ISSN: 1906-9553 (Print)
ISSN: 2651-1053 (Online)

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY

JIST

Volume 10. No.1 January - June 2020

Research Papers

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY (JIST)

ISSN 2651-1053 (Online), ISSN 1906-9553 (Print)

The Journal of Information Science and Technology (JIST) is an academic journal established by the collaboration of 21 faculties that conduct courses related to Information Technology, namely, Bangkok University, Dhurakij Pundit University, King Mongkut's Institute of Technology Ladkrabang, King Mongkut's University of Technology North Bangkok, King Mongkut's University of Technology Thonburi, Mae Fah Luang University, Mahanakorn University of Technology, Mahasarakham University, Mahidol University, Nakhon Phanom University, Panyapiwat Institute of Management, Prince of Songkla University, Rangsit University, Siam University, Silpakorn University, Sripatum University, Thai-Nichi Institute of Technology, Walailak University, Burapha University, Phayao University and Ubon Ratchathani Rajabhat University. According to the agreement of deans of all faculties (Council of IT Deans of Thailand (CITT)), Mahanakorn University of Technology was appointed as a coordinator of the journal.

The journal was established in 2010 and plans to publish 2 issues per year (JAN – JUNE and JULY – DECEMBER per year). The journal was established first print journal publication in 2010 (Vol 1. No.1) with ISSN 1906-9553 (Print) and plans to publish 2 issues per year on during January - June and July - December. Also the journal was established first online journal publication in 2010 (Vol 1. No.1) with ISSN 2651-1053 (Online) and plans to publish 2 issues per year on during January - June and July - December.

EDITOR IN CHIEF

ADVISORY BOARD

JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY (JIST)

ISSN 2651-1053 (Online), ISSN 1906-9553 (Print)

EDITORIAL BOARD

MANAGER

CONTACT ADDRESS

Council of IT Deans of Thailand (CITT),
Faculty of Information Science and Technology,
Mahanakorn University of Technology
140 Moo 1, Cheum-Sampan Road, Nongchok,
Bangkok, Thailand 10530
Tel: +(662) 988-3655 ext 4115

CALL FOR PAPER
Journal of Information Science and Technology (JIST)

https://ph02.tci-thaijo.org/index.php/JIST

Submissions
Online Journal Submission
https://ph02.tci-thaijo.org/index.php/JIST/

Publication Periodicity
Published 2 times a year in June and December.
(JAN – JUN and JULY – DEC).
(*Indexed by TCI tier 2*)

JIST Office Address

Council of IT Deans of Thailand (CITT), Faculty of Information Science and Technology, Mahanakorn University of Technology 140 Moo 1, Cheum-Sampan Rd., Nongchok, Bangkok, Thailand 10530
Tel: +(662)988-3655 ext 4115,
Fax: +(662)988-4027

Peer Review Process
All submitted manuscripts must be reviewed by at least three expert reviewers via the double-blinded review system.

The **Journal of Information Science and Technology (JIST)** aims to be the forum through which researchers, faculties, and experts of the computer and information technology and others technological related fields share and discuss their high quality research work as well as innovation. Original research articles, practical applications and innovations in the broad area of computer and information technology are suitable for publication in JIST.

- **Soft Computing:**
  Artificial Intelligence, Fuzzy and Neural Computing, Genetic Algorithms, Intelligent Agents, Neural Networks, Natural Language Processing, Robotics and Automation, Image Processing
- **Human-Computer Interaction:**
  Graphical User Interfaces, Multimedia, Interactive Systems, Computer-Supported Cooperative Work, Human Cognitive Skills, Visualization, and Computer Simulation
- **Information Assurance and Security:**
  Cryptography, Forensics and Incident Response, Biometrics, Security Policies and Procedures
- **Information Systems:**
  Databases, Information Retrieval, Transaction Processing, Distributed and Object Databases, Data Warehousing, Knowledge Management, Expert Systems, Multimedia Information Systems, Digital Libraries, Geographical Information Systems
- **Networking:**
  Advanced Computer Networks, Internet Technology and Applications, Distributed Systems, Wireless and Mobile Computing, Data Compression, Network Security, Enterprise Networking, Digital Communications
- **Programming:**
  Object-Oriented Programming, Event-Driven Programming, Functional Programming, Logic Programming, XML and other Extensible Languages, Parallel Programming
- **Platform Technologies:**
  Advanced Computer Architecture, Parallel Architectures, Cluster / Grid Computing, RFID, Embedded Systems
- **System Integration and Architecture:**
  Software Acquisition and Implementation, System Needs Assessment, Software Economics, Enterprise Systems, Computing Economics
- **Social and Professional Issues:**
  Professional Practice, Social Context of Computing, Computers Ethics, IT Law, Intellectual Property, Privacy and Civil Rights
- **Web Systems and Technologies**:
  Programming for the WWW, E-commerce, E-Learning, Content Management Systems, E-Government and E-Service
- **Multidisciplinary:**
  Bioinformatics, Biomedical Information Systems, Nano Technology
- **e-Business and m-Business:**
  e-Business Process Aspects, Intelligent e-Business System, m-Business, e-Supply Chain Management & e-Logistics, e-Customer Relationship Management, e-Negotiation, e-Payment, e-Business Design and Developments
- **Business and Information System:**
  Information management for business applications, Enterprise systems and architecture, Business systems infrastructure design for information integration, Service-oriented architecture, Service-component architecture
- **Social and Business Aspects of Convergence IT and Ubiquitous Computing:**
  Economics of emerging technologies, Home and community computing, Economic and social complexities in CIT and UC, New forms of media and communications
- **Internet of Things (IoT):**
  IoT system architecture, IoT enabling technologies, IoT communication and networking protocols such as network coding, and IoT services and applications and technology development in different standard development organizations (SDO).
- **Cloud Computing:**
  Auditing, monitoring, scheduling, resource registration/discovery, Automatic reconfiguration, self healing/monitoring, Service level agreements, integration, management.
- **Fintech and blockchain:**
  Development of Blockchain Security Enhancement Technologies, Platform Technology that Supports Safe Transactions, Financial and Economic Structures, Blockchain's Impacts on Workstyles, Fast Fintech for Smartphone Application Service Platform.

Manuscript Preparation Guide (Publication Charge: Free)

The template for writing the journal is available for downloading (jist_template_eng-2018.doc or jist_template_thai-2018.doc). Minimum length of the paper should be in between 6-16 pages of A4 size. Text should be typewritten or printed with double-spaced in 11-point of A4 white paper with margins of 1.5" for top, 1.25" for left, and 1" for bottom and right sides. All pages must be numbered sequentially. Here are some guidelines.

- Abstract with length 100-200 words in length.
- Introduction which discusses existing problem and related research
- Materials and methods used for experimentation, Results of the experiment, Discussion and conclusion, References Style is IEEE.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 1-9

ระบบตรวจวัดค่าฝุ่นละอองในอากาศบนอุปกรณ์สมาร์ทโฟน

Measuring Dust System in the Air on Smartphone Devices

*อรวรรณ พลฤทธิ์, ณัฐฐกิตติ์ อานันท์สันติ, ณัฐวัตร เหล่าตระกูลงาม และนวลรัตน์ วัฒนา**

*Orawan Phonrit, Natthakit Arnansanti, Nattawat Laotrakulngam, Nuanrath Wattana**

คณะวิทยาการจัดการ มหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง

Faculty of Management Science, Suandusit University, Trang Center

Received: October 24, 2019; **Revised:** January 20, 2020; **Accepted:** February 28, 2020; **Published:** June 23, 2020

ABSTRACT – The objective of this research was develop dust in the air measuring system on smartphone devices. Because of ours concern about the problem of air pollution from smoke was coming from Indonesia that affect people health in Trang province in March, 2019. As a consequence, ours need to develop this system, by using concept of system development life cycle approach , database design and microcontroller device for system development. For research tool in this research was dust sensor device that setup system at SuanDusit University, Trang center and present dust value on Dust@SDU mobile application on smartphone device. The result found that the $PM_{2.5}$ value from measuring dust system in 24 hours had close value with Pollution control department, since 20-22 in September, 2019 at Bankuan sub-district, Trang district in Trang province that the value from measuring dust system had higher value.

KEYWORDS: Measuring dust in the air system, Smartphone device

บทคัดย่อ -- การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อพัฒนาระบบสำหรับตรวจวัดค่าฝุ่นละอองในอากาศบนอุปกรณ์สมาร์ทโฟน ระบบปฏิบัติการแอนดรอยด์ เนื่องจากคณะผู้วิจัยเล็งเห็นปัญหาของคุณภาพอากาศที่สามารถส่งผลกระทบต่อสุขภาพของประชาชนในพื้นที่จังหวัดตรังจากภาวะควันไฟป่าจากประเทศอินโดนีเซียในปี พ.ศ.2562 จึงเป็นที่มาของเหตุผลในการพัฒนาระบบดังกล่าว โดยอาศัยแนวคิดการพัฒนาระบบ การออกแบบฐานข้อมูล และการทำงานของไมโครคอนโทรลเลอร์สำหรับเครื่องมือที่ใช้ในการวิจัย คือ อุปกรณ์สำหรับตรวจวัดค่าฝุ่น (dust sensor) ซึ่งทำการติดตั้งอุปกรณ์ตรวจวัดค่าฝุ่นละออง ณ มหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง และนำเสนอค่าปริมาณฝุ่นผ่านแอปพลิเคชัน Dust@SDU จากการตรวจวัดปริมาณฝุ่นละอองขนาดเล็ก PM2.5 จากอุปกรณ์ Dust Sensor โดยทำการเก็บตัวอย่างฝุ่นละออง PM2.5 ในอากาศต่อเนื่อง 24 ชั่วโมง ระหว่างวันที่ 20-22 กันยายน 2562 โดยในช่วงเวลาดังกล่าวเป็นช่วงที่จังหวัดตรังได้เผชิญกับปัญหาหมอกควันไฟป่าจากประเทศอินโดนีเซีย พบว่า ข้อมูลมีแนวโน้มใกล้เคียงกันกับข้อมูลจากกรมควบคุมคุณภาพซึ่งวัดคุณภาพอากาศในบริเวณพื้นที่ตำบลบ้านควน อำเภอเมืองตรัง จังหวัดตรัง โดยค่าที่วัดจากอุปกรณ์ Dust sensor มีค่าสูงกว่าเล็กน้อย

คำสำคัญ: ระบบตรวจวัดค่าฝุ่นละอองในอากาศ, อุปกรณ์สมาร์ทโฟน

**Corresponding Author: Nuanrath_wat@dusit.ac.th*

1. บทนำ

ปัญหามลพิษทางอากาศเป็นปัญหาหนึ่งที่สำคัญและทวีความรุนแรงมากขึ้น สาเหตุมาจากการเพิ่มขึ้นของจำนวนประชากร ซึ่งส่งผลต่อการใช้พลังงานที่มากขึ้นตาม ปริมาณอากาศเสียที่มาจากการผลิตของภาคอุตสาหกรรม ภาคการเกษตร และการคมนาคม เป็นต้น รวมถึงอาจเกิดขึ้นเองตามธรรมชาติ ได้แก่ ฝุ่นละอองจากลมพายุ ภูเขาไฟระเบิด แผ่นดินไหว ไฟไหม้ป่า โดยปริมาณอากาศเสียจากสาเหตุดังกล่าวนั้นถูกปล่อยออกมาในรูปของฝุ่นละอองซึ่งสามารถส่งผลกระทบต่อสิ่งแวดล้อมและสุขอนามัยของประชาชนทั้งทางตรงและทางอ้อม [1]โดยเฉพาะฝุ่นละอองขนาดเล็กกว่า 2.5 ไมครอน ($PM_{2.5}$) หากมีค่าเกินมาตรฐานคุณภาพอากาศจะส่งผลกระทบต่อระบบทางเดินหายใจของผู้คน[2]

ทุก ๆ ปี ภาคใต้ของประเทศไทยจะได้รับผลกระทบจากหมอกควันไฟป่าจากประเทศอินโดนีเซีย โดยเมื่อวันที่ 13 กันยายน 2562 กรมควบคุมมลพิษรายงานสถานการณ์ฝุ่นละอองขนาดเล็กกว่า 2.5 ไมครอน ($PM_{2.5}$) ในพื้นที่ของจังหวัดภาคใต้ โดยบริเวณ ตำบลหาดใหญ่ อ.หาดใหญ่ จังหวัดสงขลา ได้รับผลกระทบจากปริมาณฝุ่น $PM_{2.5}$ ระดับ 54 มคก./ลบ.ม.อยู่ในเกณฑ์เริ่มมีผลกระทบต่อสุขภาพ รองลงมา คือ พื้นที่บริเวณตำบลบ้านควน อำเภอเมือง จังหวัดตรัง และพื้นที่ตำบลพิมาน อำเภอเมือง จังหวัดสตูล ในระดับ 43 มคก./ลบ.ม. และ 42 มคก./ลบ.ม. ตามลำดับ โดยอยู่ในเกณฑ์ระดับคุณภาพอากาศปานกลาง จากปัญหาสถานการณ์ฝุ่นละออง $PM_{2.5}$ ที่เกิดขึ้นในพื้นที่ภาคใต้ โดยในจังหวัดตรัง ปัญหาดังกล่าวได้ส่งผลกระทบต่อสุขภาพของประชาชนในพื้นที่เป็นจำนวนมาก จากประเด็นปัญหาด้านคุณภาพอากาศจากฝุ่นละออง $PM_{2.5}$ ที่เกิดขึ้นทำให้คณะผู้วิจัยตระหนักถึงปัญหาและเห็นความสำคัญต่อสุขภาพของผู้คนในพื้นที่จังหวัดตรัง โดยการพัฒนาแอปพลิเคชันสำหรับตรวจวัดค่าฝุ่นละออง $PM_{2.5}$ บนอุปกรณ์สมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์ เพื่อสนับสนุนการจัดการการเฝ้าระวังปัญหาคุณภาพอากาศที่สามารถส่งผลกระทบต่อสุขภาพ โดยทดลองใช้ในพื้นที่มหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ความหมายและลักษณะทั่วไปของฝุ่นละออง

[3] ได้สรุปความหมายของฝุ่นละอองว่าเป็นการรวมกันของอนุภาคของแข็งและหยดละอองของเหลวที่แขวนลอยในอากาศ บางชนิดมีขนาดเล็กมาจึงมองด้วยตาเปล่าไม่เห็น และเป็นสารที่มีความหลากหลายทางด้านกายภาพ และองค์ประกอบทางเคมี

ฝุ่นละอองที่มีอยู่ในบรรยากาศรอบ ๆ ตัวเรา มีขนาดตั้งแต่ 0.002 ไมครอน ไปจนถึงฝุ่นที่มีขนาดใหญ่กว่า 500 ไมครอน โดยฝุ่นละอองที่แขวนลอยอยู่ในอากาศได้นานจะเป็นฝุ่นละอองขนาดเล็กที่มีเส้นผ่านศูนย์กลางต่ำกว่า 10 ไมครอน สำหรับฝุ่นละอองที่มีขนาดใหญ่ที่มีเส้นผ่านศูนย์กลางใหญ่กว่า 10 ไมครอน อาจแขวนลอยอยู่ในบรรยากาศได้เพียง 2-3 นาที แต่ฝุ่นละอองที่มีขนาดเล็ก โดยเฉพาะขนาดเล็กกว่า 0.5 ไมครอนอาจแขวนลอยในอากาศได้นานเป็นปี และสามารถส่งผลกระทบต่อสุขภาพอนามัยของคน สัตว์ พืช และเป็นอุปสรรคในการคมนาคมขนส่ง เป็นต้น

2.2 ประเภทของฝุ่นละออง

ฝุ่นละอองสามารถจำแนกตามขนาดได้แก่ ฝุ่นรวม (Total Suspended Particulate) มีเส้นผ่านศูนย์กลางไม่เกิน 100 ไมครอน ฝุ่นขนาดเล็ก (PM_{10}) คือ ฝุ่นที่มีขนาดเส้นผ่านศูนย์กลางไม่เกิน 10 ไมครอนลงมา และฝุ่นที่มีขนาดเส้นผ่านศูนย์กลางไม่เกิน 2.5 ไมครอน ($PM_{2.5}$) ซึ่งอนุภาคของฝุ่น $PM_{2.5}$ สามารถเข้าสู่ร่างกายไปฝังตัวในเนื้อเยื่อปอด และส่งผลกระทบอย่างรุนแรงต่อกระบวนการแลกเปลี่ยนก๊าซของร่างกาย นอกจากนี้สามารถทำให้เกิดร่องรอยแผลในเส้นเลือดและทำให้สมรรถภาพในการยืดหยุ่นของเส้นเลือดแดงลดลง ทำให้เกิดโรคหัวใจ และโรคที่เกี่ยวกับระบบการไหลเวียนโลหิต อีกทั้งอนุภาคฝุ่นที่มีขนาดเล็กกว่า 100 นาโนเมตร สามารถเคลื่อนที่เข้าสู่ร่างกายได้ในระดับเยื่อหุ้มเซลล์เข้าสู่สมอง และทำให้สมองเกิดความเสียหายได้ [3]

2.3 แนวคิดการพัฒนาระบบ (System Development Life Cycle: SDLC)

วงจรการพัฒนาระบบ ประกอบด้วยขั้นตอนการดำเนินงาน 7 ขั้นตอน ดังนี้ [4]

1) ขั้นตอนการระบุปัญหา และจุดมุ่งหมายของการพัฒนาระบบงานเป็นขั้นตอนที่มีความสำคัญมากเพราะใช้ในการกำหนดทิศทางในการพัฒนาระบบงานให้ชัดเจน สามารถแบ่งออกเป็น 2 กิจกรรมย่อย ได้แก่ การศึกษาเบื้องต้น (Initial Investigation) และ การศึกษาความเหมาะสม (Feasibility Study)

2) ขั้นตอนการวิเคราะห์ระบบ เป็นการนำสิ่งที่รวบรวมข้อมูลจากขั้นตอนที่ 1 มาทบทวนอีกครั้ง และนำมาสร้างเป็นแบบจำลองเชิงตรรกะ (Logical Model) โดยนักวิเคราะห์ระบบจะออกแบบไปตามความต้องการของผู้ใช้ว่าควรมีลักษณะการทำงานของระบบมีรูปแบบที่แสดงผลออกมาอย่างไร มีการจัดเก็บข้อมูลอะไรบ้าง วิเคราะห์ออกมาในรูปแบบของแผนภาพกระแสข้อมูล (Data Flow Diagram) และพจนานุกรมข้อมูล (Data Dictionary) โดยแบ่งออกเป็น 4 กิจกรรมย่อย ได้แก่ 1) ทบทวนระบบปัจจุบัน 2) กำหนดความต้องการระบบใหม่ 3) ออกแบบระบบใหม่ และ 4) วางแผนนำระบบไปใช้และติดตั้งระบบ

3) ขั้นตอนการออกแบบระบบงาน มีจุดมุ่งหมายเกี่ยวกับการแก้ไขปัญหานั้นว่าจะต้องทำอย่างไร ซึ่งในขั้นตอนนี้แบบจำลองเชิงตรรกะ (Logical Model) จะถูกสร้างให้เป็นแบบจำลองทางกายภาพ (Physical Model) คือ การออกแบบให้ระบบนั้นสามารถปฏิบัติงานได้จริง โดยแบ่งออกเป็น 2 กิจกรรมย่อย คือ ออกแบบด้านเทคนิค และทดสอบข้อกำหนดและวางแผน

4) ขั้นตอนการพัฒนา เป็นการทำงานร่วมกันระหว่างโปรแกรมเมอร์และนักวิเคราะห์ระบบเพื่อพัฒนาระบบงาน โดยในขั้นตอนนี้จะต้องมีการจัดทำเอกสารและฝึกอบรมผู้ใช้งานควบคู่ไปด้วย

5) ขั้นตอนการทดสอบระบบ เป็นขั้นตอนเพื่อให้แน่ใจว่าระบบที่พัฒนาขึ้นมาสามารถใช้ได้จริงและถูกต้องตามความต้องการของผู้ใช้ โดยไม่มีข้อผิดพลาดใด ๆ ซึ่งในการทดสอบควรใช้ข้อมูลที่ปฏิบัติงานจริงมาทดสอบ

6) ขั้นตอนการนำไปใช้ เป็นขั้นตอนที่มีการนำระบบงานเข้าไปใช้งานจริง หลังจากทดสอบระบบเรียบร้อยแล้วจะต้องดำเนินการติดตั้งระบบ เพื่อให้ผู้ใช้ได้ใช้งานจริง

7) ขั้นตอนการบำรุงรักษาระบบ เพื่อให้ระบบสามารถทำงานได้อย่างต่อเนื่องในระดับที่ยอมรับได้ซึ่งมีความสำคัญอย่างมาก เพราะอาจมีข้อผิดพลาดที่ไม่รู้มาก่อนขณะทำการทดสอบ หรือผู้ใช้มีความต้องการที่เปลี่ยนแปลงไป เทคโนโลยีต่าง ๆ เปลี่ยนแปลงไป ธุรกิจมีการขยายตัว หรือมีการปรับรูปแบบการบริหารงาน เป็นต้น

2.4 โปรแกรมในการพัฒนาระบบ

2.4.1 Ionic framework

Ionic framework เป็นเครื่องมือใช้ในการสร้าง HTML, CSS และ JavaScript สำหรับพัฒนาโมบายแอปพลิเคชัน ปัจจุบัน Ionic Framework มีทั้งหมด 3 เวอร์ชั่น โดยมีการใช้ Command-line interface (CLI) เข้ามาช่วยสำหรับสร้างหน้าจอการทำงานและช่วยให้การติดตั้งสำหรับนำไปใช้งานง่ายขึ้น นอกจากนี้สามารถรองรับการใช้งานในระบบปฏิบัติการต่าง ๆ โดยใช้งานร่วมกับ Framework อื่น ๆ ด้วย เช่น Angular และ Apache Condova เพื่อให้แอปพลิเคชันที่พัฒนาสามารถใช้ได้กับทุกระบบปฏิบัติการ [5]

2.4.2 โปรแกรม Arduino Software (IDE)

Arduino Software เป็นซอฟต์แวร์ที่ใช้ในการพัฒนางานสำหรับบอร์ด Arduino หรือเรียกว่า Arduino IDE โดยอาศัยโปรแกรมภาษา C สำหรับเขียนโปรแกรมและคอมไพล์ (compile) ลงบนบอร์ด IDE (Integrated Development Environment) และมีหน้าจอ Serial Monitor ใช้สำหรับแสดงผลการทำงานของโปรแกรมแทนการใช้อุปกรณ์แสดงผลอื่น ๆ ซึ่งทำให้นักพัฒนาแอปพลิเคชันเกิดความสะดวกและง่ายต่อการพัฒนาแอปพลิเคชันต่าง ๆ ได้เร็วมากขึ้น

2.5 อุปกรณ์ฮาร์ดแวร์ที่เกี่ยวข้อง

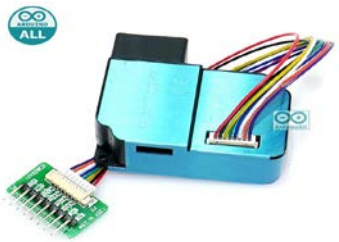
2.5.1 Laser Dust Sensor

เซนเซอร์สำหรับตรวจจับฝุ่น pm 2.5 แบบเลเซอร์ PMS3003 ผลิตโดยบริษัท Plantower ประกอบด้วยโมเดล PMSA003, PMS1003, PMS3003, PMS5003, PMS6003 และ PMS7003 ซึ่งเป็นกลุ่มของเซนเซอร์ที่อาศัยหลักการทางแสงและมีความละเอียดสูง สามารถตรวจวัดฝุ่นที่มีขนาด 0.3 ถึง 10 μm ให้การตอบสนองเร็วประมาณ 1 วินาที โดยในตัวเซนเซอร์มีอุปกรณ์คอนโทรลเลอร์ในตัวซึ่งให้สัญญาณเอาต์พุตแบบพอร์ตอนุกรม (Rx/Tx) และพัดลมดูดฝุ่นในอากาศที่ต้องการวัดเข้ามาในตัวเซนเซอร์ให้ตัดผ่านแหล่งกำเนิดแสงอินฟราเรดและโฟโตทรานซิสเตอร์ เมื่อใช้งานกับไมโครคอนโทรลเลอร์ควร

ใช้ความเร็ว Baud rate ที่ระดับ 9600 bps ลักษณะการทำงานของเซนเซอร์ตรวจจับฝุ่น pm2.5 รุ่น PSM3003 ใช้พัดลมทำหน้าที่ดูดอากาศเข้าไปในตัวเซนเซอร์ แล้วตรวจจับฝุ่นด้วยแสงเลเซอร์ ค่าที่ได้ออกมาเป็นแอนะล็อก 1-1023 ลักษณะการทำงาน คือ เซนเซอร์จะส่งแสงเลเซอร์ไปกระทบกับตัวรับ และให้อากาศผ่านในช่อง หากมีการรับแสงได้น้อย แสดงว่า ฝุ่นละอองเยอะ หากมีการรับแสงได้มากแสดงว่าฝุ่นละอองน้อย นอกจากนี้ สามารถนำไปวัดควันธูป แป้ง และเหมาะสมกับการนำไปประยุกต์ใช้กับ Air Purifier Air Conditioner และ Air monitor เป็นต้น [6]

2.5.2 ไมโครคอนโทรลเลอร์ Arduino ESP8266 (NodeMCU)

โหนด เอ็มซียู (NodeMCU) คือ แพลตฟอร์มหนึ่งสำหรับพัฒนาโปรเจคที่เกี่ยวข้องกับอินเทอร์เน็ตของสรรพสิ่ง (Internet of Things :IoT) ประกอบด้วย Development Kit (ตัวบอร์ด) และ Firmware (ซอฟต์แวร์บนบอร์ด) ซึ่งเป็นโอเพ่นซอร์สที่ถูกพัฒนาด้วยภาษาลัว (Lau Programming Language) สำหรับโมดูล WiFi (ESP8266) ในโหนดเอ็มซียู มีความสำคัญในการใช้เชื่อมต่อกับอินเทอร์เน็ต ลักษณะของโหนดเอ็มซียู จะมีความคล้ายกับบอร์ด Arduino เนื่องจากมีพอร์ต Input/Output ติดมาด้วย (built in) ทำให้สามารถเขียนโปรแกรมเพื่อควบคุมอุปกรณ์ I/O ได้โดยไม่ต้องผ่านอุปกรณ์เชื่อมต่ออื่น ๆ อีก นอกจากนี้ Arduino IDE สามารถทำงานร่วมกับโหนดเอ็มซียู จึงทำให้นักพัฒนาสามารถใช้ภาษา C/C++ เขียนโปรแกรมได้ และใช้งานได้หลากหลายมากยิ่งขึ้น [7],[8]



รูปที่ 1. แสดงภาพเซนเซอร์สำหรับตรวจจับฝุ่น pm2.5 รุ่น PMS3003

(ที่มา : https://www.arduinoall.com/product/2010/laser-dust-sensor-pm2-5-pms3003)

2.6 งานวิจัยที่เกี่ยวข้อง

[9] ศึกษาเกี่ยวกับการวิเคราะห์ปริมาณฝุ่นละอองเชิงมวล $PM_{2.5}$ และ PM_{10} ในบรรยากาศด้วยเครื่องตรวจวัดฝุ่นละอองไร้สายในพื้นที่ภาคเหนือของประเทศไทย จำนวน 4 สถานี ได้แก่ 1) สถานีโรงเรียนยุพราช 2) สถานีอำเภอดอยสะเก็ด 3)สถานีตำบลแม่เหียะ อำเภอเมือง มหาวิทยาลัยเชียงใหม่ และ 4) สถานีอำเภอนาน้อย จังหวัดน่าน โดยเก็บรวบรวมค่าฝุ่นละอองขนาดเล็กเฉลี่ยรายวันที่วัดได้จากเครื่องวัด DustDETECH มาทำการเปรียบเทียบกับข้อมูลจากสถานีตรวจวัดคุณภาพอากาศกรมควบคุมมลพิษ (สถานีโรงเรียนยุพราช) ระหว่างวันที่ 17 มีนาคม -10 เมษายน 2560 พบว่า ข้อมูลมีแนวโน้มที่ใกล้เคียงกัน และเมื่อนำข้อมูลมาเปรียบเทียบกับข้อมูลจากกรมควบคุมมลพิษในช่วงเดือนมีนาคม-เมษายน 2560 พบว่า มีแนวโน้มเป็นไปในทิศทางเดียวกัน โดยข้อมูลที่ได้จากการวัดด้วยเครื่องวัดฝุ่นละอองขนาดเล็ก DustDETECH จะถูกรวบรวมจัดเก็บในระบบฐานข้อมูล โดยบันทึกผลแบบออนไลน์อัตโนมัติ และแสดงผลผ่านเว็บไซต์ (www.cmuccdc.org)

[10] ศึกษางานวิจัย Usability Metric Framework for Mobile Phone Application โดยศึกษาทฤษฎีเรื่อง Quality in Use Integrated Measurement (QUIM) ซึ่งเป็นโมเดลของการรวบรวมการวัดการใช้งาน และดำเนินการพัฒนาตัวชี้วัดการใช้งานแอปพลิเคชัน โดยใช้ Goal Question Metric (GQM) ผลการศึกษาพบว่า มีแนวทางทั้งหมด 6 ด้าน ที่จะช่วยให้บรรลุเป้าหมายในการสร้างเฟรมเวิร์ค คือ 1) ความง่ายในการใช้งาน 2) ความถูกต้อง 3) ระยะเวลาในการเข้าถึงข้อมูล 4) ฟังก์ชันการใช้งาน 5) ความปลอดภัย และ 6) ความสวยงามและน่าใช้งาน

3. วิธีการดำเนินงานวิจัย

3.1 พื้นที่ศึกษาวิจัย

การศึกษาวิจัยครั้งนี้ต้องการศึกษาถึงปริมาณฝุ่นละอองขนาดเล็ก $PM_{2.5}$ ที่เกิดขึ้นในบริเวณมหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง

3.2 การติดตั้งเครื่องวัดฝุ่นละอองขนาดเล็ก (Dust Sensor) มีขั้นตอนดังนี้

1) ติดตั้งเครื่องวัดฝุ่นละอองขนาดเล็ก (Dust Sensor) ให้สามารถวัดฝุ่นละออง $PM_{2.5}$ และชุดอินเทอร์เฟสเพื่อควบคุมและส่งผ่านเครือข่ายอินเตอร์เน็ตแบบไร้สาย โดยมีการบันทึกลงฐานข้อมูลสำหรับระบบตรวจวัดค่าฝุ่นละอองในอากาศ และสร้างแอปพลิเคชันเพื่อตรวจสอบค่าฝุ่นละอองในอากาศ $PM_{2.5}$

2) ทดสอบการวัดค่าปริมาณฝุ่น $PM_{2.5}$ โดยทำการเก็บตัวอย่างฝุ่นละออง $PM_{2.5}$ ในอากาศอย่างต่อเนื่อง 24 ชั่วโมง เป็นเวลา 3 วัน

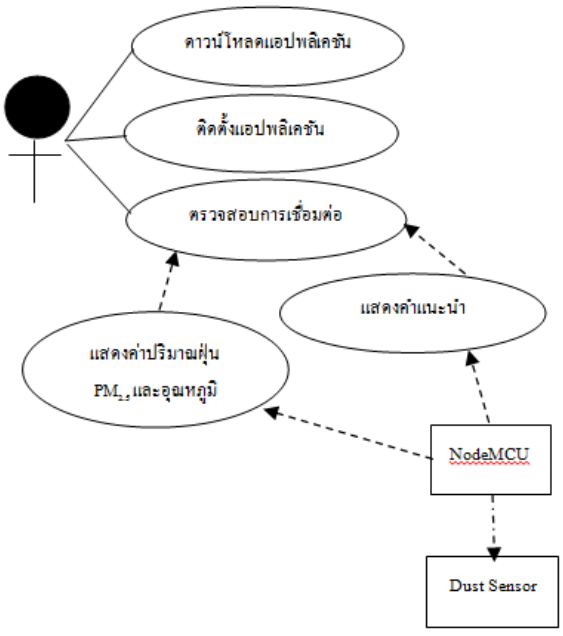
3) เปรียบเทียบค่าปริมาณฝุ่นละอองที่วัดได้จากเครื่องวัดฝุ่นละออง (Dust Sensor) กับข้อมูลจากกรมควบคุมมลพิษ (เว็บไซต์ air4thai.pcd.go.th) ระหว่างวันที่ 20 -22 กันยายน 2562 ซึ่งเปรียบเทียบกับพื้นที่การวัดบริเวณตำบลบ้านควน อำเภอเมือง จังหวัดตรัง

3.2 การวิเคราะห์และออกแบบระบบ

ผู้วิจัยได้ทำการวิเคราะห์และออกแบบระบบสำหรับตรวจวัดค่าฝุ่นละอองในอากาศด้วยแผนภาพการทำงานของผู้ใช้ระบบ (Use Case Diagram) แผนภาพขั้นตอนการทำงานของระบบ (Flow Chart) และการออกแบบฐานข้อมูลด้วย E-R ไดอะแกรม เพื่อเป็นแนวทางการพัฒนาแอปพลิเคชันสำหรับแสดงผลค่าปริมาณฝุ่นละอองในอากาศ $PM_{2.5}$ ดังนี้

3.2.1 แผนภาพแสดงการทำงานของผู้ใช้ระบบ (Use Case Diagram)

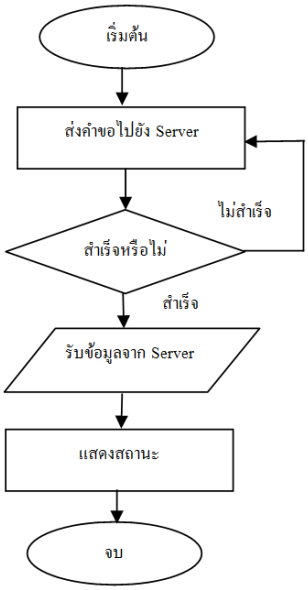
เป็นการอธิบายให้เห็นถึงการตอบสนองของระบบต่อผู้ใช้งาน โดยสามารถอธิบายการทำงานของระบบเมื่อผู้ใช้ต้องการดูสถานะของฝุ่นและอุณหภูมิ และดูคำแนะนำจากระบบ ดังรูปที่ 2



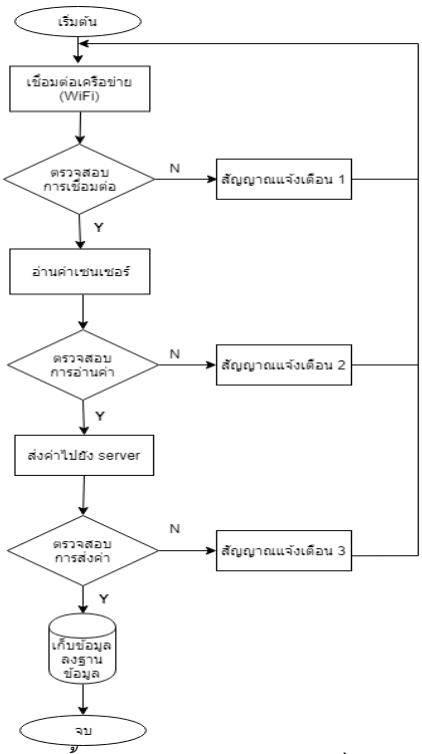
รูปที่ 2. แสดงแผนภาพการทำงานของผู้ใช้ระบบตรวจวัดค่าฝุ่นละออง $PM_{2.5}$ (Use Case Diagram)

3.2.2 แผนภาพแสดงขั้นตอนการทำงานของระบบตรวจวัดค่าฝุ่นละออง (Flowchart)

เป็นการอธิบายขั้นตอนการทำงานของระบบ และการทำงานของโหนดเอ็มซียู (NodeMCU) และดังรูปที่ 3 และ 4



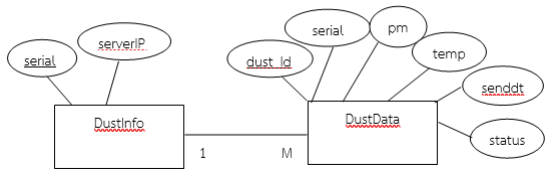
รูปที่ 3. แสดงขั้นตอนการทำงานของระบบตรวจวัดค่าฝุ่นละออง $PM_{2.5}$ (Flowchart)



รูปที่ 4. แสดงขั้นตอนการทำงานของโหนดเอ็มซียู (NodeMCU Flowchart)

3.2.3 การออกแบบฐานข้อมูลด้วยอี-อาร์ไดอะแกรม

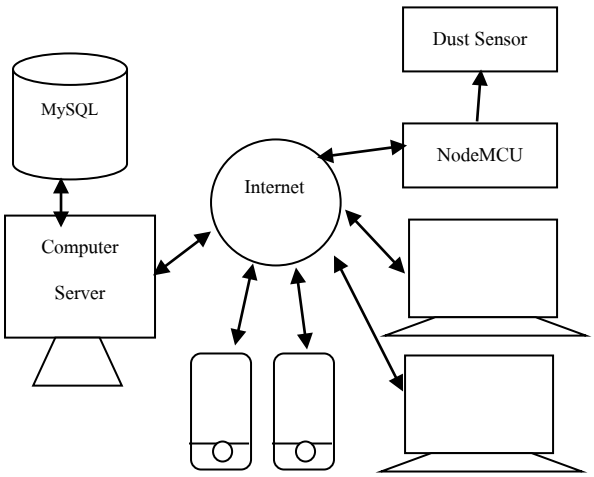
เป็นการจำลองโครงสร้างของฐานข้อมูลและความสัมพันธ์ของข้อมูล ประกอบด้วย เอนทิตี้ (Entity) และแอททริบิว (Attribute)



รูปที่ 5. แสดง E-R ไดอะแกรมของระบบตรวจวัดค่าฝุ่นละออง

3.3 องค์ประกอบการพัฒนาระบบตรวจวัดค่าฝุ่นละออง

การทำงานของระบบสำหรับตรวจวัดค่าฝุ่นละอองมีองค์ประกอบดังรูปที่ 6



รูปที่ 6. แสดงภาพจำลององค์ประกอบของการพัฒนาระบบตรวจวัดค่าฝุ่นละอองในอากาศ

จากรูปที่ 6 องค์ประกอบของการพัฒนาระบบตรวจวัดค่าฝุ่นละอองในอากาศ ประกอบด้วยภาคการทำงานที่สำคัญ ดังนี้

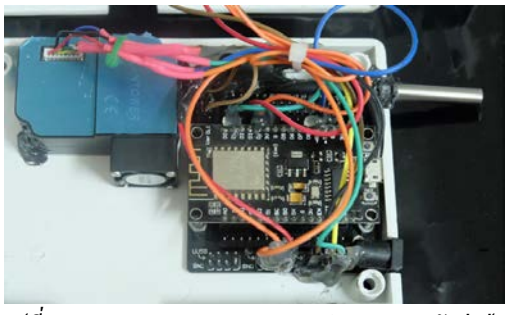
1) คอมพิวเตอร์ (Computer) เป็นอุปกรณ์ที่เป็นตัวสำคัญในการใช้งาน ทำหน้าที่ควบคุมการทำงานทั้งหมดและจัดเก็บข้อมูลตรวจวัดค่าฝุ่นละอองในอากาศลงฐานข้อมูล (Database) ของแอปพลิเคชันสำหรับตรวจวัดค่าฝุ่นละอองในอากาศ

2) โหนดเอ็มซียู (NodeMCU) เป็นอุปกรณ์ที่ทำหน้าที่เป็นตัวกลางหลักในการรับคำสั่งจากโปรแกรมสั่งงานในคอมพิวเตอร์ โดยการสั่งงานของคำสั่งนั้นจะสั่งงานโดยคอมพิวเตอร์ผ่านทางระบบเครือข่ายอินเทอร์เน็ต มายังโหนดเอ็มซียู (NodeMCU) เพื่อควบคุมการทำงานของแอปพลิเคชันสำหรับตรวจวัดค่าฝุ่นละอองในอากาศ

3) เซนเซอร์ฝุ่นละออง (Dust sensor) เป็นอุปกรณ์ที่ทำหน้าที่วัดค่าฝุ่นละอองในอากาศ โดยรับคำสั่งมาจากโหนดเอ็มซียูให้วัดค่าฝุ่นละอองในอากาศ เมื่อได้ค่าฝุ่นละอองเซนเซอร์จะส่งข้อมูลไปยังโหนดเอ็มซียู

4) MySQL เป็นฐานข้อมูลสำหรับจัดเก็บค่าฝุ่นละอองในอากาศและค่าอุณหภูมิที่ถูกส่งมาจากเซนเซอร์ฝุ่นละอองเพื่อจัดเก็บลงในฐานข้อมูล MySQLที่ถูกเชื่อมต่อกับคอมพิวเตอร์ โดยจัดเก็บค่าฝุ่นละออง และอุณหภูมิ ตามวันที่ เวลา และคำแนะนำเบื้องต้นลงในโครงสร้างตาราง (Table Schema)

5) อุปกรณ์สมาร์ทโฟน (Smartphone) เป็นอุปกรณ์ที่ใช้สำหรับติดตั้งระบบสำหรับตรวจวัดค่าฝุ่นละอองในอากาศซึ่งรองรับระบบปฏิบัติการแอนดรอยด์โดยสามารถเชื่อมต่อกับระบบอินเทอร์เน็ตเพื่ออ่านค่าฝุ่นละอองในอากาศและคำแนะนำเบื้องต้นสำหรับคุณภาพอากาศนั้น ๆ



รูปที่ 7. วงจรภาคควบคุมระบบการทำงานตรวจวัดค่าฝุ่นละอองในอากาศ

4. ผลการดำเนินงานวิจัย

4.1 การตรวจวัดปริมาณฝุ่นละออง $PM_{2.5}$ ด้วยอุปกรณ์ Dust Sensor

จากการตรวจวัดปริมาณฝุ่นละอองขนาดเล็ก $PM_{2.5}$ จากอุปกรณ์ Dust Sensor โดยทำการเก็บตัวอย่างฝุ่นละออง $PM_{2.5}$ ในอากาศต่อเนื่อง 24 ชั่วโมง ระหว่างวันที่ 20-22 กันยายน 2562 โดยในช่วงเวลาดังกล่าวเป็นช่วงที่จังหวัดตรังได้เผชิญกับปัญหาหมอกควันไฟป่าจากประเทศอินโดนีเซีย พบว่า ข้อมูลมีแนวโน้มใกล้เคียงกันกับข้อมูลคุณภาพอากาศจากกรมควบคุมคุณภาพ บริเวณพื้นที่ ตำบลบ้านควน อำเภอเมืองตรัง จังหวัด

ตรัง โดยค่าที่วัดจากอุปกรณ์ Dust sensor มีค่าสูงกว่าเล็กน้อย ดังตารางที่ 1

ตารางที่ 1 แสดงค่าเฉลี่ยฝุ่นละออง $PM_{2.5}$ 24 ชั่วโมง ระหว่างวันที่ 20-22 กันยายน 2562

แหล่งข้อมูล	20 ก.ย.	21 ก.ย.	22 ก.ย.
กรมควบคุมมลพิษทางอากาศ วัดบริเวณพื้นที่ ต.บ้านควน อ.เมือง จ.ตรัง [11]	33	30	64
Dust Sensor พื้นที่ มหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง อ.เมือง จ.ตรัง	34	35	68

หมายเหตุ : หน่วยวัด $PM_{2.5}$ ไมโครกรัมต่อลูกบาศก์เมตร (มคก.ต่อลบ.ม.)

เกณฑ์ดัชนีคุณภาพอากาศสำหรับประเทศไทย

0-50 : คุณภาพดี ไม่มีผลกระทบต่อสุขภาพ

51-100 : คุณภาพปานกลาง ไม่มีผลกระทบต่อสุขภาพ

101-200 : มีผลกระทบต่อสุขภาพ

201-300 : มีผลกระทบต่อสุขภาพมาก

มากกว่า 300 : อันตราย

เกณฑ์ดัชนีฝุ่น $PM_{2.5}$

0-25 : คุณภาพอากาศดีมาก

26-37 : คุณภาพอากาศดี

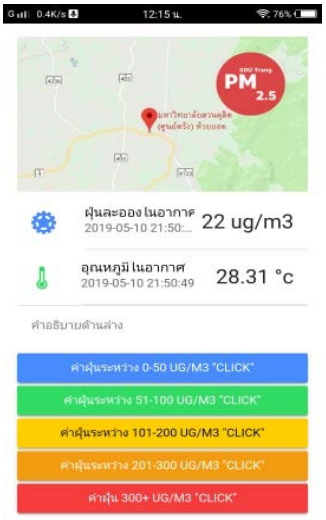
38-50 : คุณภาพอากาศปานกลาง

51-80 : เริ่มมีผลกระทบต่อสุขภาพ

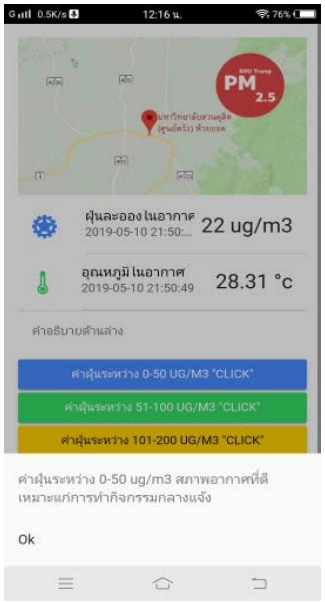
มากกว่า 80 : มีผลกระทบต่อสุขภาพ

4.2 ผลการพัฒนาแอปพลิเคชันสำหรับแสดงผลค่าปริมาณฝุ่นละอองในอากาศ

การพัฒนาแอปพลิเคชัน Dust@SDU สำหรับแสดงผลค่าปริมาณฝุ่นละอองในอากาศได้ดำเนินการออกแบบและพัฒนาตามแนวคิดการพัฒนาระบบ (SDLC) ผลการพัฒนาแอปพลิเคชันบนอุปกรณ์สมาร์ทโฟนระบบปฏิบัติการแอนดรอยด์แสดงดังรูปที่ 8-9



รูปที่ 8. แสดงหน้าจอหลักของแอปพลิเคชันตรวจวัดค่าฝุ่นละอองในอากาศ Dust@SDU



รูปที่ 9. แสดงหน้าจอเกณฑ์ระดับค่าฝุ่นละอองในอากาศและคำอธิบาย

5. บทสรุปและการอภิปราย

5.1 สรุปผลการวิจัย

จากการศึกษาวิจัย พบว่า งานวิจัยนี้สามารถศึกษาปริมาณของฝุ่นละอองในอากาศบริเวณมหาวิทยาลัยสวนดุสิต ศูนย์การศึกษานอกที่ตั้ง ตรัง โดยอาศัยระบบตรวจวัดค่าฝุ่นละอองที่พัฒนาขึ้น และทำการตรวจสอบข้อมูลที่ได้เปรียบเทียบกับข้อมูลจากกรมควบคุมมลพิษ พบว่า ข้อมูลที่ได้มีแนวโน้มใกล้เคียงกัน และทำการพัฒนาระบบฐานข้อมูลสำหรับจัดเก็บ

ข้อมูลของฝุ่นละอองในอากาศ รวมไปถึงการนำเสนอข้อมูลผ่านแอปพลิเคชันตรวจวัดค่าฝุ่นละอองในอากาศ Dust@SDU เพื่อใช้สำหรับตรวจสอบคุณภาพอากาศ นอกจากนี้ งานวิจัยมีความมุ่งหวังอยากให้เกิดองค์ความรู้ที่สามารถนำไปต่อยอดการขยายพื้นที่ติดตั้งระบบตรวจวัดค่าฝุ่นละอองในอากาศเพื่อเตือนภัยวิกฤตฝุ่นควันที่ส่งผลต่อสุขภาพของประชาชนในพื้นที่จังหวัดตรัง

5.2 ข้อจำกัดการวิจัย

1) อุปกรณ์เซนเซอร์จำเป็นต้องเชื่อมต่อกับเครือข่ายอินเทอร์เน็ต (Wi-fi) และควรมีระดับสัญญาณที่เป็นไปตามคุณสมบัติ (2.4 G) ของอุปกรณ์โหนดเอ็มซียู เพื่อให้ได้ค่าที่เที่ยงตรง

2) ในกรณีที่กระแสไฟฟ้าไปเลี้ยงอุปกรณ์ควบคุมและอุปกรณ์เซนเซอร์ไม่มีความเสถียรจะสามารถส่งผลต่อการรับค่า และการประมวลผล ซึ่งจะทำให้การแสดงค่าของฝุ่นคลาดเคลื่อนได้

5.3 ข้อเสนอแนะและการอภิปราย

1) ควรพัฒนาระบบให้สามารถแสดงค่าของพารามิเตอร์อื่น ๆ ในอุปกรณ์เซนเซอร์ เช่น ค่า PM_1 และ PM_{10} เป็นต้น

2.) ควรพัฒนาระบบตรวจวัดค่าฝุ่นละอองในอากาศให้สามารถดูข้อมูลย้อนหลัง หรือมีกราฟแสดงค่าฝุ่นละอองเพื่อให้ผู้ใช้อ่านสามารถดูรายละเอียดมากขึ้น พร้อมแสดงค่าเฉลี่ยของค่าฝุ่นละอองในอากาศในแต่ละช่วงเวลา

3) ควรพัฒนาระบบตรวจวัดค่าฝุ่นละอองในอากาศให้สามารถแจ้งเตือนค่าฝุ่นละอองอัตโนมัติที่เกินมาตรฐานให้ผู้ใช้งานทราบ

เอกสารอ้างอิง

[1] กรมควบคุมมลพิษ, “มาตรฐานฝุ่นละอองในบรรยากาศโดยทั่วไป,” กรมควบคุมมลพิษ, 2562. [Online] Available: http://www.pcd.go.th/info_serv/reg_std_airsnd01.html. [Accessed: 5 สิงหาคม 2562].

[2] กรมส่งเสริมคุณภาพสิ่งแวดล้อม, “มลพิษทางอากาศ,” กรมส่งเสริมคุณภาพสิ่งแวดล้อม, 2562. [Online] Available: http://www.deqp.go.th/knowledge. [Accessed: 5 สิงหาคม 2562].

[3] คณะสิ่งแวดล้อมและทรัพยากรศาสตร์ มหาวิทยาลัยมหิดล, “ฝุ่นละอองในบรรยากาศ,” คณะสิ่งแวดล้อมและทรัพยากรศาสตร์ มหาวิทยาลัยมหิดล, 2562. [Online]. Available: http://www.en.mahidol.ac.th/elearning/upload/Dust_Patcharawadee.pdf. [Accessed: 15 สิงหาคม 2562].

[4] สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, “วงจรการพัฒนาระบบ,” สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2562. [Online]. Available: http://www.swpark.or.th/sdlcproject/index.php/14-sample -data-articles/87-2013-08-09-08-39-48. [Accessed: 5 เมษายน 2562].

[5] ภานุพงศ์ ศุภะรัชฎเดช, “*Ionic framework*,” มิถุนายน, 2562. [Online]. Available: https://www.imwritingrich.com/what-is-ionic-framework. [Accessed: 2 มิถุนายน 2562].

[6] ArduinoAll, “Laser Dust Sensor pm2.5 PMS300,” 10 กรกฎาคม, 2562. [Online]. Available: https://www.arduinoall.com/product/2010/laser-dust-sensor-pm2-5-pms3003. [Accessed: 10 กรกฎาคม 2562].

[7] Manufacture Overhaul Rapid an Optical Co., Ltd., “ไมโครคอนโทรลเลอร์,” Manufacture Overhaul Rapid an Optical Co., Ltd., 2562. [Online]. Available: http://www.moro.co.th/ไมโครคอนโทรลเลอร์. [Accessed: 13 มิถุนายน 2562].

[8] สถิตธรรม สังข์ทอง, “Arduino ESP8266 (NodeMCU),” มิถุนายน, 2562. [Online]. Available: https://medium.com/sathittham. [Accessed: 24 มิถุนายน 2562].

[9] ฐิฏาพร สุภายี, พาณิชย์ อินต๊ะ, เสริมเกียรติ จอมจันทร์ยอง และเศรษฐ์ สัมภัตตะกุล, “การวิเคราะห์ปริมาณฝุ่นละอองเชิงมวล $PM_{2.5}$ และ PM_{10} ในบรรยากาศด้วยเครื่องตรวจวัดฝุ่นละอองไร้สายในพื้นที่ภาคเหนือของประเทศไทย,” วารสารวิจัยเทคโนโลยีนวัตกรรม, ฉบับที่ 1, ปีที่ 2, มกราคม-มิถุนายน, หน้า 69-83, 2561.

[10] A. Hussain and E. Ferneley, “Usability metric for mobile application: a goal question metric (GQM) approach,” In Proc. 10th International Conference on Information Integration and Web-based Applications & Services (iiWAS ’08), 2008, pp. 567–570.

[11] กองควบคุมมลพิษ, “รายงานสถานการณ์และคุณภาพอากาศประเทศไทย,” กองควบคุมมลพิษ, 2562. [Online]. Available: air4thai.pcd.go.th/webV2/station.php?station =m119. [Accessed: 30 กันยายน 2562].

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 10-23

วิธีการเลือกข้อมูลที่ไม่มีป้ายกำกับอย่างอัตโนมัติเพื่อการสร้างโมเดลการเรียนรู้ร่วมแบบกึ่งมีผู้สอน

An Automatic Unlabeled Selection for CO-training REGressors (AU-COREG)

ศิริขวัญ คีรีสุวรรณกุล และ เอกสิทธิ์ พัชรวงศ์ศักดา*

Sirikwan Kheereesuwannakul and Eakasit Pacharawongsakda*

สาขาวิศวกรรมข้อมูลขนาดใหญ่ วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

มหาวิทยาลัยธุรกิจบัณฑิตย์

Big Data Engineering, College of Innovative Technology and Engineering, Dhurakit Pundit University

Received: November 30, 2019; **Revised:** February 02, 2020; **Accepted:** February 27, 2020; **Published:** June 23, 2020

ABSTRACT – This research aims to improve the performance of semi-supervised learning by automatically select unlabeled data. The proposed method uses two regression models to estimate values for unlabeled data, then cluster the data into groups. Therefore, similar data are assigned in the same group and the different data are assigned into the different groups. After that, the method selects each group representative that have least error and append into training data. Then, we repeat until we have enough training data. From experimental results with three datasets, we found that the proposed method can improve performance and reduce computation time by 84%, comparing to previous work.

KEYWORDS: Co-Training, Simi-Supervised Learning

บทคัดย่อ -- งานวิจัยนี้มีวัตถุประสงค์เพื่อปรับปรุงประสิทธิภาพโมเดลพยากรณ์ ด้วยวิธีการเลือกข้อมูลที่ไม่มีป้ายกำกับอย่างอัตโนมัติ เพื่อสร้างโมเดลการเรียนรู้ร่วมแบบกึ่งมีผู้สอน (semi-supervised learning) ซึ่งเหมาะสำหรับข้อมูลที่มีป้ายกำกับ (labeled data) ปริมาณน้อยมาก โดยวิธีการที่นำเสนอนี้จะใช้ประโยชน์จากข้อมูลที่ไม่มีป้ายกำกับ (unlabeled data) ที่มีอยู่ปริมาณมากมาช่วยเพิ่มประสิทธิภาพของการสร้างโมเดลจำแนกประเภทข้อมูล (classification) หรือการประมาณค่า (regression) วิธีการที่นำเสนอนี้เริ่มด้วยการใช้โมเดล 2 โมเดลทำการกำกับค่าให้กับข้อมูลที่ไม่มีป้ายกำกับ จากนั้นนำข้อมูลเหล่านี้มาทำการจัดกลุ่ม (clustering) ให้ข้อมูลที่มีความคล้ายคลึงกันอยู่ในกลุ่มเดียวกัน และแยกข้อมูลที่ต่างกันออกให้อยู่ต่างกลุ่มกัน ถัดมาจึงเลือกตัวแทนแต่ละกลุ่มเพื่อหาข้อมูลที่ทำให้โมเดลการพยากรณ์มีความคลาดเคลื่อนน้อยที่สุดเข้าไปเป็นชุดข้อมูลสอน (training data) ในรอบถัดไปและสร้างโมเดลพยากรณ์ใหม่ ทำซ้ำจนเพิ่มข้อมูลเข้าไปในชุดสอนได้ครบ จากนั้นการพยากรณ์ขั้นสุดท้ายทำได้โดยการหาค่าเฉลี่ยของการพยากรณ์จากทั้งสองโมเดลที่สร้างขึ้น จากการทดสอบด้วยข้อมูลจำนวน 3 ชุดแสดงให้เห็นว่าวิธีการที่นำเสนอ (AU-COREG) สามารถปรับปรุงประสิทธิภาพของโมเดลได้อย่างมีนัยสำคัญและช่วยลดเวลาลงไปมากกว่า 84% เมื่อเทียบกับวิธีการเดิม

คำสำคัญ: การเรียนรู้ร่วม, เรียนรู้แบบกึ่งมีผู้สอน

**Corresponding Author: 585162020005@dpu.ac.th*

1. บทนำ

โลกปัจจุบันได้เข้าสู่ยุคดิจิทัล ดังเห็นได้จากอุปกรณ์ต่างๆ ที่มีการสร้างข้อมูลในเชิงดิจิทัลเพิ่มขึ้นอย่างเป็นจำนวนมาก โดยส่วนใหญ่เกิดจากการใช้งานอินเตอร์เน็ต จึงส่งผลให้พฤติกรรมของมนุษย์มีการเปลี่ยนแปลงจากอดีต กิจกรรมหลายๆอย่าง ถูกแทนที่ด้วยแพลตฟอร์มบนโลกออนไลน์ เช่น การซื้อสินค้า การประกาศรับสมัครงาน การดูหนังฟังเพลง การแชทหรือโซเชียลเน็ตเวิร์ค เป็นต้น แพลตฟอร์มเหล่านี้ได้สร้างข้อมูลจำนวนมหาศาลบนโลกออนไลน์ รายงานของ IBM ระบุว่า 90% ของข้อมูลทั้งหมดในโลกออนไลน์ ถูกสร้างขึ้นในช่วง 2 ปีหลังนี้เอง โดยปัจจุบันมีข้อมูลเกิดขึ้นใหม่ราว 2,500 ล้านกิกะไบต์ต่อวัน [1]

หากพิจารณาข้อมูลเหล่านั้น ข้อมูลที่มีป้ายกำกับ (Labeled Data) จะมีสัดส่วนน้อยมาก เมื่อเทียบกับข้อมูลที่ไม่มีป้ายกำกับ (Unlabeled Data) เนื่องจากการกำกับค่าให้ข้อมูลนั้น มีต้นทุนที่สูง หรือค่าที่กำกับไม่เป็นจริง หรืออาจไม่สามารถกำกับค่าได้ เพราะความต้องการใช้ข้อมูลที่เปลี่ยนแปลงไปอย่างรวดเร็ว ดังนั้นการสร้างโมเดลพยากรณ์จำเป็นต้องมีข้อมูลสอน (Training Data) ซึ่งข้อมูลสอนได้จากข้อมูลที่มีป้ายกำกับ ทว่าการที่ข้อมูลสอนนี้มีสัดส่วนน้อยมากอาจจะทำให้ความคลาดเคลื่อนสูง เมื่อเทียบกับการมีข้อมูลที่มีป้ายกำกับที่มีมากกว่า ดังนั้นการให้โมเดลเรียนรู้แบบกึ่งมีผู้สอน (Semi-supervised Learning) จากข้อมูลทั้งมีป้ายกำกับและข้อมูลที่ไม่มีป้ายกำกับจึงถูกนำมาประยุกต์ใช้ ซึ่งวิธีที่ได้รับความนิยมอย่างแพร่หลายคือ วิธีการโค เทรนนิง (Co-Training)

วิธีการ โคเทรนนิงมีการนำเสนอครั้งแรกโดย Blum และ Mitchell ในปี 1998 [2] ซึ่งเป็นการนำข้อมูลที่ไม่มีป้ายกำกับมาช่วยเพิ่มประสิทธิภาพของโมเดลพยากรณ์ การเลือกข้อมูลที่ไม่มีป้ายกำกับนี้จะใช้การสร้างโมเดล 2 โมเดลและเลือกข้อมูลที่ไม่มีป้ายกำกับที่มีความเชื่อมั่นมากที่สุดจากการพยากรณ์มาเพิ่มเข้าไปในข้อมูลสอน และสร้างโมเดลพยากรณ์ใหม่วนเรื่อยไป หลังจากนั้น Luca Didaci, Giorgio Fumera และ Fabio Roli [3] ได้ทำการวิจัยถึงผลกระทบของประสิทธิภาพของโมเดลที่เกิดจากการใช้โค เทรนนิงกับขนาดของชุดข้อมูลสอน นั่นคือลดขนาดของชุดข้อมูลสอน ให้มีขนาดน้อยที่สุด จนกระทั่งไม่สามารถนำไปใช้ได้ โดยทดสอบกับข้อมูลทั้งหมด 24 ชุดข้อมูล พบว่าขนาดของข้อมูลที่มีป้ายกำกับเพียง 1 ตัวอย่างต่อชุดข้อมูลสอน ไม่ส่งผลต่อประสิทธิภาพการโค เทรนนิง ถัดมา Ruiya Wang และ Li Li [4] ได้ทำพัฒนาอัลกอริทึมการปรับปรุงประสิทธิภาพของโค เทรนนิงโดยคณะกรรมการ (Co-Training by committee) เป็นวิธีการเรียนรู้แบบกึ่งกำกับซ้ำ ซึ่งในระหว่างการทำซ้ำ จะใช้หลายๆ โมเดล (committee) ก่อนหน้านั้นทั้งหมดหลายๆ ชุด เพื่อใช้ในการทำนายตัวอย่างที่ไม่มีป้ายกำกับในแต่ละครั้ง ซึ่งสามารถเพิ่มความแม่นยำในการทำนายได้ถึง 10% จากนั้น Ricardo Sousa และ Joao Gama [5] ทำการเปรียบเทียบระหว่างวิธีการเรียนรู้ร่วมและวิธีการเรียนรู้ด้วยตนเอง (self-learning) สำหรับการถดถอยที่มีเป้าหมายเดียวในข้อมูลแบบสตรีม ด้วยกฎการปรับโมเดลสุ่ม (Random Adaptive Model Rules) เปรียบเทียบผลจากสถานการณ์ที่ไม่นำเอาข้อมูลที่ไม่มีป้ายกำกับเข้าไปในโมเดล และสถานการณ์ที่นำเอาข้อมูลที่ไม่มีป้ายกำกับเข้าไปเพื่อปรับปรับการถดถอย ซึ่งผลลัพธ์แสดงหลักฐานที่ทำให้ประสิทธิภาพที่ดีขึ้น ในเรื่องช่วยลดความคลาดเคลื่อนในข้อมูลแบบสตรีมระดับสูง นอกจากนี้ยังมีงานวิจัยของ Fan Ma และคณะ [6] ได้นำเสนออัลกอริทึมโค เทรนนิงแบบใหม่ที่ชื่อว่า SPaCo (Self-Paced Co-training) การเรียนรู้ร่วมด้วยตัวเอง แก้ไขปัญหาการกำกับค่าของตัวอย่างที่ไม่มีป้ายกำกับที่ไม่ถูกต้องในรอบการฝึกขั้นต้น โดยการแทนที่ของตัวอย่าง (เลือกตัวอย่างเข้าและออก) ซึ่งสามารถเพิ่มประสิทธิภาพของโมเดลได้ดียิ่งขึ้น

งานวิจัยที่ใช้เทคนิคโค เทรนนิงจะเน้นที่การจำแนกประเภทข้อมูล (classification) มากกว่าการประมาณค่า (regression) ซึ่งในหลายๆ งานการประมาณค่าก็เป็นสิ่งจำเป็น ดังนั้น Zhi-Hua Zhou และ Ming Li [7] ได้ทำการวิจัยและนำเสนอวิธีการ COREG (Co-Training Regressors) การเรียนรู้ร่วมแบบกึ่งถดถอย โดยจะทำการเลือกตัวอย่างที่ไม่มีป้ายกำกับมากำกับค่าผ่านโมเดลที่ให้ค่าความคาดเคลื่อนน้อยที่สุดทั้งสองโมเดล และการพยากรณ์ในขั้นสุดท้าย โดยการหาค่าเฉลี่ยของสมการถดถอยที่สร้างขึ้นทั้งสองตัว ซึ่งอัลกอริทึมนี้สามารถใช้ประโยชน์จากข้อมูลที่ไม่มีป้ายกำกับเพื่อปรับปรุงการพยากรณ์แบบถดถอย วิธีการนี้ได้มีการนำไปใช้อย่างแพร่หลายแต่ใช้เวลาการทำงานที่นานเนื่องจากในการเลือกข้อมูลที่ไม่มีป้ายกำกับจำเป็นต้องทดสอบกับข้อมูลที่ไม่มีป้ายกำกับทีละตัวอย่าง

เนื่องจากวิธีการของ COREG ใช้เวลาการทำงานนาน คณะผู้วิจัยจึงได้นำเสนอวิธีการปรับปรุงประสิทธิภาพของโมเดลพยากรณ์ COREG ด้วยวิธีการเลือกข้อมูลที่ไม่มีป้ายกำกับอย่างอัตโนมัติ (AU-COREG) เพื่อการสร้างโมเดลการ

เรียนรู้ร่วมแบบกึ่งมีผู้สอน สำหรับข้อมูลที่มีป้ายกำกับที่มีสัดส่วนน้อย เมื่อเทียบกับข้อมูลที่ไม่มีป้ายกำกับ โดยเพื่อลดความคลาดเคลื่อนจากการพยากรณ์ และลดระยะเวลาในการสร้างโมเดล เพื่อรองรับกับข้อมูลที่หลากหลายและมีจำนวนมาก

2. แนวคิด / วิธีการที่นำเสนอ

2.1 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือการสร้างโมเดล คือ Rapid Miner Studio เวอร์ชั่น 9.2.000 บนเครื่องลูกข่าย (Client) Processor 2.7 GHz Intel Core i5 RAM 8 GB 1867 MHz DDR3 และ Rapid Miner Server CPU 4 cores RAM 8GB และ HDD 120 GB

2.2 ศึกษาและรวบรวมข้อมูล

ข้อมูลที่ใช้ในการวิจัยมีทั้งหมด 3 ชุดข้อมูล

2.2.1 ชุดข้อมูลการพยากรณ์เงินเดือน (Job Salary Prediction) ข้อมูลโฆษณาประกาศสมัครงานในประเทศอังกฤษ จัดทำโดย Adzuna (ข้อมูลที่ใช้ในการแข่งขัน Job Salary Prediction : Kaggle) ซึ่งเป็นข้อมูลที่ซึ่งประกอบด้วยข้อมูลดังนี้

- เงินเดือน (salary: US dollar)
- ชื่อตำแหน่ง (title)
- สถานที่ตั้งบริษัท (location)
- ประเภทการจ้างงาน (contact_type)
- สัญญาการจ้างงาน (contact time)
- บริษัท (company)
- ประเภทธุรกิจ (business_case)
- แหล่งข้อมูล (source)

2.2.2 ชุดข้อมูลการพยากรณ์ปริมาณการจราจร (Metro Interstate Traffic Volume Data Set)
ข้อมูลปริมาณการจราจรระหว่างรัฐโดยรถไฟฟ้ารายชั่วโมง จากทิศตะวันตกของมินนิโซตา DoT ATR สถานี 301 ซึ่งอยู่กลางระหว่างมินนิอาโปลิสและเซนต์พอลมินนิโซตา (UCI Machine Learning Repository) โดยมีข้อมูลดังนี้

- ปริมาณการจราจร (traffic_volume)
- วันหยุด (holiday)
- อุณหภูมิโดยเฉลี่ย (temp เป็น องศาเคลวิน)
- ปริมาณน้ำฝน (rain_1h (mm))
- ปริมาณหิมะ (snow_1h (mm))
- ร้อยละปริมาณหมอกที่ปกคลุม (clouds_all)
- ประเภทลักษณะอากาศ (weather_main)
- ลักษณะอากาศ (weather_description)

2.2.3 ชุดข้อมูลการพยากรณ์พลังงานไฟฟ้า (Combined Cycle Power Plant Data Set)
ข้อมูลที่รวบรวมจากโรงไฟฟ้าพลังความร้อนร่วม ในช่วง 6 ปี (2549-2554) (UCI Machine Learning Repository) โดยมีข้อมูลดังนี้

- พลังงานไฟฟ้า (EP)
- อุณหภูมิ (AT)
- ความดันบรรยากาศ (AP)
- ความชื้นสัมพัทธ์ (RH)
- ไอเสีย (V)

ตารางที่ 1. แสดงลักษณะของชุดข้อมูลและการสุ่มข้อมูล

ชุดข้อมูล	จำนวนแอตทริบิวต์	ประเภทข้อมูล	ขนาดข้อมูล (แถว)	ขนาดข้อมูลที่ใช้ (แถว)
1	8	Nominal	244,768	5,000
2	8	Nominal, Real	48,204	5,000
3	5	Real	9,568	5,000

