Journal of Engineering and Digital Technology (JEDT)

Thai-Nichi Institute of Technology

Vol.11 No.2
July - December 2023

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล
สถาบันเทคโนโลยีไทย-ญี่ปุ่น
ปีที่ 11 ฉบับที่ 2 กรกฎาคม - ธันวาคม 2566

ISSN 2774-0617 (Online)

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล

Journal of Engineering and Digital Technology (JEDT)

ปีที่ 11 ฉบับที่ 2 กรกฎาคม – ธันวาคม 2566 Vol. 11 No. 2 July - December 2023

ISSN (Online) 2774–0617

ความเป็นมา

ด้วยสถาบันเทคโนโลยีไทย-ญี่ปุ่น มีนโยบายสนับสนุนการเผยแพร่บทความจากผลงานวิจัยที่มีคุณภาพ เพื่อเป็นประโยชน์ในการพัฒนาความรู้แก่สังคม โดยเฉพาะภาคธุรกิจและอุตสาหกรรม จึงได้จัดทำวารสารวิชาการ คือ วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล (เดิมชื่อ วารสารสถาบันเทคโนโลยีไทย-ญี่ปุ่น : วิศวกรรมศาสตร์และเทคโนโลยี)

วัตถุประสงค์

1. เพื่อส่งเสริมและเผยแพร่ผลงานวิจัยสาขาวิศวกรรมศาสตร์และเทคโนโลยีที่มีคุณภาพ
2. เพื่อเป็นสื่อกลางในการแลกเปลี่ยนองค์ความรู้ทางการวิจัยในสาขาวิศวกรรมศาสตร์และเทคโนโลยี
3. เพื่อพัฒนาศักยภาพทางการวิจัยสาขาวิศวกรรมศาสตร์และเทคโนโลยี

ที่ปรึกษา

รองศาสตราจารย์รังสรรค์ เลิศในสัตย์

ขอบเขตเนื้อหา

บทความวิจัยทางด้านวิทยาศาสตร์ วิศวกรรมศาสตร์ และเทคโนโลยี ได้แก่ เทคโนโลยีวิศวกรรม เทคโนโลยีอุตสาหกรรม เทคโนโลยีมัลติมีเดีย เทคโนโลยีสารสนเทศ วิทยาศาสตร์ประยุกต์ วิทยาศาสตร์กายภาพ วิทยาศาสตร์ชีวภาพ วิทยาการคอมพิวเตอร์ วิทยาศาสตร์เคมี และสาขาอื่น ๆ ที่เกี่ยวข้อง

กำหนดออกเผยแพร่

วารสารตีพิมพ์เผยแพร่ราย 6 เดือน (ปีละ 2 ฉบับ)

- ฉบับที่ 1 มกราคม – มิถุนายน
- ฉบับที่ 2 กรกฎาคม – ธันวาคม

เจ้าของ

สถาบันเทคโนโลยีไทย-ญี่ปุ่น (Thai-Nichi Institute of Technology : TNI)

1771/1 ถนนพัฒนาการ แขวงสวนหลวง เขตสวนหลวง กรุงเทพฯ 10250

1771/1 Pattanakarn Rd., Suanluang, Bangkok 10250

บรรณาธิการ

รองศาสตราจารย์ ดร.รัตติกร วรากูลศิริพันธุ์ สถาบันเทคโนโลยีไทย-ญี่ปุ่น

กองบรรณาธิการ

1. ศาสตราจารย์ ดร.ธนารักษ์ ธีระมั่นคง สถาบันเทคโนโลยีนานาชาติสิรินธร มหาวิทยาลัยธรรมศาสตร์
2. ศาสตราจารย์ ดร.โกสินทร์ จำนงไทย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
3. ศาสตราจารย์ ดร.บุญเจริญ ศิริเนาวกุล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
4. ศาสตราจารย์ ดร.อิสระชัย งามหรู สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
5. ศาสตราจารย์ ดร.วรพงศ์ ตั้งศรีรัตน์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
6. ศาสตราจารย์ ดร.ประยุทธ อัครเอกฒาลิน มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
7. ศาสตราจารย์ ดร.มานะ ศรียุทธศักดิ์ จุฬาลงกรณ์มหาวิทยาลัย
8. รองศาสตราจารย์ ดร.จีรนุช เสงี่ยมศักดิ์ มหาวิทยาลัยขอนแก่น
9. รองศาสตราจารย์ ดร.ปริทรรศน์ พันธุบรรยงก์ สถาบันการจัดการปัญญาภิวัฒน์
10. รองศาสตราจารย์ ดร.กันต์พงษ์ วรรัตน์ปัญญา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
11. รองศาสตราจารย์ ดร.พยุง มีสัจ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

คณะกรรมการดำเนินงาน

1. รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์
2. ผู้ช่วยศาสตราจารย์ ดร.รชตะ รุ่งตระกูลชัย
3. ผู้ช่วยศาสตราจารย์ ดร.พิมพ์เพชร สระทองอุ่น
4. ดร.ประมุข บุญเสี่ยง
5. ผู้ช่วยศาสตราจารย์ ดร.ธันยมัย เจียรกุล
6. ผู้ช่วยศาสตราจารย์ อมราวดี ทัพพุน
7. นางสาวสุพิศ บายคายคม
8. นายอคิรา พรหมสุทธิ์
9. นางสาวพิมพ์รต พิพัฒนกุล
10. นางสาวอาริสา จิระเวชถาวร
11. นางสาวอนันตา ปันเด

การติดต่อกองบรรณาธิการ

1771/1 ถนนพัฒนาการ แขวงสวนหลวง เขตสวนหลวง กรุงเทพฯ 10250

Tel. 0-2763-2600 (Ext. 2704, 2752, 2782)

Website : https://ph01.tci-thaijo.org/index.php/TNIJournal

E-mail : JEDT@tni.ac.th

บทความวิจัยที่ตีพิมพ์ในวารสารฉบับนี้เป็นความคิดเห็นส่วนตัวของผู้เขียน

กองบรรณาธิการไม่มีส่วนรับผิดชอบใด ๆ ถือเป็นความรับผิดชอบของผู้เขียนแต่เพียงผู้เดียว

บทบรรณาธิการ

ขอแสดงความยินดีกับนักวิจัยทุกท่านที่บทความได้ผ่านการประเมินจากผู้ทรงคุณวุฒิ และได้รับการตีพิมพ์ลงในวารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล (JEDT) เล่มปีที่ 11 ฉบับที่ 2 นี้ ทุกบทความมีเนื้อหาที่มีคุณภาพและมีประโยชน์อย่างยิ่งทั้งในด้านวิชาการและด้านวิจัย สามารถนำไปใช้ในการพัฒนาต่อยอดไปสู่นวัตกรรม

นโยบายของวารสารยังคงเดิมที่รับบทความที่เป็นทั้งภาษาไทย และภาษาอังกฤษ ภายใต้การรับรองของ TCI (Thai-Journal Citation Index Centre) จึงได้รับความสนใจจากนักวิจัยทั้งในประเทศและจากนานาชาติมากขึ้น

JEDT ได้รับการรับรองให้อยู่กลุ่ม TCI 1 ต่อเนื่องมาจนถึงปัจจุบัน โดยทางกองบรรณาธิการจะรักษามาตรฐานนี้ให้คงอยู่ตลอดไป ทางกองบรรณาธิการหวังว่า บทความที่ได้รับการตีพิมพ์ในวารสารฉบับนี้จะเป็นประโยชน์ต่อผู้อ่านทุกท่านในการอ้างอิงเพื่อการวิจัยและพัฒนาต่อไป

รองศาสตราจารย์ ดร.รัตติกร วรากูลศิริพันธุ์
บรรณาธิการ

Editorial Message

Congratulations to all researchers whose article(s) have passed the peer review and were published in the Journal of Engineering and Digital Technology (JEDT), Vol. 11 No. 2. All articles published in this issue have best quality contents and are extremely useful. They can be used for development of innovation.

The journal's policy remains the same, accepting articles in both Thai and English under the accreditation of TCI (Thai-Journal Citation Index Centre). Therefore, JEDT received more attention from both national and international researchers.

JEDT has been certified to be in the TCI 1 continuously until the present. The editorial team will always maintain this standard. The editorial team hope that all articles that were published in this journal will be beneficial to all readers for citation to their research and development.

Assoc.Prof. Dr.Ruttikorn Varakulsiripunth
Editor-in-Chief

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล

Journal of Engineering and Digital Technology (JEDT)

สารบัญ

1 **Biochar and Wood Vinegar Production from Pineapple Leaves Using Low-emission 200-liter Kiln and the Feasibility Study of Production**
Sasithorn Sunphorka, Wuttichai Yangsawang

14 **Classification of Rice Varieties from Milled Rice Grain Images by Object Detection Method**
Wuttichai Watchararat, Jessada Tanthanuch

25 **Developing a Credit Approval Determination Model Using Principal Component Analysis with Machine Learning Techniques**
Chonlada Muangthanang, Surasak Mungsing, Nivet Chirawichichai

35 **Economic Rebar Order Quantity Model for Construction Project**
Narongrit Wongwai, Thitiporn Chamcharas, Suphawut Malaikrisanachalee

48 **Foot Traffic Pattern Analysis of Vadhana District in the Spreading of COVID-19 Periods**
Yanisa Nualanant, Thongthit Chayakula

65 **Improvement of Harmonic Detection Using SDF for Shunt Active Power Filter in Single-phase Power Systems**
Sutthinee Waosungnern, Tosaporn Narongrit, Kongpol Areerak

76 **Loss and Efficiency Improvement of EV Traction Drive System by Using SiC in Inverter**
Kaninnat Rosdee, Surapong Suwankawin

87 **Physical and Biological Properties of Silk Suture Soaked in Levofloxacin**
Pawornwan Rittipakorn, Thanaporn Chanto, Pichamon Pumsanguan, Wimonwun Juntanasut, Chonipa Khunwittaya

96 **Sentiment Classification Based on Term Weighting with Class-mutual Information**

Uraiwan Buatoom, Kanyarat Cheawchan, Vajiraporn Sriput, Sombut Foithong

114 **The Prediction Model of PM2.5 in Phayao Province**

Khomkit Meksagul, Songwut Prakaiwichien, Athikom Boonsue

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

การผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรด ด้วยเตาเผามลพิษต่ำ ขนาด 200 ลิตร และการศึกษาความเป็นไปได้ในการผลิต

ศศิธร สรรพ่อค้า[1*,3] วุฒิชัย ยังสว่าง[2,3]

[1]สาขาวิชาวิศวกรรมเคมี คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก เขตพื้นที่อุเทนถวาย, กรุงเทพมหานคร, ประเทศไทย
[2]สาขาวิชาเทคโนโลยีโลจิสติกส์และการจัดการระบบขนส่ง คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก เขตพื้นที่อุเทนถวาย, กรุงเทพมหานคร, ประเทศไทย
[3]ศูนย์นวัตกรรมก่อสร้างและเทคโนโลยีวัสดุอุเทนถวาย คณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก เขตพื้นที่อุเทนถวาย, กรุงเทพมหานคร, ประเทศไทย

*ผู้ประพันธ์บรรณกิจ อีเมล : sasithorn_su@rmutto.ac.th

รับต้นฉบับ : 19 เมษายน 2566; รับบทความฉบับแก้ไข : 28 กรกฎาคม 2566; ตอบรับบทความ : 16 กันยายน 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเตาเผาจากถังเหล็กขนาด 200 ลิตร สำหรับผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรด ในพื้นที่ตำบลเขาคันทรง อำเภอศรีราชา จังหวัดชลบุรี โดยทำการทดลองเบื้องต้นในห้องปฏิบัติการเพื่อหาอุณหภูมิและเวลาที่เหมาะสมต่อการผลิตถ่านชีวภาพ จากการทดลองพบว่าผลได้ของแข็งสูงที่สุด คือ ร้อยละ 33.02 ที่อุณหภูมิ 400 องศาเซลเซียส เวลา 1 ชั่วโมง และมีค่าความร้อนสูงสุดเท่ากับ 23.25 เมกะจูลต่อกิโลกรัม จากนั้นทำการพัฒนาเตาเผาที่ประกอบด้วย เตาเผาเพื่อผลิตถ่านชีวภาพ ชุดท่อควบแน่นเพื่อเก็บน้ำส้มควันไม้ และเตาเผาไหม้เพื่อให้ความร้อนแก่เตาเผาส่วนแรก จากการทดลองพบว่าเตาเผาที่พัฒนาขึ้นสามารถผลิตถ่านชีวภาพจากใบสับปะรดและน้ำส้มควันไม้ได้ โดยใบสับปะรด 20 กิโลกรัม สามารถผลิตถ่านชีวภาพได้ประมาณ 2 กิโลกรัม และน้ำส้มควันไม้ประมาณ 10 กิโลกรัม จากการวิเคราะห์สมบัติของถ่านชีวภาพพบว่าคาร์บอนเป็นองค์ประกอบร้อยละ 70.49 มีค่าความร้อนสูงสุด 25.97 เมกะจูลต่อกิโลกรัม ในขณะที่น้ำส้มควันไม้มีองค์ประกอบส่วนใหญ่เป็นฟีนอล (Phenol) และอนุพันธ์ของฟีนอล (Phenol Derivatives) ซึ่งรวมถึงไกวเอคอล (Guaiacol) ที่มีฤทธิ์เป็นสารต้านเชื้อราและสารต้านเชื้อแบคทีเรีย สามารถนำไปใช้ในพื้นที่การเกษตรเพื่อกำจัดโรครากเน่าและกำจัดเชื้อราในพืช รวมถึงฉีดพ่นเพื่อฆ่าเชื้อโรคและศัตรูพืชบางชนิดได้ เมื่อวิเคราะห์ต้นทุนรวมถึงความเป็นไปได้ในการสร้างผลิตภัณฑ์และการนำไปใช้ พบว่าจำนวนครั้งในการดำเนินการเพื่อคืนทุนคือ 16 ครั้ง และทำให้เกิดผลิตภัณฑ์มูลค่าเพิ่มหลากหลายรูปแบบ

คำสำคัญ : วัสดุเหลือทิ้งทางการเกษตร ถ่านชีวภาพ เตาเผามลพิษต่ำ ใบสับปะรด น้ำส้มควันไม้

Biochar and Wood Vinegar Production from Pineapple Leaves Using Low-emission 200-liter Kiln and the Feasibility Study of Production

Sasithorn Sunphorka[1*,3] Wuttichai Yangsawang[2,3]

[1*]*Chemical Engineering Program, Faculty of Engineering and Architecture, Rajamangala University of Technology, Uthen Thawai Area, Bangkok, Thailand*
[2]*Logistics Management Program, Faculty of Engineering and Architecture, Rajamangala University of Technology, Uthen Thawai Area, Bangkok, Thailand*
[3]*Uthen Thawai Construction Innovation and Material Center, Faculty of Engineering and Architecture, Rajamangala University of Technology, Uthen Thawai Area, Bangkok, Thailand*

*Corresponding Author. E-mail address: sasithorn_su@rmutto.ac.th

Received: 19 April 2023; Revised: 28 July 2023; Accepted: 16 September 2023
Published online: 27 December 2023

Abstract

This research aimed to develop a 200-liter iron kiln to produce biochar and wood vinegar from pineapple leaves provided from Khao Khan Song sub-district, Sriracha district, Chon Buri province. The preliminary experiments were conducted in the laboratory scale to determine the optimum operating temperature and time for biochar production. The experimental results showed that the highest solid yield was 33.02 wt.%, obtained at 400 °C and for 1 h, with the highest heating value as 23.25 MJ/kg. The 200-liter iron kiln was then developed consisting of 3 parts: biochar production kiln, condensation apparatus for storing wood vinegar and combustion kiln for heat production. It was found that the developed 200-liter iron kiln could produce biochar and wood vinegar from pineapple leaves. The 20 kg of pineapple leaves can produce approximately 2 kg of biochar and 10 kg of wood vinegar. The analysis of biochar revealed that biochar had 70.49 wt.% of Carbon content with the highest heating value of 25.97 MJ/kg. The generated wood vinegar is mainly composed of Phenols and Phenol derivatives, including Guaiacol which has antifungal and antibacterial properties. The wood vinegar could be used in agricultural areas to eliminate root rot disease and plant fungi. It could also be sprayed in order to inhibit germs growth and some pests as well. The analysis of costs, the possibility of product creation and implementation presented that the number of payback operations was 16 and resulted in a wide variety of value-added products.

***Keywords*:** Agricultural waste, Biochar, Low-emission kiln, Pineapple leaves, Wood vinegar

1) บทนำ

จากการสำรวจปริมาณสับปะรดโรงงานในปี พ.ศ. 2561 โดยสำนักงานเศรษฐกิจการเกษตร [1] ประเทศไทยมีผลผลิตสับปะรด 2,350,887 ตัน ในจังหวัดชลบุรีคิดเป็น 132,534 ตัน ซึ่งพันธุ์ที่ปลูกอย่างแพร่หลายคือสับปะรดศรีราชาหรือสับปะรดพันธุ์ปัตตาเวีย หลังจากแปรรูปเป็นผลิตภัณฑ์ต่าง ๆ แล้วเกิดวัสดุเหลือทิ้งประเภทต่าง ๆ ได้แก่ ใบ จุก ต้น ก้านผล และหน่อ เป็นต้น สับปะรดหนึ่งผลเมื่อเข้าแปรรูปในโรงงาน จะมีเศษเหลือใช้ 1,228.1 กรัมต่อผล จากน้ำหนักผล 1,754.4 กรัมต่อผล เมื่อคิดเป็นปริมาณเปลือกเฉลี่ยต่อไร่จะเท่ากับ 2,700 กิโลกรัม [2] ซึ่งเป็นปริมาณมากและก่อให้เกิดปัญหาและค่าใช้จ่ายในการขนส่งและการกำจัด ค่าใช้จ่ายที่เพิ่มขึ้นดังกล่าวเป็นสาเหตุหนึ่งในการเพิ่มต้นทุนการผลิตด้วย การแปรรูปวัสดุเหลือทิ้งเหล่านี้จึงมีความน่าสนใจโดยเฉพาะการแปรรูปให้เป็นผลิตภัณฑ์ที่สามารถนำกลับมาใช้เพิ่มผลผลิตทางการเกษตร หรือผลิตภัณฑ์มูลค่าเพิ่มอื่น ๆ เพื่อสร้างรายได้และอาชีพใหม่ การแปรรูปวัสดุเหลือทิ้งทางการเกษตรให้เป็นถ่านชีวภาพเป็นทางเลือกหนึ่งที่ได้รับความสนใจเนื่องจากวิธีการไม่ซับซ้อนมากและได้ประโยชน์หลากหลาย เช่น ใช้เป็นเชื้อเพลิง ตัวดูดซับของเสียในน้ำ/อากาศ/ดิน ใช้เป็นปุ๋ย นำมาผสมอาหารและผลิตภัณฑ์เสริมความงาม รวมถึงการผลิตถ่านชีวภาพเพื่อให้ได้น้ำส้มควันไม้เป็นผลิตภัณฑ์พลอยได้ซึ่งมีฤทธิ์ป้องกันราและแมลงศัตรูพืช สามารถนำมาใช้ในพื้นที่เพาะปลูกเพื่อลดการใช้ยาฆ่าแมลงได้ เป็นต้น [3]–[6] นอกจากนี้การสร้างและใช้งานเตาเผามีความไม่ยุ่งยากสามารถทำได้เองในระดับชุมชน

หลายปีที่ผ่านมามีงานวิจัยเกี่ยวกับการพัฒนาเตาเผาถ่านชีวภาพที่สามารถผลิตถ่านและน้ำส้มควันไม้ได้ในระดับภาคสนามหรือสำหรับใช้ในชุมชน [3], [7]–[11] Moonsri and Hutem [3] ศึกษาการผลิตน้ำส้มควันไม้จากไม้มะขาม ไม้สะเดา ไม้กระถิน ไม้ไผ่ และไม้ยูคาลิปตัส ทำการออกแบบและสร้างชุดเตาเผาถ่านที่สร้างจากถังน้ำมัน 200 ลิตร และท่อดักน้ำส้มควันไม้ จากการทดลองพบว่าน้ำส้มควันไม้ที่ได้มีค่าพีเอชอยู่ในช่วง 2.5–3.4 และค่าความร้อนของถ่านประมาณ 13–18 เมกะจูลต่อกิโลกรัม Sataklang *et al.* [7] ทำการปรับปรุงเตาเผาเพื่อพัฒนาประสิทธิภาพเชิงความร้อนของเตาเผาถ่านแบบไร้ควันต้นแบบซึ่งสร้างจากถังเหล็กขนาด 200 ลิตร แบ่งส่วนประกอบเป็น 3 ส่วน คือ (1) เตาเชื้อเพลิง (2) เตาเผาถ่าน และ (3) ท่อนำแก๊ส ที่มีหน้าที่นำแก๊สที่เกิดขึ้นในกระบวนการคาร์บอไนเซชัน (carbonization) เข้าสู่เตาเชื้อเพลิงเพื่อเป็นเชื้อเพลิงเสริมในการผลิตถ่านส่วนแก๊สกลุ่มที่เป็นพิษจะถูกกำจัดด้วยการเผาในบริเวณเตาเชื้อเพลิง คณะผู้วิจัยยังได้เสนอการพัฒนาเตาเพื่อแก้ไขปัญหาการลุกลามของเปลวไฟออกบริเวณด้านข้างของเตา โดยติดตั้งผนังเตาด้านนอกเพื่อควบคุมเปลวไฟ และติดตั้งครีบนำความร้อนเพื่อเพิ่มอุณหภูมิภายในเตา ผลที่ได้คือเตาแบบใหม่สามารถประหยัดพลังงานและทำให้อุณหภูมิในเตาเผาสูงขึ้น Ketsripongsa *et al.* [8] พัฒนารูปแบบเตาเผาถ่านประสิทธิภาพสูงที่เหมาะสมของชุมชนรอบป่าโคกกลาง โดยใช้ถังเหล็กเจาะรูวางแบบแนวตั้งสามารถเคลื่อนย้ายได้สะดวก จากการทดลองพบว่า มีปริมาณควันจากปากปล่องด้านบนน้อยลงเนื่องจากเกิดการเผาไหม้ที่สมบูรณ์ ระยะเวลาที่ใช้ให้ความร้อนหน้าเตาเผาถ่านจนปิดเตาเผาถ่าน 120 นาที สามารถเผาถ่านไม้ได้ 3 ครั้งต่อวัน

จากงานวิจัยที่เกี่ยวข้องจะเห็นว่าเตาเผาที่พัฒนาจากถังเหล็ก 200 ลิตร สามารถใช้ในการผลิตถ่านชีวภาพในชุมชนได้และสามารถต่อเติมท่อนำแก๊สเพื่อเก็บน้ำส้มควันไม้เป็นผลิตภัณฑ์พลอยได้อีกทางหนึ่ง ซึ่งเตาเผาที่พัฒนาขึ้นมีลักษณะที่แตกต่างกันแล้วแต่ลักษณะของวัตถุดิบที่ใช้ จากการสำรวจความต้องการเบื้องต้นโดยสอบถามจากผู้นำเกษตรกรในพื้นที่ ตำบลเขาคันทรง อำเภอศรีราชา จังหวัดชลบุรี พบว่าเกษตรกรมีความต้องการในการนำวัสดุเหลือทิ้ง โดยเฉพาะใบสับปะรด ซึ่งเป็นส่วนเหลือทิ้งที่มีปริมาณมากที่สุดในพื้นที่เพาะปลูก โดยปกติแล้วในพื้นที่ปลูกสับปะรดจะมีใบสับปะรดเหลือทิ้งปริมาณสูงกว่า 4,000 กิโลกรัมต่อไร่ [12] มาแปรรูปหรือสร้างผลิตภัณฑ์มูลค่าเพิ่ม หรือกำจัดออกจากพื้นที่ และให้ความสนใจในการนำมาผลิตถ่านชีวภาพ รวมถึงการพัฒนาเตาเผาเพื่อให้สามารถแปรรูปวัสดุเหลือทิ้งทางการเกษตรอื่น ๆ ได้ด้วย คณะผู้วิจัยจึงมีแนวคิดในการนำใบสับปะรดมาแปรรูปเป็นถ่านชีวภาพ โดยแบ่งการศึกษาออกเป็น 2 ส่วน คือ (1) การศึกษาในระดับห้องปฏิบัติการ เป็นการนำใบสับปะรดแห้งมาผ่านกระบวนการไพโรไลซิส (pyrolysis) ในเตาเผาแบบ Muffle Furnace โดยออกแบบการทดลองแบบ central composite design (CCD) และวิธีการพื้นผิวตอบสนอง (Response Surface Methodology: RSM) เพื่อศึกษาผลของอุณหภูมิและเวลาต่อร้อยละผลได้และสมบัติของถ่านชีวภาพ (2) สร้างและพัฒนาเตาเผามลพิษต่ำขนาด 200 ลิตร เพื่อผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรดในระดับภาคสนาม โดยผู้วิจัยมุ่งหวังให้สามารถนำผลการศึกษาที่ได้เป็นพื้นฐานของการศึกษาการผลิตถ่านชีวภาพจากวัสดุเหลือทิ้งทางการเกษตรอื่น ๆ เป็นต้นแบบในการพัฒนาเตาเผามลพิษต่ำที่สามารถนำมาใช้ในชุมชนเป้าหมายได้โดยส่งผล

กระทบต่อสิ่งแวดล้อมน้อยที่สุด รวมถึงศึกษาความเป็นไปได้ในการนำไปใช้งานจริงและการพัฒนาผลิตภัณฑ์อื่น ๆ ต่อไป

2) วัตถุประสงค์ของงานวิจัย

1. เพื่อสร้างเตาเผามลพิษต่ำสำหรับใช้ในชุมชน
2. เพื่อผลิตถ่านชีวภาพจากใบสับปะรดศรีราชา
3. เพื่อศึกษาประสิทธิภาพและต้นทุนของเตาเผามลพิษต่ำสำหรับการผลิตถ่านชีวภาพและผลิตภัณฑ์พลอยได้จากใบสับปะรดศรีราชา

3) วิธีการดำเนินการวิจัย

3.1) วัตถุดิบ และเครื่องมือที่ใช้ในงานวิจัย

วัตถุดิบที่ใช้ในงานวิจัยนี้ คือ ใบสับปะรดศรีราชาจากพื้นที่ตำบลเขาคันทรง อำเภอศรีราชา จังหวัดชลบุรี และเครื่องเตาเผาแบบ Muffle Furnace รุ่น Nabertherm

3.2) การผลิตถ่านชีวภาพระดับห้องปฏิบัติการ

วิจัยนี้แบ่งออกเป็น 2 ส่วน คือ การผลิตถ่านชีวภาพในระดับห้องปฏิบัติการ และการพัฒนาเตาเผามลพิษต่ำเพื่อผลิตถ่านชีวภาพระดับภาคสนาม สำหรับการผลิตถ่านชีวภาพระดับห้องปฏิบัติการนั้นทำการออกแบบการทดลองโดยวิธี CCD [13] ซึ่งเป็นวิธีการทางสถิติที่ใช้วิเคราะห์การทดลองที่มีความหลากหลายของปัจจัยต่าง ๆ พร้อมกัน โดยมีเป้าหมายคือการคัดกรองปัจจัยที่มีความสำคัญต่อผลการทดลองเพื่อหาภาวะที่เหมาะสมในการผลิตถ่านชีวภาพ โดยมีการออกแบบการทดลองดังตารางที่ 1 สภาวะการทดลองที่กำหนดในการทดลองแสดงให้เห็นถึงช่วงอุณหภูมิและเวลาที่ใช้ในการทดลอง มีการแบ่งระดับการศึกษาเป็น 5 ระดับ คือ ที่ระดับลบแอลฟา ($-\alpha$ = -1.414) ระดับต่ำ (-1) ระดับกลาง (0) ระดับสูง (+1) และระดับบวกแอลฟา ($+\alpha$ = +1.414) โดยที่ระดับต่ำของอุณหภูมิและเวลาคือ 400 องศาเซลเซียส และ 1 ชั่วโมง ระดับกลางของอุณหภูมิและเวลาคือ 500 องศาเซลเซียส และ 1.5 ชั่วโมง และระดับสูงของอุณหภูมิและเวลาคือ 600 องศาเซลเซียส และ 2 ชั่วโมง ตามลำดับ จากนั้นจะนำผลการทดลองที่ได้ไปวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) และสร้างสมการถดถอย (Regression Model) สำหรับทำนายร้อยละผลได้ของแข็งและอัตราส่วนคาร์บอนต่อไฮโดรเจนของถ่านชีวภาพที่ผลิตได้

ตารางที่ 1 : ภาวะที่ใช้ในการทดลองตามการออกแบบการทดลองแบบ CCD

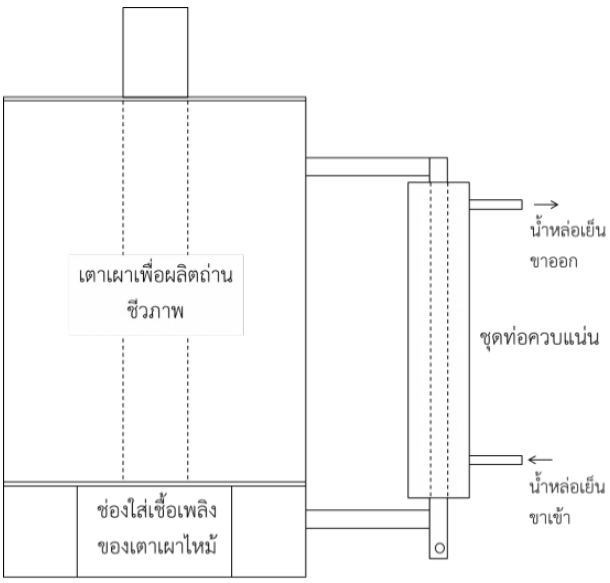
การทดลองที่	อุณหภูมิ (องศาเซลเซียส), ระดับ	เวลา (ชั่วโมง), ระดับ
1	379.29, -1.414	1.50, 0
2	400.00, -1	1.00, -1
3	400.00, -1	2.00, 1
4	450.00, 0	0.79, -1.414
5	450.00, 0	1.50, 0
6	450.00, 0	2.21, +1.414
7	500.00, 1	1.00, -1
8	500.00, 1	2.00, 1
9	520.71, +1.414	1.50, 0

วิธีการทดลองสำหรับการผลิตถ่านชีวภาพระดับห้องปฏิบัติการ คือ นำวัตถุดิบน้ำหนักประมาณ 4 กรัม มาใส่ในครูซิเบิล ชั่งน้ำหนักครูซิเบิล (W_C) และน้ำหนักที่แน่นอนของวัตถุดิบที่ใช้ (W_1) ปิดฝาครูซิเบิลและนำไปเข้าเตาเผา ตั้งเวลาและอุณหภูมิที่ต้องการ สำหรับเตาเผาที่ใช้ในการทดลองนี้มีอัตราการให้ความร้อน 15 องศาเซลเซียสต่อนาที เมื่อครบตามกำหนดเวลาแล้ว ทิ้งไว้ให้เย็นจนถึงอุณหภูมิห้อง นำออกจากเตาเผา เปิดฝา และชั่งน้ำหนัก (W_2) เก็บถ่านชีวภาพที่ได้ไว้ในตู้ดูดความชื้นจนกว่าจะมีการใช้งาน เมื่อวิเคราะห์ผลการทดลองจนได้ภาวะที่เหมาะสมแล้วจะนำไปเปรียบเทียบกับผลการทดลองจากส่วนที่สองคือการพัฒนาเตาเผามลพิษต่ำในระดับภาคสนาม

3.3) การผลิตถ่านชีวภาพระดับภาคสนาม

เตาเผามลพิษต่ำในระดับภาคสนามจะพัฒนาจากถังเหล็กขนาด 200 ลิตร โดยมีภาพร่างแสดงดังรูปที่ 1 เตาเผาที่พัฒนาขึ้นนี้ประกอบด้วย 3 ส่วนหลัก ได้แก่ เตาเผาเพื่อผลิตถ่านชีวภาพ เตาเผาไหม้เชื้อเพลิง และชุดท่อควบแน่น ผลิตภัณฑ์ที่ได้จากเตาเผามลพิษต่ำนี้จึงประกอบด้วยถ่านชีวภาพและน้ำส้มควันไม้ โดยเตาเผาไหม้เชื้อเพลิงจะมีช่องเปิดให้อากาศสามารถผ่านเข้าไปได้ เชื้อเพลิงที่ใส่เข้าไปจะถูกเผาไหม้ด้วยอากาศที่มากเกินพอผ่านกระบวนการเผาไหม้ (combustion) และคายพลังงานความร้อนอย่างเต็มที่และส่งพลังงานความร้อนต่อไปยังเตาเผาด้านบน ในขณะที่เตาเผาเพื่อผลิตถ่านชีวภาพจะทำหน้าที่บรรจุวัสดุที่เราต้องการนำมาผลิตถ่านชีวภาพ ซึ่งเตาส่วนนี้จะไม่มีช่องให้อากาศสามารถเข้าไปได้ เมื่อเตาเผาส่วนนี้ได้รับพลังงานความร้อนจากเตาเผาไหม้เชื้อเพลิงจะเกิดการสลายตัวด้วยความร้อนในที่จำกัดอากาศหรือกระบวนการไพโรไลซิส จากนั้นไอน้ำ สารระเหย และน้ำมันทาร์ จะแยกตัวออกมาจากวัตถุดิบ และระบายทางท่อด้าน

ข้างซึ่งมีการต่อชุดท่อควบแน่น ลักษณะเป็นอุปกรณ์แลกเปลี่ยนความร้อนแบบท่อสองชั้นเพื่อควบแน่นและผลิตน้ำส้มควันไม้ แก๊สที่เหลือจากการควบแน่นจะถูกป้อนเข้าเตาเผาไหม้เชื้อเพลิงเพื่อเป็นแก๊สเชื้อเพลิง ด้วยวิธีการนี้จะทำให้แก๊สหรือควันที่เกิดขึ้นจากกระบวนการมีกลิ่นฉุนแสบจมูกน้อย ช่วยลดการปลดปล่อยแก๊สเชื้อเพลิงที่เกิดจากกระบวนการผลิตถ่านชีวภาพโดยการนำมาใช้เป็นเชื้อเพลิงร่วมได้อีกด้วย



รูปที่ 1 : แบบจำลองแสดงหลักการของเตาเผามลพิษต่ำ

การผลิตถ่านชีวภาพจากใบสับปะรดด้วยเตาเผามลพิษต่ำนี้ทำได้โดยการชั่งน้ำหนักใบสับปะรด (W_3) นำใบสับปะรดใส่ในเตาเผาด้านบน จนเต็มส่วนของเตาเผาที่ผลิตถ่านชีวภาพ ซึ่งน้ำหนักรวมของใบสับปะรดที่ได้มีค่าประมาณ 20 กิโลกรัม จากนั้นปิดปากเตาเผา นำดินเหนียวอุดบริเวณฝาเตาเผาและท่อระบายควัน นำเชื้อเพลิงใส่ที่เตาเผาไหม้ด้านล่าง จุดไฟและเปิดน้ำหล่อเย็นใส่เชื้อเพลิงเข้าไปและปล่อยให้เกิดการเผาไหม้ โดยทำการวัดอุณหภูมิบริเวณด้านล่างของปล่องระบายควันซึ่งเป็นส่วนที่ติดกับฝาเตาเผา อุณหภูมิที่วัดได้ควรอยู่ระหว่าง 350–400 องศาเซลเซียส จากนั้นประมาณ 30 นาที น้ำส้มควันไม้จะเริ่มหยดออกมา สังเกตไฟในเตาเผาไหม้และเติมเชื้อเพลิงจนกว่าน้ำส้มควันไม้จะหยุดไหล จากนั้นเผาต่อไปอีก 1 ชั่วโมงแล้วปล่อยให้เตาเย็นลง เมื่อเตาเผาเย็นลงถึงอุณหภูมิห้องแล้วจึงเปิดฝา ชั่งน้ำหนักถ่านชีวภาพที่ได้ (W_4) และเก็บถ่านชีวภาพใส่ภาชนะที่เหมาะสมเพื่อนำไปวิเคราะห์ต่อไป

3.4) การวิเคราะห์สารตั้งต้นและผลิตภัณฑ์

3.4.1) การวิเคราะห์องค์ประกอบของใบสับปะรด การวิเคราะห์องค์ประกอบแบบแยกธาตุและการวิเคราะห์ธาตุด้วยเครื่อง Elemental Analyzer–CHNS/O LECO 628 Series และเครื่อง XRF (WDXRF) BRUKER S8 TIGER ตามลำดับ

3.4.2) การคำนวณค่าร้อยละผลได้ของแข็ง (%Solid)

คำนวณหาร้อยละผลได้ของแข็งได้ดังนี้

ระดับห้องปฏิบัติการ $$\%Solid = \frac{W_2 - W_c}{W_1} \times 100 \quad (1)$$

ระดับภาคสนาม $$\%Solid = \frac{W_4}{W_3} \times 100 \quad (2)$$

โดย %Solid คือ ผลได้ของของแข็ง (ร้อยละโดยน้ำหนัก)

W_c คือ น้ำหนักครูซิเบิล (กรัม)

W_1 คือ น้ำหนักใบสับปะรดก่อนทำการทดลองระดับห้องปฏิบัติการ (กรัม)

W_2 คือ น้ำหนักถ่านชีวภาพและครูซิเบิลหลังทำการทดลองระดับห้องปฏิบัติการ (กรัม)

W_3 คือ น้ำหนักใบสับปะรดก่อนทำการทดลองระดับภาคสนาม (กรัม)

W_4 คือ น้ำหนักถ่านชีวภาพหลังทำการทดลองระดับภาคสนาม (กรัม)

3.4.3) การวิเคราะห์องค์ประกอบของถ่านชีวภาพและน้ำส้มควันไม้ การวิเคราะห์องค์ประกอบโดยธาตุ (Ultimate Analysis) ของใบสับปะรดและถ่านชีวภาพวิเคราะห์ด้วยเครื่อง THERMO FLASH 2000 CHNS/O ANALYZERS และนำองค์ประกอบโดยธาตุที่ได้ไปคำนวณค่าความร้อนสูงสุดของถ่านชีวภาพจากสมการต่อไปนี้ [14]:

$$H.C.V. = 0.3856\,(C + H) - 1.6938 \text{ MJ/kg} \quad (3)$$

โดย H.C.V. คือ ค่าความร้อนสูงสุด (เมกะจูลต่อกิโลกรัม)

C คือ ร้อยละโดยน้ำหนักของธาตุคาร์บอนเฉลี่ย

H คือ ร้อยละโดยน้ำหนักของธาตุไฮโดรเจนเฉลี่ย

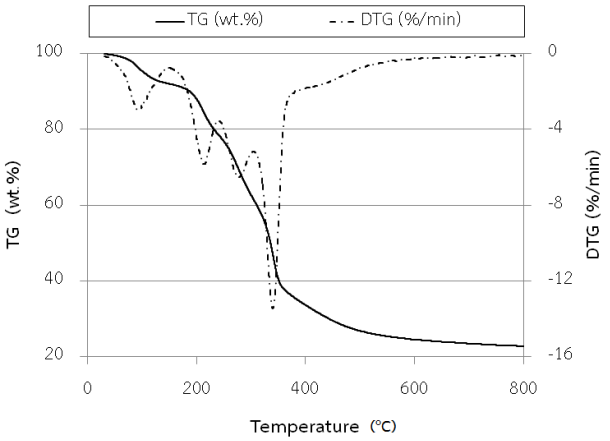
การวิเคราะห์องค์ประกอบของน้ำส้มควันไม้วิเคราะห์ด้วยเครื่อง Gas Chromatograph–Mass Spectrometer (GC-MS /MS TQ) รุ่น AGILENT 7890 GC/7000C GC/MS TRIPLE QUAD โดยใช้คอลัมน์ CP-Wax เตรียมตัวอย่างโดยการนำน้ำส้มควันไม้ 5 มล. ผสมกับแอมโมเนียมไฮดรอกไซด์ (Q RëC™ AMMONIA SOLUTION 28%) 1.5 มล. เพื่อปรับให้ pH มีค่าประมาณ 5 จากนั้นเติมไดคลอโรมีเทน (Fisher Chemical™ Dichloromethane) หรือ เอทิลอะซิเตท (RCI Labscan™ Ethyl Acetate 99.8%) 3 มล. จากนั้น

จึงนำไปกรองและเตรียมพร้อมสำหรับการวิเคราะห์องค์ประกอบต่อไป [15]

4) ผลการวิจัย

4.1) สมบัติของใบสับปะรด

ผลการวิเคราะห์ร้อยละคาร์บอน ร้อยละไฮโดรเจน ร้อยละไนโตรเจน และร้อยละออกซิเจนของใบสับปะรดพบว่ามีค่าเท่ากับ 45.88 ± 0.70, 6.60 ± 0.12, 1.15 ± 0.28 และ 46.37 ± 0.76 โดยน้ำหนัก ตามลำดับ สำหรับผลการวิเคราะห์ธาตุในเถ้าของใบสับปะรดพบว่ามีร้อยละโดยน้ำหนักของธาตุต่าง ๆ ดังนี้ โพแทสเซียม ร้อยละ 26.60 คลอรีน ร้อยละ 7.64 แคลเซียม ร้อยละ 2.89 ฟอสฟอรัส ร้อยละ 2.79 แมกนีเซียม ร้อยละ 2.35 ซิลิคอน ร้อยละ 1.44 และพบธาตุที่มีปริมาณไม่เกินร้อยละ 1 ได้แก่ ซัลเฟอร์ แมงกานีส โซเดียม รูบิเดียม และอื่น ๆ



รูปที่ 2 : การวิเคราะห์การสลายตัวทางความร้อนด้วยเทคนิค TGA

รูปที่ 2 แสดงผลการวิเคราะห์การสลายตัวทางความร้อนของใบสับปะรดด้วยเทคนิค Thermogravimetric Analysis (TGA) พบว่าใบสับปะรดมีการสูญเสียความชื้นตั้งแต่อุณหภูมิห้องถึงประมาณ 200 องศาเซลเซียส ทำให้น้ำหนักลดลงประมาณร้อยละ 10 ของน้ำหนักเริ่มต้น จากนั้นจึงเป็นการสลายตัวขององค์ประกอบหลักต่าง ๆ ได้แก่ เฮมิเซลลูโลส รวมถึงเซลลูโลสและลิกนินบางส่วน ซึ่งทำให้น้ำหนักของชีวมวลลดลงอีกประมาณร้อยละ 50 และมีการสลายตัวมากที่สุดที่อุณหภูมิประมาณ 350 องศาเซลเซียส [16] เมื่ออุณหภูมิมากกว่า 350 องศาเซลเซียส ใบสับปะรดยังมีการสลายตัวอย่างช้า ๆ จนกระทั่งถึงอุณหภูมิประมาณ 550 องศาเซลเซียส เนื่องจากการสลายตัวของลิกนินและเซลลูโลสที่เหลืออยู่ จนกระทั่งน้ำหนักของใบสับปะรดลดลงเหลือประมาณร้อยละ 25 ซึ่งมีความเป็นไปได้ว่าน้ำหนักที่เหลืออยู่นั้นเป็นน้ำหนักของคาร์บอนคงตัวและเถ้า

จากการศึกษาการสลายตัวของใบสับปะรดนี้ทำให้คณะผู้วิจัยได้เลือกช่วงอุณหภูมิในการศึกษาการผลิตถ่านชีวภาพจากใบสับปะรดที่อุณหภูมิ 400–500 องศาเซลเซียส ซึ่งเป็นอุณหภูมิที่สารระเหยมีการสลายตัวไปเป็นจำนวนมากและมีโอกาสที่จะได้ถ่านชีวภาพที่มีค่าความร้อนและความเป็นรูพรุนสูง

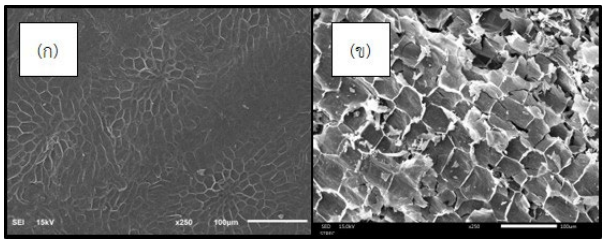
4.2) การผลิตถ่านชีวภาพระดับห้องปฏิบัติการ

จากการผลิตถ่านชีวภาพโดยการออกแบบการทดลองแบบ CCD ดังรายละเอียดในตารางที่ 2 สำหรับร้อยละผลได้ของแข็ง อัตราส่วนคาร์บอนต่อไฮโดรเจน และค่าความร้อนสูงสุด (เมกะจูลต่อกิโลกรัม) ของถ่านชีวภาพ โดยผลการทดลองแสดงให้เห็นว่าผลได้ของแข็งสูงที่สุดที่อุณหภูมิ 400 องศาเซลเซียสและเวลา 1 ชั่วโมง มีค่าร้อยละผลได้ของแข็งเท่ากับร้อยละ 33.02 สำหรับค่าความร้อนสูงสุดของถ่านชีวภาพจากใบสับปะรดผลการทดลองแสดงให้เห็นว่ามีค่าประมาณ 20–23 เมกะจูลต่อกิโลกรัม โดยที่ภาวะที่ทำให้ถ่านชีวภาพมีค่าความร้อนสูงสุดคือ อุณหภูมิ 400 องศาเซลเซียสและเวลา 1 ชั่วโมง มีค่าความร้อนสูงสุดเท่ากับ 23.25 เมกะจูลต่อกิโลกรัม ซึ่งมากกว่าค่าความร้อนของใบสับปะรดแห้ง คือ 18.54 เมกะจูลต่อกิโลกรัม

ตารางที่ 2 : ร้อยละผลได้ของแข็ง อัตราส่วน C/H และค่าความร้อนสูงสุดของถ่านชีวภาพ

Temperature (°C)	Time (h)	% Solid yield (wt.%)	C/H ratio	HHV (MJ/kg)
379.29	1.50	29.26±0.45	18.46±1.66	20.89
400.00	1.00	33.02±0.44	19.73±0.91	23.25
400.00	2.00	22.57±0.00	19.10±0.10	22.57
450.00	0.79	27.50±0.71	22.27±1.85	22.18
450.00	1.50	20.98±0.00	25.41±3.32	22.64
450.00	2.21	20.64±0.90	27.19±0.36	21.13
500.00	1.00	18.88±0.53	30.08±0.93	22.39
500.00	2.00	18.93±0.40	28.34±1.71	20.82
520.71	1.50	15.07±0.08	32.95±1.92	20.83
450.00	2.21	20.64±0.90	27.19±0.36	21.13
500.00	1.00	18.88±0.53	30.08±0.93	22.39
500.00	2.00	18.93±0.40	28.34±1.71	20.82
520.71	1.50	15.07±0.08	32.95±1.92	20.83

รูปที่ 3 : ถ่านชีวภาพจากใบสับปะรดระดับห้องปฏิบัติการ



รูปที่ 4 : (ก) ภาพถ่ายกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราดที่กำลังขยาย 250 เท่า ของใบสับปะรดก่อนการเผา และ (ข) ใบสับปะรดหลังจากการเผาที่อุณหภูมิ 400 องศาเซลเซียส เป็นเวลา 1 ชั่วโมง

ลักษณะของถ่านชีวภาพจากใบสับปะรดที่ผลิตในระดับห้องปฏิบัติการแสดงดังรูปที่ 3 ซึ่งพบว่ามีผิวดำเงาและยังมีรูปร่างคล้ายกับใบสับปะรดแห้งก่อนเผา โดยมีความกว้างของใบสูงสุดประมาณ 3 เซนติเมตร เมื่อนำใบสับปะรดและถ่านชีวภาพไปวิเคราะห์ลักษณะพื้นผิวด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราดที่กำลังขยาย 250 เท่า พบว่าผลการวิเคราะห์แสดงดังรูปที่ 4 จากรูปที่ 4(ก) พบว่าลักษณะพื้นผิวของใบสับปะรดก่อนเผาค่อนข้างเรียบ เมื่อนำไปเผาที่อุณหภูมิ 400 องศาเซลเซียส เป็นเวลา 1 ชั่วโมง สามารถสังเกตเห็นความขรุขระและรูพรุนเกิดขึ้นดังแสดงในรูปที่ 4(ข) เนื่องจากการสลายตัวขององค์ประกอบหลักบางส่วนและกลายเป็นสารระเหย ดังที่ได้กล่าวไปแล้วข้างต้น อย่างไรก็ตามรูพรุนที่เกิดขึ้นยังคงมีลักษณะไม่ชัดเจนเนื่องจากยังเกิดการสลายตัวทางความร้อนไม่สมบูรณ์ที่อุณหภูมิต่ำ ทำให้ยังคงมีสารระเหยอยู่บนผิวของถ่านชีวภาพ [17]

ตารางที่ 3 : การวิเคราะห์ความแปรปรวน (ANOVA) ของอุณหภูมิและเวลา

(ก) ต่อร้อยละผลได้ของแข็ง

Source	Sum of Squares	Degree of Freedom	Mean Square	F-value	*p*-value
Model	264.04	4	66.01	128.5	0.0002
A	179.06	1	179.06	348.57	< 0.0001*
B	50.51	1	50.51	98.33	0.0006*
AB	27.56	1	27.56	53.66	0.0018*
B^2	6.91	1	6.91	13.46	0.0214*
Residual	2.05	4	0.5137		
Total	266.09	8			

(ข) ต่ออัตราส่วน C/H

Source	Sum of Squares	Degree of Freedom	Mean Square	F-value	*p*-value
Model	203.45	2	101.73	42.85	0.0003
A-อุณหภูมิ	200.82	1	200.82	84.59	< 0.0001*
B-เวลา	2.63	1	2.63	1.11	0.3330
Residual	14.24	6	2.37		
Total	217.70	8			

โดยที่ A คือ อุณหภูมิ (องศาเซลเซียส)

B คือ เวลา (ชั่วโมง)

* *p*-value < 0.05 แสดงว่ามีผลที่ระดับความเชื่อมั่น 95% อย่างมีนัยสำคัญ

เมื่อนำผลการทดลองที่มีการออกแบบการทดลองโดยวิธี CCD ดังแสดงในตารางที่ 2 ไปวิเคราะห์ความแปรปรวน จะได้ผลการวิเคราะห์แสดงดังตารางที่ 3 โดยตารางที่ 3(ก) แสดงการวิเคราะห์ผลของปัจจัยการทดลองต่อผลได้ของแข็ง และตารางที่ 3(ข) แสดงการวิเคราะห์ผลของปัจจัยการทดลองต่ออัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจน จากการพิจารณาค่า *p*-value ที่ระดับความเชื่อมั่น 95% พบว่าค่า *p*-value ของทั้งอุณหภูมิและเวลาในตารางที่ 3(ก) และเฉพาะอุณหภูมิในตารางที่ 3(ข) มีค่าน้อย

กว่า 0.05 แสดงว่าอุณหภูมิและเวลามีผลต่อร้อยละผลได้ของแข็ง ในขณะที่มีเพียงอุณหภูมิเท่านั้นที่มีผลต่ออัตราส่วนระหว่างคาร์บอนและไฮโดรเจน นอกจากนี้อันตรกิริยาระหว่างอุณหภูมิและเวลายังมีผลต่อร้อยละผลได้ของแข็งอีกด้วย จากนั้นเมื่อนำผลการวิเคราะห์ความแปรปรวนมาสร้างสมการถดถอยสำหรับร้อยละผลได้ของแข็งและอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนสามารถแสดงได้ดังสมการที่ 4 และ 5

สมการถดถอยสำหรับร้อยละผลได้ของแข็ง

$$\%\text{Solid yield} = 21.93 - 4.73A - 2.51B + 2.63AB + 1.19B^2 \quad (4)$$

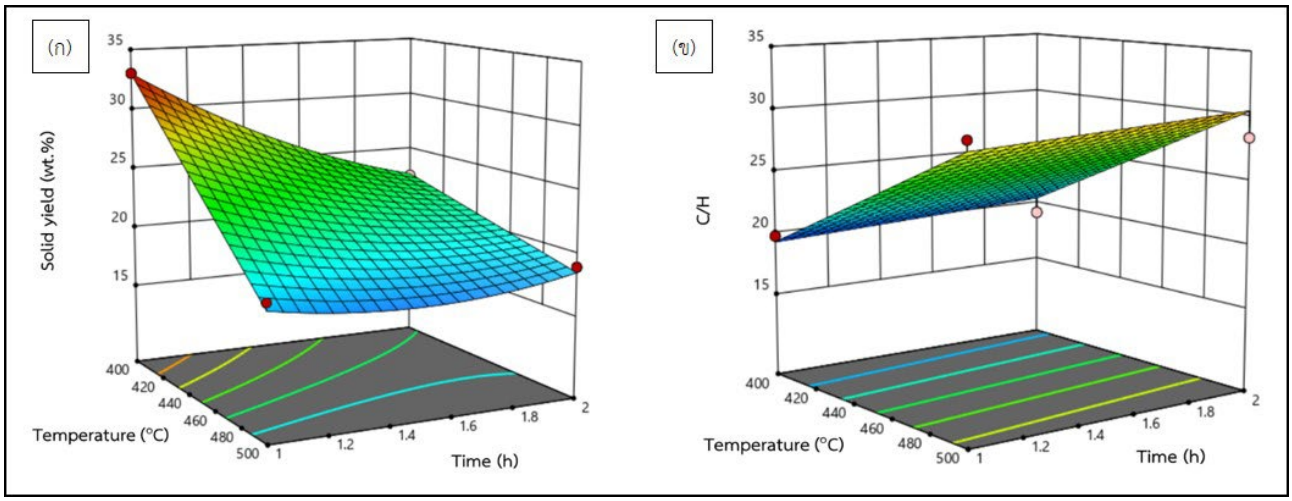
สมการถดถอยสำหรับอัตราส่วนคาร์บอนต่อไฮโดรเจน

$$\text{C/H ratio} = 24.84 + 5.01A + 0.57B \quad (5)$$

เมื่อ A คือ อุณหภูมิ (องศาเซลเซียส) และ B คือ เวลา (ชั่วโมง) โดยที่ค่า R^2 ของสมการที่ (4) และ (5) มีค่าเท่ากับ 0.99 และ 0.93 ตามลำดับ และมีพื้นผิวตอบสนองสำหรับร้อยละผลได้ของแข็งและอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนแสดงดังรูปที่ 5(ก) และ 5(ข) ตามลำดับ

จากสมการถดถอยจะเห็นว่าสัมประสิทธิ์หน้า A และ B สำหรับร้อยละผลได้ของแข็งมีค่าเป็นลบ แสดงว่าการเพิ่มอุณหภูมิและเวลาทำให้ร้อยละผลได้ของแข็งลดลง ในขณะที่สัมประสิทธิ์หน้า A สำหรับอัตราส่วนคาร์บอนต่อไฮโดรเจนมีค่าเป็นบวก แสดงว่าการเพิ่มอุณหภูมิทำให้อัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนเพิ่มขึ้น เมื่อพิจารณาร่วมกับผลการทดลองในตารางที่ 2 จะเห็นว่ามีความสอดคล้องกัน นั่นคือ เมื่ออุณหภูมิและเวลาเพิ่มขึ้น ทำให้ร้อยละผลได้ของแข็งลดลง เนื่องจากการสลายตัวทางความร้อน (thermal cracking) ของใบสับปะรดและเมื่ออุณหภูมิเพิ่มขึ้นทำให้อัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนเพิ่มขึ้น เนื่องจากการสลายตัวทางความร้อนของเซลลูโลสและลิกนินกลายเป็นของเหลวมวลโมเลกุลต่ำและแก๊ส [18] ซึ่งประกอบด้วยไฮโดรคาร์บอนและสารอินทรีย์ที่มีองค์ประกอบของไฮโดรเจนมาก หรืออีกนัยหนึ่งคือมีอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนต่ำระเหยออกไปจากโครงสร้าง และคงเหลือไว้แต่องค์ประกอบที่มีความซับซ้อนและมีคาร์บอนเป็นองค์ประกอบสูง

จากพื้นผิวตอบสนองที่แสดงดังรูปที่ 5 จะเห็นว่าพื้นผิวตอบสนองของร้อยละผลได้ของแข็งมีความโค้งเล็กน้อย อันเนื่องมาจากผลของอันตรกิริยาระหว่างอุณหภูมิและเวลา อย่างไรก็ตามจะเห็นว่าอุณหภูมิและเวลามีผลอย่างมากต่อร้อยละผลได้ของแข็งจนทำให้ลักษณะพื้นผิวตอบสนองมีความใกล้เคียงเชิงเส้น ซึ่งคล้ายผลการทดลองกับงานวิจัยก่อนหน้า [19] ในขณะที่พื้นผิวตอบสนองของอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนมีลักษณะเชิงเส้น เนื่องจากไม่มีผลของอันตรกิริยาและปัจจัยที่ไม่อยู่ในลักษณะเชิงเส้น (quadratic effects) คือ A^2 และ B^2 จากพฤติกรรมดังกล่าวจะเห็นว่าภาวะที่ทำให้ได้ร้อยละผลได้ของแข็งและอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนสูงสุดจะอยู่ที่ตำแหน่งริมพื้นผิวตอบสนอง กล่าวคือที่อุณหภูมิหรือเวลามีค่าสูงสุดหรือต่ำสุดนั่นเอง



รูปที่ 5 : (ก) พื้นผิวตอบสนองสำหรับร้อยละผลได้ของแข็ง และ (ข) อัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจน

จากผลการวิเคราะห์ดังกล่าว สามารถกล่าวได้ว่า การเพิ่มอุณหภูมิทำให้ได้ปริมาณถ่านชีวภาพลดลงแต่มีอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนเพิ่มขึ้นซึ่งส่งผลต่อการเพิ่มขึ้นของค่าความร้อนสูงสุดของถ่านชีวภาพอย่างมีนัยสำคัญ เนื่องจากอุณหภูมิที่สูงขึ้นทำให้ชีวมวลเกิดการคาร์บอไนซ์ได้อย่างสมบูรณ์มากขึ้น สารระเหยซึ่งมีอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนต่ำกว่าจะถูกกำจัดออกจากโครงสร้างของถ่านชีวภาพ ทำให้ถ่านชีวภาพมีองค์ประกอบคาร์บอนสูงขึ้น นอกจากนี้ผลการวิเคราะห์ความแปรปรวนยังแสดงให้เห็นว่าอันตรกิริยาระหว่างอุณหภูมิและเวลามีผลต่อร้อยละผลได้ของแข็ง แสดงว่าการเปลี่ยนแปลงเวลาที่ใช้ในการผลิตถ่านชีวภาพส่งผลกระทบต่อแนวโน้มที่ได้จากการเปลี่ยนแปลงอุณหภูมิ ดังนั้น การควบคุมอุณหภูมิเพื่อให้ได้ร้อยละผลได้ถ่านชีวภาพตามที่ต้องการนั้น จำเป็นต้องควบคุมเวลาที่ใช้ไปควบคู่กัน การเพิ่มอุณหภูมิอาจไม่ทำให้ร้อยละผลได้ของแข็งลดลงเสมอไปหากใช้เวลาไม่มาก ดังแสดงในผลการทดลองจากตารางที่ 2 ยกตัวอย่างเช่น การผลิตถ่านชีวภาพที่ 379.29 องศาเซลเซียส เวลา 1.50 ชั่วโมง ให้ร้อยละผลได้ของแข็งเท่ากับ 29.26 ซึ่งน้อยกว่าการผลิตถ่านชีวภาพที่ 400 องศาเซลเซียส เวลา 1 ชั่วโมง ที่ให้ร้อยละผลได้ของแข็งเท่ากับ 33.03 หรือ การผลิตถ่านชีวภาพที่ 400 องศาเซลเซียส เวลา 2 ชั่วโมง ให้ร้อยละผลได้ของแข็งเท่ากับ 22.57 ซึ่งน้อยกว่าการผลิตถ่านชีวภาพที่ 450 องศาเซลเซียส เวลา 0.79 ชั่วโมง ที่ให้ร้อยละผลได้ของแข็งเท่ากับ 27.50 เป็นต้น

4.3) การผลิตถ่านชีวภาพและน้ำส้มควันไม้ด้วยเตาเผา 200 ลิตร

การสร้างเตาเผาขนาด 200 ลิตรเพื่อใช้ผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรดนั้นเมื่อประกอบเสร็จแล้วสามารถแสดงได้ดังรูปที่ 6 เมื่อนำมาติดตั้งระบบน้ำวนสำหรับอุปกรณ์แลกเปลี่ยนความร้อนแบบท่อสองชั้นแล้ว จึงนำไปทดสอบกับใบสับปะรดที่รับมาจากพื้นที่เป้าหมาย (ตำบลเขาคันทรง) จากการทดสอบพบว่าเตาเผาจากถัง 200 ลิตรสามารถผลิตชีวภาพจากใบสับปะรดได้โดยที่ลักษณะของถ่านที่เกิดขึ้นมีรูปร่างคล้ายใบสับปะรดสดที่ใส่ลงไปตอนต้นและคล้ายกับลักษณะที่สังเกตได้จากถ่านชีวภาพที่ผลิตได้ในระดับห้องปฏิบัติการ เมื่อจับหรือขยำมีความกรอบสามารถบดให้มีขนาดเล็กลงได้โดยยังไม่ต้องใช้เครื่องในขั้นตอนแรก แม้ว่าสีดำของผงถ่านชีวภาพจะติดมือแต่สามารถล้างออกง่ายเพียงใช้น้ำสะอาดชะ แตกต่างจากถ่านปกติที่สีดำติดมือแล้วล้างออกค่อนข้างยาก โดยการเผาใบสับปะรดสด 20 กิโลกรัม จะให้น้ำส้มควันไม้และถ่านชีวภาพประมาณ 10 และ 2–4 กิโลกรัม ตามลำดับ

รูปที่ 6 : ภาพเตาเผาขนาด 200 ลิตร

ตารางที่ 4 แสดงสมบัติของถ่านชีวภาพและใบสับปะรดแห้ง พบว่าถ่านชีวภาพที่ได้มีคาร์บอนเป็นองค์ประกอบสูงถึงร้อยละ 70.49 มีอัตราส่วนคาร์บอนต่อไฮโดรเจนเท่ากับ 55.94 คิดเป็นค่าความร้อนสูงสุด 25.97 เมกะจูลต่อกิโลกรัม ซึ่งมากกว่าร้อยละคาร์บอน อัตราส่วนคาร์บอนต่อไฮโดรเจนและค่าความร้อนสูงสุดของใบสับปะรดแห้ง ซึ่งมีค่าเท่ากับ 45.88 6.95 และ 18.54 ตามลำดับ นอกจากนี้เมื่อเปรียบเทียบกับสมบัติของถ่านชีวภาพจากห้องปฏิบัติการซึ่งไม่มีการเก็บน้ำส้มควันไม้ดังที่แสดงในตารางที่ 2 พบว่าอัตราส่วนระหว่างคาร์บอนต่อไฮโดรเจนและค่าความร้อนสูงสุดของถ่านชีวภาพจากเตาเผาขนาด 200 ลิตรมีค่ามากกว่า ทั้งนี้เนื่องมาจากการทดลองในห้องปฏิบัติการนั้นไม่มีการแยกแก๊สเพื่อเก็บน้ำส้มควันไม้ ดังนั้นสารระเหยที่สลายตัวออกมาบางส่วนจึงสามารถเกิดการรวมตัวหรือเกิดกระบวนการรีพอลิเมอไรเซชัน (Repolymerization) และเกาะอยู่บนผิวของถ่านชีวภาพในขณะเย็นตัวลงได้ [20] ซึ่งสารระเหยดังกล่าวมีอัตราส่วนคาร์บอนต่อไฮโดรเจนต่ำกว่าในถ่านชีวภาพ ดังนั้นอัตราส่วนคาร์บอนต่อไฮโดรเจนรวมจึงลดลงตามไปด้วย อย่างไรก็ตามร้อยละผลได้ของแข็งของถ่านชีวภาพที่ผลิตด้วยเตาเผาขนาด 200 ลิตรนั้นน้อยกว่าการผลิตในระดับห้องปฏิบัติการเนื่องจากสารระเหยที่สลายตัวออกมาจากวัตถุดิบตั้งต้นถูกแยกออกมาควบแน่นเป็นน้ำส้มควันไม้และแก๊สที่เหลือถูกนำกลับมาใช้เป็นเชื้อเพลิงร่วมในเตาเผาไหม้แล้ว ดังที่กล่าวถึงข้างต้น ดังนั้นการพิจารณาการใช้เตาเผาขนาด 200 ลิตรนี้ในการผลิตผลิตภัณฑ์รูปแบบต่าง ๆ จึงต้องระบุ

วัตถุประสงค์หรือผลิตภัณฑ์หลักที่ต้องการ โดยคำนึงถึงวัตถุดิบตั้งต้น ยกตัวอย่างเช่น งานวิจัยนี้ใช้ใบสับปะรดซึ่งมีความชื้นสูง ทำให้ได้น้ำส้มควันไม้ปริมาณมาก ในขณะที่ถ่านชีวภาพมีปริมาณน้อย หากต้องการให้ถ่านชีวภาพมีปริมาณมากขึ้น จำเป็นต้องตากใบสับปะรดให้แห้งซึ่งต้องมีพื้นที่รองรับที่สามารถวางใบสับปะรดได้นานประมาณ 2–3 สัปดาห์

ตารางที่ 5 แสดงผลการวิเคราะห์องค์ประกอบของน้ำส้มควันไม้ โดยการสกัดด้วยไดคลอโรมีเทนหรือเอทิลอะซิเตท จากนั้นนำส่วนที่สกัดไปวิเคราะห์ด้วยเทคนิค GC/MS ผลการวิเคราะห์แสดงให้เห็นว่าองค์ประกอบหลักของน้ำส้มควันไม้ที่ถูกสกัดด้วยไดคลอโรมีเทนหรือเอทิลอะซิเตทนั้นไม่ต่างกันมากนัก องค์ประกอบส่วนใหญ่ของน้ำส้มควันไม้ คือ Furfural, Phenol, p-Cresol, 2-Methoxy-phenol (หรือ Guaiacol) และอนุพันธ์ของ Phenol อื่น ๆ ซึ่งสารประกอบประเภท furfural ที่พบในน้ำส้มควันไม้เกิดจากการสลายตัวของ Xylan ในเฮมิเซลลูโลส [21] ในขณะที่สารประกอบประเภทไกวเอคอล รวมถึงฟีนอล และอนุพันธ์ของฟีนอลอื่น ๆ นั้นได้มาจากการสลายตัวของลิกนิน มีฤทธิ์เป็นสารต้านเชื้อราและสารต้านเชื้อแบคทีเรีย [22] ทำให้มีการใช้น้ำส้มควันไม้ในการเตรียมดินปลูกเพื่อช่วยกำจัดโรครากเน่าและกำจัดเชื้อราในพืช รวมถึงฉีดพ่นเพื่อฆ่าเชื้อโรคและกำจัดศัตรูพืชบางชนิด

ตารางที่ 4 : คุณสมบัติของใบสับปะรดและถ่านชีวภาพจากใบสับปะรดที่ผลิตจากเตาเผาขนาด 200 ลิตร

Material	HHV (MJ/kg)	Ultimate analysis (wt.%)				C/H ratio
		C	H	N	O	
Pineapple leaves	18.54	45.88±0.70	6.60±0.12	1.15±0.28	46.37±6.03	6.95±6.34
Biochar	25.97	70.49±0.81	1.26±0.10	1.38±0.25	26.87±0.82	55.9±4.49

ตารางที่ 5 : องค์ประกอบของน้ำส้มควันไม้จากใบสับปะรด

No.	Compound name	Dichloromethane			Ethyl acetate		
		OC	RT	Area %	OC	RT	Area %
1	2-Cyclopenten-1-one		11.647	1.89		11.644	0.31
2	2-Cyclopenten-1-one, 2-methyl-		12.146	2.2		12.135	1.03
3	Furfural		16.346	8.03		16.335	3.16
4	Ethanone, 1-(2-furanyl)-		18.023	3.66		18.016	2.01
5	2-Cyclopenten-1-one, 3-methyl-		18.193	2.23		18.186	0.68
6	2-Cyclopenten-1-one, 2,3-dimethyl-		19.141	2.16		19.134	1.1
7	2-Furancarboxaldehyde, 5-methyl-		20.996	4.39		20.988	2.2
8	Acetophenone		24.116	1.8		24.112	1.24
9	3-Furanmethanol		25.048	2.1		25.048	2.11
10	Acetamide	*	*	*		29.517	20.32
11	Phenol, 2-methoxy-		33.192	4.74		33.199	4.09
12	2-Methoxy-5-methylphenol		37.018	1.14	*	*	*
13	Phenol		39.092	10.78		39.08	14.39
14	Phenol, 4-ethyl-2-methoxy-		39.9	2.94		39.892	2.3
15	p-Cresol		42.004	5.42		41.996	5.07
16	Phenol, 4-ethyl-		45.415	8.44		45.418	7.4
17	Phenol, 3-(1-methylethyl)-		47.05	1.76		47.05	1.75
18	Phenol, 2,6-dimethoxy-		48.504	4.29		48.504	2.79
19	Phenol, 2,4-bis(1,1-dimethylethyl)-	*	*	*		50.585	1.15
20	Phenol, 2,6-dimethoxy-4-(2-propenyl)-		83.379	3.23		57.652	0.09

