั้งหมด 500 รายการ จากนั้นแบ่งข้อมูลเป็น 2 ส่วน คือ ข้อมูล
สำหรับสร้างแบบจำลองจำนวน 350 รายการ และข้อมูลสำหรับ

ทดสอบแบบจำลองจำนวน 150 รายการ โดยนำข้อมูลมาทำการ
ทดสอบเปรียบเทียบประสิทธิภาพสำหรับการจำแนกสถานที่
ท่องเที่ยว ผลการวิจัยพบว่า ผลการเปรียบเทียบสูตรการหา
ระยะทางของอัลกอริทึม K-Nearest Neighbor ที่ใช้ในการ
จำแนกสถานที่ท่องเที่ยวภายในภาคเหนือ ประเทศไทย สรุป
ผลได้ว่า สรุปว่าการทดสอบอัลกอริทึม K-Nearest Neighbor
ปรับค่าพารามิเตอร์ k เท่ากับ 3 measure types ที่
NumericalMeasures และ ManhattanDistance ให้ประสิทธิภาพ
ที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากับ 85.57% ซึ่งสอดคล้องกับ
งานวิจัยที่ใช้อัลกอริทึมดังกล่าวมาทำการจำแนก
เช่น เกรียงศักดิ์, ปราโมทย์, ไมตรี และ วชิระ [1]
ระบบสนับสนุนการตัดสินใจเพื่อการเดินทางท่องเที่ยว ใน
จังหวัดอุบลราชธานีผ่านระบบดาวเทียมบอกพิกัด
และฉิษาภาพ จงกะสิกิจ [4] ระบบสนับสนุนการตัดสินใจ
ในการวิเคราะห์ความเสี่ยงของการเกิดโรคเรื้อรัง ภูมิของ
โรคเบาหวานและความดันโลหิตสูง เป็นต้น

5.2 ข้อเสนอแนะ

ข้อมูลสถานที่ท่องเที่ยวของผู้ใช้นั้นมีเยอะ จึงทำให้ Label
หรือคำตอบนั้นเยอะไปด้วย ทำให้อัลกอริทึมจะมีความสับสน
เนื่องจากสถานที่ท่องเที่ยวในภาคเหนือนั้นมีลักษณะคล้ายๆกัน
ทำให้ค่าระยะทางของสถานที่ต่างๆ นั้นเท่ากันหลายๆ ทำให้
Model ที่ใช้ในครั้งนี้มีค่าความถูกต้องสูง แต่ค่าระยะ และค่า
ความแม่นยำต่ำ

เอกสารอ้างอิง

- [1] เกรียงศักดิ์รักรักดี, ปราโมทย์นามวงศ์, ไมตรีริมทอง
และ วชิระ โมราชาติ. (2560). [ออนไลน์].
ระบบสนับสนุนการตัดสินใจเพื่อการเดินทางท่องเที่ยว
ในจังหวัดอุบลราชธานีผ่านระบบดาวเทียมบอกพิกัด.
[สืบค้นวันที่ 13 กุมภาพันธ์ 2566] จาก
<http://www.bba.ubru.ac.th>.
- [2] ชาลี ดวงแยม, ไกรรุ่ง เสงพระพรหม และสุพจน์
สงพระพรหม. (2550).“ระบบสนับสนุนการตัดสินใจเลือก

ประกอบอาชีพสำหรับผู้สำเร็จการศึกษาระดับปริญญาตรี
ด้วยเทคนิค Naive Bayes.” [http://e-
book.ru.ac.th/dss/files/21435149622_AbtractBeer.pdf](http://e-book.ru.ac.th/dss/files/21435149622_AbtractBeer.pdf).

- [3] Matana Wiboonyasake. (2564). [ออนไลน์]. Machine Learning คืออะไร [สืบค้นวันที่ 13 กุมภาพันธ์ 2566] จาก <https://www.aware.co.th/machine-learning-คืออะไร>
- [4] Nati Thaiyathum. (2562). [ออนไลน์]. KNN หรือ K-Nearest Neighbors คืออะไร [สืบค้นวันที่ 13 กุมภาพันธ์ 2566] จาก <https://www.glurgeek.com/education/knn/>
- [5] Achieve Plus. (2563). [ออนไลน์]. Rapidminer เสกคนไม่มีพื้นฐานให้เป็นเซียน [สืบค้นวันที่ 13 กุมภาพันธ์ 2566] จาก <https://medium.com>

ระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์

Development of Environmental Quality Measurement Systems via Monitoring and Evaluating with Real-time Notification on Application

มาริสสา แสงนวล (Marisa Sangnul)¹ มหศักดิ์ เกตุน้ำ (Mahasak Ketcham)²
 ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021857123@email.kmutnb.ac.th, ²mahasak.k@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ ประกอบไปด้วย การวัดฝุ่น PM (PM1, PM2.5, PM10) การวัดความดันน้ำ การวัดระดับน้ำ และแสดงพิกัดตำแหน่งของอุปกรณ์ ระบบนี้เป็นการพัฒนาอุปกรณ์การวัดค่าสภาพแวดล้อม โดยใช้ IOT ซึ่งเป็นอุปกรณ์อิเล็กทรอนิกส์ที่เชื่อมโยงกันและส่งค่าที่วัดได้ผ่านระบบคลาวด์เชื่อมต่อกับ Web Blynk และ Application Blynk เป็นแพลตฟอร์มที่ใช้งานได้ทั้ง iOS และ Android และการออกแบบการแสดงผล (Dashboard) บนหน้าเว็บและแอปพลิเคชัน ออกแบบแต่ละส่วนให้เหมาะสมกับค่าสภาพแวดล้อม รวมถึงมีการออกแบบการแสดงผลให้อยู่ในหน้าจอเดียวกัน เพื่อความสะดวกในการใช้งานของผู้ใช้ ในส่วนของการพัฒนาผู้วิจัยได้ใช้ภาษา C โดยพัฒนาบนโปรแกรม Arduino Software (IDE)

คำสำคัญ: ปัญหาสิ่งแวดล้อม การตรวจค่าสภาพแวดล้อม แอปพลิเคชันบนมือถือ

Abstract

The purpose of this research is to develop an environmental monitoring system, notification via real-time application, consisting of PM dust measurement (PM1, PM2.5, PM10), water pressure measurement, water level measurement and displaying the location coordinates of the device. This system is the development of

environmental measurement devices using IOT, which is an electronic device that connects and transmits measured values through the cloud connected to Web Blynk and Application Blynk is a platform that works on both iOS and Android, and display design (Dashboard). On web pages and applications, each design is suitable for the environment. Including designing the display to be on the same screen for the convenience of user use. In part of the development, the researcher used the C language to develop the Arduino Software (IDE) program.

Keywords: Environmental Issues, Environment Check Mobile Application.

1. บทนำ

ปัจจุบันสิ่งแวดล้อมโลกมีแนวโน้มถูกทำลายมากขึ้น ในขณะเดียวกันสิ่งแวดล้อมทางวัฒนธรรมกลับเพิ่มมากขึ้นเป็นลำดับเนื่องจากจำนวนประชากรเพิ่มขึ้นอย่างรวดเร็ว มีการประดิษฐ์และพัฒนาเทคโนโลยีมาใช้อำนวยความสะดวกต่อมนุษย์เพิ่มขึ้น ผลจากการทำลายสิ่งแวดล้อมทางธรรมชาติส่งผลกระทบต่อประชากร เช่น ปัญหาการแปรปรวน ของภูมิอากาศโลก กภัยพิบัติทางธรรมชาติมีแนวโน้มรุนแรงมากขึ้น มลพิษทางสิ่งแวดล้อมขยายขอบเขตกว้างขวางมากขึ้น ส่งผลกระทบต่อตรงต่อการดำรงอยู่ละการมีคุณภาพชีวิตที่ดีของมนุษย์

จากการตรวจวัดคุณภาพอากาศในเขตเมือง พบว่ามีจำนวนวันที่มลพิษทางอากาศเกินค่ามาตรฐานที่จังหวัด สมุทรปราการ (97 วัน) และกรุงเทพมหานคร (85 วัน) โดยสารมลพิษที่เป็นปัญหาสำคัญที่พบว่ามีเกิน 2 เกณฑ์มาตรฐาน คือ ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน (PM10) [1] ฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน (PM2.5) สำหรับสารอินทรีย์ระเหยง่ายสาเหตุสำคัญมาจากการพัฒนาอุตสาหกรรมในประเทศ โดยเฉพาะในพื้นที่กรุงเทพมหานครและปริมณฑล

จากประเด็นปัญหาดังกล่าว จึงทำให้เกิดแนวความคิดพัฒนาระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชัน ซึ่งสามารถดูค่าฝุ่น PM 8 ความดันน้ำ ระดับน้ำ และพิกัด เพื่อช่วยให้สามารถทราบสภาพอากาศที่อยู่รอบตัวเราได้โดยง่าย โดยการส่งค่า เมื่ออุปกรณ์วัดค่าสภาพแวดล้อม ตรวจวัดค่าได้ก็จะส่งข้อมูลทั้งหมดนี้โดยจะอาศัย Wi-Fi เป็นตัวกลางในการถ่ายโอนข้อมูล ซึ่งจะเป็นข้อมูลแบบ real-time ไปยัง web Blynk และ Application Blynk ตัวอุปกรณ์ยังมีขนาดเล็กสามารถพกพา ได้สะดวก เหมาะกับการพกติดตัวไว้เพื่อให้ทราบว่า บริเวณที่เราอยู่มีค่าฝุ่น และระดับน้ำ เป็นอย่างไร ทำให้สามารถรับมือหรือป้องกันได้ทัน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 นายพจนจสานต์ เกียรติวิชพันธุ์ [2] พัฒนาอุปกรณ์ตรวจวัดสภาพแวดล้อม ใช้เซนเซอร์ในการตรวจวัดที่หลากหลายด้าน ทั้ง เซนเซอร์วัดปริมาณแสง วัดปริมาณน้ำฝน ตรวจจับแรงสั่นสะเทือน วัดอุณหภูมิวัดความชื้นในอากาศ ตรวจวัดฝุ่นละออง เซนเซอร์วัดระยะห่าง เซนเซอร์ตรวจจับการเคลื่อนไหว และเซนเซอร์วัดความชื้น ในดิน ควบคุมโดยอุปกรณ์ Microcontroller บอร์ด Node32S ซึ่งเป็นชิปประมวลผลแบบ ESP32 และ NodeMCU ซึ่งเป็นชิปประมวลผลแบบ ESP8266 หลังจากนั้นนำข้อมูลที่ได้มาประมวลผล และ ส่งไปยังเครื่องแม่ข่ายผ่าน

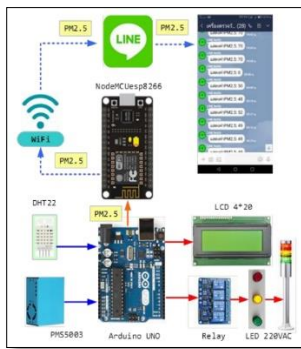
Application Blynk ด้วย ระบบ Internet ซึ่งใช้วิธีการเชื่อมต่อกับระบบ Internet โดยผ่านเครือข่ายอินเทอร์เน็ตไร้สายหรือ Wireless LAN

2.2 จิรวินัยดีเจริญชีพพงศ์, พลวัฒน์ ต่อตระกูล [3] จิรวินัยดีเจริญชีพพงศ์ , พลวัฒน์ ต่อตระกูล พัฒนาระบบตรวจจับและแจ้งเตือนฝุ่นละอองโดยใช้เทคโนโลยีอินเทอร์เน็ตในทุกสรรพสิ่ง มีการจัดเก็บข้อมูลบนระบบ เครือข่ายแบบกลุ่มเมฆ และแสดงคุณภาพของอากาศ PM2.5 โดยการพัฒนาระบบจะใช้ Arduino IDE ในการพัฒนาต้นแบบ หลังจาก การพัฒนาต้นแบบเสร็จได้มีการประเมินประสิทธิภาพของระบบด้วยวิธี Block Box Testing โดยผู้เชี่ยวชาญ 5 ท่าน ผู้ใช้งานทั่วไป จำนวน 31 ท่าน การประเมินผลแบ่งออกเป็น 4 ด้าน คือ ด้านความต้องการของผู้ใช้ ด้านหน้าที่ของระบบ ด้านการใช้งานระบบ และ ด้านความปลอดภัย โดยพบว่าหลังจากการใช้งานระบบ โดยผู้เชี่ยวชาญ มีค่าเฉลี่ย 3.78 และส่วนเบี่ยงเบนมาตรฐาน 0.33

2.3 Popoola et. Al. [4] ได้ศึกษาการใช้เครือข่ายของเซนเซอร์ต้นทุนต่ำเพื่อวัดคุณภาพอากาศ ศึกษาและวิเคราะห์ปริมาณฝุ่น ในอากาศด้วยเซนเซอร์ตรวจวัดฝุ่นละออง โดยทำการติดตั้งเซนเซอร์ต้นทุนต่ำสำหรับตรวจวัดมลพิษโดยรอบสนามบิน London Heathrow Airport (LHR) เพื่อแยกแยะการปล่อยมลพิษ บริเวณนั้นและจากนั้นก็นำข้อมูลที่นำมาทำแบบจำลองคุณภาพอากาศ และมีการทำนายผลกระทบในอนาคต

2.4 ดอนสัน ปงผาบ , ปกรณ์ สันตกิจ [5] นำเสนอการสร้างเครื่องวัดฝุ่น PM 2.5 และมีการแจ้งเตือนผลผ่านทางแอปพลิเคชันไลน์ หลักการทำงานของเครื่องจะใช้ไมโครคอนโทรลเลอร์ Arduino UNO เป็นตัวควบคุมการอ่านค่าจากเซนเซอร์ ตรวจวัดปริมาณฝุ่น PM 2.5 ค่าอุณหภูมิและความชื้นในอากาศ ในส่วนของการแจ้งเตือนจะแสดงบนจอแสดงผล และแจ้งเตือนเป็นสถานะไฟกระพริบสีเขียว สีเหลืองและสี

แดง เพื่อให้สามารถแจ้งเตือนและมองเห็นในระยะไกลได้



ภาพที่ 1: การทำงานของเครื่องวัดฝุ่นและแจ้งเตือนทางไลน์

2.5 สิทธิโชค สีนคู่ย [6] พัฒนาระบบการพัฒนา ระบบเฝ้าระวังฝุ่นละอองขนาดเล็กและคุณภาพอากาศ โดยใช้เทคโนโลยี IoT ร่วมกับระบบสารสนเทศภูมิศาสตร์ พัฒนาระบบฐานข้อมูลที่สามารถเก็บข้อมูล ปริมาณฝุ่นละอองได้แบบเรียลไทม์ และพัฒนาระบบแจ้งเตือนคุณภาพอากาศด้วยการใช้คำสั่งบน Arduino IDE ให้เมื่อปริมาณฝุ่นละอองขนาดเล็กในอากาศมี ปริมาณเกินกว่าที่กำหนด

2.6 ปรัชญา ใจสุทธี, ขนิษฐา สิทธิเทียมจันทร์, ทวีศักดิ์ สัมมา, วสุพล เผือก [7] พัฒนาระบบตรวจวัด ระดับน้ำและเตือนภัยน้ำท่วม โดยใช้เซ็นเซอร์ตรวจวัดระดับน้ำ แล้วส่งข้อมูลผ่านทางเครือข่ายไร้สาย เพื่อบันทึกข้อมูลสถิติลงสู่เซิร์ฟเวอร์แล้วแจ้งเตือนผ่านทางระบบบริการไลน์นอทีฟและแสดงผลลงบนแผนที่กูเกิ้ล ซึ่งเว็บแอปพลิเคชันที่พัฒนาขึ้นสามารถแสดงผลค่าระดับน้ำที่ตรวจวัดได้โดยเทียบกับระดับน้ำทะเลปานกลาง และแสดงพื้นที่บริเวณที่มีโอกาสน้ำท่วม นอกจากนี้ ยังเก็บข้อมูลทั้งแบบทันทีทันใดและแบบสถิติเป็นรายชั่วโมง รายวัน รายเดือน และรายปี

2.7 ไพบุลย์ บุญผา, ณฤทธิ ญิงชนิสรา, ชีระ ฤกษ์ม Iวรรณ, ทรงพล รอดทอง [11] พัฒนาระบบเครื่องวัด และแสดงผลระดับน้ำแบบไร้สาย ทำงานด้วยคลื่นอัลตราโซนิกความถี่ 40 กิโลเฮิร์ตซ์เพื่อส่งสัญญาณอัลตราโซนิกไปยังผิวน้ำเพื่อวัดระดับน้ำและสะท้อนสัญญาณกลับมายัง

ไมโครคอนโทรลเลอร์เพื่อประมวลผลและส่งข้อมูลผ่านคลื่นวิทยุแบบไร้สายความถี่ 915 เมกะเฮิร์ตซ์ไปแสดงผลที่ชุดแสดงผลแบบแอลอีดีจากการทดสอบพบว่าให้ผลการวัดที่ถูกต้องตามวัตถุประสงค์ที่กำหนดไว้

3. วิธีดำเนินงาน

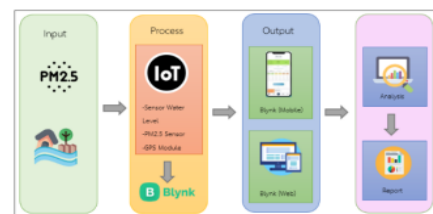
จากการศึกษาทฤษฎีต่างๆที่เกี่ยวข้องของกับงานวิจัยนั้น ผู้วิจัยได้นำเอาองค์ความรู้ต่าง ๆ มาใช้ในการพัฒนาระบบตรวจวัดค่าสภาพแวดล้อมอัจฉริยะ โดยกำหนดขั้นตอนในการดำเนินการวิจัยดังนี้

3.1 การศึกษาและรวบรวมข้อมูล

การศึกษาความต้องการของระบบตรวจวัดค่าสภาพแวดล้อมอัจฉริยะ มุ่งเน้นเรื่องการรับค่าจากอุปกรณ์เซ็นเซอร์ ที่เกี่ยวข้องกับการตรวจวัดค่าสภาพแวดล้อมต่าง ๆ และนำค่าที่ส่งผ่านระบบ Internet มายัง Application Blynk เพื่อจัดเก็บข้อมูล ซึ่งการพัฒนาระบบตรวจวัดค่าสภาพแวดล้อมจำเป็นต้องใช้บอร์ด NodeMCU8266 ซึ่งนิยมใช้ในการอ่านค่าเซ็นเซอร์ต่าง ๆ เนื่องจากราคาถูกและมีประสิทธิภาพ

3.2 วิเคราะห์และออกแบบระบบ

จากการศึกษาความต้องการของอุปกรณ์ ผู้วิจัยได้ค้นคว้าและรวบรวมข้อมูลด้านต่าง ๆ เพื่อนำมาพัฒนาเป็นระบบตรวจวัดค่าสภาพแวดล้อมอัจฉริยะ โดยมีกรอบแนวคิดดังนี้



ภาพที่ 2: กรอบแนวคิดของระบบตรวจวัดค่าสภาพแวดล้อมแจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์

จากกรอบแนวคิดระบบระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ โดยกรอบแรกจะเป็นในส่วนของ Input คือค่าฝุ่น PM และน้ำตามแม่น้ำลำคลอง ถัดมาในกรอบ

ของ Process เป็นการพัฒนาอุปกรณ์โดยใช้เซ็นเซอร์เชื่อมกับ Blynk การทำงานของตัวอุปกรณ์จะทำการวัดค่าฝุ่น PM ได้แก่ PM1, PM2.5 PM10 ค่าความดันน้ำ และระดับน้ำ ถัดมาเป็นกรอบของ Output เป็นการแสดงค่าสภาพแวดล้อมผ่านทาง Web และ Mobile โดยที่สถานีจะเป็น Dashboard ของข้อมูลสภาพอากาศที่วัดได้นำมาแสดงให้ผู้ใช้ และในส่วน of Web ยังสามารถ Export file ออกมาเพื่อนำข้อมูลมาวิเคราะห์วางแผนสภาพแวดล้อม

3.3 ออกแบบและพัฒนาระบบ

จากการศึกษาความต้องการของอุปกรณ์ ผู้วิจัยได้ค้นคว้าและรวบรวมข้อมูลด้านต่าง ๆ เพื่อนำมาพัฒนาเป็นระบบตรวจวัดค่าสภาพแวดล้อมอัจฉริยะ โดยผู้วิจัยได้แบ่งขั้นตอนในการออกแบบและพัฒนาระบบ ดังนี้

3.3.1 การเลือกใช้งานฮาร์ดแวร์ Microcontroller บอร์ด และอุปกรณ์เซ็นเซอร์ โดยเลือกใช้งานอุปกรณ์ต่าง ๆ ในระบบตรวจวัดค่าสภาพแวดล้อมอัจฉริยะ ดังนี้

3.3.1.1 Microcontroller Board รุ่น NodeMCUesp8266

3.3.1.2 Sensor Water Level

3.3.1.3 PM2.5 Sensor

3.3.1.4 GPS Module GY-NEO6MV2
Ublox

3.3.2 ซอฟต์แวร์ที่ใช้ในการเขียนโปรแกรมเพื่อเชื่อมต่อบอร์ดเข้ากับเซ็นเซอร์

3.3.2.1 Arduino IDE

3.3.2.2 ภาษา C/C++

3.3.2.3 Application Blynk

3.3.3 ออกแบบการเชื่อมต่อและการส่งค่าจากเซ็นเซอร์ ผู้วิจัยได้เลือกใช้ Application Blynk เป็นแอปพลิเคชันที่มีทั้งใน iOS และ Android เพื่อควบคุม ESP8266, NodeMCU, Arduino (any model), Raspberry (any model), Particle (any model) บนระบบอินเทอร์เน็ต ซึ่งสามารถรับค่าต่าง ๆ

จากเซ็นเซอร์เพื่อนำไปสร้างเป็นกราฟฟิก และสามารถบันทึกเป็นไฟล์ CSV เพื่อนำค่าที่ได้ไปประมวลผลต่อไปได้

3.4 จัดเก็บข้อมูลที่ได้จากอุปกรณ์

ผู้วิจัยได้พัฒนาอุปกรณ์ระบบโดยการใช้บอร์ด NodeMCUesp8266 เซ็นเซอร์ทั้งหมด 3 เซ็นเซอร์ ได้แก่ เซ็นเซอร์วัดฝุ่น PM เซ็นเซอร์วัดความดันน้ำ และเซ็นเซอร์วัดระดับน้ำ โดยใช้ภาษา C พัฒนาระบบโปรแกรม Arduino IDE ส่วนตัวจ่ายไฟให้กับอุปกรณ์จะใช้เป็นเพาเวอร์แบงค์ อุปกรณ์จะใช้เป็นกล่องอิเล็กทรอนิกส์ เป็นกล่องอเนกประสงค์ กันน้ำ เพื่อป้องกันไม่ให้น้ำไหลเข้าไปในวงจรด้านใน โดยนำอุปกรณ์ไปติดตั้งที่สถานที่จริงและเก็บข้อมูล

3.5 จัดเก็บข้อมูลที่ได้จากอุปกรณ์

การประเมินความพึงพอใจของผู้ใช้งานผู้วิจัยเลือกใช้แบบสอบถามเพื่อใช้ประเมินความพึงพอใจของระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ จากผู้ใช้งานจำนวน 15 คน ผู้ได้รับผลกระทบ ณ คลองบางเขนใหม่ จำนวน 5 ท่าน 5 คน ผู้ได้รับผลกระทบ ณ คลองวัดเสาหิน จำนวน 5 ท่าน ผู้ได้รับผลกระทบ ณ คลองบางซื่อ จำนวน 5 ท่าน มีหัวข้อดังนี้

3.5.1 ด้านที่ 1 ด้านการใช้งานของระบบ

3.5.2 ด้านที่ 2 ด้านการออกแบบ

3.5.3 ด้านที่ 3 ด้านความปลอดภัยของระบบ

3.5.4 ด้านที่ 4 ด้านประโยชน์ที่ได้รับ

สถิติที่ใช้ประเมินความพึงพอใจของผู้ใช้งานระบบที่พัฒนาขึ้น จะใช้ค่าตัวกลางเลขคณิตหรือค่าเฉลี่ยไปเทียบกับเกณฑ์การให้คะแนน และใช้ค่าส่วนเบี่ยงเบนมาตรฐานเพื่อหาความน่าเชื่อถือของข้อมูล ดังนี้

3.5.5 การหาค่าเฉลี่ย (Mean)

$$\bar{x} = \frac{\sum fx}{N}$$

3.5.6 การหาส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

$$S.D. = \frac{\sqrt{N \sum x^2 - (\sum x)^2}}{N(N-1)}$$

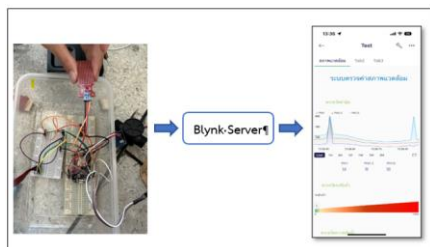
3.5.7 One-Sample Test (t-test)

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

4. ผลการดำเนินงานวิจัย

4.1. ผลการพัฒนาระบบด้านฮาร์ดแวร์

ระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ ผู้วิจัยเลือกใช้บอร์ดไมโครคอนโทรลเลอร์รุ่น NodeMCU ESP8266 และประกอบไปด้วย Sensor Water เพื่อวัดระดับน้ำ Level PM2.5 Sensor วัดฝุ่น PM โดยอุปกรณ์ที่กล่าวมาข้างต้นจะทำการวัดค่าสภาพแวดล้อมต่าง ๆ และส่งค่าผ่านแอปพลิเคชัน Blynk ซึ่งเชื่อมต่อกับ NodeMCU ESP8266 ผ่านระบบ Wi-Fi ดังภาพที่ 4-1 และระบบที่เสร็จสมบูรณ์ดังภาพที่ 4-2



ภาพที่ 3: แบบจำลองระบบวัดค่าสภาพแวดล้อมด้วย IoT



ภาพที่ 4: ระบบวัดค่าสภาพแวดล้อมแบบเสร็จสมบูรณ์

4.2 ผลการพัฒนาระบบด้านซอฟต์แวร์

หน้าจอ Application Blynk จะแสดงหน้า dashboard ได้แก่ ค่าฝุ่น PM จะใช้ dashboard แบบ

Super Chart PM ได้แก่ PM1 เส้นกราฟสีม่วง PM2.5 เส้นกราฟสีส้ม PM10 เส้นกราฟสีฟ้า ค่าระดับน้ำ โดยถ้าหากค่าต่ำกว่า 50 อยู่ที่สีเขียว สีส้ม สีแดง ตามลำดับ ตรวจวัดค่าความดันน้ำ จะแสดงตัวเลข 2 บรรทัด บรรทัดแรกคือ จำนวนรอบในการหมุนของรอบคลื่นใน 1 วินาที บรรทัดที่ 2 ตัวเลข LPM คือ หน่วยวัดที่ใช้สำหรับบอกปริมาณการไหลหรืออัตราการไหลของน้ำ และตำแหน่งของอุปกรณ์วัดค่าสภาพแวดล้อม โดยจะแสดงในรูปแบบที่ ดังภาพที่ 4-3



ภาพที่ 5: หน้าจอ Dashboard Blynk App

4.3 ผลการประเมินความพึงพอใจของระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์

ความพึงพอใจของผู้ใช้งานระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ของผู้ใช้ แสดงถึงความพึงพอใจของผู้ใช้งานระบบตรวจวัดค่าสภาพแวดล้อม แจ้งเตือนผ่านแอปพลิเคชันแบบเรียลไทม์ของจอแสดงผลพบว่ากลุ่มผู้ใช้งานมีความพึงพอใจต่อระบบนี้ ความพึงพอใจด้านอุปกรณ์ มีค่าเฉลี่ยรวม 4.70 ส่วนเบี่ยงเบนมาตรฐาน 0.36 อยู่ในระดับความพึงพอใจมากที่สุด ความพึงพอใจด้านการแสดงผล มีค่าเฉลี่ยรวม 4.93 ส่วนเบี่ยงเบนมาตรฐาน 0.25 อยู่ในระดับความพึงพอใจมากที่สุด

5. สรุปผลและอภิปราย

5.1 สรุปผลและอภิปรายผล

งานวิจัยนี้จัดทำการอุปกรณตรวจวัดค่าสภาพแวดล้อม ได้แก่ ค่าฝุ่น PM ค่าความดันน้ำ และค่าระดับน้ำ โดยมีการติดตั้งอุปกรณ์ ณ สถานที่จริง 3 สถานที่ ได้แก่ คลองบางเขนใหม่ คลองวัดเสนาหิน คลองบางซื่อ ผู้วิจัยได้นำค่าฝุ่น PM ทั้ง 3 สถานที่มาเปรียบเทียบกับค่าฝุ่นในเว็บไซต์ Air4Thai ซึ่งสรุปผลการวิจัยได้ดังนี้ คลองบางเขนใหม่ ทำการวัดสภาพแวดล้อมวันที่ 23 เดือนมีนาคม พ.ศ. 2566 ตั้งแต่เวลา 08.00 – 11.00 น. เป็นเวลา 3 ชั่วโมง โดยแบ่งเวลาออกเป็น 3 ช่วง ได้แก่ ช่วงที่ 1 08.00 – 09.00 น. ช่วงเวลาที่ 2 09.00 – 10.00 น. ช่วงเวลาที่ 3 10.00 – 11.00 น. โดยมีค่าความถูกต้องของช่วงเวลาทั้ง 3 เวลาอยู่ที่ 94.74% คลองวัดเสนาหิน ทำการวัดสภาพแวดล้อมวันที่ 24 เดือนมีนาคม พ.ศ. 2566 ตั้งแต่เวลา 08.00 – 11.00 น. เป็นเวลา 3 ชั่วโมง โดยแบ่งเวลาออกเป็น 3 ช่วง ได้แก่ ช่วงที่ 1 08.00 – 09.00 น. ช่วงเวลาที่ 2 09.00 – 10.00 น. ช่วงเวลาที่ 3 10.00 – 11.00 น. โดยมีค่าความถูกต้องของช่วงเวลาทั้ง 3 เวลาอยู่ที่ 96.03% คลองบางซื่อ ทำการวัดสภาพแวดล้อมวันที่ 25 เดือนมีนาคม พ.ศ. 2566 ตั้งแต่เวลา 08.00 – 11.00 น. เป็นเวลา 3 ชั่วโมง โดยแบ่งเวลาออกเป็น 3 ช่วง ได้แก่ ช่วงที่ 1 08.00 – 09.00 น. ช่วงเวลาที่ 2 09.00 – 10.00 น. ช่วงเวลาที่ 3 10.00 – 11.00 น. โดยมีค่าความถูกต้องของช่วงเวลาทั้ง 3 เวลาอยู่ที่ 94.52% เมื่อเปรียบเทียบกันแล้ว ค่าของอุปกรณ์มีความคลาดเคลื่อนจาก เว็บไซต์ Air4Thai เล็กน้อยเนื่องจากฝุ่นละอองลอยอยู่ในอากาศในขนาดไม่เท่ากันทำให้อุปกรณ์วัดค่าฝุ่นได้ตัวเลขที่ต่างกัน

เอกสารอ้างอิง

[1] วิระวรรณ ถิ่นยืนยง (2559). “สิ่งคุกคามทางสิ่งแวดล้อมและสุขภาพของผู้ปฏิบัติงานในรถโดยสารสาธารณะ” กรมอนามัย กระทรวงสาธารณสุข.

- [2] นายพุมจสามัค เกียรติวิชพันธุ์ (2562). “สถานีตรวจวัดสภาพแวดล้อมอัจฉริยะ” คณะเทคโนโลยีสารสนเทศมหาวิทยาลัยพระจอมเกล้าพระนครเหนือ.
- [3] จิรวิญญูดีเจริญชิตพงศ์1, พลวัฒน์ ต่อตระกูล (2563). “การพัฒนาต้นแบบระบบตรวจจับและแจ้งเตือนฝุ่นละอองโดยใช้เทคโนโลยีอินเทอร์เน็ตในทุกสรรพสิ่ง”วารสารวิชาการมหาวิทยาลัยธนบุรี ปีที่ 15 ฉบับที่ 2 : 46-57
- [4] Popoola, O. A., Carruthers, D., Lad, C., Bright, V. B., Mead, M. I., Stettler, M. E., & Jones, R. L. (2018). Use of networks of low cost air quality sensors to quantify air quality in urban settings. *Atmospheric*, 194, 58-70.
- [5] จิตอนสัน ปงผาบ1, ปกรณ์ สันตกิจ (2564). “เครื่องวัดฝุ่น PM2.5 แจ้งเตือนทางแอป environment พลิกชันไลน์” วารสารราชมงคลกรุงเทพ ปีที่ 3 ฉบับที่ 1: 34-43.
- [6] สิทธิโชค สินคู่ย (2562). “การพัฒนาระบบเฝ้าระวังฝุ่นละอองขนาดเล็กและคุณภาพอากาศโดยใช้เทคโนโลยี IoT” ร่วมกับระบบสารสนเทศภูมิศาสตร์ วิทยานิพนธ์ระดับปริญญาตรี เสนอภาควิชาทรัพยากรธรรมชาติและ.
- [7] ไพบุลย์ นุบผา, ณฤทธิ ญงชนิศรา, ชีระ ฤกษ์มวรรณ, ทรงพล รอดทอง. “เครื่องวัดและแสดงผลระดับน้ำ”. คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษมบัณฑิต.

การจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์โดยใช้เทคนิคเหมืองข้อความ

Classification of Broadband Network Devices

Using Text Mining Technique

ณัฐภัทร ศรีด้วง (Nattapat Sridoung)¹ และมหศักดิ์ เกตุคำ (Mahasak Ketcham)²
 ภาควิชาระบบสารสนเทศเพื่อการจัดการ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
 มหาวิทยาลัยมหาวิทาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021857093@email.kmutnb.ac.th, ²mahasak.k@itd.kmutnb.ac.th

บทคัดย่อ

ปัจจุบันธุรกิจบรอดแบนด์อินเทอร์เน็ตมีการแข่งขันสูง ผู้ให้บริการแต่ละรายต่างใช้งบประมาณจำนวนมากในการลงทุนเพื่อพัฒนาโครงข่ายให้สามารถตอบสนองความต้องการของผู้บริโภคและยังต้องแข่งขันด้านราคากับผู้บริการรายอื่น ๆ ทำให้ค่าบริการนั้นลดต่ำลงสวนทางกับงบประมาณที่ต้องใช้ในการลงทุนพัฒนาโครงข่ายธุรกิจบรอดแบนด์อินเทอร์เน็ตจึงต้องการการบริหารจัดการต้นทุนที่มีประสิทธิภาพเพื่อให้ธุรกิจสามารถมีกำไรในตลาดที่มีการแข่งขันสูงได้ งานวิจัยนี้เป็นการพัฒนาแบบจำลองสำหรับการจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์ โดยนำเทคนิคเหมืองข้อความมาประยุกต์ใช้กับรายการอุปกรณ์โครงข่ายบรอดแบนด์ของบริษัทที่ดำเนินธุรกิจบรอดแบนด์อินเทอร์เน็ตแห่งหนึ่งในประเทศไทย ในการสร้างเวกเตอร์คุณลักษณะจากคำอธิบายของรายการอุปกรณ์ จากนั้นใช้อัลกอริทึมจำแนกประเภทในการจำแนกอุปกรณ์ออกเป็น 3 ประเภทตามลำดับชั้นของโครงข่าย ได้แก่ Core, Access และ Last Mile โดยประเมินเปรียบเทียบประสิทธิภาพของอัลกอริทึม ได้แก่ Decision Tree, Naïve Bayes, Bayesian Network, K-Nearest Neighbor, Support Vector Machine และ Deep Neural Network พบว่าอัลกอริทึม Support Vector Machine ให้ผลการจำแนกที่ดีที่สุด โดยได้ค่าความถูกต้องเท่ากับ 90.35% จากนั้นผู้วิจัยนำผลลัพธ์ที่ได้มาแสดงในรูปแบบ Visualization ซึ่งสามารถนำไปใช้สนับสนุนการดำเนินงานธุรกิจในด้านต่าง ๆ ต่อไป

คำสำคัญ: การจำแนกข้อมูล เหมืองข้อความ โครงข่ายบรอดแบนด์ ระบบสนับสนุนการตัดสินใจ

Abstract

Broadband internet business is highly competitive. Each service provider spends a large amount of its budget investing in network development to meet the needs of consumers and has to compete on price with other operators. Broadband internet businesses require efficient cost management to remain profitable in this highly competitive market. This research develops a model for classifying broadband network equipment by applying the text mining technique to the equipment list obtained from one of the largest broadband network of a company that operates a broadband internet business in Thailand. The feature vector is generated from the description of the device manifest, Then a classification algorithm is used to classify devices into 3 categories according to the network hierarchy i.e., core, access and last mile. By comparing the efficiency of the algorithm, i.e., decision tree, naïve bayes, bayesian network, k-nearest neighbor, support vector machine, and deep neural network, it was found that the support vector machine algorithm gave the best classification results with 90.35% accuracy. Then, we presented the results in visualization to support business operations in various areas.

Keywords: Classification Model, Text Mining, Broadband Network, Decision Support System.

1. บทนำ

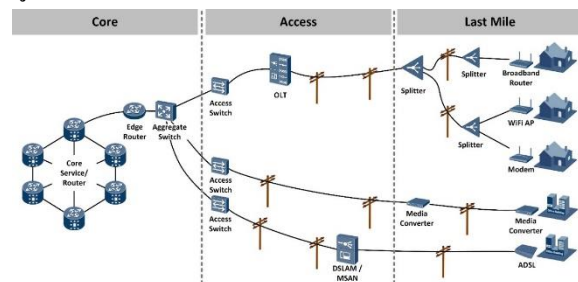
ปัจจุบันธุรกิจบรอดแบนด์เป็นธุรกิจที่มีการเติบโตแบบก้าวกระโดด ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) [1] ได้รวบรวมข้อมูล Internet Bandwidth ของประเทศไทย ระหว่างปี 2016-2020 พบว่า การเชื่อมต่อวงจรรีเตอร์เน็ต เติบโตขึ้นสู่ระดับ 14,144,148 Mbps คิดเป็นอัตราการเติบโตเฉลี่ยต่อปีร้อยละ 40.40 ซึ่งการเติบโตแบบก้าวกระโดดนี้ทำให้ผู้ให้บริการใช้เงินลงทุนพัฒนาโครงข่ายจำนวนมากเพื่อให้ทันต่อเทคโนโลยีที่มีการเปลี่ยนแปลงอย่างรวดเร็ว และเพื่อขยายโครงข่ายให้ครอบคลุมพื้นที่ในการให้บริการ มีการใช้กลยุทธ์ด้านการตลาดเพื่อจัด โปรโมชันต่าง ๆ ทำให้อัตรากำไรบริการในธุรกิจบรอดแบนด์ลดต่ำลงอย่างมีนัยสำคัญ ในช่วงระยะเวลาเพียงไม่กี่ปี [2] ดังนั้นธุรกิจบรอดแบนด์จึงต้องการการบริหารจัดการในด้านต้นทุนการให้บริการ เพื่อให้ธุรกิจสามารถมีกำไรในตลาดที่มีการแข่งขันสูงได้ โดยปัจจัยหลักที่ส่งผลต่อต้นทุนบริการบรอดแบนด์คือ อุปกรณ์ที่ใช้สำหรับการให้บริการ เนื่องจากมีราคาสูง เทคโนโลยีที่เปลี่ยนแปลงอย่างรวดเร็วทำให้ไม่สามารถใช้ทรัพยากรที่ลงทุนไปในอดีตในปัจจุบันได้ ความต้องการใช้งาน Bandwidth ที่เพิ่มขึ้นทำให้ต้องอัปเดตอุปกรณ์อยู่เป็นประจำ ไม่ว่าจะเป็นการขอทดแทนอุปกรณ์ที่ชำรุดเสียหาย หรือเพื่อรองรับผู้ใช้บริการที่เพิ่มขึ้น ในกระบวนการบันทึกการขายการอุปกรณ์ลงในฐานข้อมูลพนักงานจะต้องทราบว่าอุปกรณ์ดังกล่าวเป็นอุปกรณ์ที่โครงข่ายส่วนใดตามสถาปัตยกรรมโครงข่ายบรอดแบนด์ ซึ่งประกอบด้วย Core, Access และ Last Mile เพื่อให้หน่วยงานที่เกี่ยวข้องนำไปใช้เป็นข้อมูลประกอบการวางแผนลงทุน การวางแผนซ่อมบำรุง การบริหารต้นทุนการจัดสรรงบประมาณ และนำข้อมูลไปใช้ประโยชน์ในเชิงวิเคราะห์ด้านอื่น ๆ ต่อไป อย่างไรก็ตาม รายการอุปกรณ์ดังกล่าวมีจำนวนมาก และผู้ที่ให้ข้อมูลบางครั้งติดภาระกิจ จึงไม่สามารถให้ข้อมูลได้ในทันที ทำให้ข้อมูล

ขาดความครบถ้วน ไม่สามารถนำไปใช้ประโยชน์ในเชิงวิเคราะห์ตามที่กล่าวในข้างต้นได้ ผู้วิจัยจึงมีแนวคิดในการพัฒนาแบบจำลองจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์ โดยใช้เทคนิคการทำเหมืองข้อความร่วมกับอัลกอริทึมจำแนกประเภท เพื่อแบ่งเบาภาระของผู้ปฏิบัติงานและทำให้ข้อมูลอุปกรณ์มีความพร้อมในการนำไปวิเคราะห์ในมุมมองต่าง ๆ เพื่อเพิ่มขีดความสามารถในการแข่งขันในอุตสาหกรรม

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โครงข่ายสื่อสารบรอดแบนด์ [3]

Cisco ได้กำหนดแนวทางการออกแบบโครงข่ายแบบลำดับชั้น โดยจัดกลุ่มอุปกรณ์โครงข่ายเป็น 3 ประเภทตามภาพที่ 1 โดย Core เป็นส่วนที่เป็นแกนหลักของโครงข่าย ทำหน้าที่เป็น และ เชื่อมต่อกับ ชุมสายเคเบิ้ลภายในประเทศและต่างประเทศ, Access เป็นส่วนที่ทำหน้าที่ทำหน้าที่รวบรวมและกระจายปริมาณการใช้งานระหว่างโครงข่ายหลักกับโครงข่ายย่อยที่กระจายอยู่ตามพื้นที่ต่าง ๆ และ Last Mile เป็นส่วนที่อุปกรณ์ในระยะสุดท้ายของผู้ให้บริการเชื่อมต่อกับเครื่องคอมพิวเตอร์ของผู้ใช้งาน



ภาพที่ 1: การจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์

2.2 Text Processing Technique [4]

2.2.1 Stop Word Removal

เป็นการนำคำที่ไม่มีนัยสำคัญในการสื่อความหมายในเอกสารออก โดยที่ความหมายของคำหรือข้อความจะไม่มีการเปลี่ยนแปลง เพื่อลดขนาดของ Feature Vector

2.2.2 Stemming

เป็นการแปลงคำให้อยู่ในรูปแบบรากศัพท์ โดยแปลงคำที่อยู่ในรูปแบบ Suffix ให้เป็นคำจากรากศัพท์ เนื่องจากเป็นคำที่มีความหมายใกล้เคียงกัน

2.2.3 Tokenization

เป็นการแยกคำโดยใช้ช่องว่างระหว่างคำ (White Space) เพื่อสร้าง Feature Vector โดยแต่ละคำที่สร้างจากกระบวนการ Tokenize นั้นจะเป็น Attribute ของชุดข้อมูล

2.2.4 TF-IDF

เป็นการสร้าง Feature Vector จากการประเมินจากความสำคัญของคำที่ได้จากกระบวนการตัดคำในเอกสารออกมาเป็นค่าถ่วงน้ำหนัก โดยคำที่มีความถี่สูง (พบบ่อย) หรือคำที่มีทุกหมวด คำน้ำหนักคำนั้นจะต่ำหรือไม่มีนัยสำคัญ ส่วนคำที่พบเฉพาะในหมวดนั้น ๆ ไม่พบในหมวดอื่น จะมีคำน้ำหนักที่สูงหรือมีนัยสำคัญ ซึ่งคำน้ำหนักของ TF-IDF คำนวณจากสมการต่อไปนี้

$$\omega_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (2-1)$$

โดยที่ $tf_{i,j}$ คือ จำนวนคำ “i” ที่ปรากฏในเอกสาร j
 df_i คือ จำนวนเอกสารที่มีคำ “i”
 N คือ จำนวนเอกสารทั้งหมด

2.3 Support Vector Machine [5]

เป็นอัลกอริทึมที่ใช้ในการจำแนกข้อมูลที่อาศัยหลักการของการหาสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ดีที่สุดในช่วงขอบที่มากที่สุดระหว่างคลาสหนึ่งกับอีกคลาสหนึ่ง เป็นบรรทัดฐานในการตัดสินใจเลือกระนาบเงินที่เหมาะสมที่สุด ดังนั้นเมื่อกำหนด $\tau = \{(\vec{x}_i, y_i); i = 1, \dots, k; \vec{x}_i \in \mathbb{R}^n; y_i \in \{-1, +1\}\}$ เป็น Training Set โดยมี $\vec{w}^T \vec{x} + b$ เป็น Feature Vector จะได้ Hyperplane $\vec{w}^T \vec{x} + b = 0$ ที่สามารถแยก Training Set ได้จากเงื่อนไขดังต่อไปนี้

$$\begin{aligned} \vec{w}^T \vec{x}_i + b &\geq 0 & \text{for } y_i = +1 \\ \vec{w}^T \vec{x}_i + b &< 0 & \text{for } y_i = -1, \end{aligned}$$

เมื่อ \vec{w} ขนานกับ Hyperplane และ b เป็นระยะที่ตั้งฉากกับ Hyperplane ได้ฟังก์ชันตัดสินใจดังนี้

$$g(\vec{x}) = \vec{w}^T \vec{x}_i + b \quad (2-2)$$

2.3.1 Kernel Trick

Support Vector Machine มีข้อจำกัดคือสามารถสร้างเส้นแบ่งขอบเขตการตัดสินใจแบบเส้นตรงเท่านั้น ซึ่งอาจทำงานได้ไม่ดีหากความสัมพันธ์ของข้อมูลนั้นมีความซับซ้อน การแก้ปัญหาหนึ่งจะนำวิธีการของ Kernel มาใช้โดย Kernel Trick คือการดำเนินการทางคณิตศาสตร์ที่แปลง

ข้อมูลทดสอบให้อยู่ในมิติที่สูงขึ้นเพื่อให้ Support Vector Machine สามารถจำแนกค่าตัวแปรแบบ Polynomial ได้โดยไม่เปลี่ยนรูปแบบและความสัมพันธ์ของ Feature ดังเห็น แสดงในตารางที่ 1

ตารางที่ 1: Kernel Trick ของซัพพอร์ตเวกเตอร์แมชชีน

Kernel	Inner Product Kernel
	$K(\vec{x}, \vec{x}_i), i = 1, 2, \dots, N$
Linear Kernel	$K(\vec{x}, \vec{x}_i) = \vec{x}^T \vec{x}_i$
Second Polynomial	$K(\vec{x}, \vec{x}_i) = (\vec{x}^T \vec{x}_i + \theta)^2$
Third Polynomial	$K(\vec{x}, \vec{x}_i) = (\vec{x}^T \vec{x}_i + \theta)^3$
RBF Kernel	$K(\vec{x}, \vec{x}_i) = e^{-\gamma \ \vec{x} - \vec{x}_i\ ^2}$
Sigmoid Kernel	$K(\vec{x}, \vec{x}_i) = \tanh(\eta \vec{x} \vec{x}_i + \theta)$

2.4 งานวิจัยที่เกี่ยวข้อง

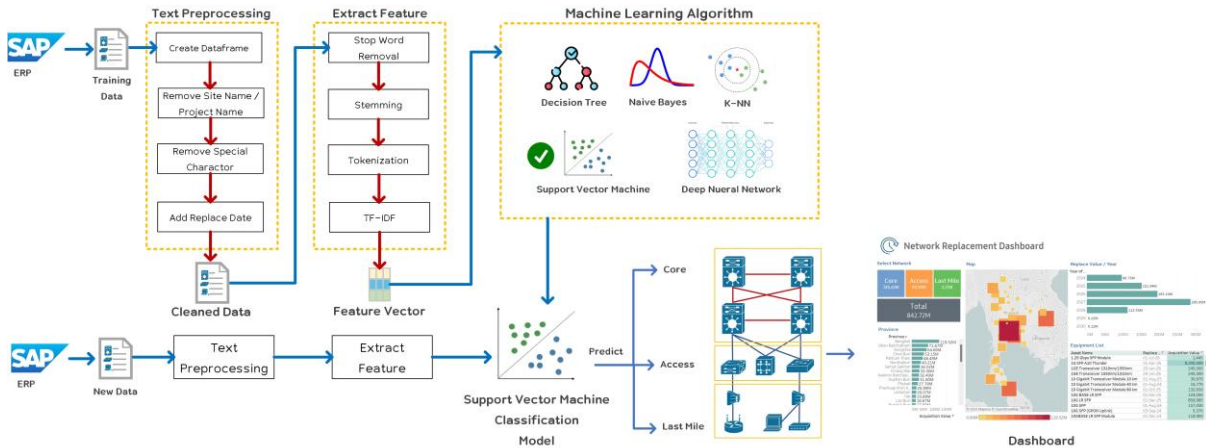
Gary M. Weiss [6] ได้ทำการศึกษาวิจัย เรื่อง Data Mining in Telecommunication โดยศึกษาข้อมูลต่าง ๆ ที่เกิดขึ้นในการดำเนินธุรกิจ เช่น ข้อมูลปริมาณการใช้งานโครงข่าย ข้อมูลส่วนประกอบของฮาร์ดแวร์และซอฟต์แวร์ในโครงข่าย โดยข้อมูลเหล่านี้สามารถนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ใช้เพื่อสกัดเอาข้อมูลที่เป็นประโยชน์มาใช้งานได้ ผลการวิจัย พบว่าเทคนิคเหมืองข้อมูลสามารถมาใช้ในการบริหารจัดการฮาร์ดแวร์และซอฟต์แวร์ในโครงข่าย ได้เป็นอย่างดี

Ramdasi, A.P. และ Sathyalakshmi, S. [7] ได้ทำการศึกษาวิจัยเรื่อง Text Mining and Machine Learning Algorithms to Identifying Diseases and Providing Repair Action Using ICD-10 Codes โดยนำเทคนิคเหมืองข้อความมาใช้ร่วมกับอัลกอริทึมจำแนกประเภท เพื่อจำแนกโรคจากรหัส ICD-10 และให้ข้อมูลแนวทางการรักษาในเบื้องต้น ผลการวิจัยพบว่าระบบที่จัดทำสามารถช่วยลดภาระของแพทย์ในการคัดกรองผู้ป่วยในเบื้องต้น ได้เป็นอย่างดี

Thorsten Joachims [8] ได้ทำการศึกษาวิจัย เรื่อง Text Categorization with Support Vector Machines โดยระบุข้อได้เปรียบของการใช้ Support Vector Machine ในการจัดประเภทเอกสารเมื่อเทียบกับอัลกอริทึมประเภทอื่น ๆ ผลการวิจัยพบว่า Support Vector Machine มีความสามารถในการจัดการข้อมูลหลายมิติ (High Dimensional Feature Spaces) ได้เป็นอย่างดีซึ่งเหมาะสมกับการใช้งานเพื่อจัด

ประเภทเอกสาร และข้อคืออีกประการคือไม่จำเป็นต้องใช้การกระบวนการ Feature Selection ที่ซับซ้อนก็สามารถได้ผลลัพธ์ที่มีความแม่นยำสูงได้

3. วิธีการดำเนินงาน



ภาพที่ 2: กรอบแนวคิดในการวิจัย

ภาพที่ 2 แสดงกรอบแนวคิดในการวิจัยงาน ซึ่งมุ่งเน้นการพัฒนาแบบจำลองจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์เป็น 3 ประเภทคือ Core, Access และ Last Mile ด้วยวิธีการทำเหมืองข้อความร่วมกับการใช้อัลกอริทึมจำแนกประเภท ได้แก่ Decision Tree, Naïve Bayes, Bayesian Network, K-Nearest Neighbor, Support Vector Machine และ Deep Neural Network เพื่อประเมินอัลกอริทึมที่มีประสิทธิภาพในการจำแนกข้อมูลสูงสุด จากนั้นนำผลลัพธ์ที่ได้มาแสดงผลแบบ Visualization โดยมีรายละเอียดในแต่ละหัวข้อดังนี้

3.1 Training Data

ผู้วิจัยคัดเลือกข้อมูลรายการอุปกรณ์โครงข่ายบรอดแบนด์ของบริษัทโทรคมนาคมแห่งหนึ่งในประเทศไทย จำนวน 100,419 รายการ ซึ่งเป็นข้อมูลที่ระบุประเภทโครงข่ายโดยผู้เชี่ยวชาญแล้วมาใช้เป็นข้อมูลฝึกสอน

3.2 Text Preprocessing

3.2.1 Create Data frame

เป็นการสร้าง Data frame จากข้อมูลในคอลัมน์ Description ซึ่งเป็นคำอธิบายอุปกรณ์ เพื่อนำมาดำเนินการทำความสะอาดข้อมูลต่อไป

3.2.2 Remove Site name/Project Name

เป็นการลบอักษรและพยัญชนะภาษาไทยออกจากคอลัมน์ Description เพื่อให้ข้อความในคอลัมน์ Description เพื่อให้ข้อความเหลือเฉพาะชื่อและประเภทของอุปกรณ์

3.2.3 Remove Special Character

เป็นการลบอักขระพิเศษออกจากคอลัมน์ Description ซึ่งประกอบด้วยอักขระ () { } [] ; : # * " ' / \ , | + - เพื่อทำให้ข้อความที่นำมาวิเคราะห์ปรากฏเฉพาะคำที่มีนัยสำคัญ

3.2.4 Add Replace Date

เป็นการระบุวันที่ต้องเปลี่ยนอุปกรณ์โดยสร้างคอลัมน์ Replace Date เพิ่มไปใน Dataframe จากข้อมูลวันที่ได้มา (Acquisition Date) บวกกับอายุการใช้งานตามสัญญา (Contract Period)

3.3 Extract Feature

ผู้วิจัยนำข้อมูลที่ได้จากการทำความสะอาดข้อมูลในหัวข้อที่ 3.2 Text Preprocessing มาผ่านกระบวนการ Extract Feature ซึ่งประกอบด้วยการใช้เทคนิค Stop Word Removal, Stemming, Tokenization และ TF-IDF ได้ผลลัพธ์เป็น Feature Vector ที่จะนำไปใช้ในกระบวนการจัดประเภทข้อมูลต่อไป

3.4 Classification

ผู้วิจัยพัฒนาแบบจำลองการแยกประเภทอุปกรณ์โครงข่ายบรอดแบนด์โดยนำ Feature Vector มาทดสอบกับอัลกอริทึม ได้แก่ Decision Tree, Naïve Bayes, Bayesian

Network, K-Nearest Neighbor, Support Vector Machine และ Deep Neural Network เพื่อพัฒนาแบบจำลองที่มีประสิทธิภาพสูงสุด โดยกำหนดสัดส่วนข้อมูลสำหรับการทดสอบเป็น Training Data : 70% และ Test Data : 30%

4. ผลการดำเนินงาน

จากการดำเนินการสร้างแบบจำลองจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์โดยใช้เทคนิคเหมืองข้อความได้ผลการดำเนินงานวิจัยในหัวข้อต่าง ๆ ดังนี้

4.1 ผลการวิเคราะห์ข้อความ

จากกระบวนการ Text Preprocessing และ Extract Feature ตามที่ได้กล่าวในวิธีการดำเนินงานงานวิจัยข้อได้ผลลัพธ์ของการวิเคราะห์ข้อความประกอบด้วย Attribute ทั้งหมด 4,683 Attributes

4.2. ผลการทดสอบประสิทธิภาพของอัลกอริทึม

จากการนำข้อมูลที่ได้กระบวนการทำเหมืองข้อความมาทดสอบโดยใช้อัลกอริทึม เพื่อจำแนกประเภทของโครงข่ายจากรายการอุปกรณ์ ได้ผลการทดสอบตามตารางที่ 2

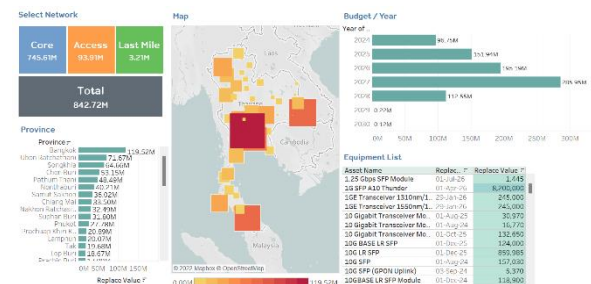
ตารางที่ 2: ผลการทดสอบประสิทธิภาพอัลกอริทึม

Algorithms	Conditions	Accuracy	Precision	Recall	F1 Score
Decision Tree, (Criterion = Gain Ratio)	Max. Dept: 5	43.43	33.30	52.65	40.80
	Max. Dept: 10	45.14	34.72	55.87	42.83
	Max. Dept: 15	46.47	35.62	58.39	44.25
	Max. Dept: 20	44.39	33.46	55.46	41.74
	Max. Dept: 25	44.39	33.46	55.46	41.74
Naïve Bayes	Estimation Mode = Full	74.61	73.93	80.84	77.23
Bayesian Networks	Estimator = Simple Estimator	67.85	69.51	72.53	70.99
K-Nearest Neighbors	K = 1	78.99	78.62	82.40	80.47
	K = 2	79.14	78.73	82.55	80.59
	K = 3	76.17	75.25	80.87	77.96
	K = 4	76.24	75.33	80.92	78.03
	K = 5	74.76	73.83	79.89	76.74
	Linear Kernel	89.68	88.45	92.13	90.25

Algorithms	Conditions	Accuracy	Precision	Recall	F1 Score
Support Vector Machine (C = 0 Epsilon = 0.001 Gamma = 0)	Second Polynomial*	90.35	90.01	91.68	90.84
	Third Polynomial	83.25	86.69	84.62	85.64
	RBF Kernel	33.26	11.09	33.33	16.64
	Sigmoid Kernel	89.68	88.45	92.13	90.25
Deep Neural Network (Epoch = 100 Learning Rate = 0.01, Hidden Layer = 3, Neural = 50)	ReLU	83.67	82.12	87.76	84.85
	Sigmoid	81.66	80.17	86.35	83.15
	Softmax	81.74	80.52	86.14	83.24
	TanH	82.55	81.01	86.91	83.86

จากผลการทดสอบพบว่าอัลกอริทึม Support Vector Machine ให้ผลการจำแนกที่ดีที่สุด เมื่อใช้ Kernel แบบ Second Polynomial โดยได้ค่าความถูกต้องเท่ากับ 90.35% ค่าความแม่นยำเท่ากับ 90.01% และค่าระลอกเท่ากับ 91.68%

4.3 การนำข้อมูลมาแสดงผลแบบ Visualization



ภาพที่ 3: หน้าจอแสดงผล

ภาพที่ 3 ผู้วิจัยนำผลลัพธ์จากแบบจำลองจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์มาแสดงผลแบบ Visualization โดยออกแบบหน้าจอให้แสดงผลข้อมูลแบบ Interactive เพื่อนำไปใช้ในเชิงการบริหารจัดการอุปกรณ์ และการจัดสรรงบประมาณตามประเภทของโครงข่ายโดยแบ่งส่วนของการแสดงผลข้อมูลเป็น 5 ส่วนดังนี้

4.3.1 Select Network เป็นส่วนที่ใช้การเลือกประเภทของโครงข่ายโดยแสดงข้อมูลที่รวบรวมจากผลลัพธ์ที่ได้จากการทำนายแบ่งเป็น Core, Access และ Last Mile พร้อมทั้งบอกงบประมาณที่ต้องใช้ในการทดแทนอุปกรณ์โครงข่ายแต่ละส่วน และมูลค่ารวมของอุปกรณ์

4.3.2 Province เป็นส่วนที่แสดงผลงบประมาณที่ต้องใช้ในการทดแทนอุปกรณ์ในแต่ละจังหวัดโดยเรียงลำดับจากมากไปน้อย

4.3.3 Map เป็นการแสดงผลข้อมูลแบบ Heatmap บนแผนที่ประเทศไทย โดยขนาดและสีของสี่เหลี่ยมจะแปรผันตามงบประมาณที่ต้องใช้ในแต่ละจังหวัด เพื่อให้ผู้ใช้งานเห็นภาพได้ง่ายขึ้นโดยไม่ต้องอ่านจากคำอธิบาย

4.3.4 Budget/Year เป็นการแสดงงบประมาณที่ต้องใช้ในแต่ละปีในรูปแบบ Bar Chart

4.3.5 Equipment List เป็นการแสดงผลรายการอุปกรณ์ประกอบด้วยชื่ออุปกรณ์ วันที่ต้องซื้อทดแทน และมูลค่าของอุปกรณ์

5. สรุป

การวิจัยในครั้งนี้ ผู้วิจัยได้พัฒนาแบบจำลองสำหรับการจำแนกประเภทอุปกรณ์โครงข่ายบรอดแบนด์ โดยนำเทคนิคเหมืองข้อความมาประยุกต์ใช้กับรายการอุปกรณ์โครงข่ายบรอดแบนด์ในการเลือกคุณลักษณะจากคำอธิบายของรายการอุปกรณ์ จากนั้นใช้อัลกอริทึมจำแนกประเภทในการจำแนกอุปกรณ์ออกเป็นประเภท Core, ประเภท Access และประเภท Last Mile จากนั้นนำผลลัพธ์ที่ได้มาแสดงบนโปรแกรมธุรกิจอัจฉริยะ สรุปผลจากการดำเนินงานวิจัยแยกเป็น 3 ส่วนดังนี้

5.1 การวิเคราะห์ข้อความ ผู้วิจัยได้นำรายการอุปกรณ์โครงข่ายบรอดแบนด์จำนวน 100,419 รายการ ที่ระบุประเภทโครงข่ายแล้วมาดำเนินการทำความสะอาดข้อมูล จากนั้นใช้เทคนิคกำจัดคำหยุด (Stop Word Removal) เทคนิคสกัดคำรากศัพท์ (Stem) เทคนิคการตัดคำ (Tokenization) และสร้างเวกเตอร์คุณลักษณะ (Feature Vector) โดยใช้เทคนิค TF-IDF พบว่ามีการสร้าง Attribute จำนวน 4,683 Attributes

5.2 การพัฒนาแบบจำลองจำแนกประเภท ผู้วิจัยนำผลลัพธ์จากกระบวนการวิเคราะห์ข้อความมาดำเนินการพัฒนาแบบจำลองร่วมกับอัลกอริทึมตามตารางที่ 4-1 พบว่าอัลกอริทึม Support Vector Machine ให้ผลการจำแนกที่ดีที่สุด เมื่อใช้ Kernel แบบ Second Polynomial

โดยได้ค่าความถูกต้องเท่ากับ 90.35% ค่าความแม่นยำเท่ากับ 90.01% ค่าระลอกเท่ากับ 91.68% และค่า F1 Score เท่ากับ 90.84%

5.3 การนำข้อมูลมาแสดงผลแบบ Visualization ทำให้ผู้ปฏิบัติงานเห็นข้อมูลในภาพรวมได้ดียิ่งขึ้น เพิ่มประสิทธิภาพในการบริหารจัดการ วางแผนพัฒนา ซ่อมบำรุง และตัดสินใจลงทุนโครงข่ายในทั้งในระดับจังหวัดและในภาพรวมของประเทศได้

สำหรับข้อเสนอแนะและแนวทางการพัฒนาต่อไปนั้น เนื่องจากเทคโนโลยีการให้บริการบรอดแบนด์มีการเปลี่ยนแปลงอย่างรวดเร็ว จึงควรมีการปรับปรุงชุดข้อมูลฝึกสอนของแบบจำลองอยู่เป็นประจำ และหากนำข้อมูลปริมาณการใช้งาน (Traffic) ตามประเภทโครงข่าย ในแต่ละจังหวัด มาใช้ประกอบการพัฒนาแบบจำลองจะสามารถพัฒนาระบบสนับสนุนการตัดสินใจที่แม่นยำยิ่งขึ้นได้

เอกสารอ้างอิง

- [1] National Electronics and Computer Technology Center (2021). Thailand Internet Bandwidth 2016-2020. Available online at <http://internet.nectec.or.th/webstats/bandwidth.iir?Sec=bandwidth>
- [2] Department of Fees and Service Rates, National Broadcasting and Telecommunications Commission (2021). The Average Service Rate of Fixed Broadband Internet Service via FTTx Connection Available online at <https://www.nbtcc.go.th/Services.aspx?>
- [3] Designing and Supporting Computer Networks, CCNA Discovery Learning Guide. Kenneth D. Stewart, Pearson Education, 2008.
- [4] Liang, J. W (2003)., Introduction to Text and Web Mining, Seminar at North Carolina Technical University.
- [5] M. Hofmann, Support Vector Machines: Kernels and the Kernel Trick, 2006.
- [6] Gary M. Weiss (2005). Data Mining in Telecommunications. Data Mining and Knowledge Discovery Handbook: Kluwer Academic Publishers.
- [7] Ramdasi, A.P., Sathyalakshmi, S. (2022). Text Mining and Machine Learning Algorithms to Identifying Diseases and Providing Repair Action

Using ICD-10 Codes. Lecture Notes in Networks and Systems, vol 321. Springer, Singapore.

- [8] T. Joachims (2015). Text categorization with support vector machines. In Proceedings of ECML-98, 10th European Conference on Machine Learning.

ระบบพยากรณ์การสต็อกน้ำมันปิโตรเลียมของคลังน้ำมันปลายทาง ด้วยเทคนิคอนุกรมเวลา

Forecasting System of Oil Refinery using Time Series Techniques

รัชชภูมิ บุนนาค (Ratchaphoom Boonnaka)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6307021856136@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีจุดประสงค์เพื่อทำการคัดเลือกแบบจำลองที่เหมาะสมในการพยากรณ์การสต็อกน้ำมันปิโตรเลียมที่ใช้เทคนิคอนุกรมเวลา ประกอบไปด้วยวิธีการสถิติ 4 วิธี แบบจำลองค่าเฉลี่ยเคลื่อนที่ แบบจำลองเอ็กซ์โปเนนเชียล แบบจำลองเอ็กซ์โปเนนเชียลกำลังสอง และแบบจำลองการแยกส่วนประกอบในการพยากรณ์และเลือกเทคนิคที่เหมาะสมโดยใช้โปรแกรม Minitab เป็นเครื่องมือช่วยในการวิเคราะห์ค่าทางสถิติทำการเปรียบเทียบประสิทธิภาพแบบจำลองในการพยากรณ์โดยใช้เกณฑ์ในการคัดเลือก คือ ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ต่ำที่สุด ซึ่งทำการเก็บรวบรวมข้อมูลการสต็อกน้ำมันปิโตรเลียม จำนวน 42 เดือน โดยน้ำมันเบนซิน E20 นั้นจำนวน 18 เดือน ส่วนน้ำมันเบนซิน 95 นั้นจำนวน 24 เดือน ในการทดสอบนั้นน้ำมันเบนซิน E20 ใช้ข้อมูลในการทดสอบทั้งหมด 6 เดือน และน้ำมันเบนซิน 95 นั้นจำนวน 12 เดือน ผลการวิจัยพบว่าข้อมูลการสต็อกน้ำมันด้วยวิธีแบบจำลองเอ็กซ์โปเนนเชียลกำลังสองให้ผลที่มีประสิทธิภาพสูงสุด โดยมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์โดย เบนซิน E20 ร้อยละ 47.95 และเบนซิน 95 ร้อยละ 38.5

คำสำคัญ: เทคนิคการวิเคราะห์อนุกรมเวลา วิธีหาค่าเฉลี่ยเคลื่อนที่ วิธีแบบจำลองเอ็กซ์โปเนนเชียล กำลังสอง วิธีแบบจำลองการแยกส่วนประกอบ

Abstract

The purpose of this research is to select a suitable model for forecasting petroleum stocks using time series techniques. It consists of 3 statistical methods, a moving average model, Quadratic exponential model and the separation model for forecasting and selecting the appropriate technique using Minitab as a tool for statistical analysis to compare the model's efficiency in forecasting. The selection criteria were The lowest mean percentage absolute error. Collected 42 months of petroleum inventory data, which is divided into two parts, consisting of gasoline E20 (Bensin E20) and gasoline 95 (Bensin 95), the first set of 18 months since January 2019. Until July 2020 to test all 3 forecasting models and use the second set of data for 24 months from January 2019 to December 2020 to verify the validity. Correct model from the research results, it was found that How to stock up on oil The quadratic exponential model yields the most efficient results. The percentage of absolute discrepancy is 47.95 percent of E20 gasoline and 38.51 percent of 95 gasoline.

Keywords: Time Series Analysis, Moving Average, Double Exponential Smoothing, Decomposition.

1. บทนำ

น้ำมันปิโตรเลียมเป็นส่วนสำคัญอย่างยิ่งต่อการขับเคลื่อนธุรกิจการคมนาคมภายในประเทศ ซึ่ง ประกอบด้วย น้ำมันปิโตรเลียมหลากหลายชนิดที่เกิดจาก กระบวนการผลิต โรงกลั่นน้ำมัน แก๊สโซฮอล์ 95 แก๊สโซฮอล์ E20 เพื่อกระจายไปเก็บในคลังสะสมน้ำมัน

การคำนวณนั้นมีหลากหลายขั้นตอน ก่อนนำมาทำการวิเคราะห์ให้ได้ผลลัพธ์ที่ดีที่สุดในการนำเสนอให้กับคลังปลายทางเพื่อประกอบการตัดสินใจ ซึ่งประกอบด้วย ใบ งานมากมายก่อนนำมาคำนวณจนเสร็จสิ้นซึ่งทำให้การทำงานนั้นช้าช้อน เนื่องจากราคาน้ำมันมีการเคลื่อนตลอด ได้มีผู้วิจัยนำเสนอ การใช้วิธีแนวโน้มเชิงเส้นของโสลท์ และ วิธีบ็อกซ์-เจนกินส์ทำการพยากรณ์ปริมาณการชำระเงินผ่านระบบแอปพลิเคชัน โมบายแบงก์กิ้ง [1] การพยากรณ์ความต้องการสินค้าและการวางแผน การผลิตสินค้าแปรรูปจาก ด้วยวิธีค่าเฉลี่ยเคลื่อนที่ วิธีการปรับให้เรียบแบบเอกซ์โปเนนเชียลสองเท่า วิธีการลดถอยเชิง และวิธีหาค่าเฉลี่ยการเคลื่อนที่ของอินทิกรัลถดถอย [2] เป็นต้น

จากปัญหาข้างต้นทำให้เงื่อนทำให้ผู้วิจัยมีแนวคิดสร้างแบบจำลองพยากรณ์การสต็อกน้ำมันเพื่อนำเสนอให้กับคลังน้ำมันปลายทาง ด้วยเทคนิคการวิเคราะห์อนุกรมเวลาด้วยแบบจำลองที่ดีที่สุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 5 หัวข้อหลัก ดังนี้

2.1 การพยากรณ์แบบอนุกรมเวลา

ได้รับการยอมรับโดยทั่วไปว่ามีความถูกต้องมากที่สุดสามารถนำข้อมูลที่มีค่าหลากหลายจำนวนมากซึ่งเป็นการนำข้อมูลจากอดีตมาคำนวณหาค่าพยากรณ์ในอนาคตได้[2]

2.2 วิธีแบบจำลองค่าเฉลี่ยเคลื่อนที่ (Moving Average)

วิธีค่าเฉลี่ยเคลื่อนที่นั้นเหมาะกับการพยากรณ์ในช่วงเวลาสั้นๆ และข้อมูลนั้นค่อนข้างมีลักษณะที่แน่นอน

ซึ่งไม่เหมาะกับข้อมูลที่มีลักษณะเป็นแนวโน้มหรือ ข้อมูลที่มีการเปลี่ยนแปลงบ่อย โดยมีลักษณะดังสมการ ต่อไปนี้ [5]

$$F_{t+1} = \frac{(Y_t + Y_{t-1} + Y_{t-2} + \dots + Y_{t-n+1})}{N} \tag{1}$$

โดย F_t = ค่าพยากรณ์ในช่วงเวลา t

t = เวลาที่ทำการพยากรณ์

A_t = ค่าความต้องการในงวดที่ t

N = จำนวนข้อมูลที่นำมาเฉลี่ย

2.3 วิธีปรับเรียบแบบเอกซ์โปเนนเชียลกำลังสอง (Double Exponential Smoothing Model: DES)

เทคนิคการพยากรณ์เป็นการนำข้อมูลในรูปแบบครั้งแต่ไม่มีความ Exponential Smoothing มาใช้งาน 2 ครั้งแต่ เป็นฤดูกาล ซึ่งเหมาะกับการพยากรณ์ในระยะสั้นข้อมูล ในการพยากรณ์นั้นควรมีอย่างน้อย 5 ชุดข้อมูลดังรูปสมการนี้[10]

$$F_t = \alpha Y_{t-1} + (1 - \alpha)F_{t-1} \tag{2}$$

$$F'_t = \alpha F_t + (1 - \alpha)F'_{t-1} \tag{3}$$

โดย F_t = ค่าพยากรณ์ความต้องการจากวิธี SES

Y_t = ค่าจริงที่เกิดขึ้น ณ เวลา t

F' = ค่าพยากรณ์จากวิธีการ DES สำหรับเวลา t

2.1.3วิธีแยกส่วนประกอบ (Decomposition)

การแยกส่วนประกอบเป็นวิธีคิดรูปแบบหนึ่งของแนวคิดเชิงคำนวณ เป็นการพิจารณาเพื่อแบ่งปัญหาหรืองานออกเป็นส่วนย่อย ทำให้สามารถจัดการกับปัญหาหรืองานได้ง่ายขึ้น การแตกปัญหาที่ซับซ้อนให้เป็นปัญหาย่อยที่มีขนาดเล็กลงและซับซ้อนน้อยลง เพื่อช่วยให้การวิเคราะห์และออกแบบวิธีการแก้ปัญหาทำได้ง่ายขึ้น โดยมีรูปแบบสมการดังนี้ [7]

$$F_t = T_t \times S_t \tag{4}$$

โดย F_t = ค่าพยากรณ์ของช่วงเวลาปัจจุบัน

Y_t = ค่าของข้อมูลจริงในช่วงเวลาปัจจุบัน

T_t = ค่าประมาณของแนวโน้ม

S_t = ค่าประมาณของการผันแปรตามฤดูกาล

C_t = ค่าประมาณของการแปรผันตามวัฏจักร

I_t = ค่าประมาณของความไม่แน่นอน

N = จำนวนข้อมูล

2.2 งานวิจัยที่เกี่ยวข้อง

ชมพูนุช [6] ข้อมูลยอดขายผลิตภัณฑ์รายเดือนจำนวน 24 เดือนแรกที่เก็บรวบรวมตั้งแต่เดือนตุลาคม พ.ศ. 2557 - เดือนกันยายน พ.ศ. 2559 ถูกนำมาใช้ในการศึกษาเปรียบเทียบประสิทธิภาพแบบจำลองทั้ง 2 วิธี สมบูรณ์เฉลี่ย และค่าเบี่ยงเบนสมบูรณ์เฉลี่ย

รุ่งนภา [7] ทำการพยากรณ์การลดปริมาณการขาดแคลนสินค้า โดยทำการเปรียบเทียบ ประสิทธิภาพด้วย 6 วิธี คือ วิธีค่าเฉลี่ยเคลื่อนที่ วิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลเดี่ยว วิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลสองเท่า วิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลด้วยการเทรนด์ วิธีวิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลกำลังสอง และวิธีการถดถอยเชิงเส้น ผลการวิจัยพบว่าวิธีเอกซ์โพเนนเชียล ให้ค่าความผิดพลาดน้อยที่สุด

กนกกาญจน์ [10] ทำการพยากรณ์ยอดขายสินค้าอุปโภคที่เหมาะสมของบริษัทเอกชนแห่งหนึ่ง โดยทำการเปรียบเทียบประสิทธิภาพด้วย 4 วิธี คือการพยากรณ์แบบแยกส่วนประกอบ วิธีถ่วงเฉลี่ยเคลื่อนที่ วิธีการปรับเรียบเอกซ์โพเนนเชียล และวิธีแบบจำลอง ผลการวิจัยพบว่าวิธีเอกซ์โพเนนเชียล ให้ค่าความผิดพลาดน้อยที่สุด

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการศึกษาคัดเลือกเปรียบเทียบแบบจำลองที่เหมาะสมในการพยากรณ์ยอดขายด้วยเทคนิคการวิเคราะห์อนุกรมเวลา โดยมีรายละเอียดและขั้นตอนการดำเนินงานดังนี้

3.1 การเก็บรวบรวมข้อมูล

การเก็บข้อมูลการสต็อกน้ำมันปีโตรเลียม ตั้งแต่เดือนมกราคม พ.ศ. 2562 ถึงเดือนธันวาคม พ.ศ. 2563 รวม 42 เดือนประกอบไปด้วย น้ำมันเบนซิน E20 และเบนซิน 95 ซึ่งข้อมูลถูกแบ่งออกเป็นจำนวน 18 เดือน และ 24 เดือน เพื่อใช้ในการสร้างแบบจำลองและ

เปรียบเทียบประสิทธิภาพในการพยากรณ์ และชุดข้อมูล Test Set เป็นข้อมูลน้ำมันเบนซิน E20 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนกรกฎาคม พ.ศ. 2563 รวม 6 เดือน และ น้ำมันเบนซิน 95 ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนธันวาคม พ.ศ. 2563 รวมเป็น 12 เดือน เพื่อใช้ข้อมูลสำหรับการทดสอบความแม่นยำของการพยากรณ์

3.2 การสร้างแบบจำลองพยากรณ์

การสร้างแบบจำลองการพยากรณ์ใช้โปรแกรม Minitab เป็นเครื่องมือ สำหรับการพยากรณ์ใช้เทคนิค Time Series Analysis 3 วิธี ได้แก่ วิธี Moving Average โดยทำการกำหนดค่า MA Length 3, 6, 9, 12 วิธี Double Exponential Smoothing กำหนดพารามิเตอร์ค่า α, γ ตั้งแต่ 0.05, 0.1 ถึง 0.9 และวิธี Decomposition โดยทำการกำหนดค่า Length 3, 6, 9, 12

3.3 การคัดเลือกแบบจำลองในการพยากรณ์

พิจารณาประสิทธิภาพของแบบจำลองพยากรณ์ทั้ง 3 วิธี ด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) และค่าเบี่ยงเบนสมบูรณ์เฉลี่ย (MAD) ที่ต่ำที่สุด

3.4 การทดสอบความถูกต้องของแบบจำลอง

ทดสอบความถูกต้องของแบบจำลองพยากรณ์ นำข้อมูลการพยากรณ์ที่ได้จากแบบจำลองมาทดสอบความถูกต้องและแม่นยำกับข้อมูลชุดที่ 2 (Test Set) จำนวน 12 เดือนพิจารณาค่าความคลาดเคลื่อนในการพยากรณ์ ด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE)

3.5 นำเสนอข้อมูลการพยากรณ์

นำเสนอข้อมูลการพยากรณ์ยอดขายในรูปแบบ Dashboard any React + Chartjs

4. ผลการดำเนินงาน

ผลการดำเนินงานวิจัยแบ่งออกเป็น 4 หัวข้อ แสดงรายละเอียดได้ดังนี้

4.1 ผลการดำเนินงานสร้างแบบจำลองพยากรณ์

ผลการสร้างแบบจำลองพยากรณ์ ยอดการสต็อก

น้ำมันปิโตรเลียม ซึ่งใช้เทคนิคอนุกรมเวลา 3 วิธี Moving Average, Double Exponential Smoothing, Decomposition โดยผลการวิจัยมีรายละเอียดดังตาราง 1 ถึง 6

ตารางที่ 1: การพยากรณ์ด้วยวิธี Moving Average ของ E20

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Moving Average	MA Length 3	23.78	0.74
	MA Length 6	18.98	0.44
	MA Length 9	16.70	0.35
	MA Length 12	18.53	0.58

ตารางที่ 2: การพยากรณ์ด้วยวิธี Moving Average ของเบนซิน95

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Moving Average	MA Length 3	75.98	1.36
	MA Length 6	79.69	1.35
	MA Length 9	79.75	1.20
	MA Length 12	83.56	1.28

จากตารางที่ 1 และ 2 เมื่อพิจารณาค่าความคลาดเคลื่อน พบว่า MA Length 9 ให้ประสิทธิภาพที่ดีที่สุด ซึ่งน้ำมัน เบนซิน E20 ได้ค่า MAPEเท่ากับ 16.70 และค่า MAD เท่ากับ 0.35 ซึ่งน้ำมันเบนซิน 95 ได้ค่า MAPE เท่ากับ 79.75 และค่า MAD เท่ากับ 1.20

ตารางที่ 3: การพยากรณ์ด้วยวิธี Double Exponential Smoothing ของ E20

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Double Exponential Smoothing	$\alpha=0.05, \gamma=0.05$	17.47	0.54
	$\alpha=0.05, \gamma=0.1$	17.52	0.54
	$\alpha=0.05, \gamma=0.2$	17.62	0.54

ตารางที่ 4: การพยากรณ์ด้วยวิธี Double Exponential Smoothing ของเบนซิน95

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Double Exponential Smoothing	$\alpha=0.05, \gamma=0.9$	77.54	1.26
	$\alpha=0.1, \gamma=0.05$	72.85	1.21
	$\alpha=0.1, \gamma=0.1$	73.43	1.21

จากตารางที่ 3 และ 4 เมื่อพิจารณาค่าความคลาดเคลื่อน ของน้ำมันเบนซิน E20 พบว่า $\alpha=0.05$ และ $\gamma=0.05$ ให้ ประสิทธิภาพที่ดีที่สุด โดยค่า MAPE เท่ากับ 17.47 และ ค่า MAD เท่ากับ 0.54 และ เบนซิน 95

พบว่า $\alpha=0.1$ และ $\gamma=0.05$ ให้ประสิทธิภาพที่ดีที่สุด โดยค่า MAPE เท่ากับ 72.85 และค่า MAD เท่ากับ 1.21

ตารางที่ 5: การพยากรณ์ด้วยวิธี Decomposition ของ E20

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Decomposition	Seasonal Length 3	17.06	0.53
	Seasonal Length 6	17.63	0.55
	Seasonal Length 9	5.64	0.17
	Seasonal Length 12	12	12

ตารางที่ 6: การพยากรณ์ด้วยวิธี Moving Average ของเบนซิน95

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Decomposition	Seasonal Length 3	55.52	1.00
	Seasonal Length 6	56.21	1.01
	Seasonal Length 9	57.81	1.01
	Seasonal Length 12	51.12	0.99

จากตารางที่ 5 และ 6 เมื่อพิจารณาค่าความคลาดเคลื่อน ของน้ำมันเบนซิน E20 พบว่าค่า Seasonal Length 9 ให้ ประสิทธิภาพที่ดีที่สุด โดยมีค่า MAPE เท่ากับ 5.64 และค่า MAD เท่ากับ 0.17 ซึ่งน้ำมันเบนซิน 95 พบว่าค่า Seasonal Length 12 ให้ประสิทธิภาพดีที่สุด โดยมีค่า MAPE เท่ากัน 51.12 และค่าเท่ากับ 0.99

4.2 การเปรียบเทียบและคัดเลือกแบบจำลอง

การเปรียบเทียบและคัดเลือกแบบจำลองนำค่าความผิดพลาดของตัวแบบการพยากรณ์ ได้แก่ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย และค่าเบี่ยงเบนสัมบูรณ์เฉลี่ย ของวิธี Moving Average วิธี Double Exponential Smoothing และ Decomposition มาเปรียบเทียบเพื่อหาแบบจำลองที่แม่นยำกับข้อมูลน้ำมันปิโตรเลียมพบว่าวิธีแบบจำลองค่าเฉลี่ยเคลื่อนที่ เป็นวิธีที่ค่าความผิดพลาดต่ำที่สุด ซึ่งเหมาะสมในการพยากรณ์ข้อมูล น้ำมันปิโตรเลียม รายละเอียดแสดงดังตาราง 7 และ 8

ตารางที่ 7: การเปรียบเทียบความแม่นยำของแบบจำลองพยากรณ์ของ E20

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Moving average	MA Length 9	16.70	0.52
Double Exponential Smoothing	$\alpha=0.05, \gamma=0.05$	17.47	0.54
Decomposition	Seasonal Length 9	5.64	0.17

ตารางที่ 8: การเปรียบเทียบความแม่นยำของแบบจำลองพยากรณ์ของเบนซิน 95

วิธีพยากรณ์	พารามิเตอร์	MAPE	MAD
Moving average	MA Length 3	79.75	1.20
Double Exponential Smoothing	$\alpha = 0.1$ $\gamma = 0.95$	72.85	1.21
Decomposition	Seasonal Length 9	51.12	0.99

จากตารางที่ 7 และ 8 แสดงการเปรียบเทียบผลการพยากรณ์ยอดการสต็อกน้ำมันปีโตรเลียมเบนซิน E20 และ เบนซิน 95 ทั้ง ทั้ง 3 วิธี พบว่าแบบจำลองพยากรณ์วิธีแบบจำลอง Decomposition ให้ประสิทธิภาพในการพยากรณ์สูงสุด โดยมีค่า MAPE เท่ากับ 5.64 และค่า MAD เท่ากับ 0.17 และมีค่า MAPE เท่ากับ 51.12 และค่า MAD เท่ากับ 0.99

4.3 การทวนสอบความถูกต้องของแบบจำลอง

ผลการทวนสอบความถูกต้องของแบบจำลอง จากการพิจารณาค่าความคลาดเคลื่อนของแบบจำลองพยากรณ์ที่ให้ผลลัพธ์ที่ดีที่สุด ได้แก่ วิธี Moving Average รองลงมาคือวิธี Double Exponential Smoothing และ วัน Decomposition นำข้อมูลการพยากรณ์ที่ได้มาทำการทวน สอบความถูกต้อง และความแม่นยำกับข้อมูลการสต็อก น้ำมันปีโตรเลียม จำนวน 12 เดือน พบว่าแบบจำลอง พยากรณ์วิธีแบบจำลองเอ็กซ์โปเนนเชียลกำลังสอง มี ความคลาดเคลื่อนต่ำที่สุดที่ร้อยละ 47.95 และ 38.51 รายละเอียดแสดงดังตารางที่ 9 และ 10

ตารางที่ 9: ผลการทวนสอบความถูกต้องของแบบจำลองของE20

เดือน	ปริมาณน้ำมัน	Double Exponential Smoothing		
	Test Set	Forecast	Error	MAPE
JAN	5.57	3.10974	2.464707	44.21
FEB	6.33	3.10974	3.221796	50.88
MAR	6.06	3.10974	2.945447	48.64
APR	5.69	3.10974	2.582179	45.37
MAY	6.66	3.10974	3.546698	53.28
JUN	5.10	3.10974	1.991821	39.04

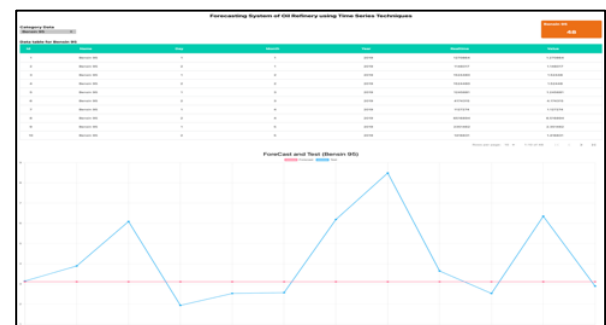
เดือน	ปริมาณน้ำมัน	Moving Average		
	Test Set	Forecast	Error	MAPE
AUG	5.70	3.10974	2.592551	45.47
SEP	5.84	3.10974	2.726763	46.72
OCT	6.46	3.10974	3.346775	51.84
NOV	5.27	3.10974	2.156422	40.95
ร้อยละ				47.95

ตารางที่ 10: ผลการทวนสอบความถูกต้องของแบบจำลองของเบนซิน 95

เดือน	ปริมาณน้ำมัน	Double Exponential Smoothing		
	Test Set	Forecast	Error	MAPE
JAN	3.14	2.18454	0.954448	30.41
FEB	3.89	2.18454	1.702651	43.80
MAR	6.09	2.18454	3.902503	64.11
APR	1.94	2.18454	-0.24396	12.57
MAY	2.53	2.18454	0.345511	13.66
JUN	2.57	2.18454	0.382133	14.89
JUL	6.19	2.18454	4.009098	64.73
AUG	8.49	2.18454	6.301692	74.26
SEP	3.64	2.18454	1.458465	40.03
OCT	2.53	2.18454	0.344714	13.63
NOV	6.35	2.18454	4.163406	65.59
ร้อยละ				38.51

4.4 การนำเสนอข้อมูล

ผู้วิจัยนำเสนอผลการพยากรณ์ โดยใช้วิธีแบบจำลองเอ็กซ์โปเนนเชียลกำลังสอง เป็นตัวแบบในการพยากรณ์ยอดการสต็อกน้ำมันปีโตรเลียมในอนาคต โดยใช้ React + Chartjs



ภาพที่ 1: แผนภูมิภาพแสดงข้อมูลยอดการสต็อกน้ำมันปีโตรเลียม

แสดงการเปรียบเทียบข้อมูลจำนวน ยอดการสต็อกน้ำมัน ที่เกิดขึ้นจริงจำนวน 12 เดือน ตั้งแต่เดือนมกราคม ถึงเดือนธันวาคม พ.ศ. 2563 ด้วยวิธีแบบจำลองเอ็กซ์โปเนนเชียลกำลังสอง พบว่าจำนวน ยอดการสต็อกน้ำมันปีโตรเลียม มีความสอดคล้องและใกล้เคียงกัน

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้ศึกษาการพยากรณ์ด้วยเทคนิคอนุกรมเวลาเพื่อให้ได้วิธีการพยากรณ์ที่ดีที่สุดสำหรับการพยากรณ์ ยอดการสต็อกน้ำมันปีโตรเลียม ผู้วิจัยได้คัดเลือกวิธีการพยากรณ์ 3 วิธี ได้แก่ วิธีแบบจำลองค่าเฉลี่ยเคลื่อนที่ วิธีปรับเรียบแบบเอ็กซ์โปเนนเชียลกำลังสอง และวิธีแยกส่วนประกอบ ทดสอบประสิทธิภาพของแบบจำลองการพยากรณ์ด้วยค่า MAPE และค่า MAD จากผลการวิจัยพบว่า วิธีแยกส่วนประกอบ มีความเหมาะสมที่สุดโดยน้ำมันเบนซิน E20 มีค่า MAPE เท่ากับ 5.64 ค่า MAD เท่ากับ 0.17 และน้ำมันเบนซิน 95 มีค่า MAPE เท่ากับ 51.12 ค่า MAD เท่ากับ 0.99 หลังจากนำข้อมูลมาทวนสอบกับข้อมูลจริงจำนวน 12 เดือน สรุปได้ว่า การสต็อกน้ำมันปีโตรเลียม โดยสร้างแบบจำลองการพยากรณ์ด้วยวิธีเอ็กซ์โปเนนเชียลกำลังสอง ซึ่งน้ำมันเบนซิน E20 พบความคลาดเคลื่อนในการพยากรณ์ที่ร้อยละ 47.95 และน้ำมันเบนซิน 95 พบความคลาดเคลื่อนในการพยากรณ์ร้อยละ 38.51

เอกสารอ้างอิง

- [1] ทศนีย์ อัครพันธ์ และจิรวุฒิ เชิญเกียรติประดับ (2563), การประยุกต์ใช้เทคนิคการวิเคราะห์ อนุกรมเวลาเพื่อพยากรณ์ปริมาณการชำระเงินผ่านระบบอิเล็กทรอนิกส์แอปพลิเคชัน โมบายแบงก์กิ้ง, วิทยานิพนธ์วิทยาศาสตร์ มหาวิทยาลัย สาขาวิชาเทคโนโลยีมหาวิทยาลัยราชพฤกษ์.
- [2] วิษรชัย อินธิปัก (2561). การพยากรณ์ความต้องการสินค้า และการวางแผนการผลิตกรณีศึกษา โรงงานแปรรูปเนื้อเป็ด. สารนิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาการจัดการทาง วิศวกรรม วิทยาลัยนวัตกรรมการด้านเทคโนโลยี และวิศวกรรมศาสตร์มหาวิทยาลัยบูรพาบุรีรัมย์.
- [3] บุญฤทธิ์ ชูประคิมฐ์ และเสาวภา ชัยพิทักษ์ (2561), “ตัวแบบพยากรณ์ปริมาณการส่งออกมะม่วง ของประเทศไทย วารสารมหาวิทยาลัยนเรศวร ปีที่ 4 ฉบับที่ 1 : 74-85. พยากรณ์ทางสถิติ” วารสารวิชาการ วิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์ ปีที่ 9 ฉบับที่ 10 กรกฎาคม-ธันวาคม พ.ศ. 2560 หน้า 127-142.
- [4] มิฉา แก้วหาวงษ์. (2558). การพยากรณ์ปริมาณการใช้ไฟฟ้าของประเทศไทยโดยใช้ตัวแบบ SARIMA และตัวแบบการถดถอยที่มีความคลาดเคลื่อนนี้เป็นตัวแบบ ARMA วารสาร วิทยาศาสตร์และเทคโนโลยี, ปีที่ 4 ฉบับที่ 1 : 24-36.
- [5] บุญทอง ทะเลโยธิน และยุทธพรณ์ อารีพงษ์ (2561). “การเปรียบเทียบตัวแบบการพยากรณ์ราคาหุ้น โดยใช้แบบจำลองอาร์มาและอาร์แม็กซ์.” วารสารวิทยาศาสตร์และเทคโนโลยีหัวเฉียวเฉลิมพระเกียรติ, ปีที่ 26 ฉบับที่ 2 : 36.44-55.
- [6] ณัฐธยาน์ มนุษย์ดี. (2553). การพยากรณ์จำนวนผู้ใช้ อินเทอร์เน็ตในประเทศไทยด้วยการวิเคราะห์ อนุกรมเวลา และการวิเคราะห์การถดถอยเชิงพหุคูณ, สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ และการสื่อสาร บัณฑิตวิทยาลัย มหาวิทยาลัยธุรกิจ บัณฑิต.
- [7] รุ่งนภา ศรีประโล. (2557). การลดปริมาณการขาดแคลนสินค้าโดยใช้เทคนิคการพยากรณ์ กรณีศึกษา บริษัท ไเซโล (ประเทศไทย) จำกัด. วิทยานิพนธ์วิศวกรรมศาสตร์ มหาบัณฑิต สาขาวิชาจัดการทางวิศวกรรม มหาวิทยาลัย ธุรกิจบัณฑิต.
- [8] ณัญ สุภานัน สุพัทธนะ (2557). การพยากรณ์ ชนราคาเหล็ก ของประเทศไทย โดยแบบ ลอง ARIMA และ แบบจำลอง ARIMAX คณะพัฒนาการเศรษฐกิจ สถาบัน บัณฑิตพัฒนา บริหารศาสตร์.
- [9] Hadizadeh, A. (2020). *Overview of Quantitative Forecasting Methods on Sales of Naphthenic coils*. Master Thesis. Department of Management and Engineering. The Institute of Technology, Linkoping University Sweden.
- [10] กนกกาญจน์ มูลผลาล. (2557). “การศึกษาเทคนิคการพยากรณ์ ยอดขายสินค้าอุปโภคที่ เหมาะสม ของบริษัทเอกชนแห่งหนึ่ง.” วารสารวิชาการบริหารธุรกิจ สมาคมสถาบันอุดมศึกษาเอกชน แห่งประเทศไทย.

การเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการวิเคราะห์ความคิดเห็นของ ผู้ใช้บริการโรงแรม

The Comparison of Algorithms Performance to Analysis Hotel Guest Reviews

สันติ สุขเกษม (Santi Sukkasem)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021910024@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์นำเสนอการเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการโรงแรม จำนวน 3 อัลกอริทึม ได้แก่ อัลกอริทึมป่าสุ่ม อัลกอริทึมเอ็กซ์ทรีมกราดิเอนบูตติ้ง การเรียนรู้เชิงลึก โดยได้ทำการรวบรวมความคิดเห็นของผู้ใช้บริการโรงแรมที่ได้รับรางวัล จำนวน 500 ข้อความ โดยแบ่งข้อความออกเป็นความคิดเห็นเชิงบวกจำนวน 250 ข้อความ และความคิดเห็นเชิงลบจำนวน 250 ข้อความ โดยผ่านกระบวนการทางเหมืองข้อความและนำข้อความที่ได้ไปสร้างเวกเตอร์และแบ่งชุดข้อมูลออกเป็น 2 ชุด อัน ได้แก่ ชุดข้อมูลสอนร้อยละ 70 และชุดข้อมูลทดสอบร้อยละ 30 จากการทดสอบการประเมินประสิทธิภาพของแต่ละอัลกอริทึมจะมีการวัด ค่าความถูกต้องของอัลกอริทึม ค่าความแม่นยำ ค่าระลึก และค่าเฉลี่ย พบว่าอัลกอริทึมที่ให้ประสิทธิภาพที่ดีที่สุด คือ อัลกอริทึมเอ็กซ์ทรีมกราดิเอนบูตติ้ง โดยมีค่าความถูกต้องอัลกอริทึมของการวิเคราะห์ผลร้อยละ 99.43 ค่าความแม่นยำเท่ากับ 0.99 ค่าระลึกเท่ากับ 0.99 และค่าการถ่วงดุลเท่ากับ 0.99

คำสำคัญ: เหมืองข้อความ อัลกอริทึมป่าสุ่ม
อัลกอริทึมเอ็กซ์ทรีมกราดิเอนบูตติ้ง
การเรียนรู้เชิงลึก

Abstract

The purpose of this research is to compare the performance of the algorithm. The comparison of Algorithm Performance to Analysis of Hotel Guest services, 3 algorithms, namely Random Forest, Extreme Gradient Boosting (XGBoost), and Deep Learning, were collected from 500 award-winning hotel users' opinions by The text was divided into 250 positive comments and 250 negative comments through the process of text mining and the text was used to create vectors and divided the dataset into 2 sets, namely training data set 70%. and the testing data set of 30% As a result of the evaluation test, the performance of each algorithm is also measured. Accuracy, Precision, Recall, and F1 Score The best algorithm were found to be the Extreme Gradient Boosting (XGBoost) algorithm with 99.43% accuracy, 0.99 precision, 0.99 recall and an F1 Score of 0.99

Keywords: Text Mining, Random Forest, Extreme Gradient Boosting (XGBoost), Deep Learning.

1. บทนำ

อุตสาหกรรมบริการเป็นแรงขับเคลื่อนสำคัญที่ช่วยขับเคลื่อนเศรษฐกิจไทยผ่านการกระตุ้นการเติบโตทางเศรษฐกิจตลอดจนส่งเสริมการลงทุนทั้งในและ

ต่างประเทศ ปี 2559 ภาคบริการขยายตัวร้อยละ 4.3 โดยร้อยละ 10.3 เป็นของโรงแรมและภัตตาคารยังคงเป็นภาคส่วนที่เติบโตเร็วที่สุดในเศรษฐกิจการเติบโต ในภาคส่วนนี้ได้รับแรงหนุนจากจำนวนนักท่องเที่ยวที่เพิ่มขึ้น ซึ่งเพิ่มขึ้นร้อยละ 8.9 ในปี 2559 ซึ่งถือว่าการท่องเที่ยวเป็นปัจจัยสนับสนุนการส่งออกสุทธิรายใหญ่ที่สุดอย่างไรก็ตามการเติบโตสูงอย่างยั่งยืนของการท่องเที่ยวจะกลายเป็นเรื่องท้าทายเนื่องจากข้อจำกัดของการบริการเป็นตัวกำหนดการขับเคลื่อนในภาคของ โรงแรม [1]

ในปี 2560 ประเทศไทยอยู่ในอันดับที่ 34 จากทั้งหมด 136 ประเทศในด้านความสามารถในการแข่งขันด้านการเดินทางและการท่องเที่ยว ตามดัชนีความสามารถในการแข่งขันด้านการเดินทางและการท่องเที่ยวความได้เปรียบในการแข่งขันของประเทศไทยในฐานะจุดหมายปลายทางด้านการท่องเที่ยวมีสาเหตุหลักมาจากทรัพยากรธรรมชาติ ความสามารถในการแข่งขันด้านราคา ด้านการบริการ นักท่องเที่ยวและโครงสร้างพื้นฐานด้านการขนส่งทางอากาศ [2] ดังนั้น เพื่อรักษาความสามารถในการแข่งขันทางการตลาด จึงจำเป็นต้องเข้าใจปัจจัยที่มีผลต่อการตัดสินใจของนักท่องเที่ยวทั้งในประเทศและนอกประเทศ ภายในภาคบริการ โดยเฉพาะ โรงแรมและที่พัก พร้อมกันนี้การท่องเที่ยวแห่งประเทศไทยยังเห็นความสำคัญของการส่งเสริมมาตรฐานโรงแรมระดับสากล จึงได้จัดทำ Thailand Tourism Awards เพื่อช่วยกระตุ้นการรับรู้ของผู้ให้บริการและผู้บริโภค ซึ่งมีผู้วิจัยได้นำเทคนิคการทำเหมืองข้อความมาประยุกต์ใช้ในการวิเคราะห์ความคิดเห็นในเรื่องต่าง ๆ เช่น อัลกอริทึม Random Forest อัลกอริทึม Bayesian Logistic Regression และอัลกอริทึม Support Vector Machine ในการจำแนกความคิดเห็นเกี่ยวกับภาพยนตร์ ร้านอาหารและสินค้า [3] อัลกอริทึม Random Forest อัลกอริทึม Neural Network และอัลกอริทึม Naïve Bayes สำหรับใช้ในการจำแนกความรู้สึกจากข้อความแสดงความคิดเห็นของผู้โดยสารที่ใช้บริการสายการบินของบริษัทในประเทศสหรัฐอเมริกา [4] อัลกอริทึม Naïve Bayes และอัลกอริทึม

k-Nearest Neighbor สำหรับการวิเคราะห์ความคิดเห็นเกี่ยวกับภาพยนตร์และโรงแรม [5]

จากปัญหาดังกล่าวผู้วิจัยได้เห็นปัญหาที่เกิดขึ้นและผลกระทบในเชิงลบที่อาจส่งผลกระทบต่อภาพรวมขององค์กรการท่องเที่ยวแห่งประเทศไทยและโรงแรม จึงศึกษาการเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการโรงแรม ด้วยเทคนิคการประมวลผลภาษาธรรมชาติ เช่น การให้นำหนักของคำ รูปประโยค เป็นต้น นำมาสร้างแบบจำลองโดยใช้ Random Forest, XGBoost, Deep learning และทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยพิจารณาจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าระลึก (Recall) และค่าการถ่วงดุล (F1 Score) เพื่อจำแนกความหมายของข้อความความคิดเห็นในเชิงบวกและเชิงลบ พร้อมทั้งนำความคิดเห็นมาปรับปรุงประสิทธิภาพการให้บริการให้เกิดภาพลักษณ์ที่ดีกับองค์กรต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการศึกษาข้อมูลที่เกี่ยวข้องต่าง ๆ ที่ใช้ในงานวิจัย โดยมีรายละเอียดดังต่อไปนี้

2.1 อัลกอริทึมป่าสุ่ม (Random Forest)

เป็นเทคนิคที่ทำการสุ่มเลือกคุณสมบัติออกมาจากชุดของตัวหลาย ๆ จุด จากนั้นนำเอาชุดของคุณสมบัติเหล่านี้มาสร้างแบบจำลองด้วยเทคนิคต้นไม้ตัดสินใจหลาย ๆ ต้น โดยเทคนิคป่าสุ่มถูกนำเสนอครั้งแรกในปีพุทธศักราช 2538 โดย Tin kam ซึ่งต่อมาเทคนิคนี้ถูกต่อยอดโดย Leo Breiman ลักษณะของต้นไม้ที่อยู่ภายในป่าของเทคนิคการสุ่มป่าไม้จะถูกควบคุมด้วย 3 ปัจจัยคือ

1. ต้นไม้แต่ละต้นจะถูกสอน (Training) โดยการใช้เซตย่อยมาจากข้อมูลกลุ่มตัวอย่าง
2. เมื่อต้นไม้โตขึ้นจะสามารถค้นหาโหนด (Node) แต่ละโหนดที่อยู่ในกิ่งที่ดีที่สุดของต้นไม้โดยใช้การสุ่มเล็กจากคุณสมบัติ N คุณสมบัติ
3. ต้นไม้แต่ละต้นจะไม่มีการตัดออก แต่จะปล่อยให้ต้นไม้โตขึ้นไปเรื่อย ๆ จนได้ผลลัพธ์ที่ดีที่สุดหลังจาก

การสร้างป่าแล้วทำการให้คะแนน (Vote) โดยต้นไม้ภายในป่า หากต้นไม้ต้นใดได้คะแนนสูงสุดก็จะเอาต้นไม้ที่ออกมาสร้างเป็นโมเดล [6]

2.2 อัลกอริทึมเอ็กซ์ทรีมกราดิเอนท์บูตติ้ง (XGBoost)

เป็นเทคนิคที่มีพื้นฐานมาจากการเรียนรู้แบบ Ensemble เช่นเดียวกับเทคนิคป่าแบบสุ่มป่าไม้ กล่าวคือเป็นการเรียนรู้จากหลายโมเดลร่วมกันแต่ Gradient Boosting จะเลือกใช้เทคนิคการเรียนรู้แบบทำซ้ำตามลำดับ ซึ่งจะเป็นการปรับปรุงโมเดลไปทีละขั้นในแต่ละรอบของการเรียนรู้ ส่งผลให้มีความแม่นยำที่สูงกว่าแต่ก็สามารถเกิดปัญหา Overfitting ได้ง่าย ในส่วนของ Extreme Gradient Boosting หรือ XGBoost เป็นเทคนิคที่ปรับปรุงมาจาก Gradient Boosting เพื่อให้สามารถทำงานได้อย่างรวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น รวมทั้งยังสามารถลดเวลาการเกิด Overfitting ได้ [7]

2.3 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึกที่เป็นส่วนหนึ่งของการเรียนรู้โครงข่ายประสาทเทียม [8] เป็นวิธีที่สร้างขึ้นเพื่อให้เครื่องสามารถเรียนรู้ได้โดยใช้ต้นแบบมาจากระบบประสาทของมนุษย์ โดยต้องใส่ข้อมูลเข้าไปในชั้นรับข้อมูล (Input Layer) จากนั้นเครื่องจะนำข้อมูลไปประมวลผลในชั้นซ่อน (Hidden Layer) แล้วจะนำเสนอข้อมูลผลลัพธ์ในชั้นแสดงผล (Output Layer) [9] การเรียนรู้เชิงลึกถือว่าเป็นวิธีการที่มีศักยภาพสูงในการจัดการกับพีเอเจอร์สำหรับการเรียนรู้แบบไม่มีผู้สอนหรือการเรียนรู้แบบกึ่งมีผู้สอน โดยจะมีชั้นของโครงข่ายประสาทเทียมหลายชั้น เพื่อทำให้เกิดประสิทธิภาพมากขึ้น

2.4 งานวิจัยที่เกี่ยวข้อง

Hirokawa, S. and Hashimoto, K. [10] นำเสนอวิธีการจำแนกประเภทการวิเคราะห์ความคิดเห็นด้วยวิธีการ SVM จำนวน 73,589 ความคิดเห็น โดยแบ่งออกเป็น 2 ประเภท คือ ความคิดเห็นในเชิงบวกและความคิดเห็นในเชิงลบ โดยแบ่งออกในเชิงบวกร้อยละ 74 และเชิงลบร้อยละ 24 ของข้อความทั้งหมด

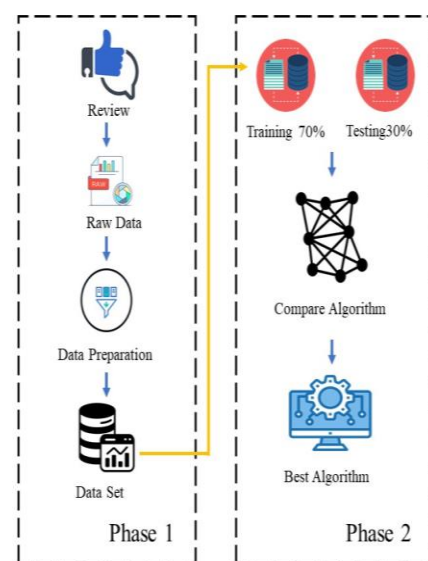
พีระสวด ทองพ่วง และคณะ [11] งานวิจัยนี้วิเคราะห์ปัจจัยที่เกี่ยวข้องกับข้อมูลทุนการศึกษาของนักศึกษาระดับ

ปริญญาตรีและเปรียบเทียบประสิทธิภาพของเทคนิคการจุดข้อมูลที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ Decision Tree, Naïve Bayes, K-Nearest Neighbors, Deep Learning และ Random Forest รวบรวมข้อมูลจากแบบรับทุนของนักศึกษาระดับปริญญาตรี ผลการวิจัยพบว่า ความแม่นยำของ Random Forest, Deep Learning, Decision Tree, Naïve Bayes และ K-Nearest Neighbors อยู่ที่ร้อยละ 94.28, 93.51, 92.64, 92.47 และ 89.70 ตามลำดับ

พวง มีสังข์ และคณะ [12] ได้นำเสนอเทคนิคทางปัญญาประดิษฐ์ที่ใช้สำหรับการตรวจข่าวปลอมภาษาไทย ผลการวิจัยพบว่า LSTM ให้ประสิทธิภาพค่าความถูกต้องร้อยละ 93.81

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์ในการเปรียบเทียบประสิทธิภาพของอัลกอริทึม สำหรับการวิเคราะห์ความคิดเห็นโรงแรม โดยผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

3.1 การรวบรวมและการจัดการข้อมูล

การเก็บข้อมูลเก็บจากชุดข้อมูลที่ให้เข้าถึงสาธารณะจากเว็บไซต์ <https://www.booking.com/index.th.html> ผู้วิจัยรวบรวมความคิดเห็นที่แสดงต่อการใช้บริการของโรงแรมรางวัลที่พักยอดเยี่ยมประเภทที่พักนักท่องเที่ยวจำนวน 500 ข้อความ จาก 6 โรงแรมที่ได้รับรางวัล

Thailand Tourism Awards ด้าน โรงแรมโดยมีการแบ่งออกเป็นความคิดเห็นเชิงบวก จำนวน 250 ข้อความ และความคิดเห็นเชิงลบ จำนวน 250 ข้อความ ดังตัวอย่าง ข้อมูลแสดงดังตารางที่ 1

ตารางที่ 1: ตัวอย่างข้อมูล

ข้อความ	ความคิดเห็น
(Y_Y)The service is not good. ☹️ ☹️	Negative
The room is smelly. !!!	Negative
(😊) Very good.	Positive
I like the service from the hotel. 🤝	Positive

จากนั้นนำข้อความมาผ่านกระบวนการจัดเตรียมข้อมูล ประกอบด้วย 4 ขั้นตอน ดังนี้

3.1.1 Text cleaning ในกระบวนการขั้นตอนนี้จะเป็นการนำประโยคที่ได้เข้ามาตรวจสอบอักษรหรืออักขระพิเศษหากประโยคนั้น ๆ หากแฝงไปด้วยสิ่งเหล่านี้ จะทำการลบออกจากประโยค

3.1.2 Stop word Removal เนื่องจากข้อมูลที่ผู้วิจัยได้เก็บนำมาใช้งานนั้นเป็นภาษาอังกฤษจึงต้องใช้คุณลักษณะการเว้นวรรคในการตัดคำออกจากประโยค โดยการนำเอาคำเชื่อมประโยคออก เช่น it, is, and, or และ for เป็นต้น

3.1.3 การสร้างดัชนีค่าสำคัญ TF-IDF Weighting อยู่ในรูปแบบของเวกเตอร์ (Word Vector) โดยค่าน้ำหนักของคำหาได้จากการเปรียบเทียบคำสำคัญกับคำในพจนานุกรม แล้วนับจำนวนคำสำคัญที่พบในคำพจนานุกรมเพื่อนำความถี่มาคำนวณหาค่าน้ำหนักดังตารางที่ 2

ตารางที่ 2: ตัวอย่างขั้นตอน TF-IDF

Text	service	good	room	smelly	Very	...	Hotel
The service is not good	0.434	0.400					
The room is smelly			0.599	1.304			
Very good		0.533			0.520		

I like the service from the hotel	0.290						0.280
-----------------------------------	-------	--	--	--	--	--	-------

3.1.4 การแบ่งข้อมูล ได้แก่ การนำข้อความความคิดเห็นมาผ่านกระบวนการเตรียมข้อมูลเรียบร้อยแล้ว จึงทำการแบ่งออกเป็น 2 ชุด ได้แก่ ชุดข้อมูลสอนร้อยละ 70 จำนวน 350 ข้อความ และชุดข้อมูลทดสอบร้อยละ 30 จำนวน 150 ข้อความ เพื่อนำเข้าสู่การสร้างแบบจำลอง

3.2 การสร้างแบบจำลองและการเปรียบเทียบประสิทธิภาพ

อัลกอริทึมที่ใช้ในการสร้างแบบจำลอง ประกอบด้วย 3 อัลกอริทึม ได้แก่ Random Forest, XGBoost, Deep learning โดยมีการกำหนดค่าพารามิเตอร์ที่ใช้สำหรับการทดสอบอัลกอริทึมทั้ง 3 อัลกอริทึม และทำการปรับค่าพารามิเตอร์ตามที่กำหนดไว้ในแต่ละอัลกอริทึมแสดงดังตารางที่ 3 ต่อไปนี้

ตารางที่ 3: ค่าพารามิเตอร์

Algorithm	Parameter
Random Forest	Number of trees = 100 Criterion = gain_ratio Maximal_depth = 10
XGBoost	Number of trees = 10 Learning_rate = 0.1 Maximal_depth = 9
Deep learning	Learning_rate = 0.01 Activation = Tanh, Rectifier, Maxout, ExpRectific Epochs = 1.0-10.0

ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึม โดยมีการพิจารณาค่าความถูกต้อง ความแม่นยำ ค่าระยะ และค่าเฉลี่ย

4. ผลการดำเนินการวิจัย

ผลการทดสอบประสิทธิภาพ โดยมีการกำหนดค่าพารามิเตอร์ของอัลกอริทึม แสดงดังตารางที่ 4 ถึง 6 ดังรายละเอียดต่อไปนี้

ตารางที่ 4: ผลการทดสอบอัลกอริทึม Random Forest

Parameter			Result
Number of trees	Maximal_depth	Criterion	Accuracy
60	8	gain_ratio	98.60
90	8	gain_ratio	98.30
90	9	gain_ratio	97.70
60	7	gain_ratio	97.10
100	3	gain_ratio	93.40
60	3	gain_ratio	90.30
40	3	gain_ratio	89.70
10	5	gain_ratio	80.40
30	2	gain_ratio	73.40
10	2	gain_ratio	61.40
30	1	gain_ratio	49.70
90	1	gain_ratio	49.10

จากตารางที่ 4 ค่าพารามิเตอร์ที่ให้ประสิทธิภาพที่ดีที่สุด มีค่าความถูกต้องเท่ากับร้อยละ 98.60 และมีพารามิเตอร์ Number of trees เท่ากับ 60, Maximal_depth เท่ากับ 8, Criterion เท่ากับ gain_ratio และมีค่าความถูกต้องร้อยละ 98.60 ตามลำดับ

ตารางที่ 5: ผลการทดสอบอัลกอริทึม XGBoost

Parameter			Result
Number of trees	Maximal_depth	Learning_rate	Accuracy
9	90	0.700	99.10
9	100	0.500	99.10
4	20	0.600	98.90
9	70	0.600	98.60
9	40	1.000	98.60
3	100	0.300	92.60

Parameter			Result
Number of trees	Maximal_depth	Learning_rate	Accuracy
4	90	0.100	87.40
1	80	0.300	86.90
1	60	1.000	82.30
1	30	0.700	80.00

จากตารางที่ 5 ค่าพารามิเตอร์ที่ให้ประสิทธิภาพที่ดีที่สุด มีค่าความถูกต้องเท่ากับร้อยละ 99.10 และมีพารามิเตอร์ Number of trees เท่ากับ 9, Maximal_depth เท่ากับ 90, Learning_rate เท่ากับ 0.700

ตารางที่ 6: ผลการทดสอบอัลกอริทึม Deep learning

Parameter			Result
Learning_rate	Activation	epochs	Accuracy
0.778	Tanh	6	99.40
0.112	Tanh	5	99.10
0.556	Tanh	6	98.30
80.89	Maxout	9	98.30
0.778	Tanh	2	98.00
0.778	Rectifier	2	97.70
0.778	Rectifier	6	97.40
1.000	Rectifier	2	96.60
0.334	Tanh	1	95.40
0.889	ExpRectific	1	93.40
0.556	Rectifier	1	90.90

จากตารางที่ 6 ค่าพารามิเตอร์ที่ให้ประสิทธิภาพที่ดีที่สุด มีค่าความถูกต้องเท่ากับร้อยละ 99.40 และมีพารามิเตอร์ Learning_rate เท่ากับ 0.778, Activation เท่ากับ Tanh, epochs เท่ากับ 6

จากการทดลองสร้างตัวแบบการวิเคราะห์ โดยใช้ชุดข้อมูลที่มีการแบ่งเป็น 70 ต่อ 30 ซึ่งมีป้ายกำกับความ

คิดเห็น 2 ระดับ ได้แก่ ความคิดเห็นเชิงบวกและความคิดเห็นเชิงลบ และมีการเปรียบเทียบประสิทธิภาพจากเทคนิคเหล่านี้ ได้แก่ Random Forest, XGBoost, Deep learning สามารถสรุปได้ดังตารางที่ 7

ตารางที่ 7: แสดงการเปรียบเทียบประสิทธิภาพดังนี้

Algorithm	Result			
	Acc	P	R	F1
Random Forest	98.57	0.98	0.98	0.98
XGBoost	99.43	0.99	0.99	0.99
Deep learning	98.57	0.98	0.98	0.98

จากตารางที่ 7 เมื่อพิจารณาผลการทดลองแล้ว อัลกอริทึมที่ให้ผลลัพธ์ในการทดลองนี้ดีที่สุด ได้แก่ XGBoost ค่าความถูกต้องอัลกอริทึมของการวิเคราะห์ผลร้อยละ 99.43 ค่าความแม่นยำเท่ากับ 0.99 ค่าระลอกเท่ากับ 0.99 ค่าการถ่วงดุลกับ 0.99

5. สรุป

จากงานวิจัยฉบับนี้ผู้วิจัยได้นำเสนอการเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้บริการ โรงแรม จำนวน 500 ข้อความ แบ่งออกเป็นข้อความคิดเห็นเชิงบวก จำนวน 250 ข้อความ และข้อความคิดเห็นเชิงลบ จำนวน 250 ข้อความ โดยการนำข้อความเหล่านั้นไปผ่านกระบวนการจัดเตรียมข้อมูลให้มีความสมบูรณ์ของข้อความมากขึ้น หลังจากนั้นนำข้อความที่ได้ไปสร้างเวกเตอร์เพื่อนำเข้าสู่แบบจำลองโดยผู้วิจัยได้แบ่งชุดข้อมูลออกเป็น 2 ชุด อันได้แก่ชุดข้อมูลสำหรับการสอนร้อยละ 70 และชุดข้อมูลสำหรับทดสอบร้อยละ 30 และทำการสร้างแบบจำลองวัดประสิทธิภาพของอัลกอริทึม โดยมีการพิจารณาจากค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าถ่วงดุล ทั้งนี้ผลการเปรียบเทียบของอัลกอริทึมพบว่าอัลกอริทึม XGBoost ให้ประสิทธิภาพดีที่สุดจากอัลกอริทึมทั้งหมด โดยมีค่าความถูกต้องอัลกอริทึมของการวิเคราะห์ผล

ร้อยละ 99.43 ค่าความแม่นยำเท่ากับ 0.99 ค่าระลอกเท่ากับ 0.99 ค่าการถ่วงดุลเท่ากับ 0.99

เอกสารอ้างอิง

- [1] World Bank Group. (2017, August 24). Thailand Economic Monitor– August 2017: Digital Transformation. World Bank. Retrieved March 15, 2023, <https://www.worldbank.org/en/country/thailand/publication/thailand-economic-monitor-august-2017-digital-transformation>
- [2] World economic forum. (n.d.). Retrieved March 15, 2023, https://www3.weforum.org/docs/WEF_TTCR_2017_web_0401.pdf
- [3] วัชรวิวัฒน์ จิตต์สกุล และสุนันทา สดสี, “การวิเคราะห์การจำแนกข้อความด้วยการเปรียบเทียบความถี่ของอัลกอริทึม” วารสารฉบับวิทยาศาสตร์และเทคโนโลยี ปีที่ 9 มกราคม – ธันวาคม 2560 หน้า 19-31.
- [4] นันทก ศุทธิเลิศ, “การวิเคราะห์ความรู้สึกผู้โดยสารที่ใช้บริการสายการบินของบริษัทในประเทศสหรัฐอเมริกา” วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยบูรพา, 2560.
- [5] Lopamudra Dey, et al., “Sentiment Analysis of Review Datasets using Naive Bayes and K-NN Classifier.” *International Journal of Information Engineering and Electronic Business*. Vol.8 No.4, pp. 54-62, 2016.
- [6] Noochanat, P. and Jaree, T. 2018. “Classifying Thai opinions on online media using text mining”, *Journal of Science & Technology MSU*, 37(3). (in Thai)
- [7] Chen, T. and C. Guestrin (2016). “Xgboost: A scalable tree boosting system.” *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 10.1145/2939672.2939785.
- [8] J. Schmidhuber, “Deep Learning in Neural Networks: An Overview,” *Neural Networks*. 61: 85–117, 2015.
- [9] B. Yoshua, “Learning Deep Architectures for AI,” *Foundations and Trends in Machine Learning*. pp. 1–127, 2009
- [10] Hirokawa, S., & Hashimoto, K. (2018, พฤศจิกายน). ความเรียบง่ายของบทวิจารณ์เชิงบวกและความหลากหลายของบทวิจารณ์เชิงลบในชื่อเสียงของโรงแรม ในปี 2018 *International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAINLP)* (หน้า 1-6) อีอีซี
- [11] พิระสุด ทองพ่วง และคณะ “การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูล เพื่อ ทำนายการได้รับทุนการศึกษาของนักศึกษาระดับปริญญาตรีโดยใช้เทคนิควิธีการ ทำเหมืองข้อมูล” *Journal of Professional Routine to Research*, 2021, 8.2: 44-52.
- [12] พยุง มีสัง และคณะ, “เทคนิคทางปัญญาประดิษฐ์สำหรับการตรวจข่าวปลอม ภาษาไทย” *The Journal of Applied Science*. vol.21, หน้า 1-19.

การเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการพยากรณ์อัตราเงินเฟ้อของไทย

Comparison of Algorithm Performance for Forecasting Thailand's Inflation

อรอมา จำเมือง (Onuma Jamuang)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการศึกษาระดับปริญญาโท โดไอสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021910032@email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการพยากรณ์อัตราเงินเฟ้อของไทย โดยใช้ชุดข้อมูลเศรษฐกิจมหภาคจากกองทุนการเงินระหว่างประเทศ ระหว่างปี พ.ศ.2523 – 2564 รวม 42 ปี จำนวน 966 ข้อมูล แบ่งข้อมูลเป็น 2 ชุด คือ ชุดข้อมูลฝึกฝนจำนวน 30 ปี และชุดข้อมูลทดสอบจำนวน 12 ปี ทำการเลือก 4 อัลกอริทึมสำหรับการทดสอบ ได้แก่ การเรียนรู้เชิงลึก ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย การวิเคราะห์การถดถอย และ เอ็กซ์ทรีมกราดิเอนท์บูตติ้ง วัดประสิทธิภาพการพยากรณ์ด้วยรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย โดยผลการวิจัย พบว่า เอ็กซ์จีบูตติ้งมีประสิทธิภาพสูงที่สุดในการพยากรณ์อัตราเงินเฟ้อ เมื่อเทียบกับอัลกอริทึมอื่นที่มีค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ดังนี้ 0.0004 และ 0.0003 การวิเคราะห์การถดถอย 0.4817 และ 0.4018 การเรียนรู้เชิงลึก 2.1396 และ 1.2314 และ ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย 3.3047 และ 1.6449 ตามลำดับ

คำสำคัญ: การพยากรณ์ เงินเฟ้อ การเรียนรู้เชิงลึก
เอ็กซ์ทรีมกราดิเอนท์บูตติ้ง การวิเคราะห์การ
ถดถอยเชิงเส้น ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

This research aims to compare machine learning algorithm performance for predicting Thailand's

Inflation using the macroeconomic data set from International Monetary Fund (IMF). Used data between 1980-2021 totaling 42 years. The data was divided into 2 sets, 30-year training set and a 12-year test set (966 records). Four algorithms were selected for testing: Deep Learning, Support Vector Machine for Regression, Linear Regression, and Extreme Gradient Boosting (XGBoost). Furthermore, we used root mean square error and mean absolute error to evaluate the algorithm's performance. In this research, the result showed that XGBoost was efficient in predicting inflation, compared to other algorithms with the value of the error RMSE = 0.0004 and MAE = 0.0003 Linear Regression RMSE = 0.4817 and MAE = 0.4018, Deep Learning RMSE = 2.1396 and MAE = 1.2314 and Support Vector Machine for Regression RMSE = 3.3047, and MAE = 1.6449 respectively.

Keywords: Prediction, Inflation, Deep learning, Extreme Gradient Boosting, Linear Regression, Support Vector Machine.

1. บทนำ

เงินเฟ้อ (Inflation) มีความสำคัญในระบบเศรษฐกิจ และมีผลกระทบต่อทั้งในด้านการดำรงชีวิต ด้านการลงทุน และการกำหนดนโยบายทางการเงินต่างๆ ซึ่งกระบวนการจัดทำข้อมูลเงินเฟ้อได้มีการสำรวจราคาสินค้าและบริการในพื้นที่ต่างๆ ผ่านแหล่งจัดเก็บข้อมูล โดยจะมีการจัดเก็บ

บันทึก และตรวจสอบสินค้าใน 5 ภูมิภาคของประเทศไทย เพื่อนำมาประมวลผลของราคาที่เกิดขึ้น เพื่อชี้วัดการเปลี่ยนแปลงหรือการเคลื่อนไหวของราคาสินค้าและบริการ ซึ่งจะมีผลกระทบต่อระบบเศรษฐกิจ

ดังนั้น หากมีเครื่องมือที่สามารถช่วยในการพยากรณ์อัตราเงินเฟ้อล่วงหน้าด้วยวิธีทางเทคโนโลยีสารสนเทศ เพื่อเป็นข้อมูลให้ประชาชน ธุรกิจ และภาครัฐบาลสามารถนำไปใช้เพื่อให้ทราบถึงแนวโน้มการเปลี่ยนแปลงของธุรกิจหรือสภาพแวดล้อมที่จะมีผลต่อองค์กรได้ในอนาคต และทำให้สามารถที่จะวางแผนกำหนดนโยบายเพื่อให้บรรลุวัตถุประสงค์ได้ จากการสำรวจงานวิจัยที่เกี่ยวข้องกับการคำนวณปัจจัยทางเศรษฐกิจต่าง ๆ ดังนี้ Serkan Aras และ Paulo J.G. Lisboa [1] ได้กล่าวว่า การพยากรณ์อัตราเงินเฟ้ออย่างแม่นยำในภาวะที่เต็มไปด้วยข้อมูลต่าง ๆ มากมาย พบว่า อัลกอริทึมของ Machine Learning มีส่วนสำคัญในการพยากรณ์ผลลัพธ์ของอัตราเงินเฟ้อได้อย่างรวดเร็วและแม่นยำได้ รวมถึง Yu-Shan Li, Ping-Feng Pai และ Ying-Lei Lin [2] ได้เสนอแนวทางการพยากรณ์อัตราเงินเฟ้อด้วยอัลกอริทึม XGBoost ซึ่งมีแนวโน้มที่ดีสำหรับนำมาใช้ในการพยากรณ์อัตราเงินเฟ้อ ส่วน Anna Almosova and Niek Andresen [3] ได้ใช้อัลกอริทึม Deep Neural Network ในการพยากรณ์อนุกรมเวลาทางเศรษฐกิจโดยทั่วไปและอัตราเงินเฟ้อได้อย่างมีประสิทธิภาพ ซึ่งอัลกอริทึมที่ใช้ในการพยากรณ์แต่ละกระบวนการนั้น พิจารณาโดยวัดจากค่าความคลาดเคลื่อนในการพยากรณ์ สำหรับค่าที่นิยมใช้ คือ ค่า RMSE และค่า MAE

จากปัญหาข้างต้นผู้วิจัยจึงมีแนวคิดในการเปรียบเทียบอัลกอริทึมสำหรับการพยากรณ์อัตราเงินเฟ้อของประเทศไทยด้วยวิธีการเรียนรู้เชิงลึก (Deep learning) ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) และเอ็กซ์ทรีมกาเดียนบูตติ้ง (XGBoost)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง

โดยแบ่งเป็น 2 หัวข้อหลัก ดังนี้

2.1 อัลกอริทึมการพยากรณ์

สำหรับงานวิจัยนี้ใช้อัลกอริทึม 4 อัลกอริทึมในการพยากรณ์ดังนี้

2.1.1 การเรียนรู้เชิงลึก (Deep learning)

การเรียนรู้เชิงลึก คือ โครงข่ายประสาทเทียมที่มีหลายชั้นระหว่างชั้นอินพุตและเอาต์พุต [4] มีโครงข่ายประสาทเทียมประเภทต่างๆ กัน แต่มักจะประกอบด้วย ส่วนประกอบเดียวกันเสมอ ได้แก่ เซลล์ประสาท (neurons) ไซแนปส์ (synapses) น้ำหนัก (weights) ความเอนเอียง (biases) และฟังก์ชัน (Function) ทางคณิตศาสตร์ ส่วนประกอบทั้งหมดเหล่านี้ทำงานคล้ายกับสมองของมนุษย์ และสามารถฝึกฝนได้เหมือนกับอัลกอริทึมอื่น ๆ ซึ่งการเรียนรู้สามารถเป็นได้ทั้งแบบการเรียนรู้แบบมีผู้สอน และการเรียนรู้แบบไม่มีผู้สอน [5] โดยจะมีชั้นของโครงข่ายประสาทเทียมหลายชั้นเพื่อทำให้เกิดประสิทธิภาพมากขึ้น

2.1.2 ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression)

ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย หรือ SVR เป็นอัลกอริทึมที่ใช้ในการจำแนกกลุ่มของข้อมูลด้วยวิธีการหาระนาบการตัดสินใจ (Decision Hyperplane) หรือไฮเปอร์เพลนที่เหมาะสมสำหรับการแบ่งข้อมูล 2 ส่วนจากกัน SVM ยังถูกนำมาประยุกต์ใช้ในการสร้างสมการในการประมาณการหาค่าฟังก์ชันเชิงเส้น $f(x)$ ในการใช้ระนาบตัดสินใจซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย หรือ SVR จะเป็นการนำข้อมูลปัจจุบัน และข้อมูลในอดีตจำนวนหนึ่งมาทำการเรียนรู้ (Training) เพื่อให้ทราบรูปแบบสำหรับพยากรณ์ผลซึ่งจะเกิดขึ้นในอนาคต ในงานวิจัยนี้ผู้วิจัยใช้วิธีสร้างสมการค่าฟังก์ชันเชิงเส้นด้วยวิธี Sequential Minimal Optimization for SVM Regression (SMOreg) [6-7]

2.1.3 การวิเคราะห์การถดถอย (Linear Regression)

เป็นวิธีการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ซึ่งเทคนิคนี้จะอาศัย

ความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรเพื่อใช้ทำนาย โดยสมการ Linear Regression จะแสดงให้เห็นค่าความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระในลักษณะของข้อมูลเชิงปริมาณ โดยสามารถเขียนได้ดังนี้

$$Y = a + bx_1 + cx_2 + \dots \quad (1)$$

เมื่อทราบค่าตัวแปรอิสระก็สามารถทำนายตัวแปรตามได้ โดยค่าตัวแปรตาม คือ Y และ x คือ ค่าตัวแปรอิสระ

2.1.4 เอ็กซ์ทรีมกราดิเอนท์บูตติ้ง (Extreme Gradient Boosting)

XGBoost มีพื้นฐานมาจาก Gradient Boosting ถูกจัดอยู่ในกลุ่มของการจำแนกประเภทแบบ Boosting ประเภทหนึ่ง ที่ได้รับการพัฒนาให้มีประสิทธิภาพมากขึ้น จึงทำให้สามารถทำงาน ได้รวดเร็วและมีประสิทธิภาพมากขึ้น สามารถจัดการหน่วยการทำงานย่อยแบบ Multithread ได้ดี และการทำให้ตัวแปร มีค่าอยู่ใน ระดับปกติ (Regularization) เป็นการป้องกันการเข้ากันได้ดีเกินไป (Over fit) ผลการเรียนรู้จะได้เป็นแบบจำลองที่นำเอาต้นไม้ตัดสินใจมาสร้างให้เกิดการเรียนรู้แบบต่อเนื่อง ต่อ ๆ กัน การเรียนรู้จากการสร้างต้นไม้ตัดสินใจหลาย ๆ ต้น จะทำให้เกิดการเรียนรู้โดยสนใจจากค่าความคลาดเคลื่อน หรือเกิดการเรียนรู้จากความ คลาดเคลื่อนของต้นไม้ตัดสินใจที่สร้างได้ก่อนหน้า ทำให้ความแม่นยำของการพยากรณ์แม่นยำมากขึ้นเรื่อย ๆ เมื่อการเรียนรู้ของต้นไม้ตัดสินใจดำเนินอย่างต่อเนื่องกันจนมีความลึกมากพอหรือไม่มีรูปแบบอื่น ๆ ให้เรียนรู้เพิ่ม ระบบจะหยุดการเรียนรู้ แบบจำลองที่ได้จากอัลกอริทึม XGBoost จะใช้แบบจำลองหลาย ๆ ตัวแบบมาประกอบกันเป็นแบบจำลองที่ซับซ้อนขึ้น ดังเช่นในงานวิจัยของ Chen & Guestrin. [8] ที่นำไปใช้ในการพยากรณ์ผู้ต้องสงสัยที่ไม่ประสงค์ต่อระบบ

2.2 งานวิจัยที่เกี่ยวข้อง

L. Zhang and J. Li [9] ได้นำเสนอการพยากรณ์อัตราเงินเพื่อมีบทบาทสำคัญในนโยบายการเงินและชีวิตประจำวัน การศึกษานี้เน้นในเรื่องการพัฒนาแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (SVR) เพื่อคาดการณ์ดัชนีราคาผู้บริโภค ใช้ข้อมูล

ย้อนหลังของ Money Gap และ CPI ในการคาดการณ์ นอกจากนี้ยังใช้ปรับปรุงคัดเลือกรามิเตอร์ของ SVR นอกจากนี้ยังตรวจสอบความเป็นไปได้ของการประยุกต์ใช้ SVR ในการพยากรณ์อัตราเงินเพื่อโดยเปรียบเทียบกับโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Backpropagation) และการวิเคราะห์การถดถอย ผลลัพธ์แสดงให้เห็นว่า SVR เป็นทางเลือกที่มีแนวโน้มดีกว่าการคาดการณ์อัตราเงินเพื่อ

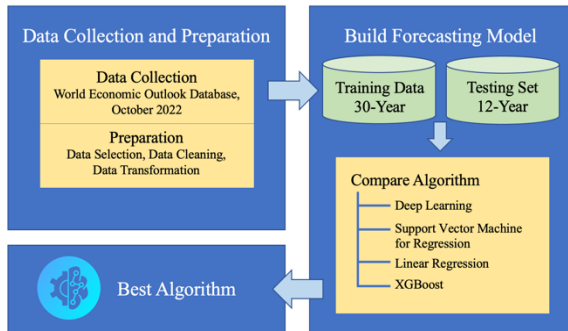
R. Zhang [10] งานวิจัยนี้ใช้เทคโนโลยีการทำเหมืองข้อมูลเพื่อตรวจสอบสาเหตุที่เป็นไปได้ของความไม่เท่าเทียมกันของรายได้ เพื่อให้เข้าใจความเชื่อมโยงระหว่างอัตราเงินเพื่อและการเติบโตของความไม่เท่าเทียมกันของรายได้ได้ดียิ่งขึ้น โดยวิเคราะห์ชุดข้อมูลในอดีตของอัตราเงินเพื่อ เช่น ดัชนีราคาผู้บริโภค ค่าสัมประสิทธิ์ Gini ดัชนีราคาผู้ผลิต อัตราการว่างงาน ค่าจ้างขั้นต่ำ ผลิตภัณฑ์มวลรวมภายในประเทศ อัตราภาษีของรัฐบาลกลาง เป็นต้น เพื่อหาข้อสรุป และเพื่อพัฒนาแบบจำลองข้อมูลที่วิเคราะห์ข้อมูล โดยได้ใช้ การเรียนรู้เชิงลึกโครงข่ายประสาทเทียมแบบหลายชั้น ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย ผลการวิจัยพบว่า การเรียนรู้เชิงลึกโครงข่ายประสาทเทียมแบบหลายชั้นสามารถพยากรณ์อัตราเงินเพื่อ ได้แม่นยำถึงร้อยละ 95.1

Araujo, G. S., and Gaglianone, W. P. [11] งานวิจัยนี้ทำการสำรวจวิธีการ Machine Learning เพื่อปรับปรุงการคาดการณ์อัตราเงินเพื่อในบราซิล ได้ออกแบบงานวิจัยด้วยฐานข้อมูลขนาดใหญ่จำนวน 501 ชุด และวิธีการพยากรณ์ 50 วิธี รวมถึงเทคนิค ML ใหม่ที่นำเสนอเพื่อทดแทนการคาดการณ์แบบจำลองเศรษฐมิติแบบดั้งเดิม และได้มีการทำการปรับปรุงตัวแปรสำคัญในการคาดการณ์อัตราเงินเพื่อ ซึ่งผลลัพธ์บ่งชี้ว่าวิธี ML สามารถทำได้ในหลายกรณี และมีประสิทธิภาพดีกว่าแบบจำลองทางเศรษฐมิติแบบดั้งเดิมจากผลของค่าเฉลี่ยความผิดพลาดกำลังสองที่น้อยที่สุด โดยผู้วิจัยได้ชี้ว่าการคาดการณ์อัตราเงินเพื่อที่ดีที่สุด จะประกอบด้วยอัลกอริทึมที่ผสมกันของวิธี Random Forest และ Xgboost โดยการค้นพบนี้มีส่วนสนับสนุนต่อการพยากรณ์เศรษฐกิจมหภาค

โดยเฉพาะอย่างยิ่งมุ่งเน้นไปที่เรื่องอัตราเงินเฟ้อของบราซิล

3. วิธีการดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์ในการเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการพยากรณ์อัตราเงินเฟ้อของไทย โดยทำการสร้างแบบจำลอง และเปรียบเทียบประสิทธิภาพ



ภาพที่ 1: กรอบงานวิจัย

3.1 การรวบรวมและการจัดการข้อมูล

งานวิจัยนี้ได้ทำการเก็บรวบรวมข้อมูลทางเศรษฐศาสตร์มหภาคจาก International Monetary Fund (IMF), World Economic Outlook Database, October 2022 [12] ผ่านทาง <https://www.imf.org/> ระหว่างปี พ.ศ.2523 – 2564 รวมทั้งสิ้น 42 ปี จำนวน 966 ข้อมูล ซึ่งกระบวนการรวบรวม และเตรียมข้อมูลสามารถแบ่งย่อยออกได้เป็น 3 กระบวนการย่อย ได้แก่ 1) การคัดเลือกข้อมูล (Data Selection) โดยมีการคัดเลือกข้อมูลที่นำมาใช้ในการวิจัย และเมื่อได้ข้อมูลมาแล้วก็นำเข้าสู่กระบวนการที่ 2) การทำความสะอาดข้อมูล (Data Cleaning) ทำการตรวจสอบและปรับปรุงข้อมูลบางส่วนที่สูญหาย (Missing value) และปรับปรุงรูปแบบข้อมูลที่สูญหาย 3) การแปลงรูปแบบข้อมูล (Data Transformation) โดยแปลงเพิ่มข้อมูลเพื่อให้เหมาะสมที่จะนำไปใช้สำหรับการพยากรณ์ ข้อมูลที่นำมาใช้ในงานวิจัยนี้คัดจากข้อมูลเศรษฐกิจที่สำคัญของประเทศไทยตามรายงานของ IMF โดยคัดเลือกข้อมูลที่มีการรายงานข้อมูลอย่างต่อเนื่องในช่วง 42 ปี และได้คัดเลือก class feature/attribute ข้อมูลดังกล่าวมาใช้ในการวิจัย

3.2 การสร้างแบบจำลองการพยากรณ์

ทำการสร้างแบบจำลองการพยากรณ์ด้วยอัลกอริทึมทั้ง 4 อัลกอริทึม คือ วิธีการเรียนรู้เชิงลึก ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย การวิเคราะห์การถดถอยเชิงเส้น และ XGBoost เริ่มจากการแยกข้อมูลแบ่งออกเป็น 2 ส่วน คือ ชุดฝึกโดยใช้ข้อมูลทั้งหมด 30 ปี และชุดทดสอบใช้ข้อมูลทั้งหมด 12 ปี เพื่อทดสอบประสิทธิภาพของอัลกอริทึมโดยใช้ข้อมูลอินพุตของ 23 แอตทริบิวต์สำหรับการสร้างแบบจำลองการพยากรณ์ กระบวนการเริ่มต้นด้วยการกำหนดค่าพื้นฐานของแต่ละอัลกอริทึม หลังจากนั้นจะปรับพารามิเตอร์ให้เหมาะสมในระดับการกำหนดค่าขั้นสูง เพื่อให้ได้ประสิทธิภาพที่ดีที่สุดของโมเดลการพยากรณ์ ในงานวิจัยนี้มีการกำหนดช่วงของพารามิเตอร์เพื่อทดสอบและคัดเลือกช่วงที่เหมาะสมสำหรับการพยากรณ์อัตราเงินเฟ้อ โดยกำหนดช่วงตามรายละเอียดดังตารางที่ 1

ตารางที่ 1: ค่าพารามิเตอร์สำหรับการทดสอบอัลกอริทึม

Algorithm	Parameter
DL	activation = Tanh, Rectifier, Maxout, ExpRectifier epochs = 1.0-10.0
SVR	svm type = epsilon-SVR, nu-SVR kernel type = rbf, linear, sigmoid degree = 1.0-100.0 gamma = 0.0-100.0 coef0= 1.0-100.0
LR	ridge = 0.0-100.0 feature selection = none, M5 prime, greedy, T-Test, Iterative T-Test min tolerance = 0.0-1.0
XGBoost	booster = tree booster, linear boots, DART rounds = 1.0-100.0 learning rate = 0.0-1.0 min split loss = 0.0-10.0

3.3 การวิเคราะห์เปรียบเทียบอัลกอริทึมการพยากรณ์

การประเมินประสิทธิภาพของอัลกอริทึมโดยใช้การวัดประสิทธิภาพของโมเดลเพื่อเปรียบเทียบผลลัพธ์ที่แท้จริงกับผลลัพธ์จากการพยากรณ์ ประกอบด้วยค่า RMSE และค่า MAE โดยผลการทดสอบจะแสดงให้เห็นว่าอัลกอริทึมใดสามารถพยากรณ์ผลลัพธ์ได้ โดยมีข้อผิดพลาดต่ำ

4. ผลการดำเนินการวิจัย

ผลการดำเนินงานวิจัยจากการปรับค่าพารามิเตอร์ เพื่อให้ได้ประสิทธิภาพที่ดีที่สุดของแต่ละอัลกอริทึม ดังนี้

ตารางที่ 2: ผลการทดสอบประสิทธิภาพอัลกอริทึม DL

Iteration	activation	epochs	RMSE	MAE
30	Rectifier	7	3.0531	2.0349
31	Maxout	7	3.1535	2.3236
32	ExpRectifier	7	2.9516	1.9915
33	Tanh	8	3.2253	1.9845
34	Rectifier	8	3.4017	2.0821
35	Maxout	8	2.1396	1.2314
36	ExpRectifier	8	3.0168	2.0862
37	Tanh	9	2.6849	1.7681
38	Rectifier	9	2.6707	1.6400
39	Maxout	9	3.5520	2.9280
40	ExpRectifier	9	3.1817	2.1197

จากตารางที่ 2 ค่าพารามิเตอร์ของ DL ที่ให้ประสิทธิภาพดีที่สุด คือ activation เท่ากับ Maxout และ epochs เท่ากับ 8 โดยให้ค่าความคลาดเคลื่อนที่ต่ำที่สุดที่ RMSE เท่ากับ 2.1396 และ MAE เท่ากับ 1.2314

ตารางที่ 3: ผลการทดสอบประสิทธิภาพอัลกอริทึม SVR

Iteration	svm type	kernel type	degree	gamma	coef0	RMSE	MAE
1000	nu-SVR	sigmoid	11	40	10.9	3.9128	2.5142
1001	epsilon-SVR	rbf	11	40	10.9	3.3047	1.6449
1002	nu-SVR	rbf	11	40	10.9	3.4272	2.0114
1003	epsilon-SVR	linear	21	40	10.9	3.8495	2.5004
1004	nu-SVR	linear	21	40	10.9	3.8085	2.4981
1005	epsilon-SVR	sigmoid	21	40	10.9	3.8758	2.5044
1006	nu-SVR	sigmoid	21	40	10.9	3.9128	2.5142
1007	epsilon-SVR	rbf	21	40	10.9	3.3047	1.6449
1008	nu-SVR	rbf	21	40	10.9	3.4272	2.0114
1009	epsilon-SVR	linear	31	40	10.9	3.8495	2.5004
1010	nu-SVR	linear	31	40	10.9	3.8085	2.4981

จากตารางที่ 3 ค่าพารามิเตอร์ของ SVR ที่ให้ประสิทธิภาพดีที่สุด คือ svm type เท่ากับ epsilon-SVR kernel type เท่ากับ rbf degree เท่ากับ 11 gamma เท่ากับ 40 และ coef0 เท่ากับ 10.9 โดยให้ค่าความคลาดเคลื่อนที่ต่ำที่สุดที่ RMSE เท่ากับ 3.3047 และ MAE เท่ากับ 1.6449

ตารางที่ 4: ผลการทดสอบประสิทธิภาพอัลกอริทึม LR

Iteration	ridge	feature selection	min tolerance	RMSE	MAE
10	90	M5 prime	0	1.2445	0.9997
11	100	M5 prime	0	1.2531	1.0043
12	0	none	0	0.4817	0.4018
13	10	none	0	0.8168	0.6724
14	20	none	0	0.9223	0.7563
15	30	none	0	0.9816	0.8081
16	40	none	0	1.0211	0.8410
17	50	none	0	1.0502	0.8643
18	60	none	0	1.0729	0.8819
19	70	none	0	1.0914	0.8958
20	80	none	0	1.1071	0.9072

จากตารางที่ 4 ค่าพารามิเตอร์ของ LR ที่ให้ประสิทธิภาพดีที่สุด คือ ridge เท่ากับ 0 feature selection เท่ากับ none และ min tolerance เท่ากับ 0 โดยให้ค่าความคลาดเคลื่อนที่ต่ำที่สุดที่ RMSE เท่ากับ 0.4817 และ MAE เท่ากับ 0.4018

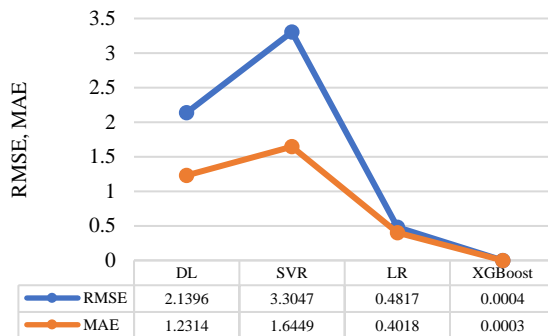
ตารางที่ 5: ผลการทดสอบประสิทธิภาพอัลกอริทึม XGBoost

Iteration	booster	rounds	learning rate	min split loss	RMSE	MAE
340	tree booster	31	1	0	0.0004	0.0003
341	linear booster	31	1	0	2.8543	1.8581
342	DART	31	1	0	0.0004	0.0003
343	tree booster	41	1	0	0.0004	0.0003
344	linear booster	41	1	0	2.8037	1.8861
345	DART	41	1	0	0.0004	0.0003
346	tree booster	51	1	0	0.0004	0.0003
347	linear booster	51	1	0	2.7739	1.9058
348	DART	51	1	0	0.0004	0.0003
349	tree booster	60	1	0	0.0004	0.0003
350	linear booster	60	1	0	2.7567	1.9187

จากตารางที่ 5 ค่าพารามิเตอร์ของ XGBoost ที่ให้ประสิทธิภาพดีที่สุด คือ booster เท่ากับ tree booster rounds เท่ากับ 60 learning rate เท่ากับ 1 และ min split loss เท่ากับ 0 โดยให้ค่าความคลาดเคลื่อนที่ต่ำที่สุดที่ RMSE เท่ากับ 0.0004 และ MAE เท่ากับ 0.0003 (โดยมีค่าต่ำกว่าในหลักทศนิยม)

ผู้วิจัยได้เปรียบเทียบประสิทธิภาพของอัลกอริทึมในการพยากรณ์อัตราเงินเฟ้อของประเทศไทยโดยปรับ

อัลกอริทึมที่กำหนดเพื่อให้ได้ผลลัพธ์ที่ดีที่สุดสำหรับแต่ละอัลกอริทึม หลังจากวิเคราะห์ผลลัพธ์แล้ว ผู้วิจัยสามารถตัดสินใจได้ว่าแบบจำลองใดให้ผลลัพธ์ที่ดีที่สุดโดยการวัดด้วยค่า RMSE และ MAE ที่ให้ผลลัพธ์ที่มีค่าน้อยที่สุดเพื่อแสดงถึงค่าความคลาดเคลื่อนต่ำ ดังนั้นอัลกอริทึมที่ให้ค่าดังกล่าวนี้คืออัลกอริทึมการพยากรณ์ที่มีประสิทธิภาพดีที่สุดแสดงในภาพที่ 2



ภาพที่ 2: เปรียบเทียบค่าความคลาดเคลื่อนของอัลกอริทึม

จากภาพที่ 2 พบว่าอัลกอริทึมที่มีประสิทธิภาพดีที่สุดจากงานวิจัยนี้ คือ XGBoost (RMSE = 0.0004, MAE = 0.0003) รองลงมา คือ Linear Regression (RMSE = 0.4817, MAE = 0.4018) Deep Learning (RMSE = 2.1396, MAE = 1.2314) และ Support Vector Machine for Regression (RMSE = 3.3047, MAE = 1.6449) ตามลำดับ

5. สรุป

งานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมสำหรับการพยากรณ์อัตราเงินเฟ้อของไทยด้วย 4 อัลกอริทึม โดยผลการวิจัยพบว่าอัลกอริทึมที่ให้ค่าความคลาดเคลื่อนน้อยที่สุดในการพยากรณ์ คือ XGBoost ซึ่งมีประสิทธิภาพมากที่สุดเมื่อเทียบกับอัลกอริทึมอื่นๆ รองลงมาเป็น Deep Learning, Linear Regression และ Support Vector Machine for Regression ตามลำดับการประเมินประสิทธิภาพของอัลกอริทึมด้วยค่า RMSE และ MAE โดยสามารถนำกระบวนการของงานวิจัยนี้ไปใช้ในการพยากรณ์อัตราเงินเฟ้อของประเทศไทยได้อย่างรวดเร็วและน่าเชื่อถือ โดยเป็นทางเลือกหนึ่งในการจัดทำข้อมูลทางเศรษฐศาสตร์มหภาคของประเทศ

สำหรับงานในอนาคต ผู้วิจัยแนะนำให้มีการปรับปรุงรายละเอียดข้อมูลให้มีความถี่มากขึ้น เช่น ใช้ข้อมูลทางเศรษฐกิจเป็นรายเดือนมาจัดทำ หรือเพิ่มข้อมูลปัจจัยทางเศรษฐกิจอื่นๆ ที่น่าสนใจมาประกอบการพยากรณ์ รวมถึงการปรับปรุงข้อมูลเพื่อให้สามารถพยากรณ์ข้อมูลเศรษฐกิจตัวอื่น ๆ ได้ รวมถึงการนำอัลกอริทึมที่น่าสนใจอื่น ๆ มาทดสอบการพยากรณ์ข้อมูลทางเศรษฐกิจต่อไป

เอกสารอ้างอิง

- [1] Aras, S., Lisboa, P.J.G., "Explainable inflation forecasts by machine learning models Expert Systems with Applications," *Expert Systems with Applications*, vol. 207, art. no. 117982, 2022.
- [2] Yu-Shan Li, Ping-Feng Pai and Ying-Lei Lin., "Forecasting inflation rates by extreme gradient boosting with the genetic algorithm," *Journal of Ambient Intelligence and Humanized Computing*, 2022.
- [3] Almosova, A., & Andresen, N., "Nonlinear inflation forecasting with recurrent neural networks," *Journal of Forecasting*, 1– 20, 2022.
- [4] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*. 61: 85–117, 2015.
- [5] B. Yoshua, "Learning Deep Architectures for AI," *Foundations and Trends in Machine Learning*. pp. 1–127, 2009
- [6] Kavitha, S., Varuna, S., & Ramya, R. "Comparative Analysis on Linear Regression and Support Vector Regression," *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, 1188 – 1193, 2016.
- [7] S. Sun, Z. Cao, H. Zhu, J. Zhao, "A Survey of Optimization Methods from a Machine Learning Perspective," 2019.
- [8] Chen, T. and C. Guestrin., "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016.
- [9] L. Zhang and J. Li, "Inflation Forecasting Using Support Vector Regression," *2012 Fourth International Symposium on Information Science and Engineering*, Shanghai, China, pp. 136-140, 2012.
- [10] R. Zhang, "Applying Data Mining Technology on Inflation Prediction in the United States," *2021 8th International Conference on Computational Science/Intelligence and Applied Informatics (CSII)*, Zhuhai, China, pp. 34-37, 2021.
- [11] Araujo, G. S., and Gaglianone, W. P., "Machine learning methods for inflation forecasting in Brazil: New contenders versus Classical Models," *Latin American Journal of Central Banking*, Volume 4, Issue 2, 2023.

- [12] World Economic Outlook Database, October 2022. (2022, October 11). IMF.
<https://www.imf.org/en/Publications/WEO/weo-database/2022/October>

แพลตฟอร์มการทำเกษตรแบบเอื้อประโยชน์ร่วมกันโดยใช้เทคโนโลยี อินเทอร์เน็ตของสรรพสิ่ง Co-Farming Platform using Internet of Things Technology

เลิศศักดิ์ หงษ์จันทร์ (Lertsak Hongjan)¹ และชัชวิน นามมัน (Chatchawin Namman)²

สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

¹lertsak.ho.64@ubu.ac.th, ²chatchawin.n@ubu.ac.th

บทคัดย่อ

แนวคิด *Sharing Economy* ที่มุ่งเน้นการใช้ทรัพยากรร่วมกันให้เกิดประโยชน์สูงสุด และความก้าวหน้าของเทคโนโลยี *Internet of Things* มีบทบาทอย่างมากในภาคการเกษตร ผู้วิจัยจึงพัฒนาแพลตฟอร์มการเกษตรแบบเอื้อประโยชน์ร่วมกันขึ้นมา โดยประกอบไปด้วย 2 ส่วนคือ แพลตฟอร์มที่เป็นตัวกลางให้บริการ และระบบ *Internet of Things* ที่สามารถติดตามสภาพแวดล้อมของพื้นที่การทำเกษตรรวมถึงควบคุมอุปกรณ์ต่าง ๆ แพลตฟอร์มพัฒนาโดยภาษา *PHP* ฐานข้อมูล *MySQL* พบว่าทำงานได้ถูกต้องครบตามความต้องการ ระบบ *IoT* สามารถทำงานร่วมกับชุดอุปกรณ์ที่ประกอบไปด้วยไมโครคอนโทรลเลอร์ โมดูลเซนเซอร์ และตัวควบคุมอุปกรณ์ได้เป็นอย่างดี โดยใช้โปรโตคอล *MQTT* สื่อสารไปยัง *Application Server* ที่สามารถปรับแต่งการแสดงผลและควบคุมได้ ผลการประเมินประสิทธิภาพจากผู้เชี่ยวชาญในระดับดีมาก ($\bar{x} = 4.72$) และผลการประเมินความพึงพอใจผู้ใช้งานในระดับดีมาก ($\bar{x} = 4.74$)

คำสำคัญ: การเกษตรแบบเอื้อประโยชน์ร่วมกัน
อินเทอร์เน็ตของสรรพสิ่ง แพลตฟอร์มเกษตร
อัจฉริยะ

Abstract

The idea of *Sharing Economy* focuses on maximizing the benefits of shared resources, and the significant advancement of *Internet of Things (IoT)* technology plays a vital role in agriculture. Therefore, researchers have developed a collaborative

agricultural (*Co-Farming*) platform consisting of two parts: a central service platform and an *IoT* system that can monitor the agricultural environment and controls various devices. The platform was developed using *PHP* language and *MySQL* database and found to work correctly and fully meet the requirements. The *IoT* system can work well with a set of devices including microcontrollers, sensor modules, and actuators, using *MQTT* protocol to communicate with the application server, which can be customized for dashboard. The performance evaluation results from experts were very good ($\bar{x} = 4.72$), and the users satisfaction evaluation results were very good ($\bar{x} = 4.74$).

Keywords: Co-Farming, IoT, Smart Farming Platform.

1. บทนำ

ในช่วงหลายปีที่ผ่านมาแนวคิดเรื่อง *Sharing Economy* [1] เริ่มเป็นที่รู้จักและถูกนำมาใช้อย่างแพร่หลาย แนวคิดนี้สนับสนุนการนำทรัพยากรที่ไม่ได้ถูกใช้งานอย่างเต็มประสิทธิภาพ (*Underutilized Assets*) มาแบ่งปันให้ผู้อื่นใช้งานเพื่อให้เกิดประสิทธิภาพและประโยชน์สูงสุดทั้งเจ้าของทรัพยากรและผู้ให้บริการ ด้านภาคการเกษตรเองก็มีการนำแนวคิดนี้มาปรับใช้เช่นกัน [2] อาทิ บริการแบ่งปันเครื่องจักรการเกษตร การให้บริการต่าง ๆ เกี่ยวกับการเกษตรตามความต้องการ (*Farm as a Service*)

เทคโนโลยีการเกษตร [3] ในปัจจุบันมีความก้าวหน้าเป็นอย่างมาก มีการประยุกต์ใช้งานเพื่อวัตถุประสงค์ต่าง ๆ อาทิ อำนวยความสะดวก ควบคุมคุณภาพ ลดต้นทุน

ลดความเสี่ยง ประกอบการตัดสินใจให้เกษตรกร โดยเฉพาะเทคโนโลยี Internet of Things ซึ่งในภาคการเกษตรกำลังได้รับความนิยมเป็นอย่างมาก

จากข้อมูลข้างต้นทำให้ผู้วิจัยมีความสนใจในการพัฒนาแพลตฟอร์มที่เป็นตัวกลางในการทำการเกษตรบนแนวคิด Sharing Economy โดยให้เจ้าของฟาร์มเปิดให้บริการต่าง ๆ แก่ผู้สนใจ และนำเทคโนโลยี IoT เข้ามาช่วยติดตามสภาพแวดล้อมและควบคุมอุปกรณ์ในพื้นที่ทำการเกษตรที่ให้บริการ เพื่อให้ผู้ใช้บริการเกิดความมั่นใจในกระบวนการผลิตและเกิดประสบการณ์ร่วม

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การเกษตรแบบแบ่งปัน

การเกษตรแบบแบ่งปัน (Shared Farm) [4] เป็นการทำการเกษตรบนพื้นฐานของแนวคิดเรื่อง Sharing Economy โดยหัวใจหลักคือจัดสรรทรัพยากรทางการเกษตรให้เกิดประโยชน์ร่วมกันสูงสุด เช่น โรงเรือน แปลงเกษตร เครื่องมือสิ่งอำนวยความสะดวก ผลผลิต ความรู้ ความสามารถ และสภาพแวดล้อมทางนิเวศวิทยา โดยผู้บริโภคไม่จำเป็นต้องเป็นเจ้าของทรัพยากรนั้น ๆ และเจ้าของทรัพยากรก็ได้ค่าตอบแทน

2.2 อินเทอร์เน็ตของสรรพสิ่ง

อินเทอร์เน็ตของสรรพสิ่ง [5] คือ เทคโนโลยีที่ทำให้อุปกรณ์อิเล็กทรอนิกส์ต่าง ๆ สามารถเชื่อมโยงหรือส่งข้อมูลถึงกันได้ผ่านเครือข่ายอินเทอร์เน็ต ทำให้เราสามารถติดตามสถานะหรือสั่งการควบคุมการทำงานของอุปกรณ์อิเล็กทรอนิกส์ต่าง ๆ ผ่านทางเครือข่ายอินเทอร์เน็ตได้

2.3 Message Queue Telemetry Transport

Message Queue Telemetry Transport (MQTT) [6] คือ โพรโตคอลที่พัฒนามาเพื่อใช้ในระบบ IoT ทำงานแบบ Broker and Clients Network สามารถส่งข้อมูลแบบเรียลไทม์ในปริมาณข้อมูลที่น้อยใช้พลังงานต่ำ ส่งข้อมูลในลักษณะ One-To-All จาก Publisher ไปยัง Broker และส่งต่อไปยังอุปกรณ์ทุกตัวที่ทำการ Subscribe ใน Topic นั้น

2.4 ตัวรับรู้และตัวกระตุ้น

ตัวรับรู้หรือเซนเซอร์ (Sensor) [7] คืออุปกรณ์ที่ใช้ใน

การตรวจจับสัญญาณหรือปริมาณทางฟิสิกส์จากสภาพแวดล้อมให้เป็นสัญญาณไฟฟ้า เช่น อุณหภูมิ แสง ความชื้น เป็นต้น ส่วนตัวกระตุ้น (Actuator) [8] คือ อุปกรณ์ที่เปลี่ยนสัญญาณไฟฟ้าเป็นพลังงานหรือการเคลื่อนที่ เช่น มอเตอร์ (Motor), รีเลย์ (Relay) เป็นต้น

2.5 ไมโครชิพ ESP8266 และ ESP32

ESP8266 [9] และ ESP32 [10] เป็นไวไฟไมโครชิพสำหรับใช้งานด้าน IoT ที่ได้รับความนิยมเป็นอย่างมาก เนื่องจากมีราคาถูก ขนาดเล็ก ประยุกต์ใช้งานกับโมดูลต่าง ๆ ได้หลากหลาย สามารถเชื่อมต่อกับ WiFi ในย่านความถี่ 2.4 GHz ทำงานกับโพรโตคอล TCP/IP ได้ สามารถประมวลผล อ่านค่าและควบคุมพอร์ตต่างๆผ่าน General Purpose Input/Output (GPIO) โดย ESP32 จะมีประสิทธิภาพมากกว่าและราคาแพงกว่า ESP8266

2.6 งานวิจัยที่เกี่ยวข้อง

สาวิตรีและนพพร [11] พัฒนาระบบเช่าพื้นที่เพาะปลูกให้ผู้สนใจทำเกษตรแต่ไม่มีพื้นที่โดยนำเทคโนโลยี IoT มาช่วยให้ผู้เช่าสามารถติดตามสภาพแวดล้อมและควบคุมอุปกรณ์ในพื้นที่เช่า พบว่าทำงานได้ถูกต้องแม่นยำ

Gao Guandong และคณะ [12] พัฒนาระบบเช่าพื้นที่ในชุมชนสำหรับทำแปลงปลูกผักให้ผู้สนใจที่อาศัยอยู่ในชุมชนเมืองได้มีประสบการณ์ในการเพาะปลูก นำเอา IoT มาช่วย โดยมี Sensor ต่าง ๆ รายงานสภาพแวดล้อมในแปลงเกษตรให้ผู้เช่ามั่นใจ และได้สัมผัสประสบการณ์

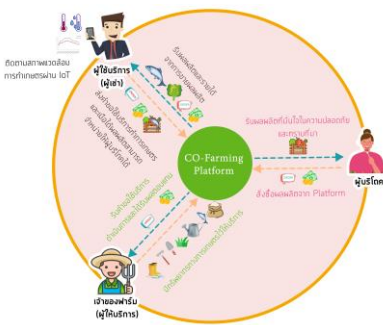
ปองพลและคณะ [13] พัฒนาโปรโตไทป์เพื่อแก้ปัญหาเนื้อทุเรียนมีรอยคล้ำและรสขมอันเนื่องมาจากสภาพดินไม่เหมาะสม โดยวัดความเป็นกรดต่างในดิน (pH) วัดความชื้นในดิน ไปยังฐานข้อมูลที่อยู่บน Server ผ่านโพรโตคอล MQTT และแสดงผลที่ Android Application พบว่าระบบสามารถรายงานผลและแจ้งเตือนได้

จากข้อมูลทฤษฎีและงานวิจัยข้างต้นเป็นข้อมูลสนับสนุนถึงความเป็นไปได้ และแนวทางในการพัฒนาแพลตฟอร์มขึ้นมาเป็นตัวกลางในการทำการเกษตร โดยใช้เทคโนโลยี Internet of Things เข้ามาเสริมการทำงาน

3. วิธีดำเนินการวิจัย

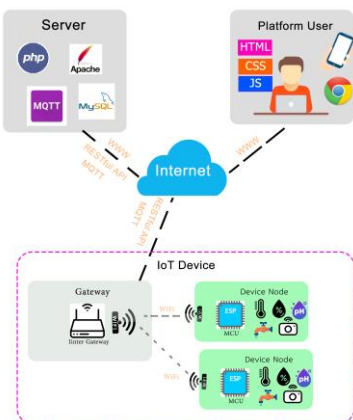
3.1 กรอบแนวคิดและสถาปัตยกรรมระบบ

แนวคิดของงานวิจัยนี้คือการพัฒนาระบบที่เป็นตัวกลางในการทำเกษตรแบบแบ่งปัน โดยมีผู้เกี่ยวข้องหลัก 4 ส่วน ได้แก่ 1) ผู้ให้บริการหรือเจ้าของฟาร์ม คือเจ้าของทรัพยากรสำหรับการทำการเกษตร 2) ผู้ใช้บริการหรือผู้เช่า คือผู้ที่ต้องการทำการเกษตรแต่ขาดทรัพยากร 3) ผู้บริโภค คือผู้ที่ต้องการซื้อผลผลิตทางการเกษตร 4) แพลตฟอร์ม คือตัวกลางการดำเนินการระหว่างส่วนต่างๆ โดยมีระบบ IoT มาช่วยเป็นเครื่องมือให้ผู้ใช้บริการทราบถึงข้อมูลสภาพแวดล้อมของพื้นที่การเกษตร เกิดประสบการณ์ร่วมในกระบวนการผลิต ดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดของแพลตฟอร์มฯ

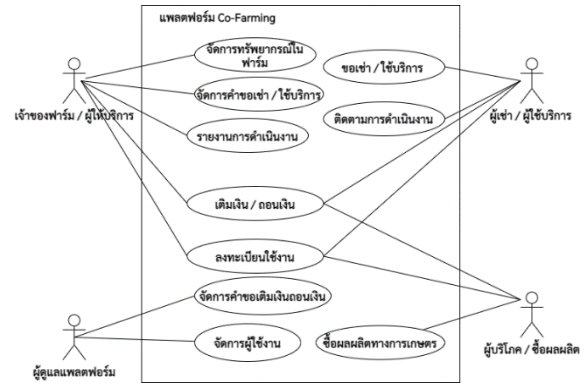
สถาปัตยกรรมของระบบมีอยู่ 3 ส่วน ได้แก่ 1) ส่วนของผู้ใช้งาน 2) ส่วนของเครื่องแม่ข่าย (Server) ให้บริการเซอร์วิสต่าง ๆ 3) ส่วนของอุปกรณ์ IoT ประกอบไป MCU ทำหน้าที่ประมวลผล, Sensor ทำหน้าที่รับค่าสภาพแวดล้อม ส่วน Actuator (Relay) ทำหน้าที่สะพานไฟเปิดปิดอุปกรณ์ตามคำสั่ง MCU ดังภาพที่ 2



ภาพที่ 2 : สถาปัตยกรรมระบบ

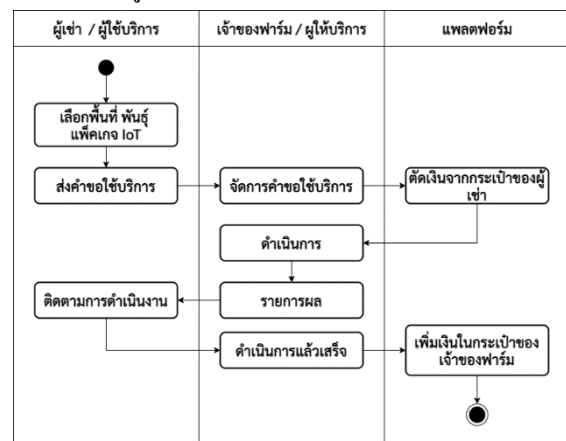
3.2 การออกแบบและพัฒนาแพลตฟอร์ม

3.2.1 การออกแบบฟังก์ชันการทำงานของแพลตฟอร์ม แบ่งผู้ใช้งานแพลตฟอร์มออกเป็น 4 กลุ่ม ซึ่งเขียนในรูปแบบ Use Case Diagram ได้ดังภาพที่ 3



ภาพที่ 3: Use Case Diagram ของแพลตฟอร์ม

3.2.2 การออกแบบขั้นตอนการให้บริการมีความเกี่ยวข้องกับผู้ใช้งานหลายส่วนตามภาพที่ 4



ภาพที่ 4: Activity Diagram ขั้นตอนการให้บริการ

3.2.3 การพัฒนาระบบ แพลตฟอร์มพัฒนาขึ้นด้วยภาษา PHP และใช้ฐานข้อมูล MySQL เครื่องมือที่ใช้พัฒนาคือ Visual Studio Code ส่วนติดต่อผู้ใช้งาน พัฒนาขึ้นโดยใช้ HTML, CSS, JavaScript ออกแบบให้แสดงผลตามขนาดหน้าจอ (Responsive Design) มีส่วนการแจ้งเตือนผู้ใช้งาน (Notification)

3.3 การออกแบบและพัฒนาชุดอุปกรณ์ IoT

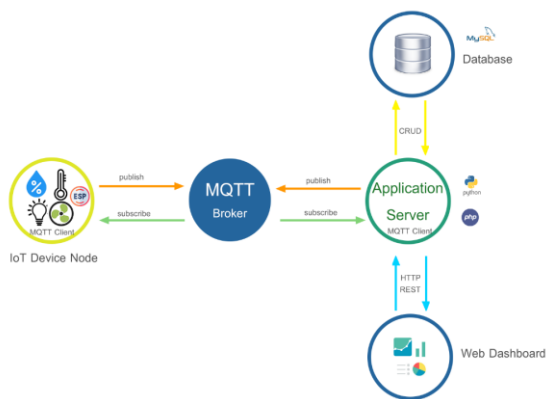
3.3.1 การออกแบบชุดอุปกรณ์ IoT อุปกรณ์ IoT มี 2 ส่วนคือ 1) Gateway ทำหน้าที่กระจายสัญญาณไวไฟให้กับอุปกรณ์ 2) ชุดอุปกรณ์ ที่ประกอบไป MCU ทำหน้าที่เป็นหน่วยประมวลผล, Sensor Module ทำหน้าที่รับ

ค่าสภาพแวดล้อม, Actuator (Relay) คอยเปิดปิดอุปกรณ์ ส่วนรายการโมดูลเซนเซอร์ที่ผู้วิจัยใช้ได้แสดงในตารางที่ 1

ตารางที่ 1: รายการโมดูลเซนเซอร์

โมดูล	การเชื่อมต่อ	การใช้งาน / การประยุกต์ใช้
1. HTU21D Temperature and humidity sensor	I2C	วัดอุณหภูมิ ความชื้น ใช้กับการเกษตรทั่วไป
2. MQ135 Air Quality Sensor Module	ADC	วัดคุณภาพของอากาศ ใช้กับการเลี้ยงแมลง
3. Soil Moisture Sensor Module	ADC	วัดความชื้นในดิน ใช้กับการปลูกผัก
4. ESP32Cam	WiFi	กล้องบันทึกภาพ ใช้กับการเกษตรทั่วไป

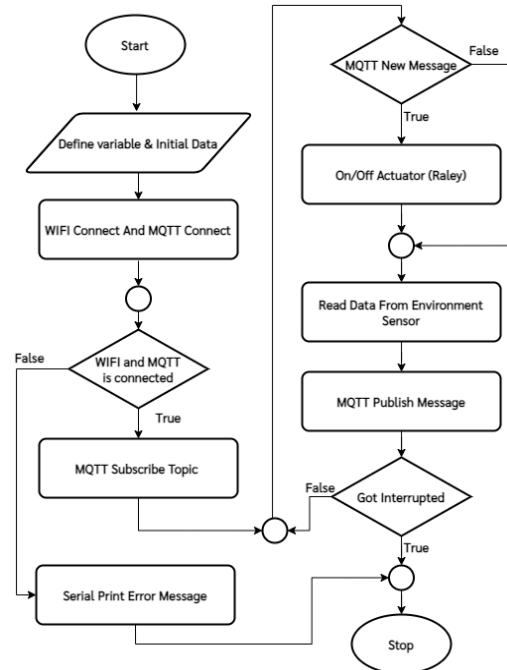
3.3.2 การออกการสื่อสารของอุปกรณ์ IoT ใช้ MQTT เป็นโพรโตคอลในการสื่อสาร โดยใช้ Mosquitto MQTT เป็น Broker คอยรับส่งข้อความไปยังอุปกรณ์ต่าง ๆ โดยใน Application Server มีสคริปต์ภาษา python ทำหน้าที่เป็น MQTT Client รอรับข้อความจากทุก ๆ Topic และเรียกใช้งาน API เพื่อส่งข้อมูลค่าสภาพแวดล้อม หรือการสั่งงานอุปกรณ์ไปเก็บยังฐานข้อมูล ส่วน Dashboard พัฒนาด้วยภาษา PHP ดึงข้อมูลจากฐานข้อมูลมาแสดง เมื่อมีการสั่งงานอุปกรณ์จากผู้ใช้สคริปต์ PHP ที่ทำหน้าที่เป็น MQTT Client จะ Publish คำสั่งเพื่อ Broker ส่งคำสั่งไปยัง IoT Device Node ดังภาพที่ 5



ภาพที่ 5: การสื่อสารของ IoT Device ด้วย MQTT

3.3.3 ขั้นตอนการทำงานของอุปกรณ์ IoT เริ่มจากการกำหนดค่าเริ่มต้นต่าง ๆ ทำการเชื่อมต่อ WiFi และ MQTT Broker ต่อด้วยการ Subscribe Topic โดยใช้ Device Key เป็นตัวอ้างอิงไปยัง Application Server เมื่อได้รับ

Message ใหม่ซึ่งเป็นคำสั่งควบคุมอุปกรณ์ MCU จะสั่งงาน On/Off Relay ตามคำสั่งที่ได้รับ ระหว่างรอรับ Message ใหม่ MCU จะอ่านค่าสภาพแวดล้อมจาก Sensor ต่าง ๆ และ Public Message ไปยัง Application Server เรื่อย ๆ ตาม Flowchart ที่แสดงตามภาพที่ 6



ภาพที่ 6: Flowchart การทำงานของอุปกรณ์ IoT

3.3.4 การพัฒนาชุดอุปกรณ์ IoT และระบบจัดการอุปกรณ์ ไอโอที ชุดอุปกรณ์ IoT พัฒนาด้วยภาษา C / C++ โดย Arduino IDE ใช้ MCU ตระกูล ESP8266 และ ESP32 ใช้โมดูลเซนเซอร์ตามลักษณะของเกษตรกรรม ใช้ไลบรารี PubSubClient สำหรับใช้งาน MQTT ไลบรารี ESP8266WiFi สำหรับเชื่อมต่อ WiFi ส่วนगेเวย์ใช้ WiFi Router ชนิด 3G/4G ด้านระบบจัดการอุปกรณ์ IoT และ Dashboard อยู่ในรูปแบบเว็บแอปพลิเคชัน โดยกำหนดรูปแบบการอ้างอิง IoT Device Node ด้วย Device Key อ้างอิงเซนเซอร์และตัวควบคุมอุปกรณ์ด้วย Param Name ดังภาพที่ 7

IoT Device Node	client.publish("COFARM/XC923TF3FD/temp", "23.45")
IoT Management & Dashboard	Device Key (ID ของ Farm Node) = XC923TF3FD Param Name (ชื่อของ sensor / actuator) = temp Value (ค่าของ sensor / actuator) = 23.45

ภาพที่ 7: การอ้างอิงข้อมูลจาก Device Node ไปยัง Application Server

3.4 การทดสอบระบบ

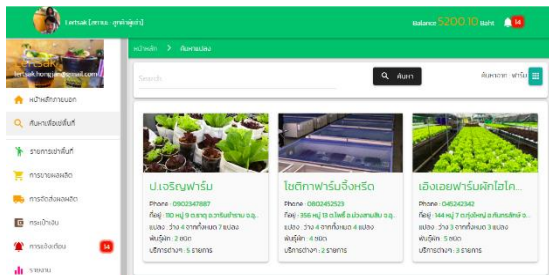
3.4.1 การทดสอบแพลตฟอร์ม ผู้วิจัยทดสอบฟังก์ชันการทำงาน (Functional Testing) ของส่วนต่างๆ และแก้ไขให้สามารถทำงานได้ถูกต้องสมบูรณ์ โดยทดสอบการทำงานในบทบาทของผู้ใช้งานแพลตฟอร์มทั้ง 4 กลุ่ม ทดสอบการแสดงผลบนอุปกรณ์ต่างๆ

3.4.2 การทดสอบระบบ IoT ผู้วิจัยทดสอบชุดอุปกรณ์และโมดูลเซนเซอร์ต่างๆ และการรับส่งข้อมูลผ่านโปรโตคอล MQTT เพื่อส่งข้อมูลไปยัง Dashboard ทดสอบการควบคุมอุปกรณ์ผ่าน Dashboard และการทำงานบนอุปกรณ์ต่างๆ ส่วนระบบจัดการอุปกรณ์ IoT ทดสอบการดำเนินการจัดการบอร์ดหรือ Device Node การจัดการเซนเซอร์และตัวควบคุมอุปกรณ์

4. ผลการดำเนินงาน

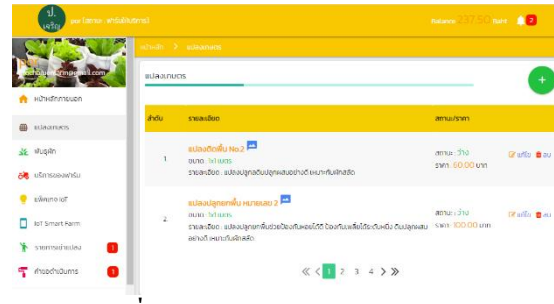
4.1 ผลการพัฒนาแพลตฟอร์ม

ภาพที่ 8 คือหน้าจอการใช้งานในกลุ่มผู้ให้บริการหรือผู้เช่าที่มีแถบเมนูต่างๆ และส่วนของข้อมูลฟาร์มต่างๆ ที่เปิดให้บริการอยู่บนแพลตฟอร์มเมื่อคลิกเลือกฟาร์มแล้วระบบจะแสดงรายการทรัพยากรต่างๆ ให้ผู้ให้บริการเลือก



ภาพที่ 8: หน้าจอการใช้งานในกลุ่มผู้ให้บริการ

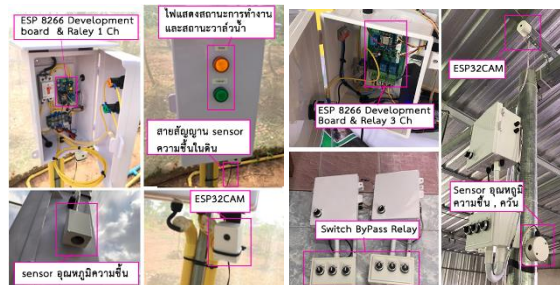
หน้าจอผู้ใช้งานในกลุ่มผู้ให้บริการหรือเจ้าของฟาร์ม นั้นคือภาพที่ 9 แสดงแถบเมนูต่างๆ เช่น เมนูจัดการพื้นที่จัดการพันธุ์ บริการของทางฟาร์ม และคำขอใช้บริการ



ภาพที่ 9: หน้าจอการใช้งานในกลุ่มผู้ให้บริการ

4.2 ผลการพัฒนาระบบ IoT

ผู้วิจัยได้พัฒนาชุดอุปกรณ์ IoT ขึ้นมา 2 ชุด เพื่อทดสอบใช้งานที่ต่างกันคือ 1) ชุดอุปกรณ์ IoT สำหรับฟาร์มผักดอง ภาพที่ 10 (ก) 2) ชุดอุปกรณ์ IoT สำหรับฟาร์มจิ้งหรีดดอง ภาพที่ 10 (ข)



(ก)

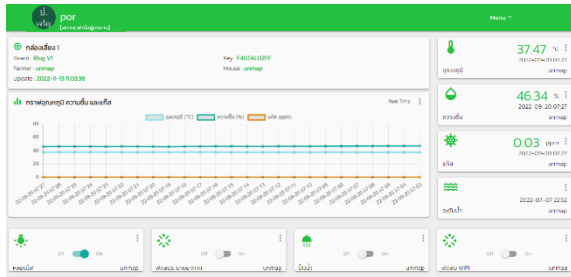
(ข)

ภาพที่ 10 : (ก) ชุดอุปกรณ์ IoT สำหรับฟาร์มผัก, (ข) ชุดอุปกรณ์ IoT สำหรับฟาร์มจิ้งหรีด

ระบบจัดการอุปกรณ์ IoT มีส่วนของรายการ IoT Device Node หรือ Board สามารถปรับแต่ง Sensor และ Actuator ตามความต้องการ และส่วนของ Dashboard สำหรับแสดงค่าของสภาพแวดล้อมและควบคุมอุปกรณ์ เมื่อทดสอบใช้งานชุดอุปกรณ์ IoT ทั้ง 2 ชุด แสดงผลบน Dashboard ได้ดังภาพที่ 11 และ 12 ซึ่งรายการ Sensor และ Actuator ต่างกันตามลักษณะของการเกษตร



ภาพที่ 11: หน้าจอ Dashboard สำหรับฟาร์มผัก



ภาพที่ 12: หน้าจอ Dashboard สำหรับฟาร์มอัจฉริค

4.3 ผลการประเมินประสิทธิภาพและความพึงพอใจ

ผลการประเมินประสิทธิภาพของระบบจากผู้เชี่ยวชาญ 5 คน พบว่ามีค่าเฉลี่ย 4.72 ส่วนเบี่ยงเบนมาตรฐาน 0.50 ซึ่งอยู่ในระดับดีมาก และผลการประเมินความพึงพอใจจากผู้ใช้งาน 20 คน อยู่ในระดับดีมากโดยมีค่าเฉลี่ย 4.74 ส่วนเบี่ยงเบนมาตรฐาน 0.55 ดังตารางที่ 2 และ 3

ตารางที่ 2: การประเมินประสิทธิภาพของระบบ

หัวข้อการประเมินประสิทธิภาพ	\bar{X}	S.D.
1. ด้านฟังก์ชันการทำงานต่าง ๆ ของระบบ	4.80	0.40
2. ด้านการออกแบบหน้าจอและส่วนติดต่อผู้ใช้	4.60	0.80
3. ด้านความปลอดภัยในการใช้งาน	4.80	0.40
4. ด้านฟังก์ชันการทำงานของระบบ IoT	4.60	0.49
5. ด้านออกแบบชุดอุปกรณ์ และรูปแบบการสื่อสาร	4.80	0.40
ค่าเฉลี่ย	4.72	0.50

ตารางที่ 3: การประเมินความพึงพอใจการใช้งาน

หัวข้อการประเมินความพึงพอใจการใช้งาน	\bar{X}	S.D.
1. ด้านความถูกต้องของฟังก์ชันการใช้งาน	4.75	0.54
2. ด้านความเข้าใจในขั้นตอนการใช้งาน	4.45	0.80
3. ด้านความสะดวกและความรวดเร็วในการใช้งาน	4.80	0.60
4. ด้านการออกแบบหน้าจอและความสวยงาม	4.85	0.36
5. ด้านการใช้งานระบบ IoT	4.85	0.48
ค่าเฉลี่ย	4.74	0.55

5. สรุป

การพัฒนาแพลตฟอร์มการทำเกษตรแบบเอื้อประโยชน์ร่วมกันตามแนวคิด Sharing Economy ซึ่งประกอบไปด้วย 2 ส่วน คือ 1) แพลตฟอร์มเป็นตัวกลางการดำเนินการ พบว่าทำงานได้ถูกต้องสมบูรณ์ 2) ระบบ IoT ประกอบด้วยชุดอุปกรณ์ IoT และระบบจัดการอุปกรณ์ที่เข้ามาช่วยเสริมให้ผู้ให้บริการสามารถติดตาม

สภาพแวดล้อมของพื้นที่การทำเกษตรรวมถึงควบคุมอุปกรณ์ต่าง ๆ พบว่าชุดอุปกรณ์สามารถทำงานร่วมกับระบบจัดการ IoT ได้เป็นอย่างดี และยังสามารถปรับแต่ง Dashboard ให้เหมาะสมกับประเภทของการเกษตร ผลการประเมินประสิทธิภาพจากผู้เชี่ยวชาญอยู่ในระดับดีมาก ($\bar{x} = 4.72$, S.D. = 0.50) และผลการประเมินความพึงพอใจผู้ใช้งานอยู่ในระดับดีมาก ($\bar{x} = 4.74$, S.D. = 0.55)

เอกสารอ้างอิง

- [1] ปราณี สุทศรี และคณะ, “Sharing Economy : นัยต่อเศรษฐกิจไทย” สืบค้นจาก https://www.bot.or.th/Thai/MonetaryPolicy/EconomicConditions/AAA/SharingEconomy_V8.pdf
- [2] สิริวัฒน์ สาครวาลี, “เศรษฐกิจแบ่งปันกับภาคการเกษตร” สืบค้นจาก <https://mjusmartfarm.wordpress.com/2018/05/22/เศรษฐกิจแบ่งปันกับภาค/>
- [3] “เทคโนโลยีการเกษตร เพื่อปรับปรุงประสิทธิภาพในภาคเกษตรกรรม” สืบค้นจาก https://www.arida.or.th/knowledge_detail.php?id=21
- [4] Jun ZHOU and Xin LIAN, “Shared Farm: an approach to achieve urban-rural reciprocity based” *54th ISOCARP Congress*, 2018.
- [5] Matana Wiboonyasake, “ทำความเข้าใจกับ Internet of Things”. สืบค้นจาก <https://www.aware.co.th/iot-คืออะไร>
- [6] “What is MQTT ? โพรโทคอลเพื่อการสื่อสารของ IOT โดยเฉพาะ”. สืบค้นจาก <https://iiot.riverplus.com/mqtt/>
- [7] “ตัวรับรู้ (Sensor)”. สืบค้นจาก <https://th.wikipedia.org/wiki/ตัวรับรู้>
- [8] “What are actuators in IoT?”. Available Online at <https://www.nabto.com/actuators-in-iot-guide/>
- [9] “What is the ESP8266” Available Online at <https://www.nabto.com/esp8266-for-iot-complete-guide/>
- [10] “What is ESP32 Available Online at <https://www.nabto.com/guide-to-iot-esp-32/>
- [11] สาวิตรี ฝางงาม และนวพร วิสิษฐพงศ์พันธ์. “ระบบเช่าพื้นที่เพาะปลูกอัจฉริยะด้วยเทคโนโลยีอินเทอร์เน็ตของสรรพสิ่ง”, *การประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 14. (NCCIT 2018)*, ณ โรงแรมแชงกรีลา เชียงใหม่, 5-6 กรกฎาคม, 2561.
- [12] Guandong Gao, et al., “An IOT-based Multi-sensor Ecological Shared Farmland Management System” *International Journal of Online and Biomedical Engineering (iJOE)*, Vol. 14, No. 3, pp. 81-95, 2018.