2.3 วิธีการสุ่มข้อมูล

การวิจัยครั้งนี้ใช้วิธีการสุ่มข้อมูล 2 วิธีดังนี้

2.3.1 วิธีการสุ่มตัวอย่างแบบชั้นภูมิ (Stratified Random Sampling) ซึ่งเป็นการสุ่มตัวอย่างจากประชากรที่มีจำนวนมาก โดยประชากรจะถูกแบ่งออกเป็นชั้นภูมิตามลักษณะอย่างใดอย่างหนึ่ง โดยไม่ให้มีหน่วยซ้ำกัน ซึ่งในชั้นภูมิเดียวกันจะประกอบด้วยหน่วยที่มีลักษณะคล้ายคลึงกันมากที่สุด และแตกต่างระหว่างชั้นภูมิมากที่สุด

2.3.2 วิธีการสุ่มตัวอย่างแบบง่าย (Simple random sampling) เป็นการสุ่มตัวอย่างโดยถือว่าทุกๆ หน่วยหรือทุกๆ สมาชิกในประชากรมีโอกาสจะถูกเลือกเท่าๆ กัน

2.4 การจัดการข้อมูล

ชุดข้อมูลที่ 2 และ 3 มีข้อมูลประเภทตัวเลขและแต่ละแอตทริบิวต์มีขนาดข้อมูลที่แตกต่างกัน ดังนั้นการแปลงข้อมูลให้อยู่ในสเกลเดียวกัน (Normalize) จึงถูกนำมาใช้ ในงานวิจัยนี้เลือกใช้วิธี Z-transformation ซึ่งทำให้เป็นค่ามาตรฐานและการกระจายของข้อมูลมีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็นหนึ่ง ดังแสดงในสมการต่อไปนี้

$$Z_i = (X_i - Mean) / Standard\ Deviation$$

ตารางที่ 2. แสดงค่าสถิติของข้อมูลชุดที่ 2

แอตทริบิวต์	ค่าน้อยที่สุด	ค่ามากที่สุด	ค่าเฉลี่ย
Temp (K)	245.62	308.43	281.24
rain_1h (mm)	0.00	25.57	0.147
snow_1h (mm)	0.00	0.44	0.00
clouds_all (%)	0.00	100.00	48.88

ตารางที่ 3. แสดงค่าสถิติของข้อมูลชุดที่ 3

แอตทริบิวต์	ค่าน้อยที่สุด	ค่ามากที่สุด	ค่าเฉลี่ย
AT (C)	1.81	35.56	19.58
V(cm Hg)	25.36	81.56	54.25
AP (milibar)	992.90	1,033.30	1,013.36
RH (MW)	25.56	100.16	48.88

2.5 ทฤษฎีที่เกี่ยวข้อง

2.5.1 การจำแนกข้อมูลด้วยวิธีการเพื่อนบ้านใกล้ที่สุดเคตัว (K-Nearest Neighbors: K-NN)

เป็นวิธีการใช้จำแนกหรือทำนายข้อมูลด้วยการเรียนรู้จากข้อมูล (Supervised Learning) ที่มีป้ายกำกับ (Labeled Data) โดยทำการเปรียบเทียบความคล้ายคลึงกับข้อมูลที่มีอยู่มากที่สุดเคตัว และกำหนดกลุ่มให้กับข้อมูลที่ไม่มีป้ายกำกับตามสมาชิกส่วนใหญ่ของกลุ่ม

(1) วิธีการเปรียบเทียบความคล้ายคลึงจะถูกกำหนดในรูปแบบของระยะทางในหลายๆ มิติ ตามขนาดของแอตทริบิวต์ในชุดข้อมูลการเรียนรู้

ขั้นตอนการหาเพื่อนบ้านเคตัว มีดังต่อไปนี้

1) กำหนดค่าเค (k) โดยปกติจะนิยมเป็นจำนวนคี่

2) คำนวณหาความคล้ายคลึงของข้อมูลที่ไม่มีป้ายกำกับกับข้อมูลที่มีป้ายกำกับ (ระยะทาง)

3) เรียงลำดับความคล้ายคลึงและเลือกข้อมูลตัวอย่างที่มีความคล้ายคลึงมากที่สุดเคตัว

4) พิจารณาข้อมูลตัวอย่างทั้งเคตัว เพื่อจัดจำแนกหรือทำนายข้อมูลแต่ละตัวว่าถูกจัดเป็นกลุ่มใด

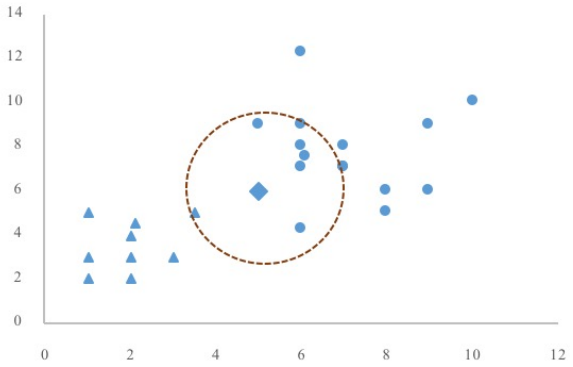
5) กำหนดกลุ่มใหม่ให้กับข้อมูลที่ไม่มีป้ายกำกับด้วยกลุ่มข้อมูลที่มีตัวอย่างมากที่สุดจากค่าเค

การวัดความคล้ายคลึงด้วยวิธีการวัดระยะห่างยูคลิเดียน (Euclidean Distant) เป็นการวัดระยะห่างระหว่าง 2 จุดในแนวเส้นตรงที่ได้มาจากทฤษฎีพีทาโกรัส ระยะห่างยูคลิเดียนระหว่างจุด p และ q คำนวณได้จาก

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

หากแอตทริบิวต์มีข้อมูลแบบอัตภาคชั้น (Nominal) การวัดระยะทางหากค่าเหมือนกันระยะทางจะเป็นศูนย์ หากต่างกันจะเป็นค่าเป็นอย่างอื่น

ตัวอย่างการจำแนกข้อมูลด้วยเพื่อนบ้านทั้ง n ตัว ดังรูปที่ 1



รูปที่ 1. แสดงตัวอย่างจำแนกข้อมูลด้วยเพื่อนบ้าน 7 ตัว

2.5.2 การจัดกลุ่ม (Clustering)

การจัดกลุ่มข้อมูลจากความคล้ายคลึงกัน (Clustering) เป็นเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยพยายามให้ระยะห่างของสิ่งที่อยู่ในกลุ่มเดียวกันให้อยู่ใกล้กันมากที่สุด (Minimize Intra-Cluster Distance) และสิ่งที่อยู่ต่างกลุ่มกันจะมีระยะห่างแตกต่างกันมากที่สุด (Maximize Inter-Cluster Distance) หรืออาจกล่าวได้ว่ากลุ่มข้อมูลที่มีคุณสมบัติและ/หรือคุณลักษณะที่คล้ายคลึงกันควรอยู่ในกลุ่มข้อมูลเดียวกัน และข้อมูลที่มีคุณสมบัติและ/หรือคุณสมบัติที่แตกต่างกันอย่างมากควรอยู่ต่างกลุ่มกัน ดังตัวอย่างการจัดกลุ่ม n กลุ่มดังรูปที่ 2

2.5.2.1 การจัดกลุ่มด้วยเทคนิค K-Mean

เป็นวิธีการจัดกลุ่มที่วิเคราะห์กลุ่มแบบไม่เป็นขั้นตอนหรือการแบ่งส่วน (Partitioning) ออกเป็นเคกลุ่ม และแทนค่าแต่ละกลุ่มด้วยค่าเฉลี่ยของกลุ่ม หรือเรียกว่าจุดศูนย์กลาง (Centroid) ของกุล่ม

2.5.2.2 การจัดกลุ่มด้วยเทคนิค K-Medoids

เป็นวิธีการจัดกลุ่มที่วิเคราะห์กลุ่มที่เหมือนกับเทคนิค K-Mean แต่การคำนวณจุดศูนย์กลางของกลุ่มจะแทนที่ด้วยค่าของข้อมูลจริงๆ ที่อยู่ในกลุ่มนั้น

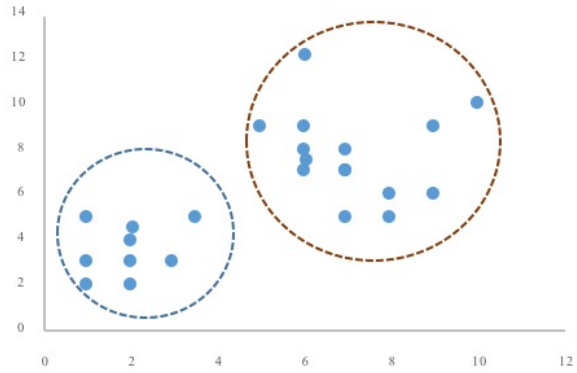
ขั้นตอนการจัดกลุ่มด้วยเทคนิค K-Mean และ K-Medoids

1) กำหนดค่าเริ่มต้นจำนวนเคกลุ่ม และกำหนดจุดศูนย์กลางเริ่มต้นทั้งเคจุด

2) พิจารณาข้อมูลที่เหลือเพื่อจัดเข้ากลุ่ม โดยการหาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลาง โดยหากข้อมูลใดใกล้ค่าจุดศูนย์กลางตัวไหน จะทำการจัดเข้ากลุ่มนั้น

3) หาจุดศูนย์กลางของแต่ละกลุ่มโดย

3.1) เทคนิค K-Mean จะทำการหาค่าเฉลี่ย (Mean) ของแต่ละกลุ่มใหม่ และกำหนดให้เป็นจุดศูนย์กลางของกลุ่มใหม่



รูปที่ 2. แสดงตัวอย่างการจัดกลุ่ม 2 กลุ่ม

3.2) เทคนิค K-Medoids จะทำการหาค่ากลางค่าใหม่ของกลุ่ม แล้วเปรียบเทียบค่าความหนาแน่น เพื่อเลือกข้อมูลที่เป็นค่ากลางที่ทำให้ค่าความหนาแน่นต่ำที่สุด

4) ทำซ้ำข้อ 2) จนกระทั่งค่าเฉลี่ยหรือจุดศูนย์กลางใหม่ในแต่ละกลุ่มจะไม่มีการเปลี่ยนแปลง

2.5.3. การวิเคราะห์การถดถอย (Regression Analysis)

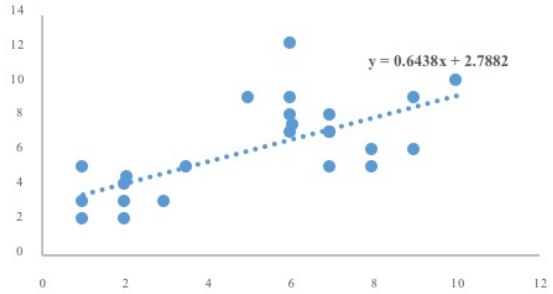
การวิเคราะห์การถดถอยเป็นเทคนิคการสร้างตัวแบบจากความสัมพันธ์ระหว่างข้อมูล 2 ตัวเป็นต้นไป หรือทำนายข้อมูลตัวหนึ่ง (ข้อมูลเชิงปริมาณ) จากข้อมูลอีกตัว (หรือมากกว่า) สามารถเขียนสมการอย่างง่ายได้ดังสมการที่ (2) ดังนี้

$$y = \alpha + \beta X + \varepsilon \qquad (2)$$

โดยที่ α เป็นค่าคงที่ที่ไม่ทราบค่าของสมการถดถอย β เป็นสัมประสิทธิ์ถดถอย (Regression Coefficient) เป็นอัตราการเปลี่ยนแปลงของค่า X ต่อค่า y และ ε เป็นค่าความคลาดเคลื่อนระหว่างค่าพยากรณ์ และค่าจริง y

ตัวอย่างการวิเคราะห์การถดถอยอย่างง่าย ดังรูปที่ 3 ซึ่งเรียกว่าตัวแบบถดถอยเชิงเส้นอย่างง่าย และมีตัววัดประสิทธิภาพของตัวแบบที่เป็นที่นิยม ได้แก่ สัมประสิทธิ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Square Error: RMSE) ดังสมการที่ (3) ซึ่งหมายถึงค่าความเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนจากการทำนาย

$$RMSE = \sqrt{\sum_{i=1}^{n}(prediction - actual)^2} \quad (3)$$



รูปที่ 3. แสดงการวิเคราะห์การถดถอยอย่างง่าย

2.6 ขั้นตอนการสร้างโมเดล AU-COREG

2.6.1 การแบ่งข้อมูล

ทำการแบ่งข้อมูลออกเป็น 2 ส่วน 1) ข้อมูลที่กำหนดให้มีป้ายกำกับ (Labeled Data) เพื่อใช้ในการสร้างและทดสอบโมเดล 2) ข้อมูลที่กำหนดให้ไม่มีป้ายกำกับ (Unlabeled Data) โดยการใช้การสุ่มข้อมูลแบบชั้นภูมิ ดังตารางที่ 4

ตารางที่ 4. แสดงการแบ่งข้อมูลเพื่อสร้างโมเดลขั้นแรก

ส่วนของข้อมูล	สัดส่วน	ขนาดข้อมูล (แถว)
ข้อมูลสำหรับการทดสอบโมเดลสุดท้าย (Testing Data)	25.0%	1,250
ข้อมูลสำหรับการสร้างโมเดลขั้นแรก (Initial Training Data) : L	7.5%	375
ข้อมูลที่ไม่มีป้ายกำกับ เพื่อเป็นส่วนที่เลือกข้อมูลเข้าใช้ในการสร้างโมเดลเพิ่มเติม : U'	2.0%	100 หรือ200
ข้อมูลที่ไม่มีป้ายกำกับ: U	65.5%	3,275

2.6.2 ขั้นตอนสร้างโมเดลขั้นต้นจากข้อมูลที่มีป้ายกำกับ (Labeled Data)

กำหนดให้ L_1 เป็นข้อมูลสำหรับสร้างโมเดลที่ 1 ในขั้นแรก และ L_2 เป็นข้อมูลสำหรับสร้างโมเดลที่ 2 ในขั้นแรก ซึ่งให้ L_1 และ L_2 มีค่าเท่ากับ L และสร้างโมเดล 2 โมเดลสมการที่ (4) และ (5)

$$h_1 \leftarrow kNN\,(L_1,k,p_1) \quad (4)$$

$$h_2 \leftarrow kNN\,(L_2,k,p_2) \quad (5)$$

2.6.3 กำหนดจำนวนรอบการทำซ้ำ (Iteration)

ในงานวิจัยกำหนดจำนวนรอบการทำซ้ำเป็น 100 รอบ (T=100) เพื่อเลือกข้อมูลในส่วน U' เข้าไปเป็นข้อมูล Training รอบละ 1 ตัวอย่าง โดยหลักการเลือกตัวอย่างเข้าไปนั้น จะทำการเลือกตัวอย่างที่ช่วยลดความคาดเคลื่อนจากการเพิ่มข้อมูลเข้าไปที่มากที่สุด จนกระทั่งครบ 100 รอบ หรือข้อมูลที่เพิ่มเข้าไปนั้นไม่สามารถช่วยลดความคาดเคลื่อนได้ หลังจากการเลือกข้อมูลเพิ่มเข้าไปทุกครั้ง (จาก U' ไป L) ต้องทำการสุ่มข้อมูลที่ไม่มีป้ายกำกับ (U) เข้าไปในส่วนของกลุ่มข้อมูล (U') ทุกครั้ง

1) พยากรณ์ข้อมูล U' ด้วยโมเดลที่ 1 (h_1) แล้วทำการจัดคลัสเตอร์ (Cluster) ข้อมูลที่ถูกกำกับค่า ด้วยเทคนิค K-Medoids โดยกำหนดจำนวนคลัสเตอร์เท่ากับ 2 (ในการทดลองจะทำการปรับค่าจำนวนคลัสเตอร์ตั้งแต่ 2 จนถึง 10 คลัสเตอร์)

- คลัสเตอร์ 1 ทำการเลือกสมาชิกที่เป็นตัวแทนคลัสเตอร์จากนั้นหาข้อมูลที่มีป้ายกำกับ (L_1) ที่มีระยะห่างกับตัวแทนของคลัสเตอร์ที่น้อยที่สุด (เพื่อนบ้าน) 5 ตัวอย่าง (มีความคล้ายคลึงกันมากที่สุด) โดยวัดระยะทางตามวิธีที่กำหนดในตารางที่ 5
- เพิ่มตัวแทนของคลัสเตอร์ที่ 1 เข้าไปใน L_1 เพื่อสร้างโมเดลใหม่ นำโมเดลที่ได้ไปทดสอบประสิทธิภาพของโมเดล ด้วยสัมประสิทธิ์ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง RMSE (Root Mean Square Error) กับเพื่อนบ้านทั้ง 5 ตัว
- ทำตามขั้นตอนข้างต้นกับคลัสเตอร์ที่เหลือจนครบ หากพบ RMSE มากกว่า 0 ให้เลือกตัวแทนของคลัสเตอร์มีค่า RMSE มีค่ามากที่สุดออกจาก U' (π_1)

2) ทำตามขั้นตอนข้างต้นกับโมเดลที่ 2 และเลือกตัวแทนจากคลัสเตอร์ที่ได้จากการพยากรณ์ด้วยโมเดลที่ 2 ที่ทำ

ให้ค่า RMSE มีค่ามากที่สุดออกจาก U' (π_2)

3) นำ π_1 เพิ่มเข้าไปใน L_1และนำ π_2 เพิ่มเข้าไปใน L_1 ดังสมการที่ (6) และ (7)

$$L_1 \leftarrow L_1 \cup \pi_2 \quad (6)$$

$$L_2 \leftarrow L_2 \cup \pi_1 \quad (7)$$

4) สร้างโมเดล 1 และ 2 จากข้อมูล L_1 และ L_2 ใหม่ และหากพบว่า L_1 และ L_2 ไม่เปลี่ยนแปลง ให้หยุดดำเนินการ

5) สุ่มเลือกตัวอย่างจาก U เพิ่มเข้ามาใน U' ให้ครบจำนวน

6) ทำตามขั้นตอนข้างต้น จนครบรอบจำนวนการทำซ้ำ (100 รอบ)

7) สร้างโมเดล AU-COREG และทดสอบประสิทธิภาพของโมเดล ดังสมการที่ (8)

$$h^*(x) = (h_1(x) + h_2(x)) / 2 \quad (8)$$

2.6.4 กำหนดจำนวนรอบการทำซ้ำ (Iteration) 200 รอบ กำหนดจำนวนรอบการทำซ้ำเป็น 200 รอบ (T =200) และทำตามขั้นตอน 2.6.3. ทั้งหมด เพื่อสร้างโมเดลใหม่มาเปรียบเทียบ

2.6.5 ทำตามขั้นตอนที่ 2.6.3 โดยเปลี่ยนเทคนิคการทำ Cluster จาก K-Medoids เป็น K-Mean และทำการเลือกสมาชิกใน Cluster โดยการระบุค่า เพื่อเป็นตัวแทน Cluster และทำการทดสอบประสิทธิภาพของโมเดลที่ได้ บันทึกผลที่ได้ จากนั้นให้ดำเนินตามขั้นตอนเดิม และเลือกสมาชิกใหม่ใน Cluster ให้เป็นตัวแทนกลุ่ม จนครบ 3 รอบ บันทึกผลเพื่อนำมาเปรียบเทียบ

2.6.6 ทำตามขั้นตอนทั้งหมดกับชุดข้อมูลทั้ง 3 โดยโมเดลที่ใช้คือ kNN โดยวิธีการวัดระยะทาง และ K ให้เหมาะสมกับประเภทของข้อมูล ดังตารางที่ 5

3. ผลการทดลองและคำอธิบายรายละเอียด

ผลการเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่า RMSE ของโมเดล Self-training, COREG และ AU-COREG ที่มีการจัดกลุ่มตั้งแต่ 2 จนถึง 10 กลุ่ม และที่มี U' เท่ากับ 100 และ 200 และระยะเวลาที่ใช้ในการสร้างโมเดลของข้อมูลชุดที่ 1 ดังตารางที่ 6-10

ตารางที่ 5. แสดงโมเดลและวิธีการวัดระยะทาง

ชุดข้อมูล	โมเดล	วิธีวัดระยะทาง (p_i)	K
1	kNN1	Mix Euclidean Distance	3
	kNN2	Nominal Distance	3
2	kNN1	Mix Euclidean Distance	5
	kNN2	Mix Euclidean Distance	9
3	kNN1	Euclidean Distance	5
	kNN2	Correlation Similarity	5

ตารางที่ 6. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล SELF-TRAINNING และ COREG ของข้อมูลชุดที่ 1

MODEL	Training Set = 100		Training Set = 200	
	RMSE	TIME	RMSE	TIME
SELF TRAINING	17,299.17		16,926.85	
COREG	16,126.82	276	16,104.93	537

ตารางที่ 7. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Medoids ของข้อมูลชุดที่ 1

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEDOIDS	2 CLUSTER	16,180.82	21	16,003.33	23
	3 CLUSTER	16,129.74	23	16,165.40	25
	4 CLUSTER	16,282.24	25	16,107.52	27

ตารางที่ 7. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Medoids ของข้อมูลชุดที่ 1 (ต่อ)

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEDOIDS	5 CLUSTER	16,279.32	28	16,257.53	31
	6 CLUSTER	16,047.12	31	16,301.17	34
	7 CLUSTER	16,279.32	34	16,128.76	37
	8 CLUSTER	16,251.33	37	16,037.87	40
	9 CLUSTER	16,279.38	40	16,124.75	43
	10 CLUSTER	16,313.28	42	16,141.95	46

ตารางที่ 8. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 1992 ของข้อมูลชุดที่ 1

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 1992	2 CLUSTER	16,195.76	17	16,368.06	18
	3 CLUSTER	16,095.27	19	16,019.71	19
	4 CLUSTER	16,241.59	21	16,303.79	20
	5 CLUSTER	16,317.01	23	16,284.17	21
	6 CLUSTER	16,186.70	22	16,159.21	22
	7 CLUSTER	16,243.52	23	16,212.69	24
	8 CLUSTER	16,110.72	25	16,132.22	24
	9 CLUSTER	16,223.59	26	16,152.23	24
	10 CLUSTER	16,186.64	26	16,363.86	25

ตารางที่ 9. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 100 ของข้อมูลชุดที่ 1

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 100	2 CLUSTER	16,203.94	18	16,119.15	18
	3 CLUSTER	16,095.27	19	16,293.71	19
	4 CLUSTER	16,309.88	21	16,174.86	20
	5 CLUSTER	16,238.26	22	16,221.03	21
	6 CLUSTER	16,354.49	22	16,156.71	22
	7 CLUSTER	16,265.92	24	16,010.29	23
	8 CLUSTER	16,242.97	26	16,264.19	24
	9 CLUSTER	16,067.44	26	16,079.38	24
	10 CLUSTER	16,180.36	26	16,246.13	25

ตารางที่ 10. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 3645 ของข้อมูลชุดที่ 1

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 3745	2 CLUSTER	16,112.53	18	16,099.72	18
	3 CLUSTER	16,331.73	21	16,303.51	20
	4 CLUSTER	16,275.75	21	16,245.74	21
	5 CLUSTER	16,193.15	22	16,202.93	21
	6 CLUSTER	16,092.36	23	16,074.08	22
	7 CLUSTER	16,193.15	25	16,187.52	22
	8 CLUSTER	16,208.41	25	16,216.09	24
	9 CLUSTER	16,166.59	27	16,084.66	26
	10 CLUSTER	16,188.76	25	16,361.12	26

3.1 ผลการพยากรณ์ข้อมูลชุดที่ 1

เมื่อทำการพยากรณ์ข้อมูลชุดทดสอบ จากตารางที่ 6-10 แสดงให้เห็นว่า

- โมเดล kNN ที่ได้จากการเรียนรู้จากชุดข้อมูลฝึก ขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 17,299.17 และ 16,926.85 ตามลำดับ
- โมเดล COREG ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 16,126.82 และ 16,104.93 ตามลำดับ
- โมเดล AU-COREG (U'= 100) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Medoids) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 16,047.12 ซึ่งได้จากการทำ Cluster จำนวน 6 Cluster ลดลง 0.49% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 276 นาที เหลือเพียง 31 นาที หรือลดลงถึง 89% และโมเดล AU-COREG (U'= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 16,003.33 ซึ่งได้จากการทำ Cluster จำนวน 2 Cluster ลดลง 0.14% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 527 นาที เหลือเพียง 23 นาที หรือลดลงถึง 96%
- โมเดล AU-COREG (U'= 200) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Mean โดยเลือกตัวแทน Cluster จาก Seed 1992 100 และ 3645) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 16,067.44 ซึ่งได้จากการทำ Cluster จำนวน 9 Cluster (Seed 100) ลดลง 0.37% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 276 นาที เหลือเพียง 26 นาที หรือลดลงถึง 91% และโมเดล AU-COREG (U'= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 16,010.29 ซึ่งได้จากการทำ Cluster จำนวน 7 Cluster (Seed 100) ลดลง 0.59% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 527 นาที เหลือเพียง 23 นาที หรือลดลงถึง 96%

3.2 ผลการพยากรณ์ข้อมูลชุดที่ 2

ผลการเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่า RMSE ของโมเดล Self-training, COREG และ AU-COREG ที่มีการจัดกลุ่มตั้งแต่ 2 จนถึง 10 กลุ่ม และที่มี U' เท่ากับ 100 และ 200 และระยะเวลาที่ใช้ในการสร้างโมเดลของข้อมูลชุดที่ 2 ดังตารางที่ 11-15 สรุปได้ดังนี้

ตารางที่ 11. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล SELF-TRAINNING และ COREG ของข้อมูลชุดที่ 2

MODEL	Training Set = 100		Training Set = 200	
	RMSE	TIME	RMSE	TIME
SELF TRAINING	2,199.14		2,120.78	
COREG	2,113.46	215	2,110.38	428

ตารางที่ 12. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Medoids ของข้อมูลชุดที่ 2

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEDOIDS	2 CLUSTER	2,121.32	12	2,120.32	17
	3 CLUSTER	2,122.24	13	2,115.78	17
	4 CLUSTER	2,119.87	15	2,108.29	19
	5 CLUSTER	2,115.01	22	2,111.66	24
	6 CLUSTER	2,118.39	25	2,114.48	28
	7 CLUSTER	2,113.61	28	2,113.61	31
	8 CLUSTER	2,112.36	31	2,115.24	33
	9 CLUSTER	2,110.84	34	2,111.19	36
	10 CLUSTER	2,110.05	36	2,120.08	39

ตารางที่ 13. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 1992 ของข้อมูลชุดที่ 2

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 1992	2 CLUSTER	2,118.82	12	2,117.95	14
	3 CLUSTER	2,114.02	12	2,109.64	17
	4 CLUSTER	2,114.52	15	2,120.36	17
	5 CLUSTER	2,117.16	20	2,118.93	20
	6 CLUSTER	2,110.68	23	2,117.87	23
	7 CLUSTER	2,112.78	26	2,110.86	26
	8 CLUSTER	2,113.20	28	2,115.21	29
	9 CLUSTER	2,117.96	31	2,111.94	31
	10 CLUSTER	2,111.36	33	2,107.28	33

ตารางที่ 14. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 100 ของข้อมูลชุดที่ 2

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 100	2 CLUSTER	2,114.02	12	2,109.80	12
	3 CLUSTER	2,119.22	14	2,120.06	15
	4 CLUSTER	2,117.78	17	2,123.85	17
	5 CLUSTER	2,115.14	20	2,116.97	20
	6 CLUSTER	2,110.68	23	2,109.16	22
	7 CLUSTER	2,116.55	25	2,114.88	25
	8 CLUSTER	2,115.95	28	2,111.49	28
	9 CLUSTER	2,117.86	30	2,109.37	30
	10 CLUSTER	2,111.38	33	2,117.51	33

ตารางที่ 15. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 3645 ของข้อมูลชุดที่ 2

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 3645	2 CLUSTER	2,114.02	12	2,125.11	12
	3 CLUSTER	2,119.22	14	2,113.98	15
	4 CLUSTER	2,117.78	17	2,108.27	17
	5 CLUSTER	2,115.14	20	2,119.51	21
	6 CLUSTER	2,110.68	23	2,114.91	23
	7 CLUSTER	2,116.55	25	2,116.81	26
	8 CLUSTER	2,115.95	28	2,116.16	29
	9 CLUSTER	2,117.86	30	2,117.70	31
	10 CLUSTER	2,111.38	33	2,115.96	34

- โมเดล kNN ที่ได้จากการเรียนรู้จากชุดข้อมูลฝึกขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 2,199.14 และ 2,120.78 ตามลำดับ

- โมเดล COREG ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 2,113.46 และ 2,110.38 ตามลำดับ

- โมเดล AU-COREG (U'= 100) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Medoids) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 2,110.05 ซึ่งได้จากการทำ Cluster จำนวน 10 Cluster ลดลง 0.16% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 215 นาที เหลือเพียง 36 นาที หรือลดลงถึง 83% และโมเดล AU-COREG (U'= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 2,108.29 ซึ่งได้จากการทำ Cluster จำนวน 4 Cluster ลดลง 0.10% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 428 นาที เหลือเพียง 19 นาที หรือลดลงถึง 96%

- โมเดล AU-COREG (*U'*= 200) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Mean โดยเลือกตัวแทน Cluster จาก Seed 1992 100 และ 3645) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 2,107.95 ซึ่งได้จากการทำ Cluster จำนวน 10 Cluster (Seed 3645) ลดลง 0.26% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดลลดลงจาก 215 นาที เหลือเพียง 34 นาที หรือลดลงถึง 84% และโมเดล AU-COREG (*U'*= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 2,107.28 ซึ่งได้จากการทำ Cluster จำนวน 10 Cluster (Seed 1992) ลดลง 0.15% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 428 นาที เหลือเพียง 33 นาที หรือลดลงถึง 92%

3.3 ผลการพยากรณ์ข้อมูลชุดที่ 3

ผลการเปรียบเทียบประสิทธิภาพของโมเดลด้วยค่า RMSE ของโมเดล Self-training, COREG และ AU-COREG ที่มีการจัดกลุ่มตั้งแต่ 2 จนถึง 10 กลุ่ม และที่มี *U'* เท่ากับ 100 และ 200 และระยะเวลาที่ใช้ในการสร้างโมเดลของข้อมูลชุดที่ 3 ดังตารางที่ 16-20 สรุปได้ดังนี้

ตารางที่ 16. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล SELF-TRAINNING และ COREG ของข้อมูลชุดที่ 3

MODEL	Training Set = 100		Training Set = 200	
	RMSE	TIME	RMSE	TIME
SELF TRAINING	10.76		10.67	
COREG	7.90	217	7.70	419

ตารางที่ 17. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Medoids ของข้อมูลชุดที่ 3

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEDOIDS	2 CLUSTER	7.63	12	7.74	12
	3 CLUSTER	7.66	15	7.70	15

ตารางที่ 17. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Medoids ของข้อมูลชุดที่ 3 (ต่อ)

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEDOIDS	4 CLUSTER	7.70	17	7.75	18
	5 CLUSTER	7.75	20	7.66	21
	6 CLUSTER	7.74	23	7.84	25
	7 CLUSTER	7.65	26	7.65	28
	8 CLUSTER	7.75	28	7.66	30
	9 CLUSTER	7.81	31	7.79	34
	10 CLUSTER	7.74	33	7.75	36

ตารางที่ 18. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 1992 ของข้อมูลชุดที่ 3

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 1992	2 CLUSTER	7.73	12	7.82	12
	3 CLUSTER	7.68	14	7.68	15
	4 CLUSTER	7.71	17	7.78	17
	5 CLUSTER	7.78	21	7.78	20
	6 CLUSTER	7.66	23	7.74	22
	7 CLUSTER	7.77	26	7.78	26
	8 CLUSTER	7.81	28	7.79	28
	9 CLUSTER	7.80	31	7.91	30
	10 CLUSTER	7.84	33	7.86	33

ตารางที่ 19. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 100 ของข้อมูลชุดที่ 3

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 100	2 CLUSTER	7.63	12	7.72	13
	3 CLUSTER	7.64	15	7.70	15
	4 CLUSTER	7.72	18	7.75	18
	5 CLUSTER	7.75	20	7.78	21
	6 CLUSTER	7.75	23	7.71	23
	7 CLUSTER	7.85	25	7.71	25
	8 CLUSTER	7.63	27	7.76	28
	9 CLUSTER	7.78	30	7.81	31
	10 CLUSTER	7.76	33	7.89	33

ตารางที่ 20. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU- COREG ที่จัด Cluster ด้วยวิธี K-Mean และ Seed = 3645 ของข้อมูลชุดที่ 3

MODEL		U'100		U'200	
AU-COREG		RMSE	TIME	RMSE	TIME
K-MEAN , SEED = 3645	2 CLUSTER	7.66	12	7.70	12
	3 CLUSTER	7.61	15	7.67	15
	4 CLUSTER	7.79	18	7.71	17
	5 CLUSTER	7.78	21	7.70	20
	6 CLUSTER	7.65	24	7.76	23
	7 CLUSTER	7.77	26	7.89	25
	8 CLUSTER	7.70	28	7.87	28
	9 CLUSTER	7.71	31	7.79	31
	10 CLUSTER	7.78	34	7.90	33

- โมเดล kNN ที่ได้จากการเรียนรู้จากชุดข้อมูลฝึกขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 10.76 และ 10.67 ตามลำดับ
- โมเดล COREG ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 100 และ 200 ตัวอย่าง จะได้ค่า RMSE เท่ากับ 7.90 และ 7.70 ตามลำดับ
- โมเดล AU-COREG (*U'*= 100) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Medoids) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 7.63 ซึ่งได้จากการทำ Cluster จำนวน 2 Cluster ลดลง 1.80% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 217 นาที เหลือเพียง 12 นาที หรือลดลงถึง 94% และโมเดล AU-COREG (*U'*= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 7.65 ซึ่งได้จากการทำ Cluster จำนวน 7 Cluster ลดลง 3.21% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 419 นาที เหลือเพียง 28 นาที หรือลดลงถึง 93%
- โมเดล AU-COREG (*U'*= 200) ที่ได้จากการเพิ่มตัวอย่างที่ถูกกำกับค่าจากข้อมูลที่ไม่มีป้ายกำกับ ขนาด 2 ตัวอย่าง จนกระทั่งถึง 10 ตัวอย่าง (ตัวแทนของ Cluster ด้วยเทคนิค K-Mean โดยเลือกตัวแทน Cluster จาก Seed 1992 100 และ 3645) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 7.61 ซึ่งได้จากการทำ Cluster จำนวน 3 Cluster (Seed 3645) ลดลง 1.58% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 217 นาที เหลือเพียง 24 นาที หรือลดลงถึง 89% และโมเดล AU-COREG (*U'*= 200) จะได้ค่า RMSE ที่มีค่าต่ำสุดเท่ากับ 7.67 ซึ่งได้จากการทำ Cluster จำนวน 3 Cluster (Seed 3645) ลดลง 2.88% (เทียบกับ COREG) ในขณะที่ระยะเวลาในการสร้างโมเดล ลดลงจาก 419 นาที เหลือเพียง 15 นาที หรือลดลงถึง 97%

4. สรุปและอภิปรายผล

การสร้างโมเดลจากการเลือกข้อมูลที่ไม่มีป้ายกำกับอย่างอัตโนมัติแบบกึ่งถดถอย (AU-COREG) ถูกพัฒนามาจากโมเดล COREG โดยลดขนาดตัวอย่างที่ไม่มีป้ายกำกับเพื่อคัดเลือกเข้าโมเดล ด้วยการทำ Cluster เพื่อช่วยลดจำนวนรอบในการคำนวณ และยังทำให้ประสิทธิภาพของโมเดลไม่ลดลง ซึ่ง

จำนวนรอบในการคำนวณ วิธี COREG

$$N = T * i * n * nU' \quad (9)$$

จำนวนรอบในการคำนวณ วิธี AU-COREG

$$N = T * i * n * nC \quad (10)$$

โดยที่ N คือจำนวนรอบในการคำนวณ

T จำนวนรอบในการทำซ้ำ

i จำนวนโมเดล

n จำนวนเพื่อนบ้าน

nU' จำนวนตัวอย่างที่ไม่มีป้ายกำกับ

nC จำนวนคลัสเตอร์

หากเพิ่มจำนวน Cluster ให้มากขึ้น จะเพิ่มจำนวนรอบในการคำนวณมากขึ้นอย่างไม่มีนัยสำคัญ ดังนั้นผู้วิจัยจึงได้ทำการทดลองเพิ่มจำนวน Cluster ตั้งแต่ 2 จนถึง 10 Cluster และเพิ่ม U' ให้มากขึ้นจาก 100 เป็น 200 ตัวอย่าง เพื่อให้สามารถทำการจัดกลุ่มให้ดียิ่งขึ้น

โมเดล AU-REG ต้องการเลือกตัวแทนของแต่ละ Cluster ที่ดีที่สุด (ลดค่าคาดเคลื่อนจากการพยากรณ์ให้มากที่สุด) เพื่อนำเข้าชุดข้อมูลเพื่อฝึกการเรียนรู้ให้กับโมเดล ดังนั้นผู้วิจัยจึงได้ใช้เทคนิคการทำ Cluster 2 วิธี ได้แก่ K-Medoids และ K-Mean โดยวิธี K-Medoids ผู้วิจัยเลือกสมาชิกที่อยู่จุดกึ่งกลางของข้อมูลใน Cluster (Centroid) เป็นตัวแทนของ Cluster และวิธี K-Mean ผู้วิจัยทำการระบุค่าสุ่ม ในการเลือกสมาชิก 3 ค่า เพื่อเปรียบเทียบประสิทธิภาพของโมเดล จากการเลือกตัวแทนของ Cluster ที่แตกต่างกัน

จากผลการทดลอง ดังตารางที่ 21-23 พบว่าโมเดล AU-COREG

- เมื่อทำการเพิ่มจำนวนข้อมูลที่ไม่มีป้ายกำกับเพิ่มจาก 100 เป็น 200 โดยส่วนใหญ่ทำให้ค่า RMSE ลดลง
- การเพิ่มจำนวนของ Cluster (แต่ไม่มากจนเกินไป) จะช่วยทำให้ค่า RMSE ลดลงได้ โดยใช้ระยะเวลาไม่แตกต่างไปจากเดิม
- การเลือกตัวแทนของ Cluster โดยใช้เทคนิคการจัด Cluster ที่แตกต่างกัน (K-Medoids และ K-Mean) มีผลต่อค่า RMSE ซึ่งพบกว่าส่วนใหญ่การจัด Cluster ด้วยวิธี K-Medoids กับข้อมูลประเภท Nominal จะให้ค่า RMSE ที่ต่ำกว่า

ตารางที่ 21. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU-COREG ที่มีค่าน้อยที่สุดของข้อมูลชุดที่ 1

Model	U'_{100}		U'_{200}	
	RMSE	Time (m)	RMSE	Time (m)
COREG	16,126.82	276	16,104.93	276
AU-COREG	16,047.12	31	16,003.33	31
ลดลง	0.49%	89%	0.14%	96%

ตารางที่ 22. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU-COREG ที่มีค่าน้อยที่สุดของข้อมูลชุดที่ 2

Model	U'_{100}		U'_{200}	
	RMSE	Time (m)	RMSE	Time (m)
COREG	2,113.46	215	2,110.38	428
AU-COREG	2,107.95	34	2,107.28	33
ลดลง	0.26%	84%	0.15%	92%

ตารางที่ 23. แสดงค่า RMSE และระยะเวลา (นาที) จากโมเดล AU-COREG ที่มีค่าน้อยที่สุดของข้อมูลชุดที่ 3

Model	U'_{100}		U'_{200}	
	RMSE	Time (m)	RMSE	Time (m)
COREG	7.77	217	7.90	419
AU-COREG	7.61	15	7.65	28
ลดลง	1.58%	93%	3.25%	93%

การวิจัยครั้งนี้นำเสนอขั้นตอนการสร้างโมเดลจากการเลือกข้อมูลที่ไม่มีป้ายกำกับอย่างอัตโนมัติแบบกึ่งถดถอย (AU-COREG) ซึ่งทำการทดลองกับข้อมูล 3 ชุด ที่มีคุณลักษณะของ

ข้อมูลที่แตกต่างกัน และใช้โมเดลพยากรณ์ kNN 2 โมเดล ที่มีวิธีการวัดระยะทางเหมือนกันแต่ต่างกันที่ค่าพารามิเตอร์เค หรือวิธีการวัดระยะทางที่แตกต่างกันขึ้นอยู่กับลักษณะของข้อมูล

ในการเรียนรู้ซ้ำในแต่ละรอบ มีการจัดกลุ่มข้อมูลที่ไม่มีป้ายกำกับ ที่ผ่านการทำนายจากโมเดลพยากรณ์ โดยและเลือกตัวแทนกลุ่มเพื่อหาตัวแทนกลุ่ม ที่ทำให้ความคลาดเคลื่อนน้อยที่สุด โดยการวิจัยนี้ ยังมีการเพิ่มข้อมูลที่ไม่มีป้ายกำกับ เพื่อให้ข้อมูลเหมาะสมกับการจัดกลุ่มที่เพิ่มขึ้น และเพิ่มวิธีการคัดเลือกตัวแทนของกลุ่ม เพื่อเพิ่มประสิทธิภาพของโมเดลให้ดียิ่งขึ้น

การทำนายขั้นสุดท้าย โดยหาค่าเฉลี่ยของการทำนายของทั้งสองโมเดล การวิจัยนี้แสดงให้เห็นว่าวิธี AU-COREG สามารถปรับปรุงประสิทธิภาพของโมเดลโดยช่วยลด RMSE ได้มากกว่า 0.14% และยังช่วยลดเวลามากกว่า 84% ผู้วิจัยยังขาดการปรับพารามิเตอร์ให้เหมาะสม กับคุณลักษณะของข้อมูล เช่น จำนวนข้อมูลที่ไม่มีป้ายกำกับที่จะเพิ่มเข้าไปในข้อมูลการเรียนรู้ จำนวนเพื่อนบ้านที่ใช้ทดสอบความคลาดเคลื่อน จำนวนรอบการทำซ้ำ เป็นต้น ซึ่งอาจช่วยปรับปรุงโมเดลให้มีประสิทธิภาพมากยิ่งขึ้น

เอกสารอ้างอิง

[1] Bizibl Marketing, “10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations,” Bizibl Marketing, June 16, 2019. [Online]. Available: https://bizibl.com/marketing/download/10-key-marketing-trends-2017-and-ideas-exceeding-customer-expectations. [Accessed: June 16, 2020].

[2] Blum A., Mitchell T., “Combining labeled and unlabeled data with co-training,” COLT' 98 Proceedings of the eleventh annual conference on Computational learning theory, July, 1998, pp. 92-100.

[3] Didaci L., Fumera, G., Roli, F., “Analysis of co-training algorithm with very small training sets,” Gimel’farb, G., et al. (eds.) SSPR/SPR 2012. LNCS., Springer, Heidelberg., vol. 7626, 2012.

[4] R. Wang and L. Li, "The performance improvement algorithm of co-training by committee," 2016 5th International Conference on Computer Science and Network Technology (ICCSNT), Changchun, 2016, pp. 407-412, doi: 10.1109/ICCSNT.2016.8070190.

[5] Sousa R., Gama J., “Co-training Semi-Supervised Learning for Single-Target Regression in Data Streams Using AMRules,” In: Kryszkiewicz M., Appice A., Ślęzak D., Rybinski H., Skowron A., Raś Z. (eds) Foundations of Intelligent Systems, ISMIS 2017, Lecture Notes in Computer Science, Vol. 10352, 2017.

[6] F Ma, D Meng, Q Xie, Z Li, X Don, “Self-paced co-training,” Proceedings of the 34th International Conference on Machine Learning, Vol. 70, pp. 2275-2284, 2017.

[7] Zhi Hua., Ming Li., “Semi-supervised regression with co-training,” IJCAI’05 proceeding of the 19th international joint conference on artificial intelligence, July, 2005, pp. 908–913.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 24-29

การประมวลผลภาพสำหรับการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึก

Image Processing for Classifying the Quality of the Chok-Anan Mango by Simulating the Human Vision using Deep Learning

นพรุจ พัฒนสาร และ ณัฐวุฒิ ศรีวิบูลย์*

Nopparut Pattansarn and Nattavut Sriwiboon*

สาขาเทคโนโลยีสารสนเทศและคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยีสุขภาพ
มหาวิทยาลัยกาฬสินธุ์

Department of Informatics and Computer, Faculty of Science and Health Technology,
Kalasin University

Received: January 08, 2020; **Revised:** April 30, 2020; **Accepted:** April 30, 2020; **Published:** June 23, 2020

ABSTRACT – This paper uses image processing technology with the deep learning methods, which can simulate the human vision to develop a model for examining and classifying the quality of the Chok-Anan mango. The research method, we have collect the image of Chok-Anan mangos and collecting quality classification data, determining quality levels into 4 levels consisting of grade A, B, C and grade D are rotten mangos. The results of the research shown that the use of deep learning by the convolutional neural network (CNN) algorithm for image processing to create a model showing the excellent accuracy at 99.79%. Then, we use the model to develop as a prototype for image classifying of Chok-Anan mangos. The result has found that the success rate of classification at 100%.

KEYWORDS: Image Processing, Deep Learning, Chok-Anan Mango, Convolutional Neural-Network

บทคัดย่อ -- งานวิจัยนี้ได้นำเทคโนโลยีการประมวลผลภาพด้วยวิธีการเรียนรู้เชิงลึกที่สามารถจำลองการมองเห็นของมนุษย์พัฒนาแบบจำลองสำหรับตรวจสอบและจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยดำเนินการเก็บรวบรวมข้อมูลภาพถ่ายมะม่วงพันธุ์โชคอนันต์และเก็บข้อมูลการจำแนกคุณภาพกำหนดระดับคุณภาพออกเป็น 4 ระดับประกอบด้วยคุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า ผลของการวิจัยแสดงให้เห็นว่าการใช้วิธีการเรียนรู้เชิงลึกด้วยอัลกอริธึมโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ในการประมวลผลภาพเพื่อสร้างแบบจำลองแสดงค่าความแม่นยำสูงสุดคือ 99.79% จากนั้นนำแบบจำลองพัฒนาเป็นระบบต้นแบบสำหรับจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์พบว่าอัตราความสำเร็จในการจำแนกคือ 100%

คำสำคัญ: การประมวลผลภาพ, การเรียนรู้เชิงลึก, มะม่วงพันธุ์โชคอนันต์, โครงข่ายประสาทเทียมแบบสังวัตนาการ

*Corresponding Author: *jistjournalsubmit@gmail.com*

1. บทนำ

มะม่วงพันธุ์โชคอนันต์เป็นไม้ยืนต้นอยู่ในวงศ์ Anacardiaceae โดยในการออกผลโตเต็มที่มีขนาดประมาณ 3-4 ผล ต่อ 1 กิโลกรัม ผลดิบมีสีเขียวเปลือกหนา เมื่อผลสุกแล้วผิวของผลมะม่วงพันธุ์โชคอนันต์จะเป็นสีเหลืองทองตลอดทั้งผลดูสวยงาม รสชาติหวานหอมถือเป็นพืชเศรษฐกิจที่สำคัญของประเทศ ในการตรวจสอบคุณภาพก่อนนำมะม่วงพันธุ์โชคอนันต์ออกสู่ตลาดนั้นจะมีการตรวจสอบโดยใช้มนุษย์ ซึ่งแบ่งระดับคุณภาพออกเป็น 4 ระดับ คือ คุณภาพระดับ A คุณภาพระดับ B คุณภาพระดับ C และระดับ D คือผลเน่า โดยคุณลักษณะที่ใช้ในการพิจารณาประกอบไปด้วย สีของผิว รูปร่างของผลเป็นต้น ซึ่งการคัดคุณภาพของมะม่วง จำเป็นต้องอาศัยประสบการณ์และทักษะของผู้ปฏิบัติงานเพื่อแยกแยะและตัดสินใจ จากการลงพื้นที่สอบถามเจ้าของสวนมะม่วงพันธุ์โชคอนันต์มีข้อผิดพลาดในการทำงานคือไม่สามารถคัดคุณภาพของสินค้าที่มีลักษณะตามความต้องการได้ ส่งผลให้เกิดปัญหาในการจำหน่ายผลิตภัณฑ์และตลาดขาดความเชื่อถือในผลิตภัณฑ์ของเจ้าของส่วนมะม่วงพันธุ์โชคอนันต์

การประมวลผลภาพ (Image Processing) เป็นอีกเทคโนโลยีด้านคอมพิวเตอร์ที่นำมาใช้ในการตรวจสอบวัตถุและผลิตภัณฑ์อย่างแพร่หลาย [1-3] ความสามารถของการประมวลผลภาพคือการประมวลผลวัตถุจากภาพถ่ายแล้วนำข้อมูลที่ได้ไปวิเคราะห์และสร้างเป็นระบบเพื่อจำแนกวัตถุที่ปรากฏในภาพ

งานวิจัยนี้จึงนำเทคโนโลยีการประมวลผลภาพมาพัฒนาระบบสำหรับการตรวจสอบและจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยดำเนินการเก็บข้อมูลภาพถ่ายมะม่วงพันธุ์โชคอนันต์จากเจ้าของสวนมะม่วง บ้านหนองผ้าอ้อม อำเภอสมเด็จ จังหวัดกาฬสินธุ์ แล้วจำแนกคุณลักษณะของสี รูปร่าง และตำหนิบนผิวของมะม่วงออกเป็น 4 ระดับคุณภาพ ประกอบด้วย คุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า จากนั้นพัฒนาเป็นระบบเพื่อจำลองการมองเห็นของมนุษย์ในคอมพิวเตอร์โดยใช้เทคนิคการประมวลผลภาพด้วยการเรียนรู้เชิงลึก (Deep Learning) เพื่อสร้างแบบจำลอง (Model) สำหรับพัฒนาระบบตรวจสอบคุณภาพมะม่วงพันธุ์โชคอนันต์ ผลของการวิจัยพบว่าการสร้างแบบจำลองมีความถูกต้อง 99.79% แสดงถึงประสิทธิภาพในการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ด้วยวิธี Deep Learning จากนั้นนำแบบจำลองพัฒนาเป็นระบบต้นแบบเพื่อจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์เพื่อเป็นตัวอย่างในการพัฒนาระบบอัตโนมัติสำหรับคัดแยกคุณภาพมะม่วง ส่งเสริมการจำหน่ายผลิตภัณฑ์และสร้างความเชื่อมั่นในตัวผลิตภัณฑ์ให้กับเจ้าของส่วนมะม่วงพันธุ์โชคอนันต์ต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 การประมวลผลภาพ (Image Processing)

การประมวลผลภาพ (Image Processing) คือเทคโนโลยีคอมพิวเตอร์ที่สามารถนำภาพมาประมวลผลผ่านกระบวนการเช่นการทำให้ภาพมีความคมชัดมากขึ้น การกำจัดสัญญาณรบกวนออกจากภาพ หรือการแบ่งส่วนของวัตถุ เป็นต้น เพื่อให้ได้ข้อมูลเชิงปริมาณไปวิเคราะห์และสร้างเป็นระบบ โดยการประมวลผลภาพสามารถนำไปใช้ประโยชน์ในงานด้านต่างๆ เช่น ระบบแยกประเภทไข่มุกด้วยวิธีการประมวลผลภาพ [1] และการตรวจจับและจดจำโมเดลรถยนต์ด้วยข้อมูลเชิงจุดภาพ [2] เป็นต้น

2.2 Deep Learning

Deep Learning หรือการเรียนรู้เชิงลึกคือการพัฒนาเทคโนโลยีคอมพิวเตอร์ให้สามารถเลียนแบบการทำงานของมนุษย์ ซึ่ง Deep Learning จะมีกระบวนการคิดคำนวณคล้ายกับระบบโครงข่ายประสาท (Neurons) ของสมองมนุษย์เรียกว่าโครงข่ายประสาทเทียม (Neural Network: NN) [4] ข้อดีของ Deep Learning คือเมื่อต้องการใช้งานอย่างเช่น การประมวลผลภาพเพื่อคัดแยกคุณภาพของมะม่วงพันธุ์โชคอนันต์การใช้งานไม่จำเป็นต้องให้ความรู้พื้นฐานกับระบบล่วงหน้า ความสามารถของ Deep Learning ก็สามารถสร้างแบบจำลองและหาคำตอบได้ ด้วยการนำ NN หลายๆ ชั้นเรียกว่า Hidden Layer มาใช้วิเคราะห์และหาคำตอบดังรูปที่ 1 ซึ่งคำว่า Deep Learning ก็มาจากการใช้ NN มากกว่า 2 ชั้นเพื่อให้เกิดการเรียนรู้และสร้างแบบจำลอง ดังนั้นจึงเปรียบเทียบได้ว่าแต่ละชั้นของ NN ยิ่งถูกใช้จำนวนมากในขั้นตอนการประมวลผล ยิ่งทำให้มีโครงสร้างการเรียนรู้ที่ลึก (Deep) มากขึ้น

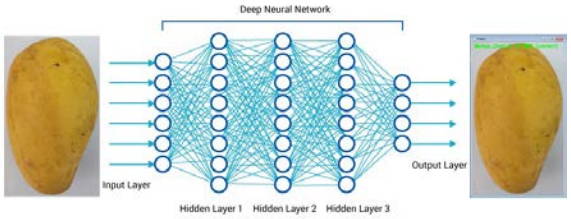
2.3 โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN)

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) [5] เป็นวิธีหนึ่งในวิธีการเรียนรู้แบบ Deep Learning เป็นการจำลองการมองเห็นของมนุษย์ที่สามารถแยกแยกคุณลักษณะ (Feature) ของวัตถุที่มองเห็น เช่น สี ขอบภาพ การตัดกันของสี แล้วนำคุณลักษณะต่างๆ มาประกอบกันเพื่อระบุคุณสมบัติของสิ่งที่มองเห็นแล้วคัดแยกว่าสิ่งนั้นมีคุณสมบัติเป็นอะไรเช่นเมื่อมองเห็นรูปภาพมะม่วงพันธุ์โชคอนันต์ การคำนวณของ CNN สามารถแสดงคำตอบได้ว่าภาพมะม่วงพันธุ์โชคอนันต์มีคุณลักษณะคือมีคุณภาพอยู่ในเกรดใดเป็นต้น ซึ่ง Deep Learning เป็นการประยุกต์ใช้วิธีการของ NN หลายๆ ชั้นเรียกว่า Hidden Layer ดังรูปที่ 1 สำหรับค้นหาคุณลักษณะและทำซ้ำหลายๆ รอบจนกระทั่งได้คำตอบของคุณคุณลักษณะที่มีความแม่นยำของการคัดแยก โดยการทำงานของ CNN จะพิจารณาจากความสัมพันธ์ของคุณลักษณะที่สกัดได้ในแต่ละชั้นกับผลลัพธ์มากที่สุด หลักการทำงานของ CNN มี 3 ส่วนดังนี้

1) Input: รับเข้าข้อมูลหรือวัตถุเหมือนกับการมองเห็นของมนุษย์ ตัวอย่างเช่นรูปภาพมะม่วงพันธุ์โชคอนันต์

2) Hidden Layer: ส่วนการประมวลผลเป็นชั้นๆ ซึ่งเหมือนกับการทำงานสมองของมนุษย์ เพื่อเรียนรู้ (Training) และและการคัดแยกประเภทของภาพ

3) Output: ส่วนแสดงผลลัพธ์การคัดแยกคุณสมบัติเป็นผลมาจากใช้ Hidden Layer จำนวนหลายชั้นมาวิเคราะห์จนได้คำตอบแสดงคุณลักษณะของแต่ละภาพที่มองเห็นเช่น มะม่วงพันธุ์โชคอนันต์มีคุณภาพอยู่ในเกรดใด เป็นต้น



รูปที่ 1. Convolutional Neural Network

จากการศึกษางานวิจัยก่อนหน้านี้ [6-8] ที่มีการประยุกต์ใช้ CNN กับการจำแนกข้อมูลโดยแสดงให้เห็นว่าข้อดีของ CNN คือสามารถวิเคราะห์วิเคราะห์และหาคำตอบได้อย่างแม่นยำ โดยความสามารถของ CNN มีความแม่นยำมากกว่า 90% เนื่องจากในขั้นตอน Feature extraction ที่เป็นการสกัดคุณลักษณะจากภาพ CNN มีความสามารถในการจัดการกับข้อมูลประเภทที่ไม่ได้มีโครงสร้างเป็นรูปแบบเฉพาะตัว (Unstructured Data) อย่างเช่น รูปภาพ (Image) เป็นต้น ดังนั้นด้วยคุณสมบัติที่ดีของ CNN งานวิจัยนี้จึงนำมาใช้สำหรับทดสอบและสร้างแบบจำลองเพื่อจำแนกคุณภาพของมะม่วงพันธุ์โชคอนันต์

2.4 งานวิจัยที่เกี่ยวข้อง

A. Paisal และ T. Kasetkasem [9] ได้ศึกษาการคัดแยกการปนของเมล็ดพันธุ์ถั่วเขียว โดยการวิเคราะห์จากภาพถ่าย ผลของงานวิจัยแสดงให้เห็นว่าการคัดแยกเมล็ดพันธุ์จากภาพถ่ายเมล็ดจำนวน 200 เมล็ด เป็นพันธุ์ชัยนาท 72 และพันธุ์กำแพงแสน 2 ระบบสามารถจำแนกภาพของเมล็ดพันธุ์ถั่วเขียวพันธุ์ชัยนาท 72 และพันธุ์กำแพงแสน 2 ที่ปนกัน ซึ่งสรุปผลการคัดแยกได้ถูกต้องมากกว่า 90 %

T. Tathawee และคณะ [6] ได้ศึกษาวิธีระบุชนิดของกล้วยไม้แล้วพัฒนาระบบการมองเห็นด้วยคอมพิวเตอร์ ผลการวิจัยพบว่าการพัฒนาเทคโนโลยีการมองเห็นด้วยระบบคอมพิวเตอร์เพื่อระบุชนิดกล้วยไม้ทั้งหมดสี่ชนิดจากสี่สกุลด้วยวิธีการเปรียบเทียบพื้นที่ contour ของสีปรากฏบนภาพดอกกล้วยไม้ การมองเห็นของคอมพิวเตอร์บ่งชี้ว่าพื้นที่ ที่ความยาวคลื่นแบบต่อเนื่อง (λ =400-700 nm) มีประสิทธิภาพสำหรับระบุชนิดกล้วยไม้ทั้งสี่ชนิดอย่างชัดเจน แต่ในช่วงความยาวคลื่นแบบไม่ต่อเนื่องที่ช่วงสีน้ำเงิน (λ = 475 nm) มีประสิทธิภาพสำหรับระบุกล้วยไม้ทั้งสี่ชนิดได้ดีที่สุด

N. Masunee [10] ได้เสนองานวิจัยการพัฒนาเทคนิคการตรวจสอบคุณภาพหมึกโดยใช้ปริมาณพื้นที่สีที่ปรากฏบนตัวหมึกเป็นลักษณะเด่นในการจำแนกระดับคุณภาพได้แก่ สีขาว สีชมพู สีแดง และสีดำ ซึ่งปริมาณพื้นที่สีที่ปรากฏบนตัวหมึกถูกคำนวณด้วยเทคนิคการประมวลผลภาพเพื่อจำแนกคุณภาพของหมึกโดยใช้การปริมาณข้อมูลพื้นที่สี

S. Aunkaew และคณะ [11] ได้เสนองานวิจัยการตรวจสอบราขาวบนผิวเนื้อยางแผ่น โดยระบบที่พัฒนาขึ้นด้วยเทคนิคการประมวลผลภาพสามารถตรวจสอบและจำแนกยางแผ่นที่มีราขาวออกจากยางแผ่นดีได้ผิดพลาดน้อยมากและมีความรวดเร็วเมื่อเทียบกับการตรวจสอบและจำแนกด้วยตามนุษย์

3. วิธีการดำเนินงานวิจัย

3.1 การเก็บรวมรวมข้อมูล

ในงานวิจัยนี้ ดำเนินการเก็บรวบรวมข้อมูลจากเจ้าของสวนมะม่วง บ้านหนองผ้าอ้อม อำเภอสมเด็จ จังหวัดกาฬสินธุ์ โดยเก็บรวบรวมรูปภาพมะม่วงพันธุ์โชคอนันต์คัดแยกเป็นระดับคุณภาพเกรดต่างๆ

3.2 การเตรียมข้อมูล

งานวิจัยได้ดำเนินการเก็บข้อมูลภาพมะม่วงพันธุ์โชคอนันต์ โดยแต่ละภาพมีขนาด 528 x 888 พิกเซล ขั้นตอนการจัดกลุ่มคุณภาพของมะม่วงพันธุ์โชคอนันต์ดำเนินการโดยสอบถามการจำแนกคุณภาพของมะม่วงในแต่ละระดับจากผู้เชี่ยวชาญเจ้าของสวนมะม่วงพันธุ์โชคอนันต์ โดยมีการจัดกลุ่มไว้ดังนี้ คุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า ดังรูปที่ 2

รูปที่ 2. ตัวอย่างการจัดกลุ่มตามคุณภาพระดับเกรด

3.3 การวัดประสิทธิภาพ

การวัดประสิทธิภาพเพื่อเปรียบเทียบค่าความแม่นยำเพื่อสร้างแบบจำลองในงานวิจัยนี้ใช้การวัดค่าความแม่นยำ (Accuracy) [7] เป็นค่าที่ได้จากวิธีการทดสอบเพื่อหาค่าพยากรณ์ความถูกต้องของข้อมูลโดยคิดเป็นค่าร้อยละ (%) ใช้สูตรการคำนวณดังนี้

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

โดย TP คือ ค่าที่พยากรณ์ถูกต้องเชิงบวก
TN คือ ค่าที่พยากรณ์ถูกต้องเชิงลบ
FP คือ ค่าที่พยากรณ์ผิดพลาดเชิงบวก
FN คือ ค่าที่พยากรณ์ผิดพลาดเชิงลบ

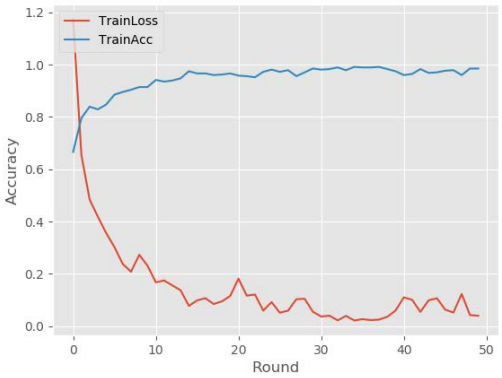
3.4 การสร้างแบบจำลอง

ในส่วนของการสร้างแบบจำลองมีการเตรียมขั้นตอน 2 ขั้นตอน คือ ขั้นตอนของการเรียนรู้ (Train) เพื่อสร้างแบบจำลองและขั้นตอนของการทดสอบ (Test) แบบจำลอง การสร้างแบบจำลองเป็นขั้นตอนการสร้างการเรียนรู้โดยใช้ข้อมูลรูปภาพมะม่วงพันธุ์โชคอนันต์ที่ได้จัดกลุ่มคุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า กลุ่มละ 150 รูป รวม 600 รูป สำหรับใช้ในการเรียนรู้ (Training Dataset) และกลุ่มละ 50 รูป สำหรับใช้ทดสอบ (Testing Dataset) แบบจำลอง โดยงานวิจัยนี้ใช้การกำหนดค่าการตรวจสอบแบบไขว้ (k-fold cross validation) [8] คือ k=10 ในการทดลองเพื่อให้ระบบสุ่มภาพในชุดข้อมูลการเรียนรู้ทั้ง 4 กลุ่มเพื่อฝึกสอนให้แบบจำลองเกิดการเรียนรู้ จำนวน 50 รอบ จากนั้นเมื่อได้แบบจำลองจึงนำภาพ 50 ภาพ ทดสอบความถูกต้องของการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์

4. ผลการวิจัย

4.1 ผลการสร้างแบบจำลอง

การพัฒนาระบบเพื่อจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ ใช้ภาษา Python ในการพัฒนา โดยใช้ไลบรารีสำหรับพัฒนา Deep Learning ร่วมกันคือ Keras และ Tensorflow เพื่อสร้างแบบจำลองและวิเคราะห์การจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ รายละเอียดสำหรับการติดตั้ง Keras และ Tensorflow แสดงใน [12] ผลของการสร้างแบบจำลองแสดงดังรูปที่ 3



รูปที่ 3. แสดงค่าความแม่นยำในการสร้างแบบจำลองจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์

จากรูปที่ 3 แสดงค่าความแม่นยำ (TrainAcc) และค่าความผิดพลาด (TrainLoss) ในการทดสอบสร้างแบบจำลองจะเห็นว่าการเรียนรู้ในแต่ละรอบ (Round) มีค่าความแม่นยำที่แตกต่างกัน โดยผลการทดลองสร้างแบบจำลองได้ค่าความแม่นยำสูงสุดคือ 99.79% แสดงถึงประสิทธิภาพในการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN

4.2 การนำแบบจำลองไปใช้งาน

จากการสร้างแบบจำลองเพื่อจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ งานวิจัยนี้ได้นำแบบจำลองที่สร้างขึ้นใช้สำหรับพัฒนาระบบเพื่อใช้ประมวลผลภาพจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ ระบบพัฒนาด้วยภาษา Python ใช้ไลบรารีสำหรับพัฒนาประกอบด้วย Keras และ Tensorflow จากนั้นนำระบบที่พัฒนาทดสอบจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยการทดลองได้นำภาพมะม่วงพันธุ์โชคอนันต์ที่จัดกลุ่มคุณภาพไว้ 4 กลุ่มข้างต้น กลุ่มละ 50 รูปคือ คุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D โดยแสดงตัวอย่างผลการคัดแยกมะม่วงพันธุ์โชคอนันต์ในแต่ละกลุ่มดังรูปที่ 4 ประกอบด้วยรูปที่ 4(a) คือตัวอย่างการทดสอบรูปภาพมะม่วงพันธุ์โชคอนันต์กลุ่มคุณภาพระดับเกรด A รูปที่ 4(b) คือตัวอย่างการทดสอบรูปภาพมะม่วงพันธุ์โชคอนันต์กลุ่มคุณภาพระดับเกรด B รูปที่ 4(c) คือตัวอย่างการทดสอบรูปภาพมะม่วงพันธุ์โชคอนันต์กลุ่มคุณภาพระดับเกรด C และรูปที่ 4(d) คือตัวอย่างการทดสอบรูปภาพมะม่วงพันธุ์โชคอนันต์กลุ่มคุณภาพระดับเกรด D และแสดงผลลัพธ์การทดสอบสรุปดังตารางที่ 1

ตารางที่ 1 สรุปผลการทดสอบใช้งานแบบจำลอง

เกรด	จำนวนภาพทดสอบ (รูป)	ผลการจำแนก (รูป)	ความถูกต้อง (%)
A	50	A = 50, B = 0, C = 0, D = 0	100
B	50	A = 0, B = 50, C = 0, D = 0	100
C	50	A = 0, B = 0, C = 50, D = 0	100
D	50	A = 0, B = 0, C = 0, D = 50	100

จากตารางที่ 1 แสดงผลลัพธ์การทดสอบแบบจำลองแสดงให้เห็นว่าแบบจำลองการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ที่ใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN มีประสิทธิภาพสูงสำหรับกระบวนการประมวลผลภาพเพื่อจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยผลลัพธ์การทดสอบรูปภาพกลุ่มคุณภาพเกรด A อัตราความสำเร็จในการจำแนกคือ 100% เกรด B อัตราความสำเร็จในการจำแนกคือ 100% เกรด C อัตราความสำเร็จในการจำแนกคือ 100% และเกรด D อัตราความสำเร็จในการจำแนกคือ 100% จากผลลัพธ์ดังกล่าวสรุปได้ว่าแบบจำลองการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ที่สร้างขึ้นโดยใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN มีความสามารถในการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ได้อย่างมีประสิทธิภาพ

(a)

(b)

(c)

(d)

รูปที่ 4. ตัวอย่างการนำรูปภาพมะม่วงพันธุ์โชคอนันต์ในแต่ละกลุ่มทดสอบในระบบ

5. บทสรุปและการอภิปราย

งานวิจัยนี้ได้นำเทคโนโลยีการประมวลผลภาพวิธี Deep Learning อัลกอริธึม CNN พัฒนาแบบจำลองเพื่อตรวจสอบและจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยได้ดำเนินการเก็บรวบรวมข้อมูลภาพถ่ายมะม่วงจากเจ้าของสวนมะม่วงและ

เก็บข้อมูลการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์โดยได้กำหนดระดับคุณภาพออกเป็น 4 ระดับประกอบด้วย คุณภาพระดับเกรด A, คุณภาพระดับเกรด B, คุณภาพระดับเกรด C และคุณภาพระดับเกรด D คือมะม่วงเน่า จากนั้นใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN ในการประมวลผลภาพเพื่อสร้างแบบจำลอง โดยผลการสร้างแบบจำลองแสดงค่าความแม่นยำสูงสุดคือ 99.79% แสดงถึงประสิทธิภาพในการสร้างแบบจำลองเพื่อจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ด้วยวิธี Deep Learning สอดคล้องกับงานวิจัย [6-8] ที่แสดงให้เห็นว่าการประยุกต์ใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN มีประสิทธิภาพในการจำแนกข้อมูลประเภทที่ไม่ได้มีโครงสร้างเป็นรูปแบบเฉพาะตัว (Unstructured Data) อย่างเช่น รูปภาพ (Image) ซึ่งสามารถสกัดคุณลักษณะเด่นจากรูปภาพได้อย่างมีประสิทธิภาพ

จากนั้นงานวิจัยนี้ได้นำแบบจำลองพัฒนาเป็นระบบต้นแบบสำหรับตรวจสอบคุณภาพมะม่วงพันธุ์โชคอนันต์ ผลการทดสอบใช้งานระบบพบว่าการทดสอบรูปภาพกลุ่มคุณภาพเกรด A อัตราความสำเร็จในการจำแนกคือ 100% เกรด B อัตราความสำเร็จในการจำแนกคือ 100% เกรด C อัตราความสำเร็จในการจำแนกคือ 100% และเกรด D อัตราความสำเร็จในการจำแนกคือ 100% ซึ่งจากผลลัพธ์การทดสอบใช้งานระบบแสดงให้เห็นว่าแบบจำลองการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ที่สร้างขึ้นโดยใช้วิธี Deep Learning ด้วยอัลกอริธึม CNN มีความสามารถในการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ได้อย่างมีประสิทธิภาพ สามารถนำผลการวิจัยเป็นตัวอย่างในการพัฒนาระบบอัตโนมัติสำหรับคัดแยกคุณภาพมะม่วงเพื่อส่งเสริมการจำหน่ายผลิตภัณฑ์และสร้างความเชื่อมั่นให้กับการจำหน่ายผลิตภัณฑ์มะม่วงพันธุ์โชคอนันต์ต่อไป

เอกสารอ้างอิง

[1] J. Jaroenjit, A. Panpanasakul, P. Chaisri, P. Promduang, and S. Prompongusawa, "Classification pearls using image processing," in Proceedings of the 9th Hatyai National and International Conference, Thailand, 2014, pp. 1679 - 1691.