หมายเหตุ : OC (Compound Occurrence) หมายถึงสารประกอบที่พบในน้ำส้มควันไม้ที่สกัดด้วยไดคลอโรมีเทนและเอทิลอะซิเตท โดยที่แถบสีแสดงสารประกอบที่พบ และเครื่องหมายดอกจันแสดงสารประกอบที่ไม่พบ

ตารางที่ 6 : การประมาณราคาต้นทุนการผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรด 20 กิโลกรัม

รายการ	ราคาต่อหน่วย	หน่วย	จำนวนที่ใช้	หน่วย	ราคารวม (บาท)
เตาเผา	5,000	บาทต่อเครื่อง	1	เครื่อง	5,000
เครื่องบดอัดถ่าน	5,000	บาทต่อเครื่อง	1	เครื่อง	5,000
รวมงบลงทุน					**10,000**
วัตถุดิบใบสับปะรด	0	บาทต่อกิโลกรัม	20	กิโลกรัม	-
ค่าเชื้อเพลิง	0	บาทต่อกิโลกรัม	20	กิโลกรัม	-
น้ำ	0.02	บาทต่อกิโลกรัม	200	กิโลกรัม	4
ไฟฟ้า	4	บาทต่อหน่วย	90	หน่วย	360
ค่าแป้งมัน	25	บาทต่อกิโลกรัม	0.2	กิโลกรัม	5
ค่าบรรจุภัณฑ์ถ่าน	2	บาทต่อกิโลกรัม	2	กิโลกรัม	4
ค่าบรรจุภัณฑ์น้ำส้มควันไม้	5	บาทต่อลิตร	10	ลิตร	50
รวมค่าดำเนินการต่อครั้ง					**393**
ขายถ่านอัดแท่ง	20	บาทต่อกิโลกรัม	2	กิโลกรัม	40
ขายน้ำส้มควันไม้	100	บาทต่อลิตร	10	ลิตร	1,000
รวมรายรับต่อครั้ง					**1,040**
จำนวนครั้งที่เผาเพื่อคืนทุน					$= \frac{10,000}{(1,040 - 393)}$ = **15.46**

หมายเหตุ : 1) ค่าใช้จ่ายอยู่ในพื้นฐานของการนำวัตถุดิบเหลือทิ้งในพื้นที่มาใช้ค่าใบสับปะรด ค่าเชื้อเพลิง และค่าขนส่งจึงไม่มี หากมีการใช้วัตถุดิบจากแหล่งอื่นอาจมีค่าใช้จ่ายในส่วนนี้เพิ่มเติม 2) ปริมาณผลิตภัณฑ์คือถ่านชีวภาพ 2 กิโลกรัม และน้ำส้มควันไม้ 10 ลิตร ต่อการใช้ใบสับปะรด 20 กิโลกรัม

รูปที่ 7 : ผลิตภัณฑ์จากใบสับปะรด ได้แก่ (ก) ถ่านผลไม้ (ข) น้ำส้มควันไม้ (ค) ผงถ่านชีวภาพ และ (ง) ถ่านอัดแท่งหรือถ่านอัดก้อน

4.4) วิเคราะห์ต้นทุนการผลิตและเสนอแนวทางการนำไปใช้

ตารางที่ 6 แสดงการประมาณราคาต้นทุนการผลิตถ่านชีวภาพและน้ำส้มควันไม้ ด้วยเตาเผามลพิษต่ำขนาด 200 ลิตร จากการทดลองด้วยการใช้ใบสับปะรดสด 20 กิโลกรัม สามารถผลิตถ่านชีวภาพได้ 2 กิโลกรัมและน้ำส้มควันไม้ 10 ลิตร เมื่อคำนวณต้นทุนพบว่างบลงทุน ได้แก่ ค่าเตาเผาและเครื่องบดอัดถ่าน คิดเป็น 10,000 บาท ค่าใช้จ่ายในการดำเนินการผลิตหนึ่งครั้งคิดเป็น 393 บาท รายได้จากการจำหน่ายถ่านชีวภาพและน้ำส้มควันไม้ต่อการผลิตหนึ่งครั้งคิดเป็น 1,040 บาท เมื่อนำมาคำนวณหาจำนวนครั้งในการเผาเพื่อคืนทุน พบว่าจะต้องเผาประมาณ 16 ครั้ง ซึ่งจากการทดสอบภาคสนามพบว่าการเผาหนึ่งครั้งใช้เวลาครึ่งวันและทิ้งไว้ให้เย็นแบบข้ามคืนจึงจะสามารถเก็บถ่านชีวภาพออกจากเตาและดำเนินการเผาครั้งต่อไปได้ ซึ่งคิดเป็นการเผาหนึ่งครั้งใช้ระยะเวลา 1 วัน แสดงให้เห็นว่าชุมชนสามารถดำเนินการเพื่อคืนทุนได้โดยใช้ระยะเวลาภายในหนึ่งเดือน

นอกจากนี้ จากการพัฒนาเตาเผาขนาด 200 ลิตร และทำการผลิตถ่านชีวภาพและน้ำส้มควันไม้จากใบสับปะรดนั้น พบว่ายังสามารถต่อยอดเป็นผลิตภัณฑ์อื่น ได้แก่ รูปที่ 7(ก) ถ่านผลไม้ รูปที่ 7(ค) ผงถ่านชีวภาพ และรูปที่ 7(ง) ถ่านอัดแท่งหรือถ่านอัดก้อน เป็นต้น แสดงให้เห็นว่าชุมชนยังมีทางเลือกในการพัฒนาผลิตภัณฑ์เพื่อต่อยอดเป็นผลิตภัณฑ์อื่นได้ เช่น การนำผลสับปะรดที่ไม่สมบูรณ์มาเผาด้วยเตาเผาดังกล่าว จะได้ถ่านชีวภาพในรูปถ่านผลไม้ ที่สามารถใช้เป็นผลิตภัณฑ์ดูดกลิ่นในครัวเรือน หรือ การนำผงถ่านชีวภาพไปผสมแป้งเปียกแล้วขึ้นรูปเป็นก้อนตามรูปร่าง

ที่ต้องการ จากนั้นตากให้แห้ง สามารถใช้เป็นถ่านดูดกลิ่นได้เช่นเดียวกัน ซึ่งผลิตภัณฑ์ดังกล่าวมีราคาสูงกว่าการขายถ่านอัดแท่งเพียงอย่างเดียว รวมถึงการนำผงถ่านชีวภาพมาผสมกับดินปลูกเพื่อปรับปรุงคุณภาพของดิน ได้แก่ ปริมาณอินทรียวัตถุในดิน ความเป็นกรด-ด่างของดิน ฟอสฟอรัสที่เป็นประโยชน์ และปริมาณโพแทสเซียม เป็นต้น [23] อย่างไรก็ตาม ปริมาณการใช้ถ่านชีวภาพในการปรับปรุงดินเพื่อการเกษตรนั้นจำเป็นต้องมีการศึกษาอัตราส่วนที่เหมาะสม เนื่องจากประสิทธิภาพของการปรับปรุงดินนั้นขึ้นกับองค์ประกอบของถ่านชีวภาพ และสภาพของดินด้วย [24] นอกเหนือจากการใช้ถ่านชีวภาพในด้านการปรับปรุงคุณภาพของดินนั้น อาจมีการนำถ่านชีวภาพไปพัฒนาเป็นผลิตภัณฑ์อื่นที่มีมูลค่ามากขึ้น เช่น การนำไปกระตุ้นเพื่อผลิตถ่านกัมมันต์ หรือการนำไปใช้เป็นตัวดูดซับซึ่งเป็นการดำเนินงานในขั้นตอนต่อไป

อย่างไรก็ตามการนำใบสับปะรดมาใช้เป็นวัตถุดิบนั้น หากนำใบสับปะรดตัดสดมาใช้พบว่าจะใช้เวลานานในการผลิต และสิ้นเปลืองเชื้อเพลิงเนื่องจากใบสับปะรดมีน้ำเป็นองค์ประกอบสูง ดังนั้นชุมชนจำเป็นต้องมีพื้นที่สำหรับนำใบสับปะรดมาเก็บไว้เพื่อให้ใบสับปะรดมีปริมาณน้ำน้อยลง ซึ่งอาจเป็นประเด็นที่ควรคำนึงถึงและมีการเตรียมความพร้อมก่อนทำการผลิต

5) สรุปผลการวิจัย

การศึกษาการผลิตถ่านชีวภาพจากใบสับปะรดศรีราชามีการดำเนินการในสองส่วนหลัก ได้แก่ การศึกษาเบื้องต้นในระดับห้องปฏิบัติการ และการพัฒนาเตาเผามลพิษต่ำขนาด 200 ลิตร ในระดับภาคสนาม สำหรับการศึกษาเบื้องต้นพบว่าถ่านชีวภาพจากการเผาที่ 400 องศาเซลเซียสและเวลา 1 ชั่วโมง มีค่าความร้อนสูงสุดเท่ากับ 23.25 เมกะจูลต่อกิโลกรัม และมีร้อยละผลได้ของแข็งเท่ากับ 33.02 จากนั้นได้พัฒนาเตาเผาจากถังน้ำมันขนาด 200 ลิตร ซึ่งประกอบด้วย 3 ส่วนหลัก ได้แก่ เตาเผาเพื่อผลิตถ่านชีวภาพ เตาเผาไหม้เชื้อเพลิง และชุดท่อควบแน่น โดยที่ด้านข้างของเตาเผาเพื่อผลิตถ่านชีวภาพมีการต่อท่อนำแก๊สเพื่อให้ไอน้ำ สารระเหย และน้ำมันทาร์ ไหลออกมาเข้าชุดท่อควบแน่น จากนั้นแก๊สที่เหลือจากการควบแน่นจะไหลเข้าเตาเผาไหม้เชื้อเพลิงเพื่อทำหน้าที่เป็นเชื้อเพลิงอีกทางหนึ่ง ทำให้แก๊สที่ปล่อยออกมามีสารประกอบไฮโดรคาร์บอนต่ำกว่าเตาเผาแบบอื่น ถ่านชีวภาพที่ได้จากเตาเผานี้มีค่าความร้อนสูงสุด 25.97 เมกะจูลต่อกิโลกรัม นอกจากนี้ยังได้น้ำส้มควันไม้เป็นผลิตภัณฑ์พลอยได้ จากการวิเคราะห์องค์ประกอบพบว่ามี ไกวเอคอล (Guaiacol) ฟีนอล (Phenol) และอนุพันธ์ของฟีนอล (Phenol Derivatives) อื่น ๆ ซึ่งมีฤทธิ์เป็นสารต้านเชื้อราและสารต้านเชื้อแบคทีเรีย โดยผลิตภัณฑ์พลอยได้นี้ตรงกับความต้องการของเกษตรในพื้นที่เป้าหมายเช่นเดียวกับถ่านชีวภาพ เนื่องจากมีเกษตรกรบางส่วนซื้อน้ำส้มควันไม้เพื่อใช้ในพื้นที่เพาะปลูก จึงมีแนวคิดในการนำใบสับปะรดที่เหลือทิ้งมาผลิตเป็นน้ำส้มควันไม้ซึ่งช่วยลดต้นทุนการผลิตลงส่วนหนึ่ง จากนั้น เมื่อวิเคราะห์ความเป็นไปได้ในสร้างผลิตภัณฑ์จากเตาเผาขนาด 200 ลิตร พบว่านอกจากถ่านชีวภาพและน้ำส้มควันไม้แล้ว ยังสามารถผลิตเป็นผงถ่านชีวภาพและถ่านผลไม้ได้ด้วย จึงเป็นทางเลือกในการสร้างผลิตภัณฑ์สำหรับชุมชน รวมถึงเตาเผานี้มีการเผาแก๊สส่วนที่เหลือจากการผลิตถ่านชีวภาพซ้ำ ลักษณะไม่ซับซ้อน ต้นทุนไม่สูงมาก ชุมชนสามารถนำไปเป็นต้นแบบหรือพัฒนาต่อยอดต่อไปได้

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยจากกองทุนส่งเสริมวิทยาศาสตร์ วิจัยและนวัตกรรม (ววน.) ประจำปีงบประมาณ พ.ศ. 2565 คณะผู้วิจัยขอขอบคุณผู้ประสานงานจากสถาบันวิจัยและพัฒนา มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก และผู้ทรงคุณวุฒิที่ให้คำแนะนำในการดำเนินโครงการ ขอขอบคุณคณะวิศวกรรมศาสตร์และสถาปัตยกรรมศาสตร์ เขตพื้นที่อุเทนถวายที่ได้ให้ความอนุเคราะห์เครื่องมือและสถานที่ในการทำวิจัยระดับห้องปฏิบัติการ รวมถึงบุคลากรสายสนับสนุนที่กรุณาช่วยเหลือด้านเอกสาร ขอขอบคุณคุณปิยะ โกสินทรจิตต์ ประธานกลุ่มเกษตรกรผู้ปลูกสับปะรดชลบุรีสำหรับความอนุเคราะห์วัตถุดิบ ขอขอบคุณ คุณนัฏฐ์ธนินทุ์ นิลโนรีและครอบครัวในการอนุเคราะห์สถานที่บ้านสวนหมอสาคร สำหรับทดสอบการใช้งานเตาเผาระดับภาคสนาม ตลอดจนผู้เกี่ยวข้องที่ไม่ได้กล่าวนามทุกท่านที่ให้ความช่วยเหลือและสนับสนุนจนงานวิจัยนี้สำเร็จลุล่วงได้

REFERENCES

[1] Office of Agricultural Economics. “Pineapple industry: Plantation area (rai), harvested area (rai), yield (ton) and yield per rai (kg) in 2018 by district.” OAE.go.th. https://www.oae.go.th/assets/portals/1/fileups/prcaidata/files/pineapple%20dit%2061.pdf (accessed Sep. 12, 2023)

[2] B. Chomsawan, P. Yodsoontorn, and P. Srisom, “Utilization of Phulae pineapple waste for cattle feed,” (in Thai), *J. Graduate Stud. Bansomdejchaopraya Rajabhat Univ.,* vol. 6, no. 2, pp. 10–16, 2021.

[3] P. Moonsri and A. Hutem, "Development of wood vinegar production from local wisdom in Phetchabun province for organic agriculture utilization," (in Thai), *J. Sci. Technol. Phetchabun Rajabhat Univ.*, vol. 3, no. 1, pp. 13–31, 2018.

[4] W. Kalasee, "A wood vinegar machine for sawdust and wood dust," (in Thai), *SDU Res. J. Sci. Technol.*, vol. 1, no. 1, pp. 45–54, 2008.

[5] S. Robb and S. Joseph, "A report on the value of biochar and wood vinegar: Practical experience of users in Australia and New Zealand Version 1.2 – April 2020," Australia New Zealand Biochar Industry Group (ANZBIG), Tyagarah, Australia, Accessed: Jun. 1, 2022. [Online]. Available: https://anzbig.org/wp-content/uploads/2020/07/ANZBI-2020-A-Report-on-the-Value-of-Biochar-and-Wood-Vinegar-v-1.2.pdf

[6] M. Hagner, O.-P. Penttinen, K. Tiilikkala, and H. Setälä, "The effects of biochar, wood vinegar and plants on glyphosate leaching and degradation," *Eur. J. Soil Biol.*, vol. 58, pp. 1–7, 2013.

[7] T. Sataklang, P. Khongkrapan, and A. Intaniwet, "The study of fumeless charcoal retort with rectangular fin installation for thermal efficiency and energy saving optimization," (in Thai), in *Proc. 12th Conf. Energy Netw. Thailand (E-NETT)*, Phitsanulok, Thailand, Jun. 8–10, 2016, pp. 632–637.

[8] U. Ketsripongsa, J. Makaew, S. Painjud, T. Yutthasarnsenee, and P. Moonwundee, "High efficiency charcoal kiln models of community in Khok Klang sub-district, Lam Plai Mat district, Buriram province," in *Proc. Innov. Technol. Conf. (ITC2017)*, Surin, Thailand, Dec. 25–26, 2017, pp. 1–11.

[9] M. F. Mail, M. H. A. Rahman, N. A. Baharom, M. S. Shahrun, and M. R. M. Daud, "Development and performance of Pyrolyzer for production of biochar and wood vinegar from domestic agriculture farm waste," *Adv. Agricultural Food Res. J.,* vol. 2, 2021, Art. no. a0000212.

[10] T. Poowadin, P. Khantikomol, and M. Polsongkram, "Producing wood vinegar from neem wood by closed circulation gas pyrolysis kiln," (in Thai), *Srinakharinwirot Univ. (J. Sci. Technol.)*, vol. 10, no. 19, pp. 50–59, 2018.

[11] S. Kaewluan, P. Sriromreun, S. N. Jansri, P. Asadamongkon, and S. Chuichulcherm, "Charcoal and wood vinegar products from hedge bamboo by using 200 liters community charcoal kiln working with gasified biomass burner," *Rajabhat Chiang Mai Res J.*, vol. 22, no. 2, pp. 229–245, 2021.

[12] T. Amornsakchai and N. Kengkhetkit, "Pineapple leaf: Source of natural fiber that should not be ignored," *Srinakharinwirot Sci. J.*, vol. 30 no. 2, Art. no. 4821.

[13] J. Sabseree. "DOE: Central composite design." (in Thai), *For Quality*, vol. 16, no. 145, pp. 72–74, Nov. 2009.

[14] A. Demirbas and A. H. Demirbas, "Estimating the calorific values of lignocellulosic fuels," *Energy Explor. Exploit.,* vol. 22, pp. 135–143, 2004.

[15] A. Pimenta, M. Fasciotti, T. Monteiro, and K. Lima, "Chemical composition of Pyroligneous acid obtained from Eucalyptus GG100 Clone," *Molecules,* vol. 23, no. 2, p. 426, Feb. 2018, doi: 10.3390/molecules23020426.

[16] J. Escalante *et al.*, "Pyrolysis of lignocellulosic, algal, plastic, and other biomass wastes for biofuel production and circular bioeconomy: A review of thermogravimetric analysis (TGA) approach," *Renew. Sustain. Energy Rev.,* vol. 169, 2022, Art. no 112914.

[17] D. Angin, "Effect of pyrolysis temperature and heating rate on biochar obtained from pyrolysis of safflower seed press cake," *Bioresour. Technol.*, vol. 128, pp. 593–597, Jan. 2013.

[18] A. Tomczyk, Z. Sokołowska, and P. Boguta, "Biochar physico chemical properties: Pyrolysis temperature and feedstock kind effects," *Rev. Environ. Sci. Biotechnol.*, vol. 19, pp. 191–215, 2020.

[19] S. Pap *et al.*, "Optimising production of a biochar made from conifer brash and investigation of its potential for phosphate and ammonia removal," *Ind. Crops. Prod.*, vol. 185, Oct. 2022, Art no. 115165.

[20] R. T. Kapoor and M. P. Shah, *Designer Biochar Assisted Bioremediation of Industrial Effluents, A Low-Cost Sustainable Green Technology*. Boca Raton, FL, USA: CRC Press, 2023.

[21] E. de Souza Araujo *et al.*, "Antibacterial and antifungal activities of pyroligneous acid from wood of *Eucalyptus urograndis* and *Mimosa tenuiflora*," *J. Appl. Microbiol.*, vol. 124, no. 1, pp. 85–96, 2017.

[22] H. A. Oramahi, T. Yoshimura, F. Diba, D. Setyawati, and Nurhaida, "Antifungal and Antitermitic activities of wood vinegar from oil palm trunk," *J. Wood Sci.*, vol. 64, pp. 311–317, 2018.

[23] S. Udpuay, C. Kueanoon, P. Chouwanapong, and P. Wiwatwittaya, "Effects of biochar from cassava root on chemical properties of clayey soil under greenhouse condition," (in Thai), in *Proc. 54th Kasetsart Univ. Annu. Conf.*, Bangkok, Thailand, Feb., 2016, pp. 328–335.

[24] S. Butnan and P. Vityakon, "Amazing biochar and its bipolar effects," (in Thai), *Khon Kaen Agr. J.*, vol. 46, no. 6, pp. 1167–1176, 2018.

การจำแนกพันธุ์ข้าวจากภาพของเมล็ดข้าวสารด้วยวิธีการตรวจหาวัตถุ

วุฒิชัย วัชรารัตน์[1] เจษฎา ตัณฑนุช[2*]

[1,2]สาขาวิชาคณิตศาสตร์ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, นครราชสีมา, ประเทศไทย*

*ผู้ประพันธ์บรรณกิจ อีเมล : jessada@g.sut.ac.th

รับต้นฉบับ : 25 เมษายน 2566; รับบทความฉบับแก้ไข : 28 พฤษภาคม 2566; ตอบรับบทความ : 29 พฤษภาคม 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

งานวิจัยนี้ได้ประยุกต์ใช้วิธีการตรวจหาวัตถุด้วยตัวแบบ YOLOv5 มาช่วยในการจำแนกพันธุ์ข้าวจากภาพของเมล็ดข้าวสารของพันธุ์ข้าวคาราก้าดาก หอมมะลิ ยิปซาลา บาสมาติ และอาโบริโอ การดำเนินการวิจัยได้แบ่งเป็นสามส่วนหลัก ได้แก่ การดำเนินการด้านวิศวกรรมข้อมูล ซึ่งเป็นการพัฒนาโปรแกรมด้วยภาษาไพทอนเพื่อใช้ในการจัดเตรียมข้อมูลเพื่อให้ระบบปัญญาประดิษฐ์เรียนรู้ การดำเนินการด้านวิทยาการข้อมูลโดยใช้โปรแกรมภาษาไพทอนร่วมกับระบบประมวลผล Google Colaboratory ทำการตรวจหาเมล็ดข้าวสาร และออกแบบการประเมินความแม่นยำของตัวแบบ ในส่วนของการจัดเตรียมข้อมูลได้นำรูปเมล็ดข้าวสารเมล็ดเดี่ยวชนิดเจเพ็กจาก https://www.muratkoklu.com/datasets/ มาทำการลดสัญญาณรบกวน กำจัดพื้นหลัง และแปลงเป็นภาพชนิดพีเอ็นจี แล้วนำมาใส่ในภาพขนาด 800 x 800 จุดภาพ จำนวนภาพละ 20–64 เมล็ด โดยมีทั้งรูปแบบไม่มีภาพเมล็ดข้าวสารซ้อนทับกัน มีภาพเมล็ดข้าวสารซ้อนทับกันร้อยละ 5, 10, 15, 20 และ 25 จากนั้นนำภาพที่ไม่มีเมล็ดข้าวสารซ้อนทับกันไปฝึกตัวแบบ YOLOv5 แล้วใช้ตัวแบบดังกล่าวจำแนกพันธุ์ข้าวและระบุตำแหน่งต่าง ๆ ของเมล็ดข้าวสารที่ปรากฏในภาพ ผลการวิจัยพบว่าตัวแบบ YOLOv5 สามารถจำแนกพันธุ์ข้าวของเมล็ดข้าวสารทั้ง 5 สายพันธุ์ได้ดี โดยการประเมินความแม่นยำของตัวแบบที่ค่าขีดแบ่ง 0.6 พบว่าตัวแบบสามารถจำแนกพันธุ์ข้าวที่ปรากฏในภาพที่มีภาพเมล็ดข้าวสารซ้อนทับกันร้อยละ 0, 5, 10, 15, 20 และ 25 ได้ถูกต้องร้อยละ 99.13, 99.00, 98.62, 98.19, 97.56 และ 96.89 ตามลำดับ

คำสำคัญ : การประมวลผลภาพ เมล็ดข้าวสาร การตรวจหาวัตถุ การจำแนกพันธุ์ข้าว

Classification of Rice Varieties from Milled Rice Grain Images by Object Detection Method

Wuttichai Watchararat[1] Jessada Tanthanuch[2*]

[1,2*]*School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand*

*Corresponding Author. E-mail address: jessada@g.sut.ac.th

Received: 25 April 2023; Revised: 28 May 2023; Accepted: 29 May 2023
Published online: 27 December 2023

Abstract

This research has applied the YOLOv5 object detection model to help classify rice varieties from images of milled rice grains from the following varieties: Karacadag, Jasmine, Ipsala, Basmati, and Arborio. The research was divided into three main parts: data engineering, which involved developing a Python program to prepare data for artificial intelligence learning; data science operations using Python programming in conjunction with Google Colaboratory for milled rice grain detection; and the development of model accuracy evaluation method. In the data preparation phase, single-grain JPEG images were obtained from https://www.muratkoklu.com/datasets/, and noise reduction, background removal, and conversion to PNG format were performed. These images were then placed into 800 x 800 pixels images, each containing 20–64 grains, with varying degrees of overlapping: no overlap, and 5, 10, 15, 20, and 25 percent overlap. The non-overlapping images were used to train the YOLOv5 model, which was then used to classify rice varieties and identify the locations of various milled rice grains in the images. The research results showed that the YOLOv5 model could effectively classify all five rice varieties. Evaluating the model's accuracy at a threshold of 0.6, it was found that the model could correctly classify rice varieties in images with 0, 5, 10, 15, 20, and 25 percent grain overlap, with accuracy rates of 99.13, 99.00, 98.62, 98.19, 97.56, and 96.89 percent, respectively.

***Keywords*:** Image processing, Milled rice grain, Object detection, Rice varieties

1) บทนำ

ในช่วงปี พ.ศ. 2565 ถึง 2566 ประเทศไทยมีการส่งออกข้าวไปยังประเทศต่าง ๆ ในโลกเป็นจำนวนถึง 8.2 ล้านตัน นับว่าเป็นผู้ส่งออกรายใหญ่มากที่สุดเป็นอันดับสองของโลก [1] โดยทั่วไปข้าวที่ส่งออกจะเป็นข้าวหอมมะลิไทย ข้าวหอมไทย ข้าวสาร 5 เปอร์เซ็นต์ ข้าวสาร 25 เปอร์เซ็นต์ ข้าวสาร 100 เปอร์เซ็นต์ ปลายข้าว ข้าวเหนียวขาว 10 เปอร์เซ็นต์ และ ข้าวนึ่ง 100 เปอร์เซ็นต์ [2] ทั้งนี้เพื่อป้องกันการปลอมปนข้าวที่มีมูลค่าต่ำกว่าเข้ากับข้าวที่มีมูลค่าสูงกว่ากระทรวงพาณิชย์ได้กำหนดมาตรฐานของข้าวชนิดต่าง ๆ ไว้โดยพิจารณาหลายองค์ประกอบ เช่น การพิจารณาจากความชื้น ลักษณะโดยทั่วไปของเมล็ดข้าว ขนาดของเมล็ด สีของเมล็ด คุณสมบัติทางเคมี รวมไปถึงร้อยละของการปลอมปนของเมล็ดข้าวชนิดอื่น ๆ [3] อย่างไรก็ตามหากมีการพัฒนาระบบการตรวจสอบคุณภาพของข้าวโดยใช้องค์ความรู้ด้านการประมวลผลภาพและปัญญาประดิษฐ์ จะช่วยให้การตรวจสอบมาตรฐานของข้าวชนิดต่าง ๆ เป็นไปได้อย่างมีประสิทธิภาพทั้งในด้านความแม่นยำและความสะดวกรวดเร็วที่มากขึ้น

งานวิจัยนี้ต้องการพัฒนาและประยุกต์ระบบการตรวจหาวัตถุเพื่อจำแนกพันธุ์ข้าวและระบุตำแหน่งเมล็ดพันธุ์ข้าวนั้น ๆ จากรูปที่ปรากฏเมล็ดข้าวสารของพันธุ์ข้าวชนิดต่าง ๆ ที่หลากหลาย แต่เนื่องด้วยไม่มีข้อมูลภาพของเมล็ดข้าวสารที่เป็นที่นิยมหรือจำหน่ายในประเทศไทยเป็นจำนวนมากพอ จึงต้องใช้ข้อมูลภาพเมล็ดสารที่เป็นข้อมูลเปิดมาใช้ในการทำวิจัยเพื่อใช้เป็นต้นแบบในการทำวิจัยเกี่ยวกับการจำแนกเมล็ดข้าวสารอื่น ๆ ต่อไป ซึ่งงานวิจัยนี้แตกต่างจากงานวิจัยอื่น ๆ ที่ใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยทั่วไป เช่น Logistic Regression, Support Vector Machine, Decision Trees หรือ K-Nearest Neighborhood เพื่อจำแนกประเภทของเมล็ดข้าว ซึ่งทั้งหมดเป็นการจำแนกเมล็ดข้าวสารเมล็ดเดี่ยวเท่านั้น แต่การดำเนินการวิจัยดังกล่าวไม่สอดคล้องกับลักษณะความเป็นจริงที่เราต้องการตรวจหาการปลอมปนของเมล็ดข้าวอื่นในภาพถ่ายที่มีรูปของเมล็ดข้าวสารที่เราพิจารณา และการใช้ระบบตรวจหาวัตถุนี้จะสามารถนำไปต่อยอดได้ในอนาคตเนื่องจากเมื่อทราบข้อมูลของตำแหน่งเมล็ดข้าวในอนาคตอาจสามารถพัฒนาเครื่องมือเพื่อใช้ในการแยกเมล็ดข้าวที่ปลอมปนนั้นได้ ทั้งนี้ในการวิจัยที่นำเสนอได้พยายามจำลองเงื่อนไขการตรวจหาเมล็ดข้าวสารทั้งในกรณีที่มีเมล็ดข้าวสารกระจายอยู่ทั่วไปรวมไปถึงกรณีการเกิดภาพของเมล็ดข้าวสารซ้อนทับกัน เพื่อให้ใกล้เคียงกับสถานการณ์จริงที่อาจจะเกิดขึ้นในธรรมชาติได้ และทำการวัดประสิทธิภาพระบบการตรวจหาเมล็ดข้าวสารที่สร้างขึ้น

2) ทบทวนวรรณกรรม

2.1) ทฤษฎีที่เกี่ยวข้อง

เนื้อหาในส่วนทฤษฎีที่เกี่ยวข้องนี้จะแบ่งเป็น 3 ส่วนหลัก ได้แก่ ทฤษฎีที่เกี่ยวกับการประมวลผลภาพดิจิทัล ตัวแบบตรวจหาวัตถุ และตัวแบบ YOLOv5

2.1.1) ทฤษฎีที่เกี่ยวข้องกับการประมวลผลภาพดิจิทัล ภาพถ่ายดิจิทัลในที่นี้จะหมายถึงภาพที่บันทึกในรูปแบบข้อมูลดิจิทัล ซึ่งในงานวิจัยนี้ ได้พิจารณาภาพถ่ายที่บันทึกในรูปแบบเจเพ็ก (Joint Photographic Experts Group: JPEG) และรูปแบบพีเอ็นจี (Portable Network Graphics: PNG) เป็นหลัก โดยภาพเจเพ็กจะเป็นภาพที่เป็นที่นิยมเนื่องจากโดยทั่วไปมีขนาดของไฟล์ภาพเล็กกว่าไฟล์ภาพชนิดอื่น ๆ เนื่องจากมีการบีบอัดข้อมูลที่มีการสูญเสียรายละเอียดของภาพในบางส่วน (lossy compression) แต่ก็ยังคงคุณภาพของภาพที่ยอมรับได้ ขณะที่ไฟล์ภาพรูปแบบพีเอ็นจีมีการบีบอัดข้อมูลแบบไม่เสียคุณภาพของภาพต้นฉบับ (loss less compression) และมีจุดเด่นในการสร้างภาพที่มีลักษณะพื้นหลังโปร่งใส (transparent background) เหมาะกับการนำภาพมาซ้อนทับกับภาพอื่น ๆ [4] ทั้งนี้เพื่อให้การดำเนินการที่เกี่ยวข้องกับการประมวลผลภาพ (image processing) เป็นไปได้อย่างมีประสิทธิภาพ เราสามารถประยุกต์ใช้ความรู้ด้าน Morphology ซึ่งเป็นแนวความคิดในการใช้ทฤษฎีเซต (set theory) ทางคณิตศาสตร์มาช่วยให้คอมพิวเตอร์พิจารณารูปเป็นวัตถุ แล้วดำเนินการด้านรูปทรงและขนาดของรูปโดยใช้ตัวดำเนินการ Morphological (morphological operation) [5]

ตัวอย่างการดำเนินการดังกล่าว เช่น

Erosion: ลดขนาดของวัตถุในภาพด้วยการลดจุดภาพบริเวณขอบของวัตถุ

Dilation: เพิ่มขนาดของวัตถุในภาพด้วยการเพิ่มจุดภาพบริเวณขอบของวัตถุ

Opening: ทำการ Erosion แล้วตามด้วย Dilation ช่วยให้สามารถลบสิ่งรบกวน (noise) บางชนิดออกได้

Closing: ทำการ Dilation แล้วตามด้วย Erosion ช่วยในการเติมเต็มจุดภาพที่ขาดหายไปบางส่วนของรูปภาพ

Contour: เป็นกระบวนการหาขอบเขตของวัตถุในภาพโดยการพิจารณาเส้น contour ของขอบภาพของวัตถุ

การนำตัวดำเนินการ Morphological มาใช้ในการกำจัดสัญญาณรบกวนและการเลือกวัตถุ มีผลต่อการจัดเตรียมภาพดิจิทัลเพื่อให้ระบบการตรวจหาวัตถุ YOLO ทำงานได้อย่างมีประสิทธิภาพต่อไป

2.1.2) ตัวแบบตรวจหาวัตถุ (Object Detectors) เป็นตัวแบบ (model) ทางด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) ที่ใช้ในงานทางด้านวิธีการรับรู้ภาพด้วยคอมพิวเตอร์ (computer vision) ซึ่งสามารถจำแนกและตรวจหาวัตถุที่อยู่ในภาพ รวมทั้งการบอกถึงตำแหน่งของวัตถุที่อยู่ในภาพนั้น ๆ โดยทั่วไปเราสามารถแบ่งตัวแบบตรวจหาวัตถุเป็น 3 ประเภทเบื้องต้นได้แก่ ประเภท single stage detector ประเภท two stage detector และประเภท transformer-based detector โดยตัวแบบตรวจหาวัตถุประเภท single stage detector และประเภท two stage detector มักใช้โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolution Neural Network: CNN) เป็นโครงสร้างหลักในการตรวจหา เนื่องด้วย CNN เป็นตัวแบบที่มีความสามารถในการจำแนกข้อมูลประเภทรูปภาพได้ดี ตัวแบบ CNN ที่ถูกใช้ในตัวแบบตรวจหาวัตถุประเภท single และ two stage detector มักมีความซับซ้อนและจำนวนพารามิเตอร์ที่น้อยกว่าตัวแบบ transformer ซึ่งถูกใช้ในตัวแบบตรวจหาวัตถุประเภท transformer-based ทำให้ใช้เวลาในการตรวจหาวัตถุน้อยแต่ก็มีความแม่นยำที่น้อยกว่าด้วย

ตัวแบบตรวจหาวัตถุประเภท two stage detector มีการแยกกระบวนการตรวจหาวัตถุออกเป็นสองขั้นตอนหลัก คือการสกัดคุณลักษณะ (feature extraction) และการทำนายพื้นที่ที่น่าจะเป็นบริเวณของวัตถุที่สนใจ (region proposals) ในขั้นตอนแรก จากนั้นทำการตรวจหาวัตถุจากบริเวณที่สนใจในขั้นตอนที่สอง ในขณะที่ตัวแบบตรวจหาวัตถุประเภท single stage detector ทำการรวบทั้งกระบวนการสกัดคุณลักษณะ การทำนายพื้นที่ และการตรวจหาวัตถุ ให้อยู่ในขั้นตอนเดียวเท่านั้น ซึ่งวิธีการนี้ทำให้มีความเร็วในการประมวลผลสูง สามารถตรวจหาวัตถุได้อย่างรวดเร็ว เหมาะสำหรับการตรวจหาวัตถุแบบเวลาจริง (real time)

ในปัจจุบันมีการปรับปรุงตัวแบบตรวจหาวัตถุประเภท single stage detector ให้มีความเร็วและความแม่นยำในการตรวจหาวัตถุมากขึ้น ตัวอย่างตัวแบบประเภทนี้ที่ถูกพัฒนา ได้แก่ ตัวแบบ Single Shot MultiBox Detector (SSD), RetinaNet, CenterNet, EfficientDet และ YOLO แต่ตัวแบบที่เป็นที่นิยมมากที่สุดและมีการพัฒนาอย่างต่อเนื่องได้แก่ตัวแบบ YOLO [6]

2.1.3) ตัวแบบ YOLOv5 ในปี ค.ศ. 2015 Joseph Redmon *et al.* ได้พัฒนาระบบตรวจหาวัตถุที่มีความรวดเร็วในการทำงานสูงกว่าระบบแบบเดิมและมีความแม่นยำที่ดี เนื่องด้วยใช้การประมวลผลโครงข่ายประสาทเทียมเดี่ยว (single neural network) ในการพยายามจำแนกวัตถุในภาพทั้งหมดในครั้งเดียว โดยตั้งชื่อขั้นตอนวิธีดังกล่าวว่า You Only Look Once Algorithm หรือ เป็นที่รู้จักในชื่อว่า YOLO algorithm ในภายหลังได้มีการพัฒนา YOLO algorithm อีกหลายครั้งจนกระทั่ง Glenn Jocher และคณะ ได้พัฒนาระบบตรวจหาวัตถุ YOLOv5 ขึ้นในปี ค.ศ. 2020 โดยต่อยอดแนวคิดมาจาก YOLO algorithm ซึ่ง YOLOv5 นี้มีประสิทธิภาพที่ดีกว่าเดิมเป็นอย่างมาก จนทำให้เป็นที่นิยมในการใช้ระบบดังกล่าวในการตรวจหาวัตถุทั้งจากภาพนิ่งและภาพเคลื่อนไหว [7]

YOLOv5 มีขั้นตอนในการทำงานเบื้องต้นดังนี้

1) การแบ่งภาพเป็นหลายส่วนย่อย (Dividing an Image into Multiple Segments) ในขั้นตอนนี้ระบบจะทำการแบ่งภาพออกเป็นส่วนย่อยที่เรียกว่า grid cells แล้วทำการตรวจหาวัตถุในแต่ละส่วนย่อยด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมแบบ CNN

2) การสกัดคุณลักษณะของวัตถุ (Feature Extraction) ในขั้นตอนนี้จะใช้ CNN ในการสกัดคุณลักษณะของวัตถุที่ปรากฏอยู่ในส่วนย่อยจากขั้นตอนที่ 1

3) การคาดเดาแบบชั้น (Class Prediction) ในการใช้ CNN จะมีชั้นเชื่อมโยงแบบสมบูรณ์ (fully-connected layer) ซึ่งชั้นนี้จะนำคุณลักษณะของวัตถุที่สกัดได้ในขั้นตอนที่ 2 ไปดำเนินการสร้างผลลัพธ์ให้เป็นเวกเตอร์ (vector) ของความน่าจะเป็นจำนวน K มิติ ตามจำนวนชั้น K ชั้น ที่ CNN ต้องการทำนาย

4) การคาดเดาตำแหน่งของวัตถุ (Bounding Box Prediction) ในขั้นตอนนี้ตัวแบบ YOLOv5 ใช้ชั้นเชื่อมโยงแบบสมบูรณ์คาดเดาตำแหน่งของวัตถุ

5) การปรับปรุงผลลัพธ์ (Result Improvement) หลังจากที่โมเดล YOLOv5 ได้ทำการคาดเดาตำแหน่งของวัตถุแล้ว ขั้นตอนต่อไปคือการดำเนินการเพื่อลดข้อผิดพลาดที่อาจเกิดขึ้นในการตรวจหาวัตถุ เช่น ปรับลดข้อมูลวัตถุที่ถูกตรวจหาที่มีความน่าจะเป็นต่ำ การลบข้อมูลรบกวน เช่น ตัวอักษรที่ไม่เกี่ยวข้อง การรวม (merging) กลุ่มวัตถุที่อยู่ใกล้เคียงกันเป็นวัตถุชิ้นเดียวกัน

6) การค้นหาวัตถุในรูปภาพ (Object Searching) ขั้นตอนนี้เป็นขั้นตอนในการตรวจหาวัตถุและระบุตำแหน่ง โดยการสร้าง

กรอบสี่เหลี่ยม (bounding box) รอบวัตถุที่ตรวจพบ นอกจากนี้สามารถแสดงข้อมูลที่เกี่ยวข้องกับวัตถุดังกล่าวตามต้องการ เช่น ชื่อของวัตถุ หรือคะแนนความเชื่อมั่นที่พบ (confidence score)

2.2) งานวิจัยที่เกี่ยวข้อง

ที่ผ่านมามีงานวิจัยหลากหลายที่ประยุกต์ใช้ระบบปัญญาประดิษฐ์เพื่อช่วยในงานที่เกี่ยวข้องกับการจำแนกพันธุ์ข้าว เช่น ในปี ค.ศ. 2015 Aki *et al.* ใช้วิธีการประมวลผลภาพร่วมกับเทคนิคการเรียนรู้เครื่อง (machine learning techniques) เพื่อจำแนกพันธุ์ข้าวจากรูปถ่ายภาพเมล็ดข้าวสาร 4 ชนิด ได้แก่ ข้าวพันธุ์บาลโด (Baldo) ข้าวพันธุ์ออสมันซิค (Osmancik) ข้าวพันธุ์ยาเซมิน (Yasemin) และกลุ่มของเม็ดข้าวหัก (broken) [8] และในปีเดียวกัน Zareiforoush *et al*. ได้ใช้วิธีการรับรู้ภาพด้วยคอมพิวเตอร์ ร่วมกับ เทคนิคเมตาฮิวริสติก (Metaheuristic Method) 4 อย่างได้แก่ เครือข่ายโครงประสาทประดิษฐ์ (Artificial Neural Network: ANN) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) ต้นไม้การตัดสินใจ (Decision Trees: DT) และ เครือข่ายเบย์เซียน (Bayesian Network) ในการวิเคราะห์เมล็ดข้าวสารที่ได้จากกรรมวิธีในการสีข้าว 4 ประเภท ได้แก่ เมล็ดข้าวสารคุณภาพดีได้จากกระบวนการสีที่ดี (high-processed sound grains) เมล็ดข้าวสารหักได้จากกระบวนการสีที่ดี (high-processed broken grains) เมล็ดข้าวสารคุณภาพดีได้จากกระบวนการสีที่คุณภาพไม่ดี (low-processed sound grains) เมล็ดข้าวสารหักได้จากกระบวนการสีที่คุณภาพไม่ดี (low-processed broken grains) [9] ในปี ค.ศ. 2019 Cinar และ Koklu ใช้เทคนิคการเรียนรู้เครื่อง 7 อย่างได้แก่ 1) การถดถอยโลจิสติก (Logistic Regression: LR) 2) การรับรู้หลายชั้น (Multilayer Perceptron: MP) 3) ซัพพอร์ตเวกเตอร์แมชชีน 4) ต้นไม้การตัดสินใจ 5) ป่าสุ่ม (Random Forest: RF) 6) นาอีฟเบส์ (Naïve Bayes: NB) และ 7) เพื่อนบ้านใกล้เคียงสุด k (k-Nearest Neighbor: kNN) ในการจำแนกข้าวสาร 2 สายพันธุ์คือ สายพันธุ์ออสมันซิคและสายพันธุ์คามีโอ (Cameo) จากรูปถ่าย 3,810 รูป โดยพิจารณาจากคุณลักษณะสัณฐาน (morphological feature) 7 อย่างได้แก่ พื้นที่ของรูปเมล็ดข้าวสารที่ปรากฏในภาพถ่าย (area) เส้นรอบรูปเมล็ดข้าวสาร (perimeter) ความยาวของแกนหลัก (major axis length) ความยาวของแกนรอง (minor axis length) ความบิดเบี้ยวของส่วนผิวเมล็ดข้าวสาร (eccentricity) พื้นที่ส่วนจมูกข้าว (convex area) และอัตราส่วนของพื้นที่รูปข้าวกับกรอบของภาพที่พิจารณา [10] และอีก 2 ปีถัดมา Cinar และ Koklu ได้พิจารณาข้าวเพิ่มเป็น 5 สายพันธุ์ ได้แก่ คาราก้าดาก (Karacadag) หอมมะลิ (Jasmine) ยิปซาลา (Ipsala) บาสมาติ (Basmati) และอาโบริโอ (Arborio) โดยเป็นข้อมูลรูปภาพเพิ่มเป็นสายพันธุ์ละ 15,000 รูป (รวม 75,000 รูป) คุณลักษณะเพิ่มขึ้นเป็น 106 คุณลักษณะโดยเป็นคุณลักษณะสัณฐาน 12 อย่าง คุณลักษณะรูปร่าง (shape feature) 4 อย่าง และคุณลักษณะด้านสี (color feature) 90 อย่าง และใช้เทคนิคการสกัดคุณลักษณะ (feature extraction) เพื่อเพิ่มประสิทธิภาพในการจำแนกพันธุ์เมล็ดข้าวสารจากรูปภาพ และเปรียบเทียบการจำแนกพันธุ์เมล็ดข้าวสารด้วย 3 ขั้นตอนวิธี ได้แก่ วิธีการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) วิธีไคกำลังสอง (Chi-square) และวิธีอัตราส่วนอัตราขยาย (gain ratio) [11] และในปีเดียวกัน Koklu *et al*. ได้พัฒนาโปรแกรมภาษาไพทอน (Python) เพื่อใช้ขั้นตอนวิธี ANN และเครือข่ายโครงประสาทเชิงลึก (Deep Neural Network: DNN) ในส่วนการจำแนกพันธุ์ข้าวจากคุณลักษณะ 106 คุณลักษณะของชุดข้อมูล เปรียบเทียบกับการใช้วิธีเครือข่ายโครงประสาทสังวัฒนาการ (Convolution Neural Network: CNN) ในส่วนของการพิจารณาคุณลักษณะและจำแนกเมล็ดข้าวสารของข้อมูลข้าวจากภาพ ซึ่งพบว่าการใช้ CNN ให้ประสิทธิภาพที่สูงกว่า [12] และในปี ค.ศ. 2022 Cinar และ Koklu ใช้เทคนิคการเรียนรู้เครื่อง 7 อย่าง ได้แก่ LR, MP, SVM, DT, RF, NB และ kNN โดยเป็นทำงานภายใต้โปรแกรม MATLAB ในการจำแนกพันธุ์ข้าวจากข้อมูลคุณลักษณะ 106 คุณลักษณะ ของข้าวทั้ง 5 สายพันธุ์ที่ได้กล่าวถึงในงานวิจัยก่อนหน้านี้ [13]

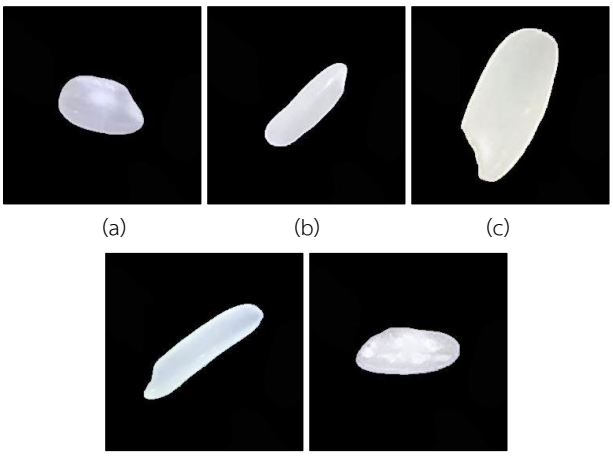
3) วิธีดำเนินการวิจัย

3.1) ข้อมูล ซอฟต์แวร์ และอุปกรณ์ที่ใช้ในการทำวิจัย

ข้อมูลภาพเมล็ดข้าวสารที่ใช้ในงานวิจัยนี้ได้มาจาก “murat koklu” ซึ่งเป็นข้อมูลสาธารณะของ Dr.Murat KÖKLÜ ซึ่งมีอยู่ในฐานข้อมูล https://www.muratkoklu.com/datasets ชุดข้อมูลประกอบด้วยภาพเมล็ดข้าวสารของพันธุ์ข้าวคาราก้าดาก หอมมะลิ ยิปซาลา บาสมาติ และอาโบริโอจำนวน 15,000 ภาพในแต่ละพันธุ์ รวมเป็นจำนวนภาพทั้งหมด 75,000 ภาพ โดยเป็นภาพชนิดเจเพ็กขนาด 250 x 250 จุดภาพ ดังแสดงในรูปที่ 1

สำหรับซอฟต์แวร์ที่ใช้ในการดำเนินการวิจัยจะประกอบไปด้วยโปรแกรมภาษาไพทอน Version 3.7 ร่วมกับระบบประมวลผล

แพลตฟอร์ม Google Colaboratory และดำเนินการในส่วนพัฒนาซอฟต์แวร์และประเมินประสิทธิภาพบนเครื่องคอมพิวเตอร์รุ่น ASUS-TUF Gaming FX505DU-AL052T หน่วยประมวลผล AMD Ryzen 7 3750H หน่วยประมวลผลกราฟิก Radeon Vega Mobile Gfx 2.30 GHz หน่วยความจำ RAM ความจุ 16GB ทำงานบนระบบปฏิบัติการ Microsoft Windows 11

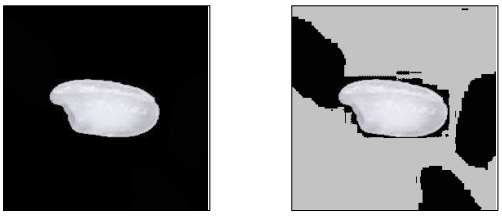


รูปที่ 1 : ตัวอย่างภาพเมล็ดข้าวสาร 5 สายพันธุ์
(a) คาราก้าดาก (b) หอมมะลิ (c) ยิปซาลา (d) บาสมาติ และ (e) อาโบริโอ

สำหรับขั้นตอนการดำเนินงานวิจัยแบ่งออกเป็น 3 ส่วนหลัก ได้แก่ 1) วิศวกรรมข้อมูล 2) การสร้างตัวแบบ และ 3) การวัดประสิทธิภาพของตัวแบบ

3.2) วิศวกรรมข้อมูล

จากการวิเคราะห์ภาพที่ได้ในการนำเข้าข้อมูล พบว่าบริเวณพื้นหลังของภาพเมล็ดข้าวสารไม่ได้มีสีดำสนิท มีการเหลื่อมของสีบางส่วนอยู่ดังแสดงในรูปที่ 2



รูปที่ 2 : ตัวอย่างภาพเมล็ดข้าวสารที่ใช้ในการทำวิจัย (ซ้าย) และ
ภาพแสดงให้เห็นว่าภาพหลังของภาพเมล็ดข้าวสารไม่เป็นสีดำสนิท (ขวา)

จากกรณีข้างต้นคณะผู้วิจัยได้ดำเนินการแก้ปัญหาด้วยการพัฒนาโปรแกรมภาษาไพทอนเพื่อใช้ตัวดำเนินการ morphological ทั้ง opening และ closing ในการกำจัดสิ่งรบกวนที่ปรากฏบนภาพและเติมเต็มในบางส่วนของภาพที่เสียหาย และใช้การดำเนินการ contour เพื่อคัดเลือกวัตถุที่มีขนาดใหญ่ที่สุดในภาพ นั่นคือภาพเฉพาะส่วนเมล็ดข้าวสารออกมา จากนั้นทำการกำจัดส่วนของพื้นหลังของภาพทำให้ได้ภาพที่มีพื้นหลังโปร่งใส แล้วทำการบันทึกภาพที่ได้ให้เป็นภาพชนิดพีเอ็นจี ดังแสดงในรูปที่ 3

รูปที่ 3 : ภาพเมล็ดข้าวสารที่ลดสิ่งรบกวนและกำจัดพื้นหลัง

จากนั้นนำภาพเมล็ดข้าวสารที่พื้นหลังโปร่งใสมาซ้อนทับลงบนภาพพื้นหลังสีดำขนาด 800 x 800 จุดภาพ ซึ่งการดำเนินการดังกล่าวจะแบ่งการดำเนินการเป็น 2 ส่วนได้แก่ การสร้างชุดข้อมูลสำหรับการฝึก (training data set) และชุดข้อมูลสำหรับการทดสอบ (test data set)

ในส่วนการสร้างข้อมูลชุดสำหรับการฝึก จะซ้อนทับภาพเมล็ดข้าวสารจำนวนสุ่มระหว่าง 20 ถึง 35 เมล็ด ลงบนภาพสีดำที่เตรียมไว้ และภาพเมล็ดข้าวสารดังกล่าวได้จากการสุ่มภาพเมล็ดข้าวสาร 5 สายพันธุ์ จากจำนวนภาพข้าวสายพันธุ์ละ 6,000 ภาพ (รวม 30,000 ภาพ) ลงบนตำแหน่งต่าง ๆ แบบสุ่ม โดยต้องไม่มีภาพเมล็ดข้าวสารซ้อนทับกัน โดยดำเนินการสร้างข้อมูลชุดฝึกจำนวน 7,000 ภาพ รวมทั้งสร้างไฟล์ข้อมูลสำหรับชุดฝึกแต่ละภาพซึ่งมีรายละเอียดว่าในภาพข้อมูลชุดฝึกจะมีตำแหน่งเมล็ดข้าวสารชนิดใด อยู่ ณ ตำแหน่งไหน มีบริเวณที่ภาพเมล็ดข้าวสารนั้นครอบคลุมพื้นที่อยู่กว้างและยาวกี่จุดภาพ โดยเรียกไฟล์ข้อมูลดังกล่าวว่าไฟล์ข้อมูลติดฉลาก (labeled data file)

รูปที่ 4 : ภาพข้อมูลชุดฝึก

```
4 0.353125 0.174375 0.09125 0.14875
0 0.864375 0.47875 0.15625 0.12
0 0.468125 0.28 0.13875 0.125
4 0.160625 0.94625 0.12625 0.105
0 0.153125 0.26 0.18375 0.1025
0 0.79 0.8575 0.1425 0.155
2 0.89375 0.159375 0.155 0.25375
3 0.31125 0.7775 0.18 0.0675
3 0.525 0.64875 0.15 0.1175
1 0.454375 0.909375 0.21125 0.17625
4 0.10875 0.63875 0.145 0.0975
0 0.2975 0.439375 0.17 0.08625
4 0.491875 0.088125 0.09625 0.13625
1 0.7125 0.454375 0.0925 0.25375
3 0.446875 0.555625 0.19125 0.05875
2 0.150625 0.06625 0.22875 0.1225
2 0.72125 0.22625 0.1675 0.1775
0 0.514375 0.42125 0.16875 0.1
0 0.64625 0.708125 0.09 0.17125
1 0.263125 0.604375 0.13875 0.24375
3 0.12875 0.38 0.15 0.1075
2 0.874375 0.644375 0.22625 0.13625
4 0.935625 0.776875 0.10625 0.12875
2 0.69625 0.071875 0.25 0.11125
```

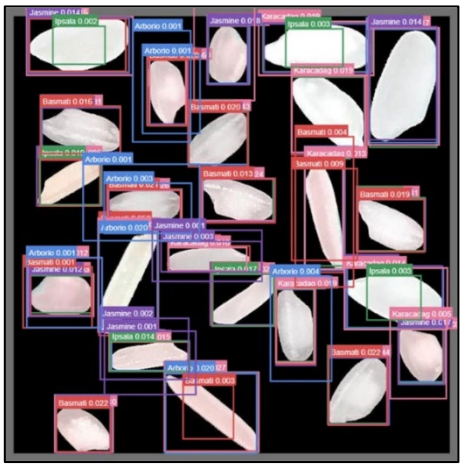
รูปที่ 5 : ภาพแสดงรายละเอียดไฟล์ข้อมูลติดฉลาก
ของภาพข้อมูลชุดฝึกจากรูปที่ 4

สำหรับในส่วนการสร้างข้อมูลชุดสำหรับการทดสอบ ได้ทำการสร้างข้อมูลในรูปแบบที่คล้ายกันกับข้อมูลชุดสำหรับการฝึก แต่เพิ่มจำนวนเมล็ดข้าวสารที่ปรากฏลงบนภาพชุดทดสอบเป็นจำนวนสุ่มตั้งแต่ 44 ถึง 64 เมล็ด และแบ่งชุดทดสอบเป็นกลุ่มชุดทดสอบที่มีภาพเมล็ดข้าวสารไม่ซ้อนทับกันเลย (ซ้อนทับกันร้อยละ 0) และกลุ่มชุดทดสอบที่อนุญาตให้มีภาพเมล็ดข้าวสารซ้อนทับกันได้ไม่เกินร้อยละ 5, 10, 15, 20 และ 25 แต่ละกลุ่มมีจำนวนข้อมูลทดสอบกลุ่มละ 500 ภาพ รวมชุดทดสอบทั้งหมดเป็น 3,000 ภาพ

รูปที่ 6 : ภาพข้อมูลชุดทดสอบที่มีภาพเมล็ดข้าวสารซ้อนทับกัน

3.3) สร้างตัวแบบตรวจหาวัตถุ

นำข้อมูลภาพข้อมูลชุดสำหรับการฝึก 7,000 ภาพ และไฟล์ข้อมูลติดฉลาก ไปฝึกระบบตรวจหาวัตถุ YOLOv5 บนระบบประมวลผลแพลตฟอร์ม Google Colaboratory โดยตั้งค่าพารามิเตอร์โดยปริยาย (default) ดังนี้ ค่าอัตราการเรียนรู้ (learning rate) ที่ 0.01 ค่า batch_size = 16 (จำนวนข้อมูลที่ส่งฝึกสอนในหนึ่งครั้ง) ค่า epoch = 5 (จำนวนรอบของการฝึกสอน)



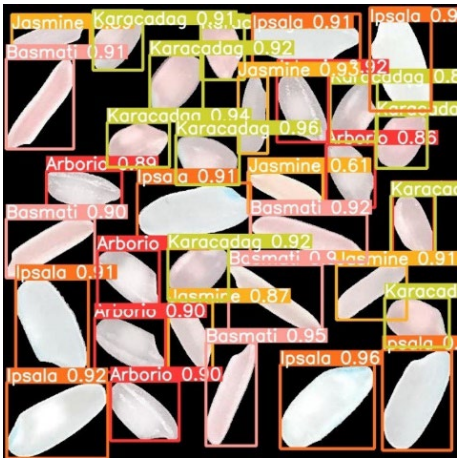
รูปที่ 7 : ภาพการฝึกเพื่อสร้างตัวแบบตรวจหาวัตถุ

ทั้งนี้เมื่อสร้างตัวแบบตรวจหาวัตถุเป็นที่เรียบร้อยแล้ว ตัวแบบดังกล่าวจะสามารถตรวจหาชนิดของเมล็ดข้าวสารและตำแหน่งในภาพข้อมูลชุดทดสอบ โดยตัวแบบจะดำเนินการสร้างไฟล์ข้อมูลติดฉลาก ซึ่งมีข้อมูลของสายพันธุ์ ตำแหน่ง และขนาดของเมล็ดข้าวสารแต่ละเมล็ดที่ปรากฏในภาพข้อมูลชุดทดสอบ

```
1 0.9625 0.718125 0.075 0.21375
3 0.94 0.9075 0.0675 0.175
0 0.4975 0.858125 0.165 0.08625
1 0.56375 0.9475 0.21 0.0675
0 0.219375 0.94875 0.15875 0.085
4 0.081875 0.299375 0.12875 0.11375
1 0.515 0.421875 0.235 0.07125
3 0.796875 0.798125 0.16875 0.07875
4 0.72125 0.939375 0.115 0.11375
0 0.25625 0.061875 0.175 0.09875
0 0.094375 0.5925 0.14875 0.115
0 0.088125 0.430625 0.14625 0.13125
2 0.734375 0.220625 0.16375 0.21375
0 0.39125 0.1575 0.165 0.115
4 0.821875 0.599375 0.13625 0.10875
0 0.3075 0.62125 0.1675 0.0875
2 0.63 0.059375 0.22 0.10875
```

รูปที่ 8 : ตัวอย่างข้อมูลบางส่วนของการตรวจหาเมล็ดข้าวสารด้วยตัวแบบ

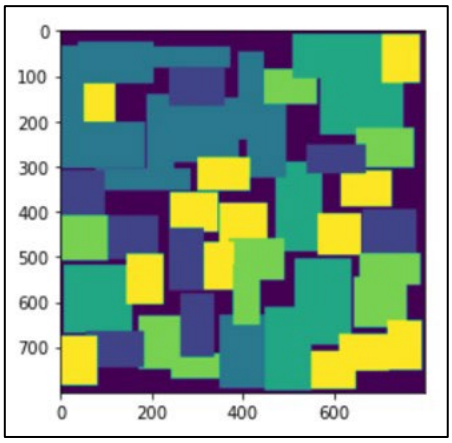
นอกจากนี้ ยังสามารถให้ตัวแบบแสดงการตรวจหาเมล็ดข้าวสารของภาพข้อมูลชุดทดสอบ โดยสามารถระบุสายพันธุ์ ตำแหน่ง ขนาด และความเชื่อมั่นที่พบ ดังแสดงในรูปที่ 9



รูปที่ 9 : แสดงการตรวจหาเมล็ดข้าวสารโดยตัวแบบตรวจหาวัตถุที่สร้างขึ้น

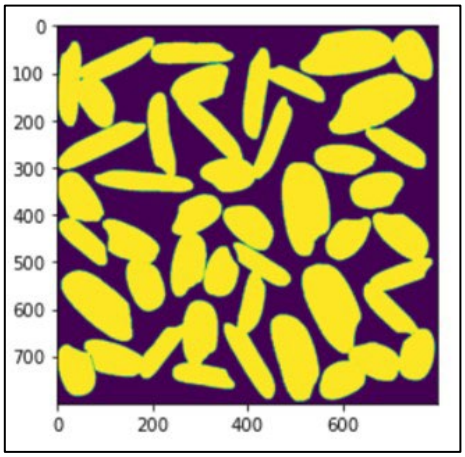
3.4) การวัดประสิทธิภาพของตัวแบบตรวจหาวัตถุ

ในการวัดประสิทธิภาพของตัวแบบตรวจหาวัตถุคณะผู้วิจัยได้ออกแบบการวัดประสิทธิภาพโดยนำข้อมูลจากไฟล์ข้อมูลติดฉลากของชุดทดสอบแต่ละภาพมาสร้างภาพเติมเต็มจากข้อมูลของสายพันธุ์ข้าวสาร ตำแหน่ง และบริเวณที่ภาพเมล็ดข้าวสารแต่ละเมล็ดนั้นครอบคลุม โดยทำการเติมเต็มบริเวณกรอบสี่เหลี่ยมผืนผ้าที่ระบุตำแหน่งของเมล็ดข้าวสารชนิดนั้น ให้แต่ละจุดภาพทั้งหมดในกรอบดังกล่าวมีค่าเป็นข้อมูลสายพันธุ์ของเมล็ดข้าวสารนั้น ๆ ทีละเมล็ด ดังแสดงในรูปที่ 10



รูปที่ 10 : ภาพเติมเต็ม

จากนั้น นำภาพข้อมูลชุดทดสอบซึ่งแสดงเมล็ดข้าวสารพันธุ์ต่าง ๆ แปลงเป็นภาพไบนารีโดยให้พื้นหลังมีค่า 0 และบริเวณที่เป็นส่วนแสดงเมล็ดข้าวสารให้มีค่าไม่เท่ากับ 0 เพื่อแยกระหว่างพื้นหลังกับส่วนที่เป็นวัตถุ ดังแสดงในรูปที่ 11



รูปที่ 11 : ภาพไบนารีของข้อมูลชุดทดสอบแสดงตำแหน่งเมล็ดข้าวสาร

ในส่วนถัดมานำข้อมูลภาพเติมเต็มและภาพไบนารีมาประมวลโดยสร้างภาพสำหรับการใช้วัดประสิทธิภาพตัวแบบโดยกำหนดให้ตำแหน่งพื้นหลังเป็น 0 และตำแหน่งที่แสดงเมล็ดข้าวสารเป็นค่าตามสายพันธุ์ของข้าวตามที่กำหนด

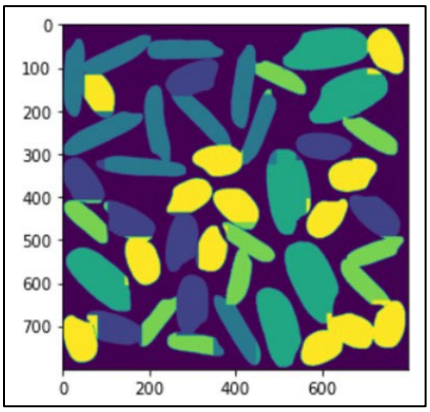
$$w(x,y) = \begin{cases} 0 & \text{, if } g(x,y) = 0 \\ h(x,y) & \text{, if } g(x,y) \neq 0 \end{cases} \quad (1)$$

เมื่อ $w(x,y)$ คือ ค่าข้อมูลของภาพสำหรับวัดประสิทธิภาพ

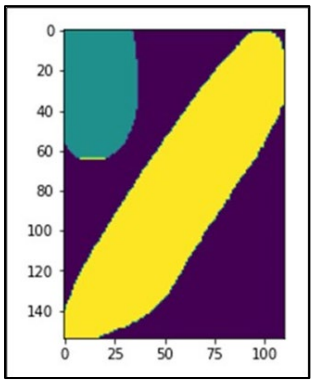
$g(x,y)$ คือ ค่าข้อมูลของภาพไบนารี และ

$h(x,y)$ คือ ค่าข้อมูลของภาพเติมเต็ม

โดยค่าต่าง ๆ พิจารณา ณ ตำแหน่งจุดภาพ (x,y)



รูปที่ 12 : ภาพสำหรับการใช้วัดประสิทธิภาพตัวแบบ



รูปที่ 13 : ตัวอย่างภาพย่อยจากภาพสำหรับการใช้วัดประสิทธิภาพตัวแบบ

สำหรับการวัดประสิทธิภาพของตัวแบบทำได้โดยการนำข้อมูลจากไฟล์ข้อมูลติดฉลากที่เป็นผลจากที่ตัวแบบทำการตรวจหาเมล็ดข้าวสาร ซึ่งมีข้อมูลพันธุ์ข้าวและบริเวณที่ตรวจหาได้บนภาพจากข้อมูลชุดทดสอบ มาทำการพิจารณาภาพย่อยจากภาพสำหรับการใช้วัดประสิทธิภาพตัวแบบซึ่งอยู่บริเวณเดียวกันดังแสดงตัวอย่างในรูปที่ 13

จะเห็นได้ว่าภาพย่อยดังกล่าวอาจมีภาพเมล็ดข้าวสารที่ตัวแบบตรวจหาวัตถุตรวจเจอและภาพบางส่วนจากเมล็ดข้าวสารเมล็ดอื่นร่วมอยู่ด้วย ดังนั้นเพื่อที่จะตรวจสอบว่า ณ ตำแหน่งดังกล่าว ตัวแบบได้ทำการตรวจหาพันธุ์ข้าวได้ถูกต้องหรือไม่ จะใช้วิธีพิจารณาอัตราส่วนของจำนวนจุดภาพในภาพย่อยของเมล็ดข้าวสารพันธุ์ที่ตัวแบบตรวจหาเทียบกับจำนวนจุดภาพที่อยู่บนตำแหน่งที่ปรากฏเมล็ดข้าวสารทั้งหมดที่ไม่ใช่ส่วนของพื้นหลังว่ามากกว่าหรือเท่ากับค่าขีดแบ่ง (threshold) ที่กำหนดหรือไม่ หากใช่จะถือว่าตัวแบบได้ตรวจหาพันธุ์เมล็ดข้าวสาร ณ ตำแหน่งดังกล่าวได้ถูกต้อง (correct)

$$\text{การตรวจหา} = \begin{cases} \text{correct}\ , & \text{if } \frac{r}{t} \geq \theta \\ \text{incorrect}, & \text{if } \frac{r}{t} < \theta \end{cases} \qquad (2)$$

เมื่อ r คือ จำนวนจุดภาพในภาพย่อยของเมล็ดข้าวสารพันธุ์ที่ตัวแบบตรวจหา

t คือ จำนวนจุดภาพทั้งหมดในภาพย่อยที่ไม่ใช่พื้นหลัง

θ คือ ค่าขีดแบ่งที่กำหนด

ทั้งนี้ค่าความแม่นยำ (accuracy) ของตัวแบบถูกนิยามโดย

$$\text{ความแม่นยำ} = \frac{\text{จำนวน correct}}{\text{จำนวน correct + จำนวน incorrect}} \qquad (3)$$

4) ผลการวิจัยและอภิปรายผล

ในการสร้างข้อมูลชุดทดสอบจำนวน 3,000 ภาพ มีข้อมูลจำนวนภาพเมล็ดข้าวสารสายพันธุ์ต่าง ๆ แบ่งตามกลุ่มชุดทดสอบที่อนุญาตให้มีภาพเมล็ดข้าวสารซ้อนทับกันได้ไม่เกินร้อยละ 0 (ไม่มีภาพเมล็ดข้าวสารซ้อนทับกันเลย), 5, 10, 15, 20 และ 25 ดังแสดงในตารางที่ 1 และตัวแบบตรวจหาวัตถุที่ได้ฝึกเพื่อตรวจหาเมล็ดข้าวสารสามารถตรวจพบเมล็ดข้าวสารได้ดังแสดงในตารางที่ 2

ตารางที่ 1 : จำนวนรวมเมล็ดข้าวสารในภาพข้อมูลชุดทดสอบทั้งหมด

	การทับซ้อน (ร้อยละ)					
	0	5	10	15	20	25
คาราก้าดาก	4,221	4,630	4,868	4,827	5,145	5,396
หอมมะลิ	4,010	4,580	4,516	4,725	5,105	5,135
ยิปซาลา	2,474	2,584	2,771	2,893	3,027	3,140
บาสมาติ	3,009	3,448	3,571	3,759	3,875	3,930
อาโบริโอ	3,799	4,145	4,295	4,348	4,637	4,882

ตารางที่ 2 : จำนวนเมล็ดข้าวสารที่ตรวจจับได้ในภาพข้อมูลชุดทดสอบทั้งหมด

	การทับซ้อน (ร้อยละ)					
	0	5	10	15	20	25
คาราก้าดาก	4,208	4,602	4,830	4,778	5,084	5,302
หอมมะลิ	4,047	4,630	4,568	4,792	5,179	5,214
ยิปซาลา	2,473	2,585	2,772	2,894	3,029	3,142
บาสมาติ	3,011	3,443	3,588	3,774	3,914	3,968
อาโบริโอ	3,889	4,259	4,447	4,562	4,877	5,194

ทั้งนี้เพื่อประเมินค่าขีดแบ่งที่เหมาะสมในการวัดความแม่นยำของตัวแบบ จะทดสอบโดยใช้ค่าขีดแบ่งทั้งหมด 6 ค่าได้แก่ 0.75, 0.70, 0.65, 0.60 และ 0.55 และผลการตรวจวัดความแม่นยำเรียงตามค่าขีดแบ่งทั้ง 6 ได้แสดงดังปรากฏในตารางที่ 3 ถึงตารางที่ 7 ตามลำดับ

ตารางที่ 3 : การวัดความแม่นยำของตัวแบบด้วยค่าขีดแบ่งเท่ากับ 0.75

การทับซ้อน (ร้อยละ)	จำนวน correct	จำนวน incorrect	ความแม่นยำ
0	17,418	210	0.9881
5	19,117	402	0.9794
10	19,579	626	0.9690
15	19,940	860	0.9586
20	20,805	1,278	0.9421
25	21,176	1,644	0.9280

ตารางที่ 4 : การวัดความแม่นยำของตัวแบบด้วยค่าขีดแบ่งเท่ากับ 0.70

การทับซ้อน (ร้อยละ)	จำนวน correct	จำนวน incorrect	ความแม่นยำ
0	17,459	169	0.9904
5	19,229	290	0.9851
10	19,757	448	0.9778
15	20,170	630	0.9697
20	21,145	938	0.9575
25	21,612	1,208	0.9471

ตารางที่ 5 : การวัดความแม่นยำของตัวแบบด้วยค่าขีดแบ่งเท่ากับ 0.65

การทับซ้อน (ร้อยละ)	จำนวน correct	จำนวน incorrect	ความแม่นยำ
0	17,471	157	0.9911
5	19,294	225	0.9885
10	19,859	346	0.9829
15	20,326	474	0.9772
20	21,390	693	0.9686
25	21,933	887	0.9611

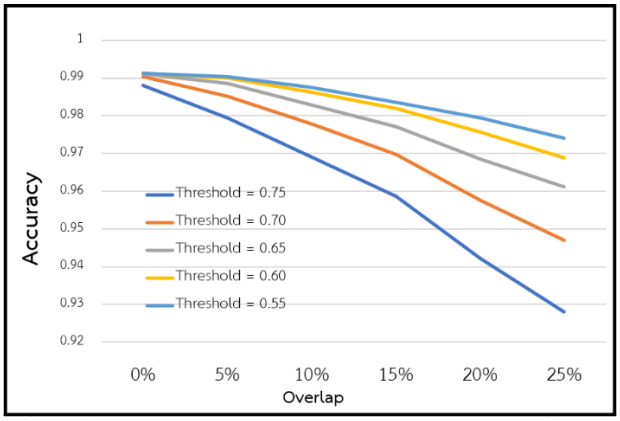
ตารางที่ 6 : การวัดความแม่นยำของตัวแบบด้วยค่าขีดแบ่งเท่ากับ 0.60

การทับซ้อน (ร้อยละ)	จำนวน correct	จำนวน incorrect	ความแม่นยำ
0	17,475	153	0.9913
5	19,324	195	0.9900
10	19,926	279	0.9862
15	20,423	377	0.9819
20	21,545	538	0.9756
25	22,111	709	0.9689

ตารางที่ 7 : การวัดความแม่นยำของตัวแบบด้วยค่าขีดแบ่งเท่ากับ 0.55

การทับซ้อน (ร้อยละ)	จำนวน correct	จำนวน incorrect	ความแม่นยำ
0	17,475	153	0.9913
5	19,332	187	0.9904
10	19,952	253	0.9875
15	20,460	340	0.9836
20	21,630	453	0.9795
25	22,228	592	0.9740

ข้อมูลความแม่นยำของตัวแบบ เมื่อประเมินด้วยค่าขีดแบ่งต่าง ๆ ที่แสดงในตารางที่ 3 ถึงตารางที่ 7 สามารถแสดงเป็นกราฟความแม่นยำได้ดังแสดงในรูปที่ 14



รูปที่ 14 : กราฟความแม่นยำของตัวแบบเมื่อตรวจวัดด้วยค่าขีดแบ่ง 0.75, 0.70, 0.65, 0.60 และ 0.55

5) สรุปผล

ในงานวิจัยนี้สามารถประยุกต์ใช้วิธีการตรวจหาวัตถุด้วยตัวแบบ YOLOv5 เพื่อใช้ในการจำแนกสายพันธุ์และตำแหน่งของเมล็ดข้าวสารที่อยู่ในภาพที่มีเมล็ดข้าวสารหลากหลายปรากฏอยู่ในภาพได้ ไม่ว่าทั้งในแบบที่ไม่มีภาพของเมล็ดข้าวสารซ้อนทับกันเลย หรือกรณีที่มีภาพเมล็ดข้าวสารซ้อนทับกันอยู่ตั้งแต่ไม่เกินร้อยละ 5, 10, 15, 20 และ 25 และเนื่องด้วยการระบุตำแหน่งของวัตถุอยู่ในรูปแบบของกรอบสี่เหลี่ยมผืนผ้า และมีเมล็ดข้าวสารที่ทับซ้อนกัน ทำให้ต้องออกแบบการประเมินประสิทธิภาพความแม่นยำใหม่ โดยพิจารณาจากอัตราส่วนจุดภาพของเมล็ดข้าวสารที่ตรวจได้เทียบกับจำนวนจุดภาพต่าง ๆ ที่ไม่ใช่ภาพพื้นหลัง ดังนั้นการหาขีดแบ่งที่เหมาะสมเพื่อระบุว่าอัตราส่วนดังกล่าวเป็นการตรวจหาที่ถูกต้องหรือไม่ ต้องพิจารณาจากผลการวัดความ

แม่นยำของตัวแบบด้วยค่าขีดแบ่งค่าต่าง ๆ เมื่อพิจารณาจากกราฟความความแม่นยำของตัวแบบที่มีค่าขีดแบ่งต่าง ๆ จะพบว่าค่าขีดแบ่งที่ 0.60 มีความเหมาะสมที่จะใช้ในการประเมินความแม่นยำ เนื่องจากเมื่อพิจารณาข้อมูลชุดทดสอบที่มีไม่มีภาพเมล็ดข้าวสารซ้อนทับกันเลย (ซ้อนทับร้อยละ 0) เมื่อวัดความถูกต้องของตัวแบบโดยใช้ค่าขีดแบ่งที่ลดลง ค่าความถูกต้องก็มากขึ้น และพบว่าที่ค่าขีดแบ่ง 0.60 และ 0.55 ตัวแบบ YOLOv5 ให้ผลการตรวจหาเมล็ดข้าวสารที่มีความถูกต้องเท่ากัน ถือว่า 0.60 เป็นค่าขีดแบ่งที่มีค่าสูงกว่าที่ไม่ทำให้ผลการตรวจหาพันธุ์เมล็ดข้าวสารที่ไม่มีภาพซ้อนทับกันเปลี่ยนแปลง และที่ค่าขีดแบ่งดังกล่าวจะถือว่าตัวแบบ YOLOv5 สามารถจำแนกพันธุ์ข้าวที่ปรากฏในภาพที่มีภาพเมล็ดข้าวสารซ้อนทับกันร้อยละ 0, 5, 10, 15, 20 และ 25 ได้ถูกต้องร้อยละ 99.13, 99.00, 98.62, 98.19, 97.56 และ 96.89 ตามลำดับ

6) ข้อเสนอแนะ

จากผลการวิจัยจะเห็นได้ว่าวิธีการจำแนกเมล็ดข้าวสารด้วยวิธีการตรวจหาวัตถุ ที่ได้นำเสนอขึ้นมา ทำงานได้ด้วยความแม่นยำสูง ทั้งนี้หากนำไปประยุกต์ใช้กับเมล็ดข้าวสารพันธุ์ต่าง ๆ ที่มีจำหน่ายในท้องตลาดของประเทศไทย ผลที่ได้จะสามารถนำไปต่อยอดในการตรวจสอบคุณภาพของข้าวที่จะส่งออกของประเทศไทยได้ในอนาคต

REFERENCES

[1] M. Shahbandeh. “Principal rice exporting countries worldwide 2022/2023.” STATISTA.com. https://www.statista.com/statistics/255947/top-rice-exporting-countries-worldwide-2011/ (accessed Apr. 10, 2023).