- [13] ปองพล นิลพฤกษ์ และคณะ, “พัฒนาโปรโตไทป์ระบบตรวจวัดและแจ้งเตือนสภาพดินในสวนทุเรียน กรณีศึกษา ต. บึงกาสาม อ.หนองเสือ จ.ปทุมธานี” วารสารวิจัย มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, ปีที่ 17 ฉบับที่ 1พ.ศ. 2561 หน้า 33-44.

การเปรียบเทียบการจำแนกภาพสิ่งปลูกสร้าง ด้วยเทคนิคการเรียนรู้เชิงลึก

Comparison of Image Classification of Buildings with Deep Learning Techniques

ภริเดช โทลานวัตร (Piradate Tolornuwat)¹ และวัชรวิวัฒน์ จิตต์สกุล (Watchareewan Jitsakul)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857280@ email.kmutnb.ac.th, ²watchareewan.j@itd.kmutnb.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบการจำแนกภาพสิ่งปลูกสร้าง ด้วยเทคนิคการเรียนรู้เชิงลึก โดยเปรียบเทียบแบบจำลองจำนวน 3 โมเดล ได้แก่ ResNet50 ResNet101 และ ResNet152 โดยมีการเปรียบเทียบในแต่ละโมเดลในการเพิ่มความถูกต้องให้กับกระบวนการฝึกสอนโดยการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกด้านภาพ วัดประสิทธิภาพแบบจำลองด้วยค่าความถูกต้อง ข้อมูลที่ใช้เป็นข้อมูลภาพสิ่งปลูกสร้าง 4 ประเภท จาก กรมธนารักษ์ จำนวนประเภทละ 200 ภาพรวม 800 ภาพ แต่ละภาพจะมีขนาด 156x156 พิกเซล และแบ่งข้อมูลเป็น 3 ชุด คือ ชุดข้อมูลฝึกสอนที่ 1 จำนวนประเภทละ 140 ภาพ รวม 560 ภาพ ชุดข้อมูลตรวจสอบจำนวนประเภทละ 18 ภาพ รวม 72 ภาพ ชุดข้อมูลทดสอบจำนวนประเภทละ 42 ภาพ รวม 168 ภาพ และชุดข้อมูลฝึกสอนที่ 2 เป็นการนำชุดข้อมูลฝึกสอนที่ 1 มาทำการเพิ่มภาพด้วยวิธีการหมุนภาพ และวิธีพลิกกลับด้านภาพ รวม 8 ทิศทาง ทำให้ได้ภาพ 4,480 ภาพ ผลการวิจัยพบว่า การเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ จะช่วยให้ประสิทธิภาพในการจำแนกภาพมีความถูกต้องสูงขึ้น โดยแบบจำลอง ResNet50 มีค่าความถูกต้องเพิ่มขึ้นสูงที่สุดจาก ร้อยละ 83.93 เป็นร้อยละ 88.69 แบบจำลอง ResNet101 มีค่าความถูกต้องสูงขึ้นจากร้อยละ 83.33 เป็นร้อยละ 86.31 และแบบจำลอง ResNet152 มีค่าความถูกต้องสูงขึ้นจากร้อยละ 86.31 เป็นร้อยละ 86.90

คำสำคัญ: การจำแนกภาพสิ่งปลูกสร้าง เทคนิคการเรียนรู้เชิงลึก การเพิ่มภาพ

Abstract

This research aimed to compare building image classification with deep learning techniques by comparing 3 models, namely ResNet50, ResNet101 and ResNet152, with each model being compared and increasing the accuracy of the training process by adding images by rotation and reverse measurement methods. Measure model performance with accuracy. The data used are four types of building images from the Treasury Department, amounting to 200 images per type, totaling 800 images. Each image is 156x156 pixels and divides the data into 3 sets, namely, Train Data Set 1 140 images per type, totaling 560 images. Validate Set 18 images per type, totaling 72 images, Test Set 42 images per type, totaling 168 images, Train Data Set 2 is the use of the train data set 1 to obtain the desired image (Rotation) and the method of inverting the image (Reverse) for a total of 8 directions to get 4,480 images. The results showed that adding images using rotation and the reverse method improved image recognition efficiency. The ResNet50 model has the greatest improvement in precision from 83.93 up to 88.69, the ResNet101 model performs better from 83.33 up to 86.31, and the ResNet152 model performs better from 86.31 up to 86.90

Keywords: Building Image Classification, Deep Learning Techniques, Image Augmentation.

1. บทนำ

กรมธนารักษ์ มีหน้าที่ดูแลที่ราชพัสดุ การบริหารเงินตราและเหรียญกษาปณ์ โรงกษาปณ์ งานด้านทรัพย์สินมีค่าของแผ่นดิน และการประเมินราคาทรัพย์สิน ซึ่งเป็นการกำหนดราคาประเมินของอสังหาริมทรัพย์และทรัพย์สินอื่น ๆ ตามกฎหมายที่เกี่ยวข้อง โดยกองประเมินราคาทรัพย์สินจะดำเนินการประเมินราคาที่ดินและสิ่งปลูกสร้างในรูปแบบบัญชีประเมินราคา แบ่งเป็น 2 ประเภท คือ บัญชีราคาประเมินราคาที่ดิน และบัญชีราคาประเมินสิ่งปลูกสร้าง โดยราคาประเมินที่ดินและสิ่งปลูกสร้างจะนำมาใช้เป็นฐานในการคำนวณค่าธรรมเนียมในการซื้อขาย ใช้อ้างอิงเพื่อพิจารณาเกี่ยวกับการให้สินเชื่อกับสถาบันการเงิน หรือการนำมาเป็นเอกสารพยานในการพิจารณาอรรถคดีของศาล การประเมินราคาที่ดินและสิ่งปลูกสร้างจะดำเนินการทบทวนใหม่ทุก ๆ 4 ปี ทั้งนี้ เนื่องจากสภาพแวดล้อมสิ่งปลูกสร้างต่าง ๆ ได้มีการเปลี่ยนแปลงตลอดเวลา และที่ผ่านมาการประเมินราคาสิ่งปลูกสร้างจะใช้คนในการสำรวจสิ่งปลูกสร้างแต่ละประเภท และนำมาจัดหมวดหมู่แยกประเภทสิ่งปลูกสร้าง เพื่อกำหนดราคาประเมินสิ่งปลูกสร้างตามกฎหมายกำหนด ทั้งนี้การดำเนินการด้วยคนมักล่าช้า เนื่องจากข้อมูลสิ่งปลูกสร้างมีเป็นจำนวนมาก ซึ่งไม่สอดคล้องกับกำลังเจ้าหน้าที่ประเมิน จากปัญหาดังกล่าวทำให้ผู้วิจัย มีแนวคิดที่จะหาวิธีช่วยวิเคราะห์ลักษณะภาพของสิ่งปลูกสร้างเพื่อให้สามารถจำแนกประเภทสิ่งปลูกสร้างได้ ซึ่งสามารถทำการประเมินราคาสิ่งปลูกสร้างได้อย่างมีประสิทธิภาพ

เพื่อให้การประเมินราคาสิ่งปลูกสร้างมีประสิทธิภาพสูงและลดปัญหาการผิดพลาด การวิเคราะห์ภาพสิ่งปลูกสร้างโดยการประมวลผลภาพด้วยเทคนิคการเรียนรู้เชิงลึก เพื่อแยกประเภทจึงมีความสำคัญ จากการศึกษาวิจัยที่ผ่านมา พบว่า มีผู้ให้ความสนใจในเทคนิคการประมวลผลภาพเป็นจำนวนมาก โดยมีเทคนิคการเรียนรู้เชิงลึกที่นิยมใช้ และมีประสิทธิภาพ คือ การใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) เช่น การพัฒนาระบบสืบค้นที่มาของลาย

ผ้าย้อมครามด้วยเทคนิคการเรียนรู้เชิงลึก เพื่อจำแนกลายผ้าย้อมครามของกลุ่มแปรรูปผ้าย้อมครามธรรมชาติในจังหวัดสกลนคร โดยประยุกต์ใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ ด้วยแบบจำลอง ResNet50 [1] การจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยเปรียบเทียบประสิทธิภาพการจำแนกพรรณไม้ ได้แก่ แบบจำลอง LeNet-5 แบบจำลอง AlexNet แบบจำลอง GoogLeNet และแบบจำลอง VGGNet [2] วิเคราะห์โรคใบมะนาวด้วยการเรียนรู้เชิงลึก โดยเปรียบเทียบการจำแนกจากภาพใบมะนาว ได้แก่ แบบจำลอง LeNet-5 แบบจำลอง VGG16 แบบจำลอง ResNet50 และแบบจำลอง VGGNet [3] เป็นต้น

จากการวิจัยที่กล่าวมาข้างต้น ผู้วิจัยมีแนวคิดที่จะศึกษาทฤษฎีการประมวลผลภาพ ด้วยเทคนิคการเรียนรู้เชิงลึก เพื่อจำแนกภาพสิ่งปลูกสร้างตามที่กองประเมินราคาทรัพย์สิน กรมธนารักษ์ กำหนด เพื่อให้ได้วิธีการดำเนินการที่รวดเร็วและมีประสิทธิภาพมากที่สุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้รวบรวมข้อมูลทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยแบ่งเป็น 4 หัวข้อหลัก ดังนี้

2.1 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ เป็นการนำภาพมาประมวลผลหรือคิดคำนวณด้วยคอมพิวเตอร์ เพื่อให้ได้ข้อมูลที่ต้องการทั้งในเชิงคุณภาพและปริมาณ การรู้จำภาพ (Image Recognition) เป็นวิธีหนึ่งของการประมวลผลภาพ ที่ถูกนำมาใช้ในการรู้จำของภาพประเภทต่าง ๆ เช่น ภาพสิ่งมีชีวิต คน สัตว์ ภาพสิ่งของ ภาพสิ่งปลูกสร้าง เป็นต้น วัตถุที่มีในภาพแต่ละส่วนแต่ละชิ้นจะผ่านกระบวนการหลายขั้นตอนก่อนจะถูกวิเคราะห์และสรุปผลออกมาเป็นภาพประเภทใด [5]

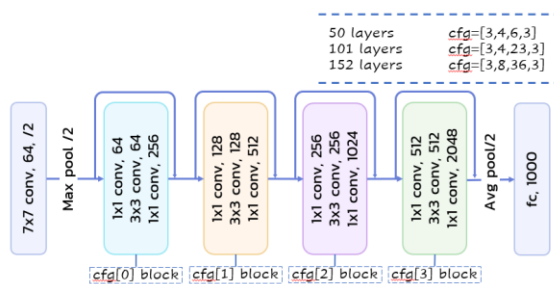
2.2 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก เป็นการเรียนรู้ของเครื่องจักร (Machine Learning) บนพื้นฐานของโครงข่ายประสาทเทียมและการเรียนรู้คุณลักษณะ การเรียนรู้สามารถเป็นได้

ทั้งแบบการเรียนรู้แบบมีผู้สอน การเรียนรู้แบบกึ่งมีผู้สอน คำว่า "ลึก" ในความหมายมาจากการที่มีชั้นของโครงข่ายหลายชั้น ที่มีประสิทธิภาพมากขึ้น การเรียนที่สะดวกขึ้น และการเข้าใจในโครงสร้างที่ชัดเจนขึ้น การเรียนรู้เชิงลึกได้รับความนิยมในการรู้จำภาพ [4]

2.3 ResNet (Deep Residual Network)

เป็นประเภทการเรียนรู้เชิงลึกที่ได้รับความนิยมสูงมากในการเรียนรู้เชิงลึก (Deep Learning) ที่อยู่บนพื้นฐานของโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) โดย ResNet เป็นสถาปัตยกรรมที่ได้รับความนิยมในปี 2015 จากการแข่งขันโครงการ ImageNet ซึ่งเป็นโครงการจัดตั้งขึ้นเพื่อแข่งขันการรู้จำภาพจากฐานข้อมูลขนาดใหญ่ โดยผู้ชนะได้แก่ Microsoft ได้นำเสนอสถาปัตยกรรมใหม่โดย "skip connection" และทำ Batch normalization เพิ่มขึ้น มีจำนวนชั้นทั้งหมด 152 ชั้นกับการเชื่อมแบบ Residual connection ซึ่งมีความซับซ้อนกว่า VGGNet ทำงานบน 8 GPU เป็นเวลา 2-3 สัปดาห์ [7] กระบวนการทำงานของ ResNet แสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการทำงานของ ResNet

2.4 งานวิจัยที่เกี่ยวข้อง

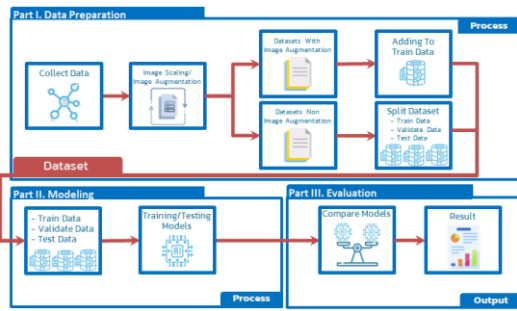
ปิยะเนตร และเอกรัฐ [8] ได้ศึกษาการจำแนกอาคารจากภาพถ่ายดาวเทียมโดยการถ่ายโอนความรู้ โดยเปรียบเทียบการจำแนกภาพอาคารประเภทต่าง ๆ ด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ ได้แก่ แบบจำลอง VGG16 แบบจำลอง InceptionV3 แบบจำลอง ResNet152V2 แบบจำลอง DenseNet201 และแบบจำลอง Xception ข้อมูลที่ใช้เป็นภาพถ่ายดาวเทียมความละเอียดสูงจากซอฟต์แวร์กูเกิล เอิร์ธโปร

เป็นข้อมูลภาพถ่ายหมวดสนามบิน หมวดสนามกีฬา หมวดโรงเรียน และหมวดวัด แบ่งข้อมูลเป็น 2 ชุด ชุดที่ 1 ชุดสำหรับฝึกสอน จำนวน 1,073 ภาพ ชุดที่ 2 ชุดสำหรับตรวจสอบความถูกต้อง จำนวน 265 ภาพ ทดสอบประสิทธิภาพด้วยค่าความถูกต้อง ผลการวิจัยพบว่าแบบจำลอง DenseNet201 มีประสิทธิภาพสูงสุด โดยมีค่าความถูกต้อง เท่ากับร้อยละ 94.4 รองลงมาแบบจำลอง VGG16 ค่าความถูกต้อง เท่ากับร้อยละ 94.05 รองลงมาแบบจำลอง InceptionV3 ค่าความถูกต้อง เท่ากับร้อยละ 94.05 รองลงมาแบบจำลอง ResNet152V2 ค่าความถูกต้อง เท่ากับร้อยละ 93.4 และแบบจำลอง Xception ค่าความถูกต้อง เท่ากับร้อยละ 90.67 ตามลำดับ

ณัฐวุฒิ [9] ได้ทำการศึกษาการปรับปรุงประสิทธิภาพการจำแนกภาพเอกซเรย์ทรวงอกด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ โดยใช้เทคนิคการเพิ่มภาพสำหรับวินิจฉัยโรคโควิด-19 โดยเปรียบเทียบเทคนิคการเพิ่มภาพในแต่ละวิธี ได้แก่ การย่อหรือขยาย การหมุนภาพ และการพลิกกลับด้าน ข้อมูลที่ใช้เป็นข้อมูลภาพเอกซเรย์สาธาณะแบบ Posteroanterior (PA) ซึ่งได้จัดกลุ่มภาพเอกซเรย์ไว้อย่างชัดเจน 2 กลุ่ม กลุ่มละ 25 ภาพ ประกอบด้วยกลุ่ม Positive คือภาพเอกซเรย์ทรวงอกกลุ่มผู้ติดเชื้อโควิด-19 และกลุ่ม Negative คือภาพเอกซเรย์ทรวงอกกลุ่มผู้ไม่ติดเชื้อโควิด-19 ทดสอบประสิทธิภาพของการจำแนกภาพด้วยค่าความถูกต้อง ผลการวิจัยพบว่า การหมุนภาพ ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้อง เท่ากับร้อยละ 99.67 รองลงมาการพลิกกลับด้าน ค่าความถูกต้อง เท่ากับร้อยละ 99.26 และการย่อหรือขยาย ค่าความถูกต้อง เท่ากับร้อยละ 98.48 ตามลำดับ

3. วิธีการดำเนินการวิจัย

การเปรียบเทียบการจำแนกภาพสิ่งปลูกสร้าง ด้วยเทคนิคการเรียนรู้เชิงลึก ผู้วิจัยมีกรอบแนวคิดแสดงดังภาพที่ 2



ภาพที่ 2: กรอบแนวคิดของงานวิจัย

จากภาพที่ 2 กรอบแนวคิดของงานวิจัยประกอบด้วย ขั้นตอนการดำเนินงานวิจัย 3 ขั้นตอน ได้แก่ การรวบรวม และการเตรียมข้อมูล (Data Preparation) การสร้างแบบจำลอง (Modeling) โดยใช้ Pretrained Models ได้แก่ ResNet50, ResNet101 และ ResNet152 และการประเมินผล (Evaluation) ดังนี้

3.1 การรวบรวมและการเตรียมข้อมูล

เก็บรวบรวมข้อมูลภาพสิ่งปลูกสร้างทั้ง 4 ประเภท จากกองประเมินราคาทรัพย์สิน กรมธนารักษ์ ได้แก่ บ้านเดี่ยว=0, ทาวน์เฮาส์=1, ห้องแถว=2 และตึกแถว=3 จำนวนประเภทละ 200 ภาพ รวม 800 ภาพ จากนั้นนำข้อมูลที่รวบรวมได้ มาจัดรูปแบบและปรับขนาดของภาพให้เป็นขนาด 156x156 พิกเซล ต่อมานำข้อมูลที่ได้แบ่งข้อมูลออกเป็น 3 ชุด ได้แก่ ชุดข้อมูลฝึกสอนที่ 1 จำนวนประเภทละ 140 ภาพ รวม 560 ภาพ ชุดข้อมูลฝึกสอนที่ 2 เป็นการนำชุดข้อมูลฝึกสอนที่ 1 มาทำการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ รวม 8 ทิศทาง ทำให้ได้ภาพ จำนวน 4,480 ภาพ เพื่อใช้ในการสร้างแบบจำลองและเปรียบเทียบประสิทธิภาพในการจำแนกของแบบจำลอง ชุดข้อมูลตรวจสอบ จำนวนประเภทละ 18 ภาพ รวม 72 ภาพ เพื่อใช้ในการตรวจสอบความถูกต้องในการฝึกสอนแบบจำลอง และชุดข้อมูลทดสอบ จำนวนประเภทละ 42 ภาพ รวม 168 ภาพ เพื่อใช้เป็นข้อมูลสำหรับทวนสอบความถูกต้องของการจำแนกภาพสิ่งปลูกสร้าง

3.2 การสร้างแบบจำลอง

สร้างแบบจำลองการจำแนกด้วย Pretrained model ได้แก่ ResNet50, ResNet101 และ ResNet152 แบ่งการดำเนินงานเป็น 3 ส่วน ดังนี้

3.2.1 กำหนดพารามิเตอร์ที่ใช้สำหรับการทดสอบแบบจำลองทั้ง 3 โมเดล ได้แก่ learning_rate = 0.0001, optimizer_fn = adam, batch_size = 16, epochs = 40 และ loss_fn = categorical_crossentropy

3.2.2 การฝึกสอนแบบจำลอง นำข้อมูลชุดฝึกสอนที่ 1 และ 2 ทั้ง 2 ชุด และชุดข้อมูลตรวจสอบความถูกต้องมาทำการการฝึกสอนแบบจำลองและตรวจสอบความถูกต้องของแบบจำลองทั้ง 3 โมเดล

3.2.3 การทวนสอบความถูกต้องของแบบจำลอง นำแบบจำลองที่ได้จากการฝึกสอนแบบจำลอง มาทวนสอบความถูกต้องด้วยชุดข้อมูลทดสอบ พิจารณาประสิทธิภาพด้วยค่าความถูกต้อง

3.3 การประเมินผล

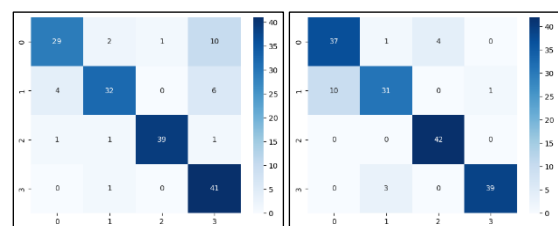
ทำการวัดประสิทธิภาพแบบจำลองการจำแนกภาพสิ่งปลูกสร้างด้วยค่าความถูกต้อง

4. ผลการดำเนินงาน

ผลการดำเนินงานวิจัยแบ่งออกเป็น 2 หัวข้อ แสดงรายละเอียดได้ดังนี้

4.1 ผลการดำเนินงานสร้างแบบจำลอง

4.1.1 ผลการสร้างแบบจำลอง ResNet50 พบว่าการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 ให้ค่าความถูกต้อง 83.93 และการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ให้ค่าความถูกต้อง 88.69 เมื่อพิจารณาค่าความถูกต้องพบว่าการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ที่ได้ทำการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ให้ประสิทธิภาพ ดีกว่า การฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 โดยค่าความถูกต้องเท่ากับร้อยละ 88.69 ซึ่งมีรายละเอียดการจำแนกภาพสิ่งปลูกสร้างดังภาพที่ 3 และตารางที่ 1

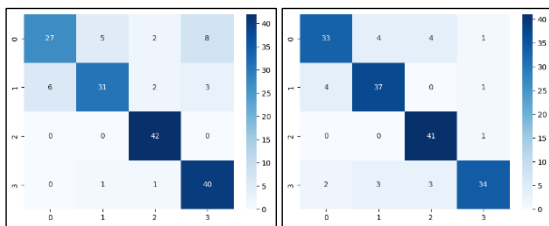


ภาพที่ 3: รายละเอียดการจำแนกภาพด้วยข้อมูลฝึกสอนชุดที่ 1 เปรียบเทียบกับข้อมูลฝึกสอนชุดที่ 2 ของ ResNet50

ตารางที่ 1: ตารางเปรียบเทียบของแบบจำลอง ResNet50

Class	ฝึกสอนชุดที่ 1		ฝึกสอนชุดที่ 2	
	Recall	F1-Score	Recall	F1-Score
0	0.69	0.76	0.88	0.83
1	0.76	0.82	0.73	0.80
2	0.92	0.95	1.00	0.95
3	0.97	0.82	0.92	0.95

4.1.2 ผลการสร้างแบบจำลอง ResNet101 พบว่าการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 ให้ค่าความถูกต้อง 83.33 และการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ให้ค่าความถูกต้อง 86.31 เมื่อพิจารณาค่าความถูกต้องพบว่าการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ที่ได้ทำการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ให้ประสิทธิภาพดีกว่า การฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 โดยค่าความถูกต้องเท่ากับร้อยละ 86.31 ซึ่งมีรายละเอียดการจำแนกภาพสิ่งปลูกสร้างดังภาพที่ 4 และตารางที่ 2



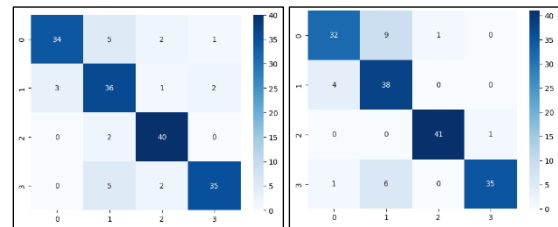
ภาพที่ 4: รายละเอียดการจำแนกภาพด้วยข้อมูลฝึกสอนชุดที่ 1 เปรียบเทียบกับข้อมูลฝึกสอนชุดที่ 2 ของ ResNet101

ตารางที่ 2: ตารางเปรียบเทียบของแบบจำลอง ResNet101

Class	ฝึกสอนชุดที่ 1		ฝึกสอนชุดที่ 2	
	Recall	F1-Score	Recall	F1-Score
0	0.64	0.72	0.78	0.81
1	0.73	0.78	0.88	0.86
2	1.00	0.94	0.97	0.91
3	0.95	0.86	0.80	0.86

4.1.3 ผลการสร้างแบบจำลอง ResNet152 พบว่าการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 ให้ค่าความถูกต้อง 83.31 และการฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ให้ค่าความถูกต้อง 86.90 เมื่อพิจารณาค่าความถูกต้องพบว่าการ

ฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 2 ที่ได้ทำการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ให้ประสิทธิภาพดีกว่า การฝึกสอนด้วยชุดข้อมูลฝึกสอนที่ 1 โดยค่าความถูกต้องเท่ากับร้อยละ 86.90 ซึ่งมีรายละเอียดการจำแนกภาพสิ่งปลูกสร้างดังภาพที่ 5 และตารางที่ 3



ภาพที่ 5: รายละเอียดการจำแนกภาพด้วยข้อมูลฝึกสอนชุดที่ 1 เปรียบเทียบกับข้อมูลฝึกสอนชุดที่ 2 ของ ResNet152

ตารางที่ 3: ตารางเปรียบเทียบของแบบจำลอง ResNet152

Class	ฝึกสอนชุดที่ 1		ฝึกสอนชุดที่ 2	
	Recall	F1-Score	Recall	F1-Score
0	0.80	0.86	0.76	0.81
1	0.85	0.80	0.90	0.80
2	0.95	0.91	0.97	0.97
3	0.83	0.87	0.83	0.89

4.2 ผลการประเมินผลและเปรียบเทียบแบบจำลอง

การประเมินผลและเปรียบเทียบแบบจำลอง โดยนำค่าความถูกต้องของตัวแบบจำลองของทั้ง 3 โมเดล ได้แก่ ResNet50, ResNet101 และ ResNet152 มาเปรียบเทียบ พบว่าการฝึกสอนด้วยข้อมูลฝึกสอนชุดที่ 2 ที่มีการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ช่วยให้ประสิทธิภาพในการจำแนกภาพมีความถูกต้องสูงขึ้น รายละเอียดแสดงดังตารางที่ 4

ตารางที่ 4: ตารางเปรียบเทียบค่าความถูกต้องของแบบจำลอง

แบบจำลอง	ค่าความถูกต้อง ที่ฝึกสอนด้วยชุดข้อมูลฝึกสอน	
	ชุดที่ 1	ชุดที่ 2
ResNet50	83.93	88.69
ResNet101	83.33	86.31
ResNet152	86.31	86.90

จากตารางที่ 4 เมื่อพิจารณาค่าความถูกต้อง พบว่าแบบจำลอง ResNet50 ที่ฝึกสอนด้วยชุดข้อมูลที่ 2 ที่มีการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ให้ประสิทธิภาพเพิ่มขึ้นสูงที่สุด โดยมีความถูกต้องเพิ่มขึ้นจากร้อยละ 83.93 เป็นร้อยละ 88.69

5. สรุป

งานวิจัยฉบับนี้ผู้วิจัยได้ศึกษาการเปรียบเทียบการจำแนกภาพสิ่งปลูกสร้าง ด้วยเทคนิคการเรียนรู้เชิงลึก และเทคนิคการเพิ่มความถูกต้องให้กับกระบวนการฝึกสอน โดยการเพิ่มภาพหลาย ๆ มุมมองด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ เพื่อเพิ่มประสิทธิภาพการจำแนกภาพสิ่งปลูกสร้าง ผู้วิจัยได้คัดเลือกแบบจำลอง Pretrained models ที่ได้รับความนิยม จำนวน 3 โมเดล ได้แก่ ResNet50, ResNet101 และ ResNet152 ทดสอบประสิทธิภาพของแบบจำลองการจำแนกด้วยค่าความถูกต้อง ผลการวิจัยพบว่าการฝึกสอนด้วยข้อมูลฝึกสอนชุดที่ 2 ที่มีการเพิ่มภาพด้วยวิธีการหมุนภาพและวิธีพลิกกลับด้านภาพ ช่วยให้ประสิทธิภาพในการจำแนกภาพมีความถูกต้องสูงขึ้น ยังมีจำนวนรูปภาพมาก และหลายหลายมุมความแม่นยำในการจำแนกรูปภาพก็สูงขึ้นด้วย โดยแบบจำลอง ResNet50 มีค่าความถูกต้องเพิ่มขึ้นสูงที่สุดจากร้อยละ 83.93 เป็นร้อยละ 88.69 รองลงมาแบบจำลอง ResNet101 มีค่าความถูกต้องสูงขึ้นจากร้อยละ 83.33 เป็นร้อยละ 86.31 และแบบจำลอง ResNet152 มีค่าความถูกต้องสูงขึ้นจากร้อยละ 86.31 เป็นร้อยละ 86.90 ทำให้เห็นได้ว่าการใช้ แบบจำลองที่มีชั้นความลึก (Layer) มาก ๆ ไม่ได้ให้ประสิทธิภาพสูงสุด เมื่อใช้ร่วมกับเทคนิคเพิ่มภาพ เนื่องจากข้อมูลที่นำมาจำแนกมีเพียง 4 ประเภท แบบจำลองที่มีชั้น (Layer) น้อยนั้นเหมาะสมกว่า การใช้แบบจำลองที่มีชั้นความลึก (Layer) มาก ๆ อีกทั้งการใช้เทคนิคเพิ่มภาพกับแบบจำลองที่มีชั้นความลึก (Layer) ไม่มากนักจะช่วยเพิ่มประสิทธิภาพได้ดี ข้อควรคำนึงถึงในการใช้เทคนิคเพิ่มภาพให้มีจำนวนมากขึ้น ส่งผลต่อเวลาที่ใช้ในการฝึกสอนก็จะนานมากขึ้นตามไปด้วยเช่นกัน

เอกสารอ้างอิง

- [1] ชารินี ไชยชนะ, สีตลา วงษ์ภาพสินธุ์ และฉัตรชัย เจียมรัมย์. (2022). “การพัฒนาระบบสืบค้นที่มาของลายผ้าข้อมครามด้วยเทคนิคการเรียนรู้เชิงลึก” *วารสารวิจัย UTK ราชวมงคลกรุงเทพ*. ปีที่ 16 ฉบับที่ 1 (มกราคม – มิถุนายน) : 68-83.
- [2] จักรินทร์ สนั่นแสน และโอพาริก สุรินดี. (2018). “โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกสำหรับการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ” *J Sci Technol MSU*. ปีที่ 2 ฉบับที่ 38 (มีนาคม – เมษายน) : 113-124.
- [3] นศพร ธรรมโชติ, ชัยสิทธิ์ ปรีชา และปิณฑน์ นุตรโคตร. (2563). “การจำแนกโรคพืชที่ปรากฏทางใบมะนาวด้วยวิธีการประมวลผลภาพบนสมาร์ตโฟน ระบบปฏิบัติการแอนดรอยด์” *รายงานการวิจัย คณะเกษตรศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย*.
- [4] ราตรี คำโหมง. (2021). “การตรวจจับมัจจุการการเรียนรู้เชิงลึก” *วารสารเทคโนโลยีสารสนเทศ*. ปีที่ 18 ฉบับที่ 1 (มกราคม – มิถุนายน) : 47-55
- [5] ณัฐวดี หงษ์บุญมี และธนวัฒน์ สิทธิโชคชัยศิริ. (2021). “การพัฒนาโปรแกรมประยุกต์วิเคราะห์สุขภาพดวงตาด้วยการรู้จำภาพและการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก” *วารสารศรีปทุมปริทัศน์ ฉบับวิทยาศาสตร์และเทคโนโลยี*. ปีที่ 13 (มกราคม – ธันวาคม) : 7-21.
- [6] จันทนา ปัญญารารักษ์, “โครงการศึกษาความเป็นไปของการวิเคราะห์โรคอัลไซเมอร์จากภาพถ่ายทางการแพทย์” มีนาคม 2561.
- [7] ศุภณัฐ จินตวัฒน์สกุล. (2562). “การจำแนกภาพชนิดแบบเซตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน” *วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ*.
- [8] ปิยะเนตร ต่วนชะเอม และเอกรัฐ รัชกาญจน์. (2022). “การจำแนกอาคารจากภาพถ่ายดาวเทียมการถ่ายโอนความรู้.” *Journal of Engineering and Digital Technology*. ปีที่ 10 ฉบับที่ 1 (มกราคม – มิถุนายน) : 28-40.
- [9] ณัฐวุฒิ ศรีวิบูลย์. (2021). “การปรับปรุงประสิทธิภาพการจำแนกภาพเอกซเรย์ทรวงอกด้วยโครงข่ายประสาทเทียมสังวัตนาการ โดยใช้เทคนิคการเพิ่มภาพสำหรับวินิจฉัยโรคโควิด-19.” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*. ปีที่ 31 ฉบับที่ 1 (มกราคม – มีนาคม) : 109-117.

การพยากรณ์จำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนน ในเขตกรุงเทพมหานครด้วยเทคนิคการวิเคราะห์อนุกรมเวลา Forecasting the Number of Victims from Road Accidents in Bangkok Using Time-Series Analysis Techniques

สุพรรณราช สีวารัชกุลวุฒิ (Suphannara Siwaratchakun)¹ และมหศักดิ์ เกตุฉำ (Mahasak Ketcham)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856046@email.kmutnb.ac.th, ²mahasak.k@itd.kmutnb.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ในการสร้างแบบจำลองเพื่อการวิเคราะห์จำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานครด้วยเทคนิคการวิเคราะห์อนุกรมเวลา โดยวิธีการเปรียบเทียบแบบจำลองในการพยากรณ์ จำนวน 6 วิธี ได้แก่ วิธีการวิเคราะห์การถดถอยเชิงเส้น วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย วิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง และวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบ Holt-Winters วัตถุประสงค์หลักของงานวิจัยคือการพยากรณ์ด้วยค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) และค่าเปอร์เซ็นต์รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSPE) ที่ต่ำที่สุด ทำการเก็บรวบรวมข้อมูลจำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร ระหว่างเดือนมกราคม 2561 - ธันวาคม 2565 จำนวน 60 เดือน โดยแบ่งข้อมูลเป็น 2 ชุด คือ ชุดข้อมูลสำหรับฝึกฝน (Train set) จำนวน 48 เดือน และชุดข้อมูลสำหรับทดสอบ (Test set) จำนวน 12 เดือน จากผลการวิจัยพบว่าจำนวนรวมของผู้ประสบภัยจำนวนของผู้ประสบภัยแบ่งตามเพศ และจำนวนของผู้ประสบภัยตามประเภทรถ ที่นำมาพยากรณ์ด้วยวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบ Holt-Winters มีประสิทธิภาพสูงสุด โดยมีค่า MAPE ร้อยละ 6.28, 5.84, 8.81, 6.06 และ 17.80 และค่า RMSPE ร้อยละ 9.50, 9.13, 10.51, 9.83 และ 20.22 ตามลำดับ

คำสำคัญ: เทคนิคการวิเคราะห์อนุกรมเวลา อุบัติเหตุบนท้องถนน การวิเคราะห์การถดถอยเชิงเส้น ค่าเฉลี่ยเคลื่อนที่ การปรับเรียบเอ็กซ์โพเนนเชียล

Abstract

The purpose of this research was to build a model for analyzing the number of road accident victims in Bangkok using time series analysis techniques. The research compares 6 forecasting models, including linear regression analysis, simple moving average, exponentially weighted moving average, single exponential smoothing, double exponential smoothing, and Holt-Winters exponential smoothing. The performance of the model was measured using the Mean Absolute Percentage Error (MAPE) and Root Mean Square Percentage Error (RMSPE), with the lowest values possible. The study collected data on the number of road accident victims in Bangkok from January 2018 to December 2022, a total of 60 months. The data was divided into 2 sets: a training set consisting of 48 months and a test set consisting of 12 months. The results found that the Holt-Winters exponential smoothing method was the most effective in predicting the total number of accident victims, the number of victims by gender, and the number of victims by vehicle type. The MAPE and RMSPE values for this method were the lowest, with a percentage of 6.28, 5.84, 8.81, 6.06, 17.80 and 9.50, 9.13, 10.51, 9.83, 20.22 respectively.

Keywords: Time Series Analysis, Linear Regression, Moving Average, Exponential Smoothing, Road Accidents.

1. บทนำ

อุบัติเหตุทางถนนเป็นปัญหาที่ต้องมีการจัดการและแก้ไข ข้อมูลสรุปจากองค์การอนามัยโลก (World Health Organization: WHO) ในปี 2561 พบว่า ประเทศไทยเป็นประเทศที่มีอัตราการเกิดอุบัติเหตุสูงเป็นอันดับ 9 ของโลก อันดับ 1 ในเอเชียและในภูมิภาคอาเซียน [1] ส่งผลให้เกิดความสูญเสียทั้งด้านร่างกาย จิตใจต่อผู้ประสบภัยและครอบครัว และยังมีผลกระทบต่อการพัฒนาทรัพยากรมนุษย์ เศรษฐกิจ และสังคมเป็นอย่างมาก

กรุงเทพมหานคร เป็นเมืองหลวงของประเทศไทย ที่มีการพัฒนาในด้านต่าง ๆ อย่างต่อเนื่อง ด้านที่เห็นได้ชัดคือ ด้านการคมนาคม ส่งผลให้มีผู้ประสบภัยจากการเกิดอุบัติเหตุเพิ่มขึ้นตามไปด้วย โดยสถิติจำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุทางถนนในปี 2561 มีจำนวน 102,551 คน ปี 2563 มีจำนวน 109,670 คน และปี 2565 มีจำนวน 119,019 คน [2] ซึ่งมีแนวโน้มพุ่งสูงขึ้นอย่างต่อเนื่อง

ปัจจุบันเทคโนโลยีทางคอมพิวเตอร์มีความก้าวหน้าเป็นอย่างมาก สามารถนำมาใช้ในการช่วยประมวลผลและวิเคราะห์ข้อมูล เพื่อนำข้อมูลที่ได้จากการวิเคราะห์มาใช้ประโยชน์ในด้านต่าง ๆ ซึ่งเทคนิคการวิเคราะห์อนุกรมเวลาเป็นเทคนิคการพยากรณ์ประเภทหนึ่ง โดยรวบรวมข้อมูลที่เกิดในอดีตตามช่วงเวลาเป็นเดือน หรือปี ที่มีความต่อเนื่องกัน มาพยากรณ์สิ่งที่คาดว่าจะเกิดขึ้นในอนาคต งานวิจัยที่ใช้เทคนิคดังกล่าว เช่น การพยากรณ์อนุกรมเวลาโดยการเปรียบเทียบวิธีแบบฉบับ และวิธีบ็อกซ์-เจนกินส์ กรณีศึกษาจำนวนการเกิดอุบัติเหตุในประเทศไทย [3] การพยากรณ์ยอดขายสินค้าด้วยเทคนิคการวิเคราะห์อนุกรมเวลา [4] เป็นต้น

จากผลการศึกษาข้างต้น ผู้วิจัยจึงมีแนวคิดในการนำข้อมูลจำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนน โดยสนใจเฉพาะข้อมูลในเขตกรุงเทพมหานครเท่านั้น มาวิเคราะห์ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา เพื่อทำการพัฒนาแบบจำลองและพยากรณ์ข้อมูลจำนวน

ผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร ได้อย่างแม่นยำ และสามารถนำข้อมูลที่พยากรณ์ไปใช้เพื่อเพิ่มประสิทธิภาพในการออกนโยบาย หรือเฝ้าระวัง เพื่อลดจำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีของการพยากรณ์ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา

ข้อมูลอนุกรมเวลา คือ ข้อมูลที่เก็บรวบรวมตามลำดับเวลาที่เกิดขึ้นอย่างต่อเนื่อง โดยมีช่วงห่างของการเก็บเท่ากัน เช่น รายปี รายเดือน และรายวัน เป็นต้น ตัวอย่างข้อมูลอนุกรมเวลาที่มีการนำมาใช้ในการประมวลผล อาทิ ข้อมูลรายได้ประชาชาติ ปริมาณเงินรายจ่ายรัฐบาล อัตราแลกเปลี่ยน เป็นต้น เพื่อศึกษาลักษณะธรรมชาติและรูปแบบการเคลื่อนไหวของข้อมูล ตลอดจนความสัมพันธ์ที่เกิดขึ้นระหว่างข้อมูล รวมทั้งการคาดคะเนแนวโน้มในอนาคต เพื่อรองรับการตัดสินใจในการดำเนินงานและการออกมาตรการต่าง ๆ [5]

2.2 วิธีการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Method : LR)

วิธีการวิเคราะห์การถดถอยเชิงเส้นเป็นวิธีการทางสถิติ ใช้เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป ว่าเป็นตัวแปรที่เป็นปัจจัยหรือเป็นเหตุผลของกันและกันหรือไม่

$$Y_i = \beta_0 + \beta_1 X_i$$

โดย Y_i = ตัวแปรตามที่ต้องการทำนาย

β_0 = ค่าคงที่ของสมการเชิงเส้น

β_1 = ค่าความชันของสมการเชิงเส้น

X_i = ตัวแปรอิสระใช้สำหรับทำนายค่า Y_i

2.3 วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย (Simple Moving Average : SMA)

วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่ายเป็นการนำข้อมูลในอดีตมาหาค่าเฉลี่ยที่เปลี่ยนไปตามช่วงเวลา โดยเหมาะกับข้อมูลที่มีลักษณะคงที่ ไม่เปลี่ยนแปลงมาก [4]

$$F_t = \frac{X_t + X_{t+1} + \dots + X_{(t-n)+1}}{n}$$

โดย F_t = ค่าพยากรณ์ในช่วงเวลา t

X_t = ค่าสังเกตในช่วงเวลา t

n = ค่าความชันของสมการเชิงเส้น

2.4 วิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพ

เนนเชียล (Exponentially Weight Moving Average : EWMA)

วิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียลเป็นเทคนิคที่พัฒนาจากวิธี SMA โดยใส่ค่าน้ำหนักของแต่ละข้อมูลตามลำดับความสำคัญ [4]

$$F_t = \alpha y_t + (1 - \alpha)F_{t-1}$$

โดย F_t = ค่าพยากรณ์ในช่วงเวลา t

y_t = ค่าสังเกตในช่วงเวลา t

α = ค่าคงที่ของการปรับเรียบ ($0 < \alpha < 1$)

2.5 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว (Single Exponential Smoothing : SES)

วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียวเป็นวิธีที่เหมาะสมกับข้อมูลที่เคลื่อนไหวอยู่ในระดับคงที่ แต่ให้ความสำคัญกับข้อมูลทุกข้อมูลที่เรียงลำดับจากอดีตจนถึงปัจจุบัน แต่ละค่าจะมีค่าคงที่ของการปรับเรียบเรียกว่า α โดยจะอยู่ในช่วงระหว่าง 0 ถึง 1

$$F_{t+1} = \alpha y_t + (1 - \alpha)F_t$$

โดย F_t = ค่าพยากรณ์ในช่วงเวลา t

y_t = ค่าสังเกตในช่วงเวลา t

2.6 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง (Double Exponential Smoothing : DES)

วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้งเป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีลักษณะเป็นแนวโน้มอย่างมีทิศทาง แต่ไม่เป็นฤดูกาล ซึ่งจะช่วยปรับค่าพยากรณ์ให้เป็นไปตามทิศทางโดยเพิ่มค่าคงที่สำหรับแนวโน้มเข้าไปในสมการ เรียกว่า β โดยอยู่ในช่วงระหว่าง 0 ถึง 1 [4]

2.6.1 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้งเชิงบวก

$$\begin{aligned} F_{t+h} &= l_t + hb_t \\ l_t &= \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \end{aligned}$$

2.6.2 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้งเชิงคูณ

$$\begin{aligned} F_{t+h} &= l_t(b_t)^h \\ l_t &= \alpha \frac{y_t}{b_{t-m}} + (1 - \alpha)(l_{t-1} * b_{t-1}) \\ b_t &= \beta \frac{l_t}{l_{t-1}} + (1 - \beta)b_{t-1} \end{aligned}$$

โดย F_{t+h} = ค่าพยากรณ์ที่เวลา $t+h$

h = ช่วงเวลาการพยากรณ์ในอนาคต

l_t = ค่าการประมาณระดับในช่วงเวลา t

b_t = ค่าการประมาณแนวโน้มในช่วงเวลา t

2.7 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ (Holt-Winters' Exponential Smoothing : HW)

วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์เป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีลักษณะของแนวโน้มและฤดูกาลที่ไม่เปลี่ยนแปลงตามเวลาที่เปลี่ยนแปลง โดยใช้ค่าปรับเรียบ α , β และ γ ที่อยู่ในช่วงระหว่าง 0 ถึง 1 [6]

2.7.1 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ เชิงบวก

$$\begin{aligned} F_{t+h} &= l_t + hb_t + s_{t+h-m(k+1)} \\ l_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \end{aligned}$$

2.6.2 วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโฮลท์-วินเทอร์ เชิงคูณ

$$\begin{aligned} F_{t+h} &= (l_t + hb_t)s_{t+h-m(k+1)} \\ l_t &= \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ s_t &= \gamma \frac{y_t}{(l_{t-1} - b_{t-1})} + (1 - \gamma)s_{t-m} \end{aligned}$$

โดย F_{t+h} = ค่าพยากรณ์ที่เวลา $t+h$

h = ช่วงเวลาการพยากรณ์ในอนาคต

m = จำนวนช่วงเวลาในหนึ่งฤดูกาล

k = จำนวนงวดที่จะพยากรณ์ล่วงหน้า

l_t = ค่าการประมาณระดับในช่วงเวลา t

b_t = ค่าการประมาณแนวโน้มในช่วงเวลา t

s_t = ค่าการประมาณฤดูกาลในช่วงเวลา t

2.8 งานวิจัยที่เกี่ยวข้อง

อดิเทพ ปานวนฉิชกิจ [4] ทำการพยากรณ์ยอดขายสินค้าด้วยเทคนิคการวิเคราะห์อนุกรมเวลา โดยใช้ 4 วิธี ได้แก่ วิธีค่าเฉลี่ยเคลื่อนที่ วิธีปรับเรียบเอ็กซ์โพเนนเชียลอย่างง่าย วิธีปรับเรียบแบบเอ็กซ์โพเนนเชียลกำลังสอง และวิธีแยกส่วนประกอบ ผลการทดลองพบว่า วิธีที่ให้ค่าการพยากรณ์ที่วนสอบที่ดีที่สุดคือ วิธีปรับเรียบแบบเอ็กซ์โพเนนเชียลกำลังสอง ให้ค่า MAPE ที่ 54.86 และ

นำเสนอข้อมูลผลการศึกษาศึกษาการพยากรณ์ผ่านทางโปรแกรม Power BI Desktop

Katleho D. Makatjane และ Ntebogang D. Moroke [7] ได้ทำการศึกษาศึกษาการพยากรณ์การขายนรถยนต์ตามฤดูกาลระยะสั้นในแอฟริกาใต้ โดยใช้ข้อมูลในช่วงเดือนมกราคม 1994 - ธันวาคม 2013 ผลการศึกษาพบว่า แบบจำลองวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ให้การคาดการณ์ตามฤดูกาลในระยะสั้นได้อย่างแม่นยำมากกว่าแบบจำลองวิธีบอซซ์-เจนกินส์

3. วิธีดำเนินการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อทำการศึกษ เปรียบเทียบ และคัดเลือกแบบจำลองที่เหมาะสมกับการพยากรณ์จำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร โดยมีขั้นตอนการดำเนินงานดังนี้

3.1 การเก็บรวบรวมข้อมูลและการจัดการข้อมูล

ในการศึกษาค้นคว้าได้ทำการเก็บรวบรวมข้อมูลจำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร จากศูนย์ข้อมูลอุบัติเหตุ เพื่อเสริมสร้างวัฒนธรรมความปลอดภัยทางถนน [2] โดยมีข้อมูลจำนวนผู้ประสบภัย เพศของผู้ประสบภัย และประเภทรถของผู้ประสบภัย ตั้งแต่เดือนมกราคม 2561 - ธันวาคม 2565 จำนวน 60 เดือน รวม 60 records ปรับรูปแบบของข้อมูลให้เหมาะสมกับแบบจำลองที่จะใช้ในการทดสอบ จากนั้นทำการแบ่งข้อมูลออกเป็น 2 ชุด โดยชุดแรกเป็นข้อมูลตั้งแต่เดือนมกราคม 2561 - ธันวาคม 2564 สำหรับใช้ในการสร้างแบบจำลองและเปรียบเทียบประสิทธิภาพในการพยากรณ์ ส่วนชุดที่สองเป็นข้อมูลตั้งแต่เดือนมกราคม - ธันวาคม 2565 สำหรับใช้ในการทดสอบความถูกต้องของแบบจำลอง

3.2 การพัฒนาแบบจำลองการพยากรณ์

ผู้วิจัยใช้ชุดข้อมูลเรียนรู้ (Training Data) ตั้งแต่เดือนมกราคม 2561 - ธันวาคม 2564 จำนวน 48 เดือนในการพัฒนาแบบจำลองการพยากรณ์ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา จำนวน 6 วิธี ได้แก่ วิธีการวิเคราะห์การถดถอยเชิงเส้น วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย วิธีหา

ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง และวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์

3.3 การวัดประสิทธิภาพและการคัดเลือกแบบจำลอง

เมื่อได้ผลจากการพัฒนาแบบจำลองแล้ว ต้องทำการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 6 วิธี และเลือกใช้แบบจำลองที่ให้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) และค่าเปอร์เซ็นต์รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSPE) ที่ต่ำที่สุด

3.4 การทวนสอบความถูกต้องของแบบจำลอง

นำแบบจำลองมาทวนสอบความถูกต้องในการนำไปใช้งาน โดยใช้ข้อมูลชุดทดสอบ (Test Data) ตั้งแต่เดือนมกราคม 2565 - ธันวาคม 2565 จำนวน 12 เดือน จากนั้นประเมินค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) และค่าเปอร์เซ็นต์รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSPE)

3.5 การนำเสนอข้อมูลพยากรณ์

นำเสนอข้อมูลการพยากรณ์จำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร ในรูปแบบรายงานด้วยโปรแกรม Power BI Desktop

4. ผลการดำเนินงาน

จากการศึกษาเพื่อพัฒนาแบบจำลองในการวิเคราะห์ข้อมูลด้วยเทคนิคการวิเคราะห์อนุกรมเวลา แสดงผลการวิจัยดังนี้

4.1 ผลการพัฒนาแบบจำลองการพยากรณ์

ผลการพัฒนาแบบจำลองด้วยเทคนิคการวิเคราะห์อนุกรมเวลา ได้แก่ วิธีการวิเคราะห์การถดถอยเชิงเส้น วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย วิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง และวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ ผลการวิจัยพบว่า MAPE และค่า RMSPE มีดังนี้

ตารางที่ 1: ประสิทธิภาพของวิธีการวิเคราะห์การถดถอยเชิงเส้น (LR) และวิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย (SMA)

ข้อมูล	LR		SMA	
	MAPE	RMSPE	MAPE	RMSPE
จำนวนรวมผู้ประสบภัย	11.99	15.92	10.26	22.19
เพศชาย	11.72	15.72	10.27	22.21
เพศหญิง	13.67	18.18	10.52	22.47
รถจักรยานยนต์	12.14	16.26	10.33	22.27
รถยนต์	16.36	14.69	10.37	22.12

ตารางที่ 2: ประสิทธิภาพของวิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล (EWMA) และวิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว (SES)

ข้อมูล	EWMA		SES	
	MAPE	RMSPE	MAPE	RMSPE
จำนวนรวมผู้ประสบภัย	9.42	13.50	11.13	15.95
เพศชาย	9.55	13.34	11.29	15.77
เพศหญิง	10.55	15.04	12.47	17.77
รถจักรยานยนต์	9.58	13.85	11.32	16.37
รถยนต์	8.78	12.18	10.37	14.39

ตารางที่ 3: ประสิทธิภาพของวิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง (DES) เริงบวกและเริงคูณ

ข้อมูล	DES เริงบวก		DES เริงคูณ	
	MAPE	RMSPE	MAPE	RMSPE
จำนวนรวมผู้ประสบภัย	7.88	11.98	7.81	12.17
เพศชาย	8.00	12.14	8.07	12.27
เพศหญิง	9.01	13.16	9.17	13.29
รถจักรยานยนต์	7.86	12.38	7.90	12.49
รถยนต์	8.85	12.05	8.95	12.15

ตารางที่ 4: ประสิทธิภาพของวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ (HW) เริงบวกและเริงคูณ

ข้อมูล	HW เริงบวก		HW เริงคูณ	
	MAPE	RMSPE	MAPE	RMSPE
จำนวนรวมผู้ประสบภัย	6.19	8.29	6.47	8.36
เพศชาย	6.45	8.66	6.77	8.91
เพศหญิง	7.08	9.56	7.16	9.76
รถจักรยานยนต์	6.27	8.44	5.98	8.13
รถยนต์	8.09	10.20	7.93	10.06

4.2 ผลการวัดประสิทธิภาพและการคัดเลือกแบบจำลอง

จากผลการวิจัย นำค่า MAPE และค่า RMSPE มาเปรียบเทียบเพื่อหาแบบจำลองที่เหมาะสม โดยพิจารณาจากค่าที่ต่ำที่สุด แสดงดังนี้

ตารางที่ 5: ตารางแสดงค่า MAPE ของแบบจำลองการพยากรณ์

Method	รวมผู้ประสบภัย	เพศผู้ประสบภัย		ประเภทรถ	
		ชาย	หญิง	รถจักรยานยนต์	รถยนต์
LR	11.99	11.72	13.67	12.14	16.36
SMA	10.26	10.27	10.52	10.33	10.37
EWMA	9.42	9.55	10.55	9.58	8.78
SES	11.13	11.29	12.47	11.32	12.18
DES เริงบวก	7.88	8.00	9.01	7.86	8.55
DES เริงคูณ	7.81	8.07	9.17	7.90	8.95
HW เริงบวก	6.19	6.45	7.08	6.27	8.09
HW เริงคูณ	6.47	6.77	7.16	5.98	7.93

ตารางที่ 6: ตารางแสดงค่า RMSPE ของแบบจำลองการพยากรณ์

Method	รวมผู้ประสบภัย	เพศผู้ประสบภัย		ประเภทรถ	
		ชาย	หญิง	รถจักรยานยนต์	รถยนต์
LR	15.92	15.72	18.18	16.26	14.69
SMA	22.19	22.21	22.47	22.27	22.12
EWMA	13.50	13.34	15.04	13.85	12.18
SES	15.95	15.77	17.77	16.37	14.39
DES เริงบวก	11.98	12.14	13.16	12.38	12.05
DES เริงคูณ	12.17	12.27	13.29	12.49	12.15
HW เริงบวก	8.29	8.66	9.56	8.44	10.20
HW เริงคูณ	8.36	8.91	9.76	8.13	10.06

จากตารางที่ 5 ถึง 6 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการพยากรณ์ พบว่า วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ เริงบวก ให้ประสิทธิภาพในการพยากรณ์จำนวนรวมผู้ประสบภัย และเพศของผู้ประสบภัยสูงสุด ส่วนวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ เริงคูณ ให้ประสิทธิภาพในการพยากรณ์ประเภทรถของผู้ประสบภัยสูงสุด โดยมีค่า MAPE และค่า RMSPE ต่ำที่สุดเมื่อเทียบกับวิธีอื่น

4.3 การทวนสอบความถูกต้องของแบบจำลอง

ทวนสอบความถูกต้องของแบบจำลองด้วยข้อมูลตั้งแต่เดือนมกราคม - ธันวาคม 2565 รวม 12 เดือน โดยใช้วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบไฮลท์-วินเทอร์ เริงบวก และเริงคูณ พบว่าการพยากรณ์จำนวนรวมผู้ประสบภัย ผู้ประสบภัยเพศชายและเพศหญิง ประเภทรถจักรยานยนต์ และรถยนต์ มีค่า MAPE ร้อยละ 6.28, 5.84, 8.81, 6.06 และ 17.80 และมีค่า RMSPE ร้อยละ 9.50, 9.13, 10.51, 9.83 และ 20.22 ตามลำดับ แสดงดังนี้

ตารางที่ 7: ผลการทวนสอบความถูกต้องของแบบจำลอง วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์เชิงบวก

ข้อมูล	HW เชิงบวก	
	MAPE	RMSPE
จำนวนรวมผู้ประสบภัย	6.28	9.50
เพศชาย	5.84	9.13
เพศหญิง	8.51	10.51

ตารางที่ 8: ผลการทวนสอบความถูกต้องของแบบจำลอง วิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์เชิงคูณ

ข้อมูล	HW เชิงคูณ	
	MAPE	RMSPE
รถจักรยานยนต์	6.06	9.83
รถยนต์	17.80	20.22

4.4 การนำเสนอข้อมูลการพยากรณ์

นำเสนอผลการพยากรณ์ที่ได้จากแบบจำลองการพยากรณ์ โดยใช้โปรแกรม Power BI Desktop สำหรับแสดงผลการพยากรณ์ในรูปแบบรายงาน แสดงดังภาพที่ 1



ภาพที่ 1: ตัวอย่างการนำเสนอผลการพยากรณ์ในรูปแบบรายงาน

จากภาพที่ 1 แสดงข้อมูลตั้งแต่เดือนมกราคม - ธันวาคม 2565 รวม 12 เดือน ที่ได้จากแบบจำลองการพยากรณ์

5. สรุปผลและข้อเสนอแนะ

งานวิจัยฉบับนี้ ผู้วิจัยได้ศึกษาการพยากรณ์ด้วยเทคนิคการวิเคราะห์อนุกรมเวลา เพื่อให้ได้วิธีพยากรณ์ที่เหมาะสมที่สุดในการพยากรณ์จำนวนผู้ประสบภัยจากการเกิดอุบัติเหตุบนท้องถนนในเขตกรุงเทพมหานคร ผู้วิจัยได้เลือกเทคนิคการพัฒนาแบบจำลองทั้งหมด 6 วิธี ได้แก่ วิธีการวิเคราะห์การถดถอยเชิงเส้น วิธีหาค่าเฉลี่ยเคลื่อนที่อย่างง่าย วิธีหาค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบเอ็กซ์โพเนนเชียล วิธีการปรับเรียบเอ็กซ์โพเนนเชียลครั้งเดียว วิธีการปรับเรียบเอ็กซ์โพเนนเชียลซ้ำ 2 ครั้ง และวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ ทดสอบประสิทธิภาพของแบบจำลองพยากรณ์ด้วยค่า MAPE และ

RMSPE ผลการวิจัยพบว่า วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์ มีความเหมาะสมที่สุดเนื่องจากมีค่า MAPE และค่า RMSPE ต่ำที่สุด หลังจากนั้นนำข้อมูลที่ได้จากการพยากรณ์มาทวนสอบกับข้อมูลจริงจำนวน 12 เดือน สรุปได้ว่า การพยากรณ์จำนวนรวมผู้ประสบภัย ผู้ประสบภัยเพศชายและเพศหญิง ประเภทรถจักรยานยนต์และรถยนต์ ด้วยวิธีวิธีการปรับเรียบเอ็กซ์โพเนนเชียลแบบโสมท์-วินเทอร์มีความแม่นยำ โดยมีค่า MAPE ร้อยละ 6.28, 5.84, 8.81, 6.06 และ 17.80 และมีค่า RMSPE ร้อยละ 9.50, 9.13, 10.51, 9.83 และ 20.22 ตามลำดับ เช่นเดียวกับงานวิจัยของ Katleho D. Makatjane และ Ntebogang D. Moroke [7]

เอกสารอ้างอิง

- [1] Natetida Bunnag. (2021). WHO เผยรายงาน ‘ไทย’ เสียชีวิตบนท้องถนนสูงสุดในเอเชียตะวันออกเฉียงใต้ แม้มีกฎหมายจราจรในเกณฑ์ดี สืบค้นเมื่อวันที่ 10 มีนาคม 2565 จาก <https://www.sdgmove.com/2021/02/10/who-global-status-report-on-road-safety-2018/>
- [2] ศูนย์ข้อมูลอุบัติเหตุ เพื่อเสริมสร้างวัฒนธรรมความปลอดภัยทางถนน. (2022). สืบค้นเมื่อวันที่ 6 มกราคม 2566 จาก <https://www.thairsc.com/thai-in-depth-data>
- [3] วราพร งามสุข. (2555). การพยากรณ์อนุกรมเวลา โดยการเปรียบเทียบวิธีแบบฉบับ และวิธีบอซซ์-เจนกินส์ กรณีศึกษาจำนวนการเกิดอุบัติเหตุในประเทศไทย. วิทยานิพนธ์ : ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา
- [4] อติเทพ ปานวนิชชกิจ. “การพยากรณ์ยอดขายสินค้าด้วยเทคนิคการวิเคราะห์อนุกรมเวลา.” *The Eighteenth National Conference on Computing and Information Technology* หน้า 394-399.
- [5] อุมาวดี เดชธำรงค์ และวิระพงษ์ จันทร์สนาม. “การพยากรณ์ข้อมูลอนุกรมเวลาราคาปิดหุ้นของบริษัทจดทะเบียนด้วยตัวแบบ ARIMA.” *วารสารบริหารธุรกิจ เศรษฐศาสตร์และการสื่อสาร* ปีที่ 13 ฉบับที่ 2 พ.ศ. - ส.ศ. 2561 หน้า 57-72.
- [6] Hyndman, R.J., & Athanasopoulos, G. *Forecasting: principles and practice, 2nd edition*, OTexts: Melbourne, 2018.
- [7] Makatjane, K. D., and Moroke, N. D. “Comparative study of holt-winters triples exponential smoothing and seasonal Arima: Forecasting short term seasonal car sales in South Africa.” *Risk Governance and Control: Financial Markets and Institutions* Page 71-82.

การพัฒนาแอปพลิเคชันแชทบอทสนับสนุนงานคอมพิวเตอร์ที่เกี่ยวข้องในองค์กร

Development of Chatbot Application Supports Computer Related Tasks in the Organization

เพิ่ม รัตนาภิรมย์ (Perm Ratanaphirom)¹ และมณเฑียร รัตนศิริวงศ์วุฒิ (Montean Rattanasiriwongwut)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857221@email.kmutnb.ac.th, ²montean.r@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์ เพื่อพัฒนาแอปพลิเคชันแชทบอทสนับสนุนงานคอมพิวเตอร์ที่เกี่ยวข้องในองค์กร ให้มีความสามารถในการตอบคำถามและให้คำแนะนำต่างๆ ในรูปแบบแชทที่สะดวกสบาย และรวดเร็ว ได้ตลอดเวลา ตามความต้องการขององค์กร การพัฒนาแอปพลิเคชันแชทบอทสนับสนุนงานคอมพิวเตอร์ที่เกี่ยวข้องในองค์กรจะใช้เทคโนโลยีปัญญาประดิษฐ์ เช่น การเรียนรู้เชิงลึก (deep learning) เพื่อให้แอปพลิเคชันสามารถเรียนรู้ และปรับปรุงการตอบคำถามและให้คำแนะนำได้อย่างมีประสิทธิภาพมากยิ่งขึ้น โดยนำข้อมูลที่ได้จากองค์กรมาทำการเปรียบเทียบกับอัลกอริทึม LSTM (Long Short-Term Memory) และ BLSTM (Bidirectional Long Short-Term Memory) เพื่อหาอัลกอริทึมที่มีประสิทธิภาพเหมาะสมกับงานมากที่สุด ผลการทดสอบพบว่า อัลกอริทึม BLSTM ให้ประสิทธิภาพสูงสุด โดยมีค่าความถูกต้องเท่ากับ 75.44% ค่าความแม่นยำเท่ากับ 79.08% ค่าระลึกเท่ากับ 75.44% และค่าความถ่วงดุลเท่ากับ 76.27% จึงเหมาะสำหรับงานในอนาคตจะเน้นการนำอัลกอริทึมนี้ไปพัฒนาระบบให้มีประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ: การเปรียบเทียบอัลกอริทึม สำหรับตัดคำ
อัลกอริทึมโครงข่ายประสาทเทียม ไลน์แอปพลิเคชัน

Abstract

The objective of this research is to develop a chatbot application that supports computer-related tasks within an organization. The application is

designed to provide convenient and quick responses to questions and recommendations based on chat format, according to the organization's needs. The development of this chatbot application involves the use of artificial intelligence technologies, such as deep learning, to enable the application to learn and improve its ability to provide efficient responses and recommendations. The data obtained from the organization is used to compare and evaluate the performance of LSTM (Long Short-Term Memory) and BLSTM (Bidirectional Long Short-Term Memory) algorithms to identify the most suitable algorithm for the task. The results of the testing show that the BLSTM algorithm provides the highest performance with an accuracy rate of 75.44%, precision rate of 79.08%, recall rate of 75.44%, and F1-score of 76.27%. This algorithm is therefore suitable for future development of a system that is highly efficient and effective.

Keywords: Algorithm Comparison for Word

Severing, Neural Network Algorithms,
Line Applications.

1. บทนำ

แชทบอทเป็นระบบจำลองการโต้ตอบตามธรรมชาติ พร้อมการสร้างแบบจำลองการสนทนาและความรู้พื้นฐานของปัญญาประดิษฐ์ เรียนรู้การตอบสนองต่อสถานการณ์ได้อย่างถูกต้องแม่นยำมากขึ้น และถูกสร้าง เพื่อให้อำนวย

สนทนาแทนมนุษย์ที่สามารถพูดคุยโต้ตอบสื่อสารกับมนุษย์ผ่านทางเสียง หรือข้อความได้

แชทบอทจัดเป็นหุ่นยนต์ประเภทหนึ่งที่สามารถทำงานได้ 24 ชั่วโมง ดังนั้นแชทบอทจึงถูกนำมาใช้ประโยชน์ทางด้านธุรกิจ เช่น ใช้สำหรับให้ข้อมูลคำแนะนำและวิธีการแก้ไขปัญหาเป็นต้น นอกจากนี้แล้วแชทบอทยังสามารถพัฒนาได้ทั้งบน LINE และบน Massager จึงทำให้องค์กรสามารถนำแชทบอทไปใช้อำนวยความสะดวกในการทำงานได้อย่างหลากหลายมากขึ้น

ปัจจุบันมีการขอคำแนะนำและวิธีการแก้ปัญหาในเรื่องเดิม ๆ อยู่ตลอดเวลา ทำให้ผู้วิจัยมีความสนใจที่จะพัฒนาแชทบอทมา เพื่อสนับสนุนการให้คำแนะนำและตอบคำถามเชิงแนะนำแก่ผู้ใช้งาน ซึ่งนวัตกรรมนี้จะเป็นเครื่องมือที่ช่วยรับทราบปัญหาและตอบสนองการให้คำแนะนำได้ทุกสถานที่ทุกเวลา

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาวิจัยครั้งนี้ผู้ทำวิจัยได้ใช้เวลาในการศึกษาทบทวนวรรณกรรมที่เกี่ยวข้องในหลายด้าน และจะมุ่งเน้นถึงงานวิจัยที่เกี่ยวข้องกับประเด็นสำคัญของเนื้อหาหัวข้อวิจัย เพื่อเป็นแนวทางสำหรับการศึกษาและพัฒนาระบบ โดยแบ่งเป็น 6 หัวข้อหลักดังนี้

2.1 การประมวลผลภาษาธรรมชาติ (Natural Language Processing)

การประมวลผลภาษาธรรมชาติ (NLP) เป็นเทคโนโลยีแมชชีนเลิร์นนิงที่ช่วยให้คอมพิวเตอร์สามารถตีความจัดการ และทำความเข้าใจภาษามนุษย์ได้ องค์กรในปัจจุบันมีข้อมูลเสียง และข้อความจำนวนมากจากช่องทางสื่อสารต่าง ๆ เช่น อีเมล ข้อความ ฟีดข่าวโซเชียลมีเดีย วิดีโอ เสียง และอื่น ๆ พวกเขาใช้ซอฟต์แวร์ NLP เพื่อประมวลผลข้อมูลนี้โดยอัตโนมัติ วิเคราะห์เจตนา หรือความเชื่อมั่นในข้อความ และตอบสนองการสื่อสารของมนุษย์แบบเรียลไทม์

2.2 กระบวนการตัดคำ (Word Segmentation)

ความมุ่งหวังที่จะให้คอมพิวเตอร์สามารถประมวลผลภาษาไทยได้อย่างมีประสิทธิภาพนั้นมีปัญหาเบื้องต้น คือ

ลักษณะการเขียนภาษาไทยจะเขียนติดต่อกันเป็นสายอักขระ โดยไม่มีเครื่องหมายวรรคตอนแสดงการแบ่งคำ ดังเช่น ภาษาอังกฤษ ซึ่งเป็นอุปสรรคอย่างหนึ่งที่ต้องการการศึกษาวิจัย และพัฒนา เพื่อให้คอมพิวเตอร์สามารถคำนวณ เพื่อแบ่งสายอักขระไทยออกเป็นคำ ๆ ซึ่งจะส่งผลให้การทำงานของคอมพิวเตอร์ในการค้นหาคำใด ๆ ทำได้อย่างถูกต้อง และแม่นยำรวมถึงการจัดขอบข่ายในโปรแกรมประมวลผลคำ (Word processor) เป็นต้น

2.3 การเรียนรู้ของเครื่อง (Machine Learning)

การใช้การเรียนรู้ของเครื่องจักร หรือการสร้าง และพัฒนาอัลกอริทึม และโมเดลคอมพิวเตอร์ ที่สามารถใช้งานได้ โดยไม่ต้องระบุ โดยตรงว่าจะทำอะไรบ้าง โดยมีวิธีการเรียนรู้แบบหลายแบบ เช่น supervised learning, unsupervised learning, semi-supervised learning และ reinforcement learning เป็นต้น โดยคอมพิวเตอร์สามารถเรียนรู้ได้จากข้อมูลที่ได้รับการเรียงลำดับ และการประมวลผลของข้อมูลเหล่านั้น เพื่อสร้างโมเดลที่สามารถใช้งานได้ในการแก้ไขปัญหาต่าง ๆ ในชีวิตประจำวัน เช่น การจัดเรียงข้อมูล เชิงพหุนาม การตัดสินใจ เป็นต้น

Machine Learning มีการนำไปใช้งานอย่างแพร่หลาย ในหลายสาขาอุตสาหกรรม เช่น การแยกแยะอักษรในการจดหมายอิเล็กทรอนิกส์ (Email Spam Classification) การตรวจจับภาพ และวิดีโอ (Image and Video Recognition) การแนะนำสินค้า (Product Recommendations) การจัดลำดับผลการค้นหา (Search Ranking) การตัดสินใจ (Decision Making) การรู้จำ (Memory-based Learning) และอื่น ๆ อีกมากมาย

2.4 อัลกอริทึมสำหรับการทำเหมืองข้อความ

อัลกอริทึมสำหรับการทำเหมืองข้อความมักใช้เทคโนโลยีปัญญาประดิษฐ์ และการเรียนรู้เชิงลึก เช่น การใช้อัลกอริทึมการเรียนรู้เชิงลึก (Deep Learning) เพื่อหาคำสำคัญ และการสร้างลกอริทึมที่สามารถจำแนก และจัดกลุ่มข้อมูลได้

BLSTM ย่อมาจาก "Bidirectional Long Short-Term Memory" ซึ่งเป็นแบบจำลอง (model) ของโครงข่ายประสาทเทียม (neural network) ที่เป็นการรวมผสาน

โครงข่าย LSTM สองส่วนเข้าด้วยกัน เพื่อให้อัลกอริทึมสามารถดู และเรียนรู้ข้อมูลจากทั้งสองทิศทาง ด้วยความสามารถของ LSTM ที่สามารถจดจำ และติดตามสถานะ (state) ของข้อมูลก่อนหน้าได้ เราจึงสามารถใช้ BLSTM ในการจำแนก และทำนายข้อมูลที่มีลำดับ หรือความสัมพันธ์กันระหว่างข้อมูล ได้ดีกว่า LSTM เดียวเพียงอย่างเดียว

LSTM ย่อมาจาก "Long Short-Term Memory" ซึ่งเป็นแบบจำลอง (model) ของโครงข่ายประสาทเทียม (neural network) ที่ใช้ในการจำแนก และทำนายข้อมูลที่มีลำดับ หรือความสัมพันธ์กันระหว่างข้อมูล เช่น ข้อมูลเสียง (speech recognition) และข้อความ (text prediction) ซึ่งมักจะมีการเปลี่ยนแปลงตามลำดับเวลา (time series data)

2.5 Line Messaging API

Line Messaging API คือ การสื่อสารระหว่างบริการของผู้ให้บริการ และผู้ใช้งานออนไลน์ เป็นการสื่อสารแบบสองฝ่าย จะทำให้สามารถให้บริการได้ในห้องแชท เพื่อให้บริการที่เหมาะสมสำหรับผู้ใช้งานออนไลน์แต่ละคน และ Messaging API จะส่ง และรับข้อมูลระหว่างเซิร์ฟเวอร์ และแอปพลิเคชันไลน์ผ่านทางเซิร์ฟเวอร์ของทางไลน์ การส่งค่าขอจะใช้ API แบบ JSON Messaging API ทำการเชื่อมต่อระหว่างผู้ใช้งานผ่านทาง LINE Official Account Messaging API นั้นเอง

2.6 งานวิจัยที่เกี่ยวข้อง

ฐิติภรณ์ และทศนันทน์ [1] การพัฒนาแชทบอทปัญหาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา” จากการศึกษาทดสอบประสิทธิภาพ โดยกำหนดบทสนทนาของนักศึกษา จำนวน 5 ข้อความ แล้วพิมพ์ให้แชทบอทตอบบทสนทนา จากนั้นทำการประเมินประสิทธิภาพด้วยพิจารณาคำตอบของแชทบอทที่แสดงออกว่ามีจำนวนคำที่ตรงไม่ตรงกับความต้องการ หรืออยู่นอกขอบเขตการสนทนา ซึ่งให้ผลลัพธ์ว่ามีความเที่ยงตรง (Precision = 88.2%) ความแม่นยำ (Recall = 100%) และการวัดประสิทธิภาพ (F-Measure = 93.6%)

สุนิสา [3] ได้ทำการศึกษาการยอมรับเทคโนโลยี Chatbot ในธุรกิจธนาคารในประเทศไทย งานวิจัยนี้เป็นงานวิจัยเชิงคุณภาพ ใช้แบบจำลองในการยอมรับทั้งหมด 6 ด้าน ได้แก่ ด้านปัจจัยภายนอก (External Variables) ด้านการรับรู้ว่าเป็นระบบที่ง่ายต่อการใช้งาน (Perceived Ease of Use หรือ PEOU) ด้านการรับรู้ถึงประโยชน์ที่จะได้รับของเทคโนโลยีนั้น (Perceived Usefulness หรือ PU) ด้านทัศนคติที่มีต่อการใช้งาน (Attitude toward Using) ด้านความตั้งใจแสดงพฤติกรรมการใช้ (Behavioral Intention to Use หรือ BI) และด้านการนำเทคโนโลยีใหม่มาใช้งาน (Actual System Use) เพื่อเป็นกรอบแนวทางในการศึกษาการยอมรับเทคโนโลยีแชทบอทในธุรกิจธนาคารในประเทศไทย โดยข้อมูลที่ได้จากการสัมภาษณ์จะเป็นลักษณะบรรยาย และใช้วิธีวิเคราะห์ข้อมูลแบบอุปนัย ผลการศึกษาพบว่า ธนาคารในประเทศไทยยอมรับเทคโนโลยีแชทบอทมาใช้ในองค์กรด้วยปัจจัยภายนอกที่สำคัญ ได้แก่ การเปลี่ยนแปลงพฤติกรรมของผู้บริโภค และการขาดแคลนเจ้าหน้าที่ลูกค้าสัมพันธ์ โดยธนาคารมองเห็นถึงโอกาสในการนำแชทบอทมาใช้เพื่อพัฒนาบริการที่ดีขึ้น รวดเร็วขึ้น ให้แก่ลูกค้า รวมถึงการทำการตลาดแบบเฉพาะเจาะจงพัชรียา ทองพูล และคณะ. (2562). “การเปรียบเทียบประสิทธิภาพในการทำนายผลการปรับความไม่สมดุลของข้อมูลในการจำแนกด้วยอัลกอริทึมการทำเหมืองข้อมูล.” วารสารวิทยาศาสตร์และเทคโนโลยี. ปีที่ 8 ฉบับที่ 6 (พฤศจิกายน - ธันวาคม) : 565-584.

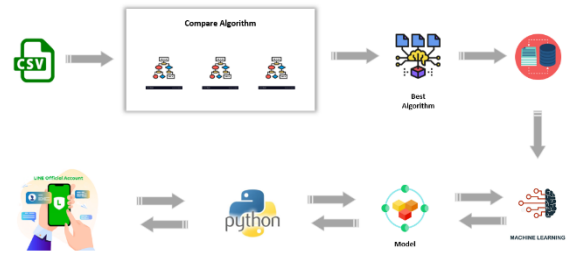
เฉลิมฤทธิ์ และณัฐวี [5] นำเสนองานวิจัยเรื่องการประยุกต์ใช้หุ่นยนต์สนทนาบนระบบ Line เพื่อลดภาระงาน Help Desk ในการให้บริการด้านสารสนเทศการใช้งานระบบบริหารจัดการงานเอกสาร โดยผู้วิจัยได้พัฒนาหุ่นยนต์สนทนามาเพื่อช่วย Help Desk ในการดูแล รับปัญหา และแก้ไขปัญหาจากผู้ใช้งาน ซึ่งนำวิธีการเรียนรู้ของเครื่องมาใช้ในการเรียนรู้แบบมีผู้สอน เรียนรู้บทสนทนา หรือข้อความที่มีปริมาณมาก โดยใช้เทคนิคในการจำแนกข้อมูลคุณลักษณะของคำ และรูปประโยค และประมวลผลภาษาธรรมชาติ และนำไปสร้างโมเดลการสนทนาเพื่อพยากรณ์ว่าผู้ใช้งานต้องการจะสื่ออะไร

ผลการวิจัยพบว่าการทดสอบประสิทธิภาพของโมเดลกับชุดข้อมูลตัวอย่างจำนวน 50 ชุด พบว่าโมเดลมีประสิทธิภาพการทำนายที่ค่าความถูกต้องร้อยละ 90 ซึ่งอยู่ในระดับที่ดีมา

แทนพงศ์ และมหศักดิ์ [6] ได้ทำการพัฒนาหุ่นยนต์สนทนาบนแอปพลิเคชันไลน์เพื่อสนับสนุนการทำงานของแผนกทรัพยากรบุคคลแบบอิเล็กทรอนิกส์ งานวิจัยนี้เป็นการประยุกต์ใช้ปัญญาประดิษฐ์กับการพัฒนาหุ่นยนต์สนทนา เพื่อที่จะช่วยลดการทำงานของแผนกทรัพยากรบุคคล ใช้เทคนิคการตัดคำ โดยใช้พจนานุกรม (Lexto) และใช้เทคนิค TF-IDF (Term Frequency-Inverse Document Frequency) เพื่อหาคำที่สำคัญในเอกสาร โดยนับจำนวนของคำในเอกสาร และความถี่กับเอกสารทั้งหมดเพื่อเพิ่มประสิทธิภาพในการสืบค้นข้อมูล และวัดประสิทธิภาพของโมเดล โดยใช้ Confusion Matrix คำนวณผลการทำนายของ โมเดล จากผลการวิจัยพบว่า จากข้อความจำนวน 20 ข้อความให้ความถูกต้องร้อยละ 78 ซึ่งสามารถนำไปใช้ช่วยแผนกบุคคลตอบคำถาม และให้ข้อมูลการลาแก่พนักงานในองค์กรได้ณัฐกิจ เจนการ และมหศักดิ์ เกตุฉ่ำ. (2562). การพัฒนาแบบจำลองในการตรวจจับข้อความภาษาไทยที่เป็นการกลั่นแกล้งทางไซเบอร์ โดยใช้วิธีชัพพอร์ตเวกเตอร์แมชชีน. The 15th National Conference and 2019-1 International Conference on Applied Computer Technology and Information Systems and 2019-1 National Conference on Business Administration, 144-149.

3. วิธีการดำเนินการวิจัย

เนื่อหางานวิจัยครั้งนี้ผู้วิจัยได้ทำการศึกษา และทำความเข้าใจ เพื่อกำหนดขั้นตอนการศึกษา และวางแนวทางในการพัฒนาระบบแชทบอทบนแพลตฟอร์มไลน์ เพื่อสนับสนุนงานการใช้ข้อมูลที่เกี่ยวข้องทางคอมพิวเตอร์ได้ถูกต้องแม่นยำ และตลอดเวลา ผู้วิจัยได้วางแผน และกำหนดขั้นตอนการดำเนินงาน เพื่อเป็นแนวทางในการปฏิบัติงาน แสดงดังภาพที่ 1



ภาพที่ 1: กรอบแนวความคิด

จากภาพที่ 1 กรอบแนวความคิดผู้วิจัยแบ่งขั้นตอนการดำเนินการวิจัยเป็น 3 ส่วนดังนี้

3.1 การรวบรวมและจัดเตรียมข้อมูล

ผู้วิจัยได้ทำการศึกษา และรวบรวมข้อมูลคำถามจากการปฏิบัติงานมาทำการวิเคราะห์ พบว่าข้อมูลคำถามเป็นคำถามเดิม ๆ บ่อยครั้ง ผู้วิจัยจึงได้ทำการรวบรวมข้อมูลที่ได้เก็บมาแบ่งออกเป็นหมวดหมู่ เพื่อแยกข้อมูลออกเป็นหมวดหมู่ และในแต่ละหมวดหมู่จะมีข้อมูลครอบคลุมต่อการทดสอบกับระบบ แล้วนำมาจัดเก็บให้อยู่ในรูปแบบของไฟล์นามสกุล .CSV เพื่อให้สามารถนำไปใช้ในการเปรียบเทียบอัลกอริทึมได้อย่างสะดวกยิ่งขึ้น

3.2 การเปรียบเทียบอัลกอริทึม

ขั้นตอนการเปรียบเทียบอัลกอริทึม ผู้วิจัยได้ทำการแบ่งการเปรียบเทียบอัลกอริทึมออกเป็น 2 ส่วน ดังนี้

3.2.1 อัลกอริทึมที่ใช้ในการตัดคำจากไลบรารี PyThaiNLP ซึ่งพัฒนาด้วยภาษา Python มาใช้สำหรับวัดประสิทธิภาพของแต่ละอัลกอริทึม ด้วยการวัดค่าความถูกต้องแม่นยำ และความเร็วในการประมวลผล โดยเลือกใช้ Colab เป็นเครื่องมือในการเปรียบเทียบอัลกอริทึม ในครั้งนี้ เนื่องจากมีการใช้งานที่ง่าย และยังสามารถ Import หรือ Export ได้สะดวกมากขึ้น ผู้วิจัยได้เลือกอัลกอริทึมที่มีความนิยมมาใช้ในการเปรียบเทียบ 3 อัลกอริทึม ได้แก่ Newmm Deepcut และ Attacut แสดงดังภาพที่ 2 และ 3

```
import csv
import time
import pandas as pd
import deepcut
import attacut
from pythainlp.tokenize import word_tokenize, Tokenizer

# โหลดไฟล์ csv และเลือกคอลัมน์ที่เราต้องการ
df = pd.read_csv('data.csv')
texts = df["Problem"].tolist()

# สร้างฟังก์ชันสำหรับตัดคำด้วย newmm
def tokenize_newmm(text):
    return word_tokenize(text, engine='newmm')

# สร้างฟังก์ชันสำหรับตัดคำด้วย deepcut
def tokenize_deepcut(text):
    return deepcut.tokenize(text)

# สร้างฟังก์ชันสำหรับตัดคำด้วย attacut
def tokenize_attacut(text):
    return attacut.tokenize(text)

# ใช้ฟังก์ชัน .apply() เพื่อประมวลผลของ DataFrame มาตรฐานด้วย deepcut, attacut, และ newmm
df["deepcut"] = df["Problem"].apply(tokenize_deepcut)
df["attacut"] = df["Problem"].apply(tokenize_attacut)
df["newmm"] = df["Problem"].apply(tokenize_newmm)

# สร้างไฟล์ CSV ใหม่โดยใช้ฟังก์ชัน .to_csv()
df.to_csv("output.csv", index=False, encoding='utf-8-sig')
```

ภาพที่ 2: ได้ผลการเปรียบเทียบอัลกอริทึมการตัดคำ

```
import csv
import timeit
from pythainlp.tokenize import word_tokenize
import deepcut
import attacut

def pythainlp_word_cut(text):
    return word_tokenize(text, engine='newmm')

def deepcut_word_cut(text):
    return deepcut.tokenize(text)

def attacut_word_cut(text):
    return attacut.tokenize(text)

# เก็บเวลาในการทำงานของแต่ละวิธี
pythainlp_times = []
deepcut_times = []
attacut_times = []

for text in texts:
    pythainlp_time = timeit.timeit(lambda: pythainlp_word_cut(text), number=100)
    deepcut_time = timeit.timeit(lambda: deepcut_word_cut(text), number=100)
    attacut_time = timeit.timeit(lambda: attacut_word_cut(text), number=100)
    pythainlp_times.append(pythainlp_time)
    deepcut_times.append(deepcut_time)
    attacut_times.append(attacut_time)

# สร้าง dataframe เพื่อเก็บผลลัพธ์
result_df = pd.DataFrame({
    'Text': texts,
    'Pythainlp Time': pythainlp_times,
    'DeepCut Time': deepcut_times,
    'Attacut Time': attacut_times
})

# บันทึกผลลัพธ์ลงไฟล์ csv
result_df.to_csv('result.csv', index=False)
```

ภาพที่ 3: ได้ผลการเปรียบเทียบอัลกอริทึมการหาค่าความเร็วในการประมวลผล

3.2.2 ในการเปรียบเทียบอัลกอริทึม BLSTM (Bidirectional Long Short-Term Memory) และ LSTM (Long Short-Term Memory) หา อัล ก อ ริ ทึม ที่มีประสิทธิภาพเหมาะสมกับงานที่สุด โดยผู้วิจัยได้กำหนดค่าพารามิเตอร์ ดังนี้

1. Cells คือ จำนวนของ units ในแต่ละ layer
2. Epochs คือ จำนวนรอบการเทรน (training iterations) ที่อัลกอริทึมจะทำการเรียนรู้
3. Batch_size คือ จำนวนตัวอย่างที่จะถูกนำมาใช้ในแต่ละรอบการเทรน แสดงดังภาพที่ 4 และ 5

```
#Bidirectional(LSTM)
model = Sequential()
model.add(Input(shape=(X_train.shape[1],)))
model.add(Embedding(MAX_NB_WORDS, 300, input_length= (X_train.shape[1],)))
model.add(SpatialDropout1D(0.2))
model.add(Bidirectional(LSTM(units=64, dropout= 0.2 ,recurrent_dropout= 0.2)))
model.add(Dense(y_train.shape[1], activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
print(model.summary())

num_units = 128
epochs = 200
batch_size = 32
history = model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size, validation_split = 0.1)
```

ภาพที่ 4: อัลกอริทึม BLSTM (Bidirectional Long Short-Term Memory)

```
#LSTM
model = Sequential()
model.add(Input(shape=(X_train.shape[1],)))
model.add(Embedding(MAX_NB_WORDS, 300, input_length= (X_train.shape[1],)))
model.add(SpatialDropout1D(0.2))
model.add(LSTM(units=128,dropout= 0.2 ,recurrent_dropout= 0.2))
model.add(Dense(y_train.shape[1], activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
print(model.summary())

num_units = 128
epochs = 200
batch_size = 128
history = model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size, validation_split = 0.1)
```

ภาพที่ 5: อัลกอริทึม LSTM (Long Short-Term Memory)

จากภาพที่ 4 และ 5 ผู้วิจัยจำนำผลลัพธ์ที่ได้มาใช้ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าระลึก (Recall) และค่าถ่วงดุล (F-Measure) โดยพิจารณาจากค่าความตามสูตรสมการ ดังนี้

- ความถูกต้อง (Accuracy) หมายถึง ค่าร้อยละของ True Positive Rate คำนวณได้จาก แสดงดังสมการที่ 6

$$\text{Accuracy} = \text{True Positives Rate} \times 100\% \quad (6)$$

- ค่าความแม่นยำ (Precision) คือ ค่าของโมเดลที่ทำนายถูกต้อง คำนวณได้จากจำนวนข้อมูลที่ทำนายถูกในคลาสนั้นหารด้วยจำนวนข้อมูลทั้งหมดที่ทำนายให้ผลลัพธ์เดียวกันในคลาสนั้น Precision ของคลาสนั้นเป้าหมาย Yes มีค่าเท่ากับ แสดงดังสมการที่ 7

$$\text{Precision (Yes)} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} \quad (7)$$

- ค่าระลึก (Recall) หมายถึง ค่าจากการทำนายด้วยโมเดลที่ตรงกับความเป็นจริง ซึ่งมีค่าเท่ากับ True Positive Rate คำนวณได้จากจำนวนข้อมูลที่ทำนายถูกในคลาสนั้นหารด้วยจำนวนข้อมูลทั้งหมดในคลาสนั้น Recall ของคลาสนี้เป้าหมาย Yes มีค่าเท่ากับ แสดงดังสมการที่ 8

$$\text{Recall (Yes)} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (8)$$

- ค่าความถ่วงดุล (F-Measure) หมายถึง ค่าที่เกิดจากการเปรียบเทียบระหว่างค่า Precision และ Recall ของแต่ละคลาสนี้เป้าหมาย F-Measure ของคลาสนี้เป้าหมาย Yes มีค่าเท่ากับ แสดงดังสมการที่ 9

$$\text{F-Measure (Yes)} = \frac{2 \times \text{Precision(Yes)} \times \text{Recall(Yes)}}{\text{Precision(Yes)} + \text{Recall(Yes)}} \quad (9)$$

3.3 การออกแบบและพัฒนาระบบ

ผู้วิจัยได้ทำการออกแบบ และพัฒนาแอปพลิเคชัน แชนทอปด้วยระบบไลน์ โดยเริ่มจากผู้ใช้งานป้อนคำถามผ่านระบบไลน์แล้วนำข้อมูลที่รับเข้าผ่านทาง Line ไปประมวลผลยังแบบจำลองที่พัฒนาขึ้น เพื่อวิเคราะห์ข้อความจากผู้ใช้งาน และนำไปเปรียบเทียบกับชุดข้อมูลบนฐานข้อมูล เพื่อส่ง Response กลับไปยังผู้ใช้ แสดงดังภาพที่ 10



ภาพที่ 10: การออกแบบและพัฒนาระบบ

4. ผลการดำเนินงาน

ผลการดำเนินงาน วิจัยในการประยุกต์ใช้ปัญญาประดิษฐ์กับหุ่นยนต์สนทนาบนระบบไลน์ เพื่อนำมาช่วยในการบริหารงาน IT Support ขึ้นตอนหลังจากการเปรียบเทียบอัลกอริทึมนั้น ผู้วิจัยได้นำอัลกอริทึมที่มีประสิทธิภาพที่ดีที่สุดของแต่ละประเภทมาทำการพัฒนาเข้ากับระบบไลน์ แล้วทำการทดสอบการทำงานของระบบ

เพื่อหาข้อผิดพลาดในการทำงานแล้วนำมาปรับปรุงประสิทธิภาพการทำงานของระบบให้ดียิ่งขึ้น ซึ่งได้ผลการดำเนินการวิจัยออกเป็น 2 ส่วน โดยมีรายละเอียด ดังนี้

4.1 ผลการเปรียบเทียบอัลกอริทึม

ตารางที่ 11: ผลการหาประสิทธิภาพการตัดคำ

Segmentation	Accuracy	Second
Attacut	94.40	1.6313
newmm	91.15	0.0112
DeepCut	90.86	15.1551

จากตารางที่ 1 สรุปได้ว่า attacut, newmm และ deepcut มีค่าความถูกต้อง 96.76%, 91.15%, 90.86% ตามลำดับ มีค่าความถูกต้องใกล้เคียงกัน และ newmm มีค่าความเร็วในการประมวลผลที่ 0.0112 วินาที ซึ่งเร็วกว่า Attacut และ Deepcut ทางผู้วิจัยจึงได้เลือกใช้ attacut ที่มีค่าความถูกต้องสูง และมีค่าความเร็วไม่แตกต่างกันมากนัก

4.2 ผลการเปรียบเทียบแบบจำลอง

ตารางที่ 12: ผลการเปรียบเทียบแบบจำลอง BLSTM และ LSTM

Parameter			Result			
Cells	Epochs	Batch size	Accuracy	Precision	Recall	F-measure
BLSTM (Bidirectional Long Short-Term Memory)						
32	100	32	73.68	76.03	73.68	74.36
64	100	32	73.68	75.88	73.68	74.3
128	100	32	73.68	77.26	73.68	74.57
32	100	64	71.05	73.54	71.05	71.9
64	100	64	71.05	73.21	71.05	71.7
128	100	64	74.56	77.44	74.56	75.32
32	100	128	74.56	76.95	74.56	75.28
64	100	128	73.68	76.88	73.68	74.61
128	100	128	74.56	76.9	74.56	75.34
32	200	32	74.56	77.58	74.56	75.41
64	200	32	74.56	78.14	74.56	75.44
128	200	32	75.44	79.08	75.44	76.27
32	200	64	73.68	76.28	73.68	74.33
64	200	64	71.93	74.09	71.93	72.65
128	200	64	73.68	77.5	73.68	74.63
32	200	128	73.68	76.79	73.68	74.62
64	200	128	72.81	76.08	72.81	73.8
128	200	128	72.81	75.08	72.81	73.48
LSTM (Long Short-Term Memory)						
32	100	32	71.05	72.62	71.05	71.65
64	100	32	71.93	73.23	71.93	72.37
128	100	32	72.81	73.9	72.81	73.22
32	100	64	72.81	74.05	72.81	73.25
64	100	64	71.93	73.4	71.93	72.42
128	100	64	70.18	71.98	70.18	70.86
32	100	128	72.81	74.33	72.81	73.36
64	100	128	73.68	74.54	73.68	74
128	100	128	71.93	73.12	71.93	72.3
32	200	32	72.81	73.86	72.81	72.92
64	200	32	70.18	71.39	70.18	70.44
128	200	32	71.05	72.3	71.05	71.45
32	200	64	71.05	72.3	71.05	71.45
64	200	64	71.93	74.18	71.93	72.44
128	200	64	73.68	75.72	73.68	74.06
32	200	128	71.05	72.19	71.05	71.25
64	200	128	72.81	74.85	72.81	73.29
128	200	128	72.81	74.85	72.81	73.29

จากตารางที่ 12 สรุปได้ว่า BLSTM มีประสิทธิภาพดีที่สุด โดยกำหนดค่า Cells = 128, Epochs = 200, Batch size = 32 ทำให้ได้ค่าความถูกต้องเท่ากับ 75.44% ค่าความแม่นยำเท่ากับ 79.08% ค่าระลึกเท่ากับ 75.44% และค่า

ถ่วงดุลเท่ากับ 76.27% ซึ่งทำให้เหมาะสมกับงานวิจัยมากที่สุด ในส่วนของแบบจำลอง LSTM มีประสิทธิภาพที่ดีที่สุด โดยกำหนดค่า Cells = 68, Epochs = 100, Batch size = 128 ทำให้ได้ค่าความถูกต้องเท่ากับ 73.68% ค่าความแม่นยำเท่ากับ 74.54% ค่าระลอกเท่ากับ 73.68% และค่าถ่วงดุลเท่ากับ 74.00%

4.3 ผลการพัฒนาระบบ

ผลลัพธ์จากการพัฒนาระบบหุ่นยนต์สนทนาบนระบบไลน์ เพื่อนำมาประยุกต์ใช้กับงานในการบริหารงาน IT Support โดยผู้ใช้งานจะทำการป้อนคำถาม และระบบจะนำคำถามไปทำการประมวลผลด้วยอัลกอริทึมที่มีประสิทธิภาพมากที่สุด และส่งคำตอบกลับไปยังผู้ใช้ แสดงดังภาพที่ 13



ภาพที่ 13: ข้อมูลผลการสอบถาม

4.4 ผลการประเมินความพึงพอใจในระบบ

สรุปผลการประเมินความพึงพอใจในการใช้งานระบบ มีค่าเฉลี่ยเท่ากับ 3.99 และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.60 แสดงให้เห็นว่าผู้ใช้งานมีความคิดเห็นในการใช้งานระบบอยู่ในระดับดี แสดงดังตารางที่ 14

ตารางที่ 14: สรุปผลการประเมินความพึงพอใจในการใช้งานระบบ

รายการประเมิน	ระดับความพึงพอใจ		
	\bar{X}	S.D.	แปลผล
1. ด้านรับรู้ประโยชน์ (PU : Perceived Usefulness)	4.12	0.61	ดี
2. ด้านการรับรู้ความง่ายในการใช้งาน (PE : Perceived Ease of Use)	4.00	0.47	ดี
3. ด้านทัศนคติต่อการใช้ (AU : Attitude Toward Using)	3.92	0.71	ดี
4. ด้านพฤติกรรมการใช้งาน (BI : Behavioral Intention)	3.92	0.63	ดี
สรุปผลการประเมิน	3.99	0.60	ดี

5. สรุป

การประยุกต์ใช้ปัญญาประดิษฐ์กับหุ่นยนต์สนทนาบนระบบไลน์ เพื่อนำมาประยุกต์ใช้งานสนับสนุนการให้บริการข้อมูลงาน IT Support ที่เกี่ยวข้องในองค์กร ซึ่งผู้วิจัยได้ทำการแบ่งการเปรียบเทียบอัลกอริทึมออกเป็น 2 ส่วน โดยมีรายละเอียดดังนี้

ส่วนที่ 1 สรุปผลการเปรียบเทียบเพื่อหาอัลกอริทึมการตัดคำที่ดีที่สุดสำหรับข้อมูลชุดนี้ โดยพบว่าอัลกอริทึม Attacut มีค่าความถูกต้องเท่ากับ 94.40% และใช้เวลาในการประมวลผลเฉลี่ย 1.6313 วินาที ดังนั้นผู้วิจัยจึงเลือกใช้ อัลกอริทึม Attacut ในการตัดคำ มาใช้ในการสร้างแบบจำลองการเปรียบเทียบ เพื่อหาอัลกอริทึมการจำแนกข้อความต่อไป

ส่วนที่ 2 สรุปผลการเปรียบเทียบเพื่อหาอัลกอริทึมการจำแนกข้อความที่ดีที่สุดสำหรับข้อมูลชุดนี้ โดยพบว่า อัลกอริทึม BLSTM (Bidirectional Long Short-Term Memory) มีประสิทธิภาพในการจำแนกสูงที่สุด โดยให้ค่าความถูกต้องเท่ากับ 75.44% ค่าความแม่นยำเท่ากับ 79.08% ค่าระลอกเท่ากับ 75.44% และค่าถ่วงดุลเท่ากับ 76.27% ซึ่งได้แบบจำลองที่เหมาะสมกับงานวิจัยมากที่สุด นำไปใช้ในการพัฒนาระบบหุ่นยนต์สนทนาบนระบบไลน์ เพื่อเพิ่มประสิทธิภาพการทำงานให้แก่องค์กร

เอกสารอ้างอิง

[1] ฐิณากัณท์ และทัศนัท ตรีนนทรัตน์. (2020). การพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทย เพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษา ระดับอุดมศึกษา, (The Development Of Thai Artificial Intelligence Chatbot For Supporting Academic Consultancy For Tertiary Students), (วิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์, วันที่ 29 สิงหาคม ปี ค.ศ.2020, วันที่ 9 ตุลาคม ปี ค.ศ.2020, วันที่ 24 ตุลาคม ปี ค.ศ. 2020)

[2] สำนักงานศูนย์วิจัย และให้คำปรึกษาแห่งมหาวิทยาลัยธรรมศาสตร์, โครงการศึกษาวิจัย เพื่อพัฒนาปัญญาประดิษฐ์. (2560). สำหรับค้นหาข้อ

- กฎหมายที่เกี่ยวข้องกับองค์กรปกครองส่วนท้องถิ่น, (สำนักงานศูนย์วิจัย และให้คำปรึกษาแห่งมหาวิทยาลัยธรรมศาสตร์)
- [3] สุณิสา ศรแก้ว, การยอมรับเทคโนโลยี Chatbot ในธุรกิจธนาคารในประเทศไทย. (2561). (Chatbot Technology Adoption In Banking Business In Thailand), (พาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์)
- [4] ธนัญญา พิรพัฒนาการ. (2562). วิธีการสำหรับการสร้างหุ่นยนต์สนทนาไทย โดยใช้หน่วยความจำระยะสั้นแบบยาวแบบสยัม และการแต่งเติมข้อมูลเชิงข้อความ. วิทยานิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
- [5] เฉลิมฤทธิ์ เฉลิมศิริกุล และณัฐวี อุดกฤษฎ์. (2562). การประยุกต์ใช้หุ่นยนต์สนทนาบนระบบ Line เพื่อลดภาระงาน Help Desk ในการให้บริการด้านสารสนเทศการใช้งานระบบบริหารจัดการงานเอกสาร. The 15th National Conference and 2019-1 International Conference on Applied Computer Technology and Information Systems and 2019-1 National Conference on Business Administration. 12-17.
- [6] แทนพงศ์ วัฒนธาดา และมหศักดิ์ เกตุจำเริญ. (2562). การพัฒนาหุ่นยนต์สนทนาบนแอปพลิเคชันไลน์ เพื่อสนับสนุนการทำงานของแผนกทรัพยากรบุคคลแบบอิเล็กทรอนิกส์. The 15th National Conference and 2019-1 International Conference on Applied Computer Technology and Information Systems and 2019-1 National Conference on Business Administration. 278-282.
- [7] ธนัท วรธกะวิกรานต์. (2561). การพัฒนาแบบจำลองการวิเคราะห์ความคิดเห็นภาษาไทยที่มีต่อโรงแรมในประเทศไทย. ปัญหาพิเศษ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาระบบสารสนเทศเพื่อ
- การจัดการ ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- [8] Pattarawat Chormai, Ponrawee Prasertsom, and Attapol Rutherford. (2019). AttaCut: A Fast and Accurate Neural Thai Word Segmenter. New in ML Workshop, NeurIPS 2019.
- [9] Lei Cui, Shaohan Huang, Furu Wei, Chuanqi Tan, Chaoqun Duan, and Ming Zhou. (2017). SuperAgent: A Customer Service Chatbot for E-commerce Websites. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics-System Demonstrations. 97-102
- [10] Pengfei Zhu, Zhuosheng Zhang, Jiangtong Li, Yafang Huang, Hai. (2018). Lingke : A Fine-grained Multi-turn Chatbot for Customer Service. School of Computer Science and Software Engineering, East China Normal University.