[2] A. Tungkastan and K. Leewun, "Pixel-Based Car Model Detection and Recognition," Engineering Journal of Siam University, vol. 19, January-June, pp. 90–102, 2018.

[3] S. Phimphisan, "Application of Data Mining for Diabetic Retinopathy Using Decision Tree," Journal of Srivanalai Vijai, vol. 3, 2016.

[4] S. Sarraf and G. Tofighi, "A hybrid sequential feature selection approach for the diagnosis of Alzheimer's Disease," in Proceedings of International Joint Conference on Neural Networks, 2016, pp. 1216-1220.

[5] E. Humphrey and J. Bello, "Rethinking Automatic Chord Recognition with Convolution Neural Networks," in Proceedings of the 11th International Conference on Machine Learning and Application, 2012.

[6] T. Tathawee, S. Prasarnpun, S. Onbua, T. Pinthong, and A. Suwannakom, "Orchid identification based on computer vision analysis," in Proceedings of the 6th National Science Research Conference, Thailand, 2014, pp. 47 - 56.

[7] B. Tilmann, "The Business Impact of Predictive Analytics," ed. IGI Global, September - December 2007.

[8] R. Kohavi, "A study of crossvalidation and bootstrap for accuracy estimation and model selection," in Proceedings of the Fourteenth International joint conference on Artificial Intelligence, Montreal, Canada 1995, pp. 1137-1143.

[9] A. Paisal and T. Kasetkasem, "Separation the mingling varieties of the mungbean seeds by image processing," Khon Kaen Agriculture Journal, vol. 3, pp. 240-247, 2011.

[10] N. Masunee, "Development of an image processing system in splendid squid quality classification," Prince of Songkla University, 2014.

[11] S. Aunkaew, T. Khaorapapong, and M. Karnjanadecha, "A Study of Inspection of White Mould on Surface Rubber Sheets Using Image Processing," Thaksin University Journal, vol. 10, July - December, pp. 50–61, 2007.

[12] A. Rosebrock, "Installing Keras with TensorFlow backend," January, 2020. [Online]. Available: https://www.pyimagesearch.com/2016/11/14/installing-keras-with-tensorflow-backend/. [Accessed: Jan 7, 2020].

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 30-38

การเปรียบเทียบประสิทธิภาพการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ระหว่าง NFS กับ FTP กรณีศึกษา การพัฒนาระบบกล้องวงจรปิดอัจฉริยะต้นทุนต่ำด้วยราสเบอร์รี่พาย

Comparing Efficiency in Copying Video Files through Powerline between NFS and FTP: A Case Study of Developing a Low-cost Smart CCTV System through Raspberry Pi

เชี่ยวชาญ ยางศิลา

Cheawchan Yangsila

สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏร้อยเอ็ด

Computer and Information Technology Faculty of Information Technology Roi Et Rajabhat University

Received: January 19, 2020; **Revised:** May 03, 2020; **Accepted:** May 09, 2020; **Published:** June 23, 2020

ABSTRACT – The objectives of this study were to develop a low-cost smart CCTV (closed circuit television) system with Raspberry Pi and to compare the efficiency in copying video files through powerline between NFS (network file system) and FTP (file transfer protocol). In the experiment, the researcher recorded and saved video files for 48 days (The video files with different lengths of 1,2,4, and 6 minutes were all recorded for 12 days). The video files had an image resolution of 1,920 × 1,080 pixels. A CCTV camera was installed in the researcher's office on the 1st floor of the Language and Computer Center building, Roi Et Rajabhat University. At 23.56 each day, the system uploaded the video files to the Raspberry Pi boards (server) and recorded the statistics. It was found that the whole system was practical. Warning messages could be sent via line notify when both Raspberry Pi boards could not communicate. From the efficiency measurement, it was found that NFS was more efficient in copying the files via powerline network than FTP. The larger the files, the bigger difference in copying time.

KEYWORDS: Low-cost Smart CCTV, Raspberry Pi, NFS, FTP

บทคัดย่อ -- การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบกล้องวงจรปิดอัจฉริยะต้นทุนต่ำด้วยราสเบอร์รี่พายและเพื่อเปรียบเทียบประสิทธิภาพการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ระหว่าง NFS กับ FTP ผู้วิจัยทำการทดลองบันทึกไฟล์วีดิโอ จำนวน 48 วัน (ความยาวของไฟล์ 1 นาที บันทึก 12 วัน, 2 นาที บันทึก 12 วัน, 4 นาที บันทึก 12 วัน และ 6 นาที บันทึก 12 วัน) ความละเอียดของภาพขนาด 1,920 × 1,080 จุด โดยได้ติดตั้งกล้องในห้องพักของผู้วิจัย ณ ชั้น 1 อาคารศูนย์ภาษาและคอมพิวเตอร์ มหาวิทยาลัยราชภัฏร้อยเอ็ด ซึ่งเวลา 23.56 น. ของแต่ละวันระบบจะทำการอัพโหลดไฟล์วีดิโอไปจัดเก็บในบอร์ดราสเบอร์รี่พาย (server) และบันทึกสถิติ จากการทดลองพบว่า ระบบสามารถใช้งานได้จริง สามารถส่งข้อความแจ้งเตือนผ่าน line notify เมื่อบอร์ดราสเบอร์รี่พายทั้งสองตัวไม่สามารถสื่อสารกันได้ จากการวัดประสิทธิภาพพบว่า NFS มีประสิทธิภาพในการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ได้ดีกว่า FTP และถ้าขนาดของไฟล์วีดิโอมีขนาดใหญ่ขึ้น ส่วนต่างของเวลาในการสำเนาไฟล์ก็จะมากขึ้นตามไปด้วย

คำสำคัญ: กล้องวงจรปิดอัจฉริยะต้นทุนต่ำ, ราสเบอร์รี่พาย, เอ็นเอฟเอส, เอฟทีพี

Corresponding Author: kasam_com@hotmail.com

1. บทนำและวัตถุประสงค์

ในปัจจุบันกล้องวงจรปิดมีความสำคัญเป็นอย่างยิ่ง หลาย ๆ คดีคลี่คลายได้อย่างรวดเร็ว แต่มีบางคดีไม่สามารถคลี่คลายได้ทั้ง ๆ ที่มีกล้องวงจรปิด สาเหตุก็เพราะว่าในขณะที่เกิดเหตุ กล้องวงจรปิดเกิดความผิดพลาดในการบันทึกข้อมูลด้วยสาเหตุใด ๆ ก็ตาม ซึ่งผู้ดูแลไม่ทราบความผิดปกติที่เกิดขึ้น

กล้องวงจรปิดที่จำหน่ายตามท้องตลาดในปัจจุบัน มี 2 แบบ ได้แก่ แบบ analog และแบบ ip camera ซึ่งกล้องวงจรปิดแบบ ip camera บันทึกวีดิโอโดยส่งข้อมูลผ่านเครือข่ายท้องถิ่น (LAN) ไปบันทึกที่เครื่อง NVR โดยกล้องแต่ละตัวต้องการแบนด์วิดท์ (bandwidth) สูงมาก ตั้งแต่ 500 kbps ถึง 1.5 Mbps [1] และต้องส่งข้อมูลตลอดเวลา ถ้ามีกล้องหลายตัวจะทำให้ระบบเครือข่ายทำงานหนักกระทบต่อการใช้งานอินเทอร์เน็ตหรืออาจจะต้องลงทุนติดตั้งเครือข่ายแยกต่างหากสำหรับระบบกล้องวงจรปิดโดยเฉพาะเพื่อไม่ให้รบกวนการใช้งานเครือข่ายอินเทอร์เน็ตแต่ก็ต้องแลกกับการต้องจ่ายเงินเพิ่มเติม ในกรณีที่เครือข่ายมีปัญหาก็จะทำให้ไม่สามารถบันทึกวีดิโอได้ผู้ดูแลต้องหมั่นตรวจสอบอยู่ตลอดเวลา ประกอบกับค่าใช้จ่ายในการติดตั้งถือว่าค่อนข้างสูง

การติดตั้งระบบเครือข่ายท้องถิ่นในปัจจุบันมีสื่อให้เลือกหลายชนิดแต่ละชนิดก็มีข้อดีข้อเสียแตกต่างกัน เช่น สาย UTP ข้อดีก็คือ ราคาถูกติดตั้งง่าย ส่วนข้อจำกัดก็คือ ระยะทางไม่เกิน 100 เมตร สาย fiber optic ข้อดีก็คือสามารถส่งข้อมูลได้ไกล ข้อมูลมีความปลอดภัย ข้อเสียก็คือ ค่าใช้จ่ายในการติดตั้งสูง และสาย power line ข้อดีเมื่อเปรียบเทียบกับสาย UTP ก็คือสามารถส่งข้อมูลได้ไกลว่า คือสามารถส่งข้อมูลระยะทาง 300 เมตรและสามารถใช้งานร่วมกับสายไฟฟ้าภายในบ้านได้ไม่ต้องลงทุนติดตั้งสายสัญญาณเพิ่มเติม มีหลายงานวิจัยได้ทดสอบประสิทธิภาพของสาย power line สามารถนำมาประยุกต์ใช้งานได้ [2,3] ส่วนข้อจำกัด ก็คือ ไม่สามารถส่งข้อมูลข้ามเบรกเกอร์ [2]

จากปัญหาที่พบข้างต้นผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบกล้องวงจรปิดอัจฉริยะต้นทุนต่ำด้วยราสเบอร์รี่พาย ส่งข้อมูลผ่านสาย power line สามารถระบุช่วงเวลาในการอัพโหลดไฟล์วีดิโอไปเก็บที่เซิร์ฟเวอร์เพื่อไม่ให้กระทบต่อการใช้งานอินเทอร์เน็ตและสามารถแจ้งเตือนผู้ดูแลผ่าน line notify เมื่อระบบเกิดความผิดพลาดใด ๆ

1.1 วัตถุประสงค์

1.1.1 เพื่อพัฒนาระบบกล้องวงจรปิดอัจฉริยะต้นทุนต่ำด้วยราสเบอร์รี่พาย

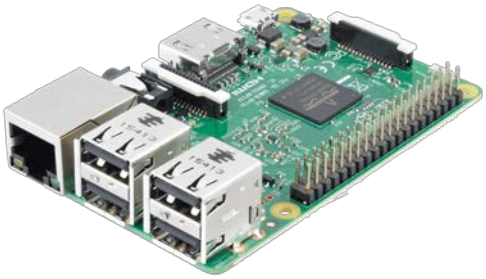
1.1.2 เพื่อเปรียบเทียบประสิทธิภาพการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ระหว่างเอ็นเอฟเอสกับเอฟทีพี

2. ทฤษฎี อุปกรณ์และงานวิจัยที่เกี่ยวข้อง

2.1 ภาษาพีเอชพี (PHP) คือ ภาษาคอมพิวเตอร์ในลักษณะเซิร์ฟเวอร์ไซด์สคริปต์ โดยลิขสิทธิ์อยู่ในลักษณะโอเพนซอร์ส ภาษาพีเอชพีใช้สำหรับพัฒนาเว็บเพจและแสดงผลออกมาในรูปแบบ HTML ในงานวิจัยนี้ ใช้ภาษาพีเอชพีในการพัฒนาเว็บเพจสำหรับอัพโหลดไฟล์วีดิโอไปเก็บที่เซิร์ฟเวอร์

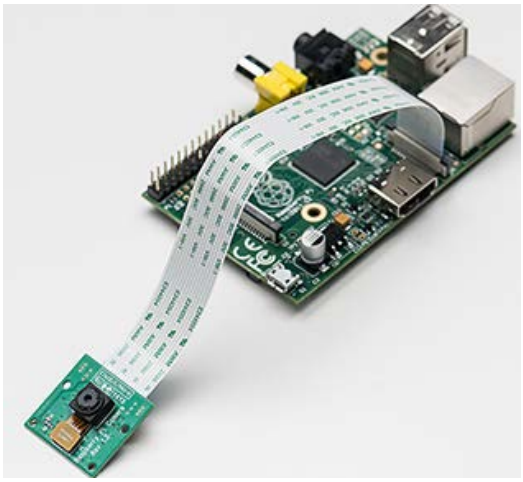
2.2 ภาษาไพทอน (python) คือ ชื่อภาษาที่ใช้ในการเขียนโปรแกรมภาษาหนึ่ง ซึ่งถูกพัฒนาขึ้นมาโดยไม่ยึดติดกับแพลตฟอร์ม กล่าวคือ สามารถรันภาษาไพทอนได้ทั้งบนระบบ Unix, Linux , Windows NT, Windows 2000, Windows XP หรือแม้แต่ระบบ FreeBSD อีกอย่างหนึ่งภาษาตัวนี้เป็นโอเพนซอร์สเหมือนภาษาพีเอชพี ทำให้ทุกคนสามารถที่จะนำภาษาไพทอนมาพัฒนาโปรแกรมได้ฟรี ในงานวิจัยนี้ได้ใช้ภาษาไพทอนเขียนโปรแกรมสั่งงานให้ราสเบอร์รี่พายบันทึกไฟล์วีดิโอ

2.3 บอร์ดราสเบอร์รี่พาย (raspberry pi) [4] คือ บอร์ดคอมพิวเตอร์ขนาดเล็กที่สามารถเชื่อมต่อกับจอมอนิเตอร์ คีย์บอร์ดและเมาส์ได้ สามารถนำมาประยุกต์ใช้ในการทำโครงงานทางด้านอิเล็กทรอนิกส์ การเขียนโปรแกรม หรือเป็นเครื่องคอมพิวเตอร์ตั้งโต๊ะขนาดเล็ก ไม่ว่าจะเป็นการทำงาน spreadsheet word processing ท่องอินเทอร์เน็ต ส่งอีเมล์ หรือเล่นเกมส์ อีกทั้งยังสามารถเล่นไฟล์วีดีโอความละเอียดสูง (high-definition) ได้อีกด้วย บอร์ดราสเบอร์รี่พายรองรับระบบปฏิบัติการลินุกซ์ (linux operating system) ได้หลายระบบ เช่น Raspbian (Debian) Pidora (Fedora) และ Arch Linux เป็นต้น โดยติดตั้งบน SD card บอร์ดราสเบอร์รี่พายนี้ถูกออกแบบมาให้มี CPU GPU และ RAM อยู่ภายในชิปเดียวกัน มีจุดเชื่อมต่อ GPIO ให้ผู้ใช้สามารถนำไปใช้ร่วมกับอุปกรณ์อิเล็กทรอนิกส์อื่น ๆ ได้อีกด้วย บอร์ดราสเบอร์รี่พาย ปัจจุบันมีด้วยกัน 2 โมเดล คือ โมเดล A และ โมเดล B ซึ่งทั้ง 2 โมเดลมีคุณสมบัติทางเทคนิคที่ใกล้เคียงกันแตกต่างกันเพียงบางส่วน ในงานวิจัยนี้ผู้วิจัยเลือกใช้ราสเบอร์รี่พาย 3 B support LAN 10/100 Mbps ดังรูปที่ 1



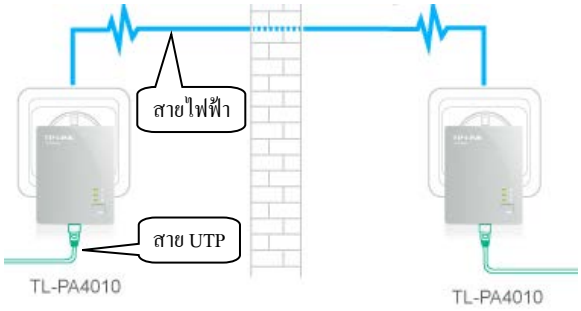
รูปที่ 1. บอร์ดราสเบอร์รี่พาย โมเดล B

2.4 Pi camera เป็นโมดูลกล้องสำหรับต่อใช้งานร่วมกับบอร์ดราสเบอร์รี่พาย ขนาดความละเอียด 5 ล้าน pixel สามารถถ่ายวิดีโอระดับ HD ที่ความละเอียด 1080p, 720p และ 640x480 ด้วยอัตราแสดงผล 30 (1080p), 60 (720p และ 640x480) และ 90 (640x480) เฟรมต่อวินาที ในงานวิจัยนี้ เลือกใช้ยี่ห้อ Element14 รุ่น RPI CAMERA BOARD ดังรูปที่ 2



รูปที่ 2. ตัวอย่าง Pi camera

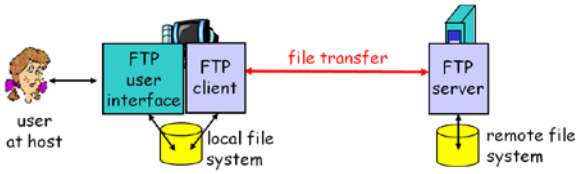
2.5 การสื่อสารผ่านสายไฟฟ้า (powerline communication : PLC) หมายถึง ระบบเครือข่ายที่ใช้สายไฟฟ้าเป็นเส้นทางในการติดต่อสื่อสารระหว่างคอมพิวเตอร์ สามารถสร้างระบบเครือข่ายสำหรับการสื่อสารข้อมูลขึ้นมาใช้งานได้ทันที เรียกได้ว่า ที่ใดมีปลั๊กที่นั้นมีเครือข่ายโดยในการส่งสัญญาณมัลติมีเดียด้วยความเร็วสูง โดยที่สัญญาณจะถูกแปลงด้วยอะแดปเตอร์ที่เชื่อมต่อแล้วส่งไปตามวงจรสายไฟฟ้า โดยใช้คลื่นความถี่ระหว่าง 2 ถึง 30 MHz ไปยังอะแดปเตอร์อีกตัวที่อยู่ปลายทาง ซึ่งจะคอยทำการแปลงสัญญาณให้กลับมาอยู่ในรูปแบบเดิม และสามารถทำงานได้ครอบคลุมในพื้นที่ที่สัญญาณเข้าไม่ถึงแต่สายไฟฟ้าเข้าถึง เพื่อมีเป้าหมายในการที่จะช่วยอำนวยความสะดวกในการใช้งานสำหรับผนังที่มีสายไฟฟ้าติดตั้งอยู่ก็สามารถทำการติดต่อสื่อสารข้อมูลกันได้ [5] หลักการทำงานของ PLC ดังภาพที่ 3 งานวิจัยนี้ใช้อุปกรณ์ PLC ยี่ห้อ tp-link รุ่น TL-PA4010 อัตราการถ่ายโอนข้อมูลสูงสุดถึง 500 Mbps บนระยะทาง 300 เมตร [6] ดังรูปที่ 4



รูปที่ 3. หลักการเชื่อมต่อ PLC

รูปที่ 4. อะแดปเตอร์ TL-PA4010

2.6 เอฟทีพี (FTP) เป็นโพรโทคอลเครือข่ายชนิดหนึ่งใช้สำหรับแลกเปลี่ยนและจัดการไฟล์บนเครือข่ายอินเทอร์เน็ต เอฟทีพีถูกสร้างขึ้นด้วยสถาปัตยกรรมแบบไคลเอนต์/เซิร์ฟเวอร์ (ดังภาพที่ 5) และใช้การเชื่อมต่อสำหรับส่วนข้อมูลและส่วนควบคุมแยกกันระหว่างเครื่องลูกข่ายกับเครื่องแม่ข่าย โปรแกรมประยุกต์เอฟทีพีเริ่มแรกโต้ตอบกันด้วยเครื่องมือรายคำสั่ง (command line) สั่งการด้วยไวยากรณ์ที่เป็นมาตรฐาน แต่ก็มีการพัฒนาส่วนต่อประสานกราฟิกกับผู้ใช้ขึ้นมาสำหรับระบบปฏิบัติการเดสก์ท็อปที่ใช้กันทุกวันนี้ เอฟทีพียังถูกใช้เป็นส่วนประกอบของโปรแกรมประยุกต์อื่นเพื่อส่งผ่านไฟล์โดยอัตโนมัติสำหรับการทำงานภายในโปรแกรม เราสามารถใช้เอฟทีพีผ่านทางการพิสูจน์ตัวจริงด้วยชื่อผู้ใช้และรหัสผ่าน หรือเข้าถึงด้วยผู้ใช้นิรนาม (anonymous)



รูปที่ 5. สถาปัตยกรรมแบบไคลเอนต์/เซิร์ฟเวอร์ของเอฟทีพี

2.7 เอ็นเอฟเอส (network file system : NFS) คือ บริการที่ทำให้เครื่องคอมพิวเตอร์สามารถเข้าถึง File และ Directory บนเครื่องคอมพิวเตอร์เครื่องอื่นได้เหมือนกับใช้งานเครื่องของตัวเอง โดยสามารถใช้บริการได้อย่างสะดวก ง่ายและมีประสิทธิภาพผ่านระบบเครือข่าย โดยระบบปฏิบัติการของเครื่องลูกข่ายไม่จำเป็นต้องเป็นระบบปฏิบัติการเดียวกันกับเครื่องแม่ข่ายที่ให้บริการ NFS

ประโยชน์ของ NFS

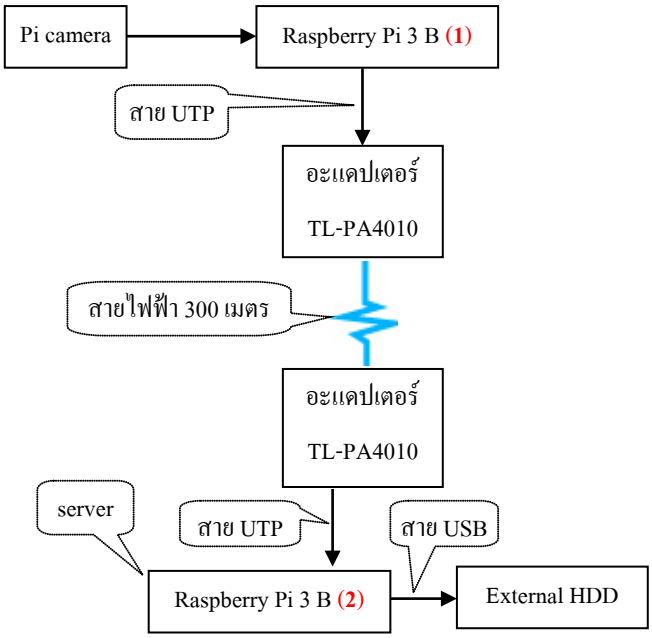
- การใช้ NFS ไม่จำเป็นจะต้องใช้ระบบปฏิบัติการเดียวกัน
- มีความปลอดภัยของข้อมูลทั้งเป็นแบบรายบุคคลและแบบรายกลุ่ม ซึ่งการเข้าใช้งานจะมีการยืนยันตัวบุคคลและรหัสผ่านก่อนเข้าใช้งาน
- สามารถเข้าใช้งานข้อมูลได้ร่วมกัน ทุกที่ ทุกเวลาที่การเชื่อมต่อเครือข่ายไปถึง
- เป็นแหล่งสำรองข้อมูล

2.8 มลฑล คล้ายสวัสดิ์ [2] ได้ศึกษาปัญหาพิเศษ เรื่อง การเปรียบเทียบประสิทธิภาพการสตรีมมิ่งไฟล์วีดีโอความละเอียดสูงระหว่างเครือข่ายไร้สายกับเครือข่ายที่ใช้สายไฟฟ้า ผลการศึกษาพบว่า ถ้าในกรณีที่ต้องการสตรีมมิ่งไฟล์วีดีโอความละเอียดสูงชนิด MKV จะสามารถเลือกใช้งานสื่อนำสัญญาณได้ทั้งชนิด Powerline หรือ WLAN ชนิด N150 และ N300 แต่ถ้าต้องการสตรีมมิ่งวีดีโอไฟล์ชนิด Blue-Ray สื่อนำสัญญาณที่สามารถใช้งานได้อย่างมีประสิทธิภาพก็คงมีเพียงสื่อนำสัญญาณชนิดใช้สายทองแดง (UTP) เท่านั้น

2.9 เกรียงไกร พงษ์มนตรี [3] ได้ศึกษาปัญหาพิเศษ เรื่องการวัดประสิทธิภาพในการโอนถ่ายข้อมูลผ่าน Powerline กรณีศึกษาในห้องเร่งอนุภาครังสีรักษา ณ โรงพยาบาลราชวิถี ผลการศึกษาพบว่า ว่าในขณะทำการฉายรังสีไม่ส่งผลกระทบต่อประสิทธิภาพการทำงานของอุปกรณ์ Powerline และในการใช้งานส่วนของระบบโรงพยาบาลราชวิถี โดยใช้แบบสอบถามเก็บข้อมูลจากผู้เชี่ยวชาญจำนวน 5 คน ได้ค่าเฉลี่ยเท่ากับ 4.05 และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.48 ดังนั้น สามารถสรุปได้ว่าความเร็วในการใช้งานต่อระบบงานของทางโรงพยาบาลราชวิถีมีระดับความพึงพอใจอยู่ในระดับมาก

3. แนวคิด/วิธีการที่นำเสนอ

3.1 การออกแบบระบบทางด้านฮาร์ดแวร์



รูปที่ 6. สถาปัตยกรรมของระบบ

จากรูปที่ 6 ระบบประกอบด้วยฮาร์ดแวร์ ได้แก่ 1) Pi camera 2) บอร์ดราสเบอร์รี่พาย จำนวน 2 ตัว ทำหน้าที่บันทึกวีดิโอและทำหน้าที่เป็นเซิร์ฟเวอร์เก็บไฟล์วีดีโอ 3) อะแดปเตอร์ TL-PA4010 ทำหน้าที่ส่งไฟล์ผ่านสายไฟฟ้า และ 4) External HDD สำหรับเก็บไฟล์วีดิโอ

3.2 การออกแบบฐานข้อมูล

ตารางที่ 1. สคีมาของตาราง video

แอททริบิวต์	ประเภท	หมายเหตุ
filename	varchar (100)	PK เก็บชื่อไฟล์
period	datetime	วัน/เวลาที่บันทึก
status	tinyint	สถานะการอัพโหลด
rec_count	bigint	ลำดับที่ของไฟล์
filesize	int	ขนาดของไฟล์

3.3 การติดตั้ง External HDD บนราสเบอร์รี่พาย (2)

ติดตั้ง External HDD ดังรูปที่ 7 แล้วติดตั้งซอฟต์แวร์

รูปที่ 7. การติดตั้ง External HDD กับราสเบอร์รี่พาย

ขั้นตอนการติดตั้งมีดังนี้

1) พิมพ์คำสั่ง df -h ถ้าการเชื่อมต่อสมบูรณ์จะพบว่ามี External HDD เพิ่มในระบบ จากตัวอย่างได้แก่ /dev/sda1 ซึ่งถูก mount ในเส้นทาง /media/pi/Elements

/dev/sda1 932G 147G 785G 16% /media/pi/Elements

2) ติดตั้งซอฟต์แวร์เพิ่มเติมเพื่อให้สามารถเขียนข้อมูลลง External HDD (ค่าเริ่มต้นสามารถอ่านได้อย่างเดียว) โดยพิมพ์คำสั่ง ดังนี้

sudo apt-get install ntfs-3g

sudo reboot

หลังจากรีสตาร์ททดลองสร้างไดเรกทอรีหรือไฟล์ใน /media/pi/Elements ถ้าสามารถสร้างได้ถือว่าผ่านขั้นตอนนี้

3) พิมพ์คำสั่งสร้างไดเรกทอรีเพื่อรองรับการ mount

sudo mkdir /media/nas

sudo chmod -R 777 /media/nas

4) แก้ไขไฟล์ /etc/fstab เพื่อติดตั้ง External HDD อย่างถาวร โดยเพิ่ม /dev/sda1 /media/nas auto defaults,user 0 1 เข้าไปท้ายไฟล์ โดยพิมพ์คำสั่ง sudo nano /etc/fstab จากนั้นสั่งรีสตาร์ทราสเบอร์รี่พายด้วยคำสั่ง sudo reboot

```
proc            /proc           proc    defaul
PARTUUID=7d842044-01  /boot           vfat    
PARTUUID=7d842044-02  /               ext4    
# a swapfile is not a swap partition, no line h
#   use  dphys-swapfile swap[on|off]  for that
/dev/sda1 /media/nas auto defaults,user 0 1
```

รูปที่ 8. การแก้ไขไฟล์ /etc/fstab

หลังจากรีสตาร์ทราสเบอร์รี่พาย พิมพ์คำสั่ง df -h จะพบว่า External HDD ถูก mount ที่เส้นทาง /media/nas

/dev/sda1 932G 147G 785G 16% /media/nas

3.4 การติดตั้ง FTP server บนราสเบอร์รี่พาย (2)

1) พิมพ์คำสั่ง sudo apt install proftpd เพื่อติดตั้ง proftpd

2) พิมพ์คำสั่ง sudo nano /etc/proftpd/proftpd.conf เพื่อแก้ไขไฟล์คอนฟิกของ proftpd แล้วแก้ไขข้อมูลดังนี้

DefaultRoot /media/nas

หมายเหตุ /media/nas คือเส้นทางที่ชี้ไปยัง External HDD

3) พิมพ์คำสั่ง sudo service proftpd reload เพื่อรีสตาร์ทเอฟทีพีเซอร์วิส

4) ทดสอบการใช้งานเอฟทีพีจากลูกข่าย

3.5 การติดตั้ง NFS server บนราสเบอร์รี่พาย (2)

เนื่องจากวัตถุประสงค์ของการวิจัยข้อที่ 2 คือต้องการเปรียบเทียบประสิทธิภาพการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ระหว่างเอฟทีพีกับเอ็นเอฟเอส จึงจำเป็นต้องติดตั้งเอ็นเอฟเอส ซึ่งมีขั้นตอนดังนี้

1) พิมพ์คำสั่ง sudo apt-get install nfs-common nfs-kernel-server

2) พิมพ์คำสั่ง sudo nano /etc/exports เพื่อกำหนดสิทธิ์สำหรับลูกข่ายที่ต้องการเชื่อมต่อเข้าใช้งานเอ็นเอฟเอส แล้วกำหนดสิทธิ์ดังนี้

/media/nas 10.39.0.25(rw,sync,no_subtree_check)

ให้สิทธิ์ลูกข่ายไอพีแอดเดรส 10.39.0.25 สามารถเรียกใช้งานเส้นทาง /media/nas บนเอ็นเอฟเอสเซิร์ฟเวอร์ มีสิทธิ์ อ่าน, เขียน

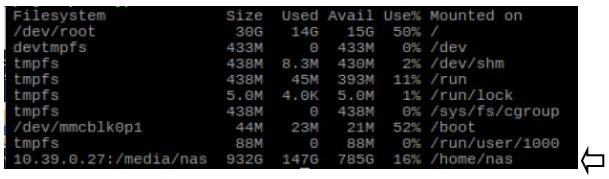
3) พิมพ์คำสั่ง sudo service nfs-kernel-server start เพื่อสตาร์ทเอ็นเอฟเอสเซิร์ฟเวอร์

3.6 การติดตั้ง NFS client บนราสเบอร์รี่พาย (1)

1) พิมพ์คำสั่ง sudo apt-get install nfs-common เพื่อติดตั้ง

2) ทดลองเชื่อมต่อกับ NFS server โดยใช้คำสั่ง sudo mount 10.39.0.27:/media/nas /home/nas

3) ตรวจสอบโดยพิมพ์คำสั่ง df -h ผลลัพธ์ดังรูปที่ 9



รูปที่ 9. ผลการติดต่อเอ็นเอฟเอสเซิร์ฟเวอร์

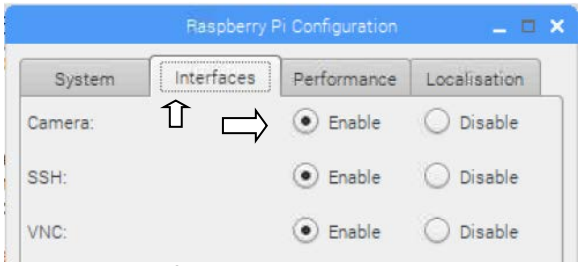
3.7 ติดตั้ง web server (apache, PHP, MySQL) บนราสเบอร์รี่พาย (1) และ (2)

1) ขั้นตอนการติดตั้งสามารถสืบค้นจากอินเทอร์เน็ต ผู้เขียนไม่ขอกล่าวถึงในบทความนี้

2) สร้างฐานข้อมูลใน MySQL และสร้างตารางตามข้อที่ 3.2 (ชื่อฐานข้อมูลและชื่อตารางดูในโค้ดภาษาไพทอน)

3.8 การเปิดใช้งาน Pi camera บนราสเบอร์รี่พาย (1)

คลิกเมนู Raspberry Pi Configuration --> Interfaces



รูปที่ 10. การเปิดใช้งาน Pi camera

3.9 การพัฒนาโปรแกรมบันทึกวีดีโอด้วยภาษา python บนราสเบอร์รี่พาย (1)

1) ติดตั้งซอฟต์แวร์สำหรับติดต่อ Pi camera

sudo apt-get update

sudo apt-get install python3-picamera

2) ติดตั้งซอฟต์แวร์สำหรับติดต่อฐานข้อมูล MySQL

sudo apt-get install python-mysql.connector

3) โค้ดภาษาไพทอน สามารถดาว์นโหลดจาก google docs https://docs.google.com/document/d/1PSghmjktavH8FKuKKjebxcwLufagT86ome_OMA-AGgM/edit?usp=sharing

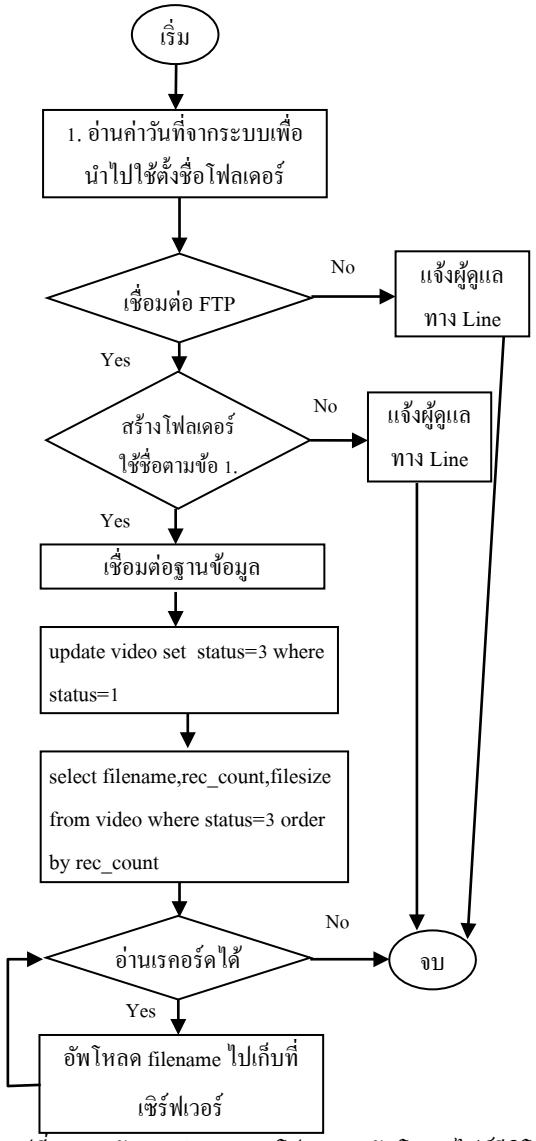
หลักการทำงานของโปรแกรมดังรูปที่ 11

รูปที่ 11. ผังงานหลักการทำงานของโปรแกรมบันทึกวีดิโอ

หมายเหตุ ต้องการให้ราสเบอร์รี่พายรันโปรแกรมอัตโนมัติให้เพิ่มคำสั่ง /usr/bin/python3 /home/pi/rec_video.py & ในไฟล์ /etc/rc.local

3.10 การพัฒนาโปรแกรมภาษา PHP สำหรับอัพโหลดวีดีโอไปเก็บบนเซิร์ฟเวอร์ ในราสเบอร์รี่พาย (1)

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพการสำเนาไฟล์ผ่านเครือข่ายเพาเวอร์ไลน์ระหว่างเอ็นเอฟเอสกับเอฟทีพี ดังนั้น การพัฒนาโปรแกรมภาษา PHP สำหรับอัพโหลดไฟล์วีดิโอไปเก็บบนเซิร์ฟเวอร์จึงพัฒนา 2 โปรแกรม ได้แก่ โปรแกรมสำหรับใช้บริการเอ็นเอฟเอส (copyvideo.php) และโปรแกรมสำหรับใช้บริการเอฟทีพี (copyvideoFTP.php) โค้ดโปรแกรมสามารถศึกษาได้จาก https://docs.google.com document/d/1VnY2gfuggWtoDZQeBDu5qtu7UAGLRYdNvMdL3l_OUM8/edit?usp=sharing หลักการทำงานของโปรแกรม copyvideoFTP.php ดังรูปที่ 12



รูปที่ 12. หลักการทำงานของโปรแกรมอัพโหลดไฟล์วีดิโอ

หมายเหตุ ตัวอย่างโค้ดส่งข้อความทาง Line ศึกษาจากโปรแกรม linenotify.php บน google docs

3.11 การตั้งเวลาอัพโหลดไฟล์อัตโนมัติ

แก้ไขไฟล์ crontab เพื่อตั้งเวลาให้ระบบปฏิบัติการ linux ทำงานโดยอัตโนมัติ โดยพิมพ์คำสั่ง sudo nano /etc/crontab แล้วพิมพ์ข้อความต่อไปนี้ต่อท้ายไฟล์

```
56 23 * * * root /usr/bin/curl http://localhost/copyvideoFTP.php
58 23 * * * root /usr/bin/curl http://localhost/copyvideo.php
```

23.56 น. ของทุกวัน ให้รันไฟล์ copyvideoFTP.php และ 23.58 น. ของทุกวัน ให้รันไฟล์ copyvideo.php

3.12 การตรวจสอบสถานะของราสเบอร์รี่พาย (1)

หน้าที่ของราสเบอร์รี่พาย (1) คือบันทึกวีดิโอ ดังนั้น ถ้าราสเบอร์รี่พาย (1) หยุดทำงานด้วยสาเหตุใดก็ตาม จะส่งผลให้การบันทึกวีดิโอมีความผิดพลาด ผู้วิจัยจึงมีแนวคิดในการมอบหน้าที่ในการตรวจสอบสถานะของราสเบอร์รี่พาย (1) ด้วยราสเบอร์รี่พาย (2) กล่าวคือ ถ้าผลการตรวจสอบพบว่าราสเบอร์รี่พาย (1) หยุดทำงานให้ราสเบอร์รี่พาย (2) ส่ง Line แจ้งผู้ดูแล วิธีการคือแก้ไขไฟล์ crontab เพื่อตั้งเวลาให้ระบบปฏิบัติการ linux ทำงานโดยอัตโนมัติ โดยพิมพ์คำสั่ง sudo nano /etc/crontab แล้วพิมพ์ข้อความต่อไปนี้ต่อท้ายไฟล์

```
56 23 * * * root /usr/bin/curl http://localhost/checkserver.php
```

หมายเหตุ 23.58 น. ของทุกวันให้รันโปรแกรม checkserver.php โดยอัตโนมัติ โค้ดโปรแกรม checkserver.php ศึกษาได้จาก google docs

4. ผลการทดลอง

ผู้วิจัยได้ทดลองบันทึกไฟล์วีดิโอ 4 ความยาว ได้แก่ ความยาวไฟล์ละ 1 นาที ไฟล์ละ 2 นาที ไฟล์ละ 4 นาที และไฟล์ละ 6 นาที แต่ละความยาวบันทึกจำนวน 12 วัน แต่ละวันทำการอัพโหลดไฟล์วีดิโอขึ้นไปเก็บบนราสเบอร์รี่พาย (2) โดยใช้บริการเอฟทีพีและเอ็นเอฟเอส ผลการทดลองดังตารางที่ 1-4

ตารางที่ 1. ผลการเปรียบเทียบประสิทธิภาพความยาว 1 นาที

วันที่	ความจุ (ต่อวัน)	จำนวนไฟล์	FTP (นาที)	NFS (นาที)
1	8,146,457,758	1,431	26.34	25.06
2	11,730,877,603	1,424	37.47	34.01
3	9,240,075,551	1,429	29.40	27.35
4	7,619,543,725	1,432	24.55	23.18
5	10,173,030,850	1,429	32.44	29.25
6	7,703,019,615	1,429	25.09	23.48
7	6,906,720,190	1,430	22.13	21.48
8	6,807,964,651	1,431	21.55	21.56
9	10,696,627,014	1,426	34.09	32.14

วันที่	ความจุ (ต่อวัน)	จำนวนไฟล์	FTP (นาที)	NFS (นาที)
10	9,149,806,031	1,429	29.33	27.52
11	15,911,374,349	1,420	50.43	45.38
12	10,371,372,300	1,428	34.24	29.06
		เฉลี่ย	**30.59**	**28.29**

จากตารางที่ 1 พบว่า วันที่ 8 NFS ใช้เวลาในการอัพโหลดไฟล์มากกว่า FTP 1 วินาที นอกนั้นพบกว่า NFS ใช้เวลาน้อยกว่า FTP

ตารางที่ 2. ผลการเปรียบเทียบประสิทธิภาพความยาว 2 นาที

วันที่	ความจุ (ต่อวัน)	จำนวนไฟล์	FTP (นาที)	NFS (นาที)
1	8,482,291,571	717	26.57	23.31
2	8,761,548,642	718	27.38	23.53
3	8,607,930,881	717	27.12	23.40
4	8,189,889,354	718	25.57	22.15
5	7,819,875,934	717	24.58	21.23
6	7,647,823,224	718	24.16	21.16
7	7,878,882,666	718	25.34	22.21
8	8,638,972,846	717	27.35	24.44
9	9,492,999,187	717	29.56	28.02
10	7,251,961,444	718	23.08	20.58
11	8,316,192,281	718	26.20	23.15
12	6,817,998,454	717	21.39	20.20
		เฉลี่ย	**25.69**	**22.78**

จากตารางที่ 2 พบว่า NFS ใช้เวลาน้อยกว่า FTP ทุกวัน

ตารางที่ 3. ผลการเปรียบเทียบประสิทธิภาพความยาว 4 นาที

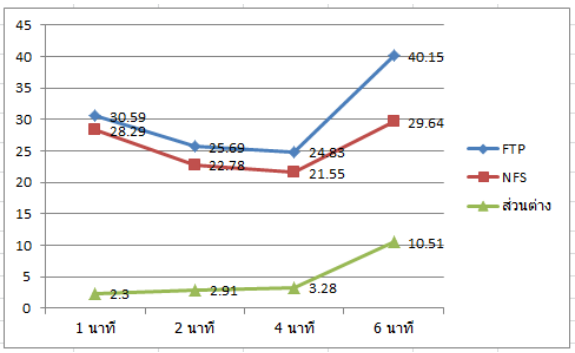
วันที่	ความจุ (ต่อวัน)	จำนวนไฟล์	FTP (นาที)	NFS (นาที)
1	6,698,633,754	359	21.14	18.35
2	8,674,018,402	359	27.33	24.29
3	8,711,801,235	359	27.29	23.47
4	7,572,569,967	360	23.57	21.03
5	7,713,161,176	359	24.18	21.14
6	8,898,969,222	359	27.55	24.22
7	8,544,552,234	360	27.06	23.35
8	6,685,499,651	359	21.04	18.44
9	5,418,057,734	359	17.11	15.14
10	8,740,200,492	360	27.32	23.25
11	9,310,276,047	359	29.27	24.51
12	8,003,329,961	359	25.07	21.40
		เฉลี่ย	**24.83**	**21.55**

จากตารางที่ 3 พบว่า NFS ใช้เวลาน้อยกว่า FTP ทุกวัน

ตารางที่ 4. ผลการเปรียบเทียบประสิทธิภาพความยาว 6 นาที

วันที่	ความจุ (ต่อวัน)	จำนวนไฟล์	FTP (นาที)	NFS (นาที)
1	7,724,526,348	240	24.40	20.16
2	8,276,782,717	239	26.14	21.54
3	6,762,632,282	240	21.20	17.57
4	7,085,316,735	240	22.29	18.23
5	7,917,416,227	240	25.29	19.45
6	9,798,376,609	239	31.11	22.00
7	15,926,896,314	239	59.00	40.36
8	17,554,358,229	240	66.00	46.24
9	13,988,773,415	239	51.41	38.58
10	15,677,782,713	240	49.49	38.50
11	13,675,995,134	239	44.25	34.04
12	16,023,143,049	240	61.16	39.01
		เฉลี่ย	**40.15**	**29.64**

จากตารางที่ 4 พบว่า NFS ใช้เวลาน้อยกว่า FTP ทุกวัน



รูปที่ 13. สรุปผลการทดลอง

5. บทสรุป

ผู้วิจัยขอสรุปผลการทดลองตามวัตถุประสงค์ของการวิจัยดังนี้

5.1 ระบบที่พัฒนาขึ้นประกอบไปด้วยฮาร์ดแวร์ ได้แก่ 1) Pi camera จำนวน 1 ตัว 2) บอร์ดราสเบอร์รี่พาย รุ่น 3 B จำนวน 2 ตัว ทำหน้าที่บันทึกไฟล์วีดิโอและทำหน้าที่เป็นเซิร์ฟเวอร์เก็บไฟล์วีดีโอ 3) อะแดปเตอร์ TL-PA4010 ทำหน้าที่ส่งไฟล์วีดิโอผ่านสายไฟฟ้า 4) External HDD สำหรับเก็บไฟล์วีดิโอ และ 5) สายไฟฟ้าแบบ VAF ขนาด 2x1 ม.ม. จำนวน 300 เมตร จากการทดลองบันทึกไฟล์วีดิโอ จำนวน 48 วัน (ความยาวของไฟล์ 1 นาที 2 นาที 4 นาที และ 6 นาที ความยาวละ 12 วัน) และอัพโหลดไปเก็บในบอร์ดราสเบอร์รี่พาย (2) สามารถใช้งานได้จริง สามารถส่งข้อความแจ้งเตือนผ่าน line notify เมื่อระบบเกิดความผิดพลาดได้

5.2 จากภาพที่ 13 พบว่า NFS มีประสิทธิภาพในการสำเนาไฟล์ผ่านเครือข่ายคอมพิวเตอร์ได้ดีกว่า FTP และถ้าขนาดของไฟล์วีดิโอมีขนาดใหญ่ขึ้น ส่วนต่างของเวลาในการสำเนาไฟล์ก็จะมากขึ้นตามไปด้วย

6. การอภิปรายผล

6.1 ผู้วิจัยได้ทำการสำรวจราคาระบบกล้องวงจรปิดแบบ IP camera ที่รองรับกล้องจำนวน 4 ตัว ความละเอียด 1,080p ที่จำหน่ายตามท้องตลาด พบว่า ราคาขั้นต่ำ 30,000 บาท (ไม่รวมค่าติดตั้ง) สำหรับงานวิจัยนี้ใช้งบประมาณดังตารางที่ 5 ซึ่งสรุปได้ว่าราคาต่ำกว่าระบบที่จำหน่ายตามท้องตลาด

ตารางที่ 5. ค่าใช้จ่ายในการติดตั้งระบบ

ที่	รายการ	จำนวน	ราคารวม
1	บอร์ดราสเบอร์รี่พาย 3 B	2	2,720
2	Adapter Power Supply 5V 2.5A	2	900
3	Micro SD 16 GB	2	260
4	Pi camera	1	1,090
5	PLC ยี่ห้อ tp-link รุ่น TL-PA4010	1	888
6	WD External HDD 1 TB	1	999
		รวม	6,857

หมายเหตุ ค่าใช้จ่ายตามตารางที่ 5 สำหรับกล้องจำนวน 1 ตัว กรณีต้องการเพิ่มจำนวนกล้อง ค่าใช้กล้องละ 1,360 + 450 + 130 + 1,090 + 888 = 3,918 บาท หรือถ้าต้องการใช้งานกล้องจำนวน 4 ตัว งบประมาณ 18,611 บาท ซึ่งราคาต่ำกว่าระบบที่จำหน่ายตามท้องตลาด

6.2 ถ้าระบบเกิดความผิดพลาด เช่น บอร์ดราสเบอร์รี่พาย (1) ไม่สามารถสื่อสารกับบอร์ดราสเบอร์รี่พาย (2) ด้วยสาเหตุใด ๆ ก็ตาม ระบบสามารถส่งข้อความแจ้งเตือนผู้ดูแลระบบผ่าน Line notify ได้ เพื่อแก้ไขปัญหาซึ่งจะทำให้การบันทึกวีดิโอไม่เกิดความผิดพลาดหรือผิดพลาดน้อยที่สุด

เอกสารอ้างอิง

[1] ม.ป.ป., “วิธีการทำงานของกล้อง IP Network Camera,” ม.ป.ป., 2562. [Online]. Available: http://www.cctvpool.com/index.php?lay=show&ac=article&Id=539898658. [Accessed: 20 ธันวาคม 2562].

[2] มลฑล คล้ายสวัสดิ์, “การเปรียบเทียบประสิทธิภาพการสตรีมมิ่งไฟล์วีดีโอความละเอียดสูงระหว่างเครือข่ายไร้สายกับเครือข่ายที่ใช้สายไฟฟ้า,” ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศ, คณะเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[3] เกรียงไกร พงษ์มนตรี, “การวัดประสิทธิภาพในการโอนถ่ายข้อมูลผ่าน Powerline กรณีศึกษาในห้องเร่งอนุภาครังสีรักษา ณ โรงพยาบาลราชวิถี,” ปัญหาพิเศษปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศ, คณะเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2554.

[4] วินัส ซัพพลาย, “ตอนที่ 1 รู้จักกับบอร์ด Raspberry Pi พร้อมเรียนรู้วิธีการติดตั้งระบบปฏิบัติการ Linux ให้กับ Respberry Pi,” 7 ตุลาคม 2559. [Online]. Available: https://www.thaieasyelec.com/article-wiki/embedded-electronics-application/บทความการพัฒนาโปรแกรมบน-raspberry-pi-ด้วย-qt.html. [Accessed: 20 ธันวาคม 2562].

[5] วโรจน์ สกุลโต, “WLAN VS Powerline,” CHIP Computer &Communications, ปีที่ 10, ฉบับที่ 6, มิถุนายน, pp. 72-74, 2554.

[6] n.d., “AV500 Nano Powerline Adapter TL-PA4010,” n.d., January, 2020. [Online]. Available: https://www.tp-link.com/us/home-networking/powerline/tl-pa4010/#overview. [Accessed: January 12, 2020].

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 39-47

ระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ

Intelligent Key Storage Management System with Automatic Identification

อรรถพล พลานนท์ และ ขนิษฐา แซ่ลิ้ม*

Attapon Palananda and Khanittha Saelim*

สาขาวิชาเทคโนโลยีคอมพิวเตอร์อุตสาหกรรม คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยราชภัฏนครปฐม

Industrial Computer Technology, Faculty of Science and Technology,

Nakhon Pathom Rajabhat University

Received: February 03, 2020; **Revised:** May 08, 2020; **Accepted:** May 09, 2020; **Published:** June 23, 2020

-

ABSTRACT – In this paper, we present the intelligent key storage management system by the automatic identification method. The device can store a maximum of 48 room keys. The borrowing key system uses automated biometric identification and a key return system uses an RFID-based automated identification method. These methods can identify the borrowed person and reference the keys of each room with identification tags to store the keys in the box. The operation system is controlled through a program window that can add, edit, and delete keys for each room. In disbursement of the keys for use, users must register to verify their fingerprints to determine the right to use the system. After successfully register in the system, Users can choose a room and borrow the keys with their fingerprints. After that, when the room has been used successfully, the user can return the key to the key return box. The system reads the room number from the identification tag and stores it in the key storage box. In our experiment, we used a questionnaire was used to collect data to survey the satisfaction of 50 users and compared the performance in two parts: measuring user satisfaction and measuring system performance. In the first part, the results of the development of this system showed that they were satisfied with the convenience and ease of use and satisfaction with the data security system and satisfaction with the design and the beauty with an average of 96.133, 88.2, and 78.4, respectively. In the second part, the proposed system compared with the staff. The results of experiments show that the developed system is faster than the staff about 3.681 seconds. In the same direction, the proposed system is faster than the staff in the average time when returning the key about 2.089 seconds.

KEYWORDS: Automatic Identification, Fingerprint, RFID, Key Management System

บทคัดย่อ -- ในงานวิจัยนี้ได้นำเสนอระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ โดยอุปกรณ์สามารถเก็บกุญแจได้สูงสุดที่ 48 ห้อง โดยในระบบการเบิกจ่ายกุญแจนั้นใช้วิธีการบ่งชี้อัตโนมัติแบบชีวมาตรเข้าช่วยในการทำงาน แต่ในส่วนของการคืนกุญแจใช้วิธีการบ่งชี้อัตโนมัติแบบอาร์เอฟไอดีที่สามารถระบุถึงบุคคลที่ทำการยืมได้ และสามารถอ้างอิงกุญแจของแต่ละห้องด้วยป้ายระบุตัวตนเพื่อทำการจัดเก็บกุญแจลงกล่องได้ โดยระบบการทำงานมีการควบคุมผ่านหน้าต่างโปรแกรมที่สามารถสามารถเพิ่ม แก้ไข และลบข้อมูลของกุญแจแต่ละห้องได้ และการเบิกจ่ายกุญแจไปใช้งานต้องผ่านการลงทะเบียนยืนยันด้วยลายนิ้วมือในครั้งแรกเพื่อสร้างสิทธิของผู้ใช้งาน ผู้ใช้งานจึงสามารถเลือกห้องและทำการเบิกจ่ายกุญแจด้วยลายนิ้วมือได้ จากนั้นเมื่อมีการใช้งานห้องเรียบร้อย ผู้ใช้สามารถคืนกุญแจในช่องคืนกุญแจ จากนั้นระบบทำการอ่านหมายเลขห้องจากป้ายระบุตัวตนและจัดเก็บในช่องเก็บกุญแจต่อไป โดยจากการนำไปใช้งานจริงมีผู้ใช้งานอยู่ในระบบจำนวน 50 ผู้ใช้งาน มีค่าระดับความพึงพอใจในส่วนการใช้งานระบบคิดเป็นค่าเฉลี่ยรวมอยู่ที่ร้อยละ 96.133 สำหรับ

*Corresponding Author: *patthamanan.s@nsru.ac.th*

ด้านความปลอดภัยของข้อมูลคิดเป็นค่าเฉลี่ยความพึงพอใจรวมอยู่ที่ร้อยละ 88.2 และด้านการออกแบบและความสวยงามคิดเป็นค่าเฉลี่ยความพึงพอใจรวมอยู่ที่ร้อยละ 78.4 ตามลำดับ สำหรับการหาค่าประสิทธิภาพด้านการทำงานเปรียบเทียบกับการทำงานของเจ้าหน้าที่ในส่วนการเบิกจ่ายกุญแจระบบที่พัฒนาขึ้นมามีค่าเวลาเฉลี่ยเวลาที่ดีกว่าอยู่ที่ 3.681 วินาที และในส่วนการคืนกุญแจระบบที่พัฒนาขึ้นมามีค่าเวลาเฉลี่ยที่ดีกว่าอยู่ที่ 2.089 วินาที ตามลำดับ

คำสำคัญ: วิธีการบ่งชี้อัตโนมัติ, เครื่องสแกนลายนิ้วมือ, คลื่นความถี่วิทยุ, ระบบบริหารจัดการกุญแจแบบอัตโนมัติ

1. บทนำ

ในปัจจุบันอาคารหรือสำนักงานมีขนาดใหญ่ มักมีปริมาณห้องจำนวนมากอยู่ภายในอาคารนั้นและมีการป้องกันการเข้าใช้งานห้อง โดยพื้นฐานของการป้องกันการเข้าใช้งานห้องคือ การปิดล็อคประตูสำหรับการเข้า-ออก ห้องนั้น ๆ การเข้าใช้งานผู้ใช้ต้องมีกุญแจสำหรับห้องนั้นจึงสามารถเปิดเพื่อใช้งานได้ หากอาคารหรือสำนักงานนั้นมีห้องจำนวนมากการเก็บรักษากุญแจหรือการค้นหากุญแจเพื่อนำไปเปิดใช้งานห้องนั้น อาจทำได้ยากขึ้นตามจำนวนห้องและกุญแจที่เพิ่มขึ้น ซึ่งปัญหาหนึ่งของการมีจำนวนห้องที่มากขึ้นของอาคารสำนักงาน คือการค้นหากุญแจของแต่ละห้องไม่เจอ ทำให้เกิดความล่าช้าในการค้นหากุญแจห้องที่ต้องการโดยกรณีที่ดีที่สุดคือ ค้นหาแล้วเจอทันที และกรณีเลวร้ายที่สุดคือค้นหาแล้วไม่เจอกุญแจที่ต้องการ ซึ่งอาจสูญหายหรือไม่ส่งคืนกุญแจ ซึ่งในกรณีนี้ยังไม่สามารถติดตามหรือหาผู้รับผิดชอบได้ ดังนั้นเพื่อแก้ปัญหาการตรวจสอบสถานะของกุญแจว่ามีอยู่หรือมีการนำออกไปใช้งานจึงมีความสำคัญ วิธีการหนึ่งที่นำมาใช้แก้ไขปัญหานี้คือการนำเอาเทคโนโลยีระบบบ่งชี้อัตโนมัติ (Auto–ID: Automatic Identification) มาใช้งานคือการระบุสิ่งของด้วยคลื่นความถี่วิทยุ (RFID: Radio Frequency Identification) [1,2] ซึ่งได้ถูกนำมาประยุกต์ใช้งานร่วมกับระบบต่าง ๆ เช่น ระบบห้องสมุดอัจฉริยะ (Smart Library) ที่มีการยืมคืนหนังสือผ่านการระบุสิ่งของด้วยคลื่นความถี่วิทยุหรือการประยุกต์ใช้งานร่วมกับสายการผลิตในโรงงาน (Production Line Automation) [3] เพื่อตรวจสอบในแต่ละขั้นตอนของการทำงานหรือการตรวจสอบบุคคลด้วยข้อมูลชีวมาตร (Biometrics) ซึ่งเป็นการพิสูจน์เอกลักษณ์ของแต่ละบุคคล

สำหรับอาคารวิศวกรรมและเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม เป็นอาคารที่มีห้องเป็นจำนวนมาก โดยในการใช้งานห้องจำเป็นต้องมีการเบิกจ่ายกุญแจและคืนกุญแจที่ห้องสโตร์ ซึ่งในบางครั้งอาจเกิดความไม่สะดวกในการใช้งานขึ้น เช่น ในกรณีที่เจ้าหน้าที่ติดภารกิจอื่นทำให้ไม่สามารถให้บริการแก่นักศึกษาได้ ส่งผลให้นักศึกษาต้องรอเจ้าหน้าที่เพื่อทำการขอเบิกกุญแจ หรือปัญหาของนักศึกษากรณีที่เบิกกุญแจไปใช้งานแต่ไม่มีการคืนกุญแจ ซึ่งกรณีที่นักศึกษาใช้งานห้องเสร็จและไม่นำกุญแจส่งคืนเจ้าหน้าที่ หรือทำกุญแจสูญหายจะไม่สามารถติดตามหาความรับผิดชอบได้

ดังนั้นในงานวิจัยนี้จึงขอเสนอระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยเทคโนโลยีระบบบ่งชี้อัตโนมัติ แบบการระบุสิ่งของจากคลื่นความถี่วิทยุและแบบชีวมาตร ร่วมกับอุปกรณ์สำหรับการจัดเก็บกุญแจ [4] สำหรับอาคารวิศวกรรมและเทคโนโลยี โดยอุปกรณ์ควบคุมทำงานร่วมกับระบบฐานข้อมูลผู้ใช้งานผ่านโปรแกรมควบคุมการทำงาน เพื่อลดปัญหาความล่าช้าในการค้นหากุญแจและการสูญหายของกุญแจ อีกทั้งยังสามารถติดตามสถานะของกุญแจได้อีกด้วย

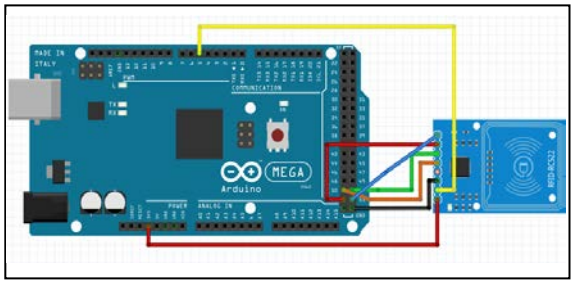
2. ทฤษฎีที่เกี่ยวข้อง

2.1 เทคโนโลยีระบบบ่งชี้อัตโนมัติ

ระบบบ่งชี้อัตโนมัติ คือระบบที่ถูกพัฒนาขึ้นมาเพื่อใช้ระบุตัวตนของมนุษย์หรือสิ่งมีชีวิตอื่น รวมถึงสินค้าและวัตถุดิบที่ใช้ในการผลิต ซึ่งระบบบ่งชี้อัตโนมัติได้พัฒนารูปแบบต่าง ๆ เพื่ออำนวยความสะดวกต่อการทำงาน โดยเฉพาะการบันทึกข้อมูลแบบอัตโนมัติ แทนการนับหรือการจดบันทึกด้วยมือของมนุษย์ซึ่งมีโอกาสในการทำงานผิดพลาดสูงกว่า

2.1.1 ระบบการระบุสิ่งของด้วยคลื่นความถี่วิทยุ

ระบบการระบุสิ่งของด้วยคลื่นความถี่วิทยุ คือ การระบุวัตถุหรือสิ่งของด้วยคลื่นความถี่วิทยุ ซึ่งประกอบด้วย อุปกรณ์อ่านและเขียน (RFID Reader and Writer) [5,6] อุปกรณ์ไมโครคอนโทรลเลอร์ และป้ายระบุตัวตน (Tag) ดังรูปที่ 1.

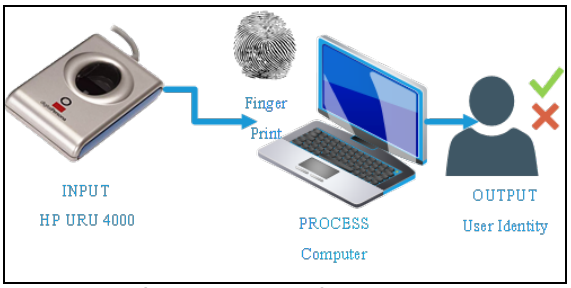


รูปที่ 1. การระบุสิ่งของด้วยคลื่นความถี่วิทยุ

จากรูปที่ 1. เป็นการต่ออุปกรณ์ไมโครคอนโทรลเลอร์ร่วมกับอุปกรณ์เครื่องอ่านและเขียน โมดูล RFID-RC522 โดยอุปกรณ์ทำการส่งผ่านข้อมูลทางการเชื่อมต่อสื่อสารแบบอนุกรมโดยอาศัยสัญญาณนาฬิกา (SPI: Serial Peripheral Interface) ไปยังไมโครคอนโทรลเลอร์เพื่อใช้สำหรับอ่านและเขียนข้อมูลสำหรับป้ายระบุตัวตน

2.1.2 ระบบชีวมาตร

ระบบชีวมาตร คือเทคโนโลยีที่มีการนำเอาลักษณะทางกายภาพของคนมาใช้ในการยืนยันตัวตน ทำให้สามารถยืนยันตัวตนของบุคคลนั้น โดยข้อมูลทางกายภาพของคนหนึ่งคน จะมีลักษณะที่ไม่เหมือนกับบุคคลอื่น ๆ เช่น ข้อมูลภาพใบหน้า ข้อมูลม่านตา หรือข้อมูลลายนิ้วมือ เป็นต้น ซึ่งจากการศึกษาพบว่าเทคโนโลยีของการตรวจวัดด้วยระบบชีวมาตรแบบการตรวจสอบข้อมูลด้วยลายนิ้วมือ เป็นวิธีที่นิยมนำมาใช้ในการยืนยันตัวบุคคลโดยรูปแบบการต่ออุปกรณ์ดังรูปที่ 2.

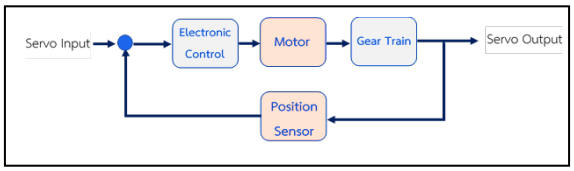


รูปที่ 2. แบบอุปกรณ์เครื่องจัดเก็บกุญแจ

การยืนยันตัวบุคคลด้วยลายนิ้วมือนั้น ได้รับการพิสูจน์และยอมรับจากงานวิจัยที่เกี่ยวข้อง เช่น การนำอุปกรณ์ตรวจสอบลายนิ้วมือมาประยุกต์ใช้ในระบบการยืมคืนหนังสือ [7] หรือนำไปใช้ในการตรวจสอบนักศึกษาในการเข้าชั้นเรียน [8] เป็นต้น ซึ่งโอกาสที่ลายนิ้วมือจะซ้ำกันได้นั้น อยู่ที่ 1 ใน 64 พันล้านคนถึงจะมีโอกาสที่ลายนิ้วมือซ้ำกัน 1 คู่เท่านั้น ซึ่งถือว่าลายเส้นบนนิ้วมือนับเป็นเอกลักษณ์เฉพาะบุคคลนั้น ๆ ได้ สำหรับการนำอุปกรณ์ตรวจสอบลายนิ้วมือมาใช้ เป็นวิธีการที่สามารถแยกแยะและนำมาใช้ยืนยันตัวบุคคลได้ โดยนำอุปกรณ์อ่านลายนิ้วมือมาต่อเข้ากับเครื่องคอมพิวเตอร์เพื่อใช้งานได้โดยตรงผ่านช่องทางสื่อสารของคอมพิวเตอร์

2.2 ระบบควบคุมเซอร์โวมอเตอร์ (Servo Motors)

อุปกรณ์เซอร์โวมอเตอร์ คือมอเตอร์ที่สามารถสั่งงานหรือตั้งค่าได้ โดยมีการหมุนเป็นมุมองศาตามแกนหมุนของมอเตอร์ ในการควบคุมการทำงานใช้ระบบการควบคุมแบบป้อนกลับ (Feedback Control) ดังรูปที่ 3.