[2] Thai Rice Exporters Association. “F.O.B. Prices.” THAIRICEEXPORTERS.or.th. http://www.thairiceexporters.or.th/ (accessed Apr. 10, 2023).

[3] *Notifications of the Ministry of Commerce: Prescribe Thai Hom Mali Rice as a Standard Product and Thai Hom Mali Rice Product Standard (no. 3), B.E. 2559*, Thai Government Gazette 133(243D), Ministry of Commerce, Oct. 21, 2016.

[4] R. Chityala and S. Pudipeddi, *Image Processing and Acquisition Using Python.* New York, NY, USA: CRC Press, 2014.

[5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* 3rd ed. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2008.

[6] S. Zaidi, S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, “A Survey of modern deep learning based object detection models,” *Digit. Signal Process.*, vol. 126, 2022, Art. no. 103514, doi: 10.1016/j.dsp.2022.103514.

[7] D. Thuan, “Evolution of YOLO algorithm and YOLOV5: The state-of-the-art object detection algorithm,” B.S. thesis, Dept. Inform. Technol., Oulu Univ. Appl. Sci., Oulum, Finland, 2021.

[8] O. Aki, A. Güllü, and E. Uçar, “Classification of rice grains using image processing and machine learning techniques,” in *Proc. Int. Sci. Conf. “UNITECH 2015”*, Gabrovo, Bulgaria, Nov. 2015, pp. 352–354.

[9] H. Zareiforoush, S. Minaei, M. R. Alizadeh, and A. Banaka, “Qualitative classification of milled rice grains using computer vision and metaheuristic techniques,” *J. Food Sci. Technol.*, vol. 53, no. 1, pp. 118–131, Jan. 2016.

[10] I. Cinar and M. Koklu, “Classification of rice varieties using artificial intelligence methods,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 7, no. 3, pp. 188–194, 2019.

[11] I. Cinar and M. Koklu, “Determination of effective and specific physical features of rice varieties by computer vision in exterior quality inspection,” *Selcuk J. Agriculture Food Sci.,* vol. 35, no. 3, pp. 229–243, 2021.

[12] M. Koklu, I. Cinar, and Y. S. Taspinar, “Classification of rice varieties with deep learning methods,” *Comp. Electron. Agriculture*, vol. 187, 2021, Art. no. 106285, doi: 10.1016/j.compag.2021.106285.

[13] I. Cinar and M. Koklu, “Identification of rice varieties using machine learning algorithms,” *J. Agricultural Sci. (Tarim Bilimleri Dergisi)*, vol. 28, no. 2, pp. 307–325, 2022.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

การพัฒนาแบบจำลองการพิจารณาการอนุมัติสินเชื่อ โดยใช้เทคนิคการวิเคราะห์องค์ประกอบร่วมกับการเรียนรู้ด้วยเครื่อง

ชลลดา ม่วงธนัง[1*] สุรศักดิ์ มังสิงห์[2] นิเวศ จิระวิชิตชัย[3]

[1,2]คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม, กรุงเทพมหานคร, ประเทศไทย*
[3]คณะวิศวกรรมศาสตร์และเทคโนโลยี สถาบันการจัดการปัญญาภิวัฒน์, นนทบุรี, ประเทศไทย

*ผู้ประพันธ์บรรณกิจ อีเมล : chonlada.mua@spulive.net

รับต้นฉบับ : 7 มีนาคม 2566; รับบทความฉบับแก้ไข: 30 เมษายน 2566; ตอบรับบทความ : 13 มิถุนายน 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอแบบจำลองและทดสอบประสิทธิภาพของแบบจำลองการอนุมัติสินเชื่อในอุตสาหกรรมการเงิน โดยใช้เทคนิคการวิเคราะห์องค์ประกอบร่วมกับการเรียนรู้ด้วยเครื่อง ทดสอบกับอัลกอริทึม 4 วิธี ได้แก่ นาอีฟเบย์ (Naïve Bayes) ต้นไม้การตัดสินใจ (Decision Tree) แรนดอมฟอร์เรส (Random Forest) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) โดยทดสอบกับชุดข้อมูลการปล่อยกู้จำนวน 441,335 ตัวอย่าง ทดสอบประสิทธิภาพของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความครบถ้วน (Recall) และค่าการวัดประสิทธิภาพโดยรวม (F-Measure)

จากการทดลองพบว่าแบบจำลองแรนดอมฟอร์เรส ให้ประสิทธิภาพค่าความถูกต้องดีที่สุดคือร้อยละ 92.90 รองลงมาเป็น ซัพพอร์ตเวกเตอร์แมชชีน และต้นไม้การตัดสินใจ คือร้อยละ 87.00 และนาอีฟเบย์ คือร้อยละ 83.40 ตามลำดับ ซึ่งพบว่าการลดมิติของข้อมูลส่งผลให้ประสิทธิภาพของแบบจำลองดีขึ้นเนื่องจากตัดคุณลักษณะที่ไม่มีนัยสำคัญทิ้งไป และสามารถแก้ไขปัญหาของแบบจำลองการอนุมัติสินเชื่อแบบเดิมที่พิจารณาจากตัวแปรคุณลักษณะแบบเก่า โดยแบบจำลองได้พัฒนาประสิทธิภาพให้ดีขึ้นจากการพิจารณาตัวแปรคุณลักษณะแบบใหม่ ผลของค่าความครบถ้วน (Recall) และค่าการวัดประสิทธิภาพโดยรวม (F-Measure) จากการทดลองพบว่าแบบจำลองแรนดอมฟอร์เรส ให้ประสิทธิภาพดีที่สุด เช่นเดียวกันกับ ค่าความถูกต้อง (Accuracy) คือร้อยละ 99.64 และ 99.35 ตามลำดับ และค่าความแม่นยำ (Precision) มีค่าสูงสุด คือ แบบจำลองแรนดอมฟอร์เรส และต้นไม้การตัดสินใจ เท่ากันที่ร้อยละ 99.07

คำสำคัญ : อนุมัติสินเชื่อ การเรียนรู้ด้วยเครื่อง วัดประสิทธิภาพ

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Developing a Credit Approval Determination Model Using Principal Component Analysis with Machine Learning Techniques

Chonlada Muangthanang[1*] Surasak Mungsing[2] Nivet Chirawichichai[3]

[1,2]School of Information Technology, Sripatum University, Bangkok, Thailand*
[3]Faculty of Engineering and Technology, Panyapiwat Institute of Management, Nonthaburi, Thailand

*Corresponding Author. E-mail address: chonlada.mua@spulive.net

Received: 7 March 2023; Revised: 30 April 2023; Accepted: 13 June 2023
Published online: 27 December 2023

Abstract

This research presents a model and tests the performance of a credit approval determination model by using principal component analysis with machine learning techniques. Four algorithms: Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine were tested with 441,335 examples of lending data. There are four model performance test results: Accuracy, Precision, Recall, and F-Measure.

From the experiment, it was found that the random forest model provides the best accuracy performance of 92.90 percent, followed by support vector machine and decision tree is 87.00 percent, and Naïve Bayes is 83.40 percent, respectively. It was found that reducing the data dimensions resulted in improved model performance by eliminating insignificant features and solve the problem of the traditional credit approval model that considers the old attribute variables. The model's performance has improved by considering new attribute variables. The results of the completeness value (Recall) and the overall performance measurement value (F-Measure) from the experiment found that Random forest model provides the best performance as well as accuracy values were 99.64 and 99.35 percent, respectively, and the highest precision value was 99.07 percent for the random forest model and decision trees.

***Keywords*:** Credit approval, Machine learning, Performance

1) บทนำ

สถาบันการเงินที่บริการสินเชื่อใช้ข้อมูลที่ลูกค้ามายื่นขอรับสินเชื่อและข้อมูลที่ได้จากองค์กรกลางภายนอกมาตัดสินการให้สินเชื่อหรือการปฏิเสธการขอรับสินเชื่อ มีการเลือกใช้เจ้าหน้าที่ในการพิจารณาสินเชื่อของลูกค้า โดยสถาบันการเงินในแต่ละแห่งต่างมีแบบจำลองคะแนนเครดิต (credit scoring) ของตนเอง [1] ในการพิจารณาอนุมัติวงเงินสินเชื่อให้แก่ลูกค้า ซึ่งพัฒนาขึ้นมาจากฐานข้อมูลลูกค้าและประสบการณ์ในอดีตของตนเอง ทำให้สถาบันการเงินแต่ละแห่งมีความเข้มข้นของเกณฑ์แตกต่างกัน หากเกณฑ์การพิจารณาเข้มจนเกินไปอาจทำให้สูญเสียโอกาสทางธุรกิจได้ และในทางกลับกันเกณฑ์ที่ประนีประนอมมากไปก็อาจทำให้รับลูกค้ากลุ่มเสี่ยงเข้ามาได้เช่นกัน ด้วยเหตุนี้จึงส่งผลให้เกิดปัญหาความล่าช้าของกระบวนการพิจารณาสินเชื่อและการอนุมัติที่ไม่เหมาะสม การพิจารณาที่ไม่สมเหตุสมผลเป็นเหตุให้ลูกค้าเปลี่ยนไปขอรับสินเชื่อจากสถาบันการเงินที่บริการสินเชื่ออื่น ดังนั้นสถาบันการเงินที่บริการสินเชื่อจึงต้องการวิธีการพิจารณาสินเชื่อที่รวดเร็ว และมีประสิทธิภาพในการอนุมัติสินเชื่อ ซึ่งในการให้คะแนนเครดิตที่จัดระดับโดยสถาบันจัดอันดับความน่าเชื่อถือ (credit rating agencies) [2] ถูกวิพากษ์วิจารณ์ว่ามีความลำเอียงในการให้คะแนนทำให้เกิดความได้เปรียบเสียเปรียบเชิงการแข่งขันทางการค้าโดยเฉพาะนักลงทุนรายย่อยจึงจำเป็นต้องได้รับการแก้ไข จำเป็นที่ต้องมีวิธีการให้คะแนนเครดิตขั้นต้นที่จะนำไปสู่การประเมินเครดิตที่ทันเวลาและมีกระบวนการทำงานที่โปร่งใสมีความเป็นกลาง เพื่อลดความเสี่ยงของระบบการเงินในอนาคต

การได้มาซึ่งการพิจารณาสินเชื่อที่มีผลอย่างสมบูรณ์ เพื่อแสดงให้เห็นถึงผลที่ดีของการวิเคราะห์เชิงคาดการณ์ (predictive analytics) ซึ่งต้องสามารถวิเคราะห์ข้อมูลในปัจจุบันและอดีตเพื่อคาดการณ์เกี่ยวกับผลลัพธ์ในอนาคต นั่นคือ การทำความสะอาดข้อมูล (clean data) และเลือกข้อมูลมาอย่างระมัดระวังภายใต้การวิเคราะห์ โดยวิธีที่ทำให้โมเดลการคาดการณ์ (predictive model) นั้นมีประสิทธิภาพมากขึ้น คือ การใช้เทคนิควิศวกรรมคุณลักษณะ (feature engineering) เพื่อให้ได้ประโยชน์สูงสุดจากข้อมูลอย่างลึกซึ้ง อีกทั้งส่งผลให้ได้รับข้อมูลเชิงลึกเพิ่มเติมและระบุโมเดลการคาดการณ์ที่เหมาะสมที่สุด ทำให้เทคนิควิศวกรรมคุณลักษณะเป็นขั้นตอนที่สำคัญ โดยเป็นเทคนิคที่ต้องใช้ความรู้เกี่ยวกับกระบวนการ และผลลัพธ์ในการแยกคุณสมบัติ (properties) หรือคุณลักษณะ (features) ต่าง ๆ ที่ทำให้โมเดลการคาดการณ์ทำงานได้ดีขึ้น สู่การได้มาซึ่งโมเดลที่เหมาะสมและประสิทธิภาพที่มีความแม่นยำ

ดังนั้น จึงสามารถกล่าวได้ว่า เทคนิควิศวกรรมคุณลักษณะ (feature engineering) เป็นกระบวนการใช้แบบจำลองความรู้หลักการเฉพาะปัญหา (domain knowledge) ในการสร้างคุณลักษณะใหม่ขึ้นมา และทำการเลือกตัดคุณลักษณะที่ไม่เกี่ยวข้องออกไป เพื่อช่วยทำให้อัลกอริทึมที่ได้มาจากเทคนิคการเรียนรู้ด้วยเครื่อง (machine learning) เรียนรู้ได้ดียิ่งขึ้น โดยมีวัตถุประสงค์ของการวิจัย เพื่อทดสอบและเปรียบเทียบประสิทธิภาพของแบบจำลอง เป็นจำนวน 4 ตัวแบบ ในการพิจารณาการอนุมัติสินเชื่อโดยใช้เทคนิคการวิเคราะห์องค์ประกอบร่วมกับแรนดอมฟอร์เรส

จากความสำคัญที่กล่าวมาข้างต้น ได้มีแนวคิดเพื่อแก้ปัญหาแบบจำลองการอนุมัติสินเชื่อแบบเดิมที่พิจารณาจากตัวแปรคุณลักษณะประชากรแบบเก่า และประวัติในการทำธุรกรรมเกี่ยวกับสินเชื่อทางการเงินในอดีต ซึ่งพบว่าจะสามารถพัฒนาให้ดีขึ้นได้ เหตุผลเพราะข้อมูลธุรกรรมการเงินเกี่ยวกับสินเชื่อในอดีต เช่น NPL ratio และข้อมูลการผ่อนชำระหนี้แบบเดิมของลูกหนี้ อาจไม่สะท้อนให้เห็นถึงคุณภาพของสินเชื่อที่แท้จริง เนื่องจากไม่ได้พิจารณาปัจจัยสภาพแวดล้อมความเสี่ยงทางการเงินที่ครบถ้วน ดังนั้น จึงเกิดแนวคิดในการพัฒนาตัวชี้วัดใหม่ที่พิจารณาจากความเสี่ยง บริบทโดยรวมของข้อมูลทางการเงินทั้งหมดโดยประยุกต์ใช้กฎพิจารณาความเสี่ยงลูกหนี้ของธนาคารแห่งประเทศไทย เพื่อนำมาใช้ร่วมกับตัวชี้วัดเดิม โดยพิจารณารายละเอียดและอัตราการผิดชำระหนี้ (default rate) เข้าร่วมด้วย จึงเป็นเหตุผลให้แสวงหาคำตอบโดยพัฒนา Credit Scoring Model แบบใหม่ ในการพิจารณาการอนุมัติสินเชื่อโดยใช้เทคนิควิศวกรรมคุณลักษณะร่วมกับเทคนิคการเรียนรู้ด้วยเครื่อง เพื่อใช้ในการพยากรณ์คุณภาพของสินเชื่อ โดยใช้ตัวชี้วัดใหม่ที่พัฒนาขึ้นมานี้ร่วมด้วย ซึ่งแบบจำลองดังกล่าวสามารถแยกแยะคุณภาพลูกหนี้ดี - เสีย ออกจากกันได้ค่อนข้างดีกว่าวิธีการดั้งเดิม โดยใช้งานร่วมกับอัลกอริทึมการเรียนรู้ของเครื่องในการพยากรณ์ เพื่อจัดการความเสี่ยงที่จะเกิดเหตุการณ์การผิดนัดชำระหนี้ขึ้น และเพื่อใช้ประกอบการตัดสินใจในการดำเนินนโยบายและกำกับดูแลการปล่อยสินเชื่อให้กับผู้กู้รายย่อย

2) ทบทวนวรรณกรรม

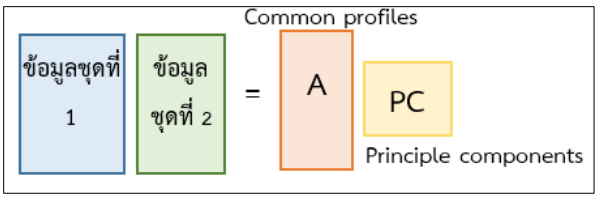
2.1) เทคนิคการวิเคราะห์หาองค์ประกอบหลัก (Principle Component Analysis: PCA)

PCA เป็นวิธีทางสถิติที่วิเคราะห์หาองค์ประกอบหลักของข้อมูลภายใต้เงื่อนไขค่าความแปรปรวน (variance) สูงสุด [3] ดังสมการ (1)

$$Y = PX \tag{1}$$

คำอธิบายสมการ เป็นกระบวนการแปลงข้อมูล X บนเมทริกซ์การแปลง P เพื่อเป็นข้อมูลเชิงเส้น และ Y แสดงความสัมพันธ์

เนื่องจากวิธี PCA เป็นวิธีที่พิจารณาความสัมพันธ์ร่วมกันของข้อมูล X หรือข้อมูลเพียงชุดเดียว หากมีชุดข้อมูลที่มากกว่าสองชุดขึ้นไป ข้อมูลเหล่านั้นจะถูกหลอมรวมก่อนในขั้นตอนแรกดังรูปที่ 1



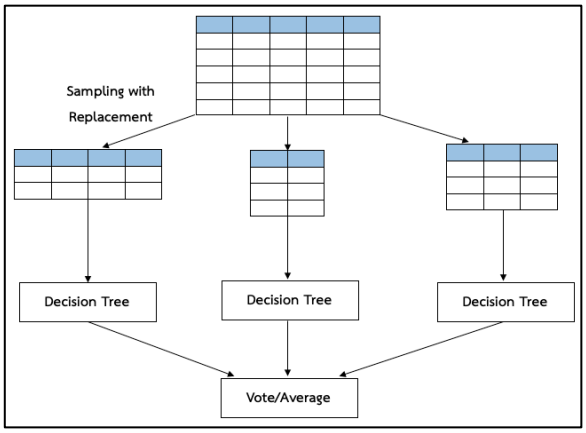
รูปที่ 1 : ขั้นตอนของชุดข้อมูลที่ถูกหลอมรวม [3]

จากนั้นข้อมูลจะหาความสัมพันธ์ร่วม Covariance และคำนวณหาองค์ประกอบหลักด้วย Eigen vector และ Eigen Value ที่สัมพันธ์กัน สำหรับการลดมิติข้อมูลด้วยวิธี PCA จะดำเนินการด้วยการจัดเรียงค่า Eigen จากมากไปหาน้อยหรือเรียกว่า diagonal ของค่า Eigen จากนั้นจะเลือกจำนวนขององค์ประกอบหลักจำนวน d ตัวเพื่อลดมิติลง การเลือกจำนวนขององค์ประกอบส่วนใหญ่จะเลือกจากค่าความเชื่อมั่นหรือค่าสะสมของ Eigen เมื่อได้จำนวนองค์ประกอบที่ จะนำค่า Eigen Vector เพื่อใช้เป็นค่า weight สำหรับการลดมิติข้อมูลก่อนเข้ากระบวนการประมวลผลต่อไป

2.2) แรนดอมฟอร์เรส (Random Forest)

Random Forest คือ อัลกอริทึมการเรียนรู้ของเครื่องที่ชนิดหนึ่งที่ใช้การสุ่มคุณลักษณะ เพื่อเพิ่มความหลายหลายของโมเดล โดยใช้โมเดลจาก Decision Tree หลาย ๆ โมเดลย่อย ๆ โดยแต่ละโมเดลจะได้รับชุดข้อมูลทดสอบที่ไม่เหมือนกันซึ่งเป็นชุดข้อมูลย่อย ของชุดข้อมูลทั้งหมด เมื่อทำการพยากรณ์ก็ให้แต่ละ Decision Tree ทำการพยากรณ์แยกเป็นคนละชุดไม่เกี่ยวข้องกันและคำนวณผลการพยากรณ์โดยใช้ Vote Output ที่ถูกเลือกโดย Decision Tree มากที่สุด (กรณี Classification) หรือหาค่า mean จาก Output ของแต่ละ Decision Tree (กรณี Regression) [4]

Random Forest เป็นขั้นตอนวิธีพัฒนาต่อยอดมาจาก Decision Tree ต่างกันที่ Random Forest เป็นการเพิ่มจำนวนต้นไม้ (tree) เป็นหลาย ๆ ต้น ทำให้ประสิทธิภาพการทำงานและพยากรณ์สูงขึ้น Random Forest มีหลักการทำงาน คือ จะแบ่งข้อมูลออกเป็นต้นไม้ตัดสินใจ (decision tree) หลาย ๆ ต้น โดยแต่ละต้นจะได้รับคุณลักษณะ (feature) และข้อมูล (data) ที่ไม่เหมือนกันทั้งหมด เพื่อทำให้ได้ต้นไม้ที่มีความหลากหลายและมีความอิสระต่อกันมากขึ้น [5] ดังรูปที่ 2



รูปที่ 2 : หลักการทำงานของ Random Forest [6]

Random Forest เป็นการเรียนรู้อีกแบบหนึ่งในกลุ่ม Ensemble Learning ที่จะนำโมเดล Decision Tree หลาย ๆ อันมาใช้ร่วมกัน ซึ่งก็เปรียบได้กับต้นไม้หลาย ๆ ต้น เมื่อรวมกันก็จะกลายเป็นป่า (forest) นั่นเอง

ดังนั้น Random Forest จึงสามารถทำนายผลได้ทั้งแบบ Classification และ Regression เช่นเดียวกับ Decision Tree ซึ่งโดยทั่วไปแล้ว Random Forest มักจะมีความแม่นยำมากกว่า Decision Tree [6]

หลักการทำงานของ Random Forest มีแนวทางดังนี้ [6]

1) เนื่องจาก Random Forest ใช้วิธีแบบ Bagging ดังนั้นข้อมูลตัวอย่างที่นำมา Train Model จึงถูกแบ่งออกเป็นกลุ่ม ๆ ย่อย โดยใช้หลักการแบบ Sampling with Replacement ตามจำนวนการเรียนรู้ที่เรากำหนด

2) ข้อมูลแต่ละชุดจะถูกนำไป Train Model ตามหลักการของ Decision Tree

3) ผลลัพธ์จากแต่ละโมเดลจะถูกดำเนินการดังนี้

- หากเป็นการทำนายแบบ Classification ก็จะทำนายผลโดยวิธีการโหวต

- หากเป็นการทำนายผลแบบ Regression ก็จะทำนายผลโดยวิธีการหาค่าเฉลี่ย

เนื่องจาก Random Forest จะดำเนินการบนพื้นฐานของ Decision Tree ซึ่งข้อกำหนดต่าง ๆ จะเป็นไปตามหลักการของโมเดล

2.3) ทฤษฎีการนอร์มัลไลเซชัน (Normalization)

การทำ Normalization หมายถึง การทำให้ข้อมูลเป็นค่ามาตรฐาน เป็นเทคนิค Feature Scaling ที่จะนำข้อมูลแต่ละ Feature อยู่ในย่านจำนวนจริงที่ Machine Learning อ่านได้ง่ายขึ้น [7]

การปรับ scale ของข้อมูล ข้อมูลที่จะนำมา Train Model นั้น หากบางคอลัมน์มีค่าแตกต่างจากคอลัมน์อื่น ๆ มากเกินไป อาจทำให้คอลัมน์นั้นมีน้ำหนักมากกว่า หรือไม่สมดุลกับข้อมูลในคอลัมน์อื่น ๆ ซึ่งอาจส่งผลต่อการทำนาย และเกิดความคลาดเคลื่อนได้ [6]

วิธี StandardScaler จะคำนวณโดยใช้สมการ (2) [6]

$$Z = \frac{X - U}{S} \tag{2}$$

คำอธิบายสมการ

Z คือ ค่าใหม่หลังการปรับสเกล

X คือ ค่าเดิม หรือข้อมูลตัวอย่างแต่ละค่า

U คือ ค่าเฉลี่ยของข้อมูลในคอลัมน์นั้น

S คือ ส่วนเบี่ยงเบนมาตรฐานของข้อมูลในคอลัมน์นั้น

2.4) ตัววัดประสิทธิภาพ

2.4.1) ค่าความถูกต้อง (Accuracy) หรือ Accuracy Score [8] เป็นการวัดความถูกต้องของโมเดล โดยคำนวณรวมทุกคำตอบ สูตรในการคำนวณ คือ

$$\text{Accuracy} = \frac{\text{True Positives + True Negatives}}{\text{Total Number of Examples}} \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

2.4.2) ค่าความแม่นยำ (Precision) [8] เป็นการคำนวณค่า Precision ของโมเดล โดยคำนวณแยกทีละคำตอบ

ค่า Precision วัดประสิทธิภาพของโมเดล ที่สนใจผลการทำนายแบบ FP (False Positive) ผลที่ได้เป็นเปอร์เซ็นต์ (%) สูตรในการคำนวณ คือ

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}} \tag{5}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

2.4.3) ค่าความครบถ้วน หรือค่าความระลึก (Recall) [8] เป็นการคำนวณค่า Recall ของโมเดลโดยคำนวณแยกทีละคำตอบ

ค่า Recall วัดประสิทธิภาพของโมเดล ที่สนใจผลการทำนายแบบ FN (False Negative) ผลที่ได้เป็นเปอร์เซ็นต์ (%) สูตรในการคำนวณคือ

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}} \tag{6}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

2.4.4) ค่าการวัดประสิทธิภาพโดยรวม (F-Measure) [8] เป็นการคำนวณค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) ของ Precision และ Recall โดยคำนวณแยกทีละลาเบลคำตอบ

ค่า F1-score ประสิทธิภาพของโมเดล ที่ไม่สนใจผลการทำนายแบบ FN (False Negative) หรือ FP (False Positive) เป็นหลัก ผลที่ได้เป็นเปอร์เซ็นต์ (%) สูตรในการคำนวณคือ

$$F_1 = 2 \text{ X } \frac{\text{Precision X Recall}}{\text{Precision + Recall}} \tag{7}$$

คำอธิบายสมการ

True Positive (TP) คือ จำนวนข้อมูลที่พยากรณ์ถูกว่าเป็นคลาส ซึ่งกำลังพิจารณา

True Negative (TN) คือ จำนวนข้อมูลที่พยากรณ์ถูกว่าเป็นคลาส ซึ่งไม่ได้พิจารณา

False Positive (FP) คือ จำนวนข้อมูลที่พยากรณ์ผิดมาเป็นคลาส ซึ่งกำลังพิจารณา

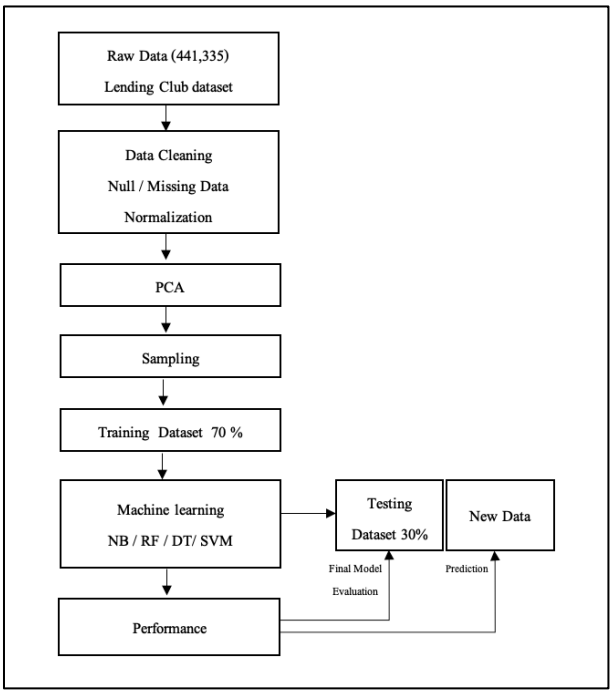
False Negative (FN) คือ จำนวนข้อมูลที่พยากรณ์ผิดมาเป็นคลาส ซึ่งไม่ได้พิจารณา

3) วิธีดำเนินการวิจัย

3.1) ประชากรและกลุ่มตัวอย่าง

กลุ่มข้อมูลเครดิตของพลเมืองประเทศสหรัฐอเมริกา (United Stated Credit) จำนวนกลุ่มตัวอย่าง 441,335 ตัวอย่าง [9] ซึ่งมีข้อมูลทั้งหมด 151 คุณลักษณะ (attributes) ประกอบด้วยตัวแปรต้น 150 คุณลักษณะ และตัวแปรตาม 1 คุณลักษณะ คือ สถานะปัจจุบันของหนี้ (loan_status)

3.2) กรอบงานวิจัย



รูปที่ 3 : กรอบงานวิจัย

3.3) เครื่องมือที่ใช้ในการวิจัย/รวบรวมข้อมูล

กระบวนการสุ่มตัวอย่าง (Sampling) แบ่งข้อมูลออกเป็น 2 ส่วน เพื่อใช้ทดสอบประสิทธิภาพโมเดล คือ

1) ข้อมูลฝึกสอน (Training Data) ที่ 70% ของข้อมูลทั้งหมด คือ กลุ่มตัวอย่าง 308,935 ตัวอย่าง

2) ข้อมูลทดสอบ (Testing Data) ที่ 30% ของข้อมูลทั้งหมด คือ กลุ่มตัวอย่าง 132,400 ตัวอย่าง

3.4) การวิเคราะห์ข้อมูล

3.4.1) กระบวนการจัดเตรียมข้อมูลกลุ่มตัวอย่าง สำหรับการประมวลผลด้วยวิธีการทำความสะอาดข้อมูล (Cleansing Data) คือ

1) การคัดกรองข้อมูลกลุ่มตัวอย่างที่มีคุณลักษณะ (Attributes) ของคุณภาพข้อมูลต่ำออก

2) แทนที่ข้อมูลที่เป็นค่าว่าง (Missing)

3) ลดมิติ (Dimension) ของข้อมูลขนาดใหญ่ด้วยวิธี การวิเคราะห์หาองค์ประกอบหลัก (Principal Component Analysis (PCA)) [3]

4) ลดความซ้ำซ้อนของข้อมูลด้วยกระบวนการ Normalization

3.4.2) กระบวนการของเทคนิควิศวกรรมคุณลักษณะ (Feature Engineering) [10]

1) ใช้เทคนิค Handling Outliers หรือ ค่าที่ผิดปกติ โดยจัดการแทนที่ข้อมูลที่มีค่าผิดปกติในข้อมูลที่มีค่าสูง หรือต่ำกว่าข้อมูลส่วนใหญ่ในชุดข้อมูลอย่างมาก ซึ่งการจัดการกับ Outlier จะพิจารณาจาก Standard Deviation และ Percentile [11]

2) ใช้เทคนิค One-Hot Encoding เป็นการเข้ารหัสข้อมูลแบบหนึ่ง โดยใช้วิธีแปลงข้อมูลที่ขาดหายไป

3.4.3) กระบวนการสกัดคุณลักษณะ (Feature Extraction) [10]

1) ใช้กระบวนการสกัดคุณลักษณะ ด้วยการเลือกเฉพาะคุณลักษณะที่สำคัญ (Selection) เพื่อให้มีมิติของข้อมูลลดลงจากคุณลักษณะทั้งหมด 151 คุณลักษณะ เหลือ 16 คุณลักษณะ โดยแบ่งออกเป็น

2) ตัวแปรต้นที่นำไปใช้งานได้ จำนวน 15 คุณลักษณะ (Attributes) คือ (1) issue_d (2) term (3) grade (4) sub_grade (5) emp_length (6) home_ownership (7) verification_status (8) purpose (9) initial_list_status (10) last_pymnt_d (11) last_credit_pull_d (12) pc_1 (13) pc_2 (14) pc_3 (15) pc_4 มีรายละเอียดคุณลักษณะ (Attributes) ดังแสดงในตารางที่ 1 ดังนี้

ตารางที่ 1 : รายละเอียดคุณลักษณะ (Attributes) ที่นำมาวิเคราะห์ข้อมูล

ข้อมูล	คำอธิบายข้อมูล
(1) issue_d	เดือนที่กู้เงิน
(2) term	จำนวนการชำระเงินกู้ ค่าเป็นเดือน และอาจเป็น 36 หรือ 60 ก็ได้
(3) grade	LC กำหนดเกรดสินเชื่อ (LC/Letter of Credit คือ สินเชื่อเพื่อเปิด)
(4) sub_grade	LC กำหนดลดเกรดเงินกู้ (LC/Letter of Credit คือ สินเชื่อเพื่อเปิด)
(5) empLength	อายุงานเป็นปี ค่าที่เป็นไปได้อยู่ระหว่าง 0 ถึง 10 โดยที่ 0 หมายถึงน้อยกว่าหนึ่งปี และ 10 หมายถึง 10 ปีขึ้นไป
(6) home_ownership	สถานะเจ้าของบ้านที่ผู้กู้ระบุระหว่างการลงทะเบียน ค่านิยมของเราคือ: เช่า, เป็นเจ้าของ, จำนอง, อื่นๆ
(7) verification_status	ระบุว่ารายได้ได้รับการยืนยันโดย LC หรือไม่ได้รับการยืนยัน หรือแหล่งที่มาของรายได้ได้รับการยืนยันหรือไม่ (LC/Letter of Credit คือ สินเชื่อเพื่อเปิด)

ตารางที่ 2 : รายละเอียดคุณลักษณะ (Attributes) ที่นำมาวิเคราะห์ข้อมูล (ต่อ)

ข้อมูล	คำอธิบายข้อมูล
(8) purpose	หมวดหมู่ที่ผู้กู้จัดเตรียมไว้สำหรับการขอสินเชื่อ
(9) initial_list_status	สถานะรายการเริ่มต้นของเงินกู้ ค่าที่เป็นไปได้คือ – W, F
(10) last_pymnt_d	ได้รับการชำระเงินเดือนล่าสุด
(11) last_credit_pull_d	เดือนล่าสุด ที่ LC ดึงเครดิตสำหรับเงินกู้นี้ (LC/Letter of Credit คือ สินเชื่อเพื่อเปิด)
(12) open_acc_6m	จำนวนการซื้อขายที่เปิดในช่วง 6 เดือนที่ผ่านมา
(13) open_il_12m	จำนวนบัญชีผ่อนชำระที่เปิดใน 12 เดือนที่ผ่านมา
(14) open_il_24m	จำนวนบัญชีผ่อนชำระที่เปิดใน 24 เดือนย้อนหลัง
(15) open_act_il	จำนวนการซื้อขายแบบผ่อนชำระที่ใช้งานอยู่ในปัจจุบัน

ตัวแปรตาม 1 คุณลักษณะ (attributes) คือ สถานะปัจจุบันของหนี้ (loan_status) ประกอบด้วย

1. สถานะชำระหนี้ปกติ (current)
2. สถานะชำระหนี้เต็มจำนวน (fully paid)
3. สถานะปิดการเรียกเก็บเงิน (charged off)

4) ผลการวิจัยและอภิปรายผล

กระบวนการพัฒนาแบบจำลอง ได้ผลการวิจัยจากการประเมินผลลัพธ์การทำนาย (Confusion Matrix) ขนาด 3×3 แสดงจำนวนที่พยากรณ์สถานะปัจจุบันของหนี้ (loan_status) ของแบบจำลองจาก 4 เทคนิค ดังนี้

ตารางที่ 2 : การประเมินผลลัพธ์การทำนาย (Confusion Matrix) แสดงจำนวนที่พยากรณ์ของตัวแบบจำลองนาอีฟเบย์ (Naïve Bayes)

	true Fully Paid	true Charged Off	true Current	class precision
pred. Fully Paid	51872	11702	8445	72.03%
pred. Charged Off	0	0	0	0.00%
pred. Current	805	453	55620	97.79%
class recall	98.47%	0.00%	86.82%	

จากตารางที่ 2 จะเห็นว่า โมเดลได้พยากรณ์ ค่าความครบถ้วน (Recall) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้เต็มจำนวน (Fully Paid) คิดเป็น 98.47% รองลงมาคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 86.82%

และ โมเดลได้พยากรณ์ ค่าความแม่นยำ (Precision) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 97.79% รองลงมาคือ สถานะชำระหนี้เต็มจำนวน (Fully Paid) คิดเป็น 72.03%

ตารางที่ 3 : การประเมินผลลัพธ์การทำนาย (Confusion Matrix) แสดงจำนวนที่พยากรณ์ของตัวแบบจำลองต้นไม้การตัดสินใจ (Decision Tree)

	true Fully Paid	true Charged Off	true Current	class precision
pred. Fully Paid	28688	3427	4847	77.61%
pred. Charged Off	223	3270	473	82.45%
pred. Current	215	69	30217	99.07%
class recall	98.50%	48.33%	85.03%	

ตารางที่ 3 จะเห็นว่า โมเดลได้พยากรณ์ ค่าความครบถ้วน (recall) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้เต็มจำนวน (fully paid) คิดเป็น 98.50% รองลงมาคือ สถานะชำระหนี้ปกติ (current) คิดเป็น 85.03%

และ โมเดลได้พยากรณ์ ค่าความแม่นยำ (precision) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้ปกติ (current) คิดเป็น 99.07% รองลงมาคือ สถานะปิดการเรียกเก็บเงิน (charged off) คิดเป็น 82.45%

ตารางที่ 4 : การประเมินผลลัพธ์การทำนาย (Confusion Matrix) แสดงจำนวนที่พยากรณ์ของตัวแบบจำลองแรนดอมฟอร์เรส (Random Forest)

	true Fully Paid	true Charged Off	true Current	class precision
pred. Fully Paid	7240	502	620	86.58%
pred. Charged Off	5	1068	5	99.07%
pred. Current	21	110	8287	98.44%
class recall	99.64%	63.57%	92.99%	

ตารางที่ 4 จะเห็นว่า โมเดลได้พยากรณ์ ค่าความครบถ้วน (Recall) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้เต็มจำนวน (Fully Paid) คิดเป็น 99.64% รองลงมาคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 92.99%

และ โมเดลได้พยากรณ์ ค่าความแม่นยำ (Precision) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะปิดการเรียกเก็บเงิน (Charged

Off) คิดเป็น 99.07% รองลงมาคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 98.44%

ตารางที่ 5 จะเห็นว่า โมเดลได้พยากรณ์ ค่าความครบถ้วน (Recall) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้เต็มจำนวน (Fully Paid) คิดเป็น 95.28% รองลงมาคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 88.83%

และโมเดลได้พยากรณ์ ค่าความแม่นยำ (Precision) สถานะปัจจุบันของหนี้ มากที่สุดคือ สถานะชำระหนี้ปกติ (Current) คิดเป็น 93.04% รองลงมาคือ สถานะชำระหนี้เต็มจำนวน (Fully Paid) คิดเป็น 83.23%

ตารางที่ 6 จะเห็นว่า ค่าความถูกต้อง (Accuracy) สูงที่สุดคือ Random Forest ที่ 92.9% และเกิดข้อผิดพลาดในการจำแนกประเภท (Classification Error) สูงที่สุดคือ Naïve Bayes ที่ 16.6%

ตารางที่ 5 : การประเมินผลลัพธ์การทำนาย (Confusion Matrix) แสดงจำนวนที่พยากรณ์ของตัวแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

	true Fully Paid	true Charged Off	true Current	class precision
pred. Fully Paid	1677	107	231	83.23%
pred. Charged Off	74	165	6	67.35%
pred. Current	9	132	1884	93.04%
class recall	95.28%	40.84%	88.83%	

ตารางที่ 6 : ประสิทธิภาพของตัวแบบจำลอง

Model	Accuracy Value	Classification Error	Standard Deviation
Naïve Bayes	83.4%	16.6%	± 0.3%
Decision tree	87.0%	13.0%	± 0.1%
Random Forest	92.9%	7.1%%	± 0.1%
Support Vector Machine	87.0%	13.0%	± 0.6%

ตารางที่ 7 : การเปรียบเทียบค่าตัววัดประสิทธิภาพที่ได้จากการพยากรณ์ของตัวแบบจำลอง จำนวน 4 ค่า

Model	Accuracy	Precision	Recall	F-Measure
Naïve Bayes	83.4%	97.79%	98.47%	98.13%
Decision tree	87.0%	**99.07%**	98.50%	98.78%
Random Forest	**92.9%**	**99.07%**	**99.64%**	**99.35%**
Support Vector Machine	87.0%	93.04%	95.28%	94.15%

ตารางที่ 7 พบว่า ตัวแบบจำลอง Random Forest มีค่าความถูกต้อง (Accuracy) ค่าความครบถ้วน หรือ ค่าความระลึก (Recall) ค่าความแม่นยำ (Precision) และค่าการวัดประสิทธิภาพโดยรวม (F-Measure) สูงที่สุด คือ 92.9%, 99.07%, 99.64%, 99.35% ตามลำดับ

ตัวแบบจำลอง Decision tree มีค่าความแม่นยำ (Precision) สูงที่สุด คือ 99.07% เช่นกันกับ ตัวแบบจำลอง Random Forest ที่มีค่าสูงที่สุดได้เท่ากัน

ตารางที่ 8 : การเปรียบเทียบเวลาที่ใช้ในการประมวลผลตัวแบบจำลอง

Model	Training Time (1,000 Rows)	Total Time
Naïve Bayes	0.002 วินาที	26 วินาที
Decision tree	0.004 วินาที	20 วินาที
Random Forest	0.046 วินาที	1 นาที 39 วินาที
Support Vector Machine	0.792 วินาที	4 นาที 8 วินาที

ผลจากตารางที่ 7 ได้พบว่า ตัวแบบจำลอง Random Forest มีค่าความถูกต้อง (Accuracy) สูงที่สุด แต่สำหรับเวลาที่ใช้ในการประมวลผลตัวแบบจำลองนั้นไม่ได้ใช้น้อยหรือมากที่สุด ซึ่งในตารางที่ 8 นั้น พบว่า แบบจำลองที่ใช้เวลาน้อยที่สุด คือ ตัวแบบจำลอง Decision tree ใช้เวลา 20 วินาที และแบบจำลองที่ใช้เวลามากที่สุด คือ ตัวแบบจำลอง Support Vector Machine ใช้เวลา 4 นาที 8 วินาที

ตารางที่ 9 : การเปรียบเทียบค่าความน่าจะเป็นเกี่ยวกับสถานะปัจจุบันของหนี้จากการจำแนกข้อมูลของแบบจำลอง

สถานะของหนี้ / แบบจำลอง	ชำระหนี้ปกติ	ชำระหนี้เต็มจำนวน	ปิดการเรียกเก็บเงิน
Naïve Bayes	98.47%	0.00%	86.82%
Decision Tree	98.50%	48.33%	85.03%
Random Forest	**99.64%**	**63.57%**	**92.99%**
Support Vector Machine	95.28%	40.84%	88.83%

จากตารางที่ 9 พบว่า แบบจำลอง Random Forest มีความน่าจะเป็นเกี่ยวกับสถานะปัจจุบันของหนี้ โดย ลูกหนี้สถานะชำระหนี้ปกติ คือ 99.64% คิดเป็น 439,746 ตัวอย่าง ลูกหนี้สถานะชำระหนี้เต็มจำนวน คือ 63.57% คิดเป็น 280,557 ตัวอย่าง และลูกหนี้สถานะปิดการเรียกเก็บเงิน คือ 92.99% คิดเป็น 410,397 ตัวอย่าง จากจำนวนกลุ่มตัวอย่างทั้งหมด 441,335 ตัวอย่าง

เมื่อได้เปรียบเทียบกับผลการทดลองจากงานวิจัยผู้อื่นที่เกี่ยวข้องได้ผล ดังนี้

จากงานวิจัยระบบการพิจารณาอนุมัติเงินกู้สวัสดิการพนักงานอัตโนมัติด้วยเทคนิควิธีการเรียนรู้ด้วยเครื่อง [12] มีการศึกษาเทคนิค 4 เทคนิค คือ Naïve Bayes, K-Nearest Neighbors, Support Vector Machine และ Random Forest เพื่อพัฒนาและเปรียบเทียบแบบจำลอง ผลการวิจัยพบว่า เทคนิค Random Forest ให้ความแม่นยำ (Precision) สูงสุดที่ 99.56% ซึ่งผลของงานวิจัยที่ได้นั้นสอดคล้องกับงานวิจัยชิ้นนี้

Chitambira [13] ได้ศึกษา 6 เทคนิค คือ Decision Tree, Artificial Neural Network, Logistic Regression, Support Vector Machine, Random Forest และ K-Nearest Neighbors เพื่อพัฒนาและเปรียบเทียบแบบจำลอง ผลการวิจัยพบว่า เทคนิค Logistic Regression ให้ความแม่นยำสูงสุดที่ 81.50% รองลงมา คือ เทคนิค Random Forest ให้ความแม่นยำ (Precision) ที่ 81.30% ซึ่งผลของงานวิจัยที่ได้นั้นสอดคล้องกับงานวิจัยชิ้นนี้

ในงานวิจัยการพัฒนาแบบจำลองการพิจารณาให้คะแนนสินเชื่อโดยใช้เทคนิคเหมืองข้อมูลของ Muangthanang *et al*. [14] ศึกษาอยู่ 5 วิธี คือ ต้นไม้การตัดสินใจ การจำแนกกลุ่ม นาอีฟเบย์ โครงข่ายประสาท การถดถอยแบบโลจิสติก และซัพพอร์ตเวกเตอร์แมชชีน เพื่อทดสอบประสิทธิภาพของแบบจำลอง ผลการวิจัยพบว่า ค่าความแม่นยำ ของแบบจำลองที่ดีที่สุดคือ โครงข่ายประสาทเทียม ที่ 90.14% และ 90.08% จากชุดข้อมูล 2 กลุ่มข้อมูล ซึ่งผลของงานวิจัยที่ได้นั้นไม่สอดคล้องกับงานวิจัยชิ้นนี้

Bai *et al*. [15] ได้ศึกษาเทคนิค 4 เทคนิค คือ Naïve Bayes, Decision tree, Logistic regression และ Support Vector Machine เพื่อเปรียบเทียบความแม่นยำในการจำแนกประเภท ได้ผล Naïve Bayes และ Decision tree นั้น มีความแม่นยำ (Precision) สูงกว่า Logistic regression และ Support Vector Machine เป็นจำนวนร้อยละ 40 และเพื่อเปรียบเทียบค่า Precision – Recall ในการจำแนกประเภท ได้ผล Naïve Bayes และ Decision tree มีค่าที่ดีกว่า 40% เช่นกัน ซึ่งผลของงานวิจัยที่ได้นั้นสอดคล้องกับงานวิจัยชิ้นนี้

5) สรุปผล

งานวิจัยนี้ได้นำเสนอแบบจำลองและทดสอบประสิทธิภาพของแบบจำลองการอนุมัติสินเชื่อโดยใช้เทคนิคการวิเคราะห์องค์ประกอบร่วมกับการเรียนรู้ด้วยเครื่อง ทดสอบกับอัลกอริทึม 4 วิธี ได้แก่ นาอีฟเบย์ (Naïve Bayes) ต้นไม้การตัดสินใจ (Decision Tree) แรนดอมฟอร์เรส (Random Forest) และ ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) โดยทดสอบกับชุดข้อมูลการปล่อยกู้จำนวน 441,335 ตัวอย่างผลการทดสอบประสิทธิภาพของแบบจำลองจากค่าความถูกต้อง (Accuracy) จากการทดลองพบว่าแบบจำลองแรนดอมฟอร์เรส ให้ประสิทธิภาพดีที่สุดคือ ร้อยละ 92.90 รองลงมาเป็นแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนและต้นไม้การตัดสินใจ ร้อยละ 87.00 และสุดท้ายนาอีฟเบย์ ร้อยละ 83.40 ตามลำดับ และจากการทดลองพบว่าการลดมิติของข้อมูลส่งผลให้ประสิทธิภาพของแบบจำลองดีขึ้นเนื่องจากตัดคุณลักษณะที่ไม่มีนัยสำคัญทิ้งไป

แบบจำลองได้พยากรณ์ความน่าจะเป็นเกี่ยวกับสถานะปัจจุบันของหนี้จากกลุ่มตัวอย่างที่ดีที่สุด คือ ลูกหนี้สถานะชำระหนี้ปกติ 439,746 ตัวอย่าง ลูกหนี้สถานะชำระหนี้เต็มจำนวน 280,557 ตัวอย่าง และลูกหนี้สถานะปิดการเรียกเก็บเงิน 410,397 ตัวอย่าง ซึ่งสามารถกล่าวได้ว่า การใช้ตัวชี้วัดใหม่ที่ได้พัฒนาขึ้น สามารถแยกแยะคุณภาพของลูกหนี้ได้ค่อนข้างดี เพื่อจัดการความเสี่ยงที่จะเกิดเหตุการณ์การผิดนัดชำระหนี้ขึ้น สถาบันการเงินสามารถนำกลุ่มตัวอย่างมาใช้ร่วมกับแบบจำลองนี้เพื่อประกอบการตัดสินใจในการดำเนินนโยบายและกำกับดูแลการปล่อยสินเชื่อให้กับผู้กู้รายย่อย

REFERENCES

[1] Committee for the Protection of Credit Information. “Basic Information about Credit Scoring.” (in Thai), CREDITINFO COMMITTEE.or.th. https://www.creditinfocommittee.or.th/Public/BasicIntormation (accessed Mar. 1, 2023).

[2] Supervision Group, Bank of Thailand. “Credit scoring, things that need to be preserved.” (in Thai), BOT.or.th. https://www.bot.or.th/th/research-and-publications/articles-and-publications/articles/Article_22Jan201.html (accessed Mar. 1, 2023).

[3] P. Moeynoi. “Dimensionality Reduction.” (in Thai), MEDIUM.com. https://medium.com/@doohpim/02-การลดมิติข้อมูล-dimensionality-reduction-a5df2adb4f92 (accessed Sep. 7, 2022).

[4] S. Teerawittayakul, “Ransomware detection study using machine learning techniques from a sample attack dataset,” (in Thai), Dep. Data Commun. Netw., King Mongkut's Univ. Technol. North Bangkok, Bangkok, Thailand, 2019.

[5] PradyaSin. "What is Random Forest." (in Thai), MEDIUM.com. https://medium.com/@pradyasin/random-forest-คืออะไร-74d2a0af3d7 (accessed Sep. 7, 2022).

[6] B. Pasilatesang, *Build Learning for AI with Python Machine Learning*. Bangkok, Thailand: SE-EDUCATION (in Thai), 2021.

[7] C. Ambi. " Difference Between Normalization and Standardization." TOWARDSAI.net. https://pub.towardsai.net/difference-between-normalization-and-standardization-745030eaf96f (accessed Sep. 7, 2022).

[8] E. Pacharawongsakda, *Practical Data Science with Rapid Miner Studio 1*. Nonthaburi, Thailand: IDC Premier (in Thai), 2022.

[9] *Lending Club Loan Data: Complete loan data for all loans issued through the 2007-2015, including status*, LendingClub 2007-2015, 2021. [Online]. Available: https://www.kaggle.com/datasets/adarshsng/lending-club-loan-data-csv

[10] N. Promrita and S. Waijanya, *Fundamental of DEEP LEARNING in Practice*. Nonthaburi, Thailand: IDC Premier (in Thai), 2021.

[11] T. Kaewsri. "Feature Engineering." (in Thai), MEDIUM.com. https://grimreaperjunior.medium.com/สกัดข้อมูลให้ตรงใจด้วยการทำ-feature-engineering-bcdba80124bc (accessed Sep. 7, 2022).

[12] S. Tidta and T. Kangkachit, "Automated system for employee welfare loan approval through machine learning techniques," (in Thai), *DPU Graduate Studies J.*, vol. 10, no. 2, pp. 93–107, 2022. [Online]. Available: https://grad.dpu.ac.th/upload/content/files/year10-2/10-7.pdf

[13] B. Chitambira, "Credit scoring using machine learning approaches," M.S. thesis, Division Math. Phys., Malardalen Univ., Vasteras, Sweden, Jun. 2022. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1664698/FULLTEXT01.pdf

[14] C. Muangthanang, S. Mungsing, and N. Chirawichitchai, "Development of a credit scoring model using data mining techniques," (in Thai), *Huachiew Chalermprakiet Sci, Technol. J.*, vol. 7, no. 1, pp. 82–93, 2021.

[15] R. Bai, H. Mo, and H. Li, "A comparison of credit rating classification models based on spark-evidence from lending-club," *Procedia Comput. Sci.*, vol. 162, pp. 811–818, 2019.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

แบบจำลองปริมาณการสั่งซื้อเหล็กเส้นที่ประหยัดสำหรับโครงการก่อสร้าง

ณรงค์ฤทธิ์ ว่องไว[1] ฐิติพร แจ่มจรัส[2*] ศุภวุฒิ มาลัยกฤษณะชลี[3]

[1]คณะวิศวกรรมศาสตร์ศรีราชา มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา, ชลบุรี, ประเทศไทย
[2,3]ภาควิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพมหานคร, ประเทศไทย*

*ผู้ประพันธ์บรรณกิจ อีเมล : thitiporn.cham@ku.th

รับต้นฉบับ : 1 มีนาคม 2566; รับบทความฉบับแก้ไข : 9 เมษายน 2566; ตอบรับบทความ : 21 เมษายน 2566
เผยแพร่ออนไลน์: 27 ธันวาคม 2566

บทคัดย่อ

การวางแผนสั่งวัสดุก่อสร้างเป็นปัญหาที่มีความซับซ้อนสูงในอุตสาหกรรมก่อสร้างเนื่องจากจะต้องสั่งซื้อวัสดุให้ถูกประเภทในปริมาณและช่วงเวลาที่เหมาะสม มิฉะนั้นจะทำให้เกิดความสูญเสียและผลกระทบ การศึกษาครั้งนี้เป็นการพัฒนาแบบจำลองปริมาณการสั่งซื้อที่ประหยัดสำหรับโครงการก่อสร้าง โดยการลดข้อจำกัดของวิธีในปัจจุบันที่คำนวณโดยการสั่งในปริมาณที่เท่ากันทุกครั้ง และมีความถี่ที่เท่ากัน ซึ่งอาจส่งผลให้มีค่าใช้จ่ายสูงขึ้น อาทิเช่น ค่าใช้จ่ายในการกองเก็บเหล็กเส้นที่เหลือจากการสั่งในปริมาณที่เกินกว่าความต้องการในแต่ละวัน ผู้วิจัยจึงได้ศึกษาและพัฒนาแบบจำลองการสั่งซื้อที่ประหยัด เพื่อคำนวณระยะเวลาในการสั่งเหล็กเส้นที่สอดคล้องกับแผนกำหนดเวลาของโครงการ เพื่อลดค่าใช้จ่ายในการสั่งซื้อ ค่าใช้จ่ายในการเก็บรักษา ค่าขนส่ง ค่าขนย้าย ค่าเสียโอกาส ค่าปรับ ค่าใช้จ่ายทางอ้อม และค่าใช้จ่ายที่เกิดจากการว่างงาน และลดความล่าช้าของกิจกรรมก่อสร้าง โดยนำวิธีการเชิงพันธุกรรมมาช่วยในการหาคำตอบที่ดีที่สุดสำหรับระยะเวลาและปริมาณที่จะต้องสั่งวัสดุก่อสร้าง โดยวิธีการดังกล่าวมีข้อดีในการช่วยหาคำตอบที่ใกล้เคียงกับคำตอบที่ดีที่สุด สามารถหาแนวโน้มคำตอบที่ถูกปรับปรุงขึ้น และช่วยสร้างแผนในการสั่งซื้อเหล็กเส้นได้หลาย ๆ รูปแบบ ซึ่งวิธีการข้างต้นได้นำมาทดลองกับโครงการกรณีศึกษาทั้งกรณีที่คำนึงถึงระยะเวลาที่เร็วที่สุด และกรณีที่คำนึงถึงค่าใช้จ่ายที่ถูกที่สุด พบว่าวิธีการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดเมื่อเปรียบเทียบกับวิธีการยอมให้มีการขาดแคลนวัสดุตามแผนงาน และวิธีการสั่งซื้อแบบมีส่วนลด พบว่าสามารถช่วยลดระยะเวลาและค่าใช้จ่ายให้น้อยลง

คำสำคัญ : แบบจำลองการสั่งซื้อที่ประหยัด วิธีการเชิงพันธุกรรม ต้นทุนการจัดซื้อวัสดุก่อสร้าง

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Economic Rebar Order Quantity Model for Construction Project

Narongrit Wongwai[1] Thitiporn Chamcharas[2*] Suphawut Malaikrisanachalee[3]

[1]*Faculty of Engineering at Sriracha, Kasetsart University Sriracha Campus, Chon Buri, Thailand*
[2*,3]*Department of Civil Engineering, Faculty of Engineering, Kasetsart University, Bangkok, Thailand*

*Corresponding Author. E-mail address: thitiporn.cham@ku.th

Received: 1 March 2023; Revised: 9 April 2023; Accepted: 21 April 2023
Published online: 27 December 2023

Abstract

Planning and ordering construction materials is a highly complex problem in the construction industry as it has to order the right type of material in the right quantity and at the right time. Otherwise it will cause losses and consequences. This study is to develop an Economical Order Quantity Model for Construction Projects. By reducing the limitation of the current method to calculate by ordering the same amount every time and have the same frequency this may result in higher expenses, for example the cost of storing the remaining rebar after ordering in excess of daily demand. Therefore, the researcher has studied and developed a Modified Economic Order Quantity Model in order to calculate the timing of ordering rebar in accordance with the project schedule to reduce the ordering cost, holding cost, delivery cost, handling cost, opportunity cost, penalty cost, indirect cost and idling cost and reduce delays in construction activities by using a Genetic Algorithm to help find the best answer for the time and quantity to order construction materials. This method has the advantage of helping to find the closest answer to the best answer can find the trend of improved answers and help create a plan for ordering steel bars in many forms the above methods have been used to experiment with case study projects, both of which consider the fastest time and the case that takes into account the cheapest cost. It was found that the modified economic order quantity model method compared to EOQ with Planned Shortage and EOQ with quantity discounts found that it can help reduce the time and cost to be less.