ระบบวิเคราะห์ท่อประปาแตกรั่วโดยใช้กฎความสัมพันธ์ และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์

An Analytic System for Pipe Breakages Using Association Rules and Web Map Services

กัณตนา พูลทอง (Kantana Pooltong)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407011858011@email.kmutnb.ac.th, ²pudsadee.b@itd.kmutnb.ac.th

บทคัดย่อ

เทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ (Web Map Services) เป็นการนำระบบสารสนเทศภูมิศาสตร์ (Geographic Information Systems: GIS) และระบบกำหนดตำแหน่งบนโลก (Global Positioning System: GPS) มาใช้ในการพัฒนาและนำเสนอข้อมูล ซึ่งสามารถประยุกต์ใช้งานได้หลากหลายด้าน เช่น การแสดงแผนที่การท่อก๊าซและการขนส่ง เป็นต้น วัตถุประสงค์ของงานวิจัยนี้ จึงได้พัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วโดยประยุกต์ใช้กฎความสัมพันธ์ (Association Rules) เพื่อค้นหารูปแบบความสัมพันธ์ของข้อมูลที่เกิดขึ้น และสีฟเลท แมพ เอพีไอ ในการแสดงแผนที่โครงข่ายท่อประปา ตำแหน่งของโครงข่าย และเส้นแนวการติดตั้งวางท่อในรูปแบบของแผนที่ผ่านเว็บไซต์ จากผลการวิจัยพบว่า ระบบที่พัฒนาขึ้นสามารถแสดงข้อมูล และวิเคราะห์ท่อประปาแตกรั่วเพื่อนำไปใช้ในการวางแผนและบำรุงรักษาได้ตรงตามความต้องการของผู้ใช้งาน อีกทั้งเป็นแนวทางในการประยุกต์ใช้และนำไปต่อยอดในการพัฒนาระบบวิเคราะห์ความเสี่ยงด้านอื่นๆ ได้ในอนาคต

คำสำคัญ: การวิเคราะห์ข้อมูล ท่อแตกรั่ว กฎความสัมพันธ์ เทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ ระบบสารสนเทศภูมิศาสตร์

Abstract

Web Map Services technology uses Geographic Information Systems (GIS) and Global Positioning

Systems (GPS) to develop and present data. It can be applied in various ways, such as map representation, tourism, and transportation. The purpose of this research is to develop a pipe breakages analysis system by applying the Association Rules to find the relationship of the data to find the pattern and applying Leaflet Map API to display a map of the water supply network, network location, and installation lines laying pipes in the form of a map through the web site. From the research results, the developed system can display information and analyze breakage pipes to use for planning and can be maintained to meet users' needs. It is also a guideline for application and can be used to develop other risk analysis systems in the future.

Keywords: Data Analytic, Pipe Breakages, Association Rules, Web Map Services, Geographic Information Systems.

1. บทนำ

การประปาส่วนภูมิภาค (กปภ.) เป็นรัฐวิสาหกิจภายใต้การกำกับดูแลของกระทรวงมหาดไทย มุ่งเน้นดูแลประชาชนด้วยบริการสาธารณูปโภคขั้นพื้นฐานด้านน้ำประปาที่มีคุณภาพ สะอาดปลอดภัย และได้มาตรฐาน โดยเฝ้าติดตามและควบคุมคุณภาพน้ำในทุกขั้นตอนของกระบวนการผลิต ตามข้อเสนอแนะขององค์การอนามัยโลก

(World Health Organization : WHO) และส่งจ่ายน้ำประปาอย่างทั่วถึง และเพียงพอในพื้นที่ 74 จังหวัดทั่วประเทศ

กปน. มีโครงข่ายท่อจ่ายน้ำกระจายอยู่ทั่วทุกภาคของประเทศ มีระบบการผลิตน้ำประปาที่มีคุณภาพ แต่ยังมีอัตราน้ำสูญเสียที่สูงขึ้นอย่างต่อเนื่อง [1] ซึ่งส่งผลกระทบต่อทำให้บริการน้ำประปา โดยสาเหตุหลักของการเกิดน้ำสูญเสียทางกายภาพ ได้แก่ การเกิดท่อประปาแตกรั่ว ซึ่งมีสาเหตุมาจากหลายปัจจัย เช่น อายุการใช้งานของท่อแรงกดทับของถนน และอุบัติเหตุจากการขุดเจาะ เป็นต้น กลยุทธ์ในการลดน้ำสูญเสียในระยะยาว ได้แก่ การดูแลบำรุงรักษาระบบท่อให้สามารถใช้งานได้อย่างต่อเนื่อง มีแผนงานในการสำรวจติดตามเพื่อเปลี่ยนท่อที่เสื่อมสภาพ และมีการวางแผนจัดลำดับความสำคัญของการซ่อมหรือเปลี่ยนท่อใหม่ เป็นต้น

จากปัญหาดังกล่าวงานวิจัยนี้จึงได้พัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วโดยใช้กฎความสัมพันธ์และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ ซึ่งประยุกต์ใช้ลิฟเลทแมพเอพีไอ ในการแสดงแผนที่โครงข่ายท่อประปา ตำแหน่งของโครงข่ายท่อประปาตามประเภท และเส้นแนวการติดตั้งวางท่อในรูปแบบของแผนที่ผ่านเว็บไซต์ โดยในบทความวิจัยแบ่งเนื้อหาออกได้เป็น 5 ส่วน ได้แก่ ส่วนที่ 1 บทนำ ส่วนที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง ส่วนที่ 3 วิธีการดำเนินการวิจัย ส่วนที่ 4 ผลการดำเนินการวิจัย และส่วนที่ 5 สรุปผลการวิจัย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การวิเคราะห์ข้อมูล และกฎความสัมพันธ์

การวิเคราะห์ข้อมูลช่วยให้ธุรกิจสามารถบรรลุวัตถุประสงค์ที่ตั้งไว้ โดยข้อมูลที่นำมาใช้อาจมีทั้งข้อมูลในอดีต ปัจจุบัน รวมไปถึงการพยากรณ์ข้อมูลในอนาคต เพื่อใช้ในการกำหนดกลยุทธ์ทางธุรกิจ ซึ่งข้อมูลเหล่านี้ถือว่ามีความสำคัญในการสร้างศักยภาพให้กับธุรกิจ โดยรูปแบบของการวิเคราะห์ข้อมูล (Data Analytics) มีรายละเอียดดังนี้

การวิเคราะห์ข้อมูลแบบพื้นฐาน (Descriptive Analytics) คือการวิเคราะห์ข้อมูลเพื่อแสดงผลที่เกิดขึ้น

หรือกำลังจะเกิดขึ้น จากข้อมูลในอดีต ในลักษณะที่เข้าใจง่าย เช่น รายงาน แผนภูมิ กราฟ และตาราง เป็นต้น

การวิเคราะห์แบบเชิงวินิจฉัย (Diagnostic Analytics) เป็นการอธิบายถึงสาเหตุของสิ่งที่เกิดขึ้น ปัจจัย และความสัมพันธ์ของปัจจัย หรือตัวแปรต่าง ๆ ที่มีความสัมพันธ์กัน ตัวอย่างเช่น ความสัมพันธ์ระหว่างยอดขายต่อกิจกรรมทางการตลาดแต่ละประเภท และการวิเคราะห์การทำงานของเครื่องคอมพิวเตอร์เพื่อตรวจจับความคิดปกติ เป็นต้น

การวิเคราะห์แบบพยากรณ์ (Predictive Analytics) เป็นการวิเคราะห์เพื่อพยากรณ์สิ่งที่กำลังจะเกิดขึ้น หรือน่าจะเกิดขึ้น โดยใช้ข้อมูลที่ได้เกิดขึ้นแล้วกับแบบจำลองทางสถิติ หรือเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) ตัวอย่างเช่น การพยากรณ์ยอดขาย และการพยากรณ์ผลประชามติ เป็นต้น

การวิเคราะห์แบบให้คำแนะนำ (Prescriptive Analytics) เป็นการวิเคราะห์ข้อมูลที่มีความซับซ้อนที่สุด เนื่องจากการพยากรณ์สิ่งที่เกิดขึ้น ข้อดี ข้อเสีย สาเหตุ และระยะเวลาของสิ่งที่เกิดขึ้น รวมถึงการให้คำแนะนำทางเลือกต่าง ๆ ที่มีอยู่และผลของแต่ละทางเลือกพร้อมด้วย

ในงานวิจัยประยุกต์ใช้กฎความสัมพันธ์ (Association Rules) ในการพัฒนาระบบสารสนเทศ เนื่องจากนำมาใช้ค้นหาความสัมพันธ์ของข้อมูลที่มีอยู่เพื่อนำไปหารูปแบบที่เกิดขึ้น บ่อยครั้ง (Frequent Pattern) และใช้ในการวิเคราะห์ความสัมพันธ์เพื่อทำนายปรากฏการณ์ต่าง ๆ [2] ข้อมูลที่ใช้ในการทำเหมืองความสัมพันธ์ (Association Mining) มักเป็นข้อมูลประเภท Transaction ผลลัพธ์ที่ได้จะอยู่ในรูปของกฎที่แสดงความสัมพันธ์ระหว่างข้อมูลในรูปแบบเหตุและผล กฎความสัมพันธ์สามารถประยุกต์ใช้ในงานที่หลากหลาย เช่น การวิเคราะห์ความสัมพันธ์ของสินค้าจากพฤติกรรมการซื้อของลูกค้า [3] และการวิเคราะห์ความสัมพันธ์ของเวชภัณฑ์ในการรักษาผู้ป่วย เป็นต้น

2.2 เทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์

เทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ (Web Map Services: WMS) เป็นระบบให้บริการแผนที่ของข้อมูลภูมิสารสนเทศผ่านเครือข่ายอินเทอร์เน็ต ซึ่งมีการกำหนดมาตรฐานโดยองค์กรความร่วมมือข้อมูลภูมิศาสตร์ระบบเปิด

(Open Geospatial Consortium: OGC) ที่ได้มีการกำหนดมาตรฐานการให้บริการข้อมูลภูมิสารสนเทศในรูปแบบต่างๆ สามารถนำมาใช้เพื่ออำนวยความสะดวกในชีวิตประจำวันของผู้คน สนับสนุนการวิเคราะห์ข้อมูลการจัดการพื้นที่ และการจำแนกประเภทของแผนที่ [4] ซึ่งช่วยให้ผู้ใช้สามารถค้นหาแผนที่ได้อย่างมีประสิทธิภาพ

2.3 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาการสร้างแบบจำลองภูมิความสัมพันธ์สำหรับฐานข้อมูลการตั้งชื่อสินค้าโดยใช้เทคนิคเอฟพี-โกรธ (FP - Growth) สร้างแบบจำลองความสัมพันธ์ของการตั้งชื่อสินค้าโดยใช้ภูมิความสัมพันธ์ โดยวิเคราะห์จากการตั้งชื่อสินค้าของลูกค้าในแต่ละราย [5] ซึ่งแยกความต้องการของลูกค้าออกมาเป็นภูมิความสัมพันธ์ โดยแต่ละรูปแบบของภูมิความสัมพันธ์จะแตกต่างกันเนื่องจากการกำหนดค่าพารามิเตอร์ Support และค่าพารามิเตอร์ Confidence ทำให้ได้จำนวนภูมิความสัมพันธ์แตกต่างกันซึ่งสามารถนำภูมิความสัมพันธ์ที่ได้ไปใช้ในการสนับสนุนการตัดสินใจนำเสนอสินค้าได้ตรงตามความต้องการของลูกค้าในแต่ละราย [5] ช่วยสร้างความพึงพอใจให้กับลูกค้าซึ่งถือเป็นกลยุทธ์หนึ่งของการวางแผนการตลาดได้อย่างเหมาะสมและมีประสิทธิภาพ

จากการศึกษาการพัฒนากระบวนสนับสนุนการตัดสินใจเลือกที่ตั้งร้านสะดวกซื้อ จากการวิเคราะห์ความน่าจะเป็นในการใช้บริการของลูกค้า ได้ประยุกต์ใช้โปรแกรมการจัดการฐานข้อมูล PostgreSQL/PostGIS ซึ่งมีความสามารถในการรองรับและบริหารจัดการภูมิสารสนเทศ [6] อีกทั้งยังมีรูปแบบการใช้งานตรงตามมาตรฐาน ISO 19125 และ Open Geospatial Consortium (OGC) ร่วมกับภาษา PHP, JavaScript, AJAX และ HTML มีการประยุกต์ร่วมกับสไลด์แมพ เอพีไอในการแสดงผลแผนที่ออนไลน์และแผนที่ผลลัพธ์ด้วย Leaflet Heat Map ผลการวิจัยพบว่า ระบบสามารถประมวลผลข้อมูลได้ถูกต้อง และสามารถแสดงผลลัพธ์ค่าความน่าจะเป็นในรูปแบบของแผนที่ Heat Map ได้เป็นอย่างดี

3. วิธีการดำเนินการวิจัย

วิธีการดำเนินการวิจัยในการพัฒนาระบบวิเคราะห์

ต่อประปาแตกรั่วโดยใช้ภูมิความสัมพันธ์ และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ สามารถแบ่งออกได้เป็น 4 ขั้นตอนหลักดังนี้

3.1 การศึกษาปัญหาและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาปัญหาที่เกี่ยวข้องกับการจัดการน้ำเสียทางกายภาพซึ่งเกี่ยวข้องกับการแตกรั่วของท่อประปาในปัจจุบันพบว่าการแตกรั่วมีสาเหตุมาจากหลายปัจจัย เช่น อายุการใช้งานของท่อแรงกดทับของถนนและอุบัติเหตุจากการขุดเจาะเป็นต้น จากปัญหาดังกล่าววัตถุประสงค์ของงานวิจัยนี้จึงได้พัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วโดยใช้ภูมิความสัมพันธ์และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ เพื่อช่วยในการแสดงผลข้อมูล และการวิเคราะห์ข้อมูลการแตกรั่วของท่อสำหรับนำไปช่วยในการวางแผน ตัดสินใจ และบำรุงรักษาท่อประปา

3.2 การออกแบบกรอบแนวคิดการวิจัย

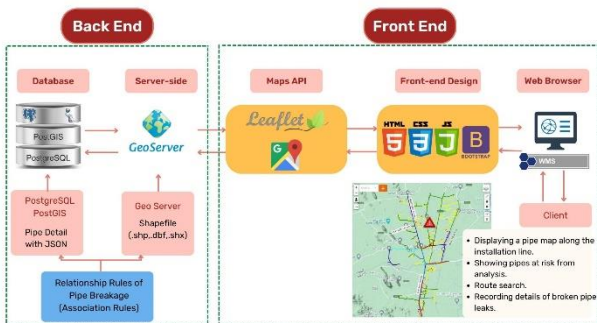
หลังจากการศึกษาปัญหาและงานวิจัยที่เกี่ยวข้อง คณะผู้วิจัยได้ทำการออกแบบกรอบแนวคิดการวิจัยของการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่ว สามารถแสดงได้ดังภาพที่ 1 ซึ่งแบ่งออกได้เป็น 2 ส่วนหลัก ได้แก่ ส่วนของการจัดการฐานข้อมูลและการประมวลผล (Back End) และส่วนต่อประสานผู้ใช้ (Front End) มีรายละเอียดดังนี้

1) ส่วนของการจัดการฐานข้อมูลและการประมวลผล

ฐานข้อมูล (Database) ที่ใช้ในการจัดเก็บข้อมูลได้แก่ โพสต์เกรสคิวแอล และทำการเพิ่มส่วนขยายโพสต์จิส (Post GIS) ที่ทำให้ฐานข้อมูล โพสต์เกรสคิวแอลสามารถรองรับข้อมูลด้านสารสนเทศภูมิศาสตร์ได้ ซึ่งทำงานร่วมกับ Geo Server ที่เป็นเซิร์ฟเวอร์สำหรับการแบ่งปันข้อมูลเชิงพื้นที่ที่มีการออกแบบเพื่อทำงานร่วมกับการเผยแพร่ข้อมูลจากแหล่งข้อมูลเชิงพื้นที่และการให้บริการข้อมูลด้านภูมิสารสนเทศในรูปแบบออนไลน์ ในส่วนการประมวลผลข้อมูลได้นำภูมิความสัมพันธ์มาประยุกต์ใช้งานเพื่อทำการวิเคราะห์ข้อมูลทางกายภาพของท่อประปาที่เคยแตกรั่วในอดีต ได้แก่ ประเภท ชั้นคุณภาพ ขนาด ความยาว ลักษณะการวางท่อ อายุการใช้งาน และหน้าที่ของท่อ มาประมวลผลเพื่อนำเสนอข้อมูลความเสี่ยงการแตกรั่วของท่อประปาบนแผนที่ออนไลน์

2) ส่วนต่อประสานผู้ใช้

ในส่วนการแสดงผลข้อมูลและแผนที่ได้ประยุกต์ใช้ ลีฟเลท แมพ เอพีไอ สำหรับแสดงแผนที่โครงข่ายแนวท่อ จากตำแหน่งการติดตั้งวางท่อ และเกิด แมพ เอพีไอ ในการนำทางเพื่อสำรวจติดตามท่อประปาแตกรั่วจาก ตำแหน่งปัจจุบันของผู้ใช้ไปยังตำแหน่งของท่อที่ต้องการสำรวจ โดยมีฟังก์ชันการทำงานของระบบ ได้แก่ การแสดงแผนที่พื้นที่ให้บริการน้ำประปาของ กปภ. การแสดง แผนที่แนวท่อทั้งหมดของ กปภ. พร้อมรายละเอียดของท่อการแสดงผลการวิเคราะห์จุดซ่อมท่อในรูปแบบ Heat Map และสามารถบันทึกผลการสำรวจท่อประปาแตกรั่วได้



ภาพที่ 1: กรอบแนวคิดการวิจัยการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่ว

3.3 การออกแบบและพัฒนาระบบสารสนเทศ

ในการออกแบบและพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วในรูปแบบเว็บแอปพลิเคชัน (Web Application) ซึ่งประยุกต์ใช้กฎความสัมพันธ์และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ โดยมีการนำข้อมูลประวัติการแตกรั่วของท่อประปา จำนวน 106,244 รายการ มาสร้างกฎความสัมพันธ์การแตกรั่วของท่อ โดยกำหนดค่าความเชื่อมั่นขั้นต่ำ คือ 0.90 ซึ่งจะได้กฎที่สามารถนำไปเป็นต้นแบบประมวลผลเพื่อแสดงแผนที่ท่อจากรูปแบบความสัมพันธ์จำนวน 17 กฎ

ในส่วนการพัฒนาระบบใช้โปรแกรมภาษา PHP ร่วมกับลีฟเลท แมพ เอพีไอ และเกิด แมพ เอพีไอ ในการแสดงแผนที่ออนไลน์ผ่านเว็บไซต์ ส่วนของการบันทึกข้อมูลใช้ฐานข้อมูลโพสต์เกรสคิวแอล ในการจัดการฐานข้อมูลเชิงสัมพันธ์ และโพสต์จิสในการจัดการข้อมูลเชิงพื้นที่ร่วมกับ GeoServer เพื่อให้บริการแผนที่ออนไลน์ในรูปแบบ WMS

จากการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วประกอบด้วย 5 ฟังก์ชันการทำงานหลัก ได้แก่ การแสดงแผนที่พื้นที่ให้บริการน้ำประปาของ กปภ. การแสดงแผนที่แนวท่อทั้งหมดของ กปภ. พร้อมรายละเอียดของท่อ การแสดงผลการวิเคราะห์จุดซ่อมท่อในรูปแบบ Heat Map การนำทางเพื่อสำรวจท่อประปาแตกรั่ว และสามารถบันทึกผลการสำรวจท่อประปาแตกรั่วได้ ซึ่งตัวอย่างฟังก์ชันการแสดงผลจุดซ่อมท่อในรูปแบบ Heat Map ได้ดังภาพที่ 2

```

Get_WMS for Heatmap

var view_leakpipes = L.tileLayer.betterWms(wms_host, {
  maxZoom: 30,
  maxNativeZoom:30,
  layers: 'PG_WEBGIS:leak_pipeall',
  transparent: true,
  format: 'image/png',
  minZoom: 9
});
map.addLayer(view_leakpipes);
    
```

ภาพที่ 2: ฟังก์ชันการแสดงผลจุดซ่อมท่อในรูปแบบ Heat Map

3.4 การทดสอบคุณภาพของระบบสารสนเทศ

ในการวิจัยนี้ได้ทำการทดสอบคุณภาพของระบบวิเคราะห์ท่อประปาแตกรั่วโดยใช้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ที่มีความรู้และประสบการณ์ทางด้านคอมพิวเตอร์ ระบบภูมิสารสนเทศและการวิเคราะห์และประมวลผลข้อมูล ซึ่งทำการประเมินคุณภาพจาก 4 ด้าน ได้แก่ ด้านความสามารถทำงานได้ตามหน้าที่ (Function Requirement Test) ด้านความถูกต้องของความสามารถทำงานได้ตามหน้าที่ (Functional Test) ด้านความง่ายต่อการใช้งาน (UX/UI Test) และด้านการรักษาความปลอดภัยของข้อมูล (Security Test)

4. ผลการดำเนินการวิจัย

4.1 ผลการวิเคราะห์ความสัมพันธ์การแตกรั่วของท่อประปา

จากผลการวิจัยที่ได้วิเคราะห์ข้อมูลเพื่อหารูปแบบความสัมพันธ์การแตกรั่วของท่อประปา พบว่า จากการกำหนดค่าความเชื่อมั่นขั้นต่ำที่ 0.90 จะได้กฎที่สามารถนำไปเป็นต้นแบบสำหรับการพัฒนาระบบสารสนเทศ จำนวน 17 กฎ ซึ่งสามารถแสดงตัวอย่างกฎความสัมพันธ์ 5 ลำดับแรกได้ดังตารางที่ 1 โดยกฎความสัมพันธ์ที่ส่งผลต่อการแตกรั่วของ

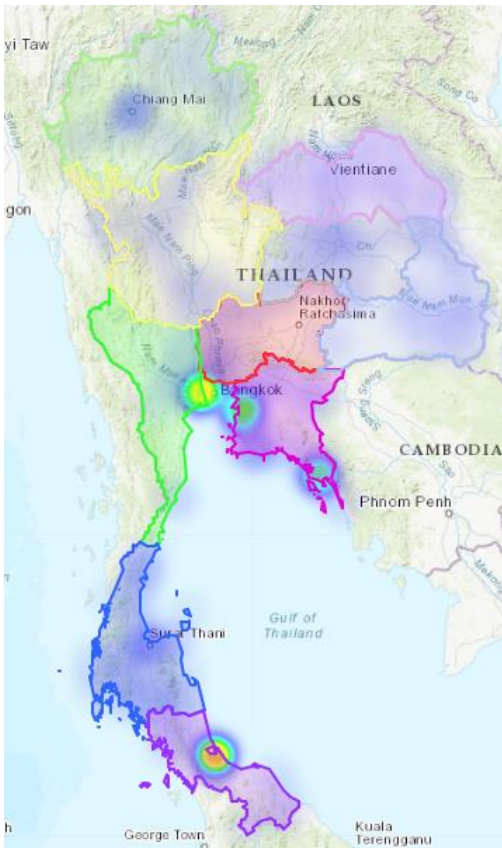
ท่อมากที่สุด ได้แก่ ท่อโพลีบิวทีลีน (Polybutylene : PB) ที่เป็นประเภทท่อจ่าย มีความสัมพันธ์กับท่อขนาด 200 มิลลิเมตร ที่ความเชื่อมั่น 100% และค่าสนับสนุนที่ 45%

ตารางที่ 1: ตัวอย่างกฎความสัมพันธ์ 5 อันดับแรก

No.	Premises	Conclusion	Support	Confidence	Lift
1	pipe_func2, pipe_type3	pipe_size0_200	0.45	1.00	1.118
2	pipe_type3	pipe_size0_200	0.43	1.00	1.118
3	pipe_size0_200, year21up	pipe_func2	0.32	0.98	1.035
4	pipe_size0_200	pipe_func2	0.87	0.97	1.029
5	year21up	pipe_func2	0.34	0.96	1.015

4.2 ผลการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่ว

ผลการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่ว สามารถแสดงตัวอย่างหน้าแรกของระบบ ซึ่งมีการแสดงแผนที่การวิเคราะห์จุดซ่อมท่อด้วย Heat Map ดังภาพที่ 3



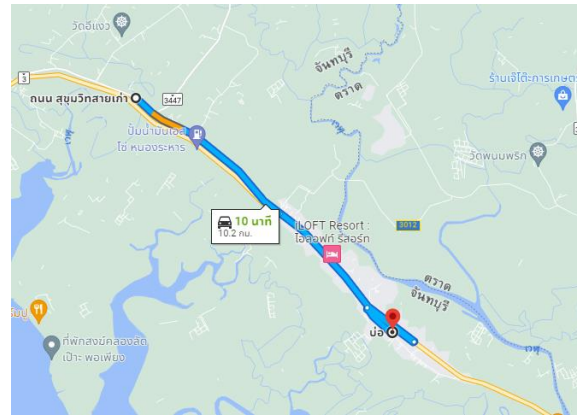
ภาพที่ 3: หน้าจอแสดงแผนที่การวิเคราะห์จุดซ่อมท่อด้วย Heat Map

ตัวอย่างหน้าจอในส่วนของการแสดงรายละเอียดท่อประปา และการแนะนำความเสี่ยงของท่อ ซึ่งสามารถแสดงได้ดังภาพที่ 4

รายละเอียดท่อประปา รหัส: 1993			
แจ้งเตือน	ท่ออายุมากกว่า 10 ปี!! เคยแตกรั่วแล้ว 2 ครั้ง ซ่อมบำรุงครั้งล่าสุด 2022-02-15		
รหัสท่อประปา	1993	ชนิด	PVC
ขนาด	100 มม.	หน้าที่	ท่อจ่ายน้ำ
ความลึก	1 ม.	ลักษณะการวาง	วางท่อใต้ดิน
ชั้นมาตรฐาน	8.5	ความยาว	6,491 เมตร
เลขที่สัญญาโครงการ	001/2500	ปีที่ติดตั้ง	2549
รหัสทรัพย์สิน	-	บริเวณที่วางท่อ	ถ.สุขุมวิท ม.9- ม.5- ม.7 ต.ป้อม
ชื่อผลิตภัณฑ์	มีดสแตนเลส	หมายเหตุ	ท่อแตกทางหลวง
วันที่บันทึกข้อมูล	511230	ปรับปรุงข้อมูลล่าสุดเมื่อ	2023-02-08T02:44:06Z

ภาพที่ 4: หน้าจอการแสดงรายละเอียดและแนะนำเกี่ยวกับท่อ

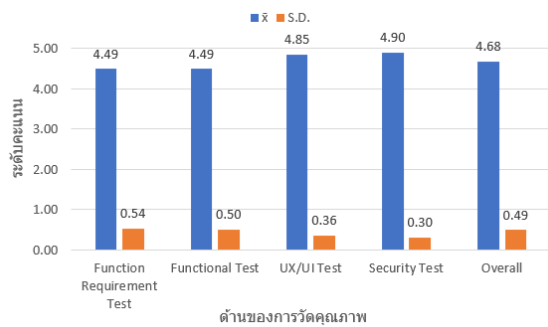
ตัวอย่างหน้าจอในส่วนของการนำทางเพื่อสำรวจท่อประปาแตกรั่วสามารถแสดงได้ดังภาพที่ 5



ภาพที่ 5: หน้าจอการนำทางเพื่อสำรวจท่อประปาแตกรั่ว

4.3 ผลการประเมินคุณภาพของระบบสารสนเทศ

การประเมินคุณภาพของระบบวิเคราะห์ท่อประปาแตกรั่วโดยผู้เชี่ยวชาญ พบว่า คุณภาพของระบบโดยภาพรวมอยู่ในระดับดีมาก ($\bar{x} = 4.68$ และ $S.D. = 0.49$) เมื่อพิจารณาทางด้าน ด้านการรักษาความปลอดภัยของข้อมูลอยู่ในระดับดีมาก ซึ่งมีค่าเฉลี่ยมากที่สุด ($\bar{x} = 4.90$ และ $S.D. = 0.30$) รองลงมา ได้แก่ ด้านความง่ายต่อการใช้งานอยู่ในระดับดีมาก ($\bar{x} = 4.85$ และ $S.D. = 0.36$) ด้านความถูกต้องของความสามารถทำงานได้ตามหน้าที่อยู่ในระดับดี ($\bar{x} = 4.49$ และ $S.D. = 0.50$) และด้านความสามารถทำงานได้ตามหน้าที่อยู่ในระดับดี ดี ($\bar{x} = 4.49$ และ $S.D. = 0.54$) ซึ่งสามารถแสดงผลการประเมินคุณภาพในรูปแบบกราฟแท่งได้ดังภาพที่ 6



ภาพที่ 6: กราฟแท่งแสดงผลการประเมินคุณภาพระบบของผู้เชี่ยวชาญ

5. สรุปผลการวิจัย

จากการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วโดยใช้กฎความสัมพันธ์และเทคโนโลยีการให้บริการแผนที่ผ่านเว็บไซต์ ซึ่งมีการประยุกต์ใช้ลีฟเลท แมพ เอพีไอในการแสดงแผนที่โครงข่ายท่อประปา ตำแหน่งของท่อประปาและแนวการติดตั้งวางท่อในรูปแบบของแผนที่ผ่านเว็บไซต์พบว่า กฎความสัมพันธ์ที่ได้จากการวิเคราะห์ข้อมูลประวัติการแตกรั่วของท่อประปาที่กำหนดค่าความเชื่อมั่นขั้นต่ำ 0.90 ได้กฎที่สามารถนำไปใช้สำหรับการพัฒนาระบบเพื่อแสดงแผนที่จากรูปแบบความสัมพันธ์ จำนวน 17 กฎ โดยพบว่ารูปแบบความสัมพันธ์ที่ส่งผลต่อการแตกรั่วของท่อประปามากที่สุด ได้แก่ ท่อโพลีบิวทีลีน ที่เป็นประเภทท่อจ่าย มีความสัมพันธ์กับท่อขนาด 200 มิลลิเมตร อยู่ที่ความเชื่อมั่น 100% และค่าสนับสนุนที่ 45%

จากผลการพัฒนาระบบวิเคราะห์ท่อประปาแตกรั่วสามารถรองรับฟังก์ชันการทำงานต่าง ๆ ได้แก่ การแสดงแผนที่พื้นที่ให้บริการน้ำประปาของ กปภ. การแสดงแผนที่แนวท่อทั้งหมดของ กปภ. พร้อมรายละเอียดของการแสดงผลการวิเคราะห์จุดซ่อมท่อในรูปแบบ Heat Map การนำทางเพื่อสำรวจท่อประปาแตกรั่ว และสามารถบันทึกผลการสำรวจท่อประปาแตกรั่วได้ จากผลการทดสอบคุณภาพของระบบโดยผู้เชี่ยวชาญ พบว่าระบบวิเคราะห์ท่อประปาแตกรั่วมีคุณภาพโดยรวมอยู่ในระดับดีมาก ซึ่งสามารถแสดงข้อมูลต่าง ๆ ผ่านแผนที่เว็บไซต์สำหรับช่วยในการวิเคราะห์ข้อมูลการแตกรั่วของท่อได้ตรงตามความต้องการของผู้ใช้งาน อีกทั้งเป็นแนวทางในการประยุกต์ใช้งานบริการแผนที่ผ่านเว็บไซต์ในการแสดงผล

แผนที่ด้านการขนส่งทางท่องเที่ยวหรืองานวิจัยอื่น ๆ ที่เกี่ยวข้อง และสามารถนำไปต่อยอดพัฒนาระบบวิเคราะห์ความเสี่ยงด้านอื่น ๆ ได้ในอนาคต

เอกสารอ้างอิง

- [1] J. Thee and J. Chatchai, "Reducing Water Loss in a Water Supply System Using a District Metering Area (DMA) : A Case Study of The Provincial Waterworks Authority (PWA), Lop Buri Branch," *Engineering and Applied Science Research*, pp. 154 – 160, 2017.
- [2] P. Paruechanon and W. Sriurai, "Applying Association Rule to Risk Analysis for Dropout Students of Information Technology Department," *Journal of Science and Science Education*, pp. 123 – 133, 2018.
- [3] Z. Xitang, Z. Lei, Y. Quanqing, and Z. Zhe, "A New Association Rule Mining Method and its Application in Mental Health Data Analysis," *Global Conference on Robotics, Artificial Intelligence and Information Technology (GCRAIT)*, 2022.
- [4] Y. Zelong, G. Zhipeng, W. Huayi and Li. Wenwen, "A Latent Feature-Based Multimodality Fusion Method for Theme Classification on Web Map Service," *IEEE Access*, Vol.8, pp. 25299 – 25309, 2020.
- [5] K. Chugh and N. Kantanatha, "Improving a Recommendation Engine for Traditional Trade Between Wholesalers and Retailers Using Association Rules," *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2022.
- [6] C. Po-Han and N. Thanh Chuong, "A BIM-WMS Integrated Decision Support Tool for Supply Chain Management" *Construction. in Automation in Construction*, Vol.98, pp. 289 – 301, 2019.

โปรแกรมความภักดีลูกค้าโดยใช้แอปพลิเคชันแบบกระจายศูนย์ Customer Loyalty Programs Using Decentralized Application

ประทีน สนนิกอร์ (Pratin Sonnekorn)¹ และพุดสดี บุญรอด (Pudsadee Boonrawd)²

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407011858071@kmutmb.ac.th, ²pudsadee.b@itd.kmutmb.ac.th

บทคัดย่อ

ปัจจุบันภาคธุรกิจมีการนำเทคโนโลยีเข้ามาเป็นส่วนสำคัญในการพัฒนาธุรกิจ เพื่อความถูกต้อง โปร่งใส ตรวจสอบได้ และเพิ่มความน่าเชื่อถือให้แก่การดำเนินธุรกิจ ซึ่งวัตถุประสงค์ของงานวิจัยนี้เพื่อพัฒนาโปรแกรมความภักดีลูกค้าโดยใช้แอปพลิเคชันแบบกระจายศูนย์ (Decentralized Application) ที่สร้างอยู่บนเครือข่ายบล็อกเชน (Blockchain) และมีรูปแบบการทำงานแบบ Peer-to-Peer (P2P) โปรแกรมที่พัฒนาขึ้นรองรับฟังก์ชันการทำงานต่าง ๆ ได้แก่ การเพิ่มคะแนนสะสมของลูกค้า การหักคะแนนสะสมของลูกค้า การแลกคะแนนสะสมของลูกค้า และการแสดงรายการประวัติการแลกคะแนนของลูกค้า จากผลการวิจัยพบว่าโปรแกรมที่พัฒนาขึ้นมีความถูกต้อง และรองรับฟังก์ชันการทำงานตามที่กำหนด อีกทั้งยังเป็นแนวทางการประยุกต์ใช้กับธุรกิจอื่น ๆ ที่เกี่ยวข้อง และสามารถนำไปต่อยอดพัฒนาโปรแกรมความภักดีลูกค้าโดยใช้แอปพลิเคชันแบบกระจายศูนย์ได้ในอนาคต

คำสำคัญ: โปรแกรมความภักดี เทคโนโลยีบล็อกเชน

การทำงานแบบ Peer-to-Peer สัญญาอัจฉริยะ

Abstract

Technology is crucial in developing businesses for accuracy, transparency, verifiability, and increased trustworthiness in business operations. This research aims to develop a customer loyalty program using a Decentralized Application (DApp) built on the Blockchain network and operates in a Peer-to-Peer (P2P) manner. The program developed supports

various functions such as adding customer loyalty points, deducting loyalty points for customers, exchanging loyalty points for customers, and displaying a history of customer point exchange transactions. The research findings indicate that the developed program is accurate and supports the specified functions. Moreover, it applied as a model for other related businesses. It can be used to develop customer loyalty programs further using DApps.

Keywords: Loyalty Program, Blockchain, Peer-to-Peer, Smart Contract.

1. บทนำ

จากการแข่งขันของภาคธุรกิจปัจจุบัน ผู้ประกอบการหลายรายมีการคิดค้น และพัฒนาธุรกิจอย่างต่อเนื่อง เพื่อให้ธุรกิจของตนสามารถแข่งขันกับธุรกิจคู่แข่งได้อย่างทัดเทียม ส่งผลให้ผู้ประกอบการเริ่มให้ความสำคัญกับการนำกลยุทธ์ทางการตลาดเข้ามาเป็นส่วนช่วยในการจูงใจลูกค้า ซึ่งหนึ่งในกลยุทธ์ทางการตลาดที่นักการตลาดให้การยอมรับว่ามีส่วนช่วยในการสร้างความสำเร็จให้ธุรกิจได้ คือ กลยุทธ์ในการสร้างความภักดี หรือ Loyalty Program นอกจากกลยุทธ์ทางการตลาดที่ดีแล้ว เครื่องมือในการนำกลยุทธ์มาใช้ประโยชน์นับเป็นสิ่งสำคัญในยุคดิจิทัลที่มีการพัฒนาความสามารถของเทคโนโลยีอย่างก้าวกระโดด การนำความสามารถของเทคโนโลยีมาช่วยในการดำเนินธุรกิจเป็นสิ่งสำคัญที่ช่วยให้ธุรกิจมีความโดดเด่นและดำเนินการได้อย่างมีประสิทธิภาพ และสร้างจุดขายได้มากขึ้น โดยหนึ่งในเทคโนโลยีที่มีความสำคัญ และน่าสนใจ

ได้แก่ เทคโนโลยีบล็อกเชน (Blockchain) ที่นำมาประยุกต์ใช้งานในหลากหลายภาคธุรกิจ

ในภาคการขนส่งมีการนำเทคโนโลยีบล็อกเชนมาช่วยในการคำนวณหาเส้นทางที่เร็วที่สุด หลีกเลี่ยงจุดก่อสร้าง และช่วยหาสถานที่จอดรถ สำหรับภาคการเกษตรและสิ่งแวดล้อมมีการนำเทคโนโลยีบล็อกเชนมาช่วยในการตรวจสอบและเก็บข้อมูลสภาพอากาศ ดิน ฟ้า และน้ำ นอกจากนี้เทคโนโลยีบล็อกเชนได้เข้ามาช่วยในการเก็บข้อมูล และเอกสารในรูปแบบของสินทรัพย์ดิจิทัล รวมถึงการแสดงความเป็นเจ้าของและการพิสูจน์ตัวตนอีกด้วย

จากความสำเร็จในการนำเทคโนโลยีบล็อกเชนเข้ามามีส่วนร่วมในงานด้านต่าง ๆ นำมาซึ่งการพัฒนาต่อ ยอด เป็น แอปพลิเคชันแบบกระจายศูนย์ (Decentralized Application: DApps) ซึ่งเป็นแอปพลิเคชันที่ไม่มีตัวกลางควบคุม ไม่ดำเนินการอยู่ภายใต้องค์กรหรือสถาบันใด สร้างอยู่บนเครือข่ายบล็อกเชน โดยใช้หลักการของ สัญญาอัจฉริยะ (Smart Contract) นอกจากนี้ยังเป็นการทำงานแบบ Peer-to-Peer (P2P) ทำให้มีความปลอดภัย และความถูกต้องสูง นับเป็นความสามารถที่ช่วยสร้างความน่าเชื่อถือ และความไว้วางใจแก่โปรแกรมความภักดี เนื่องจากโปรแกรมความภักดีนั้นต้องเก็บข้อมูลที่มีความเป็นส่วนตัว รวมไปถึงข้อมูลที่สามารถสะท้อนถึงพฤติกรรมผู้บริโภคสินค้าของลูกค้า จึงต้องการระบบในการจัดเก็บข้อมูลที่มีประสิทธิภาพ ปลอดภัย และมีความถูกต้องของข้อมูลสูง [1], [2]

จากเหตุผลดังกล่าววัตถุประสงค์ของการวิจัยนี้จึงได้พัฒนาโปรแกรมความภักดีลูกค้าด้วยเทคโนโลยีบล็อกเชน โดยประยุกต์ใช้แอปพลิเคชันแบบกระจายศูนย์ เพื่อให้ผู้ประกอบการร้านค้าสามารถนำไปปรับปรุง และพัฒนาเพื่อตอบสนองความต้องการของทั้งผู้ประกอบการและผู้บริโภคได้ตรงความต้องการ และมีประสิทธิภาพมากขึ้น

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โปรแกรมความภักดี (Loyalty Program)

โปรแกรมความภักดีจัดเป็นกิจกรรมด้านการตลาด

ที่ผู้ประกอบการธุรกิจนิยมนำมาใช้อย่างแพร่หลาย เนื่องจากสามารถลดต้นทุนในการดำเนินการทางธุรกิจ ช่วยให้ธุรกิจเป็นที่รู้จักและเป็นที่ยอมรับในวงกว้าง ช่วยสร้างความสัมพันธ์ที่ดีกับลูกค้า เพื่อให้ลูกค้าที่เคยใช้บริการกับร้านค้าเกิดความสนใจที่จะกลับมาใช้บริการอีกครั้ง หรือแนะนำคนรู้จักมาใช้บริการกับทางธุรกิจ นับเป็นส่วนหนึ่งของแนวทางในการจัดการความสัมพันธ์กับลูกค้าที่ธุรกิจให้ความสำคัญ

2.1.1 ประเภทของโปรแกรมความภักดี แบ่งออกได้ดังนี้ คือ ประเภทที่ 1 การสะสมคะแนน (Point-Based Loyalty) ลูกค้าสามารถสะสมคะแนนเพื่อแลกสิทธิพิเศษต่าง ๆ โดยเกณฑ์ในการรับคะแนนจะแตกต่างกันออกไปในแต่ละธุรกิจ คะแนนที่นำมาสะสมอาจมาจากการซื้อสินค้าและบริการ การทำกิจกรรม หรือการมีส่วนร่วมร่วมกับธุรกิจ ประเภทที่ 2 โปรแกรมคืนเงิน (Cashback) ลูกค้าจะได้รับเงินคืนเป็นเปอร์เซ็นต์ของยอดการใช้จ่าย และสามารถสะสมเพื่อใช้เป็นส่วนลดในครั้งต่อไปได้ ประเภทที่ 3 การแบ่งระดับสมาชิก (Tiered Loyalty) ที่แบ่งลูกค้าเป็นหมวดหมู่ตามปริมาณการใช้จ่ายที่ร้านค้า และประเภทที่ 4 การจ่ายเงินค่าสมาชิก (Paid Loyalty) ลูกค้าที่ทำการจ่ายเงินเข้าร่วมเป็นสมาชิกจะได้รับสิทธิประโยชน์อย่างทันทีและต่อเนื่อง โดยค่าธรรมเนียมการสมัครจะเรียกเก็บซ้ำตามระยะเวลาที่กำหนด เช่น ปีละ 1 ครั้ง เดือนละ 1 ครั้ง หรือเรียกเก็บเพียงแค่ครั้งเดียวตอนทำการสมัคร เป็นต้น

2.2 เทคโนโลยีบล็อกเชน (Blockchain)

กลไกฐานข้อมูลขั้นสูงที่ประมวลผล และจัดเก็บข้อมูลแบบกระจายศูนย์ หรือที่เรียกว่า Distributed Ledger Technology (DLT) มีรูปแบบการบันทึกข้อมูลโดยใช้วิทยาการเข้ารหัสลับ (Cryptography) ร่วมกับกลไกฉันทามติ (Consensus) โดยฐานข้อมูลบล็อกเชนจะจัดเก็บข้อมูลในบล็อกที่เชื่อมโยงกันเป็นลูกโซ่ โดยข้อมูลดังกล่าวจะมีความสอดคล้องกันตามลำดับเวลา เนื่องจากไม่สามารถลบหรือแก้ไขลูกโซ่หรือข้อมูลได้ ทำให้ข้อมูลมีความถูกต้อง และมีความน่าเชื่อถือที่ทำให้เทคโนโลยีบล็อกเชน นิยมนำมาใช้เพื่อสร้างบัญชีแยก

ประเภทที่เปลี่ยนแปลงแก้ไขไม่ได้และไม่เปลี่ยนรูปแบบเพื่อติดตามคำสั่งซื้อ การชำระเงิน บัญชี และธุรกรรมอื่น ๆ นอกจากนี้เทคโนโลยีบล็อกเชนยังมีกลไกภายในที่ป้องกันการเพิ่มธุรกรรมที่ไม่ได้รับอนุญาตอีกด้วย

2.2.1 องค์ประกอบของเทคโนโลยีบล็อกเชนแบ่งออกได้เป็น 3 ส่วน ได้แก่ ส่วนที่ 1 การจัดเก็บข้อมูลแบบกระจายศูนย์ ฐานข้อมูลที่ใช้ร่วมกันในเครือข่ายบล็อกเชนที่จัดเก็บธุรกรรม เช่น โฟล์ที่ใช้ร่วมกันซึ่งทุกคนในทีมสามารถแก้ไขได้ โดยในการแก้ไขข้อความที่เข้าร่วมกันส่วนใหญ่ทุกคนที่มีสิทธิ์แก้ไขสามารถลบไฟล์ทั้งหมดได้ อย่างไรก็ตามเทคโนโลยีการจัดเก็บข้อมูลแบบกระจายศูนย์มีการกำหนดบุคคลที่มีสิทธิ์แก้ไข รวมถึงขั้นตอนในการแก้ไขได้โดยไม่สามารถลบรายการได้เมื่อได้รับการบันทึกแล้ว ส่วนที่ 2 สัญญาอัจฉริยะ ในหลายบริษัทมีการใช้สัญญาอัจฉริยะเพื่อจัดการสัญญาธุรกิจโดยไม่จำเป็นต้องให้บุคคลที่สามหรือคนกลางช่วยเหลือ เป็นโปรแกรมที่จัดเก็บไว้ในระบบบล็อกเชน สามารถดำเนินการโดยอัตโนมัติเมื่อตรงตามเงื่อนไขที่กำหนดไว้ล่วงหน้า โดยจะทำการตรวจสอบเงื่อนไข If-Then เพื่อให้ธุรกรรมเสร็จสมบูรณ์ เช่น บริษัทขนส่งสินค้าอาจประยุกต์ใช้สัญญาอัจฉริยะในการชำระเงินโดยอัตโนมัติเมื่อสินค้ามาถึงท่าเรือ และส่วนที่ 3 การเข้ารหัสคีย์สาธารณะเป็นคุณสมบัติด้านความปลอดภัย เพื่อระบุผู้เข้าร่วมในเครือข่ายบล็อกเชน โดยกลไกนี้ทำการสร้างคีย์สองชุดสำหรับสมาชิกเครือข่าย คีย์ชุดที่หนึ่งคือ คีย์สาธารณะที่ทุกคนในเครือข่ายใช้ร่วมกัน และคีย์ชุดที่สองเป็นคีย์ส่วนตัวที่ไม่ซ้ำกันสำหรับสมาชิกทุกคน ทั้งนี้คีย์ทั้งสองชุดจะทำงานร่วมกันเพื่อใช้งานข้อมูลในบัญชีแยกประเภท

2.3 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องทำให้เห็นถึงโอกาสการนำเทคโนโลยีดังกล่าวมาประยุกต์ใช้กับโปรแกรมความภักดีได้อย่างครอบคลุม เช่น บัตรสะสมคะแนนสามารถแทนที่ด้วยกระเป๋าเงินดิจิทัล คะแนนสะสมสามารถแทนที่ด้วย ERC-20 Token โปรโมชันแคมเปญสามารถพัฒนาให้อยู่ในรูปแบบของสัญญาอัจฉริยะได้

โดยข้อมูลทั้งหมดจะถูกบันทึกในเครือข่ายบล็อกเชนซึ่งช่วยลดค่าใช้จ่ายที่เกี่ยวข้องกับระบบ เช่น เครื่องคอมพิวเตอร์เพื่อประมวลผล และเก็บข้อมูล ซึ่งสามารถสรุปทฤษฎีและงานวิจัยที่เกี่ยวข้องได้ดังตารางที่ 1 [3]

ตารางที่1: สรุปทฤษฎีและงานวิจัยที่เกี่ยวข้อง

แนวคิดที่ประยุกต์ใช้	แนวคิดหรือวิธีการ	ข้อดีและข้อเสีย
บล็อกเชน	1. เพิ่มความโปร่งใส และตรวจสอบได้ขององค์กร 2. ลดโอกาสการทำทุจริตของเจ้าหน้าที่ 3. สามารถลงทะเบียนสมาชิกทางไกลได้	<u>ข้อดี</u> ข้อมูลมีความเป็นปัจจุบัน โปร่งใส ตรวจสอบได้ และลดต้นทุนในการทำงาน <u>ข้อเสีย</u> ต้นทุนในการวางโครงสร้างระบบสูง ไม่เหมาะกับธุรกรรมที่ต้องการการตรวจสอบและประมวลผลที่รวดเร็ว
สัญญาอัจฉริยะ [4] - [7]	1. ช่วยในการทำธุรกรรมทางไกล 2. ลดขั้นตอนในการทำธุรกรรม	<u>ข้อดี</u> เป็นการทำสัญญาระหว่างบุคคลกับบุคคล ไม่มีบุคคลที่ 3 มาเกี่ยวข้อง ประหยัดเวลา และลดความสับสนที่อาจเกิดขึ้นระหว่างการทำสัญญา <u>ข้อเสีย</u> การแก้ไขทำได้ยาก หากจำเป็นต้องแก้ไข และสามารถบังคับใช้สัญญาอัจฉริยะได้เฉพาะในกรณีที่มีความชัดเจนและตรงตัวเท่านั้น

3. วิธีดำเนินงานวิจัย

3.1 การศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาปัญหาที่เกี่ยวข้องกับโปรแกรมความภักดี พบว่า โปรแกรมความภักดีของร้านค้าในปัจจุบันมีการเก็บข้อมูลของลูกค้าเป็นจำนวนมาก ทำให้ร้านค้ามีค่าใช้จ่ายสูงในการเก็บรักษาข้อมูล การดำเนินการต่าง ๆ เกี่ยวกับข้อมูล ซึ่งการนำเทคโนโลยีบล็อกเชนเข้ามาประยุกต์ใช้กับโปรแกรมความภักดีจะช่วยเพิ่มจุดขายในสินค้าและบริการ โดยอาศัยคุณสมบัติ ความถูกต้อง และความเป็นอัตโนมัติที่ไม่ต้องมีคนกลาง

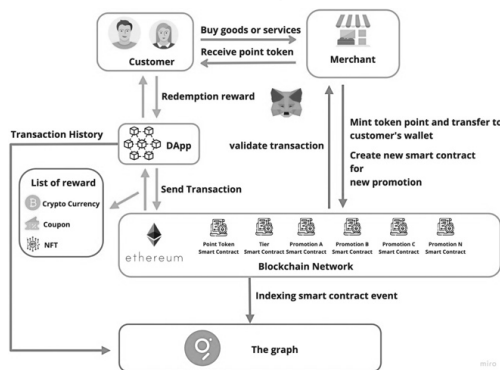
ในการให้บริการ โครงสร้างพื้นฐานของระบบ อีกทั้งการนำกระเป๋าเงินดิจิทัล (Digital Wallet) มาประยุกต์ใช้เพื่อบันทึกข้อมูล และทำงานผ่านระบบดิจิทัล ทั้งยังช่วยเพิ่มความมั่นคงปลอดภัย และทำให้การจัดการทำได้ง่าย

กระเป๋าเงินสินทรัพย์ดิจิทัลที่มีผู้ใช้บริการทั่วโลก MetaMask Wallet หรือ MetaMask ที่ให้บริการสำหรับเก็บเหรียญคริปโตบนระบบนิเวศของอีเธอเรียมทุกชนิด ซึ่งอยู่บนมาตรฐานของกลุ่ม ERC-20 รองรับการเก็บ Non-Fungible Token (NFT) ซึ่งเป็นกระเป๋าเงินสินทรัพย์ดิจิทัลที่สามารถเชื่อมต่อกับแอปพลิเคชันแบบกระจายศูนย์อื่น ๆ ได้อย่างสะดวก เช่น การเข้าใช้งานการทำธุรกรรมทางการเงินโดยไม่ผ่านตัวกลาง หรือ Decentralized Finance (DeFi) การซื้อขายเหรียญดิจิทัล รวมไปถึงการเล่นเกม NFT ได้โดยไม่ต้องผ่านตัวกลาง เป็นต้น

ดังนั้นวัตถุประสงค์ของงานวิจัยนี้จึงได้พัฒนาโปรแกรมความภักดีลูกค้าโดยประยุกต์ใช้แอปพลิเคชันแบบกระจายศูนย์เพื่อรองรับการประมวลผลและบันทึกข้อมูล โปรแกรมความภักดีลูกค้าด้วยเทคโนโลยีบล็อกเชน

3.2 การออกแบบกรอบแนวคิดการวิจัย

จากการศึกษาปัญหาและงานวิจัยที่เกี่ยวข้องของผู้วิจัย จึงได้ออกแบบกรอบแนวคิดวิจัย ซึ่งแบ่งออกเป็น 2 ส่วนหลัก ได้แก่ ส่วนที่ 1 ส่วนของร้านค้า โดยร้านค้าจะเป็นผู้กำหนดจำนวนแต้มคะแนนที่ให้กับลูกค้า และสร้างรายการแลกสิทธิประโยชน์ รวมถึงเงื่อนไขต่าง ๆ ผ่านสัญญาอัจฉริยะ และส่วนที่ 2 ส่วนของลูกค้า ลูกค้าจะเป็นผู้นำแต้มคะแนนมาแลกสิทธิประโยชน์ โดยใช้กระเป๋าเงินดิจิทัลในการทำธุรกรรมดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดการวิจัย

3.3 การออกแบบระบบโปรแกรมความภักดีลูกค้า

ขั้นตอนนี้ผู้วิจัยทำการสร้างสัญญาอัจฉริยะโดยใช้ภาษาคอมพิวเตอร์โซลิดิตี (Solidity) ซึ่งประกอบไปด้วย 2 ส่วนสำคัญดังต่อไปนี้

- 1) การออกโทเคน (Token) เพื่อใช้แทนคะแนนสะสมที่แสดงได้ดังภาพที่ 2

```
// SPDX-License-Identifier: MIT
pragma solidity ^0.8.4;

import "@openzeppelin/contracts/token/ERC20/ERC20.sol";
import "@openzeppelin/contracts/token/ERC20/extensions/ERC20Burnable.sol";
import "@openzeppelin/contracts/security/Pausable.sol";
import "@openzeppelin/contracts/access/Ownable.sol";

contract RewardPoint is ERC20, ERC20Burnable, Pausable, Ownable {
    constructor() ERC20("RewardPoint", "RWP") {}

    function pause() public onlyOwner {
        _pause();
    }

    function unpause() public onlyOwner {
        _unpause();
    }

    function mint(address to, uint256 amount) public onlyOwner {
        _mint(to, amount);
    }

    function _beforeTokenTransfer(address from, address to, uint256 amount)
        internal
        whenNotPaused
        override
    {
        super._beforeTokenTransfer(from, to, amount);
    }
}
```

ภาพที่ 2: โครงสร้างของสัญญาอัจฉริยะในออกโทเคน

- 2) ระบบแลกคะแนนสะสม (Redemption) เพื่อทำการแลกสิทธิประโยชน์แสดงได้ดังภาพที่ 3

```
pragma solidity ^0.8.4;

import "https://github.com/OpenZeppelin/openzeppelin-contracts/blob/master/contracts/token/ERC20/ERC20.sol";

contract Tier {
    IERC20 _token;

    mapping(uint => uint) public tierPrices;
    mapping(address => uint) public users;

    constructor(address token) {
        // MEMBER = 0
        // Silver = 1
        // GOLD = 2,
        // WISDOM, = 3
        tierPrices[0] = 0;
        tierPrices[1] = 10000 * 10**8;
        tierPrices[2] = 200000 * 10**8;
        tierPrices[3] = 900000 * 10**8;
        _token = IERC20(token);
    }

    modifier checkAllowance(uint amount) {
        require(_token.allowance(msg.sender, address(this)) >= amount, "Error");
        _;
    }

    function upgradeTier() external payable {
        uint nextTier = users[msg.sender]+1;
        require(nextTier < 3, "Your tier is maximum");
        _token.transferFrom(msg.sender, address(this), tierPrices[nextTier]);
        users[msg.sender] = nextTier;
    }
}
```

โปรแกรมความภักดีลูกค้า

ภาพที่ 3: ตัวอย่างคำสั่งสัญญาอัจฉริยะในการแลก

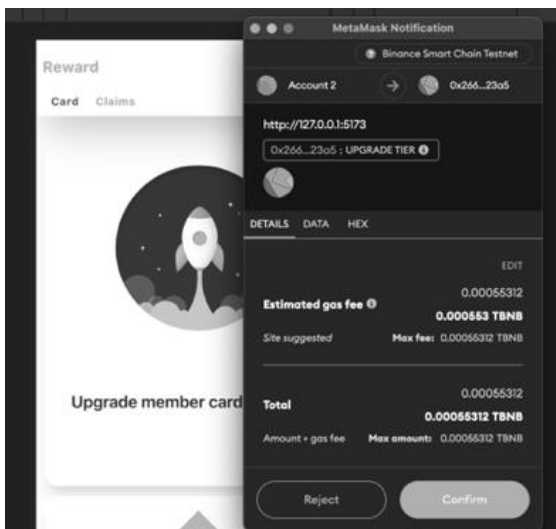
ทำการทดสอบด้วยขั้นตอนการทดสอบแบบ Test Script โดยกำหนดกรณีและเงื่อนไขในการทดสอบไว้ทั้งหมด 4 กรณี ดังต่อไปนี้

- 1) เพิ่มคะแนนสะสมของลูกค้าได้อย่างถูกต้อง และกระทำโดยร้านค้าเท่านั้น
- 2) หักคะแนนสะสมของลูกค้าได้อย่างถูกต้อง และถูกกระทำโดยสัญญาอัจฉริยะเท่านั้น
- 3) แลกคะแนนได้ตามเงื่อนไขที่กำหนดตามที่ตั้งค่าไว้ในสัญญาอัจฉริยะ
- 4) แสดงรายการประวัติการแลกคะแนนได้อย่างถูกต้อง

4. ผลการดำเนินงานวิจัย

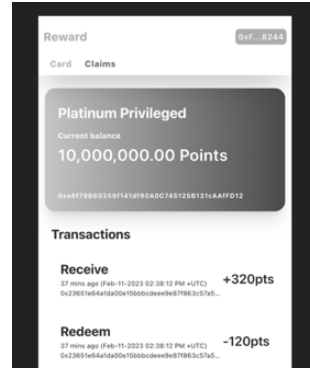
4.1 ผลการพัฒนาโปรแกรมความภักดี

ตัวอย่างหน้าจอในส่วนของการแลกรับสิทธิพิเศษ (Redeem) ที่ลูกค้าสามารถเลือกรายการสิทธิพิเศษที่ต้องการ และกดยืนยันการแลกคะแนน ซึ่งจะแสดงค่าธรรมเนียมในการทำธุรกรรมผ่านเครือข่ายบล็อกเชน ดังภาพที่ 4



ภาพที่ 4: หน้าจอการยืนยันการแลกคะแนน

ตัวอย่างหน้าจอแสดงบันทึกการปรับปรุงคะแนนและสถานะของลูกค้า โดยจะแสดงรายการคะแนนที่ผ่านการแลกสิทธิพิเศษไปแล้ว และคะแนนที่ได้รับดังภาพที่



ภาพที่ 5: หน้าจอแสดงบันทึกการปรับปรุงคะแนน

4.2 ผลการทดสอบการใช้งานโปรแกรม

ผลการทดสอบการใช้งานโปรแกรมความภักดี โดยใช้แอปพลิเคชันแบบกระจายศูนย์สามารถแสดงได้ดังตารางที่ 2 ซึ่งทดสอบจากฟังก์ชันการทำงานต่าง ๆ ที่เกี่ยวข้องกับโปรแกรม

ตารางที่ 2: ผลการทดสอบการใช้งานโปรแกรมความภักดีโดยใช้แอปพลิเคชันแบบกระจายศูนย์

ลำดับ	ฟังก์ชันการทำงาน (Features)	ผลการทดสอบการใช้งานโปรแกรม			
		ผลลัพธ์ที่คาดหวัง (Expected Result)	ผลลัพธ์จากการทดสอบ (Actual Result)	ผ่าน (Pass)	ไม่ผ่าน (Fail)
1	การเพิ่มคะแนน	เพิ่มคะแนนสะสมได้อย่างถูกต้อง	เพิ่มคะแนนสะสมได้อย่างถูกต้อง	/	-
2	การหักคะแนน	หักคะแนนสะสมได้อย่างถูกต้อง	หักคะแนนสะสมได้อย่างถูกต้อง	/	-
3	การแลกคะแนน	แลกคะแนนสะสมได้อย่างถูกต้อง	แลกคะแนนสะสมได้อย่างถูกต้อง	/	-
4	การแสดงคะแนน	แสดงรายการประวัติการแลกและรับคะแนนได้อย่างถูกต้อง	แสดงรายการประวัติการแลกและรับคะแนนได้อย่างถูกต้อง	/	-

5. สรุปผลการวิจัย

จากการพัฒนาโปรแกรมความภักดีโดยประยุกต์ใช้แอปพลิเคชันแบบกระจายศูนย์ที่สร้างอยู่บนเครือข่ายบล็อกเชน และมีรูปแบบการทำงานแบบ Peer-to-Peer โดยในงานวิจัยมีการใช้กระเป๋าเงินดิจิทัล MetaMask ในการชำระค่าธรรมเนียมการทำธุรกรรมบนบล็อกเชน โดยสามารถรองรับฟังก์ชันในการทำงานต่างๆ ที่สำคัญต่อโปรแกรมความภักดี ได้แก่ การเพิ่มคะแนนสะสมของลูกค้า การหักคะแนนสะสมของลูกค้า การแลกเปลี่ยนคะแนนของลูกค้า และการแสดงรายการประวัติการแลกเปลี่ยนของลูกค้า และจากผลการวิจัยพบว่าทุกฟังก์ชันของโปรแกรมสามารถใช้งานได้มีประสิทธิภาพ ซึ่งสามารถเป็นแนวทาง และนำไปประยุกต์ใช้ในธุรกิจอื่น ๆ ได้เป็นอย่างดี

เอกสารอ้างอิง

- [1] สำนักงานพัฒนารัฐบาลดิจิทัล. (2564). *Blockchain for Government Services การใช้เทคโนโลยีบล็อกเชนสำหรับภาครัฐ เวอร์ชัน 2.0*.
- [2] ไตรศักดิ์ ไตรเสนีย์, อมิตา มงคลปริดาไชย และธีรพงศ์ ติลา นุภาพ. (2017). “ต้นแบบสถาปัตยกรรมแบบไม่รวมศูนย์โดยใช้บล็อกเชนสำหรับข้อมูลและการวิจัยด้านการแพทย์: กรณีศึกษาโรงพยาบาลเอกชนในประเทศไทย”. *Journal of Information Science and Technology (JIST)*. 7(2): 32-43
- [3] Marchesi L., Marchesi M., Destefanis G., Barabino G, and Tigano D.. (2020). “Design Patterns for Gas Optimization in Ethereum” *In proceeding of 2020 IEEE International Workshop on Blockchain Oriented Software Engineering*. pp. 9-15
- [4] กมลลา บัวยั้งยืน. (2562). “ปัญหากฎหมายเกี่ยวกับสัญญาอัจฉริยะ”. *วารสารรามคำแหง ฉบับนิติศาสตร์*. 8(2): 109-117
- [5] Saini K., Abhishek R., Chelliah R.P., and Patel T. (2021). “Blockchain 2.0: A Smart Contract” *In proceeding of 2021 IEEE International conference on Computational Evaluation (ComPE)*. pp. 524-528
- [6] Abubashim A. and Tan C.C. (2020). “Smart Contract Designs on Blockchain Applications” *In*

proceeding of 2021 IEEE Symposium on Computers and Communications (ISCC). pp. 1-4

- [7] Ma X., Zhou J., Guo H. and Wang J., "Design of a Stored-value Card Platform Based on Smart Contract," *In proceeding of 2019 3rd International Conference on Circuits, System and Simulation (ICCSS)*, pp. 178-182

ระบบเลเซอร์มาร์คโดยประยุกต์ใช้การรู้จำอักขระด้วยแสง ร่วมกับส่วนการประสานโปรแกรมประยุกต์ Laser Marking System Using Optical Character Recognition and Application Program Interface

ปานุมাত্র อินทรีย์ (Panumart Inchee)¹ และผุสดี บุญรอด (Pudsadee Boonrawd)²

ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407011858097@email.kmutnb.ac.th, ²pudsadee.b@itd.kmutnb.ac.th

บทคัดย่อ

การทำเครื่องหมายบนผลิตภัณฑ์ด้วยเลเซอร์มาร์คได้รับความนิยมในการผลิตของภาคอุตสาหกรรมอิเล็กทรอนิกส์ เพื่อตอบสนองความต้องการของลูกค้าที่หลากหลาย ปัจจุบันการทำงานร่วมกับเครื่องเลเซอร์มาร์คมีความซับซ้อนในการนำเข้าข้อมูล และการแปลงค่าข้อมูล โดยต้องใช้วิธีการคำนวณหาค่าข้อมูลที่ต้องการในการทำเลเซอร์มาร์คของแต่ละผลิตภัณฑ์ ซึ่งต้องใช้ผู้ชำนาญงานทำงานร่วมกับเครื่องจักรเท่านั้น จากปัญหาดังกล่าว วัตถุประสงค์ของงานวิจัยนี้จึงได้พัฒนาระบบเลเซอร์มาร์คโดยประยุกต์ใช้การรู้จำอักขระด้วยแสงร่วมกับส่วนการประสานโปรแกรมประยุกต์ เพื่อให้ได้ระบบสารสนเทศที่ทันสมัย สามารถนำเข้าข้อมูลจากที่เป็นเอกสารแปลงค่าสู่ฐานข้อมูล รองรับการรับส่งข้อมูลระหว่างเครื่องจักรบนเครือข่ายร่วมกัน ทำให้เกิดความสะดวกในการใช้งานลดปัญหาข้อมูลที่ผิดพลาด และสามารถนำไปพัฒนาต่อยอดเป็นระบบอัตโนมัติ เพื่อช่วยยกระดับธุรกิจ ในภาคอุตสาหกรรมอิเล็กทรอนิกส์ให้เข้าสู่ยุคของการนำเทคโนโลยีดิจิทัลมาเปลี่ยนแปลงกระบวนการ (Digital Transformation)

คำสำคัญ: ระบบเลเซอร์มาร์ค การรู้จำอักขระด้วยแสง ส่วนการประสานโปรแกรมประยุกต์

Abstract

Laser marking on products is popular in the production

of the electronics sector to meet the needs of various customers. Nowadays, working with laser marking machines is complicated in data import and data conversion. It must calculate the correct data value in the laser marking of each product, which only requires a specialist to work with the machine. The purpose of this research is to develop a laser marking system by applying optical character recognition with an application interface to achieve a modern information system, able to import data from a converted document into a database, supporting the Transmission of data between machines on a shared network makes it easy to use. And reduce the problem of misinformation. It can also be further developed into an automated system to help elevate businesses in the electronics industry into Digital Transformation.

Keywords: Laser Marking System, Optical Character Recognition, Application Program Interface.

1. บทนำ

บริษัท สตาร์ส ไมโครอิเล็กทรอนิกส์ (ประเทศไทย) จำกัด (มหาชน) เป็นบริษัทที่ผลิตและประกอบชิ้นส่วนอิเล็กทรอนิกส์ชั้นนำของประเทศไทย โดยมีฐานลูกค้าจากต่างประเทศเป็นหลัก การทำเครื่องหมายบนผลิตภัณฑ์ด้วยเลเซอร์มาร์คได้รับความนิยมเป็นอย่างมากสำหรับ

การผลิตภาคอุตสาหกรรมอิเล็กทรอนิกส์ เพื่อตอบสนองความต้องการของลูกค้าที่หลากหลาย ซึ่งการผลิตชิ้นส่วนอิเล็กทรอนิกส์ให้กับลูกค้าจะมีเอกลักษณ์เฉพาะตัว และมีความแตกต่างกัน ในส่วนการทำงานร่วมกับเครื่องเลเซอร์มาร์ค ผู้ใช้งานจะทำการพิมพ์ตัวอักษรหรืออักขระต่าง ๆ เข้าไปที่เครื่องเลเซอร์มาร์ค โดยตรง เครื่องเลเซอร์จะสแกนแสงเลเซอร์ตามข้อมูลที่เกิดจากการพิมพ์ของผู้ใช้งาน ซึ่งอาจทำให้มีโอกาสเกิดข้อผิดพลาดและส่งผลกระทบต่อผลิตภัณฑ์ของลูกค้ายิ่งขึ้นทั้งการจัดเก็บเอกสารที่อยู่ในรูปแบบพีดีเอฟ (Portable Document Format: PDF) ยังถือว่าเป็นอุปสรรคในกระบวนการทำงาน เพราะต้องใช้เวลานานในการแปลงข้อมูล และการนำเข้าฐานข้อมูลเพื่อนำไปใช้ในการประมวลผล

จากปัญหาดังกล่าวผู้วิจัยจึงได้ทำการพัฒนาระบบเลเซอร์มาร์คโดยประยุกต์ใช้การรู้จำอักขระด้วยแสง (Optical Character Recognition: OCR) เพื่อทำการแปลงรูปภาพของข้อความที่ได้จากการพิมพ์ หรือเขียนด้วยลายมือให้เป็นข้อมูลทางดิจิทัลที่เครื่องคอมพิวเตอร์สามารถนำไปประมวลผลได้โดยอัตโนมัติ ซึ่งเป็นแนวทางที่ผสมผสานกันกับการประมวลผลเอกสารอัจฉริยะ (Intelligent Document Processing: IDP) อีกทั้งยังสามารถทำการกำบัง (Masking) เพื่อปกปิดข้อมูลส่วนบุคคลที่มีความอ่อนไหว (Sensitive Data) และเลือกเฉพาะข้อมูลที่ต้องการ และได้ทำการพัฒนาส่วนการประสานโปรแกรมประยุกต์ (Application Program Interface: API) เพื่อให้ได้ระบบสารสนเทศที่รองรับความต้องการในการใช้งานของภาคธุรกิจ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

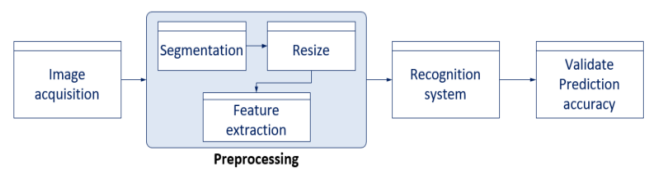
2.1 การรู้จำอักขระด้วยแสง

การรู้จำอักขระด้วยแสงเป็นเทคโนโลยีที่ใช้เพื่อรู้จำหรืออ่านข้อความที่ปรากฏในไฟล์ภาพ โดยเครื่องจะทำการแปลงภาพที่มีข้อความซึ่งพิมพ์ด้วยมือ เขียน หรือพิมพ์ด้วยเครื่องพิมพ์จากไฟล์ รวมถึงเอกสารที่ผ่านการสแกนรูปภาพเอกสาร รูปภาพที่มีสัญลักษณ์หรือฉากที่มีเครื่องหมาย หรือคำบรรยายได้ภาพในภาพให้กลายเป็น

ข้อความ ซึ่งเครื่องจักรสามารถประมวลผล และค้นหาได้ [4] โดยสามารถแสดงกระบวนการทำงานการรู้จำอักขระด้วยแสงดังภาพที่ 1 ซึ่งนำภาพเข้าสู่ระบบ และทำการประมวลผลเพื่อปรับขนาดภาพ ลดสัญญาณรบกวน จากนั้นเข้าสู่กระบวนการรู้จำ และกระบวนการวิเคราะห์ประสิทธิภาพของโมเดล เพื่อให้ได้ผลลัพธ์ที่มีความถูกต้องแม่นยำ

2.1.1 การจำแนกตัวหนังสือ หรืออักขระในแบบภาพรวม (Pattern Recognition) หมายถึง การเปรียบเทียบรูปแบบขนาดทุกตัวอักษร ความกว้างเท่ากัน และแบบลายเส้นเหมือนกัน

2.1.2 การตรวจจับในแต่ละบรรทัด และลายเส้นของตัวหนังสือ (Feature Detection) หมายถึง การจำแนกตัวหนังสือด้วยองค์ประกอบของลายเส้นที่เป็นเส้นตรง เส้นเฉียง เส้นแนวนอน และเส้นคั่นข้าง โดยสามารถประมวลผลทั้งองค์ประกอบของลายเส้น เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง



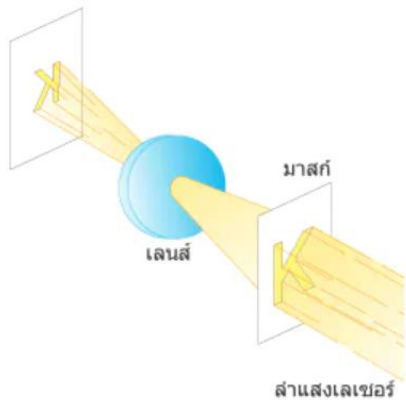
ภาพที่ 1: กระบวนการทำงานการรู้จำอักขระด้วยแสง

2.2 การทำเลเซอร์มาร์คลงบนชิ้นส่วนอิเล็กทรอนิกส์

การทำเลเซอร์มาร์คลงบนชิ้นส่วนอิเล็กทรอนิกส์โดยการใช้แสงเลเซอร์ฉายแสงลงบนวัตถุให้เกิดรอยบนพื้นผิวหรือให้เกิดเครื่องหมายอย่างถาวรบนพื้นผิวชิ้นส่วนอิเล็กทรอนิกส์ ซึ่งสามารถทำงานได้กับวัตถุหลายชนิดทั้งโลหะและอโลหะ โดยแสงเลเซอร์มีความละเอียดอ่อนมาก ซึ่งจะไม่ทำลายชิ้นส่วนทางอิเล็กทรอนิกส์ และวงจรไฟฟ้า การใช้เลเซอร์มาร์คเพื่อให้เกิด โลโก้ ชื่อผลิตภัณฑ์ หมายเลขผลิตภัณฑ์ หมายเลขรุ่น และข้อมูลอื่นๆ [1] โดยเลเซอร์มาร์คสามารถแบ่งออกได้เป็น 2 ชนิด ได้แก่

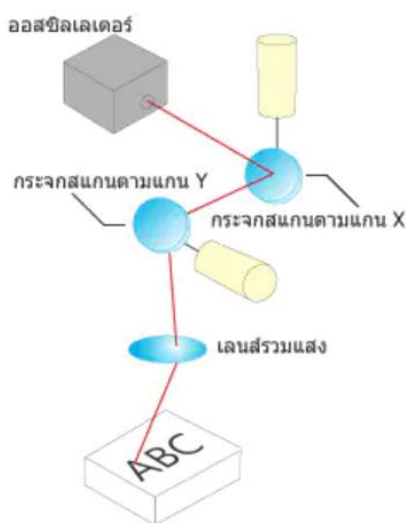
2.2.1 การมาร์คด้วยลำแสงคงที่ ซึ่งแสงเลเซอร์จะถูกฉายผ่านพื้นผิวของมาร์ค โดยจะเป็นการมาร์คในบริเวณที่แสงเลเซอร์ส่องผ่านเท่านั้น เลเซอร์ชนิดมาร์คมีความเร็วและความละเอียดสูง แต่จะต้องเตรียมมาร์คที่มีเนื้อหาดัง

กับสิ่งที่จะมีมาร์คสว่างหน้าดังภาพที่ 2 ซึ่งจากภาพแสดงถึงการมาร์คลำแสงคงที่ [2]



ภาพที่ 2: การมาร์คลำแสงคงที่

2.2.2 การสแกนลำแสงเคลื่อนที่ การสแกนจะฉายแสงขึ้นงานด้วยเลเซอร์ลำแสงเดี่ยว จากนั้นจะทำการมาร์คด้วยการเคลื่อนที่แบบสโตรค กระจกสแกนจะจับลำแสงเลเซอร์ที่ถูกยิงออกมาจากออสซิลเลเตอร์ จากนั้นจะยิงไปที่พื้นผิวของชิ้นงานเพื่อทำการมาร์ค ซึ่งวิธีการสแกนทั่วไปสามารถควบคุมได้เพียงแกน X และ Y เท่านั้น จึงสแกนแสงได้บนพื้นผิวเรียบเพียงอย่างเดียว เลเซอร์มาร์คเกอร์ในปัจจุบันสามารถควบคุมความสูงในแนวแกน Z โดยใช้กระจกสแกนได้ ทำให้มาร์คชิ้นงานที่มีรูปทรงหลากหลายได้มากขึ้นดังภาพที่ 3 ซึ่งแสดงถึงการสแกนลำแสงเคลื่อนที่ [2]



ภาพที่ 3: การสแกนลำแสงเคลื่อนที่

2.3 งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษางานวิจัยต่าง ๆ ที่เกี่ยวข้องกับระบบเลเซอร์มาร์ค ซึ่งมีรายละเอียดดังนี้

การสร้างแบบจำลองแบบใหม่สำหรับการอนุมานที่ตีความได้ และจำแนกประเภทโดยใช้กรอบแนวคิดการให้เหตุผลเชิงประจักษ์สูงสุด [3] วิธีการนี้รวมการวิเคราะห์ทางสถิติ การรวมหลักฐานแบบผสม และการอนุมานตามกฎความเชื่อการเรียนรู้ของเครื่อง การวิเคราะห์ทางสถิติเพื่อใช้หาหลักฐานจากข้อมูลการอนุมานสำหรับวิเคราะห์ความสัมพันธ์ระหว่างอินพุตและเอาต์พุตของระบบ คณิตการพึ่งพาอาศัยกันถูกใช้เพื่อหาปริมาณการพึ่งพาอาศัยกันระหว่างตัวแปรที่มีการใช้อัลกอริทึมทางพันธุกรรมที่ดัดแปลง เพื่อฝึกแบบจำลองสามารถสร้างชุดกฎในระดับเอทริบิวต์เป็นการวิเคราะห์เพื่อสร้างกฎเกณฑ์ของการจำแนกประเภทให้ได้โมเดลที่มีความแม่นยำ [5]

การพัฒนาบริการเครือข่ายสังคมออนไลน์ โดยใช้เทคโนโลยีของเว็บเซอร์วิสมาสนับสนุน เพื่อลดการทำงานที่ซ้ำซ้อนของแต่ละส่วน ทั้งยังช่วยจัดการและควบคุมข้อมูลต่าง ๆ ขององค์กรที่จะส่งผ่านเว็บเซอร์วิสไปยังเครือข่ายสังคม การพัฒนาเว็บเซอร์วิสเป็นศูนย์กลางในการเชื่อมต่อผ่าน Web Services Description Language (WSDL) และ Simple Object Access Protocol (SOAP) [9] ซึ่งในงานวิจัยนี้ได้ใช้เว็บเซอร์วิส เพื่อแลกเปลี่ยนข้อมูลระหว่างเครื่องจักร และเครื่องแม่ข่าย ให้สามารถติดต่อและสามารถทำงานข้ามแพลตฟอร์ม Morphology Technique การค้นหาขอบภาพ Sobel edge detection และการวิเคราะห์ไบนารี ขั้นตอนที่ 2 ใช้เทคนิคการรู้จำตัวอักษรด้วยแสงร่วมกับการหาค่าที่ใกล้เคียงเพื่อการรู้จำตัวอักษรบนแผ่นป้ายทะเบียน การวิจัยนี้ได้นำเทคนิคการประมวลผลล่วงหน้าในขั้นตอนและแง่มุมต่าง ๆ เช่น การตรวจสอบลายเซ็น และการตัดแยกไฟล์เอกสารเว็บเซอร์วิสบริการเครือข่ายสังคมกรณศึกษาเฟสบุ๊ค [10]

3. วิธีการดำเนินการวิจัย

วิธีการดำเนินงานวิจัยของระบบเลเซอร์มาร์คโดย

ประยุกต์ใช้การรู้จำอักขระด้วยแสงร่วมกับส่วนการประสานโปรแกรมประยุกต์ที่มีรายละเอียดดังนี้

3.1 การศึกษาข้อมูลและวิเคราะห์ปัญหา

จากการศึกษาข้อมูลและวิเคราะห์ปัญหาในปัจจุบันพบว่าการทำงานร่วมกับเครื่องจักรมีกระบวนการทำงานที่ซับซ้อน ใช้เวลาในการดำเนินงาน และอาจทำให้เกิดข้อผิดพลาดขึ้นได้ ดังนั้นงานวิจัยนี้จึงทำการพัฒนาระบบเลเซอร์มาร์ค โดยประยุกต์ใช้การรู้จำอักขระด้วยแสงร่วมกับส่วนการประสานโปรแกรมประยุกต์เพื่อแก้ปัญหาดังกล่าว

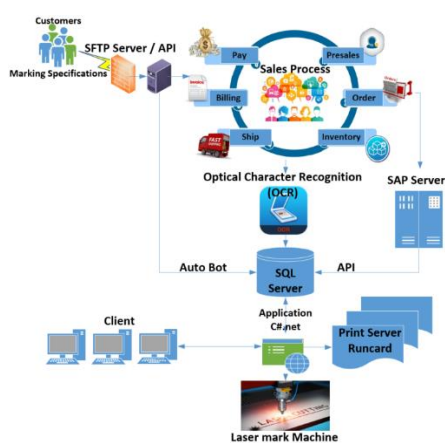
3.2 การออกแบบระบบเลเซอร์มาร์ค

การออกแบบระบบเลเซอร์มาร์คนั้น ผู้วิจัยได้ออกแบบกรอบแนวคิด และกระบวนการทำงานต่าง ๆ ดังนี้

3.2.1 การรับส่งข้อมูล (File Transfer Protocol : FTP) ในส่วนกระบวนการนี้ทำการรันโปรแกรมแบบอัตโนมัติ (Automatic) โดยผู้วิจัยได้ทำการเขียนโปรแกรมเพื่อใช้ในการรับส่งข้อมูลจากเครื่องแม่ข่ายที่เชื่อมต่อภายนอกองค์กรผ่านระบบเครือข่ายเข้าสู่ฐานข้อมูลภายในองค์กร (Database SQL Server) โดยทำการติดตั้งโปรแกรมให้ทำงานแบบอัตโนมัติ เพื่อการรับส่งข้อมูลที่รวดเร็ว ดังภาพที่ 4

3.2.2 การจัดเก็บไฟล์เอกสารผ่านเครื่องอ่านการรู้จำอักขระด้วยแสง โดยการนำเอกสารเข้าเครื่องอ่านการรู้จำอักขระด้วยแสงตามตำแหน่งที่ต้องการ ซึ่งเครื่องฯ จะทำการอ่านข้อมูล และบันทึกเข้าสู่ฐานข้อมูลดังภาพที่ 4

3.2.3 การสร้างแบบจำลอง ผู้วิจัยได้ทำการศึกษา และรวบรวมข้อมูลการทำเลเซอร์มาร์คของแต่ละผลิตภัณฑ์ โดยทำการจัดกลุ่มตามความแตกต่าง เพื่อการกำหนดเงื่อนไขในการพัฒนาโปรแกรม



ภาพที่ 4: กรอบแนวคิดการพัฒนาเลเซอร์มาร์ค

3.2.4 การทำงานร่วมกับเครื่องจักร โดยผ่าน Dynamic Link Library (DLL) ซึ่งเป็น โปรแกรมขนาดเล็กที่รวมชุดคำสั่งต่าง ๆ ที่สามารถแยกทำงานได้อย่างอิสระตามหน้าที่ของแต่ละ DLL สามารถเรียกใช้งานเมื่อจำเป็น และแบ่งกันใช้งานทรัพยากรได้อย่างมีประสิทธิภาพ ในการเรียกใช้ DLL เดียวกันจากหลายระบบพร้อมกัน ทำให้ใช้พื้นที่ในการจัดเก็บลดลง จากคุณสมบัตินี้ทำให้ประหยัดหน่วยความจำ ง่ายต่อการพัฒนาแก้ไข และปรับปรุง ทำให้การตอบสนองของเครื่องจักรมีความเร็วสูงและเป็นไปอย่างอัตโนมัติมากยิ่งขึ้น

3.3 การเตรียมข้อมูล เครื่องจักร และการพัฒนาระบบ

ในขั้นตอนการเตรียมข้อมูล เครื่องจักร และพัฒนาระบบสารสนเทศ มีรายละเอียดดังนี้

3.3.1 การพัฒนาระบบเลเซอร์มาร์ค ผู้วิจัยจัดทำส่วนการประสานโปรแกรมประยุกต์ (Application Program Interface : API) ซึ่งเป็นส่วนเชื่อมต่อประสานกันระหว่างแอปพลิเคชัน เพื่อแลกเปลี่ยนข้อมูลระหว่างเครื่องลูกข่ายและเครื่องแม่ข่าย โดยงานวิจัยนี้ใช้ Web Socket API ที่มีรูปแบบ JSON Format ในการรับส่งข้อมูลให้มีประสิทธิภาพ ซึ่งแสดงได้ดังภาพที่ 5

3.3.2 ข้อมูลที่ใช้ในการพัฒนาและประเมินระบบสารสนเทศ ผู้วิจัยได้นำข้อมูลตามรูปแบบ Marking Specifications จำนวน 1,000 ชุด ซึ่งแบ่งกลุ่มตามรูปแบบการกำหนดค่าข้อมูลได้ 20 รูปแบบ ที่มีเงื่อนไขที่แตกต่างกัน

```

//POST api/value
[HttpPost("create")]
References
public ActionResult<string> Post([FromBody] ZRoot request)
{
    using var db = this._dbContext.CreateConnection();
    using var tx = new TransactionScope();

    var lotData = new lotdata()
    {
        testerid = request.lotdata.testerid,
        mode = request.lotdata.mode,
        product = request.lotdata.product,
        datecode = request.lotdata.datecode,
        lot = request.lotdata.lot,
        temperature = request.lotdata.temperature,
        lot_operator = request.lotdata.lot_operator,
        program = request.lotdata.program,
        flow = request.lotdata.flow,
        tpversion = request.lotdata.tpversion
    };
    db.Insert(lotData);

    var units = new Munits();
    db.Insert(units); //OK

    foreach (var item in request.units.siteresult)
    tx.Complete();

    return Content("{\"Success\":true, \"message\":\"OK\"}");
}
    
```

ภาพที่ 5: ส่วนประสานโปรแกรมประยุกต์

3.4 การประเมินคุณภาพระบบเลเซอร์มาร์ค

การประเมินคุณภาพระบบเลเซอร์มาร์ค ซึ่งทำการประเมินแบบ User Acceptance Test (UAT) โดยผู้เชี่ยวชาญ ซึ่งมีรายละเอียดดังนี้

3.4.1 ด้านความสามารถทำงานได้ตามหน้าที่ (Function Requirement Test)

3.4.2 ด้านความถูกต้องของความสามารถทำงานได้ตามหน้าที่ (Functional Test)

3.4.3 ด้านความง่ายต่อการใช้งาน (Usability)

3.4.4 ด้านการรักษาความปลอดภัยของข้อมูล (Security Test)

4. ผลการวิจัย

ผลการวิจัยจากการพัฒนาระบบเลเซอร์มาร์คสามารถแสดงตัวอย่างหน้าจอระบบสารสนเทศ โดยภาพที่ 6 แสดงส่วนที่เป็นข้อมูลหลักที่ได้ข้อมูลจากการเชื่อมต่อผ่านส่วนการประสาน โปรแกรมประยุกต์ และภาพที่ 7 แสดงผลจากการคำนวณค่าข้อมูลตามรูปแบบเงื่อนไขที่แตกต่างกัน และแสดงส่วนของผู้อนุมัติ



ภาพที่ 6: หน้าจอแสดงส่วนที่เป็น Data Master



ภาพที่ 7: หน้าจอแสดงส่วนที่เป็นการแสดงผลและแสดงส่วนการอนุมัติหน้ามาร์ค

ผลการประเมินระบบเลเซอร์มาร์คจากผู้เชี่ยวชาญจาก 3 หน่วยงานที่เข้าร่วม ได้แก่

4.1 หน่วยงานวิศวกรผลิตภัณฑ์ (Product Engineer) หน่วยงานที่สร้างผลิตภัณฑ์ และกำหนดเงื่อนไขของการทำเลเซอร์มาร์ค

4.2 หน่วยงานวิศวกรกระบวนการผลิต (Process Engineer) หน่วยงานที่สร้างขั้นตอนการทำงาน และกระบวนการทำงานของระบบ

4.3 หน่วยงานตรวจสอบคุณภาพ (Quality Engineer) หน่วยงานที่ตรวจสอบคุณภาพการผลิต และคุณภาพของผลิตภัณฑ์

ผลการประเมินการพัฒนาระบบเลเซอร์มาร์คโดยประยุกต์ใช้การรู้จำอักขระด้วยแสงและส่วนการประสานโปรแกรมประยุกต์โดยภาพรวมอยู่ในระดับดีมาก โดยด้านความสามารถทำงานได้ตามหน้าที่อยู่ในระดับดีมาก ซึ่งมีค่าเฉลี่ยอยู่ที่ 4.51 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.55 ระดับคุณภาพดีมาก ด้านความถูกต้องของความสามารถทำงานได้ตามหน้าที่อยู่ในระดับดีมาก ซึ่งมีค่าเฉลี่ยอยู่ที่ 4.54 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.50 ด้านความง่ายต่อการใช้งานอยู่ในระดับดี ซึ่งมีค่าเฉลี่ยอยู่ที่ 4.48 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.37 และด้านการรักษาความปลอดภัยของข้อมูลอยู่ในระดับดีมาก ซึ่งมีค่าเฉลี่ยอยู่ที่ 4.75 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.43

5. สรุปผลการวิจัย

จากการพัฒนาระบบเลเซอร์มาร์คโดยประยุกต์ใช้การรู้จำอักขระด้วยแสงร่วมกับส่วนการประสานโปรแกรมประยุกต์เพื่อทำการแปลงข้อมูลเอกสารให้เป็นข้อมูลดิจิทัลที่เครื่องคอมพิวเตอร์สามารถนำไปใช้ในการประมวลผลและทำการพัฒนาในส่วนการประสานโปรแกรมประยุกต์สำหรับเชื่อมต่อข้อมูลระหว่างเครื่องจักรบนเครือข่ายให้สามารถแลกเปลี่ยนข้อมูลระหว่างเครื่องลูกข่ายและเครื่องแม่ข่ายได้อย่างมีประสิทธิภาพ ช่วยลดข้อผิดพลาดจากการคำนวณของผู้ใช้งาน โดยตรง จากผลการวิจัยระบบสารสนเทศที่พัฒนาขึ้นอยู่ในระดับดีมาก ซึ่งรองรับความต้องการในการใช้งาน สามารถค้นหาข้อมูลที่ต้องการได้อย่างแม่นยำ อำนวยความสะดวก ลดระยะเวลาในการดำเนินงานและยังสามารถนำไปพัฒนาต่อยอดระบบสารสนเทศอื่น ๆ ที่เกี่ยวข้องในอนาคตได้

เอกสารอ้างอิง

- [1] เทคโนโลยีสารสนเทศ, แผนก. (2564). "MES Systems 2021" พระนครศรีอยุธยา: บริษัท สดาร์ส ไมโครอิเล็กทรอนิกส์ (ประเทศไทย) จำกัด (มหาชน), 2564.
- [2] ศูนย์รวมการมาร์คด้วยเลเซอร์, แผนก. (2566). "การมาร์คด้วยเลเซอร์" กรุงเทพฯ : บริษัท คีย์เอ็นซ์ (ไทยแลนด์) จำกัด, 2566.
- [3] Shuaiyu Yao, Jian-Bo Yang and Dong-Ling Xu. (2021). "A probabilistic modeling approach for interpretable data inference and classification" School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China Alliance Manchester Business School, The University of Manchester, Manchester, United Kingdom Journal of Intelligent & Fuzzy Systems 40 (2021) 5101–5117.
- [4] Sholom M. Weiss, Nitin Ind Urkhya, (1995) "Rule-based Machine Learning Methods of Functional Prediction" Journal of Artificial Intelligence Research 3 (1995) 383-403.
- [5] Laszlo Kovacs, (2020) "Concept Lattice-Based Classification in NPL" Proceedings of 14th International Conference INTER-ENG 2020 Interdisciplinarity in Engineering, Mures, Romania, 8–9 October 2020.
- [6] Raghunath Dey, Rakesh Chandra Balabantaray, Surajit Mohanty, Debabrata Singh, Marimuthu Karuppiah, Debabrata Samanta, (2022) "Approach for Preprocessing in

Offline Optical Character Recognition (OCR)" Interdisciplinary Research in Technology and Management (IRTM)|978- 1-6654-7886-1/22/\$31.00 ©2022 IEEE | DOI: 10.1109.