รูปที่ 3. การทำงานของเซอร์โวมอเตอร์

โดยที่ Motor คือ มอเตอร์

Electronic Control คือ ส่วนควบคุมและประมวลผลมอเตอร์

Gear Train คือ ชุดเกียร์สำหรับทดแรงขับ

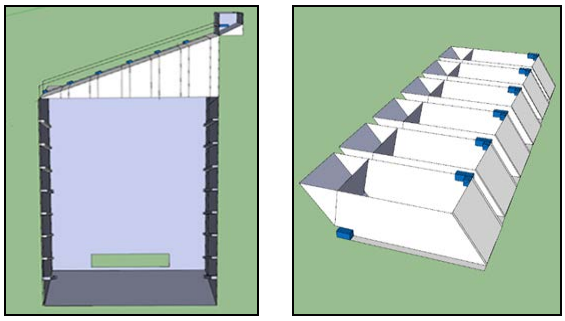
Position Sensor คือ ตัวรับรู้ตรวจจับตำแหน่งการหมุนเพื่อหาองศาการหมุน

หลักการทำงานของเซอร์โวมอเตอร์ประกอบด้วย เมื่อมีสัญญาณพัลซ์ส่งเข้ามายังส่วนควบคุมการทำงานภายในเซอร์โวมอเตอร์

2.3 อุปกรณ์เครื่องจัดเก็บกุญแจ

งานชิ้นนี้เป็นการพัฒนาต่อยอดจากงานวิจัยระบบการจัดเก็บกุญแจภายในอาคารวิศวกรรมและเทคโนโลยี มหาวิทยาลัยราช-ภัฏนครปฐม โดยพัฒนาจากเครื่องต้นแบบที่สามารถเก็บกุญแจได้เพียง 4 ช่องเก็บ และไม่สามารถเพิ่มจำนวนช่องเก็บกุญแจได้ จากนั้นจึงนำเครื่องต้นแบบมาพัฒนาให้สามารถเก็บกุญแจได้เพิ่มขึ้นเป็น 48 ช่องเก็บ และสามารถส่งมอบกุญแจได้รวดเร็วกว่าเครื่องต้นแบบที่เป็นสายพานลำเลียง ซึ่งใช้เวลาในการส่งมอบกุญแจที่ช้า โดยเครื่องจัดเก็บกุญแจที่พัฒนาขึ้นใหม่มีจำนวน 8 ชั้น ดังรูปที่ 4(ก) และในแต่ละชั้นมีจำนวนช่องเก็บ

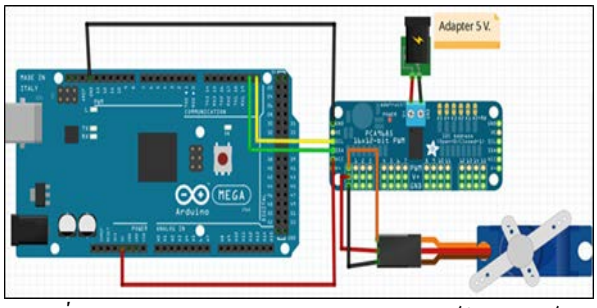
กุญแจได้ 6 ช่อง ดังรูปที่ 4(ข) สามารถเก็บกุญแจได้ทั้งหมด 48 ช่อง



(ก) ตู้ใส่กุญแจ (ข) ช่องเก็บกุญแจในตู้ใส่กุญแจ

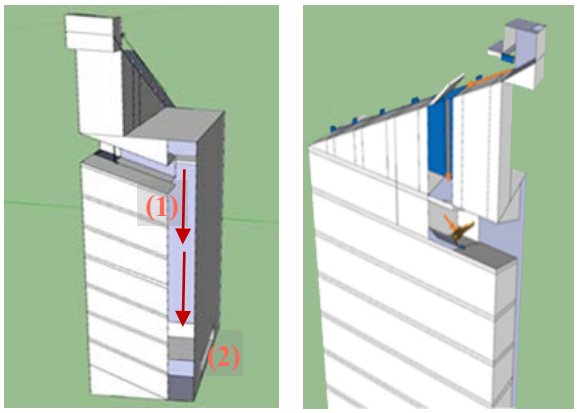
รูปที่ 4. แบบอุปกรณ์เครื่องจัดเก็บกุญแจ

จากรูปที่ 4(ข) ใช้เซอร์โวมอเตอร์จำนวน 2 ตัว ต่อ 1 ช่อง ดังนั้นเครื่องจัดเก็บกุญแจใช้จำนวนเซอร์โวมอเตอร์ทั้งหมด 96 ตัว ในการควบคุมการเปิดปิดทางเข้าและทางออกของกุญแจทั้งหมด โดยมีวงจรควบคุมการทำงานในการเปิดปิดทางเข้าและทางออกในแต่ละช่องที่เก็บกุญแจ ซึ่งควบคุมการทำงานผ่านไมโครคอลโทรลเลอร์ ที่ต่อร่วมกับโมดูล PCA 9685 ในการควบคุมเซอร์โวมอเตอร์ ดังรูปที่ 5.



รูปที่ 5. การต่อวงจรควบคุมการทำงานของเซอร์โวมอเตอร์.

โดยการทำงานของมอเตอร์แต่ละตัวถูกควบคุมจากการอ่านป้ายระบุตัวตน RFID ซึ่งในตำแหน่งของมอเตอร์ด้านหน้าแสดงได้ดังรูปที่ 6 (ก) คือมอเตอร์ตัวที่ 1 ใช้สำหรับการสั่งปล่อยกุญแจ และสำหรับมอเตอร์ตัวที่ 2 ใช้สำหรับเปิดรับกุญแจสู่ช่องเก็บกุญแจแสดงได้ดังรูปที่ 6 (ข) เพื่อเก็บกุญแจไว้ในช่องเก็บกุญแจที่ต้องการโดยการควบคุมมอเตอร์ได้พัฒนาโปรแกรมในการเปิดการทำงานและการปิดการทำงานผ่านการอ่านป้ายระบุตัวตน



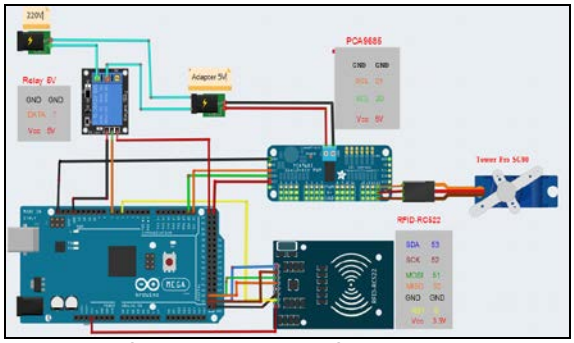
(ก) ตำแหน่งการปล่อยกุญแจ (ข) การไหลของกุญแจ

รูปที่ 6. ตำแหน่งในการติดตั้งมอเตอร์

3. วิธีดำเนินงานวิจัย

3.1 ภาพรวมของระบบ

ระบบที่นำเสนอเป็นระบบที่ทำงานร่วมกันหลายส่วน ซึ่งประกอบไปด้วย ส่วนของการควบคุมอุปกรณ์เครื่องจัดเก็บกุญแจ ส่วนของพัฒนาระบบด้วยโปรแกรมภาษาซี เพื่อควบคุมบอร์ดไมโครคอลโทรลเลอร์อาดุยโน่ในการอ่านค่าเซนเซอร์ และส่วนควบคุมการทำงานของเซอร์โวมอเตอร์ร่วมกับอุปกรณ์ RFID ในการอ่านและเขียนป้ายระบุตัวตน แสดงได้ดังรูปที่ 7.



รูปที่ 7. วงจรภายในของเครื่องจัดเก็บกุญแจ

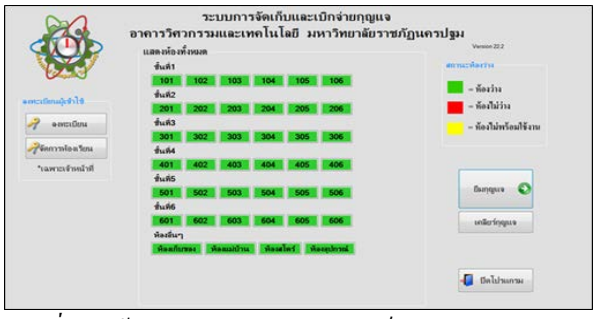
ซึ่งในขั้นตอนการทำงานเมื่อมีการคืนกุญแจในช่องคืนกุญแจระบบทำการอ่านค่าจากป้ายระบุตัวตน ด้วยการอ่านค่าสัญญาณจาก RFID เพื่อตรวจสอบหมายเลขห้อง จากนั้นนำกุญแจมาจัดเก็บตามหมายเลขห้องที่อ่านจากป้ายระบุตัวตนได้ไปยังช่องเก็บกุญแจ แสดงได้ดังรูปที่ 8.

รูปที่ 8. อุปกรณ์เครื่องจัดเก็บกุญแจ.

สำหรับส่วนของการบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ ถูกพัฒนาส่วนติดต่อผู้ใช้และการเชื่อมต่อกับอุปกรณ์ต่าง ๆ ด้วย โปรแกรม Microsoft Visual Basic 2013 ร่วมกับเครื่องตรวจสอบลายนิ้วมือยี่ห้อ HIP รุ่น URU 4000 และระบบฐานข้อมูล SQL Server 2008 โดยมีขั้นตอนในการทำงาน แสดงได้ดังรูปที่ 9.

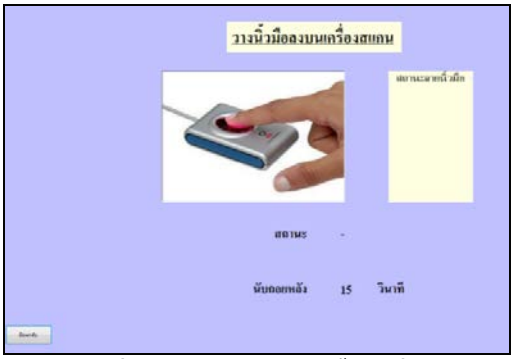
รูปที่ 9. ขั้นตอนการเบิกกุญแจไปใช้งาน

โดยระบบการบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ แบ่งออกเป็น 2 ส่วนการทำงานได้แก่ ส่วนผู้ดูแลระบบ (Admin) และส่วนผู้ใช้งาน (User) โดยในส่วนผู้ดูแลระบบทำหน้าสร้างบัญชีผู้ใช้งานและดูแลจัดการห้องสำหรับบัญชีผู้ใช้งานเริ่มต้นครั้งแรกในการขอเบิกจ่ายกุญแจต้องผ่านการสร้างผู้ใช้งานขึ้นมาก่อน ซึ่งบัญชีผู้ใช้งานมีการจัดเก็บลายนิ้วและประวัติข้อมูลส่วนตัวเบื้องต้นลงฐานข้อมูล โดยการเก็บข้อมูลเพื่อใช้ในการติดตามผู้ใช้งานในกรณีไม่นำส่งกุญแจคืนตู้เก็บกุญแจ สำหรับส่วนการดูแลจัดการห้องนั้นผู้ดูแลระบบสามารถเพิ่มหรือลดจำนวนห้องสำหรับใช้งานได้ รวมถึงสามารถกำหนดรหัสหรือแก้ไขรหัสของป้ายระบุตัวตนที่ใช้งานคู่กับกุญแจห้องได้ และผู้ดูแลระบบสามารถเข้าถึงหน้าต่างของโปรแกรมในส่วนผู้ดูแลระบบได้ โดยการป้อนรหัสก่อนการเข้าใช้งานเพื่อป้องกันบุคคลอื่นเข้ามาแก้ไขระบบการทำงาน โดยหน้าต่างหลักของโปรแกรมการทำงาน แสดงได้ดังรูปที่ 10.

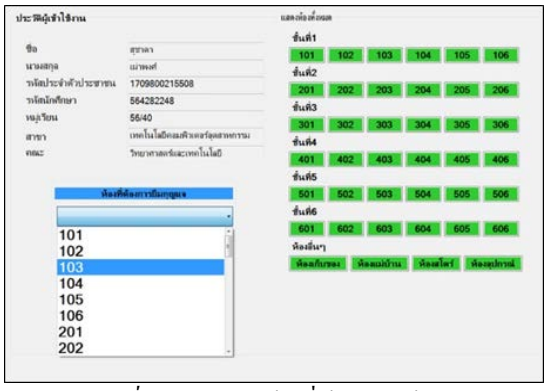


รูปที่10. หน้าต่างหลักของระบบการจัดเก็บและเบิกจ่ายกุญแจ

สำหรับในส่วนผู้ใช้งาน ต้องผ่านการสร้างบัญชีผู้ใช้งานก่อน โดยในขั้นตอนการสร้างบัญชีผู้ใช้งานนั้น ระบบต้องทำการเก็บข้อมูลพื้นฐานของผู้ใช้งาน เช่น ชื่อ นามสกุล เลขบัตรประชาชน รหัสนักศึกษา และข้อมูลลายนิ้วมือ เป็นต้น เพื่อใช้เป็นฐานข้อมูลสำหรับการเบิกจ่ายกุญแจ ในครั้งต่อไป สำหรับผู้ใช้งานที่ผ่านการสร้างบัญชีผู้ใช้งานเรียบร้อยแล้ว ผู้ใช้งานสามารถสแกนลายนิ้วมือเพื่อทำการเบิกจ่ายกุญแจได้ทันที และผู้ใช้งานสามารถคืนกุญแจห้องเมื่อมีการใช้งานเสร็จแล้วที่ช่องคืนกุญแจได้ โดยในการเบิกจ่ายกุญแจใช้วิธีการตรวจสอบลายนิ้วมือ [9] ของผู้ใช้งานดังรูปที่ 11(ก) จากนั้นระบบแสดงหน้าต่างของผู้ใช้งานและทำการเลือกห้องที่ต้องการ เพื่อเบิกกุญแจ โดยข้อกำหนดให้ผู้ใช้งาน 1 คน สามารถเบิกกุญแจได้เพียงห้องเดียวเท่านั้น ไม่สามารถเบิกกุญแจห้องพร้อมกันได้มากกว่า 1 ห้อง หน้าต่างสำหรับใช้งานแสดงได้ดังรูปที่ 11(ข)



รูปที่ 11. (ก) ตรวจสอบลายนิ้วมือเพื่อใช้งาน



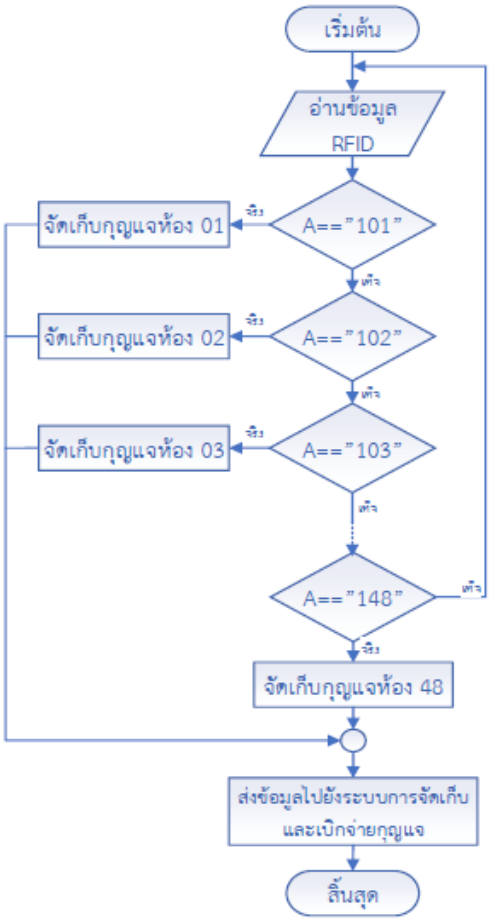
รูปที่ 11. (ข) เลือกห้องที่ต้องการใช้งาน

รูปที่ 11. หน้าต่างผู้ใช้งาน.

สำหรับวิธีในการคืนกุญแจนั้น เริ่มต้นจากผู้ใช้งานใส่กุญแจในกล่องรับคืนกุญแจ จากนั้นระบบทำการอ่านข้อมูลในป้ายระบุตัวตนที่ติดไว้กับกุญแจ เมื่อระบบอ่านหมายเลขห้องจากป้ายระบุตัวตนด้วยอุปกรณ์อ่านสัญญาณ RFID ได้ ขั้นตอนต่อมาไมโครคอลโทรลเลอร์อาดุยโน่ จะทำการประมวลผลตามคำสั่งของโปรแกรมที่ได้ตั้งไว้ โดยสั่งควบคุมการหมุนของเซอร์โวมอเตอร์ให้หมุนสายพานทำให้กล่องรับคืนกุญแจเปิดออก จากนั้นลูกกุญแจจะเลื่อนผ่านช่องต่าง ๆ ไปยังตำแหน่งของช่องการเก็บกุญแจที่อ่านได้ จากนั้นสั่งควบคุมการหมุนของเซอร์โวมอเตอร์ให้หมุนเปิดกล่องเก็บกุญแจทำให้ลูกกุญแจไหลตามแรงโน้มถ่วงเข้าไปยังกล่องเก็บกุญแจ โดยเมื่อกุญแจลงสู่กล่องเก็บกุญแจเรียบร้อย ไมโครคอลโทรลเลอร์ส่งข้อมูลไปยังระบบการบริหารการจัดเก็บกุญแจอัจฉริยะ ผ่านการสื่อสารแบบอนุกรม เพื่อปรับปรุงข้อมูลการเบิกจ่ายกุญแจของระบบตามขั้นตอนวิธี แสดงได้ดังรูปที่ 12.

4. การทดลองและผลการทดลอง

สำหรับระบบการบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัตินั้น ใช้สถานที่ภายในอาคารวิศวกรรมและเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม เพื่อเก็บผลการทดลองการใช้งานระบบการจัดเก็บและเบิกจ่ายกุญแจแบบอัตโนมัติ โดยติดตั้งไว้บนชั้น 5 หน้าห้องสโตร์ใกล้กับตำแหน่งของการเบิกจ่ายกุญแจเดิม โดยมีเจ้าหน้าที่ผู้ดูแลห้องสโตร์จำนวน 1 คน และเปรียบเทียบผู้ใช้งานอยู่ในระบบจริงทั้งหมด 50 คน โดยแบ่งเป็นกลุ่มที่ 1 นักศึกษาที่มีการเรียนอยู่ภายในตึกวิศวกรรมและเทคโนโลยีเป็นประจำทั้งหมด 28 คน, กลุ่มที่ 2 นักศึกษา



รูปที่ 12. ผังการทำงานส่วนการจัดเก็บกุญแจ

นอกสาขาที่มีเรียนภายในตึกเป็นครั้งคราวทั้งหมด 16 คน และกลุ่มที่ 3 เจ้าหน้าที่ที่เกี่ยวข้องกับการดูแลรักษาห้องทั้งหมด 6 คน จากการวัดผลด้วยวิธีการประเมินค่าความพึงพอใจจากผู้ใช้งานแบ่งออกเป็น 3 ด้าน โดยแบ่งค่าน้ำหนักของแต่ละด้านดังนี้ ด้านรูปแบบการใช้งานระบบมีค่าน้ำหนักอยู่ที่ร้อยละ 50, ด้านความปลอดภัยของระบบอยู่ร้อยละ 30 และด้านการออกแบบความสวยงามของระบบอยู่ที่ร้อยละ 20 โดยหลักจากเก็บผลการประเมินของผู้ใช้งานระบบการบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัตินั้น นำคะแนนของแต่ละกลุ่มมาหาค่าเฉลี่ย โดยพิจารณาจากสมการ (1)

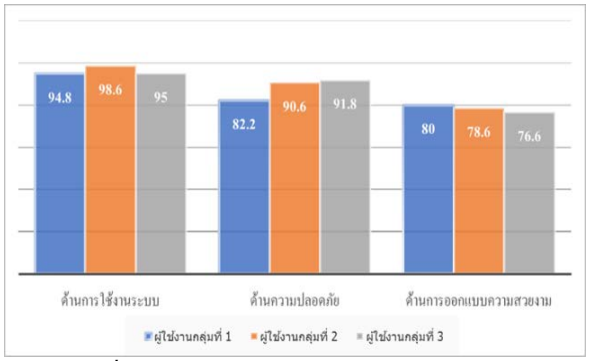
$$mean\ x = (w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n)/n \quad (1)$$

เมื่อ w หมายถึง ค่าน้ำหนักคะแนนที่ให้

x หมายถึง ค่าระดับคะแนนความคิดเห็น

n หมายถึง จำนวนผู้ร่วมประเมิน

ผลที่ได้จากการนำค่าน้ำหนักในแต่ละข้อของแต่ละด้านและนำผลที่ได้นำมาหาค่าเฉลี่ยความพึงพอใจของแต่ละกลุ่มผู้ใช้งาน ผลคะแนนค่าเฉลี่ยที่ได้จากการประเมินของ กลุ่มที่ 1 ให้ค่าความพึงพอใจในด้านการใช้งานระบบอยู่ที่ร้อยละ 94.8 ซึ่งอยู่ในเกณฑ์ ดีมาก สำหรับด้านความปลอดภัยของระบบอยู่ที่ร้อยละ 82.2 ซึ่งอยู่ในเกณฑ์ ดี เนื่องจากผู้ใช้งานระบบบางคนไม่มั่นใจในระบบให้ระบุเลขที่ของบัตรประชาชนลงไปในการสมัครสมาชิก และด้านการออกแบบความสวยงามของระบบอยู่ที่ร้อยละ 80 ซึ่งอยู่ในเกณฑ์ ดี ซึ่งอธิบายได้ว่าสามารถทำงานแต่หน้าต่างการทำงานควรมีการปรับปรุงให้สวยงามกว่าเดิม สำหรับผลคะแนนค่าเฉลี่ยที่ได้จากการประเมินของ กลุ่มที่ 2 ให้ค่าความพึงพอใจในด้านการใช้งานระบบอยู่ที่ร้อยละ 98.6 อยู่ในเกณฑ์ ดีมาก สำหรับด้านความปลอดภัยของระบบอยู่ที่ร้อยละ 90.6 ซึ่งอยู่ในเกณฑ์ ดีมาก และด้านการออกแบบความสวยงามของระบบอยู่ที่ร้อยละ 78.6 ซึ่งอยู่ในเกณฑ์ พอใช้ ซึ่งผลคะแนนที่ได้เนื่องจาก ขอให้มีการปรับปรุงหน้าต่างโปรแกรมเพิ่มเติม และผลคะแนนค่าเฉลี่ยที่ได้จากการประเมินของ กลุ่มที่ 3ให้ค่าความพึงพอใจในด้านการใช้งานระบบอยู่ที่ร้อยละ 95 ซึ่งอยู่ในเกณฑ์ ดีมาก ซึ่งสำหรับด้านความปลอดภัยของระบบอยู่ที่ร้อยละ 91.8 อยู่ในเกณฑ์ ดี ส่วนและด้านการออกแบบความสวยงามของระบบอยู่ที่ร้อยละ76.6 ซึ่งอยู่ในเกณฑ์ พอใช้ ซึ่งผลคะแนนที่ได้เนื่องจาก หน้าต่างการใช้งานดูเรียบ ๆ ขาดความสวยทำให้ขาดความน่าสนใจในส่วนของผู้ใช้งาน โดยสรุปคะแนนเฉลี่ยได้ดังรูปที่ 13



รูปที่ 13. เปรียบเทียบความพึงพอใจของผู้ใช้งาน.

จากรูปที่ 13 อธิบายได้ถึงความพึงพอใจของระบบที่พัฒนาขึ้นนั้นสามารถใช้งานได้จริงตรงตามวัตถุประสงค์ของการพัฒนาระบบ แต่ขาดความสวยงามซึ่งผู้ใช้งานได้อธิบายถึงการออกแบบพัฒนาโปรแกรมที่ขาดความเป็นมิตรกับผู้ใช้งาน และยังขาดการพัฒนาในด้านความปลอดภัยของข้อมูลโดยแนะนำการเข้ารหัสข้อมูลส่วนสำคัญสำหรับการใช้งานก่อนการจัดเก็บข้อมูลจึงส่งผมให้ค่าการประเมินในด้านดังกล่าวน้อยกว่าด้านอื่น

สำหรับการเปรียบเทียบเวลาที่ใช้สำหรับส่วนการจัดเก็บและส่วนการเบิกจ่ายกุญแจ จำนวน 48 ดอก กับเจ้าหน้าที่ดูแลห้องสโตร์ โดยแบ่งการเปรียบเทียบเวลาออกเป็นเวลาที่ใช้ในการเบิกจ่ายกุญแจ และเวลาที่ใช้ในการจัดเก็บกุญแจ โดยทำการสุ่มเรียกกุญแจจำนวน 10 ครั้ง ในแต่ละวันโดยทำการทดสอบเป็นเวลา 10 วัน จากนั้นนำมาหาค่าเฉลี่ยด้านเวลามาเปรียบเทียบเพื่อหาค่าประสิทธิภาพในการทำงานดังตารางที่ 1.

ตารางที่ 1. การเปรียบเทียบเวลาในการทำงานในส่วนการเบิกจ่ายกุญแจ

ครั้งที่ / วิธีการเบิกกุญแจ	ผลการทดลอง (ค่าเฉลี่ยของเวลา) (s)	
	เจ้าหน้าที่	เครื่องเบิกจ่ายกุญแจ
1	16.40	12.27
2	16.81	13.12
3	15.70	12.14
4	15.92	12.77
5	14.80	13.54
6	16.22	12.14
7	16.80	12.12
8	15.88	12.25
9	16.67	12.57
10	16.67	12.14

หลังจากการทดสอบการทำงานในส่วนการเบิกจ่ายกุญแจแสดงให้เห็นถึงเวลาในการทำงานของระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ ในส่วนการเบิกจ่ายกุญแจสามารถทำเวลาโดยเฉลี่ยดีกว่าเจ้าหน้าที่อยู่ที่ 3.681 วินาที จากนั้นทำการทดลองในส่วนการคืนกุญแจโดยนำกุญแจที่เบิกไปกลับมาคืนในตำแหน่งเดิมที่เก็บกุญแจ สำหรับการทำงานของระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ เป็นการรับกุญแจที่ต้องการคืนจากช่องคืนกุญแจจากนั้นประมวลผลและแยกการจัดเก็บตามเลขของกุญแจที่อ่านได้เก็บลงกล่องเก็บกุญแจที่เตรียมไว้ โดยทำการทดลอง

จัดเก็บกุญแจจำนวน 10 ครั้ง เป็นเวลา 10 วัน ค่าที่ได้นำมาหาค่าเฉลี่ยด้านเวลาและนำมาเปรียบเทียบเพื่อหาค่าประสิทธิภาพในการทำงานดังตารางที่ 2

ตารางที่ 2. การเปรียบเทียบเวลาในการทำงานในส่วนการคืนกุญแจ

ครั้งที่ / วิธีการเบิกกุญแจ	ผลการทดลอง (ค่าเฉลี่ยของเวลา) (s)	
	เจ้าหน้าที่	เครื่องเบิกจ่ายกุญแจ
1	10.7	9.44
2	11.32	9.52
3	11.74	9.60
4	11.90	9.43
5	10.97	9.55
6	11.95	9.62
7	11.24	9.84
8	12.60	9.24
9	11.87	9.25
10	11.77	9.68

ผลการทดลองที่ได้แสดงให้เห็นว่าระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ ในส่วนของการคืนกุญแจนั้น สามารถทำเวลาได้ดีกว่าเจ้าหน้าที่ ซึ่งเวลาเฉลี่ยของการจัดเก็บกุญแจโดยเจ้าหน้าที่ 11.606 วินาที และเวลาเฉลี่ยในการทำงานของอุปกรณ์ในการจัดเก็บกุญแจใช้เวลาเฉลี่ย 9.517 วินาที ซึ่งระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัตินั้น ให้ผลการทดลองด้านเวลาที่ดีกว่าเจ้าหน้าที่อยู่ 2.089 วินาที

5. บทสรุป การอภิปราย และข้อเสนอแนะ

งานวิจัยฉบับนี้ได้นำเสนอระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ โดยทำงานร่วมกับอุปกรณ์ฮาร์ดแวร์ที่ได้ประดิษฐ์คิดค้นขึ้น โดยสามารถเก็บกุญแจและเบิกจ่ายกุญแจได้อย่างมีระบบ โดยนำเทคโนโลยี RFID และป้ายระบุตัวตนมาช่วยในการระบุกุญแจ และมีโปรแกรมระบบบริหารการจัดเก็บกุญแจอัจฉริยะ โดยมีระบบสมาชิกของผู้ใช้งานที่ช่วยให้ทราบถึงบุคคลที่มาขอเบิกกุญแจไปใช้งาน โดยจากการทดลองระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัติ แสดงให้เห็นว่าใช้เวลาในการเบิกจ่ายกุญแจและจัดเก็บกุญแจน้อยกว่าการใช้งานเจ้าหน้าที่ดูแลห้องสโตร์โดยทำเวลาได้ดีกว่า 3.681 วินาที ในการเบิกจ่ายกุญแจและ 2.089 วินาที สำหรับการจัดเก็บกุญแจตามลำดับ โดยมีค่าความพึงพอใจจากผู้ใช้งานทั้งหมด โดยให้ในส่วนการใช้งานระบบดีที่สุด โดยคิดเป็นค่าเฉลี่ยรวมอยู่ที่ร้อยละ 96.133 ซึ่งแสดงให้เห็นว่าระบบบริหารการจัดเก็บกุญแจอัจฉริยะด้วยวิธีการบ่งชี้อัตโนมัตินั้น สามารถนำไปใช้งานได้จริง

เอกสารอ้างอิง

[1] A.Pavithra, M.Kalavathi, S.Keerthi, and SK.Sabirunnaisa, “Access Control Using RFID and Arduino,” B.Tech. Project Report (Electronics and communication engineering), Jawaharlal Nehru Technological University, Hyderabad, 2013.

[2] Praharshin M. Senadeera, and Numan S. Dogan, “Emerging Applications in RFID Technology,” Inter. J. of Computer Science and Electronics Engineering (IJCSEE), vol. 4, no. 2, pp. 75-79, 2016.

[3] H. F. Abdulsada, “Design and Implementation of Smart Attendance System Based On Raspberry pi,” Journal of Babylon University, Engineering Sciences, vol. 25, no. 5, pp. 1610-1618, 2017.

[4] สุนีรัตน์ อินตะมะ และวันดี มงคงกาย, “ไขความลับ...กับระบบจัดการกุญแจ,” May, 2019. [Online]. Available: https://www.excise.go.th/cs/groups/public/documents/document/mjaw/mdy2/~edisp/webportal1 6 2 0 0 0 6 6 8 0 5 .pdf. [Accessed: May 20, 2019].

[5] E. Michael, Y. Isah, Bako Hussaini, O.F. Hassan, and V.C. Ezika, “Arduino Based RFID Line Switching Using SSR,” Int. J. of Scientific and Technology Research, vol. 6, no. 10, pp.98-100, 2017.

[6] EZ. Orji, CV. Oleka, UI. Nduanya, “Automatic Access Control System using Arduino and RFID,” Journal of Scientific and Engineering Research, vol. 5, no. 4, pp. 333-340, 2018.

[7] พิทย์พิมล ชูรอด, เนาวลักษณ์ แสงสนิท, และสุพิริยา ผลนาค , “การพัฒนาระบบยืมหนังสือด้วยเครื่องสแกนลายนิ้วมือของสำนักหอสมุด มหาวิทยาลัยทักษิณ วิทยาเขตพัทลุง,” วารสารมหาวิทยาลัยทักษิณ, ปีที่ 16, ฉบับที่ 2, กรกฎาคม-ธันวาคม, หน้า 1-8, 2556.

[8] อาจารี นาโค, “การประยุกต์เครื่องอ่านลายนิ้วมือเพื่อตรวจสอบการเข้าชั้นเรียน,” วารสารมหาวิทยาลัยทักษิณ, ปีที่ 16 ฉบับที่ 3 ฉบับพิเศษ, หน้า 11-20, 2556.

[9] K. S. Myint, C.M.M. Nyein, “*Fingerprint Based Attendance System Using Arduino,”* Int. J. of Scientific and Research Publications, vol. 8, no. 8, July, pp. 422-426, 2018.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 48-58

การประยุกต์และวิเคราะห์ข้อมูลสีด้วยการประมวลผลภาพถ่ายสถานที่ท่องเที่ยวเพื่อการออกแบบบรรจุภัณฑ์ของที่ระลึก

Application and Analysis of Color Information using Tourist Attraction Image Processing for Souvenir Packaging Design

*คณากาญจน์ รักไพฑูรย์ เกษม ทิพย์ธาราจันทร์ และ ฐิติพร เลิศรัตน์เดชากุล**

*Kanakarn Ruxpaitoon, Kasem Thiptarajan, and Thitiporn Lertrusdachakul**

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีไทย-ญี่ปุ่น

Faculty of Information Technology, Thai-Nichi Institute of Technology

Received: February 26, 2020; **Revised:** May 28, 2020; **Accepted:** May 28, 2020; **Published:** June 23, 2020

ABSTRACT – Packaging is the external appearance that creates the first impression and stimulates the needs of customers. The beautiful packaging with unique and consistent design particularly for souvenir of tourism will assist in product recognition and attract customers' attention to create the story and increase the product value. As the color has greatly affected the psychological perception, this research therefore proposes the color analysis by image processing and applies to the design of souvenir packaging to communicate the meaning and convey the emotions of embedded image of the tourist attraction. The design concept and recommended colors are obtained from the analysis of representative colors, color image scale and place information. The representative colors are based on k-means clustering and the investigation from an art expert. The packaging designs of plai balm, plai oil and aroma oil; souvenirs of famous tourist attraction in Thailand, are used as the case studies in this research. The proposed design integrates the concept extracted from the image color analysis of overall exterior architecture and color tone of harmonious meaning. The evaluation results show that most respondents are more interested in buying the proposed packaging designs than the original ones. The results of design evaluation in perception of embedded feeling, desire of purchase and the suitability of color usage are all in high levels. The comments and suggestions from respondents are introduced to improve for more perfect packaging design.

KEYWORDS: Color Analysis, Color Image Scale, Packaging Design, Souvenir

บทคัดย่อ -- บรรจุภัณฑ์นับเป็นรูปลักษณ์ภายนอกที่สร้างความประทับใจแรกและกระตุ้นความต้องการของผู้ซื้อ โดยเฉพาะอย่างยิ่งผลิตภัณฑ์ของฝาก ของที่ระลึกในการท่องเที่ยว บรรจุภัณฑ์ที่สวยงาม มีความเป็นเอกลักษณ์ สอดคล้องกับตัวสินค้า จะช่วยสร้างการจดจำและดึงดูดความสนใจของนักท่องเที่ยว สร้างเรื่องราวและมูลค่าเพิ่มให้แก่ผลิตภัณฑ์ และเนื่องด้วยสีมีผลต่อการรับรู้ทางจิตวิทยาเป็นอย่างมาก งานวิจัยนี้จึงได้นำเสนอการวิเคราะห์ข้อมูลสีด้วยการประมวลผลภาพถ่ายสถานที่ท่องเที่ยวและประยุกต์สู่การออกแบบบรรจุภัณฑ์ของฝาก ของที่ระลึก เพื่อสื่อสารความหมาย ถ่ายทอดอารมณ์ ความรู้สึก และเข้าถึงภาพลักษณ์ที่แฝงอยู่ของสถานที่ท่องเที่ยว โดยแนวคิดของการออกแบบและกลุ่มสีที่แนะนำเกิดจากการวิเคราะห์ข้อมูลตัวแทนสีของภาพที่ได้จากการแบ่งกลุ่มแบบเคมีนร่วมกับการกรองข้อมูลสีจากผู้เชี่ยวชาญทางศิลปะ เพื่อนำมาวิเคราะห์ตามหลัก Color Image Scale เชื่อมโยงกับข้อมูลสถานที่ ซึ่งงานวิจัยนี้ได้ศึกษากรณีตัวอย่างการประยุกต์และวิเคราะห์ข้อมูลสีสำหรับการออกแบบบรรจุภัณฑ์ยาหม่องไพล น้ำมันไพล และน้ำมันอโรมา ของที่ระลึกจากสถานที่ท่องเที่ยวที่มีชื่อเสียงของประเทศไทย โดยการออกแบบได้ผสมผสานแนวคิดจากการวิเคราะห์ข้อมูลสีของภาพถ่าย

**Corresponding Author: thitiporn@tni.ac.th*

โดยรวมของสถาปัตยกรรมภายนอก และเลือกใช้โทนสีที่มีความกลมกลืนสื่อความหมาย ซึ่งจากผลประเมินการออกแบบพบว่า ผู้ตอบแบบสอบถามส่วนใหญ่มีความสนใจเลือกซื้อบรรจุภัณฑ์รูปแบบใหม่มากกว่ารูปแบบเดิม และมีผลประเมินในด้านการให้อารมณ์ ความรู้สึกที่แฝงอยู่ ความรู้สึกอยากซื้อผลิตภัณฑ์ และความเหมาะสมในการใช้สีของบรรจุภัณฑ์อยู่ในระดับมาก โดยข้อเสนอแนะต่าง ๆ ได้ถูกนำมาพัฒนาการออกแบบบรรจุภัณฑ์ให้มีความสมบูรณ์มากยิ่งขึ้น

คำสำคัญ: การวิเคราะห์ข้อมูลสี, Color Image Scale, การออกแบบบรรจุภัณฑ์, ของที่ระลึก

1. บทนำ

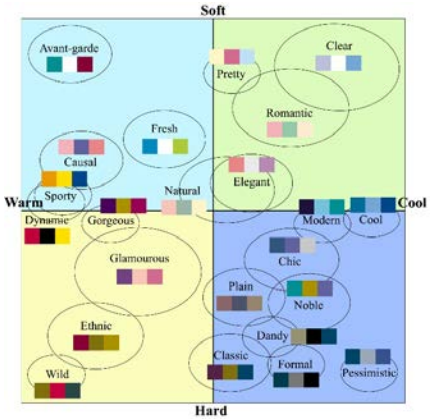
การพัฒนาประเทศไทยภายใต้โมเดล “Thailand 4.0” เพื่อขับเคลื่อนให้ประเทศมีความมั่นคง มั่งคั่ง และยั่งยืนอย่างเป็นรูปธรรม ได้เล็งเห็นถึงการวิจัยและพัฒนาในกลุ่มอุตสาหกรรมสร้างสรรค์ วัฒนธรรม และบริการที่มีมูลค่าสูง (Creative, Culture & High Value Services) โดยจากสถิติของการท่องเที่ยวแห่งประเทศไทย (ททท.) พบว่าในปี 2561 มีจำนวนนักท่องเที่ยวต่างชาติถึง 38 ล้านคน และมีแนวโน้มเติบโตขึ้น [1] ดังนั้นการเพิ่มมูลค่าเพิ่มของของฝาก ของที่ระลึก ซึ่งเป็นส่วนหนึ่งในห่วงโซ่การกระจายรายได้ในกระบวนการท่องเที่ยว ซึ่งนอกจากจะถูกซื้อหาเป็นของฝากแทนไมตรีจิตแล้ว ยังช่วยในการเล่าเรื่องราว ภูมิวัฒนธรรม วิถีแห่งสถานที่นั้น ๆ สร้างการสื่อสาร เผยแพร่จุดเด่น ความน่าสนใจ นำไปสู่การเชิญชวนให้มาเยี่ยมเยียน ช่วยส่งเสริมการท่องเที่ยวและเพิ่มมูลค่าทางเศรษฐกิจให้กับประเทศ การเล่าเรื่องราวผ่านผลิตภัณฑ์ดังกล่าวจัดเป็นส่วนหนึ่งของ Kotozukuri ของประเทศญี่ปุ่น [2] เพื่อใช้ในการสร้างคุณค่าทางจิตใจและความประทับใจให้กับลูกค้า ในการออกแบบผลิตภัณฑ์ของฝาก ของที่ระลึกจำเป็นต้องมีการบูรณาการศาสตร์ต่าง ๆ (Interdisciplinary) เช่น Kotozukuri วัฒนธรรม ภูมิปัญญาท้องถิ่น ศิลปศาสตร์และเทคโนโลยี โดยการออกแบบผลิตภัณฑ์ บรรจุภัณฑ์และสื่อโฆษณาประชาสัมพันธ์นั้น ถือเป็นเครื่องมืออันทรงพลังและสำคัญในการสร้างมูลค่าและการรับรู้ของสินค้าแก่ผู้พบเห็น โดยเฉพาะอย่างยิ่งในยุคแห่งการแข่งขันสูงทางเทคโนโลยี การออกแบบที่สามารถสื่อสารแนวคิดของผลิตภัณฑ์ให้เข้าถึงและสร้างความประทับใจให้กับลูกค้า นอกจากจะช่วยส่งเสริมการท่องเที่ยวแล้ว ยังช่วยส่งเสริมการตลาดดิจิทัลและการตลาดออนไลน์ได้อย่างมีประสิทธิภาพ การออกแบบที่ประสบความสำเร็จนอกจากจะคำนึงถึงประสบการณ์ในการใช้งานผลิตภัณฑ์ของผู้บริโภคแล้ว องค์ประกอบทางศิลป์ซึ่งมีผลทางจิตวิทยาต่อผู้ใช้ก็เป็นส่วนสำคัญในการสร้างความรู้สึกเชิงบวกให้กับผลิตภัณฑ์ องค์ประกอบทางศิลปะได้แก่ เส้น รูปร่าง รูปทรง พื้นผิว ลวดลาย แสงเงา สี ความสมดุล การเคลื่อนไหว โดยสีถือเป็นองค์ประกอบสำคัญที่ส่งผลกระทบต่อจิตใจและการรับรู้ทางอารมณ์ ทำให้งานออกแบบสามารถสื่อสารแนวคิด ความรู้สึกไปยังผู้พบเห็นได้อย่างเข้าถึงทางจิตใจมากยิ่งขึ้น

ในงานออกแบบทั่วไปนั้น มักมีจุดกำเนิดจากแนวคิด (Concept) ของผลิตภัณฑ์หรือสิ่งที่ต้องการจะออกแบบ ซึ่งมักแสดงอยู่ในรูปของข้อความ หรือคำสำคัญต่าง ๆ ที่สื่อถึงรายละเอียดของแนวคิด โดยงานวิจัยนี้มีวัตถุประสงค์เพื่อช่วยนักออกแบบค้นหาแนวคิดที่นอกเหนือจากคำจำกัดความโดยตรงที่ได้รับมอบหมาย แต่เป็นแนวคิดที่เกิดจากภาพถ่ายเพื่อช่วยในการพัฒนาการออกแบบให้มีความสมบูรณ์ยิ่งขึ้น โดยใช้หลักการของการวิเคราะห์ข้อมูลสีและตัวแทนสี (Representative Color) จากการประมวลผลภาพถ่าย ร่วมกับ Color Image Scale เพื่อกลั่นกรองเรื่องราว อารมณ์ ความรู้สึกที่แฝงอยู่ ซึ่งได้นำเสนอวิธีการวิเคราะห์ข้อมูลสีของภาพสู่แนวคิดและประยุกต์ใช้แนวคิดเพื่อการออกแบบ โดยได้เลือกการออกแบบบรรจุภัณฑ์ของฝาก ของที่ระลึกจากวัดพระเชตุพนวิมลมังคลารามราชวรมหาวิหารหรือวัดโพธิ์เป็นตัวอย่างกรณีศึกษา เนื่องด้วยวัดโพธิ์จัดได้ว่าเป็นสถานที่ท่องเที่ยวของไทยที่เป็นที่รู้จักกันดีของนักท่องเที่ยวชาวต่างประเทศ และมีการถ่ายทอดศิลปวัฒนธรรมไทยให้เห็นถึงการใช้สีของวัดไทย มีเรื่องราวทางประวัติศาสตร์และเปรียบเสมือนมหาวิทยาลัยแห่งแรกของประเทศไทย โดยแนวคิดที่ได้จากภาพถ่ายสามารถนำไปประยุกต์ ผสมผสานกับแนวคิดหลักของการออกแบบ หรือสามารถใช้เป็นจุดเริ่มต้นของการออกแบบสำหรับผู้ที่ยังไม่มีแนวคิดหรือยังคิดแนวทางการออกแบบไม่ออกได้

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

Image Scale เป็นเครื่องมือการทำแผนที่ความรู้สึกของมนุษย์ที่วิจัยและพัฒนาโดยสถาบันวิจัยและการออกแบบสีของญี่ปุ่น

โดย Color Image Scale คือแนวความคิดที่นำเสนอชุดสีหรือ Color Combination ซึ่งเป็นแนวคิดเพื่อใช้เทียบชุดสีกับค่าสีในตารางเพื่อทราบความหมายของสีในเชิงจิตวิทยาสีกับความรู้สึก และเป็นแนวความคิดที่นำเสนอชุดสี ซึ่งแต่ละชุดสีนั้นสามารถตีความเป็นคำสื่ออารมณ์ที่ใช้กับงานออกแบบสื่อดิจิทัล โดยโครงสร้างของเครื่องมือนี้ถูกพัฒนามาจาก Munsell Color System โดยสถาบันวิจัยสีและการออกแบบนิปปอน (Nippon Color & Design Research Institute) ซึ่ง Kobayashi [3] - [4] ผู้พัฒนาแนวความคิด Color Image Scale ได้ทำการจัด Color Combination แบ่งออกเป็นกลุ่มตามโทนสี Soft – Hard หรือโทนสีอ่อน – เข้ม และโทนสี Warm – Cool หรือโทนสีร้อน – เย็น ดังแสดงในรูปที่ 1 ซึ่งการรวมตัวของสีต่าง ๆ เป็นกลุ่มสีเพื่อเป็นตัวแทนสีของภาพนั้น สามารถแสดงอารมณ์ ความรู้สึกของภาพที่ซับซ้อนและละเอียดอ่อนได้ดีกว่าการใช้ตัวแทนสีเดียว โดย Color Image Chart ในรูปที่ 1 ประกอบด้วย 23 กลุ่มชุดอารมณ์ความรู้สึก เช่น น่ารัก (Pretty), โรแมนติก (Romantic), ทันสมัย (Modern), เป็นทางการ (Formal) ซึ่งในแต่ละกลุ่มชุดอารมณ์ความรู้สึกจะแบ่งย่อยออกเป็น คำศัพท์สำคัญแสดงภาพลักษณ์ต่าง ๆ หรือที่เรียกว่า Key Image Word และในแต่ละ Key Image Word จะมีชุดสีประมาณ 16 สีเป็นองค์ประกอบ ตัวอย่างเช่น กลุ่มชุดอารมณ์ความรู้สึก น่ารัก (Pretty) ประกอบด้วย 6 Key Image Words คือ น่ารัก (Pretty), ที่เป็นที่รัก (Endearing), น่าเอ็นดูเหมือนเด็ก (Childish), สวยงาม (Lovely), หวาน (Sweet), น่ารักสดใส (Cute) โดยในแต่ละ Key Image Word เหล่านี้ จะมีชุดสีผสมผสานเป็นองค์ประกอบเฉพาะตัว แต่อาจมีการใช้สีบางสีร่วมกันได้ ซึ่ง 23 กลุ่มชุดอารมณ์ความรู้สึก ประกอบด้วย Key Image Words ทั้งสิ้นจำนวน 160 คำศัพท์ [5]



รูปที่ 1. Color Image Chart

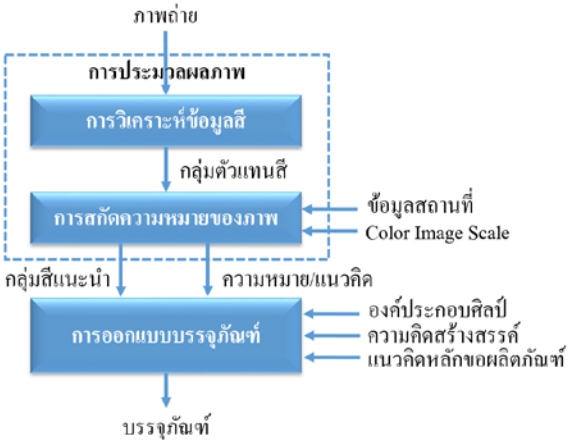
แนวคิด Color Image Scale พบมากในประเทศญี่ปุ่น ซึ่งถือได้ว่าเป็นประเทศที่ใส่ใจในทุก ๆ แขนงของงานออกแบบ โดยมีการศึกษาที่เกี่ยวข้องกับ Color Image Scale และการวิเคราะห์สีจากภาพ รวมถึงการประยุกต์ใช้งานในด้านต่าง ๆ เช่น การศึกษาพื้นฐานเกี่ยวกับ Color Image Scale ของทิวทัศน์สะพาน [6] โดยในงานวิจัยดังกล่าวได้ศึกษาวิเคราะห์ความสัมพันธ์ระหว่างสีตามระบบ H&T (Hue and Tone) ของทัศนียภาพสะพานตามประเภทโครงสร้าง สถานที่ก่อสร้างและความยาว กับ Color Image Scale เพื่อค้นหาภาพลักษณ์ของสะพาน สร้างเป็น Color Image Scale สำหรับทัศนียภาพสะพาน เพื่อประยุกต์ใช้ในการออกแบบสีสะพานที่เหมาะสม นอกจากนั้นที่จังหวัด Aomori ได้มีการจัดทำคู่มือการแนะนำสีภูมิทัศน์จังหวัด Aomori ขึ้น ชื่อว่า Aomori Prefecture Landscape Color Guide Plan [7] ซึ่งได้ศึกษาวิเคราะห์เกี่ยวกับสีภูมิทัศน์ธรรมชาติของพื้นที่ต่าง ๆ ในจังหวัด สีของเมือง อาคาร สิ่งก่อสร้าง ผสมผสานกับความคิดเห็นของคนในชุมชนเกี่ยวกับภาพลักษณ์ที่เหมาะสมของพื้นที่ที่สำคัญในจังหวัด และลักษณะของแหล่งท่องเที่ยวและภูมิอากาศ เพื่อแนะนำแนวทางการใช้สีที่เหมาะสมและแสดงภาพลักษณ์ของท้องถิ่น

ในเชิงธุรกิจได้มีการนำ Color Image Scale มาช่วยในการออกแบบสี การผสมผสานและจัดวางรูปแบบของสีให้กับที่อยู่อาศัย เช่น รูปแบบการออกแบบที่อยู่อาศัยในเมืองที่เรียบง่ายจะแฝงไปด้วยสีที่ให้อารมณ์ความรู้สึกของความเรียบง่ายทันสมัย มีความเป็นธรรมชาติและสวยงามแบบทันสมัย [8] ในด้านของความรู้สึกกับการประยุกต์ใช้ในอุตสาหกรรมในเชิงของ Human Media Engineering ได้มีการนำเสนอแนวทางการวางแผนการออกแบบผลิตภัณฑ์จากการวิเคราะห์สถานการณ์ผลิตภัณฑ์ของบริษัท ทำให้ได้ทราบถึงการเปรียบเทียบกับผลิตภัณฑ์ของคู่แข่ง กลุ่มลูกค้าเป้าหมาย เพื่อกำหนดจุดยืนของผลิตภัณฑ์ วิเคราะห์แนวโน้มของยุคสมัยและผลิตภัณฑ์อื่น ๆ ที่เกี่ยวข้อง สร้างเป็น Color Image Scale เพื่อใช้ในการวางแผนการตลาดและการผลิต [9] ในส่วนของงานวิจัยที่มีการประยุกต์ใช้ Color Image Scale กับเสียง ได้แก่ งานวิจัยของ Yamawaki และ Shiizuka [10] ที่ได้ศึกษาเกี่ยวกับการนำเสียงเพลงเทียบกับ Color Image Scale โดยการนำเสียงกับสีมาประสานความสัมพันธ์โดยใช้เสียงมาทดสอบเทียบเป็นสีที่มีการแบ่งความรู้สึกผ่านตาราง Color Image Scale ซึ่งมีคำศัพท์ของตารางเป็นตัวช่วยในการเลือกอารมณ์เพลงที่ได้ยิน Namba [11]

ได้นำเสนองานวิจัยเกี่ยวกับการปรับปรุงเสียงฟังก์ชันของเครื่องใช้ในครัวเรือน เช่น เสียงแสดงการทำงาน เสียงแจ้งข้อมูล เสียงแจ้งเตือนเหตุผิดปกติ โดยนำลักษณะการทำงานของ Color Image Scale มาประยุกต์ใช้ เนื่องจากอุปกรณ์เครื่องใช้ในครัวเรือนและสารที่เสียงฟังก์ชันต้องการจะสื่อ ต่างก็ให้อารมณ์ความรู้สึกแก่ผู้รับฟัง ซึ่งจะเป็นตัวเชื่อมไปยังลักษณะของเสียงเพื่อสื่อสารข้อมูลโดยที่ไม่เห็นอุปกรณ์เครื่องใช้และสถานการณ์ทำงาน คณะผู้วิจัยได้นำเสนอการต่อยอดการประยุกต์ใช้งาน Color Image Scale สำหรับการออกแบบบรรจุภัณฑ์ของฝากของที่ระลึก ด้วยการประมวลผลภาพถ่ายของสถานที่ท่องเที่ยวและวิเคราะห์ข้อมูลสีร่วมกับ Color Image Scale เพื่อสกัดความหมายที่แฝงอยู่ ออกมาเป็นแนวคิดในการผสมผสานกับแนวคิดของผลิตภัณฑ์ เพื่อเลือกองค์ประกอบศิลป์ สี แนวทางความคิดสร้างสรรค์ในการออกแบบที่นอกจากจะสื่อแนวคิดหลักของผลิตภัณฑ์แล้ว ยังแฝงไว้ซึ่งอารมณ์ ความรู้สึกและความเป็นสถานที่ท่องเที่ยวแห่งนั้น ๆ

3. วิธีการดำเนินงานวิจัย

ในการประยุกต์และวิเคราะห์ข้อมูลสีเพื่อสกัดความหมายจากภาพถ่ายสำหรับการออกแบบบรรจุภัณฑ์ของฝาก ของที่ระลึกนั้น มีขั้นตอนดังแสดงในรูปที่ 2 ซึ่งประกอบด้วย การนำภาพถ่ายมาประมวลผลภาพเพื่อเป็นข้อมูลสำหรับการวิเคราะห์สีและค้นหาตัวแทนสี ซึ่งจะถูกนำมาประมวลผลร่วมกับข้อมูลของสถานที่ท่องเที่ยวและ Color Image Scale เพื่อค้นหาความหมายที่แฝงอยู่ และกลุ่มสีที่แนะนำ หลังจากนั้นจึงนำแนวคิดที่ได้ ประยุกต์ร่วมกับกลุ่มสีที่แนะนำ หลักการทางศิลปะ ความคิดสร้างสรรค์ และแนวคิดหลักของผลิตภัณฑ์เพื่อออกแบบบรรจุภัณฑ์ที่สื่อเรื่องราว อารมณ์ ความรู้สึกแก่ผู้พบเห็น ซึ่งในงานวิจัยนี้ได้เลือกผลิตภัณฑ์ยาหม่องไพล น้ำมันไพล และน้ำมันอโรมา เป็นกรณีศึกษาการออกแบบบรรจุภัณฑ์ของฝาก ของที่ระลึกจากวัดโพธิ์ เนื่องจากเป็นผลิตภัณฑ์ที่ได้รับความนิยมของโรงเรียนแพทย์แผนโบราณวัดพระเชตุพน (วัดโพธิ์) ซึ่งเปิดให้บริการการนวดไทยตำรับวัดโพธิ์ที่มีชื่อเสียงและได้รับการยอมรับทั้งจากชาวไทยและชาวต่างประเทศ โดยมีสาขาเปิดให้บริการแก่นักท่องเที่ยวบริเวณศาลาในวัดโพธิ์ด้วย วิธีการดำเนินงานวิจัยในแต่ละขั้นตอนมีดังนี้



รูปที่ 2. ขั้นตอนการออกแบบบรรจุภัณฑ์จากการประมวลผลภาพถ่าย

3.1 การเตรียมภาพถ่าย

เลือกบริเวณเป้าหมายของสถานที่ท่องเที่ยว เพื่อกำหนดขอบเขตที่ชัดเจนและครอบคลุมตามวัตถุประสงค์ โดยในกรณีศึกษานี้ คณะผู้วิจัยเลือกสถาปัตยกรรมภายนอกของวัดโพธิ์เป็นบริเวณที่สนใจเนื่องจากเป็นจุดที่สร้างการจดจำและความประทับใจแรกให้แก่ผู้เยี่ยมชม หลังจากนั้นดำเนินการถ่ายภาพในจุดสำคัญ จุดสนใจต่าง ๆ ของนักท่องเที่ยว ด้วยวิธีการถ่ายภาพพร้อมกับใช้ Color Checker ในการวัดค่าสีเพื่อแก้ไขสภาวะแสงและสีที่คลาดเคลื่อนจากความเป็นจริง ด้วยเทคนิคการปรับแต่งสภาวะแสงและสีจาก Color Checker [12] แล้วทำการเลือกภาพที่จะนำมาใช้โดยให้ครอบคลุมสามารถแสดงภาพรวมของสถานที่ได้ เมื่อได้ภาพครบถ้วนแล้วจึงนำมาไดคัท ตัดส่วนที่ไม่เกี่ยวข้องในภาพออก เช่น ต้นไม้ ท้องฟ้า

3.2 การวิเคราะห์ข้อมูลสี

ในขั้นตอนนี้ภาพสี RGB ที่ได้จากการเตรียมภาพถ่ายในข้อ 3.1 จะถูกนำมาประมวลผล โดยปรับขนาดภาพให้ด้านที่ยาวที่สุดมีขนาดเท่ากับ 500 พิกเซล และรักษาอัตราส่วนเดิมของภาพไว้ จากนั้นแปลงค่าสีเป็นค่ามาตรฐาน CIE $L^*a^*b^*$ ซึ่งเป็นระบบสีที่สัมพันธ์กับการมองเห็นของมนุษย์ โดย L^* หมายถึง ความสว่าง ส่วน a^* และ b^* ใช้กำหนดค่าสี กล่าวคือ $-a^*$ หมายถึง อยู่ในทิศของสีเขียว $+a^*$ หมายถึง อยู่ในทิศของสีแดง $-b^*$ หมายถึง อยู่ในทิศของสีน้ำเงิน และ $+b^*$ หมายถึง อยู่ในทิศของสีเหลือง ซึ่งค่า CIE $L^*a^*b^*$ ดังกล่าวจะถูกนำมาใช้ในการค้นหาตัวแทนสี (Representative Color) หรือ RC โดยใช้หลักการแบ่งกลุ่มข้อมูล

สีแบบเคมีน (*k*-Means Clustering) [13] - [14] และใช้การคำนวณระยะห่างระหว่างข้อมูลสีกับจุดศูนย์กลางของกลุ่ม แบบ Squared Euclidean Distance ซึ่งจะทำการแบ่งข้อมูลสีทุกพิกเซลออกเป็น 128 กลุ่ม และดึงสีที่มีระยะห่างใกล้กับจุดศูนย์กลางของกลุ่มมากที่สุด เป็นตัวแทนสีของภาพ หลังจากนั้นจึงแปลงค่าสีที่ได้ทั้ง 128 สี เป็นระบบสี HSV หรือ Hue (ค่าสีของสีหลัก) Saturation (ความบริสุทธิ์ของสี) และ Value (ความสว่างของสี) เพื่อทำการวิเคราะห์ข้อมูลความแตกต่างของค่าสี (Difference in Hue Values) ความบริสุทธิ์ของสี โดยสีที่มีความแตกต่างของค่าสีสูง หรือมีความบริสุทธิ์ของสีสูงในแต่ละ Hue ของ Munsell Color System จะถูกดึงออกมาเป็นตัวแทนสี ในขณะที่สีส่วนที่เหลือจะถูกแบ่งและจัดกลุ่มใหม่ด้วยเคมีน เพื่อลดจำนวนตัวแทนสีทั้งหมดลงเหลือ 32 สี สำหรับใช้ในการกรองข้อมูลจากการเลือกสีโดยผู้เชี่ยวชาญทางศิลปะ ซึ่งรายละเอียดกระบวนการหาตัวแทนสี (RC) อ้างอิงจากผลงานวิจัยของ Ruxpaitoon และ Lertrusdachakul [15] ที่ได้ทำการศึกษาวิเคราะห์สีของสถาปัตยกรรมภายนอกของวัดประจำรัชกาลแห่งราชวงศ์จักรีเพื่อหาความหมาย อารมณ์ ความรู้สึกที่แฝงอยู่ โดยผู้เชี่ยวชาญทางศิลปะจะเลือกสีที่เป็นตัวแทนสีที่ดีที่สุดของภาพพิจารณาจากความโดดเด่นต่อการรับรู้ อยู่ในบริเวณที่เป็นจุดสนใจของภาพ ปริมาณและการกระจายตัวของสีในภาพ โดยอาศัยโปรแกรมแสดงผลช่วยในการคัดเลือก ซึ่งจะแสดงค่าสีทั้งแบบระบบสี RGB และระบบสี CIE L*a*b* ของพิกเซลที่สนใจในภาพและของตัวแทนสีที่เลือกจาก 32 สี รวมถึงแสดงภาพสีประกอบเฉพาะพิกเซลในกลุ่มสีของตัวแทนสีที่เลือกดังกล่าวเพื่อใช้ในการพิจารณาเปรียบเทียบและตัดสินใจเลือกตัวแทนสีที่เหมาะสม เมื่อได้ตัวแทนสีของแต่ละภาพของภาพทั้งหมดแล้วจึงนำมาประมวลผลอีกครั้งเพื่อเลือกตัวแทนสีแสดงภาพรวมของสถาปัตยกรรมภายนอกของวัดโพธิ์ ด้วยการแปลงค่าสีเป็นระบบสี Munsell และแบ่งกลุ่มสีย่อยในแต่ละ Hue ตามการกระจายตัวของระดับความสว่างของสีด้วยค่า Threshold จากวิธีการของ Otsu [16] โดยสีที่มีความบริสุทธิ์ของสีสูงที่สุดและมีความถี่ของค่าความสว่างของสีมากที่สุดในแต่ละกลุ่มสีย่อยนี้ จะถูกเลือกให้เป็นตัวแทนสีของภาพทั้งหมด [15]

3.3 การสกัดความหมายของภาพ

เมื่อได้ตัวแทนสี (RC) ที่ใช้แสดงภาพรวมของสถานที่ท่องเที่ยวแล้ว จึงนำข้อมูลการกระจายตัวของตัวแทนสีและสีทั้งหมด ค่า CIE L*a*b* และอัตราส่วนของตัวแทนสี การกระจายตัวของค่าความเข้มสี (C^*_{ab}) และค่าสี (h_{ab}) ของตัวแทนสี ซึ่งคำนวณได้จากสมการที่ 1 และ 2 มาเทียบกับ Color Image Scale และ Color Image Chart [3] - [5] ซึ่งได้แบ่งสีออกเป็น 160 ชุดสี พร้อมคำศัพท์เพื่อช่วยในการจำกัดคำนิยาม โดยคัดเลือกชุดสีที่มีคุณลักษณะใกล้เคียงกับตัวแทนสี และสกัดความหมายจากคำศัพท์ของชุดสีและสีที่เป็นองค์ประกอบ โดยอาศัยข้อมูลของสถานที่ เช่น ข้อมูลทางประวัติศาสตร์ ศิลปวัฒนธรรม ลักษณะและรูปแบบการก่อสร้าง มาประกอบในการสกัดความหมายของภาพ โดยกลุ่มสีที่แนะนำสำหรับใช้ในการออกแบบคือ ตัวแทนสีของภาพ (RC) และ ชุดสีที่ถูกคัดเลือกจาก Color Image Scale และ Color Image Chart

$$C^*_{ab} = \sqrt{a^{*2} + b^{*2}} \quad (1)$$

$$h_{ab} = tan^{-1}\frac{b^*}{a^*} \quad (2)$$

โดย a^* และ b^* คือ ค่าสี a* และ b* ของระบบสี CIE L*a*b*

3.4 การออกแบบบรรจุภัณฑ์และการประเมินผล

งานวิจัยนี้ได้นำเสนอการประยุกต์แนวคิดที่ได้จากการสกัดความหมายด้วยการประมวลผลและการวิเคราะห์ข้อมูลสีของภาพถ่ายสถานที่ท่องเที่ยว เพื่อการออกแบบบรรจุภัณฑ์ของฝากของที่ระลึก ร่วมกับแนวคิดหลักของผลิตภัณฑ์ โดยใช้องค์ประกอบศิลป์ที่สอดคล้อง เหมาะสม และความคิดสร้างสรรค์เพื่อสร้างความประทับใจให้แก่ผู้พบเห็น ด้วยกลุ่มสีที่แนะนำเพื่อสื่อสารเรื่องราว อารมณ์ ความรู้สึกของสถานที่ท่องเที่ยวนั้น ๆ โดยมีการดำเนินการออกแบบ และจัดทำแบบสอบถามออนไลน์ไปยังคนไทยและคนญี่ปุ่น (ตัวแทนชาวต่างชาติที่มีความละเอียดอ่อนเกี่ยวกับเรื่องบรรจุภัณฑ์) เพื่อนำผลประเมินและข้อเสนอแนะมาปรับปรุงการออกแบบ ซึ่งได้ทำการสอบถามเกี่ยวกับข้อมูลทั่วไป ได้แก่ เพศ อายุ สัญชาติ และสอบถามในหัวข้อต่อไปนี้ของแต่ละบรรจุภัณฑ์ รวม 3 ผลิตภัณฑ์

- เปรียบเทียบความสนใจเลือกซื้อเป็นของฝาก ของที่ระลึก ระหว่างบรรจุภัณฑ์รูปแบบเดิมกับรูปแบบใหม่
- ระดับอารมณ์ ความรู้สึกของบรรจุภัณฑ์รูปแบบใหม่

- ความรู้สึกอยากซื้อผลิตภัณฑ์ของบรรจุภัณฑ์รูปแบบใหม่
- ความเหมาะสมในการใช้สีของบรรจุภัณฑ์รูปแบบใหม่
- ข้อเสนอแนะเกี่ยวกับการออกแบบบรรจุภัณฑ์รูปแบบใหม่

โดยในแต่ละหัวข้อคำถามจะมีภาพบรรจุภัณฑ์ประกอบการพิจารณา และทำการพิจารณาค่าเฉลี่ยของผลประเมินแบ่งเป็น 5 ระดับดังนี้

4.51-5.00 มากที่สุด

3.51-4.50 มาก

2.51-3.50 ปานกลาง

1.51-2.50 น้อย

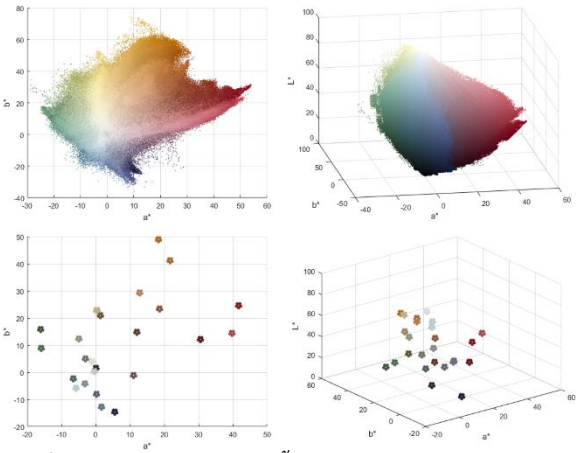
1.00-1.50 น้อยที่สุด

4. ผลการวิจัยและการอภิปรายผล

จากการเตรียมภาพถ่ายสถาปัตยกรรมภายนอกของวัดโพธิ์เพื่อแสดงจุดสนใจและภาพรวมของสถานที่ ภาพที่ถูกคัดเลือกสำหรับใช้ในการประมวลผลจำนวน 24 ภาพ ถูกไดคัทให้เหลือเฉพาะบริเวณที่สนใจ และปรับสีให้เป็นสีตามที่ตามนุษย์มองเห็นจริงดังแสดงในรูปที่ 3 ซึ่งผลจากการวิเคราะห์ข้อมูลสีของภาพ [15] ทำให้ได้ตัวแทนสีจำนวน 25 สีตามในรูปที่ 4 โดยรูปที่ 5 แสดงการกระจายตัวของตัวแทนสีและสีทั้งหมดใน CIE $L^*a^*b^*$ Color Space ซึ่งค่า CIE $L^*a^*b^*$ และอัตราส่วนของตัวแทนสีเรียงจากมากไปน้อย สรุปได้ดังตารางที่ 1

รูปที่ 3. ภาพถ่ายสถาปัตยกรรมภายนอกของวัดโพธิ์ที่ผ่านการไดคัทและปรับแก้ไขสีสำหรับการประมวลผล

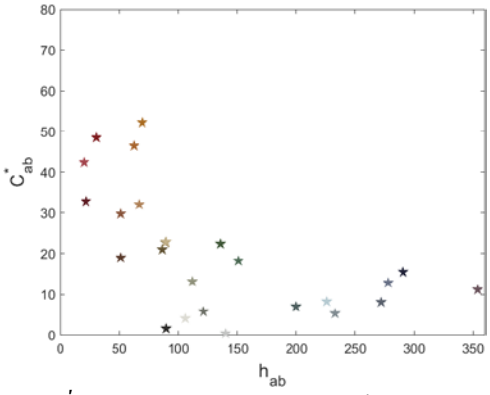
รูปที่ 4. ตัวแทนสี (RC) ของภาพถ่ายในรูปที่ 3