***Keywords*:** Economic order quantity model (EOQ), Genetic algorithm (GA), The cost of purchasing construction materials

1) บทนำ

การวางแผนสั่งวัสดุก่อสร้างเป็นปัญหาที่มีความซับซ้อนสูงในอุตสาหกรรมก่อสร้างเนื่องจากจะต้องสั่งซื้อวัสดุให้ถูกประเภทในปริมาณและช่วงเวลาที่เหมาะสม มิฉะนั้นจะทำให้เกิดความสูญเสียและผลกระทบ โดยการสั่งซื้อวัสดุในแต่ละครั้งจะต้องตรวจสอบปริมาณจากแบบก่อสร้าง รวมถึงการตรวจสอบรายละเอียดคุณสมบัติของวัสดุที่ต้องจัดซื้ออย่างละเอียดรอบคอบจากข้อกำหนดงานก่อสร้าง อย่างไรก็ดีการสั่งซื้อวัสดุบางอย่างอาจมีความจำเป็นต้องเผื่อปริมาณวัสดุที่อาจสูญเสียในกระบวนการทำงานจริง เช่น การสั่งซื้อเหล็กเส้นจะต้องเผื่อปริมาณเศษเหล็กที่สูญเสียจากการตัดเหล็ก เป็นต้น การสั่งซื้อวัสดุในปริมาณที่น้อยเกินไปจะทำให้มีวัสดุไม่เพียงพอต่อการใช้งานทำให้การทำงานล่าช้า และยังส่งผลให้คนงานและเครื่องจักรกลอยู่ในภาวะว่างงาน ซึ่งส่งผลให้เกิดความสูญเสียอื่น ๆ ตามมา การสั่งซื้อวัสดุเร็วเกินไปเป็นการสร้างภาระที่จะต้องจัดหาพื้นที่จัดเก็บวัสดุ ทำให้เกิดค่าใช้จ่ายในการเก็บรักษา (holding cost) ที่ยังไม่มีความจำเป็นต้องใช้งานในเวลานั้น นอกจากนั้นวัสดุบางอย่างอาจเกิดความเสียหายถ้าจัดเก็บไม่เหมาะสมหรืออาจหมดอายุใช้งาน อีกทั้งยังเป็นการเพิ่มต้นทุนค่าเสียโอกาส (opportunity cost) เพราะใช้เงินไปกับการซื้อวัสดุที่ยังไม่มีความจำเป็นต้องใช้งาน อย่างไรก็ดีวัสดุบางประเภทที่ราคามีความผันผวนสูงหรือเป็นวัสดุที่หายากและอาจขาดแคลนในอนาคต ผู้จัดการโครงการอาจตัดสินใจจัดซื้อและกักตุนวัสดุก่อสร้างเหล่านั้นล่วงหน้า โดยยอมเสียค่าจัดเก็บ และค่าเสียโอกาสในการลงทุน เพื่อประกันความเสี่ยงและควบคุมต้นทุนของโครงการก่อสร้าง การสั่งซื้อวัสดุช้าเกินไปจะทำให้ไม่มีวัสดุก่อสร้างใช้งาน ทำให้ไม่สามารถทำงานตามแผนงานที่วางไว้ นอกจากนั้นยังส่งผลให้คนงานและเครื่องจักรกลอยู่ในภาวะว่างงานและส่งผลให้โครงการล่าช้า ดังนั้นผู้จัดการโครงการจะต้องรู้กำหนดเวลาที่จะต้องใช้วัสดุแต่ละประเภทโดยการตรวจสอบจากกำหนดเวลาก่อสร้าง (scheduling) และรู้ระยะเวลาการสั่งซื้อล่วงหน้า (lead time) ของวัสดุแต่ละประเภท เพื่อที่จะสามารถสั่งวัสดุในช่วงเวลาที่เหมาะสมเพื่อลดต้นทุนและความสูญเสียที่จะเกิดขึ้นในโครงการก่อสร้าง

จากปัญหาข้างต้น โครงการจำเป็นต้องมีรูปแบบการบริหารสินค้าคงคลังที่เป็นระบบ รวมถึงการกำหนดปริมาณ การสั่งซื้อวัสดุที่เหมาะสม เพื่อรักษาปริมาณวัสดุไม่ให้มีน้อยหรือมากเกินความต้องการและลดค่าใช้จ่าย จากการศึกษาที่ผ่านมาได้นำแนวทางการสั่งซื้อสินค้าที่ประหยัดที่สุด (Economic Order Quantity) มาวางแผนการสั่งวัสดุก่อสร้างเข้าโครงการด้วยวิธีการเพิ่มประสิทธิภาพ (optimization) ซึ่งพบว่าสามารถหาคำตอบได้ดีในมุมของค่าใช้จ่ายที่ถูกลง แต่วิธีการหาคำตอบค่อนข้างนานและไม่ได้คำนึงถึงระยะเวลาการก่อสร้าง หรือบางคนใช้วิธีการข้างต้นมาคำนวณหาปริมาณที่สั่งวัสดุในแต่ละครั้งให้สอดคล้องกับแผนงานก่อสร้าง แต่ความถี่ยังเท่ากันอยู่

ดังนั้นผู้วิจัยจึงพัฒนารูปแบบวิธีการแก้ไขแบบจำลองการสั่งซื้อที่ประหยัด (modified economic order quantity model) เพื่อลดข้อจำกัดของวิธีการสั่งซื้อสินค้าที่ประหยัดที่สุด (Economic Order Quantity) แบบดั้งเดิมและข้อจำกัดของงานวิจัยที่ผ่านมาที่ยังไม่ได้ศึกษาในประเด็นการเพิ่มความรวดเร็วในการสั่งวัสดุก่อสร้างเข้าโครงการให้เหมาะสมกับงานโครงสร้างและสถาปัตยกรรม ซึ่งส่งผลให้ลดระยะเวลาและค่าใช้จ่ายของโครงการ และยังไม่ได้ศึกษาในส่วนของค่าใช้จ่ายในโครงการก่อสร้างจริงทั้งหมด ได้แก่ ค่าใช้จ่ายในการสั่งซื้อ (ordering cost) ค่าใช้จ่ายในการเก็บรักษา (holding cost) ค่าขนส่ง (delivery cost) ค่าใช้จ่ายทาวเวอร์เครน (handling cost) ค่าเสียโอกาส (opportunity cost) ค่าปรับ (penalty cost) ค่าใช้จ่ายทางอ้อม (indirect cost) และค่าใช้จ่ายที่เกิดจากการว่างงาน (idling cost)

2) ทบทวนวรรณกรรม

การวางแผนแบบ CPM (Critical Path Method) [1] เป็นเทคนิคที่ใช้ในการบริหารโครงการ โดยการนำกิจกรรมต่าง ๆ มาเขียนเครือข่ายแสดงความสัมพันธ์ และคำนวณหาเวลาเริ่มต้นของกิจกรรม และเวลาสิ้นสุดของแต่ละกิจกรรม เพื่อทำให้โครงการสำเร็จและใช้ระยะเวลาสั้นที่สุด เวลาดังกล่าว ประกอบด้วย เวลาที่เริ่มต้นเร็วที่สุด (Early Starting: ES) เวลาที่เสร็จเร็วที่สุด (Early Finishing: EF) เวลาที่เสร็จช้าที่สุด (Late Finishing: LF) เวลาที่เริ่มต้นช้าที่สุด (Late Starting: LS) เวลาพอเพียง (Total Float: TF) เวลาลอยตัว (Free Float: FF) และสายงานวิกฤต

การบริหารสินค้าคงคลัง [2], [3] คือการดูแล วางแผน และจัดการสินค้าคงคลัง วัตถุดิบ และสต๊อกสินค้า โดยที่เป้าหมายของการบริหารสินค้าคงคลังคือการดูการไหลเวียนของสินค้าจากกระบวนการผลิตไปกระบวนการจัดเก็บ ต้นทุนของวัสดุคงคลัง (inventory cost) ประกอบด้วย ค่าสั่งซื้อ (ordering costs) ค่าขนส่ง (delivery costs) ค่าใช้จ่ายในการเก็บรักษา (holding cost) ค่าใช้จ่ายเนื่องจากวัสดุขาดแคลน (shortage cost) และค่าเสียโอกาส (opportunity costs)

การสั่งซื้อที่ประหยัดที่สุด (Economic Order Quantity: EOQ) [4], [5] เป็นการสั่งซื้อวัสดุที่เหมาะสมคุ้มค่า เพื่อไม่ให้เกิดต้นทุนจมจากการสต็อกวัสดุในคลังไม่ว่าจะเป็นการเคลื่อนย้าย การจัดเก็บ ต้นทุนเสื่อมสภาพ ฯลฯ ช่วยให้รู้ว่าควรสั่งของมาปริมาณมากเท่าไร โดยคำนวณจากความต้องการวัสดุ (demand) ก่อนที่จะนำเข้ามาสต็อกไว้ดังสมการที่ (1) แต่บางทีก็ยอมให้วัสดุขาดแคลนตามแผนงานดังสมการที่ (2) และเมื่อซื้อวัสดุจำนวนมาก ทางฝ่ายจัดซื้อจะต่อรองให้ราคาต่อหน่วยของวัสดุลดลง ยิ่งสั่งซื้อจำนวนมาก ราคาต่อหน่วยยิ่งลดลง มีผลทำให้ต้นทุนการเก็บรักษาเปลี่ยนแปลงตามไปด้วยดังสมการที่ (3)

$$Q^* = \sqrt{2DC_o/C_h} \tag{1}$$

$$Q^* = \sqrt{2DC_o/C_h}\sqrt{(C_h + C_b)/C_b} \tag{2}$$

$$TC_i = C_o + C_h + DC_i \tag{3}$$

เมื่อ Q^* คือ ปริมาณการสั่งซื้อที่ประหยัดต่อครั้ง (Economic Order Quantity)

C_o คือ ต้นทุนการสั่งซื้อต่อครั้ง

C_h คือ ต้นทุนการเก็บรักษาต่อครั้ง

C_b คือ ต้นทุนการสั่งซื้อที่ค้างส่งต่อครั้ง

D คือ ปริมาณวัสดุที่ต้องการใช้ต่อครั้ง

TC คือ ต้นทุนรวม (total cost)

C_i คือ ต้นทุนส่วนลดตามปริมาณการสั่งซื้อ

วิธีเชิงพันธุกรรม (Genetic Algorithm: GA) [6], [7] เป็นวิธีการค้นหาคำตอบที่ดีที่สุดโดยใช้หลักการคัดเลือกแบบธรรมชาติเลียนแบบวิวัฒนาการของสิ่งมีชีวิต โดยอาศัยแนวคิดที่ว่า ยีนของสิ่งมีชีวิตที่แข็งแรงจะมีโอกาสในการสร้างทายาทจากรุ่นสู่รุ่นมากกว่ายีนของสิ่งมีชีวิตที่ไม่แข็งแรง ประกอบด้วย ประชากร (population) พ่อและแม่พันธุ์ (parent) ทายาท (offspring) การสุ่มเลือก (selection) ผสมข้ามสายพันธุ์ (crossover) ดังตารางที่ 1 กลายพันธุ์ (mutation) ดังตารางที่ 2 ความแข็งแรง (fitness value) และฟังก์ชันความแข็งแรง (fitness function) โดยการสุ่มเลือกจะใช้วิธี Roulette wheel ดังแสดงในสมการที่ (4)

$$P_i = f_i / \sum_j (f_i) \text{ for } j = 1, \ldots, N \tag{4}$$

เมื่อ P_i คือ ประชากร (population)

f_i คือ ความแข็งแรง (fitness value)

N คือ จำนวนของประชากร (number of population)

ตารางที่ 1 : ตัวอย่างการผสมข้ามสายพันธุ์ (Crossover)

ประชากร (Population)						
Parent 1	1	2	6	5	3	4
Parent 2	1	6	2	4	5	3
ผสมข้ามสายพันธุ์ (Crossover)						
Offspring 1	1	2	6	4	5	3
Offspring 2	1	6	2	5	3	4

ตารางที่ 2 : ตัวอย่างการกลายพันธุ์ (Mutation)

ประชากร (Population)						
Parent 1	1	2	6	5	3	4
กลายพันธุ์ (Mutation)						
Offspring 1	1	2	5	6	3	4

การใช้อัลกอริทึมเชิงพันธุกรรม [8] ได้รับการพิสูจน์แล้วว่าเป็นแนวทางที่น่าพึงพอใจสำหรับการเพิ่มประสิทธิภาพกำหนดการจัดส่งวัสดุและการควบคุมระดับสินค้าคงคลังที่เกี่ยวข้อง และทำให้แผนการจัดส่งวัสดุมีต้นทุนลดลงเมื่อเทียบกับคู่สัญญาจริง นอกจากนี้เวลาในการประมวลผลของคอมพิวเตอร์สำหรับการพัฒนาแผนเพิ่มประสิทธิภาพนั้นค่อนข้างน้อยซึ่งส่งเสริมกับการใช้งานจริง

3) วิธีดำเนินการวิจัย

การทำวิจัยครั้งนี้ผู้วิจัยได้มีการทบทวนทฤษฎีและงานวิจัยที่เกี่ยวข้อง แล้วนำความรู้ที่ได้มาพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดที่ใช้ในการสั่งเหล็กเส้นเข้าโครงการก่อสร้าง และนำมาพัฒนา Excel Speed Sheet เพื่อทำการคำนวณด้วยวิธีการเชิงพันธุกรรม จากกนั้นนำมาทดลองกับโครงการกรณีศึกษาโครงสร้างขนาดเล็ก 2 ชั้น ซึ่งประกอบด้วย 7 กิจกรรม แล้วรวบรวมข้อมูลผลการทดลองมาวิเคราะห์ เพื่อหาผลสรุปและข้อเสนอแนะ โดยมีกระบวนการจัดทำดังต่อไปนี้

3.1) การพัฒนา Excel Speed Sheet ที่ใช้ในการคำนวณ

เป็นการพัฒนาเพื่อใช้ในการคำนวณปัญหาแบบอัตโนมัติด้วยวิธีการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัด และวิธีการเชิงพันธุกรรม โดยสามารถนำมาใช้กับแผนกำหนดเวลาโครงการก่อสร้างได้ทุกขนาดแล้วแต่จะป้อนข้อมูล งานวิจัยนี้ผู้วิจัยเลือกจำลองปัญหาแผนกำหนดเวลาโครงการก่อสร้างขนาดเล็ก 2 ชั้น โดยมีระยะเวลาโครงการตามสัญญา 20 วันดังแสดงในตารางที่ 3 และ 4 และกำหนดค่าสั่งซื้อและค่าขนส่งวัสดุดังตารางที่ 5 และค่าใช้จ่ายที่เกี่ยวข้องในโครงการก่อสร้างดังตารางที่ 6

ตารางที่ 3 : ตัวอย่างการพัฒนา Excel speed sheet ที่ใช้ในการคำนวณ

1	2	3	4	5	6	1	7	1
Act.	Dur.(day)	Pred.	Delay Start (day)	ST.(day)	FN.(day)	Total Mat. Req.(ton)	Daily Mat. Req.(ton)	
A	3		2	2	5	81.6	27.2	0
B	4	A	2	7	11	13.5	3.375	0
C	5	A	1	6	11	33.6	6.72	0
D	4	B, C	0	11	15	17	4.25	0
E	3	C	2	13	16	72.9	24.3	0
F	2	D, E	1	17	19	56.7	28.35	0
G	3	E, F	1	20	23	40.5	13.5	0
Daily Demand of Materials 7								0
Cumulative Daily Demand of Materials 8								0
The Amount of Ordering Material 4								83
Cumulative Daily Supply of Materials 9								83
Cumulative Inventory of Materials 10								83
Ordering Cost 11							17,545,000	830,000
Delivery Cost 12							28,424.00	1,369.50
Holding Cost 13							463,164.00	2,490
Handling Cost 14							43,976.36	2,393.31
Opportunity Cost 15							4,041,887.52	
Penalty Cost 16							3,000.00	
Idling Cost 17							4,500.00	
Total Cost 18							22,129,951.88	
Project Duration 19							23.00	

ตารางที่ 4 : ตัวอย่างคำตอบที่ได้จากการสุ่มใน Excel speed sheet

Obj1 (Dur.) (day)	Obj2 (Total Cost) (baht)	20 Obj3= 1+2	21 fitness value	22 Proba -bility	23 Cum. Prob.
23	23,752,150	46.75	223.245	0.037	0.037
23	22,391,992	45.39	224.108	0.037	0.074
23	24,678,836	47.68	68.610	0.011	0.086
23	21,018,078	44.02	211.803	0.035	0.121
23	19,693,756	42.69	227.680	0.038	0.159
23	25,596,844	48.60	218.111	0.036	0.195

ตารางที่ 5 : ค่าสั่งซื้อ และค่าขนส่งวัสดุ

ต้นทุนค่าใช้จ่าย	ปริมาณ (ตัน)	ราคาต่อหน่วย (บาท/ตัน)
7 ค่าสั่งซื้อ (Ordering Cost)	0 - 50	15,000
	50 - 100	10,000
8 ค่าขนส่ง (Delivery Cost)	0 - 30	5.5
	30 - 60	11
	60 - 90	16.50

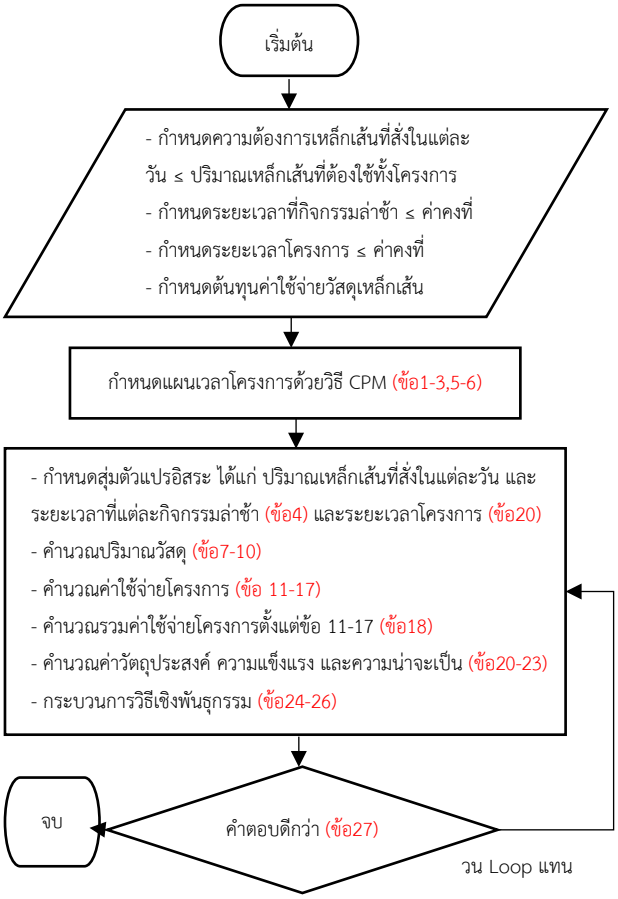
แหล่งที่มา : บัญชีราคาค่าวัสดุและค่าแรงงาน [9] กรมบัญชีกลาง [10]

ตารางที่ 6 : ค่าใช้จ่ายที่เกี่ยวข้องในโครงการก่อสร้าง

ต้นทุนค่าใช้จ่าย	จำนวนเงิน	หน่วย
9 อัตราดอกเบี้ยรายวัน (daily interest rate)	0.01	%/วัน
10 ค่าปรับงานล่าช้า (late schedule penalty	1,000	บาท/วัน
11 ค่าใช้จ่ายในการเก็บรักษา (holding cost rate)	30	บาท/ตัน-วัน
12 ราคาต่อหน่วยทาวเวอร์เครน (tower crane unit price)	4,000	บาท/วัน
13 อัตราการทำงานทาวเวอร์เครน (tower crane productivity)	17.34	ตัน/ชั่วโมง
14 ค่าใช้จ่ายจากการว่างงาน (idling cost)	500	บาท/วัน

แหล่งที่มา : สัดส่วนต้นทุนงานก่อสร้าง [11] อัตราดอกเบี้ย [12] ค่าเช่าทาวเวอร์เครน [13] ค่าแรงคนงานขั้นต่ำ [14]

โดยมีขั้นตอนการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดดังรูปที่ 1



รูปที่ 1 : ขั้นตอนการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัด

3.1.1) กำหนดข้อมูลนำเข้า ได้แก่ ความต้องการเหล็กเส้นที่สั่งในแต่ละวัน ≤ ปริมาณเหล็กเส้นที่ต้องใช้ทั้งโครงการ และระยะเวลาที่กิจกรรมล่าช้า ≤ ค่าคงที่ ดังข้อที่ 4 และระยะเวลาโครงการ ≤ ค่าคงที่ ดังข้อที่ 20 และต้นทุนค่าใช้จ่ายเหล็กเส้นดังตารางที่ 5 และ 6

3.1.2) กำหนดแผนเวลาโครงการด้วยวิธี CPM จากข้อมูลตารางที่ 3 กำหนดกิจกรรม และปริมาณเหล็กเส้นของแต่ละกิจกรรมดังข้อที่ 1 กำหนดระยะเวลาของแต่ละกิจกรรมดังข้อที่ 2 และกำหนดขั้นตอนกิจกรรมก่อนหน้าดังข้อที่ 3 จากนั้นคำนวณระยะเวลาของแต่ละกิจกรรมด้วยวิธี CPM ดังข้อที่ 5 และ 6

3.1.3) คำนวณปริมาณวัสดุและค่าใช้จ่ายโครงการ จากข้อมูลตารางที่ 3 กำหนดสุ่มตัวแปรอิสระโดยใช้วิธี Roulette wheel ดังสมการที่ (4) ได้แก่ ปริมาณการสั่งซื้อวัสดุเหล็กเส้นที่สั่งในแต่ละวัน ≤ ปริมาณเหล็กเส้นที่ต้องใช้ทั้งโครงการ (315.8 ตัน) และระยะเวลาที่แต่ละกิจกรรมล่าช้า ≤ ค่าคงที่ (5 วัน) ดังข้อที่ 4 และระยะเวลาโครงการกำหนดยอมให้ล่าช้า ≤ ค่าคงที่ (35 วัน) ดังข้อที่ 20 จากนั้นนำข้อมูลที่ได้มาคำนวณหาปริมาณวัสดุและค่าใช้จ่ายโครงการ ดังนี้

1) ปริมาณความต้องการวัสดุในแต่ละวัน ดังข้อที่ 7

= ปริมาณวัสดุของกิจกรรม ÷ ระยะเวลาของกิจกรรมนั้น

2) ปริมาณความต้องการวัสดุสะสมรายวัน ดังข้อที่ 8

= ความต้องการวัสดุสะสมก่อนหน้า 1 วัน + ความต้องการวัสดุในวันนั้น

3) ปริมาณการสั่งซื้อวัสดุสะสมรายวัน ดังข้อที่ 9

= ปริมาณการสั่งซื้อวัสดุที่สั่งสะสมก่อนหน้า 1 วัน + ปริมาณการสั่งซื้อวัสดุที่สั่งในวันนั้น

4) ปริมาณกองเก็บวัสดุสะสม ดังข้อที่ 10

= ปริมาณการสั่งซื้อวัสดุสะสมรายวัน - ปริมาณความต้องการวัสดุสะสมรายวัน

5) ค่าสั่งซื้อ ดังข้อที่ 11

= ปริมาณการสั่งซื้อวัสดุ x ค่าสั่งซื้อ

6) ค่าขนส่ง ดังข้อที่ 12

= ปริมาณการสั่งซื้อวัสดุ x ค่าขนส่ง

7) ค่าใช้จ่ายในการเก็บรักษา ดังข้อที่ 13

= ปริมาณกองเก็บวัสดุสะสม x ค่าใช้จ่ายในการเก็บรักษา

8) ค่าใช้จ่ายทาวเวอร์เครน ดังข้อที่ 14

= (ความต้องการวัสดุของแต่ละกิจกรรมต่อวัน + ปริมาณการสั่งซื้อ) ÷ อัตราการทำงานทาวเวอร์เครน x ราคาต่อหน่วยทาวเวอร์เครน

9) ค่าเสียโอกาส ดังข้อที่ 15

= (ค่าสั่งซื้อ + ค่าขนส่ง) x อัตราดอกเบี้ยรายวัน x ระยะเวลาโครงการ

10) ค่าปรับงานล่าช้า ดังข้อที่ 16

= (ระยะเวลาโครงการ - ระยะเวลาโครงการตามสัญญา) x ค่าปรับงานล่าช้า

11) ค่าใช้จ่ายจากการว่างงาน ดังข้อที่ 17

= ผลรวมของวันล่าช้า x ค่าใช้จ่ายจากการว่างงาน

12) รวมค่าใช้จ่ายโครงการ ดังข้อที่ 18

= ผลรวมของค่าใช้จ่ายโครงการ ตั้งแต่ข้อที่ 11 ถึง 17

จากข้อมูลตารางที่ 4 เป็นการคำนวณดังต่อไปนี้

13) ค่าวัตถุประสงค์ ดังข้อที่ 20

= (เปอร์เซ็นต์ x (ผลรวมของค่าใช้จ่ายทั้งหมด ÷ 1,000,000)) + (เปอร์เซ็นต์ x ระยะเวลาโครงการ)

14) ค่าความแข็งแรงของคำตอบ ดังข้อที่ 21

= 10,000 ÷ ค่าวัตถุประสงค์

15) ค่าความน่าจะเป็นที่จะปรับปรุง ดังข้อที่ 22

= ความแข็งแรงในวันนั้น ÷ ผลรวมของความแข็งแรง

16) ค่าความน่าจะเป็นสะสม ดังข้อที่ 23

= ค่าความน่าจะเป็นก่อนหน้า 1 วัน + ค่าความน่าจะเป็นในวันนั้น

3.1.4) คำนวณหาคำตอบที่ดีที่สุดด้วยวิธีเชิงพันธุกรรม

1) สุ่มตัวแปรอิสระโดยใช้วิธี Roulette wheel ดังแสดงในสมการที่ (4) ถ้าอยู่ระหว่างความน่าจะเป็นสะสมของค่าไหนก็จะนำมาเป็นพ่อพันธุ์และแม่พันธุ์ ดังข้อที่ 24

2) นำพ่อพันธุ์และแม่พันธุ์ที่ได้จากการสุ่มมาผสมข้ามสายพันธุ์ (Crossover) กันให้เกิดทายาท ดังข้อที่ 25

= (พ่อพันธุ์และแม่พันธุ์ x (วันล่าช้าหรือปริมาณการสั่งซื้อวัสดุเหล็กเส้น – 2) + 2)

3) สุ่มตำแหน่งที่จะสลับคำตอบของทายาทจนกลายพันธุ์เป็นอีกรุ่น (Mutation) ดังข้อที่ 26

= การสุ่มทายาทจาก Crossover x (วันล่าช้าหรือปริมาณการสั่งซื้อวัสดุเหล็กเส้น – 1) + 1

4) ถ้าได้คำตอบที่ดีกว่าจะวนกลับไปแทนค่าในข้อที่ 4 แล้วได้คำตอบชุดใหม่ในข้อที่ 5 ถึง 27 แล้ววนคำตอบจนกว่าจะได้คำตอบที่ดีที่สุด ดังข้อที่ 27

3.2) กรณีศึกษาโครงการก่อสร้างขนาดเล็ก 2 ชั้น

จากข้อมูลแบบจำลองปัญหากรณีศึกษาแผนกำหนดเวลาโครงการก่อสร้างขนาดเล็ก 2 ชั้น ผู้วิจัยได้ใช้ VBA (Visual Basic for Application) ในการพัฒนา Excel speed sheet เพื่อทำการเขียนชุดคำสั่งวิธีการเชิงพันธุกรรมดังขั้นตอนที่กล่าวข้างต้นหาคำตอบของแต่ละรุ่น โดยองค์ประกอบของชุดคำสั่งแสดงตัวอย่างการเขียนในรูปที่ 2, 3 และ 4

```
Const ACT = 7
Const DAYS = 35
Const POPULATION = 30
Const ITERATION = 100

Const Murate = 0.1 'mutation rate
Const MPOINT = 2 'number of points for mutation
```

รูปที่ 2 : การกำหนดตัวแปรอิสระ

```
Math.Randomize
For p = 1 To POPULATION
    For a = 1 To ACT
        GEN(p, a) = Math.Rnd * 5 'Delay Start
        Sheets("Generation").Cells(1 + p, a).Value = GEN(p, a)
    Next a
    For d = 1 To DAYS
        GEN2(p, d) = Math.Rnd * 315.8 'Amount of Ordering
        Sheets("Generation").Cells(1 + p, 7 + d).Value = GEN2(p, d)
```

รูปที่ 3 : การกำหนดค่าคงที่วันล่าช้าและความต้องการเหล็กเส้น

```
Public Sub FindObj3(ByVal p As Integer)
    Obj3(p) = Obj(p) + (Obj2(p) / 1000000)
    Sheets("Generation").Cells(1 + p, 43).Value = Obj3(p)
End Sub

Private Sub FitnessValue()
Dim p As Integer
For p = 1 To POPULATION
    fVAL(p) = 10000 / Obj3(p)
    Sheets("Generation").Cells(1 + p, 44).Value = fVAL(p)
    Next p
End Sub

Private Sub CalProb()
    Dim p As Integer
    Dim total As Single
    Dim cum As Single
    total = 0


    For p = 1 To POPULATION
        total = total + fVAL(p)
    Next p

    For p = 1 To POPULATION
        Prob(p) = fVAL(p) / total
        cum = cum + Prob(p)
        CumProb(p) = cum
        Sheets("Generation").Cells(1 + p, 45).Value = Prob(p)
        Sheets("Generation").Cells(1 + p, 46).Value = CumProb(p)
```

รูปที่ 4 : การเขียนสมการวัตถุประสงค์ ค่าแข็งแรงและความน่าจะเป็น

จากข้อมูลข้างต้นผู้วิจัยได้ทำการวิเคราะห์ และเปรียบเทียบโดยแบ่งกรณีศึกษาออกเป็นดังนี้

3.2.1) วิเคราะห์และเปรียบเทียบกรณีการสั่งซื้อวัสดุก่อสร้างเข้าโครงการ

1) กรณียอมให้มีการขาดแคลนวัสดุตามแผนงาน กรณีนี้จะให้ความสำคัญกับระยะเวลาโครงการ (Project Duration) 100% โดยเร่งระยะเวลาโครงการไม่ให้ล่าช้า ซึ่งไม่ได้คำนึงถึงปริมาณการสั่งซื้อวัสดุ และค่าใช้จ่ายโครงการ

2) กรณีไม่ยอมให้มีการขาดแคลนวัสดุตามแผนงาน กรณีนี้จะให้ความสำคัญกับค่าใช้จ่ายโครงการ (Project Cost) 100% โดยให้ความสำคัญกับปริมาณการสั่งซื้อวัสดุกับค่าใช้จ่ายโครงการ ซึ่งไม่ได้คำนึงถึงระยะเวลาโครงการ

3) การหาค่าระยะเวลาโครงการที่เหมาะสมที่สุด (Optimal Project Duration) และค่าใช้จ่ายโครงการที่เหมาะสมที่สุด (Optimal Project Cost) โดยใช้วิธีการประมาณค่าในช่วง (Interpolation)

แบ่งการให้ความสำคัญระยะเวลาโครงการ และค่าใช้จ่ายโครงการออกเป็น 5 กรณี ได้แก่

1) Project Duration 50% + Project Cost 50%

2) Project Duration 70% + Project Cost 30%

3) Project Duration 30% + Project Cost 70%

4) Project Duration 60% + Project Cost 40%

5) Project Duration 40% + Project Cost 60%

3.2.2) วิเคราะห์และเปรียบเทียบหาค่าความเหมาะสมที่สุดของการวนหาคำตอบ การหาค่าความเหมาะสมที่สุดจะใช้วิธีการวนหาคำตอบของปริมาณการสั่งซื้อวัสดุ ระยะเวลาโครงการ และค่าใช้จ่ายโครงการ โดยแบ่งรุ่นออกเป็น 4 กรณี ได้แก่ 10 รุ่น, 30 รุ่น, 70 รุ่น และ 100 รุ่น

3.2.3) วิเคราะห์และเปรียบเทียบหาผลลัพธ์คำตอบที่ดีที่สุด การเปรียบเทียบระหว่างคำตอบที่ได้จากวิธีการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัด โดยการใช้วิธีเชิงพันธุกรรม และสมการการสั่งซื้อที่ประหยัด ดังสมการที่ (1), (2) และ (3)

4) ผลการวิจัยและอภิปรายผล

4.1) วิเคราะห์และเปรียบเทียบกรณีการสั่งซื้อวัสดุก่อสร้างเข้าโครงการ

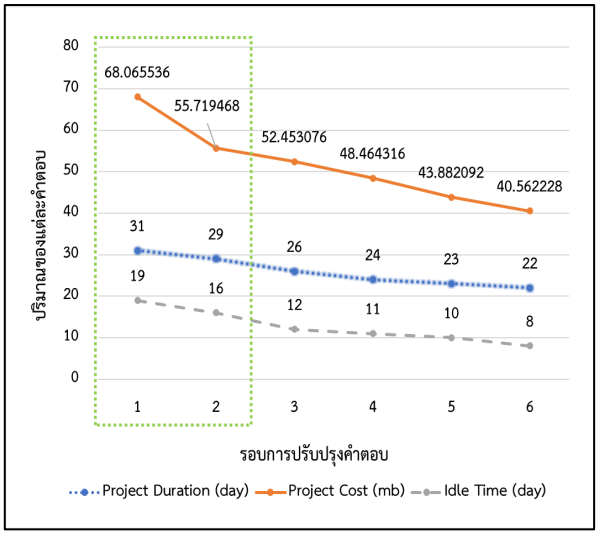
4.1.1) กรณียอมให้มีการขาดแคลนวัสดุตามแผนงาน ผู้วิจัยได้ทำยกตัวอย่างกรณีศึกษาของการปรับปรุงคำตอบรอบที่ 1 และ 2 มาวิเคราะห์ โดยจะให้ความสำคัญกับ Project Duration 100% เพื่อหาระยะเวลา และค่าใช้จ่ายโครงการรวมต่ำที่สุดดังแสดงในตารางที่ 7 และ 8 และรูปที่ 5 สามารถวิเคราะห์ได้ดังต่อไปนี้

ตารางที่ 7 : เปรียบเทียบค่าใช้จ่ายการปรับปรุงคำตอบรอบที่ 1 และ 2

Inventory Cost	รอบที่ 1	รอบที่ 2	Diff.
Project Duration (day)	31	29	2
Project Cost (baht)	68,065,536	55,719,468	12,346,068
Ordering Costs (baht)	49,865,000	41,600,000	8,265,000
Delivery Cost (baht)	107,618.50	87,367.50	20,251
Holding cost (baht)	2,436,051	1,800,879	635,172

ตารางที่ 7 : เปรียบเทียบค่าใช้จ่ายการปรับปรุงคำตอบรอบที่ 1 และ 2 (ต่อ)

Inventory Cost	รอบที่ 1	รอบที่ 2	Diff.
Handling Cost (baht)	144,858.13	124,886.68	19,971.45
Opportunity Cost (baht)	15,491,512	12,089,337	3,402,175
Penalty Cost (baht)	11,000	9,000	2,000
Idling Cost (baht)	9,500	8,000	1,500



รูปที่ 5 : การเปรียบเทียบระยะเวลาโครงการ ค่าใช้จ่ายโครงการ และระยะเวลาว่างงาน

จากการปรับปรุงคำตอบรอบที่ 1 และ 2 ในรูปที่ 5 พบว่าเมื่อระยะเวลาโครงการลดลง ระยะเวลาว่างงานและค่าใช้จ่ายโครงการก็จะลดลงตามไปด้วย เนื่องจากปริมาณการสั่งซื้อในแต่ละรอบสั่งเข้ามาเป็นทีละจำนวนมาก และไม่ได้สั่งเป็นจำนวนถี่ จึงทำให้ได้รับส่วนลดของราคาตามปริมาณการสั่งซื้อดังตารางที่ 8 จึงทำให้ค่าสั่งซื้อ ค่าขนส่ง ค่าใช้จ่ายในการเก็บรักษา ค่าใช้จ่ายทาวเวอร์เครน ค่าดอกเบี้ยที่กู้ยืม ค่าปรับ ค่าใช้จ่ายทางอ้อมและค่าใช้จ่ายจากการว่างงานก็จะลดลงตามไปด้วยดังตารางที่ 7

ตารางที่ 8 : ปริมาณการสั่งซื้อที่ทำให้ได้รับส่วนลดของราคา 2 วัน

Improve	Pro. Dur. (day)	Pro. Cost (baht)	Ord Cost (baht)	1	2
1	31	68,065,536	49,865,000	178	193
2	29	55,719,468	41,600,000	150	178

4.1.2) กรณีไม่ยอมให้มีการขาดแคลนวัสดุตามแผนงาน ผู้วิจัยได้ทำยกตัวอย่างกรณีศึกษาของการปรับปรุงคำตอบรอบที่ 2, 3 และ 4 มาวิเคราะห์ โดยจะให้ความสำคัญกับ Project Cost 100% เพื่อหาระยะเวลา และค่าใช้จ่ายโครงการรวมต่ำที่สุดดังแสดงในตารางที่ 9, 10 และ 11 และรูปที่ 6 สามารถวิเคราะห์ได้ดังต่อไปนี้

ตารางที่ 9 : เปรียบเทียบค่าใช้จ่ายการปรับปรุงคำตอบรอบที่ 2 และ 3

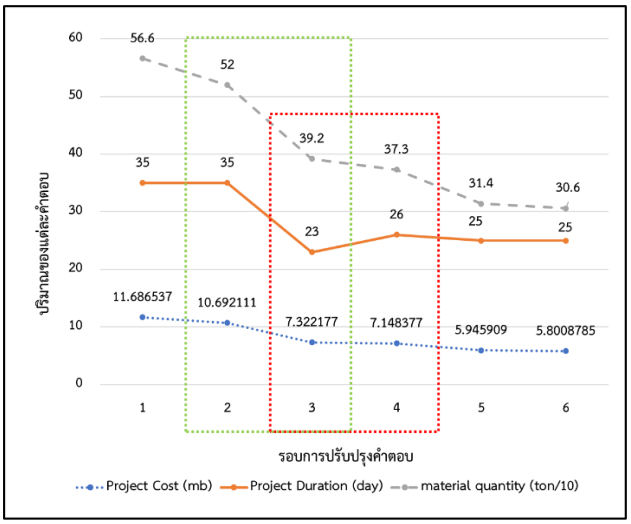
Inventory Cost	รอบที่ 2	รอบที่ 3	Diff.
Project Duration (day)	35	23	12
Project Cost (baht)	10,692,111	7,322,177	3,369,934
Ordering Costs (baht)	7,800,000	5,880,000	1,920,000
Delivery Cost (baht)	2,860	2,156	704
Holding cost (baht)	78,675	40,818	37,857
Handling Cost (baht)	53,075.26	38,307.09	14,768.17
Opportunity Cost (baht)	2,731,001	1,352,895.88	1,378,105.13
Penalty Cost (baht)	15,000	3,000	12,000
Idling Cost (baht)	11,500	5,000	6,500
Material quantity (ton)	520	392	128

ตารางที่ 10 : เปรียบเทียบค่าใช้จ่ายการปรับปรุงคำตอบรอบที่ 3 และ 4

Inventory Cost	รอบที่ 3	รอบที่ 4	Diff.
Project Duration (day)	23	26	-3
Project Cost (baht)	7,322,177	7,148,377	173,800
Ordering Costs (baht)	5,880,000	5,595,000	285,000
Delivery Cost (baht)	2,156	2,051.50	104.50
Holding cost (baht)	40,818	42,696	-1,878
Handling Cost (baht)	38,307.09	40,896.19	-2,589.10
Opportunity Cost (baht)	1,352,895.88	1,455,233.38	-102,337.50
Penalty Cost (baht)	3,000	6,000	-3,000
Idling Cost (baht)	5,000	6,500	-1,500
Material quantity (ton)	392	373	19

ตารางที่ 11 : ปริมาณการสั่งซื้อที่ไม่ได้รับส่วนลดของราคา 2 วัน

Improve	Pro. Cost (baht)	Pro. Dur. (day)	Ord Cost (baht)	1	2
2	10,692,111	35	7,800,000	20	2
3	7,322,177	23	5,880,000	26	23
4	7,148,377	26	5,595,000	12	19



รูปที่ 6 : การเปรียบเทียบค่าใช้จ่ายโครงการ ระยะเวลาโครงการ และปริมาณการสั่งซื้อวัสดุ

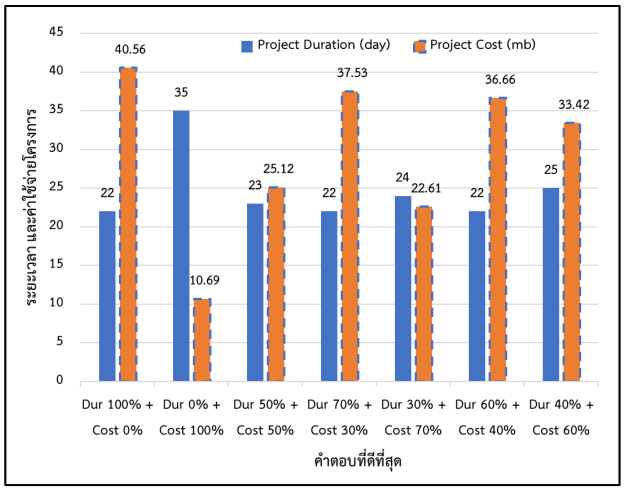
จากการปรับปรุงคำตอบรอบที่ 2 และ 3 ในรูปที่ 6 พบว่าค่าใช้จ่ายโครงการลดลง ปริมาณการสั่งซื้อวัสดุและระยะเวลาโครงการก็จะลดลงตามไปด้วย จึงทำให้ค่าสั่งซื้อ ค่าขนส่ง ค่าใช้จ่ายในการเก็บรักษา ค่าใช้จ่ายทาวเวอร์เครนและค่าใช้จ่ายจากการว่างงานจะลดลงตามปริมาณการสั่งซื้อในแต่ละรอบ โดยสั่งเข้ามาเป็นทีละจำนวนน้อย และสั่งเป็นจำนวนถี่ จึงทำให้ไม่ได้รับส่วนลดของราคาตามปริมาณการสั่งซื้อดังตารางที่ 11 ส่วนค่าดอกเบี้ยที่กู้ยืม ค่าปรับ และค่าใช้จ่ายทางอ้อมก็จะลดลงตามระยะเวลาโครงการดังตารางที่ 9

จากการปรับปรุงคำตอบรอบที่ 3 และ 4 ในรูปที่ 6 พบว่าค่าใช้จ่ายโครงการลดลง ปริมาณการสั่งซื้อวัสดุก็จะลดลงตามไปด้วย แต่ระยะเวลาโครงการกลับเพิ่มขึ้น จึงทำให้มีค่าดอกเบี้ยที่กู้ยืม ค่าปรับ ค่าใช้จ่ายทางอ้อมและค่าใช้จ่ายจากการว่างงานเพิ่มขึ้นตามระยะเวลาโครงการ แต่ค่าสั่งซื้อ ค่าขนส่งและค่าใช้จ่ายทาวเวอร์เครนจะลดลงตามปริมาณการสั่งซื้อในแต่ละรอบ โดยสั่งเข้ามาเป็นทีละจำนวนน้อย และสั่งเป็นจำนวนถี่ จึงทำให้ไม่ได้รับส่วนลดของราคาตามปริมาณการสั่งซื้อดังตารางที่ 11 แต่ค่าใช้จ่ายในการเก็บรักษากลับเพิ่มขึ้นซึ่งเกิดจากการสั่งวัสดุในปริมาณที่มากกว่าปริมาณเหล็กทั้งโครงการ 315.8 ตัน ดังตารางที่ 10

4.1.3) วิเคราะห์และเปรียบเทียบหาค่าความเหมาะสมที่สุดของระยะเวลาโครงการและค่าใช้จ่ายโครงการ โดยใช้วิธีการประมาณค่าในช่วง (Interpolation) แบ่งการให้ความสำคัญระยะเวลาโครงการ และค่าใช้จ่ายโครงการออกเป็น 5 กรณี ได้แก่ Project Duration 50% + Project Cost 50%, Project Duration 70% + Project Cost 30%, Project Duration 30% + Project Cost 70%, Project Duration 60% + Project Cost 40% และ Project Duration 40% + Project Cost 60% แล้วนำคำตอบที่ได้มาเปรียบเทียบกับกรณียอมให้มีการขาดแคลนวัสดุตามแผนงาน Project Duration 100% และกรณีไม่ยอมให้มีการขาดแคลนวัสดุตามแผนงาน Project Cost 100% ดังตารางที่ 12 และรูปที่ 7

ตารางที่ 12 : การเปรียบเทียบหาค่าความเหมาะสมที่สุด

กรณีการสั่งซื้อ	Project Duration (day)	Project Cost (baht)
Dur 100% + Cost 0%	22	40,562,228
Dur 0% + Cost 100%	35	10,692,111
Dur 50% + Cost 50%	23	25,118,760
Dur 70% + Cost 30%	22	37,530,592
Dur 30% + Cost 70%	24	22,606,224
Dur 60% + Cost 40%	22	36,664,856
Dur 40% + Cost 60%	25	33,417,384

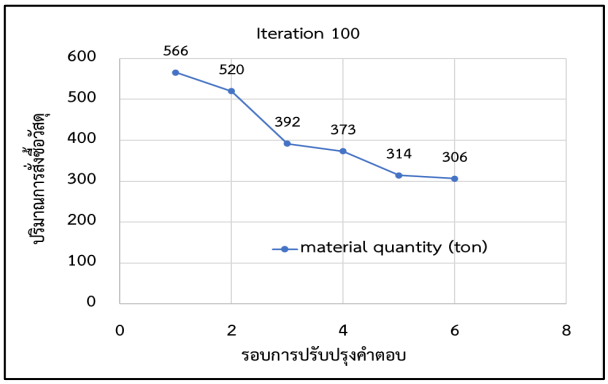


รูปที่ 7 : การเปรียบเทียบหาค่าความเหมาะสมที่สุด

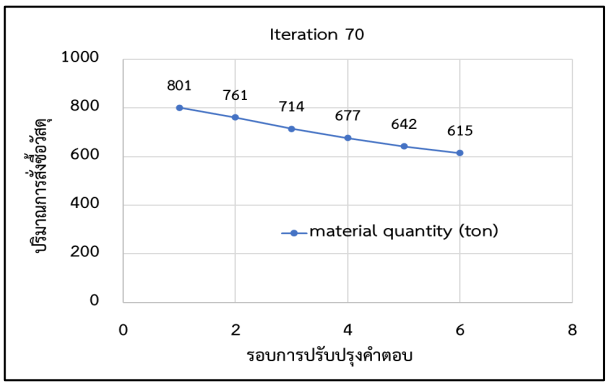
จากรูปที่ 7 ระยะเวลาโครงการที่เหมาะสมที่สุดและต้นทุนโครงการที่เหมาะสมที่สุดคือกรณี Project Cost 100% เพราะเมื่อเปรียบเทียบค่าใช้จ่ายโครงการมีค่าน้อยกว่ากรณีอื่น ๆ แต่ระยะเวลาโครงการจะมากกว่ากรณีอื่น ๆ แต่ไม่เกินวันล่าช้าที่ยอมรับได้ดังตารางที่ 12

4.2) วิเคราะห์และเปรียบเทียบหาค่าความเหมาะสมที่สุดของการหาวนหาคำตอบ (Iteration)

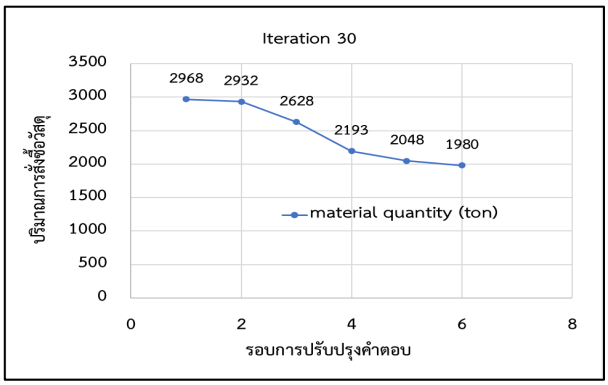
4.2.1) เปรียบเทียบการวนหาคำตอบของปริมาณการสั่งซื้อวัสดุ ดังรูปที่ 8, 9, 10 และ 11 สามารถวิเคราะห์ได้ดังนี้



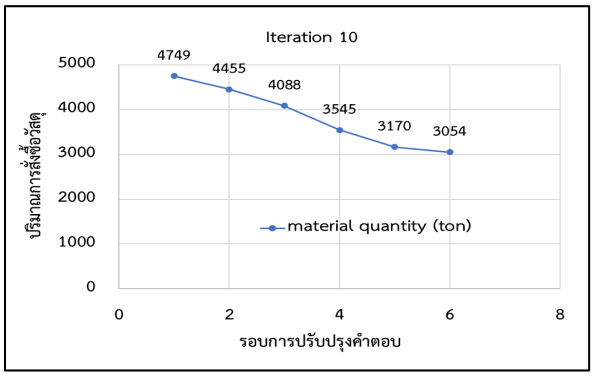
รูปที่ 8 : การวนหาคำตอบ 100 รุ่นของปริมาณการสั่งซื้อวัสดุ



รูปที่ 9 : การวนหาคำตอบ 70 รุ่นของปริมาณการสั่งซื้อวัสดุ



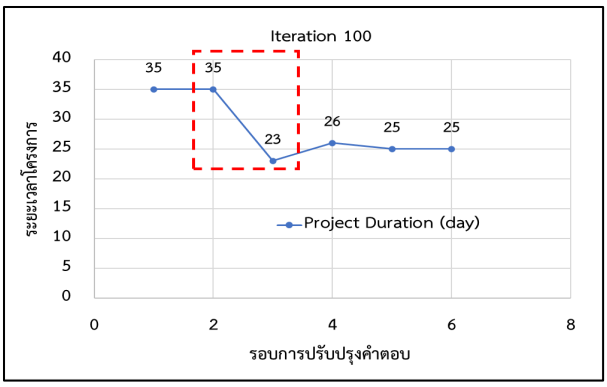
รูปที่ 10 : การวนหาคำตอบ 30 รุ่นของปริมาณการสั่งซื้อวัสดุ



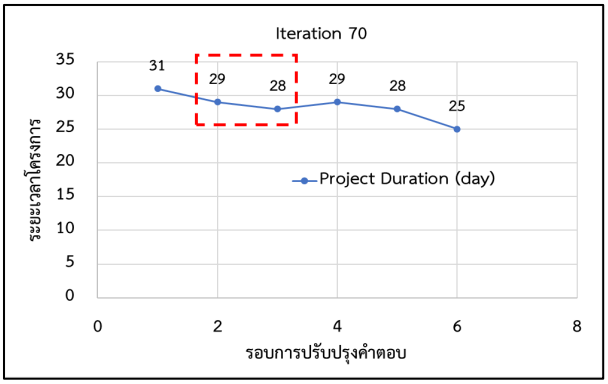
รูปที่ 11 : การวนหาคำตอบ 10 รุ่นของปริมาณการสั่งซื้อวัสดุ

รูปที่ 8, 9, 10 และ 11 พบว่า คำตอบที่ดีที่สุดคือ 100 รุ่น ยิ่งวนหาคำตอบทั้งหมดของรุ่นมากเท่าไร ยิ่งได้ปริมาณการสั่งซื้อวัสดุน้อยลงเท่านั้น โดยคาดว่าปริมาณของวัสดุจะลดลงเรื่อย ๆ จนได้ค่าใกล้เคียงคำตอบที่ดีที่สุด

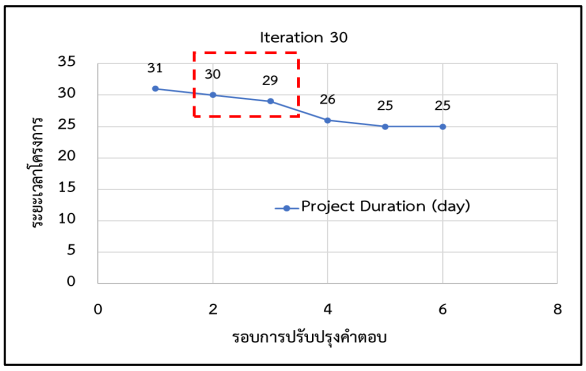
4.2.2) เปรียบเทียบการวนหาคำตอบของระยะเวลาโครงการ ดังรูปที่ 12, 13, 14 และ 15 สามารถวิเคราะห์ได้ดังนี้



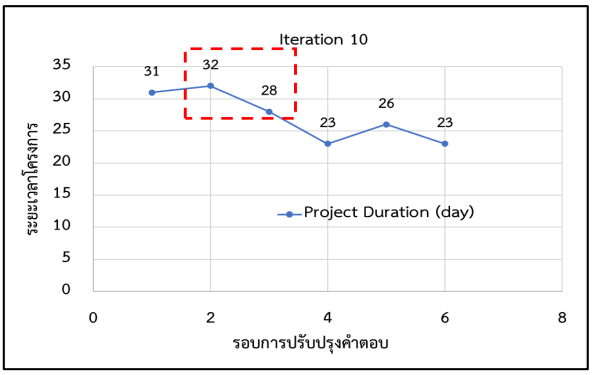
รูปที่ 12 : การวนหาคำตอบ 100 รุ่นของระยะเวลาโครงการ



รูปที่ 13 : การวนหาคำตอบ 70 รุ่นของระยะเวลาโครงการ



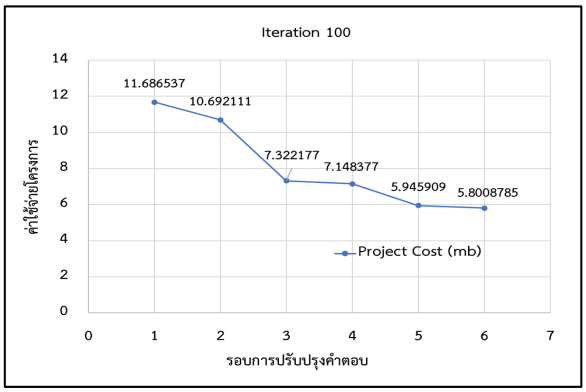
รูปที่ 14 : การวนหาคำตอบ 30 รุ่นของระยะเวลาโครงการ



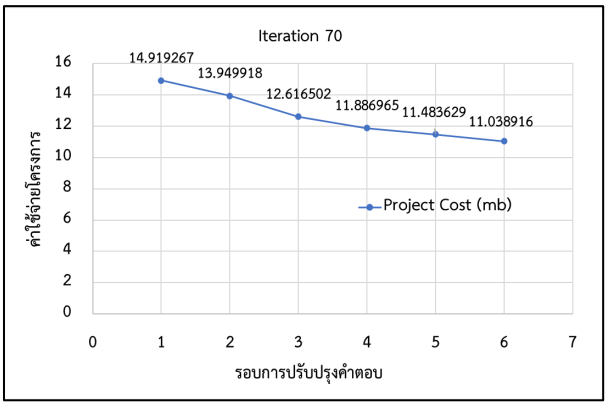
รูปที่ 15 : การวนหาคำตอบ 10 รุ่นของระยะเวลาโครงการ

จากรูปที่ 12, 13, 14 และ 15 พบว่า คำตอบที่ดีที่สุดคือ 100 รุ่น ยิ่งวนหาคำตอบทั้งหมดของรุ่นมากเท่าไร ยิ่งได้ระยะเวลาโครงการน้อยลงเท่านั้น สังเกตได้จากการปรับปรุงคำตอบครั้งที่ 2 และ 3 ของ 100 รุ่น ระยะเวลาที่ได้จะน้อยลงถึง 12 วัน เมื่อเทียบกับ 10, 30 และ 70 รุ่นระยะเวลาที่ได้จะน้อยลงเฉลี่ย 2-5 วันซึ่งจะค่อย ๆ ลดลง

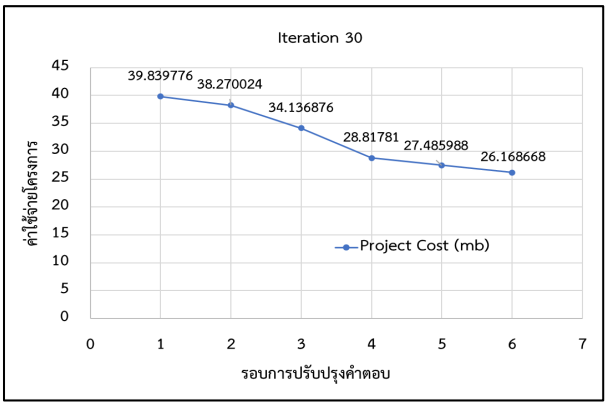
4.2.3) เปรียบเทียบการวนหาคำตอบของต้นทุนค่าใช้จ่ายรวม ดังรูปที่ 16, 17, 18 และ 19 สามารถวิเคราะห์ได้ดังนี้



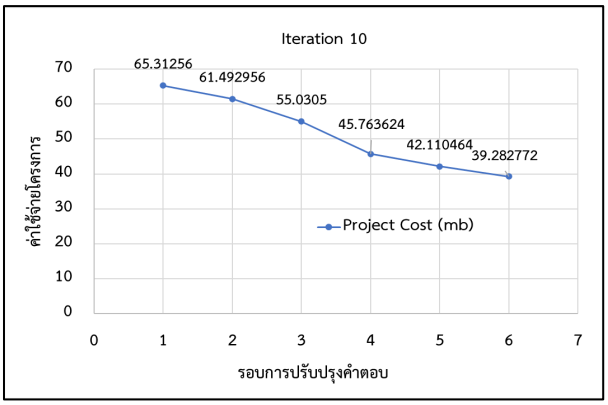
รูปที่ 16 : การวนหาคำตอบ 100 รุ่นของต้นทุนค่าใช้จ่ายรวม



รูปที่ 17 : การวนหาคำตอบ 70 รุ่นของต้นทุนค่าใช้จ่ายรวม



รูปที่ 18 : การวนหาคำตอบ 30 รุ่นของต้นทุนค่าใช้จ่ายรวม



รูปที่ 19 : การวนหาคำตอบ 10 รุ่นของต้นทุนค่าใช้จ่ายรวม

จากรูปที่ 16, 17, 18 และ 19 พบว่า คำตอบที่ดีที่สุดคือ 100 รุ่น ยิ่งวนหาคำตอบทั้งหมดของรุ่นมากเท่าไร ยิ่งได้ค่าใช้จ่ายโครงการน้อยลงเท่านั้น ซึ่งจะลดลงตามปริมาณการสั่งซื้อรูปที่ 8, 9, 10 และ 11

4.3) วิเคราะห์และเปรียบเทียบหาผลลัพธ์คำตอบที่ดีที่สุด

เป็นการเปรียบเทียบปริมาณการสั่งซื้อ ค่าสั่งซื้อ ค่าใช้จ่ายในการเก็บรักษา ดังแสดงในตารางที่ 13

ตารางที่ 13 : การเปรียบเทียบระหว่าง GA และสมการ EOQ

Inventory Cost	GA	EOQ
Material quantity (ton)	306	542.91
Ordering costs (baht)	11,016,000	19,544,607.44
Holding cost (baht)	11,827.50	16,287.17
Total cost (baht)	11,027,827.50	19,560,894.61

จากข้อมูลตารางที่ 13 เป็นคำตอบที่ดีที่สุดที่ได้จากการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดซึ่งใช้วิธีเชิงพันธุกรรมมาเปรียบเทียบกับสมการในการสั่งซื้อที่ประหยัดกรณียอมให้มีการขาดแคลนวัสดุตามแผนงาน และกรณีการสั่งซื้อแบบมีส่วนลดดังสมการที่ (1), (2) และ (3) พบว่า คำตอบที่ดีที่สุดที่ได้จาก GA ได้คำตอบที่ดีกว่าคำตอบที่ได้จากสมการ EOQ เนื่องจากการสั่งวัสดุของ GA เป็นการสั่งมาแล้วใช้เลย ซึ่งปริมาณนั้นสั่งตามความต้องการของกิจกรรมนั้น ๆ โดยไม่ได้สต็อกไว้ แต่การสั่งวัสดุของสมการ EOQ เป็นการสั่งปริมาณซ้ำ ๆ ในปริมาณที่เท่า ๆ กัน และเก็บสต็อกไว้จึงส่งผลให้เกิดค่าใช้จ่ายที่มากกว่า

5) สรุปผล

คำตอบที่ดีที่สุดที่ได้จากการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดและวิธีเชิงพันธุกรรม คือคำตอบที่ได้จากกรณีไม่ยอมให้มีการขาดแคลนวัสดุตามแผนงาน Project Cost 100% ซึ่งเกิดจากการวนหาคำตอบ 100 รุ่น เพราะยิ่งวนหาคำตอบได้มากเท่าไร ก็ยิ่งได้คำตอบที่ดีที่สุดมากเท่านั้น และได้คำตอบที่ดีกว่าสมการปริมาณการสั่งซื้อที่ประหยัด เนื่องจากวิธีการพัฒนาแบบจำลองการสั่งซื้อที่ประหยัดและวิธีเชิงพันธุกรรมคำนึงถึงค่าใช้จ่ายที่เกี่ยวกับงานก่อสร้างทั้งหมด ได้แก่ ค่าสั่งซื้อ ค่าใช้จ่ายในการจัดเก็บ ค่าขนส่ง ค่าขนย้าย ค่าเสียโอกาส ค่าปรับ ค่าใช้จ่ายทางอ้อม และค่าใช้จ่ายที่เกิดจากการว่างงาน อีกทั้งยังคำนึงถึงระยะเวลาโครงการและยังคำนึงถึงปริมาณการสั่งซื้อวัสดุในแต่ละครั้ง ซึ่งขึ้นอยู่กับกิจกรรมงานนั้น ๆ โดยปริมาณที่สั่งจะไม่เท่ากันทุกรอบ และเป็นการสั่งมาแล้วใช้เลยโดยไม่ได้สต็อกไว้

6) ข้อเสนอแนะ

การศึกษาครั้งนี้ได้ศึกษาเฉพาะวัสดุประเภทซื้อขาดเท่านั้น แต่ไม่ได้มีศึกษาวัสดุที่มีทั้งการเช่า และซื้อขาย เช่น ไม้แบบ ดังนั้นอาจมีผลลัพธ์ที่แตกต่างกัน แต่สามารถนำไปประยุกต์ใช้ได้ และยังสามารถนำไปใช้ได้ทั้งในงานโครงการขนาดเล็กจนไปถึงขนาดใหญ่ เพื่อเป็นเครื่องมือในการเพิ่มความรวดเร็วในการทำงานและลดระยะเวลาการก่อสร้าง และยังสามารถวางแผนกำหนดสั่งวัสดุก่อสร้างเป็นจำนวนความถี่ที่เท่าไรก็ได้ตามความต้องการของผู้ใช้ งานศึกษาครั้งนี้ไม่ได้กำหนดเวลาการสั่งซื้อวัสดุล่วงหน้า ซึ่งอาจส่งผลให้เกิดความเสี่ยงในเรื่องของราคาต่อหน่วยที่เปลี่ยนแปลงตลอดเวลา การผลิตวัสดุอาจไม่ทันตามความต้องการ การขนส่ง และอาจเกิดการว่างงานของคนงาน และการกำหนดการวนหาคำตอบ 100 รอบอาจไม่ได้การันตีว่าเป็นคำตอบที่ดีที่สุด การเพิ่มการวนหาคำตอบจะทำให้ได้คำตอบที่ใกล้เคียงคำตอบที่ดีที่สุดมากกว่าเดิม

REFERENCES

[1] T. Chupakdee. “Construction Project Management (CPM).” (in Thai), WORDPRESS.com. https://cmteerawut.wordpress.com (accessed Jan. 1, 2023).

[2] N. Charoenchewakul, S. Kantabutra, and P. Sirisawat, “The comparative study of logistics cost in supply chain of bottled water factory and bottled mineral water factory,” (in Thai), *J. Manage. Sci. Chiangrai Rajabhat Univ.*, vol. 10, no. 1, pp. 66–92, 2015.

[3] C. Pinyopummin, D. Rinchumphu, and K. Arunothayanan, “Inventory management in small construction business,” (in Thai), in *Proc. 25th Nat. Conv. Civil Eng.*, Jul. 2020, pp. CEM29-1–CEM29-7. [Online]. Available: https://conference.thaince.org/index.php/ncce25/article/view/669/148

[4] D. R. Anderson, D. J. Sweeney, and T. A. Williams, *An Introduction to Management Science Quantitative Approaches to Decision Making,* 10th ed. Cincinnati, OH, USA: South-Western College Pub., 2003.

[5] P. V. H. Son, N. H. C. Duy, and P. T. Dat, “Optimization of construction material cost through logistics planning model of dragonfly algorithm — Particle swarm optimization,” *KSCE J. Civil Eng.*, vol. 25, pp. 2350–2359, Apr. 2021.

[6] Newcastle Engineering Design Centre. “Roulette wheel selection.” EDC.NCL.ac.uk. http://www.edc.ncl.ac.uk/highlight/rhjanuary2007g02.php (accessed Mar. 16, 2023).

[7] H. Said and K. El-Rayes, “Optimizing material procurement and storage on construction sites,” *J. Constr. Eng. Manage.*, vol. 137, no. 6, pp. 421–431, 2011.

[8] M. Georgy and S. Y. Basily, “Using genetic algorithms in optimizing construction material delivery schedules,” *Constr. Innov.*, vol. 8, no. 1, pp. 23–45, 2008.

[9] Design and Construction Group, Office of the Basic Education Commission (OBEC). “Price list of construction materials and labor costs 2021.” (in Thai), YOTATHAI.com. https://www.yotathai.com/yotanews/cost-build-64 (accessed Dec. 31, 2022).

[10] The Comptroller General's Department. “Middle price for construction.” (in Thai), CGD.go.th. https://saraban-law.cgd.go.th/easinetimage/inetdoc?id=show_CGD.A.18981_1_BCS_1_pdf (accessed Feb. 1, 2023).

[11] ADMIN BUILK. “Construction Cost Proportion.” (in Thai), BUILK.com. https://www.builk.com/th/สัดส่วนต้นทุนก่อสร้าง_2018/ (accessed Jan. 1, 2023).

[12] P. Thanatitri. “Interest Rate 2021.” (in Thai), KRUNGSRI.com. https://www.krungsri.com/Krungsri2020/media/Banking-Rate/rates-loans/th/interest-rates-loans-01042564-th.pdf (accessed Dec. 31, 2022).

[13] M. Laotrakul. “Tower Crane Rental.” (in Thai), BTSTOWER CRANE.com. http://www.btstowercrane.com/uncategorized/tower-crane-wolff (accessed Jan. 31, 2023).

[14] B. Suttamanatwong. “Minimum Wage rate 2022.” (in Thai), MOL.go.th. https://www.mol.go.th/อัตราค่าจ้างขั้นต่ำ (accessed Dec. 31, 2022).

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

การวิเคราะห์รูปแบบการสัญจรด้วยเท้าของธุรกิจในเขตวัฒนา ในช่วงการระบาดของโรคโควิด-19

ญาณิศา นวลอนันต์[1*] ธงทิศ ฉายากุล[2]

[1,2]ภาควิชาวิศวกรรมสำรวจ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพมหานคร, ประเทศไทย*

*ผู้ประพันธ์บรรณกิจ อีเมล : eyee_yns@hotmail.com

รับต้นฉบับ : 28 กุมภาพันธ์ 2566; รับบทความฉบับแก้ไข : 16 มิถุนายน 2566; ตอบรับบทความ : 10 กรกฎาคม 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

ในปี ค.ศ. 2020–2022 ประเทศไทยมีผู้ติดเชื้อและเสียชีวิตจำนวนมากจากการแพร่ระบาดของโควิด-19 การประกาศใช้มาตรการเพื่อควบคุมสถานการณ์การแพร่ระบาดส่งผลให้หลายธุรกิจขาดทุนสะสมและปิดกิจการ จึงมีความจำเป็นอย่างยิ่งที่จะต้องศึกษาผลกระทบที่เกิดขึ้นเพื่อรับมือสถานการณ์โรคระบาดในอนาคต ข้อมูลการสัญจรด้วยเท้าที่ใช้ในงานวิจัยนี้ได้จากการติดตามตำแหน่งโทรศัพท์เคลื่อนที่แบบไม่เปิดเผยข้อมูลส่วนบุคคล โดยใช้ระบบดาวเทียมช่วยระบุตำแหน่งบนพื้นโลกเพื่อวิเคราะห์ผลกระทบธุรกิจในเขตวัฒนา จังหวัดกรุงเทพฯ โดยเป้าหมายของงานวิจัย คือ ศึกษาและเปรียบเทียบความเหมือน-ต่างของรูปแบบการสัญจรด้วยเท้าของธุรกิจแต่ละประเภท โดยจัดกลุ่มขนาดสถานที่ตามจำนวนการสัญจรด้วยเท้าเฉลี่ยต่อวันแล้วเปรียบเทียบด้วยเงื่อนไขความแตกต่างระหว่างอัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าของกลุ่มและสถานที่, วิเคราะห์ผลกระทบประกาศมาตรการจากรัฐบาลต่อปริมาณการสัญจรด้วยเท้า และวิเคราะห์การฟื้นตัวธุรกิจ โดยการเปรียบเทียบความแตกต่างปริมาณการสัญจรด้วยเท้าด้วยสถิติทดสอบ Kolmogorov-Smirnov และ Wilcoxon Signed-Rank จากงานวิจัยนี้พบว่าธุรกิจแต่ละประเภทมีรูปแบบการสัญจรด้วยเท้าที่เป็นเอกลักษณ์และสัมพันธ์กับจำนวนการสัญจรด้วยเท้าเฉลี่ยต่อวัน ผลการเปรียบเทียบสถานที่ตัวอย่างที่ผ่านเงื่อนไขต่อสถานที่ตัวอย่างทั้งหมดของธุรกิจบาร์และสถานบันเทิง 130/204, ธุรกิจศูนย์การแพทย์และโรงพยาบาล 143/204, ธุรกิจอาคารสำนักงาน 199/282, ธุรกิจร้านอาหาร 304/480 และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า 145/180 การประกาศใช้พระราชกำหนดการบริหารราชการในสถานการณ์ฉุกเฉิน (พ.ร.ก.ฉุกเฉิน), เคอร์ฟิว และห้ามนักท่องเที่ยวต่างชาติเดินทางเข้าประเทศไทย ส่งผลให้จำนวนการสัญจรด้วยเท้าลดลงและคงที่อยู่ในระดับต่ำที่สุดในช่วงระยะเวลาที่ศึกษา การสั่งปิดธุรกิจบาร์และสถานบันเทิงครั้งที่ 2 ใช้เวลาถึงสองเดือนเพื่อควบคุมจำนวนการสัญจรด้วยเท้าให้อยู่ในระดับต่ำ โดยระหว่างที่ประกาศใช้มาตรการนี้มีบางช่วงเวลาที่จำนวนการสัญจรด้วยเท้าเพิ่มขึ้นซึ่งสวนทางกับจุดประสงค์ของมาตรการ และจากสถิติทดสอบพบว่าในเดือนตุลาคม ค.ศ. 2022 ไม่มีธุรกิจประเภทใดฟื้นตัวเมื่อเทียบกับช่วงก่อนเกิดการแพร่ระบาดโควิด-19

คำสำคัญ : ธุรกิจเขตวัฒนา การฟื้นตัวธุรกิจจากการระบาดโควิด-19 มาตรการโควิด-19 รูปแบบการสัญจรด้วยเท้า ผลกระทบจากโควิด-19

Foot Traffic Pattern Analysis of Vadhana District in the Spreading of COVID-19 Periods

Yanisa Nualanant[1*] Thongthit Chayakula[2]

[1*,2]*Department of Survey Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand*

*Corresponding Author. E-mail address: eyee_yns@hotmail.com

Received: 28 February 2023; Revised: 16 June 2023; Accepted: 10 July 2023
Published online: 27 December 2023

Abstract

Between 2020 and 2022, several businesses in Thailand had deficits and wound up due to the government declarations for controlling the epidemic situation with large number of COVID-19 casualties. In this way, it is essential to analyze the impact on such businesses and prepare for the outbreak that may occur in the future. This research aimed to study the effects of businesses in Vadhana district of Bangkok, Thailand, using foot traffic data obtained by anonymized cell phone GPS location. The goals of this research are to study and differentiate foot traffic patterns of business types by percent difference condition between average foot traffic ratio of the group and the place. Moreover, foot traffic can evaluate the impact of government declaration and business recovery during the sample periods using statistical methods, Kolmogorov-Smirnov and Wilcoxon Signed-Rank, which performed to analyze foot traffic amounts. The study indicates that foot traffic pattern of each business type related to the number of foot traffic average per day. The results of the number of sample places which passed condition/total sample places are bars and night club 130/204, medical center and hospital 143/204, office building 199/282, restaurant 304/480, and retail shop and department store 145/180. Specifically, the government declarations, emergency decree, curfew, and preventing foreign tourists from entering the country, caused foot traffic reduction at the lowest level of enforcement periods. The second measure of closing bars and entertainment venues resulted in foot traffic slowly decreased to low level for two months and peaked in some period of this reduction that went against declaration's objectives. Finally, foot traffic statistical analyses in October 2022 showed that all business types had not yet recovered from pre-pandemic period.