- [7] ศตวรรษ อรุณรัตน์ และยุวธิดา ชิวปรีชา, (2559). "ระบบการบริหารจัดการรับ-ส่งไฟล์โดยใช้โพรโตคอลเอฟทีพี" การประชุมวิชาการระดับชาติ วิทยาศาสตร์และเทคโนโลยีระหว่างสถาบัน ครั้งที่ 4, 2559.
- [8] ศุภกร รัศมีมณฑล, (2541). "การพัฒนาโปรแกรมส่งถ่ายข้อมูลบนโพรโตคอล FTP และ HTTP โดยวิธีการเชิงวัตถุ" วิทยานิพนธ์ วิทยาศาสตรมหาบัณฑิต มหาวิทยาลัยสงขลานครินทร์, 2541.
- [9] วิรัชพัชร พรหมนิวงค์, (2546). "การพัฒนาเว็บเซอร์วิสเพื่อแลกเปลี่ยนข้อมูลระหว่างเครื่องแม่ข่ายต่างแพลตฟอร์มการศึกษา ระบบตรวจสอบผลการศึกษาศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ" สารนิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัยสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2546.
- [10] นริพร พิรุฬห์ทรัพย์, (2552). "การลดสิ่งรบกวนในไทยโอซีอาร์ โดยการเรียนรู้แบบกึ่งสอน" มหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2552.

การพัฒนาแนวทางการทดสอบระบบโดยใช้เทคนิคแบล็กบ็อกซ์เทสดีง สำหรับฟังก์ชันมาตรฐาน

The Development of Guideline for Software Testing by Using Black-Box Techniques for Standard Functions

ศราวดี ตระกุลสุน (Sarawadee Trakunhun)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856101@email.kmutnb.ac.th, ²sakchai.t@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแนวทางการทดสอบระบบโดยใช้เทคนิคแบล็กบ็อกซ์เทสดีงสำหรับฟังก์ชันมาตรฐาน ซึ่งมีวิธีการดำเนินการวิจัยประกอบด้วยการศึกษาฟังก์ชันระบบที่ใช้พื้นฐานจากการค้นคว้างานวิจัยที่เกี่ยวข้องและวิเคราะห์ความต้องการของผู้ใช้งาน นำมาเปรียบเทียบฟังก์ชันมาตรฐานที่มีการพัฒนาแนวทางมากที่สุด และได้นำฟังก์ชันดังกล่าวมาสร้างกรณีทดสอบระบบด้วยเทคนิคแบล็กบ็อกซ์เทสดีง โดยนำกรณีทดสอบดังกล่าวมาผ่านการหาความเที่ยงตรงตามเนื้อหาจากผู้เชี่ยวชาญ (IOC) และนำมาให้ผู้ใช้งานทดสอบกับระบบเพื่อการใช้งานจริง พร้อมทั้งประเมินความเหมาะสมของแนวทางการทดสอบระบบดังกล่าวเพื่อให้ตรงตามวัตถุประสงค์ ผลการวิจัยพบว่าแนวทางการทดสอบระบบโดยใช้เทคนิคแบล็กบ็อกซ์เทสดีงสำหรับฟังก์ชันมาตรฐาน เมื่อนำไปใช้งานทดลองโดยกลุ่มผู้ใช้งานจำนวน 10 คน ได้ผลค่าเฉลี่ยความเหมาะสมเท่ากับ 4.94 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 ซึ่งสามารถสรุปได้ว่าแนวทางการทดสอบระบบดังกล่าวมีความเหมาะสมในระดับเหมาะสมมากที่สุด และสามารถนำไปใช้งานได้จริง ถูกต้องตรงตามความต้องการผู้ใช้งานและมาตรฐานการประกันคุณภาพ

คำสำคัญ: แบล็กบ็อกซ์เทสดีง กรณีทดสอบ ทดสอบระบบ

Abstract

The aim of this research is to develop of guideline for software testing by using black-box techniques for standard functions. The methodology is to research and analyzing requirement. To compare the standard functions with the most developed guidelines and brought the above function to create test cases for software testing by using black-box techniques. These test cases are passed through an expert: Index of item-objective congruence (IOC). As well as assessing the suitability of the method for testing such systems to meet the objectives. The evaluation of test cases for software testing by using black-box techniques was conducted by 10 users. The result of appropriate level was found to have an average of 4.94 with 0.25 standard deviation. This guideline for software testing by using black-box techniques for standard functions can be put into practice requirements and quality assurance standards.

Keywords: Black-Box Techniques, Testcases, Software Testing.

1. บทนำ

ในปัจจุบันคอมพิวเตอร์และเทคโนโลยีสารสนเทศได้มีการพัฒนาอย่างมากและแพร่หลาย อีกทั้งยังนำมาประยุกต์ในอุตสาหกรรมด้านต่าง ๆ อาทิเช่น ด้านเศรษฐกิจ ด้านการเงินบัญชี ด้านบริหารทรัพยากรมนุษย์ เป็นต้น

จึงทำให้มีการพัฒนาเทคโนโลยีสารสนเทศที่หลากหลาย มีคุณภาพ ทำงานได้ตรงต่อความต้องการและอำนวยความสะดวกในการทำงานได้มีประสิทธิภาพมากยิ่งขึ้น การพัฒนาซอฟต์แวร์ในปัจจุบันนั้นมีเป้าหมายคือการพัฒนาซอฟต์แวร์ให้มีคุณภาพตรงความต้องการใช้งานและไม่มีข้อผิดพลาดภายในซอฟต์แวร์ จึงได้มีการทดสอบระบบซอฟต์แวร์เป็นหนึ่งในขั้นตอนกระบวนการพัฒนาซอฟต์แวร์ การทดสอบซอฟต์แวร์เป็นการค้นหาข้อผิดพลาดที่มีอยู่ในระบบซอฟต์แวร์ ตรวจสอบความถูกต้องทั้งในแง่ของกระบวนการทำงาน และความถูกต้องตรงต่อความต้องการของลูกค้า ช่วยให้ซอฟต์แวร์ที่พัฒนาขึ้นมีความถูกต้อง สมบูรณ์ ปลอดภัยและมีประสิทธิภาพที่ดี ผู้ที่ทำการทดสอบจะถูกเรียกว่าทีมทดสอบ (Test team) โดยทั่วไปจะประกอบด้วย นักทดสอบมืออาชีพ นักวิเคราะห์ นักออกแบบระบบ ผู้เชี่ยวชาญการจัดการ โครงแบบ และผู้ใช้งานระบบ การทดสอบซอฟต์แวร์จะมีให้เลือก 2 เทคนิควิธี คือการทดสอบแบบกล่องขาว (Whitebox Testing) และการทดสอบแบบกล่องดำ (Blackbox Testing)

อีกทั้งเพื่อให้ซอฟต์แวร์ระบบมีคุณภาพมากยิ่งขึ้น จำเป็นที่จะต้องได้รับการทดสอบซอฟต์แวร์ที่เป็นไปตามมาตรฐานการประกันคุณภาพซอฟต์แวร์ โดยส่วนใหญ่การควบคุมคุณภาพของการทดสอบซอฟต์แวร์เป็นไปด้วยความยากลำบาก นักพัฒนาบางรายจึงได้มีการเลือกที่จะทดสอบในรูปแบบสุ่ม ซึ่งการเลือกทดสอบด้วยวิธีนี้อาจจะทำให้เกิดข้อผิดพลาด เนื่องจากข้อมูลที่จะก่อให้เกิดปัญหาในบางครั้งอาจจะไม่ได้อยู่ในกรณีชุดทดสอบ ทำให้ซอฟต์แวร์แจ้งกรณีทดสอบผ่าน ทั้งที่การทดสอบดังกล่าวไม่ครอบคลุมทุกกรณี หรือนักพัฒนาบางรายเลือกใช้การประมวลผลแบบลำดับ (Sequential Processing) เพื่อให้ง่ายต่อการตรวจแก้จุดบกพร่อง (Defect) แต่ก็พบว่า การทดสอบระบบทำงานได้ล่าช้า และไม่สะท้อนพฤติกรรมของทดสอบระบบจริง

ในงานวิจัยนี้ ผู้วิจัยได้นำเทคนิควิธีการทดสอบ Blackbox Testing มาเป็นเครื่องมือช่วยในการพัฒนาแนวทางทดสอบซอฟต์แวร์ระบบ โดยการสร้างกรณีทดสอบระบบสำหรับฟังก์ชันมาตรฐาน โดยพัฒนาแนวทางการทดสอบระบบ (Guideline) ขึ้นมาใช้ในงานในแผนการจัดการควบคุมคุณภาพ

(Quality Management) จึงได้มีการศึกษาฟังก์ชันที่ใช้พื้นฐานจากงานวิจัยต่าง ๆ และส่วนใหญ่ได้มีการพัฒนาการสร้างกรณีทดสอบขึ้นเป็นลักษณะในรูปแบบเดิม ไม่มีแบบแผนหรือมาตรฐานที่แน่นอน อีกทั้งยังมีข้อผิดพลาดไปถึงผู้ใช้งาน (User) โดยผู้วิจัยจึงต้องการให้งานวิจัยนี้ได้มีส่วนร่วมช่วยพัฒนาการทดสอบซอฟต์แวร์ให้มีประสิทธิภาพ ช่วยการทำงานให้สะดวกและรวดเร็วมากยิ่งขึ้น ลดระยะเวลาในการทดสอบซอฟต์แวร์ และส่งผลให้ภาพรวมของการพัฒนาซอฟต์แวร์มีคุณภาพและมั่นใจในประสิทธิภาพได้มากยิ่งขึ้น โดยอยู่ภายใต้ขอบเขตระยะเวลาที่กำหนด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การทดสอบซอฟต์แวร์ [1]

การทดสอบซอฟต์แวร์ หรือ Software Testing เป็นขั้นตอนหลักของการพัฒนาซอฟต์แวร์ โดยเป็นขั้นตอนกระบวนการค้นหาข้อผิดพลาดที่มีอยู่ในระบบ ช่วยให้ซอฟต์แวร์ที่พัฒนาขึ้นมีความถูกต้อง สมบูรณ์ ปลอดภัย และมีประสิทธิภาพที่ดี ซึ่งขั้นตอนการทดสอบความผิดพลาดของซอฟต์แวร์แบ่งเป็น 2 ขั้นตอน ดังนี้

2.1.1 การทดสอบระดับโปรแกรม

มีวัตถุประสงค์เพื่อตรวจสอบโปรแกรมที่มีการสร้างขึ้นโดยนักพัฒนา เมื่อรวมกันเป็นระบบการทำงานแล้วสามารถทำงานได้ตามที่ออกแบบไว้ ความผิดพลาดที่อาจจะเกิดขึ้นในระดับนี้ เช่น ความผิดพลาดของขั้นตอนวิธี ความผิดพลาดด้านไวยากรณ์ ความผิดพลาดทางด้านการคำนวณ เป็นต้น

2.1.2 การทดสอบระดับระบบ

มีวัตถุประสงค์เพื่อตรวจสอบว่าระบบที่พัฒนาขึ้นสามารถแก้ไขปัญหาได้ตรงตามความต้องการที่ระบุในเอกสารกำหนดความต้องการ ความผิดพลาดที่อาจเกิดขึ้นในระดับนี้ได้แก่ การวิเคราะห์ ความต้องการไม่ชัดเจน การแปลความหมายไม่ถูกต้อง การสื่อสารระหว่างทีมงานพัฒนาผิดพลาด การออกแบบระบบ ความผิดพลาดสามารถเกิดขึ้นจากลูกค้า นักออกแบบระบบ นักออกแบบ โปรแกรมเมอร์ ทีมงานทดสอบ รวมทั้งทีมงานในการบำรุงรักษาระบบได้

2.2 การทดสอบฟังก์ชัน (Function testing) [1]

การทดสอบฟังก์ชัน หรือ Function Testing เป็นการทดสอบมุ่งเน้นไปที่การทำงานตามที่ลูกค้าต้องการ โดยตรวจสอบจากเอกสารระบุความต้องการ มีการเปรียบเทียบระหว่างระบบที่สร้างขึ้นกับเอกสารที่ระบุความต้องการ ทีมงานทดสอบจะไม่สนใจโครงสร้างของระบบ แต่สนใจเฉพาะกิจกรรมที่ระบบสามารถกระทำได้นั้น เป็นลักษณะกล่องปิด ดังนั้นในการทดสอบจึงจำเป็นต้องมีสิ่งที่ต้องทราบถึงหน้าที่และกิจกรรมต่าง ๆ ที่ต้องการ

2.3 ระดับการทดสอบซอฟต์แวร์ [2]

การทดสอบซอฟต์แวร์ คือ ตรวจสอบคุณสมบัติ และทดสอบความสามารถในการทำงานในด้านต่าง ๆ ของซอฟต์แวร์ให้เป็นไปตามรูปแบบและข้อกำหนดต่าง ๆ ที่ได้ออกแบบไว้ โดยระดับการทดสอบซอฟต์แวร์ มีดังนี้

2.3.1 การทดสอบระดับหน่วย

การทดสอบหลังจากที่ผู้พัฒนาโปรแกรมได้เขียนโปรแกรมตามที่ได้รับมอบหมายเสร็จสิ้นโดยจะตรวจสอบหาจุดบกพร่อง ข้อบกพร่องและโครงสร้างของโปรแกรมทำงานได้ถูกต้อง ครบถ้วน สมบูรณ์ ตรงตามความต้องการ การเขียน Unit test เป็นอีกวิธีการที่จะทำให้มั่นใจได้ว่า ส่วนของโปรแกรมที่เล็กนั้น สามารถทำงานได้ถูกต้อง มีโปรแกรมเมอร์หรือผู้พัฒนาระบบเป็นผู้ทดสอบ

2.3.2 การทดสอบการทำงานร่วมกัน

การทดสอบการทำงานของโมดูลโปรแกรมทั้งหมด โดยทำการทดสอบในแต่ละฟังก์ชันหรือ โมดูล สามารถทำงานได้ตรงตามความต้องการ ไม่เกิดผลกระทบ การทดสอบระดับนี้จะแสดงให้เห็นบุคคลอื่นที่ไม่ใช่ นักพัฒนา ได้เห็นว่าระบบทั้งหมดสามารถทำงานร่วมกันได้ โดยมีผู้ทดสอบระบบเป็นผู้ทดสอบ

2.3.3 การทดสอบทั้งระบบ

การทดสอบระบบหรือโปรแกรม โดยดูจากภาพรวมระบบการทำงานมีการตอบสนองต่อความต้องการทั้งใน ส่วนของฟังก์ชันการทำงานและประสิทธิภาพการทำงาน สอดคล้องกับความต้องการของลูกค้า (Requirement specification) โดยมีผู้ทดสอบระบบเป็นผู้ทดสอบ และในระดับนี้ทางทีมผู้พัฒนาได้นำใช้ในการพัฒนาระบบ

2.3.4 การทดสอบเพื่อการรับมอบ

กระบวนการทดสอบระบบก่อนใช้งานจริง เพื่อเปรียบเทียบประสิทธิภาพของระบบกับความต้องการของลูกค้า โดยในระดับนี้ ผู้ใช้งานจะเข้ามามีส่วนร่วม ในการทดสอบระบบก่อนที่จะใช้งานจริง

2.4 วงจรการพัฒนากระบวนงาน [3]

วงจรการพัฒนากระบวนงาน (System Development Life Cycle : SDLC) มีขั้นตอนในการพัฒนา 7 ขั้นตอน มีดังนี้

2.4.1 ค้นหาปัญหา โอกาสและเป้าหมาย

ระบบสารสนเทศจะเกิดขึ้นได้ก็ต่อเมื่อผู้บริหารหรือ ผู้ใช้ตระหนักว่าต้องการระบบสารสนเทศหรือต้องการแก้ไข

2.4.2 ศึกษาความเป็นไปได้

กำหนดปัญหา โดยนักวิเคราะห์และออกแบบระบบ ต้องกำหนดให้ได้ว่าการแก้ปัญหานั้น มีความเป็นไปได้ทางเทคนิคหรือไม่

2.4.3 วิเคราะห์ความต้องการของระบบ

วิเคราะห์และกำหนดความต้องการระบบใหม่ กำหนด เครื่องมือข้อมูลที่ใช้ งาน บุคลากรและหน้าที่ โดยผู้ใช้ต้องให้ความร่วมมือ ศึกษาเอกสารที่มีอยู่ และศึกษาระบบเดิมเพื่อให้ เข้าใจขั้นตอนการทำงานของระบบ

2.4.4 การออกแบบระบบ

การออกแบบระบบใหม่เพื่อให้สอดคล้องกับความต้องการของผู้ใช้งานและผู้บริหาร

2.4.5 พัฒนาซอฟต์แวร์และจัดทำเอกสาร

การเขียนโปรแกรม จัดทำคู่มือการใช้งานโปรแกรม และรวมถึงการฝึกอบรมผู้ที่มีความเกี่ยวข้องในระบบ โดย มีการกำหนดบุคลากรในด้านต่างๆ

2.4.6 ทดสอบและบำรุงรักษาระบบ

นักพัฒนาระบบและทีมงานทดสอบโปรแกรม มีการตรวจสอบระบบ จะต้องสามารถทำงานได้ตามที่ต้องการ

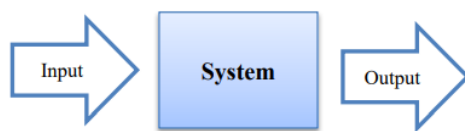
2.4.7 ดำเนินงานและประเมินผล

มีการติดตั้งระบบให้พร้อมการใช้งาน นำระบบใหม่ มาใช้แทนระบบเดิม โดยนักวิเคราะห์และออกแบบระบบ ทำการประเมินผลเพื่อให้ทราบถึงความพอใจของผู้ใช้ ระบบหรือสิ่งที่ต้องแก้ไขปรับปรุงหรือปัญหาที่พบ

2.5 เทคนิคแบล็กบ็อกซ์เทสติ้ง [4]

สำหรับเทคนิคในการทดสอบซอฟต์แวร์ที่นิยมใช้ได้แก่ Black-box testing และ White-box testing

เทคนิค Blackbox Testing คือการทดสอบการทำงานของซอฟต์แวร์ที่ไม่ได้สนใจกลไกภายในของระบบ การทำการทดสอบเน้นเพื่อการตรวจสอบผลการทำงานของระบบในแต่ละหน้าที่ตามข้อกำหนดความต้องการว่าถูกต้องตามความต้องการ ผู้ที่ทำการทดสอบนั้น ไม่ทำการเข้าถึงรหัสต้นฉบับ (Source code) โดยรหัสต้นฉบับจะถูกพิจารณาให้เป็น “Big black box” ที่ผู้ทดสอบระบบไม่สามารถเห็นภายในได้ โดยผู้ทดสอบระบบจะรู้เพียงอย่างเดียวว่าสามารถให้ Input แก่ระบบและระบบส่ง Output ดังภาพที่ 1



ภาพที่ 1: เทคนิค Blackbox Testing

โดยเทคนิค Blackbox Testing สามารถทดสอบโดยที่ไม่จำเป็นต้องมีความรู้เฉพาะแต่จะเป็นวิธีการทดสอบจากมุมมองของผู้ใช้งานจริง ไม่ใช่ นักพัฒนาหรือผู้ออกแบบระบบ ซึ่งในบางครั้งการออกแบบ อาจจะไม่ได้ออกมาตามความต้องการของผู้ใช้งานจริง ผู้ทดสอบสามารถออกแบบแนวทางการทดสอบหรือสร้างกรณีทดสอบได้ทันที หลังจากที่รับเอกสารความต้องการลูกค้า โดยไม่ต้องรอให้มีการพัฒนาระบบเสร็จก่อน หรือ ถ้าหากมีการเปลี่ยนแปลงการเขียน Code กรณีทดสอบเดิมที่เคยเขียน ก็ยังสามารถนำมาใช้ได้ เพราะผลลัพธ์ที่ได้ยังคงเดิม

2.6 ทฤษฎีที่เกี่ยวกับการประกันคุณภาพซอฟต์แวร์ Software Quality ISO/IEC 9126 [5]

คุณภาพซอฟต์แวร์ หมายถึง การดำเนินงานที่มีประสิทธิภาพ ซึ่งเป็นไปตามข้อกำหนด ระดับที่ผลิตภัณฑ์ซอฟต์แวร์ สามารถตอบสนองต่อความต้องการของผู้ใช้งาน ซอฟต์แวร์ได้อย่างครบถ้วน ถึงจะสามารถเรียกได้ว่าซอฟต์แวร์มีคุณภาพ โดยคุณลักษณะของซอฟต์แวร์ที่มีคุณภาพตามมาตรฐานนานาชาติ ISO/IEC9126 เป็นสิ่งที่ถูกนำมาใช้ในการกำหนดระดับคุณภาพซอฟต์แวร์ มีการกำหนดไว้ดังนี้

2.6.1 ประโยชน์ใช้สอย (Functionality)

ซอฟต์แวร์ต้องมีประโยชน์ตรงตามความต้องการของลูกค้า มีคุณลักษณะที่มีกำหนดไว้ ตอบสนองความต้องการที่ระบุไว้ชัดเจน

2.6.2 ความน่าเชื่อถือ (Reliability)

ซอฟต์แวร์สามารถทำงานได้อย่างสมบูรณ์ ลูกค้าสามารถใช้งานได้อย่างสบายใจ โดยถ้าหากซอฟต์แวร์ที่มีการพัฒนาและผ่านการใช้งานมากเท่าไร ซอฟต์แวร์ดังกล่าวนั้นจะผ่านการปรับปรุงแก้ไขให้สมบูรณ์มากขึ้นเท่านั้น

2.6.3 ความสามารถในการใช้งาน (Usability)

ซอฟต์แวร์จะต้องสะดวกและง่ายต่อการใช้งาน สามารถเสริมสร้างการเรียนรู้ได้รวดเร็ว

2.6.4 ประสิทธิภาพ (Efficiency)

ซอฟต์แวร์จะต้องก่อให้เกิดประสิทธิภาพในการใช้งาน โดยจะต้องใช้ทรัพยากรต่าง ๆ ได้อย่างคุ้มค่า และเหมาะสมในระดับที่ไม่เกินขีดความสามารถของทรัพยากร

2.6.5 ความสามารถในการบำรุงรักษา (Maintainability)

ซอฟต์แวร์จะต้องง่ายต่อการบำรุงรักษา สามารถเปลี่ยนแปลง ปรับเปลี่ยนให้เหมาะสมและตอบสนอง ได้อย่างรวดเร็วและทันท่วงที

2.6.6 ความสามารถในการโอนย้ายระบบ (Portability)

สามารถโอนย้ายระบบตามเทคโนโลยีใหม่ได้ โดยมีความสอดคล้องกับมาตรฐาน เป็นคุณสมบัติที่สำคัญเมื่อจำเป็นต้องโอนย้ายระบบ

2.7 งานวิจัยที่เกี่ยวข้อง

กฤษณา [6] ได้จัดทำงานวิจัยเรื่องการพัฒนาแนวทางการทดสอบซอฟต์แวร์ด้วยเทคนิคแบล็กบ็อกซ์เทสติ้งสำหรับระบบลงทะเบียนสมัครสอบ โดยได้สร้างกรณีทดสอบสำหรับระบบและได้นำมาผ่านการรองรับตามมาตรฐานการประกันคุณภาพซอฟต์แวร์ ISO/IEC 9126

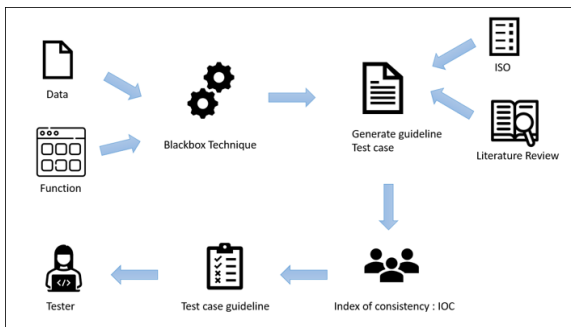
สกรณ [7] ได้จัดทำงานวิจัยเกี่ยวกับการสร้างกรณีทดสอบสำหรับการทดสอบระดับหน่วยแบบเพิ่มทีละหน่วย โดยอัตโนมัติจากกรณีทดสอบระดับหน่วย ผลลัพธ์ที่ได้จากการวัดประสิทธิภาพโดยการทดสอบ ทำให้เห็นว่าเครื่องมือดังกล่าวสามารถแจกแจงโครงสร้างความสัมพันธ์ระหว่าง

โมดูลและสร้างกรณีทดสอบได้ถูกต้อง ซึ่งเวลาที่ใช้สร้างน้อยมากเมื่อเทียบกับการสร้างกรณีทดสอบแบบดั้งเดิม

สุชาติ [8] ได้นำเสนองานวิจัยเกี่ยวกับเรื่องการสร้างกรณีทดสอบสำหรับโปรแกรมประยุกต์บนเว็บด้วยเทคนิคการทดสอบแบบแบล็กบ็อกซ์ ซึ่งได้มีการนำเสนอวิธีการสร้างกรณีทดสอบและพบว่าเครื่องมือสามารถสร้างกรณีทดสอบจากแฟ้มเอกสารเอชทีเอ็มแอลและเอ็กซ์เอ็มแอลสคีมาออกมาได้อย่างถูกต้องตรงตามหลักการออกแบบกรณีทดสอบ

3. วิธีการดำเนินงาน

งานวิจัยนี้เป็นงานวิจัยเชิงประยุกต์ โดยงานวิจัยมีกรอบแนวคิดของงานวิจัยได้ดังภาพที่ 1



ภาพที่ 2: กรอบแนวคิดของงานวิจัย

ผู้วิจัยได้แบ่งขั้นตอนในการดำเนินงานวิจัยออกเป็นขั้นตอน ดังนี้

3.1 ศึกษาและค้นคว้างานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ทำการศึกษา ค้นคว้า และเก็บรวบรวมข้อมูลจากงานวิจัยที่เกี่ยวข้องกับการสร้างแนวทางกรณีทดสอบ โดยได้มีการนำข้อมูลฟังก์ชันมาเปรียบเทียบเพื่อเป็นแนวทางในการสร้างกรณีทดสอบ โดยสามารถสรุปการเปรียบเทียบฟังก์ชันได้ดังตารางที่ 1

ตารางที่ 1 : ตารางการเปรียบเทียบงานวิจัยที่เกี่ยวข้องกับการสร้างกรณีทดสอบ

ฟังก์ชันมาตรฐาน	เรณุกา บุตรดา, หทัยชนก กุพันธ์นา [1]	กฤษณา [6]	สริตา [9]	สุชาติ [8]
การเข้าใช้งานระบบ	✓	✓	✓	

การลงทะเบียน		✓		✓
การตั้งค่าการใช้งาน		✓		
การจัดการข้อมูลส่วนตัว	✓	✓		
การจัดการรายการ	✓	✓	✓	
การค้นหาข้อมูลและการจัดลำดับ	✓	✓	✓	
การแจ้งเตือน	✓			
รายงาน	✓	✓	✓	✓

3.2 วิเคราะห์ข้อมูล

เมื่อได้ฟังก์ชันมาตรฐานจากการเปรียบเทียบงานวิจัยที่มีการสร้างแนวทางทดสอบมากที่สุด พร้อมทั้งมีการเก็บรวบรวมข้อมูลความต้องการลูกค้าที่ต้องการพัฒนาระบบมาวิเคราะห์เพื่อสร้างกรณีทดสอบ ซึ่งโดยส่วนใหญ่ไปในทิศทางเดียวกัน ฟังก์ชันมาตรฐานที่ได้จะประกอบด้วย 5 ฟังก์ชัน ได้แก่ การเข้าใช้งานระบบ, การจัดการรายการ, รายงาน, การค้นหา และการจัดลำดับ

3.3 ออกแบบและพัฒนาแนวทางทดสอบ

เมื่อได้มีการวิเคราะห์ข้อมูลจากความต้องการลูกค้าและฟังก์ชันมาตรฐานจากงานวิจัย ผู้วิจัยได้มีการสร้างกรณีทดสอบในรูปแบบของ Manual Testing หรือการทดสอบที่ดำเนินการโดยไม่ได้ใช้เครื่องมืออัตโนมัติ (Automated Tool) หรือสคริปต์ (Script) โดยผู้ทดสอบจะทำการสร้างกรณีทดสอบในแต่ละฟังก์ชันให้เป็นไปตามมาตรฐานการประกันคุณภาพซอฟต์แวร์ โดยการพัฒนาแนวทางทดสอบ (Guideline Testcase) โดยใช้เทคนิค Blackbox Testing ได้กรณีทดสอบทั้งหมด 55 ข้อ และมีหัวข้อแนวทางทดสอบคือ เลขที่ ฟังก์ชัน กรณีทดสอบ (Testcase) ลำดับขั้นตอนการทดสอบ (Testcase Step) การป้อนข้อมูลและผลคาดว่าจะได้รับ (Expected Result)

3.4 ประเมินผล

ผู้วิจัยได้นำกรณีทดสอบ 55 ข้อ จาก 5 ฟังก์ชันที่ได้มีการพัฒนามาผ่านการหาความเที่ยงตรงตามเนื้อหาจากผู้เชี่ยวชาญหรือ IOC และ นำมาให้ผู้ใช้งานทดสอบกับ

ระบบเพื่อการใช้งานจริง พร้อมทั้งประเมินความเหมาะสมของแนวทางการทดสอบระบบดังกล่าว โดยใช้เทคนิค Linkert Scale แบ่งระดับความเหมาะสมเป็น 5 ระดับ

เมื่อได้ผลประเมินความเหมาะสมจากผู้ใช้งาน จะทำการวิเคราะห์หาค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐานและค่าความเชื่อมั่นเป็นรายชื่อ และแปลผล

4. ผลการดำเนินงาน

4.1 ผลประเมินหาความเที่ยงตรง IOC

แนวทางการทดสอบระบบ โดยใช้เทคนิค Blackbox Testing สำหรับฟังก์ชันมาตรฐาน ได้ผ่านการหาค่าความเที่ยงตรงตามเนื้อหาจากผู้เชี่ยวชาญหรือค่า IOC

ผลการวิเคราะห์หาความเที่ยงตรงตามเนื้อหา ปรากฏว่าแนวทางการทดสอบระบบทุกข้อผ่านเกณฑ์มีค่าดัชนีความสอดคล้อง ตั้งแต่ .50 ขึ้นไป

ผลจากการวิเคราะห์หาค่าความเชื่อมั่น (Reliability) แนวทางการทดสอบระบบมีค่าเท่ากับ .963 ซึ่งมีค่ามากกว่า .60 เพราะฉะนั้นแนวทางการทดสอบระบบอยู่ในเกณฑ์ที่สามารถนำไปใช้งานได้จริง

4.2 ผลประเมินความเหมาะสมจากผู้ใช้งาน

ผลลัพธ์ที่ได้จากการประเมิน จำนวนผู้ใช้งาน 10 คน โดยนำกรณีทดสอบทำการทดสอบระบบและใช้งานจริง ผู้วิจัยได้ทำการวิเคราะห์หาค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐานทั้งหมด 55 ข้อ โดยแบ่งตามฟังก์ชัน และมีรายละเอียดดังตารางที่ 2

ตารางที่ 2: ผลการประเมินความเหมาะสม

ฟังก์ชัน	จำนวนข้อ	\bar{X}	S.D.	ระดับ
1.การเข้าใช้งานระบบ	9	4.99	0.11	เหมาะสมมากที่สุด
2.การจัดการรายการ	14	4.93	0.26	เหมาะสมมากที่สุด
3.รายงาน	7	4.94	0.23	เหมาะสมมากที่สุด
4.การค้นหา	4	4.78	0.47	เหมาะสมมากที่สุด
5.การจัดลำดับ	21	4.96	0.21	เหมาะสมมากที่สุด
รวม		4.94	0.25	

5. สรุป

งานวิจัยนี้ผู้วิจัยได้มีวัตถุประสงค์เพื่อพัฒนาแนวทางการทดสอบระบบ โดยใช้เทคนิค Blackbox Testing สำหรับฟังก์ชันมาตรฐาน โดยนำมาวิเคราะห์หาความเที่ยงตรงตามเนื้อหาจากผู้เชี่ยวชาญและความเหมาะสมจากผู้ใช้งาน จากผลที่ได้รับกรณีทดสอบดังกล่าวมีความเหมาะสมในระดับที่เหมาะสมมาก และสามารถนำไปใช้งานได้จริง ถูกต้องตรงตามความต้องการผู้ใช้งานและมาตรฐานการประกันคุณภาพซึ่งมีความสอดคล้องกับงานวิจัย [6] ที่มีทิศทางเดียวกันและได้ตรงตามมาตรฐานการประกันคุณภาพซอฟต์แวร์และงานวิจัย [8] กรณีทดสอบสามารถทดสอบได้อย่างถูกต้องตรงตามหลักการออกแบบกรณีทดสอบ ซึ่งผู้ทดสอบไม่จำเป็นต้องสร้างกรณีทดสอบเอง จึงเป็นผลให้ค่าใช้จ่ายแรงงาน และเวลาในการสร้างกรณีทดสอบ รวมทั้งเวลาที่ใช้ในการทดสอบระบบซอฟต์แวร์ลดลง

เอกสารอ้างอิง

- [1] เรณูภา บุตรคา และ หทัยชนก ภูพานา, “ระบบรายงานข้อผิดพลาดซอฟต์แวร์,” *โครงการงานวิทยาการคอมพิวเตอร์* พ.ศ. 2564.
- [2] อุไร ทองหัวไผ่ “การทดสอบซอฟต์แวร์,” *วารสารเกษมบัณฑิต* ปีที่ 16 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2558 หน้า 145-153.
- [3] ภักศุก ธีรานูวรรตน์ “การพัฒนาแบบสารสนเทศผลิตภัณฑ์เพื่อบริการด้านการขาย กรณีศึกษา บริษัท แครเรียร์ (ประเทศไทย) จำกัด” *วิทยานิพนธ์ ศิลปศาสตรมหาบัณฑิต มหาวิทยาลัยสุโขทัยธรรมาธิราช* พ.ศ. 2546.
- [4] นันทินี ช่วยชู “การสร้างกรณีทดสอบโดยใช้ขั้นตอนวิธีเชิงพันธุกรรม” *วิทยานิพนธ์ วิทยาศาสตร์มหาบัณฑิต สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยสงขลานครินทร์* พ.ศ. 2558.
- [5] อภาพร พิมสาร “การประเมินผลการนำโปรแกรมแคส มาใช้ในบริษัท นิคโมเตอร์ จำกัด” *การค้นคว้าอิสระ บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่* พ.ศ. 2557.
- [6] กฤษณา นิมา, “การพัฒนาแนวทางการทดสอบซอฟต์แวร์ด้วยเทคนิคแบล็คบ็อกซ์เทสตั้งสำหรับระบบลงทะเบียนสมัครสอบ” *ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสารสนเทศเพื่อการจัดการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ* พ.ศ. 2556.

- [7] สกรณ์ นุชบง “การสร้างกรณีทดสอบสำหรับการทดสอบระดับหน่วยแบบเพิ่มทีละหน่วยโดยอัตโนมัติจากกรณีทดสอบระดับหน่วย” วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต มหาวิทยาลัยเทคโนโลยีสุรนารี พ.ศ. 2556.
- [8] สุชาดา ศุภผล “การสร้างกรณีทดสอบสำหรับโปรแกรมประยุกต์บนเว็บด้วยเทคนิคการทดสอบแบบแบล็กบ็อกซ์” วิทยานิพนธ์ วิทยาศาสตรมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2547.
- [9] สรिता คุปตยานนท์ “แนวทางการทดสอบเรสต์ฟูลเว็บเซอร์วิสในการพัฒนาแบบอไจล์” วิทยานิพนธ์ วิทยาศาสตรมหาบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2556

**ระบบการเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพด้วย
กลยุทธ์ลงทุนแบบถัวเฉลี่ยต้นทุนและกลยุทธ์การลงทุนครั้งเดียวต่อปีโดยใช้ธุรกิจอัจฉริยะ
Retirement Mutual Fund Investment Performance Comparison System
with Dollar Cost Averaging Strategy and Lump Sum Strategy
using Business Intelligence**

อภิชาติ เกียรติยุทธชาติ (Apichart Kiatyudthachart)¹ และณัฐวี อุตกฤษณ์ (Nattavee Utakrit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856119@email.kmutnb.ac.th, ²nattavee.u@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อแสดงผลการเปรียบเทียบประสิทธิภาพการลงทุนระหว่างกลยุทธ์ถัวเฉลี่ยการลงทุน (Dollar Cost Averaging) และกลยุทธ์การลงทุนแบบเงินก้อน (Lump Sum) ในการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพ จึงได้ทำการพัฒนาระบบเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพด้วยกลยุทธ์ลงทุนแบบถัวเฉลี่ยต้นทุนและกลยุทธ์การลงทุนครั้งเดียวต่อปีโดยใช้ธุรกิจอัจฉริยะ ซึ่งแบ่งออกเป็น 6 ขั้นตอน ได้แก่ การทำความเข้าใจและกำหนดปัญหา การเก็บรวบรวมข้อมูล การออกแบบโครงสร้างข้อมูล การเตรียมข้อมูล การพัฒนาระบบการเปรียบเทียบและการออกแบบหน้าจอแสดงผลกระบวนการทดสอบเพื่อประเมินประสิทธิภาพของระบบ โดยงานวิจัยประยุกต์ใช้ธุรกิจผลลัพธ์ที่ได้จะแสดงผลมุมมองของการวิเคราะห์และนำเสนอในรูปแบบแดชบอร์ด ที่สามารถสื่อสารให้ผู้ใช้งานเข้าใจง่าย ใช้ได้กับทุกอุปกรณ์ สามารถแสดงผลในเงื่อนไขต่าง ๆ ได้หลายมิติ เป็นประโยชน์ในการคัดเลือกกองทุนรวมเพื่อการเลี้ยงชีพที่มีประสิทธิภาพการบริหารและกลยุทธ์การลงทุนที่ดีส่งผลต่อการเพิ่มขึ้นของเงินออมและเงินในยามเกษียณ บรรลุวัตถุประสงค์การลงทุน รวมถึงเป็นแนวทางในการใช้ข้อมูลเพื่อให้เกิดประโยชน์กับประชาชนวงกว้างให้มากที่สุด ด้านการประเมินประสิทธิภาพของระบบโดยผู้เชี่ยวชาญพบว่าระบบมีความถูกต้องร้อยละ 100

คำสำคัญ: กองทุนรวมเพื่อการเลี้ยงชีพ กลยุทธ์การลงทุน
ธุรกิจอัจฉริยะ

Abstract

The purpose of this research is to compare the efficiency of investment strategy between Dollar Cost Averaging : DCA and Lump Sum : LS in Retirement Mutual Fund Investment Therefore, we have developed a system to compare investment efficiency in Retirement Mutual Fund Investment Performance Comparison System with Dollar Cost Averaging Strategy and Lump Sum Strategy using Business Intelligence. It is divided into 6 steps: Business Understanding Data Acquisition, Schema design, Data Preparation, System Development and Display Design. This research also uses Business Intelligence. The results will show the view of analytics in the form a Dashboard that can be communicated to users for easy understanding can be displayed in many formats Able to show fund comparison results for the benefit of showing results in various conditions in many dimensions. It is beneficial for investors to select Retirement Mutual Funds with good management performance and the best investment strategy In order to affect the increase of savings and money to be used in retirement, Including being a guideline for using information to maximize the benefit of investors and the public. The evaluation of the system performance

by experts found that the system was 100 percent accurate.

Keywords: Retirement Mutual Funds, Investment Strategy, Business Intelligence.

1. บทนำ

ในปัจจุบันการลงทุนในกองทุนรวม [1] ถือเป็นทางเลือกหนึ่งที่มีความนิยมเป็นอย่างมาก เนื่องจากการลงทุนที่ให้ผลตอบแทนที่ดีในความเสี่ยงระดับหนึ่ง มีผู้จัดการกองทุนที่มีความรู้ความสามารถทำการลงทุนผ่านการวิเคราะห์และบริหารความเสี่ยงอย่างมืออาชีพ ณ มิถุนายน 2565 ประเทศไทยมีบริษัทหลักทรัพย์จัดการกองทุน ทำหน้าที่จัดการกองทุนรวม 23 แห่ง มีกองทุนรวมทั้งหมด 1,811 กองทุน มีมูลค่าสินทรัพย์สุทธิจำนวนภายใต้การบริหารจัดการ 3.96 ล้านล้านบาท มีบุคคลที่ทำหน้าที่บริหารกองทุนหรือที่เรียกว่าผู้จัดการกองทุนจำนวน 878 คน [2] กองทุนรวมเพื่อการเลี้ยงชีพเป็นกองทุนรวมรูปแบบหนึ่งที่รัฐส่งเสริมให้มีขึ้น โดยมีวัตถุประสงค์เพื่อเป็นหลักประกันแก่ตนเองหรือครอบครัวเมื่อเกษียณอายุ โดยมีข้อกำหนดพิเศษ คือ ผู้ลงทุนที่ได้รับสิทธิประโยชน์ทางภาษีจะไถ่ถอนเงินลงทุนได้เมื่อผู้ลงทุนอายุไม่น้อยกว่า 55 ปี และมีการลงทุนมาแล้วไม่น้อยกว่า 5 ปี

การลงทุนในกองทุนรวมมีกลยุทธ์การลงทุนอยู่ 2 แบบด้วยกัน คือ กลยุทธ์ถัวเฉลี่ยการลงทุน (Dollar Cost Averaging : DCA) และ กลยุทธ์การลงทุนแบบเงินก้อน (Lump Sum : LS) [3] จากผลการศึกษางานวิจัยที่เกี่ยวข้อง [4]-[6] มีการใช้ตัวอย่างข้อมูลที่แตกต่างกันไปแต่ละช่วงเวลา ผู้วิจัยจึงมีแนวคิดที่จะประยุกต์ใช้ธุรกิจอัจฉริยะเพื่อจัดทำระบบเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพของกลยุทธ์ DCA และ LS เพื่อศึกษากลยุทธ์การลงทุนที่เหมาะสมในการลงทุนกองทุนรวมเพื่อการเลี้ยงชีพในแต่ละกองทุน รวมทั้งมีการวัดประสิทธิภาพการลงทุนซึ่งจะใช้มาตรวัด Sharpe Ratio [8] เพื่อเป็นประโยชน์ต่อผู้ลงทุนในการคัดเลือกกองทุนรวมเพื่อการเลี้ยงชีพที่มีประสิทธิภาพการบริหารที่ดีและคัดเลือกกลยุทธ์การลงทุนที่ดีที่สุดได้ เพื่อให้ส่งผลต่อการเพิ่มขึ้นของเงินออมและเงินที่จะไว้ใช้ในยามเกษียณซึ่งจะ

บรรลุวัตถุประสงค์การลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพ รวมถึงเป็นแนวทางในการใช้ข้อมูลเพื่อให้เกิดประโยชน์กับผู้ลงทุนและประชาชนวงกว้างให้มากที่สุด

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การลงทุนด้วยกองทุนรวม [1]

คือการระดมเงินลงทุนจากคนจำนวนมากและนำไปจดทะเบียนให้มีฐานะเป็นนิติบุคคล เพื่อตั้งเป็น กองทุนขึ้นมา โดยเงินที่ได้รับนั้นจะมีผู้จัดการกองทุน ที่เป็นมืออาชีพ นำเงินไปลงทุนในสินทรัพย์ต่าง ๆ ตามนโยบายของแต่ละกองทุน ให้ได้รับผลตอบแทนที่คงอยู่แล้วนำมาเฉลี่ยคืนให้กับผู้ลงทุนแต่ละรายตามสัดส่วนที่ลงทุน

2.2 กลยุทธ์การลงทุนด้วยเงินจำนวนเดียวทั้งหมด (Lump Sum : LS) [3]

เป็นกลยุทธ์การลงทุนที่นำเงินลงทุนทั้งจำนวนซื้อหลักทรัพย์ในครั้งเดียวทั้งก้อน ซึ่งแนวทางนี้เหมาะกับนักลงทุนที่มีประสบการณ์และสามารถคาดการณ์แนวโน้มตลาดที่เกิดขึ้นด้วยการจับจังหวะการลงทุนเพื่อซื้อหลักทรัพย์ในขณะที่มีราคาต่ำทำให้ได้รับผลตอบแทนสูง หากสามารถคาดการณ์ทิศทางและแนวโน้มตลาดได้อย่างแม่นยำ แต่ในทางกลับกันหากลงทุนผิดจังหวะอาจสร้างความเสียหายต่อพอร์ตการลงทุนเป็นอย่างมากและทำให้นักลงทุนสูญเสียโอกาสในการลงทุน

2.3 กลยุทธ์การลงทุนแบบถัวเฉลี่ยต้นทุน (Dollar - cost Averaging Investment : DCA) [3]

เป็นการลงทุนในจำนวนที่เท่า ๆ กันในทุกช่วงระยะเวลา หัวใจสำคัญคือการทยอยลงทุนเป็นงวด ๆ ด้วยจำนวนเงินที่เท่า ๆ กันทุกงวด โดยไม่สนใจราคาของหลักทรัพย์ว่าจะขึ้นหรือลงมากเพียงใด เพื่อขจัดอารมณ์ที่ใช้ในการลงทุนและสร้างวินัยการลงทุนที่มีประสิทธิภาพ

2.4 การวัดประสิทธิภาพการบริหารพอร์ตการลงทุน ผลตอบแทนและความเสี่ยงด้วย Sharpe Ratio [7]

ศาสตราจารย์ William F. Sharpe ผู้คิดค้น จากผลตอบแทนที่คาดหวัง (Expected Return) จากการลงทุนลบด้วยผลตอบแทนที่ปราศจากความเสี่ยง (Risk-Free Rate) และนำไปเทียบต่อหน่วยกับค่าเบี่ยงเบนมาตรฐาน

(Standard Deviation) หาก Sharpe ratio มีค่าที่สูงหรือเป็นบวกแสดงว่าผลการดำเนินงานอยู่ในเกณฑ์ที่ดีความเสี่ยงต่ำ และเมื่อความเสี่ยงจากการวัดผลดำเนินงานต่ำหรือลบแสดงว่าผลการดำเนินงานอยู่ในเกณฑ์ที่ไม่ดี

2.5 ธุรกิจอัจฉริยะ (Business intelligence : BI)

คือ แนวคิดและเทคโนโลยีสำหรับการสร้างระบบบริหารจัดการข้อมูล การจัดเก็บข้อมูลหลายมิติ และการนำเสนอข้อมูลในรูปแบบรายงาน ตารางและแผนภูมิ ซึ่งสามารถวิเคราะห์ข้อมูลในระบบเชิงลึกและภาพรวม ทำให้เกิดองค์ความรู้ใหม่หรือตอบคำถามความต้องการจากฐานข้อมูลภายในองค์กร เพื่อใช้สำหรับวางแผนและสนับสนุนการตัดสินใจให้ผู้บริหารและเจ้าหน้าที่ได้อย่างรวดเร็ว [8] – [9]

2.6 ฐานข้อมูลเชิงสัมพันธ์ (Relational Database)

ฐานข้อมูลเชิงสัมพันธ์ คือ การเก็บข้อมูลในรูปแบบของตาราง (Table) หลายๆตารางที่มีความสัมพันธ์กัน ในแต่ละตารางแบ่งออกเป็นแถวๆ และในแต่ละแถวจะแบ่งเป็นคอลัมน์ (Column) ในทางทฤษฎีจะมีคำศัพท์เฉพาะแตกต่างกันออกไป เนื่องจากแบบจำลองแบบนี้เกิดจากทฤษฎีทางคณิตศาสตร์เรื่องเซต (Set)

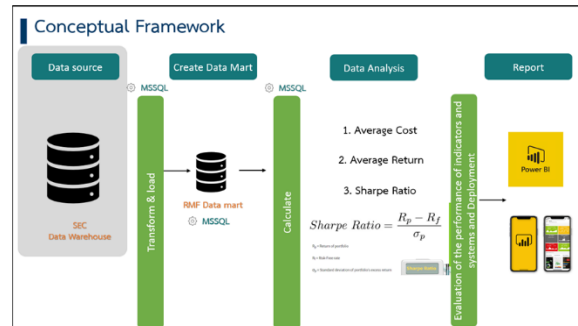
2.7 แดชบอร์ด (Dashboard)

คือการนำข้อมูลในรายงานหรือข้อมูลใหม่มาสรุปให้สามารถเห็นภาพได้ในหน้าเดียวและเป็นข้อมูลที่อัปเดตสม่ำเสมอ เพื่อให้สามารถวิเคราะห์ตัดสินใจได้ทันเวลา [10]

3. วิธีดำเนินการวิจัย

ผู้วิจัยทำการพัฒนาระบบบนพื้นฐานกระบวนการวิทยาศาสตร์ข้อมูล (Data Science) [11] โดยใช้เทคนิคการจัดการข้อมูลตามแนวคิดธุรกิจอัจฉริยะ และคำนวณมูลค่าต่างๆ ด้วยการเขียนโปรแกรมจัดการข้อมูลเพื่อให้ระบบมีประสิทธิภาพ ผู้วิจัยได้กำหนดขั้นตอนการดำเนินงาน 6 ขั้นตอน ดังนี้ การทำความเข้าใจและกำหนดปัญหา (Business Understanding) การเก็บรวบรวมข้อมูล (Data Acquisition) การออกแบบโครงสร้างข้อมูล (Schema design) การเตรียมข้อมูล (Data Preparation) การพัฒนาระบบการเปรียบเทียบและการออกแบบหน้าจอแสดงผล

(System Development and Display Design) กระบวนการทดสอบเพื่อประเมินประสิทธิภาพของระบบ(System Development and Display Design) โดยกรอบแนวคิดสามารถแสดงได้ดังภาพที่ 1



ภาพที่ 1: กรอบแนวคิดงานวิจัยของระบบการเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพด้วยกลยุทธ์ลงทุนแบบถัวเฉลี่ยต้นทุนและกลยุทธ์การลงทุนครั้งเดียวต่อปีโดยใช้ธุรกิจอัจฉริยะ

3.1 การทำความเข้าใจและกำหนดปัญหา

จากผลการศึกษางานวิจัยที่เกี่ยวข้อง พบว่าในแต่ละงานวิจัยมีการใช้กลุ่มตัวอย่างข้อมูลกองทุนรวมที่แตกต่างกัน และเลือกกลุ่มตัวอย่างเพื่อนำมาศึกษา ทำให้ผลการศึกษาที่ได้แตกต่างกัน ผู้วิจัยจึงมีแนวคิดที่จะประยุกต์ใช้ธุรกิจอัจฉริยะเพื่อจัดทำระบบเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพของกลยุทธ์ DCA และ LS โดยใช้ข้อมูลของกองทุนรวมเพื่อการเลี้ยงชีพทุกกองทุนที่จดทะเบียนอยู่ จำนวน 251 กองทุน เพื่อศึกษากลยุทธ์การลงทุนที่เหมาะสมในการลงทุนกองทุนรวมเพื่อการเลี้ยงชีพในแต่ละกองทุน รวมทั้งวัดประสิทธิภาพการลงทุนซึ่งจะใช้มาตรวัด Sharpe Ratio

3.2 การเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในการวิจัยนำมาจากคลังข้อมูล (Data Warehouse) ของสำนักงานคณะกรรมการ ก.ล.ต. ที่ทำการเก็บรวบรวมจากแบบรายงานต่าง ๆ ที่บริษัทหลักทรัพย์จัดการกองทุนมีภาระหน้าที่ต้องรายงานตามประกาศกำหนดและข้อมูลอื่น ๆ ที่เกี่ยวข้อง ซึ่งอยู่ในรูปแบบฐานข้อมูลเชิงสัมพันธ์

3.3 การออกแบบโครงสร้างข้อมูล

ออกแบบโครงสร้างข้อมูลและการแสดงผลให้ผู้ใช้งานอ่านข้อมูลและนำไปวิเคราะห์เชิงลึกได้อย่างรวดเร็วตาม

หลักการฐานข้อมูลเชิงสัมพันธ์ ด้วยการจัดทำคลังข้อมูลขนาดเล็ก (Data Mart) โดยใช้การสร้าง view สกัดและปรับเปลี่ยนข้อมูลภายใน table จากคลังข้อมูล (Data Warehouse) ด้วยโปรแกรม MSSQL ในการจัดการฐานข้อมูล การออกแบบโครงสร้างแบ่งเป็น 4 ตาราง ดังนี้

ตารางที่ 1: ข้อมูลมิติกองทุนรวมเพื่อการเลี้ยงชีพ

column name	data type	primary key	foreign key	description
fund_id	varchar (15)	/	/	รหัสกองทุน
valid_from	date	/	/	วันที่เริ่ม
valid_to	date			วันที่สิ้นสุด
regist_id	varchar (10)			เลขที่จดทะเบียนกองทุน
fund_abbr_name	varchar (30)			ชื่อย่อกองทุน
fund_name.th	varchar (255)			ชื่อกองทุนภาษาไทย
fund_name.en	varchar (255)			ชื่อกองทุนภาษาอังกฤษ
amc_id	char (10)	/ (dim_company)		รหัสบริษัทจัดการ
policy.thai.desc	varchar (255)			นโยบายภาษาไทย
policy.eng.desc	varchar (255)			นโยบายภาษาอังกฤษ
invest_country	varchar (255)			นโยบายการลงทุนต่างประเทศ
fund_key	varchar (24)			รหัสกองทุน
dw_modified_date	datetime			วันที่ปรับปรุงข้อมูล

ตารางที่ 2: ข้อมูลมิติบริษัทจัดการลงทุน

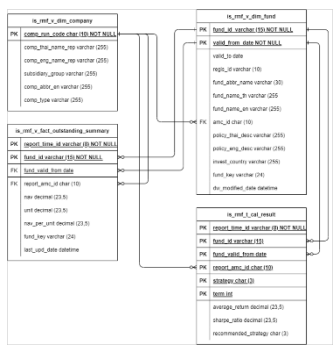
column name	data type	primary key	foreign key	description
comp_run_code	char (10)	/		รหัสบริษัทจัดการ
comp_thai_name_rep	varchar (255)			ชื่อบริษัทภาษาไทย
comp_eng_name_rep	varchar (255)			ชื่อบริษัทภาษาอังกฤษ
subsidiary_group	varchar (255)			ประเภทบริษัทจำแนกตามผลิตภัณฑ์
comp_abbr.en	varchar (255)			ชื่อย่อบริษัท
comp_type	varchar (255)			ประเภทบริษัท

ตารางที่ 3: ข้อมูลแสดงมูลค่าสินทรัพย์สุทธิ ตัวเลขทางการเงินของกองทุนรวมเพื่อการเลี้ยงชีพ

column name	data type	primary key	foreign key	description
report_time_id	varchar (8)	/	/	date format YYYYMMDD
fund_id	varchar (15)	/	/ (dim_fund)	รหัสกองทุน
fund_valid_from	date	/	/ (dim_fund)	วันที่เริ่ม
report_amc_id	char (10)	/	/ (dim_company)	รหัสบริษัทจัดการ
nav	decimal (23,5)			มูลค่าสินทรัพย์สุทธิ
unit	decimal (23,5)			จำนวนหน่วยลงทุน
nav_per_unit	decimal (23,5)			มูลค่าสินทรัพย์สุทธิต่อหน่วย
fund_key	varchar (24)			รหัสกองทุน
last_upd_date	datetime			วันที่ปรับปรุงข้อมูล

ตารางที่ 4: ตารางแสดงผลลัพธ์จากการคำนวณ

column name	data type	primary key	foreign key	description
report_time_id	varchar (8)	/	/ (fact_outstanding_summary)	date format YYYYMMDD
fund_id	varchar (15)	/	/ (fact_outstanding_summary, dim_fund)	รหัสกองทุน
fund_valid_from	date	/	/ (dim_fund)	วันที่เริ่ม
report_amc_id	char (10)	/	/ (dim_company)	รหัสบริษัทจัดการ
strategy	char (3)	/		IS / DCA
beta	int	/		จำนวนปีการลงทุน
average_cost	decimal (23,5)			ต้นทุนเฉลี่ย
average_return	decimal (23,5)			ผลตอบแทนเฉลี่ย
sharpe_ratio	decimal (23,5)			มาตราวัด Sharpe Ratio
recommended_strategy	char (3)			กลยุทธ์ที่แนะนำ



ภาพที่ 2: แสดงแบบจำลองที่ใช้อธิบายโครงสร้างของฐานข้อมูล

3.4 การเตรียมข้อมูล

ในงานวิจัยจะต้องศึกษาและเตรียมข้อมูลที่ต้องการจากคลังข้อมูลเพื่อตอบโจทย์การจัดทำระบบ จากนั้นนำข้อมูลเข้าสู่คลังข้อมูลขนาดเล็ก (Data Mart) ที่ได้จัดเตรียมรวมทั้งทำความเข้าใจข้อมูลเพื่อให้พร้อมใช้ในการคำนวณต่อไป

3.5 การพัฒนาระบบการเปรียบเทียบและการออกแบบหน้าจอแสดงผล

พัฒนาระบบในรูปแบบ Stored Procedures ด้วยโปรแกรมจัดการข้อมูล MSSQL ใช้ภาษา SQL ในการจัดการฐานข้อมูล เพื่อคำนวณมูลค่าต้นทุนเฉลี่ย อัตราผลตอบแทน และประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยงรายกองทุน โดยมีการคำนวณ ดังนี้

3.5.1 การคำนวณต้นทุนต่อหน่วยลงทุนเฉลี่ย

3.5.1.1 กลยุทธ์ DCA กำหนดให้ซื้อกองทุน ณ วันทำการสุดท้ายของเดือน เดือนละ 1,000 บาท จากนั้นนำมาเฉลี่ยรายปีในแต่ละกองทุน

3.5.1.2 กลยุทธ์ LS กำหนดให้ซื้อกองทุน ณ วันทำการสุดท้ายของปี ดังนั้นต้นทุนเฉลี่ยในแต่ละปี จะเท่ากับมูลค่าต่อหน่วยลงทุนของกองทุนนั้น ๆ

3.5.2 การคำนวณอัตราผลตอบแทน

เป็นการคำนวณผลตอบแทนเฉลี่ยต่อปี (Annualized Return) แบบปัดทศนิยม สิ้นปี 1 ปี 3 ปี 5 ปี 7 ปี และ 10 ปี โดยมีสูตรการคำนวณ ดังนี้

$$\{(1 + \text{ผลตอบแทนรวม})^{\frac{1}{\text{จำนวนปีที่ลงทุน}}}\} - 1$$

3.5.3 การคำนวณประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยง

ใช้มาตราวัด Sharpe Ratio

$$\text{Sharpe Ratio} = (\text{Rate of Return} - \text{Risk-Free Rate}) / \text{Standard Deviation}$$

3.5.4 การออกแบบหน้าจอแสดงผล

ผลลัพธ์ที่ได้จะแสดงผลมุมมองของการวิเคราะห์ในรูปแบบข้อมูลจินตภาพ (Data Visualization) และนำเสนอในรูปแบบแดชบอร์ด (Dashboard) แสดงผลผ่านโปรแกรม PowerBI

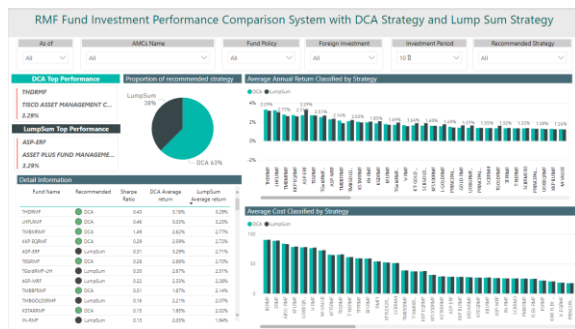
3.6 กระบวนการทดสอบเพื่อประเมินประสิทธิภาพของระบบ

ในการประเมินประสิทธิภาพของระบบจะจัดให้มีผู้เชี่ยวชาญ 3 ท่านประเมินประสิทธิภาพ โดยการตรวจสอบความถูกต้องของข้อมูล

กระบวนการประเมินประสิทธิภาพกำหนดให้ผู้ตรวจสอบความถูกต้องของข้อมูลท่านละ 5 รายการเปรียบเทียบผลการคำนวณต้นทุนต่อหน่วยลงทุนเฉลี่ย อัตราผลตอบแทนเฉลี่ยต่อปี และ Sharpe Ratio ใช้แหล่งข้อมูลจากคลังข้อมูล คำนวณโดยใช้เครื่องมือ Microsoft excel ใช้แหล่งข้อมูลจากคลังข้อมูล เพื่อเปรียบเทียบความถูกต้องของผลลัพธ์ที่ได้จากระบบ

4. ผลการดำเนินงานวิจัย

ผลการดำเนินการวิจัยสามารถแสดงตัวอย่างผลการพัฒนาระบบการเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพด้วยกลยุทธ์ลงทุนแบบถัวเฉลี่ยต้นทุนและกลยุทธ์การลงทุนครั้งเดียวต่อปีโดยใช้ธุรกิจอัจฉริยะ ในรูปแบบรายงานอัตโนมัติและเชื่อมโยงกัน (Automated and Interactive Dashboard) แสดงได้ดังภาพที่ 3 - 8



ภาพที่ 3: แสดงหน้าจอแสดงภาพรวมของรายงาน

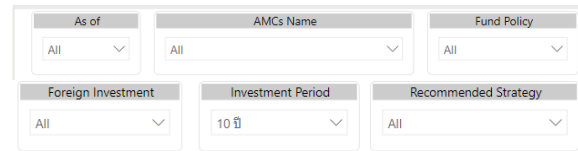
ตัวอย่างหน้าจอแสดงภาพรวมของระบบการเปรียบเทียบกลยุทธ์การลงทุน สามารถแสดงได้ดังภาพที่ 6

จากผลการพัฒนาระบบได้นำข้อมูลการเปรียบเทียบกลยุทธ์การลงทุน และประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยง นำเสนอ 5 ส่วน ดังนี้

1) ส่วนการเลือกเงื่อนไข

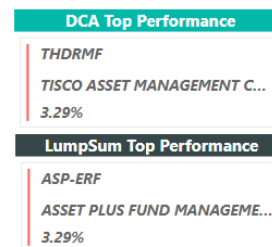
ผู้ใช้งานสามารถเลือกเงื่อนไขของการเปรียบเทียบเพื่อแสดงผลได้จากตัวเลือกเงื่อนไข ได้แก่ ข้อมูล ณ วันที่ซื้อบริษัทหลักทรัพย์จัดการกองทุน ชื่อกองทุนรวมเพื่อการ

เลี้ยงชีพ นโยบายการลงทุน ความเสี่ยงในการลงทุนต่างประเทศ ระยะเวลาย้อนหลังในการคำนวณ และหากผู้ใช้งานมีกลยุทธ์ที่ต้องการลงทุนแล้ว สามารถเลือกตัวเลือกเงื่อนไขกลยุทธ์ที่แนะนำ เพื่อแสดงผลเฉพาะกองทุนที่กลยุทธ์ของผู้ใช้งานที่มีประสิทธิภาพดีกว่า แสดงได้ดังภาพที่ภาพที่ 4



ภาพที่ 4: แสดงตัวเลือกเงื่อนไขของรายงาน

2) ส่วนสรุปข้อมูลกองทุนรวมเพื่อการเลี้ยงชีพที่มีผลการดำเนินงานย้อนหลังดีที่สุดในแต่ละกลยุทธ์การลงทุน แสดงได้ดังภาพที่ 5



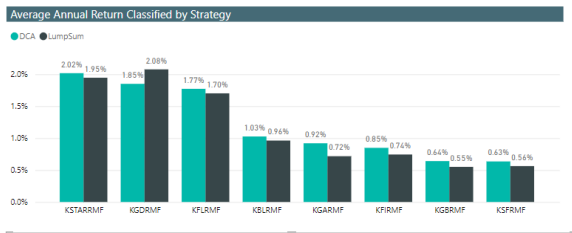
ภาพที่ 5: แสดงผลกองทุนรวมเพื่อการเลี้ยงชีพที่มีผลการดำเนินงานย้อนหลังดีที่สุดในแต่ละกลยุทธ์การลงทุน

3) ส่วนแสดงรายละเอียดข้อมูลการเปรียบเทียบกลยุทธ์การลงทุน ผลตอบแทนเฉลี่ยย้อนหลัง และประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยงรายกองทุน แสดงได้ดังภาพที่ภาพที่ 6

Fund Name	Recommended	Sharpe Ratio	DCA Average return	LumpSum Average return
THDRMF	DCA	0.40	3.16%	3.29%
LHFLRMF	DCA	0.46	3.03%	3.23%
TMBMRMF	DCA	1.49	2.62%	2.77%
KKP EQRMF	DCA	0.29	2.59%	2.72%
ASP-ERF	LumpSum	0.31	3.29%	2.71%
TEGRMF	DCA	0.28	2.68%	2.70%
TGoldRMF-UH	LumpSum	0.30	2.67%	2.51%
ASP-MRF	LumpSum	0.22	2.33%	2.28%
TMBBFRMF	DCA	0.51	1.87%	2.14%
TMBGOLDSRMF	LumpSum	0.16	2.21%	2.07%
KSTARRMF	DCA	0.15	1.95%	2.02%
IN-RMF	LumpSum	0.13	2.03%	1.94%

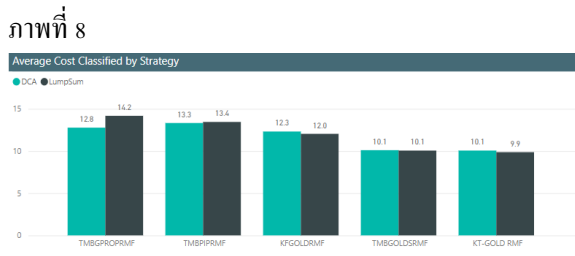
ภาพที่ 6: แสดงรายละเอียดข้อมูลการเปรียบเทียบกลยุทธ์การลงทุน ผลตอบแทนเฉลี่ยย้อนหลัง และประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยงรายกองทุน

4) ส่วนแสดงการเปรียบเทียบผลตอบแทนเฉลี่ยย้อนหลัง แสดงได้ดังภาพที่ 7



ภาพที่ 7: แสดงการเปรียบเทียบผลตอบแทนเฉลี่ยย้อนหลัง

5) ส่วนแสดงการเปรียบเทียบต้นทุนเฉลี่ย แสดงได้ดัง



ภาพที่ 8: แสดงการเปรียบเทียบต้นทุนเฉลี่ย

5. สรุปและอภิปรายผล

จากการพัฒนาระบบการเปรียบเทียบประสิทธิภาพการลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพด้วยกลยุทธ์ลงทุนแบบถัวเฉลี่ยต้นทุนและกลยุทธ์การลงทุนครั้งเดียวต่อปี โดยใช้ธุรกิจอัจฉริยะ โดยใช้ข้อมูลกองทุนรวมเพื่อการเลี้ยงชีพที่จดทะเบียนอยู่ ประยุกต์ใช้แนวคิดธุรกิจอัจฉริยะที่มีการจัดการข้อมูลตั้งแต่นำเข้าถึงออกรายงานอย่างเป็นระบบและอัตโนมัติ เพื่อนำเสนอผลลัพธ์ของระบบในรูปแบบรายงานอัตโนมัติและเชื่อมโยงกัน แสดงผลในเงื่อนไขต่าง ๆ ได้หลายมิติ เป็นประโยชน์ต่อผู้ลงทุนในการคัดเลือกกองทุนรวมเพื่อการเลี้ยงชีพที่มีประสิทธิภาพการบริหารที่ดี และคัดเลือกกลยุทธ์การลงทุนที่ดีที่สุด เพื่อให้ส่งผลต่อการเพิ่มขึ้นของเงินออมและเงินที่จะไว้ใช้ในยามเกษียณซึ่งจะบรรลุวัตถุประสงค์การลงทุนในกองทุนรวมเพื่อการเลี้ยงชีพ

ผลการทดสอบประเมินประสิทธิภาพของระบบด้วยวิธีการสุ่มตรวจสอบความถูกต้องของข้อมูล โดยผู้เชี่ยวชาญพบว่า ระบบมีการคำนวณผลตอบแทนเฉลี่ยย้อนหลัง ต้นทุนเฉลี่ย และประสิทธิภาพอัตราผลตอบแทนหลังปรับด้วยความเสี่ยง ถูกต้องร้อยละ 100

สิ่งที่ผู้วิจัยค้นพบจากการวิจัยครั้งนี้คือ แนวทางการจัดการข้อมูลเพื่อการวิเคราะห์และคำนวณมูลค่าต่าง ๆ

ตลอดจนการพัฒนาเฉพาะเรื่องตามแนวคิดธุรกิจอัจฉริยะ สำหรับวางแผนและสนับสนุนการตัดสินใจให้ผู้ใช้งานตัดสินใจได้อย่างรวดเร็วสอดคล้องกับงานวิจัย[8]-[9] ซึ่งอาจสามารถนำไปปรับใช้กับการวิเคราะห์ข้อมูลในด้านอื่น ๆ ในบริบทใกล้เคียงกัน และออกแบบการแสดงผลให้เหมาะสมกับการใช้งานเหล่านั้นได้

เอกสารอ้างอิง

- [1] ธนาคารแห่งประเทศไทย “กองทุนรวม (Mutual Fund)” [สืบค้น วันที่ 5 สิงหาคม 2565]. จาก <https://www.1213.or.th/th/others/investments/Pages/mutualfund.aspx>
- [2] สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์ “สถิติตลาดทุน” [สืบค้น วันที่ 5 สิงหาคม 2565]. จาก <https://www.sec.or.th/TH/Pages/MarketData/MarketStatisticsDashboard.aspx>
- [3] Investinganswers. “What Is Dollar Cost Averaging?” [สืบค้น วันที่ 5 สิงหาคม 2565]. จาก <https://investinganswers.com/dictionary/d/dollar-cost-averaging>
- [4] อนันต์ กฤษฎาภรณ์. (2559). การศึกษากลยุทธ์การลงทุนแบบ LUMP SUM และ DOLLAR COST AVERAGE ในกองทุน LTF. การค้นคว้าอิสระ บริหารธุรกิจมหาบัณฑิต คณะพาณิชยศาสตร์และการบัญชีมหาวิทยาลัยธรรมศาสตร์.
- [5] พชร นิมิตรสุโข. (2556). กลยุทธ์การลงทุนแบบ Lump sum และ Dollar cost average ในกองทุน LTF และ RMF. การค้นคว้าแบบอิสระปริญญา บัณฑิตมหาบัณฑิต สาขาวิชาการบัญชีธุรกิจแบบบูรณาการ คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์.
- [6] จันทรา เหล่ากุลประสิทธิ์. (2559). กลยุทธ์การลงทุนแบบ Dollar Cost Average กับ Value Average ในการลงทุนหลักทรัพย์ที่ใช้คำนวณใน SETHD. การค้นคว้าอิสระ บริหารธุรกิจมหาบัณฑิต คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์.
- [7] Sharpe, W.F. (1994). The Sharpe Ratio. The Journal of Portfolio Management, 21(1), 49–58.
- [8] ธนิยา สังข์แสง. (2564). “ระบบการวิเคราะห์อูบัตินเหตุทางถนนโดยใช้ธุรกิจอัจฉริยะ.” ใน The Eighteenth National Conference on Computing and Information Technology, กรุงเทพมหานคร, ประเทศไทย, 19-20 พฤษภาคม 2565
- [9] ดลยา นามโพธิ์ชัย และผยอง มีสัง. (2563). “ระบบธุรกิจอัจฉริยะด้านการจัดการรายงานภายในองค์กร.” ใน The Seventeenth National Conference on Computing and

Information Technology, กรุงเทพมหานคร, ประเทศไทย
, 13-14 พฤษภาคม 2564

- [10] ประภาพร กุลลิมรัตน์ชัย (2018). การเข้าสู่ยุคของวิทยาศาสตร์
ข้อมูล.วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเซีย ฉบับ
วิทยาศาสตร์และเทคโนโลยี. 12(2) พฤษภาคม –สิงหาคม
2561.
- [11] Techsauce. “What Is Dashboard? [สี่ บ คั น วั น ที่ 5
สิงหาคม 2565]. จาก <https://techsauce.co/tech-and-biz/what-is-dashboard>

การพัฒนาระบบบริหารจัดการทรัพยากรโครงการ A Development of Project Resource Management

ณัฐพล พันอินากุล (Nattapol Puninagoon)¹ และณัฐวี อุตกฤษฎ์ (Nattavee Utakrit)²
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021857085@email.kmutnb.ac.th, ²nattavee.u@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบบริหารจัดการทรัพยากรโครงการ ประกอบไปด้วยระบบวางแผนจัดการพัสดุ และครุภัณฑ์ รวมไปถึงทรัพยากรบุคคล ระบบการแจ้งเตือนการร้องขอทรัพยากร และระบบการออกรายงานของข้อมูลทรัพยากรโครงการ ระบบนี้พัฒนาขึ้นโดยการประยุกต์ใช้หลักการและทฤษฎี ด้านการบริหารโครงการ ร่วมกับการออกแบบการเชื่อมต่อประสานกับผู้ใช้งาน และการออกแบบการแสดงผลให้เหมาะสมบนหน้าจอขนาดต่าง ๆ (Web Responsive) การประยุกต์ใช้บริการไลน์แจ้งเตือน (Line Notify) ในการแจ้งเตือนรายการเบิกทรัพยากรต่างๆ เพื่อช่วยให้การสื่อสารเป็นปัจจุบันและ เข้าใจกัน ได้ตรงกัน ลดความผิดพลาดในการสื่อสาร และบริหารทรัพยากรในโครงการ ได้อย่างมีประสิทธิภาพ รวมทั้งมีการออกแบบการแสดงผลหน้าจอที่สรุปข้อมูลทุกอย่างให้อยู่ในหน้าจอเดียว (Dashboard Reporting) เพื่อให้ผู้ใช้ทราบข้อมูลแล้วเข้าใจได้ทันที ในส่วนของการพัฒนาผู้วิจัยได้ใช้ภาษา HTML, CSS และ JavaScript โดยพัฒนาบน โปรแกรม Visual Studio Code ผลการทดสอบประสิทธิภาพของระบบ พบว่าสามารถทำงานได้อย่างถูกต้อง โดยมีค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ยเท่ากับ 98% และผลการประเมินความพึงพอใจของผู้ใช้งานอยู่ในระดับมากที่สุด ค่าเฉลี่ย 4.46 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.72 สรุปได้ว่าระบบที่พัฒนาขึ้นนี้บรรลุตามวัตถุประสงค์ของการวิจัย และสามารถนำไปประยุกต์ใช้ในการทำงานจริงได้

คำสำคัญ: การบริหารโครงการ การบริหารทรัพยากรโครงการ การแสดงผล บริการไลน์แจ้งเตือน

Abstract

The objective of this research is to develop a project resource management system. It consists of a planning system for managing supplies and equipment, including human resources. This system was developed by applying project management principles and theories with user interface design and display design optimized for various screen sizes. In notification of various resource picking lists to help keep communication up-to-date and understand each other, reduce errors in communication and manage project resources efficiently. Including designing a screen display that sums up everything in one screen so that users know the information and understand it immediately. In development the researcher has used HTML, CSS and JavaScript languages developed on the Visual Studio Code program. The performance testing results found that the system work properly. The average accuracy percentage was 98% and the results of the user satisfaction assessment were at a high level with an average of 4.46 and a standard deviation of 0.72. It can be concluded that this developed system achieves the objectives of research and can be usage in work.

Keywords: Project Management, Project Resources, Management, Dashboard Reporting, Line Notify.