รูปที่ 5. การกระจายตัวของสีทั้งหมด (บน) และตัวแทนสี (ล่าง) ใน CIE $L^*a^*b^*$ Color Space

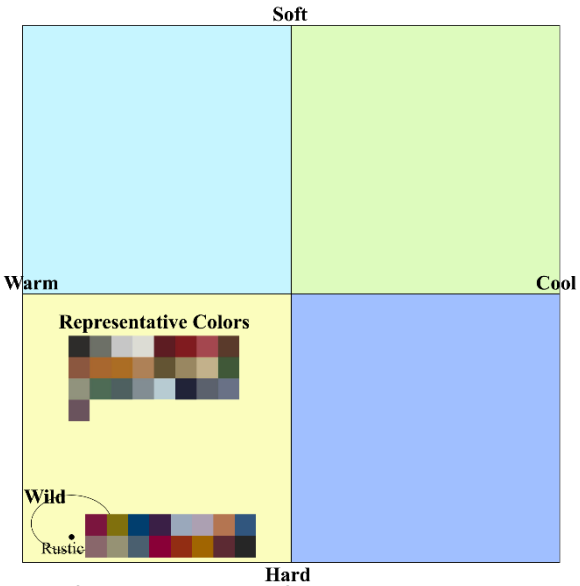
ตารางที่ 1 ค่า CIE $L^*a^*b^*$ และอัตราส่วนของตัวแทนสี

No.	L^*	a^*	b^*	RC	Ratio
1	34.71	-16.04	15.72		7.07
2	79.65	-0.35	0.29		6.88
3	37.95	11.04	-1.20		6.41
4	80.35	-5.65	-5.88		6.36
5	46.83	-2.97	4.93		6.10
6	87.65	-1.11	3.91		5.40
7	27.75	11.94	14.74		5.19
8	18.18	0.01	1.58		5.19
9	36.45	1.30	20.96		4.73
10	20.97	30.49	12.17		4.15
11	60.26	-4.85	12.25		4.10
12	39.38	-6.57	-2.38		4.07
13	57.99	-3.17	-4.21		4.01
14	47.76	1.75	-12.78		3.99
15	57.02	0.43	22.74		3.50
16	49.92	21.68	41.13		3.19
17	28.01	41.74	24.63		3.12
18	42.38	-15.94	8.86		2.96
19	42.30	18.63	23.30		2.95
20	57.44	12.76	29.35		2.84
21	72.88	0.25	22.79		2.05
22	51.53	18.37	48.93		1.73
23	41.21	0.26	-8.10		1.50
24	14.55	5.48	-14.51		1.50
25	42.84	39.76	14.45		1.00



รูปที่ 6. การกระจายตัวของค่าความเข้มสี (C^*_{ab}) และค่าสี (h_{ab}) ของตัวแทนสี

ผลการเปรียบเทียบข้อมูลสีและตัวแทนสีกับ Color Image Scale และ Color Image Chart ร่วมกับการพิจารณาข้อมูลทางประวัติศาสตร์ของสถานที่ ได้ชุดสีที่มีความใกล้เคียงกับตัวแทนสี (RC) อยู่ในช่วงโทนสีร้อน – เข้มของ Color Image Chart และคำศัพท์ที่สื่อความหมายของภาพคือ ความมีพลัง ความกล้าหาญ มีสุนทรียภาพทางความงาม โดยรูปที่ 7 แสดงตัวแทนสีทั้ง 25 สี (เรียงลำดับตาม Munsell Color System) และชุดสีที่ใกล้เคียงกับตัวแทนสี (จำนวน 16 สี เรียงตามความถี่ของการใช้งาน) ซึ่งทั้งหมดเป็นกลุ่มสีที่แนะนำเพื่อประกอบการพิจารณาในการออกแบบ [15]



รูปที่ 7. คำศัพท์และกลุ่มสีที่แนะนำสำหรับการออกแบบ

สำหรับการออกแบบบรรจุภัณฑ์ของผลิตภัณฑ์ยาหม่องไพล น้ำมันไพล และน้ำมันอโรมานั้น (รูปที่ 8 แสดงผลิตภัณฑ์ที่วางขายอยู่ในปัจจุบัน) จากข้อมูลทางประวัติศาสตร์และสิ่งที่มีชื่อเสียง เป็นที่รู้จักหรือเป็นจุดสนใจของนักท่องเที่ยว บวกกับอารมณ์ ความรู้สึกที่สกัดจากสีของภาพรวมสถาปัตยกรรมภายนอกของวัด ทำให้ยักษ์และพระเจดีย์ถูกเลือกให้เป็นตัวแทนสื่อความเป็นไทยและความเป็นวัดโพธิ์ โดยยักษ์แสดงถึงความมีพลัง ซึ่งปัจจุบันยักษ์วัดโพธิ์มีปรากฏอยู่ 2 คู่ คือ พญาไมยราพณ์กับพญาแสงอาทิตย์ และพญาขรกับพญาสัทธาสูร ตั้งอยู่ที่ซุ้มประตูทางเข้าพระมณฑป มีลักษณะคล้ายยักษ์ในวรรณคดีเรื่องรามเกียรติ์ สำหรับองค์พระเจดีย์นั้น วัดโพธิ์ถือว่าเป็นวัดที่มีพระเจดีย์มากที่สุดในประเทศไทย โดยมีพระเจดีย์ที่สำคัญ คือ พระมหาเจดีย์สี่รัชกาล ซึ่งเป็นพระมหาเจดีย์ประจำรัชกาลที่ 1-4 ได้แก่ พระมหาเจดีย์ศรีสรรเพชดาญาณ พระมหาเจดีย์ดิลกธรรมกรกนิทาน พระมหาเจดีย์มุนีบัตบริขาร และพระมหาเจดีย์ทรงพระศรีสุริโยทัย ซึ่งพระมหาเจดีย์แต่ละองค์นั้นเป็นเจดีย์ย่อไม้สิบสอง ประดับด้วยกระเบื้องเคลือบและมีความงดงามทางศิลปะ โดยได้นำหน้ายักษ์ทั้ง 4 ตน มาออกแบบเป็นลวดลายสำหรับบรรจุภัณฑ์ยาหม่องไพลและน้ำมันไพล และนำพระเจดีย์มาออกแบบเป็นลวดลายสำหรับบรรจุภัณฑ์น้ำมันอโรมา และเลือกใช้สีประกอบ ตามกลุ่มสีที่แนะนำจากการสกัดความหมายของภาพสถาปัตยกรรมภายนอกโดยรวมของวัด สำหรับหลักการทางศิลปะที่นำมาประยุกต์ใช้ นอกจากลวดลายของยักษ์และพระเจดีย์ที่ออกแบบให้มีรายละเอียดเรียบง่าย แต่แฝงไว้ซึ่งความมีเอกลักษณ์เฉพาะตัวแล้วนั้น ได้มีการเลือกใช้โทนสี ความสว่างเพื่อสร้างมิติให้แก่ลวดลาย และผสมผสานการใช้งานศิลปะแนวคอลลาจ (Collage) ด้วยการนำหน้ายักษ์และพระเจดีย์มาต่อกันเป็นลวดลาย สร้างความน่าสนใจให้กับกล่องของผลิตภัณฑ์ นอกจากนั้นการออกแบบยังได้คำนึงถึงพฤติกรรมการใช้สอย เนื่องจากยาหม่องไพลมีสรรพคุณทาเพื่อบรรเทาอาการเคล็ด ขัดยอก ปวดเมื่อยตามร่างกาย แมลงกัดต่อย จึงออกแบบให้มีลักษณะเป็นหลอดบีบ เพื่อให้สามารถพกพาและใช้งานง่าย โดยทุกบรรจุภัณฑ์ยังคงรักษาแนวคิดหลักของความเป็นผลิตภัณฑ์ธรรมชาติเพื่อสุขภาพ ซึ่งได้มีการปรับโทนสีโดยรวมให้มีความกลมกลืนเข้ากัน เรียบอย่างมีเอกลักษณ์ และสามารถจัดทำเป็นชุด (Collection) สำหรับเทศกาลหรือโอกาสพิเศษได้

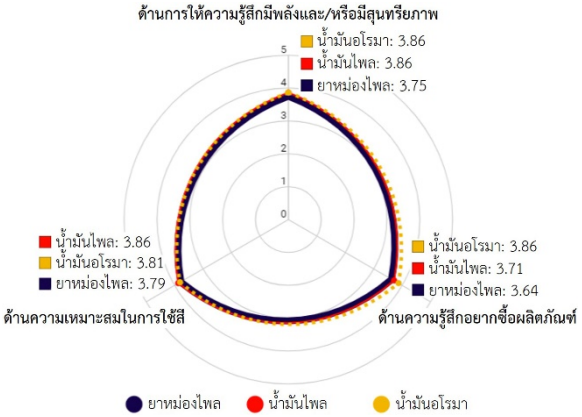
รูปที่ 8. ผลิตภัณฑ์ต้นแบบของกรณีศึกษา

ผลประเมินจากแบบสอบถามออนไลน์โดยผู้ตอบแบบสอบถามทั้งสิ้น 80 คน ประกอบด้วยชาย 21 คน และหญิง 59 คน โดยเป็นชาวไทย 49 คน และชาวญี่ปุ่น 31 คน พบว่า ส่วนใหญ่มีความสนใจเลือกซื้อบรรจุภัณฑ์ที่ออกแบบใหม่เพื่อเป็นของฝาก ของที่ระลึกมากกว่ารูปแบบเดิม โดยสนใจเลือกซื้อคิดเป็นร้อยละ 68.75, 76.25 และ 73.75 สำหรับบรรจุภัณฑ์ยาหม่องไพล น้ำมันไพล และน้ำมันอโรมา ตามลำดับ ดังรายละเอียดแสดงในตารางที่ 2 โดยผลประเมินการออกแบบบรรจุภัณฑ์รูปแบบใหม่ในด้านการให้ความรู้สึกมีพลังและ/หรือมีสุนทรียภาพ ความรู้สึกอยากซื้อผลิตภัณฑ์ และความเหมาะสมในการใช้สีของบรรจุภัณฑ์ มีค่าเฉลี่ยของแต่ละบรรจุภัณฑ์อยู่ในเกณฑ์มากทุกด้าน โดยมีค่าเฉลี่ยอยู่ในช่วงตั้งแต่ 3.64 ถึง 3.86 ดังรายละเอียดแสดงในรูปที่ 9

ตารางที่ 2 ผลประเมินการเปรียบเทียบความสนใจเลือกซื้อระหว่างบรรจุภัณฑ์รูปแบบเดิมกับรูปแบบใหม่เพื่อเป็นของฝากของที่ระลึก

ผลิตภัณฑ์	รูปแบบบรรจุภัณฑ์	ชาย	หญิง	ไทย	ญี่ปุ่น	รวม
ยาหม่องไพล	เดิม	19.05%	35.59%	20.41%	48.39%	31.25%
	ใหม่	80.95%	64.41%	79.59%	51.61%	68.75%
น้ำมันไพล	เดิม	23.81%	23.73%	16.33%	35.48%	23.75%
	ใหม่	76.19%	76.27%	83.67%	64.52%	76.25%
น้ำมันอโรมา	เดิม	23.81%	27.12%	18.37%	38.71%	26.25%
	ใหม่	76.19%	72.88%	81.63%	61.29%	73.75%

ซึ่งมีข้อเสนอแนะที่สำคัญ คือ ควรปรับเพิ่มการแสดงรายละเอียดเกี่ยวกับวัดหรือลวดลายเพื่อให้ผู้ที่ไม่รู้จักหรือไม่เคยมาเมืองไทยได้ทราบ ผู้ตอบแบบสอบถามชาวญี่ปุ่นผู้หญิงบางส่วนรู้สึกว่ายักษ์มีความน่ากลัว บรรจุภัณฑ์น้ำมันอโรมาควรปรับให้แสดงถึงกลิ่นและความเป็นผลิตภัณฑ์ธรรมชาติชัดเจนขึ้น นอกจากนั้นทุกบรรจุภัณฑ์ควรระบุส่วนประกอบและสรรพคุณอย่างชัดเจน จากข้อเสนอแนะดังกล่าว คณะผู้วิจัยได้ทำการปรับปรุงการออกแบบ และปรับให้มีข้อมูลและรูปแบบการสื่อสารที่เข้าใจง่ายมากขึ้น โดยผลของการปรับปรุงการออกแบบบรรจุภัณฑ์ได้ถูกรวบรวมไว้ในรูปที่ 10 - 14



รูปที่ 9. ผลประเมินการออกแบบบรรจุภัณฑ์รูปแบบใหม่

รูปที่ 10 แสดงบรรจุภัณฑ์ยาหม่องไพล โดยมีพญาขร (ยักษ์สีเขียว) และพญาสัทธาสูร (ยักษ์สีหงเสนหรือสีอิฐ-ชมพู) เป็นลวดลายให้เลือกสะสม การออกแบบมีลักษณะเป็นหลอดบีบเพื่อสะดวกต่อการใช้งาน โทนสีหลัก ได้แก่ สีเขียว สีหงเสน โทนสีเนื้อและสีเหลืองอ่อน

รูปที่ 10. บรรจุภัณฑ์ยาหม่องไพล

รูปที่ 11 แสดงบรรจุภัณฑ์น้ำมันไพล โดยมีพญาแสงอาทิตย์ (ยักษ์สีแดง) และพญาไมยราพณ์ (ยักษ์สีม่วงอ่อน) แบบครึ่งหน้าเป็นลวดลาย ซึ่งสามารถนำมาต่อเป็นรูปหน้ายักษ์

แบบเต็มหน้า ทั้งแบบประเภทเดียวกันหรือแบบผสมผสานได้ดังแสดงในรูปที่ 12 สามารถเลือกซื้อเป็นคู่เพิ่มความน่าสนใจในการเป็นของฝาก ของที่ระลึกได้ โทนสีหลักที่ใช้ ได้แก่ โทนสีแดง สีม่วงอ่อน และสีเหลืองอ่อน

รูปที่ 11. บรรจุภัณฑ์น้ำมันไพล

รูปที่ 12. บรรจุภัณฑ์แบบคู่สำหรับผลิตภัณฑ์น้ำมันไพล

รูปที่ 13 แสดงบรรจุภัณฑ์น้ำมันอโรมา ซึ่งมีหลายกลิ่น โดยได้เลือกกลิ่นตะไคร้และกลิ่นมะลิมาออกแบบเนื่องจากเป็นที่นิยมและรู้จักกันดีในหมู่ชาวต่างประเทศ ลวดลายประกอบด้วยพระเจดีย์แบบครึ่งภาพและภาพสัญลักษณ์แสดงกลิ่น โดยใช้หลักการของศิลปะแนวคอลลาจเช่นเดียวกับบรรจุภัณฑ์น้ำมันไพล สามารถเลือกซื้อเป็นคู่เพื่อสร้างพระเจดีย์ที่สมบูรณ์ ดังแสดงในภาพด้านล่างของรูปที่ 13 โดยเป็นกลิ่นเดียวกันหรือต่างกันก็ได้ ซึ่งจะมีชื่อกลิ่นแสดงอย่างชัดเจนอยู่ด้านหน้า

รูปที่ 13. บรรจุภัณฑ์น้ำมันอโรมา

รูปที่ 14 แสดงบรรจุภัณฑ์แบบชุดพิเศษ (Special Collection) ประกอบด้วยผลิตภัณฑ์ 3 ประเภท จัดเรียงเป็นคู่ โดยแต่ละคู่ออกแบบให้รองรับการวางสลับตำแหน่งซ้าย-ขวาได้ ตกแต่งกล่องบรรจุภัณฑ์ด้วยภาพร่างที่เกี่ยวข้องกับวัดโพธิ์และความเป็นไทย เช่น ภาพของใบโพธิ์ พระพุทธไสยาส (พระนอนวัดโพธิ์) พระมหาเจดีย์ รถตุ๊ก ๆ เป็นต้น ด้านข้างสันขวาประกอบด้วยข้อมูลส่วนประกอบและสรรพคุณของแต่ละผลิตภัณฑ์ และด้านข้างสันซ้ายแสดงหน้ายักษ์แต่ละตนพร้อมชื่อประกอบ ด้านหลังเป็นข้อมูลเกี่ยวกับวัดโพธิ์ องค์ประกอบโดยรวมเน้นความเป็นไทยแบบร่วมสมัย

รูปที่ 14. บรรจุภัณฑ์ชุดพิเศษ (Special Collection)

โดยผลประเมินจากแบบสอบถามออนไลน์สามารถนำไปประยุกต์ใช้เป็นแนวทางในพัฒนาผลิตภัณฑ์ต่อยอดสู่การจัดจำหน่ายจริง ที่ต้องพิจารณาในเรื่องของต้นทุนค่าใช้จ่ายการตลาด กลุ่มเป้าหมาย การควบคุมปัจจัยด้านราคา และปัจจัยแวดล้อมต่าง ๆ อย่างถี่ถ้วนต่อไป

5. บทสรุปและการอภิปราย

งานวิจัยนี้ได้นำเสนอวิธีการประยุกต์ใช้ภาพถ่ายให้เป็นแนวคิด (Concept) ประกอบการออกแบบบรรจุภัณฑ์ของฝาก ของที่ระลึก โดยเป็นการผสมผสานระหว่างภาพลักษณ์ (Image) ของสถานที่ท่องเที่ยวที่ถ่ายทอดผ่านการใช้สี กับแนวคิดหลักของผลิตภัณฑ์ เพื่อสื่อสารอารมณ์ความรู้สึกที่แฝงอยู่ โดยใช้หลักการวิเคราะห์ข้อมูลสีด้วยการประมวลผลภาพถ่ายสถานที่ท่องเที่ยว ร่วมกับ Color Image Scale และข้อมูลเกี่ยวกับสถานที่เพื่อค้นหาความหมาย แนวคิด และกลุ่มสีที่แนะนำในการช่วยเลือกองค์ประกอบศิลป์ โทนสี แนวความคิดสร้างสรรค์สำหรับการออกแบบ ซึ่งผลจากกรณีศึกษาการออกแบบบรรจุภัณฑ์ยาหม่องไพล น้ำมันไพล และน้ำมันอโรมา เพื่อเป็นของฝาก ของที่ระลึกของวัดโพธิ์นั้น จากแนวคิดและกลุ่มสีแนะนำที่ได้จากการวิเคราะห์ข้อมูลสีของภาพถ่ายโดยรวมของวัดโพธิ์ ทำให้ยักษ์ซึ่งนำการรับรู้ทางความรู้สึกมีพลัง และพระเจดีย์ซึ่งแสดงถึงความงามทางศิลปะถูกเลือกให้เป็นลวดลายบนบรรจุภัณฑ์ โดยต่างก็เป็นที่รู้จักในหมู่ชาวต่างประเทศและมีเรื่องราวเชื่อมโยงกับวัดโพธิ์มาอย่างยาวนาน การออกแบบเน้นความกลมกลืนของสีที่เป็นเอกลักษณ์ของลวดลาย ความเป็นไทยร่วมกับโทนสีที่แนะนำ ปรับระดับความสว่างของสีเพื่อเพิ่มมิติและความสวยงามให้กับลวดลาย มีรูปทรงที่พกพาและใช้งานง่าย และมีประเภทบรรจุภัณฑ์ชุดคู่และชุดพิเศษที่ใช้ศิลปะแนวคอลลาจช่วยเพิ่มความน่าสนใจ น่าสะสมและซื้อหา โดยผลประเมินการออกแบบบรรจุภัณฑ์รูปแบบใหม่จากแบบสอบถามออนไลน์พบว่ามีค่าเฉลี่ยในด้านการให้ความรู้สึกมีพลังและ/หรือมีสุนทรียภาพ ความรู้สึกอยากซื้อผลิตภัณฑ์ และความเหมาะสมในการใช้สีของบรรจุภัณฑ์อยู่ในระดับมาก ผู้ตอบแบบสอบถามส่วนใหญ่มีความสนใจในการเลือกซื้อบรรจุภัณฑ์ที่ออกแบบใหม่เพื่อเป็นของฝาก ของที่ระลึกมากกว่าบรรจุภัณฑ์รูปแบบเดิม ซึ่งนอกจากการประยุกต์ใช้งานของแนวคิดที่ได้จากภาพถ่ายร่วมกับแนวคิดหลักของการออกแบบแล้ว ผู้ที่กำลังหาแนวคิดให้งานออกแบบ สามารถนำเอาภาพที่เกี่ยวข้องกับเรื่องราวที่จะใช้ในงาน มาทำการประมวลผล วิเคราะห์ข้อมูลสีเพื่อดึงค่าตัวแทนสีของภาพ สร้างเป็นกลุ่มโทนสีและกลุ่มคำศัพท์เพื่อการออกแบบ ซึ่งสามารถต่อยอดไปยังการประยุกต์ใช้งานร่วมกับการออกแบบเทคโนโลยีอัจฉริยะต่าง ๆ เพื่อการเข้าถึงข้อมูลเชิงลึกหรือข้อมูลที่มีความซับซ้อนของผลิตภัณฑ์ต่อไปได้

กิตติกรรมประกาศ

ขอขอบคุณทุนสนับสนุนการวิจัยจาก Takahashi Industrial and Economic Research Foundation และ นางสาวพาขวัญ ยุพิน ผู้ช่วยงานด้านกราฟิก

เอกสารอ้างอิง

[1] Ministry of Tourism and Sports, “Tourism Economic Review,” mots.go.th, 2019. [Online]. Available: https://www.mots.go.th/download/article/article_20191025094442.pdf. [Accessed: Feb. 23, 2020].

[2] M. Ballé, D. Powell, and K, Yokozawa, “Monozukuri, Hitozukuri, Kotozukuri,” Planet Lean, 2019. [Online]. Available: https://planet-lean.com/monozukuri-hitozukuri-kotozukuri/. [Accessed: Feb. 23, 2020].

[3] S. Kobayashi, “Color Image Scale,”. Kodansha, 1990.

[4] S. Kobayashi, “Colorist. Kodansha,” 1997.

[5] H. Nagumo, “New Color Image Chart,” Graphic-sha Publishing, 2016.

[6] H. Isami and A. Yasuoka, “A fundamental study on color image scale of bridge landscapes,” Journal of Applied Computing in Civil Engineering, vol. 12, pp. 21-32, 2003. [Online]. Available: https://www.jstage.jst.go.jp/article/journalac2003/12/0/12_0_21/_pdf/-char/en. [Accessed : Feb. 23, 2020].

[7] Aomori Prefectural Government, “Aomori prefecture landscape color guide plan,” 2000. [Online]. Available: http://www.city.mutsu.lg.jp/index.cfm/38,1244,c,html/1673/shikisai-guideplan1.pdf. [Accessed: Feb. 23, 2020].

[8] Asahi Tostem Exterior Building Materials Co., Ltd., "Utilizing the image scale for color planning," asahitostem.co.jp, [Online]. Available: https://www.asahitostem.co.jp/product/bijoux/sheathing_sec023.php. [Accessed: Feb. 23, 2020].

[9] T. Kato, "Human media engineering," indsys.chuo-u.ac.jp, [Online]. Available: http://www.indsys.chuo-u.ac.jp/~kato/HM/HM-1-handout.pdf. [Accessed: Feb. 23, 2020].

[10] K. Yamawaki and H. Shiizuka, "Characteristic extraction of music with color image scale," Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, vol. 17, no. 5, pp. 615-621, 2005. [Online]. Available: https://www.jstage.jst.go.jp/article/jsoft/17/5/17_KJ00003659690/_pdf. [Accessed: Feb. 23, 2020].

[11] S. Namba, "Improvement of recognizability of home appliance function sounds," Toshiba Review, vol. 55, no. 7, 2000. [Online]. Available: https://www.toshiba.co.jp/tech/review/2000/07/f03.pdf. [Accessed: Feb. 23, 2020].

[12] X-Rite, "ColorChecker Passport," [Online]. Available: https://www.xrite.com/-/media/xrite/files/manuals_and_userguides/c/o/colorcheckerpassport_user_manual_en.pdf. [Accessed: Feb. 23, 2020].

[13] S. Lloyd, "Least squares quantization in PCM," *IEEE* Transactions on Information Theory, vol. 28, no. 2, pp. 129-137, March 1982.

[14] The MathWorks, Inc., "kmeans," mathworks.com, [Online]. Available: https://www.mathworks.com/help/stats/kmeans.html. [Accessed: Feb. 23, 2020].

[15] K. Ruxpaitoon and T. Lertrusdachakul, "Color analysis on the symbolic temple of king in Chakri dynasty of Thailand," 2019 11th International Conference on Knowledge and Smart Technology (KST2019), Phuket, Thailand, Jan. 23-26, 2019, pp. 17-22.

[16] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, Jan. 1979

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 59-69

การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกเพื่อวัดระดับความหวานของแตงโมผ่านสมาร์ทโฟน

Apply of Deep Learning Techniques to Measure the Sweetness Level of Watermelon via Smartphone

ณัฐวดี หงษ์บุญมี และ ณัฐพงศ์ จันต๊ะวงศ์*

Nattavadee Hongboonmee and Nutthapong Jantawong*

ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

Department of Computer Science and Information Technology, Faculty of Science, Naresuan University

Received: April 11, 2020; **Revised:** June 03, 2020; **Accepted:** June 03, 2020; **Published:** June 23, 2020

ABSTRACT – This research proposed the development of automated system for analyzing sweetness and watermelon varieties from photos using deep learning techniques for use on smartphones. To help the public who would like to know the names of varieties and sweetness of watermelons. The main components of the system include: (1) Modeling, classifying varieties and sweetness levels of watermelons with deep learning neural network through the TensorFlow library, using the InceptionV3 and MobileNet algorithms to compare image classification. In which the trainers are able to classify 4 types of images, each type of 100 images, and training our model for 500 rounds. The result shows that the model from the InceptionV3 algorithm has the same accuracy as the model from the MobileNet algorithm. The accuracy is 97.20%. Therefore considering the model size obtained from learning, it found that MobileNet model size is smaller than InceptionV3, so choose the model from MobileNet to develop the system further. (2) Using the model from MobileNet algorithm to develop application on smartphones, which developed by Android Studio program. Results of the user satisfaction test, found that the average satisfaction is 4.34, standard deviation 0.62, it at good level. In conclusion, this application is effective and can use in real life.

KEYWORDS: Watermelon Sweetness, Image Classification, Deep Learning, Convolution Neural-Network, Smartphone

บทคัดย่อ -- งานวิจัยนี้นำเสนอการพัฒนาระบบอัตโนมัติสำหรับวิเคราะห์ความหวานและพันธุ์แตงโมด้วยภาพถ่ายโดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับใช้งานบนสมาร์ทโฟนเพื่ออำนวยความสะดวกสำหรับบุคคลทั่วไปที่ต้องการทราบชื่อพันธุ์และความหวานแตงโม ส่วนประกอบหลักของระบบประกอบด้วย (1) การสร้างโมเดลจำแนกพันธุ์และระดับความหวานของแตงโมด้วยโครงข่ายประสาทเทียมการเรียนรู้เชิงลึกผ่านไลบรารี TensorFlow โดยนำอัลกอริทึม InceptionV3 และ MobileNet มาทำการทดลองเปรียบเทียบการจำแนกภาพ ซึ่งฝึกสอนให้สามารถจำแนกภาพจำนวน 4 ประเภท ประเภทละ 100 ภาพ ฝึกสอนจำนวน 500 รอบ ผลการทดลองพบว่าโมเดลจากอัลกอริทึม InceptionV3 มีค่าความ

*Corresponding Author: *nattavadeeho@nu.ac.th*

ถูกต้องที่เท่ากับโมเดลจากอัลกอริทึม MobileNet ค่าความถูกต้องเท่ากับ 97.20% แต่จากการพิจารณาขนาดของโมเดลที่ได้จากการเรียนรู้ พบว่า MobileNet มีขนาดของโมเดลเล็กกว่า InceptionV3 ดังนั้นจึงเลือกโมเดลจาก MobileNet ไปพัฒนาระบบต่อไป (2) การนำโมเดลจากอัลกอริทึม MobileNet ไปพัฒนาเป็นแอปพลิเคชันบนสมาร์ทโฟน ดำเนินการพัฒนาด้วยโปรแกรม Android Studio ผลการทดสอบความพึงพอใจการใช้แอปพลิเคชันจากผู้ใช้พบว่าความพึงพอใจเฉลี่ยเท่ากับ 4.34 ส่วนเบี่ยงเบนมาตรฐาน 0.62 อยู่ในเกณฑ์ระดับดี สามารถสรุปได้ว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถนำไปใช้งานจริงได้

คำสำคัญ: ความหวานแตงโม, การจำแนกหมวดหมู่ภาพ, การเรียนรู้เชิงลึก, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, สมาร์ทโฟน

1. บทนำ

ประเทศไทยเป็นประเทศเกษตรกรรมที่อุดมสมบูรณ์ไปด้วยผลไม้นานาชนิด แตงโมเป็นผลไม้ชนิดหนึ่งที่นิยมเพาะปลูกและสามารถให้ผลผลิตได้ตลอดทั้งปีทำให้แตงโมสามารถรับประทานได้ทุกฤดูกาล [1] แต่การตรวจสอบคุณภาพรสชาติแตงโมให้ได้มาตรฐานเป็นเรื่องยากสำหรับผู้ที่ไม่มีความเชี่ยวชาญ เนื่องจากในอดีตและปัจจุบันยังใช้วิธีการตรวจสอบความหวานของแตงโมจากประสบการณ์หรือวิธีการทดสอบปฏิกิริยาทางเคมีในห้องปฏิบัติการซึ่งต้องใช้เวลาในการประมวลผลนาน จึงทำให้ผู้บริโภคมักมีปัญหาเมื่อแตงโมที่ซื้อมามีคุณภาพรสชาติด้อยกว่าที่ควรอันเนื่องมาจากการผสมแตงโมหวานปะปนมากับแตงโมไม่หวาน ซึ่งสิ่งที่บ่งชี้ถึงคุณภาพความหวานแตงโมได้แก่ ลักษณะผล สี น้ำหนัก ความแก่ของแตงโม นอกจากนี้การตรวจสอบคุณภาพของแตงโมในด้านความหวานนั้นยังสามารถใช้การเคาะที่แตงโมได้ วิธีนี้ต้องอาศัยผู้ที่มีประสบการณ์และผู้ที่มีความชำนาญเท่านั้นจึงจะสามารถแยกแยะความหวานได้

จากปัญหาข้างต้นจึงเกิดแนวคิดในการวัดระดับความหวานของแตงโมโดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกด้วยการถ่ายภาพจากกล้องสมาร์ทโฟน วัตถุประสงค์ของงานวิจัยคือ (1) เพื่อสร้างโมเดลจำแนกพันธุ์และระดับความหวานของแตงโมโดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก (2) เพื่อพัฒนาแอปพลิเคชันแสดงชื่อพันธุ์และระดับความหวานของแตงโมด้วยการถ่ายภาพผ่านกล้องสมาร์ทโฟน แอปพลิเคชันที่พัฒนาจะช่วยเพิ่มความสะดวกรวดเร็วในการวัดความความหวานของแตงโมและช่วยให้ผู้ที่ไม่มีความเชี่ยวชาญสามารถแยกแยะความหวานของแตงโมได้ด้วยตนเอง นอกจากนี้ยังเป็นแนวทางที่จะช่วยพัฒนาเทคโนโลยีด้านการตรวจสอบความหวานของผลไม้ช่วยลดต้นทุนการผลิตเครื่องตรวจสอบคุณภาพของผลไม้ได้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 แตงโม

งานวิจัยนี้เลือกศึกษาแตงโมพันธุ์ที่นิยมบริโภคในปัจจุบัน 2 พันธุ์ ได้แก่ พันธุ์กินรีและพันธุ์ตอปิโด

พันธุ์กินรี [2] เป็นพันธุ์มาตรฐานที่นิยมปลูกในประเทศไทย การเลือกซื้อแตงโมพันธุ์กินรีที่ควรเลือกคือลูกที่มีผิวเรียบแตงโมสีเขียวสดแสดงว่ายังไม่แก่จัด ถ้ามีฝ้าขาวนวลขึ้นแสดงว่าแก่จัดหรือถ้าขั้วแตงโมงอลงแสดงว่าแก่จัดเช่นกัน

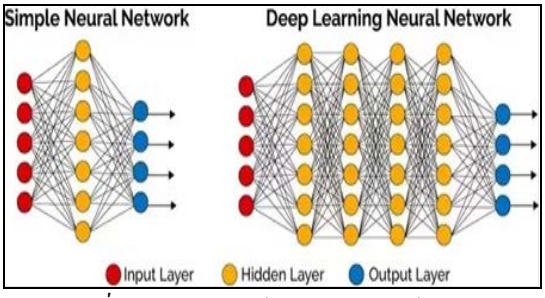
รูปที่ 1. แสดงแตงโมพันธุ์กินรี

พันธุ์ตอปิโด [2] เป็นพันธุ์ที่เนื้อเนียนแน่น เปลือกบางละเอียดรสชาติดี ลักษณะผลยาวรีทรงคล้ายหมอน เปลือกบางผิวสีเขียวเข้มสลับลายสีชัดเจนเนื้อสีแดงหวานกรอบ

รูปที่ 2. แสดงแตงโมพันธุ์ตอปิโด

2.2 การเรียนรู้เชิงลึก

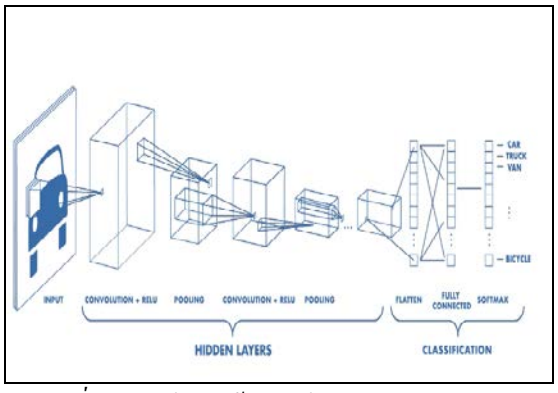
การเรียนรู้เชิงลึก (Deep Learning) [3] คือชุดคำสั่งที่สร้างขึ้นมาเพื่อการเรียนรู้ของเครื่องคอมพิวเตอร์โดยชุดคำสั่งนี้จะทำให้เครื่องสามารถประมวลผลข้อมูลจำนวนมากที่พยายามเรียนรู้วิธีการแทนข้อมูลอย่างมีประสิทธิภาพ หลักการของการเรียนรู้เชิงลึกเป็นโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ที่เป็นโหนดหลายๆ ชั้นและใช้การประมวลผลแบบขนานทำให้สามารถประมวลผลได้ครั้งละจำนวนมากช่วยให้การเรียนรู้ของเครื่องสามารถให้ผลลัพธ์ในการตัดสินใจและคาดการณ์ได้ดีมากยิ่งขึ้น



รูปที่ 3. แสดงโครงสร้างของการเรียนรู้เชิงลึก

2.3 โครงข่ายประสาทเทียมแบบคอนโวลูชัน

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Networks: CNN) [4] เป็นรูปแบบหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning) ที่จำลองรูปแบบประเภทโครงข่ายประสาทเทียมให้เป็นรูปแบบการเรียนรู้เชิงลึก (Deep Learning) ที่เป็นลักษณะโครงข่ายประสาทเทียมเชิงลึก (Deep Learning Neural Network) โดยมีกระบวนการสกัดคุณลักษณะของภาพโดยใช้ Convolutional Layer ผ่านทางฟิลเตอร์ในแต่ละ Convolutional Layer ผู้ใช้ไม่จำเป็นต้องกำหนดรูปแบบในการสอนให้กับโมเดลเพียงแค่ทำการเตรียมข้อมูลภาพตัวอย่างที่ต้องการใช้งานเอาไว้แล้วนำภาพเหล่านั้นป้อนเข้ากระบวนการ CNN จะทำเรียนรู้โดยอัตโนมัติและหากผู้ใช้ต้องการที่จะทำนายรูปภาพก็เพียงแค่ป้อนรูปภาพที่ต้องการทำนายเข้าไป CNN จะทำการเรียนรู้เพื่อเปรียบเทียบรูปภาพที่ต้องการทำนายกับข้อมูลรูปภาพที่มีอยู่เพื่อแสดงผลลัพธ์ออกมา



รูปที่ 4. แสดงโครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชัน

โดยทั่วไป CNN ประกอบด้วยส่วนประกอบแยกเป็นชั้น (Layer) ดังนี้ (รูปที่ 4)

1. Input Layer สำหรับอ่าน Input Image เข้ามาในโครงข่ายประสาทเทียม
2. Convolutional Layer ถูกออกแบบเพื่อสกัด Features จาก การดำเนินการในระดับพิกเซลออกจาก Input Image ผลลัพธ์ ของชั้นคอนโวลูชันคือ Convolution Feature Map
3. ReLU (Rectified Linear Unit) คือชั้นของ Non-linear Activation Function
4. Pooling Layer มีจุดประสงค์ในการ Subsample Rectified Feature Map เพื่อลดมิติเชิงพื้นที่และสร้าง Feature Representation ขนาดเล็ก
5. Softmax Layer คือ Layer สุดท้ายของโครงข่าย เพื่อให้ Output ออกมาเป็น Multiclass Logistic Classifier
6. Output Layer คือชั้นของการนำเสนอผลลัพธ์ของการ Classification ที่ได้มาจากชั้น Softmax Layer ก่อนหน้า

2.4 MobileNet

MobileNet [5] เป็นโครงข่ายประสาทเทียมแบบคอนโวลูชันรูปแบบหนึ่งที่ถูกพัฒนาขึ้นโดย Andrew G. Howard และคณะในปีพ.ศ.2560 โดยรูปแบบของ MobileNet ถูกพัฒนาให้เป็น

โครงข่ายประสาทเทียมที่เล็กกว่าโครงข่ายประสาทเทียมแบบเชิงลึกทั่วๆ ไป เพื่อให้สามารถนำโมเดลไปใช้งานบนอุปกรณ์คอมพิวเตอร์ขนาดเล็ก เช่น สมาร์ทโฟนและยังสามารถรักษาประสิทธิภาพในการทำงานไว้ได้ในระดับที่ใกล้เคียงกับโครงข่ายประสาทเทียมแบบเชิงลึกขนาดใหญ่

2.5 InceptionV3

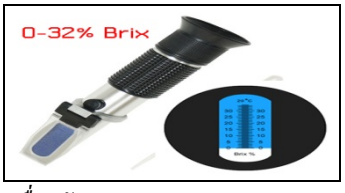
ปัจจุบันการประยุกต์ใช้การเรียนรู้เชิงลึกในด้านการจำแนกวัตถุในภาพหรือการรู้จำรูปภาพ (Image Recognition) ได้ถูกพัฒนาไปอย่างกว้างขวางมีผู้พัฒนาหลายกลุ่มได้สร้างโมเดลของโครงข่ายประสาทเทียมแบบคอนโวลูชันในแบบของตนเองขึ้นมา เช่น InceptionV3 ของทีมผู้พัฒนา GoogLeNet [6] เป็นโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ถูกพัฒนาให้มีจำนวน 22 ชั้น (Layers) เพื่อให้สามารถทำการตรวจสอบและจำแนกองค์ประกอบโดยรวมของวัตถุแต่ละชนิดได้ดียิ่งขึ้น

2.6 TensorFlow

TensorFlow [7] คือไลบรารี่ที่มีการทำงานของเทคโนโลยีการเรียนรู้เชิงลึก พัฒนาโดยกูเกิลสามารถนำมาประยุกต์ใช้งานได้อย่างยืดหยุ่นสามารถนำไปใช้งานได้แบบ Portable มีการใช้งานง่ายและเป็น Open Source สามารถรองรับการทำงานบนเครื่องเดสก์ท็อปและโมบายได้

2.7 เครื่องวัดความหวาน

เครื่องวัดความหวาน 0-32% Brix เป็นเครื่องมือวัดช่วง 0-32% Brix [8] จุดเด่นอยู่ที่การใช้หักเหของแสงจากปริซึมอ่านค่าในสเกลทำให้วัดค่าได้แม่นยำเหมาะสำหรับการทดสอบความหวานหรือระดับน้ำตาลในอาหารและผลไม้ต่างๆ



รูปที่ 5. เครื่องวัดความหวานแบบ Brix Refractometer

ระดับความหวานของแตงโมสามารถวัดได้จากเครื่องวัดค่าความหวาน 0-32% Brix โดยเฉลี่ยแล้วแตงโมกินรีและแตงโมตอปิโดจะมีความหวานเฉลี่ยอยู่ที่ 10 – 14 บริกซ์ ดังนั้นงานวิจัยนี้จึงกำหนดระดับความหวานแตงโมไว้ดังตารางที่ 1

ตารางที่ 1. การกำหนดระดับความหวานแตงโมในงานวิจัย

ค่าจากเครื่อง *Brix Refractometer*	กำหนดระดับความหวาน
0-9 บริกซ์	ไม่หวาน
>= 10 บริกซ์	หวาน

2.8 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการจำแนกภาพด้วยเทคนิคการเรียนรู้เชิงลึกพบว่า

ชิดชัยและคณะ [9] ศึกษาเรื่องระบบแจ้งเตือนตรวจสอบและระบุสุนัขที่สูญหายโดยใช้เทคนิคโครงข่ายประสาทเทียมบน TensorFlow นอกจากนั้นยังพัฒนาแอปพลิเคชันบนโทรศัพท์เคลื่อนที่โดยใช้ Android Studio ผลการทดสอบความแม่นยำของการแยกสายพันธุ์สุนัขจำนวน 10 สายพันธุ์ พบว่าค่าความแม่นยำเท่ากับ 86.50%

พรรษดิชัยและคณะ [10] ศึกษาเรื่องการตรวจจับและจำแนกยานพาหนะจากวิดีโอรักษาความปลอดภัยด้วยเทคนิคการเรียนรู้เชิงลึก งานวิจัยนี้ได้นำเทคนิค InceptionV3 และ MobileNet มาทำการทดลองเปรียบเทียบการจำแนกรถจักรยานยนต์ด้วยข้อมูลภาพและวิดีโอของยานพาหนะ ผลการวิจัยพบว่าเทคนิค InceptionV3 มีค่าความถูกต้องที่ใกล้เคียงกันกับ MobileNet แต่จากการทดลองกับวิดีโอพบว่ายังไม่สามารถจำแนกภาพรถจักรยานยนต์ได้ทั้งหมดโดยเฉพาะที่เป็นภาพระยะไกล

อรรถพลและคณะ [11] นำเสนอการจำแนกพระเครื่องโดยใช้เทคนิคการเรียนรู้เชิงลึกโดยใช้โครงข่ายประสาทเทียมแบบ CNN ประเภท Xception เป็นต้นแบบในการสกัดคุณลักษณะและสร้างโมเดล ชุดข้อมูลสำหรับการทำวิจัยถูกสร้างจากการดาวน์โหลดรูปพระเครื่องจากอินเทอร์เน็ตโดยทำ

การคัดเลือกรูปพระเครื่องที่เป็นที่นิยมสะสมและมีมูลค่าสูง จำนวน 10 ประเภท ผลการทดลองพบว่าโมเดลมีค่าความแม่นยำ 96.90%

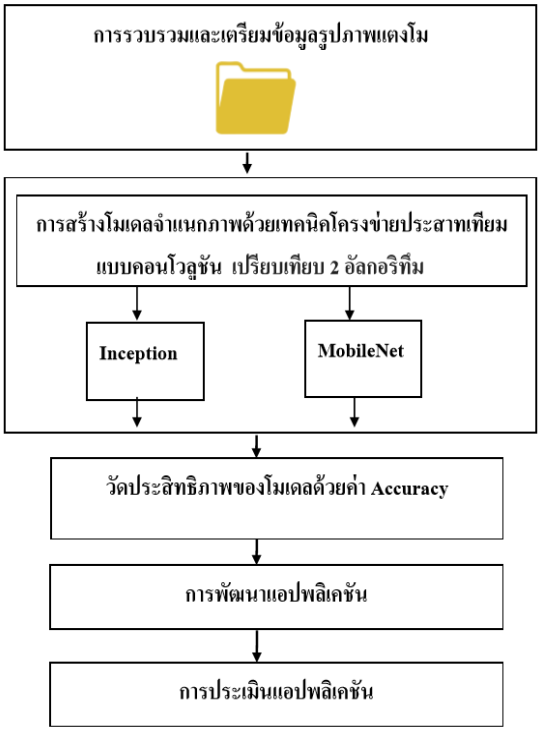
DeepFood [12] ได้นำเสนอการจำแนกวัตถุดิบในการทำอาหารจากภาพโดยใช้เทคนิคการเรียนรู้เชิงลึกประเภทโครงข่ายประสาทเทียมแบบ CNN โดยแบบจำลองเรียนรู้วัตถุดิบจำนวน 41 ประเภทจากภาพวัตถุดิบประเภทละ 100 ภาพ ซึ่งความแม่นยำของแบบจำลองประเภทต่าง ๆ เฉลี่ยอยู่ที่ 80.00%

Singh [13] นำวิธี CNN ที่ใช้โครงสร้าง LeNet-5 มาใช้เพื่อจำแนกประเภทข้อมูลที่อยู่ในชุดข้อมูลย่อยของ SUN ซึ่งมีรูปภาพจำนวน 3,000 รูปภาพที่ประกอบด้วย 8 คลาสได้แก่ น้ำ รถ ภูเขา พื้นดิน ต้นไม้ ตึก หิมะและท้องฟ้า ในการทดลองได้แบ่งข้อมูลออกเป็น 80% สำหรับข้อมูลชุดเรียนรู้และ 20% สำหรับข้อมูลชุดทดสอบ ผลการทดลองพบว่าวิธี CNN มีความผิดพลาด (Error Rate) ในการจำแนกประเภทข้อมูล 27.97% ซึ่งน้อยกว่าวิธี Bag of Words ที่มีความผิดพลาดสูงถึง 47.44%

จากศึกษางานวิจัยได้ข้อสรุปว่าเทคนิคการเรียนรู้เชิงลึกประเภทโครงข่ายประสาทเทียมแบบ CNN มีความสามารถในการจำแนกและรู้จำภาพที่มีผลลัพธ์การทำนายที่แม่นยำและน่าพอใจ ดังนั้นงานวิจัยนี้จึงประยุกต์ใช้การเรียนรู้เชิงลึกประเภทโครงข่ายประสาทเทียมแบบ CNN มาประยุกต์ใช้กับการจำแนกพันธุ์และระดับความหวานแตงโมเพื่อประสิทธิภาพที่แม่นยำมากยิ่งขึ้น

3. วิธีดำเนินการศึกษา

การจำแนกพันธุ์และระดับความหวานของแตงโมด้วยภาพถ่ายโดยใช้เทคโนโลยีการเรียนรู้เชิงลึก มีการดำเนินงาน 5 ขั้นตอน ได้แก่ 1) การรวบรวมและเตรียมข้อมูล 2) การสร้างโมเดลจำแนกภาพ 3) การวัดประสิทธิภาพโมเดล 4) การพัฒนาแอปพลิเคชันและ 5) การประเมินแอปพลิเคชัน แสดงรายละเอียดดังรูปที่ 6



รูปที่ 6. แสดงขั้นตอนการดำเนินงาน

3.1 การรวบรวมและเตรียมข้อมูล

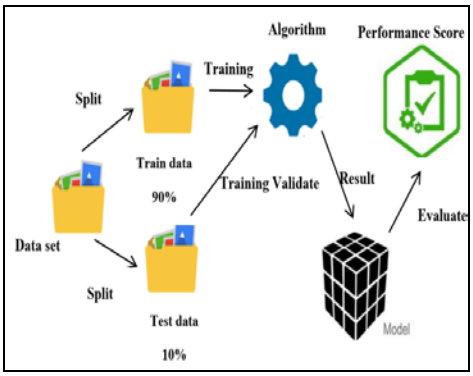
การเก็บรวบรวมและเตรียมข้อมูล ในการศึกษานี้ใช้กลุ่มตัวอย่างแตงโมลูกเดี่ยวจำนวน 400 ลูก แบ่งข้อมูลภาพออกเป็น 4 คลาส ได้แก่ 1) พันธุ์กินรี รสชาติหวาน 2) พันธุ์กินรี รสชาติไม่หวาน 3) พันธุ์ตอปิโด รสชาติหวานและ 4) พันธุ์ตอปิโด รสชาติไม่หวาน คลาสละ 100 รูป รวมทั้งหมดจำนวน 400 รูป ทำการถ่ายภาพแตงโมด้วยกล้องถ่ายภาพแบบควบคุมแสง กำหนดขนาดภาพทุกภาพเท่ากับ 224x224 พิกเซล ดำเนินการแบ่งข้อมูล 400 ตัวอย่างเป็น 2 ส่วนคือ ข้อมูลส่วนที่หนึ่ง 90% จำนวน 360 ตัวอย่างใช้ในการฝึกสอนสร้างโมเดล ข้อมูลส่วนที่สอง 10% จำนวน 40 ตัวอย่างนำมาใช้ในการทดสอบโมเดล ส่วนการกำหนดระดับความหวานใช้การเทียบความหวานกับเครื่องวัดค่าความหวาน 0-32% Brix ข้อมูลตัวอย่างทั้งหมดนี้นำมาใช้เป็นข้อมูลฝึกสอนของโครงข่ายประสาทเทียมทั้งแบบ InceptionV3 และ MobileNet ตัวอย่างภาพแสดงรายละเอียดดังตารางที่ 2

ตารางที่ 2. จำนวนรูปภาพแตงโมที่นำมาทดลอง

พันธุ์/ ระดับความหวาน	ข้อมูลฝึกสอน (ภาพ)	ข้อมูลทดสอบ (ภาพ)	จำนวนทั้งหมด (ภาพ)
 กินรี หวาน	90	10	100
 กินรี ไม่หวาน	90	10	100
 ตอปิโด หวาน	90	10	100
 ตอปิโด ไม่หวาน	90	10	100
รวม	**360**	**40**	**400**

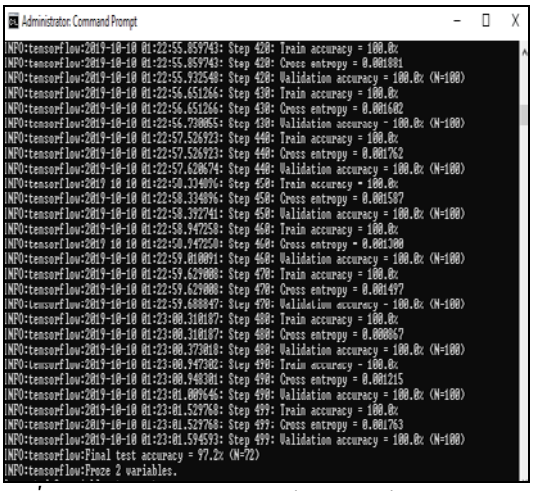
3.2 การสร้างโมเดลจำแนกภาพ

การสร้างโมเดลจำแนกพันธุ์และระดับความหวานแตงโม มีขั้นตอนดังรูปที่ 7



รูปที่ 7. ภาพรวมขั้นตอนการสร้างโมเดล

การสร้างโมเดลเริ่มจากนำรูปภาพที่เตรียมไว้มาทำการแบ่งข้อมูลเป็นสองส่วนคือข้อมูลสำหรับการฝึกสอน (Train Dataset) 90% และข้อมูลสำหรับทดสอบ (Test Dataset) 10% ข้อมูลภาพนำเข้าขนาด 224x224 พิกเซล ทำการฝึกสอนและสร้างโมเดลด้วยภาษา Python และไลบรารี่ TensorFlow ซึ่งเป็นไลบรารี่สำหรับฝึกสอนและสร้างโมเดลที่มีการใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน เลือกทดลองเปรียบเทียบ 2 อัลกอริทึมจาก CNN ได้แก่อัลกอริทึม MobileNet และ InceptionV3 โดยทำการฝึกสอนจำนวน 500 รอบ



รูปที่ 8. แสดงการฝึกสอนและสร้างโมเดลด้วยภาษา Python

รูปที่ 8 แสดงการฝึกสอนและสร้างโมเดลรู้จำภาพด้วยภาษา Python โดยเมื่อฝึกสอนเรียบร้อยจะได้ผลการสร้างโมเดลประกอบด้วยไฟล์ retrained_graph.pb และ retrained_labels.txt ซึ่งจะนำไปใช้ในขั้นตอนต่อไป

3.3 การวัดประสิทธิภาพโมเดล

การวัดประสิทธิภาพใช้ตัวชี้วัด ได้แก่ ค่าความถูกต้อง (Accuracy) [14] เป็นตัววัดผลประสิทธิภาพของโมเดล มีวิธีการคำนวณดังสมการที่ 1 ดังนี้

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

โดยที่

True Position (TP) คือ จำนวนข้อมูลที่ทำนายถูกต้องและผลลัพธ์ถูกต้อง

True Negative (TN) คือ จำนวนข้อมูลที่ทำนายไม่ถูกต้องและผลลัพธ์ไม่ถูกต้อง

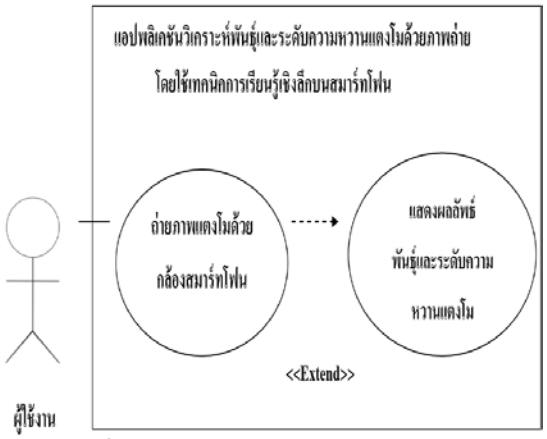
False Positive (FP) คือ จำนวนข้อมูลที่ทำนายถูกต้องและผลลัพธ์ไม่ถูกต้อง

False Negative (FN) คือ จำนวนข้อมูลที่ทำนายไม่ถูกต้องและผลลัพธ์ถูกต้อง

การทดสอบวัดประสิทธิภาพการเรียนรู้ของแต่ละโมเดลงานวิจัยนี้ทำการเปรียบเทียบประสิทธิภาพทั้งในแง่ของความถูกต้องและขนาดของโมเดลที่ได้จากการเรียนรู้ โดยผลที่ได้จากการทดลองทั้งหมดถูกนำเสนอในหัวข้อถัดไป

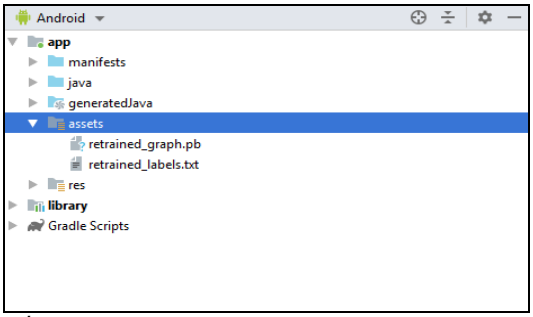
3.4 การพัฒนาแอปพลิเคชัน

การพัฒนาแอปพลิเคชันวัดความหวานและจำแนกพันธุ์แตงโมด้วยภาพถ่าย ดำเนินการออกแบบระบบด้วยแผนภาพ Use Case Diagram (รูปที่ 9) การทำงานของระบบจะสามารถถ่ายภาพแตงโมด้วยกล้องสมาร์ทโฟนจากนั้นทำการรู้จำและจำแนกภาพด้วยโมเดลที่ผ่านการฝึกสอนเรียบร้อยแล้วและทำการแสดงผลด้วยการระบุพันธุ์และความหวานแตงโมที่ตรวจพบ ซึ่งมีการประมวลผลในแบบเรียลไทม์



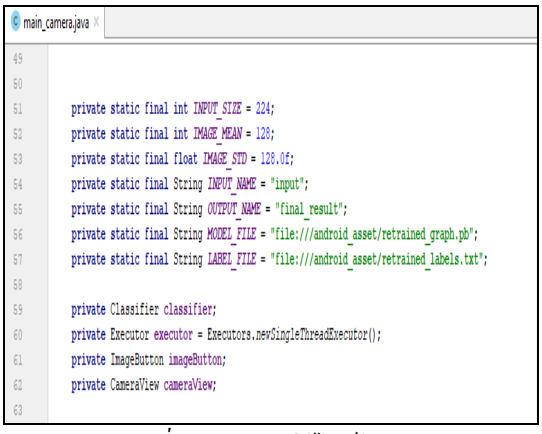
รูปที่ 9. แสดง Use Case Diagram ของระบบ

การพัฒนาระบบดำเนินการพัฒนาระบบในรูปแบบโมบายแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์เพื่อให้สะดวกต่อการใช้งาน โดยใช้โปรแกรม Android Studio และภาษา JAVA เป็นภาษาหลักในการพัฒนา โดยนำโมเดลจำแนกภาพที่ผ่านการฝึกสอนเรียบร้อยแล้วเข้าโปรแกรม Android Studio เพื่อเรียกใช้ประมวลผลในแอปพลิเคชัน ซึ่งไฟล์ที่ได้จากการฝึกสอนจะมีจำนวนสองไฟล์ คือไฟล์ retrained_graph.pb และไฟล์ retrained_labels.txt



รูปที่ 10. การนำโมเดลเข้าใช้งานในโปรแกรม Android Studio

การเรียกใช้งานโมเดลในโปรแกรม Android Studio โดยเรียกใช้ในหน้า main_camera.java เรียกใช้โดยการ Set Path ไปยังไฟล์โมเดลที่อยู่โฟลเดอร์ assets ที่สร้างขึ้นในโปรแกรม Android Studio ดังรูปที่ 10-12



รูปที่ 11. การเรียกใช้ไฟล์โมเดล

รูปที่ 12. การดึงข้อมูลมาแสดงผล

แอปพลิเคชันจะทำการประมวลผลภาพจำแนกพันธุ์และระบุความหวานแตงโมว่าเป็นแตงโมหวานหรือแตงโมไม่หวาน พร้อมทั้งบอกค่าความถูกต้องเป็นจำนวนเปอร์เซ็นต์ (%)

3.5 การประเมินแอปพลิเคชัน

การประเมินแอปพลิเคชันใช้การประเมินแบบแบล็กบ็อกซ์ (Black-box Testing) เพื่อตรวจสอบการทำงานของแอปพลิเคชันในแต่ละส่วนว่ามีการทำงานที่ถูกต้องตามเงื่อนไขที่กำหนดไว้ และเพื่อหาข้อผิดพลาดที่อาจจะเกิดขึ้นกับแอปพลิเคชัน

4. ผลการศึกษาและการอภิปรายผล

ผลการศึกษาเปรียบเทียบประสิทธิภาพโมเดลและผลการพัฒนาแอปพลิเคชัน มีรายละเอียดดังต่อไปนี้

4.1 ผลการวัดประสิทธิภาพโมเดล

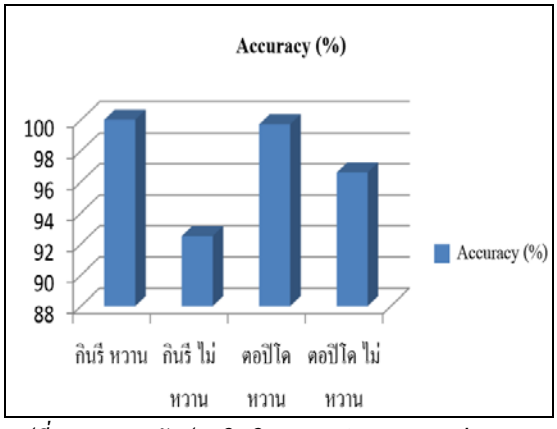
การศึกษานี้ได้ทดลองเปรียบเทียบการสร้างโมเดลรู้จำภาพด้วยเทคนิค CNN โดยทำการเปรียบเทียบ 2 อัลกอริทึมจากวิธี CNN ได้แก่ อัลกอริทึม MobileNet และ InceptionV3 ผลการเปรียบเทียบประสิทธิภาพโมเดล พิจารณาจากความถูกต้องและขนาดของโมเดลที่ได้จากการเรียนรู้ ดังตารางที่ 3

ตารางที่ 3. แสดงผลการเปรียบเทียบโมเดล

Model	Final Accuracy (%)	Model Size
InceptionV3	97.20	85.40 MB
MobileNet	**97.20**	**16.73 MB**

จากผลการทดลองที่ได้นำรูปภาพทั้งหมด 400 ภาพ มาทำการเรียนรู้กับโครงข่ายประสาทเทียมแบบ InceptionV3 และ MobileNet เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกรูปภาพ ทดสอบการเรียนรู้ของโมเดลด้วยการแบ่งข้อมูลออกมาเป็นข้อมูลฝึกสอนจำนวน 90% และข้อมูลทดสอบจำนวน 10% ของข้อมูลทั้งหมด พบว่าความถูกต้องในการจำแนกรูปภาพความหวานและพันธุ์แตงโมมีความถูกต้อง (Final Accuracy) ที่เท่ากันในทั้งสองโครงข่ายคือ 97.20% ซึ่งถือว่าทั้งสองโครงข่ายมีประสิทธิภาพเท่ากัน แต่ถ้าดูจากขนาดของโมเดลที่ได้จากการเรียนรู้ของโครงข่ายจะเห็นว่า MobileNet มีขนาดของโมเดลเล็กกว่า InceptionV3 อย่างเห็นได้ชัดโดยมีขนาดเพียงแค่ 16.73 เมกะไบต์ (MB) หรือว่าเพียงหนึ่งในห้าส่วนของ InceptionV3 เท่านั้น

ดังนั้นจึงเลือกใช้โมเดลจากโครงข่าย MobileNet ที่มีขนาดเล็กแต่มีค่า Final Accuracy ที่สูงเท่ากับ InceptionV3 เพื่อให้เหมาะสมกับการนำไปประยุกต์ใช้ประมวลผลกับสมาร์ทโฟนที่มีขนาดหน่วยความจำที่มีขนาดเล็ก



รูปที่ 13. ผลการวัดประสิทธิภาพการจำแนกของแต่ละคลาสจากอัลกอริทึม MobileNet

รูปที่ 13 แสดงผลการวัดประสิทธิภาพการจำแนกแต่ละคลาสจากอัลกอริทึม MobileNet พบว่าโมเดลมีการจำแนกภาพได้อย่างมีประสิทธิภาพที่ดี คลาสที่สามารถจำแนกได้ค่าความถูกต้องมากที่สุด ได้แก่ คลาสแตงโมกินรี หวาน มีค่าความถูกต้อง 100.00% รองลงมาคือคลาสตอปิโด หวาน 99.70% คลาสตอปิโด ไม่หวาน 96.60% คลาสที่มีค่าความถูกต้องน้อยที่สุด ได้แก่ คลาสแตงโมกินรี ไม่หวาน มีค่าความถูกต้อง 92.50% ส่วนค่าความถูกต้องเฉลี่ยรวม คือ 97.20% สรุปได้ว่าโมเดลจากอัลกอริทึม MobileNet มีความแม่นยำที่สูงสามารถนำไปพัฒนาแอปพลิเคชันวัดระดับความหวานของแตงโมได้

4.2 ผลการพัฒนาแอปพลิเคชัน

ผลการพัฒนาแอปพลิเคชันจำแนกพันธุ์และวัดความหวานแตงโม มีรายละเอียดดังรูปที่ 14-16 หน้าจอหลักแอปพลิเคชัน (รูปที่ 14) ผู้ใช้งานสามารถเลือกเมนูถ่ายรูปแตงโมเพื่อวิเคราะห์พันธุ์และความหวานของแตงโม

รูปที่ 14. แสดงหน้าจอหลักของแอปพลิเคชัน

เมื่อผู้ใช้คลิกเมนูถ่ายภาพ แอปพลิเคชันจะแสดงหน้าจอดังรูปที่ 15 โดยแอปพลิเคชันจะประมวลผลภาพและวิเคราะห์ภาพแตงโมว่าเป็นแตงโมพันธุ์กินรีหรือตอปิโด และแสดงระดับความหวานว่าเป็นแตงโมหวานหรือไม่หวานพร้อมทั้งบอกค่าความถูกต้องเป็นจำนวนเปอร์เซ็นต์(%)

รูปที่ 15. แสดงหน้าจอถ่ายรูป

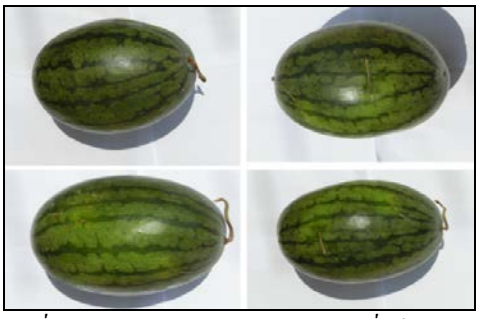
รูปที่ 16. แสดงหน้าจอผลลัพธ์

ผลการทดลองการจำแนกพันธุ์และระดับความหวานแตงโมจากแอปพลิเคชันจริง โดยได้ทดลองการจำแนกแตงโมแต่ละคลาส คลาสละ 20 รูป แตงโมที่ใช้ทดสอบเป็นแตงโมคนละชุดกับภาพแตงโมที่ใช้ฝึกสอน (Unseen Data) โดยข้อกำหนดของการถ่ายภาพแตงโม คือ 1) การถ่ายภาพต้องเห็นส่วนของแตงโมทั้งลูก 2) ภาพถ่ายจะต้องไม่เบลอจนเกินไป จากตารางที่ 4 แสดงผลการทดสอบการจำแนกพันธุ์และความหวานแตงโม โดยสรุปค่าเฉลี่ยความแม่นยำในการจำแนกพันธุ์และความหวานแตงโมเท่ากับร้อยละ 87.50%

ตารางที่ 4. แสดงผลการทดสอบการจำแนกพันธุ์และความหวานของแอปพลิเคชัน

พันธุ์/ ความหวาน	จำแนกถูกต้อง	จำแนกผิด	Accuracy (%)
กินรี หวาน	19	1	95.00
กินรี ไม่หวาน	16	4	80.00
ตอปิโด หวาน	18	2	90.00
ตอปิโด ไม่หวาน	17	3	85.00
เฉลี่ย			**87.50**

การทดสอบการใช้งานจริงแอปพลิเคชัน พบว่าแสงสว่างและฉากหลังของรูปภาพมีส่วนในความถูกต้องของการจำแนกภาพ คลาสที่จำแนกได้ค่าความถูกต้องน้อยที่สุด ได้แก่ กินรี ไม่หวาน เนื่องจากภาพแตงโมกินรี ไม่หวาน และกินรีหวาน มีความคล้ายคลึงกันทำให้แอปพลิเคชันประมวลผลผิดพลาดได้ ดังตัวอย่างรูปที่ 17-18



รูปที่ 17. ตัวอย่างรูปในคลาสกินรีหวานที่ใช้ฝึกสอน

รูปที่ 18. ตัวอย่างรูปในคลาสกินรี ไม่หวานที่ใช้ฝึกสอน

4.3 ผลการประเมินความพึงพอใจจากผู้ใช้งานแอปพลิเคชัน

การประเมินความพึงพอใจในการใช้งานแอปพลิเคชันจากผู้ใช้งานจำนวน 30 คน โดยแบ่งการประเมินผลออกเป็น 3 ด้าน คือ ด้านการทำงานตรงตามความต้องการของผู้ใช้ ด้านความถูกต้องในการทำงานและด้านการใช้งาน ซึ่งผลการประเมินในภาพรวมอยู่ในเกณฑ์ดี โดยการประเมินความพึงพอใจมีค่าเฉลี่ยเท่ากับ 4.34 ส่วนเบี่ยงเบนมาตรฐาน 0.62 รายละเอียดผลการประเมินดังตารางที่ 5

ตารางที่ 5. ผลการประเมินความพึงพอใจจากผู้ใช้งาน

รายการประเมิน	Mean	S.D.	การแปลผล
1.ด้านการทำงานตรงตามความต้องการของผู้ใช้ (Functional Requirement Test)	4.37	0.60	ดี
2.ด้านความถูกต้องในการทำงาน (Functional Test)	4.20	0.54	ดี
3.ด้านการใช้งาน (Usability Test)	4.47	0.72	ดี
เฉลี่ยรวม	**4.34**	**0.62**	**ดี**

5. บทสรุป

การศึกษานี้นำเสนอการจำแนกพันธุ์และความหวานของแตงโมด้วยภาพถ่ายโดยประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก ซึ่งเทคนิคนี้จะให้ผลลัพธ์ที่รวดเร็วเพียงผู้ใช้งานแอปพลิเคชันใช้กล้องสมาร์ทโฟนถ่ายรูปแตงโมที่ต้องการจะวัดความหวานจากนั้นแอปพลิเคชันจะประมวลผลจำแนกพันธุ์และความหวานของแตงโมได้ การดำเนินงานเริ่มจากการรวบรวมรูปภาพกลุ่มตัวอย่างสำหรับทดลอง โดยแบ่งคลาสการจำแนกเป็น 4 คลาส ได้แก่ พันธุ์กินรีหวาน กินรีไม่หวาน พันธุ์ตอปิโดหวาน ตอปิโดไม่หวาน คลาสละ 100 ภาพ รวมทั้งหมด 400 ภาพ เปรียบเทียบความหวานแตงโมโดยเทียบกับเครื่องวัดความหวานแบบ Brix Refractometer จากนั้นนำภาพตัวอย่างไปดำเนินการฝึกสอนและสร้างโมเดลการจำแนกภาพด้วยโครงข่ายประสาทเทียมแบบ CNN ภายใต้ไลบรารี่ TensorFow เลือกเปรียบเทียบ 2 อัลกอริทึม คือ MobileNet และ InceptionV3 รอบการฝึกสอนจำนวน 500 รอบ จากการเปรียบเทียบพบว่าโมเดลจากอัลกอริทึม MobileNet มีค่า Final Accuracy ที่เท่ากับโมเดลจากอัลกอริทึม InceptionV3 คือ 97.20% แต่ขนาดของโมเดลจะแตกต่างกัน ดังนั้นจึงเลือกใช้โมเดล MobileNet ที่มีขนาดเล็กแต่มีค่า Final accuracy ที่สูงเท่ากับ InceptionV3 เพื่อให้เหมาะสมกับการนำไปประยุกต์ใช้ประมวลผลกับสมาร์ทโฟนที่มีขนาดหน่วยความจำที่มีขนาดเล็ก

ผลการทดสอบประสิทธิภาพการจำแนกพันธุ์และความหวานแตงโมจากการใช้งานจริงแอปพลิเคชัน โดยวัดจากร้อยละความถูกต้องจากการทดลองจำแนกแต่ละคลาส จำนวนคลาสละ 20 ครั้ง สามารถจำแนกได้คิดเป็นร้อยละความถูกต้องเฉลี่ย 87.50% คลาสที่จำแนกได้ถูกต้องมากที่สุด ได้แก่ แตงโมกินรีหวาน จำแนกได้ถูกต้อง 95.00% ส่วนคลาสที่จำแนกได้ถูกต้องน้อยที่สุด ได้แก่ แตงโมกินรี ไม่หวาน จำแนกได้ถูกต้อง 80.00% อาจเนื่องมาจากภาพแตงโมกินรี ไม่หวาน และกินรีหวาน มีความคล้ายคลึงกันทำให้แอปพลิเคชันประมวลผลผิดพลาด นอกจากนี้แสงและภาพพื้นหลังก็มีผลกับการประมวลผลของแอปพลิเคชันเช่นเดียวกัน

ผลการประเมินความพึงพอใจแอปพลิเคชันจากผู้ใช้งานพบว่าผู้ใช้มีความพึงพอใจเฉลี่ยเท่ากับ 4.34 ส่วนเบี่ยงเบนมาตรฐาน 0.62 อยู่ในเกณฑ์ระดับดี สามารถสรุปได้ว่าแอปพลิเคชันนี้มีประสิทธิภาพสามารถนำไปใช้ประโยชน์ในการเพิ่มความสะดวกรวดเร็วให้กับผู้ที่ไม่มีความเชี่ยวชาญในการวิเคราะห์พันธุ์และความหวานแตงโม นอกจากนี้ยังเป็นแนวทางที่จะช่วยพัฒนาเทคโนโลยีด้านการตรวจสอบความหวานของผลไม้ช่วยลดต้นทุนการผลิตเครื่องตรวจสอบคุณภาพของผลไม้ได้

ข้อเสนอแนะเพิ่มเติมสำหรับงานครั้งต่อไปคือการเพิ่มจำนวน Train Data Set ทั้งนี้เนื่องจากจำนวนของรูปภาพที่ใช้ในการทดสอบครั้งนี้มีจำนวนจำกัดและการควบคุมสภาวะแวดล้อมในการทดสอบเช่นแสงและภาพพื้นหลังซึ่งจะส่งผลให้โมเดลมีประสิทธิภาพและได้มาตรฐานที่สูงขึ้น

เอกสารอ้างอิง

[1] E. MuengKasem, “Watermelon Secrets We Never Knew: I Love Nature Set,” 1st ed. Bangkok: Nanmeebooks Kiddy, 2016.