***Keywords*:** Business in Vadhana district, Business recovery from COVID-19, COVID-19 declaration, Foot traffic pattern, Impact from COVID-19

1) บทนำ

ข้อมูลการสัญจรด้วยเท้า (foot traffic) มีบทบาทสำคัญต่อการติดตามจำนวนและตำแหน่งของผู้เข้าใช้บริการในบริเวณที่สนใจ เช่น ร้านค้า, ห้างสรรพสินค้า และสถานที่ต่าง ๆ เพื่อวิเคราะห์พฤติกรรมของมนุษย์ โดยข้อมูลนี้ถูกประยุกต์ใช้อย่างแพร่หลายในงานที่เกี่ยวข้องกับการศึกษารูปแบบการเดินของผู้ใช้บริการภายในอาคาร (indoor) ผ่านการติดตามตำแหน่งด้วยอุปกรณ์ตรวจจับสัญญาณซึ่งไม่สามารถบันทึกข้อมูลได้หากออกนอกขอบเขตการตรวจจับ [1] การวางแผนดำเนินธุรกิจด้วยโมเดลธุรกิจอัจฉริยะจากแบบจำลองการสัญจรด้วยเท้าเพื่อศึกษารูปแบบการเดินภายในอาคาร [2] สามารถศึกษาลักษณะการเดินเลือกซื้อสินค้าแบบทวนและตามเข็มนาฬิกาภายในร้านสะดวกซื้อร่วมกับเซ็นเซอร์ตรวจจับลักษณะการมองของสายตาเพื่อให้เข้าใจพฤติกรรมของลูกค้า [3] นอกจากธุรกิจร้านค้าปลีกและห้างสรรพสินค้าแล้วข้อมูลการสัญจรด้วยเท้ายังถูกใช้ในธุรกิจโรงพยาบาล เช่น ติดตั้งเครื่องนับจำนวนบุคลากรทางการแพทย์ที่เดินผ่านประตูห้องผ่าตัดวิเคราะห์ร่วมกับเครื่องตรวจจับแบคทีเรียในอากาศเพื่อออกแบบจำนวนคนในทีมลดโอกาสที่ผู้ป่วยติดเชื้อขณะผ่าตัด [4] การติดตามตำแหน่งเพื่อเก็บข้อมูลการสัญจรด้วยเท้าภายนอกอาคาร (outdoor) มีขอบเขตพื้นที่ศึกษาบริเวณกว้างเพื่อสังเกตพฤติกรรมของคนจำนวนมาก ได้มีการพัฒนาโมเดลสำหรับตรวจจับคนเดินบนถนนและแยกชนิดยานพาหนะแบบเรียลไทม์ด้วยอัลกอริทึม YOLOv3 ซึ่งเก็บข้อมูลผ่านกล้องวงจรปิดที่ติดตั้งรอบถนน [5] การเก็บข้อมูลการสัญจรด้วยเท้าแบบเรียลไทม์ด้วยเซ็นเซอร์ติดตั้งในพื้นที่สาธารณะเพื่อนับจำนวนคนเดินผ่านสามารถนำไปปรับแก้และประเมินความถูกต้องของโมเดลพยากรณ์ปริมาณการสัญจรด้วยเท้าที่อาจเปลี่ยนแปลงไปตามปัจจัยต่าง ๆ เช่น เวลา และสภาพอากาศ [6] อีกทั้งสามารถสร้างโมเดลพยากรณ์กิจวัตรประจำวันจากการจำแนกรูปแบบกิจกรรมของคนในเมืองด้วยอัลกอริทึม Gibbs Sampling โดยระบุตำแหน่งผู้ใช้โทรศัพท์ที่เช็คอินผ่านแอปพลิเคชัน Foursquare [7] จะเห็นได้ว่าข้อมูลการสัญจรด้วยเท้าสามารถอธิบายลักษณะการเดินได้เป็นอย่างดี เมื่อทำงานร่วมกับอุปกรณ์ตรวจจับและประยุกต์ใช้การประมวลผลที่ซับซ้อนจะสามารถศึกษาและแก้ไขปัญหาต่าง ๆ ได้ทั้งระดับมหภาค (macro level) และระดับจุลภาค (micro level)

เมื่อเทคโนโลยีถูกพัฒนาซับซ้อนมากขึ้น การรังวัดตำแหน่งการสัญจรด้วยเท้าติดตามผ่านโทรศัพท์เคลื่อนที่แบบนิรนามร่วมกับระบบระบุตำแหน่งบนพื้นโลกด้วยดาวเทียม (Global Navigation Satellite System: GNSS) การเก็บข้อมูลภายนอกอาคารจึงมีลักษณะเป็นฐานข้อมูลขนาดใหญ่และมีความถูกต้องเชิงตำแหน่งสูง [8] สามารถสร้างโมเดลพยากรณ์ปริมาณการสัญจรด้วยเท้าในอนาคตด้วยวิธีการทาง Machine Learning เพื่อศึกษาการขยายของตัวเมือง [9] ประยุกต์ใช้กับการจำแนกพฤติกรรมของผู้ใช้บริการโทรคมนาคมร่วมกับระบบแจ้งหนี้และเรียกเก็บเงิน และระบบลูกค้าสัมพันธ์ เพื่อหาพฤติกรรมต้องสงสัยที่อาจก่อความเสียหายต่อองค์กรนำไปสู่การวางแผนกลยุทธ์ทางการตลาดและระบบตรวจจับการฉ้อโกงและอาชญากร [10] ถึงแม้การระบุตำแหน่งด้วยดาวเทียมจะให้พิกัดที่มีความแม่นยำสูงแต่การรับสัญญาณในพื้นที่ไม่เปิดโล่งหรือภายในอาคารอาจทำให้พิกัดเกิดความคลาดเคลื่อน จึงอาศัยอุปกรณ์เสริมช่วยเก็บข้อมูลการสัญจรด้วยเท้า เช่น อุปกรณ์ Wi–Fi [11] เพื่อให้สามารถบอกตำแหน่งของผู้ใช้บริการในอาคารได้ถูกต้องมากขึ้น

ข้อมูลการสัญจรด้วยเท้าถูกนำมาวิเคราะห์รูปแบบการสัญจรด้วยเท้าที่เปลี่ยนแปลงไปเนื่องจากผลกระทบโรคระบาดโควิด-19 สามารถคำนวณความน่าจะเป็นจากความเสี่ยงติดเชื้อด้วยแบบจำลองรูปแบบการเดินภายในโถงอาคารเรียนเพื่อออกแบบเส้นทางการเดินในสถานการณ์ที่มีผู้คนแออัดเพื่อลดโอกาสติดเชื้อ [12] เมื่อจำนวนผู้ป่วยโควิด-19 รายวันเพิ่มสูงขึ้นรัฐบาลหลายประเทศประกาศใช้มาตรการเว้นระยะห่างทางสังคม (social distancing) เพื่อควบคุมปริมาณผู้ติดเชื้อซึ่งส่งผลกระทบต่อธุรกิจหลายประเภท จากการศึกษาพบว่าข้อมูลการสัญจรด้วยเท้าสามารถใช้วิเคราะห์ผลกระทบต่อเศรษฐกิจระดับมหภาคภายในประเทศ [13] กรณีศึกษาประเทศแคนาดา ข้อมูลการสัญจรด้วยเท้าทำให้ทราบว่าประชาชนทั่วไปไม่สามารถเข้าถึงเครื่องกระตุกหัวใจแบบอัตโนมัติ (AEDs) ที่ติดตั้งตามจุดต่าง ๆ ในพื้นที่สาธารณะ เนื่องจากรัฐบาลประกาศใช้มาตรการสั่งปิดสถานที่สาธารณะซึ่งถือเป็นปัจจัยหนึ่งที่ทำให้ผู้ป่วยฉุกเฉินเสียชีวิต [14] อีกทั้งสามารถศึกษาพฤติกรรมของผู้ใช้บริการธุรกิจร้านอาหารที่เปลี่ยนแปลงไปเนื่องจากการประกาศใช้มาตรการขอความร่วมมือให้ประชาชนอยู่ภายในบ้านเพื่อลดความเสี่ยงติดเชื้อโควิด-19 ซึ่งเป็นช่วงที่เกิดการแพร่ระบาดอย่างรุนแรงในประเทศสหรัฐอเมริกา ทำให้ทราบพฤติกรรมประชาชนที่พักอาศัยในย่านรายได้ต่ำการเดินทางจะลดลงช่วงล็อกดาวน์รอบแรก ส่วนประชาชนที่พักอาศัยในย่านรายได้สูงจะใช้บริการขนส่งอาหารมากขึ้นจึงรับประทานอาหารในร้านน้อยลง [15]

งานวิจัยนี้จะวิเคราะห์การสัญจรด้วยเท้าของธุรกิจ 5 ประเภทในเขตวัฒนา กรุงเทพมหานคร ได้แก่ ธุรกิจบาร์และสถานบันเทิง, ธุรกิจอาคารสำนักงาน, ธุรกิจศูนย์การแพทย์และโรงพยาบาล,

ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า โดยการวิจัยประกอบด้วย 3 ประเด็น ดังนี้ 1) การศึกษาและเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้าของธุรกิจแต่ละประเภท 2) วิเคราะห์ผลกระทบต่อปริมาณการสัญจรด้วยเท้าจากประกาศมาตรการที่เกี่ยวข้องกับโควิด-19 และ 3) วิเคราะห์การฟื้นตัวของธุรกิจ

2) ทฤษฎีที่เกี่ยวข้อง

2.1) สถิติทดสอบ Kolmogorov-Smirnov

การทดสอบลักษณะการแจกแจงข้อมูล สามารถคำนวณได้จากฟังก์ชันแจกแจงความถี่สะสมของกลุ่มทดสอบ $(F_n x)$ ดังนี้

$$F_n(x) = \frac{number\ of\ (elements\ in\ the\ sample \le x)}{n} \quad (1)$$

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(X_i) \quad (2)$$

$$D_n = sup_x|F_n(x) - F(x)| \quad (3)$$

โดย X_i คือ ตัวอย่างสุ่มจากประชากร, x คือ ข้อมูลทดสอบ, n คือ จำนวนประชากร, $F(x)$ คือ ความถี่สะสมสัมพัทธ์ที่สังเกตได้, $F_n(x)$ คือ ความถี่สะสมสัมพัทธ์ที่คาดหวัง, sup_x คือ ค่าสัมบูรณ์ระยะห่างที่มากที่สุดระหว่างความถี่สะสมสัมพัทธ์ที่คาดหวังและความถี่สะสมสัมพัทธ์ที่สังเกตได้ และ D คือ ค่าที่นำไปเปรียบเทียบกับค่าวิกฤต [16], [17] โดยในงานวิจัยนี้จะใช้เพื่อวิเคราะห์การฟื้นตัวธุรกิจ และทดสอบที่ระดับความเชื่อมั่น 95% มีสมมติฐานหลัก (H_0) คือ ข้อมูลแจกแจงแบบปกติ

2.2) สถิติทดสอบ Wilcoxon Signed-Rank

ทดสอบแบบไม่อิงพารามิเตอร์สำหรับข้อมูลที่มีการแจกแจงแบบไม่ปกติ เป็นการเปรียบเทียบค่ากลางจากผลต่างระหว่างข้อมูล 2 กลุ่ม แล้วจัดอันดับค่าสัมบูรณ์ของผลต่างซึ่งในกรณีที่อันดับซ้ำกันจะใช้ค่าเฉลี่ยอันดับของผลต่างที่ซ้ำทั้งหมด จากนั้นเปรียบเทียบผลรวมอันดับของผลต่างบวกและลบนำค่าที่มากกว่าเทียบกับค่าวิกฤต [18] ถ้าหากผลต่างระหว่าง 2 กลุ่ม มีการแจกแจงแบบไม่ปกติจะส่งผลให้อัตราความคลาดเคลื่อนประเภทที่ 1 (Type 1 error rate) เพิ่มขึ้นตามขนาดของกลุ่มตัวอย่างและดีกรีสมมาตร [19] สำหรับงานวิจัยนี้จะใช้เพื่อวิเคราะห์การฟื้นตัวธุรกิจ และทดสอบที่ระดับความเชื่อมั่น 95% มีสมมติฐานหลัก (H_0) คือ สถานการณ์แพร่ระบาดโควิด-19 ไม่ส่งผลให้ค่ามัธยฐานของข้อมูลทดสอบทั้ง 2 ช่วงเวลาแตกต่างกัน

3) วิธีดำเนินการวิจัย

3.1) พื้นที่ศึกษา

ข้อมูลการสัญจรด้วยเท้าบริเวณเขตวัฒนา กรุงเทพมหานคร ประเทศไทย (รูปที่ 1)

รูปที่ 1 : พื้นที่ศึกษา

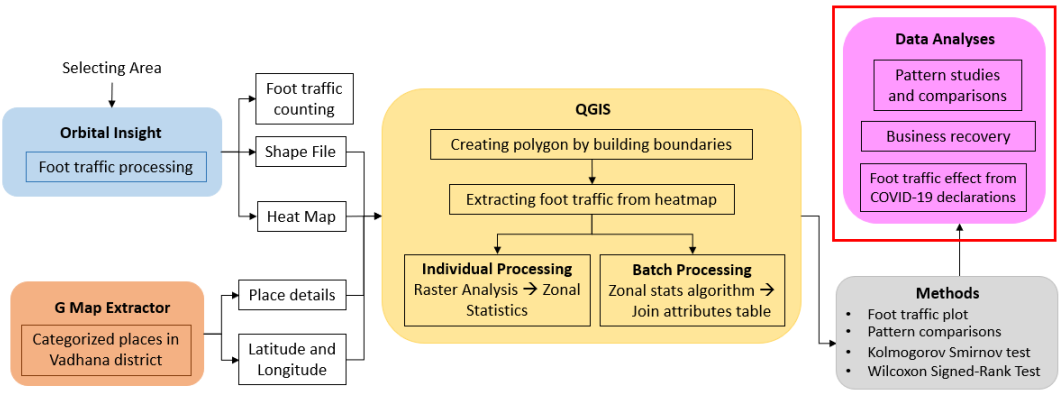
3.2) ข้อมูลที่ใช้ในงานวิจัย

ข้อมูลการสัญจรด้วยเท้าในงานวิจัยนี้แบ่งเป็น 3 ส่วน ประกอบด้วย 1) การศึกษาและเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้า ได้แก่ ข้อมูลที่จัดเก็บด้วยความถี่รายวัน ตั้งแต่วันที่ 1 เมษายน – 31 ธันวาคม ค.ศ. 2019 2) ศึกษาผลกระทบการสัญจรด้วยเท้าจากมาตรการที่เกี่ยวข้องกับโควิด-19 ได้แก่ ข้อมูลที่จัดเก็บด้วยความถี่รายวัน ตั้งแต่วันที่ 1 มีนาคม ค.ศ. 2020 – 31 มีนาคม ค.ศ. 2021 3) วิเคราะห์การฟื้นตัวธุรกิจ ได้แก่ ข้อมูลที่จัดเก็บด้วยความถี่รายชั่วโมงเดือนเมษายน ค.ศ. 2019, ค.ศ. 2020, ค.ศ. 2021 และตุลาคม ค.ศ. 2022

3.3) วิธีการเก็บข้อมูล

3.3.1) การรวบรวมข้อมูลการสัญจรด้วยเท้า เขตวัฒนามีความหลากหลายทางธุรกิจสูงและเป็นแหล่งท่องเที่ยวที่นิยมของชาวต่างชาติ ประเทศไทยเกิดการแพร่ระบาดโควิด-19 ตั้งแต่ต้น ค.ศ. 2020 จึงเลือกช่วงเวลาที่เหมาะสมและสอดคล้องกับจำนวนผู้ติดเชื้อใหม่รายวันในประเทศไทย คือ 1 เมษายน ค.ศ. 2019 – 31 มีนาคม ค.ศ. 2021 และตุลาคม ค.ศ. 2022 ประมวลผลบนแพลตฟอร์ม Orbital Insight ด้วย Foot traffic algorithm เพื่อนับจำนวนโทรศัพท์เคลื่อนที่ภายในพื้นที่ที่สนใจ (Area of Interest: AOI) [20] เป็นระยะเวลา 0–60 นาที

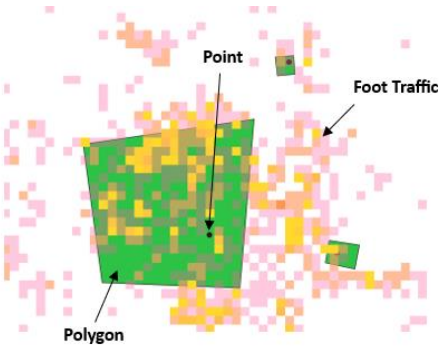
3.3.2) การรวบรวมรายละเอียดสถานที่ในเขตวัฒนา สำรวจสถานที่ในเขตวัฒนาด้วย Google Map ใช้คำค้นหาที่สอดคล้องกับธุรกิจประเภทนั้น ๆ เพื่อให้ผลลัพธ์ใกล้เคียงกับความเป็นจริงมากที่สุด จากนั้นใช้แพลตฟอร์ม G Map Extractor รวบรวมรายละเอียดของแต่ละสถานที่



รูปที่ 2 : ขั้นตอนการดำเนินงาน

3.4) การวิเคราะห์ข้อมูล

3.4.1) ประมวลผลด้วยโปรแกรม QGIS พล็อตพิกัดสถานที่ (point) จากหัวข้อ 3.2.2 สร้างโพลิกอน (polygon) รอบสิ่งปลูกสร้างที่ปรากฏบนแผนที่เพื่อกำหนดบริเวณที่ต้องการสกัด (extract) ปริมาณการสัญจรด้วยเท้า โดยสร้างชั้นข้อมูล (layer) แยกสำหรับธุรกิจแต่ละประเภท จากนั้นซ้อนทับแผนที่ความร้อน (heat map) ที่ได้จาก Orbital Insight เพื่อเก็บรวบรวมปริมาณการสัญจรด้วยเท้าภายในเขตวัฒนาบนขอบเขตที่สร้างขึ้น (รูปที่ 3) แล้วสกัดการสัญจรด้วยเท้าของแต่ละสถานที่ด้วย Zonal Stats Algorithm



รูปที่ 3 : การซ้อนทับแผนที่ความร้อนบนขอบเขตสถานที่

3.4.2) เปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้า จัดกลุ่มขนาดสถานที่ด้วยค่าเฉลี่ยจำนวนการสัญจรด้วยเท้าต่อวัน จากนั้นจับคู่วันประกอบด้วย 6 คู่ ได้แก่ Mon–Tue, Tue–Wed, Wed–Thu, Thu–Fri, Fri–Sat และ Sat–Sun โดยใช้เงื่อนไขการเปรียบเทียบเพื่อทดสอบความเหมือน-ต่างระหว่างรูปแบบการสัญจรด้วยเท้าของกลุ่มและสถานที่ ซึ่งเงื่อนไขนี้ได้จากการทดสอบหาระดับความแตกต่างที่เหมาะสมด้วยสายตาระหว่างอัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าต่อวันของกลุ่มและสถานที่ โดยตรวจสอบแนวโน้มค่าเฉลี่ยการสัญจรด้วยเท้าแต่ละคู่แบ่งเป็น 2 กรณี คือ แนวโน้มเพิ่ม (+) หมายถึง ค่าเฉลี่ยการสัญจรด้วยเท้าวันถัดมามากกว่าวันก่อนหน้าซึ่งอัตราส่วนจะมีค่ามากกว่าศูนย์ และแนวโน้มลด (-) หมายถึง ค่าเฉลี่ยการสัญจรด้วยเท้าวันถัดมาน้อยกว่าวันก่อนหน้าซึ่งอัตราส่วนจะมีค่าน้อยกว่าศูนย์

<u>ขั้นตอนการพิจารณาเงื่อนไขการเปรียบเทียบ</u>

1) คำนวณอัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าของสถานที่ตัวอย่างประเภทธุรกิจบาร์และสถานบันเทิง ดังตารางที่ 1

Mon–Tue = 26.775/29.050 = 0.92 (-)

Tue–Wed = 28.436/26.775 = 1.06 (+)

Wed–Thu = 28.872/28.436 = 1.02 (+)

Thu–Fri = 41.077/28.872 = 1.42 (+)

Fri–Sat = 49.179/41.077 = 1.20 (+)

Sat–Sun = 33.974/49.179 = 0.69 (-)

ตารางที่ 1 : ค่าเฉลี่ยการสัญจรด้วยเท้าของสถานที่ตัวอย่าง

Name	Foot traffic average						
	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Sample Bar	29.050	26.775	28.436	28.872	41.077	49.179	33.974

2) เนื่องจากสถานที่ตัวอย่างมีปริมาณเฉลี่ยการสัญจรด้วยเท้าต่อวันอยู่ในช่วง 10–100 traffic/day ซึ่งจัดอยู่ในกลุ่มธุรกิจบาร์และสถานบันเทิงขนาดกลาง คำนวณอัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าจากตารางที่ 2

Mon–Tue = 18.248/16.717 = 1.09 (+)

Tue–Wed = 18.966/18.248 = 1.04 (+)

Wed–Thu = 19.735/18.966 = 1.04 (+)

Thu–Fri = 24.327/19.735 = 1.23 (+)

Fri–Sat = 26.761/24.327 = 1.10 (+)

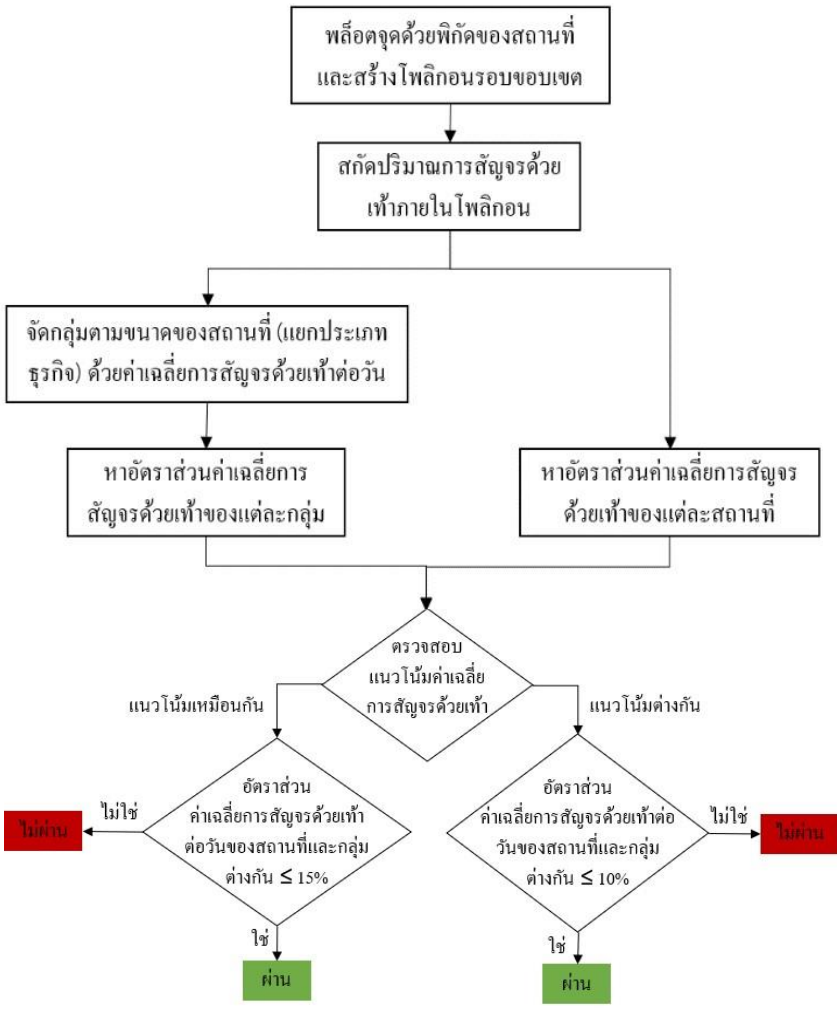
Sat–Sun = 20.468/26.761 = 0.76 (-)

3) ตรวจสอบแนวโน้มค่าเฉลี่ยการสัญจรด้วยเท้าก่อนพิจารณาตามเงื่อนไขรูปที่ 4 จากตารางที่ 3 แบ่งเป็น 2 กรณี คือ

กรณีที่ 1 : แนวโน้มค่าเฉลี่ยการสัญจรด้วยเท้าของสถานที่ตัวอย่างเหมือนกับธุรกิจบาร์และสถานบันเทิงกลุ่มขนาดกลาง แต่ละคู่จะผ่านเงื่อนไขก็ต่อเมื่ออัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าต่อวันของสถานที่และกลุ่มต่างกันไม่เกิน 15%

ตารางที่ 2 : ค่าเฉลี่ยการสัญจรด้วยเท้าธุรกิจบาร์และสถานบันเทิงขนาดกลาง

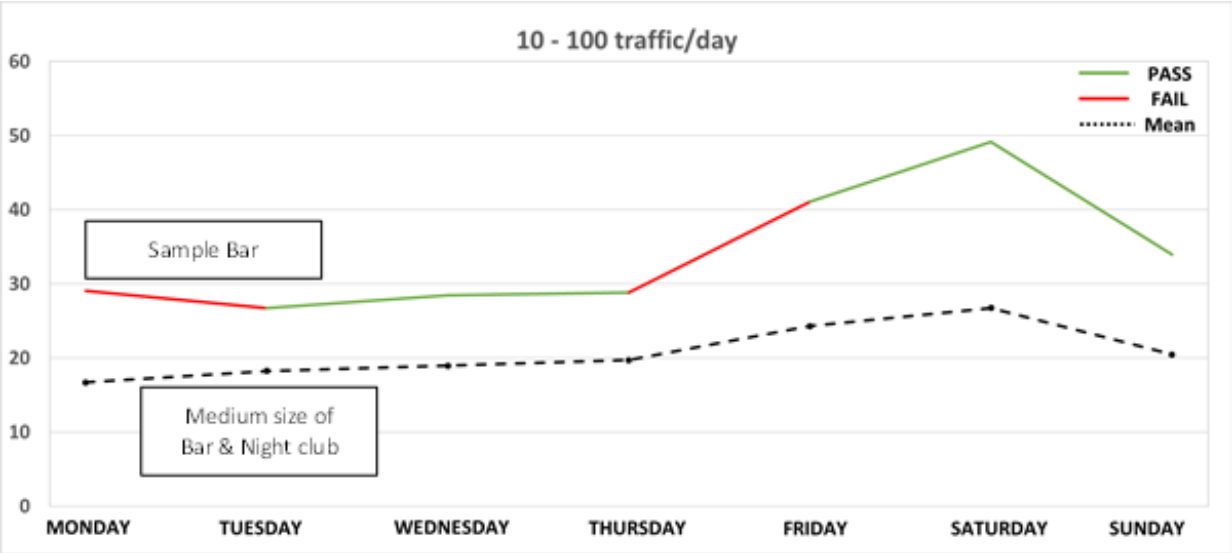
Name	Foot traffic average						
	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Medium size of Bar & Night Club average	16.717	18.248	18.966	19.735	24.327	26.761	20.468



รูปที่ 4 : เงื่อนไขการเปรียบเทียบ

ตารางที่ 3 : ผลการเปรียบเทียบรูปแบบการสัญจรด้วยเท้าตามเงื่อนไข

	Trend Check					
	Mon-Tue	Tue-Wed	Wed-Thu	Thu-Fri	Fri-Sat	Sat-Sun
Sample Bar	-	+	+	+	+	-
Medium size of Bar & Night Club average	+	+	+	+	+	-
	Trend similarity check					
ผ่านเงื่อนไข	✗	✔	✔	✔	✔	✔
	Percent difference of foot traffic ratio					
ไม่ผ่านเงื่อนไข	15.57%	2.18%	2.42%	15.42%	8.84%	9.68%



รูปที่ 5 : ผลการเปรียบเทียบรูปแบบการสัญจรด้วยเท้าตามเงื่อนไข

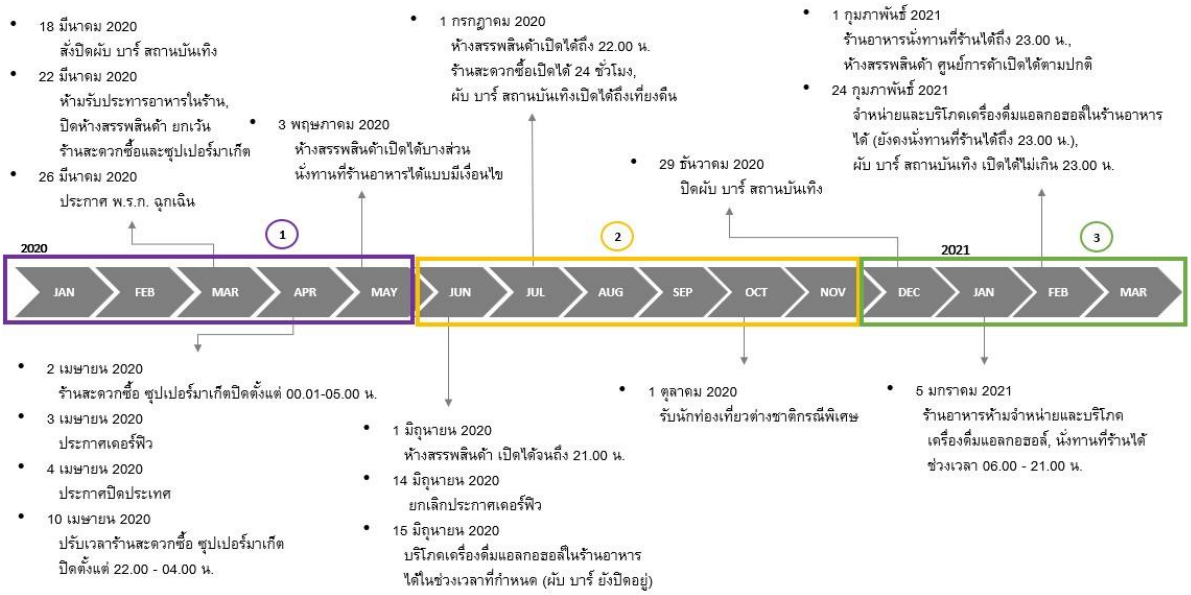
กรณีที่ 2 : แนวโน้มค่าเฉลี่ยการสัญจรด้วยเท้าของสถานที่ตัวอย่างต่างกับธุรกิจบาร์และสถานบันเทิงกลุ่มขนาดกลาง แต่ละคู่จะผ่านเงื่อนไขก็ต่อเมื่ออัตราส่วนค่าเฉลี่ยการสัญจรด้วยเท้าต่อวันของสถานที่และกลุ่มต่างกันไม่เกิน 10%

ดังนั้น จากสถานที่ตัวอย่างพบว่า Tue–Wed, Wed–Thu, Fri–Sat และ Sat–Sun ผ่านเงื่อนไขการเปรียบเทียบรูปแบบการสัญจรด้วยเท้า (รูปที่ 5)

3.4.3) วิเคราะห์ผลกระทบจากมาตรการที่เกี่ยวข้องกับโควิด-19 ช่วงเวลาประกาศมาตรการจากรัฐบาลดังรูปที่ 6 รวบรวมจากประกาศกรุงเทพมหานคร เรื่องสั่งปิดสถานที่เป็นการชั่วคราว ฉบับที่ 1-24 และราชกิจจานุเบกษา ฉบับที่ 2 และ 10 โดยธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้าจะวิเคราะห์ด้วยการสัญจรด้วยเท้าที่เก็บความถี่รายวัน ส่วนธุรกิจศูนย์การแพทย์และโรงพยาบาล และธุรกิจอาคารสำนักงานไม่ได้รับผลกระทบจากมาตรการโดยตรงจึงใช้ค่าเฉลี่ยการสัญจรด้วยเท้าต่อวันซึ่งครอบคลุมช่วงการระบาดระลอกแรก–ระลอกที่สอง [21]

3.4.4) สถิติทดสอบปริมาณการสัญจรด้วยเท้า จัดเรียงข้อมูล 2 แบบ คือ เรียงตามวันในสัปดาห์ (day of week) และเรียงตามวันที่ (date) เลือกเฉพาะช่วงเวลาที่ธุรกิจเปิดทำการตามตารางที่ 4

การประมวลผลด้วยโปรแกรม SPSS แบ่งออกเป็น 2 ส่วนคือ ฟังก์ชัน Descriptive Statics อธิบายลักษณะข้อมูลด้วยค่าทางสถิติ, ลักษณะการแจกแจงความถี่สะสม และ Boxplot เพื่อดูลักษณะการแจกแจงของผลต่าง และฟังก์ชัน 2 Related Samples สำหรับสถิติทดสอบ Wilcoxon Signed-Rank



รูปที่ 6 : ประกาศมาตรการเกี่ยวข้องกับโรคโควิด-19 จากรัฐบาล

ตารางที่ 4 : ช่วงเวลาประมวลสถิติทดสอบปริมาณการสัญจรด้วยเท้า

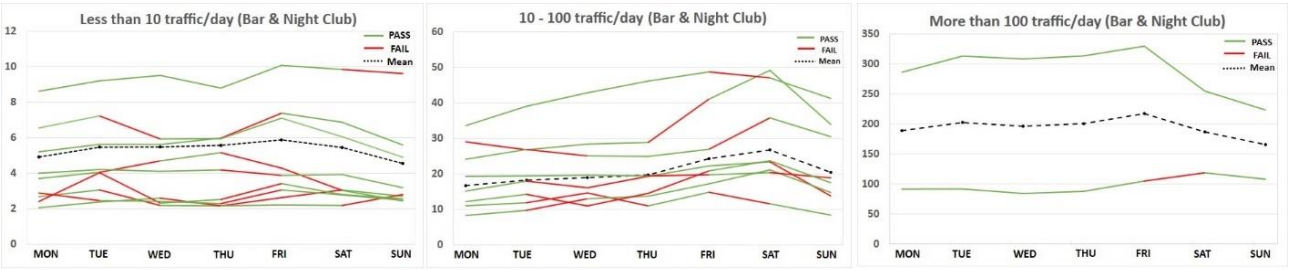
Business Type	Group by	Group	Time
Bar & Night Club	Day of Week	Week1, Week2, Week3, Week4, Week5	19:00 – 02:00
Medical Center & Hospital	Date	1st – 7th, 8th – 14th, 15th – 21st, 22nd – 30th	00:00 – 23:00
Office Building	Day of Week	Week1, Week2, Week3, Week4, Week5	08:00 – 18:00
Restaurant	Date	1st – 7th, 8th – 14th, 15th – 21st, 22nd – 30th	10:00 – 23:00
Retail Shop & Department Store	Date	1st – 7th, 8th – 14th, 15th – 21st, 22nd – 30th	08:00 – 23:00

4) ผลการวิจัย

4.1) ผลการศึกษาและเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้าของธุรกิจแต่ละประเภท

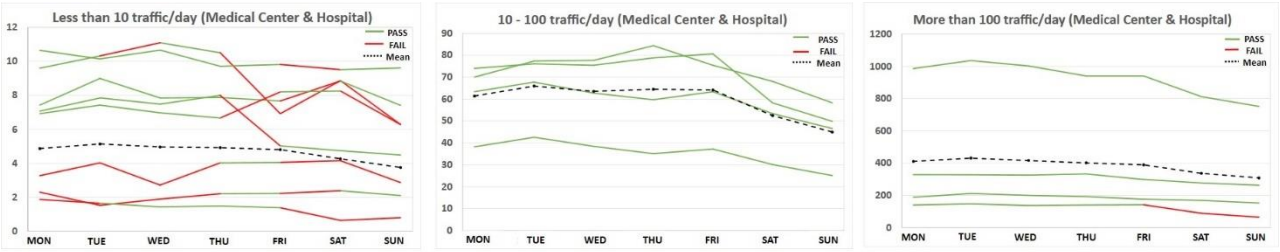
เมื่อเปรียบเทียบตามเงื่อนไขที่กำหนดในหัวข้อที่ 3.4.2 พล็อตกราฟระหว่างวันและค่าเฉลี่ยการสัญจรด้วยเท้าเพื่อแสดงรูปแบบการสัญจรด้วยเท้าของแต่ละสถานที่ในกลุ่ม (รูปที่ 7–11) ผลการเปรียบเทียบแสดงในตารางที่ 5 พบว่าธุรกิจบาร์และสถานบันเทิงมีรูปแบบการสัญจรด้วยเท้าของกลุ่มขนาดเล็กค่อนข้างหลากหลาย ผ่านเงื่อนไขการเปรียบเทียบ 73 คู่ จากทั้งหมด 120 คู่, กลุ่มขนาดกลางมีจำนวนการสัญจรด้วยเท้าคงที่วันจันทร์–พฤหัสบดี เพิ่มขึ้นวันศุกร์ สูงสุดวันเสาร์ และลดลงวันอาทิตย์ ผ่านเงื่อนไข 47 คู่ จากทั้งหมด 72 คู่ ส่วนกลุ่มขนาดใหญ่จำนวนการสัญจรด้วยเท้าคงที่วันจันทร์–พฤหัสบดี และเพิ่มขึ้นในวันศุกร์ โดยช่วงวันจันทร์–ศุกร์ สถานที่ส่วนมากมีรูปแบบคล้ายกับค่าเฉลี่ยการสัญจรด้วยเท้าของกลุ่มสอดคล้องกับผลการเปรียบเทียบที่ผ่านเงื่อนไขทุกคู่, ธุรกิจศูนย์การแพทย์และโรงพยาบาล กลุ่มขนาดเล็กผ่านเงื่อนไขการเปรียบเทียบ 97 คู่ จากทั้งหมด 156 คู่, กลุ่มขนาดกลางการสัญจรด้วยเท้าแต่ละวันค่อนข้างคงที่วันจันทร์–ศุกร์ ลดลงวันเสาร์และอาทิตย์สอดคล้องกับผลการเปรียบเทียบที่ผ่านเงื่อนไขทุกคู่ ส่วนกลุ่มขนาดใหญ่มีรูปแบบเหมือนค่าเฉลี่ยการสัญจรด้วยเท้าของกลุ่มเพราะผ่านเงื่อนไขทั้งหมด ยกเว้นวันศุกร์–เสาร์ และเสาร์–อาทิตย์ ที่ผ่านเงื่อนไข 3 คู่ จาก 4 คู่, ธุรกิจอาคารสำนักงาน พบว่าทั้งสองกลุ่มมีรูปแบบคล้ายกัน คือ วันจันทร์–ศุกร์จำนวนการสัญจรด้วยเท้าค่อนข้างคงที่ส่งผลให้ทั้ง 4 ช่วง มีคู่ที่ผ่านเงื่อนไขค่อนข้างมาก โดยกลุ่มขนาดเล็กผ่านเงื่อนไข 99 คู่จากทั้งหมด 150 คู่ และกลุ่มขนาดใหญ่ 100 คู่ จากทั้งหมด 132 คู่ แต่ในทางกลับกันการที่จำนวนการสัญจรด้วยเท้าลดลงอย่างฉับพลันวันเสาร์และวันอาทิตย์ พบว่ารูปแบบการสัญจรด้วยเท้าแตกต่างกับค่าเฉลี่ยทั้งสองกลุ่มสอดคล้องกับผลการเปรียบเทียบของทั้งสองกลุ่ม ซึ่งรูปแบบการสัญจรด้วยเท้าของธุรกิจประเภทนี้สอดคล้องกับ [21] ที่ว่าวันจันทร์–ศุกร์เป็นช่วงการทำงานมาตรฐานของพนักงาน

บริษัท, ธุรกิจร้านอาหาร กลุ่มขนาดเล็กรูปแบบการสัญจรด้วยเท้าหลากหลายและแตกต่างกับรูปแบบของค่าเฉลี่ยซึ่งสอดคล้องกับผลการเปรียบเทียบอยู่ระหว่าง 33–39 คู่ จาก 61 คู่ ส่วนกลุ่มขนาดใหญ่จำนวนการสัญจรด้วยเท้าค่อนข้างคงที่แต่ละวันและรูปแบบคล้ายกับค่าเฉลี่ยกลุ่ม สอดคล้องกับผลที่ผ่านเงื่อนไข 87 คู่ จากทั้งหมด 114 คู่ และธุรกิจร้านค้าปลีกและห้างสรรพสินค้าสถานที่ในกลุ่มขนาดเล็กมีจำนวนการสัญจรด้วยเท้าต่อวันใกล้เคียงกัน ผ่านเงื่อนไข 103 คู่ จากทั้งหมด 132 คู่, กลุ่มขนาดกลาง จำนวนการสัญจรด้วยเท้าค่อนข้างคงที่และรูปแบบคล้ายกับค่าเฉลี่ยของกลุ่มสอดคล้องกับผ่านเงื่อนไข 30 คู่ จากทั้งหมด 36 คู่ ส่วนกลุ่มขนาดใหญ่สามารถผ่านเงื่อนไขการเปรียบเทียบได้หมดทุกคู่ โดยผลการเปรียบเทียบจำนวนคู่ที่ผ่านเงื่อนไข/คู่ทั้งหมดของธุรกิจบาร์และสถานบันเทิง 130/204, ธุรกิจศูนย์การแพทย์และโรงพยาบาล 143/204, ธุรกิจอาคารสำนักงาน 199/282, ธุรกิจร้านอาหาร 304/480 และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า 145/180



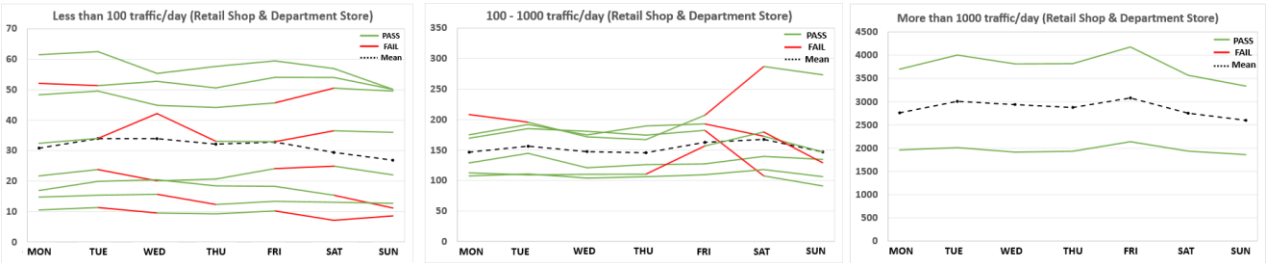
(ก) (ข) (ค)

รูปที่ 7 : รูปแบบการสัญจรด้วยเท้าธุรกิจบาร์และสถานบันเทิง (ก) กลุ่มขนาดเล็ก (ข) กลุ่มขนาดกลาง (ค) กลุ่มขนาดใหญ่



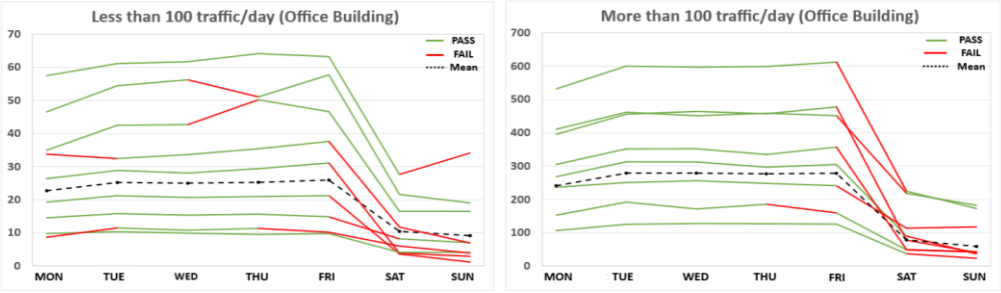
(ก) (ข) (ค)

รูปที่ 8 : รูปแบบการสัญจรด้วยเท้าธุรกิจศูนย์การแพทย์และโรงพยาบาล (ก) กลุ่มขนาดเล็ก (ข) กลุ่มขนาดกลาง (ค) กลุ่มขนาดใหญ่



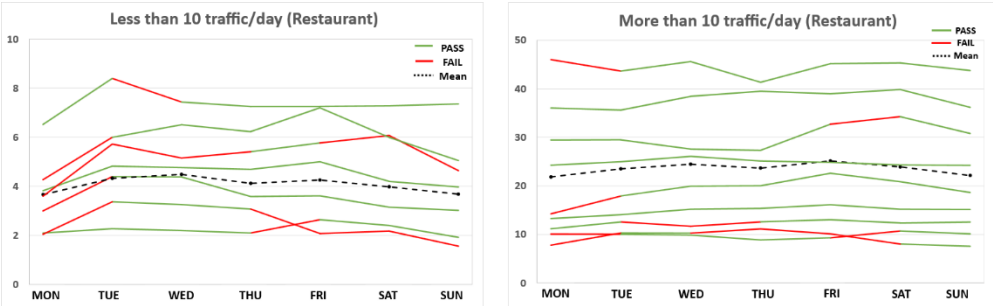
(ก) (ข) (ค)

รูปที่ 9 : รูปแบบการสัญจรด้วยเท้าธุรกิจร้านค้าปลีกและห้างสรรพสินค้า (ก) กลุ่มขนาดเล็ก (ข) กลุ่มขนาดกลาง (ค) กลุ่มขนาดใหญ่



(ก) (ข)

รูปที่ 10 : รูปแบบการสัญจรด้วยเท้าธุรกิจอาคารสำนักงาน (ก) กลุ่มขนาดเล็ก (ข) กลุ่มขนาดใหญ่



(ก) (ข)

รูปที่ 11 : รูปแบบการสัญจรด้วยเท้าธุรกิจร้านอาหาร (ก) กลุ่มขนาดเล็ก (ข) กลุ่มขนาดใหญ่

ตารางที่ 5 : ผลการเปรียบเทียบด้วยเงื่อนไข

Business Type	Group	Number of sample	Number of pass condition						Total pass	Total Sample
			Mon-Tue	Tue-Wed	Wed-Thu	Thu-Fri	Fri-Sat	Sat-Sun		
Bar & Night Club	Small (< 10 traffic/day)	20	13	12	15	9	12	12	73	120
	Medium (10 -100 traffic/day)	12	8	7	8	8	8	7	46	72
	Large (> 100 traffic/day)	2	2	2	2	2	1	2	11	12
Total pass/total sample of Bar & Night Club = 130/204										
Medical Center & Hospital	Small (< 10 traffic/day)	26	17	15	21	19	11	14	97	156
	Medium (10 -100 traffic/day)	4	4	4	4	4	4	4	24	24
	Large (> 100 traffic/day)	4	4	4	4	4	3	3	22	24
Total pass/total sample of Medical Center & Hospital = 143/204										
Office Building	Small (< 100 traffic/day)	25	17	21	21	22	6	12	99	150
	Large (> 100 traffic/day)	22	22	22	22	21	4	12	100	132
Total pass/total sample of Office Building = 199/282										
Restaurant	Small (< 10 traffic/day)	61	36	39	38	37	33	34	217	366
	Large (> 10 traffic/day)	19	12	18	15	15	13	14	87	114
Total pass/total sample of Restaurant = 304/480										
Retail Shop & Department Store	Small (< 100 traffic/day)	22	19	16	19	21	12	16	103	132
	Medium (100 - 1000 traffic/day)	6	5	6	6	5	3	5	30	36
	Large (> 1000 traffic/day)	2	2	2	2	2	2	2	12	12
Total pass/total sample of Retail Shop & Department Store = 145/180										

4.2) ผลกระทบต่อปริมาณการสัญจรด้วยเท้าจากการประกาศมาตรการที่เกี่ยวข้องกับโควิด-19

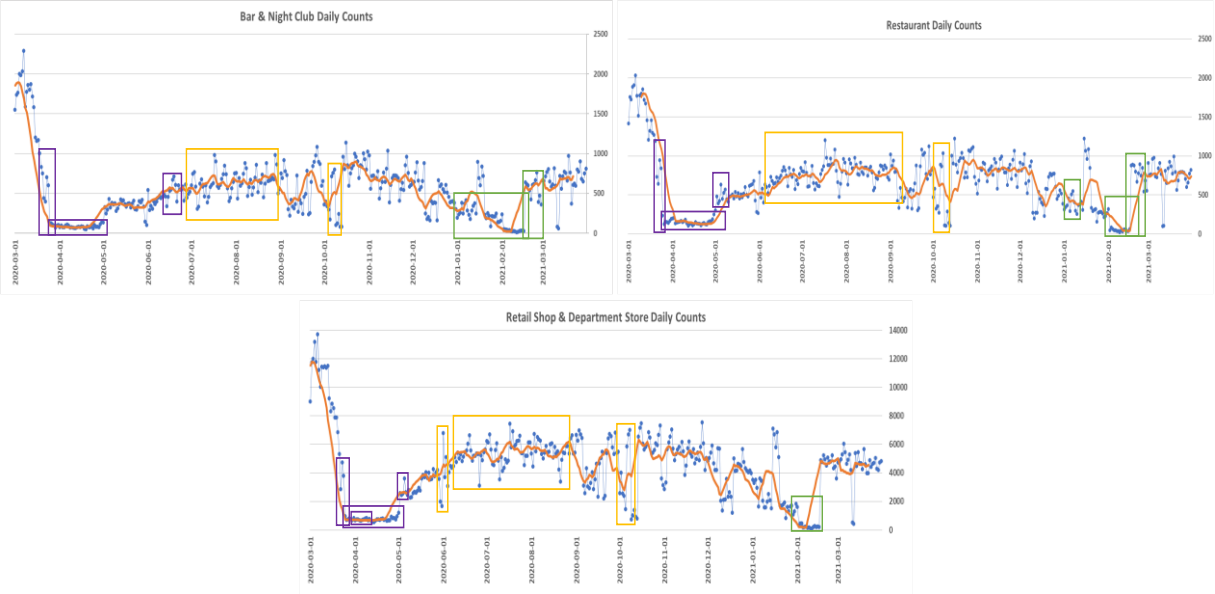
จากรูปที่ 6 และ 12 วิเคราะห์มาตรการที่เกี่ยวข้องกับธุรกิจร่วมกับข้อมูลการสัญจรด้วยเท้า กำหนดให้ traffic หมายถึงหน่วยของจำนวนการสัญจรด้วยเท้า โดยแบ่งเป็น 3 ช่วงดังนี้

1) ช่วงระบาดระลอกที่ 1 ปลายเดือนมีนาคม ค.ศ. 2020 มีการสั่งปิดบาร์และสถานบันเทิง, ห้างสรรพสินค้า และห้ามรับประทานอาหารภายในร้าน วันถัดมาพบว่าธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้าจำนวนการสัญจรด้วยเท้าลดลงจาก 1003 traffic เหลือ 830 traffic, เพิ่มขึ้นจาก 638 traffic เป็น 941 traffic และเพิ่มขึ้นจาก 2833 traffic เป็น 4708 traffic ตามลำดับ จากนั้นวันที่ 25 มีนาคม. ค.ศ. 2020 ก่อนประกาศใช้ พ.ร.ก. ฉุกเฉิน 1 วันจำนวนการสัญจรด้วยเท้าของทั้ง 3 ธุรกิจลดลงอย่างฉับพลันเหลือ 127 traffic, 232 traffic และ 958 traffic ตามลำดับ หลังประกาศใช้ พ.ร.ก. ฉุกเฉิน, เคอร์ฟิว และห้ามนักท่องเที่ยวต่างชาติเข้าประเทศไทย จำนวนการสัญจรด้วยเท้าอยู่ในช่วง 46–109 traffic, 102–239 traffic และ540–802 traffic ตามลำดับ เป็นระยะเวลาประมาณ 1 เดือน โดยระหว่างที่ประกาศใช้ทั้ง 3 มาตรการนี้ วันที่ 2 และ 10 เมษายน. ค.ศ. 2020 มีการกำหนดช่วงเวลาปิดร้านสะดวกซื้อและซุปเปอร์มาเก็ตตั้งแต่ 00.01–05.00 น. และ 22.00–04.00 น. ตามลำดับ พบว่าธุรกิจร้านค้าปลีกและห้างสรรพสินค้าหลังประกาศจำนวนการสัญจรด้วยเท้าคงที่ จากนั้น 3 พฤษภาคม ค.ศ. 2020 อนุญาตให้นั่งรับประทานในร้านได้แบบมีเงื่อนไขแต่ยังคงห้ามจำหน่ายและบริโภคเครื่องดื่มแอลกอฮอล์ภายในร้าน และห้างสรรพสินค้าสามารถเปิดทำการได้บางส่วน 2 วันถัดมาพบว่าธุรกิจร้านอาหารและธุรกิจร้านค้าปลีกและห้างสรรพสินค้า จำนวนการสัญจรด้วยเท้าเพิ่มขึ้นจาก 390 traffic เป็น 627 traffic และเพิ่มขึ้นจาก 2448 traffic เป็น 3598 traffic ตามลำดับ ในส่วนค่าเฉลี่ยต่อวัน (ตารางที่ 6 และรูปที่ 13) เป็นช่วงที่ค่าเฉลี่ยสูงที่สุดยกเว้นธุรกิจศูนย์การแพทย์และโรงพยาบาลมีค่าเฉลี่ยต่ำกว่าช่วงจำนวนผู้ติดเชื้อต่อวันคงที่อยู่ 68.58 traffic

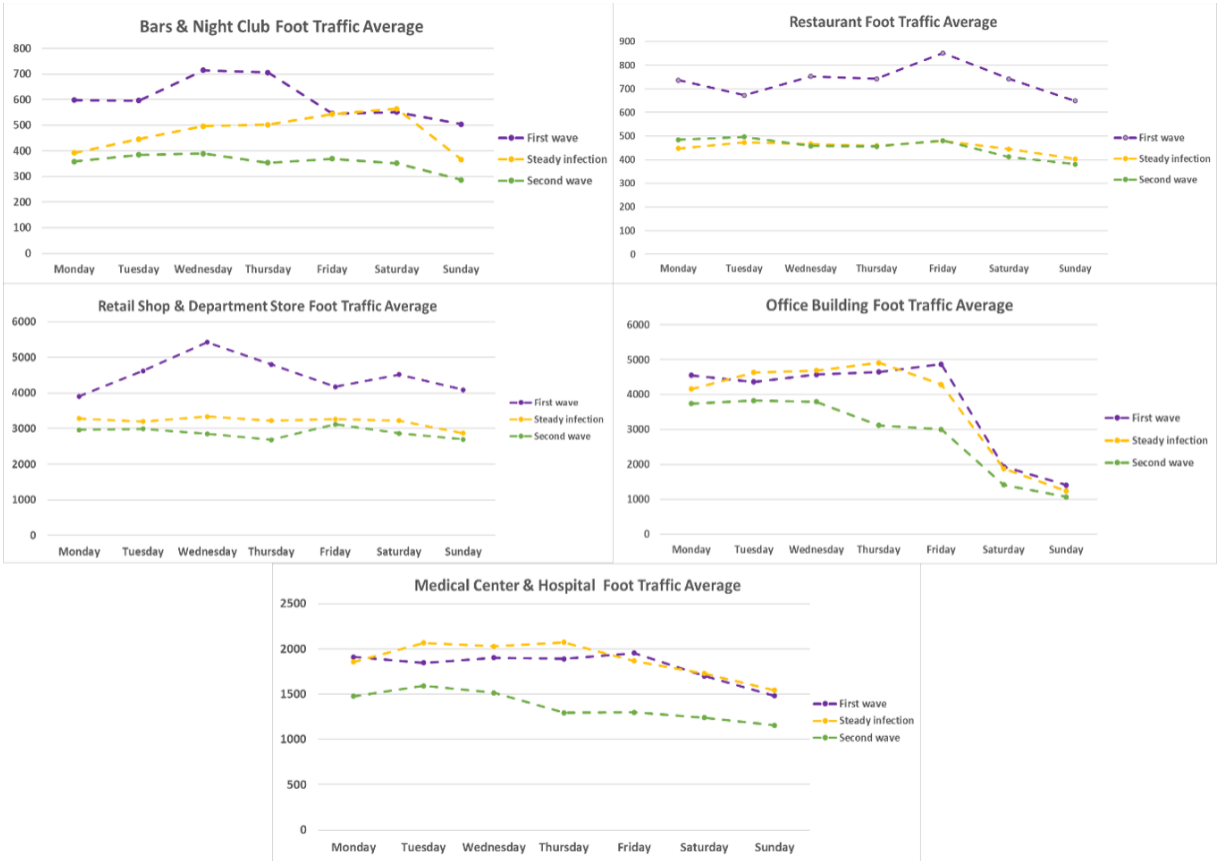
2) ช่วงจำนวนผู้ติดเชื้อต่อวันคงที่ การยกเลิกเคอร์ฟิวในวันที่ 14 มิถุนายน. ค.ศ. 2020 ธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า วันถัดมาจำนวนการสัญจรด้วยเท้าเพิ่มขึ้นจาก 337 traffic เป็น 453 traffic, 664 traffic เป็น 712 traffic และ 4882 traffic เป็น 5133 traffic ตามลำดับ โดย 5 วันหลังประกาศจำนวนการสัญจรด้วยเท้าสูงขึ้นเป็น 714 traffic, 810 traffic และ 6633 traffic ตามลำดับ การอนุญาตให้จำหน่ายและบริโภคเครื่องดื่มแอลกอฮอล์ภายในร้านอาหารได้ในช่วงเวลาที่กำหนด, บาร์และสถานบันเทิงเปิดได้ถึงเที่ยงคืน และห้างสรรพสินค้าเปิดได้ถึง 22.00 น. และร้านสะดวกซื้อเปิดได้ 24 ชั่วโมง จำนวนการสัญจรด้วยเท้ามีแนวโน้มเพิ่มสูงขึ้นเล็กน้อยจนถึงเดือนสิงหาคม ในส่วนค่าเฉลี่ยต่อวันธุรกิจบาร์และสถานบันเทิงวันศุกร์และเสาร์ใกล้เคียงกับช่วงระบาดระลอกแรกต่างกันเพียง 1.84 traffic และ 12.93 traffic ตามลำดับ ส่วนธุรกิจอาคารสำนักงาน และธุรกิจศูนย์การแพทย์และโรงพยาบาล ค่าเฉลี่ยต่างกับช่วงระบาดระลอกแรกเล็กน้อยอยู่ที่ 80.10 traffic และ 68.58 traffic ตามลำดับ

3) ช่วงระบาดระลอกที่ 2 การเปิดรับนักท่องเที่ยวต่างชาติกรณีพิเศษ สัปดาห์ต่อมาธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า จำนวนการสัญจรด้วยเท้าเพิ่มจาก 360 traffic เป็น 798 traffic, 467 traffic เป็น 1031 traffic และ 2567 traffic เป็น 6989 traffic ตามลำดับ การสั่งปิดบาร์และสถานบันเทิงอีกครั้งในเดือนธันวาคม ค.ศ. 2020 ใช้เวลาประมาณสองเดือนหลังประกาศอยู่ระดับต่ำกว่า 50 traffic ซึ่งวันที่ 15–19 มกราคม ค.ศ. 2021 จำนวนการสัญจรด้วยเท้าเพิ่มขึ้นอย่างฉับพลัน การห้ามจำหน่ายและบริโภคเครื่องดื่มแอลกอฮอล์และนั่งทานอาหารที่ร้านได้ช่วง 06.00–21.00 น. พบว่าวันถัดมาธุรกิจร้านอาหารมีจำนวนการสัญจรด้วยเท้าเพิ่มขึ้นจาก 290 traffic เป็น 504 traffic ต่อมาอนุญาตให้นั่งรับประทานที่ร้านได้ถึง 23.00 น. วันถัดมาจำนวนลดลงจาก 315 traffic เหลือ 44 traffic และจำนวนคงที่ในช่วง 20–80 traffic เป็นระยะเวลาประมาณ 2 สัปดาห์ ซึ่งในวันเดียวกันนี้ห้างสรรพสินค้าและศูนย์การค้าทั้งหมดสามารถเปิดบริการได้ตามปกติ พบว่าวันถัดมาธุรกิจร้านค้าปลีกและห้างสรรพสินค้าจำนวนการสัญจรด้วยเท้าลดลงจาก 1607 traffic เหลือ 289 traffic และอยู่ระหว่าง 66–449 traffic เป็นระยะเวลาประมาณ 2 สัปดาห์ และวันที่ 24 กุมภาพันธ์ ค.ศ. 2021 การอนุญาตให้ธุรกิจบาร์และสถานบันเทิงเปิดได้ไม่เกิน 23.00 น. และจำหน่ายและบริโภคเครื่องดื่มแอลกอฮอล์ที่ร้านอาหารได้ในช่วงเวลาที่กำหนด พบว่าก่อนประกาศ 1 สัปดาห์ ธุรกิจบาร์และสถานบันเทิง และธุรกิจร้านอาหาร จำนวนการสัญจรด้วยเท้าเพิ่มขึ้นอยู่ก่อนแล้วจาก 26 traffic เป็น 637

traffic และ 40 traffic เป็น 874 traffic ตามลำดับ วันถัดมาหลังประกาศเพิ่มขึ้นจาก 662 traffic เป็น 762 traffic และ 742 traffic เป็น 844 traffic ตามลำดับ ส่วนค่าเฉลี่ยต่อวันพบว่าช่วงระบาดระลอกที่ 2 ธุรกิจทุกประเภทมีค่าเฉลี่ยการสัญจรด้วยเท้าน้อยที่สุด ซึ่งธุรกิจศูนย์การแพทย์และโรงพยาบาลต่างกับช่วงผู้ติดเชื้อต่อวันคงที่ 0.71 traffic



รูปที่ 12 : การสัญจรด้วยเท้าช่วงเวลาที่เกี่ยวข้องกับมาตรการโควิด-19



รูปที่ 13 : ค่าเฉลี่ยการสัญจรด้วยเท้าต่อวัน

ตารางที่ 6 : ค่าเฉลี่ยการสัญจรด้วยเท้าต่อวันของแต่ละช่วงเวลา

	Period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Average
Bar & Night Club	1	598.25	596.14	714.59	705.93	545.44	551.34	503.49	602.17
	2	392.38	445.69	496.51	501.56	543.60	564.27	366.54	472.94
	3	358.80	384.80	389.74	353.43	369.11	351.92	286.23	356.29
Restaurant	1	736.47	673.12	753.24	742.18	851.94	742.41	649.59	735.56
	2	448.35	473.27	466.96	459.08	480.42	444.89	403.38	453.76
	3	485.05	497.20	458.20	456.32	481.00	412.74	380.90	453.06
Retail Shop & Department Store	1	3903.18	4615.82	5421.06	4799.59	4169.59	4515.94	4086.35	4501.65
	2	3279.69	3195.46	3331.85	3217.96	3263.15	3219.58	2871.15	3196.98
	3	2959.15	2991.10	2848.30	2683.70	3116.79	2864.79	2692.32	2879.45
Office Building	1	4552.76	4366.18	4575.29	4644.76	4869.88	1932.24	1406.76	3763.98
	2	4156.96	4634.23	4684.50	4910.58	4285.92	1880.59	1234.42	3683.89
	3	3742.10	3828.15	3795.95	3120.05	3005.89	1411.00	1066.10	2852.75
Medical Center & Hospital	1	1910.35	1844.41	1900.94	1887.82	1952.06	1700.06	1480.47	1810.87
	2	1855.04	2066.35	2027.62	2070.42	1866.15	1727.41	1543.19	1879.45
	3	1476.25	1591.45	1513.95	1293.47	1299.11	1238.95	1155.10	1366.90

ตารางที่ 7 : ผลการทดสอบด้วย Wilcoxon Signed-Rank

	Mean				Wilcoxon Signed-Rank Test (P value)					
Bar & Night Club	Apr2019	Apr2020	Apr2021	Oct2022	Apr2019 - Apr2020	Apr2019 - Apr2021	Apr2019 - Oct2022	Apr2020 - Apr2021	Apr2020 - Oct2022	Apr2021 - Oct2022
Week 1	118.63	2	44.19	22.88	0.002	0.002	0.002	0.004	0.002	0.033
Week 2	83.29	2.3	10.93	13.14	<.001	<.001	<.001	<.001	<.001	0.805
Week 3	89.11	2.27	6.23	13.14	<.001	<.001	<.001	<.001	<.001	<.001
Week 4	94.41	2.43	6.23	14.27	<.001	<.001	<.001	<.001	<.001	<.001
Week 5	95.81	1.63	1.25	13.88	0.008	0.008	0.008	0.414	0.007	0.007
Overall mean	91.86	2.25	10.19	13.48						
Oct2022 recovery (%)	-85.33%	499.11%	32.29%	-						
Office Building										
Week 1	173.23	25.32	85.82	48.68	<.001	0.003	<.001	0.042	<.001	0.289
Week 2	587.42	85.03	280.35	88.38	<.001	<.001	<.001	<.001	0.465	<.001
Week 3	535.88	86.87	68.51	68.92	<.001	<.001	<.001	0.048	<.001	0.855
Week 4	1029.32	108.57	159.42	78.92	<.001	<.001	<.001	0.038	<.001	0.016
Week 5	1401.68	160.41	43.68	58.59	<.001	<.001	<.001	<.001	<.001	0.198
Overall mean	728.73	93.39	152.68	74.72						
Oct2022 recovery (%)	-89.75%	-19.99%	-51.06%	-						
Medical Center & Hospital										
1st-7th	132.3	32.67	75.23	31.74	<.001	<.001	<.001	<.001	0.462	<.001
8th -14th	124.607	30.625	42.006	21.0357	<.001	<.001	<.001	0.154	<.001	0.002
15th – 21st	135.149	30.3988	78.5417	23.125	<.001	<.001	<.001	<.001	<.001	<.001
22nd -30th	164.125	39.1944	12.7407	19.7454	<.001	<.001	<.001	<.001	<.001	<.001
Overall	140.72	33.62	49.5	23.63						
Oct2022 recovery (%)	-83.21%	-29.71%	-52.26%	-						
Restaurant										
1st-7th	59.3	13.29	32.8	23.19	<.001	<.001	<.001	<.001	<.001	0.347
8th -14th	57.49	9.68	19.52	18.59	<.001	<.001	<.001	<.001	<.001	0.676
15th – 21st	61.95	9.33	32.45	18.18	<.001	<.001	<.001	<.001	<.001	<.001
22nd -30th	70.72	13.45	8.24	15.98	<.001	<.001	<.001	<.001	<.001	<.001
Overall mean	62.92	11.57	22.25	18.79						
Oct2022 recovery (%)	-70.14%	62.40%	-15.55%	-						
Retail Shop & Department Store										
1st-7th	705.56	48.41	235.3	139.71	<.001	<.001	<.001	<.001	<.001	<.001
8th -14th	661.09	49.56	122.89	121.35	<.001	<.001	<.001	<.001	<.001	0.913
15th – 21st	660.43	51.7	201.73	109.9	<.001	<.001	<.001	<.001	<.001	<.001
22nd -30th	832.21	58.08	29.81	106.19	<.001	<.001	<.001	<.001	<.001	<.001
Overall mean	722.65	52.35	139.59	118.41						
Oct2022 recovery (%)	-83.61%	126.19%	-15.17%	-						

4.3) ผลสถิติทดสอบความแตกต่างปริมาณการสัญจรด้วยเท้า

กำหนดให้ Z หมายถึง สถิติทดสอบซี

p หมายถึง ค่านัยสำคัญ

จากสถิติทดสอบ Kolmogorov-Smirnov พบว่ามีข้อมูลที่แจกแจงแบบไม่ปกติ หากใช้สถิติทดสอบแบบพาราเมตริกผลที่ได้จะขาดความน่าเชื่อถือจึงเลือกใช้สถิติทดสอบแบบนอนพาราเมตริกซึ่งเป็นสถิติที่ลักษณะการแจกแจงข้อมูลไม่ส่งผล จากตารางที่ 7 ผลการใช้สถิติทดสอบ Wilcoxon Signed-Rank เปรียบเทียบข้อมูลทีละคู่ที่ระดับนัยสำคัญ 0.05 พบว่าคู่ที่ยอมรับสมมติฐานหลัก คือ สถานการณ์แพร่ระบาดโควิด-19 ไม่ส่งผลให้ข้อมูลทั้งสองช่วงเวลาแตกต่างกัน ดังนี้ ธุรกิจบาร์และสถานบันเทิง สัปดาห์ที่ 2 เม.ย. 2021 เทียบกับ ต.ค. 2022 และสัปดาห์ที่ 5 เม.ย. 2020 เทียบกับ เม.ย. 2021 ($Z = -0.246,\ p = 0.805$) และ ($Z = -0.816,\ p = 0.414$) ตามลำดับ, ธุรกิจอาคารสำนักงานสัปดาห์ที่ 1, 3, 5 เม.ย. 2021 เทียบกับ ต.ค. 2022 และสัปดาห์ที่ 2 เม.ย. 2020 เทียบ ต.ค. 2022 ($Z = -1.060,\ p = 0.289$), ($Z = -0.183,\ p = 0.855$), ($Z = -1.288,\ p = 0.198$) และ ($Z = -0.730,\ p = 0.465$) ตามลำดับ, ธุรกิจศูนย์การแพทย์และโรงพยาบาลวันที่ 1–7 เม.ย. 2020 เทียบกับ ต.ค. 2022 โดย ($Z = -0.735,\ p = 0.462$) และวันที่ 8–14 เม.ย. 2020 เทียบ เม.ย. 2021 ($Z = -1.424,\ p = 0.154$), ธุรกิจร้านอาหารวันที่ 1–7 และ 8–14 เม.ย. 2021 เทียบกับ ต.ค. 2022 โดย ($Z = -0.941,\ p = 0.347$) และ ($Z = -0.417, p = 0.676$) ตามลำดับ และธุรกิจร้านค้าปลีกและห้างสรรพสินค้าวันที่ 8–14 เม.ย. 2021 เทียบกับ ต.ค. 2022 ($Z = -0.109,\ p = 0.913$)