1. บทนำ

การบริหารโครงการ เป็นการจัดการ การใช้ทรัพยากรต่าง ๆ ที่มีอยู่ให้เหมาะสม เพื่อให้การดำเนินโครงการบรรลุ

วัตถุประสงค์ของโครงการ ภายใต้ขอบเขตที่กำหนดไว้ [1] จากการศึกษาข้อมูลพบว่า มีโครงการจำนวนมากไม่น้อยที่ไม่ประสบความสำเร็จ สาเหตุสำคัญประการหนึ่งของสถานะความล้มเหลวของโครงการมีหลายปัจจัย เช่น ขาดการบริหารทรัพยากรที่มีประสิทธิภาพ โดยเฉพาะอย่างยิ่งการวางแผน และจัดการทรัพยากรในการทำโครงการซึ่งเป็นที่จำเป็น และต้องอาศัยผู้ที่มีความรู้ทักษะในการทำงานในด้านนี้เป็นอย่างมาก นอกจากนี้ยังเกิดจากการสื่อสารระหว่างบุคลากร ไม่ทั่วถึงและขาดความชัดเจนไม่เป็นปัจจุบัน และไม่เพียงพอ [2]

ปัจจุบันมีการพัฒนาซอฟต์แวร์สำหรับการจัดการโครงการไว้ให้บริการที่หลากหลายรูปแบบ เช่น แอปพลิเคชันสำหรับบริหารโครงการชื่อ Asana ที่ทีมสามารถทำงานร่วมกันได้ในรูปแบบออนไลน์ [3] และซอฟต์แวร์ Microsoft Project ซึ่งเป็นโปรแกรมบริหารจัดการโครงการติดตามงาน กำหนดทรัพยากรให้กับงานต่าง ๆ จาก Microsoft สำหรับธุรกิจเป็นที่นิยม [4] แต่ระบบบริหารจัดการโครงการที่มีอยู่ในปัจจุบันบางระบบยังมีข้อจำกัดในเรื่องการใช้งาน เพราะมีการใช้งานที่ค่อนข้างยากและซับซ้อน เหมาะกับผู้ใช้งานที่มีทักษะหรือผ่านการใช้งาน จึงจะสามารถใช้งานได้ นอกจากนี้ อีกทั้งบางระบบยังมีข้อจำกัดที่มอบหมายงานได้แค่หนึ่งคนต่อหนึ่งงาน และยังไม่มีการแจ้งเตือนที่ตอบสนองโดยทันที และบางระบบมีราคาค่อนข้างสูง

จากเหตุผลดังกล่าว ผู้วิจัยเห็นได้ว่าควรมีการศึกษาพัฒนาระบบบริหารจัดการโครงการที่สามารถใช้งานได้อย่างรวดเร็ว และไม่ซับซ้อน โดยงานวิจัยนี้ได้มุ่งเน้นไปที่ระบบบริหารจัดการทรัพยากรโครงการ เนื่องจากเป็นส่วนหนึ่งที่สำคัญ ซึ่งการระบบบริหารจัดการทรัพยากรนี้ประกอบไปด้วย ทางด้านการวางแผนจัดการทรัพยากรพัสดุ ทรัพยากรบุคคล การจัดการทรัพยากรด้านการสื่อสาร โดยประยุกต์ใช้บริการออนไลน์แจ้งเตือนเข้ามาช่วยในการแจ้งเตือนรายการเบิกจ่ายทรัพยากรต่าง ๆ เพื่อช่วยให้การสื่อสาร ด้านการแสดงผล และการออกแบบส่วนแสดงผลข้อมูลในรูปแบบ Dashboard เพื่อแสดงค่าใช้จ่าย

ในการรายการเบิกพัสดุและทรัพยากรบุคคลเพื่อให้ใช้ทรัพยากรเหล่านี้ได้อย่างมีประสิทธิภาพ

2. ทฤษฎี หลักการ และงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีการบริหารโครงการ

การบริหารโครงการ หมายถึง การวางแผน ดำเนินการ ตรวจสอบควบคุม และปิดโครงการ ของงานโครงการ ผ่านทางเครื่องมือ ความรู้ ทรัพยากร และกิจกรรมต่าง ๆ ให้บรรลุวัตถุประสงค์ของโครงการภายในระยะเวลาและข้อจำกัดที่กำหนดไว้ [5] ทรัพยากรเป็นปัจจัยหนึ่งที่สำคัญในโครงการ โดยในส่วนกระบวนการบริหารทรัพยากรโครงการนั้นมีขั้นตอนเริ่มต้นที่สำคัญที่ผู้จัดการโครงการต้องคำนึงถึง เช่น การวางแผนการบริหารจัดการทรัพยากร รวมถึงขั้นตอนการตรวจสอบทรัพยากรในระหว่างการค้าโครงการ [6]

2.2 การสื่อสารอย่างมีประสิทธิภาพ

คือการที่ผู้ส่งสาร สามารถสื่อความต้องการของตัวเองออกไปยังผู้รับสารให้รับรู้ถึงสิ่งที่กำลังสื่อสารและเข้าใจสิ่งที่สื่อออกมาได้อย่างชัดเจน ถูกต้อง ในเวลาที่เหมาะสม

2.3 ทฤษฎีการวิเคราะห์ปริมาณงานและปริมาณคน

เป็นการพยากรณ์โดยวิเคราะห์ว่าทรัพยากรมนุษย์ที่ต้องการในอนาคตขององค์กรมีจำนวนเท่าใด และมีคุณสมบัติเป็นอย่างไร โดยนำดุลยพินิจ และประสบการณ์มาใช้ ซึ่งส่วนใหญ่แล้วนำเอาหลายวิธีการมาผสมผสานกัน

2.4 ทฤษฎี Safety Stock และ Re-order Point

Safety Stock คือ สินค้าคงคลังขั้นต่ำที่ต้องมีติดคลังไว้ เพื่อให้มีสินค้าพร้อมส่งมอบให้ลูกค้าได้ทันทีระหว่างขอการสั่งซื้อหรือรอการผลิตรอบถัดไป

Re-order Point คือ การกำหนดปริมาณสินค้าไว้ส่วนหนึ่งเพื่อเป็นสินค้าที่มีในคลัง ป้องกันสินค้าขาดแคลน

2.5 หลักการพัฒนาระบบแอปพลิเคชัน

เว็บแอปพลิเคชัน คือแอปพลิเคชันที่ถูกเขียนขึ้นมาให้สามารถเปิดใช้ใน Web browser ได้โดยตรง ไม่ต้องโหลดแอปพลิเคชัน แบบเต็มๆ ลงเครื่อง ทำให้โดยรวมแล้วกินทรัพยากรค่อนข้างต่ำ สามารถเปิดใช้งานได้ไวกว่าการเปิดแอปพลิเคชัน แบบปกติจึงมีหน้าตาที่เป็นมิตรและใช้งาน

ค่อนข้างง่าย ซึ่งในปัจจุบันมี เว็บแอปพลิเคชันที่หลากหลายรูปแบบ [7]

2.6 หลักการ Responsive Web Design

เป็นเทคนิคการออกแบบเว็บไซต์แบบใหม่ ซึ่งจะมีการปรับเปลี่ยนขนาดของเว็บไซต์ให้เหมาะสมกับการแสดงผลบนหน้าจอขนาดต่างๆ และความละเอียดของหน้าจอในอุปกรณ์ที่แตกต่างกัน เช่น คอมพิวเตอร์ โน้ตบุ๊ก โทรศัพท์มือถือ แท็บเล็ต

2.7 ทฤษฎีการออกแบบ UX/UI

UI คือ กระบวนการสร้างส่วนต่อประสานสำหรับผู้ใช้ และระบบเว็บหรือแอปพลิเคชันเป็นส่วนที่ผู้ใช้งานมองเห็นได้ เช่น สี รูปทรง ตัวหนังสือ

UX คือ กระบวนการปรับแต่งประสบการณ์สำหรับผู้ใช้ให้พึงพอใจ

2.8 บริการไลน์แจ้งเตือน (Line Notify)

บริการ ไลน์แจ้งเตือน เป็นการแจ้งเตือนจากบัญชีทางการของไลน์ในรูปแบบของ API สามารถนำไปเชื่อมต่อกับเซิร์ฟเวอร์ได้ จะสร้างการแจ้งเตือนแบบข้อความไปยังกลุ่มหรือบัญชีส่วนตัวโดยไม่เสียค่าใช้จ่าย

2.9 งานวิจัยที่เกี่ยวข้อง

ฉัฐวดี โพธิ์กะสังข์ [9] นำเสนอเรื่องการพัฒนาระบบสารสนเทศสำหรับควบคุมและติดตามผลการดำเนินงานโครงการ นำเอาบริการไลน์แจ้งเตือนเข้ามาประยุกต์ใช้กับเว็บแอปพลิเคชันและมีรายงานแสดงความก้าวหน้าผลการดำเนินงานโครงการ ช่วยให้ผู้ใช้สามารถติดตามความก้าวหน้าโครงการได้อย่างสะดวกรวดเร็วและเข้าใช้งานได้ทุกที่ทุกเวลาโครงการ เพื่อให้สามารถบริหารจัดการโครงการได้อย่างมีประสิทธิภาพ แต่อย่างไรก็ตามงานวิจัยนี้ยังขาดในส่วนของการวางแผนการใช้ทรัพยากรในโครงการ

ธนภัทร จรรย์เวชประเสริฐ [10] นำเสนอเรื่องการพัฒนาสารสนเทศเพื่อการบริหารโครงการสำหรับธุรกิจขนาดเล็ก นำเอาทฤษฎีการออกแบบส่วนต่อประสานงานกับผู้ใช้มาออกแบบฟังก์ชันการใช้งานสามารถใช้งานง่ายและครอบคลุม ซึ่งระบบจะเป็นเว็บแอปพลิเคชันที่สามารถเปิดได้ทุกอุปกรณ์ เพื่อพัฒนาระบบ

สารสนเทศการบริหารโครงการแบบกระต๊อดเหมาะสำหรับธุรกิจขนาดเล็ก เพื่อช่วยอำนวยความสะดวกในการบริหารโครงการแต่ยังขาดในส่วนของการแจ้งเตือนการใช้ทรัพยากรไปยังโทรศัพท์มือถือของผู้ใช้งาน

3. สมมติฐานการวิจัย

3.1 ประสิทธิภาพแบบแบล็คบ็อกซ์ (Black Box) ในส่วนของระบบบริหารจัดการทรัพยากรโครงการ ด้านการวางแผน ด้านการสื่อสาร, ด้านการแสดงผล มีความถูกต้องมากกว่า 80%

3.2 ความพึงพอใจของผู้ใช้ระบบที่พัฒนาขึ้นใหม่อยู่ในระดับมาก มีค่าเฉลี่ยความพึงพอใจ > 3.51

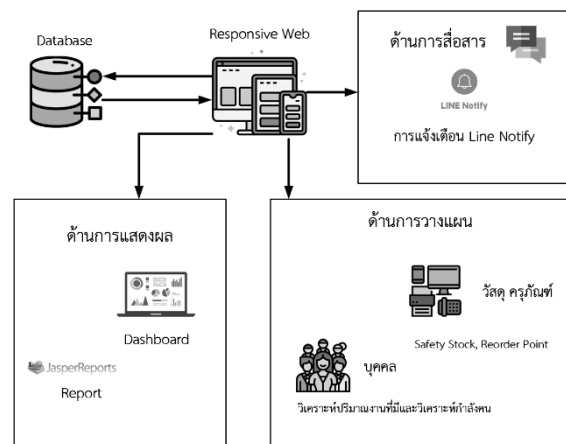
4. วิธีดำเนินการวิจัย

4.1 การศึกษา รวบรวมข้อมูลและสรุปปัญหา

จากการศึกษาปัญหาพบว่าสาเหตุของความล้มเหลวของโครงการประกอบด้วย การสื่อสารระหว่างทีมต่างๆ ไม่ทั่วถึงและไม่ชัดเจนรวมถึงทรัพยากรสำคัญเป็นปัจจัยหลักในความสำเร็จของโครงการ ในปัจจุบันมีผู้คนพัฒนาซอฟต์แวร์สำหรับการจัดการโครงการแต่ระบบมีหน้าตาการใช้งานที่ค่อนข้างยากและซับซ้อน เหมาะกับผู้ใช้งานที่มีทักษะหรือผ่านการใช้งาน

4.2 การวิเคราะห์และออกแบบระบบ

ผู้วิจัยได้ออกแบบภาพรวมในการทำงานของระบบดังภาพที่ 1



ภาพที่ 1: ภาพรวมในการทำงานของระบบ

4.2.1 ในส่วนของเว็บแอปพลิเคชันเป็นการพัฒนาด้วย Angular ในการสร้างหน้าเว็บส่วนติดต่อกับผู้ใช้โดยใช้รูปแบบของ HTML, CSS และ JavaScript ส่วนที่ให้ติดต่อกับฐานข้อมูลหรือ Service จะใช้ภาษา Java ในการพัฒนาฐานข้อมูลคือ MySQL และส่วนที่เกี่ยวข้องกับฟังก์ชันจะประยุกต์ใช้ทฤษฎีการวิเคราะห์ปริมาณงานและปริมาณคนมาใช้คำนวณจำนวนปริมาณงานและปริมาณคนที่มีในโครงการทางด้านพัสดุ และครุภัณฑ์ที่มีทฤษฎี Safety Stock และ Re-order Point มาคำนวณว่าพัสดุ และครุภัณฑ์ที่เคยขอเบิกมีใกล้จะหมดหรือในคลังเหลือน้อย ระบบจะประกอบด้วยฟังก์ชันที่สำคัญคือ 1) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลผู้ใช้งาน 2) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลหน่วยพัสดุและพัสดุ 3) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลพนักงาน 4) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลโครงการ 5) ฟังก์ชันการร้องขอพนักงานและพัสดุ 6) ฟังก์ชันการแจ้งเตือนการร้องขอทรัพยากรผ่านบริการไลน์แจ้งเตือน 7) ฟังก์ชันออกรายงานแสดงผลข้อมูลค่าใช้จ่ายทางด้าน พัสดุ ครุภัณฑ์ และพนักงาน

4.2.2 การวิเคราะห์ระบบบริหารจัดการทรัพยากรโครงการ ผู้วิจัยใช้ Context Diagram และ Data Flow Diagram ในการอธิบายการทำงานของระบบ

4.2.3 การออกแบบตารางฐานข้อมูลทั้งหมดในระบบ ประกอบด้วยตารางทั้งหมด 9 ตาราง

4.2.4 การออกแบบจำลองความสัมพันธ์ของข้อมูลโดยใช้แบบจำลอง (Entity Relationship Diagram)

4.3 การพัฒนาระบบ

ผู้วิจัยได้สร้างฐานข้อมูลและสร้าง Service สำหรับติดต่อกับฐานข้อมูลเป็นภาษา Java โดยใช้โปรแกรม IntelliJ เป็นเครื่องมือสำหรับพัฒนา ส่วนด้านการแสดงผลจะใช้ภาษา HTML, CSS และ JavaScript พัฒนาระบบโปรแกรม Visual Studio Code เป็นเครื่องมือในการพัฒนาระบบ

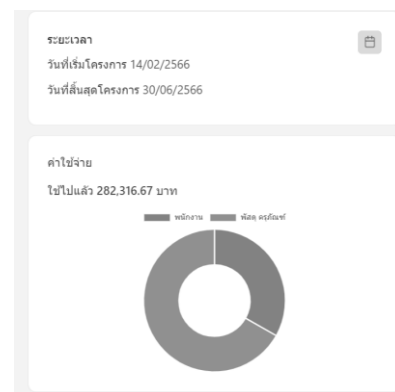
4.4 การทดสอบและประเมินผลระบบ

ผู้วิจัยทดสอบโดยใช้กระบวนการทดสอบแบบแบล็คบ็อกซ์ เป็นการทดสอบระบบโดยสนใจแค่กระบวนการภายนอก เช่นการบันทึกข้อมูลและผลลัพธ์ที่

ได้โดยไม่สนใจ Source Code โดยผู้ทดสอบได้กรอกข้อมูลนำเข้า Input ในระบบ จากนั้นระบบจะประมวลผลและส่งข้อมูลออกจากระบบ Output ให้ผู้ทดสอบได้ทราบว่ามี ความถูกต้องและตรงตามความต้องการหรือไม่ [8] หลังจากการทำสอบจะนำค่าสถิติที่ได้จากการทดสอบมาหาประสิทธิภาพของตัวระบบ คิดเป็นร้อยละ (%) โดยมีฟังก์ชันดังนี้ 1) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลผู้ใช้งาน 2) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลหน่วยพัสดุและพัสดุ 3) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลพนักงาน 4) ฟังก์ชัน ค้นหา/เพิ่ม/แก้ไข/ลบ ข้อมูลโครงการ 5) ฟังก์ชันการร้องขอพนักงานและพัสดุ 6) ฟังก์ชันการแจ้งเตือนการร้องขอทรัพยากรผ่านบริการไลน์แจ้งเตือน 7) ฟังก์ชันออกรายงานแสดงผลข้อมูลค่าใช้จ่ายทางด้าน พัสดุ ครุภัณฑ์ และพนักงาน

5. ผลการดำเนินงาน

ระบบวางแผนจัดการ รายงานแสดงผล และแจ้งเตือนทรัพยากรของโครงการ ประกอบด้วย 7 ส่วนหลักๆ มีภาพตัวอย่างของระบบที่พัฒนาขึ้น ดังภาพที่ 2 ถึงภาพที่ 7 ดังนี้

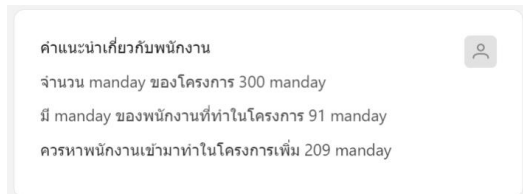


ภาพที่ 2: หน้าจอการแสดงผล Dashboard

ภาพที่ 3: หน้าจอการร้องขอพนักงาน



ภาพที่ 4: หน้าจอการแนะนำพัสดุ ครุภัณฑ์เล็กน้อย



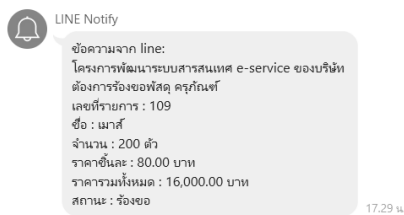
ภาพที่ 5: หน้าจอการแนะนำพนักงานไม่พอกับจำนวนงานที่มี

วันที่ 24 มิถุนายน 2566

รายงานการร้องขอพนักงาน
โครงการพัฒนาระบบสารสนเทศ e-service ของบริษัท

ชื่อพนักงาน	ตำแหน่ง	จำนวน manday	เริ่มวันที่	สิ้นสุดวันที่	ค่าใช้จ่าย
นายสังข์ ท่างาน	โปรแกรมเมอร์	30	01/03/2566	31/03/2566	25,000.00 บาท
นางสาวระวีใจ ไฉ่เรียนธุ์	นักวิเคราะห์ระบบ	15	13/02/2566	28/02/2566	15,000.00 บาท
นางญาณีตา สุโข	นักวิเคราะห์ระบบ	46	13/02/2566	31/03/2566	53,666.67 บาท
รวม					93,666.67 บาท

ภาพที่ 6: รายงานค่าใช้จ่ายพนักงานในโครงการ



ภาพที่ 7: แจ้งเตือนการร้องขอทรัพยากรในโครงการ

การทดสอบความถูกต้องจะเลือกส่วนของฟังก์ชันหลักที่มีผลสรุปดังนี้

ตารางที่ 1: ผลการทดสอบประสิทธิภาพ

ฟังก์ชัน	จำนวนครั้งที่ทดสอบ	% ความถูกต้อง
1. การจัดการข้อมูลต่างๆ ค้นหา/บันทึก/แก้ไข/ลบ	30	100
2. การแจ้งเตือนการร้องขอทรัพยากรโครงการ	30	93
3. การออกรายงานของข้อมูลการร้องขอทรัพยากรโครงการ	30	100

จากตารางที่ 1 พบว่าผลการวัดประสิทธิภาพการทำงานมีค่าเท่ากับ 98% สรุปได้ว่าระบบที่พัฒนาขึ้นสามารถดำเนินการตามฟังก์ชันหลัก ได้อย่างถูกต้องและมีประสิทธิภาพ การประเมินความพึงพอใจจะใช้แบบสอบถามแบบมาตราส่วนประเมินค่า 5 ระดับซึ่งแบ่งออกเป็น 3 ด้านดังตารางที่ 2

ตารางที่ 2: ผลการประเมินความพึงพอใจ

รายการประเมิน	\bar{X}	S.D.	ความพึงพอใจ
การทำงานของระบบ	4.60	0.58	มากที่สุด
การใช้งานของระบบ	4.28	0.88	มาก
ประโยชน์ที่ได้รับ	4.53	0.64	มากที่สุด
รวมทั้งหมด	4.46	0.72	มาก

จากตารางที่ 2 พบว่าผลการประเมินความพึงพอใจของผู้ใช้จำนวน 15 คน มีผลการทดสอบอยู่ในระดับมาก โดยมีค่าเฉลี่ยเท่ากับ 4.46 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.72

6. สรุปและอภิปรายผล

งานวิจัยนี้ได้ประยุกต์แนวทางการบริหารทรัพยากรโครงการ มาพัฒนาระบบจัดการทรัพยากรในโครงการ โดยได้เริ่มศึกษางานวิจัยที่เกี่ยวข้องต่างๆ จากนั้นจึงได้ทำการวิเคราะห์และออกแบบระบบ สามารถช่วยในการวางแผนการแสดงผลและการสื่อสารภายในโครงการได้อย่างรวดเร็วเป็นปัจจุบันอีกทั้งยังเข้าถึงข้อมูลได้ทุกที่ทุกเวลา และยังสามารถแสดงผลบนอุปกรณ์ต่างๆ ได้อย่างเหมาะสม

ผลการประเมินประสิทธิภาพโดยใช้การทดสอบแบบลิคิ์บ็อกซ์ด้วยวิธี UAT โดยให้ผู้เชี่ยวชาญทั้ง 3 คนมาทดสอบพบว่าระบบที่พัฒนาขึ้นมีประสิทธิภาพในระดับ 98% อีก 2% ที่ไม่ถูกต้องเกิดจากอินเทอร์เน็ตในขณะที่ทำการทดสอบ ไม่เสถียรจึงทำให้เกิดความผิดพลาดในฟังก์ชันการส่งไลน์แจ้งเตือน ซึ่งผลการประเมินที่ได้

สอดคล้องกับการประเมินผลการใช้งานระบบสารสนเทศ สำหรับควบคุมและติดตามผลการดำเนินงานโครงการ [9] และระบบสารสนเทศเพื่อการบริหารโครงการสำหรับ ธุรกิจขนาดเล็ก [10] ถึงแม้จะเป็นระบบที่มีความแตกต่างกันแต่ใช้หลักการ และทฤษฎีการออกแบบในลักษณะเดียวกันจึงมีผลลัพธ์เป็นไปในทิศทางเดียวกัน

ผลการประเมินความพึงพอใจ โดยใช้แบบสอบถาม แบบมาตราส่วนประเมินค่า 5 ระดับ ของผู้ใช้งาน 15 คน แบ่งเป็น 3 ด้านดังนี้ 1) ความพึงพอใจในการทำงานของระบบ ได้ค่าเฉลี่ยเท่ากับ 4.60 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.58 2) การใช้งานของระบบ ได้ค่าเฉลี่ยเท่ากับ 4.28 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.88 3) ความพึงพอใจในประโยชน์ที่ได้รับ ได้ค่าเฉลี่ยเท่ากับ 4.53 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.64

จากการวิจัยในครั้งนี้ผู้วิจัยค้นพบว่า ระบบบริหารจัดการทรัพยากรโครงการ ที่ออกแบบให้ใช้งานง่าย มีการแสดงผลที่ชัดเจนมองเห็นง่าย และออกแบบให้สามารถ การปรับเปลี่ยนขนาดของหน้าจอให้เหมาะสมกับการ แสดงผลบนหน้าจอขนาดต่าง ๆ เป็นที่นิยมในปัจจุบัน เช่น สมาร์ทโฟน หรือแท็บเล็ต มีระบบการแจ้งเตือนผู้ใช้งาน แบบสามารถตอบสนอง โดยทันทีจะช่วยอำนวยความสะดวกให้ผู้ใช้งานได้เป็นอย่างดี ในส่วนข้อเสนอแนะเพื่อ การนำงานวิจัยไปต่อยอด ควรเพิ่มการจัดการทรัพยากรใน เรื่องของงบประมาณเข้าไปเพื่อสามารถจัดการและ พยากรณ์เกี่ยวกับการใช้งบประมาณของโครงการได้และ สามารถแสดงผลความเสี่ยงที่โครงการจะล้มเหลวได้เพื่อ เพิ่มประสิทธิภาพในการบริหารโครงการได้ดียิ่งขึ้น

เอกสารอ้างอิง

- [1] มหาวิทยาลัยสุโขทัยธรรมาธิราช. (2558). [ออนไลน์]. ความหมายของโครงการและความสำคัญของโครงการ [สืบค้น วันที่ 6 กรกฎาคม 2565]. จาก <https://www.stou.ac.th/stouonline/lom/data/sec/Lom14/01-01-01.html>
- [2] iPlus. (2564). [ออนไลน์]. สาเหตุใดบ้างที่อาจจะทำให้โครงการล้มเหลวได้ [สืบค้นเมื่อวันที่ 6 กรกฎาคม 2565]. จาก [https://www.ipluscenter.com/5สาเหตุใดบ้างที่](https://www.ipluscenter.com/5สาเหตุใดบ้างที่อาจจะทำ) อาจจะทำ
- [3] Content Shifu. (2564). [ออนไลน์]. Asana คืออะไร [สืบค้นเมื่อวันที่ 27 กรกฎาคม 2565]. จาก <https://contentshifu.com/tools/asana>
- [4] Microsoft (2565). [ออนไลน์]. Microsoft Project. [สืบค้น เมื่อวันที่ 27 กรกฎาคม 2565]. จาก <https://www.microsoft.com/th-th/microsoft-365/project/project-management-software?rtc=1>
- [5] Tiger. (2564). [ออนไลน์]. การบริหารโครงการ คืออะไร [สืบค้นเมื่อวันที่ 27 กรกฎาคม 2565]. จาก <https://thaiwinner.com/project-management/>
- [6] Project Management Institute. (2017). A GUIDE TO THE PROJECT MANAGEMENT BODY OF KNOWLEDGE PMBOK Guide 6 Edition. Inc. 14
- [7] Pichaya Srifar. (2561). [ออนไลน์]. กระบวนการพัฒนา Application [สืบค้นเมื่อวันที่ 28 กรกฎาคม 2565]. จาก <https://medium.com/@pichayas/5-ขั้นตอน-กระบวนการพัฒนา-mobile-application-ตั้งแต่ต้นจนจบ-5926634aeabf>
- [8] Laurie Williams. (2006). [ออนไลน์]. Testing Overview and Black-Box Testing Techniques. [สืบค้น เมื่อวันที่ 28 กรกฎาคม 2565]. จาก <http://agile.csc.ncsu.edu>
- [9] ณัฐวดี โพธิ์กะสังข์. (2563). การพัฒนาระบบสารสนเทศ สำหรับควบคุมและติดตามผลการดำเนินงานโครงการ. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. สำนักหอสมุดกลาง
- [10] ชนภัทร จรรย์เวชประเสริฐ. (2562). การพัฒนาระบบสารสนเทศเพื่อการบริหารโครงการสำหรับธุรกิจขนาดเล็ก. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. สำนักหอสมุดกลาง

การพัฒนา รูปแบบของระบบบริหารจัดการการขายหน้าร้าน และสินค้าคงคลัง สำหรับร้านค้าโชวห่วยขนาดเล็กบนโทรศัพท์มือถือสมาร์ทโฟน Development of Purchase and Inventory Management System on Smartphones for a Small Grocery Store

สุรชาติ อุ่นสอน (Surachat Aunsorn)¹ และณัฐวี อดตฤกษ์ (Natavee Utakrit)²

¹ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021810038@email.kmutnb.ac.th, ²nattavee.u@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนารูปแบบของระบบบริหารจัดการการขาย และสินค้าคงคลังสำหรับร้านค้าโชวห่วยขนาดเล็กที่มีข้อจำกัดในเรื่องงบประมาณในการจัดซื้อระบบและอุปกรณ์สมัยใหม่เพื่อช่วยบริหารจัดการร้านค้า รวมถึงการใช้งานที่ซับซ้อนเกินความจำเป็น ผู้วิจัยจึงได้ศึกษาพัฒนาและประยุกต์ใช้หลักการของจุดขายหน้าร้าน (POS: Point of Sale) ร่วมกับทฤษฎีการบริหารสินค้าคงคลัง และใช้บริการบนแอปพลิเคชันไลน์ (LINE) ซึ่งเป็นสื่อสังคมออนไลน์ที่เป็นที่นิยมอย่างมากในปัจจุบัน เพื่อความสะดวกในการใช้งาน และการเชื่อมโยงข้อมูลผ่านอุปกรณ์มือถือประเภทสมาร์ทโฟนที่มีต้นทุนที่ต่ำกว่าสามารถเรียนรู้การใช้งานได้ง่าย และมีผู้ใช้งานกันมากในปัจจุบัน ระบบดังกล่าวนี้พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชัน โดยใช้หลักการออกแบบ Responsive Web Design มีระบบสแกนบาร์โค้ดเพื่อขายสินค้า และมีระบบบริหารจัดการคลังสินค้าสำหรับเจ้าของร้าน รวมถึงระบบแจ้งเตือนข้อมูลการยอดขายผ่านทาง แอปพลิเคชันไลน์ ผลการทดสอบประสิทธิภาพในการบริหารจัดการการขาย และสินค้าคงคลังพบว่ามีประสิทธิภาพอยู่ที่ 94% และผลการประเมินความพึงพอใจของผู้ใช้งานระบบอยู่ที่ระดับดีมากที่ค่าเฉลี่ย 4.88 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.37 จึงสรุปได้ว่าระบบพัฒนาขึ้นนี้มีการทำงานที่มีประสิทธิภาพ และผู้ใช้งานมีความพึงพอใจมาก

คำสำคัญ: สมาร์ทโฟน ร้านโชวห่วย การสแกนบาร์โค้ด ระบบบริหารการขายหน้าร้าน ระบบสินค้าคงคลัง เว็บแอปพลิเคชัน

Abstract

The objective of this project is to create a sales and inventory management system for small grocery stores with limited budget that struggle to manage the complex operations of the shop and purchase expensive equipment. The theory of inventory management was integrated by the researcher with POS (Point of Sale). Moreover, use the service via the LINE application, a well-known social network platform that free and very popular nowadays. This system is developed as a web application using Responsive Web Design. It has a barcode scanning system for selling products and a product inventory management system for the store owner. Additionally, there is a sales report notification system and a notification system for sellers when products are running low via LINE application. The system's efficiency score is 94%, and the S.D. for user satisfaction is 0.37, indicating that both the system's effectiveness and users' level of satisfaction are high.

Keywords: Smartphone, Small Grocery Store, Barcode Scanning, Purchase Management System, Inventory Management System, Web Application.

1. บทนำ

เทคโนโลยีได้เข้ามามีบทบาทในธุรกิจมากขึ้น ทำให้เกิดการเปลี่ยนแปลงอย่างรวดเร็วส่งผลให้ ธุรกิจส่วนใหญ่จึงต้องมีการปรับตัวตามให้ทันต่อการเปลี่ยนแปลงของเทคโนโลยีเหล่านั้น จะต้องนำเทคโนโลยี มาใช้ในการดำเนินธุรกิจนั้นมืออยู่มากมาย เช่น ธุรกิจ โรงแรม ธุรกิจร้านอาหาร หรือแม้แต่ธุรกิจขายสินค้า โดยจะมีธุรกิจด้านการขายสินค้าอยู่ประเภทหนึ่งที่ยังคงไม่ใช้ เทคโนโลยีในการดำเนินธุรกิจนั่นก็คือร้านโชวห่วย ที่เป็นกิจการเล็ก ๆ ในครัวเรือน ซึ่งโดยมากผู้ประกอบการร้านโชวห่วยส่วนใหญ่จะเป็นผู้สูงอายุ

ร้านค้าแบบสะดวกประเภทโชวห่วย ที่ไม่มีการใช้ เทคโนโลยีเข้ามาช่วยในการจัดการสินค้า หรือสรุป รายงานการขายต่าง ๆ ส่งผลให้บางครั้งการเติมสินค้าเข้าร้าน หรือการสรุปรายงานการขายของร้านโชวห่วยจะมีข้อผิดพลาดเกิดขึ้นได้ ธุรกิจการค้าส่วนใหญ่มักใช้ระบบการขายหน้าร้าน (POS: Point of Sale) ในการขายสินค้า ระบบ ดังกล่าวนี้นี้จะพบเห็นได้ตามห้างสรรพสินค้า มินิมาร์ท และร้านค้าทั่วไปได้ที่จุดแคชเชียร์ชำระเงิน ระบบระบบการขายหน้าร้านจำเป็นต้องอาศัยอุปกรณ์ในการคิดเงินที่มีตัวเครื่องติดกับเครื่องสแกนบาร์โค้ด มีการเชื่อมต่องานข้อมูลคอมพิวเตอร์ และใช้ซอฟต์แวร์ในการบริหารจัดการ [1] ซึ่งราคาของเครื่อง POS ในตลาด มีราคาตั้งแต่ 10,000 บาทขึ้นไป ซึ่งถือว่ามีราคาค่อนข้างสูงสำหรับร้านโชวห่วยที่มีขนาดเล็ก

จากผลการสำรวจการใช้ ICT ของประชาชนในประเทศไทยปี 2565 พบว่า คนไทยมีแนวโน้มการใช้คอมพิวเตอร์และโทรศัพท์มือถือประจำครัวเรือนเพิ่มขึ้นอย่างต่อเนื่อง และประชาชนส่วนใหญ่มีโทรศัพท์สมาร์ทโฟนถึง 94.1% ซึ่งส่วนใหญ่คนไทยใช้ Android มากถึง 74.8% [2]

แอปพลิเคชันไลน์ มีจำนวนผู้ใช้งานมากกว่า 50 ล้านคน มากกว่า 2 ใน 3 ของจำนวนประชากร ทำให้ LINE เป็นโซเชียลที่มีคอมมูนิตี้กว้างขวางมากในประเทศไทย โดยคนไทยมีการใช้งาน application LINE ค่อนข้างเยอะ ผู้สูงอายุมีการปรับตัวและมีแนวโน้มเปิดรับกับเทคโนโลยี

ใหม่ ๆ มากขึ้น [4] โดยเฉพาะอุปกรณ์สื่อสารเคลื่อนที่ (Mobile devices) เช่น สมาร์ทโฟน แท็บเล็ต ซึ่งมีราคาถูกลง [3]

ในการศึกษาวิจัยครั้งนี้ได้เลือกร้านโชวห่วยขนาดเล็กแห่งหนึ่งเป็นกรณีศึกษา โดยในการดำเนินธุรกิจขายของร้านโชวห่วยแห่งนี้เมื่อมีลูกค้าเข้ามาจับจ่ายซื้อของมีการทำงานที่ยังไม่เป็นระบบ และเมื่อมีการซื้อของจำนวนมากยากต่อการจกรายการสินค้าที่ลูกค้าซื้อ และเกิดความล่าช้า และไม่มีการใช้งานเครื่อง POS ที่เป็นอุปกรณ์ที่คอยอำนวยความสะดวกในการคิดเงินอย่างเป็นระบบ ไม่มีระบบข้อมูลการขาย ข้อมูลสินค้าคงคลัง ข้อมูลสมาชิก ยอดซื้อสะสม และมีราคาที่ยกขึ้นสูงเครื่อง ต้องมีการจ้างช่างมาติดตั้ง และดูแลรักษาตลอดเวลา อีกทั้งผู้ใช้งานต้องมีการอบรมวิธีใช้งาน ซึ่งไม่เหมาะกับขนาดของร้าน และเจ้าของร้านโชวห่วยส่วนใหญ่เป็นผู้สูงอายุที่ไม่ทันเทคโนโลยี นั้นทำให้ร้านโชวห่วยส่วนใหญ่ไม่นิยมใช้งานเครื่อง POS

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ระบบจุดขายหน้าร้าน (POS)

ระบบ POS คือ ระบบขายหน้าร้าน ย่อมาจาก Point of Sale ก็คือ ระบบที่ตั้งอยู่ตรงบริเวณแคชเชียร์และต่อพ่วงกับเครื่องสแกนบาร์โค้ด ในสมัยก่อนร้านค้ามักใช้งานระบบที่สามารถใช้คิดเงินได้เพียงอย่างเดียว และบอกเพียงยอดขาย ทำให้ไม่สามารถตอบโต้การขายสำหรับผู้ประกอบการและเจ้าของร้านค้าได้ ต่อมาเครื่องคิดเงินสินค้าก็ได้ถูกพัฒนาใส่โปรแกรมลงไปจนกลายเป็นเครื่องคิดเงินในชื่อของ “ระบบ POS” และเมื่อระบบนี้อยู่ในอุปกรณ์อื่น เราก็จะเรียกมันว่า “เครื่อง POS” [4]

2.2 สมาร์ทโฟน

สำนักงานสถิติแห่งชาติ เผยผลการสำรวจการใช้ ICT ของประชาชนในประเทศไทยปี 2565 พบว่า คนไทยมีการใช้โทรศัพท์มือถือสมาร์ทโฟนเพิ่มขึ้นอย่างต่อเนื่อง โดยผู้มีโทรศัพท์มือถือ 57.5 ล้านคน คิดเป็น 87.9% แบ่งเป็นโทรศัพท์มือถือสมาร์ทโฟนถึง 94.1% [5]

2.3 การใช้งาน Line ในผู้สูงอายุ

เนื่องจากในปัจจุบัน อินเทอร์เน็ตเป็นสิ่งทีทุกคนเข้าถึงได้ ส่งผลให้ผู้สูงอายุมีการปรับตัวให้เข้ากับโลกสมัยใหม่

จึงทำให้ผู้สูงอายุโซเซียลมีเดียเพิ่มขึ้น ปัจจุบันประเทศไทย มีประชากรสูงวัยอายุ 60 ปีขึ้นไป คิดเป็น 18.3% ของ ประชากรไทยทั้งหมด โดยแบ่งเป็น ช่วงอายุ 60-69 ปี จำนวน 6,843,300 คน คิดเป็น 56.5% ของผู้สูงอายุทั้งหมด ช่วงอายุ 70-79 ปี มีจำนวน 3,522,778 คน คิดเป็น 29.1% และ อายุ 80 ปีขึ้นไป จำนวน 1,750,121 คน คิดเป็น 14.4% และพบว่าโซเซียลมีเดียที่ผู้สูงอายุในประเทศไทยนิยมเล่น มากที่สุดอันดับหนึ่งคือ ไลน์ (LINE) [6]

2.3 LINE Notify

LINE Notify คือสิ่งที่จะสามารถทำให้ระบบ หรือ โปรแกรมของเราสามารถได้รับข้อความแจ้งเตือนจาก Line โดยทำการเชื่อมต่อกับทาง Line เว็บเซอร์วิส Line notify นั้นทำให้เราสามารถได้รับการแจ้งเตือนจาก LINE Notify ผ่านทาง LINE ของเรา โดยเราสามารถเลือก เชื่อมต่อกับเซอร์วิสที่เราต้องการ และนอกจากนี้ Line notify ยังสามารถแจ้งเตือนผ่าน Line กลุ่มได้อีกด้วย [7]

2.4 บาร์โค้ด (barcode)

บาร์โค้ด คือ เลขหมายประจำตัวสินค้า ใ้แทนด้วย แท่งบาร์ขาว-ดำ ประกอบด้วยตัวเลข 8-13 หลัก สามารถ อ่านได้ด้วยเครื่องสแกนเนอร์ นิยมใช้กับสินค้าอุปโภค บริโภคแทบทุกชนิด และสินค้าสำเร็จรูปต่าง ๆ สำหรับ ตัวเลขรหัสสินค้าของบาร์โค้ดสินค้าที่มีจำหน่ายใน ประเทศไทยส่วนใหญ่จะมีขึ้นของจดจากสถาบันสัญลักษณ์ รหัสแห่งประเทศไทยเพื่อให้ได้เลขรหัสบาร์โค้ดที่เป็นมาตรฐาน เป็นที่เรียบร้อยแล้ว ซึ่งงานวิจัยนี้จะศึกษาเฉพาะสินค้าที่มีการขึ้นจกรหัสบาร์โค้ดที่เป็นมาตรฐาน โดยบาร์โค้ดจะมี ลักษณะเป็นเลขชุด 13 หลัก

2.5 การออกแบบหน้าจอ Responsive Web Design

Responsive Web Design เป็นเทคนิคการออกแบบ เว็บไซต์ โดยจะมีการปรับเปลี่ยนขนาดของเว็บไซต์ให้ เหมาะสมกับการแสดงผลบนหน้าจอในขนาดต่างๆ และความละเอียดของหน้าจอของอุปกรณ์ที่แตกต่างกัน เช่น คอมพิวเตอร์ โน้ตบุ๊ก โทรศัพท์มือถือ แท็บเล็ต การ ออกแบบเว็บไซต์แบบ Responsive จะสามารถปรับขนาด ของเว็บไซต์ได้อัตโนมัติตามขนาดของอุปกรณ์ที่ใช้งานอยู่ ทำให้หน้าเว็บไซต์จะมีเพียง 1 URL โดยที่ไม่จำเป็นต้อง

แยกเว็บไซต์เป็นเวอร์ชัน Desktop และ Mobile หากเปิด เว็บไซต์ด้วยโทรศัพท์มือถือ เว็บไซต์ก็จะปรับขนาดให้ เหมาะสมกับความกว้างของจอ ทำให้มีความสะดวกใน การใช้งานมากขึ้น และนอกจากนี้ขนาดของตัวอักษรก็สามารถปรับให้เหมาะสมเพื่อให้สามารถมองตัวอักษรได้ อย่างชัดเจน [9]

2.6 ทฤษฎีบริหารสินค้าคงคลัง

การบริหารสินค้าคงคลัง คือวางแผน และจัดการ ทรัพย์สินที่ยังไม่ได้ขายหรือ สินค้าคงคลัง ที่ยังเก็บไว้เพื่อ รอการขาย เพื่อให้การบริหารสินค้าที่เก็บอยู่มีปริมาณที่ เหมาะสมเพียงพอต่อการขาย ในสภาพที่เหมาะสม จาก ขบวนการรับสินค้าเข้า กระบวนการจัดเก็บ จนไปถึง สนับสนุนกระบวนการขาย เพื่อไม่ให้สินค้าขาด [10]

2.7 งานวิจัยที่เกี่ยวข้อง

วิศปต์ย์ [11] ศึกษาการใช้ไลน์ของผู้สูงอายุ: การศึกษา เิงปรากฏการณ์วิทยา การใช้งานไลน์ในมุมมอง ของ ผู้สูงอายุ ทราบถึงพฤติกรรมในการใช้งานไลน์ของ ผู้สูงอายุที่มีความชัดเจนและสมบูรณ์ยิ่งขึ้น อันจะเป็น ประโยชน์ต่อการกำหนดแนวทางการส่งเสริมการเข้าถึง เทคโนโลยีสารสนเทศของผู้สูงอายุ ณัฐธยาน์และคณะ [12] โปรแกรมอัตโนมัติเพื่อจัดการคลังสินค้าใน สหกรณ์และร้านค้า ได้ศึกษาปัญหาที่เกิดขึ้นภายในร้านค้า สหกรณ์เกษตรศาสตร์ จำกัด และร้านขายของชำ วริษฐา [13] ประสิทธิภาพในการจัดการคลังสินค้าด้วยระบบ บาร์โค้ดผ่าน โปรแกรม WMS ของพนักงานบริษัท สปอร์ตส เอ็นจิเนียริ่ง แอนด์ รีกรีเอชั่น เอเชีย จำกัด ทำการศึกษาประสิทธิภาพการจัดการคลังสินค้าด้วยระบบ บาร์โค้ดผ่าน โปรแกรม WMS การนำระบบการจัดการ คลังสินค้าเข้ามาใช้ในการจัดการคลังสินค้า เพื่อให้เกิดการ ทำงานอย่างเป็นระบบและคุ้มค่าแก่การลงทุน เป็นการ ควบคุมการรับสินค้า การจัดเก็บ และการเคลื่อนย้ายสินค้า ซึ่งจะสามารถช่วยให้การทำงานสะดวก รวดเร็ว และมี ประสิทธิภาพ เอก และ ธนชัย [14] เว็บแอปพลิเคชัน ระบบ จัดการสินค้า และการขายของร้านคิมแอนด์จิวมีนิมาร์ท เป็นการพัฒนาระบบในรูปแบบเว็บแอปพลิเคชัน ประกอบด้วยส่วนระบบการจัดการข้อมูลร้านค้า และ

ระบบการขายสินค้า โดยตัวแอปพลิเคชันจะมีความสามารถในการทำงานที่เกี่ยวข้อง ได้แก่ ระบบบันทึกการขายสั่งซื้อ คำนวณเงินและออกใบเสร็จให้ลูกค้าได้ถูกต้อง และมีระบบตัดสต็อกสินค้า มีระบบการแจ้งเตือนสินค้าที่มีจำนวนเหลือน้อยหรือรายงานสินค้าที่ขาย

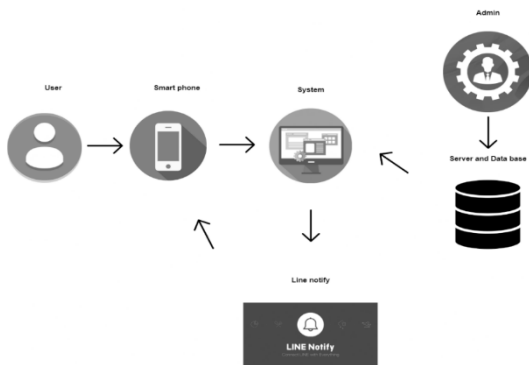
3. วิธีดำเนินการวิจัย

3.1 ศึกษาและวิเคราะห์ปัญหาเบื้องต้น

การศึกษาและวิเคราะห์ปัญหาเบื้องต้นจากเจ้าของร้านขายของโชวห่วยขนาดเล็กที่เลือกเป็นกรณีศึกษา ซึ่งพบว่า

- 1) ไม่มีการบันทึกจำนวนสินค้าเข้า – ออก
- 2) การนับจำนวนสินค้าคงคลังด้วยตัวเอง ไม่มีการจัดการที่เป็นระบบ
- 3) เมื่อขายสินค้าในบางครั้ง จำราคาสินค้าไม่ได้
- 4) เครื่อง POS ส่วนใหญ่ราคาสูง ใช้งานยาก มีหลายระบบที่ไม่จำเป็นสำหรับร้านขนาดเล็ก

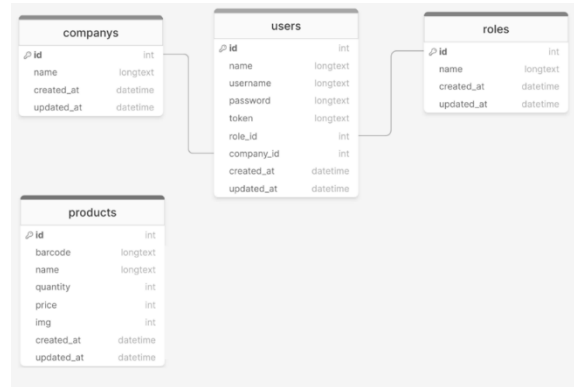
3.2 วางแผนและพัฒนาระบบ



ภาพที่ 1: กรอบแนวคิดของระบบ

ทำการออกแบบระบบ โดยนำข้อมูลจากการศึกษาวิเคราะห์ปัญหา ประยุกต์เข้ากับทฤษฎี และงานวิจัยที่เกี่ยวข้องเพื่อสร้างกรอบแนวคิดในการพัฒนาระบบ ดังภาพที่ 1 จากนั้นเริ่มพัฒนาระบบโดยเครื่องมือในการพัฒนาจะใช้ Angular framework ในการพัฒนา Front end ส่วนของ Website และใช้ Android Studio Code ในการพัฒนา Front end ในส่วนของ Android Application ที่เชื่อมต่อไปยัง Website ภายนอกได้โดยเริ่มจากวาง โครงสร้างของระบบก่อน ในส่วนความปลอดภัยของระบบ จะมีการเข้าสู่ระบบและกำหนดสิทธิ์การเข้าใช้งาน และทำการสร้างระบบในส่วนติดต่อกับระบบจัดการข้อมูล ในส่วนของฐานข้อมูลมีการใช้ MYSQL เพื่อส่งข้อมูลไป

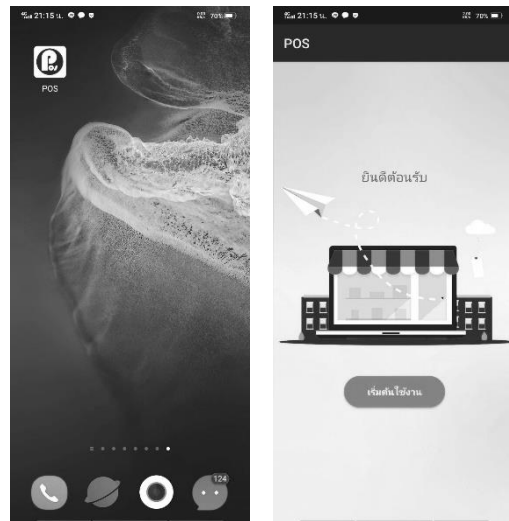
ประมวลผลและส่งข้อมูลกลับมา ผู้วิจัยได้ทำการวางผังของระบบฐานข้อมูลเอาไว้ก่อนจะพัฒนาระบบจริงโดยผังของระบบจะมีลักษณะดังภาพที่ 2



ภาพที่ 2: ER diagram ของระบบ

การเชื่อมต่อ API ก็ต้องเข้ากับระบบเพื่อทำให้สามารถใช้งานกล้องในการสแกนบาร์โค้ด และมีระบบในการเพิ่ม Line Token เพื่อให้ผู้ใช้สามารถเชื่อมต่อการแจ้งเตือนผ่าน Line Notify ได้

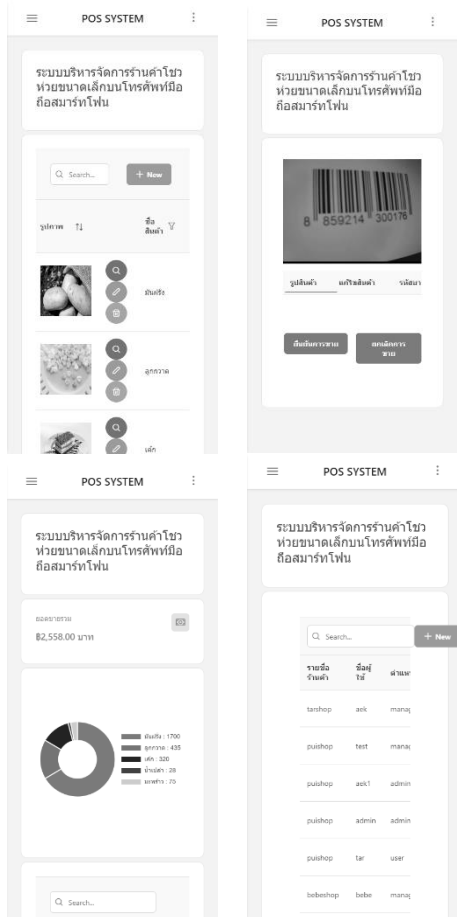
ในด้านการออกแบบส่วนติดต่อกับผู้ใช้ เนื่องจากกลุ่มเป้าหมายคือผู้สูงอายุที่เปิดร้านโชวห่วยผู้วิจัยจึงออกแบบหน้า Application ที่จะนำผู้ใช้ไปยัง Web application ของระบบ ให้ใช้งานง่ายดังภาพที่ 3



ภาพที่ 3: Application บนโทรศัพท์มือถือสมาร์ทโฟน

ข้อดีของการพัฒนาระบบในรูปแบบเว็บแอปพลิเคชันคือสามารถใช้งานได้ทั้งบนมือถือสมาร์ทโฟน และบนเว็บไซต์ ไม่ต้องดาวน์โหลดแอปพลิเคชัน ผู้ใช้งานไม่ต้องคอยอัปเดตระบบด้วยตนเอง ลดค่าใช้จ่ายในการซื้อโปรแกรมที่มีราคาแพง โดยมีฟังก์ชันการทำงานของระบบ

ดังนี้ ค้นหารายการสินค้า 1)เพิ่ม ลบ และทำการแก้ไข รายการสินค้า 2)สแกนบาร์โค้ด ขายสินค้า 3)แสดงยอดขาย และสรุปผลสินค้ารายวัน 4)การแจ้งเตือนผ่านไลน์แจ้งเตือน (Line notify) 5)แบ่งสิทธิการใช้งานในแต่ละผู้ใช้ เพื่อจำกัดการเข้าถึง



ภาพที่ 4: หน้าต่างการใช้งานระบบ

4. ผลการทดสอบระบบ

ทำการทดสอบระบบลองทดสอบการใช้งานในแต่ละบทบาทหน้าที่ เช่นผู้ขาย หรือเจ้าของร้าน รวมถึงผู้ดูแลระบบ เพื่อทดสอบใช้งานตามฟังก์ชันต่าง ๆ อาทิ การแสดงรายการสินค้า การเพิ่มสินค้าการลบสินค้า แก้ไขรายการสินค้า ทดสอบการสแกนขายสินค้าจากป้าบบาร์โค้ด ผ่านโทรศัพท์มือถือสมาร์โฟน และการทดสอบการแจ้งข้อมูลการขาย และสินค้าคงคลังผ่านแอปพลิเคชันไลน์ ประเมินโดยผู้เชี่ยวชาญจำนวน 3 คน จากเจ้าของร้านค้าขนาดเล็ก พนักงานฝ่ายขายของบริษัทนำเข้าสินค้า และนักพัฒนาระบบ ผลออกมาดังตารางที่ 1

ตารางที่ 1: ตารางแสดงผลการประเมินความถูกต้องของระบบ

หัวข้อ	จำนวนครั้งที่ทดสอบ	จำนวนครั้งที่ถูก	จำนวนครั้งที่ผิด	ค่าเฉลี่ย
การจัดการคลังสินค้า	30	30	0	100%
การเช็คยอดขาย	30	30	0	100%
การสแกนบาร์โค้ด	30	21	9	70%
การแจ้งยอดขายผ่าน line	30	30	0	100%
ความถูกต้องของสิทธิ์การใช้งาน	30	30	0	100%
รวม	150	47	3	94%

จากตารางที่ 1 พบว่าผลการประเมินความถูกต้องของการทำงานของระบบมีค่าเฉลี่ยรวมอยู่ที่ 94%

ในส่วนของความพึงพอใจ ประเมินโดยผู้ใช้งานรวมทั้งหมด 20 คน ผลออกมาดังตารางที่ 2

ตารางที่ 2: ตารางแสดงผลความพึงพอใจของระบบ

หัวข้อ	X bar	SD	ระดับความพึงพอใจ
ความง่ายในการใช้งาน	4.95	0.22	มากที่สุด
ความเร็วในการทำงานของระบบ	4.95	0.22	มากที่สุด
ที่ภาษาที่เข้าใจง่าย	4.80	0.51	มากที่สุด
มีการเลือกใช้ตัวอักษรที่เหมาะสมและอ่านง่าย	4.85	0.48	มากที่สุด
มีการแสดงยอดขายที่เหมาะสมและเข้าใจง่าย	4.90	0.30	มากที่สุด
ระบบสามารถใช้งานบนแพลตฟอร์มโทรศัพท์มือถือได้อย่างลื่นไหล	4.85	0.48	มากที่สุด

จากตารางที่ 2 มีผลทดสอบอยู่ในระดับดีมาก โดยมีค่าเฉลี่ยเท่ากับ 4.88 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.37

5. สรุปและข้อเสนอแนะ

ระบบที่พัฒนาขึ้น มีฟังก์ชันการทำงานที่มีความซับซ้อนที่น้อยกว่า เหมาะสมสำหรับร้านค้าขนาดเล็ก และใช้ฟังก์ชันที่มีอยู่ในโทรศัพท์มือถือทำให้ไม่ต้องใช้อุปกรณ์เสริมภายนอก เช่นเครื่องสแกนบาร์โค้ด ทำให้มีราคาถูก ผลการทดสอบประสิทธิภาพการทำงานของระบบบริหารจัดการร้านค้าสามารถทำงานได้ถูกต้องครบถ้วน 100% ยกเว้นในส่วนของการสแกนสินค้าที่ยังพบการผิดพลาดอยู่บ้าง เนื่องจาก ในขณะที่สแกนบาร์โค้ดแล้วผู้ใช้งานเกิดขยับกล้องขณะสแกน หรือบาร์โค้ดบริเวณสินค้าเกิดการชำรุดเสียหาย ก็จะก่อให้เกิดความผิดพลาดได้ แต่อย่างไรก็ตามผู้วิจัยมีข้อเสนอแนะในการปรับปรุงว่าผู้สมควรใช้อุปกรณ์เสริมของสมาร์ทโฟนอย่างเช่นขาตั้งเพื่อป้องกันความคลาดเคลื่อนเมื่อสแกนบาร์โค้ด และในส่วนความพึงพอใจของผู้ใช้ พบว่ามีค่าเฉลี่ยในระดับดีมากสำหรับข้อเสนอแนะเพื่อการต่อยอดการวิจัยในอนาคต ผู้วิจัยเห็นว่าควรมีการเพิ่มเติมฟังก์ชันการสแกนในรูปแบบ QR Code หรือการประยุกต์ใช้เทคนิคการประมวลผลด้วยภาพเพื่อสแกนสินค้าจากป้ายยี่ห้อสินค้าได้โดยตรง จะช่วยเพิ่มวิธีการสแกนสินค้าในการขายที่หลากหลาย และอำนวยความสะดวกแก่ผู้ขายได้มากขึ้น

เอกสารอ้างอิง

- [1] STOREHUB. (2565). ระบบ POS. (ออนไลน์). จาก <https://shorturl.asia/vMCSy>
- [2] RYT9. (2565). สถิติการใช้ ICT ในประเทศไทย. (ออนไลน์). จาก <https://www.ryt9.com/s/nso/3295274>
- [3] Brandinside. (2565). ยอดผู้ใช้งาน LINE ในปี 2565. (ออนไลน์). จาก <https://brandinside.asia/line-thailand-10-yrs/>
- [4] STOREHUB. (2565). ระบบ POS. (ออนไลน์). จาก <https://shorturl.asia/vMCSy>
- [5] สำนักงานสถิติแห่งชาติ. (2565). สถิติการใช้งาน สมาร์ทโฟน. (ออนไลน์). จาก <https://marketeeronline.co>
- [6] The Bangkok Insight. (2561). การใช้ไลน์ในผู้สูงอายุ. (ออนไลน์). จาก <https://www.thebangkokinsight.com/news/opinions/44260/>
- [7] LineNotify. การแจ้งเตือนโดยใช้ linenotify . (ออนไลน์). จาก <https://notify-bot.line.me/th/>
- [8] BKKPAPERBOX (มปป.) บาร์โค้ดคืออะไร. (ออนไลน์). จาก www.bkkpaperbox.com
- [9] Responsive Web design การใช้เทคนิค Responsive Web design ในการพัฒนา web application . (ออนไลน์). จาก <https://shorturl.asia/rvESJ>
- [10] MyAccount. (2560). การบริหารจัดการสต็อกคลัง. (ออนไลน์). จาก www.myaccount-cloud.com/
- [11] วิสปีดย์ ชัยช่วย. (2560). การใช้LINEของผู้สูงอายุ: การศึกษาเชิงปรากฏการณ์วิทยา. (บทความวิจัย). มหาวิทยาลัยขอนแก่น
- [12] ณัฐชยาน์ วิมลจิตรานนท์ และคณะ (2565) โปรแกรมอรรถประโยชน์เพื่อการ การจัดการคลังสินค้า ในสหกรณ์และร้าน.
- [13] วิริษฐา จิตราภิรมย์. (2562) ประสิทธิภาพในการจัดการคลังสินค้าด้วยระบบบาร์โค้ดผ่านโปรแกรม WMS ของพนักงานบริษัท สपोर्टส เอ็นจิเนียริง แอนด์ รีทรีเอชั่น เอเชียจำกัด. (บทความวิจัย). คณะโลจิสติกส์ มหาวิทยาลัยบูรพา
- [14] เอก บำรุงศรี และ ธนชัย เกียรติทองศิริ. (2563) เว็บแอปพลิเคชัน ระบบจัดการสินค้า และการขายของร้านคิมแอนด์จีวมินิมาร์ท. (บทความวิจัย). คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม

เกมสำหรับการลดความเสี่ยงในการเกิดภาวะสมองเสื่อมสำหรับผู้สูงอายุ

A Game for Dementia Risk Reduction in Elders

สุนิตต์ยา เย็นทั่ว (Sunitta Yentour)¹ นที ยงยุทธ (Nathee Yongyut)²

กงกฤษ ปีตานนท์ (Konggrit Pitanon)³ และเดชาวุฒิ วานิชสรรพ (Dechawut Wanichsan)⁴

¹ภาควิชาเทคโนโลยีและนวัตกรรมทางการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี

^{2,3,4}สาขาวิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏรำไพพรรณี

¹sunitta.y@rbru.ac.th, ²nathee.y@rbru.ac.th, ³konggrit.p@rbru.ac.th, ⁴dechawut.w@rbru.ac.th

บทคัดย่อ

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาเกมฝึกบริหารสมองเพื่อช่วยลดความเสี่ยงในการเกิดภาวะสมองเสื่อมในผู้สูงอายุ 2) เพื่อเปรียบเทียบผลการทดสอบสภาพสมองเบื้องต้นก่อนและหลังการใช้เกมฝึกบริหารสมอง และ 3) เพื่อศึกษาความพึงพอใจของผู้สูงอายุที่มีต่อเกมฝึกบริหารสมอง ซึ่งได้ทำการทดสอบผู้สูงอายุจำนวน 30 คน จากโรงเรียนผู้สูงอายุ ต. พลับพลา อ. เมือง จ. จันทบุรี ด้วยวิธีการสุ่มอย่างง่าย ผลการทดลองพบว่าเกมที่พัฒนาขึ้นมีประสิทธิภาพตามเกณฑ์อยู่ในระดับดีมาก ผู้ร่วมวิจัยมีผลการทดสอบสภาพสมองเบื้องต้นหลังการใช้งานเกมสูงกว่าก่อนใช้งานเกม และระดับความพึงพอใจของผู้สูงอายุต่อเกมอยู่ในระดับดีมาก

คำสำคัญ: ภาวะสมองเสื่อม ผู้สูงอายุ เกม สังคมผู้สูงอายุ

Abstract

The purpose of this study were 1) to develop a brain-training game for demetia risk reduction in elders, 2) to compare brain testing score between before using and after using the game, and 3) to study satisfaction level of participants. After that, measurements were carried out, using simple random sampling, on a total of 30 elders from Elderly school, Plubpla, Muang, Chanthaburi. The experimental results revealed that efficienciency of the game was in very high level, the post-test brain-testing score of the participants was higher than the pre-test score of

them, lastly the elders felt satisfied after using the developed game at very high level.

Keywords: Dementia, Elders, Game, Aging Society.

1. บทนำ

จากข้อมูลการจำแนกจำนวนประชากรรายอายุ ในทะเบียนราษฎร ของสำนักบริหารการทะเบียน กรมการปกครอง ที่สรุปผลในเดือนมกราคม ปี 2566 พบว่า สัดส่วนของประชากรที่มีอายุ 60 ปีขึ้นไป คิดเป็นร้อยละ 19.28 (12,738,238 / 66,081,169) ของจำนวนประชากรทั้งหมด และสัดส่วนของประชากรที่มีอายุ 65 ปีขึ้นไป คิดเป็นร้อยละ 13.11 (8,664,897 / 66,081,169) ของจำนวนประชากรทั้งหมด [1] ในทางกลับกัน อัตราการเกิดโดยเฉลี่ยห้าปีซ้อนหลัง ช่วงปี 2561-2565 คือ 577,550 คน เท่านั้น ซึ่งสวนทางกับจำนวนประชากรที่กำลังก้าวเข้าสู่ อายุ 60 ปี อายุ 55 ถึง 59 ที่มีค่าเฉลี่ยของประชากรในแต่ละอายุอยู่ที่ 967,228 คน [1] จากข้อมูลเหล่านี้แสดงให้เห็นว่า โครงสร้างประชากรของประเทศไทยกำลังเปลี่ยนผ่านจากการเป็นสังคมผู้สูงอายุ (Aging society) ไปยังการเป็นสังคมสูงอายุโดยสมบูรณ์ (Aged society) ในอีกไม่กี่ปีข้างหน้า

ภาวะสมองเสื่อมเป็นชื่อกลุ่มอาการที่มีการทำงานของสมองในด้านความจำ ความคิด การใช้เหตุผล การใช้ภาษา และการรับรู้สิ่งแวดล้อม ผิดปกติไป ทำให้เกิดปัญหาทางความคิด การตัดสินใจ พฤติกรรม และ อารมณ์ผิดปกติไป มีผลต่อการใช้ชีวิต หรือการทำกิจวัตรประจำวัน จนใน

ที่สุดผู้ป่วยหลายคนช่วยเหลือตัวเองไม่ได้ ต้องมีผู้ดูแลอย่างใกล้ชิด ผลกระทบที่เกิดขึ้นที่เกิดจากที่ต้องดูแลผู้ป่วยที่มีภาวะสมองเสื่อมในบ้าน เกิดทั้งในด้านสังคมและด้านเศรษฐกิจ เนื่องจากผู้ป่วยสมองเสื่อมจำเป็นต้องมีผู้ดูแลอย่างใกล้ชิดตลอดเวลา รวมถึงค่าใช้จ่ายทางตรงและทางอ้อมที่ต้องใช้ในการดูแลรักษาผู้ป่วยอีกด้วย อัตราการเกิดภาวะสมองเสื่อมมีความสัมพันธ์กับอายุที่เพิ่มมากขึ้น โดยเฉพาะอย่างยิ่งกับผู้มีอายุ 60 ปีขึ้นไป [2] อันทำให้การเฝ้าระวังผู้ป่วยก่อนเกิดภาวะสมองเสื่อมมักกระทำกับผู้สูงอายุเป็นหลัก

จากปัญหาดังกล่าว ผู้วิจัยได้ตระหนักและให้ความสำคัญกับการลดความเสี่ยงในการเกิดภาวะสมองเสื่อมในผู้สูงอายุ ผู้วิจัยจึงได้ออกแบบและพัฒนาเกมฝึกบริหารสมองเพื่อช่วยชะลอการเกิดภาวะสมองเสื่อมในผู้สูงอายุ โดยได้ทำการทดลองกับผู้สูงอายุที่มีช่วงอายุที่แตกต่างกันจำนวน 30 คนในจังหวัดจันทบุรี เกมที่พัฒนาขึ้นได้รับการออกแบบเป็นอย่างดี มีการเลือกใช้สีและแบบอักษรที่มีความเหมาะสมในการใช้งานของผู้สูงอายุ การชะลอการเกิดภาวะสมองเสื่อมในผู้สูงอายุลงได้ จะเป็นแนวทางในการเสริมสร้างความเข้มแข็งให้ผู้สูงอายุสามารถใช้ชีวิตในสังคมปัจจุบันได้อย่างปกติสุข โดยไม่รู้สึกว่าตนเองเป็นภาระให้ลูกหลานและประเทศ

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 สังคมผู้สูงอายุ และสังคมผู้สูงอายุโดยสมบูรณ์

สังคมผู้สูงอายุ (Aging Society) จะเป็นสังคมที่มีประชากรอายุ 60 ปีขึ้นไปที่อยู่จริงในพื้นที่ต่อประชากรทุกช่วงอายุในพื้นที่เดียวกัน และ $\geq 10\%$ ขึ้นไป หรือมีประชากรอายุ 65 ปีขึ้นไปที่อยู่จริงในพื้นที่ต่อประชากรทุกช่วงอายุในพื้นที่เดียวกัน $\geq 7\%$ ขึ้นไป สังคมผู้สูงอายุโดยสมบูรณ์ (Aged Society) จะเป็นสังคมที่มีประชากรอายุ 60 ปีขึ้นไปที่อยู่จริงในพื้นที่ต่อประชากรทุกช่วงอายุในพื้นที่เดียวกัน $\geq 20\%$ ขึ้นไป หรือมีประชากรอายุ 65 ปีขึ้นไปที่อยู่จริงในพื้นที่ต่อประชากรทุกช่วงอายุในพื้นที่เดียวกัน $\geq 14\%$ ขึ้นไป



© 2022 United Nations, DESA, Population Division, Licensed under Creative Commons License CC BY 3.0 IGO. United Nations, DESA, Population Division, World Population Prospects 2022. <http://population.un.org/wpp/>

ภาพที่ 1: [3] แผนภูมิแสดงจำนวนประชากรไทยปีพุทธศักราช 2566

จากข้อมูลในภาพที่ 1 แสดงให้เห็นว่าประชากรที่อยู่ในช่วงอายุ 40-59 ปีที่เตรียมเข้าสู่ประชากรผู้สูงวัยนั้นมีอัตราส่วนที่เพิ่มขึ้นเป็นเท่าตัวเมื่อเทียบกับประชากรที่อยู่ในช่วงอายุ 60 ปีขึ้นไป อีกทั้งจำนวนประชากรที่อยู่ในช่วงอายุ 0-20 ปี ที่มีแนวโน้มที่ลดลง จากข้อมูลดังกล่าวนี้ทำให้ประเทศไทยจะต้องเตรียมตัวเพื่อรองรับสังคมผู้สูงอายุโดยสมบูรณ์ (Aged Society) ให้ทัน่วงที

2.2 แบบทดสอบสภาพสมองเบื้องต้น

เครื่องมือที่ใช้คัดกรองภาวะสมองเสื่อมคือแบบทดสอบสภาพสมองเบื้องต้นฉบับภาษาไทย (MMSE-Thai 2002) ของสมาคมโรคสมองเสื่อมแห่งประเทศไทย ประกอบด้วยคำถาม 11 ข้อ การแปลผลการทดสอบเพื่อวัดการมีภาวะสมองเสื่อมหรือไม่ โดยขึ้นอยู่กับผลรวมคะแนนของผู้ถูกสัมภาษณ์ โดยแบ่งเป็นต่าง ๆ ดังนี้ กรณีที่ไม่ได้เรียนหนังสือ คะแนนรวม ≤ 14 กรณีจบการศึกษาระดับประถมศึกษา คะแนนรวม ≤ 17 กรณีจบการศึกษากว่าระดับประถมศึกษา คะแนนรวม ≤ 22 ซึ่งมีการแปลผลได้ว่า สงสัยว่ามีภาวะสมองเสื่อม

2.3 เครื่องมือที่ใช้ในการพัฒนาสื่อเกม

โปรแกรมอะโดบี แอนิเมท (Adobe Animate) เป็นเครื่องมือสำหรับการสร้างเนื้อหาเชิงโต้ตอบบนเว็บเพจและบนสมาร์ตโฟนที่สามารถสร้างสภาพแวดล้อมของเกมสร้างภาพ เคลื่อนไหว และสามารถใส่เสียงภายในสื่อได้ พร้อมทั้งสามารถออกแบบ และเขียนโปรแกรมด้วยภาษาแอ็กชันสคริปต์ อีกทั้งสามารถเผยแพร่ไปยังแพลตฟอร์มต่าง ๆ เช่น HTML5, Canvas, WebGL, Flash/Adobe AIR และ SVG ได้อย่างรวดเร็ว ซึ่งทำให้เข้าถึงผู้ใช้งานได้มากขึ้น

2.4 งานวิจัยที่เกี่ยวข้อง

ธีรภพ บุญชนะธนาธิญ และคณะ [4] ได้ทำการศึกษาการพัฒนาสื่อเกมฝึกสมองสำหรับการป้องกันภาวะสมองเสื่อมในผู้สูงอายุ โดยได้คำนึงถึงทฤษฎีของสื่อที่เหมาะสมต่อการรับรู้ของผู้สูงอายุ และกระตุ้นการทำงานของสมองพร้อมด้วยการศึกษาลักษณะฟอนต์ที่เหมาะสมต่อประสิทธิภาพสำหรับผู้สูงวัยในบริบทตัวอักษรไทยบนหน้าจอแท็บเล็ตคอมพิวเตอร์ของริตติ นโซติ เทียนมงคล [5] พบว่าฟอนต์ TH Fah Kwang มีคุณสมบัติเด่น ดังนี้ 1) เป็นฟอนต์ที่มองเห็นได้ชัดเจนที่สุดในขนาดที่เล็กที่สุดที่ 4.54 พอยต์ 2) ผลการประเมินค่าความสามารถในการอ่านได้ถูกต้องมากที่สุด อยู่ที่ร้อยละ 59.0 และ 3) สามารถอ่านทั้งประโยคได้ถูกต้องในค่าความเร็วเฉลี่ย 3.05 คะแนน/วินาที ซึ่งเป็นการใช้เวลาเฉลี่ยในการอ่านสั้นที่สุดในการทดสอบ

วันเฉลิม พรหมพร และคณะ [6] ได้ศึกษาวิจัย วิเคราะห์ สังเคราะห์ ข้อมูลการออกแบบการออกแบบแอปพลิเคชันเกมออกกำลังกายสำหรับผู้สูงอายุในประเทศไทยใช้ป้องกันภาวะสมองเสื่อมชนิดอัลไซเมอร์ พบว่าประกอบโดยรวมของเกมมีเนื้อเรื่องที่ไม่ซับซ้อน การออกแบบภาพการใช้ สัญลักษณ์ สี และฟอนต์ที่เหมาะสม สามารถช่วยกระตุ้นให้กลับมาเล่นเกมอีกครั้ง นอกจากนี้ อุดม พงษ์วุฒิชัยธรรม. [7] กล่าวว่าหลักวิธีการช่วยจำในผู้สูงอายุ เช่น ฝึกจำรูปภาพ ฝึกจำความเหมือนหรือความต่างของภาพ สามารถฝึกการทำงานสมองส่วนความจำได้

3. วิธีดำเนินการวิจัย

3.1 การสร้างและหาคุณภาพเครื่องมือ (ดังภาพที่ 2)

- 1) ศึกษารูปแบบเกมที่เหมาะสมกับผู้สูงอายุ
- 2) ออกแบบองค์ประกอบของเกม ได้แก่ ผังงาน การออกแบบหน้าจอ การเลือกใช้ฟอนต์ ขนาดตัวอักษร การใช้สี รูปภาพและรูปแบบการเล่นเกม
- 3) พัฒนาเกมโดยการเตรียมทรัพยากรต่าง ๆ ได้แก่ รูปภาพ เนื้อหา เสียง และโปรแกรมพัฒนาเกม จากนั้นทำการพัฒนาด้วยโปรแกรมอะโดบี แอนิเมท ด้วยภาษาแอ็คชันสคริปต์ 3.0 (Action script 3.0) ในรูปแบบ 2 มิติ

4) นำเกมที่สร้างเสร็จแล้วไปประเมินคุณภาพสื่อโดยผู้เชี่ยวชาญ จำนวน 5 ท่าน ด้วยแบบประเมินมาตราส่วนประมาณค่า 5 ระดับ แล้วทำการปรับแก้ตามคำแนะนำของผู้เชี่ยวชาญ

3.2 วิธีดำเนินการเก็บรวบรวมข้อมูล

1) นำเกมที่พัฒนาเสร็จแล้วไปทำกิจกรรมกับกลุ่มตัวอย่างซึ่งเป็นผู้สูงอายุที่มีอายุตั้งแต่ 60 ปีขึ้นไป ในโรงเรียนผู้สูงอายุสังกัดเทศบาลตำบลพลับพลา อ.เมือง จ.จันทบุรี จำนวน 30 ท่าน ดังนี้

1.1) ก่อนเริ่มทำกิจกรรมให้ผู้สูงอายุทำแบบทดสอบสภาพสมองเสื่อมเบื้องต้น ฉบับภาษาไทย MMSE-Thai 2002

1.2) ทำกิจกรรมกับผู้สูงอายุจำนวน 3 ครั้ง ในวันเสาร์อาทิตย์ โดยใช้เวลาดำเนินกิจกรรมประมาณ 2 ชั่วโมงต่อครั้ง โดยเริ่มจากแนะนำวิธีการเล่นเกมให้ผู้สูงอายุเล่นเกมที่พัฒนาขึ้น

1.3) ภายหลังจากทำกิจกรรมในครั้งที่ 3 เสร็จสิ้นให้ผู้สูงอายุทำแบบทดสอบสภาพสมองเสื่อมเบื้องต้น ซึ่งเป็นชุดเดียวกับก่อนทำกิจกรรมอีกครั้งหนึ่ง

1.4) ให้ผู้สูงอายุทำแบบประเมินความพึงพอใจที่มีต่อเกม

3.3 การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลโดยใช้สถิติค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและการทดสอบค่าที

เครื่องมือที่ใช้ในการวิจัย คือ 1) เกมฝึกบริหารสมองเพื่อช่วยชะลอการเกิดภาวะสมองเสื่อมในผู้สูงอายุ 2) แบบทดสอบสภาพสมองเสื่อมเบื้องต้น ฉบับภาษาไทย MMSE-Thai 2002 ซึ่งเป็นแบบทดสอบ จัดทำโดย สถาบันเวชศาสตร์ผู้สูงอายุ กรมการแพทย์ กระทรวงสาธารณสุข และ 3) แบบประเมินความพึงพอใจของผู้สูงอายุที่มีต่อเกมฝึกบริหารสมองฯ

การออกแบบกลไกการทำงานของเกม ทำโดยกำหนดให้มีความยาก 3 ระดับ คือ ระดับง่ายมีภาพให้จับคู่จำนวน 3 คู่ ระดับปานกลางมีภาพให้จับคู่จำนวน 4 คู่ ระดับยากมีภาพให้จับคู่จำนวน 5 คู่ (ดังตารางที่ 1) ในแต่ละระดับจะประกอบด้วยด่านย่อยจำนวน 5 ด่านและมีภาพที่แตกต่างกัน เงื่อนไขการเล่นคือผู้เล่นต้องพลิกการ์ด



ภาพที่ 2 : วิธีดำเนินการวิจัย

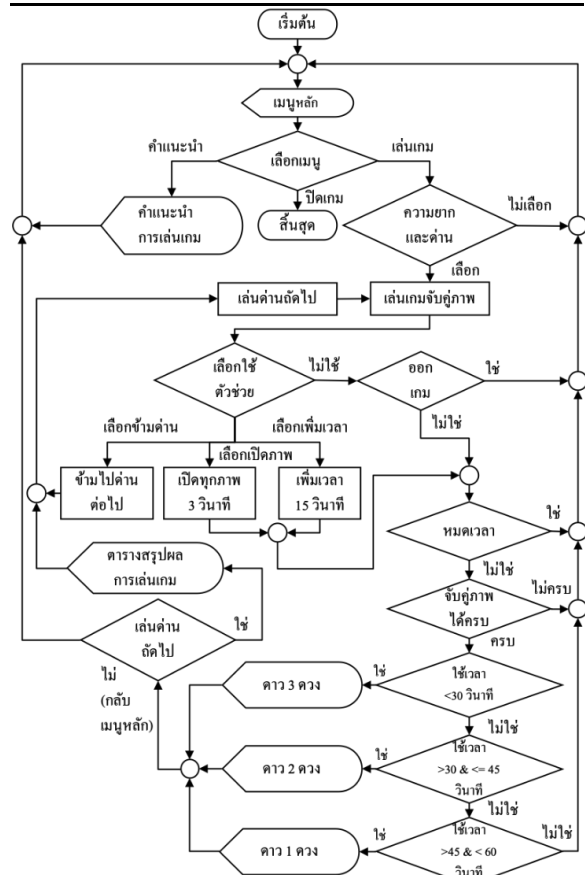
เพื่อแสดงภาพให้ตรงกันเป็นคู่ หากปรากฏภาพที่เหมือนกันจะทำให้ภาพคู่ นั้นหายไป เวลาในการเล่น 60 วินาทีต่อด่าน มีตัวช่วย 3 วิธี คือ 1) ตัวช่วยเพิ่มเวลา 15 วินาที 2) ตัวช่วยเปิดทุกภาพ 3 วินาที และ 3) ตัวช่วยข้ามไปด่านต่อไป ในกรณีที่จับคู่ภาพไม่ครบและหมดเวลาให้กลับไปเล่นด่านนั้นอีกครั้ง หากจับคู่ภาพได้ครบจะแสดงหน้าต่างจำนวนดาวที่ได้รับ ซึ่งขึ้นอยู่กับความเร็วในการเล่น ถ้าใช้เวลาน้อยกว่า 30 วินาที ได้รับดาว 3 ดวง หรือใช้เวลามากกว่า 30 วินาที และน้อยกว่า 45 วินาที ได้รับ

ดาว 2 ดวง และใช้เวลามากกว่า 45 วินาที และไม่ถึง 60 วินาที ได้รับดาว 1 ดวง หลังจากนั้นจะไปสู่นำตารางสรุปผลการเล่นรวมว่าได้ผ่านด่านใดมาบ้างแล้วและด่านนั้นได้รับดาวจำนวนกี่ดวง ผู้เล่นสามารถเล่นซ้ำในแต่ละด่านได้ตามความต้องการเพื่อให้มีผลการเล่นดีที่สุดและสามารถออกจากเกมได้ตลอดเวลา ดังภาพที่ 3

วิจัยเรื่องนี้ได้รับทุนสนับสนุนจาก กองทุนวิจัยมหาวิทยาลัยราชภัฏรำไพพรรณี ประจำปีงบประมาณ พ.ศ. 2565 เครื่องมือวิจัยได้รับการรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยราชภัฏรำไพพรรณี เลขที่ IRB-39/2565 วันที่ 27 ตุลาคม

ตารางที่ 1 ระดับความยากของเกม

ระดับ	คำอธิบาย
ง่าย	จับคู่ภาพที่แตกต่างกันจำนวน 3 คู่
ปานกลาง	จับคู่ภาพที่แตกต่างกันจำนวน 4 คู่
ยาก	จับคู่ภาพที่แตกต่างกันจำนวน 5 คู่



ภาพที่ 3 : ผังงานกลไกการทำงานของเกม

4. ผลการดำเนินการวิจัย

4.1 ผลการพัฒนาเกม

เกมที่พัฒนาขึ้นมีการออกแบบที่คำนึงถึงการใช้งานสำหรับผู้สูงอายุโดยเฉพาะ ได้แก่ การเลือกใช้ฟอนต์ TH Fah Kwang ซึ่งมีส่วนสัดการอ่านถูกต้องมากที่สุด อยู่ที่ร้อยละ 59.0 [5] ขนาดตัวอักษรมีขนาดใหญ่ สามารถอ่านผ่านหน้าจอสมาร์ทโฟนได้อย่างชัดเจน การเลือกใช้สีที่มีมองเห็นสบายตาอยู่ในโทนสีชมพูเขียว ซึ่งเป็นสีประจำมหาวิทยาลัยราชภัฏรำไพพรรณีอีกด้วย นอกจากนี้ภาพที่ใช้สำหรับการ์ดแต่ละใบเป็นภาพที่วาดขึ้นมาเองมีสีสันที่สดใส สวยงาม น่ารัก ทำให้ผู้สูงอายุที่เล่นเกมมีความผ่อนคลายขณะเล่นเกม ลักษณะการออกแบบหน้าต่างของเกมแสดงดังภาพที่ 2-6 ส่งผลให้ผลการประเมินคุณภาพสื่อโดยผู้เชี่ยวชาญ ประกอบด้วยด้านเนื้อหาและการนำเสนอ ด้านการออกแบบจอภาพและด้านการจัดการในเกมคอมพิวเตอร์ มีค่าเฉลี่ยรวมเท่ากับ 4.67 ค่าเบี่ยงเบนมาตรฐาน 0.48 อยู่ในระดับดีมาก ดังตารางที่ 2

ตารางที่ 2 ผลการประเมินคุณภาพสื่อโดยผู้เชี่ยวชาญ 5 ท่าน

รายการประเมิน	ค่าเฉลี่ย	SD	ระดับ
ด้านเนื้อหาและการนำเสนอ			
1. การจัดลำดับเนื้อหาจากง่ายไปหายาก	5.00	0.00	ดีมาก
2. ความถูกต้องของเนื้อหา	4.67	0.58	ดีมาก
3. เนื้อหาเหมาะสมกับผู้เล่นเกม	4.67	0.58	ดีมาก
ด้านการออกแบบจอภาพ			
1. ความชัดเจนของอักษร	4.67	0.58	ดีมาก
2. ความชัดเจนของรูปภาพ	5.00	0.00	ดีมาก
3. เสียงประกอบในเกมมีความเหมาะสม	4.67	0.58	ดีมาก
4. การใช้สีที่ดึงดูดความสนใจ	4.67	0.58	ดีมาก
ด้านการจัดการในเกมคอมพิวเตอร์			
1. ความชัดเจนของคำแนะนำการเล่น	4.67	0.58	ดีมาก
2. ความสะดวกเมื่อเล่นกับคอมพิวเตอร์	4.00	0.00	ดี
3. ความเหมาะสมของการมีปฏิสัมพันธ์ระหว่างผู้เล่นกับเกม	4.67	0.58	ดีมาก

รายการประเมิน	ค่าเฉลี่ย	SD	ระดับ
4. ความเหมาะสมของระยะเวลาเล่นเกม	4.33	0.58	ดี
5. สอดคล้องกับการประยุกต์ใช้ใน	5.00	0.00	ดีมาก
การจัดการเรียนรู้			
ค่าเฉลี่ยทั้งหมด	4.67	0.48	ดีมาก

1) ตัวอย่างหน้าต่างเกม แสดงดังภาพที่ 4-9 โดยหน้าแรก (ภาพที่ 4) ของเกมจะมีชื่อเกมและปุ่มให้กดเข้าไปเล่นเกมหรืออ่านคำแนะนำการเล่นเกม ถ้ากดปุ่มคำแนะนำจะแสดงหน้าต่างคำอธิบายกติกาและเงื่อนไขการเล่น คำอธิบายการใช้ตัวช่วย (ภาพที่ 5 และ 6) ถ้ากดปุ่มเริ่มเล่นเกมจะเข้าสู่หน้าต่างการเล่นตามระดับความยากที่เลือก (ภาพที่ 7 และ 8) และหน้าต่างสรุปผลการเล่นของแต่ละด่านที่ผ่านมา (ภาพที่ 9)



ภาพที่ 4 : หน้าแรกของเกม



ภาพที่ 5: หน้าต่างอธิบายกติกาและเงื่อนไขการเล่น



ภาพที่ 6: แสดงหน้าต่างคำอธิบายการใช้ตัวช่วย



ภาพที่ 7: หน้าต่างของด่านที่ 1



ภาพที่ 8: แสดงตัวอย่างการจับคู่ภาพระดับยาก



ภาพที่ 9: แสดงหน้าต่างสรุปผลการเล่นเกม

4.2 ผลการเปรียบเทียบการทดสอบสภาพสมองเบื้องต้น

ตารางที่ 3: ผลการเปรียบเทียบการทดสอบสภาพสมองเบื้องต้น

การทดสอบ	n	\bar{X}	SD	t	p
ก่อนใช้เกม	30	24.83	3.47	-4.267	.000*
หลังใช้เกม	30	25.73	2.89		

* มีนัยสำคัญทางสถิติที่ระดับ .05

จากตารางที่ 3 พบว่า หลังใช้เกมผู้สูงอายุมีคะแนนผลการทดสอบสภาพสมองเบื้องต้นสูงกว่าก่อนใช้ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยคะแนนเฉลี่ยหลังใช้เกม 25.73 คะแนน และก่อนใช้เกม 24.83 คะแนน

4.3 ผลการประเมินความพึงพอใจ

ตารางที่ 4: ผลการประเมินความพึงพอใจของผู้สูงอายุที่มีต่อเกม

รายการประเมิน	ค่าเฉลี่ย	SD	ระดับ
1. มีวัตถุประสงค์ชัดเจน	4.67	0.61	ดีมาก
2. เกมมีสีสันสดใส	4.50	0.82	ดีมาก
3. ตัวอักษรอ่านง่าย ชัดเจน	4.50	0.73	ดีมาก
4. ภาษาที่ใช้เข้าใจง่าย	4.53	0.68	ดีมาก
5. ภาพและเสียงประกอบชัดเจน	4.63	0.49	ดีมาก
6. เกมมีความน่าสนใจ น่าติดตาม	4.73	0.52	ดีมาก
7. กฎ กติกา เงื่อนไขต่าง ๆ เข้าใจง่าย	4.47	0.78	ดี
8. วิธีการเล่นไม่ยุ่งยาก ซับซ้อน	4.70	0.60	ดีมาก
9. เวลาแต่ละด่านมีความเหมาะสม	4.67	0.61	ดีมาก
10. ได้รับ ความ สนุก ส นาน เพลิดเพลิน	4.73	0.52	ดีมาก
11. ได้รับการฝึกสมองด้านความจำ	4.83	0.46	ดีมาก
12. ช่วยให้มีความจำที่ดีขึ้น	4.80	0.41	ดีมาก
13. การปฏิสัมพันธ์ ผู้เล่นกับเกม	4.73	0.52	ดีมาก
14. การรายงานสรุปผลการเล่นเกม	4.63	0.56	ดีมาก
15. ภาพรวมที่มีต่อเกม	4.83	0.38	ดีมาก
ค่าเฉลี่ยทั้งหมด	4.66	0.59	ดีมาก

5. สรุป

ประชากรของประเทศไทยกำลังเปลี่ยนผ่านจากการเป็นสังคมผู้สูงอายุ ไปยังการเป็นสังคมสูงอายุโดยสมบูรณ์ ภาวะสมองเสื่อมเป็นสิ่งที่บั่นทอนความสามารถในการใช้ชีวิตประจำวันของผู้สูงอายุ การเตรียมความพร้อม

เพื่อลดผลกระทบดังกล่าวด้วยการสร้างประชากรผู้สูงอายุที่มีคุณภาพทั้งด้านร่างกาย และจิตใจจึงเป็นสิ่งสำคัญ ผู้วิจัยจึงพัฒนาเกมฝึกบริหารสมองที่มีประสิทธิภาพอยู่ในระดับดีมาก ซึ่งจะเป็นเครื่องมือหนึ่งที่สามารถชะลอการเกิดภาวะสมองเสื่อมในผู้สูงอายุได้ โดยผลการเปรียบเทียบสะท้อนให้เห็นสภาพสมองที่ดีขึ้นหลังการฝึกสมองโดยใช้เกมกับผู้สูงอายุทั้ง 30 คน นอกจากนี้เกมยังสร้างความพึงพอใจให้ผู้สูงอายุในระดับมากที่สุดด้วย

เอกสารอ้างอิง

- [1] กรมการปกครอง. "สถิติประชากรทางการทะเบียนราษฎร(รายเดือน)". Internet: <https://stat.bora.dopa.go.th/stat/statnew/statMONTH/statmonth/>. [Jan 10, 2023]
- [2] C.P. Ferri., M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, et.al., "Global prevalence of dementia : a delphiconsensus study," *Lancet*, vol 366, no.9503, pp.2112-2117, Dec.2005
- [3] United Nations, Department of Economic and Social Affairs, Population Division . "World Population Prospects: The 2022 Revision". Internet: <http://population.un.org/wpp/>. [Mar 10, 2023]
- [4] ชีรภพ บุญยะธนาธัญ, พงษ์ดนัย วิชัยสุนทร และเชษฐาภาณี วาณิชสรพรพ์. "เกมฝึกสมองสำหรับการป้องกันภาวะสมองเสื่อมในผู้สูงอายุ". *The Sixteenth National Conference on Computing and Information Technology (NCCIT 2020)*, Bangkok, May 2020.
- [5] รัตนาโชติ เทียนมงคล. "การศึกษาลักษณะฟอนต์ที่เหมาะสมต่อประจักษ์ภาพสำหรับผู้สูงวัยในบริบทตัวอักษรไทยบนหน้าจอแท็บเล็ตคอมพิวเตอร์". *Veridian E-Journal, Silpakorn University*, Vol. 10, No. 3 (2017), pp.1066-1082, Jun 2017.
- [6] วันเฉลิม พรหมสร, บุญชู บุญลิขิตศิริ, ปรัชญา แก้วแก่น. "การออกแบบแอปพลิเคชันเกมส้อมอกกำลังกายสำหรับผู้สูงอายุในประเทศไทยใช้ป้องกันภาวะสมองเสื่อมชนิดอัลไซเมอร์". *วารสารศิลปกรรมบูรพา*, Vol. 24, No. 2 (2021): กรกฎาคม – ธันวาคม 2564, pp.155-171, Dec 2022.
- [7] อุดม พงษ์วุฒิธรรม. "การพยาบาลเพื่อการฟื้นฟูสภาพสมองผู้สูงอายุไทย". *วารสารการพยาบาล*, Vol. 14, No. 4 (2542) : ตุลาคม – ธันวาคม 2542, pp. 64-71.

สมาร์ตไลน์แชทบอทและเทคนิคการแก้ไขคำผิดแบบอัตโนมัติ

A Smart LINE Chatbot and Techniques for Auto Correct

จิรศักดิ์ พิพัฒน์วารากุล (Jirasak Pipatwarakul)¹ พยุง มีสังข์ (Phayung Meesad)²
ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
¹s6407021857026@email.kmutnb.ac.th, ²phayung.m@itd.kmutnb.ac.th

บทคัดย่อ

การให้บริการลูกค้าหลังการขายผ่านช่องทางไลน์แอปพลิเคชันจำเป็นต้องมีพนักงานคอยตอบคำถามให้กับลูกค้าอยู่ตลอดเวลา และพบว่าบ่อยครั้งลูกค้ามีการพิมพ์ภาษาไทยไม่ถูกต้อง และใช้ภาษาแชทในการสื่อสารจึงทำให้ไลน์แชทบอทไม่เข้าใจข้อความในการสนทนา โดยงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อ 1) นำเสนอโมเดลสมาร์ตไลน์แชทบอทและเทคนิคการแก้ไขคำผิดแบบอัตโนมัติ 2) เปรียบเทียบประสิทธิภาพโมเดลสมาร์ตไลน์แชทบอทและเทคนิคการแก้ไขคำผิดอัตโนมัติ 3) สร้างไลน์แชทบอทสนับสนุนระบบตอบคำถาม โดยใช้เทคนิคการแก้ไขคำผิดอัตโนมัติ โดยในการวิจัยครั้งนี้ได้แบ่งงานวิจัยออกเป็น 2 ส่วนดังนี้ 1) ส่วนของการเปรียบเทียบและคัดเลือกเทคนิคในการจัดการข้อมูลได้แก่ เทคนิคการสะกดคำ, เทคนิคการตัดคำ, เทคนิคการกำจัดคำหยุด, และการปรับความไม่สมดุลของชุดข้อมูล โดยรวบรวมคำถามที่ลูกค้าได้เคยมีการสอบถามเข้ามาจำนวน 5,000 คำถาม จากนั้นนำข้อมูลที่ได้มาเรียนรู้เชิงลึกด้วยโมเดล LSTM และ GRU ทำการทดสอบเพื่อเปรียบเทียบประสิทธิภาพพบว่า โมเดล LSTM ให้ประสิทธิภาพดีที่สุด โดยมีค่าความถูกต้องเท่ากับร้อยละ 86.23 ค่าความแม่นยำเท่ากับร้อยละ 97.54 ค่าระลึกเท่ากับร้อยละ 86.23 และค่าถ่วงดุลเท่ากับร้อยละ 87.34 2) พัฒนาระบบไลน์แชทบอทสำหรับทดสอบการทำงาน โดยนำเทคนิคและโมเดลที่ได้จากส่วนที่ 1 มาใช้ในการพัฒนาและเชื่อมต่อกับระบบไลน์แชทบอทเพื่อใช้แสดงผลในการรับ-ส่งข้อความ

คำสำคัญ: ไลน์แอปพลิเคชัน เทคนิคการสะกดคำ เรียนรู้เชิงลึก

Abstract

Providing after-sales customer service through the LINE Application channel requires employees to answer questions for customers. Often, customers did not type sentences correctly and used chat language to communicate, causing LINE Chatbots to misunderstand real messages. The objectives of this research were 1) to propose a Smart LINE Chatbots model using an automatic error-correction technique, 2) to measure the performance of the proposed technique based on accuracy, precision, recall, and F-measure, and 3) to design and develop LINE Chatbot application based on the proposed automatic error correction techniques. This research consists of two parts. Part 1 was the comparison of data management techniques, spelling techniques, segmentation techniques, stop word techniques, and imbalanced datasets. By performing so, this research collected 5,000 questions that customers had previously asked. The data were taken through deep learning with the LSTM and GRU models and tested to compare the performance. The results showed that the LSTM model was the best performance with the best result accuracy of 86.23%, a precision of 97.54%, a recall of 86.23%, and an F-measure of 87.34%. Part 2 was to develop the LINE Chatbot system to test the functionality and apply techniques and models obtained from Part 1 to develop and connect to the LINE Chatbot system to display results in sending and receiving messages.

Keywords: LINE Application, Spelling Techniques, Deep Learning.