[2] T. Chaerueangyot, “Growing crops and making good money,” 1st ed. Bangkok: Agricultural Knowledge Distribution Club, 2015.

[3] Q. Wu, Y. Liu, Q. Li, S. Jim and F. Li, “The Application of Deep Learning in Computer Vision,” in Proceeding of 2017 Chinese Automation Congress *(CAC)*, Jinan, China, 2017, pp. 6522–6527.

[4] P. Chandran, B. Byju, R. Deepak, K. Nishakumari, P. Devanand and P. Sasi, “Missing Child Identification System using Deep Learning and Multiclass SVM,” in Proceeding of 2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Thiruvananthapuram, India, 2018, pp. 113–116.

[5] B. Debnath, M. O’Brien, M. Yamaguchi and A. Behera, “Adapting MobileNets for Mobile Based Upper Body Estimation,” in Proceeding of 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 2018.

[6] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, Rethinking the Inception Architecture for Computer Vision. In Proceeding of 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, Nevada, 2016, pp. 2818-2826.

[7] F. Ertam and G. Aydin, “Data Classification with Deep Learning using Tensorflow,” in Proceeding of 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 2018, pp. 755–758.

[8] KBThaiscale, “Brix Refractometer 0-32%,” 2019. [Online]. Available: http://www.kbthaiscale.com/product/489. [Accessed: April 11, 2020]

[9] C. Sarawong, S. Somabut, P. Imtongkhum, C. Pimson and C. So-In, “Notification Validation and Identification Systems of Lost Dog,” in Proceeding of 14th National Conference on Computing and Information Technology (NCCIT), King Mongkut’s University of Technology North Bangkok, Chiang Mai, 2018, pp. 678–685.

[10] P. Wairotchanaphutha, N. Boonsirisumpun, W. Puarungroj, “Detection and Classification of Vehicles using Deep Learning Algorithm for Video Surveillance Systems,” in Proceeding of 14th National Conference on Computing and Information Technology (NCCIT), King Mongkut’s University of Technology North Bangkok, Chiang Mai, 2018, pp. 402–407.

[11] A. Rangsuk, T. Kungkajit, “Classification of Amulets using Deep Learning Techniques,” in Proceeding of 10th National Conference on Information Technology (NCIT), Mahasarakhan University, Khon Kaen, 2018, pp. 190–194.

[12] L. Pan, S. Pouyanfar, H. Chen, J. Qin and S. Chen, "DeepFood: Automatic Multi-Class Classification of Food Ingredients Using Deep Learning", 2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC), San Jose, CA, 2017, pp. 181-189.

[13] A. V. Singh, Content-based Image Retrieval using Deep Learning, New York: Rochester Institute of Technology, 2015.

[14] E. Pacharawongsakda, “An Introduction to Data Mining Techniques,” 2nd ed. Bangkok: Data cube, 2014.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 70-74

Monitoring System Platform for Agricultural Research and Analysis

Somluthai Wongprasadpetch[1,2], Patcharin Manowan[1,2], Chanatip Srisot[1,2], Tossapon Boongoen[1,2], Thongchai Yooyativong[2] and Suppakarn Chansareewittaya[2]*

[1]*Center of Excellence in AI and Emerging Technologies*

[2]*School of Information Technology, Mae Fah Luang University, Chiang Rai, Thailand*

Received: May 13, 2020; **Revised:** June 11, 2020; **Accepted:** June 11, 2020; **Published:** June 23, 2020

ABSTRACT – This research article contents the study and development in the topic of monitoring system platform for agricultural research and analysis. This platform is developed to manage and analyze the data of the initial environment. The Chinese kale is used as an example plant in this research. The IoT devices and sensors are used to monitoring and collecting the data. The IoT controllers are raspberry Pi and Arduino Uno. The sensors are temperature sensors, relative humidity sensors, light intensity sensors, soil humidity sensors, and Raspberry Pi camera. The data of the environment are such as temperature, relative humidity, light intensity, soil humidity, and Chinese kale leaf color. The data will be collected and displayed on the thingspeak platform. This platform is still running and on the way to be improved and developed.

KEYWORDS: Monitoring System, Monitoring Platform, Internet of Thing, Smart farm, Agricultural

1. Introduction

The office of Agricultural Economics (Thailand) has revealed the number of farmers in 2017, there are more than 980,000 farmers throughout Thailand and the future population of the agricultural sector may be tended to increase. Although the population of the agricultural sector increases every year. However, agricultural production tends to decrease. Agricultural Economic Report: 2nd Quarter 2019 and Outlook for 2019 shows that the crop production index is at 100.1, down 3.3 percent from the same period in 2018, Since the weather is hot and dry, including the amount of water that is less than last year.

According to Thai farmers, the main problems are Thai farmers still lack the knowledge of technology and analysis of factors affecting the growth of vegetables and also still have to deal with unfavorable weather conditions. These problems are the main causes of unexpected productivity of the agriculture of both farmers and consumers. Some researcher developed the machine and system which can help the farmer [1-3].

Therefore, a lot of Thai farmers need a platform that helps in monitoring the environment to make better agricultural results. Thus, this platform is designed and developed to serve this demand. The internet of things (IoT) devices such as raspberry Pi and Arduino Uno. Also, the sensors are applied for monitoring and collecting temperature values, humidity values, and soil moisture values. These data are useful for farmers to follow up with the environment of plants and can use for data analysis in the future.

2. MATERIALS AND DEVICES

2.1 Sample plant

Chinese kale (Brassica alboglabra L.H. Bailey) is a vegetable that has different characteristics of each species. Three species are popular in Thailand, including broadleaf, pointed leaf and the last one is long petiole (Chiang Mai Royal Agricultural Research Center, 2014) [4]. All three kale species can be grown

[a]*Corresponding Authors : suppakarn.cha@mfu.ac.th*

throughout the year, but all three kale species will grow well from October until April. Although kale is a vegetable that has a growth period until the harvest period is 45 to 55 days. Chinese kale is also weather-resistant as well. The optimum temperature for growing kale yields that the temperature range is 18 to 32 °C (The World Vegetable Center, 2012). Therefore, Chinese kale is a suitable option for this research [5].

2.2 DHT11 and DHT22

DHT11 is the sensor module that measures the temperature and humidity. This sensor includes a resistive-type of humidity measurement and NCT temperature measurement with a small size component. There are 4-pins that are the low voltage at 3 to 5.5 voltage used. The measurement range of DHT11 is 0 to 50-degree Celsius, humidity is in the range 20 to 80 %RH and The measurement range of DHT22 are -40 to 80-degree Celsius, humidity is in range 0 to 100 %RH. DHT22 can also display the value in the form of a decimal point, which has more accuracy than DHT11 [6].

2.3 LDR (Light Dependent Resistor)

LDR or light dependent resistor the module to detect the environmental light and measuring the light intensity. The module is 3-pins connecting that are a low voltage that at 3.3 to 5 voltages used.

2.4 Soil moisture sensor

The soil moisture sensor is used to detect the moisture of the soil. The component has two probes to pass current in the soil that is a low voltage that at 3.3 to 5 voltages used. The depth detection in the soil is about 37 mm.

2.5 Webcam

The OKER webcam model 088 is used in this research. This webcam is high quality at 2 megapixels up to 10 megapixels depended on support software. The frame rate of this webcam is 30fps. Moreover, a high-quality glass lens focus range at 30mm up to-infinitive is included.

2.6 Raspberry Pi 3 Model B+

Raspberry Pi 3 Model B+ is the new version of Raspberry Pi that boasting core processor at a 64-bit system and the processor is running at 1.4GHz. The built in dual-band Wi-Fi content in 2.4GHz and 5GHz. The built in Bluetooth is 4.2/BLE [9].

2.7 Arduino UNO R3

The microcontroller board with ATmega328 that have digital I/O for 14 pins, analog input for 6 pins and PWM Digital I/O for 6 pins. The operating voltage for this board is 5V with the recommended input voltage at 7 – 12V.

3. SYSTEM DIAGRAM

This system diagram is designed based on automated farming by using raspberry pi3 model B+ and Arduino UNO is used as IoT devices and sends the data of sensors to the database via Wi-Fi. The sensors that are used in this research are DHT11, DHT22, LDR sensor module, soil moisture sensor. For DHT11 and DHT22, two of the sensors use to compare the same values to get more detailed and the accuracy of real data. Also, the camera that is an OKER Webcam model 088 is used to capture the Chinese kale leaf.

The data will be stored in the database at thingspeak.com which is an IoT platform. The thinkspeak can use to analyze and perform all data of this platform. The data is set into three parts [7-8]. First is the air part. This part uses DHT11 module, DHT22 module to collect air temperature and air humidity. The LDR sensor module is used to collect light intensity. Second, the OKER Webcam model 088 is used to capture pictures of plant growth and the color of plant leaves. The third is the ground part. The soil moisture sensor is used to collect the value of ground humidity. The overview of the system diagram is shown in fig. 1.

For the analysis of the color value of Chinese kale leaf from the picture, the diagram is shown in fig. 2. The system is set up to perform color analysis every day, the system cropping the image size 54 * 54 pixels in the green area, which is the area of a Chinese kale leaf. After getting the image, the system will start to analyze the color value by converting the RGB color model to the HSV color model and display the H or Hue value which is the degree of color value and take the result to be displayed on the thingspeak platform [10].

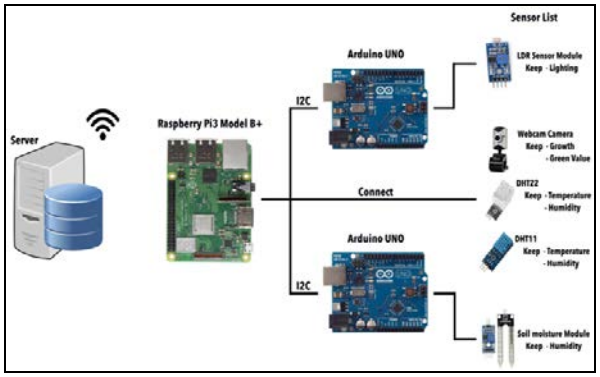


Figure 1. The overview of system diagram.

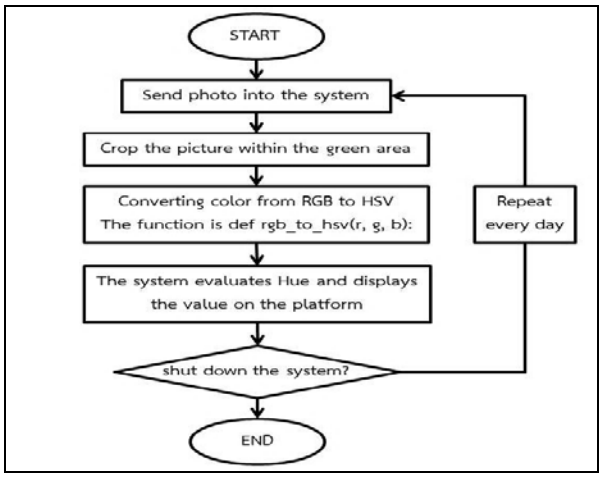


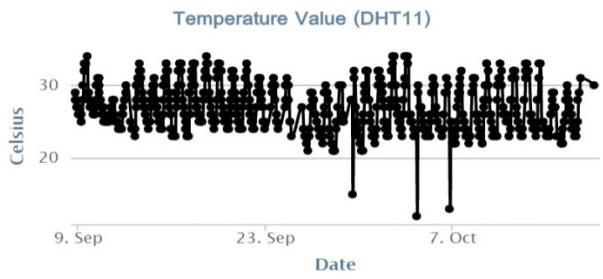
Figure 2. The overview of converting color diagram.

4. RESULTS AND DISCUSSION

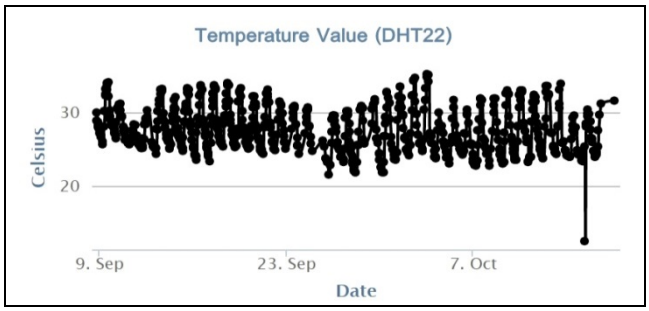
For data collection from the sensors mentioned in Section II, the system is set to collect ambient weather data on the experiment plots every 1 hour and take the values displayed on the Thingspeak.com platform. By setting the system to store the first value as of 8 September 2019 at 8:00 pm until October 17, 2019, at 3.00 pm, which is an example of the data set presented in this research. The system recorded values are as follows.

4.1 Temperature

For the values shown in the graph below, both graphs are the air temperature values in the experimental area. The first graph (a) shows the temperature values from DHT11 and the second graph (b) shows the temperature values from DHT22. From both graphs, there are some very low points on the graphs. The cause is a failure from the equipment and electrical system from the experimental area causing an error.



(a)

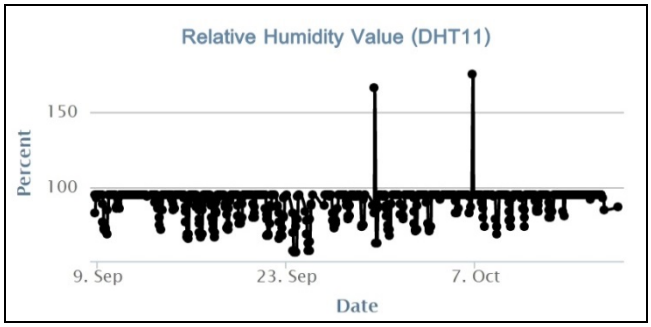


(b)

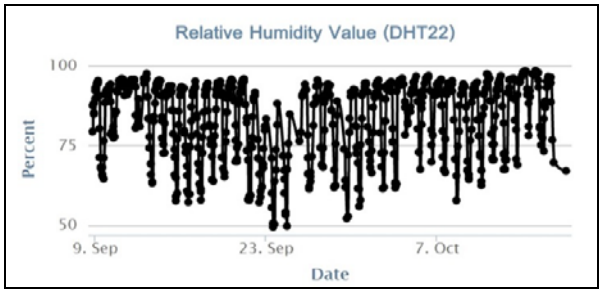
Figure 3. (a) The temperature value from DHT11 and (b) The temperature value from DHT22.

4.2 Relative humidity

For the values shown in the graph below, the two graphs are the relative humidity of the air in the experimental area. The first graph (a) shows the relative humidity from the DHT11 and the second graph (b) shows the relative humidity from the DHT22. The error values that appear on the graph are caused by the same reasons mentioned above.



(a)



(b)

Figure 4. (a) The relative humidity value from DHT11 and (b) The relative humidity value from DHT22.

4.3 Light intensity

For the values shown in the graph below is the light intensity in the experimental area by the graph showing values from the light dependent resistor (LDR). In the case of the graph showing the light intensity, the highest light intensity is about 254 and the lowest is 2. Due to the location of this system, the experimental area is located in the uncontrolled environment that is in the shade but the sunlight can illuminate that area and from collecting every 1 hour, causing the high-low values throughout the period shown in the graph.

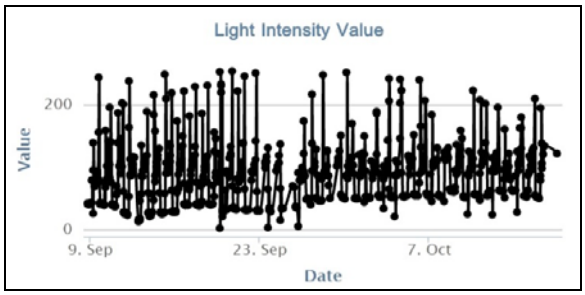


Figure 5. The light intensity value graph.

4.4 Soil humidity

The values shown in the graph below are the soil humidity values in the experimental area. The graph shows the relative humidity from the Soil Moisture Sensor. After the system starts working with an error occur, the sensor cannot be operated. After the initial solution, the error values shown in the graph are a group of data that is collected at the lowest value of the graph and the system resumes normal operation within a few hours afterward.

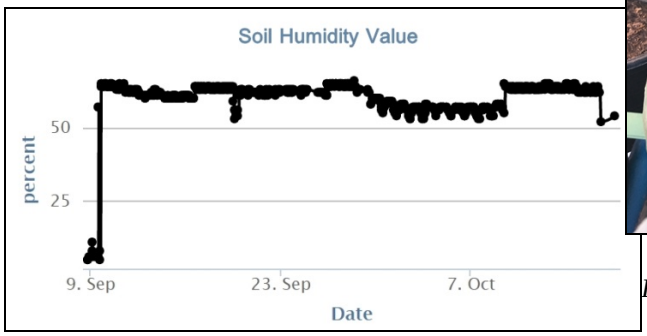


Figure 6. The humidity value graph.

4.5 Chinese kale leaf color

The values shown in the graph below are the color values of Chinese kale leaf to assess whether the Chinese kale is still growing well. The graph below shows the Hue value in the HSV color model. The HSV is highly accurate color theories and the color that is closely related to human eyes. The HSV color model that are 1) H: Hue is the primary color that varies according to the frequency of light, with values ranging from 0 to 360 degrees, such as green at 120 degrees, yellow at 60 degrees, etc. 2) S: Saturation is the saturation of that color with values ranging from 0 to 100 percent. 3) V: Value is the brightness of color which will have values from 0 to 100 percent as well.

According to the HSV color model, the system has been set up to analyze the color of the Chinese kale leaf using the values of Hue as a reference to the HSV color model. The degrees of green are in the range of 70 degrees to 160 degrees. The system will analyze the color from the image in the area of Chinese kale leaf (Sample area) from the RGB color model to the HSV color model and take the color values shown in the graph. However, the value may be higher than the human eye can analyze because the initial analysis from that system will have other factors involved such as the light and the atmosphere around vegetable plots, etc. The preliminary testing of the webcam, referring to the image from below (b) of the Chinese kale leaf is not clear as there is a lot of light reflected in the image resulting in unable to analyze the color precisely. To initially solve this problem, there are temporarily used images from iPhone S6 as the image below (a). The system will analyze the color value from the image and show the received value on the thingspeak.

(a) (b)

Figure 7. The image of Chinese kale leaf from from iPhone S6 and (b) The image of Chinese kale leaf from Oker webcam.

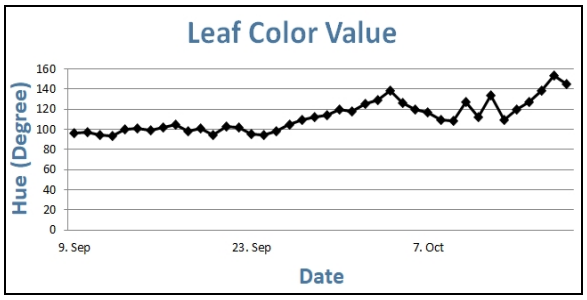


Figure 8. The Chinese kale leaf color value which is Hue in HSV color model.

5. CONCLUSION

For this research article, a set of devices for vegetable plot installation has been developed to facilitate the collection of the environmental values in vegetable plot areas. All values shown on the platform can be checked at any time. Even if the equipment is working at a satisfactory level but there are still device limitations mentioned in Section 3 that need to be improved. Also, this kit is the prototype of the development of basic equipment set for the research team to be able to further study and develop at the machine learning and data analysis. In terms of color value analysis from Chinese kale leaves. Although the value obtained from the analysis has factors of light impact which cause color analysis to be inaccurate, the values that the system has collected and shown in Figure 9. The graph shows that the leaves of Chinese kale are green, which increases with the growth of Chinese kale. (At the red straight line), which is an acceptable color value that the Chinese kale is growing all the time.

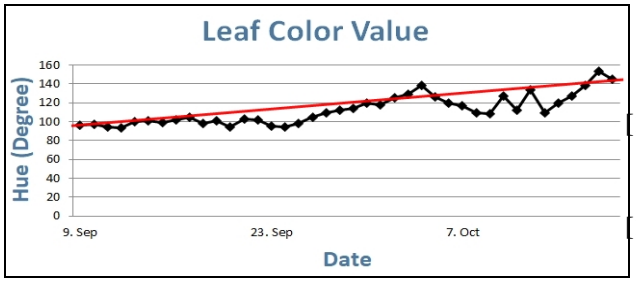


Figure 9. The Chinese kale leaf color value which is Hue in HSV color model with the red straight line that show the growth of Chinese kale.

References

[1] S. Chansareewittaya, P. Paramee, and, S. Iempongsai, "Automatic Hydroponic Harvesting Robot," Thai Society of Agricultural Engineering Journal, vol.25, no.1 , pp.56-63, 2018.

[2] S. Khoukitpaisal, P. Rattanasin, P. Singpant, T. Laohapensaeng, and S. Chansareewittaya, "Smart System for NFT Hydroponics Farm," The 44th Congress on Science and Technology of Thailand (STT45), 2018.

[3] Wongsakorn Chaiwongsa, Widchanin Keardsamer, Teeravisit Laohapensaeng, and Suppakarn Chansareewittaya, "Smart Hydroponic Farm using IoT," TSAE National Conference2018 Thai Society of Agricultural Engineering, 2561.

[4] Chiang Mai Royal Agricultural Research Center, "The Maternal Line Selection of Chainese Kale, Choy-Sum and Pak-Choy for Open-pollinated Varieties," Chiang Mai Royal Agricultural Research Center, 2014. [Online]. Available : http://www.doa.go.th/research/attachment.php?aid=2462. [Accessed: 15 April 2019].

[5] Office of Agricultural Economics, 2019. Agricultural Economic Report: 2nd Quarter 2019 and Outlook for 2019. [Online]. Available: http://www.oae.go.th/assets/portals/1/fileups/bappdata/files/%E0%B8%A0%E0%B8%B2%E0%B8%A7%E0%B8%B0%E0%B8%AF%20%E0%B9%84%E0%B8%95%E0%B8%A3%E0%B8%A1%E0%B8%B2%E0%B8%AA%202_2562%20(v1).pdf. [Accessed: May 25, 2019].

[6] The World Vegetable Center, "Grow Chinese Kale", 2012. [Online]. Available: https://avrdc.org/wpfbfile/grow%20chinese%20kale-pdf/. [Accessed: May 25, 2019].

[7] Adafruit. "DHT11, DHT22 and AM2302 Sensors", [Online]. Available: https://learn.adafruit.com/dht. [Accessed: August 16, 2019.

[8] Sunrom Technologies. 2008. "Light Dependent Resistor", Available: https://www.sunrom.com/get/443700. [Accessed: August 16, 2019].

[9] Kumsawat P, "Environment Reporting System in Agriculture Farm Using Low-cost Android-based Wireless Sensor Network," PhD dissertation, Nakhon Ratchasima: Institute of engineering, Suranaree University of Technology, 2018.

[10] Lumnian T, "Design and application of a smart farm base on IoT," Proceeding of the 7th Conference of Electrical Engineering Network of Rajamangala University of Technology (EENET), Chon Buri, Thailand, May 27-29, 2015, pp. 456-459.

[11] Raspberrypi Org. "Raspberry Pi 3 Model B+," [Online]. Available: https://docs.rs-online.com/2ab2/0900766b8162cdf1.pdf. [Accessed August 16, 2019].

[12] FEC, 2019 "Arduino Uno R3," [Online]. Available:https://www.fecegypt.com/uploads/datasheet/1522237550_arduino%20uno%20r3.pdf. [Accessed: August 16, 2019].

[13] Intel software, 2018. "Color Models," [Online]. Available: https://software.intel.com/en-us/ipp-dev-reference-color-models. [Accessed: April 7, 2019].

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 75-84

การศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิคการคัดเลือกคุณลักษณะบนชุดข้อมูลที่ไม่สมดุล

The Study of Factors Affecting for On-time Graduation of Ungraduated Student Using Feature Selection Technique on Imbalanced Datasets

วรการ ใจดี และ นพณัฐ วรรณภีร์*

Worakarn Jaidee and Noppanat Wannapee

สาขาวิชาระบบสารสนเทศทางธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

Department of Business Information System, Faculty of Business Administration and Liberal Art, Rajamangala University of Technology Lanna, Changpuak, Muang, Chiang Mai 50300

Received: May 13, 2020; **Revised:** June 01, 2020; **Accepted:** June 01, 2020; **Published:** June 23, 2020

ABSTRACT – The objectives of this research were (1) to study the factors affecting the students' graduation according to the plan, and (2) to study the guidelines and improvements in teaching and learning problems of each subject affecting the graduation according to the plan of the student. Main factors or data in this research, the researcher used the grade point in each subject, GPA, the gender of the first-year and second-year students by collecting data from Computer Information System, Faculty of Business Administration and Liberal Art, Rajamangala University of Technology Lanna, Chiang Mai with a total of 358 records. As a result, it revealed that datasets were imbalanced with a higher number of students who graduated as their plan than students who did not graduate as their plan completely. Therefore, the researcher solved the problem using the method of balancing the data set by Synthetic Minority Over-Sampling Technique: SMOTE before entering the process of analyzing factors by using feature selection techniques which can divide as 3 techniques as follows: (1) Chi-Square Feature Selection (2) Information Gain Feature Selection and (3) Correlation Based Feature Selection. From the result, it was found that the most important and topnotch factor in every feature selection was Computer Programming 1 Subject This subject is an elementary and compulsory subject in the Computer Information System and most students received quite low points in this subject. Which may cause this subject to be topnotch in every feature selection techniques and maybe the main factor affecting graduation's plan of students significantly.

KEYWORDS: Imbalanced Datasets, Feature Selection, SMOTE

*Corresponding Author: *worakarn.jaidee@gmail.com*

บทคัดย่อ -- **งานวิจัยนี้ได้จัดทำขึ้นโดยมีวัตถุประสงค์ 2 ประการคือ (1) เพื่อที่ศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษา และ (2) เพื่อศึกษาแนวทางและข้อปรับปรุงในการแก้ไขปัญหาการเรียนการสอนในแต่ละรายวิชาที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษา โดยปัจจัยหรือข้อมูลหลักในการวิจัยนี้ ผู้วิจัยจะใช้ข้อมูลระดับคะแนนในแต่ละวิชา ข้อมูลเกรดเฉลี่ย ข้อมูลเพศของนักศึกษาชั้นปีที่ 1-2 โดยผู้วิจัยได้เก็บรบรวบรวมข้อมูลนักศึกษาจากสาขาวิชาระบบสารสนเทศทางคอมพิวเตอร์ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ จำนวน 358 คน จากการรวบรวมข้อมูลเบื้องต้นพบว่า ชุดข้อมูลมีความไม่สมดุลกัน (Imbalanced Datasets) โดยมีจำนวนกลุ่มนักศึกษาที่สำเร็จการศึกษาตามแผนมากกว่ากลุ่มนักศึกษาที่ไม่สำเร็จการศึกษาตามแผน ดังนั้นผู้วิจัยจึงได้ทำการแก้ปัญหาโดยใช้วิธีการปรับสมดุลให้กับชุดข้อมูลโดยใช้เทคนิคการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (Synthetic Minority Over-sampling Technique : SMOTE) ก่อนที่จะนำเข้าสู่กระบวนการวิเคราะห์ปัจจัยโดยจะใช้เทคนิคการคัดเลือกคุณลักษณะ (Feature Selection) 3 เทคนิคคือ (1) Chi-Square Feature Selection (2) Information Gain Feature Selection และ (3) Correlation Based Feature Selection จากการทดลองพบว่าปัจจัยที่มีความสำคัญมากที่สุดและเป็นอันดับหนึ่งในทุกเทคนิคการคัดเลือกคุณลักษณะคือ วิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งวิชานี้เป็นวิชาพื้นฐานและเป็นวิชาบังคับของสาขาระบบสารสนเทศทางคอมพิวเตอร์และเป็นวิชาที่นักศึกษาส่วนใหญ่ได้ระดับคะแนนในวิชานี้ค่อนข้างน้อย ซึ่งอาจส่งผลให้รายวิชานี้ได้ลำดับความสำคัญมาเป็นอันดับหนึ่งในทุกเทคนิคการคัดเลือกคุณลักษณะและอาจเป็นปัจจัยหลักที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาอย่างมีนัยสำคัญ**

คำสำคัญ: ชุดข้อมูลที่ไม่สมดุล, การคัดเลือกคุณลักษณะ, การสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย

1. บทนำ

โดยทั่วไปแล้วปัญหาการไม่สำเร็จการศึกษาตามแผนของนักศึกษา เป็นปัญหาที่พบมากในกลุ่มนักศึกษาในระดับปริญญาตรี โดยเฉพาะสาขาวิชาทางคอมพิวเตอร์ที่มีอัตราการลาออกค่อนข้างสูง ซึ่งคิดเป็นร้อยละ 55.21% [1] โดยสาเหตุมาจากหลากหลายปัจจัย เช่น ปัญหาด้านครอบครัว ปัญหาด้านการเรียน ปัญหาด้านสุขภาพ เป็นต้น โดยปัญหาที่พบเจอมากที่สุดในปัจจุบันก็คือปัญหาด้านการเรียนซึ่งนักศึกษาได้ระดับคะแนนในแต่ละวิชาน้อยกว่าเกณฑ์ที่กำหนดโดยเฉพาะระดับคะแนนในวิชาพื้นฐานและวิชาบังคับ ซึ่งปัญหานี้อาจทำให้นักศึกษารู้สึกท้อและเบื่อหน่ายกับการเรียนและทำให้นักศึกษาหาไม่ทราบว่าจะดำเนินการอย่างไรหรือดำเนินการวางแผนการเรียนไปในทิศทางที่ดีขึ้นได้อย่างไร จนส่งผลให้นักศึกษาไม่สามารถที่จะสำเร็จการศึกษาตามแผน ถูกรีไทร์หรือลาออกกลางคันได้ ในงานวิจัยนี้ผู้วิจัยมีจุดมุ่งหมายและวัตถุประสงค์ 2 ประการคือ (1) เพื่อศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษา และ (2) เพื่อปรับปรุงรูปแบบการเรียนการสอนในแต่ละรายวิชาที่ส่งผลต่อการไม่สำเร็จการศึกษาของนักศึกษา โดยผู้วิจัยได้เก็บรวบรวมข้อมูลระดับคะแนนของนักศึกษาชั้นปีที่ 1-2 จากสาขาวิชาระบบสารสนเทศทางคอมพิวเตอร์ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ จำนวน 358 คน [2] จากการรวบรวมข้อมูลเบื้องต้นพบว่า ชุดข้อมูลมีความไม่สมดุลกัน (Imbalanced Datasets) โดยมีจำนวนกลุ่มนักศึกษาที่สำเร็จการศึกษาตามแผนมากกว่ากลุ่มนักศึกษาที่ไม่สำเร็จการศึกษาตามแผน

ปัญหาความไม่สมดุลของคลาส เป็นปัญหาที่พบมากในชุดข้อมูลทั่วไปและสามารถเกิดขึ้นได้ในหลากหลายโดเมนข้อมูล [3] เช่น ข้อมูลทางการแพทย์ ข้อมูลด้านการศึกษา ข้อมูลด้านการจัดการความเสี่ยง ข้อมูลด้านการรู้จำใบหน้า เป็นต้น ปัญหาความไม่สมดุลของคลาสโดยทั่วไปแล้วอาจเกิดจากสาเหตุที่ข้อมูลมีการทับซ้อนกันหรือมีความซับซ้อนของข้อมูลสูง (Class Overlapping) ซึ่งทำให้ยากต่อการจำแนก ปัญหาข้อมูลในการเรียนรู้ไม่เพียงพอ (Small Sample Size) และปัญหาที่พบอย่างแพร่หลายในงานวิจัยคือปัญหาคลาสรองมีสัดส่วนน้อยเมื่อเทียบกับคลาสหลัก (Class Distribution) ตัวอย่างเช่น [4] ชุดข้อมูลการสำเร็จการศึกษาหนึ่งมีจำนวนข้อมูลนักศึกษา

ทั้งหมด 100 คน โดยมีข้อมูลของคลาสที่สำเร็จการศึกษาตามแผนจำนวน 85 คน และมีข้อมูลคลาสที่ไม่สำเร็จการศึกษาตามแผนจำนวน 15 คน โดยจะเรียกข้อมูลคลาสที่สำเร็จการศึกษาตามแผนที่ซึ่งมีขนาดใหญ่กว่าว่าคลาสส่วนมาก (Majority class) และเรียกคลาสที่ไม่สำเร็จการศึกษาตามแผนที่มีขนาดเล็กกว่าว่าคลาสส่วนน้อย (Minority class) ซึ่งปัญหาความไม่สมดุลของคลาสเหล่านี้อาจทำให้การนำข้อมูลมาใช้การทำเหมืองข้อมูลด้วยอัลกอริทึมพื้นฐานต่างๆ เช่น การคัดเลือกคุณลักษณะ (Feature Selection) การจำแนกประเภทข้อมูล (Data Classification) ไม่มีประสิทธิภาพหรือมีประสิทธิภาพต่ำ

การคัดเลือกคุณลักษณะ (Feature Selection) เป็นวิธีการในการคัดเลือกคุณลักษณะที่มีความสำคัญและใช้เพื่อลดขนาดหรือมิติของข้อมูลลง การคัดเลือกคุณลักษณะจึงเป็นหัวใจสำคัญในการทำเหมืองข้อมูล (Data Mining) หรือการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งในปัจจุบันมีเทคนิคหรือวิธีการที่หลากหลายที่นักวิจัยได้นำมาใช้ในงานวิจัย ตัวอย่างเช่น [5] เทคนิค Correlation Based Feature Selection เป็นเทคนิคที่ใช้การหากลุ่มของคุณลักษณะที่ได้รับการประเมินค่าจากความสามารถในการคาดการณ์ โดยคุณลักษณะที่ถูกคัดเลือกจะถูกนำไปใช้ในการในขั้นตอนต่อไป เช่น การจำแนกประเภทข้อมูล (Data Classification) เป็นต้น เทคนิคนี้ยังสามารถจัดการกับคุณลักษณะที่ไม่เกี่ยวข้องโดยจะจัดอันดับกลุ่มย่อยของมิติข้อมูลและทำการคัดเลือกกลุ่มย่อยของมิติข้อมูลที่มีความสัมพันธ์กันสูงกับคลาส และไม่มีความสัมพันธ์กับคลาสอื่นๆ เทคนิค chi-square เป็นเทคนิคการวัดโดยการใช้สถิติประมาณความสัมพันธ์ร่วมระหว่างคุณลักษณะเฉพาะกับคลาสของคุณลักษณะเฉพาะ เทคนิค Information Gain เป็นเทคนิคการพิจารณาจากค่าความน่าจะเป็นของแต่ละคุณลักษณะที่เป็นไปได้แล้ววัดค่าความไร้ระเบียบ (Entropy) เพื่อคัดเลือกคุณลักษณะที่มีความสำคัญในการจำแนกกลุ่มได้ดีที่สุด โดยทั่วไปแล้วเทคนิคการคัดเลือกคุณลักษณะด้วย Information Gain มักจะนำมาใช้ร่วมกับอัลกอริทึมหรือตัวแบบต้นไม้ตัดสินใจ (Decision Tree)

ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงเล็งเห็นว่าการใช้เทคนิคในการจัดการชุดข้อมูลที่ไม่สมดุลโดยใช้เทคนิคการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (Synthetic Minority Over-sampling Technique : SMOTE) [6] ร่วมกับวิธีการคัดเลือกคุณลักษณะด้วยเทคนิค Chi-Square, Information Gain และเทคนิค Correlation Based Feature Selection จะทำให้ได้ปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาได้อย่างถูกต้องและมีประสิทธิภาพ

2. วิธีการวิจัย

ในการศึกษาวิจัยครั้งนี้ ผู้วิจัยได้ทำการศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิคการคัดเลือกคุณลักษณะ บนชุดข้อมูลที่ไม่สมดุล โดยผู้วิจัยได้มีการดำเนินการ 5 ขั้นตอนดังต่อไปนี้

2.1 ขั้นตอนการรวบรวมข้อมูล (Data Gathering)

งานวิจัยนี้ผู้วิจัยได้รวบรวมข้อมูลประวัติผลการเรียนของนักศึกษาสาขาวิชาระบบสารสนเทศทางคอมพิวเตอร์ มหาวิทยาลัยราชมงคลล้านนา เชียงใหม่ ตั้งแต่ปี พ.ศ. 2557 – 2559 จำนวน 358 คน โดยใช้ข้อมูลผลการเรียนและระดับคะแนนในแต่ละวิชาใน 4 เทอมแรก (ชั้นปีที่ 1 – 2) สาเหตุที่ใช้ข้อมูลผลการเรียนและระดับคะแนนในแต่ละวิชาใน 4 เทอมแรก เนื่องจากในช่วงก่อนขึ้นชั้นปีที่ 3 นักศึกษาส่วนใหญ่จะมีอัตราการย้ายสาขา การลาออกหรือพ้นสภาพการเป็นนักศึกษาค่อนข้างมาก จึงทำให้ข้อมูลของนักศึกษาบางรายไม่ครบหรือขาดหายไป ซึ่งอาจส่งผลต่อประสิทธิภาพในขั้นตอนของการวิจัยได้ โดยข้อมูลจะประกอบไปด้วย ระดับคะแนนในแต่ละวิชา เกรดเฉลี่ย เพศ โดยมีจำนวนคุณลักษณะทั้งหมด 27 คุณลักษณะ ดังตารางที่ 1

ตารางที่ 1. แสดงรายละเอียดคุณลักษณะข้อมูล

ลำดับ	ชื่อคุณลักษณะ	คำอธิบาย
1	Gender	เพศนักศึกษา
2	เศรษฐศาสตร์จุลภาค	ระดับคะแนน
3	การวิเคราะห์ธุรกิจเชิงปริมาณ	ระดับคะแนน
4	การเงินธุรกิจ	ระดับคะแนน
5	การบัญชีชั้นต้น	ระดับคะแนน
6	องค์การและการจัดการ	ระดับคะแนน
7	การจัดการการผลิตและการปฏิบัติการ	ระดับคะแนน
8	หลักการตลาด	ระดับคะแนน
9	การใช้งานระบบสารสนเทศในธุรกิจ	ระดับคะแนน
10	การเขียนโปรแกรมคอมพิวเตอร์ 1	ระดับคะแนน

ตารางที่ 1. แสดงรายละเอียดคุณลักษณะข้อมูล (ต่อ)

ลำดับ	ชื่อคุณลักษณะ	คำอธิบาย
11	โครงสร้างข้อมูลและอัลกอริทึม	ระดับคะแนน
12	ระบบจัดการฐานข้อมูล	ระดับคะแนน
13	การเขียนโปรแกรมบนเว็บ	ระดับคะแนน
14	การวิเคราะห์และออกแบบระบบ	ระดับคะแนน
15	ระบบสารสนเทศในองค์กร	ระดับคะแนน
16	พลศึกษา	ระดับคะแนน
17	ภาษาอังกฤษเพื่ออาชีพ	ระดับคะแนน
18	ภาษาอังกฤษเพื่อการสื่อสาร	ระดับคะแนน
19	ภาษาอังกฤษผ่านสื่อและเทคโนโลยี	ระดับคะแนน
20	ภาษาอังกฤษในชีวิตประจำวัน	ระดับคะแนน
21	ภาษาไทยเพื่อการสื่อสาร	ระดับคะแนน
22	การพัฒนาคุณภาพชีวิตและสังคม	ระดับคะแนน
23	ปรัชญาเศรษฐกิจพอเพียงเพื่อการพัฒนาที่ยั่งยืน	ระดับคะแนน
24	สถิติพื้นฐาน	ระดับคะแนน
25	การคิดและการตัดสินใจเชิงวิทยาศาสตร์	ระดับคะแนน
26	GPA	เกรดเฉลี่ย
27	Status	0 = สำเร็จการศึกษา 1 = ไม่สำเร็จการศึกษา

2.2 ขั้นตอนการเตรียมข้อมูล (Data Pre-processing)

จากการรวบรวมข้อมูลเบื้องต้น ผู้วิจัยพบว่าข้อมูลมีความไม่สมบูรณ์และมีค่าที่ขาดหายไป (Missing Value) จึงทำให้ไม่สามารถนำข้อมูลเข้าสู่กระบวนการคัดเลือกคุณลักษณะได้ ผู้วิจัยจึงได้ทำการแก้ปัญหาข้อมูลที่ขาดหายไปโดยการเติมค่าที่ขาดหายไปด้วยวิธีการหาเพื่อนบ้านใกล้เคียงที่สุด (k-Nearest Neighbor Imputation) [7] เทคนิคการเติมข้อมูลที่ขาดหายไปด้วยวิธีเพื่อนบ้านใกล้เคียงเป็นเทคนิคที่มีประสิทธิภาพสูงและเป็นการประยุกต์อัลกอริทึมเพื่อนบ้านใกล้เคียงที่สุดมาใช้ในการเติมข้อมูลที่ขาดหายไป โดยข้อมูลจะถูกเติมจากข้อมูลใกล้เคียงของแต่ละเรคคอร์ดนั้นๆ และยังสามารถแทนที่ข้อมูลที่หายไปด้วยค่าที่เป็นไปได้ที่ใกล้เคียงที่สุดกับค่าจริง โดยวิธีการเติมข้อมูลที่ขาดหายไปด้วยวิธีนี้จะมีหลักการทำงานเหมือนกับอัลกอริทึมเพื่อนบ้านใกล้เคียงที่สุดคือ จะใช้วิธีการวัดระยะห่างแบบ Euclidean Distance ซึ่งเกิดจากรากที่สองของผลต่างระหว่างแอตทริบิวต์ต่างๆ ยกกำลังสอง ดังสมการที่ 1

$$D_{Euclidean}(p,q) = D_{Euclidean}(q,p)$$

$$= \sqrt{(p_i - q_i)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

$$= \sqrt{\sum_{n=1}^{k} (p_i - q_i)^2} \quad (1)$$

โดยที่ p คือข้อมูลที่ต้องการจำแนก
q คือข้อมูลในแต่ละกลุ่ม

2.3 ขั้นตอนการเพิ่มความสมดุลของข้อมูลด้วยวิธีการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (SMOTE)

หลักจากขั้นตอนการเตรียมข้อมูลเรียบร้อยแล้ว ผู้วิจัยพบว่าข้อมูลมีความไม่สมดุลกัน โดยมีจำนวนกลุ่มนักศึกษาที่สำเร็จการศึกษาตามแผนมากกว่ากลุ่มนักศึกษาที่ไม่สำเร็จการศึกษาตามแผน ดังนั้นผู้วิจัยจึงแก้ปัญหาโดยการนำวิธีการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย เพื่อช่วยในการปรับสมดุลของข้อมูลเพื่อให้จำนวนของคลาสรอง (Minority Class) มีจำนวนเทียบเท่าหรือใกล้เคียงกับจำนวนข้อมูลที่เป็นคลาสหลัก (Majority Class) ก่อนนำข้อมูลเข้าสู่กระบวนการคัดเลือกคุณลักษณะต่อไป

เทคนิคการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (SMOTE: Synthetic Minority Over-Sampling Technique) มีหลักการทำงานเหมือนกับอัลกอริทึมเพื่อนบ้านใกล้เคียงที่สุด โดยจะทำการสุ่มเลือกแล้วเก็บในตัวแปร x_i โดยทั่วไปแล้วตัวแปร x_i และตัวแปร $x^{'}$ จะถูกเรียกว่า (Seed Sample) จากนั้นจะทำการสุ่มค่าตัวแปร σ ให้มีค่า 0 หรือ 1 เพื่อใช้ในการคำนวณ โดยมีสมการในการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อยดังสมการที่ 2

$$x_{new} = x_i + (x^{'} - x_i) * \sigma \quad (2)$$

ตารางที่ 2. ตารางเปรียบเทียบข้อมูลที่ผ่านกระบวนการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย

กลุ่ม	original	smote
กลุ่มที่สำเร็จการศึกษา	203	203
กลุ่มที่ไม่สำเร็จการศึกษา	155	203
จำนวนทั้งหมด	358	406

2.4 ขั้นตอนการคัดเลือกคุณลักษณะข้อมูล (Feature Selection)

ในขั้นตอนการคัดเลือกคุณลักษณะเป็นขั้นตอนในการสำคัญในงานวิจัยนี้เพื่อที่จะศึกษาว่าปัจจัยใดส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษามากที่สุด โดยในขั้นตอนนี้ผู้วิจัยจะใช้เทคนิค (1) Chi-Square (2) Information Gain และ (3) Correlation Based Feature Selection

(1) Chi-Square (x^2) เป็นเทคนิคการประเมินค่าของคุณลักษณะโดยการใช้การคำนวนค่า chi-square ทางสถิติ เพื่อศึกษาว่าการแจกแจงความถี่ของตัวแปรคุณลักษณะเป็นไปตามรูปแบบที่กำหนดไว้หรือไม่ ดังแสดงในสมการที่ 3

$$x^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (3)$$

โดยที่ O1 , O2 … On เป็นความถี่ของตัวแปรที่ได้จากการศึกษา

E1 , E2 … En เป็นความถี่ที่คาดหวัง (หรือความถี่ที่ควรจะเป็น)

(2) Information Gain เป็นเทคนิคที่ใช้ในการชี้วัดเพื่อเลือกคุณลักษณะข้อมูลที่ต้องการที่น้อยที่สุดที่เหมาะสมในการนำไปใช้ระบุ โดยการคำนวณหาค่า Gain สำหรับแต่ละมิติข้อมูล ซึ่งมิติข้อมูลใดมีค่า Gain สูงสุด จะถูกคัดเลือกเพื่อนำไปใช้ระบุ สมการที่ 4 แสดงการคำนวณค่า Information Gain หรือค่า Entropy ของชุดข้อมูลทั้งหมด สมการที่ 5 แสดงการคำนวณค่า Entropy ของชุดมิติข้อมูลในแต่ละคุณลักษณะ สมการที่ 6 เป็นสมการในการคำนวณค่า Information Gain สำหรับพิจารณามิติข้อมูลคุณลักษณะ A

$$E(D) = -\sum_{i=1}^{n} p_i \log_2(p_i) \quad (4)$$

$$E_A(D) = \sum_{j=1}^{m} \frac{|D_j|}{D} x\ E(D_j) \quad (5)$$

$$Gain(A) = E(D) - E_A(D) \quad (6)$$

โดยที่ p_i คือค่าความน่าจะเป็นที่เรคคอร์ดหนึ่งๆ จะมีหมวดหมู่ของข้อมูล หรือกล่าวว่าการคำนวณค่า Information Gain หรือการวัดค่า Entropy ก่อนที่จะมีการแบ่งข้อมูลออกตามมิติข้อมูลและหลักการแบ่งว่ามีประสิทธิภาพดีขึ้นหรือไม่ ถ้ามีประสิทธิภาพดีขึ้นค่า information gain ก็จะมีค่าสูง

(3) (Correlation Based Feature Selection : CFS) เป็นเทคนิคในการเลือกคุณสมบัติของคุณลักษณะโดยใช้การพิจารณาบนพื้นฐานความสัมพันธ์ของกลุ่มคุณลักษณะที่ได้จากการประเมินค่าความสามารถในการคาดการณ์ และยังสามารถจัดการกับคุณลักษณะที่ไม่เกี่ยวข้อง โดย CFS จะจัดอันดับกลุ่มย่อยของมิติข้อมูลและทำการคัดเลือกกลุ่มย่อยของมิติข้อมูลที่มีความสัมพันธ์กันสูงกับคลาสและไม่มีความสัมพันธ์กับคลาสอื่นๆ สมการประเมินกลุ่มย่อยของมิติข้อมูลแบบ CFS แสดงในสมการที่ 7

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{cf}}}} \quad (7)$$

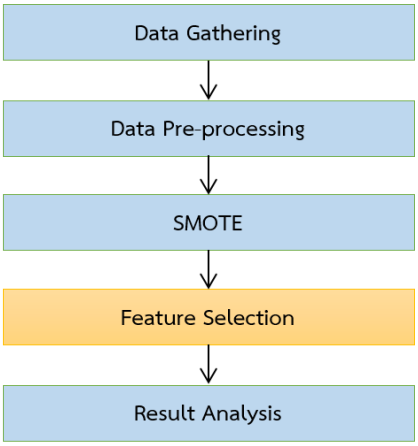
โดยที่ M_s คือค่าที่ค้นหาได้ของมิติข้อมูลกลุ่มย่อย S ซึ่งประกอบด้วยมิติข้อมูล K

$\overline{r_{cf}}$ คือค่าเฉลี่ยความสัมพันธ์ของตัวแปรกับคลาส $(f \in S)$

$\overline{r_{ff}}$ คือค่าเฉลี่ยความสัมพันธ์ระหว่างมิติข้อมูล

2.5 ขั้นตอนการวิเคราะห์และสรุปผล (Result Analysis)

ในขั้นตอนของการวิเคราะห์และสรุปผล ผู้วิจัยจะทำการสรุปผลการใช้เทคนิคการการคัดเลือกคุณลักษณะทั้ง 3 เทคนิค (1) Chi-Square (2) Information Gain และ (3) Correlation Based Feature Selection โดยทำการหาข้อสรุปว่าปัจจัยใดที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษามากที่สุด



รูปที่ 1. แสดงขั้นตอนการทำงานของการวิจัย

2.6 เครื่องมือที่ใช้ในการวิจัย

ในงานวิจัยนี้ผู้วิจัยได้ใช้ซอฟต์แวร์ Jupyter Notebook โดยเขียนด้วยภาษาไพธอน (Python) ในขั้นตอนของกระบวนการทดลองตั้งแต่ขั้นตอนของการเตรียมข้อมูล (Data Pre-processing) จนถึงขั้นตอนของการคัดเลือกคุณลักษณะ (Feature Selection)

3. ผลการทดลอง

ผลการวิจัยการศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาโดยใช้เทคนิคการคัดเลือกคุณลักษณะทั้งหมด 3 เทคนิค คือ (1) Chi-Square (2) Information Gain และ (3) Correlation Based Feature Selection โดยผลการวิจัยของแต่ละเทคนิคหลังจากผ่านกระบวนการในการจัดการกับข้อมูลที่ขาดหายไป (Missing Value) และขั้นตอนการเพิ่มความสมดุลของข้อมูลด้วยวิธีการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (SMOTE) เรียบร้อยแล้ว สามารถสรุปผลการทดลองได้ดังนี้

3.1 การใช้เทคนิค Chi-Square ในการคัดเลือกคุณลักษณะ

จากผลการทดลองโดยใช้เทคนิค Chi-Square ในการคัดเลือกคุณลักษณะ โดยผู้วิจัยได้จัดอันดับคุณลักษณะที่มีค่าความสำคัญที่คำนวณได้ 10 อันดับแรก ดังตารางที่ 3

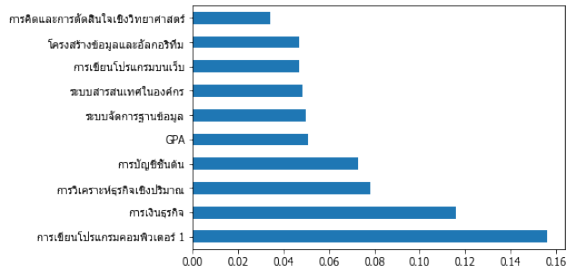
ตารางที่ 3. คุณลักษณะ 10 อันดับแรกที่มีความสำคัญมากที่สุดโดยการใช้เทคนิค Chi-Square

อันดับ	วิชา	ค่าความสำคัญที่คำนวณได้
1	การเขียนโปรแกรมคอมพิวเตอร์ 1	3506.270
2	การเงินธุรกิจ	1892.362
3	การวิเคราะห์ธุรกิจเชิงปริมาณ	1805.420
4	การบัญชีขั้นต้น	1802.137
5	ภาษาอังกฤษผ่านสื่อและเทคโนโลยี	1045.797
6	สถิติพื้นฐาน	963.232
7	ระบบสารสนเทศในองค์กร	799.03
8	ระบบจัดการฐานข้อมูล	708.430
9	ภาษาอังกฤษเพื่อการสื่อสาร	657.455
10	การเขียนโปรแกรมบนเว็บ	550.181

จากตารางที่ 3 คุณลักษณะ 10 อันดับแรกที่มีความสำคัญมากที่สุด โดยการใช้เทคนิค Chi-Square พบว่า วิชาที่เป็นปัจจัยหลักที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาคือ วิชา การเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งเป็นวิชาหลักของสาขาวิชาระบบสารสนเทศทางคอมพิวเตอร์มีค่าความสำคัญที่ 3506.270 อันดับที่สองถึงอันดับที่สี่คือวิชาการเงินธุรกิจมีค่าความสำคัญที่ 1892.362 วิชาการวิเคราะห์ธุรกิจเชิงปริมาณมีค่าความสำคัญที่ 1805.420 และวิชาการบัญชีขั้นต้นมีค่าความสำคัญที่ 1802.137 ซึ่งทั้งสามวิชานี้เป็นวิชาหลักของทางคณะบริหารธุรกิจและศิลปศาสตร์ อันดับที่ห้าวิชาภาษาอังกฤษผ่านสื่อและเทคโนโลยีมีค่าความสำคัญที่ 1045.797 อันดับที่หกวิชาสถิติพื้นฐานมีค่าความสำคัญที่ 963.232 อันดับที่เจ็ดวิชาระบบสารสนเทศในองค์กรมีค่าความสำคัญที่ 799.03 อันดับที่แปดวิชาระบบจัดการฐานข้อมูลมีค่าความสำคัญที่ 708.430 อันดับที่เก้าภาษาอังกฤษเพื่อการสื่อสารมีค่าความสำคัญที่ 657.455 และอันดับที่สิบการเขียนโปรแกรมบนเว็บมีค่าความสำคัญที่ 550.181

3.2 การใช้เทคนิค Information Gain ในการคัดเลือกคุณลักษณะ

จากผลการทดลองโดยใช้เทคนิค Information Gain ในการคัดเลือกคุณลักษณะ โดยผู้วิจัยได้จัดอันดับคุณลักษณะที่มีความสำคัญ 10 อันดับแรก ดังรูปที่ 2

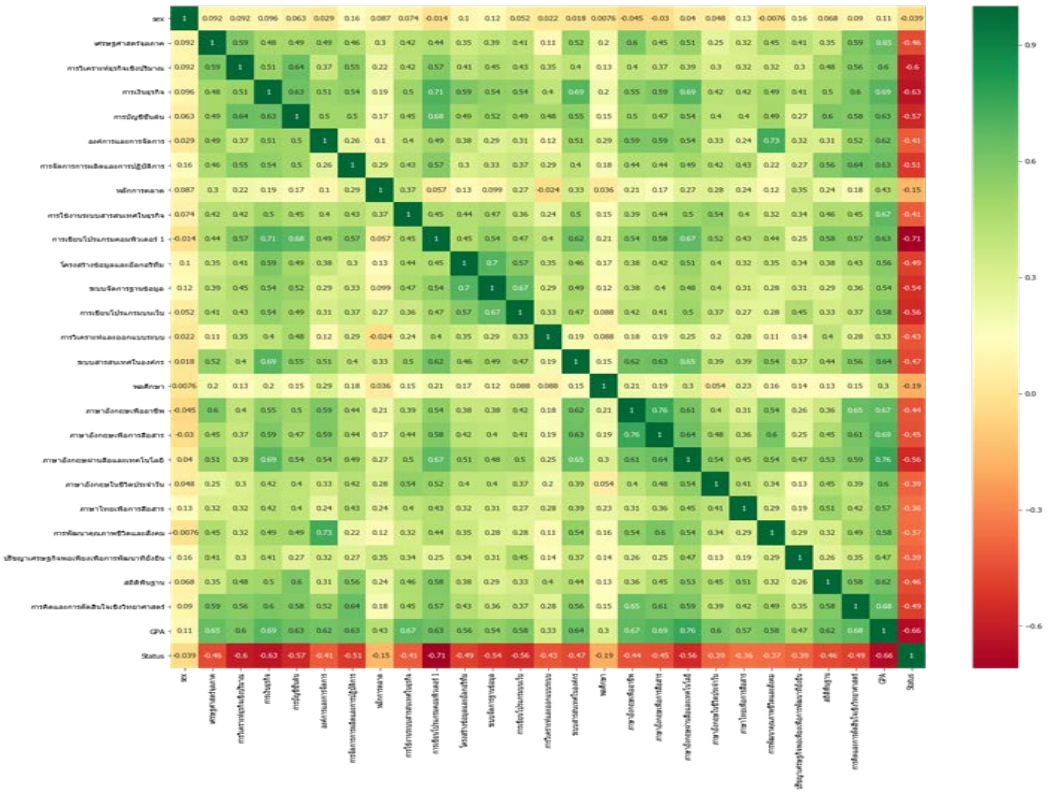


รูปที่ 2. แสดงคุณลักษณะ 10 อันดับแรกที่มีความสำคัญมากที่สุด โดยการใช้เทคนิค Information Gain

จากภาพที่ 2 คุณลักษณะ 10 อันดับแรกที่มีความสำคัญมากที่สุด โดยการใช้เทคนิค Information Gain พบว่าวิชาที่เป็นปัจจัยหลักที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษา คือวิชา การเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งเป็นวิชาหลักของสาขาวิชาระบบสารสนเทศทางคอมพิวเตอร์ และอันดับสองถึงอันดับสี่ คือวิชาวิชาการเงินธุรกิจ วิชาการวิเคราะห์ธุรกิจเชิงปริมาณ และการบัญชีขั้นต้น ซึ่งทั้งสามวิชานี้เป็นวิชาหลักของทางคณะบริหารธุรกิจและศิลปศาสตร์ และในอันดับที่ห้า คือระดับคะแนนเฉลี่ย (GPA) และในอันดับที่หกถึงอันดับที่เก้า ได้แก่วิชาระบบจัดการฐานข้อมูล วิชาระบบสารสนเทศในองค์กร วิชาการเขียนโปรแกรมบนเว็บ วิชาโครงสร้างข้อมูลและอัลกอริทึม ซึ่งทั้งสี่วิชานี้เป็นวิชาหลักและวิชาเลือกของหลักสาขาระบบสารสนเทศทางคอมพิวเตอร์ ส่วนในอันดับที่สิบคือวิชาการคิดและการตัดสินใจเชิงวิทศาสตร์ที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาน้อยที่สุดใน 10 อันดับแรก

3.3 การใช้เทคนิค Correlation Based Feature Selection ในการคัดเลือกคุณลักษณะ

ในการคัดเลือกคุณลักษณะด้วยเทคนิค Correlation Based Feature Selection ผู้วิจัยได้ใช้วิธีการแสดงผลข้อมูลของผลการทดลองในรูปแบบของ Correlation Matrix ร่วมกับ Heatmap ซึ่งได้ผลการทดลองดังรูปที่ 3



รูปที่ 3. ผลการทดลองโดยการใช้เทคนิค Correlation Matrix ร่วมกับ Heatmap

จากรูปที่ 3 การคัดเลือกคุณลักษณะโดยการวิเคราะห์สหสัมพันธ์ (Correlation Based Feature Selection) จะเห็นว่าวิชาที่มีความสัมพันธ์กับสถานการณ์สำเร็จการศึกษาตามแผนของนักศึกษา (Status) มากที่สุดคือวิชา การเขียนโปรแกรมคอมพิวเตอร์ 1 มีค่าสหสัมพันธ์ -0.71 ส่วนอันดับที่สองคือระดับคะแนนเฉลี่ย (GPA) ของนักศึกษาในสองปีแรกมีค่าสหสัมพันธ์ -0.66 และในอันดับที่สามถึงอันดับที่ห้าคือวิชาวิชาการเงินธุรกิจที่มีค่าสหสัมพันธ์ -0.63 วิชาการวิเคราะห์ธุรกิจเชิงปริมาณมีค่าสหสัมพันธ์ -0.6 และวิชาการบัญชีขั้นต้นมีค่าสหสัมพันธ์ -0.57 ตามลำดับ ซึ่งทั้งสามวิชานี้เป็นวิชาหลักของทางคณะบริหารธุรกิจและศิลปศาสตร์

อันดับที่หกและเจ็ดมีค่าสหสัมพันธ์เท่ากันที่ -0.56 ได้แก่ วิชาการเขียนโปรแกรมบนเว็บและวิชาภาษาอังกฤษผ่านสื่อเทคโนโลยี อันดับที่แปดวิชาระบบจัดการฐานข้อมูลมีค่าสหสัมพันธ์ -0.54 อันดับที่เก้าและอันดับที่สิบมีค่าสหสันธ์เท่ากันที่ -0.49 ได้แก่วิชาโครงสร้างข้อมูลและอัลกอริทึมและวิชาการคิดและการตัดสินใจเชิงวิทยาศาสตร์

ผลการวิจัยพบว่าปัจจัยที่มีความสำคัญต่อการส่งผลการต่อการสำเร็จการศึกษาตามแผนของนักศึกษามากที่สุดคือ ระดับคะแนนในรายวิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งจากผลการทดลองด้วยการใช้เทคนิคการคัดเลือกคุณลักษณะทั้ง 3 เทคนิคพบว่ารายวิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 มีความสำคัญมาเป็นอันดับหนึ่งในทุกเทคนิคของการคัดเลือกคุณลักษณะ ซึ่งอาจส่งผลให้รายวิชานี้ได้ลำดับความสำคัญมาเป็นอันดับหนึ่งในทุกเทคนิคการคัดเลือกคุณลักษณะและอาจเป็นปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนอย่างมีนัยสำคัญ ด้วยเหตุนี้ผู้วิจัยได้สรุปสถิติ รวมถึงความถี่ของแต่ละปัจจัยใน 10 อันดับแรกจากการคัดเลือกคุณลักษณะทั้งสามเทคนิค ดังตารางที่ 4

ตารางที่ 4. แสดงสรุปสถิติ ความถี่ของแต่ละปัจจัยใน 10 อันดับแรกจากการคัดเลือกคุณลักษณะทั้งสามเทคนิค

ลำดับ	วิชา	ความถี่
1	การเขียนโปรแกรมคอมพิวเตอร์ 1	**3**
2	การเงินธุรกิจ	3
3	การวิเคราะห์ธุรกิจเชิงปริมาณ	3
4	การบัญชีขั้นต้น	3
5	ภาษาอังกฤษผ่านสื่อและเทคโนโลยี	2
6	สถิติพื้นฐาน	1
7	ระบบสารสนเทศในองค์กร	2
8	ระบบจัดการฐานข้อมูล	3
9	ภาษาอังกฤษเพื่อการสื่อสาร	1
10	การเขียนโปรแกรมบนเว็บ	3
11	การคิดและการตัดสินใจเชิงวิทยาศาสตร์	2
12	GPA	2
13	โครงสร้างข้อมูลและอัลกอริทึม	2

จากตารางที่ 4 สามารถสรุปได้ว่าปัจจัยที่ติดอันดับหนึ่งถึงสิบครบทุกเทคนิคการคัดเลือกคุณลักษณะได้แก่ วิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 วิชาการเงินธุรกิจ วิชาการวิเคราะห์ธุรกิจเชิงปริมาณ วิชาการบัญชีขั้นต้น วิชาระบบจัดการฐานข้อมูล และวิชาการเขียนโปรแกรมบนเว็บ ซึ่งวิชาที่ติดอันดับหนึ่งถึงสิบครบทุกเทคนิคการคัดเลือกคุณลักษณะล้วนแต่เป็นวิชาพื้นฐานและเป็นวิชาบังคับของคณะบริหารธุรกิจและศิลปศาสตร์และสาขาระบบสารสนเทศทางคอมพิวเตอร์ และวิชาที่ติดอันดับทั้งหมดนี้ นักศึกษาส่วนใหญ่จะได้ระดับคะแนนในวิชานี้ค่อนข้างน้อย ซึ่งอาจส่งผลให้แต่ละรายวิชาได้ลำดับความสำคัญในอันดับหนึ่งถึงสิบครบทุกเทคนิคอย่างมีนัยสำคัญ ส่วนปัจจัยที่ติดอันดับหนึ่งถึงสิบเพียงสองเทคนิคการคัดเลือกคุณลักษณะได้แก่ วิชาภาษาอังกฤษผ่านสื่อและเทคโนโลยี วิชาระบบสารสนเทศในองค์กร วิชาการคิดและการตัดสินใจเชิงวิทยาศาสตร์ วิชาโครงสร้างข้อมูลและอัลกอริทึม และระดับคะแนนเฉลี่ย (GPA) และปัจจัยที่ติดอันดับหนึ่งถึงสิบเพียงหนึ่งเทคนิคการคัดเลือกคุณลักษณะได้แก่ วิชาภาษาอังกฤษเพื่อการสื่อสารและวิชาสถิติพื้นฐาน

จะเห็นได้ว่าทั้งสามเทคนิคที่ใช้ให้ผลลัพธ์ที่ใกล้เคียงกันและมีจำนวนวิชาที่ติดมาเป็นอันดับหนึ่งถึงสิบมากถึงหกวิชาที่เหมือนกัน ซึ่งรายวิชาเหล่านี้ล้วนแล้วแต่เป็นวิชาหลักของทางคณะฯ และทางสาขาวิชาฯ ซึ่งอาจเป็นปัจจัยหลักที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาอย่างมีนัยสำคัญ

4. สรุปและการอภิปราย

การวิจัยนี้มีวัตถุประสงค์เพื่อที่ศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผน และเพื่อปรับปรุงรูปแบบการเรียนการสอนในแต่ละรายวิชาที่ส่งผลต่อการไม่สำเร็จการศึกษาของนักศึกษา จากผลการวิจัยนี้สามารถสรุปได้ว่าจากการใช้เทคนิคการคัดเลือกคุณลักษณะทั้งสามเทคนิค (1) chi-square feature selection (2) information gain feature selection และ (3) correlation based feature selection ปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษามากที่สุดได้แก่ วิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งมีความสำคัญมาเป็นอันดับหนึ่งในทุกเทคนิคของการคัดเลือกคุณลักษณะ และในส่วนของปัจจัยอื่นๆ ที่ติดอันดับหนึ่งถึงสิบในทุกเทคนิคของการคัดเลือกคุณลักษณะได้แก่ วิชาการเงินธุรกิจ วิชาการวิเคราะห์ธุรกิจเชิงปริมาณ วิชาการบัญชีขั้นต้น วิชาระบบจัดการฐานข้อมูล และวิชาการเขียนโปรแกรมบนเว็บ ซึ่งวิชาที่ติดอันดับหนึ่งถึงสิบครบทุกเทคนิคของการคัดเลือกคุณลักษณะล้วนแต่เป็นวิชาพื้นฐานและเป็นวิชาบังคับของคณะบริหารธุรกิจและศิลปศาสตร์และสาขาระบบสารสนเทศทางคอมพิวเตอร์ และนักศึกษาส่วนใหญ่ได้ระดับคะแนนในแต่ละวิชาค่อนข้างน้อย ซึ่งอาจส่งผลให้รายวิชาเหล่านี้ติดอันดับหนึ่งถึงสิบในทุกเทคนิคของการคัดเลือกคุณลักษณะและอาจส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาอย่างมีนัยสำคัญ

ปัจจัยที่ติดอันดับหนึ่งถึงสิบเพียงสองเทคนิคของการคัดเลือกคุณลักษณะได้แก่ วิชาภาษาอังกฤษผ่านสื่อและเทคโนโลยี วิชาระบบสารสนเทศในองค์กร วิชาการคิดและการตัดสินใจเชิงวิทยาศาสตร์ วิชาโครงสร้างข้อมูลและอัลกอริทึม และระดับคะแนนเฉลี่ย (GPA) โดยวิชาระบบสารสนเทศในองค์กรและวิชาโครงสร้างข้อมูลและอัลกอริทึมเป็นวิชาพื้นฐานและเป็นวิชาบังคับของทางสาขาฯ และนักศึกษาส่วนใหญ่ได้ระดับคะแนนในแต่ละวิชาค่อนข้างน้อยเช่นกัน ซึ่งอาจส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาอย่างมีนัยสำคัญ และปัจจัยที่ติดอันดับหนึ่งถึงสิบเพียงหนึ่งเทคนิคการคัดเลือกคุณลักษณะได้แก่ วิชาภาษาอังกฤษเพื่อการสื่อสารและวิชาสถิติพื้นฐาน

5. ข้อเสนอแนะ

เทคนิคการคัดเลือกคุณลักษณะทั้งสามเทคนิคที่ผู้วิจัยได้นำเสนอนั้นเป็นเพียงบางเทคนิคของการคัดเลือกคุณลักษณะจากหลากหลายเทคนิคที่มีอยู่ในปัจจุบัน และยังมีข้อจำกัดในเรื่องของปัจจัยข้อมูลที่มีเพียงข้อมูลระดับคะแนนรายวิชา เกรดเฉลี่ย เพศของนักศึกษาเท่านั้น ซึ่งอาจจะมีปัจจัยในด้านอื่นๆ ที่อาจส่งผลต่อการสำเร็จการศึกษาของนักศึกษาได้ เช่น ข้อมูลด้านสถานภาพทางครอบครัว สถานการณ์กู้ยืมเงิน ข้อมูลการศึกษาในอดีต เป็นต้น ในอนาคตหากมีการพัฒนางานวิจัยให้มีประสิทธิภาพมากขึ้น อาจจะต้องมีการเก็บรวมรวมข้อมูลเพื่อให้ได้ปัจจัยที่เพิ่มมากขึ้นและใช้เทคนิคในการคัดเลือกคุณลักษณะในหลายๆ เทคนิคมากขึ้นเพื่อสามารถนำผลของการทดลองไปใช้ได้อย่างมีประสิทธิภาพสูงสุด และในงานวิจัยนี้ผู้วิจัยยังไม่ได้มีกำดำเนินการในการนำคุณลักษณะที่คัดเลือกได้ไปหาค่าความแม่นยำในความถูกต้องของแต่ละเทคนิคการคัดเลือกคุณลักษณะที่ใช้ ซึ่งผู้วิจัยจะมีการดำเนินการในหัวข้อดังกล่าวในงานวิจัยถัดไป

เอกสารอ้างอิง

[1] วรการ ใจดี, “ตัวแบบการทำนายผลการสำเร็จการศึกษาตามแผน โดยใช้เทคนิคเอสเอ็มโอทีอีและตัวแบบการเรียนรู้แบบรวมกลุ่ม บนชุดข้อมูลที่ไม่สมดุล สำหรับข้อมูลการสำเร็จการศึกษา,” รายงานสืบเนื่องจากการประชุมสัมมนาทางวิชาการ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาตะวันออก ระดับชาติครั้งที่ 12, 2562, หน้า 325-335.