5) อภิปรายผลการวิจัย

5.1) อภิปรายผลการศึกษาและเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้าของธุรกิจแต่ละประเภท

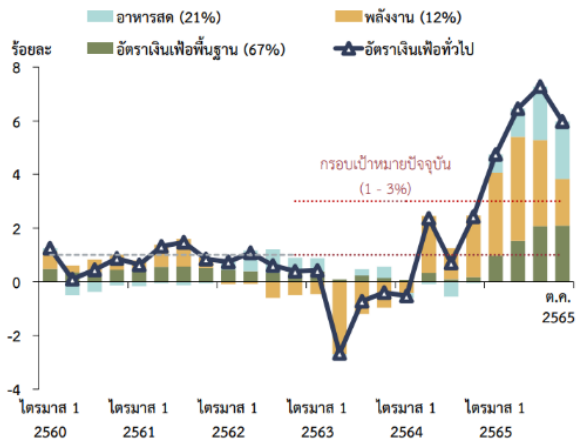
การจัดกลุ่มสถานที่ด้วยจำนวนการสัญจรด้วยเท้าเฉลี่ยต่อวันก่อนพิจารณาเปรียบเทียบด้วยเงื่อนไขเป็นวิธีการที่ไม่ซับซ้อนคล้ายกับวิธีการของ [21] โดยการจัดกลุ่มส่งผลต่อรูปแบบการสัญจรด้วยเท้าของธุรกิจ กล่าวคือ ยิ่งจำนวน traffic ต่อวันสูงจะยิ่งแสดงรูปแบบที่มีเอกลักษณ์ชัดเจนมากขึ้น โดยเฉพาะธุรกิจบาร์และสถานบันเทิง, ธุรกิจศูนย์การแพทย์และโรงพยาบาลและธุรกิจร้านอาหาร จะเห็นได้ว่าทั้ง 3 ธุรกิจนี้กลุ่มขนาดเล็กมีรูปแบบการสัญจรด้วยเท้าที่หลากหลาย ซึ่งทุกธุรกิจที่นำมาทดสอบกลุ่มขนาดใหญ่จะมีจำนวนคู่ที่สามารถผ่านเงื่อนไขการเปรียบเทียบมากกว่ากลุ่มขนาดเล็กที่อยู่ในธุรกิจประเภทเดียวกันในทำนองเดียวกันกลุ่มที่มีจำนวนการสัญจรด้วยเท้าต่อวันสูงที่สุดของธุรกิจแต่ละประเภทจะมีรูปแบบการสัญจรด้วยเท้าของแต่ละสถานที่คล้ายกันส่งผลให้มีจำนวนคู่ที่สามารถผ่านเงื่อนไขการเปรียบเทียบมากกว่ากลุ่มขนาดเล็กที่อยู่ในธุรกิจประเภทเดียวกันในทำนองเดียวกันกลุ่มที่มีจำนวนการสัญจรด้วยเท้าต่อวันสูงที่สุดของธุรกิจแต่ละประเภทจะมีรูปแบบการสัญจรด้วยเท้าของแต่ละสถานที่คล้ายกันส่งผลให้มีจำนวนคู่ที่สามารถผ่านเงื่อนไขการเปรียบเทียบสูงกว่า 70% ของจำนวนคู่ทั้งหมด โดยธุรกิจที่มีรูปแบบการสัญจรด้วยเท้าเป็นเอกลักษณ์ชัดเจน คือ ธุรกิจบาร์และสถานบันเทิงที่จำนวนการสัญจรด้วยเท้าเพิ่มขึ้นวันศุกร์และเสาร์ และธุรกิจอาคารสำนักงานที่จำนวนการสัญจรด้วยเท้าลดลงอย่างฉับพลันในวันเสาร์และวันอาทิตย์

5.2) อภิปรายผลกระทบประกาศมาตรการเกี่ยวข้องกับโควิด-19

ธุรกิจบาร์และสถานบันเทิง ธุรกิจร้านค้าปลีกและห้างสรรพสินค้าและธุรกิจร้านอาหาร ได้รับผลกระทบสูงจากประกาศมาตรการครั้งแรกเกี่ยวกับสั่งปิดทำการและห้ามรับประทานอาหารในร้านช่วงเดือนมีนาคม ค.ศ. 2020 ซึ่งธุรกิจร้านค้าปลีกและห้างสรรพสินค้าเป็นที่น่าสังเกตว่าถึงแม้จะสั่งปิดห้างสรรพสินค้าแต่ในวันต่อมาจำนวนการสัญจรด้วยเท้าเพิ่มสูงขึ้นเกือบ 2000 traffic อาจเป็นเพราะมีการออกมาจับจ่ายเพื่อกักตุนสินค้าจึงแนะนำรัฐบาลหรือผู้ที่เกี่ยวข้องว่าควรประกาศมาตรการที่ค่อย ๆ ลดจำนวนผู้ใช้บริการลง เพื่อให้ผู้ประกอบการสามารถปรับตัวและช่วยบรรเทาความเสียหายจากทั้งความเสี่ยงแพร่เชื้อโควิด-19 และสินค้าขาดแคลนจากการกักตุน การประกาศ พ.ร.ก.ฉุกเฉิน, เคอร์ฟิว และห้ามนักท่องเที่ยวต่างชาติเข้าประเทศไทย ส่งผลให้ทั้ง 3 ธุรกิจ จำนวนการสัญจรด้วยเท้าลดลงและคงที่ในระดับต่ำที่สุดตลอดระยะเวลา 1 เดือน ซึ่งถือว่ามาตรการนี้สามารถควบคุมให้ประชาชนออกมาใช้น้อยลง โดยหลังจากยกเลิก พ.ร.ก.ฉุกเฉินและเริ่มมีนโยบายผ่อนคลายตามมา แนวโน้มการสัญจรด้วยเท้าปรับตัวสูงขึ้นและคงที่ระยะเวลาประมาณ 2 เดือน หลังจากเปิดรับนักท่องเที่ยวต่างชาติกรณีพิเศษพบว่าจำนวนการสัญจรด้วยเท้ามีแนวโน้มเพิ่มสูงขึ้น และการสั่งปิดธุรกิจบาร์และสถานบันเทิงครั้งที่ 2 พบว่าช่วงวันที่ 15–19 มกราคม ค.ศ. 2021 จำนวนการสัญจรด้วยเท้าเพิ่มสูงขึ้นอย่างฉับพลัน ซึ่งเป็นที่น่าสนใจว่าอาจมีการเปิดให้บริการในช่วงที่ประกาศปิดบาร์และสถานบันเทิง

5.3) อภิปรายผลการฟื้นตัวธุรกิจ

เมื่อพิจารณาการฟื้นตัวปริมาณการสัญจรด้วยเท้าเดือนตุลาคม ค.ศ. 2022 เทียบกับเมษายน ค.ศ. 2021 ธุรกิจอาคารสำนักงาน, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า ค่าเฉลี่ยการสัญจรด้วยเท้ายังคงลดลงยกเว้นธุรกิจบาร์และสถานบันเทิงที่เริ่มมีการฟื้นตัว โดยธุรกิจร้านอาหารวันที่ 1–14, ธุรกิจร้านค้าปลีกและห้างสรรพสินค้าวันที่ 8–14, ธุรกิจอาคารสำนักงานสัปดาห์ที่ 1, 3, 5 และธุรกิจบาร์และสถานบันเทิงสัปดาห์ที่ 2 พบว่าสถานการณ์การระบาดโควิด-19 ไม่ส่งผลกระทบให้ปริมาณการสัญจรด้วยเท้าแตกต่างกันอย่างมีนัยสำคัญ ส่วนธุรกิจศูนย์การแพทย์และโรงพยาบาลที่ค่าเฉลี่ยการสัญจรด้วยเท้าลดลง 52.26% และปริมาณการสัญจรด้วยเท้าแตกต่างกันอย่างมีนัยสำคัญทุกช่วงเวลา เมื่อพิจารณาเดือนตุลาคม ค.ศ. 2022 เทียบกับเมษายน ค.ศ. 2020 พบว่าธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้ามีภาพรวมค่าเฉลี่ยการสัญจรด้วยเท้าเพิ่มขึ้นสวนทางกับค่าเฉลี่ยการสัญจรด้วยเท้าธุรกิจอาคารสำนักงานและธุรกิจศูนย์การแพทย์และโรงพยาบาลที่ลดลง โดยธุรกิจอาคารสำนักงานสัปดาห์ที่ 2 และธุรกิจศูนย์การแพทย์และโรงพยาบาลวันที่ 1–7 สถานการณ์การระบาดโควิด-19 ไม่ส่งผลกระทบให้ปริมาณการสัญจรด้วยเท้าแตกต่างกันอย่างมีนัยสำคัญ และไม่มีธุรกิจประเภทใดที่สามารถฟื้นตัวได้ใกล้เคียงหรือมากกว่าช่วงก่อนการระบาดในเดือนเมษายน ค.ศ. 2019 ซึ่งสอดคล้องกับสถานการณ์เงินเฟ้อภายในประเทศไทยที่ทำจุดสูงสุดช่วงปลายไตรมาสที่ 3 ปี ค.ศ. 2022 (รูปที่ 14) [22]



รูปที่ 14 : อัตราเงินเฟ้อในประเทศไทย [22]

ส่งผลให้ประชาชนเกิดความระมัดระวังในการใช้จ่ายกันมากขึ้นจึงเป็นปัจจัยหนึ่งที่ส่งผลให้ปริมาณการสัญจรด้วยเท้าของประเภทธุรกิจที่เลือกศึกษาไม่เกิดการฟื้นตัวมากนักในเดือนตุลาคม ค.ศ. 2022

6) สรุปผลการวิจัย

การศึกษาและเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้า เมื่อจัดกลุ่มสถานที่ตัวอย่างด้วยจำนวนการสัญจรด้วยเท้าต่อวันทำให้เห็นรูปแบบการสัญจรด้วยเท้าที่เป็นเอกลักษณ์ของธุรกิจแต่ละประเภทซึ่งสัมพันธ์กับจำนวนการสัญจรด้วยเท้าต่อวัน โดยรูปแบบการสัญจรด้วยเท้าจะชัดเจนมากขึ้นหากจำนวนการสัญจรด้วยเท้าต่อวันสูง และยังสามารถระบุประเภทธุรกิจของสถานที่ได้ด้วยเงื่อนไขที่ไม่ซับซ้อน

เมื่อวิเคราะห์ผลกระทบจากมาตรการที่เกี่ยวข้องกับโควิด-19 จากรัฐบาล พบว่าหลังการประกาศครั้งแรกเกี่ยวกับการสั่งปิดบริการและห้ามรับประทานอาหารในร้านอาหารจนถึงก่อนประกาศใช้ พ.ร.ก.ฉุกเฉิน ส่งผลกระทบอย่างมากต่อจำนวนการสัญจรด้วยเท้าธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านค้าปลีกและห้างสรรพสินค้า และธุรกิจร้านอาหาร ลดลง 84.70%, 79.65% และ 75.34% ตามลำดับ การประกาศใช้พ.ร.ก.ฉุกเฉิน, เคอร์ฟิว และห้ามนักท่องเที่ยวต่างชาติเดินทางเข้าประเทศไทยสามารถควบคุมปริมาณการสัญจรด้วยเท้าให้คงที่และอยู่ในระดับต่ำที่สุดของช่วงเวลาที่เลือกศึกษา แต่การสั่งปิดธุรกิจบาร์และสถานบันเทิงครั้งที่ 2 ในเดือนธันวาคม ค.ศ. 2020 ใช้เวลาถึงสองเดือนเพื่อควบคุมปริมาณการสัญจรด้วยเท้าของธุรกิจบาร์และสถานบันเทิงให้คงที่อยู่ในระดับต่ำ โดยการสัญจรด้วยเท้าเพิ่มสูงขึ้นอย่างฉับพลันในบางช่วงถึงแม้จะมีการประกาศสั่งปิดแล้วก็ตาม ซึ่งมาตรการที่กล่าวมาไม่เพียงช่วยลดปริมาณการสัญจรด้วยเท้าแต่ยังสามารถทำให้เกิดการฟื้นตัวได้เช่นกัน การยกเลิก พ.ร.ก. ฉุกเฉิน เห็นได้ชัดว่าปริมาณการสัญจรด้วยเท้าเพิ่มสูงขึ้นและคงที่เป็นระยะเวลานาน อีกทั้งหลังจากเปิดรับนักท่องเที่ยวต่างชาติกรณีพิเศษปริมาณการสัญจรด้วยเท้ามีแนวโน้มเพิ่มสูงขึ้นเช่นกัน

การฟื้นตัวปริมาณการสัญจรด้วยเท้าในเดือนตุลาคม ค.ศ. 2022 เทียบกับเดือนเมษายน ค.ศ. 2021 ธุรกิจอาคารสำนักงาน, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า ค่าเฉลี่ยการสัญจรด้วยเท้าลดลง 51.06%, 15.55% และ 15.17% ตามลำดับ ยกเว้นธุรกิจบาร์และสถานบันเทิงที่เริ่มมีการฟื้นตัวและมีค่าเฉลี่ยการสัญจรด้วยเท้าเพิ่มขึ้น 32.29% เมื่อ

เทียบเดือนตุลาคม ค.ศ. 2022 กับเดือนเมษายน ค.ศ. 2020 ธุรกิจบาร์และสถานบันเทิง, ธุรกิจร้านอาหาร และธุรกิจร้านค้าปลีกและห้างสรรพสินค้า ค่าเฉลี่ยการสัญจรด้วยเท้าเพิ่มขึ้น 499.11% 62.40% และ 126.19% ตามลำดับ ยกเว้นธุรกิจอาคารสำนักงาน และธุรกิจศูนย์การแพทย์และโรงพยาบาล ค่าเฉลี่ยลดลง 19.99% และ 29.71% ตามลำดับ และไม่มีธุรกิจประเภทใดสามารถฟื้นตัวได้ใกล้เคียงหรือมากกว่าช่วงก่อนการระบาดในเดือนเมษายน ค.ศ. 2019

7) ข้อเสนอแนะ

1. การเปรียบเทียบความเหมือน-ต่างรูปแบบการสัญจรด้วยเท้ากลุ่มขนาดกลางและใหญ่ควรเพิ่มจำนวนสถานที่ตัวอย่างเพื่อให้รูปแบบค่าเฉลี่ยของกลุ่มใกล้เคียงกับความเป็นจริงมากขึ้น
2. เพิ่มช่วงเวลาที่นอกเหนือจากช่วงที่มีปัจจัยเงินเฟ้อเข้ามาเกี่ยวข้องสำหรับวิเคราะห์การฟื้นตัวธุรกิจเพื่อให้เห็นลักษณะการเปลี่ยนแปลงปริมาณการสัญจรด้วยเท้าในสภาวะเศรษฐกิจปกติ

8) งานวิจัยในอนาคต

ประยุกต์ใช้ Machine Learning สำหรับจำแนกธุรกิจด้วยรูปแบบการสัญจรด้วยเท้า จากการทดสอบเบื้องต้นด้วยอัลกอริทึม Dynamic time warping, Random forest classifier และ Gradient boosting classifier ที่จำนวน Training data และ Test data เท่ากันคือ 80 และ 12 ตามลำดับ พบว่าความถูกต้องโดยรวมของการจำแนก (Overall accuracy) ของทั้ง 3 วิธีอยู่ที่ 0.25, 0.50 และ 0.58 ตามลำดับ ซึ่งการวิจัยในอนาคตจะทดสอบกับจำนวนข้อมูลที่มากขึ้น, ปรับไฮเปอร์พารามิเตอร์ให้เหมาะสมกับการทดสอบ และวิเคราะห์ความสัมพันธ์ร่วม (Correlation) ระหว่างข้อมูลก่อนนำไปประมวลผลเพื่อเพิ่มประสิทธิภาพในการจำแนกให้ดียิ่งขึ้น

กิตติกรรมประกาศ

ข้อมูล Foot traffic ที่ใช้ในงานวิจัยนี้ได้รับความอนุเคราะห์จากบริษัท ไทยคม จำกัด (มหาชน) ขอขอบคุณสำหรับคำปรึกษาตลอดระยะเวลาตั้งแต่เริ่มวิจัย และขอขอบคุณ ดร.ธงทิศ ฉายากุล ที่ปรึกษาวิทยานิพนธ์ ที่ให้อิสระในการวิจัยภายใต้คำแนะนำและให้คำปรึกษาแก้ไขปัญหาต่าง ๆ จนงานวิจัยนี้เสร็จสมบูรณ์

REFERENCES

[1] X. Wang, J. Liono, W. Mcintosh, and F. D. Salim, “Predicting the city foot traffic with pedestrian sensor data,” in *Proc. 14th EAI Int. Conf. Mobile and Ubiquitous Syst.: Comput., Netw. and Services*, Melbourne, Australia, 2017, pp. 1–10.

[2] T. Chao, R. Du, J. Gluck, H. Maidasani, K. Wills, and B. Shneiderman, “C-flow: visualizing foot traffic and profit data to make informative decisions,” Department of Computer Science, University of Maryland, College Park, MD, USA, 2013. [Online]. Available: https://www.researchgate.net/publication/340935443_C-Flow_Visualizing_Foot_Traffic_and_Profit_Data_to_Make_Informative_Decisions

[3] W. J. Ladeira, F. de Oliveira Santini, and D. C. Pinto, “Clockwise versus counterclockwise turning bias: Moderation effects of foot traffic and cognitive experience on visual attention,” *J. Retailing Consumer Services*, vol. 67, 2022, Art. no. 102965.

[4] G. Alizo, A. Onayemi, J. D. Sciarretta, and J. M. Davis, “Operating room foot traffic: a risk factor for surgical site infections,” *Surgical Infections*, vol. 20, no. 2, pp. 146–150, 2019.

[5] J. Barthélemy, N. Verstaevel, H. Forehead, and P. Perez, “Edge-computing video analytics for real-time traffic monitoring in a smart city,” *Sensors*, vol. 19, no. 9, p. 2048, 2019.

[6] A. Sevtsuk, R. Basu, and B. Chancey, “We shape our buildings, but do they then shape us? A longitudinal analysis of pedestrian flows and development activity in Melbourne,” *PloS one*, vol. 16, no. 9, 2021, Art. no. e0257534.

[7] S. Hasan and S. V. Ukkusuri, “Urban activity pattern classification using topic models from online geo-location data,” *Transp. Res. Part C: Emerging Technol.*, vol. 44, pp. 363–381, 2014.

[8] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, “Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment,” *Transp. Res. Part C: Emerg. Technol.*, vol. 18, no. 4, pp. 568–583, 2010.

[9] S. Abrishami, P. Kumar, and W. Nienaber, “Smart stores: A scalable foot traffic collection and prediction system,” in *Proc. 17th Indus. Conf. Data Mining (ICDM 2017),* Leipzig, Germany, Jul. 2017, pp. 107–121.

[10] F. Zheng and Q. Liu, “Anomalous telecom customer behavior detection and clustering analysis based on ISP’s operating data,” *IEEE Access*, vol. 8, pp. 42734–42748, 2020.

[11] J. Roetman, "Using WiFi Signals in Combination with GPS to Track Human Traffic in a Space," *EasyChair,* no. 3902, 2020. [Online]. Available: https://easychair.org/publications/preprint/vzfQ

[12] V. Romero, W. D. Stone, and J. D. Ford, "COVID-19 indoor exposure levels: An analysis of foot traffic scenarios within an academic building," *Transp. Res. Interdisciplinary Perspectives*, vol. 7, 2020, Art. no. 100185.

[13] C. J. Cronin and W. N. Evans, "Private precaution and public restrictions: what drives social distancing and industry foot traffic in the COVID-19 era?," National Bureau of Economic Research, Cambridge, U.K., 2020. [Online]. Available: https://www.nber.org/system/files/working_papers/w27531/w27531.pdf

[14] K. B. Leung, R. Alam, S. C. Brooks, and T. C. Chan, "Public defibrillator accessibility and mobility trends during the COVID-19 pandemic in Canada," *Resuscitation*, vol. 162, pp. 329–333, 2021.

[15] A. Kulkarni, M. Kim, J. Bhattacharya, and J. Bhattacharya, "Businesses in high-income zip codes often saw sharper visit reductions during the COVID-19 pandemic," *Humanities Social Sci. Commun.*, vol. 10, 2022, Art. no. 713.

[16] K. M. Ramachandran and C. P. Tsokos, *Mathematical Statistics with Applications in R*. Cambridge, MA, USA: Academic Press, 2020.

[17] NIST. "Kolmogorov-Smirnov Goodness-of-Fit Test." NIST.gov. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm (accessed Jan 1, 2023).

[18] S. Chansakul and S. Bowarnkitiwong, "Nonparametric statistics and its application in nursing research," (in Thai), *EAU Heritage J. Sci. Technol.*, vol. 11, no. 1, pp. 38–48, 2017.

[19] E. Kasuya, "Wilcoxon signed-ranks test: symmetry should be confirmed before the test," *Animal Behaviour*, vol. 3, no. 79, pp. 765–767, 2010.

[20] Orbital Insight. "Geolocation - Foot Traffic." ORBITAL INSIGHT.com. https://docs.orbitalinsight.com/docs/geolocation-foot-traffic (accessed Nov. 29, 2022).

[21] S. Y. Cheung, S. C. Ergen, and P. Varaiya, "Traffic surveillance with wireless magnetic sensors," University of California, Berkeley, CA, USA, 2005. [Online]. Available: https://people.eecs.berkeley.edu/~varaiya/papers_ps.dir/sensors_its_final9.pdf

[22] Bank of Thailand, "Monetary Policy Report November 2022," 2022. [Online]. Available: https://www.bot.or.th/content/dam/bot/documents/en/our-roles/monetary-policy/mpc-publication/monetary-policy-report/MPR_2022_Q4.pdf

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

การปรับปรุงการตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดี สำหรับวงจรกรองกำลังแอกทีฟแบบขนานในระบบไฟฟ้ากำลังหนึ่งเฟส

สุทธิณี ว่าวสูงเนิน[1] ทศพร ณรงค์ฤทธิ์[2*] กองพล อารีรักษ์[3]

[1,2,3]กลุ่มวิจัย PEMC สาขาวิชาวิศวกรรมไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, นครราชสีมา, ประเทศไทย*

*ผู้ประพันธ์บรรณกิจ อีเมล : tosaporn@sut.ac.th

รับต้นฉบับ : 13 มีนาคม 2566; รับบทความฉบับแก้ไข : 31 สิงหาคม 2566; ตอบรับบทความ : 28 กันยายน 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

บทความนี้นำเสนอการปรับปรุงสมรรถนะการตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดี (SDF) เพื่อใช้คำนวณหากระแสอ้างอิงของวงจรกรองกำลังแอกทีฟแบบขนานในระบบไฟฟ้ากำลังหนึ่งเฟส โดยการปรับปรุงดังกล่าวจะประยุกต์ใช้วิธีการตรวจจับแรงดันลำดับเฟสบวกมูลฐาน (PSVD) ร่วมกับวิธีฟูริเยร์เอสดีเพื่อเพิ่มความถูกต้องในการคำนวณกระแสอ้างอิงในกรณีที่แรงดันที่แหล่งจ่ายมีความเพี้ยนของรูปคลื่น โดยทั้งนี้มีวัตถุประสงค์เพื่อเพิ่มประสิทธิผลการกำจัดฮาร์มอนิกสำหรับระบบไฟฟ้ากำลังหนึ่งเฟส การทดสอบสมรรถนะการตรวจจับฮาร์มอนิกจะใช้เทคนิคการจำลองสถานการณ์แบบฮาร์ดแวร์ในลูปที่ใช้โปรแกรม Simulink ร่วมกับบอร์ด DSP รุ่น Experimenter Kit เพื่อจำลองผลระบบกำจัดฮาร์มอนิก โดยการจำลองดังกล่าวจะแบ่งออกเป็น 2 กรณี ได้แก่ กรณีแรงดันที่แหล่งจ่ายเป็นรูปคลื่นไซน์บริสุทธิ์ และกรณีแรงดันที่แหล่งจ่ายมีความเพี้ยนของรูปคลื่นเนื่องจากมีฮาร์มอนิกปะปน ผลการจำลองสถานการณ์แสดงให้เห็นว่าการตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดีร่วมกับการตรวจจับแรงดันลำดับเฟสบวกมูลฐาน (SDF + PSVD) สามารถให้ค่าเปอร์เซ็นต์ความเพี้ยนฮาร์มอนิกรวม (%THD) ของกระแสที่แหล่งจ่ายภายหลังการชดเชยลดลงโดยในกรณีที่แรงดันที่แหล่งจ่ายเป็นรูปไซน์มีค่าเท่ากับ 3.07% และในกรณีที่แรงดันที่แหล่งจ่ายมีความเพี้ยนของรูปคลื่นมีค่าเท่ากับ 2.00% ซึ่งมีค่าน้อยกว่าเมื่อเปรียบเทียบกับการใช้เฉพาะวิธีฟูริเยร์เอสดี (SDF) และวิธีการตรวจจับซิงโครนัส (SD) จากผลดังกล่าวยืนยันได้ว่าวิธีฟูริเยร์เอสดีร่วมกับการตรวจจับแรงดันลำดับเฟสบวกมูลฐานมีสมรรถนะการตรวจจับฮาร์มอนิกที่ดีกว่า จึงส่งผลให้วงจรกรองกำลังแอกทีฟแบบขนานสามารถชดเชยกำจัดกระแสฮาร์มอนิกในระบบไฟฟ้ากำลังหนึ่งเฟสได้อย่างมีประสิทธิผล นอกจากนี้ค่า %THD ของกระแสที่แหล่งจ่ายภายหลังการชดเชยมีค่าอยู่ในกรอบมาตรฐาน IEEE std. 519-2022 ด้วยเช่นกัน

คำสำคัญ : วงจรกรองกำลังแอกทีฟ การตรวจจับฮาร์มอนิก การกำจัดฮาร์มอนิก การตรวจจับแรงดันลำดับเฟสบวกมูลฐาน วิธีฟูริเยร์เอสดี

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Improvement of Harmonic Detection Using SDF for Shunt Active Power Filter in Single-phase Power Systems

Sutthinee Waosungnern[1] Tosaporn Narongrit[2*] Kongpol Areerak[3]

[1,2,3]PEMC Research Group, School of Electrical Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand*

*Corresponding Author. E-mail address: tosaporn@sut.ac.th

Received: 13 March 2023; Revised: 31 August 2023; Accepted: 28 September 2023
Published online: 27 December 2023

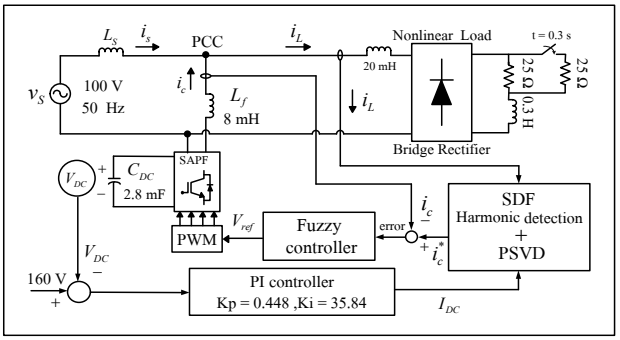
Abstract

This paper presents the performance improvement of harmonic detection using the Synchronous Detection with Fourier analysis (SDF) method to calculate the reference current of a shunt active power filter (SAPF) for single-phase power systems. The Positive Sequence Voltage Detector (PSVD) is applied to the SDF method to improve the accuracy of the reference current calculation in the case of a distorted voltage source. Where the objective is to enhance the effectiveness of harmonic elimination for single-phase power systems. For the harmonic detection testing, the hardware-in-the-loop simulation technique of the Simulink/MATLAB program and the TMS320C2000TM Experimenter Kit DSP board are used to simulate the harmonic elimination system. The simulation is divided into two cases of voltage sources: a pure sinusoidal waveform and a distorted waveform caused by harmonics. The simulation results show that harmonic detection using SDF in cooperation with the PSVD method (SDF+PSVD) can reduce the percentage of the total harmonic distortion (%THD) value of the source current to 3.07% for the case of a pure sinusoidal voltage waveform and to 2.00% for the case of a distorted voltage waveform. These %THD values are lower when compared with the SDF method and synchronous detection (SD). From the result, it confirms that the SDF cooperated with the PSVD method can provide better performance for harmonic detection. Therefore, the shunt active power filter can effectively eliminate current harmonics in the single-phase power system. Moreover, %THD of the source current after compensation is also satisfied under the IEEE std. 519-2022.

***Keywords*:** Active power filter, Harmonic detection, Harmonic elimination, Positive sequence voltage detector, Synchronous detection with Fourier analysis

1) บทนำ

ในปัจจุบันโรงงานอุตสาหกรรม สำนักงาน อาคาร ที่พักอาศัยที่ใช้ระบบไฟฟ้ากำลังหนึ่งเฟส มีการต่อใช้งานอุปกรณ์ไฟฟ้าที่ทำงานไม่เป็นเชิงเส้น (non-linear) เช่น คอมพิวเตอร์ อุปกรณ์สำรองไฟฟ้า วงจรสวิตชิ่งเพาเวอร์ซัพพลาย [1] วงจรชาร์จแบตเตอรี่รถไฟฟ้า [2] วงจรซัพพลายจากแผงโซล่าเซลล์ [3] เป็นต้น ทำให้เกิดฮาร์มอนิกในระบบไฟฟ้า ซึ่งฮาร์มอนิกส่งผลเสียต่อคุณภาพกำลังไฟฟ้าและผลเสียอื่น ๆ หลายประการ เช่น เกิดความร้อนและเกิดกำลังสูญเสียที่สายส่ง กำลังสูญเสียที่หม้อแปลง อุปกรณ์ป้องกัน เบรกเกอร์ ฟิวส์ ทำงานผิดพลาด และอุปกรณ์ไฟฟ้ามีอายุการใช้งานสั้นลง เป็นต้น [1]–[5] ดังนั้นเพื่อเพิ่มคุณภาพกำลังไฟฟ้า จึงจำเป็นต้องกำจัดฮาร์มอนิกหรือทำให้ลดลงอยู่ภายใต้มาตรฐาน ปัจจุบันมีวิธีการกำจัดฮาร์มอนิกที่นิยมใช้ ได้แก่ วงจรกรองกำลังแอกทีฟ (Active Power Filter: APF) วงจรกรองกำลังพาสซีฟ (Passive Power Filter: PPF) และวงจรกรองกำลังไฮบริด (Hybrid Power Filter: HPF) [6] โดยบทความนี้ได้เลือกใช้วงจรกรองกำลังแอกทีฟแบบขนาน (Shunt Active Power Filter: APF) เนื่องจากเป็นวิธีที่มีประสิทธิผลในการกำจัดฮาร์มอนิก สามารถยืดหยุ่นต่อการเปลี่ยนแปลงค่าพารามิเตอร์ของระบบไฟฟ้า และไม่ทำให้เกิดปัญหาเรโซแนนซ์ [7], [8]



รูปที่ 1 : ระบบกำจัดฮาร์มอนิกด้วยวงจรกรองกำลังแอกทีฟแบบขนานในระบบไฟฟ้ากำลังหนึ่งเฟส

วงจรกรองกำลังแอกทีฟแบบขนานในระบบที่พิจารณากำจัดฮาร์มอนิกดังรูปที่ 1 [9], [10] นอกจากนี้วงจรกรองกำลังแอกทีฟแบบขนานมีส่วนประกอบที่สำคัญ คือ การตรวจจับฮาร์มอนิกและระบบควบคุม [11] โดยการตรวจจับฮาร์มอนิกจะทำหน้าที่คำนวณหากระแสฮาร์มอนิกสำหรับใช้เป็นกระแสอ้างอิงให้กับวงจรกรองกำลังแอกทีฟแบบขนาน ดังนั้นการตรวจจับฮาร์มอนิกจึงจำเป็นต้องมีความถูกต้องแม่นยำในการคำนวณ โดยในอดีตที่ผ่านมาได้มีการนำเสนอเปรียบเทียบสมรรถนะการตรวจจับฮาร์มอนิกของวิธีการตรวจจับซิงโครนัส (Synchronous Detection: SD) กับวิธีฟูริเยร์เอสดี (Synchronous Detection with Fourier analysis: SDF) ซึ่งการตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดีถูกพัฒนามาจากวิธีการตรวจจับซิงโครนัส ด้วยการนำการเทคนิคการวิเคราะห์ฟูริเยร์แบบวินโดว์เลื่อน (Sliding Window with Fourier Analysis: SWFA) มาทำหน้าที่แทนการใช้ตัวกรอง LPF ในขั้นตอนที่ 3 ดังนั้นจึงมีรูปแบบการคำนวณที่ง่ายและมีขั้นตอนการคำนวณขั้นอื่น ๆ เหมือนกัน ซึ่งจากผลการทดสอบพบว่าวิธีฟูริเยร์เอสดีมีสมรรถนะในการคำนวณตรวจจับฮาร์มอนิกที่ถูกต้องแม่นยำกว่าวิธีการตรวจจับซิงโครนัส [12] ดังนั้นบทความนี้จึงเลือกใช้วิธีฟูริเยร์เอสดี อย่างไรก็ตามวิธีดังกล่าวมีข้อจำกัดการคำนวณในกรณีที่ระบบไฟฟ้ามีแรงดันที่แหล่งจ่ายไม่เป็นรูปคลื่นไซน์บริสุทธิ์ [13], [14] ซึ่งจะทำให้ความถูกต้องในการตรวจจับฮาร์มอนิกลดลง ด้วยเหตุนี้บทความจึงนำเสนอการปรับปรุงการตรวจจับฮาร์มอนิก ด้วยการนำวิธีการตรวจจับแรงดันลำดับเฟสบวกมูลฐาน (Positive Sequence Voltage Detector: PSVD) [15]–[17] มาทำงานร่วมกับวิธีฟูริเยร์เอสดีเพื่อให้การคำนวณตรวจจับฮาร์มอนิกหรือการคำนวณกระแสอ้างอิงมีความถูกต้องมากยิ่งขึ้น ซึ่งส่งผลให้วงจรกรองกำลังแอกทีฟแบบขนานมีประสิทธิผลการกำจัดฮาร์มอนิกเพิ่มขึ้น สำหรับระบบควบคุมการทำงานของวงจรกรองกำลังแอกทีฟแบบขนานจะประกอบไปด้วย ระบบควบคุมกระแสชดเชยทำหน้าที่ควบคุมกระแสชดเชยให้คล้อยตามกระแสอ้างอิง ซึ่งบทความนี้เลือกใช้ตัวควบคุมฟัซซี เนื่องจากเป็นตัวควบคุมที่ให้สมรรถนะที่ดีในการควบคุมกระแสชดเชย [18], [19] และระบบควบคุมแรงดันบัสไฟตรง ทำหน้าที่ควบคุมแรงดันให้มีค่าคงที่ตามค่าแรงดันอ้างอิงที่ได้ออกแบบไว้ โดยบทความนี้ได้เลือกใช้ตัวควบคุมพีไอ เนื่องจากมีสมรรถนะที่ดีเพียงพอในการควบคุมค่าแรงดันกระแสตรง [20]

สำหรับการนำเสนอของบทความนี้จะประกอบไปด้วย ในหัวข้อที่ 2 จะนำเสนอการตรวจฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดี หัวข้อที่ 3 จะนำเสนอการปรับปรุงการตรวจฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดีที่ทำงานร่วมกับวิธี PSVD หัวข้อที่ 4 จะนำเสนอการควบคุมกระแสชดเชยด้วยตัวควบคุมฟัซซี หัวข้อที่ 5 จะนำเสนอการจำลองสถานการณ์การกำจัดฮาร์มอนิกด้วยเทคนิคฮาร์ดแวร์ในลูป หัวข้อที่ 6 จะนำเสนอผลการจำลองสถานการณ์และการ

2) การตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดี

การตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดีหรือวิธี SDF เป็นวิธีที่สามารถคำนวณหากระแสอ้างอิงได้อย่างถูกต้องแม่นยำ และมีรูปแบบการคำนวณที่ง่าย ซึ่งสามารถอธิบายรายละเอียดขั้นตอนการคำนวณได้ดังนี้ [12]–[14]

<u>ขั้นที่ 1</u> : แปลงแรงดันที่แหล่งจ่าย (v_S) และกระแสโหลด (i_L) จากแกนเฟสเป็นแรงดันที่แหล่งจ่ายบนแกน $\alpha\beta$ ($v_{S,\alpha}, v_{S,\beta}$) และกระแสโหลดบนแกน $\alpha\beta$ ($i_{L,\alpha}, i_{L,\beta}$) แสดงดังสมการที่ (1) และ (2) ตามลำดับ

$$\begin{bmatrix} v_{S,\alpha} \\ v_{S,\beta} \end{bmatrix} = \begin{bmatrix} v_S(\theta) \\ v_S(\theta - \frac{\pi}{2}) \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} i_{L,\alpha} \\ i_{L,\beta} \end{bmatrix} = \begin{bmatrix} i_L(\theta) \\ i_L(\theta - \frac{\pi}{2}) \end{bmatrix} \tag{2}$$

<u>ขั้นที่ 2</u> : คำนวณหาค่ากำลังแอกทีฟ (p) แสดงได้ดังสมการที่ (3)

$$p = v_{S,\alpha} i_{L,\alpha} \times v_{S,\beta} i_{L,\beta} \tag{3}$$

<u>ขั้นที่ 3</u> : เนื่องจากค่ากำลังแอกทีฟประกอบด้วยปริมาณกำลังแอกทีฟมูลฐาน (P_{dc}) และปริมาณกำลังแอกทีฟฮาร์มอนิก (P_{ac}) แสดงดังสมการที่ (4) จึงนำค่ากำลังแอกทีฟมาทำการแยกปริมาณ P_{dc} ออกจากปริมาณ P_{ac} ด้วยการใช้เทคนิคการวิเคราะห์ฟูริเยร์แบบวินโดว์เลื่อน (SWFA) [12]

$$p = P_{dc} + P_{ac} \tag{4}$$

<u>ขั้นที่ 4</u> คำนวณหาค่ายอดกระแสที่แหล่งจ่ายมูลฐาน (I_s) แสดงดังสมการที่ (5)

$$I_S = \left[\frac{P_{dc}}{V_S} \right] + I_{DC} \tag{5}$$

โดยที่ V_S คือ ค่ายอดแรงดันที่แหล่งจ่าย

I_{DC} คือ ค่ากระแสตรงที่ได้จากเอาต์พุตของตัวควบคุมพีไอสำหรับควบคุมแรงดันบัสไฟตรง (V_{DC})

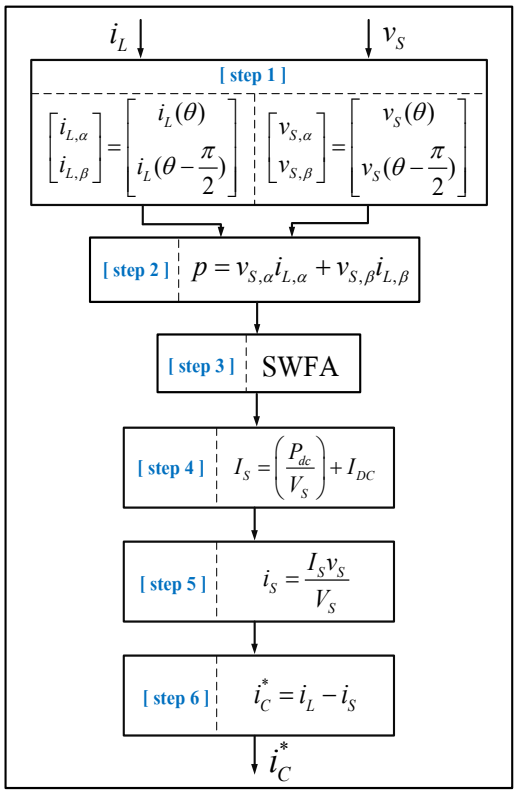
<u>ขั้นที่ 5</u> คำนวณค่ากระแสที่แหล่งจ่ายมูลฐาน (i_s) ด้วยสมการที่ (6)

$$i_S = \frac{I_S v_S}{V_S} \tag{6}$$

<u>ขั้นที่ 6</u> คำนวณหากระแสอ้างอิง (i_C^*) สำหรับวงจรกรองกำลังแอกทีฟแบบขนาน แสดงดังสมการที่ (7)

$$i_c^* = i_L - i_s \tag{7}$$

จากขั้นตอนการคำนวณหากระแสอ้างอิงของวิธี SDF ในข้างต้นสามารถสรุปเป็นแผนภาพแสดงดังรูปที่ 2



รูปที่ 2 : แผนภาพการตรวจจับฮาร์มอนิกด้วยวิธี SDF

3) การปรับปรุงการตรวจจับฮาร์มอนิกด้วยวิธีฟูริเยร์เอสดี

การตรวจจับฮาร์มอนิกด้วยวิธี SDF เป็นวิธีที่สามารถคำนวณกระแสอ้างอิงได้อย่างแม่นยำ อย่างไรก็ตามวิธี SDF ยังมีข้อจำกัดในกรณีที่แรงดันที่แหล่งจ่ายของระบบไฟฟ้ามีสัญญาณที่ไม่เป็นรูปไซน์บริสุทธิ์ ซึ่งจะทำให้การตรวจจับฮาร์มอนิกไม่ถูกต้องและจะส่งผลต่อการกำจัดฮาร์มอนิกของวงจรกรองกำลังแอกทีฟแบบขนาน ด้วยเหตุนี้จึงทำการปรับปรุงการตรวจจับฮาร์มอนิกด้วยการนำวิธีการตรวจจับแรงดันลำดับเฟสบวกมูลฐานหรือวิธี PSVD มาทำงานร่วมกับวิธี SDF เพื่อแก้ปัญหาดังกล่าว โดยสามารถแสดงรายละเอียดขั้นตอนการคำนวณของวิธี PSVD [15]–[17] ได้ดังนี้

<u>ขั้นที่ 1</u> : แปลงแรงดันที่แหล่งจ่าย (v_S) จากแกนเฟสให้อยู่บนแกน $\alpha\beta$ ($v_{S,\alpha}, v_{S,\beta}$) แสดงดังสมการที่ (1)

<u>ขั้นที่ 2</u> : คำนวณหาค่ามุมของแรงดันที่แหล่งจ่าย (θ_{v_S}) ด้วยวงจรเฟสล็อกลูป (Phase Lock Loop: PLL) [16]

<u>ขั้นที่ 3</u> : คำนวณหาค่ากระแสบนแกน $\alpha\beta$ (i_α, i_β) ด้วยสมการที่ (8)

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \sqrt{\frac{3}{2}} \begin{bmatrix} \sin\theta_{v_S} \\ -\cos\theta_{v_S} \end{bmatrix} \quad (8)$$

<u>ขั้นที่ 4</u> : คำนวณหาค่ากำลังไฟฟ้า (p', q') ตามสมการที่ (9)

$$\begin{bmatrix} p' \\ q' \end{bmatrix} = \begin{bmatrix} v_{S,\alpha} \cdot i_\alpha & v_{S,\beta} \cdot i_\beta \\ v_{S,\beta} \cdot i_\alpha & v_{S,\alpha} \cdot i_\beta \end{bmatrix} \quad (9)$$

<u>ขั้นที่ 5</u> : เนื่องจากค่ากำลังไฟฟ้าประกอบด้วยปริมาณกำลังมูลฐาน (p'_{dc}, q'_{dc}) และปริมาณกำลังฮาร์มอนิก (p'_{ac}, q'_{ac}) จึงทำการแยกปริมาณ p'_{dc} และ q'_{dc} ออกจากปริมาณ p'_{ac} และ q'_{ac} ด้วยการใช้วงจรกรองผ่านต่ำ (Low Pass Filter: LPF)

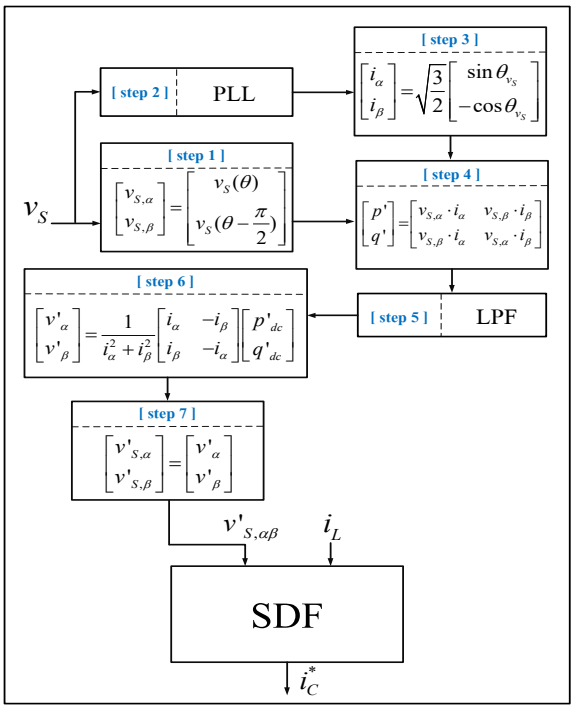
<u>ขั้นที่ 6</u> : คำนวณหาค่าแรงดันสำหรับอ้างอิงที่อยู่บนแกน $\alpha\beta$ (v'_α, v'_β) ด้วยสมการที่ (10)

$$\begin{bmatrix} v'_\alpha \\ v'_\beta \end{bmatrix} = \frac{1}{i_\alpha^2 + i_\beta^2} \begin{bmatrix} i_\alpha & -i_\beta \\ i_\beta & -i_\alpha \end{bmatrix} \begin{bmatrix} p'_{dc} \\ q'_{dc} \end{bmatrix} \quad (10)$$

<u>ขั้นที่ 7</u> : กำหนดค่า $v'_{S,\alpha}$ และ $v'_{S,\beta}$ มีค่าเท่ากับ v'_α และ v'_β ตามลำดับ ดังสมการที่ (11) เพื่อนำไปใช้เป็นอินพุตแทนค่า $v_{S,\alpha}, v_{S,\beta}$ ในขั้นตอนแรกของการคำนวณหาค่ากระแสอ้างอิงด้วยวิธี SDF ต่อไป

$$\begin{bmatrix} v'_{S,\alpha} \\ v'_{S,\beta} \end{bmatrix} = \begin{bmatrix} v'_\alpha \\ v'_\beta \end{bmatrix} \quad (11)$$

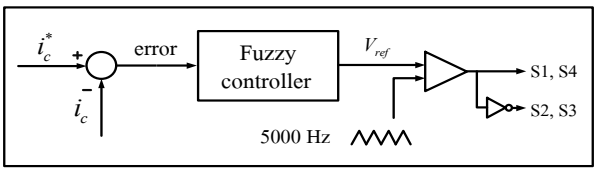
จากขั้นตอนการคำนวณวิธี PSVD ในข้างต้นสามารถสรุปเป็นแผนภาพการตรวจจับฮาร์มอนิกด้วยวิธี SDF ร่วมกับวิธี PSVD (SDF + PSVD) แสดงได้ดังรูปที่ 3



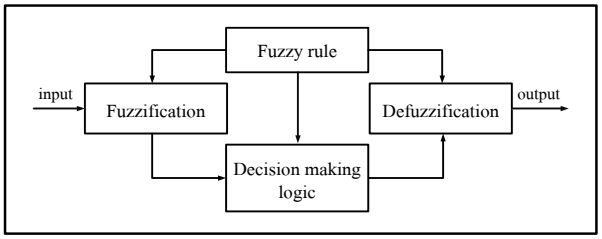
รูปที่ 3 : แผนภาพการตรวจจับฮาร์มอนิกด้วยวิธี SDF ร่วมกับวิธี PSVD (SDF + PSVD)

4) การควบคุมกระแสชดเชยด้วยตัวควบคุมฟัซซี

การควบคุมกระแสชดเชยด้วยตัวควบคุมฟัซซีซึ่งเป็นตัวควบคุมที่สามารถทำงานในระบบที่มีความคลุมเครือเหมาะสำหรับควบคุมกระแสชดเชยที่มีลักษณะสัญญาณเป็นฮาร์มอนิกโดยบล็อกไดอะแกรมตัวควบคุมฟัซซีสามารถแสดงดังรูปที่ 4 จากรูปดังกล่าวพบว่าสัญญาณอินพุต คือ ค่าผลต่างระหว่างกระแสอ้างอิง (i_C^*) กับกระแสชดเชย (i_C) หรือค่าความผิดพลาด (error) และสัญญาณเอาต์พุต คือ ค่าแรงดันอ้างอิง (V_{ref}) สำหรับเป็นอินพุตให้กับเทคนิคการสวิตช์แบบ PWM (Pulse Width Modulation) เพื่อสร้างพัลส์ควบคุมการทำงานของสวิตช์ไอจีบีที (IGBT) ของวงจรกรองกำลังแอกทีฟแบบขนานต่อไป [19]



รูปที่ 4 : แผนภาพบล็อกระบบควบคุมกระแสชดเชยด้วยตัวควบคุมฟัซซี



รูปที่ 5 : กระบวนการของตัวควบคุมฟัซซี

ตัวควบคุมฟัซซีมีกระบวนการควบคุมดังรูปที่ 5 โดยสามารถอธิบายการออกแบบส่วนต่าง ๆ ได้ดังนี้ [18]

4.1) การทำฟัซซี (Fuzzification) บทความนี้ออกแบบให้ตัวแปรภาษาของฟังก์ชันสมาชิกอินพุตเป็นค่า error ที่มีค่าเชิงภาษา 3 ค่า ได้แก่ “neg” (มีค่าเป็นลบ) “zero” (มีค่าเป็นศูนย์) และ “pos” (มีค่าเป็นบวก) โดยขอบเขตฟังก์ชันสมาชิกอินพุตกำหนดใช้เท่ากับ 0.1 A ดังในแสดงรูปที่ 6 และออกแบบฟังก์ชันสมาชิกของเอาต์พุตเป็นค่า V_{ref} ที่ค่าเชิงภาษา 3 ค่า ได้แก่ “dec” (มีค่าลดลง) “cons” (มีค่าคงที่) และ “inc” (มีค่าเพิ่มขึ้น) โดยกำหนดขอบเขตฟังก์ชันสมาชิกเอาต์พุตกำหนดใช้เท่ากับ 160 V ตามค่า V_{DC} ดังแสดงในรูปที่ 7

4.2) กฎฟัซซี (Fuzzy Rules) ได้ทำการออกแบบกฎฟัซซีสำหรับควบคุมกระแสชดเชย 3 ข้อดังนี้

กฎข้อที่ 1 IF error = neg THEN voltage = dec

กฎข้อที่ 2 IF error = zero THEN voltage = cons

กฎข้อที่ 3 IF error = pos THEN voltage = inc

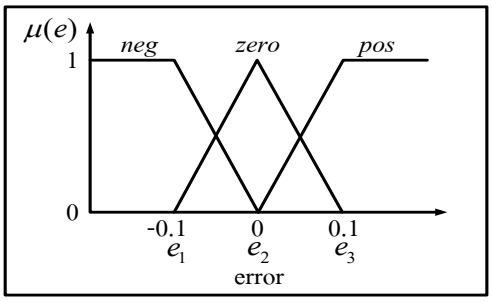
4.3) ลอจิกเพื่อการตัดสินใจ (Decision Making Logic) เลือกใช้การอนุมานแบบ Mamdani

4.4) การทำดีฟัซซี (Defuzzification) ได้เลือกใช้วิธีหาจุดศูนย์ถ่วง (Center of Gravity: COG) ซึ่งสามารถคำนวณหาค่า V_{ref} ได้จากสมการที่ (12)

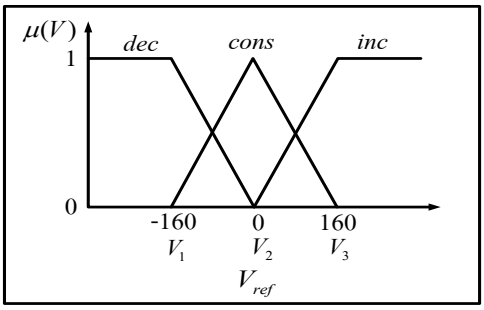
$$V_{ref} = \frac{\sum_{x=a}^{b} \mu(V_x)V_x}{\sum_{x=a}^{b} \mu(V_x)} \qquad (12)$$

โดยที่ $\mu(V_x)$ คือ ค่าความเป็นสมาชิกของเอาต์พุตสมาชิก x ใด ๆ

a,b คือ ขอบเขตฟังก์ชันสมาชิกเอาต์พุต



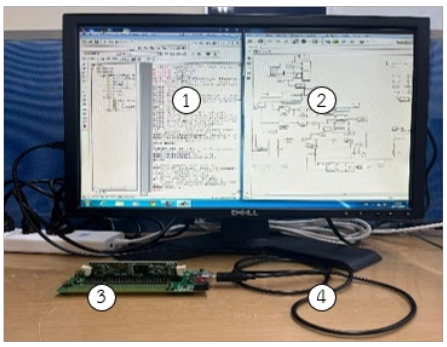
รูปที่ 6 : ตำแหน่งฟังก์ชันสมาชิกอินพุต error



รูปที่ 7 : ตำแหน่งฟังก์ชันสมาชิกเอาต์พุต V_{ref}

5) การจำลองสถานการณ์การกำจัดฮาร์มอนิกด้วยเทคนิคฮาร์ดแวร์ในลูป

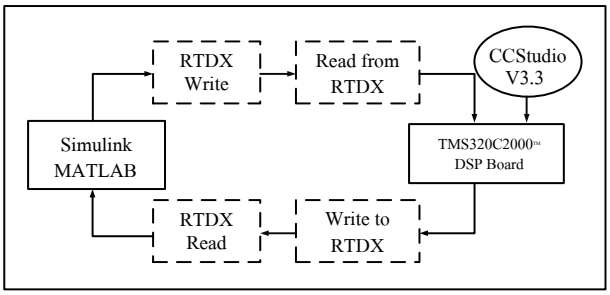
การจำลองสถานการณ์การกำจัดฮาร์มอนิกของวงจรกรองกำลังแอกทีฟแบบขนานด้วยเทคนิคฮาร์ดแวร์ในลูป (Hardware In Loop: HIL) เป็นการทำงานร่วมกันระหว่างโปรแกรม Simulink/MATLAB และบอร์ด DSP รุ่น TMS320C2000™ Experimental Kit โดยแลกเปลี่ยนข้อมูลผ่านสาย USB JTAG emulation ดังแสดงในรูปที่ 8



1. CCStudio v3.3
2. Simulink/MATLAB program
3. TMS320C2000TM Experimental Kit
4. USB JTAG emulation

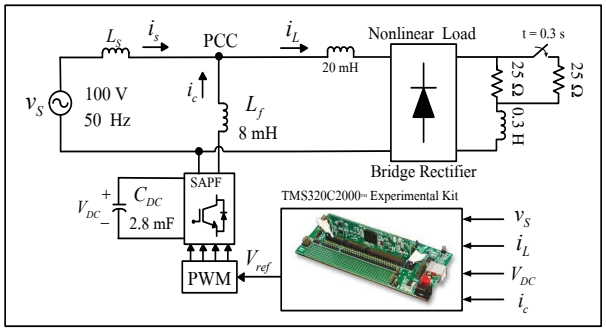
รูปที่ 8 : การเชื่อมต่อระหว่างโปรแกรม Simulink/MATLAB กับบอร์ด DSP รุ่น TMS320C2000™ Experimental Kit

กระบวนการการทำงานของเทคนิคฮาร์ดแวร์ในลูปสามารถอธิบายได้ดังรูปที่ 9 โดยโปรแกรม Simulink/MATLAB จะส่งข้อมูลผ่านบล็อก RTDX Write ไปยังบอร์ด DSP ขณะเดียวกันบอร์ด DSP สามารถรับข้อมูลด้วยคำสั่ง Read from RTDX จากนั้นข้อมูลจะถูกนำไปคำนวณตามโปรแกรมที่เขียนลงใน CCStudio V3.3 และส่งข้อมูลออกจากบอร์ด DSP ผ่านคำสั่ง Write to RTDX ซึ่งโปรแกรม Simulink/MATLAB สามารถรับข้อมูลด้วยบล็อก RTDX Read จากกระบวนการดังกล่าว คือ การคำนวณครบหนึ่งรอบช่วงเวลาชักตัวอย่าง (sampling time) และจะทำงานวนซ้ำอย่างนี้ไปเรื่อย ๆ ตลอดระยะเวลากำหนดจำลองสถานการณ์



รูปที่ 9 : แผนภาพการทำงานของระบบฮาร์ดแวร์ในลูป

ระบบจำลองสถานการณ์การกำจัดฮาร์มอนิกในรูปที่ 1 สามารถสร้างเป็นระบบจำลองด้วยเทคนิคฮาร์ดแวร์ในลูปได้ดังรูปที่ 10 โดยส่วนของการตรวจจับฮาร์มอนิกการควบคุมแรงดันบัสไฟตรง และการควบคุมกระแสชดเชยจะถูกโปรแกรมลงบนบอร์ด DSP ส่วนระบบไฟฟ้ากำลังหนึ่งเฟส และวงจรกรองกำลังแอกทีฟแบบขนานจะถูกสร้างจากชุดบล็อกบนโปรแกรม Simulink/MATLAB โดยที่ข้อมูลอินพุตของบอร์ด DSP คือ ค่าแรงดันที่แหล่งจ่าย (v_S) ค่ากระแสโหลด (i_L) ค่าแรงดันบัสไฟตรง (V_{DC}) และค่ากระแสชดเชย (i_C) ของวงจรกรองกำลังแอกทีฟแบบขนาน และข้อมูลเอาต์พุตของบอร์ด DSP คือ ค่าแรงดันอ้างอิง (V_{ref}) สำหรับส่งไปยังเทคนิค PWM บนโปรแกรม Simulink/MATLAB เพื่อสร้างพัลส์ควบคุมสวิตซ์ไอจีบีที (IGBT) ของวงจรกรองกำลังแอกทีฟ



รูปที่ 10 : ระบบจำลองสถานการณ์การกำจัดฮาร์มอนิกด้วยฮาร์ดแวร์ในลูป

6) ผลการจำลองสถานการณ์และการอภิปรายผล

การจำลองสถานการณ์การกำจัดฮาร์มอนิกในระบบไฟฟ้ากำลังหนึ่งเฟสด้วยเทคนิคฮาร์ดแวร์ในลูปดังรูปที่ 10 เพื่อทดสอบเปรียบเทียบสมรรถนะของการตรวจจับฮาร์มอนิกที่ใช้วิธี SD วิธี SDF และการตรวจจับฮาร์มอนิกที่ปรับปรุงขึ้นเรียกว่าวิธี SDF + PSVD จะแบ่งการทดสอบเป็น 2 กรณี ได้แก่ กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์ และกรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน โดยการจำลองสถานการณ์การกำจัดฮาร์มอนิกจะพิจารณาโหลดในสภาวะปกติ และสภาวะที่โหลดมีเปลี่ยนแปลงด้วยการเพิ่มขนาดกระแส

กรณีที่ 1 กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์ ผลการจำลองสถานการณ์ที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD สามารถแสดงดังรูปที่ 11 วิธี SDF สามารถแสดงดังรูปที่ 12 และที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SDF + PSVD แสดงดังรูปที่ 13 จากทั้งสามรูปดังกล่าวในช่วงเวลาตั้งแต่ 0 ถึง 0.1 s เป็นช่วงเวลาก่อนการฉีดกระแสชดเชย (i_C) พบว่ากระแสที่แหล่งจ่าย (i_S) มีลักษณะสัญญาณผิดเพี้ยนไม่เป็นรูปไซน์เช่นเดียวกับสัญญาณของกระแสโหลด (i_L) เมื่อคำนวณค่า %THD ของกระแสที่แหล่งจ่ายดังกล่าวพบว่ามีค่าเท่ากับ 26.9% จากนั้นเมื่อทำการชดเชยฮาร์มอนิกตั้งแต่เวลา 0.1 s เป็นต้นไป จะเห็นได้ว่าวงจรกรองกำลังแอกทีฟแบบขนานที่ใช้ตัวควบคุม ฟัซซีสามารถฉีดกระแสชดเชยให้มีลักษณะคล้อยตามกระแสอ้างอิง (i_C^*) ได้ ทำให้สัญญาณของกระแสที่แหล่งจ่ายกลับมาเป็นรูปไซน์ โดยค่า %THD ของกระแสที่แหล่งจ่ายหลังการชดเชยที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD (รูปที่ 11) ในสภาวะโหลดปกติ (ช่วงเวลา 0.1 ถึง 0.3 s) มีค่าเท่ากับ 4.04% และในสภาวะโหลดมีการเปลี่ยนแปลงเพิ่มขนาด (ช่วงเวลา 0.3 ถึง 0.5 s) มีค่าเท่ากับ 5.51% สำหรับที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SDF (รูปที่ 12) ในสภาวะโหลดปกติ มีค่าเท่ากับ 1.89% และในสภาวะโหลดมีการเปลี่ยนแปลงเพิ่มขนาด มีค่าเท่ากับ 3.08% และผลการตรวจจับฮาร์มอนิกด้วยวิธี SDF + PSVD (รูปที่ 13) พบว่าค่า %THD ในสภาวะโหลดปกติเท่ากับ 1.88% และในสภาวะโหลดมีการเปลี่ยนแปลงเท่ากับ 3.07%

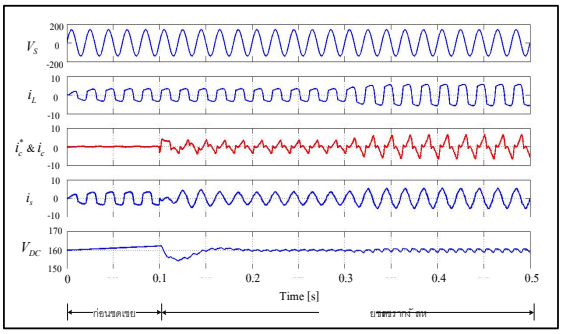
กรณีที่ 2 กรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปนโดยมีค่า %THD เท่ากับ 9.11% [10] ผลการจำลองสถานการณ์การกำจัดฮาร์มอนิกที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD สามารถแสดงดังรูปที่ 14 วิธี SDF สามารถแสดงดังรูปที่ 15 และที่ใช้วิธี SDF + PSVD แสดงได้ดังรูปที่ 16 ซึ่งพบว่าในช่วงเวลาก่อนการฉีดกระแสชดเชย (ตั้งแต่เวลา 0 ถึง 0.1 s) กระแสที่แหล่งจ่ายมีความเพี้ยนที่ค่า %THD เท่ากับ 29.08% แต่เมื่อวงจรกรองกำลังแอกทีฟแบบขนานร่วมกับตัวควบคุมฟัซซีทำการฉีดกระแสชดเชย (ตั้งแต่เวลา 0.1 s เป็นต้นไป) พบว่าค่า %THD ของกระแสที่แหล่งจ่ายภายหลังการชดเชยที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD (รูปที่ 14) ในสภาวะโหลดปกติ และโหลดมีการเปลี่ยนแปลงเพิ่มขนาด มีค่าเท่ากับ 11.32% และ 13.52% ตามลำดับ ส่วนเมื่อใช้วิธี SDF (รูปที่ 15) ในสภาวะโหลดปกติและโหลดมีการเปลี่ยนแปลงเพิ่มขนาด มีค่าเท่ากับ 10.87% และ 12.25% ตามลำดับ และวิธี SDF + PSVD (รูปที่ 16) พบว่าค่า %THD ที่โหลดปกติ และโหลดมีการเปลี่ยนแปลงเพิ่มขนาดมีค่าเท่ากับ 2.34% และ 2.00% ตามลำดับ โดยสามารถแสดงสรุปผลค่า %THD ของกระแสที่แหล่งจ่ายได้ดังตารางที่ 1

นอกจากนี้ จากรูปที่ 11 ถึง 16 แสดงให้เห็นว่าตัวควบคุมพีไอมีสมรรถนะที่สามารถควบคุมแรงดันบัสไฟตรง (V_{DC}) ให้มีค่าคงที่เท่ากับ 160 V

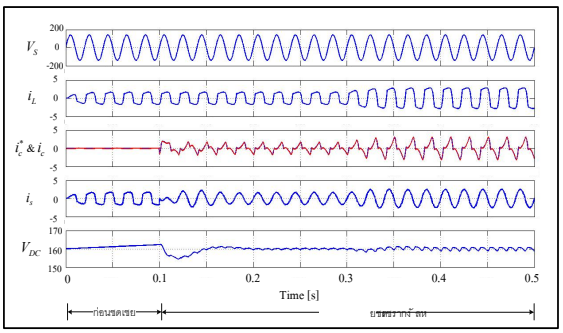
ตารางที่ 1 : ผลการกำจัดฮาร์มอนิกด้วยฮาร์ดแวร์ในลูป

การตรวจจับฮาร์มอนิก	**แรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์**		**แรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน**	
	%THDi	%THDv	%THDi	%THDv
ก่อนกระแสชดเชย				
วิธี SD	26.9%	0%	29.08%	9.11%
วิธี SDF				
วิธี SDF+PSVD				
หลังกระแสชดเชย กรณีโหลดปกติ				
วิธี SD	4.04%	0%	11.32%	9.11%
วิธี SDF	1.89%		10.87%	
วิธี SDF+PSVD	1.88%		2.34%	
หลังกระแสชดเชย กรณีโหลดมีการเปลี่ยนแปลง				
วิธี SD	5.51%	0%	13.52%	9.11%
วิธี SDF	3.08%		12.25%	
วิธี SDF+PSVD	3.07%		2.00%	

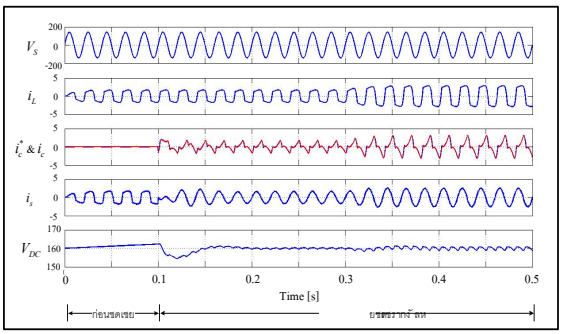
จากตารางที่ 1 ในกรณีการทดสอบแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์พบว่าค่า %THD ของกระแสที่แหล่งจ่ายหลังการชดเชยที่ใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD โหลดมีการเปลี่ยนแปลงมีค่าเกินมาตรฐาน IEEE std. 519-2022 (< 5%) ส่วนวิธี SDF และวิธี SDF + PSVD มีค่าประมาณเท่ากันและอยู่ในกรอบมาตรฐาน IEEE ทั้งช่วงโหลดปกติและโหลดมีการเปลี่ยนแปลง แสดงว่าวิธี SDF และวิธี SDF + PSVD มีสมรรถนะการคำนวณหากระแสอ้างอิงที่ดีกว่าวิธี SD แบบดั้งเดิม อย่างไรก็ตามในกรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน พบว่าการคำนวณตรวจจับฮาร์มอนิกของวิธี SDF + PSVD สามารถคำนวณได้ถูกต้องแม่นยำมากกว่าที่ใช้วิธี SD และวิธี SDF โดยพิจารณาจากค่า %THD ของกระแสที่แหล่งจ่ายที่มีค่าน้อยกว่าและอยู่ภายใต้มาตรฐานในทุกช่วงโหลด



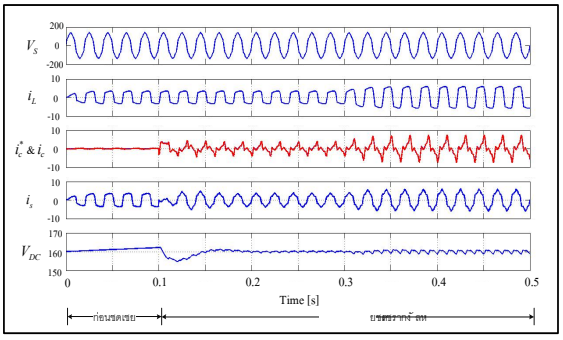
รูปที่ 11 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SD กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์



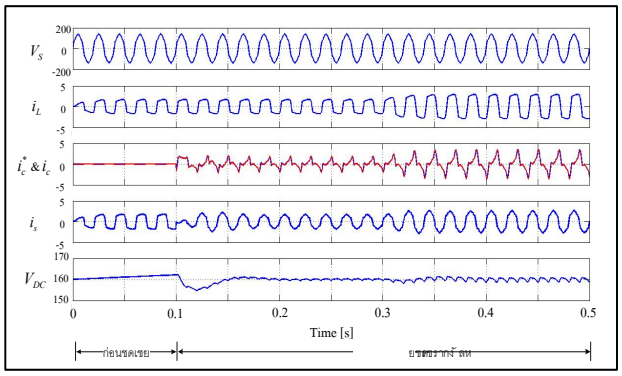
รูปที่ 12 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SDF กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์