1. บทนำ

ปัจจุบันการสื่อสารได้ก้าวเข้าสู่ยุคดิจิทัลเทคโนโลยีการสื่อสารสังคมออนไลน์ได้เข้ามามีบทบาทมากขึ้นในชีวิตประจำวันและการดำเนินธุรกิจ พฤติกรรมของผู้คนในปัจจุบันมีการปรับเปลี่ยนไปตามการนำเสนอสินค้าและบริการของผู้ประกอบการที่ต้องเดินทางตามเทคโนโลยีที่เติบโตในอุตสาหกรรมส่งผลให้เทคโนโลยีในการสื่อสารสังคมออนไลน์ได้มีการปรับเปลี่ยนไปอย่างรวดเร็ว ซึ่งสามารถเห็นข่าวสารได้จากสื่อโฆษณาดิจิทัล สื่อสังคมออนไลน์ไม่ว่าจะเป็น เฟสบุ๊ก ยูทูบ เว็บไซต์บนอินเทอร์เน็ต และไลน์แอปพลิเคชัน เป็นต้น

เมื่อไลน์แอปพลิเคชันเข้ามามีบทบาทมากขึ้นถือเป็นช่องทางในการทำการตลาดที่สำคัญในยุคดิจิทัล มีการใช้งานติดต่อสื่อสารกันผ่านไลน์แอปพลิเคชันเป็นจำนวนมาก รวมถึงองค์กรต่าง ๆ นิยมใช้เป็นการประชาสัมพันธ์ธุรกิจในการเสนอกิจกรรมส่งเสริมการขายและการให้บริการหลังการขาย เพื่ออำนวยความสะดวกให้กับลูกค้า และในบางธุรกิจได้ใช้ไลน์แอปพลิเคชันเป็นช่องทางหลักในการติดต่อสื่อสารระหว่างเจ้าของธุรกิจและลูกค้าด้วยเช่นกัน

การให้บริการหลังการขายกับลูกค้าผ่านช่องทางไลน์แอปพลิเคชัน จำเป็นต้องมีการพัฒนาไลน์แอปพลิเคชัน [1] ขึ้นมา ซึ่งขั้นตอนการพัฒนาแอปพลิเคชันต้องอาศัยเทคนิคและวิธีการที่หลากหลาย ได้แก่ การประมวลผลภาษาธรรมชาติ [2] ทำให้ไลน์แชทบอทสามารถเข้าใจภาษามนุษย์ได้มากขึ้น และได้ตอบข้อความได้ถูกต้องมากขึ้น แต่ในการสนทนาผ่านไลน์แอปพลิเคชันเป็นการพิมพ์ข้อความเพื่อการสนทนาบ่อยครั้งลูกค้ายังมีการพิมพ์ภาษาไทยที่ผิดและใช้ภาษาแชทในการสื่อสาร [3] หากผู้ใช้งานมีการพิมพ์ตัวอักษรผิดส่งผลให้ความหมายในประโยคไม่ตรงกับคำสำคัญที่ต้องการและไลน์แชทบอทไม่เข้าใจข้อความในการสนทนา

จากปัญหาดังกล่าวผู้วิจัยได้เห็นถึงปัญหาการพิมพ์ข้อความโดยมีคำผิดเกิดขึ้นจึงได้นำวิธีการแก้ไขคำผิด [4] เข้ามาใช้งาน ซึ่งเป็นการพยากรณ์คำผิดเพื่อสนับสนุนให้สมาร์ทโฟนแชทบอทให้เข้าใจบทสนทนาได้มากขึ้น อีกทั้งยังมีความถูกต้อง ความแม่นยำในการตอบคำถาม กระบวนการเทคนิคต่าง ๆ ที่ได้มีการนำมาใช้จะช่วยทำให้เข้าใจข้อมูลภาษาธรรมชาติและสามารถนำเสนอคำตอบที่เหมาะสมให้กับสมาร์ทโฟนแชทบอท

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 เทคนิคการสะกดคำ (Spelling)

เป็นเทคนิคที่ใช้ในการตรวจตรวจคำผิดจากข้อความหรือประโยคว่ามีคำเหล่านั้นอยู่ในพจนานุกรมหรือไม่ โดยนำคำใกล้เคียงที่สุดในพจนานุกรมเป็นคำที่ถูกต้องและเลือกให้โดยอัลกอริทึม [4] เรียกว่า Spelling Correction ยกตัวอย่างเช่น

2.1.1 อัลกอริทึมของ Peter Norvig เป็นการค้นหารายการคำใกล้เคียงจาก Dictionary โดยใช้จำนวนอักษรที่ผิดผสมกับความน่าจะเป็นจากความถี่ของคำนั้นที่ปรากฏในคลังข้อมูล โดยอัลกอริทึมของ Peter Norvig นั้นไม่ได้ใช้ Context หรือคำแวดล้อมที่มาก่อนหน้าหรือต่อจากนั้น และไม่ได้ใช้ตำแหน่งปุ่มใกล้เคียงบนคีย์บอร์ดมาคำนวณความน่าจะเป็น

2.1.2 การตรวจการสะกดคำด้วย Custom Dictionary และ Word Frequency เป็นการใช้พจนานุกรมจาก Thai National Corpus (TNC) ซึ่งมีคำศัพท์ในคลังคำศัพท์จำนวน 4 หมื่นคำ โดยมีการกำหนดพารามิเตอร์เพื่อกรองคำศัพท์ที่มีความถี่และความยาวนานเกินไปที่ส่งก่อนนำไปใช้งาน

2.1.3 Spelling Correction คือการเลือกคำที่มีความน่าจะเป็นสูงสุด โดยเป็นคำแรกของรายการที่ได้จากการตรวจการตรวจสอบคำ (Spellchecker) ทำงานโดยชี้ไปยังตำแหน่งที่สะกดผิด [5]

2.2 เทคนิคการตัดคำ (Segmentation)

เป็นเทคนิคที่ใช้ในการแบ่งตัวอักษรจากข้อความที่ต่อเนื่องกันออกเป็นหน่วยคำ เนื่องจากรูปแบบของภาษาไทยมีการเขียนในลักษณะที่คำติดกัน โดยไม่มีการใช้

เครื่องหมายวรรคตอนคั่นระหว่างคำเหมือนภาษาอังกฤษที่มีการใช้ช่องว่างคั่นระหว่างคำ แต่ภาษาไทยจะมีการเว้นวรรคเป็นระยะเพื่อให้ผู้อ่านทำความเข้าใจกับความหมายของข้อความ [6]

2.3 เทคนิคการกำจัดคำหยุด (Stop word)

เป็นเทคนิคที่ใช้ในการนำคำที่ไม่มีนัยสำคัญออกโดยไม่ทำให้ความหมายของประโยคข้อความมีการเปลี่ยนแปลงไป เป็นคำที่ไม่มีนัยสำคัญในที่นี้หมายถึงคำที่ใช้กันโดยทั่วไปไม่มีความหมายสำคัญต่อรูปของประโยคข้อความทำให้สามารถลบคำเหล่านั้นออกไปจากรายการคำศัพท์ของประโยคข้อความได้ โดยไม่ทำให้ใจความสำคัญเปลี่ยนแปลงไปถือได้ว่าคำหยุดเป็นคุณลักษณะที่ไม่เกี่ยวข้องหรือไม่มีประโยชน์ในการนำมาใช้ร่วมกับการประมวลผล โดยการนำคำไปเปรียบเทียบกับคลังข้อมูลภาษาไทย (Thai Stopwords Corpus) ซึ่งช่วยกำจัดคุณลักษณะที่ไม่เป็นประโยชน์ ส่งผลให้ประหยัดทั้งพื้นที่, เวลาในการประมวลผล, และลดขนาดของดัชนีลง ตัวอย่างเช่น “หนอย”, “ครับ” [7]

2.4 การปรับความไม่สมดุลของชุดข้อมูล

เป็นการปรับเพิ่มข้อมูลด้วยวิธีการสุ่มเพื่อช่วยในการปรับสมดุลข้อมูลโดยเพิ่มข้อมูลของกลุ่มที่มีจำนวนข้อมูลน้อยและทำการสร้างข้อมูลเทียมเพื่อให้ได้จำนวนข้อมูลของแต่ละกลุ่มมีจำนวนเท่ากับจำนวนของข้อมูลของกลุ่มที่มากที่สุด [8]

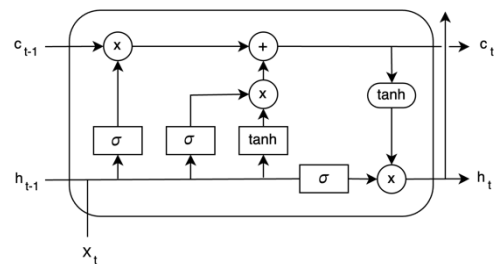
2.5 การเรียนรู้เชิงลึก (Deep learning)

การเรียนรู้เชิงลึกเป็นส่วนหนึ่งของการเรียนรู้ของเครื่องจักร (Machine Learning) มีพื้นฐานมาจากโครงข่ายประสาทเทียม (Artificial Neural Network : ANN) ซึ่งเป็นวิธีที่สร้างขึ้นเพื่อทำให้เครื่องจักรสามารถเรียนรู้ได้ โดยใช้ต้นแบบมาจากการทำงานของโครงข่ายประสาทของมนุษย์ และต้องใส่ข้อมูลเข้าไปในชั้นรับข้อมูล (Input Layer) จากนั้นนำข้อมูลที่ได้ออกไปประมวลผลที่มีการซ้อนกันหลาย ๆ ชั้น (Hidden Layer) และนำเสนอข้อมูลผลลัพธ์ในชั้นแสดงผล (Output Layer) โครงข่ายประสาทเทียมได้มีการพัฒนาอย่างต่อเนื่องซึ่งวิธีที่นิยมใช้ใน

ปัจจุบันคือ Long Short-Term Memory (LSTM) และ Gated Recurrent Unit (GRU) [9]

2.5.1 Long Short-Term Memory

เป็นโครงข่ายประสาทเทียมแบบหนึ่งที่ถูกออกแบบมาสำหรับการประมวลผลลำดับ (sequence) โดยมีพื้นฐานมาจาก Recurrent Neural Network (RNN) ซึ่งได้นำมาช่วยในการแก้ไขปัญหา Vanishing Gradient ที่ถูกพบเมื่อใช้วิธี RNN กับข้อมูลที่มีความยาว โดย LSTM ประกอบด้วย input gate, output gate และ forget gate ซึ่งเป็นสิ่งที่ควบคุมการไหลของข้อมูล โดยที่ LSTM ทำการรับข้อมูลมาจากชั้น input เป็นครั้งแรกและเข้าสู่ input gate ไปจนถึง output gate เพื่อตัดสินใจว่าจะเก็บค่าที่ได้แล้วนำไปวนซ้ำใน LSTM หรือแสดงผลข้อมูล หากเลือกที่วนซ้ำจะนำข้อมูลกลับมาเพื่อเข้าสู่ forget gate ซึ่งเป็นขั้นตอนในการตัดสินใจว่าจะลบค่าที่เก็บไว้ทิ้งไปหรือยังคงเก็บค่าไว้ หากเก็บไว้จะปรอการอัปเดตจาก input gate ซึ่งจะตัดสินใจว่าจะอัปเดตค่านั้นหรือไม่และจะอัปเดตด้วยค่าอะไร เพื่อนำค่าที่ได้ส่งไป output gate และตัดสินใจอีกครั้งว่าจะนำข้อมูลนั้นออกไปแสดงหรือนำข้อมูลกลับไปวนซ้ำอีกครั้ง ดังนั้น LSTM จึงสามารถเรียนรู้จากข้อมูลที่เป็นลำดับและเก็บหรือลบข้อมูลทิ้งถ้าข้อมูลนั้นไม่จำเป็น [10] ดังภาพที่ 1

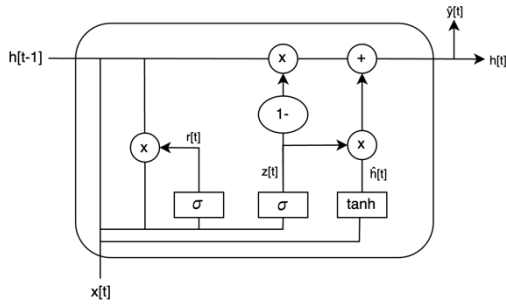


ภาพที่ 1: การทำงานของ LSTM

2.5.2 Gated Recurrent Unit

GRU มีหลักการการทำงานคล้ายกับ LSTM แต่มีความเร็วที่มากกว่าเพราะมีการตัด input gate กับ output gate ออกในการวนซ้ำและเปลี่ยนมาใช้ reset gate กับ update gate ซึ่งช่วยลดการทำงานที่มีความซับซ้อนของโครงข่ายประสาทเทียม โดยเริ่มจาก input เข้ามาครั้งแรกจะทำการเก็บค่านั้นไว้แล้วตัดสินใจว่าจะนำข้อมูลไปวนซ้ำใน GRU

หรือ แสดงผลหากนำข้อมูลไปวนซ้ำจะเข้าสู่ reset gate เพื่อตัดสินใจว่าต้องลบข้อมูลไหนและเก็บข้อมูลไหนไว้ หลังจากนั้นจะเข้าสู่ update gate เพื่อตัดสินใจว่าจะแสดงผลข้อมูลส่วนไหนแล้วนำข้อมูลไปตัดสินใจว่าจะแสดงผลหรือนำข้อมูลกลับไปวนซ้ำอีกครั้ง [10] ดังภาพที่ 2



ภาพที่ 2: การทำงานของ GRU

2.6 การประเมินประสิทธิภาพของแบบจำลอง

มาตรฐานที่ใช้ในการประเมินหรือการวัดประสิทธิภาพของแบบจำลอง (Model) มีการแบ่งหมวดหมู่ของผลลัพธ์โดยนำข้อมูลผลลัพธ์จริงและผลลัพธ์ที่ได้จากแบบจำลองมาเปรียบเทียบกับเพื่อพิจารณาค่าความถูกต้อง (Accuracy) ดังสมการ 2-1, ค่าความแม่นยำ (Precision) ดังสมการ 2-2, ค่าระลึก (Recall) ดัง สมการ 2-3, และค่าถ่วงดุล (F-measure) ดังสมการ 2-4 [11]

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2-1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2-2)$$

$$Recall = \frac{TP}{TP+FN} \quad (2-3)$$

$$F\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2-4)$$

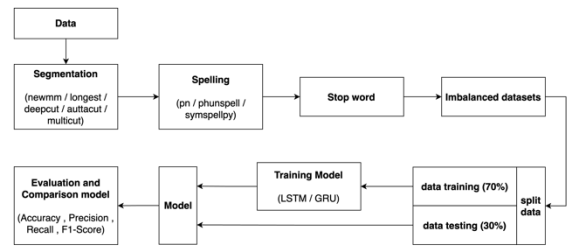
โดยที่

- TP คือ แบบจำลองทำนายว่า “จริง” และมีค่าเป็น “จริง”
- FP คือ แบบจำลองทำนายว่า “จริง” และมีค่าเป็น “ไม่จริง”
- FN คือ แบบจำลองทำนายว่า “ไม่จริง” และมีค่าเป็น “จริง”
- TN คือ แบบจำลองทำนายว่า “ไม่จริง” และมีค่าเป็น “ไม่จริง”

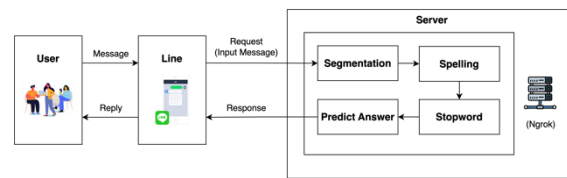
3. วิธีดำเนินการวิจัย

การดำเนินงานวิจัยในครั้งนี้ได้แบ่งออกเป็น 2 ส่วนคือ

- 1) Training phase เป็นขั้นตอนในการเปรียบเทียบ และคัดเลือกเทคนิคต่าง ๆ ที่มีประสิทธิภาพที่สุด ดังภาพที่ 3
- 2) Deployment phase เป็นขั้นตอนในการพัฒนาระบบ ดังภาพที่ 4



ภาพที่ 3: กรอบแนวคิด Training phase



ภาพที่ 4: กรอบแนวคิด Deployment phase

3.1 การรวบรวมข้อมูลและเตรียมข้อมูล

ผู้วิจัยได้รวบรวมเก็บข้อมูลคำถามและคำตอบจากระบบ กรณีศึกษาร้านชิปปี้สไมล์ เซอร์วิส สาขาภูเก็ต ซึ่งเป็นร้านที่ให้บริการรับส่งพัสดุแบบครบวงจร โดยมีจำนวนข้อมูลทั้งหมด 5,000 คำถาม และเก็บไว้ในรูปแบบไฟล์นามสกุล .csv ซึ่งเป็นชุดข้อมูลที่ได้จากลูกค้าสอบถามเข้ามาผ่านทาง LINE Official ของทางร้าน

3.2 การเปรียบเทียบและเลือกใช้เทคนิคต่าง ๆ เพื่อจัดการข้อมูล

ผู้วิจัยนำข้อมูลที่ได้จาก 3.1 มาจัดการข้อมูลต่อโดยใช้เทคนิคดังต่อไปนี้ 1) เทคนิคการตัดคำโดยเปรียบเทียบอัลกอริทึม newmm, longest, deepcut, auttcut, และ multicut เพื่อหาอัลกอริทึมที่ดีที่สุดในการตัดคำ 2) นำข้อความที่ได้จากการตัดคำมาใช้เทคนิคการสะกดคำเพื่อแก้ไขคำผิด โดยเปรียบเทียบอัลกอริทึม pn, phunspell, และsymspellpy เพื่อหาอัลกอริทึมที่ดีที่สุดในการสะกดคำ

3) นำข้อมูลที่ได้จากการสะกดคำมาจำกัดคำหยุด โดยนำคำไปเทียบจากคลังข้อมูลคำภาษาไทย thai stopwords ของ pythaiNLP 4) ปรับความไม่สมดุลของชุดข้อมูล โดยใช้ SMOTE ปรับเพิ่มข้อมูลแต่ละกลุ่มให้มีจำนวนเท่ากับจำนวนของข้อมูลของกลุ่มที่มากที่สุด

3.3 การเรียนรู้แบบจำลอง

การเรียนรู้แบบจำลองจะแบ่งข้อมูลออกเป็น 2 ส่วน คือ ข้อมูลชุดการเรียนรู้และข้อมูลชุดการทดสอบ โดยมีอัตราส่วนเป็น 70:30 และนำข้อมูลชุดการเรียนรู้เข้ามาเรียนรู้ในแบบจำลอง ซึ่งแบบจำลองได้มีการกำหนดค่าพารามิเตอร์ Input shape = 25, embedding = 16 (Weight ที่อยู่ใน Word vectors model), dropout = 0, 0.1 (การสุ่มปิดนิวรอน), activation = “softmax”, optimizer = “adam” (เป็น Optimizer ที่นิยมมากและรวมจุดเด่นของ RMSProp และ momentum แล้วลบจุดด้อยออกไปด้วย Loss Function = categorical_crossentropy), batch size = 32, 128, epochs = 100, 200, 300 (จำนวนรอบในการเรียนรู้)

3.4 การเปรียบเทียบประสิทธิภาพ

นำข้อมูลชุดทดสอบเข้ามาทดสอบในแบบจำลองเพื่อเปรียบเทียบประสิทธิภาพ โดยเปรียบเทียบค่าความถูกต้อง ค่าความแม่นยำ ค่าระลอก และค่าการถ่วงดุล

3.5 การพัฒนาระบบ

เป็นขั้นตอน Deployment phase ดังภาพที่ 4 โดยใช้เทคนิคและแบบจำลองที่มีประสิทธิภาพดีที่สุดที่ได้จากการศึกษาในขั้นตอน Training phase มาใช้งานและเชื่อมต่อกับไลน์แอปพลิเคชัน โดยใช้ Ngrok [12] ในการสร้างระบบเครือข่ายจำลองสำหรับใช้ในการเชื่อมต่อกับ LINE Messaging API ให้Ik,ki5ใช้งานข้อมูลคำสั่งโค้ดและได้ระบบไลน์แชทบอทเพื่อสนับสนุนระบบตอบคำถามโดยใช้เทคนิคการแก้ไขคำผิด

4. ผลการดำเนินงานวิจัย

4.1 ผลการหาประสิทธิภาพการตัดคำดังตารางที่ 1

ตารางที่ 1: ผลการหาประสิทธิภาพการตัดคำ

Segmentation	ค่าความถูกต้อง (%)	ความเร็วในการประมวลผล (วินาที)
newmm	88.53%	0.0013

longest	83.64%	0.0017
deepcut	87.04%	0.2296
attacut	89.56%	0.1175
multicut	83.92%	0.0001

จากตารางที่ 1 สรุปได้ว่า newmm, deepcut, และ attacut มีค่าความถูกต้องร้อยละ 88.53, 87.04, และ 89.56 ตามลำดับมีค่าความถูกต้องใกล้เคียงกัน และ newmm มีค่าความเร็วในการประมวลผลที่ 0.0013 วินาทีซึ่งเร็วกว่า deepcut และ attacut จึงเลือกใช้ newmm ที่มีค่าความถูกต้องสูงและใช้เวลาในการประมวลผลน้อยที่สุด

4.2 ผลการหาประสิทธิภาพการสะกดคำดังตารางที่ 2

ตารางที่ 2: ผลการหาประสิทธิภาพการสะกดคำ

Spelling	ค่าความถูกต้อง (%)	ความเร็วในการประมวลผล (วินาที)
pn	71.43%	0.335821
phunspell	78.57%	0.000885
symspellpy	67.32%	0.000367

จากตารางที่ 2 สรุปได้ว่า phunspell มีค่าความถูกต้องและค่าความเร็วในการประมวลผลดีที่สุดสำหรับงานวิจัยนี้ จึงเลือกใช้ phunspell

4.3 ผลการเปรียบเทียบแบบจำลองดังตารางที่ 3

ตารางที่ 3: ผลการเปรียบเทียบแบบจำลอง

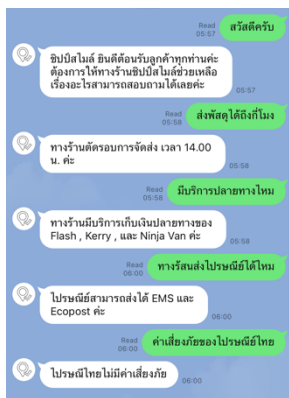
Parameter			Result			
Epochs	Batch size	Drop out	Accuracy	Precision	Recall	F-measure
LSTM						
100	32	-	85.73	97.56	85.73	86.96
100	128	-	85.47	97.67	85.47	86.79
100	32	0.1	85.77	97.59	85.77	87.00
100	128	0.1	84.07	97.04	84.07	84.95
200	32	-	85.60	97.67	85.60	86.97
200	128	-	85.80	97.83	85.80	87.23
200	32	0.1	86.23	97.54	86.23	87.35
200	128	0.1	85.77	97.80	95.77	87.20
300	32	-	84.93	97.61	84.93	86.24
300	128	-	85.87	98.13	85.87	87.23
300	32	0.1	84.73	97.54	84.73	86.09
300	128	0.1	84.07	97.64	84.07	85.67
GRU						
100	32	-	83.37	96.70	83.37	84.09
100	128	-	83.00	96.58	83.00	83.75
100	32	0.1	82.97	96.81	82.97	83.90
100	128	0.1	82.93	96.88	82.93	83.83
200	32	-	85.23	97.77	85.23	86.69
200	128	-	83.97	96.62	83.97	95.47

200	32	0.1	83.00	96.84	83.00	83.94
200	128	0.1	82.90	96.84	82.90	83.80
300	32	-	82.53	96.44	82.53	83.07
300	128	-	82.60	96.87	82.60	83.69
300	32	0.1	82.63	96.86	82.63	83.94
300	128	0.1	82.61	96.86	82.61	93.89

จากตารางที่ 3 สรุปได้ว่า LSTM ที่มีประสิทธิภาพดีที่สุด กำหนด Epochs = 200, Batch size = 32, Drop out = 0.1 โดยมีค่าความถูกต้องร้อยละ 86.23 ค่าความแม่นยำร้อยละ 97.54 ค่าร้อยละ 86.23 และค่าถ่วงดุลร้อยละ 87.35 เป็นแบบจำลองที่เหมาะสมกับงานวิจัยมากที่สุด ในส่วนของแบบจำลอง GRU ที่มีประสิทธิภาพดีที่สุดกำหนด Epochs = 200, Batch size = 32, Drop out = 0 โดยมีค่าความถูกต้องร้อยละ 85.23 ค่าความแม่นยำร้อยละ 97.77 ค่าร้อยละ 85.23 และค่าถ่วงดุลร้อยละ 86.69

4.4 ผลการพัฒนาระบบ

ผลจากการพัฒนาระบบไลน์แชทบอทเพื่อสนับสนุนระบบตอบคำถาม โดยใช้เทคนิคการแก้ไขคำผิด โดยทดสอบส่งข้อความ และโต้ตอบกลับดังภาพที่ 5



ภาพที่ 5: ผลจากการพัฒนาระบบไลน์แชทบอท

5. สรุปผลการวิจัย

ในบทความนี้ได้นำเสนอขั้นตอนการเปรียบเทียบและคัดเลือกเทคนิคต่าง ๆ ที่มีประสิทธิภาพดีที่สุด เลือกใช้เทคนิคการสะกดคำ phunspell โดยมีค่าความถูกต้องร้อยละ 78.75 และมีค่าความเร็วในการประมวลผล 0.000885 วินาที เลือกใช้เทคนิคการตัดคำ newmm โดยมีค่าความถูกต้องร้อยละ 88.53 และมีค่าความเร็วในการประมวลผล 0.0013 วินาที เลือกใช้แบบจำลอง LSTM โดยกำหนดค่า

Epochs = 200, Batch size = 32, Drop out = 0.1 ได้ค่าความถูกต้องร้อยละ 86.23 ค่าความแม่นยำร้อยละ 97.54 ค่าร้อยละ 86.23 และค่าถ่วงดุลร้อยละ 87.35 เป็นแบบจำลองที่เหมาะสมกับงานวิจัยมากที่สุด ในส่วนของ การพัฒนาไลน์แชทบอทเพื่อสนับสนุนระบบตอบคำถาม โดยใช้เทคนิคการแก้ไขคำผิด เมื่อนำมาทดลองใช้งานพบว่าสามารถนำมาช่วยลดการทำงานให้กับพนักงานได้ แต่ก็ยังพบว่าแชทบอทยังไม่ตอบโจทยข้อคำถามใช้งาน เนื่องจากยังไม่สามารถตอบคำถามในส่วนที่ยังไม่เคยเรียนรู้มาก่อนและยังตอบคำถามได้ไม่ถูกต้องในบางครั้ง

เอกสารอ้างอิง

- [1] ชัญญกร ดีพร้อม, คณพล ค่ายหนองสวาง, และ วิระพงษ์ จันทร์สนาม, “การพัฒนาระบบตอบคำถามอัตโนมัติ: กรณีศึกษาบริษัท ท่าอากาศยานไทย จำกัด (มหาชน)”, 2563.
- [2] Pimpa P, “Natural Language Processing”, 2018. <https://medium.com/@pimpapermp/natural-language-processing-3fde7c502b9e>
- [3] สุจิตรา ประชามัง, ณัฐวุฒิ พิมชาติ, และ นิธินาถ อุดมสันต์, “การใช้ภาษาแชทในการสื่อสารผ่านแอปพลิเคชันไลน์”, 2561.
- [4] Ratchakrit Arreerard, Stephen Mander, และ Scott Piao, “Survey on Thai NLP Language Resources and Tools”, 2022.
- [5] Surapong Kanoktipsatharporn, “Spell Checker ภาษาไทย ตรวจสอบการสะกดคำภาษาไทย ด้วย PyThaiNLP”. <https://www.bualabs.com/archives/3895/what-is-spell-checker-thai-language-spell-checker-pythainlp-spelling-correction-python-pythainlp-ep-3/>
- [6] วิภาดา ผาพันธุ์, “การพยากรณ์ทิศทางของราคาหุ้นรายวัน จากข้อความข่าวภาษาไทย โดยใช้วิธีการประมวลผลภาษาธรรมชาติ”, 2563.
- [7] แทนพงศ์ วัชรชาติ, “การพัฒนาหุ่นยนต์สนทนาบนระบบส่งข้อความของแอปพลิเคชันไลน์ เพื่อสนับสนุนการทำงานของแผนกทรัพยากรบุคคล”, 2561.
- [8] พุทธิพร ธนธรรมเมธี และ เขียวเรศ ศิริสถิตย์กุล, “เทคนิคการจำแนกข้อมูลที่พัฒนาสำหรับ ชุดข้อมูลที่ไม่สมดุลของภาวะข้อเข่าเสื่อมในผู้สูงอายุ”, 2018.
- [9] ธนดล สิงขรอาสน์, “การเรียนรู้เชิงลึกสำหรับการตรวจจับและรู้จำคาบรยายในวิดีโอ”, 2564.
- [10] สุดสรณ์ อยู่มาก, “การพยากรณ์ราคาหุ้นระยะสั้นและระยะยาวด้วยวิธีการเรียนรู้เชิงลึก”, 2564.
- [11] ninenox, “ทำความเข้าใจ accuracy, precision, recall, f1-score”. <http://www.ninenox.com/2020/09/24/ทำความเข้าใจ>

เข้าใจ-accuracyprecisionrecallf1-score/

- [12] สโรชา คุรุชัจฉานาค, “การศึกษาและพัฒนาเซตบอทเพื่อการสื่อสารบริการการแพทย์ฉุกเฉินบริเวณพื้นที่เมืองพิษณุโลก”, 2564.

ระบบสารสนเทศข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏ เพื่อการพัฒนาท้องถิ่น: มหาวิทยาลัยราชภัฏภูเก็ต

The Information System for the Strategic Project of Rajabhat University for Local Development: Phuket Rajabhat University

ดลรัตน์ คงหาญ (Donrat Khongharn)¹ และศักดิ์ชาย ตั้งวรรณวิทย์ (Sakchai Tangwannawit)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021857204@email.kmutnb.ac.th, ²sakchai.t@itd.kmutnb.ac.th

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบฐานข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่นมหาวิทยาลัยราชภัฏภูเก็ตและประเมินการยอมรับระบบฐานข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อพัฒนามหาวิทยาลัยราชภัฏภูเก็ต กลุ่มตัวอย่างคือผู้เชี่ยวชาญจำนวน 3 ท่าน และผู้ใช้เทคโนโลยี 30 คน โดยใช้การเลือกแบบเจาะจงในการวิจัย มีดังนี้ 1) แบบประเมินการยอมรับเทคโนโลยีระบบรายงานผลข้อมูลโครงการ 2) แบบประเมินความพึงพอใจระบบรายงานผลข้อมูลโครงการ ผลการวิจัยพบว่า 1) ผลการประเมินการยอมรับเทคโนโลยีระบบรายงานผลข้อมูลโครงการ พบว่าแอปพลิเคชันระบบรายงานผลข้อมูลโครงการมีประสิทธิภาพโดยรวมอยู่ในระดับดี ($\bar{x} = 4.62$ SD. = 0.76) 2) ผลการประเมินความพึงพอใจระบบรายงานผลข้อมูลโครงการ พบว่ามีความพึงพอใจต่อแอปพลิเคชันระบบรายงานผลข้อมูลโครงการมีประสิทธิภาพโดยรวมอยู่ในระดับดี ($\bar{x} = 4.27$ SD. = 0.81)

คำสำคัญ: การพัฒนาแอปพลิเคชัน โครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น แอปชีต พัฒนาระบบ

Project for Local Development, Phuket Rajabhat University. and assessing the acceptance of the database system for the Rajabhat University Strategic Project for the Development of Phuket Rajabhat University. The sample group was 3 experts and technology users, 30 people using the selection of the research tools, as follows: 1) the project data reporting system technology acceptance assessment form; 2) the project data reporting system satisfaction assessment form. The results showed that 1) Assessment result of technology acceptance of project information reporting system It was found that the application of the project information reporting system had overall efficiency at a good level ($\bar{x} = 4.62$ SD. = 0.76). 2) The satisfaction assessment results of the project information reporting system. It was found that they were satisfied with the application of the project information reporting system with overall efficiency at a good level ($\bar{x} = 4.27$ SD. = 0.81).

Keywords: Application, Development Rajabhat University Strategic Project for Local Development, AppSheet.

Abstract

The purpose of this research is to develop a database system for the Strategic Rajabhat University

1. บทนำ

มหาวิทยาลัยราชภัฏภูเก็ต ดำเนินกิจกรรมโครงการภายใต้โครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น ตั้งแต่ปีงบประมาณ พ.ศ. 2560 โดยดำเนินกิจกรรมโครงการในพื้นที่ที่รับผิดชอบเพื่อพัฒนาและยกระดับคุณภาพชีวิตของคนในพื้นที่ให้ครอบคลุมทั้ง 4 ด้าน คือ เศรษฐกิจ สังคม สิ่งแวดล้อม และการศึกษา [1] ดังนั้นในแต่ละปีที่มีดำเนินกิจกรรมโครงการต่าง ๆ นั้น ฝ่ายงานที่รับผิดชอบการจัดเก็บข้อมูลมีปัญหาในการเก็บข้อมูลเป็นอย่างมาก เช่น หาข้อมูลไม่เจอหรือบางไฟล์เก็บข้อมูลใน E-mail บางไฟล์เก็บข้อมูลไว้ใน E-mail ของเจ้าหน้าที่ จึงทำให้มีปัญหาเกี่ยวกับการจัดการข้อมูลเป็นอย่างมาก จึงมีความต้องการที่จะจัดทำระบบจัดเก็บฐานข้อมูลของหน่วยงาน และในปัจจุบันระบบเทคโนโลยีสารสนเทศเข้ามามีบทบาทในการบริหารงานและการวางแผนเพื่อพัฒนาการศึกษา และการ กำกับติดตามการประเมินการดำเนินงาน แผนงานและโครงการต่าง ๆ ระบบสารสนเทศที่ถูกต้องแม่นยำและเชื่อถือได้นั้นจะเป็นประโยชน์อย่างมากในการบริหารงานองค์กรซึ่งจำเป็นต้องใช้ระบบเทคโนโลยีสารสนเทศในการตัดสินใจเพื่อให้องค์กรบรรลุวัตถุประสงค์และเป้าหมายที่วางไว้ ระบบเทคโนโลยีสารสนเทศนั้นยังเปรียบเสมือนกับเครื่องมือที่ใช้ในการตรวจสอบ การทำงานของผู้บริหารและบุคลากรทุกระดับขององค์กรเป็นเครื่องมือ และเป็นข้อมูลประกอบ การตัดสินใจรวมทั้งช่วยให้สามารถคาดการณ์เหตุการณ์ที่จะเกิดขึ้นได้ในอนาคต ส่งผลให้การวางแผนการเตรียม ความพร้อมได้อย่างถูกต้องทันต่อเหตุการณ์และมีประสิทธิภาพ การพัฒนาสารสนเทศนั้นต้องพัฒนาขึ้นมาอย่างมี ระบบ มีการนำข้อมูลสารสนเทศในปีที่ผ่านมา มาศึกษาวิเคราะห์อย่างละเอียดภายใต้ผู้มีความรู้และเชี่ยวชาญเฉพาะทาง พร้อมให้การชี้แนะแนวทางการดำเนินการระบบสารสนเทศที่มีประสิทธิภาพมากขึ้น ในปีถัดไป จึงกล่าวได้ว่า ระบบสารสนเทศ (Information system) เป็นระบบที่นำข้อมูลอื่นให้เกิดสารสนเทศเป็นประโยชน์สำหรับการบริหารจัดการ

ดังนั้น ผู้วิจัยจึงได้ศึกษา รูปแบบ วิธีการสร้าง และประยุกต์ใช้เทคโนโลยี App Sheet ซึ่ง AppSheet คือ No-code Platform ที่ผู้พัฒนาไม่จำเป็นต้องมีทักษะเรื่องการเขียนโปรแกรม ก็สามารถสร้าง Application ขึ้นมาได้ ซึ่งช่วยให้สร้างแอปพลิเคชันบนอุปกรณ์เคลื่อนที่และเว็บแอปพลิเคชันได้โดยไม่ต้องเขียนโค้ด และแอปทำงานแบบไดนามิกและใช้ได้ผ่านอุปกรณ์เคลื่อนที่หรือเบราว์เซอร์ต่างๆ ออกแบบอินเทอร์เฟซแอปพลิเคชันโดยใช้เทมเพลต UX ซึ่งสามารถตอบโจทย์ความต้องการของผู้วิจัย ดังนั้นผู้วิจัยจึงใช้ AppSheet สร้าง แอปพลิเคชัน เพื่อนำมาพัฒนาเป็นระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏ เพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต เพื่อให้การจัดเก็บและรายงานผลข้อมูลมีความครบถ้วนสมบูรณ์

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 โครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต

มหาวิทยาลัยราชภัฏภูเก็ต ในฐานะสถาบันอุดมศึกษาเพื่อการพัฒนาท้องถิ่น จึงกำหนดมาตรการเพื่อขับเคลื่อนภารกิจด้านการพัฒนาท้องถิ่น [2] โดยการสนับสนุนองค์ความรู้ทางวิชาการและร่วมดำเนินงานกับหน่วยงานและองค์กรภาคี ผ่านการบูรณาการกับการจัดการเรียนการสอน การบริการวิชาการ วิจัย การทำนุบำรุงศิลปวัฒนธรรม และการมีส่วนร่วมของอาจารย์และนักศึกษาลงพื้นที่ถ่ายทอดองค์ความรู้แขนงต่างๆ ตามศาสตร์ ที่เชี่ยวชาญแก่ประชาชนและชุมชนท้องถิ่นในพื้นที่ให้บริการคือ ภูเก็ต พังงาและกระบี่ โดยมีเป้าหมายการพัฒนาและแก้ไขปัญหาที่ครอบคลุมทั้ง ด้านการศึกษา เศรษฐกิจ สังคม และสิ่งแวดล้อม ผลการดำเนินโครงการ ประจำปีงบประมาณ พ.ศ. 2565 จำนวน หมู่บ้าน/ ชุมชน ในเขตพื้นที่บริการ (จังหวัดภูเก็ต จังหวัดพังงา และจังหวัดกระบี่) ที่ได้รับการพัฒนาครบทั้ง 4 ด้าน คือ ด้านการศึกษา ด้านเศรษฐกิจ ด้านสังคม และด้านสิ่งแวดล้อม จำนวน 7 หมู่บ้าน/ ชุมชน เกิดเครือข่ายความ

ร่วมมือ 34 เครื่องข่าย ประชาชนกลุ่มเป้าหมายที่เข้าร่วมโครงการมีรายได้เพิ่มขึ้น ร้อยละ 20

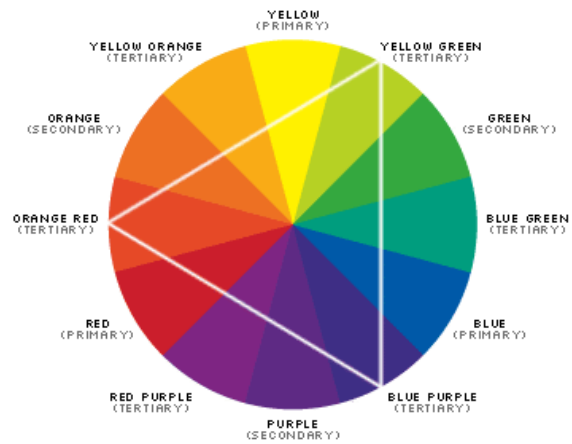
จากผลการดำเนินโครงการพัฒนาคุณภาพชีวิตและยกระดับเศรษฐกิจฐานราก ดังกล่าว เพื่อให้การดำเนินการกิจกรรมโครงการเกิดความต่อเนื่อง และยั่งยืนของชุมชนท้องถิ่น มหาวิทยาลัยจึงจำเป็นต้องดำเนินโครงการ / กิจกรรม ต่อเนื่องในปีงบประมาณพ.ศ. 2566 เพื่อให้เกิดประสิทธิภาพ และประสิทธิผลของโครงการ ซึ่งการดำเนินการดังกล่าว เป็นการขับเคลื่อนการบรรลุเป้าหมายผลสัมฤทธิ์ที่สำคัญ (Key Result) ระยะ 5 ปี พ.ศ. 2566-2570 ของแผนยุทธศาสตร์มหาวิทยาลัยราชภัฏภูเก็ตเกิดยุทธศาสตร์ที่ 1 พันธกิจสัมพันธ์เพื่อการพัฒนาท้องถิ่นอันตามัน คือ คริวเรือนตกเกณฑ์ที่พื้นที่ความยากจนและมีคุณภาพชีวิตที่ดีขึ้น ไม่น้อยกว่า 1,500 คริวเรือน และเกิดนวัตกรรมชุมชนที่มหาวิทยาลัยพัฒนาขึ้นสามารถช่วยเพิ่มรายได้หรือพัฒนาคุณภาพชีวิตให้แก่ชุมชน ไม่น้อยกว่า 350 นวัตกรรม

2.2 AppSheet

AppSheet เป็นแพลตฟอร์มการพัฒนาแอปพลิเคชันที่ไม่จำเป็นต้องเขียนโค้ด ซึ่งช่วยให้ทุกคนสร้างแอปพลิเคชันบนอุปกรณ์เคลื่อนที่และเว็บแอปพลิเคชันได้ แม้จะไม่เคยเขียนโค้ด [3] และสร้างแอป AppSheet ขึ้นได้จากแหล่งข้อมูลต่างๆ เช่น Google ชีต, Excel, Cloud SQL, Salesforce และเครื่องมือเชื่อมต่อกันๆ ที่คล้ายกัน กิจกรรมของผู้ใช้แอปจะซิงค์กับแหล่งข้อมูลที่เชื่อมต่อ [4]

2.3 ทฤษฎีสี

สี (COLOUR) หมายถึง ลักษณะกระทบต่อสายตาให้เห็นเป็นสีมีผลถึงจิตวิทยา [5] คือมีอำนาจให้เกิดความเข้มของแสงที่อารมณ์และความรู้สึกได้ การที่ได้เห็นสีจากสายตาสายตาจะส่งความรู้สึกไปยังสมองทำให้เกิดความรู้สึกต่างๆตามอิทธิพลของสี เช่น สดชื่น ร้อน ตื่นเต้น เศร้า



ภาพที่ 1: วงจรสี

2.4 ความหมายของ UX/UI Design

UX ย่อมาจาก User Experience คือ การออกแบบประสบการณ์ใช้งานเพื่อให้ผู้ใช้เกิดความพึงพอใจสูงสุด เช่น ใช้งานง่าย มีลำดับขั้นตอนชัดเจน เป็นการสร้างความรู้สึกจดจำให้ผู้ใช้จนอยากกลับมาใช้งานอีก

UI ย่อมาจาก User Interface คือ การออกแบบที่เชื่อมประสานระหว่างผู้ใช้งานกับระบบหรือผลิตภัณฑ์ ซึ่งมุ่งเน้นไปที่หน้าตาการออกแบบ เช่น การวางภาพ ปุ่ม หรือขนาดตัวอักษร เป็นต้น [6]

2.5 องค์ประกอบที่สำคัญของ UI design มีดังนี้

การทำ UI Design [7] ประกอบด้วย 5 องค์ประกอบที่สำคัญ ดังนี้

1. Information Design เป็นกระบวนการออกแบบที่เน้นนำเสนอข้อมูล เพื่อสร้างความเข้าใจและเกิดการเรียงลำดับข้อมูลได้อย่างครบถ้วน
2. Interaction Design กระบวนการออกแบบโดยคำนึงถึงปฏิสัมพันธ์ระหว่างผู้ใช้งานที่สามารถไปถึงเป้าหมายได้อย่างรวดเร็ว หรือการใช้คำที่สื่อความหมายทำให้ผู้ใช้งาน ทราบได้ทันทีว่าต้องการจะสื่อสารความหมายอะไร หรือ ให้ผู้ใช้ทำอะไร
3. Information Architecture เป็นการจัดเรียงข้อมูลอย่างเป็นระบบ และการจัดกลุ่มข้อมูลให้มีความถูกต้อง เพื่อให้ผู้ใช้งานรู้ลำดับการเข้าถึงข้อมูลได้อย่างรวดเร็ว
4. Visual Design เน้นการนำเสนอภาพเพื่อให้เกิดความสวยงาม ครบทุกองค์ประกอบ เช่น การจัดวาง

ตำแหน่ง การใช้สี ซึ่งการวิจัยในครั้งนี้ได้นำหลักทฤษฎีสี เข้ามาใช้ในการออกแบบแผนการใช้ตัวอักษรในการ นำเสนอให้สวยงาม ให้ผู้อ่านเข้าใจได้ง่าย

5. Human Computer Interaction เป็นการ ออกแบบ ปฏิสัมพันธ์ระหว่างมนุษย์กับคอมพิวเตอร์ ซึ่งเป็น พื้นฐาน ในการเรียนรู้การออกแบบ UX/UI ที่ดี

2.6 องค์ประกอบการออกแบบ UX Design มีดังนี้

การออกแบบ UX Design [8] มี 4 องค์ประกอบดังนี้

1. Design Thinking คือ การคิดเชิงออกแบบ เป็น กระบวนการคิดที่ให้ความสำคัญกับการทำความเข้าใจ ปัญหาของผู้ใช้ เพื่อให้สิ่งที่ออกแบบมาตอบสนองต่อ ความ ต้องการของผู้ใช้งานสูงสุด

2. Service Design เป็นการออกแบบการบริการที่จะ เน้นแต่การบริการเท่านั้น โดยจะคำนึงถึงประสบการณ์ ของ ผู้ที่เกี่ยวข้องในงานทั้งหมด เพื่อสร้างระบบที่มี ประสิทธิภาพ และมีความต่อเนื่องกันตั้งแต่ต้นจนจบ กระบวนการ

3. Design Sprint เป็นกระบวนการเพื่อแก้ไข ปัญหาที่ เกิดขึ้นกับผลิตภัณฑ์หรือบริการที่ออกสู่ตลาดแล้ว ได้อย่าง เป็นระบบ จุดเด่นในส่วนนี้คือความรวดเร็ว โดยจะใช้เวลา เพียงไม่กี่วันในแต่ละ Sprint เพื่อแก้ไขปัญหที่เกิดขึ้น

4. User Centered Design เป็นกระบวนการ ออกแบบที่ ให้ความสำคัญกับผู้ใช้งานเป็นหลัก โดยผู้ใช้งาน จะเข้ามา มีส่วนร่วมในการประเมินผลการออกแบบร่วมกัน [9]

3. วิธีดำเนินการวิจัย

กรอบแนวความคิดระบบรายงานผลข้อมูล โครงการ ยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต



ภาพที่ 2 : กรอบแนวความคิดการดำเนินงานวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

ประชากร กลุ่มที่ใช้ศึกษาการยอมรับเทคโนโลยี มหาวิทยาลัยราชภัฏภูเก็ตโดยวิธีการเลือกกลุ่มตัวอย่าง แบบเจาะจง (Purposive Sampling)

กลุ่มตัวอย่าง กลุ่มที่ทดลองใช้ศึกษาการยอมรับ เทคโนโลยี โดยวิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling) ได้แก่ อาจารย์ เจ้าหน้าที่ มหาวิทยาลัยราชภัฏภูเก็ต จำนวน 10 คน

3.2 การเก็บรวบรวมข้อมูล

เก็บรวบรวมข้อมูล โดยการแจกแบบสอบถาม ออนไลน์ (Google Form) แบบสอบถามที่ใช้งานวิจัยครั้งนี้ แบ่งออกเป็น 4 ส่วน ได้แก่ 1. ข้อมูลทั่วไปของผู้ตอบ แบบสอบถาม 2. รายการปัจจัยรับรู้ประโยชน์ (PU: Perceived Usefulness) 3. รายการปัจจัยด้านการรับรู้ความ ง่ายในการใช้งาน (PE: Perceived Ease of Use) 4. รายการ ปัจจัยด้านทัศนคติต่อการใช้ (AU: Attitude Toward Using) 5. รายการปัจจัยความตั้งใจที่จะนำไปใช้ โดยใช้มาตร ประเมินค่า 5 ระดับ โดยการวัดข้อมูลประเภทอันตรภาค ชั้นแบ่งเป็น 5 ระดับ ได้แก่ 1 = น้อยที่สุด 2 = น้อย 3 = ปาน กลาง 4 = มาก และ 5 = มากที่สุด

3.3 การวิเคราะห์ข้อมูลและออกแบบระบบ (Data Analyze and System Design)

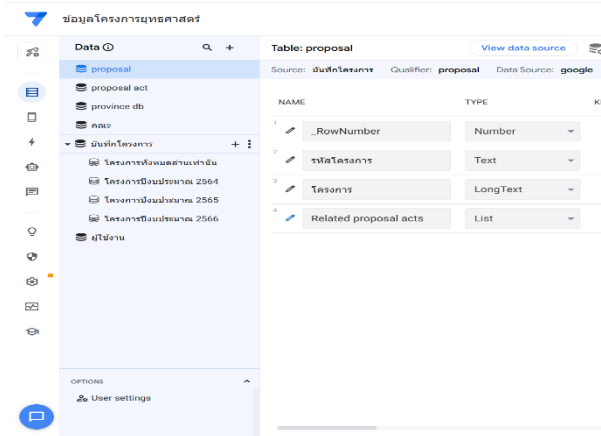
ผู้วิจัยได้นำข้อมูลที่ผ่านการตรวจสอบความถูกต้อง และความสมบูรณ์ของข้อมูลเป็นที่เรียบร้อยแล้ว

ลำดับ	ชื่อโครงการ	หน่วยงาน	ผู้จัด	สถานที่	วันที่	จำนวนผู้ลงทะเบียน
1	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	5
2	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	3
3	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	5
4	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	5
5	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	1
6	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	1
7	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	1
8	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	4
9	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	1
10	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	1
11	2566 โครงการพัฒนาระบบสารสนเทศเพื่อใช้ในการเรียนการสอน		อุทิศ	ฉะเชิงเทรา	1	3

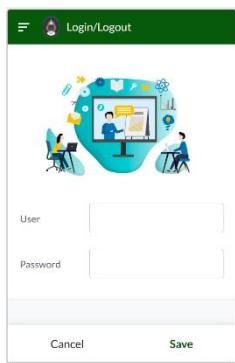
ภาพที่ 3: ข้อมูล

3.4 พัฒนาระบบ (Implementation)

นำการออกแบบมาใช้ในการพัฒนา ระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต โดยการจัดเก็บ Data



ภาพที่ 4 : ออกแบบระบบ



ภาพที่ 5 : หน้าจอ Login

3.5 การประเมินการยอมรับ (Evaluation)

3.5.1 แบบประเมินการยอมรับเทคโนโลยีระบบ

ฐานข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต

3.5.2 สถิติ ใช้สถิติพื้นฐาน ใช้ค่าเฉลี่ย (Average)

และส่วนเบี่ยงเบนมาตรฐาน (Standard Division)

4. ผลการดำเนินการวิจัย

ผลการวิจัยพบว่า

1) ผลการประเมินการยอมรับเทคโนโลยีระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ตเกิดพบว่าแอปพลิเคชันระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต มีประสิทธิภาพโดยรวมอยู่ในระดับดี ($\bar{x} = 4.62$ SD. = 0.76)

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ระดับ
ระบบช่วยให้สะดวกในการกรอกข้อมูล	4.33	1.15	ดี
ระบบมีส่วนช่วยลดระยะเวลาในการกรอกข้อมูล	4.67	0.58	ดี
ระบบสามารถเก็บไฟล์ในรูปแบบ Excel	4.33	1.15	ดี
ระบบช่วยให้ผู้บริหารพิจารณาข้อมูลที่เกี่ยวข้องเพื่อวางแผนการดำเนินงานได้	5.00	0.00	ดี
ขั้นตอนการเรียกดูระบบมีความชัดเจนและเข้าใจง่าย	4.33	4.33	ดี
ระบบมีการใช้ขนาดตัวอักษร และสีของตัวอักษรได้อย่างชัดเจน	4.33	0.58	ดี
การออกแบบหน้าจอและมุมมองรายงานมีความชัดเจน	4.67	0.58	ดี
ระบบช่วยให้ผู้บริหารพิจารณาข้อมูลที่เกี่ยวข้องเพื่อวางแผนการดำเนินงานได้	5.00	0.00	ดี
ระบบช่วยให้ท่านเกิดความรู้สึกง่ายต่อการทำงาน	5.00	1.00	ดี
ระบบช่วยลดภาระงานของท่านได้	4.67	0.58	ดี
ระบบทำให้การทำงานมีความสะดวกมากขึ้น	5.00	0.00	ดี
ระบบทำให้การทำงานรวดเร็วมากขึ้น	4.33	1.15	ดี
ระบบทำให้การทำงานมีประสิทธิภาพมากขึ้น	4.33	1.15	ดี
เฉลี่ย	4.62	0.76	ดี

ภาพที่ 6 : ผลการประเมินการยอมรับเทคโนโลยีระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่นมหาวิทยาลัยราชภัฏภูเก็ต

รายการประเมิน	ผลการประเมิน		
	\bar{X}	S.D.	ระดับ
ระบบช่วยให้สะดวกในการกรอกข้อมูล	4.80	0.63	ดี
ระบบมีส่วนช่วยลดระยะเวลาในการกรอกข้อมูล	4.20	0.42	ดี
ระบบสามารถเก็บไฟล์ในรูปแบบ Excel	4.10	0.57	ดี
ระบบช่วยให้ผู้บริหารพิจารณาข้อมูลที่เกี่ยวข้องเพื่อวางแผนการดำเนินงานได้	4.10	0.57	ดี
ขั้นตอนการเรียกดูระบบมีความชัดเจนและเข้าใจง่าย	4.10	0.57	ดี
ระบบมีการใช้ขนาดตัวอักษร และสีของตัวอักษรได้อย่างชัดเจน	4.10	0.32	ดี
การออกแบบหน้าจอและมุมมองรายงานมีความชัดเจน	4.50	0.53	ดี
ระบบช่วยให้ผู้บริหารพิจารณาข้อมูลที่เกี่ยวข้องเพื่อวางแผนการดำเนินงานได้	4.40	0.52	ดี
ระบบช่วยให้ท่านเกิดความรู้สึกง่ายต่อการทำงาน	4.33	1.15	ดี
ระบบช่วยลดภาระงานของท่านได้	4.50	0.53	ดี
ระบบทำให้การทำงานมีความสะดวกมากขึ้น	4.50	0.53	ดี
ระบบทำให้การทำงานรวดเร็วมากขึ้น	4.40	0.70	ดี
ระบบทำให้การทำงานมีประสิทธิภาพมากขึ้น	4.40	0.70	ดี
เฉลี่ย	4.34	0.81	ดี

ภาพที่ 6 : ผลการประเมินความพึงพอใจระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่นมหาวิทยาลัยราชภัฏภูเก็ต

2) ผลการประเมินความพึงพอใจระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต พบว่ามีความพึงพอใจต่อแอปพลิเคชันระบบรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต มีประสิทธิภาพโดยรวมอยู่ในระดับดี ($\bar{x} = 4.34$ SD. = 0.81)

5. สรุปผล

การพัฒนากระบวนการรายงานผลข้อมูลโครงการยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น มหาวิทยาลัยราชภัฏภูเก็ต ประกอบด้วยโมดูลบันทึกข้อมูล โมดูลการสมัครสมาชิก โมดูลการค้นหาข้อมูล

ข้อเสนอแนะสำหรับการพัฒนาต่อไปในอนาคตควรทำระบบรายงานที่เชื่อมต่อไปยัง Google data studio และ AppSheet ไม่เหมาะกับการสร้าง Application ที่ซับซ้อน หากต้องการการรายงานที่มีความซับซ้อนอาจจะไม่สามารถรายงานผลได้ครบถ้วน

เอกสารอ้างอิง

- [1] นงรัตน์ อิศโร. (2561). ยุทธศาสตร์มหาวิทยาลัยราชภัฏเพื่อการพัฒนาท้องถิ่น ระยะ 20 ปี (พ.ศ. 2560 - 2579) มปท : 2561
- [2] กองนโยบายและแผน. แผนยุทธศาสตร์มหาวิทยาลัยราชภัฏระยะ 5 ปี พ.ศ. 2566 – 2570. มปท : มปป.
- [3] สหสวรรค์ งามทรง และคณะ. (2564). การสร้างโมบายแอปพลิเคชันแบบไม่ต้องเขียนโปรแกรมเพื่อการสำรวจข้อมูลจากภาคสนาม สำหรับงานบริหารจัดการภาษีขององค์กรปกครองส่วนท้องถิ่น. วารสารวิชาการเพื่อพัฒนานวัตกรรมเชิงพื้นที่, ปีที่ 2 (ฉบับที่ 3). หน้า 55 – 68
- [4] นพรัตน์ ประทุมนอก และคณะ. (2565). การพัฒนาแอปพลิเคชันตรวจสอบการเข้าร่วมกิจกรรมด้วยแพลตฟอร์มแอปซิท. วารสารวิจัยและนวัตกรรมทางวิทยาศาสตร์และเทคโนโลยี, ปีที่ 3 (ฉบับที่ 2).
- [5] ทฤษฎีสี. สืบค้น 15 มกราคม 2566, จาก <https://www.xn12ca0ezbc4ai2ee1bzl.com/17/02/2022/%E0%B8%9A%E0%B8%A3%E0%B8%A3%E0%B8%93%E0%B8%B2%E0%B8%99%E0%B8%B8%E0%B8%81%E0%B8%A3%E0%B8%A1%E0%B9%80%E0%B8%A7%E0%B9%87%E0%B8%9A%E0%B9%84%E0%B8%8B%E0%B8%95%E0%B9%8C/>
- [6] Blockfint.com. (2021). สืบค้น 15 มกราคม 2566 UX Design vs UI Design vs Interaction Design vs Visual Design. Retrieved from <https://blockfint.com/blog/confusion-between-ux-ui-interaction-design-visual-design.> (in Thai)
- [7] Interaction-design.org. (2002). สืบค้น 15 มกราคม 2566. *User Interface Design*. Retrieved from <https://www.interaction-design.Org/literature/topics/ui-design>
- [8] Imsudjai, S. (2022). สืบค้น 15 มกราคม 2566. *UX Design*. Retrieved from <https://www.borntodev.com/2022/01/04/ux-Design.> (in Thai)
- [9] Iconext Writer. (2021). สืบค้น 15 มกราคม 2566. What UX/UI Design. Retrieved from <https://iconext.co.th/th/2021/19/ux-uidesign/> (in Thai)

การยืนยันตัวตนอย่างต่อเนื่องสำหรับ การใช้บริการห้องสมุดโดยใช้การจดจำใบหน้า Continuous Authentication for Library Services Using Facial Recognition

เปรมากร ทวีหันท (Premagorn Thaweehan)¹ และปราโมทย์ กัวเจริญ (Pramote Kuacharoene)²

^{1,2}ภาควิชาวิทยาการคอมพิวเตอร์และระบบสารสนเทศ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

¹premagorn.tha@stu.nida.ac.th, ²pramote@as.nida.ac.th

บทคัดย่อ

มีเหตุการณ์บนโลกเกิดขึ้นมากมาย ทำให้เกิดความผันผวน ความไม่แน่นอน เรียกว่า “VUCA World” การรับมือกับ สถานการณ์คือการเรียนรู้และปรับตัว หนึ่งในสถานที่ ที่รวบรวมความรู้ที่หลากหลายและทันสมัยอยู่เสมอ นั่นก็คือ “ห้องสมุด” ซึ่งควรยกระดับไปสู่ “ห้องสมุดอัจฉริยะ” การมีระบบที่ยืนยันตัวตนด้วยใบหน้าถือเป็นเทคโนโลยีที่มีความสำคัญในด้านความปลอดภัย และความเป็นส่วนตัว ของผู้ใช้งานเพื่อป้องกันการถูกสวมรอยจากบุคคลอื่นที่ไม่ได้รับอนุญาต การใช้ใบหน้ายืนยันตัวตนมีความปลอดภัยสูงเนื่องจากมีความจำเพาะเจาะจงแต่ละบุคคล และลอกเลียนแบบได้ยาก เมื่อเทียบกับการยืนยันตัวตน โดยใช้รหัสผ่านเพียงอย่างเดียว ถ้าหากข้อมูลรหัสผ่านถูกขโมยจะถูกสวมรอยได้ทันที ทางผู้วิจัยเห็นถึงความสำคัญนี้ จะออกแบบระบบซึ่งสามารถแบ่งเป็น 5 ระบบย่อยคือ ระบบลงทะเบียนใบหน้า, ระบบยืนยันตัวตนด้วยการ ล็อกอิน, ระบบการจดจำใบหน้า, ระบบการอ่านบาร์โค้ด, ระบบการจองทรัพยากรห้องสมุด จากการทดสอบพบว่า การจดจำใบหน้ามีความแม่นยำ แม้อยู่ในสถานการณ์ของ สภาพแวดล้อมที่แตกต่าง สามารถทำงานได้อย่างรวดเร็ว แม้มีข้อมูลใบหน้าที่บันทึกไว้ปริมาณมาก

คำสำคัญ: การจดจำใบหน้า การยืนยันตัวตน ห้องสมุด

อัจฉริยะ โลก VUCA

Abstract

In today's uncertain world, known as the "VUCA World," it is important to learn and adapt to changing

situations. The library, as a source of diverse knowledge, should be upgraded to a "Smart Library" to better navigate this volatility and uncertainty. One key technology that can enhance security and protect user privacy is facial authentication. This method of identification is highly secure as it is unique to each individual and difficult to imitate. To create an effective system, the researcher has divided the smart library into five sub-systems: facial registration system, login authentication system, facial recognition system, barcode reading system, and library resource reservation system. Testing shows that facial recognition is accurate even in different environmental situations and can work quickly even with a large amount of saved face data.

Keywords: Facial Recognition, Authentication, Smart Library, VUCA World.

1. บทนำ

เกิดเหตุการณ์บนโลกขึ้นมากมาย ทำให้เกิดความผันผวน ความไม่แน่นอน ความซับซ้อน ความคลุมเครือ หรือเรียกเป็นคำจำกัดความว่า “VUCA World” [8] ส่งผลให้ ผู้คนปรับเปลี่ยนพฤติกรรมไปจากเมื่อก่อนไม่ว่าจะเป็น การ Work From Home, Study From Home เปลี่ยนรูปแบบการทำงาน เกิดการเปลี่ยนแปลงจากโลกออฟไลน์สู่โลกที่เป็นออนไลน์มากขึ้น

การรับมือกับสถานการณ์ของ “VUCA World” [8] นั้นคือการปรับตัวให้ทันต่อการเปลี่ยนแปลง ปรับแนวคิด

ใหม่ และการเรียนรู้ตลอดชีวิต หรือที่เรียกว่า “Lifelong Learning” [6] และหัวใจสำคัญของการเรียนรู้ตลอดชีวิตก็คือสถานที่ที่รวบรวมความรู้ที่หลากหลายและทันสมัยอยู่เสมอ นั่นก็คือ “ห้องสมุด” ต้องยกระดับสู่ “ห้องสมุดอัจฉริยะ” เพื่อรับมือต่อการเปลี่ยนแปลง

การสร้างระบบห้องสมุดอัจฉริยะ [9] ให้มีประสิทธิภาพนั้นมีหลายปัจจัยที่ต้องคำนึงถึง บริการยืม-คืนทรัพยากรห้องสมุดโดยใช้ใบหน้ายืนยันตัวตน ครอบคลุมไปถึงห้องสมุดแบบออนไลน์ การยืนยันตัวตนเพื่อใช้บริการห้องสมุดออนไลน์ เพื่อกำหนดสิทธิในการเข้าถึงทรัพยากรสารสนเทศในห้องสมุดออนไลน์ให้กับบุคคลนั้นได้อย่างถูกต้อง หรือเป็นเรื่องความปลอดภัยของผู้ใช้งานที่จะไม่เกิดการถูกสวมรอยเพื่อใช้งานระบบโดยผู้ใช้งานอื่น หรือผู้ที่ไม่ได้รับอนุญาต โดยผู้วิจัยได้เห็นถึงความสำคัญของการยืนยันตัวตนโดยใช้วิธีการจดจำใบหน้า [3], [4], [7] ที่ระบบจะสามารถจดจำใบหน้าของผู้ใช้งานได้เพื่อยืนยันตัวตน เนื่องจากใบหน้าของบุคคลนั้นมีความจำเพาะเจาะจงสูงในแต่ละบุคคล และสามารถลอกเลียนแบบได้ยาก

ระบบการยืนยันตัวตนแบบต่อเนื่องเป็นการยืนยันตัวตนผู้ใช้งานตลอดเวลาที่ใช้งาน มีแนวคิดมาจาก Zero Trust Architecture [10] เป็นสถาปัตยกรรมที่ออกแบบระบบเชิงกลยุทธ์ในการรักษาความปลอดภัยทางไซเบอร์ โดยตั้งสมมติฐานลดความไว้วางใจใด ๆ ให้เท่ากับศูนย์ และตรวจสอบความถูกต้องอย่างต่อเนื่องในทุกขั้นตอนของการโต้ตอบทางดิจิทัล ซึ่งมีหลักการ 3 ประการคือ จำกัดสิทธิการเข้าถึงของผู้ใช้งานให้น้อยที่สุด, ทุกการกระทำมีความไม่ปลอดภัย, มีการตรวจสอบหรือยืนยันอย่างต่อเนื่อง

2. วรรณกรรมที่เกี่ยวข้อง

การจดจำใบหน้าเกี่ยวข้องกับปัญหาหลายประการ [5] อย่างเช่น ค้นหาใบหน้าที่อยู่ในรูปหรือในกล้อง ไฟล์สที่ใบหน้าและทำความเข้าใจว่าแม้ว่าใบหน้าจะหันไปทางอื่นหรือในที่ที่มีแสงไม่ดี ก็ยังคงเป็นบุคคลคนเดิม และสามารถเลือกลักษณะเฉพาะของใบหน้าที่คุณสามารถใช้

เพื่อแยกแยะความแตกต่างจากคนอื่น ๆ เช่น ดวงตาโตแค่นั้น ใบหน้ายาวแค่นั้น เป็นต้น และสุดท้ายเปรียบเทียบคุณลักษณะเฉพาะของใบหน้านั้นกับบุคคลทั้งหมดที่คุณรู้จักแล้วเพื่อกำหนดชื่อของบุคคลนั้น

เริ่มต้นที่การตรวจหาใบหน้า ใช้วิธีการที่คิดค้นขึ้นในปี 2548 ที่เรียกว่า Histogram of Oriented Gradients หรือ HOG [5] จะแบ่งภาพออกเป็นสี่เหลี่ยมเล็กๆ ขนาด 16x16 พิกเซลต่อภาพ ในแต่ละช่องสี่เหลี่ยม เราจะนับจำนวนการไล่ระดับสีที่ชี้ไปในแต่ละทิศทาง จากนั้นเราจะแทนที่สี่เหลี่ยมในภาพด้วยลูกศรชี้ทิศทางที่เด่นที่สุด ผลลัพธ์ที่ได้คือเราเปลี่ยนภาพต้นฉบับให้เป็นการนำเสนอที่เรียบง่ายมาก ต่อมาคือการจัดการความบิดเบี้ยวของใบหน้าให้ใบหน้ามาอยู่ตรงกลาง คือการจุดเฉพาะ 68 จุด ที่มีอยู่บนใบหน้าทุกหน้า จากนั้นก็หมุน ปรับขนาด และปรับภาพเพื่อให้ตาและปากอยู่กึ่งกลางที่ดีที่สุดเท่าที่จะเป็นไปได้ หลังจากที่ได้ใบหน้ามาแล้ว

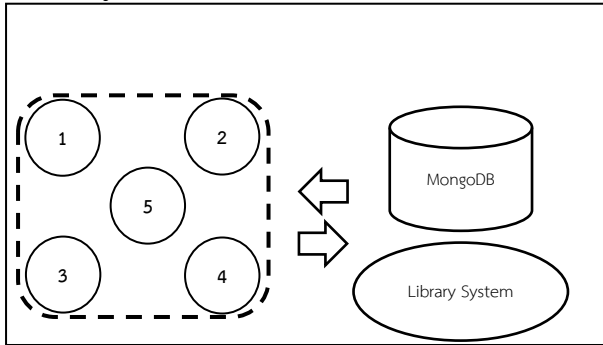
สิ่งต่อมาคือแยกความแตกต่างของใบหน้า โดยให้คอมพิวเตอร์เป็นคนแยกเองผ่านวิธี Deep Learning หากว่าส่วนใดของใบหน้าที่มีความสำคัญต่อการแยกความแตกต่าง โดยมี Input ทั้งหมด 128 จุด โดยการฝึกฝนแบบจำลองจะใช้ Loss function คือ Triple Loss [2], [5] โดยใช้ภาพบุคคลเดียวกันขึ้นมา 2 ภาพที่ต่างกันและใบหน้าคนอื่นขึ้นมาอีก 1 ภาพ ซึ่งเมื่อฝึกฝนโมเดลไปเรื่อย ๆ ค่า Loss ของภาพที่เป็นคนเดียวกันจะเข้าใกล้กันมากขึ้น ส่วนหน้าของคนอื่นจะห่างออกไปเรื่อย ๆ จนมีความแม่นยำสูงถึง 99.38% ซึ่งทางผู้สร้างแบบจำลองนี้ขึ้นมา เปิดให้ใช้ได้ผ่านเครื่องมือที่ชื่อ DLib [1] ซึ่งรองรับการใช้ด้วยภาษา Python

3. วิธีการดำเนินการวิจัย

การพัฒนาการจดจำใบหน้าเพื่อยืนยันตัวตนในการใช้บริการห้องสมุด เป็นการพัฒนาโดยใช้เว็บแอปพลิเคชัน (Web Application) เป็นพื้นฐานของระบบ โดยการใช้ภาษาโปรแกรมคือ ภาษา Python ซึ่งเป็นภาษาที่เหมาะสมกับการพัฒนาเว็บแอปพลิเคชัน ซึ่งมีเครื่องมือที่ช่วยให้การพัฒนา เช่น Flask Framework ที่ช่วยพัฒนา

เว็บไซต์ได้อย่างรวดเร็ว และ ภาษา Python ยังเหมาะแก่การพัฒนาาระบบที่เกี่ยวข้องกับการเรียนรู้เชิงลึก (Deep Learning) ซึ่งเป็นเทคโนโลยีที่ใช้ในการจดจำใบหน้า โดยสามารถแบ่งออกเป็น 6 ส่วน ดังนี้

3.1 System Architecture



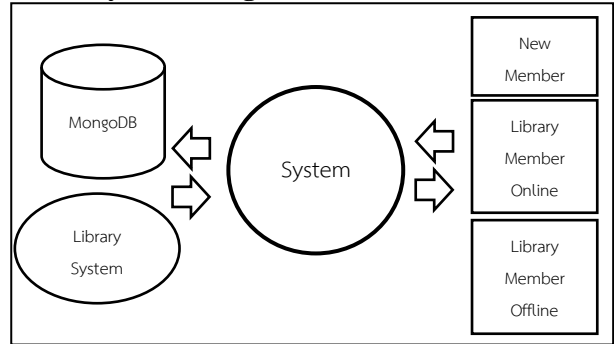
ภาพที่ 1: สถาปัตยกรรมของระบบ

จากภาพที่ 1 ระบบการยืนยันด้วยใบหน้า มีทั้งหมด 5 ระบบย่อย ดังนี้

1. ระบบลงทะเบียนใบหน้า เป็นระบบที่ใช้สร้างข้อมูลของใบหน้าใหม่ที่ยังไม่มีบนฐานข้อมูลหรือเป็นการอัปเดตใบหน้าเก่าให้เป็นใบหน้าปัจจุบัน
2. ระบบยืนยันตัวตนด้วยการล็อกอิน เป็นการยืนยันตัวตนโดยใช้ username และ password ซึ่งสามารถใช้อีเมล Gmail, Hotmail หรืออีเมลของมหาวิทยาลัย
3. ระบบการจดจำใบหน้า ถือว่าเป็นส่วนสำคัญของระบบ เป็นระบบที่ใช้เปรียบเทียบความใกล้เคียงของใบหน้าที่ถ่ายจากกล้อง กับใบหน้าที่เก็บไว้ในฐานข้อมูล
4. ระบบอ่านบาร์โค้ด เป็นระบบใช้อ่าน Barcode หรือ QR Code ที่ติดอยู่กับหนังสือ หรือทรัพยากรอื่นๆ
5. ระบบการจองทรัพยากรห้องสมุด เป็นระบบที่บันทึกข้อมูลของผู้ใช้บริการไม่ว่าจะเป็นช่องทางออนไลน์หรือออฟไลน์ เช่น การบันทึกข้อมูลของผู้ใช้บริการยืมหนังสือ ซึ่งจะส่งข้อมูล ไปยังระบบห้องสมุด

นอกเหนือจากระบบย่อยทั้ง 5 ระบบ จะมีระบบหลักของห้องสมุด และฐานข้อมูล MongoDB Atlas ซึ่งเป็นฐานข้อมูลคลาวด์แบบ NoSQL ซึ่งมีความยืดหยุ่นทั้งการเก็บข้อมูลหลากหลายรูปแบบ และความยืดหยุ่นในการรองรับปริมาณการใช้งาน

3.2 System Design

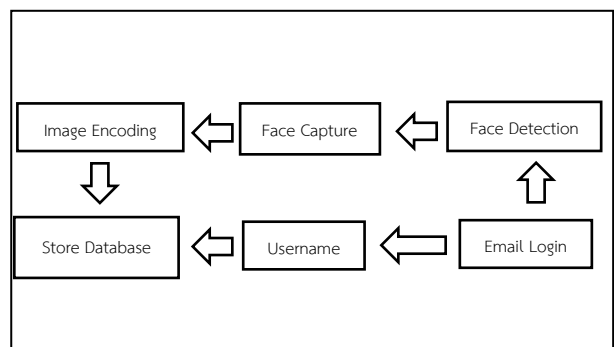


ภาพที่ 2: การออกแบบระบบ

จากภาพที่ 2 สามารถแบ่งตามการใช้งานของผู้ใช้งาน 3 กลุ่ม ดังนี้

1. ผู้ใช้งานใหม่ เป็นกลุ่มผู้ใช้งานที่เข้ามาใช้งานระบบครั้งแรก ซึ่งต้องมีการลงทะเบียนซึ่งข้อมูลที่ลงทะเบียนประกอบด้วยชื่อบัญชีผู้ใช้และรูปถ่าย โดยข้อมูลการลงทะเบียนจะถูกส่งไปยังฐานข้อมูลเพื่อเก็บข้อมูลของผู้ใช้งาน
2. ผู้ใช้งานออนไลน์ เป็นกลุ่มผู้ใช้งานที่ต้องยืนยันตัวตนด้วยการเข้าระบบผ่านอีเมลและยืนยันด้วยใบหน้า เมื่อยืนยันใบหน้าสำเร็จจะเข้าสู่ระบบห้องสมุด และขณะใช้งานจะมีการยืนยันด้วยใบหน้าอยู่ตลอดเวลา
3. ผู้ใช้งานออฟไลน์ เป็นกลุ่มผู้ใช้งานที่เข้ามาใช้บริการภายในห้องสมุด และต้องการจองหรือยืมทรัพยากรของห้องสมุด ยกตัวอย่างเช่น หนังสือ สามารถใช้ใบหน้าและรหัสผ่านเพื่อยืนยันตัวตนได้

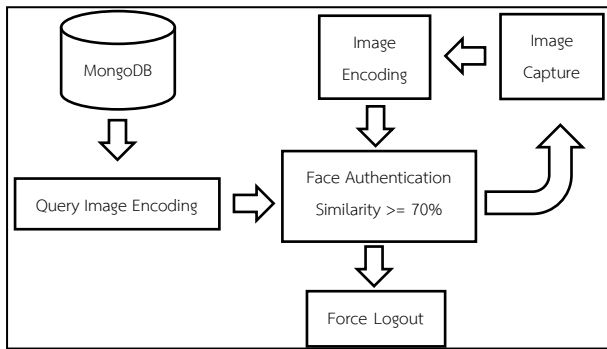
3.3 Face Registration Process



ภาพที่ 3: กระบวนการลงทะเบียนใบหน้า

จากภาพที่ 3 กระบวนการลงทะเบียนใบหน้าสำหรับระบบเทคโนโลยีการจดจำใบหน้าเพื่อยืนยันตัวตน เริ่มต้นที่การเข้าสู่ระบบด้วยอีเมล ซึ่งสามารถใช้อีเมล Gmail เพื่อเข้าสู่ระบบ เมื่อเข้าสู่ระบบสำเร็จจะทำการเก็บข้อมูลผู้ใช้ จากนั้นจะทำการถ่ายรูปโดยจะมีระบบค้นหาใบหน้า ถ้าไม่พบใบหน้าจะไม่สามารถถ่ายรูปได้ เมื่อตรวจพบใบหน้าแล้วถ่ายรูป ภาพที่ได้จะทำการแปลงข้อมูลรูปภาพเป็นรูปแบบตัวเลขเพื่อตอกลงข้อมูลมาใช้เพื่อระบุตัวตนจะสามารถประมวลผลได้อย่างรวดเร็ว จากนั้นจะนำข้อมูลรูปภาพและข้อมูลผู้ใช้บันทึกไปยังฐานข้อมูล

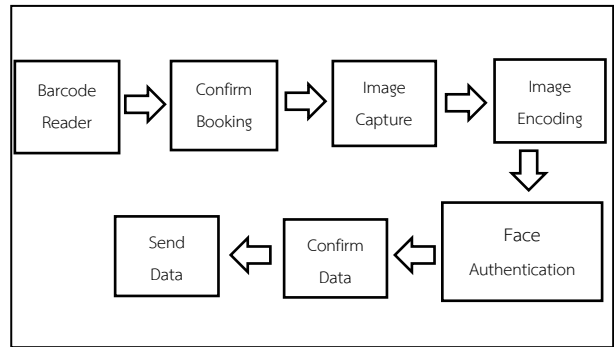
3.4 Face Authentication Process (Online)



ภาพที่ 4: กระบวนการยืนยันตัวตนด้วยใบหน้า (ออนไลน์)

จากภาพที่ 4 กระบวนการยืนยันตัวตนด้วยใบหน้าระบบจะดึงใบหน้าของผู้ใช้งานที่เก็บไว้ตอนลงทะเบียนในฐานข้อมูลเพื่อนำมาใช้ในการยืนยันตัวตนซึ่งเป็นข้อมูลที่อยู่ในรูปแบบตัวเลข ขณะเดียวกันจะบันทึกรูปภาพจากกล้อง แล้วนำรูปถ่ายที่ได้ไปส่งกระบวนการแปลงข้อมูลรูปภาพเป็นรูปแบบตัวเลข เมื่อได้ข้อมูลรูปจากทั้งสองฝั่งคือสิ่งที่บันทึกไว้ และรูปที่ถ่ายจากปัจจุบัน เข้าสู่ Face Recognition Algorithm เมื่อเปรียบเทียบค่าความคล้ายคลึงกันของรูปทั้งสอง โดยตั้งเงื่อนไขไว้ว่าถ้าค่าที่ได้มีความคล้ายคลึงมากกว่าหรือเท่ากับ 70 เปอร์เซ็นต์ ซึ่งถือว่ายืนยันตัวตนสำเร็จ เมื่อเวลาผ่านไปประมาณ 10 วินาทีจะวนกระบวนการกลับไปยังการบันทึกรูปภาพจากกล้องเพื่อเป็นการยืนยันตัวตนอยู่เสมอ แต่ถ้ากรณีมีค่าต่ำกว่าจะถูกบังคับออกจากระบบทันที

3.5 Face Authentication Process (Offline)



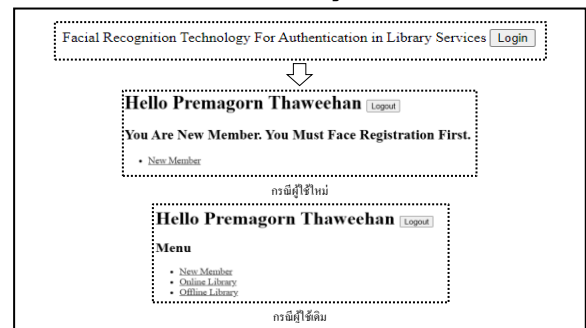
ภาพที่ 5: กระบวนการยืนยันตัวตนด้วยใบหน้า(ออฟไลน์)

จากภาพที่ 5 กระบวนการยืนยันตัวตนด้วยใบหน้า (ออฟไลน์) เริ่มต้นที่การนำ Barcode หรือ QR Code ที่ติดอยู่กับสิ่งของที่ต้องการจะยืมมาสแกนที่กล้องเพื่ออ่านค่ารหัส เมื่อทำการสแกนและยืนยันการจองเรียบร้อยแล้ว จะทำการถ่ายรูปเพื่อบันทึกหน้าของผู้ยืม จากนั้นจะนำรูปถ่ายที่ได้ไปส่งกระบวนการแปลงข้อมูลรูปภาพเป็นรูปแบบตัวเลข แล้วนำข้อมูลไปตรวจสอบใบหน้าที่ใกล้เคียงในฐานข้อมูล เมื่อพบใบหน้าที่ใกล้เคียงที่สุดจะต้องยืนยันด้วยข้อมูลอีกครั้ง เมื่อยืนยันรหัสถูกต้องจะทำการส่งข้อมูลไปยังระบบห้องสมุด

4. ผลการดำเนินงาน

ตัวอย่างหน้าเว็บไซต์ของระบบและการทดสอบการใช้งานสามารถแบ่งออกเป็น 4 ส่วน ดังนี้

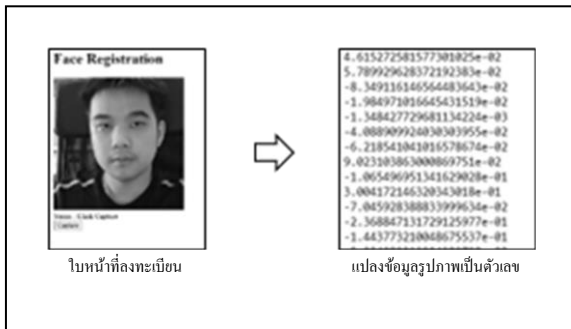
4.1 หน้าลงชื่อเข้าใช้และหน้าเมนูหลัก



ภาพที่ 6: กระบวนการยืนยันตัวตนด้วยใบหน้า(ออนไลน์) หน้าลงชื่อเข้าใช้และหน้าเมนูหลัก

จากภาพที่ 6 หน้าแรกเริ่มที่การเข้าสู่ระบบ โดยคลิกที่ปุ่ม Login จากนั้นจะเข้าสู่หน้าการลงชื่อเข้าใช้โดยใช้อีเมลของ Gmail ในการเข้าสู่ระบบ เป็นการยืนยันตัวตนโดยใช้รหัสผ่าน หลังจากที่ทำกรเข้าสู่ระบบสำเร็จ จะมีการตรวจสอบว่ามีข้อมูลของผู้ใช้งานเก็บอยู่ในระบบ ซึ่งหน้าจอจะมี 2 ลักษณะคือ หน้าจอของผู้ใช้งานใหม่กับหน้าจอผู้ใช้งานปัจจุบัน

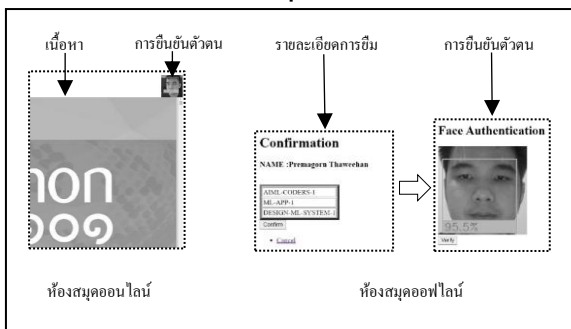
4.2 หน้าลงทะเบียนผู้ใช้งาน



ภาพที่ 7: หน้าลงทะเบียนผู้ใช้งาน

จากภาพที่ 7 เมื่อเลือก New Member จะเข้าสู่ระบบลงทะเบียนใบหน้า ระบบจะเปิดกล้องขึ้นมาแล้วถ่ายรูปโดยคลิกที่ปุ่ม Capture เมื่อเราร่างรูปสำเร็จ ระบบจะทำการแปลงข้อมูลจากรูปเป็นตัวเลขแล้วบันทึกข้อมูล

4.3 หน้ายืนยันตัวตนห้องสมุดออนไลน์และออฟไลน์



ภาพที่ 8: หน้ายืนยันตัวตนห้องสมุดออนไลน์และออฟไลน์

จากภาพที่ 8 ห้องสมุดออนไลน์ ในมุมมองด้านขวาจะเป็นการบันทึกหน้าของผู้ใช้งานจากกล้องหากระบบจดจำใบหน้า ใบหน้าเรามีความใกล้เคียงกับใบหน้าที่ยังบันทึกไว้ในฐานข้อมูลมากกว่าหรือเท่ากับ 70% จะถือว่า Verify ไม่มีอะไรเกิดขึ้น และทำการตรวจสอบใบหน้าเราอยู่เรื่อยๆ แต่หากต่ำกว่า 70% ระบบจะบังคับเราออกจากระบบทันที

ห้องสมุดออฟไลน์ หลังจากทีสแกนหนังสือหรือทรัพยากรอื่น ๆ เสร็จสิ้นจะทำการยืนยันตัวตนโดยใช้ใบหน้าหากต่ำกว่า 70% ระบบจะไม่สามารถยืนยันการยืมได้

4.4 หน้ายืนยันตัวตนห้องสมุดออนไลน์และออฟไลน์

ทดสอบวัดความแม่นยำของประสิทธิภาพการจดจำใบหน้าในสภาพแสงที่แตกต่างกัน ผลทดสอบได้ว่าแสงปกติ, แสงมาก, แสงน้อย มีความแม่นยำ 99.22%, 98.55%, 95.44% ตามลำดับ

จากผลการทดสอบเห็นได้ว่า สภาพแสงปกติมีประสิทธิภาพสามารถยืนยันตัวตนได้ดีที่สุด รองลงมาเป็นสภาพแสงมาก และสุดท้ายเป็นสภาพแสงน้อย ซึ่งทั้ง 3 สภาพแสงยังสามารถมีประสิทธิภาพในการยืนยันตัวตน แต่ถ้าหากแสงน้อยจนกระทั่งสูญเสียรายละเอียดบนใบหน้าจะส่งผลต่อประสิทธิภาพความสามารถการยืนยันตัวตนได้เช่นกัน

ทดสอบวัดระยะเวลาที่ใช้ในการค้นหาใบหน้าทีใกล้เคียงใน Database โดยมีจำนวนใบหน้าที่ไม่ซ้ำกันตั้งแต่ 500 หน้า, 1000 หน้า, 10000 หน้า, 100000 หน้า ผลทดสอบพบว่าใช้เวลาในการค้นหาใบหน้า 0.01, 0.03, 0.04, 0.22 วินาที ตามลำดับ

จากผลการทดสอบพบว่าระยะเวลาที่ใช้ตั้งแต่ 1 รูป ไปจนถึง 500 รูป มีระยะเวลาที่ใช้ใกล้เคียงกัน จากนั้นตั้งแต่ 1,000 รูปเป็นต้นไปมีการเปลี่ยนแปลงของระยะเวลาที่ใช้ อย่างเห็นได้ชัด โดยจำนวนใบหน้าสูงสุดคือ 100,000 หน้า มีระยะเวลาที่ใช้ประมาณ 0.22 วินาที ซึ่งระยะเวลาที่เพิ่มขึ้นนั้นยังไม่ถือว่าส่งผลต่อการใช้งานมาก ดังนั้นถือว่าระบบมีประสิทธิภาพในการยืนยันตัวตน

5. สรุป

การพัฒนาาระบบเทคโนโลยีการจดจำใบหน้าเพื่อยืนยันตัวตนในการใช้บริการห้องสมุด พบว่าการจดจำใบหน้าใบหน้าที่มีความแม่นยำค่อนข้างสูง แม้อยู่ในสถานการณ์ของสภาพแวดล้อมที่แตกต่างกัน สามารถทำงานได้อย่างรวดเร็วแม้มีข้อมูลใบหน้าที่บันทึกไว้ปริมาณมาก

มีประสิทธิภาพเพียงพอต่อการนำไปประยุกต์ใช้เพื่อยืนยันตัวตนหรือบริการด้านอื่น ๆ ของห้องสมุด

การพัฒนาระบบเทคโนโลยีการจดจำใบหน้าเพื่อยืนยันตัวตนสามารถพัฒนาไปสู่ระบบห้องสมุดอัจฉริยะต่อไปได้อีกหลากหลายบริการ และมีอุปกรณ์สำคัญคือ กล้องซึ่งมีผลต่อประสิทธิภาพต่อความแม่นยำต่อการจดจำใบหน้าและการใช้สแกน QR Code, Barcode นอกจากนี้ระบบสามารถจดจำใบหน้ายังมีจุดอ่อนในส่วนที่สามารถใช้รูปถ่ายมาแทนใบหน้าคนๆนั้นจริงๆ ได้ทำให้มีโอกาสถูกปลอมแปลงใบหน้าได้เช่นกัน และสามารถเพิ่มวิธีการยืนยันตัวตนด้วยวิธีการอื่นๆ เพิ่มได้อีก เพื่อเพิ่มประสิทธิภาพด้านความปลอดภัยให้กับระบบ เช่น การยืนยันตัวตนด้วยเสียง เป็นต้น รวมไปถึงการออกแบบหน้าจอผู้ใช้ให้มีความสวยงามและใช้งานง่ายสำหรับผู้ใช้งาน

เอกสารอ้างอิง

- [1] Raj Dhillon, Quynh Chi Nguyen, "Strategies to Respond to a VUCA World," *Lund University Publication Student Papers*, vol. 1, 2021.
- [2] Aylin Kaplan, "LIFELONG LEARNING: CONCLUSIONS FROM A LITERATURE REVIEW," *International Online Journal of Primary Education*, vol. 5, no. 2, 2016.
- [3] Sotonye Orji, Isaac Echezonam Anyira, "What is "Smart" About Smart Libraries," *International Journal of Research in Library Science*, vol. 7, no. 4, 2021.
- [4] Upala, M., "IoT Solution for Smart Library Using Facial Recognition," *IOP Conference Series: Materials Science and Engineering*, vol. 495, 2019.
- [5] Admane, A., Sheikh, A., Paunekar, S., Jawade, S., Wadbude, S., & Sawarkar, M. J., "Library Management System Based on Face Recognition Technique," *2019 IJSRST*, vol. 6, no. 2, pp. 214-217, 2019.
- [6] Yalagi, D. P. S., & Mane, P. V., "Smart library automation using face recognition," *Journal of Physics: Conference Series*, vol. 1854, 2020.
- [7] Scott Rose, Oliver Borchert, Stu Mitchell, Sean Connelly, "Zero Trust Architecture," *National*

Institute of Standards and Technology Special Publication, 2020.

- [8] Geitgey, A., "Machine Learning is Fun Part 4: Modern Face Recognition with Deep Learning," <https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cfc121d78>, 2016.
- [9] Schroff, F., Kalenichenko, D., & Philbin, J., "FaceNet: A Unified Embedding for Face Recognition and Clustering," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 815-823, 2015.
- [10] King, D. E., "Dlib-ml: A Machine Learning Toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755-1758, 2009.

กลไกการจัดสรรแบบเปลี่ยนแปลงได้สำหรับมาตรฐาน 802.11ah

Dynamic Allocation Algorithm for 802.11ah

กณุตม์ อินทร์ใจเอื้อ (Kanut Injaiea)¹ อภิรักษ์ จันทร์สร้าง (Aphirak Jansang)² ชัยพร ใจแก้ว (Chaiporn Jaikaeo)³
 วิรัช ตั้งตรงไพโรจน์ (Withawat Tangtrongpairoj)⁴ และอนันต์ ผลเพิ่ม (Anan Phonphoem)⁵

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ (วิทยาเขตบางเขน)

¹kanut.in@ku.th, ²aphirak.j@ku.ac.th, ³chaiporn.j@ku.ac.th, ⁴withawat.t@ku.ac.th, ⁵anan.p@ku.ac.th

บทคัดย่อ

IEEE 802.11ah เผยแพร่ออกมาเพื่อรองรับอินเทอร์เน็ตในทุกสรรพสิ่ง (IoT) ที่เติบโตอย่างก้าวกระโดด Restricted Access Window (RAW) ช่วยลดจำนวนอุปกรณ์ที่จะเข้าแย่งใช้ช่องสัญญาณ ถึงแม้ว่าจะช่วยลดการชนที่อาจเกิดขึ้นได้ แต่การจัดกลุ่ม RAW ที่ไม่เหมาะสมในสถานะที่มีอุปกรณ์หนาแน่นจะส่งผลให้ประสิทธิภาพลดลงไปอย่างมาก ในงานวิจัยนี้เสนอกระบวนการจัดการ RAW ที่เรียกว่า priority RAW ที่ออกแบบมาเพื่อให้เกิดการชนน้อยและมีการเข้าใช้ช่องสัญญาณอย่างมีประสิทธิภาพ โดยการออกแบบ RAW ที่เหมาะสมพิจารณาจากความน่าจะเป็นที่เกิดการชน จากนั้นออกแบบกระบวนการลำดับ priority ที่เปลี่ยนแปลงไปตามประสิทธิภาพของอุปกรณ์เพื่อที่จะถูกเลือกเข้าไปใน RAW กลุ่มนี้ วิธีการที่นำเสนอสามารถลดอัตราแพ็กเก็ตสูญหายได้ จากผลการทดลองพบว่าวิธีที่นำเสนอมีอัตราแพ็กเก็ตสูญหายน้อยกว่ามาตรฐาน IEEE 802.11ah

คำสำคัญ: อินเทอร์เน็ตในทุกสรรพสิ่ง, หน้าต่างการเข้าถึงที่จำกัด มาตรฐาน 802.11ah เครือข่ายหนาแน่น

Abstract

IEEE 802.11ah was published to support the exponential growth of the Internet of Things (IoT). A Restricted Access Window (RAW) reduces the number of devices contending for channels, although it reduces the potential collision. But the improper grouping of RAW in a dense environment will result in significant performance degradation. This paper proposes a RAW management process called priority RAW, designed to achieve fewer collisions and

efficient channel utilization by designing a suitable RAW based on the collision probability. Then create a dynamic priority that changes according to device performance to be selected into this RAW groups. The proposed method can reduce the packet loss rate. The experimental results found that the proposed method had less packet loss rate than IEEE 802.11ah standard.

Keywords: IoT, RAW, 802.11ah, Dense Network.