[2] พิชัย ระเวงวัน, “โมเดลเพื่อการพยากรณ์สถานภาพทางการศึกษาของนักศึกษา,” รายงานสืบเนื่องจากการประชุมวิชาการเสนอผลงานวิจัยบัณฑิตศึกษา ระดับชาติและนานาชาติ มหาวิทยาลัยขอนแก่น, 2560, หน้า 273-283.

[3] เบญจภรณ์ จันทรกองกุล, “วิธีการที่เหมาะสมสำหรับการแบ่งกลุ่มข้อมูลที่ไม่สมดุลสูง,” วิทยาศาสตร์ดุษฎีบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์. มหาวิทยาลัยบูรพา, 2557.

[4] กิตติพงศ์ ชมบุญ, “เทคนิคการจำแนกประเภทข้อมูลส่วนน้อยบนข้อมูลไม่สมดุลด้วยวิธีการแบ่งข้อมูล,” วิศวกรรมศาสตร์ดุษฎีบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, 2558.

[5] อัจจิมา มณฑาพันธุ์, “การเปรียบเทียบวิธีการคัดเลือกคุณลักษณะที่สำคัญในการปรับปรุงการพยากรณ์มะเร็งเต้านม,” Royal Thai Air Force Medical Gazette, 2562, หน้า 49-56.

[6] Zheng Z, Cai Y, Li Y, “Oversampling Method for Imbalanced Classification,” Computing and Informatics, Vol. 34, 2015, pp. 1017-1037.

[7] Beretta L, Santaniello A, “ Nearest Neighbor Imputation Algorithm: A Critical Evaluation,” The 5th Translational Bioinformatics Conference, 2015, pp. 197-208.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 85-97

การประเมินการใช้งานแอปพลิเคชัน 7-Eleven TH บนพื้นฐานของทฤษฎีการยอมรับและการใช้เทคโนโลยี

An Evaluation of 7-Eleven TH Application Based on Technology Adoption and Acceptance Theory

*สุอัมพร ปานทรัพย์ และ ดัชกรณ์ ตันเจริญ**

*Su-amporn Parnsup and Datchakorn Tancharoen**

คณะวิศวกรรมศาสตร์และเทคโนโลยี สถาบันการจัดการปัญญาภิวัฒน์

Faculty of Engineering and Technology, Panyapiwat Institute of Management

Received: May 30, 2020; **Revised:** June 15, 2020; **Accepted:** June 16, 2020; **Published:** June 23, 2020

ABSTRACT – This paper presents the evaluation of 7-Eleven TH application based on technology adoption and acceptance theory (UTAUT). In order to analyze the relationship model with 413 users in Bangkok and metropolitan region using research tools such as the 7-level rating scale online questionnaire, technical analysis and structural equation model. The research model consists of 9 components including the expectation for application performance, the expectation for application efforts, the social influence on application usage, the condition of the convenience for application usage, the entertainment of application, the value of application usage, the familiarity with application usage, the intentional behavior, and the behavior of application. By analyzing the causal relationship model analysis, it is found that the model is consistent with the empirical data. There is a direct and indirect influence of the path coefficient with a statistical significance of 0.001. It is found that the factors that affect the behavior of the intention to use the 7-Eleven TH application consist of the familiarity, entertainment, operating conditions, value, price, performance expectations, social influence and the expectation of the efforts, respectively. The results can be used to formulate policies, adjust strategies and promote access to consumer groups including government organizations, private sectors, retail and wholesale business. It is benefit in the application development for the organization.

KEYWORDS: Technology Adoption, Acceptance Theory, UTAUT, 7-Eleven Application

บทคัดย่อ -- บทความนี้นำเสนอการประเมินการใช้งานแอปพลิเคชัน 7-Eleven TH บนพื้นฐานของทฤษฎีการยอมรับและการใช้เทคโนโลยี เพื่อวิเคราะห์รูปแบบความสัมพันธ์กับผู้ใช้งานในเขตกรุงเทพมหานครและปริมณฑล จำนวน 413 คน โดยใช้เครื่องมือการวิจัย ได้แก่ แบบสอบถามออนไลน์แบบมาตรประมาณค่า 7 ระดับ เทคนิคการวิเคราะห์องค์ประกอบและโมเดลสมการโครงสร้างโดยใช้โปรแกรมสำเร็จรูป ซึ่งมีโมเดลการวิจัยประกอบด้วย 9 องค์ประกอบ ได้แก่ ความคาดหวังในประสิทธิภาพการใช้แอปพลิเคชัน ความคาดหวังในความพยายามใช้แอปพลิเคชัน อิทธิพลของสังคมต่อการใช้งานแอปพลิเคชัน สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน แรงจูงใจด้านความบันเทิงของแอปพลิเคชัน มูลค่าราคาในการใช้แอปพลิเคชัน ความเคยชินในการใช้แอปพลิเคชัน พฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน และพฤติกรรม

**Corresponding Author: datchakorntan@pim.ac.th*

การใช้งานแอปพลิเคชัน โดยผลการวิเคราะห์รูปแบบความสัมพันธ์เชิงสาเหตุ พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์เป็นอย่างดี มีค่าอิทธิพลทางตรงและทางอ้อมของสัมประสิทธิ์เส้นทางอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ทั้งนี้ พบว่า ปัจจัยที่ส่งผลต่อพฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH ประกอบไปด้วยความเคยชินในการใช้งาน รองลงมา คือ แรงจูงใจด้านความบันเทิง สภาพสิ่งอำนวยความสะดวกในการใช้งาน มูลค่าราคา ความคาดหวังในประสิทธิภาพ อิทธิพลทางสังคม และความคาดหวังในความพยายาม ตามลำดับ โดยสามารถนำเอาผลการศึกษานี้ไปใช้ในการวางนโยบาย ปรับกลยุทธ์ และส่งเสริมการเข้าถึงกลุ่มผู้บริโภคในองค์กรภาครัฐ ภาคเอกชน ทั้งกลุ่มธุรกิจด้านการค้าปลีกและค้าส่ง ให้เกิดประสิทธิภาพได้มากที่สุดในด้านการพัฒนาระบบแอปพลิเคชันสำหรับองค์กร

คำสำคัญ: การใช้เทคโนโลยี, ทฤษฎีการยอมรับ, UTAUT, แอปพลิเคชัน 7-Eleven

1. บทนำ

ปัจจุบันโลกมีการเปลี่ยนแปลงไปอย่างรวดเร็ว เทคโนโลยีได้เข้ามามีบทบาทสำคัญต่อการเปลี่ยนแปลงสังคม เทคโนโลยีใหม่ถูกนำเข้ามาใช้ และมีบทบาทมากขึ้นในหลาย แง่มุม เช่น การติดต่อสื่อสารเพื่อสนับสนุนการทำงาน ตลอดจนช่วยอำนวยความสะดวกในการใช้ชีวิตประจำวัน ในแวดวงธุรกิจก็ถูกขับเคลื่อนเศรษฐกิจด้วยนวัตกรรม เกิดการแข่งขันกันมากยิ่งขึ้น ประเทศไทยถูกจัดอยู่ในอันดับต้น ๆ ของประเทศที่มีการใช้โทรศัพท์มือถือมากที่สุดในโลก จะเห็นได้ว่าประชากรไทยในยุคปัจจุบันมีอัตราการใช้งานเครื่องคอมพิวเตอร์ลดน้อยลง แต่ในขณะเดียวกันการใช้โทรศัพท์มือถือกลับมีอัตราการใช้งานเพิ่มมากขึ้นอย่างต่อเนื่อง เช่นเดียวกับอัตราการใช้งานอินเทอร์เน็ต โดยอัตราการใช้งานอินเทอร์เน็ตของคนไทยเพิ่มสูงขึ้นเรื่อย ๆ

ดังนั้นเมื่อพฤติกรรมการดำเนินชีวิตของประชากรในสังคมมีการเปลี่ยนแปลงไป รูปแบบการดำเนินธุรกิจในสังคมไทยจึงเปลี่ยนแปลงตามไปด้วย การดำเนินธุรกิจจำเป็นที่จะต้องตามกระแส รวมถึงพฤติกรรมและปัจจัยต่าง ๆ ของผู้บริโภคที่เปลี่ยนแปลงไปตามยุคสมัย มีธุรกิจใหม่เกิดขึ้นมากมาย องค์กรธุรกิจเดิมมีการปรับเปลี่ยนกลยุทธ์เพื่อความอยู่รอดและการเติบโต กิจกรรมด้านการขายและการตลาดถือเป็นหัวใจสำคัญที่ส่งผลต่อการเติบโตของธุรกิจ จึงมีการเปลี่ยนแปลงอย่างเห็นได้ชัด โดยการนำเอาเทคโนโลยี เข้ามาช่วยในลักษณะการทำการตลาดดิจิทัล (Digital Marketing) สื่อดิจิทัล และเทคโนโลยี เพื่อสนับสนุนการตลาดสมัยใหม่ เช่น การทำโฆษณาผ่านสื่อสังคมออนไลน์ การใช้เว็บไซต์เพื่อเผยแพร่ข้อมูลและสร้างความเชื่อมั่นให้กับองค์กร หรือการใช้แอปพลิเคชันเพื่อเพิ่มและส่งเสริมกิจกรรมทางการตลาด เพื่อให้เข้าถึงและเจาะกลุ่มผู้บริโภคหรือลูกค้ามากยิ่งขึ้น การทำตลาดออนไลน์จึงเป็นสิ่งที่กำลังได้รับความนิยมเพิ่มมากขึ้น โดยวิธีการทำการตลาดออนไลน์นั้นมีหลากหลายช่องทาง ขึ้นอยู่กับแต่ละองค์กร แต่ละกลุ่มธุรกิจ และกลุ่มลูกค้าเป้าหมาย เพื่อให้สอดคล้องกับแนวโน้มการใช้งานโทรศัพท์มือถือที่เพิ่มมากขึ้นอย่างต่อเนื่อง จึงเกิดการพัฒนาไปสู่ Mobile Marketing

การตลาดออนไลน์มีการแข่งขันกันเป็นอย่างมาก เครื่องมือด้านเทคโนโลยีสารสนเทศต่าง ๆ ถูกนำมาใช้เพื่อช่วยกระตุ้นการตลาดให้เกิดความน่าสนใจ เทคโนโลยีที่น่าจับตามองในยุคนี้อย่างหนึ่ง คือ Mobile Applications ซึ่งถูกสร้างขึ้นมาเพื่อช่วยอำนวยความสะดวกในด้านต่าง ๆ จึงเป็นหนึ่งในตลาดดิจิทัลที่ผู้พัฒนาบริการและองค์กรธุรกิจให้ความสนใจ เพราะนอกจากจะเป็นการใช้เพื่อสนับสนุนภาพลักษณ์ขององค์กร และเพิ่มความสะดวกรวดเร็วในการให้บริการขององค์กรแล้ว การมีแอปพลิเคชันเป็นหนึ่งในองค์ประกอบของธุรกิจ ถือว่าเป็นองค์กรที่พร้อมจะให้บริการต่อกลุ่มผู้บริโภคและสามารถเข้าถึงกลุ่มผู้บริโภคได้มากขึ้น อันเป็นการสร้างความประทับใจให้กับลูกค้า ดังนั้นกลุ่มธุรกิจต่าง ๆ ในประเทศไทยจึงหันมาให้ความสำคัญกับแอปพลิเคชันบนสมาร์ทโฟนมากยิ่งขึ้น โดยผู้ใช้งานก็ให้การตอบรับต่อการตลาดในรูปแบบนี้อย่างมากเช่นเดียวกัน

จากการจัดอันดับ 20 แอปพลิเคชันยอดนิยมของคนไทยในรอบ 10 ปีของแอปสโตร์จะเห็นได้ว่าแอปพลิเคชัน 7-Eleven TH เป็นแอปพลิเคชันร้านค้าปลีกเดียวที่ติดอันดับแอปพลิเคชันที่มีการดาวน์โหลดมากที่สุด ผู้วิจัยจึงมีความสนใจที่จะศึกษาการยอมรับและการใช้งานแอปพลิเคชัน 7-Eleven TH ซึ่งจะทำให้ทราบถึงปัจจัยที่ส่งผลต่อการยอมรับและการใช้เทคโนโลยีที่มี

ผลต่อพฤติกรรมของผู้บริโภค การวิจัยนี้จะเป็นประโยชน์ต่อองค์กรภาครัฐและเอกชน ตลอดจนกลุ่มธุรกิจด้านการค้าปลีกและค้าส่ง ในการวางนโยบาย ปรับกลยุทธ์ และส่งเสริมการเข้าถึงกลุ่มผู้บริโภคให้เกิดประสิทธิผลและมีประสิทธิภาพมากที่สุด

2. แนวคิด และทฤษฎีที่เกี่ยวข้อง

การวิจัยครั้งนี้ได้กำหนดสมมติฐานของการวิจัยและ ปัจจัยที่มีอิทธิพลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ประกอบด้วย 7 องค์ประกอบ ดังนี้

1. ความคาดหวังในประสิทธิภาพของแอปพลิเคชัน (Performance expectancy)
2. ความคาดหวังในความพยายามใช้แอปพลิเคชัน (Effort expectancy)
3. อิทธิพลของสังคมต่อการใช้แอปพลิเคชัน (Social influence)
4. สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน (Facilitating conditions)
5. แรงจูงใจด้านความบันเทิงแอปพลิเคชัน (Hedonic motivation)
6. มูลค่าราคาในการใช้แอปพลิเคชัน (Price value)
7. ความเคยชินในการใช้แอปพลิเคชัน (Habit)

รวมทั้ง 3 ตัวแปรเสริม/ตัวผันแปรที่ส่งผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ได้แก่ เพศ อายุ และประสบการณ์

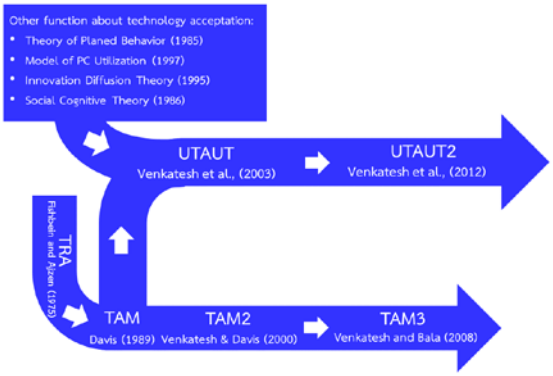
2.1 ข้อมูลเกี่ยวกับแอปพลิเคชัน 7-Eleven TH

แอปพลิเคชัน 7-Eleven TH คือ โปรแกรมประยุกต์ที่ถูกพัฒนาขึ้นมาสำหรับสนับสนุนธุรกิจ บริษัท ซีพี ออลล์ จำกัด (มหาชน) เพื่อช่วยอำนวยความสะดวกต่อผู้บริโภคในการใช้จ่ายในร้านเซเว่น-อีเลฟเว่น รวมถึงการสะสมคะแนน การสะสมแสตมป์ การประชาสัมพันธ์ข้อมูลข่าวสาร โปรโมชั่น กิจกรรม และส่วนลดพิเศษอื่น ๆ อีกมากมาย ดังรูปที่ 1

2.2 ทฤษฎีการยอมรับการใช้เทคโนโลยี (UTAUT)

Venkatesh (2012) ได้กล่าวว่า ทฤษฎีการยอมรับการใช้เทคโนโลยี Unified Theory of Acceptance and Use of Technology หรือ UTAUT2 คือ ทฤษฎีที่อธิบายถึงการยอมรับเทคโนโลยี จนนำไปสู่การใช้เทคโนโลยีของผู้บริโภค [1] ก่อนจะมีทฤษฎี UTAUT2 นั้น ได้มีทฤษฎีที่อธิบายถึงการยอมรับและการใช้เทคโนโลยี โดยได้สรุปวิวัฒนาการของทฤษฎีการยอมรับและใช้เทคโนโลยี โดยทฤษฎี UTAUT2 เป็นทฤษฎีล่าสุดที่ได้รวมทฤษฎีต่าง ๆ เข้าไว้ด้วยกัน ดังรูปที่ 2

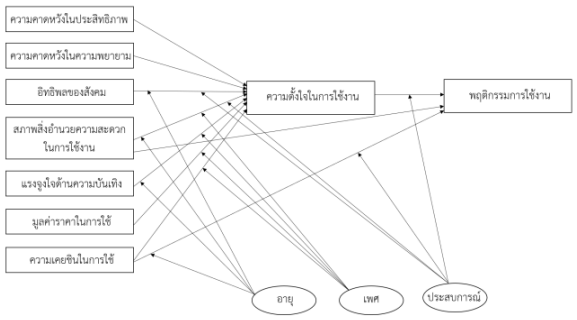
รูปที่ 1. รูปแบบของแอปพลิเคชัน 7-Eleven TH



รูปที่ 2. วิวัฒนาการของทฤษฎีการยอมรับการใช้เทคโนโลยี

สิงหะ ฉวีสุข และสุนันทา วงศ์จตุรภัทร (2555) ได้กล่าวเกี่ยวกับทฤษฎี UTAUT2 ว่าเป็นทฤษฎีที่ถูกพัฒนามาเพื่อลดข้อจำกัดของการศึกษาของการนำเทคโนโลยีมาใช้ในองค์การธุรกิจ และแสดงให้เห็นถึงปัจจัยสำคัญที่สามารถให้คำอธิบายการยอมรับหรือไม่ยอมรับเทคโนโลยีได้อย่างน่าเชื่อถือและเพียงพอ [2] หลักการของ UTAUT 2 ศึกษาพฤติกรรมการใช้ ที่ได้รับแรงขับเคลื่อนจาก ความตั้งใจแสดงพฤติกรรม โดยปัจจัยที่มีอิทธิพลต่อความตั้งใจแสดงพฤติกรรมประกอบด้วยปัจจัยหลัก 7 ประการ ได้แก่ (1) ความคาดหวังในประสิทธิภาพ (2) ความคาดหวังในความพยายาม (3) อิทธิพลของสังคม (4) สภาพสิ่งอานวยความสะดวกในการใช้งาน (5) แรงจูงใจด้านความบันเทิง (6) มูลค่าราคา และ (7) ความเคยชิน ส่วนตัวแปรเสริม/ตัวผัน

แปร จานวน 3 ตัวแปร ได้แก่ (1) เพศ (2) อายุ และ (3) ประสบการณ์ ยกเว้นตัวแปรความสมัครใจในการใช้งาน ไม่ได้ถูกนามาศึกษา เพราะเนื่องจากกลุ่มตัวอย่างคือกลุ่มผู้บริโภคที่ใช้ Mobile internet โดยสมัครใจ ความสัมพันธ์ระหว่างปัจจัยหลักและตัวแปรเสริม/ตัวผันแปรตามทฤษฎี UTAUT2 แสดงในรูปของแบบจำลอง ดังรูปที่ 3



รูปที่ 3. แบบจำลองแสดงความสัมพันธ์ของปัจจัยใน UTAUT2

พฤติกรรมการใช้งานเทคโนโลยีจะประกอบด้วยปัจจัยหลัก 7 ประการ ซึ่งได้แสดงถึงความสัมพันธ์ระหว่างความตั้งใจแสดงพฤติกรรม และ/หรือ พฤติกรรมการใช้ได้รับอิทธิพลจาก 7 ปัจจัยหลัก ดังนี้

(1) ความคาดหวังด้านสมรรถภาพ (Performance Expectancy : PE) หมายถึง ความเชื่อของแต่ละคนว่าการใช้เทคโนโลยีสามารถช่วยเพิ่มประสิทธิภาพการปฏิบัติงานให้กับผู้ใช้ได้ เป็นปัจจัยที่มีอิทธิพลโดยตรงต่อความตั้งใจแสดงพฤติกรรม และสามารถใช้อธิบายความสัมพันธ์ระหว่างความตั้งใจแสดงพฤติกรรม และพฤติกรรมการใช้เทคโนโลยีได้ โดยมีปัจจัยที่เกี่ยวข้อง คือ การรับรู้ถึงประโยชน์ที่ได้รับจากเทคโนโลยี ความสามารถของเทคโนโลยีที่แต่ละบุคคลเชื่อว่าการใช้งานเทคโนโลยีจะเพิ่มประสิทธิภาพการทางานได้ แรงจูงใจภายนอก ความคาดหวังในผลลัพธ์ของการทำงาน

(2) ความคาดหวังจากความพยายาม (Effort Expectancy: EE) หมายถึง ความง่ายในการใช้งาน มีปัจจัยที่เกี่ยวข้อง คือ การรับรู้ว่าเป็นระบบที่ง่ายต่อการใช้งาน นวัตกรรมมีความง่ายต่อการใช้งาน และความง่ายต่อการใช้งานตามที่กฎในแบบจำลอง

(3) อิทธิพลทางสังคม (Social Influence: SI) หมายถึง การรับรู้ของแต่ละคนว่ากลุ่มคนอื่นที่มีความสำคัญต่อตัวเองได้ให้ความหวังว่า แต่ละคนควรใช้เทคโนโลยีใหม่ ถือเป็นปัจจัยที่ส่งผลทางตรงต่อความตั้งใจแสดงพฤติกรรม สาหรับปัจจัยที่เกี่ยวข้อง คือ บรรทัดฐาน และปัจจัยทางสังคม

(4) สภาพสิ่งอานวยความสะดวก (Facilitating Conditions: FC) หมายถึง ความเชื่อของแต่ละคนว่าโครงสร้างพื้นฐานองค์การที่ดีจะช่วยส่งเสริม และเพิ่มความสะดวกในการใช้งานได้ ปัจจัยที่มีความเกี่ยวข้อง ได้แก่ การรับรู้ถึงการควบคุมพฤติกรรมของตนเอง สภาพสิ่งอานวยความสะดวก และความสอดคล้องเหมาะสมกับผู้ใช้งาน

(5) แรงจูงใจด้านความบันเทิง (Hedonic Motivation: HM) หมายถึง ความสนุก ความพอใจที่ได้รับจากการใช้เทคโนโลยี

(6) มูลค่าตามราคา (Price Value: PV) หมายถึง ความรู้ความคิดและทักษะการเปรียบเทียบของผู้บริโภคเกี่ยวกับประโยชน์ที่จะได้รับ และค่าใช้จ่ายเพื่อการใช้ประโยชน์นั้น ๆ

(7) อุปนิสัยส่วนบุคคล (Habit: H) หมายถึง การที่บุคคลมีแนวโน้มแสดงพฤติกรรมโดยอัตโนมัติ เนื่องด้วยสิ่งที่เรียนรู้มาในอดีต และได้เคยปฏิบัติอย่างสม่ำเสมอมาจนกลายเป็นอุปนิสัยส่วนตัว

2.3 งานวิจัยที่เกี่ยวข้อง

วอนชนก ไชยสุนทร (2558) ได้ศึกษาปัจจัยในด้านมูลค่าตามราคาพบว่า ปัจจัยด้านมูลค่าตามราคาไม่ส่งผลต่อการยอมรับระบบการชำระเงินในประเทศไทย [3] ส่วนปัจจัยด้านราคามีความสอดคล้องกับการยอมรับการบริการ นอกจากนั้นค่าบริการส่งผลต่อการใช้งานอิเล็กทรอนิกส์ทางการเงินอย่างต่อเนื่อง และส่งผลต่อทัศนคติอันดีต่อ Mobile Banking

เกวรินทร์ ละเอียดดีนันท์ และนิตนา ฐานิตธนกร (2559) ได้ศึกษาเรื่องการยอมรับเทคโนโลยีและพฤติกรรมผู้บริโภคทางออนไลน์ที่มีผลต่อการตัดสินใจซื้อหนังสืออิเล็กทรอนิกส์ของผู้บริโภคในเขตกรุงเทพมหานคร [4] ผลการศึกษาพบว่า การยอมรับเทคโนโลยี ด้านการนำมาใช้งานจริงส่งผลต่อการตัดสินใจซื้อหนังสืออิเล็กทรอนิกส์ของผู้บริโภคในเขตกรุงเทพมหานครมากที่สุด รองลงมา คือ พฤติกรรมผู้บริโภคออนไลน์ ด้านทัศนคติที่มีต่อสื่อออนไลน์ การยอมรับเทคโนโลยี ด้านความง่ายในการใช้งาน พฤติกรรมผู้บริโภคออนไลน์ ด้านความบันเทิงทางออนไลน์ ด้านการรับรู้ทางออนไลน์และการยอมรับเทคโนโลยี ด้านความตั้งใจที่จะใช้ตามลำดับ

ชวิศา พุ่มดนตรี (2559) ได้ศึกษาเกี่ยวกับปัจจัยที่ส่งผลต่อการยอมรับการใช้บริการพร้อมเพย์ของประชาชนในเขตกรุงเทพมหานครและปริมณฑล [5] พบว่าปัจจัยที่ส่งผลต่อการยอมรับการใช้บริการพร้อมเพย์ของประชาชน ในเขตกรุงเทพมหานคร และปริมณฑล มีทั้งหมด 8 ปัจจัย โดยแบ่งเป็นปัจจัยประชากรศาสตร์ 2 ปัจจัย คือ ปัจจัยด้านอาชีพ และรายได้ และปัจจัยการยอมรับเทคโนโลยี และการรับรู้ความเสี่ยง 6 ปัจจัย ซึ่งมีผลในทางลบ 1 ปัจจัย คือ ปัจจัยการรับรู้ความเสี่ยง ส่วน 5 ปัจจัยที่เหลือมีผลทางบวก โดยทั้ง 6 ปัจจัยสามารถเรียงลาดับปัจจัยที่ส่งผลต่อการยอมรับจากมากไปหาน้อย คือ ปัจจัยการรับรู้ความเสี่ยง ปัจจัยด้านมูลค่าตามราคา ปัจจัยด้านความคาดหวังด้านสมรรถภาพ ปัจจัยด้านความคาดหวังจากความพยายาม และสภาพสิ่งอำนวยความสะดวก

พรชนก พลาบูลย์ (2560) ได้ศึกษาเรื่องการยอมรับนวัตกรรมและเทคโนโลยี การใช้เทคโนโลยีและพฤติกรรมผู้บริโภคที่ส่งผลต่อความตั้งใจของประชาชนในการใช้บริการธุรกรรมทางการเงินผ่านระบบพร้อมเพย์ [6] ซึ่งพักอาศัยอยู่ในกรุงเทพมหานคร จำนวน 370 คน และสถิติเชิงอนุมานที่ใช้ทดสอบสมมติฐาน ได้แก่ การวิเคราะห์ความถดถอยเชิงพหุ กลุ่มตัวอย่างที่ศึกษาส่วนใหญ่เป็นเพศหญิง มีอายุ 20-25 ปี มีการศึกษาในระดับปริญญาตรี มีอาชีพข้าราชการ ส่วนใหญ่รู้จักบริการพร้อมเพย์จากโทรทัศน์

ฉัตรเฉลิม เกิดสวัสดิ์ (2560) ได้ศึกษาเกี่ยวกับรูปแบบความสัมพันธ์เชิงสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านโซเชียลมีเดียของข้าราชการโรงเรียนนายร้อยพระจุลจอมเกล้า [7] เพื่อตรวจสอบความสอดคล้องของรูปแบบความสัมพันธ์เชิงสาเหตุที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ โดยมีกลุ่มตัวอย่างจำนวน 400 คน ได้มาโดยการสุ่มแบบง่ายจากข้าราชการโรงเรียนนายร้อยพระจุลจอมเกล้า เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ได้แก่ แบบสอบถามแบบมาตรประมาณค่า 7 ระดับ ใช้เทคนิคการวิเคราะห์องค์ประกอบและโมเดลสมการโครงสร้างโดยใช้โปรแกรมสำเร็จรูป ผลการวิเคราะห์รูปแบบความสัมพันธ์เชิงสาเหตุ พบว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์เป็นอย่างดี

พิมพ์พรรณ สุวรรณศิริศิลป์ (2559) ได้ศึกษาปัจจัยที่ส่งผลต่อการยอมรับและใช้งานบริการชำระเงินทางอิเล็กทรอนิกส์ เพื่อศึกษาหาปัจจัยที่ส่งผลต่อการยอมรับและใช้งานบริการ [8] โดยการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างจำนวน 178 ราย ใช้รู้แบบการวิจัยเชิงสำรวจ ด้วยข้อคำถามจำนวน 62 ข้อ พบว่าปัจจัยด้านความเคยชินเป็นปัจจัยที่มีอิทธิพลต่อความตั้งใจในการใช้บริการแบบพร้อมเพย์มากที่สุด ส่วนปัจจัยด้านการรับรู้ความเสี่ยงมีอิทธิพลเชิงลบต่อความตั้งใจในการใช้งานบริการแบบพร้อมเพย์ นอกจากนี้ ปัจจัยด้านความตั้งใจในการใช้งานบริการแบบพร้อมเพย์และความเคยชินยังมีอิทธิพลในเชิงบวกต่อการใช้งานบริการแบบพร้อมเพย์โดยมีความสำคัญต่อการใช้งานบริการแบบพร้อมเพย์ตามลำดับ

ยงยุทธ บุญกิจ (2562) ได้ศึกษารูปแบบความสัมพันธ์เชิงสาเหตุการยอมรับและการใช้เทคโนโลยีที่มีต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านสื่อสังคมออนไลน์ของบุคลากรสายสนับสนุนในมหาวิทยาลัยธรรมศาสตร์ [9] เพื่อพัฒนารูปแบบความสัมพันธ์เชิงสาเหตุการยอมรับและการใช้เทคโนโลยีที่มีต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านสื่อสังคมออนไลน์ และเพื่อตรวจสอบความสอดคล้องของรูปแบบความสัมพันธ์เชิงสาเหตุของโมเดลที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ จากกลุ่มตัวอย่างจำนวน 400 คน พบว่าพฤติกรรมความตั้งใจใช้งาน สภาพสิ่งอำนวยความสะดวกในการใช้งาน และความเคยชินในการใช้งาน มีอิทธิพลทางตรงพฤติกรรมการสื่อสารภายในองค์กรผ่านสื่อผ่านแอปพลิเคชันไลน์ ในขณะที่ความเคยชินในการใช้งานแอปพลิเคชันไลน์ มีอิทธิพลทางอ้อมต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านแอปพลิเคชันไลน์โดยผ่านพฤติกรรมความตั้งใจใช้งาน

สำหรับงานวิจัยนี้จะนำเสนอการประเมินการใช้งานแอปพลิเคชัน 7-Eleven TH บนพื้นฐานของทฤษฎีการยอมรับและการใช้เทคโนโลยี เพื่อวิเคราะห์รูปแบบความสัมพันธ์กับผู้ใช้งานในเขตกรุงเทพมหานครและปริมณฑล

3. วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยแบบวิจัยเชิงสำรวจ เพื่อศึกษารูปแบบความสัมพันธ์เชิงสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH จากกลุ่มตัวอย่างการศึกษาของประชากรในเขตกรุงเทพมหานครและปริมณฑล ที่มีประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH จำนวน 400 คน โดยค่าขนาดความเหมาะสมของกลุ่มตัวอย่างด้วยวิธีการใช้สูตรการหาขนาดตัวอย่างที่ระดับความเชื่อมั่น 95% โดยการวิจัยนี้มีวัตถุประสงค์เพื่อศึกษารูปแบบ

ความสัมพันธ์เชิงสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ของประชากรในเขตกรุงเทพมหานครและปริมณฑล เพื่อตรวจสอบความสอดคล้องของรูปแบบความสัมพันธ์เชิงสาเหตุที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์โดยเน้นปัจจัยหลัก 7 อย่างได้แก่ ความคาดหวังในประสิทธิภาพ (Performance Expectancy) ความคาดหวังในความพยายาม (Effort Expectancy) อิทธิพลของสังคมต่อการใช้งาน (Social Influence) สภาพสิ่งอำนวยความสะดวกในการใช้งาน (Facilitating Conditions) แรงจูงใจด้านความบันเทิง (Hedonic Motivation) มูลค่าราคา (Price Value) และ ความเคยชิน (Habit) รวมทั้ง 3 ตัวแปรเสริม/ตัวผันแปรที่ส่งผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ได้แก่ เพศ อายุ และประสบการณ์

3.1 ประชากรและกลุ่มตัวอย่าง

3.1.1 ประชากร

ประชากรของการวิจัยครั้งนี้ คือ ผู้ใช้งานแอปพลิเคชัน 7-Eleven TH ในเขตกรุงเทพมหานครและปริมณฑล ซึ่งไม่ทราบจำนวนประชากรที่แน่นอน

3.1.2 กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการศึกษาเป็น ประชากรในเขตกรุงเทพมหานครและปริมณฑล ที่มีประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH จำนวน 400 คน ผู้วิจัยได้ทำการหาขนาดความเหมาะสมของกลุ่มตัวอย่างจึงได้กำหนดขนาดกลุ่มตัวอย่างโดยใช้สูตรการหาขนาดตัวอย่างที่ระดับความเชื่อมั่น 95% [10] จะได้กลุ่มตัวอย่างที่ใช้ในการศึกษาจำนวนมากกว่า 400 คนเพื่อป้องกันความคาดเคลื่อนของข้อมูล ผู้วิจัยจึงใช้กลุ่มตัวอย่างประมาณ 420 คนเป็นอย่างน้อย ตัวอย่างที่ใช้ในการศึกษา ได้มาจากการสุ่มกลุ่มตัวอย่างโดยไม่คำนึงถึงความน่าจะเป็น (Non-Probability Sampling) ซึ่งผู้วิจัยเลือกใช้วิธีการสุ่มกลุ่มตัวอย่างแบบสะดวก (Convenience Sampling) โดยการเก็บกลุ่มตัวอย่างแบบออนไลน์ เพื่อต้องการให้ได้ผู้ที่มีประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH อย่างแท้จริง

3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในเก็บรวบรวมข้อมูลสำหรับการวิจัยในครั้งนี้เป็นแบบสอบถามเกี่ยวกับสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ซึ่งผู้วิจัยได้ค้นคว้างานวิจัยต่าง ๆ ที่เกี่ยวข้อง และทำการรวบรวมข้อมูลเพื่อจัดทำแบบสอบถาม โดยแบ่งออกเป็น 5 ส่วน ดังนี้

3.2.1 การศึกษาแนวคิดและทฤษฎี

ผู้วิจัยดำเนินการศึกษาแนวคิด ทฤษฎีและเอกสารการวิจัย ที่เกี่ยวข้องกับสาเหตุของการยอมรับและการใช้เทคโนโลยี และพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH

3.2.2 การสร้างแบบสอบถาม

ผู้วิจัยดำเนินการสร้างแบบสอบถามแบบมาตรประมาณค่า 7 ระดับขึ้น ให้สอดคล้องกับกรอบแนวคิดและวัตถุประสงค์ของการวิจัย

3.2.3 การนำเสนอแบบสอบถาม

ผู้วิจัยสร้างนำเสนอแบบสอบถามที่สร้างขึ้น เสนอต่ออาจารย์ที่ปรึกษาพิจารณาตรวจสอบความถูก ต้องเสนอแนะเพิ่มเติม ปรับปรุง แก้ไข เพื่อให้ผู้อ่านเกิดความเข้าใจง่าย ชัดเจน ตรงตามวัตถุประสงค์ของการวิจัย

3.2.4 การตรวจสอบความถูกต้องของแบบสอบถาม

นำแบบสอบถามที่ผ่านการพัฒนาและปรับปรุงแล้วให้ผู้เชี่ยวชาญ จำนวน 3 ท่าน ตรวจสอบความตรงด้านเนื้อหาของแบบสอบถาม (Content Validity) รวมทั้งตรวจสอบความถูกต้องเหมาะ สมของเนื้อหา ภาษาที่ใช้ จากนั้นนำมาคำนวณหาค่าดัชนีความสอดคล้องของข้อคำถาม กับวัตถุประสงค์ของการวิจัย (Index of Congruence หรือ IOC)

โดยมีเกณฑ์ในการพิจารณาค่าสอดคล้องของข้อคำถามกับวัตถุประสงค์ของการวิจัย คือค่า IOC ต้องมากกว่า 0.50 (IOC > 0.50) จึงจะถือว่าข้อคำถามนั้นมีความสอดคล้องกับเนื้อหาและวัตถุประสงค์ของการวิจัย [11] และมีการปรับข้อคำถามให้สอดคล้องตามที่ผู้เชี่ยวชาญแนะนำ จากนั้นจึงให้อาจารย์ที่ปรึกษาตรวจสอบและให้ความเห็นชอบอีกครั้ง ทั้งนี้จากการที่ผู้วิจัยได้นำแบบสอบถามที่ผ่านการพัฒนาและปรับปรุงแล้วให้ผู้เชี่ยวชาญจำนวน 3 ท่าน

ผลการตรวจสอบความตรงด้านเนื้อหาของแบบสอบถาม (Content Validity) มีค่าอยู่ที่ 0.67-1.00 รวมทั้งการตรวจสอบความถูกต้อง ความเหมาะสมของเนื้อหา ภาษาที่ใช้แล้ว พบว่าข้อคำถามเหล่านั้นมีความสอดคล้องกับเนื้อหาและวัตถุประสงค์ของการวิจัย

3.2.5 การทดลองใช้แบบสอบถามกับกลุ่มตัวอย่าง

การนำแบบสอบถามที่ผ่านจากการตรวจสอบของผู้เชี่ยวชาญมาทดลองใช้ (Try Out) กับกลุ่มตัวอย่าง จำนวน 30 ชุด เพื่อหาค่าความเชื่อมั่นโดยใช้วิธีหาค่าสัมประสิทธิ์อัลฟ่าโดยวิธีการคำนวณของครอนบัค (Cronbach) [12] ค่าอัลฟ่าที่ได้จะแสดงถึงระดับความคงที่ของแบบสอบถาม โดยจะมีค่าระหว่าง $0 \leq \alpha \leq 1$ ค่าที่ใกล้เคียงกับ 1 มาก แสดงว่ามีความเชื่อมั่นสูง ได้ค่าสัมประสิทธิ์อัลฟ่า ทั้งฉบับเท่ากับ 0.98 สอดคล้องกับเกณฑ์คุณภาพของเครื่องมือที่ควรจะมีค่าความเชื่อมั่น 0.70 ขึ้นไป แสดงให้เห็นว่าข้อคำถามในแบบสอบถามนั้นมีความน่าเชื่อถือในระดับสูง [13]

3.2.6 การเผยแพร่แบบสอบถามกับกลุ่มตัวอย่างจริง

การนำแบบสอบถามที่ได้ทำการแก้ไขโดยสมบูรณ์ในรูปแบบของแบบสอบถามออนไลน์แล้วจึงนำมาใช้กับกลุ่มตัวอย่างจริง

3.3 การเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้เป็นการวิจัยแบบวิจัยเชิงสำรวจ เพื่อศึกษารูปแบบความสัมพันธ์เชิงสาเหตุที่ส่งผลต่อการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้ 7-Eleven TH โดยผู้วิจัยได้ดำเนินการเก็บข้อมูลโดยวิธีการเผยแพร่แบบสอบถามออนไลน์ (Questionnaires Online) จากประชากรในเขตกรุงเทพมหานครและปริมณฑล ที่มีประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH โดยเก็บข้อมูลในช่วงเดือนพฤศจิกายน พ.ศ. 2562 ถึง เดือนกุมภาพันธ์ พ.ศ. 2563 รวมระยะเวลาในการเก็บแบบสอบถามทั้งสิ้น 4 เดือน มีผู้ตอบแบบสอบถามมีจำนวน 400 คน หลังจากนั้นผู้วิจัยได้ทำการคัดเลือกแบบสอบถามที่มีความสมบูรณ์ได้จำนวนผู้ตอบแบบสอบถาม 400 คนมาทำการตรวจสอบความถูกต้องของข้อมูลที่จะนำไปวิเคราะห์ข้อมูลทางสถิติต่อไป

3.4 การวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ผู้วิจัยได้ใช้หลักการสถิติเชิงบรรยายในการบรรยายข้อมูลและใช้หลักการสถิติเชิงอนุมานในการบรรยายผลการวิเคราะห์ข้อมูลที่เก็บรวบรวมมาได้โดยมีรายละเอียดดังนี้

3.4.1 วิเคราะห์ค่าสถิติเชิงบรรยาย

กลุ่มตัวอย่างที่ได้ทำการเก็บข้อมูล เพื่ออธิบายลักษณะทั่วไปของผู้ตอบแบบสอบถามโดยวิธีการคำนวณหาค่าความถี่ (Frequency) ร้อยละ (Percentage) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ค่าความเบ้ (Skewness) และค่าความโด่ง (Kurtosis)

3.4.2 วิเคราะห์ข้อมูลสถิติเชิงอนุมาน

เพื่อนำมาใช้ในการวิเคราะห์ข้อมูลในการหารูปแบบความสัมพันธ์เชิงสาเหตุของตัวแปรที่กำหนดไว้ตามโมเดลสมมติฐานโดยใช้แนวคิดทฤษฎีพฤติกรรมการยอมรับและการใช้เทคโนโลยี (Unified Theory of Acceptance and Use of Technology: UTAUT) [14] มีการพัฒนาตั้งแต่ ปี ค.ศ. 2003 โดยใช้โปรแกรมสำเร็จรูป ในการปรับโมเดล และทำการทดสอบความสอดคล้องของโมเดลสมมติฐานกับข้อมูลที่ได้จากกลุ่มตัวอย่าง จนผลการทดสอบไม่มีความแตกต่างทางสถิติ

3.4.3 วิเคราะห์ความสัมพันธ์เชิงสาเหตุ

ใช้วิธีการวิเคราะห์เทคนิคสมการโครงสร้างเพื่อหาเส้นทางอิทธิพลเชิงสาเหตุของตัวแปร หาขนาดอิทธิพลว่ามีมากน้อยเพียงใด และมีทิศทางแบบใด จากแนวคิดและทฤษฎีที่ผู้วิจัยใช้อ้างอิงโดยมีการทดสอบความกลมกลืนระหว่างโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์

3.4.4 ดำเนินการปรับโมเดล

หากโมเดลไม่มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์

3.4.5 สรุปผลและอภิปรายผลการวิเคราะห์

เมื่อโมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์แล้วจึงสรุปและอภิปรายผลการวิเคราะห์ตามลำดับต่อไป

4. ผลการวิจัย

การวิจัยนี้ได้แบ่งการนำเสนอผลการวิเคราะห์ข้อมูลดังนี้

4.1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

4.1.1 การวิเคราะห์จำนวนของผู้ตอบแบบสอบถาม

หลังจากที่ผู้วิจัยดำเนินการเก็บข้อมูลโดยวิธีการเผยแพร่แบบสอบถามออนไลน์ (Online Questionnaires) จากประชากรในเขตกรุงเทพมหานครและปริมณฑล ที่มีประสบการณ์การใช้งานแอปพลิเคชั่น 7-Eleven TH ได้แบบสอบถามจำนวนทั้งสิ้น 450 ฉบับ มาคัดเลือกแบบสอบถามที่มีความสมบูรณ์ จำนวน 413 ฉบับ และเมื่อทำการตรวจสอบความถูกต้องของข้อมูลแล้วสามารถนำมาวิเคราะห์ข้อมูลเบื้องต้นของกลุ่มตัวอย่าง แสดงในตาราง 4.1

ตารางที่ 1. จำนวนและร้อยละของผู้ตอบแบบสอบถาม

สถานภาพ	จำนวน	ร้อยละ
พักอาศัย		
กรุงเทพมหานคร	196	47.5
ปริมณฑล	217	52.5
รวม	413	100.00
เพศ		
ชาย	158	38.3
หญิง	255	61.7
รวม	413	100.00
อายุ		
ต่ำกว่า 20 ปี	37	9.0
อายุ 21 - 25 ปี	80	19.4
อายุ 26 – 30 ปี	122	29.5
อายุ 31 - 35 ปี	84	20.3
อายุ 36 - 40 ปี	38	9.2
อายุ 41 - 45 ปี	18	4.4
อายุ 45 ปีขึ้นไป	34	8.2
รวม	413	100.00

จากตารางที่ 1 ผลการวิเคราะห์ข้อมูลทั่วไปของผู้ตอบแบบสอบถามทุกคนมีประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH จากจำนวนผู้ตอบแบบสอบถามทั้งหมด 413 คน พบว่าเป็นคนที่พักอาศัยอยู่ในกรุงเทพมหานครจำนวน 196 คน คิดเป็นร้อยละ 47.5 และพักอาศัยอยู่ในเขตปริมณฑล จำนวน 217 คน คิดเป็นร้อยละ 52.5 ส่วนใหญ่ของผู้ตอบแบบสอบถามเป็นเพศหญิง จำนวน 255 คน คิดเป็นร้อยละ 61.7 และเป็นเพศชายจำนวน 158 คน คิดเป็นร้อยละ 38.3 อายุเฉลี่ยของผู้ตอบแบบสอบถามอยู่ระหว่างอายุ 26-30 ปี จำนวน 122 คน คิดเป็นร้อยละ29.5 ผู้ตอบแบบสอบถามส่วนใหญ่มีระดับการศึกษาในระดับปริญญาตรี จำนวน 311 คน คิดเป็นร้อยละ 75.3 ด้านระยะเวลาในการใช้อินเทอร์เน็ตบนโทรศัพท์มือถือระหว่าง 5-6 ชั่วโมงต่อวันมากถึง 176 คน คิด เป็นร้อยละ 42.6 มากกว่า 6 ชั่วโมง จำนวน 142 คน คิดเป็นร้อยละ 34.4 ระหว่าง 3-4 ชั่วโมง จำนวน 76 คน คิดเป็นร้อยละ 18.4 และระหว่าง 1-2 ชั่วโมง มีเพียง 19 คนคิดเป็นร้อยละ 4.6 เท่านั้น ด้านประสบการณ์การใช้แอปพลิเคชัน 7-Eleven TH ระหว่าง 7 เดือน - 1 ปี มีจำนวนมากถึง 165 คน คิดเป็นร้อยละ 40.0 น้อยกว่า 6 เดือน จำนวน 137 คน คิดเป็นร้อยละ 33.2 ระหว่าง 1-2 ปี จำนวน 87 คน คิดเป็นร้อยละ 21.1 และมากกว่า 2 ปี จำนวน 24 คน คิดเป็นร้อยละ 5.8 ด้านความถี่ในการเข้าไปซื้อสินค้า/บริการ ที่ร้านสะดวกซื้อ 7-Eleven ผู้ตอบแบบสอบถามเข้าไปซื้อสินค้า/บริการ ที่ร้านสะดวกซื้อ 7-Eleven ทุกวัน มากถึง 216 คน คิดเป็นร้อยละ 52.3 สัปดาห์ละ 2-3 ครั้ง จำนวน 150 คน คิดเป็นร้อยละ 36.3 สัปดาห์ละครั้ง จำนวน 24 คน คิดเป็นร้อยละ 5.8 เดือนละ 2-3 ครั้ง จำนวน 9 คน คิดเป็นร้อยละ 2.2 เดือนละครั้ง จำนวน 8 คน คิดเป็นร้อยละ 1.9 และน้อยกว่าเดือนละครั้ง จำนวน 6 คน คิดเป็นร้อยละ 1.5 ตามลำดับ

4.1.2 ผลการวิเคราะห์ลักษณะตัวแปรในการวิจัย

ผลจากการวิเคราะห์ลักษณะตัวแปรในการวิจัย พบว่าส่วนของค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความโด่งและความหมายของข้อคำถามในแต่ละข้อ แต่ละด้าน ดังนี้

(1) ความคาดหวังในประสิทธิภาพการใช้แอปพลิเคชัน

ความคาดหวังในประสิทธิภาพการใช้แอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.38 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านคิดว่าแอปพลิเคชัน 7-Eleven TH มีประโยชน์ต่อการรับรู้ข้อมูลข่าวสาร โปรโมชั่น และสิทธิพิเศษอื่น ๆ ของท่านมีค่าเฉลี่ยมากที่สุดเท่ากับ 5.55 รองลงมา คือ ท่านคิดว่าการใช้แอปพลิเคชัน 7-Eleven TH ช่วยให้เกิดความสะดวกต่อการเข้าไปซื้อสินค้า/บริการในร้านสะดวกซื้อ 7-Eleven ของท่าน มีค่าเฉลี่ยเท่ากับ 5.48

(2) ความคาดหวังในความพยายามใช้แอปพลิเคชัน

ความคาดหวังในความพยายามใช้แอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.54 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านคิดว่าการเพิ่มทักษะความชำนาญในการใช้แอปพลิเคชัน 7-Eleven TH ทำได้ง่าย มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.62 รองลงมา คือ ท่านคิดว่าแอปพลิเคชัน 7-Eleven TH สามารถเรียนรู้การใช้งานได้ง่าย มีค่าเฉลี่ยเท่ากับ 5.60

(3) อิทธิพลของสังคมต่อการใช้งานแอปพลิเคชัน

อิทธิพลของสังคมต่อการใช้งานแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.12 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านคิดว่าการใช้แอปพลิเคชัน 7-Eleven TH ช่วยให้ท่านก้าวทันเทคโนโลยี มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.34 รองลงมา คือ ท่านคิดว่าการใช้ แอปพลิเคชัน 7-Eleven TH ช่วยเสริมสร้างภาพลักษณ์ของท่านว่าเป็นคนทันสมัยและทันต่อเทคโนโลยี มีค่าเฉลี่ยเท่ากับ 5.21

(4) สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน

สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.52 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านมีความรู้ทางเทคโนโลยีมากพอที่จะใช้แอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.79 รองลงมา คือ ท่านคิดว่าการชำระค่าสินค้าและบริการผ่าน

แอปพลิเคชัน 7-Eleven TH ง่ายกว่าการใช้เงินสดมีค่าเฉลี่ยเท่ากับ 5.59

(5) แรงจูงใจด้านความบันเทิงในการใช้แอปพลิเคชัน

แรงจูงใจด้านความบันเทิงแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.36 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านรู้สึกพึงพอใจที่ได้ใช้คูปองส่วนลดจากแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.70 รองลงมา คือ ท่านรู้สึกดีทุกครั้งที่ได้สะสมหรือแลกคะแนนผ่านแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.59

(6) มูลค่าราคาในการใช้แอปพลิเคชัน

มูลค่าราคาในการใช้แอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 5.64 หรือในระดับมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านคิดว่าคูปองส่วนลดจากแอปพลิเคชัน 7-Eleven TH ช่วยประหยัดค่าใช้จ่าย มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.74 รองลงมา คือ ท่านคิดว่าการชำระค่าสินค้า/บริการผ่านแอปพลิเคชัน 7-Eleven TH ช่วยให้ท่านได้รับคะแนนสะสมหรือสิทธิพิเศษอื่น ๆ เพิ่มขึ้น มีค่าเฉลี่ยเท่ากับ 5.72

(7) ความเคยชินในการใช้แอปพลิเคชัน

ความเคยชินในการใช้แอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 4.87 หรือในระดับค่อนข้างมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านใช้แอปพลิเคชัน 7-Eleven TH ในการรับข้อมูล ข่าวสาร รายการส่งเสริมการขายของร้านสะดวกซื้อ 7-Eleven เป็นประจำอย่างต่อเนื่อง มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.36 รองลงมา คือ ท่านใช้แอปพลิเคชัน 7-Eleven TH เพื่อตรวจสอบคะแนนสะสมเป็นประจำ มีค่าเฉลี่ยเท่ากับ 4.85

(8) พฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน

พฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 4.83 หรือในระดับค่อนข้างมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า เมื่อท่านเข้าถึง แอปพลิเคชัน 7-Eleven TH ท่านมีความตั้งใจที่จะทำรายการใดรายการหนึ่ง เช่น ตรวจสอบคะแนนสะสม เป็นต้น มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.18 รองลงมา คือ ท่านมีความตั้งใจในการใช้แอปพลิเคชัน 7-Eleven TH เพื่อร่วมรายการส่งเสริมการขายของร้านสะดวกซื้อ 7-Eleven มีค่าเฉลี่ยเท่ากับ 4.97

(9) พฤติกรรมการใช้งานแอปพลิเคชัน

พฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 4.88 หรือในระดับค่อนข้างมาก เมื่อจำแนกเป็นรายข้อปรากฏว่า ท่านมีการสะสมและแลกคะแนนผ่านแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยมากที่สุดเท่ากับ 5.16 รองลงมา คือ ท่านรับข้อมูล ข่าวสาร รายการส่งเสริมการขายของร้านสะดวกซื้อ 7-Eleven ผ่านแอปพลิเคชัน 7-Eleven TH มีค่าเฉลี่ยเท่ากับ 4.91

4.2 ผลการวิเคราะห์ความสัมพันธ์เชิงสาเหตุ

ผลการวิเคราะห์ความสัมพันธ์เชิงสาเหตุโดยใช้วิธีการวิเคราะห์เทคนิคสมการโครงสร้างเพื่อหาเส้นทางอิทธิพลเชิงสาเหตุของตัวแปร ผลจากการวิเคราะห์การศึกษาการยอมรับและการใช้งานแอปพลิเคชั่น 7-Eleven TH กรณีศึกษาของประชากรในเขตกรุงเทพมหานครและปริมณฑล โดยใช้วิธีการวิเคราะห์เทคนิคสมการโครงสร้างเพื่อหาเส้นทางอิทธิพลเชิงสาเหตุของตัวแปร โดยการทดสอบความกลมกลืนระหว่างโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์ พบตัวแปรแฝง และตัวแปรสังเกตได้

เมื่อพิจารณาผลจากการวิเคราะห์ความสัมพันธ์เชิงสาเหตุโดยใช้วิธีการวิเคราะห์เทคนิคสมการโครงสร้างเชิงเส้นเพื่อหาเส้นทางอิทธิพลเชิงสาเหตุของตัวแปร โดยการทดสอบความกลมกลืนระหว่างโมเดลสมมติฐานกับข้อมูลเชิงประจักษ์ ค่าสถิติทดสอบไคสแควร์ (Chi-Square) ของการประมาณค่าพารามิเตอร์แบบ Maximum Likelihood (ML) ควรมีค่า p-Value มากกว่า 0.05 ค่าสถิติไคสแควร์สัมพัทธ์ (CMIN/DF) และใช้ค่าสถิติอื่น ๆ ประกอบการพิจารณา เช่น RMR GFI AGFI และ PGFI ในงานวิจัยนี้ ได้ค่า p-value เท่ากับ 0.07 พบค่าสถิติ Chi-square เท่ากับ 621.264 กับองศาอิสระ เท่ากับ 538 ค่า CMIN/DF เท่ากับ 1.16 ค่า GFI เท่ากับ 0.93 ค่า AGFI เท่ากับ 0.90 ค่า SRMR เท่ากับ 0.04 และค่า RMSEA เท่ากับ 0.02 สรุปได้ว่าโมเดลมีความสอดคล้องและกลมกลืนกับข้อมูลเชิงประจักษ์เป็นอย่างดี [15] ดังตารางที่ 2

ตารางที่ 2. ความสอดคล้องจากการวิเคราะห์สมการโครงสร้าง

ค่าสถิติที่ใช้ในการตรวจสอบ	เกณฑ์การพิจารณา	ค่าสถิติที่ได้	การพิจารณา
CMIN/DF	< 2.00	1.86	ผ่านเกณฑ์
GFI	≥ 0.90	0.91	ผ่านเกณฑ์
AGFI	≥ 0.90	0.90	ผ่านเกณฑ์
RMSEA	≤ 0.08	0.02	ผ่านเกณฑ์
SRMR	≤ 0.08	0.04	ผ่านเกณฑ์
TLI	≥ 0.90	0.99	ผ่านเกณฑ์
CFI	≥ 0.90	0.99	ผ่านเกณฑ์
Hoelter	> 200	394	ผ่านเกณฑ์

ซึ่งจากผลการวิเคราะห์ค่าสถิติตามตารางที่ 2 นั้นหมายถึงโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์ ผู้วิจัยจึงไม่จำเป็นต้องทำการปรับแต่งโมเดลเพื่อให้มีความสอดคล้องกับข้อมูลเชิงประจักษ์

ตารางที่ 3. ค่าอิทธิพลต่อพฤติกรรมการใช้งานแอปพลิเคชัน

ตัวแปรแฝงภายนอก	ตัวแปรแฝงภายใน						หมายเหตุ
	พฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH			พฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH			
	อิทธิพลทางตรง	อิทธิพลทางอ้อม	อิทธิพลรวม	อิทธิพลทางตรง	อิทธิพลทางอ้อม	อิทธิพลรวม	
พฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH	-	-	-	1.007***	-	1.007***	ผ่านเกณฑ์
ความคาดหวังในประสิทธิภาพการใช้แอปพลิเคชัน 7-Eleven TH	0.120	-	0.120	-	0.121	0.121***	ผ่านเกณฑ์
ความคาดหวังในความพยายามใช้แอปพลิเคชัน 7-Eleven TH	0.024	-	0.024	-	0.024	0.024***	ผ่านเกณฑ์
อิทธิพลของสังคมต่อการใช้งานแอปพลิเคชัน 7-Eleven TH	-0.059	-	-0.059	-	-0.059	-0.059***	ผ่านเกณฑ์
สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน 7-Eleven TH	0.160	-	0.160	-	0.161	0.161***	ผ่านเกณฑ์
แรงจูงใจด้านความบันเทิงแอปพลิเคชัน 7-Eleven TH	0.214	-	0.214	-	0.215	0.215***	ผ่านเกณฑ์
มูลค่าราคาในการใช้แอปพลิเคชัน 7-Eleven TH	0.152	-	0.152	-	0.153	0.153***	ผ่านเกณฑ์
ความเคยชินในการใช้แอปพลิเคชัน 7-Eleven TH	1.191***	-	1.191***	-	1.199***	1.199***	ผ่านเกณฑ์
R^2	0.87			1.01			

หมายเหตุ: *** หมายถึง มีนัยสำคัญทางสถิติที่ระดับ .001

ตารางที่ 3 แสดงค่าอิทธิพลทางตรง อิทธิพลทางอ้อม และอิทธิพลรวมที่มีผลต่อพฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH แสดงขนาดเส้นทางอิทธิพลของการยอมรับและการใช้งานแอปพลิเคชัน 7-Eleven TH กรณีศึกษาของประชากรในเขตกรุงเทพมหานครและปริมณฑล ได้การวัดภาพรวมของตัวแปรภายนอก และตัวแปรภายในโดยการวัดขนาดอิทธิพลระหว่างตัวแปรในโมเดล มีค่าขนาดอิทธิพลทางตรง (Direct Effects) อิทธิพลทางอ้อม (Indirect Effects) และอิทธิพลรวม (Total Effects) โดยพฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH ได้รับอิทธิพลทางตรงจากตัวแปรแฝงภายนอก 7 ตัวแปรดังนี้

1. ความคาดหวังในประสิทธิภาพการใช้แอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงบวกและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.120
2. ความคาดหวังในความพยายามใช้แอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงบวกและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.024
3. อิทธิพลของสังคมต่อการใช้งานแอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงลบและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.059
4. สภาพสิ่งอำนวยความสะดวกในการใช้งานแอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงบวกและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.160
5. แรงจูงใจด้านความบันเทิงแอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงลบและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.214
6. มูลค่าราคาในการใช้แอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงลบและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 0.152
7. ความเคยชินในการใช้แอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงบวกและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 1.191 แสดงให้เห็นว่าผู้ใช้งานให้ความสำคัญต่อความเคยชินในการใช้แอปพลิเคชัน 7-Eleven TH ซึ่งส่งผลโดยตรงต่อพฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH

พฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH ได้รับอิทธิพลทางตรงจากตัวแปรแฝงภายนอก 1 ตัวแปร คือ ความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH มีอิทธิพลเชิงบวกและมีค่าสัมประสิทธิ์เส้นทาง เท่ากับ 1.007 แสดงให้เห็นว่าความตั้งใจที่จะใช้งานแอปพลิเคชันนั้นส่งผลต่อพฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH แสดงให้เห็นว่าถึงแม้ความคาดหวังในประสิทธิภาพ ความคาดหวังในความพยายาม อิทธิพลทางสังคม สภาพสิ่งอำนวยความสะดวก แรงจูงใจด้านความบันเทิง มูลค่าราคา และความเคยชินในการใช้ จะไม่ส่งผลทางตรงแต่กลับส่งผลทางอ้อมต่อพฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH

5. บทสรุปและการอภิปราย

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษารูปแบบความสัมพันธ์เชิงสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH ของประชากรในเขตกรุงเทพมหานครและปริมณฑล และเพื่อตรวจสอบความสอดคล้องของรูปแบบความสัมพันธ์เชิงสาเหตุที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ โดยมมีขนาดของกลุ่มตัวอย่างโดยการวิเคราะห์รูปแบบความสัมพันธ์โครงสร้างเชิงสาเหตุ ได้ขนาดของกลุ่มตัวอย่างจำนวน 400 ตัวอย่าง วิธีการเผยแพร่แบบสอบถามออนไลน์ (Questionnaires Online) ที่มีประสบการณ์การใช้งานแอปพลิเคชั่น 7-Eleven TH มีผู้ตอบแบบสอบถามมีจำนวน 450 คน หลังจากนั้นผู้วิจัยได้ทำการคัดเลือกแบบสอบถามที่มีความสมบูรณ์ได้จำนวนผู้ตอบ

แบบสอบถาม 413 คน มาทำการตรวจสอบความถูกต้องของข้อมูลที่จะนำไปวิเคราะห์ข้อมูลทางสถิติต่อไป

เครื่องมือที่ใช้ในการวิจัย ลักษณะของเครื่องมือวิจัยในครั้งนี้เป็นแบบสอบถามจำนวน 1 ชุด จำนวน 53 ข้อ โดยแบบสอบถามที่นำมาใช้ในงานวิจัยนี้ได้มาจากการศึกษาเอกสารงานวิจัย รวมถึงพัฒนาแบบสอบถามมาจากเอกสารงานวิจัยที่เกี่ยวข้องเพื่อสรุปได้ถึงด้านต่าง ๆ ที่เป็นสาเหตุของการยอมรับและการใช้งานแอปพลิเคชั่น 7-Eleven TH ที่จะทำการศึกษา โดยได้แบ่งออกเป็น 3 ตอน ได้แก่ ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ได้แก่ ประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH ที่อยู่ เพศ อายุ พฤติกรรมการใช้อินเทอร์เน็ตบนโทรศัพท์มือถือ รวมถึงประสบการณ์การใช้งานแอปพลิเคชัน 7-Eleven TH จำนวนทั้งสิ้น 8 ข้อ ตอนที่ 2 ความคิดเห็นด้านปัจจัยของทฤษฎีการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการใช้แอปพลิเคชัน 7-Eleven TH จำนวน 7 ด้าน ได้แก่ ด้านความคาดหวังในประสิทธิภาพ ด้านความคาดหวังในความพยายามใช้ ด้านอิทธิพลของสังคมต่อการใช้งาน ด้านสภาพสิ่งอำนวยความสะดวกในการใช้งาน แรงจูงใจด้านความบันเทิงในการใช้งาน มูลค่าราคาในการใช้งาน ความเคยชินในการใช้งาน โดยเป็นแบบมาตรประมาณค่า 7 ระดับ จำนวน 35 ข้อ และ ด้านทัศนคติต่อการใช้งานแอปพลิเคชัน 7-Eleven TH จำนวน 2 ด้าน ได้แก่ ด้านพฤติกรรมความตั้งใจในการใช้งาน และด้านพฤติกรรมการใช้งาน โดยคำตอบเป็นแบบมาตราส่วนประมาณค่า 7 ระดับ จำนวน 10 ข้อ โดยกำหนดระดับพฤติกรรมออกเป็น 7 ระดับ จากระดับน้อยสุดที่ 1 ถึงระดับมากที่สุดที่ 7 จากนั้นจึงได้นำแบบสอบถามที่ผ่านการพัฒนาและปรับปรุงแล้วให้ผู้เชี่ยวชาญ ตรวจสอบความตรงด้านเนื้อหาและตรวจสอบความถูกต้องเหมาะสมของเนื้อหา ภาษาที่ใช้ พบว่าข้อคำถามทั้งหมดมีค่าความสอดคล้องของข้อคำถามตรงกับวัตถุประสงค์ของการวิจัย จากนั้นจึงนำมาตรวจสอบความน่าเชื่อถือด้วยวิธีการหาค่าสัมประสิทธิ์แอลฟา โดยค่าความเที่ยงที่ได้มีค่าเท่ากับ 0.98 แสดงให้เห็นว่าข้อคำถามในแบบสอบถามนั้นมีความน่าเชื่อถือในระดับสูง

ปัจจัยที่มีส่งผลต่อการวิจัย พบว่าปัจจัยที่มีอิทธิพลต่อพฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH ประกอบด้วย 7 ปัจจัยหลัก ได้แก่ 1. ความคาดหวังในประสิทธิภาพ (Performance Expectancy) หมายถึง ความเชื่อของแต่ละบุคคลว่าแอปพลิเคชัน 7-Eleven TH มีประโยชน์ต่อการรับรู้ข้อมูลข่าวสาร โปรโมชั่น และสิทธิพิเศษอื่นๆ ช่วยให้เกิดความสะดวกต่อการเข้าไปซื้อสินค้า/บริการในร้านสะดวกซื้อ 7-Eleven ช่วยให้เกิดความรวดเร็วต่อการเข้าไปซื้อสินค้า/บริการในร้านสะดวกซื้อ 7-Eleven และ เหมาะสำหรับใช้ในชีวิตประจำวัน 2. ความคาดหวังในความพยายาม (Effort Expectancy) หมายถึง แอปพลิเคชัน 7-Eleven TH เป็นระบบที่เข้าใจได้ง่าย เป็นระบบที่ไม่ซับซ้อน สามารถใช้งานผ่านแอปพลิเคชัน 7-Eleven TH ได้อย่างถูกต้อง สามารถเรียนรู้การใช้งานได้ง่าย และการเพิ่มทักษะความชำนาญในการใช้แอปพลิเคชัน 7-Eleven TH ทำได้ง่าย 3. อิทธิพลทางสังคม (Social Influence) หมายถึง การรับรู้ของแต่ละบุคคลว่ากลุ่มบุคคลที่มีความสำคัญต่อบุคคลได้ให้ความคาดหวัง หรือ เชื่อว่าแต่ละบุคคลควรใช้เทคโนโลยีสารสนเทศใหม่ เช่น เพื่อน ครอบครัว สื่อโฆษณามีผลทำให้ท่านใช้แอปพลิเคชัน 7-Eleven TH รวมถึงการใช้แอปพลิเคชัน 7-Eleven TH ช่วยให้ท่านก้าวทันเทคโนโลยี และช่วยเสริมสร้างภาพลักษณ์ว่าเป็นคนทันสมัยและทันต่อเทคโนโลยี 4. สภาพสิ่งอำนวยความสะดวกในการใช้งาน (Facilitating Conditions) หมายถึง ความเชื่อ ของแต่ละบุคคลว่าโครงสร้างพื้นฐานที่มี จะช่วยส่งเสริมหรืออำนวยความสะดวกให้เกิดการใช้งาน แอปพลิเคชัน 7-Eleven TH ได้ เช่น มีอุปกรณ์เชื่อมต่อการใช้งานแอปพลิเคชัน 7-Eleven TH ได้ตลอดเวลา มีความรู้ทางเทคโนโลยีมากพอที่จะใช้ ความเร็วของระบบอินเทอร์เน็ตที่ท่านใช้มีผลต่อการใช้ และการชำระค่าสินค้าและบริการผ่านแอปพลิเคชัน 7-Eleven TH ง่ายกว่าการใช้เงินสด 5. แรงจูงใจด้านความบันเทิง (Hedonic Motivation) หมายถึง ความเชื่อของแต่ละบุคคลว่าความสนุกหรือความพึงพอใจที่ได้รับจากการใช้งานแอปพลิเคชัน 7-Eleven TH ได้ เช่น รู้สึกเพลิดเพลิน พึงพอใจ รู้สึกสบายตา สบายใจ รู้สึกดีทุกครั้งที่ได้สะสมหรือแลกคะแนนผ่านแอปพลิเคชัน 7-Eleven TH และรู้สึกพึงพอใจที่ได้ใช้คูปองส่วนลด 6. มูลค่าราคา (Price Value) หมายถึง ความเชื่อของแต่ละบุคคลว่าการใช้งานแอปพลิเคชัน 7-Eleven TH ทำให้ผู้ใช้ได้รับความรู้และทักษะการคิดเปรียบเทียบของผู้บริโภคที่จะได้รับ และส่วนลดหรือสิทธิพิเศษ ได้แก่ มีคูปองส่วนลดจากแอปพลิเคชัน 7-Eleven TH ช่วยประหยัดค่าใช้จ่าย การชำระค่าสินค้า/บริการผ่านแอปพลิเคชัน 7-Eleven TH ช่วยให้ท่านได้รับคะแนนสะสมหรือสิทธิพิเศษอื่นๆเพิ่มขึ้น การใช้แอปพลิเคชัน 7-Eleven TH ช่วยให้ท่านสามารถซื้อสินค้า/บริการในร้านสะดวกซื้อ 7-Eleven ได้ในราคาพิเศษ และ

รายการส่งเสริมการขายในแอปพลิเคชัน 7-Eleven TH ช่วยประหยัดค่าใช้จ่าย 7. ความเคยชิน (Habit) หมายถึง การที่บุคคลมีแนวโน้มที่จะแสดงพฤติกรรมการใช้งานแอปพลิเคชัน 7-Eleven TH โดยอัตโนมัติ เพราะสืบเนื่องจากสิ่งที่เรียนรู้มาในอดีตที่เคยปฏิบัติอย่างสม่ำเสมอจนกลายเป็นความเคยชิน ได้แก่ ใช้แอปพลิเคชัน 7-Eleven TH ในการรับข้อมูล ข่าวสาร รายการส่งเสริมการขายของร้านสะดวกซื้อ 7-Eleven เป็นประจำอย่างต่อเนื่อง ใช้แอปพลิเคชัน 7-Eleven TH ในการสะสมหรือแลกคะแนนเป็นประจำเมื่อซื้อสินค้า/บริการร้าน 7-Eleven ใช้แอปพลิเคชัน 7-Eleven TH ตรวจสอบรายการส่งเสริมการขายของร้านสะดวกซื้อ 7-Eleven เป็นประจำ และใช้แอปพลิเคชัน 7-Eleven TH เพื่อตรวจสอบคะแนนสะสมเป็นประจำ

จากการวิเคราะห์ค่าความสัมพันธ์ระหว่างตัวแปรสังเกตได้ที่ใช้ในโมเดลรูปแบบความสัมพันธ์เชิงสาเหตุ พบว่าปัจจัยที่ส่งผลต่อพฤติกรรมความตั้งใจในการใช้งานแอปพลิเคชัน 7-Eleven TH ประกอบไปด้วยความเคยชินในการใช้งาน รองลงมาคือ แรงจูงใจด้านความบันเทิง สภาพสิ่งอำนวยความสะดวกในการใช้งาน มูลค่าราคา ความคาดหวังในประสิทธิภาพ อิทธิพลทางสังคม และความคาดหวังในความพยายาม ตามลำดับ ส่วนปัจจัยที่ส่งผลต่อพฤติกรรมในการใช้งานแอปพลิเคชัน 7-Eleven TH ประกอบด้วย ความเคยชินในการใช้งาน รองลงมาคือ ความตั้งใจในการใช้งาน แรงจูงใจด้านความบันเทิง สภาพสิ่งอำนวยความสะดวกในการใช้งาน มูลค่าราคา ความคาดหวังในประสิทธิภาพ อิทธิพลทางสังคม และความคาดหวังในความพยายาม ตามลำดับ

เอกสารอ้างอิง

[1] Venkatesh, V., Thong, J. Y. L., & Xu, X., “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology,” MIS Quarterly, Vol.36, No.1, 2012, pp. 157-178.