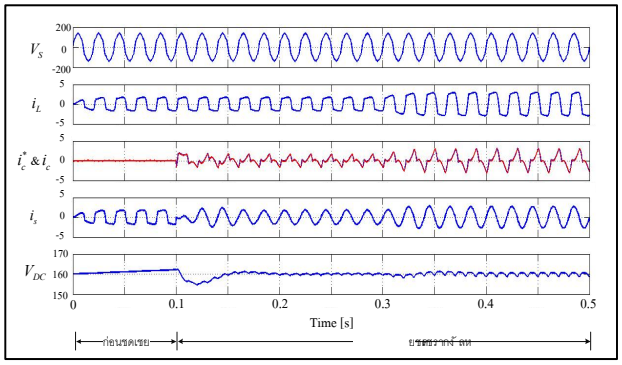
รูปที่ 13 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SDF + PSVD กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์



รูปที่ 14 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SD กรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน

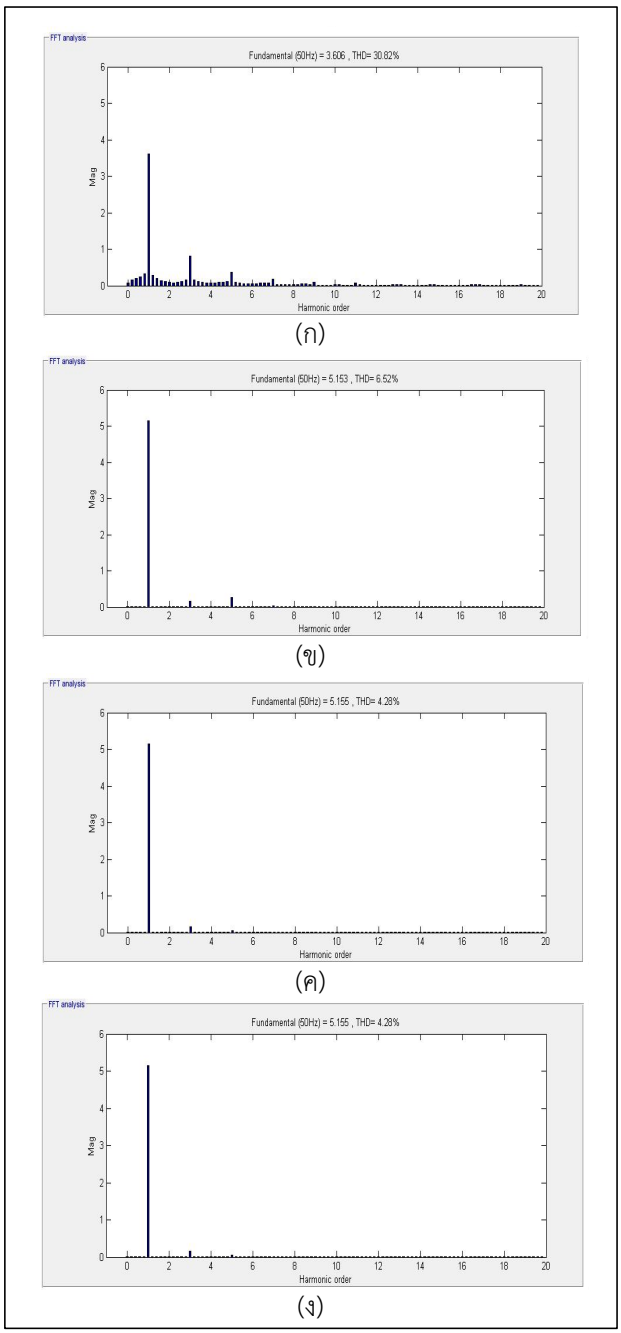


รูปที่ 15 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SDF กรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน

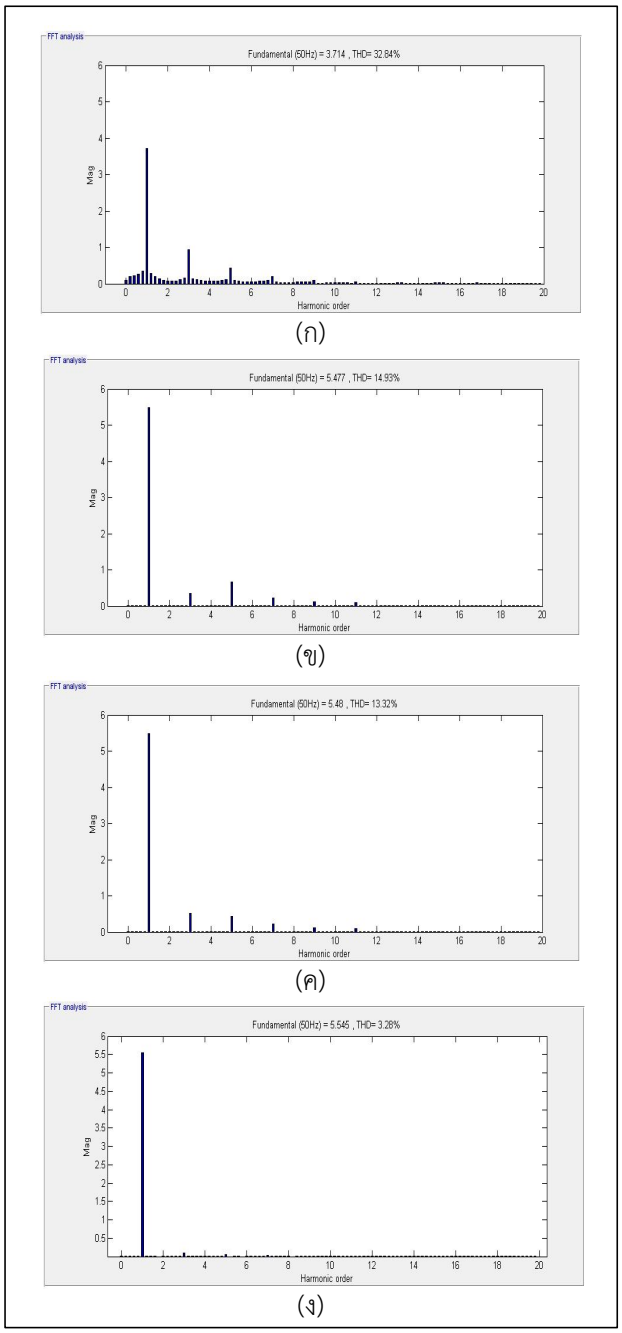


รูปที่ 16 : ผลการกำจัดฮาร์มอนิกที่ใช้วิธี SDF + PSVD กรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน

นอกจากนี้เมื่อพิจารณาผลจากการวิเคราะห์ FFT ของรูปสัญญาณกระแสที่แหล่งจ่าย (i_S) ในกรณีที่แรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์ดังรูปที่ 17 พบว่าสเปกตรัมของรูปสัญญาณกระแสที่แหล่งจ่ายก่อนการชดเชย (รูปที่ 17 (ก)) มีปริมาณฮาร์มอนิกอันดับที่ 3, 5 และ 7 ปรากฏอย่างชัดเจน อย่างไรก็ตามภายหลังการชดเชยพบว่าวงจรกรองกำลังแอกทีฟแบบขนานที่ใช้การตรวจจับฮาร์มอนิกของวิธี SD (รูปที่ 17 (ข)) วิธี SDF (รูปที่ 17 (ค)) และวิธี SDF+ PSVD (รูปที่ 17 (ง)) มีปริมาณฮาร์มอนิกลดน้อยลงมากโดยปรากฏเหลือสเปกตรัมอันดับ 3 และ 5 เพียงเล็กน้อย ส่วนในกรณีที่แรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปนสามารถแสดงสเปกตรัมของรูปสัญญาณกระแสที่แหล่งจ่ายได้ดังรูปที่ 18 โดยก่อนการชดเชย (รูปที่ 18 (ก)) ปรากฏปริมาณฮาร์มอนิกอันดับที่ 3 5 และ 7 ส่วนภายหลังการชดเชยพบว่ากรณีใช้การตรวจจับฮาร์มอนิกด้วยวิธี SD (รูปที่ 18 (ข)) และวิธี SDF (รูปที่ 18 (ค)) ยังคงพบปริมาณฮาร์มอนิกอันดับ 3 5 และ 7 เหลืออยู่ ในขณะที่วิธี SDF + PSVD (รูปที่ 18 (ง)) แทบไม่ปรากฏปริมาณฮาร์มอนิก ซึ่งหมายความว่าการตรวจจับฮาร์มอนิกทั้ง 3 วิธีมีสมรรถนะที่ดีในระบบที่แรงดันที่แหล่งจ่ายเป็นรูปไซน์ ส่วนในกรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน วิธี SDF + PSVD สามารถลดความผิดพลาดการคำนวณหากระแสอ้างอิงได้ทำให้มีความแม่นยำมากกว่าวิธี SD และวิธี SDF ส่งผลให้วงจรกรองกำลังแอกทีฟแบบขนานสามารถกำจัดฮาร์มอนิกได้อย่างมีประสิทธิผลที่ดีกว่า



รูปที่ 17 : ผลการวิเคราะห์ FFT กรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์ (ก) ก่อนการชดเชย (ข) วิธี SD (ค) วิธี SDF (ง) วิธี SDF+PSVD



รูปที่ 18 : ผลการวิเคราะห์ FFT กรณีแรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน (ก) ก่อนการชดเชย (ข) วิธี SD (ค) วิธี SDF (ง) วิธี SDF+PSVD

7) สรุป

บทความนี้ได้นำเสนอการปรับปรุงการตรวจจับฮาร์มอนิกของวิธี SDF ด้วยการนำวิธี PSVD เข้ามาช่วยคำนวณแรงดันมูลฐานลำดับเฟสบวกเพื่อใช้เป็นอินพุตในการคำนวณหากระแสอ้างอิงให้กับวงจรกรองกำลังแอกทีฟแบบขนานในระบบไฟฟ้ากำลังหนึ่งเฟสสำหรับกรณีที่แรงดันที่แหล่งจ่ายมีฮาร์มอนิกปะปน เพื่อให้การคำนวณหากระแสอ้างอิงมีความถูกต้องมากยิ่งขึ้น โดยผลการจำลองสถานการณ์ด้วยเทคนิคฮาร์ดแวร์ในลูปพบว่าการตรวจจับฮาร์มอนิกด้วยวิธี SDF ร่วมกับวิธี PSVD สามารถคำนวณกระแสอ้างอิงได้อย่างถูกต้องแม่นยำมากขึ้น เมื่อเปรียบเทียบกับการใช้วิธี SD และวิธี SDF ส่งผลให้วงจรกรองกำลังแอกทีฟแบบขนานสามารถกำจัดฮาร์มอนิกได้อย่างมีประสิทธิผลทั้งในกรณีแรงดันที่แหล่งจ่ายเป็นรูปไซน์บริสุทธิ์และมีฮาร์มอนิกปะปน โดยค่า %THD ของกระแสที่แหล่งจ่ายหลังการชดเชยมีค่าลดลงและอยู่ภายใต้กรอบมาตรฐาน IEEE std. 519-2022

REFERENCES

[1] V. E. Wagner *et al.*, “Effects of harmonics on equipment,” *IEEE Trans. Power Delivery*, vol. 8, no. 2, pp. 672–680, Apr. 1993, doi: 10.1109/61.216874.

[2] L. Li, B. Wang, and Y. Deng, “Model establishment and harmonic analysis of electric vehicle charger,” in *Proc. 13th IEEE Conf. Ind. Electron. and Appl. (ICIEA)*, Wuhan, China, 2018, pp. 2204–2209, doi: 10.1109/ICIEA.2018.8398076.

[3] D. Schwanz, T. Busatto, M. Bollen, and A. Larsson, “A stochastic study of harmonic voltage distortion considering single-phase photovoltaic inverters,” in *Proc. 18th Int. Conf. Harmonics and Quality of Power (ICHQP)*, Ljubljana, Slovenia, May 2018, pp. 1–6, doi: 10.1109/ICHQP.2018.8378889.

[4] G. C. Jaiswal, M. S. Ballal, D. R. Tutakne, and H. M. Suryawanshi, “Impact of power quality on the performance of distribution transformers: A fuzzy logic approach to assessing power quality,” *IEEE Indus. App. Magazine*, vol. 25, no. 5, pp. 8–17, 2019, doi: 10.1109/MIAS.2018.2875207.

[5] D. M. Said and K. M. Nor, “Effects of harmonics on distribution transformers,” in *Proc. Australasian Universities Power Eng. Conf.*, Sydney, Australia, 2008, pp. 1–5.

[6] J. Gong, D. Li, T. Wang, W. Pan, and X. Ding, “A comprehensive review of improving power quality using active power filters,” *Electric Power Syst. Res.*, vol. 199, pp. 1–15, Oct. 2021, doi: 10.1016/j.epsr.2021.107389.

[7] L. Motta and N. Faúndes, “Active / passive harmonic filters: Applications, challenges & trends,” in *Proc. 17th Int. Conf. Harmonics and Quality of Power (ICHQP)*, Belo Horizonte, Brazil, 2016, pp. 657–662, doi: 10.1109/ICHQP.2016.7783319.

[8] M. Izhar, C. M. Hadzer, M. Syafrudin, S. Taib, and S. Idris, “Performance for passive and active power filter in reducing harmonics in the distribution system,” in *Proc. Nat. Power and Energy Conf. (PECon)*, Kuala Lumpur, Malaysia, Nov. 2004, pp. 104–108, doi: 10.1109/PECON.2004.1461625.

[9] S. Waosungnern, T. Narongrit, and K. Areerak, “Design of a shunt active power filter for single-phase power systems,” (in Thai), *TNI Journal of Eng. and Technol.*, vol. 9, no. 1, pp. 27–36, 2021.

[10] D. C. Bhonsle and R. B. Kelkar, "Design and simulation of single phase shunt active power filter using MATLAB," in *Proc. Int. Conf. Recent Advancements in Elect., Electron. and Control Eng.*, Sivakasi, India, 2011, pp. 237–241, doi: 10.1109/ICONRAEeCE.2011.6129786.

[11] R. Griñó, R. Costa-Castelló, and E. Fossas, “Digital repetitive control of a single-phase current active filter,” in *Proc. Euro. Control Conf. (ECC),* Cambridge, U.K., Sep. 2003, pp. 3494–3497, doi: 10.23919/ECC.2003.7086583.

[12] S. Waosungnern, T. Narongrit, and K. Areerak, “Harmonic detection using synchronous detection with fourier analysis for shunt active power filter in single - phase power systems,” (in Thai), in *Proc. 45th Elect. Eng. Conf.*, Nakhon Nayok, Thailand, Nov. 2022, pp. 142–145.

[13] M. P. Kumar *et al.*, “Power quality improvement using shunt active power filter with synchronous detection method,” *Complexity Int. J.*, vol. 25, no. 2, pp. 1637–1645, 2021.

[14] M. A. Kabir and U. Mahbub, “Synchronous detection and digital control of shunt active power filter in power quality improvement,” in *Proc. IEEE Power and Energy Conf.,* Urbana, IL, USA, 2011, pp. 1–5, doi: 10.1109/PECI.2011.5740499.

[15] C. Panpean, K. Areerak, and P. Santiprapan, “A harmonic voltage elimination in electric railway system using series active power filter,” in *Proc. 25th Int. Conf. Elect. Machines and Syst. (ICEMS)*, Chiang Mai, Thailand, 2022, pp. 1–5, doi: 10.1109/ICEMS56177.2022. 9983452.

[16] V. Kaura and V. Blasko, “Operation of a phase locked loop system under distorted utility conditions,” *IEEE Trans. Indus. App.*, vol. 33, no. 1, pp. 58–63, 1997.

[17] Y. G. Jung, “The current synchronous detection method combined with positive sequence detector for active power filters” *J. Elect. Eng. Technol.*, vol. 18, pp. 431–440, 2023.

[18] S. Waosungnern, T. Narongrit, K. Areerak, and A. Srikaew, “The compensating current control using fuzzy logic for shunt active power filter in single-phase power systems,” (in Thai), in *Proc. 44th Elect. Eng. Conf.*, Nan, Thailand, Nov. 2021, pp. 517–520.

[19] A. Srikaew, *Fuzzy Logic: Computational Intelligence*. Bangkok, Thailand: Charansanitwong Printing (in Thai), 2009.

[20] N. A. Kedar, A. P. Yadav, and V. B. Saruk, “Space vector modulation based control technique for shunt active power filter,” *Int. Res. J. Eng. Technol. (IRJET)*, vol. 7, no. 8, pp. 70–77, 2020.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Loss and Efficiency Improvement of EV Traction Drive System by Using SiC in Inverter

Kaninnat Rosdee[1*] Surapong Suwankawin[2]

[1*,2]*Department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand*

*Corresponding Author. E-mail address: 6370391321@student.chula.ac.th

Received: 26 August 2023; Revised: 3 September 2023; Accepted: 18 October 2023
Published online: 27 December 2023

Abstract

This paper investigates the effectiveness of employing SiC MOSFETs compared to Si-IGBTs in the two-level inverter for the traction motor drives in electric vehicles. The study focuses on the differences in structure and material between SiC MOSFETs and Si-IGBTs by which the conduction and switching losses can be reduced by over 40% and 60%, respectively, when a SiC MOSFET inverter is employed. This advantage enables the SiC MOSFET inverter to drive the traction motor at a higher switching frequency, surpassing traditional Si-IGBT inverters' capabilities. To assess the performance of the systems, simulations are conducted using the PLECS simulation platform. The comparison includes evaluating the traction drive system's efficiency and losses with SiC MOSFET and Si-IGBT inverters on the IPMSM traction motor torque-speed curve; the result shows that the efficiency is improved by over 1% on the entire torque-speed curve. Moreover, the paper also explores the trade-off between the switching frequency and motor harmonic core loss of the motor. This evaluation offers a comprehensive understanding of the interplay between these factors and aids in optimizing the performance of the traction motor system.

***Keywords*:** Conduction loss, Electric vehicles, Harmonic core loss, Switching loss, Wide bandgap

I. INTRODUCTION

Electric vehicles (EVs) are gaining popularity due to their numerous advantages, such as high efficiency, low emissions, and reduced dependency on fossil fuels. However, an EV's performance highly depends on the performance of the power electronics components, particularly the inverter that converts the DC (direct current) from the battery to the AC (Alternating Current) that drives the traction motor.

Power losses in EV drive trains are accounted for 18% of the total electrical power conversion in the system [1]. Two significant losses in the drive trains are 1) the traction motor's power losses and 2) the inverter's power losses. The traction motor's power losses consist of power loss at the conductor (copper loss) and power loss at the iron core (core loss), which are 70% of the drive trains losses. The other 30% of the drive train losses happen in the inverter, which is the loss when the switching devices are conducting the current (conduction loss) and the power loss when the switching devices change state (switching loss).

Power losses in EV drive trains can be reduced by a variety of methods, e.g., motor design [2]–[6], enhancement of motor control methods [3], [5], [7]–[9], development of inverter topologies [10]–[13], improvement of invertzers pulse width modulation (PWM) schemes [14]–[16] and the use of wide band gap (WBG) power switching devices for inverters [17]–[20].

WBG materials are semiconductor materials with a wider energy gap between the valence and conduction band compared to silicon (Si); the widely known WBG materials are, SiC (Silicon Carbide) and GaN (Gallium Nitride). The WBG materials can withstand higher voltage. It also has higher thermal conductivity and saturated electron drift velocity, resulting in better heat dissipation and the ability to operate at higher switching frequencies. Hence, the power-switching devices fabricated from WBG materials are used in the high-performance system. SiC and GaN are both generally fabricated as sub-type of FETs (Field effect transistors). SiC is usually fabricated as MOSFET (Metal-oxide silicon field effect transistor), and GaN is usually fabricated as HEMT (High electron mobility transistor). In [17] presents the application of GaN HEMT devices to two-level inverters for electric vehicles. The drive train efficiency was evaluated with the drive cycle testing. The test results indicated that the drive train efficiency could be increased by over 1% compared with IGBT (Insulated-gate bipolar transistor). However, this work was tested with a low DC bus voltage of 48V, which is too low for the typical electric vehicle application rating. In [18] presents a test of a two-level inverter based on GaN HEMT at 300V DC bus voltage, which is the voltage level for compact EVs. The study shows that inverters can reduce both conduction and switching power losses. The research also investigated the reduction of iron core loss by increasing the switching frequency. The optimum switching frequency must be selected between the switching loss power of the inverter and the power loss at the iron core caused by the switching carrier harmonics. Subsequent research [19] presents a suitable switching frequency control method with a switching frequency map obtained from the test. Article [20] studied the application of a combination of IGBT and GaN HEMT power switches for three-level inverters. It was tested on a system with a DC bus voltage of 800V and various motor driving conditions.

This paper compares the performance and efficiency of two-level three-phase inverters in an EV drive system. Specifically, this study compares the power loss of two-level inverters under different switching frequencies, motor driving conditions, and different types of power switching devices: IGBT and SiC MOSFET. The study is conducted on simulation software that simulates the operation of an EVs motor driving system. The results of this study will provide insight into the suitability of these power-switching devices for EV applications and

help engineers design more efficient and reliable EV drive systems.

The later sections of this paper are organized as follows. Section II describes the losses evaluation of the power switching devices. Motor core loss estimation due to PWM supply is presented in section III. Sections IV and V present the simulation parameters and methods. Simulation results and their analysis are discussed in Section VI. Finally, Section VII provides the study's conclusions.

II. LOSSES MODELLING AND EVALUATION OF THE POWER SWITCHING DEVICES

The power-switching device has two significant losses: conduction loss and switching loss [21]. These losses happen because of the non-ideal switching behaviors of the switching devices, as depicted in Figure 1. Since the loss model of IGBT, MOSFET, and anti-paralleled diode are different. Therefore, the different characteristics are discussed in this section.

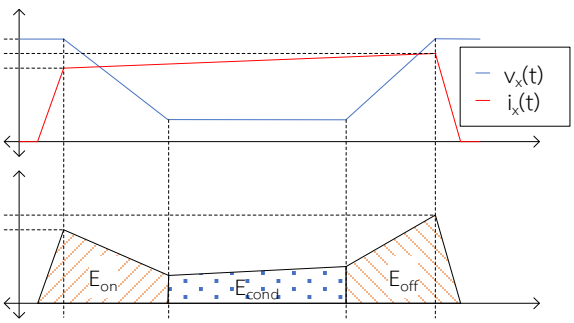


Figure 1: Non-ideal switching behavior of switching devices

A. Conduction Loss Evaluation

The conduction loss of a power-switching device depends on the current it is conducting and the voltage across it during the conduction state. Therefore, the conduction loss is evaluated according to equation (1), where $v_x\ t$, $i_x\ t$ and T is the voltage across the device, the current through the device, and the period of fundamental frequency, respectively. In addition, the subscript x denoted the device types: CE when the device is an IGBT (stands for collector-emitter), DS when the device is a MOSFET (stands for drain-source), or d when the device is a diode.

$$P_{cond} = \frac{1}{T}\int_0^T v_x\ t\ \cdot i_x\ t\ dt \qquad (1)$$

The voltage across the device during the conduction state $v_x\ t$ varies depending on the device's current $i_x\ t$ and junction temperature T_j, which differ according to the device's characteristics, as shown in Figures 2 and 3. These data are obtained via the device datasheets.

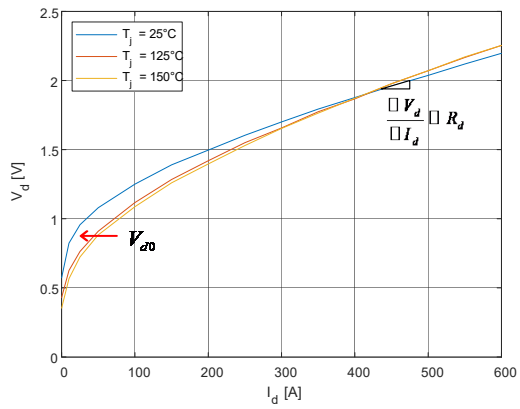


Figure 2: Output characteristic of an IGBT

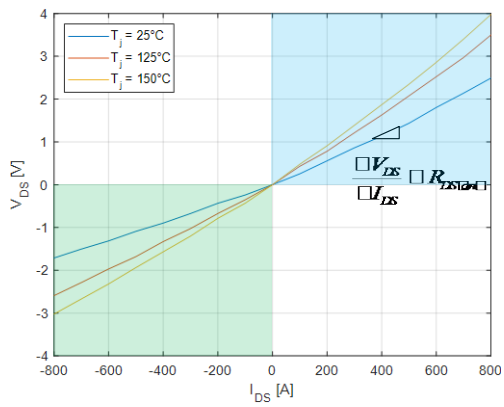


Figure 3: Output characteristic of a MOSFET

Due to the presence of the output bipolar junction transistor, the IGBT features a significantly lower voltage drop when the switch conducts high current. However, as penalties, IGBT has a diode-like voltage drop when it conducts low current. Furthermore, reverse conduction

is impossible because of the additional P-N junction, so the anti-paralleled diode is required, as depicted in Figure 4.

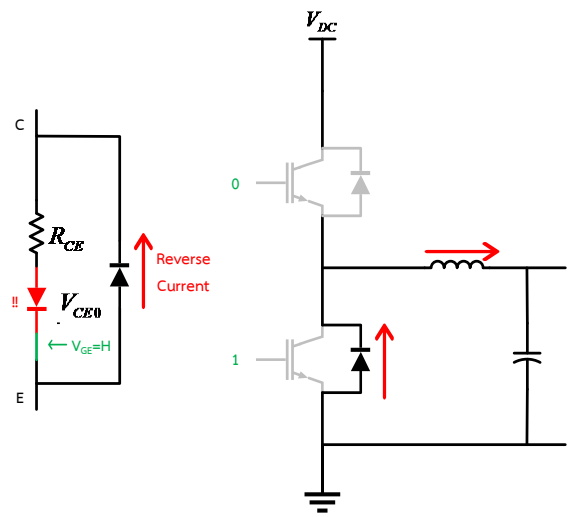


Figure 4: IGBT equivalent circuit and reverse conducting state in a half-bridge circuit

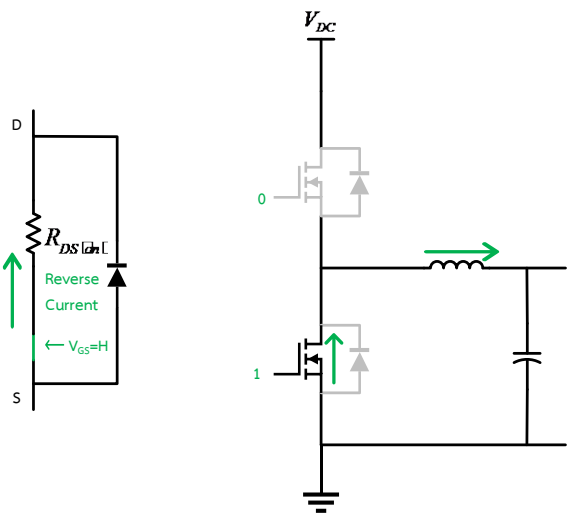


Figure 5: MOSFET equivalent circuit and reverse conducting state in a half-bridge circuit

Therefore, the losses during reverse conduction depend on the anti-paralleled diode's performance and characteristics. Unlike MOSFET, it can be represented as a resistor, so MOSFET can conduct both forward and reverse current, as shown in Figure 5. Moreover, the turn-on state resistance $R_{DS(on)}$ is small in the range of around 100 milliohms, the reverse voltage drop across source and drain $v_{SD}(t)$ will typically not be high enough to forward bias the anti-paralleled diode. Hence, the anti-paralleled diode is required for the conduction during dead time.

B. Switching Loss Evaluation

The switching loss is evaluated from switching energy consisting of turn-on loss and turn-off loss, which the manufacturer provides as a function of the current through the device i_x, the voltage across the device v_x, and the device's junction temperature T_j. Hence, the switching power loss is calculated according to (2)

$$P_{sw} = \frac{1}{T}\int_t^{t+T} \frac{E_{on}(v_x,i_x,T_j)+E_{off}(v_x,i_x,T_j)}{T_s} \cdot dt \quad (2)$$

where T_s is the switching period, and T is the fundamental period, $E_{on}(v_x,i_x,T_j)$ and $E_{off}(v_x,i_x,T_j)$ are the turn-on and turn-off loss at the voltage and current where the state changing occurs. Nevertheless, as shown in Figure 1, for a two-level converter, the voltage at the state-changing points is always equal to the DC bus voltage V_{DC} for either the upper switches (S1, S2, S3) or lower switches (S4, S5, S6). Hence, the switching power loss becomes

$$P_{sw} = \frac{1}{T}\int_t^{t+T} \frac{E_{on}(V_{DC},i_x,T_j)+E_{off}(V_{DC},i_x,T_j)}{T_s} \cdot dt \quad (3)$$

The IGBT and MOSFET have both the turn-on and turn-off loss; however, the turn-on energy loss of a diode or the energy consumed when the diode change state from reverse bias to forward bias is typically low enough to be neglectable. The dominant loss is the turn-off loss, also called reverse recovery loss $E_{rr}(v_x,i_x,T_j)$, which happens because of the charge at the depletion region Q_{rr}. Hence, the switching power loss of a diode is evaluated as follows:

$$P_{rr} = \frac{1}{T}\int_t^{t+T} \frac{E_{rr}(V_{DC},i_x,T_j)}{T_s} \cdot dt \quad (4)$$

III. INTERIOR PERMANENT MAGNET MOTOR CORE LOSS ESTIMATION WITH PWM SUPPLY

The IPMSM's core loss depends on various variables, e.g., the input voltage waveform, structure, motor geometry, stator, and rotor materials. Hence, these details are necessary for motor core loss analysis. By normal means, the famous and accurate method is FEM (Finite element method). However, to simplify the simulation model, this paper used the hysteretic core inductor with a ring shape which is similar to the motor stator shape and the BF (Building Factor) [22]. BF is the ratio between the ring-shaped inductor core loss and actual motor core loss, as shown in Figure 6.

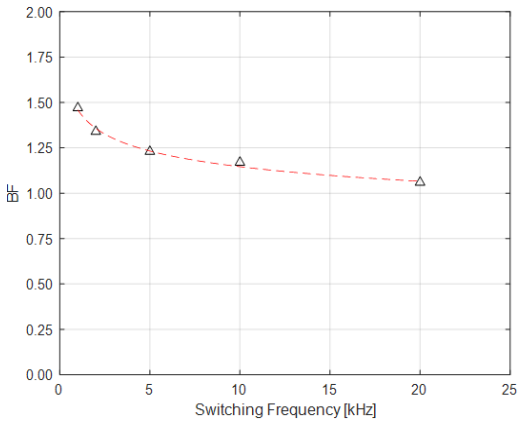


Figure 6: IPMSM building factor vs. supply switching frequency

As widely known, core loss consists of two significant parts; the hysteresis loss P_{hys} and the eddy current loss P_{eddy}, and to model these losses in the inductor core accurately, the non-uniformly discretized model is applied. The advantages of the non-uniformly discretized model are that it includes the skin effect when the core is supplied by a high-frequency supply resulting in an absolute error of less than 5% compared to the actual test results when the frequency of the supply is under 20kHz by implementing only two sections in the non-uniform ladder [23].

The non-uniformly discretized model is constructed by dividing a steel lamination into 2n sub-laminations, where n is the section's number in the non-uniform ladder, and the thickness of the sub-lamination is doubled progressively from the lamination's surface toward the lamination's center. Thus, the sub-lamination thickness and cross-sectional area are calculated as follows:

$$\triangle d_k = \frac{1}{\sum_{j=1}^{n} 2^{(j-k)}}\left(\frac{d}{2}\right) \tag{5}$$

$$A_k = 2 \cdot \triangle d_k \cdot w \tag{6}$$

where d and w are the lamination's thickness and width. The subscript k denotes the kth section in the non-uniform ladder.

The equivalent resistance is used to evaluate the eddy current loss: R_k is the equivalent resistance in the kth section, calculated by the following equation

$$R_k = \frac{2\rho \cdot w}{\triangle d_k \cdot l_e} \tag{7}$$

where ρ and l_e are the lamination's resistivity and core's magnetic path length.

Finally, the core is constructed by stacking the K layers of steel lamination. Hence, equation (8) and (9) becomes:

$$A_k = 2 \cdot \triangle d_k \cdot w \cdot K \cdot k_s \tag{8}$$

$$R_k = \frac{2\rho \cdot w \cdot K}{\triangle d_k \cdot l_e} \tag{9}$$

where k_s is the lamination's stacking factor.

The hysteresis loss P_{hys} can be calculated according to equation (10).

$$P_{hys} = \frac{1}{T}\int_t^{t+T} H \cdot dB \tag{10}$$

where H, B, T, are the magnetic field strength, magnetic flux density, and fundamental period, respectively. In other words, the hysteresis loss is the area enclosed by the B-H hysteresis loop. Under PWM supply, there are minor B-H hysteresis loops created by the switching pulses in addition to the fundamental B-H hysteresis loop, and these minor loops are causing the additional hysteresis loss, as shown in Figure 7. Therefore, the switching frequency of the PWM supply has a significant impact on these minor hysteresis loops and core loss.

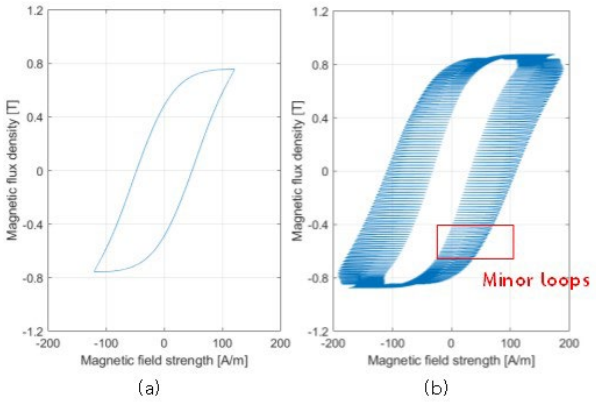


Figure 7: (a) B-H hysteresis loop of 50 Hz sinusoidal supply (b) B-H hysteresis loop of 50 Hz PWM supply $(f_s = 10\text{kHz})$

IV. EVALUATION OF LOSSES REDUCTION OF THE THREE-PHASE INVERTER BY PLECS SIMULATION

This section explains how losses of both the Si-IGBTs and SiC MOSFETs inverters are evaluated using the circuit simulator PLECS.

Figure 8 shows the simulation block diagram for evaluating the inverter loss of an IPMSM (Interior permanent magnet synchronous motor) drive system using PLECS. In the simulation, the three-phase voltages of the IPMSM are supplied by a two-level VSI (Voltage source inverter), which generates the PWM (Pulse width modulation) phase voltages at the switching frequency f_s. The PWM signals are calculated using the 3-arms carrier-based SVPWM (Space vector PWM) technique. For the motor control algorithm, the FOC (Field oriented control) with MTPA scheme (Maximum torque per ampere) and VCLMT scheme (Voltage-current limited maximum torque) are used in the inner control loop. MTPA calculates d-q axis currents that minimize the copper loss in the constant torque region, and VCLMT calculates the d-q axis current in the field weakening region. The outer control loop uses a PI controller as the motor speed regulator.

In PLECS, the device loss data are obtained via datasheets and experimental results as reference data, and these device loss data $E_{on}^{ref}\left(V_{x_ref}, I_{x_ref}, T_{j_ref}\right)$, $E_{off}^{ref}\left(V_{x_ref}, I_{x_ref}, T_{j_ref}\right)$, $E_{rr}^{ref}\left(V_{x_ref}, I_{x_ref}, T_{j_ref}\right)$, $v_{CE}^{ref}(t)$, $v_{DS}^{ref}(t)$ and $v_{d}^{ref}(t)$ are defined as Look-up tables (LUTs). So, when the operating points of the devices are not on the tables, PLECS performs the linear inter/extrapolation. From these device loss data, conduction and switching losses can be evaluated by loss equations in section II.

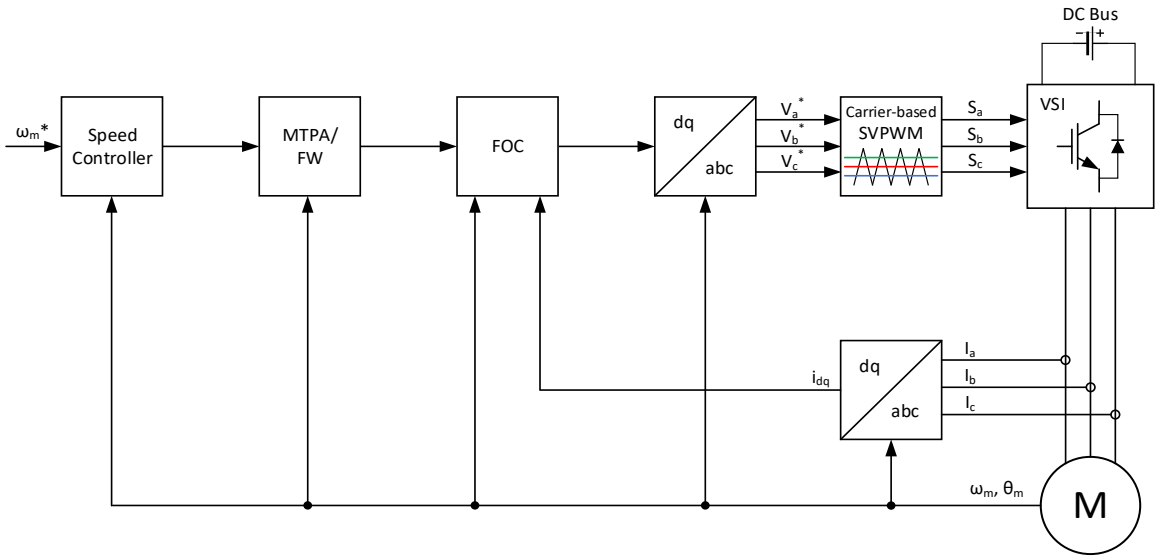


Figure 8: PLECS simulation block diagram

These simulations in PLECS carry out loss evaluations with LUTs; therefore, they can simplify complex parameter identification and loss estimations of power-switching devices. Furthermore, the junction temperature T_j is fixed in the simulations to simplify and reduce simulation time—the following parameters in Table 1, 2, and 3 are used to simulate the traction drive system.

Table 1: Switching device specifications

Classification	Product name	Ratings
SiC MOSFET	FS03MR12A6MA1B	1200V/400A
Si IGBT	FS380R12A6T4B	1200V/380A

Table 2: Simulation setting for losses evaluation

Parameter	Symbol	Value
DC bus voltage	V_{DC}	800V
Switching frequency	f_s	10 kHz, 20 kHz
Dead time	t_{dead}	1.0 μs (IGBT) 0.5 μs (SiC MOSFET)
Junction Temp.	T_j	125°C

Table 3: IPMSM Parameters

Parameter	Symbol	Value
Power Rating	P_{rated}	165 kW
Pole Pairs	p	3
Stator Resistance	R_s	28 mΩ
d-axis Inductance	L_d	0.41 mH
q-axis Inductance	L_q	0.725 mH
Permanent magnet flux	l_{pm}	0.17 Wb
Rotor Inertia	J_r	0.4 kgm^2

V. EVALUATION OF CORE LOSS REDUCTION OF THE IPMSM BY PLECS SIMULATION

The IPMSM core loss is evaluated based on the ring core analogy, as mentioned in section III. The PLECS magnetic domain, which simulates the magnetic circuit based on the permeance-capacitance analogy approach, is used to simulate the losses in the ring core inductor, and the ring core model is based on the non-uniformly discretized model with two sections in the non-uniform ladder. Figure 9 and Table 4 show the non-uniform ladder model and the ring core material and geometry, respectively.

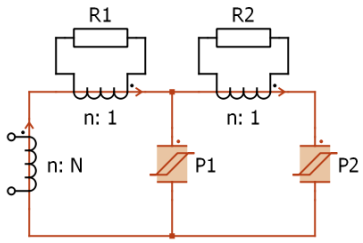


Figure 9: Per phase non-uniformly discretized model in PLECS

Table 4: Ring core material and geometry

Parameter	Symbol	Value
Steel lamination	-	35H300
Height	h	70 mm.
Outer diameter	D_{out}	127 mm.
Inner diameter	D_{in}	102 mm.
Magnetic path length	l_e	360 mm.
Turn number	N	283

VI. SIMULATION RESULTS

A. Comparison of Switching and Conduction Losses and Efficiency on the Ipmsm Torque-Speed Curve

Figures 10 and 11 show each inverter's conduction loss and switching loss at a rotor speed of 5000 r/min (80% of the rated speed). SiC MOSFET inverter can reduce the conduction loss by 40% and 69% at load torque $T_L =$ 225 and 75 Nm at the same carrier frequency and because of the reverse conduction capability; therefore, conduction loss in the anti-paralleled diodes is reduced by over 85%. Also, the SiC MOSFET inverter can decrease the switching loss drastically; as shown in Figure 11, the switching loss of the SiC MOSFET inverter was reduced by 62% at $T_L =$ 225 Nm and 71% at $T_L =$ 75 Nm at the same switching frequency. Even when the SiC MOSFET inverter operates at the 20kHz switching frequency, the switching loss is still lower than the Si-IGBT inverter; hence, the SiC MOSFET inverter is capable of operating

at the double switching frequency and still achieves better efficiency compared to the Si-IGBT inverter. For conduction loss, it does not increase much because the load current is not changed. Nevertheless, the conduction loss of anti-paralleled diodes is doubled because when the switching frequency doubles, it also doubles the conduction time of the diodes since the dead time counts per fundamental period increased. Furthermore, as the dead time increases, the conduction time of the SiC MOSFET decreases; thus, the conduction loss is reduced slightly.

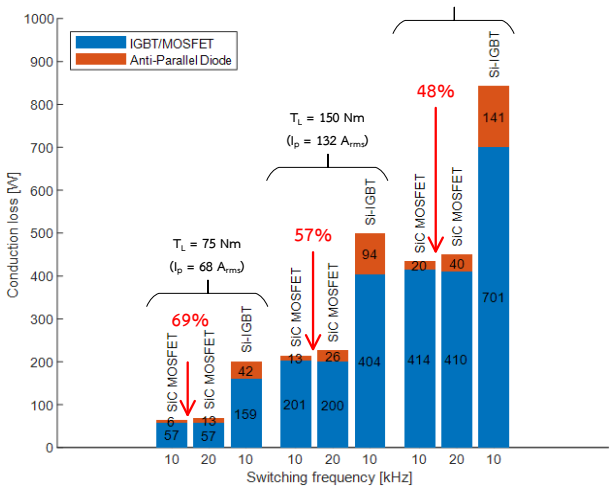


Figure 10: Comparison of conduction loss at w_m = 5000 r/min

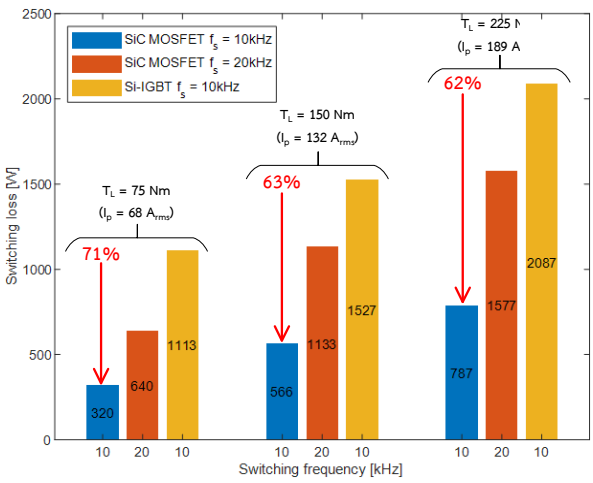


Figure 11: Comparison of switching loss at w_m = 5000 r/min

Si-IGBT and SiC MOSFET inverter efficiency maps on the motor torque-speed curve are shown in Figures 12 and 13. The efficiency improvement is presented in Figure 14. The SiC MOSFET inverter can achieve over 99% efficiency and get over 1% better than the Si-IGBT in a wide area on the torque-speed curve, specifically the high-torque high-speed area. On the other hand, the Si-IGBT can achieve around 98%. Furthermore, in the low-load torque in the field-weakening region near the rated speed (6300 r/min), the efficiency improves by over 20% and slowly decreases. Because, in this region, the inverter must feed the d-axis current continuously to weaken the magnetic flux linkage, which induces the back EMF (Electromotive force). Hence there is always conduction loss occurs. For example, when the motor runs at a speed of 7000 r/min, the d-axis current of around 153A_p must be injected. The voltage across Si-IGBT is around 1.5V_p; conversely, the voltage across SiC MOSFET is only 0.62V_p. As described in Section II, the big difference in voltage drop is the effect of the diode-like voltage drop when Si-IGBT conducts low current. Hence, the conduction loss played a significant part in the efficiency improvement of the FW region's low-torque area. And then, the d-axis current increases with the increasing rotor speed to counteract the back EMF, which reduces the effect of the diode-like voltage drop; therefore, the efficiency improvement slowly decreases.

B. Motor Core Loss Reduction by Increasing the Switching Frequency of the Inverter

As mentioned in the previous section, utilizing the SiC MOSFET allows the inverter to operate at a higher switching frequency without significantly increasing inverter conduction and switching loss. Furthermore, increasing the switching frequency improves motor efficiency because the harmonic core loss of the motor decreases at a higher switching frequency.

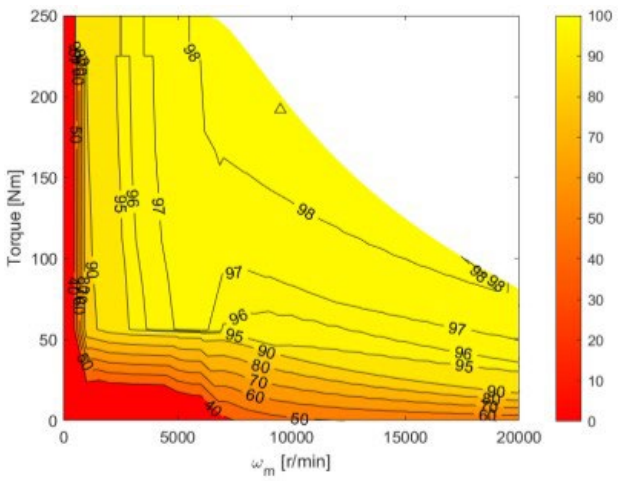


Figure 12: Two-level Si-IGBT inverter efficiency map

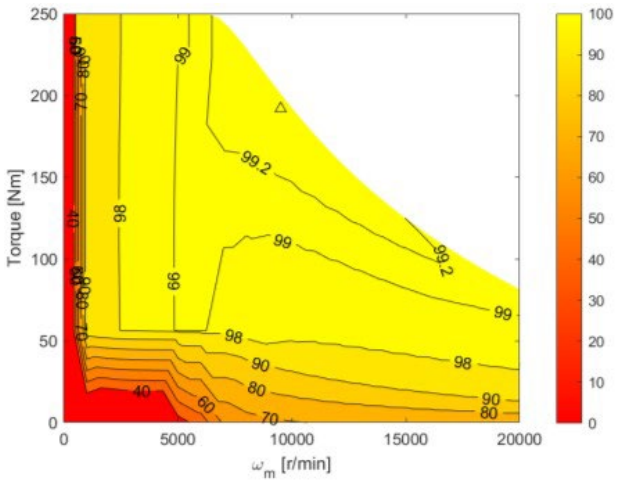


Figure 13: Two-level SiC MOSFET inverter efficiency map

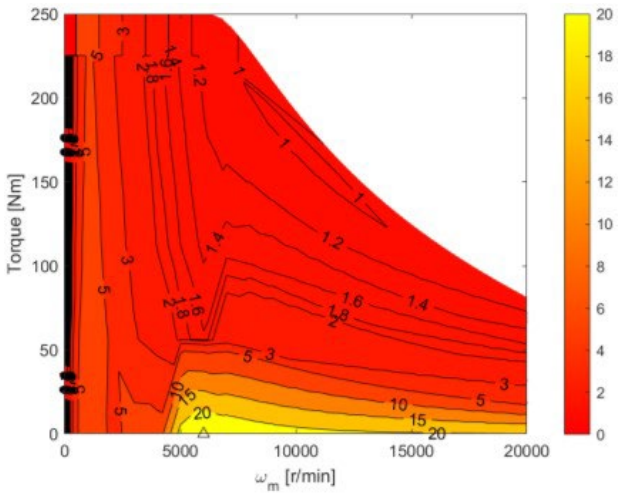


Figure 14: Difference in inverter efficiency

Figure 15 shows the relative core loss reduction of the inverter-fed iron core at switching frequencies of 10 kHz vs. 20 kHz.

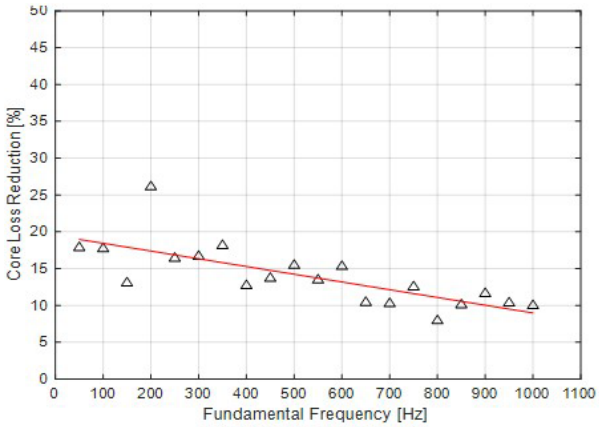


Figure 15: Relative core loss reduction of 10 kHz vs. 20 kHz

Core loss is reduced by up to 20% in the low-speed region and around 10% in the high-speed region. However, the overall efficiency of the drive train may not increase because the increases in the switching frequency mean increases in switching loss of the inverter resulting in inverter efficiency decreases. As shown in Figures 16–18, increasing the switching frequency at every motor's operating point is not always beneficial. Hence, optimizing between inverter switching loss and motor core loss will improve the traction drive system's overall efficiency.

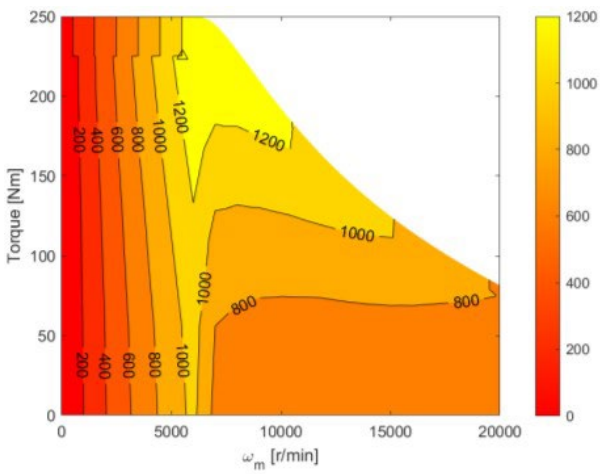


Figure 16: Core loss reduction of 10 kHz vs. 20 kHz (W)

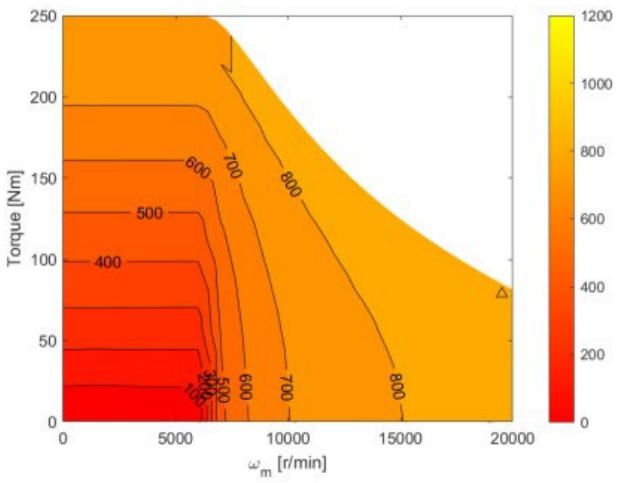


Figure 17: Difference of switching loss 10 kHz vs. 20 kHz (W)

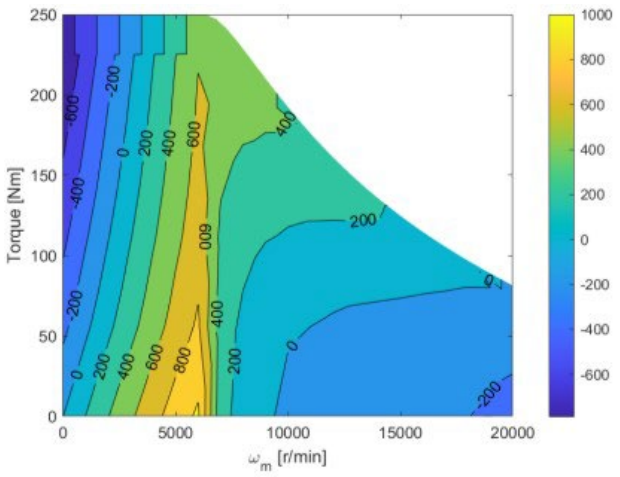


Figure 18: Difference between core loss reduction vs. switching loss increase (W)

VII. CONCLUSION

This paper demonstrates the effectiveness of utilizing a SiC MOSFET inverter—the advantages of employing the SiC MOSFET inverter and IPMSM as the EV drive train was discussed. The SiC MOSFET inverter significantly reduces power losses compared to the Si-IGBT inverter, with over 40% conduction loss reduction and over 60% switching loss reduction. As a result, the inverter and traction drive system efficiency is improved, and core loss reduction of the motor due to the increased switching frequency is evaluated.

REFERENCES

[1] U.S. Department of Energy. "Where the Energy Goes: Electric Cars." FUELECONOMY.gov. https://www.fueleconomy.gov/feg/atv-ev.shtml (accessed Dec. 1, 2022).

[2] I. Husain *et al.*, "Electric drive technology trends, challenges, and opportunities for future electric vehicles," *Proc. IEEE*, vol. 109, no. 6, Jun. 2021, pp. 1039–1059.

[3] V. Ruuskanen, J. Nerg, M. Rilla, and J. Pyrhönen, "Iron loss analysis of the permanent-magnet synchronous machine based on finite-element analysis over the electrical vehicle drive cycle," *IEEE Trans. Ind. Electron.*, vol. 63, no. 7, pp. 4129–4136, Jul. 2016, doi: 10.1109/TIE.2016.2549005.

[4] M. Younkins, P. Carvell, and J. Fuerst, "Dynamic motor drive: Optimizing electric motor controls to improve efficiency," presented at *42nd Int. Vienna Motor Symp.*, Vienna, Austria, Apr. 29-30, 2021.

[5] L. Shao, A. E. H. Karci, D. Tavernini, A. Sorniotti, and M. Cheng, "Design approaches and control strategies for energy-efficient electric machines for electric vehicles—A review," *IEEE Access*, vol. 8, pp. 116900–116913, 2020.

[6] Y. Miyama, M. Hazeyama, S. Hanioka, N. Watanabe, A. Daikoku, and M. Inoue, "PWM carrier harmonic iron loss reduction technique of permanent-magnet motors for electric vehicles," *IEEE Trans. Ind. Appl.*, vol. 52, no. 4, pp. 2865–2871, 2016, doi: 10.1109/TIA.2016.2533598.

[7] O. Wallscheid, M. Meyer, and J. Böcker, "An open-loop operation strategy for induction motors considering iron losses and saturation effects in automotive applications," in *Proc. IEEE 11th Int. Conf. Power Electron. and Drive Syst.*, Sydney, Australia, 2015, pp. 981–985.

[8] S. Vaez, V. I. John, and M. A. Rahman, "Energy saving vector control strategies for electric vehicle motor drives," in *Proc. Power Convers. Conf. - PCC '97*, Nagaoka, Japan, Aug. 1997, pp. 13–18.

[9] M. Alnajjar and D. Gerling, "Minimization of energy losses in the traction drive of hev using optimized adaptive control," in *Proc. IEEE Veh. Power and Propulsion Conf. (VPPC),* Coimbra, Portugal, 2014, pp. 1–6.

[10] A. Choudhury, P. Pillay, and S. S. Williamson, "Comparative analysis between two-level and three-level DC/AC electric vehicle traction inverters using a novel DC-Link voltage balancing algorithm," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 2, no. 3, pp. 529–540, Sep. 2014.

[11] N. Sorokina *et al.,* "Inverter and battery drive cycle efficiency comparisons of multilevel and two-level traction inverters for battery electric vehicles," in *Proc. IEEE Int. Conf. Environ. and Elect. Eng. and IEEE Ind. & Commercial Power Syst. Eur. (EEEIC / I&CPS Europe)*, Bari, Italy, Nov. 2021, pp. 1–8.

[12] X. Zhang, J. Y. Gauthier, and X. Lin-Shi, "A composite converter with reduced power electronics for electric powertrain applications," in *Proc. IEEE Energy Convers. Congr. and Expo. (ECCE)*, Vancouver, Canada, Oct. 2021, pp. 1468–1475.

[13] E. Libbos, E. Krause, A. Banerjee, and P. T. Krein, "Inverter design considerations for variable-pole induction machines in electric vehicles," *IEEE Trans. Power Electron.*, vol. 37, no. 11, pp. 13554–13565, Nov. 2022.

[14] J.-I. Itoh and T. Ogura, "Evaluation of total loss for an inverter and motor by applying modulation strategies," in *Proc. 14th Int. Power Electron. and Motion Control Conf. EPE-PEMC*, Ohrid, Macedonia, Sep. 2010, pp. S12-21–S12-28, doi: 10.1109/epepemc.2010.5606664

[15] R. Menon, N. A. Azeez, A. H. Kadam, and S. S. Williamson, "Energy loss analysis of traction inverter drive for different PWM techniques and drive cycles," in *Proc. IEEE Int. Conf. Ind. Electron. Sustain. Energy Syst. (IESES)*, Hamilton, New Zealand, 2018, pp. 201–205.

[16] A. Choudhury, P. Pillay, and S. S. Williamson, "A hybrid PWM-Based DC-Link voltage balancing algorithm for a three-level NPC DC/AC traction inverter drive," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 3, no. 3, pp. 805–816, Sep. 2015, doi: 10.1109/JESTPE.2015.2439296.

[17] K. Kumar and S. B. Santra, "Performance analysis of a three-phase propulsion inverter for electric vehicles using GaN semiconductor devices," *IEEE Trans. Ind. Appl.,* vol. 54, no. 6, pp. 6247–6257, 2018, doi: 10.1109/TIA.2018.2862400.

[18] Y. Nakayama *et al.,* "Efficiency improvement of motor drive system by using a GaN three phase inverter," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT),* Melbourne, Australia, Feb. 2019, pp. 1599–1604, doi: 10.1109/ICIT.2019.8755018.

[19] K. Ohta *et al.,* "Variable switching frequency control for efficiency improvement of motor drive system by using GaN three phase inverter," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT),* Buenos Aires, Argentina, Feb. 2020, pp. 119–123, doi: 10.1109/ICIT45562.2020.9067266.

[20] J. Lu, R. Hou, P. Di Maso, and J. Styles, "A GaN/Si hybrid T-type three-level configuration for electric vehicle traction inverter," in *Proc. IEEE 6th Workshop Wide Bandgap Power Devices and Appl. (WiPDA),* Atlanta, GA, USA, 2018, pp. 77–81.

[21] J. Guo, "Modeling and design of inverters using novel power loss calculation and DC-Link current/voltage ripple estimation methods and bus bar analysis," Ph.D. dissertation, Dept. Elect. Comput. Eng., McMaster Univ., Hamilton, ON, Canada, 2017.

[22] N. Denis, S. Odawara, and K. Fujisaki, "Attempt to evaluate the building factor of a stator core in Inverter-Fed permanent magnet synchronous motor," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2424–2432, Mar. 2017.

[23] E. J. Tarasiewicz, A. S. Morched, A. Narang, and E. P. Dick, "Frequency dependent eddy current models for nonlinear iron cores," *IEEE Trans. Power Syst.*, vol. 8, no. 2, pp. 588–597, May 1993.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

คุณสมบัติทางกายภาพและชีวภาพของไหมเย็บแผลที่ชุบด้วยลีโวฟลอกซาซิน

ปวรวรรณ ฤทธิพากร[1*] ธนพร จันทร์โท[2] พิชญ์มณฑ์ พุ่มสงวน[3] วิมลวรรณ จันทนะสุต[4] ชนม์นิภา คุณวิทยา[5]

[1]สาขาวิชาสุขภาพช่องปากผู้ใหญ่ สำนักวิชาทันตแพทยศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, นครราชสีมา, ประเทศไทย*
[2,3,4,5]สำนักวิชาทันตแพทยศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, นครราชสีมา, ประเทศไทย

*ผู้ประพันธ์บรรณกิจ อีเมล : r.paw@g.sut.ac.th

รับต้นฉบับ : 27 เมษายน 2566; รับบทความฉบับแก้ไข : 6 ตุลาคม 2566; ตอบรับบทความ : 9 ตุลาคม 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

ไหมเย็บแผล (silk suture) เป็นไหมเย็บแผลที่ทางทันตแพทย์นิยมใช้นำมาเย็บแผลในช่องปาก ซึ่งจะคงสภาพแรงดึงและสอดคล้องกับการหายของแผล โดยจะเย็บคงสภาพเพื่อให้เกิดการหายของแผลเป็นระยะเวลาอย่างน้อย 3–5 วัน คุณสมบัติของเส้นไหมแบบ non-resorbable suture ที่นิยมใช้ในทันตกรรม สามารถทำให้มีเศษอาหารเข้ามาสะสม และส่งผลให้เกิดการติดเชื้อมากกว่าแบบไหมละลาย โดยเฉพาะในผู้ป่วยที่มีโรคประจำตัว เช่น เบาหวาน หรือโรคที่มีภูมิคุ้มกันบกพร่อง ปัจจุบันมีการพัฒนาให้ไหมเย็บแผลสามารถเคลือบด้วยยาปฏิชีวนะ เพื่อลดความเสี่ยงในการติดเชื้อแต่ยังไม่พบการศึกษาที่ใช้ไหมที่ทำด้วย silk การศึกษานี้มีวัตถุประสงค์เพื่อประเมินคุณสมบัติทางกายภาพและชีวภาพของไหมเย็บแผลชนิดไม่ละลายแบบพันเกลียว (non-absorbable and multifilament) ที่แช่ด้วยยาปฏิชีวนะลีโวฟลอกซาซิน ความเข้มข้น 3 มก./มล. ด้วย 2 วิธี คือแบบแห้งและแบบเปียก จากนั้นนำไปทดสอบเปรียบเทียบคุณสมบัติทางกายภาพโดยการทดสอบแรงดึง ปริมาณรูพรุน การมีอยู่ของยาลีโวฟลอกซาซิน และการปลดปล่อยยา โดยเปรียบเทียบกับไหมเย็บแผลปกติ ผลการศึกษาพบว่าไหมเย็บแผลที่แช่ยาลีโวฟลอกซาซินมีความแข็งแรง ไม่แตกต่างอย่างมีนัยสำคัญทางสถิติกับไหมเย็บแผลปกติ และการปลดปล่อยสารละลายลีโวฟลอกซาซินจากไหมทั้งสองกลุ่ม ไม่แตกต่างกันและสามารถปลดปล่อยยาได้ต่อเนื่องเป็นระยะเวลา 7 วัน ในกลุ่มเปียกพบว่าค่ารูพรุนมีปริมาณที่น้อยกว่าและขนาดเส้นผ่านศูนย์กลางรูพรุนมีขนาดที่ใหญ่กว่ากลุ่มแห้งและกลุ่มควบคุม ดังนั้นการใช้ไหมเย็บแผลปกติที่นำไปชุบด้วยยาปฏิชีวนะ มีความแข็งแรงที่ใกล้เคียงกับจากไหมเย็บแผลปกติ และสามารถใช้ประโยชน์ในการปลดปล่อยยาเพื่อลดการติดเชื้อบริเวณแผล

คำสำคัญ : การติดเชื้อ ลีโวฟลอกซาซิน ไหมเย็บแผล ค่าทนแรงดึง

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Physical and Biological Properties of Silk Suture Soaked in Levofloxacin

Pawornwan Rittipakorn[1*] Thanaporn Chanto[2] Pichamon Pumsanguan[3] Wimonwun Juntanasut[4]
Chonipa Khunwittaya[5]

[1*]*School of Adult Oral Health, Institute of Dentistry, Suranaree University of Technology, Nakorn Ratchasima, Thailand*
[2,3,4,5]*Institute of Dentistry, Suranaree University of Technology, Nakorn Ratchasima, Thailand*

*Corresponding Author. E-mail address: r.paw@g.sut.ac.th

Received: 27 April 2023; Revised: 6 October 2023; Accepted: 9 October 2023
Published online: 27 December 2023

Abstract

Silk suture was a type of suture material commonly used by dentists to close wounds in the oral cavity. It provides strong tensile strength and was consistent with wound healing by maintaining suture integrity to promote wound healing for at least 3-5 days. However, non-resorbable sutures made of silk were known to accumulate food debris and increase the risk of infection, particularly in patients with underlying conditions such as diabetes or immune deficiencies. There has been recent development in coating silk sutures with antimicrobial agents to reduce the risk of infection, but no studies had been conducted using silk sutures coated with such agents. This study aimed to evaluate the physical and biological properties of non-absorbable and multifilament braid silk sutures that have been coated with levofloxacin. The physical properties of strength, amount of pore with SEM and BET analysis, and the presence of levofloxacin solution were assessed, as were the biological properties of drug release profile. All groups were compared between dry and wet conditions of the coated silk sutures and non-coated silk sutures. The results showed that the levofloxacin-coated silk sutures had no significant difference tensile strength between non-coated silk sutures, and the release of levofloxacin was sustained over a period of 7 days. BET analysis showed that the wet group had less porosity and wider pore diameter than the other groups. Therefore, the use of levofloxacin-coated silk sutures was provided comparable suture strength to non-coated silk sutures while reducing the risk of infection.