1. บทนำ

ในปัจจุบัน อินเทอร์เน็ตในทุกสรรพสิ่ง (IoT) เข้ามามีบทบาทสำคัญและหลากหลายมากขึ้นในชีวิตประจำวัน ซึ่ง IoT ไม่ใช่แค่คอมพิวเตอร์ที่เชื่อมต่อกันเป็นเครือข่ายหรือโทรศัพท์มือถือเท่านั้น แต่ยังรวมไปถึงเซ็นเซอร์และแอคทูเอเตอร์และสิ่งของต่างๆที่ใช้ในชีวิตประจำวันอีกด้วย แนวโน้มการเติบโตของ IoT ที่เพิ่มขึ้นอย่างก้าวกระโดดอาจส่งผลกระทบต่อประสิทธิภาพของระบบเครือข่าย เช่น ปริมาณข้อมูลที่ส่งได้ลดลง อัตราข้อมูลสูญหาย หรือความล่าช้าที่เกิดขึ้น ในระบบเวลาจริง หรือในระบบควบคุม ปัญหาเหล่านี้ส่งผลกระทบเป็นอย่างมาก ตลอดหลายปีที่ผ่านมา IEEE คิดค้นมาตรฐาน IEEE 802.11 ออกมาอย่างมากมาย เพื่อพัฒนาระบบเครือข่ายให้มีความหนาแน่นมากขึ้น และมาตรการสำหรับอุปกรณ์ที่มีความเฉพาะเจาะจง ด้วยธรรมชาติของ IoT ที่มีแพ็กเก็ตขนาดเล็ก รับส่งข้อมูลบ่อยและมีอายุการใช้งานที่สั้น IEEE ได้มีการออกแบบมาตรฐาน IEEE 802.11ah เพื่อรองรับและเพิ่มประสิทธิภาพการทำงานของระบบ IoT

IEEE 802.11ah หรือในที่รู้จักอีกชื่อว่า Wi-Fi Halow เป็นมาตรฐานที่ใช้งานอยู่บนย่านความถี่ต่ำกว่า 1 กิกะเฮิรตซ์ ที่ไม่ได้ขึ้นทะเบียน ถูกออกแบบมาเพื่อรองรับการใช้งานของ IoT มาตรฐาน IEEE 802.11ah สืบทอดคุณลักษณะหลายอย่างของ IEEE 802.11n ทั้งชั้น Physical (PHY) และชั้น Data link (MAC) ซึ่งในชั้น Physical มาตรฐานนี้ทำงานอยู่บนย่านความถี่ต่ำกว่า 1 กิกะเฮิรตซ์ ที่ไม่ได้ขึ้นทะเบียน ทำให้ 802.11ah มีระยะครอบคลุมได้ถึง 1 กิโลเมตร โดยไม่ต้องมีตัวกลางในการส่งข้อมูล และกินพลังงานน้อยกว่า ในชั้น MAC มีคุณสมบัติหลักมากมาย เช่น ส่วนเสกเตอร์ของ MAC แบบสั้น, Traffic Indication Map (TIM) และ Restricted Access Window (RAW) เพื่อแก้ไขปัญหาอุปกรณ์ IoT ที่เชื่อมต่อกันอย่างหนาแน่น

หนึ่งในกลไกการทำงานหลักของ IEEE 802.11ah คือ RAW ซึ่งจะทำการแบ่งอุปกรณ์ที่เข้าใช้งานออกเป็นกลุ่ม และจะทำการจำกัดการเข้าถึงช่องสัญญาณให้เฉพาะกับอุปกรณ์ที่อยู่ในกลุ่มเพียงเท่านั้น เพื่อลดจำนวนอุปกรณ์ที่แย่งเข้าใช้ช่องสัญญาณ ซึ่งในแต่ละอุปกรณ์จะมีกลไกการเข้าใช้งานอยู่สองกลไกคือ Distributed Coordination Function (DCF) และ Enhanced Distributed Channel Access (EDCA) โดยอุปกรณ์จะใช้กลไก DCF เมื่ออยู่นอกการทำงานของ RAW และจะใช้ EDCA เมื่ออยู่ในการทำงานของ RAW

ถึงแม้ว่า IEEE 802.11ah ถูกออกแบบมาเพื่อรองรับกับอุปกรณ์ IoT แต่มีปัญหาหลายอย่างที่มิงานวิจัยออกมามากมายเพื่อแก้ปัญหาและพัฒนาประสิทธิภาพของมาตรฐานนี้ หนึ่งในนั้นคือปัญหาที่เกี่ยวข้องกับกลไกการทำงาน ของ RAW ในระบบเครือข่ายที่มีขนาดใหญ่ ถึงแม้ว่าถูกออกแบบมาเพื่อรองรับการใช้งานของอุปกรณ์ที่รวมตัวอย่างหนาแน่น การออกแบบขนาด RAW ก็ส่งผลอย่างมากต่อประสิทธิภาพโดยรวมของระบบเครือข่าย ในบทความนี้เสนอกระบวนการจัดการ RAW โดยแบ่งออกเป็น priority RAW และ non-priority RAW ซึ่งใน priority RAW ออกแบบมาเพื่อให้อุปกรณ์ที่อยู่ในกลุ่มนี้เกิดการชนน้อยที่สุด และได้รับทรัพยากรอย่างเต็มที่ และออกแบบค่าคะแนนประจำตัวอุปกรณ์ที่คำนวณจาก

ประสิทธิภาพของอุปกรณ์เพื่อนำมาใช้ในการพิจารณาคัดเลือกเข้า priority RAW

2. วรรณกรรมที่เกี่ยวข้อง

มาตรฐาน IEEE 802.11ah มีจุดมุ่งหมายเพื่อแก้ไข ปัญหาหลักของ IoT อย่างเช่น การประหยัดพลังงาน และการเข้าแย่งใช้ช่องสัญญาณในระบบเครือข่ายที่หนาแน่น ถึงแม้ว่ามาตรฐานนี้ออกแบบมาสำหรับ IoT แต่ยังมีปัญหาหลายอย่างที่ต้องแก้ไขและพัฒนาเพิ่มขึ้น มีงานมากมายที่มุ่งเน้นไปที่กลไกการทำงานของ RAW ซึ่งเป็นการจัดกลุ่มอุปกรณ์เพื่อลดโอกาสการชนของ RAW มีกลุ่มตัวแปรที่ใช้ในการจัดกลุ่มเรียกว่า RAW parameter set (RPS) ซึ่งเป็นกลุ่มตัวแปรที่ใช้ในการออกแบบลักษณะของกลุ่ม RAW ใน [1] กล่าวว่ากลไกการจัดกลุ่มส่งผลต่อประสิทธิภาพของ RAW เป็นอย่างมาก และได้มีการศึกษาผลกระทบของพารามิเตอร์ต่างๆของ RAW และพิจารณาว่าการจัดกลุ่มที่เหมาะสมตามพารามิเตอร์หลายตัวและกลยุทธ์แบบปรับเปลี่ยนกลุ่มจำเป็นต่อการเพิ่มประสิทธิภาพภายใต้สภาวะที่เปลี่ยนแปลงตลอดเวลา

เพื่อที่จะพัฒนาประสิทธิภาพกลไกการทำงาน ของ RAW มีกลุ่มของงานวิจัยที่มุ่งไปที่การหาขนาดของ RAW ที่ดีที่สุด ใน [2] คาดการณ์จำนวนอุปกรณ์เพื่อใช้งาน uplink จากความน่าจะเป็นที่ส่งข้อมูลได้สำเร็จและนำมาใช้ในการหาขนาดของ RAW ที่เหมาะสม มีหลายงานวิจัยที่เริ่มหาขนาด RAW โดยใช้คุณภาพของเครือข่ายเป็นตัวกำหนด ใน [3] ได้มีการใช้เงื่อนไขของข้อมูลที่อุปกรณ์ต้องการเพื่อปรับขนาด RAW แบบเวลาจริง ถึงแม้ว่างานวิจัยส่วนใหญ่จะเริ่มอ้างอิงตามเงื่อนไขในการหาขนาด แต่ในสถานการณ์จริงมีความซับซ้อนมากกว่าที่เกิดขึ้นบนทฤษฎี ในสถานการณ์จริง ระบบเครือข่ายสามารถมีอุปกรณ์ที่มีเงื่อนไขความต้องการและ Quality of Service (QoS) ที่หลากหลาย ใน [4] ได้สร้างโมเดลที่สามารถปรับใช้งานตามสถานการณ์ในโลกจริงได้หลากหลายรูปแบบ แต่ในบางสถานการณ์เช่น ในระบบควบคุมที่ระบบต้องการข้อมูลแบบเวลาจริงอาจเป็นปัญหาได้ใน [5] ได้มีการสร้างการจัดกลุ่มที่มีความสำคัญสูงขึ้นมาสำหรับเครือข่ายที่

อุปกรณ์มี QoS ที่แตกต่างกัน แต่ในงานนี้ อุปกรณ์บางตัวมี QoS ที่สูงมากกว่าอุปกรณ์อื่น ถ้าทุกอุปกรณ์มี QoS ที่เท่ากันในสภาวะที่มีทรัพยากรไม่เพียงพอ การกระจายทรัพยากรอย่างเหมาะสมจึงมีความท้าทายเพิ่มเข้ามา

ถ้าอุปกรณ์ส่วนใหญ่มีความต้องการเท่ากัน ปัญหาหลักที่เกิดขึ้นคือการจัดสรรทรัพยากรให้เหมาะสม และอีกปัญหาหลักของกลไกการทำงานของ RAW คือ DCF/EDCA ซึ่งเป็นกลไกในการแย่งเข้าใช้สื่อ ถึงแม้ว่า RAW จะแบ่งอุปกรณ์เป็นกลุ่มขนาดเล็ก แต่ก็มีการชนเกิดขึ้นได้ ในกรณีที่เวลาเป็นสิ่งสำคัญ การชนสามารถส่งผลกระทบต่อประสิทธิภาพเป็นอย่างมาก ในงาน [6] ทำนายรูปแบบของการส่งข้อมูล และจัดสรรตารางไว้ล่วงหน้า และใน [7] เสนอจัดขนาด RAW ตามการขอเข้าใช้งานของอุปกรณ์ แต่การจัดสรรทรัพยากรไว้ล่วงหน้าอาจจะไม่เหมาะสมถ้ามีการส่งข้อมูลที่กระชั้นกันเกิดขึ้น

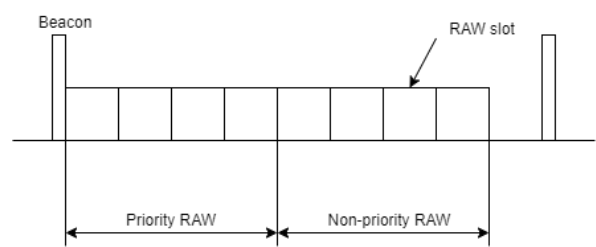
3. วิธีดำเนินงานวิจัย

จากการศึกษาพบว่า การจัดกลุ่ม RAW ส่งผลต่อประสิทธิภาพของมาตรฐาน 802.11ah ในบทความนี้เสนอกระบวนการจัดการ RAW โดยการสร้างกลุ่ม priority RAW และ non-priority RAW ซึ่งในกลุ่ม priority RAW ถูกออกแบบมาให้ตัวอุปกรณ์ที่ถูกจัดอยู่กลุ่มนี้มีจำนวนที่เหมาะสมเพื่อที่จะได้รับทรัพยากรอย่างเต็มที่และเพิ่มประสิทธิภาพของอุปกรณ์ที่ถูกเลือกเข้ามาอยู่ในกลุ่ม อุปกรณ์จะมีค่าคะแนนประจำตัวที่ออกแบบขึ้นเพื่อใช้เป็นเกณฑ์ในการพิจารณา ซึ่งค่าคะแนนจะมีการเปลี่ยนแปลงตลอดเวลาขึ้นอยู่กับประสิทธิภาพของอุปกรณ์ขณะนั้น โดยตัวงานทั้งหมดจะแบ่งออกได้เป็นสองส่วน คือ การจัดปริมาณของอุปกรณ์ให้เหมาะสมกับกลุ่ม priority RAW และการเลือกเกณฑ์ที่จะนำมาใช้ในการจัดกลุ่ม

3.1 รูปแบบการจัดกลุ่มและวิธีการเปลี่ยนกลุ่ม RAW

การออกแบบ RAW เป็นไปตามภาพที่ 1 ในหนึ่งช่วง beacon กำหนดให้มีแค่ 2 กลุ่ม RAW คือ priority RAW และ non-priority RAW ซึ่งในแต่ละกลุ่มมีจำนวน slot ที่เท่ากัน และคำนวณช่วงเวลาของแต่ละ slot ให้มากที่สุดเพื่อลดช่วงที่ไม่มีการส่งข้อมูลระหว่าง beacon เริ่มต้น

อุปกรณ์ที่ใช้ 802.11ah แต่ละอุปกรณ์จะถูกจัดให้อยู่ใน priority RAW ในจำนวนที่เหมาะสม ส่วนอุปกรณ์ที่เหลือจะถูกจัดลงใน non-priority RAW ซึ่งคะแนนของอุปกรณ์ในตอนเริ่มต้นจะมีค่าเท่ากัน เมื่อระยะเวลาผ่านไปคะแนนจะมีการเปลี่ยนแปลงไปตามประสิทธิภาพของตัวอุปกรณ์ ยิ่งตัวอุปกรณ์มีประสิทธิภาพต่ำ คะแนนยิ่งมากตามไปด้วย อุปกรณ์ที่มีคะแนนมากจะถูกเลือกเข้าไปอยู่ใน priority RAW เพื่อเพิ่มประสิทธิภาพ ส่วนอุปกรณ์ใน priority RAW ที่มีคะแนนน้อยจะถูกสลับให้ไปอยู่ใน non-priority RAW



ภาพที่ 1: รูปแบบการจัด RAW ที่นำเสนอ

3.2 วิธีเลือกจำนวนอุปกรณ์ที่เหมาะสมกับ priority RAW

อุปกรณ์ที่อยู่ใน RAW จะถูกแบ่งอย่างเท่ากันลงใน slot เพื่อลดจำนวนอุปกรณ์การเข้าแย่งใช้ทรัพยากร ยังมีอุปกรณ์ใน RAW มาก จำนวนอุปกรณ์ที่แย่งใช้ทรัพยากรในหนึ่ง slot ยิ่งมากตามไปด้วย ก่อให้เกิดการชนกันเกิดขึ้น ยิ่งชนกันมากทำให้ประสิทธิภาพโดยรวมของระบบลดลง ฉะนั้นการคำนวณเพื่อเลือกจำนวนอุปกรณ์ที่เหมาะสมเพื่อให้เกิดการชนน้อยที่สุดใน priority RAW จึงมีความสำคัญ ในการที่อุปกรณ์หนึ่งตัวจากอุปกรณ์ทั้งหมด (N) ในหนึ่ง slot สามารถส่งข้อมูลได้สำเร็จเกิดจากอุปกรณ์ตัวใดตัวหนึ่งเลือกเวลารอที่น้อยที่สุด (CW_{min}) ฉะนั้นการคำนวณความน่าจะเป็นที่อุปกรณ์ส่งได้สำเร็จ ($P_{success}$) ได้ตามสมการที่ 3-1

$$P_{success} = \sum_{i=1}^{CW_{min}-1} \frac{(CW_{min}-i)^{N-1}}{CW_{min}^N} \quad (3-1)$$

ความน่าจะเป็นที่อุปกรณ์มีการชนเกิดขึ้น (P_{col}) จะมีค่าเท่ากับ $1-P_{success}$ ซึ่งใน RAW หนึ่งกลุ่มมี slot อยู่ L ช่อง แต่ละ slot มีโอกาสถูกเลือกเท่ากันคือ $1/L$ ความน่าจะเป็น

เป็นที่เกิดการชน (P_{RAW}) คิดคำนวณจากความน่าจะเป็นที่เกิดการชนใน slot i ใดๆ (P_i) ได้ดังสมการที่ 3-2

$$P_{RAW} = \sum_{i=1}^L \left(\frac{1}{L}\right) p_i \tag{3-2}$$

โดยปกติอุปกรณ์ที่อยู่ในกลุ่มจะถูกแบ่งลง slot เท่ากัน y อุปกรณ์ แต่กรณีที่แบ่งอุปกรณ์ไม่ลงตัวทำให้บาง slot มีจำนวนอุปกรณ์มากกว่า slot อื่นๆอยู่ m อุปกรณ์ เมื่อคิดรวมสองกรณีเข้าด้วยกันแล้ว P_{RAW} จะได้ดังสมการที่ 3-3

$$P_{RAW} = \left(\frac{m}{L}\right) \left[1 - \sum_{i=1}^{CW_{min}^{-1}} \frac{(CW_{min} - i)^y}{CW_{min}^{y+1}} \right] + \left(1 - \frac{m}{L}\right) \left[1 - \sum_{i=1}^{CW_{min}^{-1}} \frac{(CW_{min} - i)^{y-1}}{CW_{min}^y} \right] \tag{3-3}$$

การหาจำนวนอุปกรณ์ที่เหมาะสมกับ priority RAW สามารถคำนวณได้จากการหาความน่าจะเป็นที่จะเกิดการชนใน RAW หนึ่งกลุ่ม เพราะการใส่ปริมาณที่มากเกินไปทำให้เกิดการชนเยอะและเสื่อน้อยเกินไปทำให้ priority RAW ทำงานโดยไม่มีประสิทธิภาพ

3.3 เกณฑ์ที่ใช้ในการพิจารณาเข้ากลุ่ม priority RAW

จากที่กล่าวมาข้างต้น เกณฑ์ที่ใช้ในการพิจารณาคือค่าคะแนนประจำตัวอุปกรณ์ ซึ่งคำนวณมาจากประสิทธิภาพของตัวอุปกรณ์ขณะนั้น ในบทความนี้เลือกเกณฑ์ที่จะนำมาใช้ในการพิจารณาสามค่าคือ ดีเลย์ เปรอร์เซ็นต์แพ็กเก็ตที่สูญหาย และ ความยาวของคิว เพราะเป็นค่าที่ส่งผลต่อประสิทธิภาพของระบบ โดยมีการปรับสมดุลและเพิ่มความเท่าเทียมด้วย $w_d w_l w_q$ ซึ่งเป็นค่าน้ำหนักของเกณฑ์ทั้งสามค่าตามลำดับ คะแนนจะคำนวณดังสมการที่ 3-4

$$Score = w_d D + w_l L + w_q Q \tag{3-4}$$

อุปกรณ์ที่มีประสิทธิภาพต่ำจะส่งผลให้คะแนนสูงและจะถูกเลือกเข้าไปใน priority RAW เพื่อเพิ่มประสิทธิภาพ ในขณะที่อุปกรณ์ที่มีคะแนนน้อยกว่าจะถูกสลับมาอยู่ใน non-priority RAW โดยกระบวนการสลับกลุ่มเป็นไป

ตาม Algorithm 1 ซึ่งจะมีการคำนวณคะแนนทุกช่วงเวลาหนึ่ง ทำให้เกิดการสลับเปลี่ยนกลุ่มอย่างต่อเนื่อง ส่งผลให้เกิดการกระจายทรัพยากรไปยังอุปกรณ์อย่างเท่าเทียมในกรณีที่มียุอุปกรณ์รวมตัวกันหนาแน่น

Algorithm 1: NewAID Calculation Function

```

Input: List of score  $S$ , Number of devices  $N$ 
Output: List of new AID  $A$ 
1 Sort  $S$  in descending order and assigned to  $X$ 
2 Initialize  $Y$  as empty list
3 for  $i=1,2,\dots,N$  do
4   for  $j=1,2,\dots,N$  do
5     if index  $i$  of list  $S =$  index  $j$  of list  $X$ 
6       then
7         Get  $j$  in list  $Y$ 
8       end
9   end
10 return list  $Y$ 
    
```

4. ผลการดำเนินงาน

ในบทความนี้ได้มีการวัดผลของมาตรฐาน IEEE 802.11ah กับวิธีการที่นำเสนอขึ้น ด้วยโปรแกรมจำลองระบบเครือข่าย Network Simulator 3 (NS3) มีการตั้งค่าตามตารางที่ 1 โดยจะทำการแบ่งการทดลองออกเป็นสามส่วน

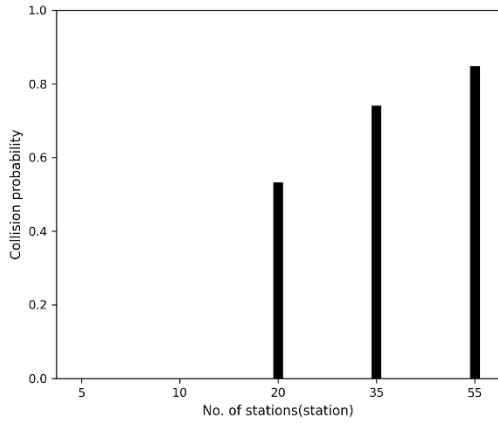
ตารางที่ 1: พารามิเตอร์ที่ใช้ในการทดลอง

ตัวแปร	ค่า
Payload	256 bytes
Beacon interval	100 ms
Bandwidth	2 MHz
Data rate	2 Kbps, 4Kbps, 6Kbps
Traffic type	UDP
Simulation time	300 s

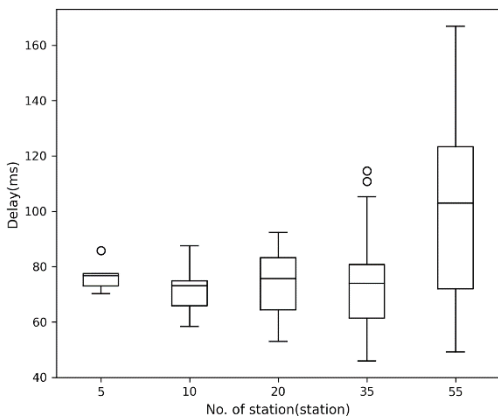
4.1 ความสัมพันธ์ระหว่างความน่าจะเป็นในการชนกับผลที่เกิดขึ้นจริงจากการจำลอง

ในการทดลองนี้ออกแบบให้เป็น RAW 1 กลุ่ม มีขนาด 10 slot และให้ CW_{min} เท่ากับ $15 \mu s$ เป็นการดูแลแนวโน้มของความน่าจะเป็นที่ได้จากสมการที่ 3-3 กับผลที่เกิดจากการจำลอง จากภาพที่ 2 จะเห็นว่ากรณีอุปกรณ์ 20 อุปกรณ์ มีจำนวนอุปกรณ์เป็นสองเท่าของจำนวน slot มีความน่าจะเป็น

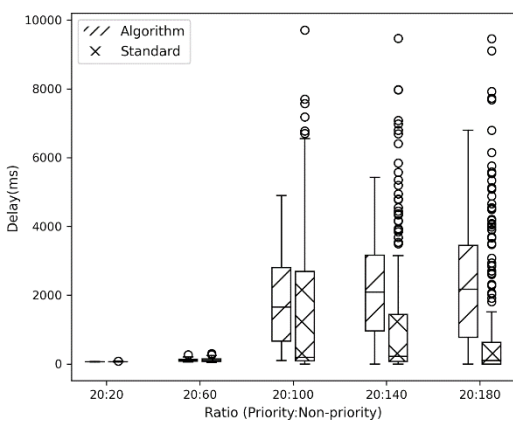
เป็นในการชนถึง 50 เปอร์เซ็นต์ แต่ดูจากภาพที่ 3 จะเห็นว่ามีตัวเลขไม่ต่างกันมากกับกรณีของ 5 อุปกรณ์และ 10 อุปกรณ์ แต่ถ้ามีการชนมากกว่า 50 เปอร์เซ็นต์ขึ้นไป ค่าตัวเลขจะมีความต่างชัดเจนมากขึ้น



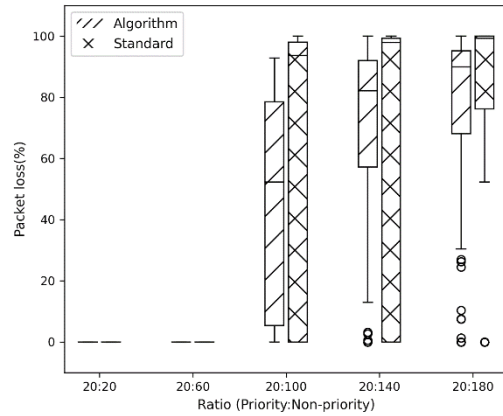
ภาพที่ 2: ผลความน่าจะเป็นที่เกิดการชน



ภาพที่ 3: ผลตัวเลขที่เกิดขึ้นจากการจำลอง



ภาพที่ 4: ตัวเลขเมื่อจำนวนอุปกรณ์เพิ่มขึ้น



ภาพที่ 5: อัตราการสูญเสียแพ็คเกจเมื่อจำนวนอุปกรณ์เพิ่มขึ้น

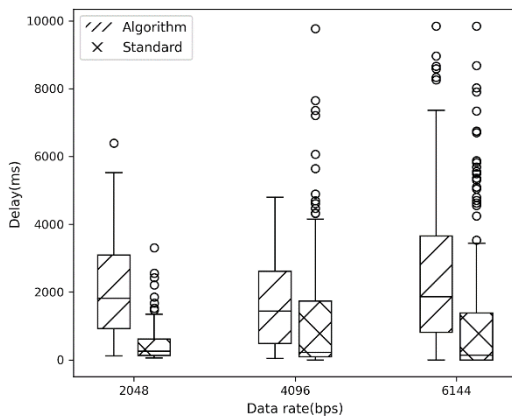
4.2 การเปรียบเทียบระหว่างวิธีการที่เสนอกับมาตรฐาน 802.11ah ที่มีการจัดรูปแบบ RAW เหมือนกัน

การทดลองนี้มีการแบ่ง RAW เป็น 2 กลุ่มคือ priority RAW และ non-priority RAW ในแต่ละ RAW มี 10 slot โดย priority RAW จะถูกจัดอุปกรณ์ลงแค่ 20 ตัว ส่วนอุปกรณ์ที่เหลือจะถูกจัดให้อยู่ใน non-priority RAW ซึ่งวิธีการที่เสนออุปกรณ์จะมีการเปลี่ยนกลุ่มเกิดขึ้น ในขณะที่อุปกรณ์ที่ใช้มาตรฐานปกติจะไม่มีเปลี่ยนกลุ่ม จากภาพที่ 4 จะเห็นว่าในกรณีที่มีอุปกรณ์ในระบบ 40 กับ 80 อุปกรณ์ ทั้งสองวิธีจะมีตัวเลขไม่ต่างกัน เมื่อมีอุปกรณ์มากกว่า 120 อุปกรณ์ขึ้นไป วิธีที่เสนอมีตัวเลขน้อยกว่าวิธีมาตรฐาน โดยมีความต่างของค่ามัธยฐานอยู่ที่ 1,458.22 1,867.05 และ 2,071.23 ms ตามลำดับ แต่ถ้าพิจารณาภาพที่ 5 จะเห็นว่าในกรณีที่มี 120 อุปกรณ์ขึ้นไป วิธีที่เสนอมีอัตราแพ็คเกจสูญหายน้อยกว่าวิธีมาตรฐาน โดยมีความต่างของค่ามัธยฐานอยู่ที่ 41.37 15.82 และ 9.37 เปอร์เซ็นต์ ตามลำดับ ฉะนั้นเมื่อมีอุปกรณ์ในระบบมากขึ้น วิธีที่นำเสนอจะมีอัตราการสูญเสียแพ็คเกจน้อยกว่าวิธีของมาตรฐานปกติ

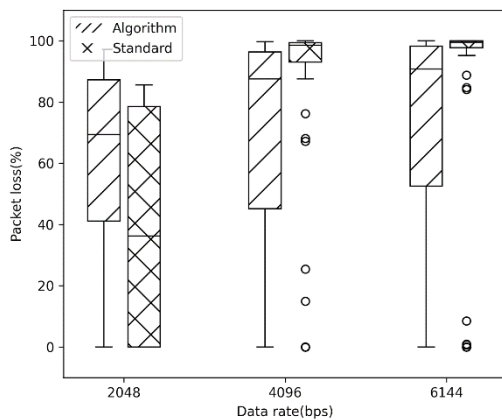
4.3 การเปรียบเทียบระหว่างวิธีการที่เสนอกับมาตรฐาน 802.11ah ที่มีการจัดรูปแบบ RAW ต่างกัน

ในวิธีมาตรฐานมีการจัด RAW เป็นแบบกลุ่มเดียว 20 slot ส่วนวิธีที่นำเสนอยังคงรูปแบบการจัด RAW เดิมไว้ซึ่งในการทดลองนี้จะทำการคงที่จำนวนอุปกรณ์ แต่จะเปลี่ยนความถี่ของข้อมูลที่ถูกส่งจากตัวอุปกรณ์ จากภาพที่ 6 อัตราการส่งข้อมูลที่สูงขึ้นส่งผลให้แนวโน้มตัวเลขของวิธีการที่เสนอมากกว่าวิธีมาตรฐานทุกกรณี โดยมีความต่างของค่ามัธยฐาน

อยู่ที่ 1,554.31 1,216.49 และ 1,729.25 ms ตามลำดับ แต่ถ้าพิจารณาภาพที่ 7 ร่วมกันจะเห็นว่าในกรณีที่อัตราการส่งข้อมูล 2,048 bps วิธีการที่เสนอมีค่ามัธยฐานของอัตราแพ็กเก็ตสูญหายมากกว่า 33.13 เปอร์เซ็นต์ ส่วนกรณีที่เหลือวิธีการที่เสนอมีค่ามัธยฐานของอัตราแพ็กเก็ตสูญหายน้อยกว่า 11.02 และ 8.78 เปอร์เซ็นต์ ตามลำดับ ฉะนั้นเมื่อมีอัตราการส่งข้อมูลที่สูงกว่า 2,048 bps ขึ้นไป วิธีที่นำเสนอจะมีอัตราการสูญเสียบั๊กเก็ตน้อยกว่าวิธีของมาตรฐานปกติ



ภาพที่ 6: ค่าเฉลี่ยจากการเปลี่ยนอัตราการส่งข้อมูล



ภาพที่ 7: อัตราการสูญเสียบั๊กเก็ตจากการเปลี่ยนอัตราการส่งข้อมูล

5. สรุป

ในงานนี้ได้มีการนำเสนอกระบวนการจัดการ priority RAW ขึ้นมา โดยมีการออกแบบขนาด RAW ที่เหมาะสมตามจำนวนอุปกรณ์ที่ใส่เข้าไป ซึ่งอ้างอิงจากความน่าจะเป็นในการชน และออกแบบค่าคะแนนประจำตัวอุปกรณ์เพื่อใช้ในการพิจารณาเลือกอุปกรณ์เข้าไปอยู่ใน priority RAW

จากการทดลองเปรียบเทียบประสิทธิภาพกับมาตรฐานปกติ จะเห็นว่าวิธีการที่นำเสนอมีอัตราการสูญเสียบั๊กเก็ตที่น้อยกว่าวิธีมาตรฐาน แต่ทางด้านดีเลย์ยังไม่สามารถเปรียบเทียบได้อย่างชัดเจน เพราะวิธีการมาตรฐานมีอัตราการสูญเสียบั๊กเก็ตมาก จึงทำให้ดีเลย์จากบั๊กเก็ตที่สูญหายไม่ถูกนำมาคำนวณผลด้วย แต่จากผลที่ได้ทำให้เห็นว่าการสร้างกลุ่มที่ออกแบบให้มีการชนเกิดขึ้นน้อยและมีทรัพยากรเพียงพอ และมีการสลับอุปกรณ์ที่มีประสิทธิภาพต่ำเข้ามาอยู่ในกลุ่มสามารถเพิ่มประสิทธิภาพให้กับระบบเครือข่ายได้

เอกสารอ้างอิง

- [1] Tian, L., J. Famaey, and S. Latré. "Evaluation of the IEEE 802.11ah Restricted Access Window mechanism for dense IoT networks." in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. 2016.
- [2] Park, C.W., D. Hwang, and T.J. Lee, "Enhancement of IEEE 802.11ah MAC for M2M" *Communications. IEEE Communications Letters*, 2014. **18**(7): p. 1151-1154.
- [3] Tian, L., et al., "Real-Time Station Grouping under Dynamic Traffic for IEEE 802.11ah". *Sensors*, 2017. **17**.
- [4] Tian, L., et al. "IEEE 802.11ah Restricted Access Window Surrogate Model for Real-Time Station Grouping." in *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*. 2018.
- [5] Ahmed, N., D. De, and M.I. Hussain. "A QoS-aware MAC protocol for IEEE 802.11ah-based Internet of Things." in *2018 Fifteenth International Conference on Wireless and Optical Communications Networks (WOCN)*. 2018.
- [6] Ahmed, N. and M.I. Hussain, "Periodic Traffic Scheduling for IEEE 802.11ah Networks." *IEEE Communications Letters*, 2020. **24**(7): p. 1510-1513.
- [7] Lei, X. and S.H. Rhee, "Performance Improvement of Sub 1 GHz WLANs for Future IoT Environments". *Wireless Personal Communications*, 2017. **93**(4): p. 933-947.

ความตระหนักรู้ ความพร้อมและตัวชี้วัดด้านความมั่นคงปลอดภัยไซเบอร์ ในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย Cybersecurity Awareness, Readiness and Indicator in the Armed Force Security Center, Royal Thai Armed Forces Headquarters

ทศพร เครือวัลย์ (Todsaporn Klearwun)¹ สุรศักดิ์ มั่งสิงห์ (Surasak Mungsing)²

ประสาน วงษ์กิติโสภณ (Prasan Wongkitisopon)³ ชนะภัทร ภัทรเมธากุล (Chanapatt Pattaramaetakul)⁴

และ ประสงค์ ปราณีตพลกรัง (Prasong Praneetpolgrang)⁵

^{1,2,3,4} สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม

⁵ กองการศึกษา โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

¹k.todsaporn52@gmail.com, ²surasak.mu@spu.ac.th, ³prasan@gmesh.com, ⁴chanapatt.p@gmail.com, ⁵prasong.pr@spu.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อวิเคราะห์ภัยคุกคามทางไซเบอร์ที่มีต่อศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย 2) เพื่อศึกษาระดับความตระหนักรู้ ความพร้อม และตัวชี้วัดด้านความมั่นคงปลอดภัยไซเบอร์ในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย งานวิจัยนี้เป็นการวิจัยแบบผสมผสาน โดยทำการสัมภาษณ์เชิงลึกผู้เชี่ยวชาญ จำนวน 9 คน และใช้แบบสอบถามกลุ่มตัวอย่างเป็นข้าราชการ พนักงานราชการและลูกจ้าง สังกัดศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย จำนวน 350 ตัวอย่าง ผลการวิจัยพบว่า ปัญหาและภัยคุกคามทางไซเบอร์ที่มีต่อหน่วยงานของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย ได้แก่ มัลแวร์ นโยบายและแนวปฏิบัติที่ล้าสมัย ความซับซ้อนของเครือข่ายและทรัพยากรที่จำกัด สำหรับด้านความพร้อมทางไซเบอร์ในภาพรวมของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย อยู่ที่ 80.27% โดยมีความพร้อมด้านกระบวนการอยู่ที่ 82.00% รองลงมาคือความพร้อมด้านบุคลากร 79.60% และความพร้อมด้านเทคโนโลยี 79.10% ในขณะที่ตัวชี้วัดด้านความมั่นคงปลอดภัยไซเบอร์ พบว่ามีจำนวนทั้งหมด 7 ตัวชี้วัด

คำสำคัญ: ความตระหนักรู้ ความมั่นคงปลอดภัยไซเบอร์

Abstract

This research has the following objectives: 1) to analyze the cyber threats to the Armed Force Security Center, Royal Thai Armed Forces Headquarters. 2) to study the readiness, awareness, and cybersecurity indicators of the Armed Force Security Center, Royal Thai Armed Forces Headquarters. This research used a mixed methodology. In-depth interviews with 9 experts were conducted, and a questionnaire was distributed to a sample of 350 civil servants, government employees, and employees of the Armed Force Security Center (AFSC), Royal Thai Armed Forces Headquarters (RTARF). The research results indicated that cybersecurity problems and cyber threats to the Armed Forces Security Center, Royal Thai Armed Forces Headquarters include malware attacks, outdated policies and practices, network complexity, and limited resources. The overall cybersecurity readiness of (AFSC) was at 80.27%, process readiness at 82.00%, personnel readiness at 79.60%, and technological readiness at 79.10%. Research results also showed that there are 7 cybersecurity indicators.

Keywords: Awareness, Cybersecurity.

1. บทนำ

ปัจจุบันภัยคุกคามทางไซเบอร์กลายเป็นความกังวลสำหรับหน่วยงานรัฐบาลทั่วโลก ซึ่งผู้โจมตีมีเป้าหมายทั้งภาครัฐและเอกชน เช่น โรงพยาบาล ธนาคาร และหน่วยงานทางทหาร โดยเป็นการโจมตีที่มีเป้าหมายเพื่อให้ได้มาซึ่งข้อมูลสำคัญ ทำให้ข้อมูลหรือระบบสารสนเทศไม่สามารถเข้าถึงได้ หรือทำลายชื่อเสียงขององค์กร

บุคลากรถือเป็นสาเหตุหลักของช่องโหว่ในโครงสร้างพื้นฐานที่สำคัญ การขาดความตระหนักรู้ของบุคลากรมักเป็นสาเหตุหลักของการบุกรุกเครือข่าย ทำให้จำเป็นต้องลดพฤติกรรมเสี่ยงของบุคลากรเพื่อเพิ่มประสิทธิภาพด้านการรักษาความมั่นคงปลอดภัยไซเบอร์ขององค์กร

ศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย มีการใช้ระบบสารสนเทศในการเพิ่มประสิทธิภาพในการทำงาน ทำให้มีความเสี่ยงต่อภัยคุกคามทางไซเบอร์ เพื่อเตรียมพร้อมรับมือกับภัยคุกคามที่อาจเกิดขึ้นในอนาคต หน่วยงานต้องมีมาตรการป้องกันภัยคุกคามทางไซเบอร์ที่เข้มงวด รัดกุม และเพียงพอตามความเสี่ยงทางไซเบอร์ที่มีอยู่

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความมั่นคงปลอดภัยไซเบอร์

ความมั่นคงปลอดภัยไซเบอร์ เป็นการปกป้องข้อมูลจากการโจมตีทางไซเบอร์ที่หลากหลาย เพื่อรักษาความลับ ความครบถ้วน และความพร้อมใช้งานของเทคโนโลยีดิจิทัล โดยการดำเนินการนั้นจะต้องอาศัย บุคลากร กระบวนการทำงาน และเครื่องมือที่ทันสมัย [3]

2.2 ภัยคุกคามทางไซเบอร์

ภัยคุกคามทางไซเบอร์ คือ เหตุการณ์หรือสถานการณ์ใด ๆ ซึ่งส่งผลเสียและสร้างความเสียหายต่อบุคคลด้านการรักษาความลับ การรักษาความครบถ้วนและสภาพความพร้อมใช้งานของระบบสารสนเทศหรือข้อมูลของทั้งส่วนบุคคล องค์กรหรือประเทศชาติ [4]

2.3 ความผิดพลาดจากบุคลากร

บุคลากรถือเป็นช่องโหว่ด้านความปลอดภัยไซเบอร์ที่ใหญ่ที่สุด โดยบุคลากรเป็นทั้งผู้ใช้งานคอมพิวเตอร์และ

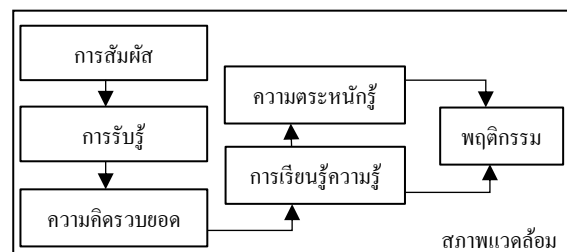
นักวิเคราะห์ระบบ ซึ่งมักพบพฤติกรรมที่ทำให้เกิดความผิดพลาดในการรักษาความมั่นคงปลอดภัยไซเบอร์ขององค์กร เช่น การแชร์รหัสผ่าน การเข้าถึงเว็บไซต์ที่น่าสงสัยและไม่อัปเดตซอฟต์แวร์ เป็นต้น [5]

2.4 บทบาทของพฤติกรรมมนุษย์ในการรักษาความปลอดภัยทางไซเบอร์

พฤติกรรมมนุษย์เป็นปัจจัยสำคัญที่มีอิทธิพลต่อความสำเร็จของการรักษาความมั่นคงปลอดภัยไซเบอร์ในองค์กร เพื่อให้เข้าใจความสัมพันธ์ระหว่างพฤติกรรมมนุษย์กับความมั่นคงปลอดภัยไซเบอร์ องค์กรควรจัดทำนโยบายที่มีประสิทธิภาพ เพื่อส่งเสริมการรับรู้และการศึกษาเกี่ยวกับความมั่นคงปลอดภัยไซเบอร์และควรมีการใช้เทคโนโลยีเพื่อลดความเสี่ยงทางไซเบอร์ [6]

2.5 ทฤษฎีความตระหนักรู้

ความตระหนักรู้ (Awareness) เป็นแนวคิดเชิงจิตวิทยา (Psychological Approach) ที่เกี่ยวข้องกับการรับรู้และเข้าใจสิ่งเร้าจากสิ่งแวดล้อมส่งผลให้เกิดความคิดรวบยอดซึ่งนำไปสู่การเรียนรู้และส่งผลต่อพฤติกรรมของบุคคลต่อสิ่งเร้านั้น ๆ ในที่สุด [7] แสดงดังภาพที่ 1



ภาพที่ 1: กระบวนการของการเกิดความตระหนักรู้

2.6 การฝึกอบรมการรับรู้ความมั่นคงปลอดภัยไซเบอร์

การฝึกอบรมการรับรู้ความมั่นคงปลอดภัยไซเบอร์เป็นองค์ประกอบที่สำคัญของกลยุทธ์ความมั่นคงปลอดภัยไซเบอร์ในภาพรวมขององค์กร โดยเป็นเครื่องมือที่มีประสิทธิภาพในการลดความเสี่ยงภัยคุกคามทางไซเบอร์ที่เกิดจากความผิดพลาดของมนุษย์ แต่ควรมีการใช้ร่วมกับมาตรการอื่น ๆ เช่น การมีนโยบายและระเบียบปฏิบัติที่เข้มแข็ง [8] เพื่อส่งเสริมวัฒนธรรมความปลอดภัย

2.7 ความพร้อมทางไซเบอร์

ความพร้อมทางไซเบอร์ หมายถึง ความสามารถขององค์กรในการระบุ ป้องกัน ตรวจสอบ และกู้คืนจากการโจมตีทางไซเบอร์ ซึ่งการมีกรอบการทำงานจะช่วยให้องค์กรสามารถประเมินความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์และระบุด้านที่ต้องปรับปรุงได้อย่างถูกต้อง โดยจะช่วยองค์กรลดโอกาสและผลกระทบของเหตุการณ์ทางไซเบอร์ ซึ่งบุคลากร กระบวนการ และเทคโนโลยีถือเป็นประเด็นที่ต้องให้ความสำคัญ [9]

2.8 งานวิจัยที่เกี่ยวข้อง

ศุภาพร พรหมโส และคณะ [1] ศึกษาสถานภาพความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์ของมหาวิทยาลัยราชภัฏ พบว่า มหาวิทยาลัยราชภัฏมีปัญหาเกี่ยวกับความมั่นคงปลอดภัยไซเบอร์ ได้แก่ ปัญหาเกี่ยวกับมาตรการความมั่นคงปลอดภัย, นโยบายองค์กร และบุคลากร สำหรับอุปสรรคที่ส่งผลกระทบต่อพัฒนาความมั่นคงปลอดภัยไซเบอร์ ได้แก่ นโยบายผู้บริหาร, ทักษะความรู้, วัฒนธรรมองค์กร, ความไม่เสถียรของระบบ

สุทธิพันธุ์ ขวลิตเลขา และคณะ [2] ศึกษาสภาพปัญหาและพฤติกรรมการใช้งานบนโลกไซเบอร์ของบุคลากรในบริษัทวิทยุการบินแห่งประเทศไทย พบว่า บุคลากรในบริษัทวิทยุการบินฯ มีการกำหนดกลยุทธ์ด้านภัยคุกคามทางไซเบอร์, ผู้บังคับบัญชาให้ความสำคัญกับนโยบายและมีการระบุภัยคุกคามทางไซเบอร์ที่ส่งผลกระทบต่อ การปฏิบัติงานโดยตรงในภาพรวมความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์อยู่ในระดับ มาก นอกจากนี้ ผู้วิจัยได้พัฒนาแนวทางการเสริมสร้างความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ด้วยวิธีการเรียนรู้แบบออนไลน์และใช้วีดิทัศน์ผสมผสาน

3. วิธีดำเนินงานวิจัย

งานวิจัยนี้ใช้รูปแบบการวิจัยแบบผสมผสาน (Mixed Method) [11] ที่มีทั้งการวิจัยเชิงปริมาณและเชิงคุณภาพ การวิจัยเชิงปริมาณ ผู้วิจัยใช้แบบสอบถามปลายเปิด ในส่วนของการวิจัยเชิงคุณภาพ จะใช้การสัมภาษณ์เชิงลึก (In-Depth Interview) กับผู้บริหารเพื่อศึกษาสภาพปัญหา

และภัยคุกคามทางไซเบอร์ที่มีต่อหน่วยงานของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย โดยมีกรอบแนวคิดในการวิจัยแสดง ดังภาพที่ 2

INPUT	PROCESS	OUTPUT
ศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องต่าง ๆ	รวบรวมและวิเคราะห์ ข้อมูลจากแบบสอบถาม	ความตระหนักรู้ ความพร้อมและตัวชี้วัดด้านความมั่นคงปลอดภัยไซเบอร์ในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย
	รวบรวมและวิเคราะห์ ข้อมูลการสัมภาษณ์เชิงลึกจากผู้ที่เกี่ยวข้อง	

ภาพที่ 2: กรอบแนวคิดในการวิจัย

3.1 ประชากรและกลุ่มตัวอย่าง

กลุ่มประชากรที่ทำกรวิจัย ได้แก่ ข้าราชการทหาร ข้าราชการ และลูกจ้าง สังกัดศูนย์รักษาความปลอดภัย กองบัญชาการ กองทัพไทย จำนวน 2,000 คน โดยมีกลุ่มตัวอย่าง 350 คน กำหนดโดยใช้สูตรของเครซีและมอร์แกน (Krejcie and Morgen 1970) ที่มีระดับความเชื่อมั่น 95% และค่าความผิดพลาด 5% การเลือกกลุ่มตัวอย่างโดยใช้การสุ่มตัวอย่างแบบไม่คำนึงถึงความน่าจะเป็นในการสุ่ม (Non-Probability Sampling) นอกจากนี้ยังได้ การสัมภาษณ์เชิงลึกกับผู้บริหารและผู้ดูแลระบบเครือข่ายที่มีความเชี่ยวชาญด้านความมั่นคงปลอดภัยไซเบอร์ในหน่วยงานจำนวน 9 คน

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 ผู้วิจัยได้ศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับสภาพปัญหาทางไซเบอร์และวิเคราะห์ภัยคุกคามทางไซเบอร์ที่มีต่อหน่วยงานของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย จากนั้นได้ทำการสร้างแบบสอบถามหาคำตอบของสภาพปัญหาทางไซเบอร์และวิเคราะห์ภัยคุกคามทางไซเบอร์ที่มีต่อหน่วยงาน

3.3 การวิเคราะห์ข้อมูล

3.3.1 ข้อมูลจากแบบสัมภาษณ์เชิงลึก จะใช้เทคนิคการวิเคราะห์เนื้อหา ประกอบบริบทและประมวลผลเข้าด้วยกัน

รวมทั้งใช้แนวคิดจากเอกสารและผลงานวิจัยที่เกี่ยวข้องมาสนับสนุนงานวิจัย

3.3.2 ข้อมูลที่เป็นแบบสอบถาม [12] ที่เก็บรวบรวมข้อมูลได้จะถูกนำไปวิเคราะห์ด้วยโปรแกรมคอมพิวเตอร์ทางสถิติเพื่อให้ได้สถิติเชิงพรรณนา (Descriptive Analysis) โดยทำการการแจกแจงค่าเฉลี่ย (Average) และค่าร้อยละ (Percentage)

4. ผลการดำเนินการ

4.1 จากการวิเคราะห์ข้อมูลที่ได้จากการสัมภาษณ์เชิงลึกของกลุ่มตัวอย่าง โดยใช้วิธีการวิเคราะห์แก่นสาระ (Thematic Analysis) สามารถสรุปผลการวิจัยได้ดังนี้

4.1.1 ผลวิเคราะห์แก่นสาระ

จากการศึกษาพบว่า ปัจจัยที่ส่งผลต่อปัญหาและภัยคุกคามทางไซเบอร์ต่อศูนย์รักษาความปลอดภัยกองบัญชาการกองทัพไทย สรุปได้ดังนี้ 1) การโจมตีด้วยมัลแวร์ เกิดจากบุคลากรเข้าถึงเว็บไซต์ที่ไม่ปลอดภัยและใช้ซอฟต์แวร์ที่ล้าสมัย 2) ระเบียบ นโยบาย และแนวปฏิบัติที่ไม่มีการพัฒนาให้ทันกับสถานการณ์ปัจจุบัน 3) ทรัพยากรจำกัด เช่น ขาดงบประมาณในการจัดหาบำรุงรักษา หรือต่อลิขสิทธิ์ 4) ผู้บริหารระดับสูงยังไม่เห็นความสำคัญ 5) ปัจจัยทางวัฒนธรรมที่ทำให้ยากต่อการแบ่งปันและทำงานร่วมกันเป็นทีม

4.1.2 ผลการศึกษาพบว่า ปัจจัยที่ส่งผลต่อระดับความพร้อมทางไซเบอร์ของกำลังพลสังกัดศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย ได้แก่ วัฒนธรรมทัศนคติ การใช้ทรัพยากรร่วมกัน และการสื่อสาร อุปสรรคในการพัฒนามูลกรด้านความมั่นคงปลอดภัยไซเบอร์ ได้แก่ ความแตกต่างด้านทักษะ ความรู้ ความสามารถในการจดจำ และความเชี่ยวชาญ ปัจจัยที่ส่งผลต่อความพร้อมของกระบวนการ ได้แก่ การมีแผนระบุและจัดการความเสี่ยง การควบคุมการเข้าถึงและการตอบสนองความเสี่ยง ในขณะที่อุปสรรค ได้แก่ การขาดแรงกดดันจากผู้บริหารระดับสูง ในด้านเทคโนโลยี และพบว่าอุปกรณ์เคลื่อนที่เป็นปัจจัยที่ความกังวลของเจ้าหน้าที่รักษาความปลอดภัย เนื่องจากมีหลากหลายรูปแบบและยากในการควบคุม

4.2 การวิเคราะห์ข้อมูลที่ได้จากแบบสอบถามของกลุ่มตัวอย่าง สามารถสรุปผลการวิจัยได้ดังนี้

4.2.1 ผลการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับปัจจัยส่วนบุคคลของกลุ่มตัวอย่าง พบว่า ผู้ตอบแบบสอบถามส่วนใหญ่เป็นเพศชาย จำนวน 243 คน คิดเป็นร้อยละ 69.40 ของกลุ่มตัวอย่างที่ตอบแบบสอบถามในครั้งนี้ พบว่า ผู้ตอบแบบสอบถามส่วนใหญ่มีอายุอยู่ในช่วง 30-40 ปี จำนวน 137 คน คิดเป็นร้อยละ 39.10 ระดับชั้นยศของผู้ตอบแบบสอบถาม พบว่า ส่วนใหญ่เป็นข้าราชการต่ำกว่าชั้นสัญญาบัตร จำนวน 191 คน คิดเป็นร้อยละ 54.60

4.2.2 ผลการวิเคราะห์เกี่ยวกับพฤติกรรมกรรมการใช้งานระบบอินเทอร์เน็ตใน พบว่า พฤติกรรมการใช้งานระบบอินเทอร์เน็ตโดยรวมอยู่ในระดับมาก มีค่าเฉลี่ยอยู่ที่ 3.72 โดยมีรายละเอียด ตามตารางที่ 1

ตารางที่ 1: ผลการวิเคราะห์พฤติกรรมกรรมการใช้งานอินเทอร์เน็ต

รายการประเมินพฤติกรรม	ค่าเฉลี่ย	ระดับพฤติกรรม	ลำดับที่
1. ด้านการใช้งานอินเทอร์เน็ต	3.89	มาก	1
2. ด้านการใช้งานผ่านแอปพลิเคชัน	3.47	มาก	3
3. ด้านการป้องกันตัวเองจากภัยคุกคามอินเทอร์เน็ต	3.80	มาก	2
ค่าเฉลี่ย	3.72	มาก	

4.2.3 ผลการวิเคราะห์เกี่ยวกับพฤติกรรมกรรมการใช้งานสื่อสังคมออนไลน์ พบว่า โดยรวมอยู่ในระดับมากที่สุด มีค่าเฉลี่ยอยู่ที่ 4.30 รายละเอียดตามตารางที่ 2

ตารางที่ 2: ผลการวิเคราะห์พฤติกรรมกรรมการใช้งานสื่อสังคมออนไลน์

รายการประเมินพฤติกรรม	ค่าเฉลี่ย	ระดับพฤติกรรม	ลำดับที่
1. ด้านการใช้งานสื่อสังคมออนไลน์	4.13	มาก	2
2. ด้านการเข้าถึงสื่อสังคมออนไลน์	4.48	มากที่สุด	1
ค่าเฉลี่ย	4.30	มากที่สุด	

4.2.4 ผลการวิเคราะห์ระดับความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของบุคลากร พบว่ามีระดับความตระหนักรู้โดยรวมอยู่ในระดับมาก มีค่าเฉลี่ยอยู่ที่ 4.18 รายละเอียดตามตารางที่ 3

ตารางที่ 3: ระดับความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของบุคลากร

รายการประเมินความตระหนักรู้	ค่าเฉลี่ย	ระดับความตระหนักรู้	ลำดับที่
1. ด้านการใช้คอมพิวเตอร์	4.49	มากที่สุด	1
2. ด้านการใช้พาสเวิร์ด	3.97	มาก	4
3. ด้านการใช้อีเมล	4.22	มากที่สุด	2
4. ด้านการเข้าเว็บไซต์	4.07	มาก	3
ค่าเฉลี่ย	4.18	มาก	

4.2.5 ผลการวิเคราะห์ระดับความพร้อมทางไซเบอร์ของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย สามารถสรุปได้ดังนี้ คือ สภาพความพร้อมทางไซเบอร์โดยรวม มีระดับความพร้อมอยู่ที่ 80.27% โดยมีลำดับดังนี้ ลำดับที่ 1 คือ ความพร้อมด้านกระบวนการ มีความพร้อมอยู่ที่ 82.00% ลำดับที่ 2 คือ ความพร้อมด้านบุคลากร มีความพร้อมอยู่ที่ 79.60% และลำดับที่ 3 ความพร้อมด้านเทคโนโลยี มีระดับความพร้อมอยู่ที่ 79.10% แสดงรายละเอียด ตามตารางที่ 4

ตารางที่ 4: ผลการวิเคราะห์ความพร้อมทางไซเบอร์ในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย

รายการประเมินความพร้อม	ค่าเฉลี่ย	ความพร้อม (%)	ลำดับที่
1. ความพร้อมด้านบุคลากร	3.98	79.60	2
2. ความพร้อมด้านกระบวนการ	4.10	82.00	1
3. ความพร้อมด้านเทคโนโลยี	3.96	79.10	3
ค่าเฉลี่ย	4.01	80.27	

4.3 ผลการสังเคราะห์หาตัวชี้วัดความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์สำหรับศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย

ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลเชิงปริมาณที่ได้มาจากคำตอบของข้อคำถามปลายปิดในประเด็นความพร้อมทางไซเบอร์ของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย โดยพิจารณาในข้อคำตอบย่อยในประเด็นบุคลากร กระบวนการ และเทคโนโลยี ทั้งนี้เพื่อสกัดและสังเคราะห์หาตัวชี้วัดความพร้อมทางไซเบอร์ ซึ่งผู้วิจัยได้กำหนดเกณฑ์การคัดเลือกตัวชี้วัดของข้อคำถามที่มีค่าคะแนนเฉลี่ย (Mean) สูงกว่า 3.50 ขึ้นไปอย่างมีนัยสำคัญทางสถิติมาพิจารณา หลังจากนั้นจึงนำตัวชี้วัดที่ได้ไปให้ผู้เชี่ยวชาญจำนวน 9 ท่านยืนยันให้ความเห็นชอบอีกครั้ง ซึ่งพบว่าผลลัพธ์ที่ได้ คือ ตัวชี้วัดจำนวน 7 ตัวชี้วัด ได้แก่ 1. ความสามารถด้านความมั่นคงปลอดภัยไซเบอร์ของกำลังพล 2. ความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของกำลังพล 3. ภาวะผู้นำด้านไซเบอร์ของหน่วยงาน 4. นโยบายความมั่นคงปลอดภัยไซเบอร์ของหน่วยงาน 5. การบริการจัดการการรักษาความมั่นคงปลอดภัยไซเบอร์ ในหน่วยงาน 6. โครงสร้างพื้นฐานสำคัญทางด้านไซเบอร์ของหน่วยงาน และ 7. การวิจัยและพัฒนานวัตกรรมด้านไซเบอร์ของหน่วยงาน

5. สรุปผลการวิจัย

ผลการศึกษาพฤติกรรม ความตระหนักรู้ ความพร้อม และตัวชี้วัดด้านความมั่นคงปลอดภัยไซเบอร์ในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย สรุป ดังนี้

กำลังพลในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย มีพฤติกรรมการและความตระหนักรู้ความปลอดภัยทางไซเบอร์ โดยรวมอยู่ในระดับมาก และค้นพบปัจจัยที่ส่งผลต่อสภาพปัญหา คือ บุคลากรบางส่วนยังขาดทักษะและความรู้เกี่ยวกับการใช้งานในโลกไซเบอร์ เนื่องจากปัจจัยทางวัฒนธรรมองค์กรและมีทรัพยากรที่จำกัดในการบำรุงรักษาเครื่องมือหรืออุปกรณ์การรักษาความปลอดภัยทำให้เกิดเป็นสาเหตุสำคัญของการถูกโจมตีทางไซเบอร์

ในงานวิจัยนี้ พบว่าระดับความพร้อมทางไซเบอร์ของบุคลากร กระบวนการและเทคโนโลยีของหน่วยงาน โดยมีความพร้อมทางไซเบอร์โดยรวมอยู่ที่ 80.27% ปัจจัยสำคัญที่ส่งผลต่อความพร้อมทางไซเบอร์มีดังนี้ 1) การฝึกอบรมอย่างต่อเนื่อง 2) วัฒนธรรมและทัศนคติ 3) ทรัพยากรและเครื่องมือ 4) การสื่อสารร่วมกัน 5) การมีแผนในการระบุและบริหารความเสี่ยง 6) การควบคุมการเข้าถึง และ 7) การตอบสนองต่อความเสี่ยง และพบว่าปัญหาอุปสรรคที่ส่งผลต่อความพร้อมทางไซเบอร์ คือ 1) ความรู้ ความสามารถในการรับรู้และความชำนาญของกำลังพลมีความแตกต่างกัน 2) ผู้บริหารระดับสูงยังไม่เห็นถึงความสำคัญ และ 3) อุปสรรคพวก เนื่องจากขาดการควบคุมความปลอดภัยและมีประเภทที่หลากหลาย

นอกจากนั้น ผู้วิจัยยังได้สังเคราะห์ตัวชี้วัดความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์ของศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทย ได้จำนวน 7 ตัวชี้วัด ได้แก่ 1. ความสามารถด้านความมั่นคงปลอดภัยไซเบอร์ของกำลังพล 2. ความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ของกำลังพล 3. ภาวะผู้นำด้านไซเบอร์ของหน่วยงาน 4. นโยบายความมั่นคงปลอดภัยไซเบอร์ของหน่วยงาน 5. การบริการจัดการการรักษาความมั่นคงปลอดภัยไซเบอร์ในหน่วยงาน 6. โครงสร้างพื้นฐานสำคัญทางด้านไซเบอร์ของหน่วยงาน และ 7. การวิจัยและพัฒนานวัตกรรมด้านไซเบอร์ของหน่วยงาน

6. ข้อเสนอแนะ

ผู้วิจัยเสนอแนวทางการพัฒนาความสามารถในการเสริมสร้างความตระหนักรู้ทางไซเบอร์สำหรับบุคลากรในศูนย์รักษาความปลอดภัย กองบัญชาการกองทัพไทยให้มีประสิทธิภาพมากยิ่งขึ้น โดยควรทำการพัฒนาตัวแบบวุฒิภาวะความสามารถในการสร้างความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์สำหรับผู้ปฏิบัติงาน และ ถือเป็นแนวทางการพัฒนาศักยภาพบุคลากรด้านไซเบอร์ในกองทัพไทย ให้มีขีดความสามารถด้านไซเบอร์ที่ได้มาตรฐานและเป็นที่ยอมรับ ในอันที่จะสามารถปฏิบัติงานได้อย่างมีประสิทธิภาพต่อไป

เอกสารอ้างอิง

- [1] สุภาพร พรหมโส และคณะ, “สถานภาพความพร้อมและดัชนีความพร้อมด้านความมั่นคงปลอดภัยไซเบอร์ของมหาวิทยาลัยราชภัฏ,” *วารสารวิทยาศาสตร์และเทคโนโลยี นายเรืออากาศ* ปีที่ 17 ฉบับที่ 2 กรกฎาคม-ธันวาคม พ.ศ. 2564 หน้า 17-30.
- [2] สุทธิพันธุ์ ชาลิตเลขา และคณะ, “การเสริมสร้างความตระหนักรู้ด้านความมั่นคงปลอดภัยไซเบอร์ สำหรับบุคลากรในบริษัทวิทยุการบินแห่งประเทศไทย จำกัด,” *การประชุมวิชาการระดับชาติและนานาชาติ ครั้งที่ 11 พ.ศ. 2565* หน้า 602-612.
- [3] ศิวลีย์ สิริโรจน์บริรักษ์, “การพัฒนามาตรฐานการรักษาความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity) ของกระทรวงกลาโหม,” *วารสารสถาบันวิชาการป้องกันประเทศ* ปีที่ 6 ฉบับที่ 3 พฤษภาคม-สิงหาคม 2558 หน้า 21.
- [4] Y. Li and Q. Liu, “A comprehensive review study of cyber-attacks and cyber security: Emerging trends and recent developments,” *Energy Reports*, vol. 7, pp. 8176-8186, 2021.
- [5] P. Watters, et al., *Cognition, Behavior and Cybersecurity*, Frontiers Media SA, 2021.
- [6] M. Sailio, et al., “Cyber Threat Actors for the Factory of the Future,” *Applied Sciences*, vol. 10, no. 12, p. 4334, 2020.
- [7] Goog CV, Merkel WR. *Dictionary of education*, McGraw-Hill; 1973.
- [8] I. Corradini, *Building a Cybersecurity Culture in Organizations*. Springer Nature, 2020.
- [9] C. Boletsis, et al., “Cybersecurity for SMEs: Introducing the Human Element into Socio-technical Cybersecurity Risk Assessment,” *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 3, p.266, 2021.
- [10] P. R. Trim and Y.-I. Lee, “The Global Cyber Security Model: Counteracting Cyber Attacks through a Resilient Partnership Arrangement,” *Big Data and Cognitive Computing*, vol. 5, no. 3, p. 32, 2021.
- [11] J. W. A. Witsenboer, et al., “Measuring cyber secure behavior of elementary and high school students in the Netherlands,” *Computers & Education*, vol. 186, p. 104536, 2022.

การพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเต้นผิดจังหวะ

Development of an ECG Machine to Diagnose Arrhythmias

สุพีรวงศ์ โพธิกนิษฐ (Supeerawong Pothikanit)¹ และมหศักดิ์ เกตุฉ่ำ (Mahasak Ketcham)²

ภาควิชาการจัดการเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹s6407021856054@email.kmutnb.ac.th, ²mahasak.k@itd.kmutnb.ac.th

บทคัดย่อ

โรคหัวใจและหลอดเลือด (Cardiovascular disease; CVD) เป็นสาเหตุอันดับหนึ่งของการเสียชีวิตของประชากรโลก โดยมีผู้เสียชีวิตจากกลุ่มโรคนี้ประมาณ 17.9 ล้านคน และจากสถิติข้อมูลการเสียชีวิตของคนไทยในกลุ่มโรคหัวใจและหลอดเลือด พบว่า ร้อยละ 80 เสียชีวิตด้วยกล้ามเนื้อหัวใจขาดเลือดเฉียบพลัน ค่าใช้จ่ายในการรักษาพยาบาลเฉลี่ยของผู้ป่วยโรคหัวใจสูงถึง 6,906 ล้านบาทต่อปี ส่งผลกระทบต่อคุณภาพชีวิตของประชากรทั้งในระดับบุคคล ครอบครัว สังคม และประเทศชาติ แสดงให้เห็นว่า โรคหัวใจและหลอดเลือดเป็นโรคที่มีความรุนแรงและจำเป็นต้องได้รับการดูแลอย่างเร่งด่วน การพัฒนาองค์ความรู้และเทคโนโลยีสำหรับวินิจฉัยโรคและกำกับติดตามอาการของผู้ป่วยจึงมีความสำคัญเป็นอย่างมาก การวิจัยครั้งนี้จึงมีวัตถุประสงค์เพื่อพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเต้นผิดจังหวะ โดยใช้ข้อมูลคลื่นไฟฟ้าหัวใจจากฐานข้อมูล MIT-BIH Arrhythmia โดยใช้แบบจำลองที่สร้างด้วยอัลกอริทึมโครงข่ายประสาทเทียม อัลกอริทึมโครงข่ายประสาทเทียมคอนโวลูชัน อัลกอริทึมโครงข่ายประสาทเทียมแบบวนกลับ และอัลกอริทึม Long Short-Term Memory ผลการวิจัยพบว่าแบบจำลองที่สร้างด้วยอัลกอริทึมโครงข่ายประสาทเทียมคอนโวลูชันมีประสิทธิภาพในการวิเคราะห์สูงที่สุด

คำสำคัญ: โรคหัวใจเต้นผิดจังหวะ โครงข่ายประสาทเทียมโครงข่ายประสาทเทียมคอนโวลูชัน โครงข่ายประสาทเทียมแบบวนกลับ Long Short-Term Memory

Abstract

Cardiovascular disease (CVD) is the leading cause of death in the world population, with approximately 17.9 million deaths attributed to this group of diseases. Statistics on mortality rates for Thai people in the group of heart and vascular diseases reveal that 80% of deaths are caused by sudden myocardial infarction. The average healthcare expenditure for heart disease patients is up to 6,906 million baht per year, impacting the quality of life for individuals, families, communities, and nations. These facts demonstrate that cardiovascular disease is a severe and urgent health issue that requires close monitoring and rapid care. Therefore, the development of knowledge and technology for diagnosis and patient monitoring is of paramount importance. The purpose of this research is to develop an electrocardiogram (ECG) analysis device using data from the MIT-BIH Arrhythmia database and a model constructed with neural network algorithms, convolutional neural network algorithms, recurrent neural network algorithms, and long short-term memory algorithm. The research found that the model constructed with the convolutional neural network algorithm was the most effective for analyzing ECG signals.

Keywords: Arrhythmias, Neural Network, Convolutional Neural Network Algorithms, Recurrent Neural Network, Long Short-Term Memory.

1. บทนำ

โรคหัวใจและหลอดเลือด (Cardiovascular disease; CVD) เป็นสาเหตุอันดับหนึ่งของการเสียชีวิตของประชากรโลก [1] โดยมีผู้เสียชีวิตจากกลุ่มโรคนี้นับประมาณ 17.9 ล้านคน และจากสถิติข้อมูลการเสียชีวิตของคนไทยในกลุ่มโรคหัวใจและหลอดเลือด พบว่า ร้อยละ 80 เสียชีวิตด้วยกล้ามเนื้อหัวใจขาดเลือดเฉียบพลัน [2]

จากข้อมูลทะเบียนการเสียชีวิตของสำนักบริหารทะเบียน กระทรวงมหาดไทย พบว่า อัตราการตายในช่วงอายุ 30-69 ปี จากโรคไม่ติดต่อ 4 โรคสำคัญคือ โรคหลอดเลือดสมอง โรคหัวใจขาดเลือดโรคเบาหวาน และโรคทางเดินหายใจอุดกั้นเรื้อรัง และมีแนวโน้มสูงเพิ่มขึ้น โดยอัตราการตายจากโรคหลอดเลือดสมองมีการเพิ่มขึ้นมากที่สุดเป็นอันดับที่ 1 จาก 33.4 ต่อแสนประชากร เพิ่มขึ้นเป็น 40.9 ต่อแสนประชากร อันดับที่ 2 คือ โรคหัวใจขาดเลือด จาก 22.4 ต่อแสนประชากร เป็น 27.8 ต่อแสนประชากร อันดับที่ 3 คือ โรคเบาหวาน จาก 13.2 ต่อแสนประชากร เพิ่มขึ้นเป็น 17.8 ต่อแสนประชากร และอันดับสุดท้าย คือ โรคทางเดินหายใจอุดกั้นเรื้อรังจาก 3.8 ต่อแสนประชากร เป็น 4.5 ต่อแสนประชากร นอกจากนี้ ยังพบว่าอัตราผู้ป่วยในของโรคหัวใจขาดเลือดมีแนวโน้มเพิ่มสูงขึ้นอย่างต่อเนื่อง ค่าใช้จ่ายในการรักษาพยาบาลเฉลี่ยของผู้ป่วยโรคหัวใจสูงถึง 6,906 ล้านบาทต่อปี ส่งผลกระทบต่อคุณภาพชีวิตของประชากร ทั้งในระดับบุคคล ครอบครัว สังคม และประเทศชาติ แสดงให้เห็นว่า โรคหัวใจและหลอดเลือดเป็นโรคที่มีความรุนแรงและจำเป็นต้องได้รับการดูแลอย่างเร่งด่วน การพัฒนาองค์ความรู้และเทคโนโลยีสำหรับวินิจฉัยโรค และกำกับติดตามอาการของผู้ป่วยจึงมีความสำคัญเป็นอย่างมาก

การวิเคราะห์คลื่นไฟฟ้าหัวใจโดยใช้คอมพิวเตอร์ประมวลผลได้รับความสนใจมานานหลาย 10 ปี และมีแนวโน้มเพิ่มสูงขึ้นอย่างต่อเนื่อง [3] งานวิจัยส่วนใหญ่จะมุ่งเน้นการวิเคราะห์รูปแบบและการจำแนกประเภทของคลื่นไฟฟ้าหัวใจ รวมถึงเพิ่มประสิทธิภาพและความถูกต้องในการจำแนกสัญญาณคลื่นไฟฟ้าหัวใจโดยการนำเทคนิคใหม่ ๆ มาประยุกต์ใช้ ซึ่งการตรวจจับความผิดปกติ

ของคลื่นไฟฟ้าหัวใจที่รวดเร็วและการวิเคราะห์ที่ถูกต้องมีความสำคัญมากต่อชีวิตผู้ป่วย [4] ผลการวิเคราะห์ดังกล่าวสามารถช่วยสนับสนุนแพทย์ในการตัดสินใจเลือกแนวทางการรักษาผู้ป่วยได้อย่างถูกต้องและปลอดภัยมากที่สุด

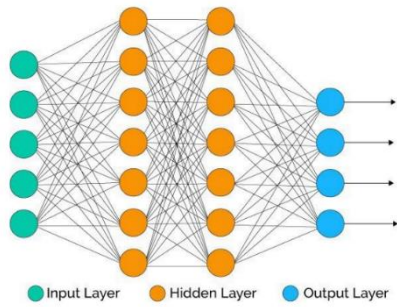
งานวิจัยในปัจจุบัน มีผู้วิจัยให้ความสนใจเกี่ยวกับการจำแนกประเภทผู้ป่วยโรคหัวใจขาดเลือดกันอย่างแพร่หลาย แต่ยังไม่พบงานวิจัยที่พัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเด่นชัดจังหวะที่มีราคาถูก และสามารถใช้งานได้ง่าย โดยผู้ใช้ไม่จำเป็นต้องมีความรู้ทางด้านทางการแพทย์ก็สามารถใช้งานเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเด่นชัดจังหวะได้ ผู้วิจัยจึงสนใจที่จะพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเด่นชัดจังหวะโดยปรับ Baseline และกำจัดสัญญาณรบกวนด้วยเทคนิค Wavelet Transform จากนั้นระบุตำแหน่งของ R-Peak เพื่อสกัดคุณลักษณะที่สำคัญและจำแนกประเภทของผู้ป่วยโรคหัวใจและหลอดเลือดด้วยอัลกอริทึมต่าง ๆ จากนั้นประเมินประสิทธิภาพของเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเด่นชัดจังหวะด้วยค่าความถูกต้อง ค่าความเร็วในการประมวลผล ซึ่งจะทำได้สามารถนำเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเด่นชัดจังหวะไปใช้งานได้ อย่างมีประสิทธิภาพ

2. วรรณกรรมที่เกี่ยวข้อง

2.1 โครงข่ายประสาทเทียม (Artificial Neural

Network: ANN) [5]

วิธีการสร้างแบบจำลองข้อมูลหรือโมเดลข้อมูล (Data modelling) โดยวิธีโครงข่ายประสาทเทียมเป็นวิธีการที่มีพื้นฐานมาจากการเลียนแบบการทำงานของสมองมนุษย์ ซึ่งประกอบด้วยนิวรอน (Neurons) ที่ใช้ในการเชื่อมโยงเพื่อแก้ปัญหาต่าง ๆ โดยมีโครงสร้างการทำงานตามภาพที่ 1 ซึ่งประกอบด้วยข้อมูลเข้า (Input) 1 ชั้น และข้อมูลออก (Output) 1 ค่าจะเรียกว่าเพอร์เซพตรอน (Perceptron)



ภาพที่ 1: ส่วนประกอบของโครงข่ายประสาทเทียม

จากภาพที่ 1 จะเห็นว่าส่วนประกอบของโครงข่ายประสาทเทียมประกอบไปด้วย 3 ส่วนดังนี้

- 1) ชั้นข้อมูลเข้า (Input Layer) คือ ชั้นที่เป็นข้อมูลเข้า จำนวนของโหนดในชั้นนี้ขึ้นอยู่กับจำนวนของข้อมูลเข้า ว่ามีข้อมูลอะไรบ้างที่นำเข้ามาคิดในแบบจำลอง ซึ่งอาจจะเรียกปัจจัยที่นำมาวิเคราะห์เหล่านี้ว่าคุณลักษณะ (Feature)
- 2) ชั้นซ่อน (Hidden Layer) คือ ชั้นที่อยู่ระหว่างกลาง ซึ่งจะมีผลอย่างมากต่อประสิทธิภาพในการเรียนรู้ของแบบจำลอง ซึ่งในชั้นนี้จะมีกี่ชั้นก็ได้ และแต่ละชั้นจะมีจำนวนของนิวรอนจำนวนเท่าไรก็ได้เช่นกัน ซึ่งการเพิ่มชั้นและจำนวนนิวรอนจะส่งผลกระทบต่อการทำงานของโมเดล ในส่วนของชั้นซ่อน มีการทำงานเปรียบเสมือนส่วนที่เรียนรู้เชิงลึก (Deep Learning) โดยทุก ๆ โหนดต้องประกอบด้วยฟังก์ชันแบบไม่เป็นเชิงเส้น
- 3) ชั้นข้อมูลออก (Output Layer) คือชั้นที่จะนำเอาข้อมูลจากการคำนวณ ไปใช้ และจำนวนของโหนดในชั้นนี้ขึ้นอยู่กับรูปแบบของข้อมูลที่จะนำไปใช้

2.2 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN)

โครงข่ายประสาทเทียมคอนโวลูชัน [6] เป็นอัลกอริทึมที่ถูกนำมาใช้อย่างกว้างขวางทั้งในทางคณิตศาสตร์ สถิติ การประมวลผลสัญญาณ (Signal processing) รวมถึงการประมวลผลภาพ (computer vision)

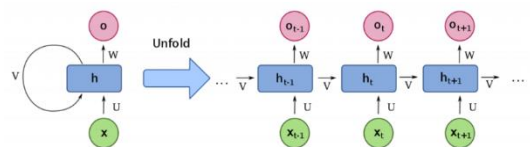
โครงข่ายประสาทเทียมคอนโวลูชันถูกนำมาใช้ใน รูปแบบโมเดลที่สามารถเรียนรู้และเลือกใช้ลักษณะเด่น (Feature extraction) ของรูปภาพได้ด้วยตัวเอง ข้อดีคือ โมเดลจะเรียนรู้และเลือกลักษณะเด่นของภาพได้ดีกว่ามนุษย์ ดังนั้นจะทำให้ได้การประมวลผลที่แม่นยำมากกว่า

ข้อแตกต่างจากโครงข่ายประสาทเทียม คือ คำนวณที่ใช้เป็นค่าชุดเดียวกันหมด

2.3 โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN)

โครงข่ายประสาทเทียมแบบวนกลับ [7] เป็น อัลกอริทึมที่ถูกนำมาใช้ในการวิจัยเกี่ยวกับการรู้จำเสียง (Speech recognition) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing) เป็นอัลกอริทึมที่มีการเอาผลลัพธ์ที่ได้จากการคำนวณย้อนกลับมาใช้เป็นข้อมูล ขาเข้าอีกครั้ง ซึ่งมีประโยชน์อย่างมากในข้อมูลที่มีความต่อเนื่อง เช่น ข้อมูลเสียง ข้อความ หรือรูปภาพ

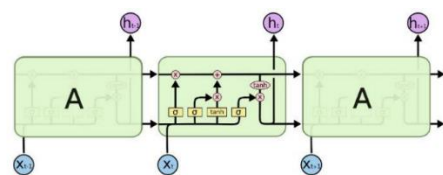
โครงข่ายประสาทเทียมแบบวนกลับถูกออกแบบมา เพื่อแก้ไขปัญหาสำหรับงานที่ข้อมูลมีลำดับ โดยใช้หลักการนำสถานะภายในของโมเดลกลับมาเป็นข้อมูลเข้าใหม่คู่กับข้อมูลเข้าแบบปกติ เรียกว่า สถานะซ่อน (Hidden State) หรือสถานะภายใน (Internal State) ช่วยให้โมเดล รู้จำรูปแบบ (Pattern) ของลำดับข้อมูลเข้า (Input Sequence) โดยแสดงดังภาพที่ 2



ภาพที่ 2: โครงสร้างของโครงข่ายประสาทเทียมแบบวนกลับ

2.4 Long Short-Term Memory (LSTM)

คืออัลกอริทึมที่ถูกพัฒนามาจาก Recurrent Neural Network (RNN) โดยพัฒนาให้มีความเสถียรและมี ประสิทธิภาพมากขึ้น โดยมีหลักการทำงานคือ สามารถ เก็บสถานะ หรือข้อมูลของแต่ละโหนดเอาไว้เพื่อที่เวลา ย้อนกลับ ไปจะได้ทราบถึงที่มาของข้อมูลค่าดังกล่าวว่า เดิมเป็นค่าอะไร โดยมีโครงสร้างการทำงานตามภาพที่ 3



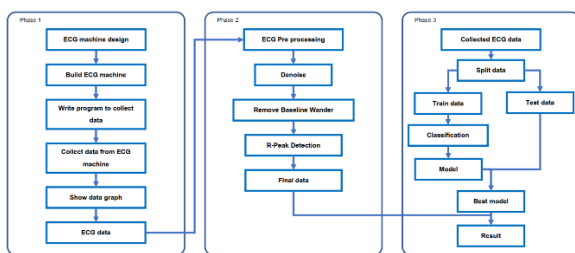
ภาพที่ 3: โครงสร้างของ Long Short-Term Memory

3. วิธีการดำเนินการวิจัย

3.1 การพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อ

วิเคราะห์โรคหัวใจเต้นผิดจังหวะ

การพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเต้นผิดจังหวะแบ่งออกเป็น 3 ส่วน ได้แก่ ส่วนออกแบบและพัฒนาระบบเครื่องตรวจคลื่นไฟฟ้าหัวใจ ส่วนจัดเตรียมข้อมูลเพื่อใช้ในการวิเคราะห์โรคหัวใจเต้นผิดจังหวะ และส่วนวิเคราะห์และทำนายโรคหัวใจเต้นผิดจังหวะ โดยมีกรอบแนวคิดแสดงดังภาพที่ 4



ภาพที่ 4: กรอบแนวคิดการวิจัยวิธีการพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเต้นผิดจังหวะ

3.2 การวิเคราะห์คลื่นไฟฟ้าหัวใจเพื่อทำนายโรคหัวใจเต้นผิดจังหวะ

3.2.1 กระบวนการเตรียมข้อมูลคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์ (Preprocessing)

ข้อมูลคลื่นไฟฟ้าหัวใจที่ใช้ในการวิเคราะห์เป็นข้อมูลที่นำมาจาก MIT-BIH Arrhythmia Database โดยข้อมูลคลื่นไฟฟ้าหัวใจที่จะนำไปวิเคราะห์จะต้องมีการเตรียมข้อมูลโดยการกำจัดสัญญาณรบกวนด้วยวิธี Wavelet Transform และปรับ Baseline เพื่อให้ข้อมูลคลื่นไฟฟ้าหัวใจมีความชัดเจนและมีคุณภาพมากขึ้น โดยมีขั้นตอนดังนี้

3.2.1.1 การกำจัดสัญญาณรบกวน (De-noising)

สัญญาณรบกวน (Noise) ส่งผลต่อค่าความถูกต้องแม่นยำในการวิเคราะห์คลื่นไฟฟ้าหัวใจ ในขั้นตอนนี้จึงต้องมีการนำข้อมูลคลื่นไฟฟ้าหัวใจมากำจัดสัญญาณรบกวนด้วยการ Wavelet Transform ด้วยวิธี Biorthogonal wavelet [8] คำนวณโดยใช้สมการที่ (1)

$$\int_{-\infty}^{\infty} u_i(t) \tilde{u}_j(t) dt = \delta(i-j) \quad (1)$$

3.2.1.2 การปรับ Baseline (Remove Baseline Wander)

Baseline ของข้อมูลคลื่นไฟฟ้าหัวใจยังมีลักษณะที่ไม่เป็นเส้นตรง ส่งผลถึงความถูกต้องในการวิเคราะห์คลื่นไฟฟ้าหัวใจในขั้นตอนของการหา R peak ได้ จะใช้วิธีการปรับ Baseline โดยใช้เทคนิค median filtering [9] คำนวณโดยใช้สมการที่ (2)

$$R_{xy} = \frac{\sum_{i=1}^N X_i Y_i}{\sqrt{\sum_{i=1}^N X_i^2 \sum_{i=1}^N Y_i^2}} \quad (2)$$

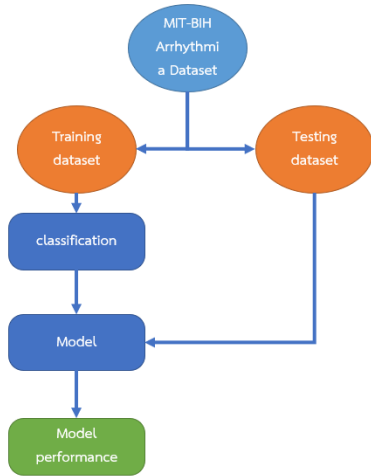
3.2.2 การระบุตำแหน่ง R peak (R peak detection)

ข้อมูลที่ใช้ในการวิเคราะห์โรคหัวใจเต้นผิดจังหวะจะต้องทำการระบุตำแหน่ง R peak เพื่อที่จะสามารถแบ่งข้อมูลคลื่นไฟฟ้าหัวใจมาใช้ในการทำนายโรคหัวใจเต้นผิดจังหวะได้ โดยอัลกอริทึม Christov [10] ในการระบุตำแหน่ง R peak ซึ่งเป็นการระบุตำแหน่งของ R peak โดยใช้สมการที่ (3)

$$F = F + \frac{\max(Y_{\text{lastes } 50 \text{ ms in } 350 \text{ ms interv}}) - \max(Y_{\text{earliest } 50 \text{ ms in } 350 \text{ ms interv}})}{150} \quad (3)$$

3.2.3 การจำแนกโรคหัวใจเต้นผิดจังหวะ (Arrhythmia Classification)

นำชุดข้อมูล MIT-BIH Arrhythmia Dataset แบ่งออกเป็น 2 ชุดข้อมูล โดยแบ่งเป็นชุดข้อมูลสำหรับฝึกสอน (training dataset) จำนวน 87554 record และชุดข้อมูลสำหรับทดสอบ (testing dataset) จำนวน 21892 record ซึ่งชุดข้อมูลนี้จะเป็นข้อมูลที่มีผลในการจำแนกผู้ป่วย 5 ประเภท ดังนี้ 1. จังหวะปกติ 2. Supraventricular Ectopic Beat 3. Ventricular ectopic beats 4. Fusion Beats 5. ไม่สามารถระบุได้ จากนั้นนำชุดข้อมูลสำหรับฝึกสอนมาทำการจำแนก (Classification) ด้วยอัลกอริทึมโครงข่ายประสาทเทียม อัลกอริทึมโครงข่ายประสาทเทียมคอนโวลูชัน อัลกอริทึมโครงข่ายประสาทเทียมแบบวนกลับ และอัลกอริทึม Long Short-Term Memory จะได้ผลลัพธ์เป็นแบบจำลอง (Model) ของแต่ละอัลกอริทึม จากนั้นนำชุดข้อมูลสำหรับทดสอบมาทำการทำนายโดยใช้แบบจำลองทั้ง 2 แบบ และทำการเปรียบเทียบประสิทธิภาพในการทำนายของแต่ละแบบจำลอง โดยลำดับการจำแนกโรคหัวใจเต้นผิดจังหวะแสดงดังภาพที่ 5



ภาพที่ 5: ลำดับการจำแนกโรคหัวใจเต้นผิดจังหวะ

3.2.4 การประเมินผลการวิเคราะห์

ในการทดสอบอัลกอริทึมที่ใช้ในการจำแนกนั้น จะมีผลลัพธ์ที่เป็นไปได้อยู่ 4 กรณี คือ ผลบวกจริง (True Positive) ผลบวกปลอม (False Positive) ผลลบจริง (True Negative) และผลลบปลอม (False Negative) เมื่อได้ผลลัพธ์ทั้ง 4 แล้ว นำข้อมูลที่ได้มาประเมินผลการวิเคราะห์ของอัลกอริทึมด้วยค่าทางสถิติ 3 ค่า คือ ค่าความถูกต้อง (Accuracy) ค่าความไว (Sensitivity) และค่าความจำเพาะ (Specificity) รายละเอียดของผลลัพธ์แสดงตามตารางที่ 1

ตารางที่ 1: ความสัมพันธ์ระหว่างค่าทำนายผลบวก ค่าทำนายผลลบ ค่าความไว และค่าความจำเพาะ

	พบโรค	ไม่พบโรค	
ผลการวิเคราะห์เป็นบวก	ผลบวกจริง	ผลบวกปลอม	ค่าทำนายผลบวก = ผลรวมผลบวกจริง / ผลทั้งหมดที่เป็นบวก
ผลการวิเคราะห์เป็นลบ	ผลลบปลอม	ผลลบจริง	ค่าทำนายผลลบ = ผลรวมผลลบจริง / ผลทั้งหมดที่เป็นลบ
	ค่าความไว = ผลรวม ผลบวกจริง / ผู้ป่วย ทั้งหมดที่พบโรค	ค่าความจำเพาะ = ผลรวมผลลบจริง / ผู้ป่วยทั้งหมดที่ไม่พบ โรค	

4. ผลการดำเนินการวิจัย

ผลการวิเคราะห์คลื่นไฟฟ้าหัวใจเพื่อสร้างแบบจำลองในการทำนายโรคหัวใจเต้นผิดจังหวะ

ผลการวิจัยที่นำเสนอนี้เป็นผลที่ได้จากการวิเคราะห์คลื่นไฟฟ้าหัวใจจากฐานข้อมูล MIT-BIH Arrhythmia ผ่านกระบวนการเตรียมข้อมูลเพื่อวิเคราะห์ (Pre-processing) โดยข้อมูลที่ได้จากฐานข้อมูล MIT-BIH Arrhythmia แบ่งออกเป็นข้อมูลชุดฝึกสอน (Training dataset) จำนวน 87,554 records และข้อมูลชุดทดสอบ (Testing dataset) จำนวน 21,892 records จากนั้นนำข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบมาทำการวิเคราะห์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (ANN) โครงข่ายประสาทเทียมคอนโวลูชัน (CNN) โครงข่ายประสาทเทียมแบบวนกลับ (RNN) และ Long Short-Term Memory (LSTM) จากนั้นนำผลการวิเคราะห์มาประเมินประสิทธิภาพในการทำนายโดยคำนวณค่าความถูกต้องในการจำแนก (Accuracy) ค่าความจำเพาะ (Specificity) และค่าความไว (Sensitivity) ผลลัพธ์ที่ได้แสดงดังตารางที่ 2

ตารางที่ 2: ค่าความถูกต้องในการจำแนก ค่าความจำเพาะ และค่าความไวของแบบจำลองที่ได้จากอัลกอริทึมต่าง ๆ

อัลกอริทึม	ค่าความถูกต้องในการจำแนก	ค่าความจำเพาะ	ค่าความไว
โครงข่ายประสาทเทียม	ร้อยละ 95.93	ร้อยละ 87.83	ร้อยละ 97.95
โครงข่ายประสาทเทียมคอนโวลูชัน	ร้อยละ 99.84	ร้อยละ 99.60	ร้อยละ 99.90
โครงข่ายประสาทเทียมแบบวนกลับ	ร้อยละ 99.52	ร้อยละ 98.78	ร้อยละ 99.70
Long Short-Term Memory	ร้อยละ 82.84	ร้อยละ 26.46	ร้อยละ 96.94

จากตารางที่ 2 พบว่าอัลกอริทึมโครงข่ายประสาทเทียมคอนโวลูชันมีประสิทธิภาพในการวิเคราะห์เพื่อทำนายโรคหัวใจเต้นผิดจังหวะสูงที่สุด ผู้วิจัยจึงเลือกใช้แบบจำลองที่ได้จากการวิเคราะห์ด้วยอัลกอริทึมโครงข่ายประสาทเทียมคอนโวลูชันมาใช้วิเคราะห์โรคหัวใจเต้นผิดจังหวะ

5. สรุป

งานวิจัยครั้งนี้ผู้วิจัยได้ทำการพัฒนาเครื่องตรวจคลื่นไฟฟ้าหัวใจเพื่อวิเคราะห์โรคหัวใจเต้นผิดจังหวะพร้อมทำการวัดประสิทธิภาพของอัลกอริทึมที่ใช้ในการทำนายผลตรวจคลื่นไฟฟ้าหัวใจ

ผลการวัดประสิทธิภาพพบว่าโครงข่ายประสาทเทียมคอนโวลูชันมีค่าความถูกต้องในการจำแนกสูงถึงร้อยละ 99.84 ซึ่งสามารถใช้เป็นผลการวิเคราะห์อ้างอิงเบื้องต้นให้ผู้ที่อยู่ในกลุ่มเสี่ยงหรือผู้ป่วยสามารถที่จะเข้ารับการรักษาได้อย่างทันท่วงที

ข้อเสนอแนะในการวิจัยครั้งต่อไป ในการวิจัยนี้มุ่งเน้นไปที่อุปกรณ์ประมวลผลขนาดเล็กอย่าง Raspberry Pi 4 Model B ซึ่งทำให้ความเร็วในประมวลผลการตรวจวัดคลื่นไฟฟ้าหัวใจทำได้ช้า แต่สามารถพกพาเพื่อใช้งานในสถานที่ต่าง ๆ ได้สะดวก รวดเร็ว กว่า หากต้องการเพิ่มความเร็วในการประมวลผลการตรวจวัดคลื่นไฟฟ้าหัวใจสามารถเปลี่ยนอุปกรณ์ประมวลผลเป็นเครื่องคอมพิวเตอร์พกพาได้แต่จะสูญเสียในเรื่องของความสะดวกในการพกพาเครื่องตรวจคลื่นไฟฟ้าหัวใจ และในกรณีที่ต้องการเพิ่มระยะเวลาในการใช้งานเครื่องตรวจคลื่นไฟฟ้าหัวใจให้นานขึ้นสามารถทำได้ โดยการเพิ่มความจุของ Power Bank โดยในการวิจัยนี้ใช้ Power Bank ขนาด 10,000 mAh ซึ่งสามารถใช้งานได้ยาวนานสุด 6 ชั่วโมง อย่างไรก็ตามงานวิจัยครั้งนี้เป็นการวิเคราะห์ข้อมูลที่ได้จากจากฐานข้อมูล MIT-BIH Arrhythmia เท่านั้น ไม่ได้มีการวัดค่าหรือวิเคราะห์ผลการทดลองจากมนุษย์ ทำให้การนำไปใช้งานจริงอาจจะต้องผ่านการวิจัย และทดลองกับมนุษย์เพื่อให้ผลที่ได้มีความถูกต้อง เทียบตรง และปลอดภัยที่สุดก่อนนำไปใช้งานทั่วไป

เอกสารอ้างอิง

- [1] World Health Organization (WHO). "Cardiovascular diseases (CVDs)" [cited 5 March 2022]. Available from <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-cvds>, 2021.
- [2] กรมควบคุมโรค. "รอบรู้เรื่องโรคและภัยสุขภาพ : สคร.6 ชลบุรี ชวนปกป้องหัวใจ ในวันหัวใจโลก" [สืบค้นวันที่ 5

มีนาคม 2565]. จาก https://ddc.moph.go.th/odpc6/news.php?news=14902&deptcode=odpc6&news_views=265,2565

- [3] Kundu, Nasipuri et al. "Knowledge-based ECG interpretation: a critical review" Department of Computer Science & Engineering Jadavpur University Calcutta India, 2000
- [4] Martis, Acharya et al. "Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals" International Journal of Neural Systems, Vol. 22, No. 06, 1250027, 2012
- [5] S Agatonovic-Kustrin, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research" Journal of Pharmaceutical and Biomedical Analysis, Volume 22, Issue 5, June 2000, Pages 717-727, 2000
- [6] Saad, Tareq et al. "Understanding of a convolutional neural network" 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017
- [7] Alex Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network" Physica D: Nonlinear Phenomena, Volume 404, 2020
- [8] Mukesh Dalal, "Evaluation of orthogonal and biorthogonal wavelets for video steganography" Information Security Journal: A Global Perspective, Volume 29, Pages 40-50, 2020
- [9] Hongshen Tang, Rongrong Ni, Yao Zhao et al., "Median filtering detection of small-size image based on CNN", Journal of Visual Communication and Image Representation 51, pg. 162, 2018
- [10] Christov, I., Neycheva, T., Schmid, R. et al. "Pseudo-real-time low-pass filter in ECG, self-adjustable to the frequency spectra of the waves." Med Biol Eng Comput 55, 1579–1588, 2017

AUTHOR INDEX

	Pages
Aekapop Bupeng	22
Alongkron Klaiklueng	1
Anan Phonphoem	91, 423
Aphirak Jansang	91, 423
Aphorn Chiawchankaset	8
Apichart Kiatyudthachart	378
Apisak Suno	263
Apisara Onthao	145
Athakorn Kengpol	1
Atiwong Suchato	212
Chaiporn Jaikaeo	91, 423
Chamaiphon Yeamklin	29
Chamaiporn Peesiri	219
Chanapatt Pattaramaetakul	429
Charun Sanrach	185
Chatchawin Namman	326
Chawarus Kerot	270
Chonlada Chuchomchuen	151
Danupon Kumpanya	42
Dechawut Wanichsan	397
Dittaya Wanvarie	212
Donrat Khongharn	411
Duangrudee Phetphirun	239
Fathip Wanhom	276
Jare Sophanpanitchakul	179
Jedsada Arunruerk	110
Jirasak Pipatwarakul	404
Kamonchanok Rungjaroen	257
Kankawin Chamchoi	61
Kantana Pooltong	353
Kanut Injaiea	423
Khanittha Sittitiamjan	85
Konggrit Pitanon	397
Kwannaree Klaprabchone	197
Lertsak Hongjan	197
Mahasak Ketcham	123, 197, 300, 339, 435
Marisa Sangnul	294

AUTHOR INDEX

	Pages
Mintra Mesawong	55
Montean Rattanasiriwongwut	67, 151, 232, 288, 345
Monvorath Phongpaibul	15
Napatkanok Srikanjanodom	138
Narongdech Rattananonsathien	197
Naruemon Pratanwanich	212
Natdanai Kongkhunthod	36
Nathee Yongyut	397
Nattapat Sridoung	300
Nattapol Puninagoon	385
Nattapong Jantarasorn	251
Nattavee Utakrit	61, 116, 378, 385, 391
Natthanicha Suriyamongkol	79
Nichapha Saeweang	232
Nuengwong Tuaycharoen	49, 172, 212
Nuttaya Sangnark	282
Onuma Jamuang	319
Orawan Chunhapran	191
Paitoon Janreung	42
Panumart Inchee	365
Pasit Thong-on	67
Patrapee Sunantapot	110
Pattadon Prongpreecha	49
Pattarapong Jarrusrojoka	191
Payap Sirinam	36, 103
Perm Ratanaphirom	345
Phayung Meesad	226, 404
Piradate Tolornuwat	333
Pramote Kuacharoene	417
Pramote Luenam	130, 138
Prasan Wongkitisopon	429
Prasong Praneetpolgrang	36, 103, 429
Pratin Sonnekorn	359
Preeyarwat Tamawimok	116
Premagorn Thaweehan	417
Pudsadee Boonrawd	353, 359, 365

AUTHOR INDEX

	Pages
Puwanart Poopankul	226
Rachata Sriaroon	130
Ratchaphoom Boonnaka	307
Rath Burirat	191
Sakchai Tangwannawit	165, 219, 282, 371, 411
Samatachai Jantarat	110
Sanit Sirisawatvatana	8
Santi Sukkasem	313
Sarawadee Trakunhun	371
Sathit Suwannawach	85
Siriwarn Seemo	29
Sittichoke Sakon	204
Sittichop Ratrojkij	191
Sittisak Saechueng	22
Somchoke Kittisithipong	123
Sujaree Thairat	15
Sukree Sinthupinyo	79
Sunitta Yentour	397
Supatsara Krasaesat	288
Supeerawong Pothikanit	435
Suphannara Siwaratchakun	339
Supichaya Jitarom	158
Surachat Aunsorn	391
Surasak Mungsing	429
Sutat Gammanee	197
Tanapon Jensuttiwetchakul	55, 145, 158, 276
Taratip Suwannasart	179
Tatchanon Srisamudh	212
Thanakrit Pengkian	103
Thanaphon Yochana	191
Thanawat Yochanang	97
Theerapat Deeprasert	91
Thipwan Fufaung	85
Todsapon Banklongsi	185
Todsaporn Klearwun	429
Ungsumalee Suttapakti	22

AUTHOR INDEX

	Pages
Unhawa Ninrutsirikun	8
Uraivan Inyaem	29
Vithida Chongsuphajaisiddhi	8
Walailak Gatthaisong	165
Wannakorn Puenguthaisri	245
Waraporn Jirapanthong	204
Watchareewan Jitsakul	73, 239, 245, 251, 257, 263, 270, 307, 313, 319, 333
Wipada Klungtes	172
Withawat Tangtrongpaioj	91, 423
Wittawas Puntumchinda	22
Yongyut Intarpsit	73



NCCIT2024

The 20th National Conference on Computing and Information Technology

Organized by

Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok

WWW.NCCIT.NET

www.nccit.net

Conference Organizers



ECU



OSU



HNUE



KMUTNB



FernUni



TU Chemnitz



Gesellschaft für Informatik



MSU



KRU



NPRU



RMUTT



NIDA