[2] สิงหะ ฉวีสุข และ สุนันทา วงศ์จตุรภัทร, “ทฤษฎีการยอมรับการใช้เทคโนโลยีสารสนเทศ,” วิทยานิพนธ์ปริญญามหาบัณฑิต, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ, 2555.

[3] วอนชนก ไชยสุนทร, “ปัจจัยที่ส่งผลต่อการยอมรับการใช้บริการพร้อมเพย์ของประชาชนในเขตกรุงเทพมหานครและปริมณฑล,” ทุนสนับสนุนงานวิจัย ประจำปี 2558, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2558.

[4] เกวรินทร์ ละเอียดดีนันท์ และนิตนา ฐานิตธนกร, “การยอมรับเทคโนโลยและพฤติกรรมผู้บริโภคออนไลน์ที่มีผลต่อการตัดสนใจซื้อหนังสืออิเล็กทรอนิกส์ของผู้บริโภคในเขตกรุงเทพมหานคร,” มหาวิทยาลัยกรุงเทพ, 2559.

[5] ชวิศา พุ่มดนตรี, “ปัจจัยที่ส่งผลต่อการยอมรับการใช้บริการพร้อมเพย์ของประชาชนในเขตกรุงเทพมหานครและปริมณฑล,” วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์, 2559.

[6] พรชนก พลาบูลย์, “การยอมรับนวัตกรรมและเทคโนโลยีการใช้เทคโนโลยีและพฤติกรรมผู้บริโภคที่ส่งผลต่อความตั้งใจของประชาชนในการใช้บริการธุรกรรมทางการเงินผ่านระบบพร้อมเพย์ของรัฐบาลไทย,” วิทยานิพนธ์มหาบัณฑิต, มหาวิทยาลัยกรุงเทพ, 2560.

[7] ฉัตรเฉลิม เกิดสวัสดิ์, “รูปแบบความสัมพันธ์เชิงสาเหตุของการยอมรับและการใช้เทคโนโลยีที่มีผลต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านโซเชียลมีเดียของข้าราชการโรงเรียนนายร้อยพระจุลจอมเกล้า,” วิทยานิพนธ์มหาบัณฑิต, มหาวิทยาลัยรังสิต, 2560.

[8] พิมพ์พรรณ สุวรรณศิริศิลป์, “ปัจจัยที่ส่งผลต่อการยอมรับและใช้งานบริการแบบพร้อมเพย์,” วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์, 2559.

[9] ยงยุทธ บุญกิจ, “รูปแบบความสัมพันธ์เชิงสาเหตุการยอมรับและการใช้เทคโนโลยีที่มีต่อพฤติกรรมการสื่อสารภายในองค์กรผ่านสื่อสังคมออนไลน์ของบุคลากรสายสนับสนุนในมหาวิทยาลัยธรรมศาสตร์,” วิทยานิพนธ์มหาบัณฑิต, มหาวิทยาลัยรังสิต, 2562.

[10] กัลยา วานิชย์บัญชา, “การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย,” กรุงเทพฯ: พิมพ์ครั้งที่ 6, คณะพาณิชยศาสตร์ฯ จุฬาลงกรณ์มหาวิทยาลัย, 2545.

[11] ธานินทร์ ศิลป์จารุ, “การวิจัยและวิเคราะห์ข้อมูลทางสถิติด้วย SPSS,” กรุงเทพฯ : พิมพ์ครั้งที่ 11, บิสซิเนสอาร์แอนด์ดี, 2555.

[12] กัลยา วานิชย์บัญชา, “หลักสถิติ,” กรุงเทพฯ : พิมพ์ครั้งที่ 8, ธรรมสาร, 2549.

[13] กัลยา วานิชย์บัญชา, “สถิติสำหรับงานวิจัย,” กรุงเทพฯ: พิมพ์ครั้งที่ 6, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2555.

[14] Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D., “User Acceptance of Information Technology: Toward a Unified View,” MIS Quarterly, Vol.27, No.3, 2003, pp. 425–478.

[15] กริช แรงสูงเนิน, “การวิเคราะห์ปัจจัยด้วย SPSS และ AMOS เพื่อการวิจัย,” กรุงเทพฯ: ซีเอ็ดยูเคชั่น, 2554.

JIST Journal of Information Science and Technology
Volume 10, NO 1 | JAN – JUN 2020 | 98-113

Ontology-based Knowledge Management System with a Case Study on The East-West Economic Corridor

ปัทมา เจริญพร

Pattama Charoenporn

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Department of Computer Science, Faculty of Science,

King Mongkut's Institute of Technology Ladkrabang

Received: May 30, 2020; **Revised:** June 18, 2020; **Accepted:** June 20, 2020; **Published:** June 23, 2020

ABSTRACT – This research aims to develop a knowledge management system with an ontology case study of the East-West Economic Corridor. The operation began with data collection from a sample group using purposive sampling. The sample group consisted of 400 people living in Mukdahan, Kalasin and Khon Kaen province, Thailand with a good knowledge of tourism. The data used to create a database for tourism ontology is derived from tacit and explicit data, divided into 3 Domain classes. The domain class consists of 4 levels of class hierarchy. The relationship between the ontology is divided into 3 types. The knowledge creation process uses the SECI Model and the creation of knowledge data using Taxonomy relationship to establish relationship with the knowledge management system database. The knowledge management system created by the research can be used on both web application and mobile application. In system testing, there are two methods. One is to test the functions by using the User Acceptance Test and test the data recall by finding the F-measure level. The F-measure is 87% and the average time of the data recall is 0.11 seconds. The system is considered to have a good search performance. As for the results of the User Acceptance Test, the results passed all the criteria. The final step is to test the user satisfaction with the QUIS questionnaire. The result of the satisfaction assessment that is over level 5 in all items makes it possible to conclude that the developed knowledge management system provides satisfactory results to the users of the system. In the future, there may be additions to improve knowledge in other areas or added prediction of tourists' interest in choosing tourist destinations to cover more of the development of knowledge management systems.

KEYWORDS: Knowledge management processes, Ontology, Questionnaire for user interaction-satisfaction, The SECI Model, Tourism, The East-West Economic Corridor

Corresponding Author: pattama.ch@kmitl.ac.th

1. Introduction

Given the current situation in which the World Economy is staggering, inevitably affecting Thailand's economy as a whole. The most important thing that helps keep the economy from falling further is Thailand's emphasis on the Tourism Industry which helped stimulate the entire nation's economy. From the Tourism Authority of Thailand [1] which has assigned an economic objective to gain 10 percent growth of tourism profits from 2019 in 2020. From the World Economic Forum's report [2], there has been a ranking on the competitive potential of tourism businesses worldwide. The ranking of 2019 has ranked Thailand on 31st out of 140 countries. When compared to only ASEAN countries, Thailand is ranked 3rd, following Malaysia and Singapore. Therefore, supporting the tourism industry is a vital component for continuous development of Thailand's potential in the tourism industry's international competition.
The East-West Economic Corridor is a project to develop travel routes to connect with neighboring countries such as Vietnam, Laos, Myanmar, and Cambodia, as illustrated in Figure 1.

Figure 1. East-West Economic Corridor [3].

However, the entrepreneur confidence index survey on tourism businesses in 2019's fourth quarter [4] revealed that entrepreneurs feel greatly worried about the industry's situation in said quarter since overall profits have declined from what it was normally. The cause was assumed to be the reduced number of tourists due to several recent crises throughout the world. Therefore, to help the tourism industry handle the ever-changing worldwide context, there should be more technologies to lend a hand, increasing the potential for global competition. Knowledge creation from rural communities to urban communities becomes a way to broadcast Thailand's tourism information worldwide.

In this research, the researcher has developed an Ontology-based Knowledge Management System with a Case Study on The East-West Economic Corridor to serve as a place for exchanging and searching for knowledge about tourism, giving birth to new knowledge points through technology. In this research, a case study about economic development along the East-West Economic Corridor has been cited. The corridor consists of 3 provinces: Muk Dahan, Kalasin, and Khon Kaen. The research also utilizes the SECI Knowledge creation process, linking that knowledge to each other with ontology. After completing Development, the system operation will be assessed and all comments and instructions will be compiled to help plan Knowledge Management regarding other issues to be complete, according to each area's specific situations and the business sector's demands.

2. Background and Related Works

2.1 Knowledge Management System

The knowledge management system is a process to gather knowledge from people or documents and develop them as a system, giving organizations access to the knowledge, allowing them to educate themselves and work efficiently. The SECI Model has 4 main components [5] as the following:
1.Socialization is the process of creating and transferring one's personal knowledge to another person (Tacit to Tacit). The method of sharing one's knowledge to another is based on experience and must take into account factors of appropriate time and environment, such as information obtained from creating a discussion board system to exchange knowledge between experts and the general public.
2.Externalization is a collection of specific knowledge that is transmitted through experience in the social exchange process, transformed into explicit knowledge, which is written in the desired format (Tacit to explicit), such as the collection of documents from research or recordings, creating a research storage system related to tourism, crafting a database system that experts can enter to add or edit information.
3. Combination is the application of explicit knowledge that is gathered from both inside and outside the organization to select and synthesize it into new knowledge (Explicit to New Explicit) to increase the complexity, such as creating a system for analyzing tourism information from tourism behavior of Thai and foreign tourists.
4. Internalization is a result of hands-on learning until it becomes the specific skill and expertise of the person which becomes a new unique knowledge (Explicit to New Tacit), resulting in practices for

organizations to apply and ultimately become a culture, resulting in the process of creating knowledge and entering the social exchange process. (Socialization), such as the results of measuring the number of visits to the knowledge management system, the results of relevant video traffic statistics, and the results of statistics collected on tourist sites that users have searched for, etc.

For this research, the SECI model is used as a knowledge creation process for the data collection process. The process of data collection can be described as follows:

Step 1; Socialization uses sharing and creating Tacit knowledge from communication between each other. For this research, data collection from questionnaires is used to collect tacit knowledge. Step 2; Externalization process Creating and sharing knowledge from tacit knowledge and explicit knowledge. For this research, using the method of information dissemination on the web board to display the knowledge for the general public After that, we will go to step 3 which is the Combination method, further studying the teaching techniques from many places, then summarize and disseminate it as a new teaching technique. Which is caused by the collection of knowledge from various sources and one's own knowledge. Then step 4 is Internationalization. It is a part of studying other techniques in addition to textbooks. or existing books, then applying it to real work until the skills and expertise become their own tacit knowledge, and when knowledge comes, it exchanges with other people. Next, a process called Socialization itself is a continuous process.

2.2 Concept of Ontology

Ontology [6] refers to concepts that are applied for knowledge management and the creation of groups of meanings within the scope of interest in a hierarchical style using the Web Ontology Language (OWL) to build the infrastructure. According to research by Grigoris Antoniou and Frank van Harmelen. (2004) [7] found that the Web Ontology Language (OWL) was created by the W3C (Web Ontology Working Group) organization and was developed as an extension of the Resource Description Framework: RDF. OWL was developed in the second version on 11 December 2012 [8] and has more syntax added to describe complex semantic information. And to be able to define the data structure and explain the data (Metadata) that is related in the database system better. The description is in the form of a class. Class properties and class relationships explain the relationships that occur. Examples of applying ontology is, for example; the system of Search Engines for finding and gaining access to the information that users need by allowing the system to search for information from the ontology knowledge database so that the system can find related words with words that have the same meaning as the related words, etc.

For this research, the ontology creation method for database creation is used to group the hierarchical data consistently. The ontology work process is divided into 6 steps: 1. Document collection, 2. Word omission, 3. Word frequency count, 4. Keyword selection, 5. Creating occurrent frequency table, 6. Present the ontology in the form of Taxonomy relationship.

2.3 The Questionnaire for User Interaction Satisfaction (QUIS 7.0)

The Questionnaire for User Interaction Satisfaction 7.0 [9] is a user satisfaction survey using open-ended questions. This questionnaire is used to assess satisfaction and comments of users towards the user interface. The questionnaire consists of 5 parts: Screen, Terminology and System Information, Learning, System Capabilities, Usability and User Interface. The Likert Scale is used to measure satisfaction in 9 levels from 0, which is "hard to read" up to 9, which is "easy to read".

2.4 Related Works

1. S. Mouhim's research (2010) [10] presents knowledge management concepts and guidelines to building knowledge management systems, semantic webs, and ontologies. The data collection process, the research utilized data collection from internet users searching about tourism in Morocco. Then using the information to create relationships in the form of ontology. Data collection test subjects are chosen from various website services, such as tourism portals, websites, blogs, existing ontologies, vocabulary, and then studying patterns. Of tourist interest and Moroccan culture. Then, searching terms that are diverse and complex are created. Finally, a presentation on tourism knowledge management for Morocco and the architecture of KMS knowledge management systems for managing tourism information in Morocco. In the future, the researcher recommends investigating the complexity of the ontology by allowing experts to examine and improve their ability to respond more accurately and efficiently which may need inquiries from travelers'

international tourism data analysis and the development of research ontologies so that they can be used with other existing ontologies.

2. Prantner and others's research (2007) [11] presents the initial analysis of the OnTourism project with the objectives of (1) creating and evaluating Semantic web technologies such as creating ontologies, semantic enhancement of content and semantic searches to informationally and economically important tourism domains. (2) Creating an ontology to identify the meaning of words and structure of the development of the tourism industry with similar definitions. And (3) presenting the concept of the tourism industry in Austria with the OnTourism work model created for OnTourism. The ontology will be created by presenting the Tourism Domain with the ontology creation tools in the presentation section. OnTourism uses Web 2.0 technology and folksonomy for optimizing the search of terms in the Travel Domain.

From the above research, it is found that the ontology building for relation of definitions can be applied to knowledge management system smoothly. But there should be an increase to the process of data analysis from experts in each area to get information for creating ontology that is repetitive and suitable for in-depth search.

In terms of studying the scope of the research area, the researcher has selected the research area from the study of literature boundary, literature review to determine the area for determining the scope of information in which the researcher has conducted studies from various related research by showing details as follows

3. Phusrisom's (2004) [12] on benefits about economic growth, the research shows that Thailand will be affected by the economic stimulus from the construction of The Second Thai–Lao Friendship Bridge (Mukdahan-Savannakhet), which was found that the provinces in the northeast along the route (East-West Economic Corridor: EWEC) is Kalasin, Khon Kaen and Mukdahan, and they will benefit from the increased rate of economic growth since there are both domestic and foreign investors's increase in investment. Being the center of agricultural industry production and developing to be a center for human resource and human resource development in the Greater Mekong Sub-region. Convenient transportation will also be developed to help efficiently solve the poverty problem of the people in the northeast region along the EWEC route.

4. Apichatvullop, Y (2007) [13] has conducted research with educational objectives regarding definitions and value on the economic terrace as perceived by each stakeholder group and to cope with route changes. The data collected from sample groups which was selected from a group of stakeholders on the EWEC route, including organizations in business operators and residents of the EWEC route area across the study. It was found that the organizations within areas that will be affected by the EWEC route include education agencies, health agency, tourism agency, Local Administration Organization, social development, transportation, and communication. For those who operate private businesses, they will be involved in an industrial unit Services for wholesale, retail, and agro-industry businesses As for the people who live in the area, it was found that business operators in the area that the EWEC route passes including 3 provinces in the north-eastern region of Khon Kaen, Kalasin and Mukdahan had an opinion and expected that the EWEC route would have a positive impact on the businesses.

From the above research, it was found that although some research findings suggest that it is expected that the EWEC route will have a positive impact on trade of various businesses, especially Khon Kaen, Kalasin and Mukdahan which will benefit from the increased economic growth rate. Due to being a neighboring province to The Second Thai–Lao Friendship Bridge (Mukdahan-Savannakhet), as well as both domestic and foreign investors are expected to invest more, it can be seen that most of the research focuses on studying the effects in Areas such as transportation, International trade, but no research has yet been studied on the impact of tourism economy. Therefore, supporting tourism through information technology, focusing on 3 provinces in the northeast, Khon Kaen, Kalasin and Mukdahan, to be model provinces for building a knowledge management system on the EWEC route that should be studied and fulfilled in terms of knowledge Information and knowledge for academic purposes and benefits in terms of bringing research results to organizational development and related agencies.

3. METHODS

There are 5 procedures in the development process. 1. Data collection, 2. Designing and Developing the Ontology, 3. Building the Knowledge Management System, 4. Knowledge Management System Testing, 5. Evaluate user satisfaction towards the Knowledge Management System. The details to each procedure are as follows:

3.1 Data Collection

This step is to collect data from 400 participants chosen by purposive sampling [14] from the

provinces in the northeast region along the East-West Economic Corridor.: EWEC), namely Khon Kaen, Kalasin and Mukdahan. The total of 400 samples must live in the research area and they must also have knowledge in tourism as well and are able to use the information to create knowledge for creating an ontology.

The class ontology is based on the reference of the Entrepreneurs Sentiment Index Quarter 1/2020 [15]. Tourism class is divided into three areas, which can be used to create relationships in Domains and scopes which can be as follows:

1.Domain of Historical tourism

2.Domain of Natural attractions

3.Domain of Tourist attractions in arts, culture, traditions and activities, and methods of exchanging data from sample groups are to be compared with the data from the Folksonomies method and the web board obtained from the system created by the research with the steps in data collection as follows;

1. Studying information about tourism industry along the East-West Economic Corridor from the sample area of 3 provinces together by studying from documents, related research, brochures, notes, and books to collect various details for the data cleaning process.

2. Acquiring information from research, consisting of data survey about the tourism industry, using questionnaires which will consist of both open and closed-end questions. Interviews and exchanges will also happen between the researcher and the sample group, referring to SECI model for data collection, with details shown on Figure 2.

From the figure 2, it can be explained that.

Step 1: Knowledge creation; is to collect data from the sample group. The information will be in the

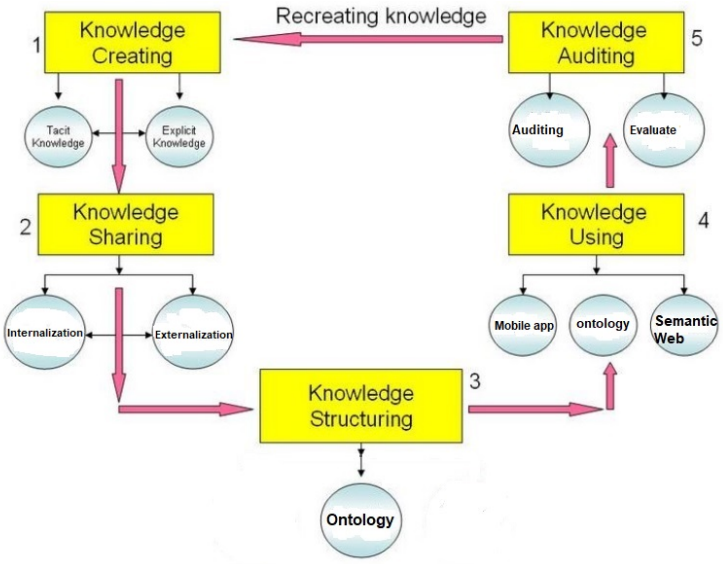


Figure 2. Data Analysis Method.

form of Tacit Knowledge, including data from brainstorming of respondents about community tourist attractions, historical sites, natural attractions, arts and cultural sites, traditions and activities, individual experience information about local culture. By dividing into each scope of the research area, then analyzing for comparison with external data using Folksonomies and the web board of the research. With methods to categorize by content, allowing users to share opinions based on various categories. Once the information is received, the Tacit knowledge will then be developed into Explicit knowledge.

Step 2: Knowledge Sharing. This step is to design explicit knowledge in the form of a structure that can be distributed by dividing into 2 structures:

1. Internalization: The structure is to format the knowledge that is available in the organization in a standard format, such as creating websites, documents, records

2. Externalization: It is the process of integrating tacit knowledge obtained from the extraction of knowledge with a link between internal knowledge and external knowledge.

From the study, it was found that Tourism in various forms in all 3 provinces, using external knowledge gained, such as research results, customer needs and modern technology, applied and combined with prior knowledge to improve the means of tourism to meet the needs of tourists even more. When designing explicit knowledge of tourism information, granting 3 sub nodes of

tourism in the following provinces: tourist attractions, accommodation, restaurants, accommodation types and tourist activities

Step 3. Creating a Knowledge Structure, Bringing the whole structural information from both Internalization and Externalization to design relationships between structural data using ontology methods.

Step 4. Knowledge Using; is the process of designing the ontology database and designing the knowledge management system into a semantic web format to create a mobile application for the knowledge management system that the research created. Ontology design is explained in part 4.

Step 5. Knowledge auditing is a process for auditing and validating content from knowledge management systems that are created which the system will present tourism knowledge at various structural levels by organizing it into a knowledge base, so that those who want to search for knowledge from the knowledge management system can study for information and create knowledge in the form of tacit knowledge. For the newly created information, such as from a blog or from a web board that has users filling in information, the system will lead to the next knowledge creating process.

The process model for the process of transforming tacit knowledge into explicit knowledge can be described in Figure 3.

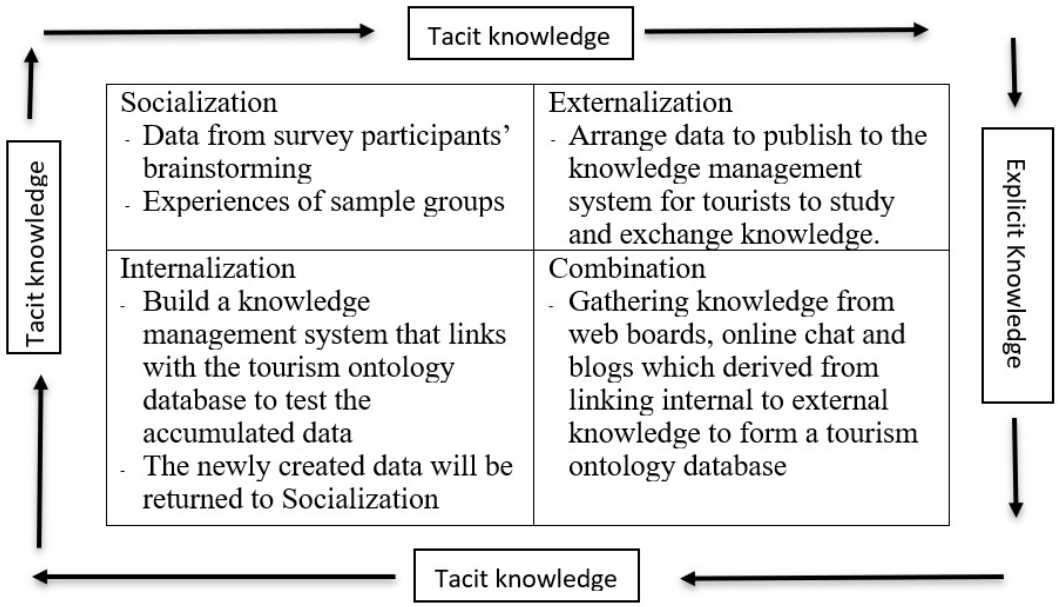


Figure 3. Transforming tacit knowledge into explicit knowledge.

3.2 Designing and Developing the Ontology

This step is finding the data's relationship by Taxonomy Relationship, done by creating the ontology via combination which divides 70% of the data to build a model and the other 30% is going to be used as test data for the model in development. After that, the relationship will be built via the Protégé app 5.5 [16] to find the relationship of tourism groups which are classified according to content types to build the ontology. There are 3 steps to the development:

1. Building a thesaurus using text files to contain vocabularies and search for information from the sample group to be used as database for related vocabularies used to compare with words that the user types into the system, then design the ontology knowledge base which consists of the main classes such as attraction information, archaeology, images, saved as OWL files.

2. Definitive data recall has 3 working steps; 1. Keyword Selection, for example: choosing from tourism groups, adding conditions for complex searches 2. Word group selection for data return 3. Recall data from knowledge database.

3. Test data recall from the ontology knowledge database by measuring the F-measure value, measured from the relationship value between Precision and Recall values. The equation for finding F-measure value [17] is shown in Equation 1.

$$F-measure = \frac{2 \times precision \times recall}{precision + recall} \times 100 \quad (1)$$

The Precision value is the value which measures the ability to rule out all irrelevant documents.
The Recall value is the value that evaluates the system's efficiency in pulling out relevant documents and gauges the Processing Time [18] is shown in Equation 2, which is the time elapsed from each recall to find the average recall time.

$$Processing\ Time = \sum_{i=1}^{N} elapsedTimeMin/N \quad (2)$$

elapsedTimeMin is the time elapsed in each retrieve relevant documents return
N is the number of retrieve relevant documents returns.

3.3 Building the Knowledge Management System

This part is the design of the Knowledge Management System's infrastructure, divided into 2 sections as shown in Figure 4.

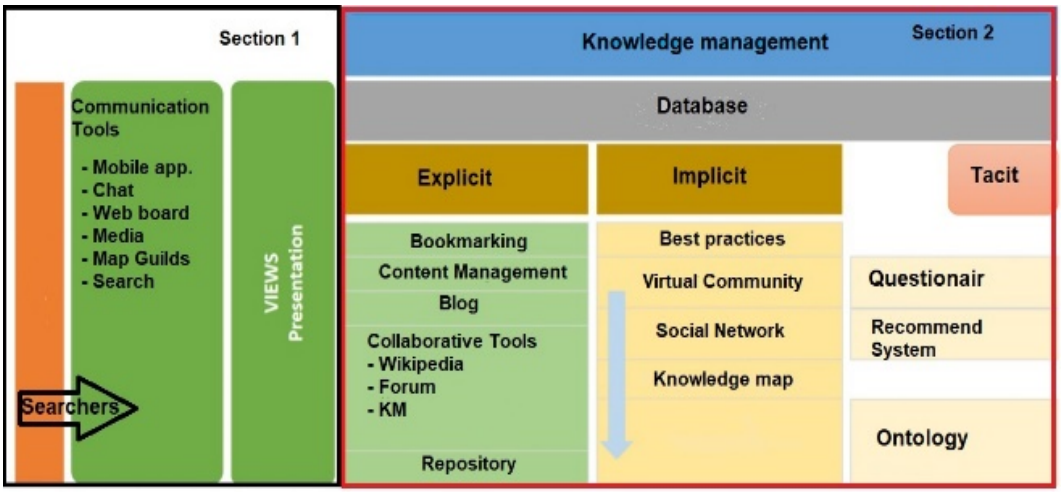


Figure 4. Infrastructure of the Knowledge Management System.

In figure 4, Section 1 is the user (searcher) interface. From here, users of both the web app and mobile app who wishes to look for information from the Knowledge Management System, for example, find routes to an attraction, which can be done by searching for information from the web board. The system's functional are as shown in Table 1.

Table 1. The system's functional

No	Functional Requirement
1	Several topics of knowledge in the database. For example, searching from the attraction name, coordinates or popularity.
2	Online chat with tourism experts
3	Web boards for each topic, allowing interested people to write posts and exchange knowledge.

Table 1. The system's functional (Continued)

No	Functional Requirement
4	Central Database for storing various types of data such as articles, photos, videos, including 3-D photos.
5	Map Guide Planning which can assign destinations from searches.
6.	Newsletters from mobile application

Section 2 is about the Knowledge Management System. This section is used to store and analyze data from building the ontology knowledge database. The system functional are as illustrated in Table 2.

Table 2. The knowledge management's functional

NO.	functional Requirement
1	A system to analyze and classify Explicit / Implicit / Tacit knowledge to build credibility for each knowledge type.
2	Access Authority Check
3	Voting and rating for knowledge from users and tourism experts.

3.4 Knowledge Management System Testing

Testing of the system will be carried out under real System Environment by testing the system parallel to outside sources such as the web service accompanying the system. The method of this test is User Acceptance Test will be used to test usage data before real use.

3.5 Evaluate User Satisfaction towards the Knowledge Management System.

This step is to assess user satisfaction by the QUIS 7.0 questionnaire, used by 400 people from the sample group with system design knowledge. Concurrent "Thinking aloud" system is used to test and assess satisfaction statistically.

4. Results

4.1 Results from data collection

After gathering data for building the knowledge management system, the next step will be building the ontology with details as follows:
1. Creating a dictionary. The researcher has created a dictionary of terms with various definitions of information to be used as a vocabulary database for use in comparison with search terms that users enter. The vocabulary in the dictionary is divided into 6 categories, namely tourism information, tourism type, tourism business, trading, support from various agencies and tourism development.
2. The concept design of the tourism ontology database in combination style consisting of Domain and scope in 3 aspects as follows:
1. Historical tourism domain
2. Natural attractions domain
3. Tourist attractions in arts, culture, traditions and activities domain
As for the super node, it is divided into types of each domain as follows.
1. Historical tourism domain consisting of 3 super classes which are 1. Province class 2. Historical sites class 3. Antiques class
Sub node class of Historical tourism domain consists of 1. Province class divided into Khon Kaen, Mukdahan, Kalasin 2. Historic Site class, which is divided into Religious, Historical Park and City Wall, Monument, Ancient Community 3. Antiques class divided into naturally occurring, artificial. Details are shown as figure 5.

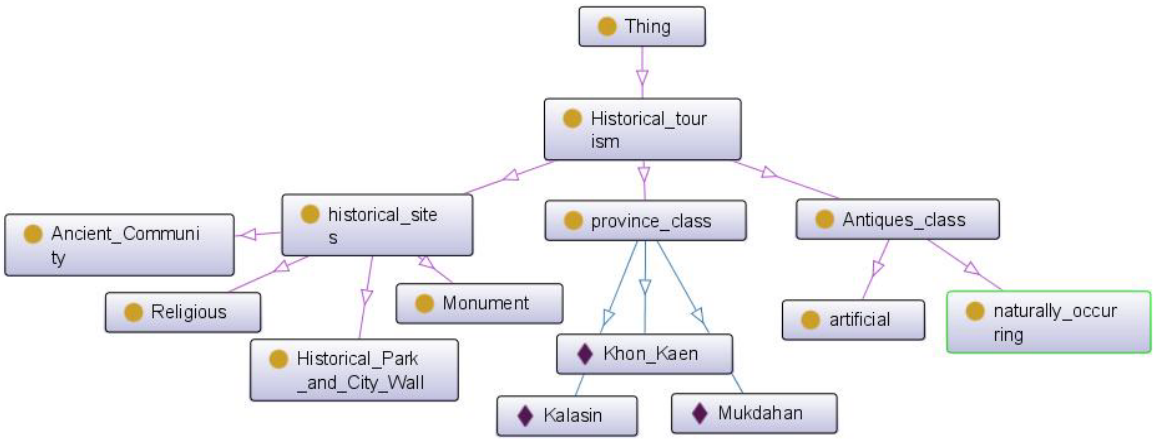


Figure 5. Historical tourism domain.

2. Natural attractions domain consists of 3 super classes which are 1. Province class 2. Naturally occurring class and Man-made attraction class
Sub node class of the Natural attractions domain consists of 1. Province class divided into Khon Kaen, Mukdahan, and Kalasin. 2. Naturally-occurring class are divided into Mountain, Forest, Waterfall, Beach, Cataracts 3. Man-made attraction class in divided into park, dam, reservoir with details shown figure 6.

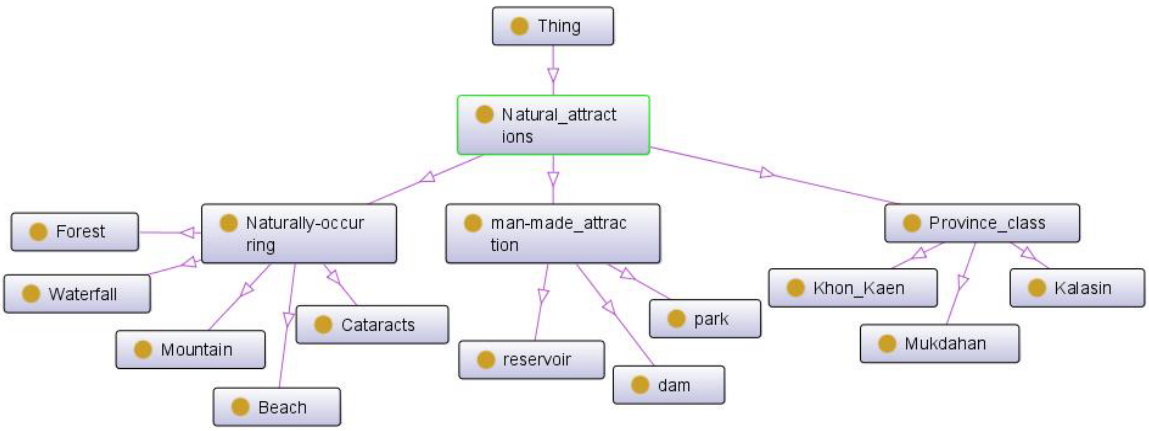


Figure 6. Natural attractions domain.

3. Tourist attractions in arts, culture, traditions and activities domain consists of 6 super classes; 1. Province class 2. Rural lifestyle tourism class 3. Recreational attraction class 4. Arts and culture class, 5. Traditional class, and 6. Activities class.
Sub node class of Tourist attractions in arts, culture, traditions and activities domain consists of 1. Provinces class, divided into Khon Kaen, Muk Daharn and Kalasin. 2. Rural lifestyle tourism class is divided into Village and Community, Cultural Center, Markets and Farms. 3. Recreational attraction class is divided into parks and stadiums 4. Art and culture class divided into ways of life 5. Traditions class divided into Traditional Festivals. 6.Activities class are divided into cultural exhibitions and welcome participants. Details shown in figure 7.

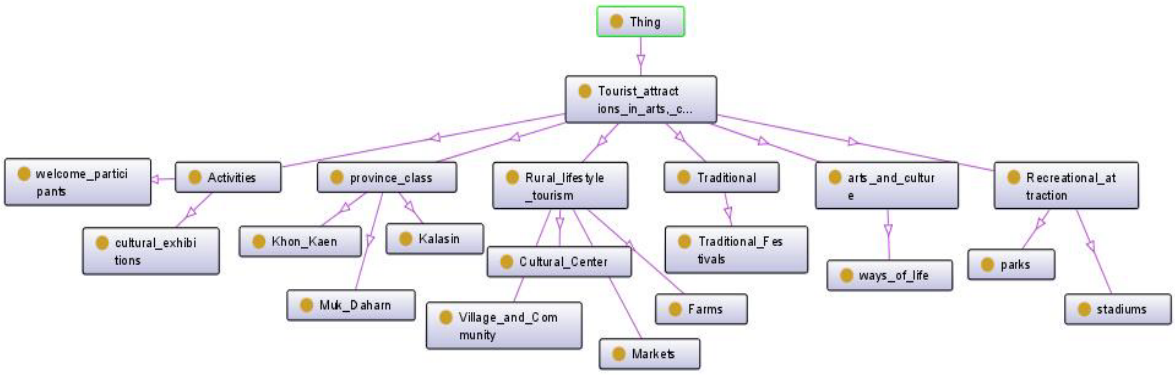


Figure 7. Natural attractions domain.

The final of sub node class level 4 is organized into every sub class, divided into 7 classes which are 1. accommodation, 2. festival, 3. district, 4. route, 5. subdistrict, 6. restaurant, 7. shops and souvenirs, respectively. Details shown in figure 8.

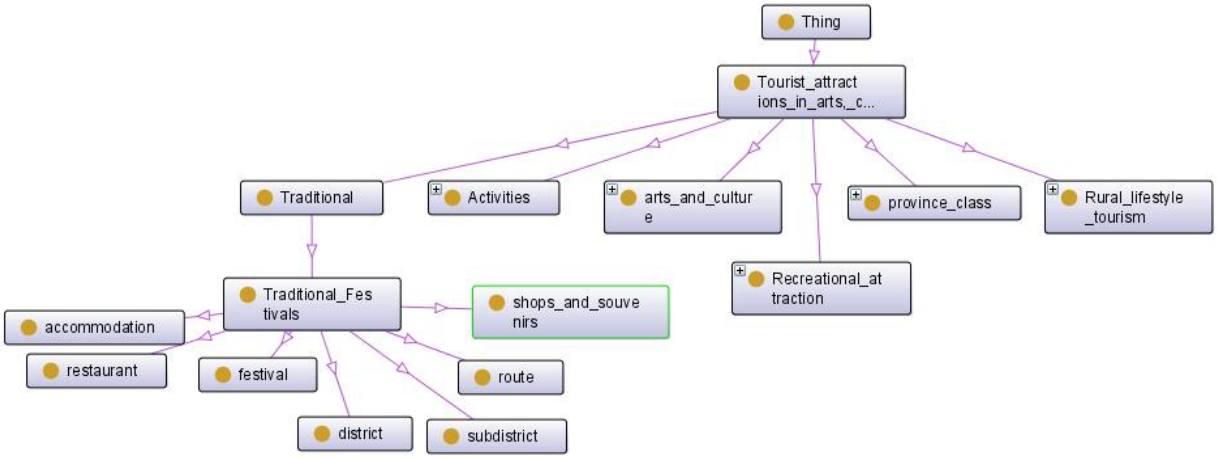


Figure 8. The final sub node class level 4.

The relationship between ontology is divided into 3 types as follows:

- Is - a refers to a relationship classified as
- Part-of (p / o) means “is a component of”
- Attribute-of (a / o) means “the attribute of”

The design of the tourism ontology database consists of 3 Domain classes, 4 levels of class Hierarchy and the relationship between ontology divided into 3 types. The researcher describe the data in the OWL (Ontology Web Langrage) and connected the database of the knowledge management system to the tourism ontology database using the Ontology Application Management program (OAM) Framework [19] to enable the knowledge management system that the research created to find the definition of terms according to the research objectives.

For example, the relationship between tourism ontology database and knowledge management system database can be seen in the table 3.

Table 3. Relationship between tourism ontology database and knowledge management system

Tourism ontology database	Relation	Knowledge management system database
Accommodation Class	a/o	Accommodation_Type Acc_ID Acc_Name
District Class	a/o	District _Type Dist_ID District _Name
Festival Class	a/o	Festival_Type Fest_ID Fest_Name
Route class	a/o	Route_Type Route_ID Route_Name

4.2 Results of analysis data

In the efficiency test by assessing the precision of data searching, the researcher has assigned basic conditions regarding the data and search words by randomizing data from the database and having 400 users recall data with 50 different search words, 50 times each. From the experiment, it can be seen that the precision has 87% F-measure value and average processing time of 0.11 seconds since the ontology’s infrastructure is built to be the combination time, which is very repetitive, resulting in longer search times but yields better results than building the ontology with “Top down, Bottom UP” method.

By allowing users to search for tourist attractions, activity types and choose the relationship as a component of Khon Kaen province. The program will extract the information from the tourism ontology database and display the data as in the figure 9(a) searching page and figure 9 (b) the detail of searching page.

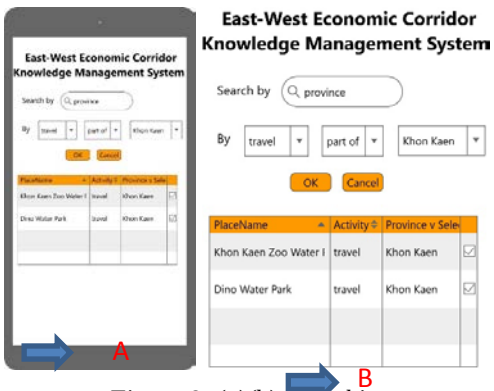


Figure 9. (a)(b) Searching page.

4.3 Designing Knowledge Management System with Ontology

In the design of knowledge management systems, we can describe the system in 2 parts.
1.Explaining the process of knowledge management system using a Flowchart diagram as shown in Figure 10.
2. Explaining the operation procedures of system users with the Use case diagram as shown in Figure 11. The details are as follows:
From Figure 10, the knowledge management system can be described as follows: When the system receives data from the data admin or data science, the system will be group Explicit and Tacit data by considering data recall from the management system database. Then, the data will be stored in the store data. When the user enters the keyword into the system, the system will retrieve the data from the store data to create data modeling. At this stage, the system will check the accuracy of finding words from the tourism ontology database which has set the search accuracy to be more than 80% to pass the validation process. In the process of data management, Tacit data will start from Expert user filling in the system in the web board section. Then, the tacit data is stored in the tourism ontology database and will be connected to the knowledge management system database to be used for creating a database of research knowledge management systems. As for the explicit data from the user, gained from filling in the information on the web board and online chat. The data will enter the Content Management System database so that the system can help manage the content and data related to the key words that need to be stored. After that, all information will be sent to link to the tourism ontology database.

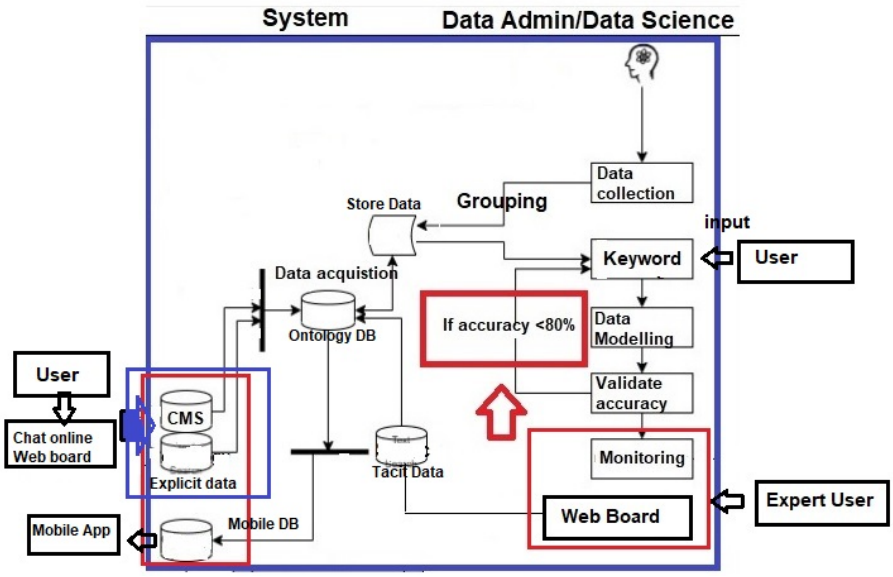


Figure 10. Flow of Knowledge management system.

In Figure 11 illustrates the use case diagram, consisting of users which is broken down into 4 parts, user, expert user, data admin and data science. The user's duty is to find knowledge and information and communicate online with tourism experts and provide web boards for tourism-related topics for interested users to post in and exchange. Data admin and Data science will take care of central database and storage according the ontology pattern. Lastly, the expert user adds information and answers the questions in the web board. They also keep track and add new knowledge by processing them under the SECI Model.

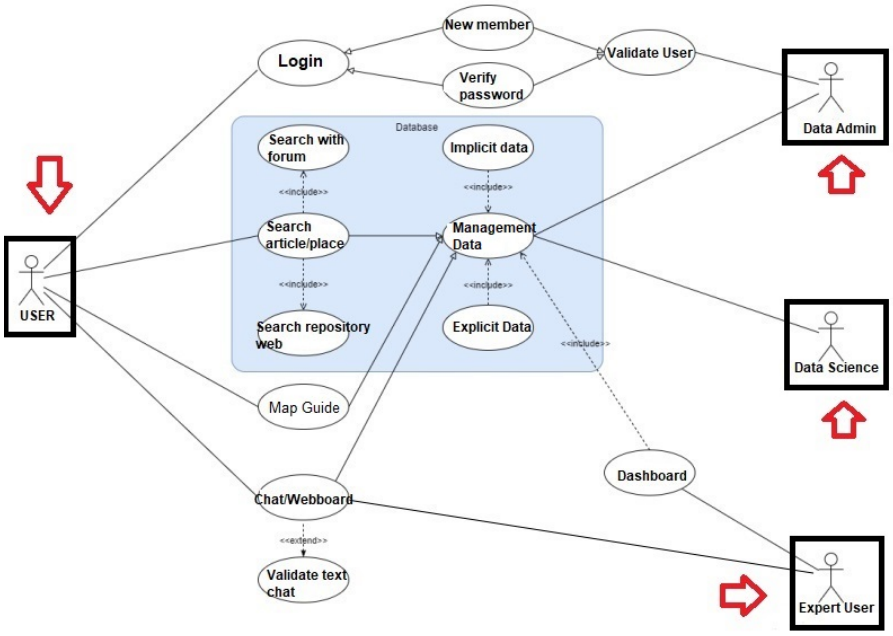


Figure 11. Use case diagram

4.4 Knowledge Management System Design

The designed Knowledge Management System can be used on both web apps and mobile apps, with the login screen and main menu shown in Figure 12(a) and 12(b)

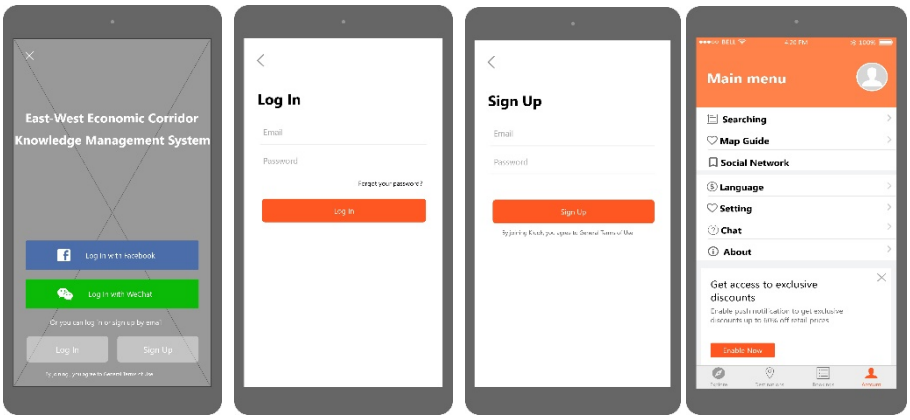


Figure 12(a) Login and Main page

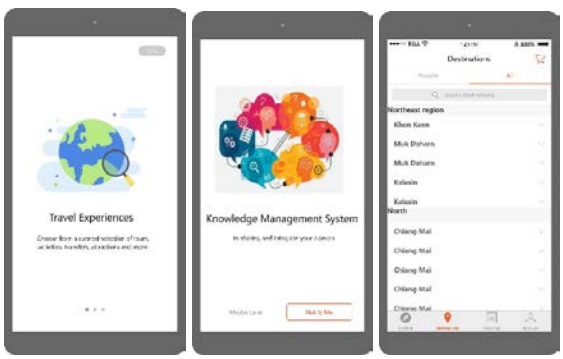


Figure 12(b) Writing Information and searching

Figure 12(a)(b) Illustrates the mobile app's user interface with 7 pages: main page, Login page, sign in page, create travel experiences, sharing information, searching destination

4.5 Knowledge Management System Test Results

As for the testing process, the testing utilizes the User Acceptance Test, with a total of 400 system testers showing the test percentage. By referring to the method from the Descriptive Statistics [20]. The test values can be shown in Table 4.

Table 4. Showing test results from user acceptance

Method	Requirement	UAT Checklist	Results	
			% pass	% Not pass
1	Provides a search system which covers a variety of knowledge, such as attraction names, coordinates or popularity	- Can search from a variety of information, such as attraction info, coordinates or popularity - Cannot Input Numerical information	95%	5%
2	Provides online chat via messaging for communication with tourism experts	- Capable of group and personal online chat - Has a system for checking online chat users	100%	0
3	Provides a web board for interested people to post and exchange knowledge	- Capable of posting, editing, adding to the posts and comments - Capable of verifying users in the comment section	100%	0

Table 4. Showing test results from user acceptance (Continued)

Method	Requirement	UAT Checklist	Results	
			% pass	% Not pass
4	Includes a central data storage system to store different types of data including data, articles, pictures, videos, and 3-d pictures.	- Articles, pictures, videos, and 3-d pictures can be added, edited or deleted.	100%	0
5	Includes Map Guide Planning that can pinpoint the location from the search	- Able to connect to Google maps to find travel routes	100%	0
6	Includes newsletters for interesting news on mobile application	- Users can send and receive news on mobile applications.	100%	0
7	Includes a knowledge analysis system to classify Explicit and Tacit data	- Able to classify users and expert users when user login in the system - Present username in the top of page	100%	0
8	Includes a system to check for usage rights	- Able to log in and check for usage.	100%	0
9	Includes a system to vote and rate knowledge for users and tourism experts	- Data classification can be rated on web boards.	100%	0

From the table, you can see that A total of 400 system testers tested the system from 9 functional requirements of the system. There were 1 item that 95% of the participants answered because some testers did not understand the process of filling the data into the search menu due to the complicated

process. Therefore, the researcher has redesigned the user interface by presenting the user interface as in section 4.2. As for the remaining test results, all system testers can test out every function.

4.6 Evaluating User Satisfaction Towards the Knowledge Management System with QUIS 7.0

User Satisfaction Assessment with QUIS 7.0 from 400 users are illustrated in Table 5. It presents in percentage.

Table 5. Results of satisfaction assessment with Quis 7.0

Question / Scale	1	2	3	4	5	6	7	8	9
Q1 Screen							5	95	
Q2 Terminology and system information							6	94	
Q3 Learning								15	85
Q4 System Capabilities								5	95
Q5 Usability and User interface								5	95

From Table 5 shows the average result of each question. Out of 400 respondents, the results show that the Q1 questions about the Screen, 95%of the users answered “easy to read” level 8 and 5% on level 7. Question Q2 on Terminology and system information, 94% of the users’ answers are consistent at level 8, leaving Level 7 at 6%. Q3 on Learning, 15% of the users answered at easy-Level 8 and 85% for Level 9. Q4 questions about System Capabilities; 95% of users answered Reliable level 9, 5% Level 8. Questions Q5 Usability and User interface, 95% of users answered good level 9, 5% level 8, respectively, based on the average of each question. It shows that the result of the satisfaction assessment over level 5 in all items makes it possible to conclude that the developed knowledge management system provides test results that are satisfactory to the system users.

5. Summary, Suggestions and Discussion

5.1 Summary

In this research, the objective is to develop a knowledge management system using an ontology method with a case study on the tourism industry based on the EWEC economy corridor. By using an ontology database management method, users can search for tourist information, Meeting the needs for the data used in creating the knowledge management model, namely tourist information, divided into 3 domains: 1. Historical tourism domain 2. Natural attractions domain3. Tourist attractions in arts, culture, traditions and activities domain. The work process of knowledge management system is divided into 2 parts which are (1) creating a database of the knowledge management system by importing data from users and expert users (2) the creation of tourism ontology database which is the part that creates the relationship between data created in the form of ontology, then analyzing for data accuracy from the users’ search. In this section is the use of tourist information separated by various criteria obtained from experts and tourists that have used the system. The results of the knowledge management system analysis concluded that when the knowledge management system was tested with 400 system users, data retrieval was performed by specifying different search terms. The test found that the F-measure level is 87% and the average time of data recall is 0.11 seconds. The system test using the User Acceptance Test gives the total test result value at 100%. With the QUIS 7.0 questionnaire, the result of the satisfaction assessment is over level 5 in all items thus making it possible to conclude that the developed knowledge management system provides test results that are satisfactory to the test users.

5.2 Suggestions and Discussion

Developing a knowledge management system using an ontology in the next research comes with suggestions as follows.

1. Adding factors that affect the selection of tourist attractions of both domestic and foreign tourists to help the system more clearly identify the various types of tourism needs based on personal characteristics.
2. More study should be conducted on the data for tourist groups forecasts based on personal characteristics to better predict interest in selecting destinations and supporting the tourism industry.
3. Additional vocabulary from Thesauruses should be studied in the tourism domain of the World

Tourism Organization (WTO) [21] for the e-library that has collected synonyms, autonyms, and other word forms

4. More study should be conducted on tourism behavior of tourists that changes according to time to add more functionality to the system.

6. ACKNOWLEDGMENT

This research is funded by the faculty of science, King Mongkut's Institute of Technology Ladkrabang. Thank you for your guidance to all companies and government organization in Mukdahan, Kalasin and Khon Kaen province, Thailand who provided participants for this study.

References

[1] TAT News, "TAT target 10% growth in 2020 with quality markets and responsible tourism after turning 60 in 2020," TAT News, 2019. [Online]. Available: https://www.tatreviewmagazine.com/article/tourism-direction-2020/. [Accessed: January 16, 2020].

[2] TAT News, "The Travel and Tourism Competitiveness Report 2019," TAT News, 2019. [Online]. Available: http://reports.weforum.org/travel-and-tourism-ompetitiveness-report-2019/. [Accessed: December 10, 2019].

[3] HKTDC Research, "A Practical Guide to Doing Business in Thailand," HKTDC Research, 2017. [Online]. Available: https://economists-pick-research.hktdc.com/en/. [Accessed: December 5, 2019].

[4] TAT News, "Thailand tourism confidence index 4/2562," TAT News, 2019. [Online]. Available: http://www.thailandtourismcouncil.org/wp-content/uploads/2020/01/Newsletter-Q4.2562.pdf. [Accessed: December 10, 2019].

[5] Nonaka, L., Takeuchi, H., & Umemoto, K., "A Theory of Organizational Knowledge Creation," International Journal of Technology Management, vol. 11, pp. 833-845, 1996.

[6] Staab, S., & Studer, R. (Eds.), "Handbook on ontologies," Springer Science & Business Media, DOI:10.1007/978-3-540-92673-3, 2010.

[7] Parsons, Simon, "The Semantic Web Primer by Grigoris Antoniou and Frank van Harmelen," MIT Press, 238 pages, ISBN. The Knowledge Engineering Review, 2004, pp. 278-279. DOI: 10.1017/S0269888905230205.

[8] W3C OWL Working Group, "OWL 2 Web Ontology Language Document Overview (Second Edition)," W3C OWL Working Group, 2012. [Online]. Available: https://www.w3.org/TR/owl2-overview/. [Accessed: October 16, 2019].

[9] Moumane, Karima & Idri, Ali & Abran, Alain, "Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards," Springer Plus, 5 pages, Doi:10.1186/s40064-016-2171-z, 2016. [Online]. Available: https://www.researchgate.net/publication/301720411_Usability_evaluation_of_mobile_applications_using_ISO_9241_and_ISO_25062standards. [Accessed: October 10, 2019].

[10] S. Mouhim; A. El Aoufi; C. Cherkaoui; El. Megder; D. Mammass, "Towards a knowledge management system for tourism based on the semantic web technology," In 2011 IEEE International Conference on Multimedia Computing and Systems, 2011, pp. 1-6.

[11] Prantner, K., Ding, Y., Luger, M., Yan, Z., & Herzog, C., "Tourism ontology and semantic management system: state-of-the-arts analysis," IADIS International Conference, 2007. [Online]. Available: www.iadis.net/dl/final_uploads/200712C070.pdf. [Accessed: October 10, 2019.]

[12] Phusrisom , K."การวิเคราะห์ผลประโยชน์ที่คาดว่าประเทศไทยจะได้รับ จากการก่อสร้างสะพานข้ามแม่น้ำโขง ไทย – ลาว แห่งที่ 2," Independent Study, Faculty of Economics Khon Kaen University, Khon Kaen, Thailand, ISBN : 9746595326, 2547.

[13] Apichatvullop, Y., "Contest Meanings and Expectations of the East-West Economic Corridor in Thailand," Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand, 2007.

[14] Saunders, M., Lewis, P. & Thornhill, A., "Research Methods for Business Students," 6th edition, Pearson Education Limited, pp. 288, 2012.

[15] Tourism council of Thailand, "Tourism Situation And Forecast Of Tourists' Behavior and Confidence index of Thailand's Tourism Industry," Tourism council of Thailand, 2020. [Online]. Available: http://www.thailandtourismcouncil.org/. [Accessed: January10, 2020].

[16] Tudorache, Tania & Vendetti, Jennifer & Noy, Natasha, "Web-Protege: A Lightweight OWL Ontology Editor for the Web," Proceedings of the Fifth OWLED Workshop on OWL: Experiences and Directions, collocated with the 7th International Semantic Web Conference (ISWC-2008), Karlsruhe, Germany, October 26-27, 2008.
[17] Y. Sasaki, "The truth of the f-measure. 2007," 2007. [Online]. Available: https: //www.toyota-ti.ac.jp/Lab/Denshi/COIN/people/yutaka.sasaki/F-measure-YS-26Oct07.pdf. [Accessed: October 01, 2019].
[18] T. Kulthida, J. Adam, R. Edie, "The construction of an ontology using rms for tourism," The Emergence of Digital Libraries -- Research and Practices: 16th International Conference on Asia-Pacific Digital Libraries, ICADL 2014, Chiang Mai, Thailand, November 5-7, 2014, Proceedings. Lecture Notes in Computer Science 8839, Springer 2014, ISBN 978-3-319-12822-1.
[19] M. Buranarach, Y. Myat and T. Supnithi, "A CommunityDriven Approach to Development of an Ontology-Based Application Management Framework," Proceeding of the Second Joint International Conference (JIST 2012), 2012, pp.306-312.
[20] Yamane, T., "Statistics An Introductory Analysis," 2nd Edition, New York. Harper & Row, 1967.
[21] World Tourism Organization, "UNWTO Elibrary," World Tourism Organization, 2019. [Online]. Available: https://www.e-unwto.org/. [Accessed: October 10,2019].

CONTRIBUTORS' FORM

(to be modified as applicable and one signed copy attached with the manuscript)

Manuscript Title:

_____

Author(s):

_____

I/we certify that I/we have participated sufficiently in the intellectual content, conception and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript, to take public responsibility for it and have agreed to have my/our name listed as a contributor. I/we believe the manuscript represents valid work. Neither this manuscript nor one with substantially similar content under my/our authorship has been published or is being considered for publication elsewhere, except as described in the covering letter. I/we certify that all the data collected during the study is presented in this manuscript and no data from the study has been or will be published separately. I/we attest that, if requested by the editors, I/we will provide the data/information or will cooperate fully in obtaining and providing the data/information on which the manuscript is based, for examination by the editors or their assignees. Financial interests, direct or indirect, that exist or may be perceived to exist for individual contributors in connection with the content of this paper have been disclosed in the cover letter. Sources of outside support of the project are named in the cover letter.

I/We hereby transfer(s), assign(s), or otherwise convey(s) all copyright ownership, including any and all rights incidental thereto, exclusively to the Journal, in the event that such work is published by the Journal. The Journal shall own the work, including 1) copyright; 2) the right to grant permission to republish the article in whole or in part, with or without fee; 3) the right to produce preprints or reprints and translate into languages other than English for sale or free distribution; and 4) the right to republish the work in a collection of articles in any other mechanical or electronic format.

We give the rights to the corresponding author to make necessary changes as per the request of the journal, do the rest of the correspondence on our behalf and he/she will act as the guarantor for the manuscript on our behalf.

All persons who have made substantial contributions to the work reported in the manuscript, but who are not contributors, are named in the Acknowledgment and have given me/us their written permission to be named. If I/we do not include an Acknowledgment that means I/we have not received substantial contributions from non-contributors and no contributor has been omitted.

	Name	**Signature**	**Date signed**	
1				
2				
3				
4				(up to 4 contributors for case report/images/review)
5				
6				(up to 6 contributors for original studies)

INFORMATION FOR AUTHORS

The Journal of Information Science and Technology (JIST) aims to be the forum through which researchers, faculties, graduate students and experts of the computer and information technology and other technological related fields share and discuss their high quality research work as well as innovation. Original research articles, practical applications and innovations in the broad area of computer and information technology are suitable for publication in JIST. Periodicity (Publication) is 2 issues per year (JAN - JUN and JULY - DECEMBER).

Accepted papers will be Double-blind peer reviewed by minimum 3 reviewers with "Accept" result.

1. Soft Computing
2. Human-Computer Interaction
3 Information Assurance and Security
4. Information Systems
5. Networking
6. Programming
7. Platform Technologies
8. System Integration and Architecture
9. Social and Professional Issues
10. Web Systems and Technologies
11. Multidisciplinary
12. e-Business and m-Business
13. Business and Information System
14. Social and Business Aspects of Convergence IT and Ubiquitous Computing
15 Internet of Things (IoT)
16 Cloud Computing
17 Fintech and blockchain

Manuscript Preparation Guide

The template for writing the journal is available for downloading (http://www.ist-journal.org). Length of the paper should be in between 6-16 pages of A4 size. Text should be typewritten or printed with double-spaced in 11-point of A4 white paper with margins of 1.5" for top, 1.25" for left, and 1" for bottom and right sides. All pages must be numbered sequentially. Here are some guidelines.

- Abstract which length 100-200 words.
- Introduction which talks about the problem and related research.
- Materials and methods use for experimentation.
- Results of the experiment.
- Discussion and Conclusion, References

Submissions (Publication Charge: Free) (Indexed by TCI tier 2)

Please submit paper in MS Word via https://ph02.tci-thaijo.org/index.php/JIST/
(Online Journal Submission)

Contact Address

Faculty of Information Science and Technology
Mahanakorn University of Technology
140 Moo 1, Cheumsampan Road, Nongchok
Bangkok, Thailand 10530
Tel: 02-988-3655 ext 4115 Fax: 02-988-4027
E-mail: jist@mut.ac.th

สคทส. CITT
สมาคมสภาคณบดีคณะเทคโนโลยีสารสนเทศ
The Association of Council of Information Technology Deans of Thailand