Keywords: Infection, Levofloxacin, Suture materials, Tensile strength

1) บทนำ

ไหมเย็บแผล (silk suture) เป็นเครื่องมือทางการแพทย์ที่มีความสำคัญในการเย็บปิดแผล โดยในทางทันตกรรม ต้องอาศัยความแข็งแรง และทนต่อแรงดึงที่เหมาะสมของแผล บริเวณนั้น ซึ่งไหมเย็บแผลควรมีคุณสมบัติทางกายภาพที่เหมาะสม เข้ากันได้ดีกับเนื้อเยื่อของร่างกาย และไม่ส่งเสริมให้แบคทีเรียมาเกาะบริเวณแผลผ่าตัด [1] ไหมเย็บแผลจะต้องมีความแข็งแรงที่เพียงพอและสอดคล้องกับกระบวนการหายของแผล จึงจะสามารถนำไหมออกจากบริเวณแผล โดยจะเย็บคงสภาพเพื่อให้เกิดการหายของแผลเป็นระยะเวลาอย่างน้อย 3 วันขึ้นกับลักษณะของแผลและความแข็งแรงที่ต้องการบริเวณแผลผ่าตัด ที่ส่งผลต่อการเลือกใช้ไหมเย็บแผล

ไหมเย็บแผลในปัจจุบันที่ใช้ทางทันตกรรมมีหลายประเภท เช่น ไหมละลาย (adsorbable) และ ไหมไม่ละลาย (non-adsorbable) หรือแบ่งตามชนิดต้นกำเนิด อาทิเช่น ไหมสังเคราะห์ และไหมธรรมชาติ [2] ในทางทันตกรรมปัจจุบัน นิยมใช้ไหมสังเคราะห์ชนิดไม่ละลาย ลักษณะหลายเส้นใย (multifilament silk suture) เนื่องจากราคาไม่แพงและหาซื้อง่าย คุณสมบัติทางกายภาพเหมาะสม เมื่อเทียบกับไหมชนิดอื่น จากการศึกษาพบว่าไหมเย็บแบบพันเกลียวหลายเส้นใย (multifilament) พบเชื้อแบคทีเรียสแตฟิโลค็อกคัส ออเรียส (*Staphylococcus aureus*) เมื่อส่องด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องผ่าน (Transmission Electron Microscopy: TEM) [3] ทำให้บริเวณเนื้อเยื่อรอบ ๆ ไหมเย็บแผล มีเชื้อแบคทีเรียอยู่บริเวณร่องไหม ทำให้เซลล์เม็ดเลือดขาว (Leukocyte) ที่สามารถซึมผ่านไหมเย็บและทำให้เกิดขั้นตอนการกำจัดเซลล์แบคทีเรีย (Phagocytosis) ได้ค่อนข้างยาก เมื่อไหมเย็บแผลอยู่ในช่องปากและสัมผัสกับ oral fluid จะมีกระบวนการที่เรียกว่าการซึมผ่านรูเล็ก (capillary action) ซึ่งเกิดการบวมของเส้นไหม ทำให้แบคทีเรียสามารถซึมผ่านทางเส้นไหมสู่บริเวณแผลได้ และคงทนได้ถึง 30 วัน [4]

หนึ่งในหลายปัจจัยของการติดเชื้อ และอักเสบของบริเวณแผลผ่าตัด (Surgical Site Infection: SSI) คือวัสดุไหมเย็บแผลโดยเฉพาะแบบหลายเส้นใย (braided) และการมัดปมไหม ซึ่งสามารถเป็นที่สะสมของเชื้อแบคทีเรีย โดยเฉพาะอย่างยิ่งในคนไข้ที่มีโรคประจำตัวที่เสี่ยงในการติดเชื้อแทรกซ้อน เช่น โรคหัวใจและโรคเบาหวาน เป็นต้น ปัจจุบันในท้องตลาดมีไหมเย็บแผลที่เคลือบด้วยสารต้านจุลชีพไตรโคลซานจำหน่าย (Vicryl Plus®, Ethicon Inc., Someville, NJ, USA) ซึ่งเป็นไหมละลาย มีการศึกษาพบว่าสามารถลดการเกาะของแบคทีเรียกลุ่ม Staphylococci ได้ [5] แต่มีรายงานว่ามีอัตราการเกิดการแตกของแผล (wound dehiscence) ได้มากกว่าไหมแบบธรรมดา [6] มีการศึกษาที่ใช้ไหมแช่ในสารละลายคลอเฮกซิดีน กลูโคเนตความเข้มข้นร้อยละ 0.12 สามารถยับยั้งเชื้อแบคทีเรีย และปลดปล่อยตัวยาได้รวดเร็วใน 24 ชั่วโมงแรกจากนั้นลดลงจนถึง 7 วัน [7] อีกการศึกษาได้ใช้ไหมเย็บแผลเคลือบด้วยเตตราไซคลิน พบว่าไม่ทำให้ความแข็งแรงของไหมลดลงเมื่อเทียบกับไหมธรรมดา และสามารถยับยั้งแบคทีเรียชนิด *Escherichia coli* ได้มากกว่า *S. aureus* [8] แต่อย่างไรก็ตามมีรายงานว่ามีโอกาสดื้อยาและ มีโอกาสเกิดผลข้างเคียงที่ไม่พึงประสงค์จากการใช้ยา เช่น ไข้ และ ผื่นจากการใช้ยา [9]

ลีโวฟลอกซาซินเป็นยาต้านจุลชีพในกลุ่มฟลูออโรควิโนโลน (Fluoroquinolones) ออกฤทธิ์ฆ่าเชื้อแบคทีเรีย (Bactericidal) โดยยับยั้งกระบวนการสร้าง DNA โดยจับที่ Topoisomerase II (DNA gyrase) และ Topoisomerase IV [10] ได้ทั้งแบคทีเรียแกรมบวกและลบ และผลข้างเคียงไม่พึงประสงค์ที่ค่อนข้างต่ำและโอกาสดื้อยาที่น้อย ซึ่งเชื้อในช่องปากที่มักพบคือ กลุ่ม Staphylococci เช่น *S. aureus* และกลุ่ม Coliform คือ *E. coli* มีการศึกษาของ Chen *et al.* [11] ใช้ไหมหลายเส้นใยแบบพันเกลียวนำไปเคลือบกับสารละลายยาลีโวฟลอกซาซินไฮโดรคลอไรด์ (Levofloxacin Hydrochloride) ผสมกับ พอลิคาโปรแลกโตน (Poly ε-caprolactone: PCL) ใช้สารละลายความเข้มข้น 3 มิลลิกรัม/มิลลิลิตร พบว่ามีความสามารถในการยับยั้งเชื้อแบคทีเรีย *S. aureus* และ *E. coli* ได้ต่อเนื่องจนถึง 7 วัน และปลดปล่อยยาได้ต่อเนื่อง 120 ชั่วโมง งานวิจัยทางศัลยกรรมทางตาได้ใช้พอลิคาโปรแลกโตนเคลือบด้วยยาลีโวฟลอกซาซินร้อยละ 8 ด้วยวิธีอิเล็กโตรสปันทำเป็นไหมขนาดเล็กพิเศษพบว่าไม่ทำให้ความแข็งแรงของไหมลดลง [12]

เนื่องจากปัจจุบันไหมเคลือบสารที่มีคุณสมบัติฆ่าเชื้อแบคทีเรียมักจะเป็นไหมละลายและมีราคาสูง และยังไม่มีการใช้ยาต้านจุลชีพที่มีผลข้างเคียงน้อย อีกทั้งลีโวฟลอกซาซินเหมาะกับการนำมาใช้ร่วมกับไหมแบบหลายเส้นใยที่นิยมใช้ในทางทันตกรรม [13] จุดประสงค์ของการศึกษานี้ต้องการทดสอบคุณสมบัติทางกายภาพของไหมเย็บแผลที่เคลือบด้วยสารต้านจุลชีพลีโวฟลอกซาซิน และทดสอบความสามารถในการปลดปล่อยยา เปรียบเทียบกับไหมเย็บแผลปกติ เพื่อนำไปต่อยอดในการใช้งานในอนาคต และเหมาะสมกับการใช้งานสำหรับแผลในช่องปาก อีกทั้งยังสามารถ

นำไปใช้ได้ง่ายและราคาถูกในผู้ป่วยที่มีโรคประจำตัวที่เสี่ยงต่อการติดเชื้อจากแผลสูง

2) วิธีดำเนินการวิจัย

2.1) ขั้นตอนการเตรียมไหมเย็บแผลและ เตรียมสารละลายลีโวฟลอกซาซิน

เตรียมไหมเย็บแผล Silk braided suture 3-0 (UNIK® Braided Silk Wax US3/0) ซึ่งเป็นไหมเย็บแผลทั่วไปที่ใช้ทางการแพทย์ชนิดไม่ละลายแบบพันเกลียว (non-absorbable and multi-filament) นำมาตัดให้ได้ความยาว 30 เซนติเมตร จำนวน 51 เส้น 15 กรัม จากนั้นนำไปทำให้ปราศจากเชื้อด้วยหม้อนึ่งอัดไอน้ำ อุณหภูมิ 121 องศาเซลเซียส เป็นเวลา 15 นาที 15 ปอนด์ต่อตารางนิ้ว

การเตรียมยาลีโวฟลอกซาซิน (Levofloxacin Hydrochloride, Siam Pharmaceutical®, Bangkok, Thailand) ความเข้มข้นที่ 3 มก./มล. ในรูปของสารละลาย

2.2) กลุ่มตัวอย่าง

แบ่งไหมเย็บแผลขนาด 3-0 ออกเป็น 3 กลุ่ม ได้แก่กลุ่มควบคุม คือไหมที่ไม่ได้ชุบสารละลายลีโวฟลอกซาซิน กลุ่มเปียกและกลุ่มแห้ง โดยกลุ่มเปียก คือนำไหมเย็บไปแช่สารละลายลีโวฟลอกซาซินความเข้มข้นดังข้างต้น เป็นเวลา 24 ชั่วโมง และทิ้งให้แห้งที่อุณหภูมิห้อง กลุ่มแห้ง คือนำไหมเย็บไปแช่สารละลายลีโวฟลอกซาซิน เป็นเวลา 24 ชั่วโมง และผ่านกระบวนการทำแห้งแบบระเหิด (freeze dry) ด้วยเครื่องทำให้แห้งภายใต้ความเย็นและสุญญากาศ (Labconco 195, USA) จนแห้งสนิท

2.3) การศึกษาลักษณะทางกายภาพของไหมเย็บแผล (Physical Properties)

2.3.1) การศึกษาค่าทนแรงดึงของไหมเย็บแผล ทดสอบค่าทนแรงดึงของไหมเย็บแผลด้วยเครื่อง Universal Testing Machine (รุ่น WAW-200E) ตามมาตรฐาน ASTM D2556 แบ่งเป็นกลุ่มควบคุม และกลุ่มศึกษารวม 3 กลุ่ม กลุ่มละ 10 ชิ้น ทำการมัดปมไหมทุกเส้นในการทดลอง ใช้ผู้ทดลอง 1 คนในการมัดปม ให้อยู่ตรงกลางเส้นไหมพอดีโดยใช้ไม้บรรทัดวัด (ปมอยู่ห่างจากปลายทั้งสองข้าง ข้างละ 15 ซม.) จากนั้นให้แรงดึงต่อเนื่องด้วยอัตรา 20 มล./นาที [7] บันทึกค่าการเปลี่ยนแปลงที่เกิดขึ้นและแรงที่ใช้จนไหมเย็บขาดออกจากกัน

2.3.2) การศึกษาลักษณะพื้นผิวและรูปร่างโครงสร้าง 3 มิติ ศึกษาลักษณะพื้นผิวของไหมเย็บแผลด้วยกล้องจุลทรรศน์ชนิดส่องกราด (Scanning Electron Microscope: SEM, FEI Quanta 400) นำไหมเย็บแผลจากกลุ่มควบคุมและกลุ่มศึกษา มากลุ่มละ 1 ชิ้น นำไปบันทึกภาพพื้นผิวที่กำลังขยาย 300x, 1000x และ 5000x และศึกษาขนาดเส้นผ่านศูนย์กลางค่าเฉลี่ยรูพรุนของวัสดุ (mean pore diameter) และรูพรุนทั้งหมด (total pore diameter) ด้วยเครื่องวิเคราะห์พื้นผิวและความพรุนของวัสดุ (Brunauer-Emmett-Teller: BET) โดยใช้ไหมเย็บแผลจากกลุ่มควบคุมและกลุ่มศึกษา มากลุ่มละ 5 กรัม โดยไหมเย็บแผลจะถูกนำมาทำกระบวนการคายแก๊ส (degassing) ที่อุณหภูมิ 80 องศาเซลเซียส เป็นเวลา 24 ชั่วโมง เพื่อให้ไหมอยู่ในสภาวะสุญญากาศ และแทนที่รูพรุนของไหมด้วยแก๊สไนโตรเจน จากนั้นทำการวัดปริมาณแก๊สไนโตรเจนที่ถูกดูดซับบนผิวของไหมเย็บแผล

2.3.3) การศึกษาการมีอยู่ของยาลีโวฟลอกซาซินในไหมเย็บแผล ทำการทดสอบหมู่ฟังก์ชันของสารโดยใช้ฟูเรียร์ทรานส์ฟอร์มอินฟราเรดสเปกโตรมิเตอร์ (Fourier Transform Infrared Spectrometer: FTIR รุ่น Tensor 27) และไมโครสโคป (Microscope รุ่น Hyperion 3000) โดยแสดงองค์ประกอบทางเคมีของตัวอย่างที่เปลี่ยนแปลงจากเดิมด้วยสเปกตรัมการมีอยู่ของหมู่พันธะคาร์บอนิล (Carbonyl C=O), หมู่พันธะ aromatic C-H และหมู่พันธะ O-H ของกลุ่มคาร์บอกซิล ของตัวยาลีโวฟลอกซาซินบนไหมเย็บ ซึ่งใช้จำนวนตัวอย่างกลุ่มละ 3 ชิ้น

2.4) การศึกษาคุณสมบัติทางชีวภาพของไหมเย็บแผล (Biological Properties)

2.4.1) การทดสอบการปลดปล่อยยา (Drug Releasing Profile) ศึกษาปริมาณยาโดยใช้ลีโวฟลอกซาซิน (Levofloxacin) ความเข้มข้น 3 มก./มล. ในรูปของสารละลายที่ปลดปล่อยออกจากไหมเย็บแผล โดยนำไหมเย็บแผลไปแช่ในน้ำปราศจากไอออน (deionized water) ปริมาตร 10 มล. ที่อุณหภูมิ 37 องศาเซลเซียส ทำการเก็บตัวอย่างสารละลายปริมาตร 1 มล. จากนั้นทำการเติมน้ำปราศจากไอออนกลับเข้าไปให้มีปริมาตรเท่าเดิม และทำการวัดความเข้มข้นของสารละลายลีโวฟลอกซาซินโดยใช้เครื่องวัดการดูดกลืนแสง (UV visible spectrometer) ที่ความยาวคลื่น 280 นาโนเมตร ทำการบันทึกผลโดยเก็บตัวอย่างสารละลายในช่วงเวลาที่ 0, 6, 12, 18, 24, 30, 36, 48 ชั่วโมง และต่อเนื่องไปจนถึง 7 วัน โดยมีกลุ่มตัวอย่าง กลุ่มละ 4 ชิ้น วัดซ้ำ 3 ครั้ง ทำการบันทึก

ผล ค่าที่ได้นำมาเทียบกับกราฟมาตรฐานของยาลีโวฟลอกซาซิน เพื่อคำนวณปริมาณยาต่อความยาวของไหมที่ใช้ (มล./30 ซม.)

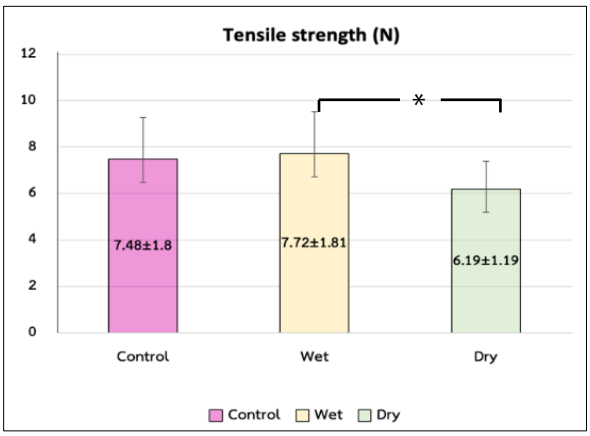
2.5) การวิเคราะห์ข้อมูล

วิเคราะห์ข้อมูลทางสถิติระหว่างกลุ่มด้วยโปรแกรม Stata/SE เวอร์ชัน 17 ทดสอบการแจกแจงแบบปกติด้วยสถิติ Shapiro-Wilk test จากนั้นนำมาวิเคราะห์ค่าความแตกต่างเฉลี่ยระหว่างทั้ง 3 กลุ่มด้วยความแปรปรวนทางเดียว (One-way ANOVA) และทดสอบความแตกต่างระหว่างกลุ่มด้วย Least Significant Difference (LSD) ทดสอบเปรียบเทียบระหว่างกลุ่มศึกษาด้วยสถิติทดสอบที (T-test) ที่ระดับนัยสำคัญ 0.05

3) ผลการวิจัยและอภิปรายผล

3.1) ผลของการทนแรงดึงของไหมเย็บแผลที่ชุบด้วยยาลีโวฟลอก-ซาซิน

จากรูปที่ 1 จะพบว่าไหมเย็บแผล ในกลุ่มควบคุม (control) มีค่าทนแรงดึงเฉลี่ย (tensile strength) ที่ 7.48 ± 1.8 นิวตัน ไม่แตกต่างอย่างมีนัยสำคัญทางสถิติเทียบกับกลุ่มที่มีค่าทนแรงดึงด้วยวิธีเปียก (wet) ที่ 7.72 ± 1.81 นิวตัน ค่าแรงดึงเฉลี่ยในไหมในกลุ่มแห้ง (dry) มีค่าทนแรงดึงที่ 6.19 ± 1.19 นิวตันซึ่งมีค่าเฉลี่ยน้อยกว่ากลุ่มเปียก ($p < 0.05$)



รูปที่ 1 : ผลการทนแรงดึงของไหมเย็บแผลที่ชุบด้วยยาลีโวฟลอกซาซิน (*=p<0.05)

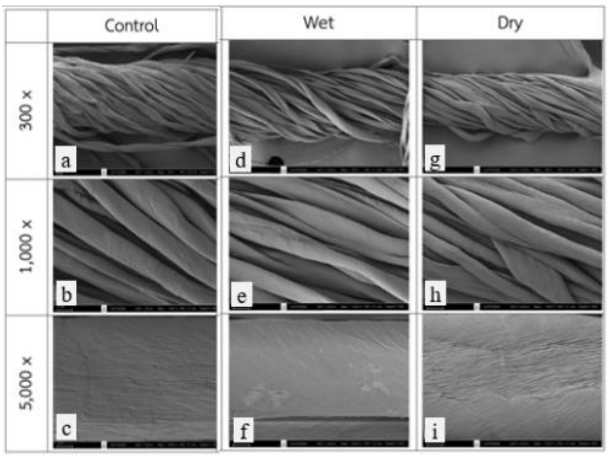
จากตารางที่ 1 จะแสดงค่าเฉลี่ยมอดูลัส (Young's modulus) ของไหมในกลุ่มควบคุมและกลุ่มศึกษา ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ จากตารางที่ 1 พบว่าในส่วนของค่าเฉลี่ยร้อยละของการยืดไหมจนขาด (break elongation) ในกลุ่มควบคุมจะสูงกว่าในกลุ่มทดลองแต่ไม่แตกต่างอย่างมีนัยสำคัญทางสถิติ เช่นเดียวกับค่าการยืดตัวของไหมเย็บ (elongation) จะพบว่าคุณสมบัติการทนแรงดึงของไหมกลุ่มทดลองทั้ง 2 กลุ่มมีคุณสมบัติทนแรงดึงที่ใกล้เคียงกับกลุ่มทดลอง

ตารางที่ 1 : ผลของการวัดค่าเฉลี่ยมอดูลัสของไหมเย็บแผลในกลุ่มควบคุมและกลุ่มที่ชุบด้วยยาลีโวฟลอกซาซิน ($p<0.05$)

กลุ่ม	ค่าเฉลี่ย Young's Modulus (GPa)	ค่าเฉลี่ย Break Elongation (%)	ค่าเฉลี่ย E- Elongation (GPa)
กลุ่มควบคุม	5.411 (±2.67)	9.589 (±2.88)	7.40 (±2.31)
กลุ่มเปียก	4.561 (±2.03)	7.38 (±3.12)	6.824 (±2.41)
กลุ่มแห้ง	4.928 (±2.74)	6.68 (±2.00)	6.565 (±1.68)

3.2) ผลการศึกษาลักษณะพื้นผิวและรูปร่าง 3 มิติของไหมเย็บ

จากการบันทึกภาพด้วยกล้องจุลทรรศน์แบบส่องกราด (Scanning Electron Microscope) ที่กำลังขยายต่าง ๆ ดังรูปที่ 2 จะพบว่าไหมเย็บแสดงลักษณะเป็นวัสดุหลายเส้นใยชนิดเกลียวแบบหยาบ ลักษณะพื้นผิวในกลุ่มแห้ง (dry) จะมีความหยาบขึ้นเล็กน้อย ซึ่งจะเห็นเป็นลักษณะเกลียวที่คลายตัวและมีลักษณะพื้นผิวที่ขรุขระมากขึ้น (รูปที่ 2i) เมื่อเทียบกับกลุ่มควบคุม (control) และกลุ่มเปียก (wet) ที่กำลังขยาย 5000x โดยลักษณะพื้นผิวโดยรวมของไหมที่ชุบด้วยสารละลายลีโวฟลอกซาซินไม่มีความเปลี่ยนแปลงเมื่อเทียบกับไหมเย็บปกติ



รูปที่ 2 : ภาพจากกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราดแสดงลักษณะพื้นผิวของไหมกลุ่มควบคุม (a-c); ไหมกลุ่มที่ชุบด้วยลีโวฟลอกซาซินกลุ่มเปียก (d-f); และไหมกลุ่มแห้ง (g-i)

3.3) ผลการศึกษาปริมาตรรูพรุนของไหมเย็บด้วยเครื่อง BET Analysis

จากผลค่าเฉลี่ยรูพรุน (mean pore diameter) และปริมาตรรูพรุน (total pore volume) ของไหมเย็บแผลดังตารางที่ 2 พบว่าค่าเฉลี่ยรูพรุนของไหมเย็บแผลกลุ่มเปียกมีค่ามากกว่ากลุ่มควบคุมและกลุ่มแห้ง และปริมาตรรูพรุนจะมีค่าลดลง สอดคล้องกับค่าเฉลี่ยเส้นผ่านศูนย์กลางที่เพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติเมื่อเทียบกับกลุ่มควบคุมและกลุ่มแห้ง เนื่องจากกลุ่มเปียกมีการเกิดการดูดซับทางกายภาพ (physisorption) กับสารละลายลีโว-ฟลอกซาซิน ทำให้ไหมเย็บมีการบวมและขนาดรูพรุนเพิ่มขึ้นตามการเกิดกระบวนการซึมผ่านรูเล็ก (capillary action) ที่ได้กล่าวในบทนำ ซึ่งอาจจะส่งผลให้มีการผ่านของเซลล์แบคทีเรียได้ง่ายขึ้น จากผลการทดลอง เนื่องจากไหมกลุ่มเปียกมีการดูดซับน้ำและปล่อยให้มีการแพร่ของเหลวแบบพื้นผิว (surface diffusion)

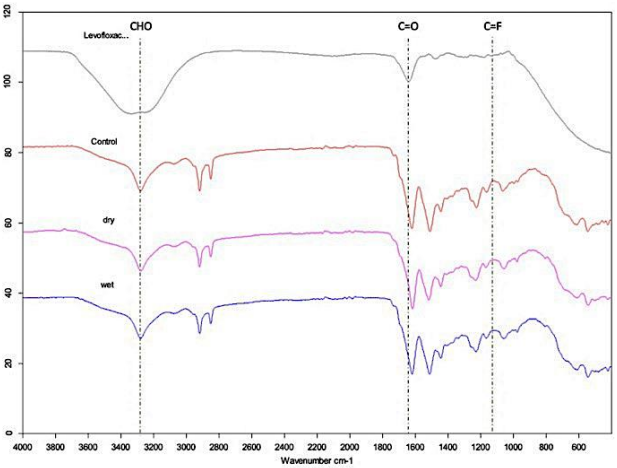
ตารางที่ 2 : ผลการศึกษาปริมาตรและขนาดรูพรุนของไหมเย็บแผล

กลุ่ม	ปริมาตรรูพรุน (Total pore volume) (10^3 cm^3/g-1)	ค่าเฉลี่ยรูพรุน (Mean pore diameter) (nm)
กลุ่มควบคุม	6.7964	6.9895
กลุ่มเปียก	3.6280*	10.229*
กลุ่มแห้ง	7.0804	5.775

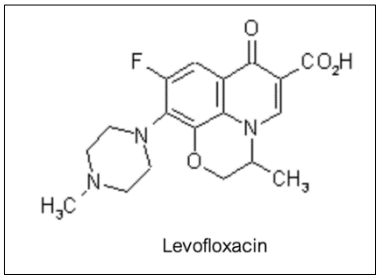
(= $p < 0.05$)*

3.4) ผลของการมีอยู่ของยาลีโวฟลอกซาซินในไหมเย็บแผล

ผลของการศึกษาค่าการดูดกลืนรังสีอินฟาเรดที่ความยาวคลื่นช่วง 400-4000 cm^{-1} พบว่าหมู่พันธะของยาลีโวฟลอกซาซินตามโครงสร้างในรูปที่ 4 จะมีลักษณะค่าดูดกลืนสูงสุดของหมู่คาร์บอนิล C=O จะอยู่ที่ช่วง 1690-1760cm^{-1} หมู่พันธะแอลเคน C-H จะมีค่าช่วงความถี่สูงที่ 3300 cm^{-1} และหมู่พันธะ O-H ของกลุ่มคาร์บอกซิล จะสูงสุดที่ 3265.81 cm^{-1} [14] สอดคล้องกับไหมเย็บทางทันตกรรมที่มีส่วนประกอบของไกลโคไซด์ (glycoside) เป็นหลัก [15] จากรูปที่ 3 จะเห็นลักษณะการดูดกลืนรังสีที่ใกล้เคียงกันในทั้ง 3 กลุ่ม เนื่องจากลักษณะหมู่พันธะของตัวยาลีโวฟลอกซาซินที่มีหมู่พันธะคาร์บอนิล C=O และ C=F จะมีค่าพีคของช่วงสเปคตรัมที่ค่อนข้างทับซ้อนกับช่วงสเปคตรัมของหมู่ฟังก์ชันไหมเย็บแผลทำให้มองเห็นการขยับของพีคในช่วงสเปคตรัมที่ตำแหน่งในกราฟที่ไม่เด่นชัดและตัวยาลีโวฟลอกซาซินอาจจะมีค่าความเข้มข้นที่น้อยเกินไป จึงต้องทำการทดสอบรูปแบบการปลดปล่อยยาเพิ่มเติม



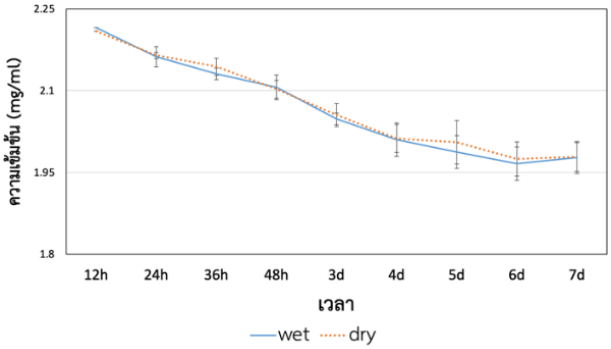
รูปที่ 3 : สเปคตรัมของ FTIR ของลีโวฟลอกซาซิน และไหมเย็บแผล (control) ที่ชุบด้วยลีโวฟลอกซาซินแบบเปียก (wet) และแบบแห้ง (dry)



รูปที่ 4 : โครงสร้างของยาลีโวฟลอกซาซิน [10]

3.5) ผลการศึกษารูปแบบการปลดปล่อยยา

จากรูปที่ 5 เนื่องจากไหมดูดซับสารละลายวิธี Physisorption มีแรงพันธะแวนเดอร์วาลที่ยึดกันแบบไม่มีทิศทางและแรงยึดต่ำ ไหมเย็บแผลในกลุ่มศึกษา (กลุ่มเปียกและกลุ่มแห้ง) จึงมีการปลดปล่อยยาลีโวฟลอกซาซินอย่างรวดเร็ว (burst release) ในช่วง 12–24 ชั่วโมงแรก และตามด้วยการปลดปล่อยอัตราคงที่ในปริมาณที่ต่อเนื่อง และลดลงเล็กน้อยจนถึงระยะเวลา 7 วัน การปลดปล่อยความเข้มข้นในไหมเย็บแผลกลุ่มเปียกและกลุ่มแห้งมีความเข้มข้นใกล้เคียงกัน (ตารางที่ 3) จากกราฟแสดงรูปแบบการปลดปล่อยยา จะพบว่าสารละลายลีโวฟลอกซาซินจะปล่อยความเข้มข้นที่คงที่ต่อเนื่องตั้งแต่ 12 ชั่วโมงแรกจนถึงระยะเวลา 7 วัน สอดคล้องกับการคงสภาพของไหมไว้ในช่องปากระยะเวลา 7 วัน และพบว่าไหมเย็บแผลในกลุ่มศึกษาไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ



รูปที่ 5 : รูปแบบการปลดปล่อยสารละลายลีโวฟลอกซาซิน (3 มก./มล.) จากไหมเย็บแผลในกลุ่มเปียก (wet = เส้นทึบ) และกลุ่มแห้ง (dry = เส้นประ, แกน x เวลา ; h= ชั่วโมง, d= วัน)

ตารางที่ 3 : ค่าเฉลี่ยความเข้มข้นของการปลดปล่อยสารละลายลีโวฟลอก-ซาซินจากไหมเย็บแผลในกลุ่มเปียก (wet) และกลุ่มแห้ง (dry) ที่เวลาต่างกันเป็นระยะเวลา 7 วัน

เวลา	กลุ่มเปียก (wet) Mean± SD (mg/ml)	กลุ่มแห้ง (dry) Mean± SD (mg/ml)
12 ชั่วโมง	2.216 ± 0.019	2.209 ± 0.006
24 ชั่วโมง	2.163 ± 0.012	2.165 ± 0.015
36 ชั่วโมง	2.131 ± 0.023	2.144 ± 0.016
48 ชั่วโมง	2.106 ± 0.013	2.103 ± 0.021
3 วัน	2.048 ± 0.031	2.055 ± 0.025
4 วัน	2.01 ± 0.034	2.012 ± 0.041
5 วัน	1.988 ± 0.031	2.005 ± 0.031
6 วัน	1.966 ± 0.029	1.97 ± 0.026
7 วัน	1.978 ± 0.009	1.979 ± 0.015

4) สรุปและวิจารณ์ผล

ไหมเย็บแผลในทางทันตกรรมมีความสำคัญในการเย็บและคงสภาพแผลหลังทำหัตถการผ่าตัด เนื่องจากเป็นต้องเจอกับสภาพแวดล้อมในช่องปากที่มีน้ำลาย และมีการส่งเสริมให้มีคราบต่าง ๆ จากการรับประทานอาหารเข้าไปเกาะได้ง่าย ปกติในช่องปากจะมีแบคทีเรียมาสะสม 750 ล้านเซลล์/มล. ซึ่งสามารถเกาะที่ไหมเย็บแผลได้ถึง 200 พันล้านเซลล์/กรัม [16], [17] ทำให้ จึงมีหลายการศึกษา [11], [18]–[20] ได้ลดข้อเสียดังกล่าว โดยการนำไหมเย็บแผลไปเคลือบกับสารต้านจุลชีพ เช่น เตตราไซคลินไฮโดรคลอไรด์ (Tetracycline Hydrochloride) หรือคลอเฮกซิดีน (Chlorhexidine) ปัจจุบันมีไหมละลายที่ผลิตจำหน่ายยี่ห้อ ETHICON coated vicryl® ที่เคลือบด้วยไตรโคลซาน (Triclosan) จำหน่าย รวมไปถึงการวิจัยในห้องปฏิบัติการที่ใช้ไหมชุบไตรโคลซานมีฤทธิ์ในการใช้ฆ่าเชื้อแบคทีเรียในห้องปฏิบัติการและในสัตว์ทดลอง [21], [22] นอกจากนี้ยังมีการวิจัยสังเคราะห์ไหมด้วยโพลีคาแลกโตนเคลือบด้วยยาต้านจุลชีพแอมพิซิลินด้วยวิธีอิเล็กโตรสปัน (Electrospun) เพื่อลดการติดเชื้อในแผลผ่าตัด [23] แต่ยังไม่มีงานวิจัยในคลินิกที่ใช้ไหมชุบยาปฏิชีวนะทางทันตกรรม ส่วนยาปฏิชีวนะลีโวฟลอก-ซาซินผู้วิจัยเลือกนำมาชุบกับไหมเย็บแผล เนื่องจากมีคุณสมบัติในการฆ่าเชื้อแบคทีเรีย และลดโอกาสในการดื้อยา และสามารถคงตัวได้นาน 3 วัน [11] งานวิจัยทางห้องปฏิบัติการพบว่ามีการใช้ลีโวฟลอกซาซินเคลือบกับโพลีเอทธิลีนไกลคอล (PEG) ด้วยวิธีอิเล็กโตรสปันแบบเปียกทำไหมเย็บแผลในงานวิจัยพบว่ามีความแข็งแรงและมีฤทธิ์ฆ่าเชื้อในงานจักษุศัลยกรรม รวมถึงสามารถมีการปลดปล่อยยาได้นาน 60 วัน [24]

จากผลการศึกษาพบว่าเมื่อนำมาทดสอบแรงดึงในกลุ่มควบคุมไม่มีความแตกต่างจากกลุ่มที่ชุบด้วยสารละลายลีโวฟลอกซาซินด้วยวิธีเปียก ซึ่งสามารถรักษาคุณสมบัติความแข็งแรงคงทนของไหมเย็บแผลได้ ในขณะที่กลุ่มที่เตรียมผิวแบบแห้งจะส่งผลต่อค่าความทนแรงดึงที่ลดลง สอดคล้องกับงานวิจัย Shuttleworth *et al.* [25] ที่เมื่อนำไปทำให้ปราศจากเชื้อด้วยการแช่ด้วยสารละลายกลูตารัลดีไฮด์เป็นเวลา 24 ชั่วโมง ที่คล้ายกับกลุ่มเปียกพบว่าค่าทนแรงดึงแตกต่างจากไหมเย็บแผลปกติ ส่วนงานวิจัยของ Nagaraja and Shetty [26] นำไหมเย็บไปทำให้ปราศจากเชื้อด้วยวิธีอบแก๊สเอธิลีนออกไซด์ ซึ่งคล้ายกับกลุ่มแห้งพบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติกับกลุ่มปกติ เมื่อนำไปวิเคราะห์ลักษณะพื้นผิวด้วยกล้องจุลทรรศน์อิเล็กตรอนแบบส่องกราด จะพบว่าลักษณะพื้นผิวในกลุ่มแห้งมีความหยาบขึ้นเล็กน้อยเมื่อดูที่กำลังขยาย 5000x ซึ่งอาจจะส่งผลให้เกิดการเกาะของแบคทีเรียที่เพิ่มขึ้น ผลการศึกษาปริมาตรของรูพรุนพบว่า เมื่อนำไหมเย็บไปชุบกับสารละลายลีโวฟลอกซาซิน แล้วนำมาเข้าสู่กระบวนการทำแห้งแบบระเหิดแบบกลุ่มแห้งนั้น ไม่ได้เพิ่มค่าเฉลี่ยรูพรุนมากขึ้น

เมื่อนำไปทดสอบการมีอยู่ของสารละลายลีโวฟลอกซาซินโดยดูจากหมู่ฟังก์ชันของยาพบว่าในกลุ่มแห้งมีการเปลี่ยนแปลงค่าดูดกลืนสูงสุดเล็กน้อย เนื่องจากข้อจำกัดของเครื่องมือกรณีวัสดุตัวอย่างมีการทับซ้อนของพีคในสเปคตรัมตัวยาซึ่งอาจจะเกิดจากความเข้มข้นที่น้อยของตัวยาหรือวัสดุมีน้ำเป็นองค์ประกอบทางผู้วิจัยจึงได้ศึกษารูปแบบการปลดปล่อยสารละลายลีโวฟลอก-ซาซินเพิ่มเติม พบว่าลักษณะการปลดปล่อยยาเป็นไปตามลักษณะ

จากกระบวนการเกิดการดูดซับทางกายภาพ (physisorption) คือช่วงแรกมีการปล่อยออกมามากที่สุดใน 12 ชั่วโมงแรก จากนั้นค่อย ๆ ลดลงต่อเนื่องเป็นระยะเวลา 7 วัน การปลดปล่อยสารละลายที่มีลีโวฟลอกซาซินในระยะเวลาที่คงที่ 7 วันสอดคล้องกับระยะเวลาของการคงอยู่ของไหมเย็บในปาก และช่วยป้องกันการเกาะของแบคทีเรีย ก่อนที่จะมีการสร้างสิ่งป้องกันบาดแผลของร่างกาย [11], [20]. ข้อจำกัดของงานวิจัยนี้คือยังไม่ได้ทดสอบทางชีวภาพคือประสิทธิภาพการฆ่าเชื้อแบคทีเรียในห้องปฏิบัติการ แต่งานวิจัยหลายชิ้นมีการทดสอบยาลีโวฟลอกซาซินหรือยาในกลุ่มฟลูโอโรควิโนโลน (Fluoroquinolones) เคลือบไหมเย็บแผลสามารถมีฤทธิ์ฆ่าเชื้อแบคทีเรียกลุ่ม สแตฟิโลค็อกคัส ออเรียส (*S. aureus*) และ สแตฟิโลค็อกคัส เอพิเดอร์มิดิส (*S. epidermidis*) ได้ [24], [27], [28]

ดังนั้นการใช้ไหมเย็บแผลปกติที่นำไปชุบด้วยยาปฏิชีวนะลีโวฟลอกซาซิน มีความแข็งแรงที่ใกล้เคียงกับจากไหมเย็บแผลปกติ และสามารถใช้ประโยชน์ในการปลดปล่อยยาเพื่อลดการติดเชื้อบริเวณแผลได้ในระยะที่บาดแผลเริ่มกระบวนการหายและส่งเสริมความแข็งแรงที่สอดคล้องกับระยะเวลาในการเย็บ ซึ่งไหมเย็บแผลปกติสามารถหาได้สะดวก นิยมใช้อย่างแพร่หลายและราคาไม่แพง ซึ่งไหมเย็บแผลที่ผลิตในท้องตลาดที่ชุบด้วยยาปฏิชีวนะปัจจุบัน มีราคาแพงและไม่ได้มีในทุกโรงพยาบาล ซึ่งจากการใช้ไหมเย็บแผลแช่ด้วยยาปฏิชีวนะในงานวิจัยนี้สามารถนำไปต่อยอดทางคลินิกเพื่อประโยชน์ในผู้ป่วยที่เสี่ยงต่อการติดเชื้อที่มากกว่าคนปกติได้

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์ของเจ้าหน้าที่และผู้ช่วยวิจัย ประจำหน่วยงานอาคารเครื่องมือ 10 - 12 มหาวิทยาลัยเทคโนโลยีสุรนารี สำหรับคำแนะนำ และเอื้อเฟื้อสถานที่ รวมไปถึงมารดาของข้าพเจ้าที่ช่วยเหลือและให้กำลังใจในการศึกษาวิจัยนี้

REFERENCES

[1] J. W. Alexander, J. Z. Kaplan, and W. A. Altemeier, “Role of suture materials in the development of wound infection,” *Ann. Surg.*, vol. 165, no. 2, pp. 192–199. Feb. 1967, doi: 10.1097/00000658-196702000-00005.

[2] C. K. S. Pillai and C. P. Sharma, “Review paper: Absorbable polymeric surgical sutures: Chemistry, production, properties, biodegradability, and performance,” *J. Biomater. Appl.*, vol. 25, no. 4, pp. 291–366, Nov. 2010, doi: 10.1177/0885328210384890.

[3] B. Blomstedt and B. Österberg, “Physical properties of suture materials which influence wound infection,” in *Moderne Nahtmaterialien und Nahttechniken in der Chirurgie*, A. Thiede and H. Hamelmann, Eds., Heidelberg, Germany: Springer Berlin Heidelberg, 1982, pp. 39–46, doi: 10.1007/978-3-642-68736-5_6.

[4] T. Grigg, F. Liewehr, W. Patton, T. Buxton, and J. Mcpherson, “Effect of the wicking behavior of multifilament sutures,” *J. Endod.*, vol. 30, no. 9, pp. 649–652, Sep. 2004, doi: 10.1097/01.DON.0000121617.67923.05.

[5] E. Laas *et al.*, “Antibacterial-Coated suture in reducing surgical site infection in breast surgery: A prospective study,” *Int. J. Breast Cancer*, vol. 2012, pp. 1–8, 2012, doi: 10.1155/2012/819578.

[6] A. E. Deliaert *et al.*, “The effect of triclosan-coated sutures in wound healing. A double blind randomised prospective pilot study,” *J. Plast. Reconstr. Aesthet. Surg.*, vol. 62, no. 6, pp. 771–773, Jun. 2009, doi: 10.1016/j.bjps.2007.10.075.

[7] S. Suttapreyasri *et al.*, “Physicochemical and Antimicrobial Properties of Silk Suture Soaked in Chlorhexidine Gluconate,” (in Thai), *J. Dent. Assoc. Thai.*, vol. 67, no. 1, pp. 77–90, 2017, doi: 10.14456/JDAT.2017.7.

[8] S. Viju and G. Thilagavathi, “Characterization of tetracycline hydrochloride drug incorporated silk sutures,” *J. Text. Inst.*, vol. 104, no. 3, pp. 289–294, Mar. 2013, doi: 10.1080/00405000.2012.720758.

[9] B. G. Katzung, S. B. Masters, and A. J. Trevor, *Basic & Clinical Pharmacology*, 12th ed. New York, NY, USA: McGraw-Hill Medical, 2012.

[10] D. N. Fish and A. T. Chow, “The Clinical Pharmacokinetics of Levofloxacin,” *Clin. Pharmacokinet.*, vol. 32, no. 2, pp. 101–119, Feb. 1997, doi: 10.2165/00003088-199732020-00002.

[11] X. Chen, D. Hou, L. Wang, Q. Zhang, J. Zou, and G. Sun, “Antibacterial surgical silk sutures using a high-performance slow-release carrier coating system,” *ACS Appl. Mater. Interfaces*, vol. 7, no. 40, pp. 22394–22403, Oct. 2015, doi: 10.1021/acsami.5b06239.

[12] K. S. Parikh *et al.*, "Ultra-thin, high strength, antibiotic-eluting sutures for prevention of ophthalmic infection," *Bioeng. Transl. Med.*, vol. 6, no. 2, p. e10204, May 2021, doi: 10.1002/btm2.10204.

[13] C. Dennis, S. Sethu, S. Nayak, L. Mohan, Y. Y. Morsi, and G. Manivasagam, "Suture materials - Current and emerging trends: Suture Materials," *J. Biomed. Mater. Res. A*, vol. 104, no. 6, pp. 1544–1559, Jun. 2016, doi: 10.1002/jbm.a.35683.

[14] H. Gevariya, N. Patel, and S. Gami, "Formulation and characterization of levofloxacin-loaded biodegradable nanoparticles," *Asian J. Pharm.*, vol. 5, no. 2, p. 114, 2011, doi: 10.4103/0973-8398.84552.

[15] A. Faris *et al.*, "Characteristics of suture materials used in oral surgery: Systematic review," *Int. Dent. J.*, vol. 72, no. 3, pp. 278–287, Jun. 2022, doi: 10.1016/j.identj.2022.02.005.

[16] C. E. Edmiston *et al.*, "Bacterial adherence to surgical sutures: Can Antibacterial-Coated sutures reduce the risk of microbial contamination?," *J. Am. Coll. Surg.*, vol. 203, no. 4, pp. 481–489, Oct. 2006, doi: 10.1016/j.jamcollsurg.2006.06.026.

[17] G. Banche *et al.*, "Microbial adherence on various intraoral suture materials in patients undergoing dental surgery," *J. Oral Maxillofac. Surg.*, vol. 65, no. 8, pp. 1503–1507, Aug. 2007, doi: 10.1016/j.joms.2006.10.066.

[18] J.-E. Otten, M. Wiedmann-Al-Ahmad, H. Jahnke, and K. Pelz, "Bacterial colonization on different suture materials—A potential risk for intraoral dentoalveolar surgery," *J. Biomed. Mater. Res. B Appl. Biomater.*, vol. 74B, no. 1, pp. 627–635, Jul. 2005, doi: 10.1002/jbm.b.30250.

[19] S. Rothenburger, D. Spangler, S. Bhende, and D. Burkley, "In Vitro Antimicrobial evaluation of coated Vicryl* plus antibacterial suture (coated polyglactin 910 with triclosan) using zone of inhibition assays," *Surg. Infect.*, vol. 3. no. s1, pp. s79–s87, Dec. 2002.

[20] X. Ming, S. Rothenburger, and M. M. Nichols, "In vivo and in vitro antibacterial efficacy of PDS plus (polidioxanone with triclosan) Suture," *Surg. Infect.*, vol. 9, no. 4, pp. 451–457, Aug. 2008, doi: 10.1089/sur.2007.061.

[21] J. McCagherty *et al.*, "Investigation of the in vitro antimicrobial activity of triclosan-coated suture material on bacteria commonly isolated from wounds in dogs," *Am. J. Vet. Res.*, vol. 81, no. 1, pp. 84–90, Jan. 2020, doi: 10.2460/ajvr.81.1.84.

[22] J. M. Adkins, R. A. Ahmar, H. D. Yu, S. T. Musick, and A. M. Alberico, "Comparison of antimicrobial activity between bacitracin-soaked sutures and triclosan coated suture," *J. Surg. Res.*, vol. 270, pp. 203–207, Feb. 2022, doi: 10.1016/j.jss.2021.09.010.

[23] H. Liu, K. K. Leonas, and Y. Zhao, "Antimicrobial properties and release profile of ampicillin from Electrospun poly (εepsilon;-caprolactone) Nanofiber Yarns," *J. Eng. Fibers Fabr.*, vol. 5, no. 4, pp. 10–19, Dec. 2010, doi: 10.1177/155892501000500402.

[24] F. Kashiwabuchi *et al.*, "Development of Absorbable, Antibiotic-Eluting Sutures for Ophthalmic Surgery," *Transl. Vision Sci. Technol.*, vol. 6, no. 1, pp. 1–8, Jan. 2017, doi: 10.1167/tvst.6.1.1.

[25] G. N. Shuttleworth, L. F. Vaughn, and H. B. Hoh, "Material properties of ophthalmic sutures after sterilization and disinfection," *J. Cataract Refractive Surg.*, vol. 25, no. 9, pp. 1270–1274, Sep. 1999, doi: 10.1016/S0886-3350(99)00156-X.

[26] P. A. Nagaraja and D. Shetty, "Effect of re-sterilization of surgical sutures by ethylene oxide," *Aust. N. Z. J. Surg.*, vol. 77, pp. 80–83, Jan. 2007, doi: 10.1111/j.1445-2197.2006.03560.x.

[27] P. Janiga, B. Elayarajah, R. Rajendran, R. Rammohan, B. Venkatrajah, and S. Asa, "Drug-eluting silk sutures to retard post-operative surgical site infections," *J. Ind. Text.*, vol. 42, no. 2, pp. 176–190, Oct. 2012, doi: 10.1177/1528083711432948.

[28] M. Sriyai *et al.*, "Development of an antimicrobial-coated absorbable monofilament suture from a medical-grade poly (L -lactide- *co* -ε-caprolactone) copolymer," *ACS Omega*, vol. 6, no. 43, pp. 28788–28803, Nov. 2021, doi: 10.1021/acsomega.1c03569.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

การจำแนกความพึงพอใจบนหลักการของการถ่วงน้ำหนักคำด้วยคลาสมิวชวลอินฟอร์เมชัน

อุไรวรรณ บัวตูม[1] กัลยารัตน์ เชี่ยวชาญ[2] วชิราภรณ์ ศรีพุทธ[3] สมบัติ ฝอยทอง[4*]

[1,4]สาขาปัญญาประดิษฐ์ประยุกต์และเทคโนโลยีอัจฉริยะ คณะวิทยาศาสตร์และศิลปศาสตร์, มหาวิทยาลัยบูรพา วิทยาเขตจันทบุรี, จันทบุรี, ประเทศไทย*

[2]สาขาเทคโนโลยีสารสนเทศและวิทยาการข้อมูล คณะวิทยาศาสตร์และศิลปศาสตร์, มหาวิทยาลัยบูรพา วิทยาเขตจันทบุรี, จันทบุรี, ประเทศไทย

[3]สาขาบริหารธุรกิจ คณะวิทยาศาสตร์และศิลปศาสตร์, มหาวิทยาลัยบูรพา วิทยาเขตจันทบุรี, จันทบุรี, ประเทศไทย

*ผู้ประพันธ์บรรณกิจ อีเมล : sombut@buu.ac.th

รับต้นฉบับ : 30 พฤษภาคม 2566; รับบทความฉบับแก้ไข : 4 สิงหาคม 2566; ตอบรับบทความ : 5 สิงหาคม 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

การพัฒนาอย่างรวดเร็วของเทคโนโลยีสารสนเทศและแพลตฟอร์มออนไลน์ทำให้การสั่งซื้อสินค้าผ่านระบบออนไลน์จึงเป็นที่นิยม ในปัจจุบันนี้เนื้อหาหรือข้อความของการโพสต์ของลูกค้าจึงมีความสำคัญต่อผู้ผลิตสินค้าในการปรับปรุงคุณภาพของสินค้าให้ตรงกับความต้องการของลูกค้า การวิเคราะห์ความพึงพอใจจากข้อความถือว่าเป็นกระบวนการอย่างหนึ่งของการประมวลผลภาษาธรรมชาติที่สามารถใช้ในการวิเคราะห์เพื่อสะท้อนความรู้สึกและทัศนคติของผู้ใช้ที่มีต่อสินค้าเชิงอารมณ์ได้ทั้งด้านบวกหรือลบ งานวิจัยทางด้านการจำแนกประเภทความพึงพอใจและการจำแนกประเภทเอกสารโดยส่วนใหญ่จะนิยมใช้การกำหนดค่าน้ำหนักคำ ด้วยค่าความถี่เอกสารผกผัน (idf) อย่างไรก็ตามการกำหนดค่าน้ำหนักคำด้วยค่าความถี่เอกสารผกผันอย่างเดียวอาจจะมีประสิทธิภาพไม่เพียงพอในการจำแนกประเภทความพึงพอใจเนื่องจากค่าน้ำหนักนี้ไม่ได้พิจารณามุมมองการจำแนกประเภทข่าวประกอบด้วย งานวิจัยนี้จึงนำเสนอการปรับค่าน้ำหนักคำแบบมีผู้สอนด้วยการนำค่ามิวชวลอินฟอร์เมชันของคลาสคำนวณร่วมกับความถี่ของคำและความถี่เอกสารผกผัน ผลการทดลองแสดงให้เห็นว่าวิธีการที่นำเสนอมีประสิทธิภาพที่ดีกว่าการแจกแจงเทอมและการกำหนดค่าน้ำหนักคำด้วยค่าความถี่เอกสารผกผันอย่างเดียว เมื่อพิจารณาจากค่าตัวชี้วัดประสิทธิภาพ: ค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึก และค่าความถ่วงดุล

คำสำคัญ : มิวชวลอินฟอร์เมชัน การทำเหมืองความคิดเห็น การวิเคราะห์ความพึงพอใจ การปรับค่าน้ำหนักคำ การจำแนกประเภทเอกสาร

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

Sentiment Classification Based on Term Weighting with Class-mutual Information

Uraiwan Buatoom[1] Kanyarat Cheawchan[2] Vajiraporn Sriput[3] Sombut Foithong[4*]

[1,4*]*Applied Artificial Intelligent and Smart Technology Program, Faculty of Science and Arts, Burapha University, Chanthaburi Campus, Chanthaburi, Thailand*
[2]*Information Technology and Data Science Program, Faculty of Science and Arts, Burapha University, Chanthaburi Campus, Chanthaburi, Thailand*
[3]*Business Administration Program, Faculty of Science and Arts, Burapha University, Chanthaburi Campus, Chanthaburi, Thailand*

*Corresponding Author. E-mail address: sombut@buu.ac.th

Received: 30 May 2023; Revised: 4 August 2023; Accepted: 5 August 2023
Published online: 27 December 2023

Abstract

Online platforms and information technology are developing quickly, which boosts the popularity of online e-commerce. Nowadays, posting content from client sentiments is vital for product makers to improve product quality as much as possible to exceed customers' expectations. Sentiment analysis is a process of natural language processing that finds out the sentiment and attitudes of users towards a product, whether positive or negative. Most sentiment and text classification research uses term weighting with inverse document frequency (idf). However, assigning term weights using the idf method alone may not be effective enough to classify sentiment because this weight does not consider vital information classification views. This paper presents a supervised term weighting using the class mutual information calculated with the term frequency and the inverse document frequency. Experimental results show that the proposed method performs more effectively than term distribution and the term weighting that uses only the inverse document frequency when considering by the performance indicator value: Accuracy, Precision, Recall and F1-Measure.

***Keywords*:** Mutual information, Opinion mining, Sentiment analysis, Term weighting, Text classification

1) บทนำ

ด้วยการพัฒนาอย่างรวดเร็วของเทคโนโลยีสารสนเทศและแพลตฟอร์มออนไลน์ในปัจจุบัน การสั่งซื้อสินค้าผ่านระบบออนไลน์จึงเป็นที่นิยมอย่างแพร่หลายในปัจจุบัน ผู้สั่งซื้อสินค้าสามารถโพสต์แสดงความคิดเห็นคุณภาพของสินค้าและบริการของผู้จัดจำหน่าย ดังนั้นเนื้อหาการโพสต์แสดงความรู้สึกของลูกค้าจึงมีความสำคัญอย่างยิ่งต่อผู้ผลิตสินค้าและผู้สั่งซื้อสินค้าเพื่อใช้ในการปรับปรุงคุณภาพของสินค้าและบริการให้ตรงกับความต้องการของลูกค้ามากที่สุด นอกจากนี้บทวิจารณ์สินค้าและบริการที่ลูกค้าโพสต์ไว้จำนวนมากตามเว็บไซต์การทำธุรกิจผ่านสื่ออิเล็กทรอนิกส์ได้กลายเป็นแหล่งข้อมูลอันมีค่าสำหรับการวิเคราะห์ทางการตลาด บทวิจารณ์เหล่านี้เมื่อนำมาวิเคราะห์ความพึงพอใจ (เป็นเชิงบวก หรือ เชิงลบ) สามารถนำมาใช้เพื่อกำหนดกลยุทธ์ทางธุรกิจของเว็บไซต์การทำธุรกิจผ่านสื่ออิเล็กทรอนิกส์ต่าง ๆ เช่น Amazon.com และ Epinion.com [1] ซึ่งผู้ใช้ออนไลน์สามารถได้รับประโยชน์จากการอ่านความคิดเห็นที่มีประโยชน์ของผู้อื่นผ่านบทวิจารณ์ต่าง ๆ ที่โพสต์ได้

การวิเคราะห์ความพึงพอใจนั้นจะเป็นการศึกษาข่าวสารในส่วนของข้อความ อารมณ์ ความคิดเห็น หรือทัศนคติที่มีต่อสินค้าและบริการ ซึ่งเป็นข่าวสารที่มีความเป็นส่วนตัว [2] โดยทำการจัดหมวดหมู่ข้อความว่าเป็นเชิงบวก หรือ เชิงลบ ต่อผู้ผลิตและผู้ให้บริการ การวิเคราะห์ความพึงพอใจนั้นสามารถแบ่งออกได้เป็น 3 ระดับใหญ่ ๆ ได้แก่ ระดับเอกสาร ระดับประโยค และระดับแง่มุมเชิงลึก [3] โดยในระดับเอกสาร [4] จะเป็นการพิจารณาดูว่าข้อความคิดเห็นที่อยู่ในเอกสารนั้นเป็นความคิดเห็นที่เป็นเชิงบวก หรือ เชิงลบ ซึ่งในระดับนี้จะมองว่าข้อมูลในเอกสารเป็นหน่วยของข่าวสาร [5] ที่จะนำมาวิเคราะห์ ส่วนในระดับประโยค [6] นั้นเป้าหมายเพื่อจัดกลุ่มความพึงพอใจในระดับที่เล็กลงกว่าระดับเอกสารโดยพิจารณาแต่ละประโยคของผู้โพสต์ สิ่งสำคัญประการแรกที่ต้องดำเนินการ คือพิจารณาดูว่าความคิดเห็นของผู้โพสต์นั้นเป็นแบบมุมมองส่วนตัว (subjective) หรือ มุมมองที่อยู่บนข้อเท็จจริง (objective) ซึ่งถ้าประโยคนั้นเป็นความคิดเห็นในมุมมองส่วนตัว แล้วการวิเคราะห์ความพึงพอใจในระดับประโยคก็จะพิจารณาดูว่าเป็นความคิดเห็นเชิงบวก หรือ ลบ [7] ในขณะที่ระดับแง่มุมเชิงลึกก็จะแตกต่างต่างจากทั้ง 2 ระดับโดยจะมีความซับซ้อนมากกว่าในการวิเคราะห์ความพึงพอใจซึ่งมีเป้าหมายเพื่อพิจารณาจัดกลุ่มประเภทความพึงพอใจว่าเป็นแง่มุมที่เป็นเป้าหมายของเนื้อหาหรือไม่ [8]

การจำแนกประเภทความพึงพอใจเป็นงานของการจำแนกประเภทเอกสารภาษาธรรมชาติแบบอัตโนมัติเพื่อระบุหมวดหมู่ที่ถูกต้องของเอกสาร โดยการเปลี่ยนข้อความที่แสดงของเอกสารไปเป็นรูปแบบที่มีความกระชับที่สามารถทำให้เอกสารเหล่านั้นสามารถจัดหมวดหมู่โดยใช้คอมพิวเตอร์ได้ ซึ่งโดยทั่วไปแล้วนิยมใช้โมเดลของปริภูมิเวกเตอร์ (Vector space model) ในการแสดงเนื้อหาของเอกสารในรูปของเวกเตอร์ในปริภูมิของคำ (term space) ตัวอย่างเช่น $d = (w_1, w_2, w_3,..., w_k)$ โดยที่ k เป็นจำนวนของคำในเอกสาร โดยทั่วไปคำที่แตกต่างกันก็จะมีความสำคัญที่แตกต่างกันในข้อความ ดังนั้นตัวบ่งชี้ที่สำคัญ w_i จะเป็นเครื่องแสดงว่าคำ t_i มีอิทธิพลมากน้อยแค่ไหนต่อเอกสาร d วิธีการกำหนดค่าน้ำหนักคำเป็นขั้นตอนที่สำคัญที่จะปรับปรุงประสิทธิภาพการจำแนกประเภทความพึงพอใจ และการจัดหมวดหมู่ของเอกสารด้วยการกำหนดค่าน้ำหนักที่เหมาะสมให้กับคำ ถึงแม้ว่าการจัดหมวดหมู่ของเอกสารมีการศึกษาอย่างมากในช่วงหลายทศวรรษที่ผ่านมาก็ตาม แต่การปรับค่าน้ำหนักคำโดยปกติเป็นการยืมหลักการมาจากการสืบค้นเอกสาร (information retrieval) เช่น รูปแบบง่ายที่สุด คือการนำเสนอแบบทวิภาคและที่นิยมใช้กันมากที่สุดคือ ความถี่ของคำ-ความถี่เอกสารผกผัน (tf.idf) ซึ่งทั้งสองหลักการนี้เป็นการกำหนดค่าน้ำหนักคำแบบไม่มีผู้สอน (unsupervised term weighting)

ในการจำแนกประเภทเอกสารนั้นการปรับค่าน้ำหนักคำด้วย tf.idf ถือว่าประสบความสำเร็จมากที่สุดมีการนำมาประยุกต์ใช้อย่างแพร่หลาย [9]–[12] ปัญหาการจำแนกประเภทความพึงพอใจจัดว่าเป็นหัวข้อหนึ่งของปัญหาการจำแนกประเภทเอกสารบนหลักการพิจารณาเป็นหัวข้อ ซึ่งอัลกอริทึมการจำแนกประเภทเอกสารต่าง ๆ สามารถนำมาประยุกต์ใช้กับการจำแนกประเภทความพึงพอใจได้เช่นกัน ในการศึกษาการจำแนกประเภทความพึงพอใจมีงานวิจัยจำนวนมากได้นำหลักการ tf.idf [11]–[13] มาทำการปรับค่าน้ำหนักคำเช่นเดียวกับการจำแนกประเภทเอกสารเพื่อศึกษาถึงประสิทธิภาพของวิธีการต่าง ๆ ที่นำเสนอ บางงานวิจัยชี้ให้เห็นว่าหัวข้อหนึ่ง ๆ มีแนวโน้มที่จะถูกเน้นด้วยคำหลักบางคำที่เกิดขึ้น และความรู้สึกโดยรวมอาจไม่ถูกเน้นผ่านการใช้คำเดิม ๆ ซ้ำ ๆ ก็ได้ [13] แสดงว่าการปรับค่าน้ำหนักคำด้วยวิธีการ tf.idf อย่างเดียวอาจจะมีประสิทธิภาพไม่เพียงพอในการ

จำแนกประเภทความพึงพอใจ เนื่องจากทั้งความถี่ของคำและความถี่เอกสารผกผันจะพิจารณาเฉพาะความถี่ของคำที่เกิดขึ้นเป็นสำคัญ ซึ่งความถี่ของคำที่บ่งชี้ถึงทัศนคติด้านบวก หรือ ด้านลบ ของข้อความก็อาจจะมีความถี่โดยส่วนใหญ่ไม่สูงมากนัก

มีหลายงานวิจัย [9], [10] ใช้วิธีการปรับค่าน้ำหนักคำแบบมีผู้สอน (supervised term weighting) โดยการใช้การแจกแจงเทอมเพื่อเพิ่มประสิทธิภาพการจำแนกประเภทเอกสารให้มีประสิทธิภาพสูงขึ้นโดยนำหลักการทางสถิติมาประยุกต์ใช้ซึ่งประกอบด้วยการคำนวณ ค่าเบี่ยงเบนมาตรฐานในเอกสารที่รวบรวม (in-collection deviation) ค่าเบี่ยงเบนภายในคลาส (intra-class deviation) และค่าเบี่ยงเบนระหว่างคลาส (inter-class deviation) อย่างไรก็ตามการแจกแจงเทอมค่อนข้างใช้เวลานานในการคำนวณเนื่องจากมีสูตรที่ซับซ้อน [9], [10] นอกจากนี้เมื่อนำเทคนิคเหล่านี้มาประยุกต์ใช้กับการวิเคราะห์ความพึงพอใจที่แต่ละบทวิจารณ์มีความยาวคำที่ไม่มากและไม่ค่อยซ้ำกัน เหมือนกับเอกสารอื่น ๆ อย่างเช่น หนังสือ หรือ หนังสือพิมพ์ จึงทำให้ประสิทธิภาพการจำแนกประเภทความพึงพอใจไม่สูงมาก [13]

มิวชวลอินฟอร์เมชัน (Mutual Information) เป็นทฤษฎีที่สำคัญในการวัดข่าวสาร นิยมนำมาประยุกต์ใช้ในการหาความสัมพันธ์ระหว่างคุณลักษณะ และคลาสของข้อมูล เพื่อพิจารณาว่าคุณลักษณะใดของข้อมูล มีพลังอำนาจในการจำแนกประเภทข้อมูลสูง คุณลักษณะใดที่ไม่จำเป็นในการจำแนกข้อมูล ข้อดีของมิวชวลอินฟอร์เมชันคือ มีความทนทานต่อสัญญาณรบกวนและการแปลงพิกัด รวมทั้งมีรูปแบบการคำนวณที่ไม่ซับซ้อน นักวิจัยจำนวนมากจึงนิยมนำมาใช้ในงานด้านการเลือกลักษณะเด่นของข้อมูล (feature selection) [14]–[17] นอกจากนี้ในงานวิจัยด้านการจำแนกประเภทเอกสารยังนิยมนำมิวชวลอินฟอร์เมชันมาช่วยเพิ่มประสิทธิภาพในการสืบค้นเอกสารให้มีค่าความถูกต้องในการทำนายและค่าตัวชี้วัดต่าง ๆ สูงขึ้น โดยนำมาใช้ในการลดจำนวนคำที่ไม่จำเป็นซึ่งทำให้มิติของเอกสารลดลง และช่วยลดเวลาในการประมวลผลการสืบค้นเอกสาร [15]–[17]

จากข้อดีของมิวชวลอินฟอร์เมชันที่กล่าวมาข้างต้น ในงานวิจัยจึงนี้ได้นำเสนอแนวคิดใหม่ในการนำมิวชวลอินฟอร์เมชันมาประยุกต์ใช้ในการปรับค่าน้ำหนักคำในการสืบค้นเอกสารบทวิจารณ์ความพึงพอใจของลูกค้าโดยพิจารณาให้ความสำคัญกับคำต่าง ๆ ที่มีพลังอำนาจในการจำแนกประเภทความพึงพอใจ ซึ่งการปรับค่าน้ำหนักคำด้วยวิธีการ tf.idf เพียงอย่างเดียวนั้นอาจจะมีประสิทธิภาพไม่เพียงพอในการประยุกต์ใช้กับการจำแนกประเภทความพึงพอใจบทวิจารณ์ของผู้ใช้ที่มีต่อผลิตภัณฑ์และบริการ เนื่องจากเป็นหลักการที่ให้ความสำคัญกับความถี่ของคำที่เกิดขึ้นเท่านั้น จึงมองไม่เห็นข่าวสารที่สำคัญที่ใช้ในการจำแนกประเภทบทวิจารณ์ความพึงพอใจของลูกค้าออนไลน์ ดังนั้นงานวิจัยนี้จึงนำเสนอการปรับค่าน้ำหนักคำแบบมีผู้สอนโดยใช้คลาสของมิวชวลอินฟอร์เมชัน [14] ในการเพิ่มประสิทธิภาพการจำแนกประเภทความพึงพอใจโดยการปรับค่าน้ำหนักคำ tf.idf โดยคำนวณร่วมกับคลาสของมิวชวลอินฟอร์เมชัน หลักการที่นำเสนอนี้จะใช้เวลาในการคำนวณที่น้อยกว่าหลักการแจกแจงเทอมอย่างมาก นอกจากนี้ยังมีประสิทธิภาพที่ดีกว่าในการจำแนกความถูกต้องของการจัดประเภทความพึงพอใจว่าเป็นด้านบวก หรือ ด้านลบ

ส่วนที่เหลือของงานวิจัยมีดังนี้ คือ ส่วนที่ 2 เป็นบทสรุปของมิวชวลอินฟอร์เมชันและการแจกแจงเทอมทั้ง 3 ชนิด รวมถึงวิธีการอื่น ๆ ที่นำมาเปรียบเทียบ ส่วนที่ 3 แสดงปัญหาที่เกิดขึ้นของ idf และอธิบายหลักการสำหรับวิธีการที่นำเสนอ ส่วนที่ 4 เปรียบเทียบผลลัพธ์ของวิธีการที่นำเสนอกับวิธีการอื่น ๆ ด้วยการเปรียบเทียบค่าตัวชี้วัด ค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึก และ ค่าความถ่วงดุล และส่วนของการสรุปผลและหัวข้อการวิจัยในอนาคตที่เป็นไปได้จะกล่าวถึงในหัวข้อที่ 5

2) ทฤษฎีพื้นฐาน

2.1) เอนโทรปีและมิวชวลอินฟอร์เมชัน (Entropy and Mutual Information)

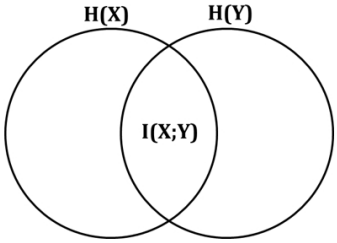
เครื่องมือที่นักวิจัยนิยมใช้วัดความไม่แน่นอน หรือ ความยุ่งเหยิงของตัวแปรสุ่มคือ เอนโทรปี ถ้ากำหนดให้ตัวแปรสุ่มแบบไม่ต่อเนื่อง X และ Y ที่มีเซ็ตของผลลัพธ์ที่เป็นไปได้ทั้งหมด คือ Ω และ Ψ ตามลำดับ และกำหนดฟังก์ชั่นความหนาแน่นความน่าจะเป็นของ X คือ $p(x) = \Pr\{X = x\}, x\epsilon\Omega$ ดังนั้นการคำนวณหาเอนโทรปีของตัวแปรสุ่ม X สามารถนิยามเป็น

$$H(X) = -\sum_{x\epsilon\Omega} p(x)\, log\, p(x) \qquad (1)$$

สำหรับตัวแปรสุ่มแบบไม่ต่อเนื่องสองตัว คือ X และ Y ทีมีการแจกแจงความน่าจะเป็นร่วมกัน p(x, y) ดังนั้นในการวัดผลรวมของข่าวสารที่ตัวแปรสุ่มหนึ่งเกี่ยวข้องกับอีกตัวแปรสุ่มจะคำนวณหาโดยมิวชวลอินฟอร์เมชัน ซึ่งสามารถคำนวณได้ดังนี้

$$I(X;Y) = \sum_{x\epsilon\Psi}\sum_{y\epsilon\Omega} p(x,y)\log\frac{p(x,y)}{p(x)\cdot p(y)} \qquad (2)$$

ถ้ามิวชวลอินฟอร์เมชันระหว่าง 2 ตัวแปรมีค่ามาก (น้อย) หมายถึงว่า ตัวแปรทั้งสองจะมีความสัมพันธ์ที่ใกล้ชิดกัน (ไม่มีความสัมพันธ์ใกล้ชิดกัน) ถ้ามิวชวลอินฟอร์เมชันมีค่าเป็นศูนย์แล้วตัวแปรสุ่มทั้งสองตัวจะไม่เกี่ยวข้องกันอย่างสมบูรณ์ ในขณะที่ รายละเอียดความสัมพันธ์ระหว่างเอนโทรปีและมิวชวลอินฟอร์เมชัน สามารถพิจารณาได้จากรูปที่ 1



รูปที่ 1 : ความสัมพันธ์ระหว่างเอนโทรปีและมิวชวลอินฟอร์เมชัน

2.2) การปรับค่าน้ำหนักคำแบบดั้งเดิม (Traditional Term Weighting)

ในงานการจำแนกประเภทเอกสารนั้นแต่ละคำของเวกเตอร์เอกสารจะเกี่ยวข้องกับค่า (น้ำหนัก) ซึ่งใช้วัดความสำคัญของคำและบ่งชี้ว่าคำนั้นมีศักยภาพในการจัดหมวดหมู่เอกสารมากน้อยแค่ไหน การกำหนดค่าน้ำหนักคำมาตรฐานที่นิยมใช้ได้แก่ ความถี่ของคำที่ปรากฏในเอกสาร (tf) แต่การใช้งาน tf อย่างเดียวอาจจะมีพลังอำนาจในการจำแนกไม่มากพอในการสืบค้นเอกสารที่ต้องการจากเอกสารที่ไม่ต้องการอื่น ๆ ดังนั้นความถี่เอกสารผกผัน (idf) จึงถูกนำมาใช้เพื่อเพิ่มพลังในการจำแนกของคำสำหรับการสืบค้นเอกสาร กำหนดให้ $D = \{d_1, d_2, d_3, \dots, d_N\}$ เป็นเซตของเอกสารทั้งหมด N เอกสาร (Corpus), $T = \{t_1, t_2, t_3, \dots, t_k\}$ เป็นเซตทั้งหมดของ k คำ ดังนั้นแต่ละเอกสารสามารถเขียนแสดงในรูปในรูปของปริภูมิเวกเตอร์ของคำ $d_j = (w_{j1}, w_{j2}, w_{j3}, \dots, w_{ji})$ เมื่อ j = 1,2,3,…., N และ i = 1,2,3, …, k โดยที่ w_{ij} แสดงถึงค่าน้ำหนักของคำ t_i ในเอกสาร d_j โดยความถี่ของคำ (tf) ในรูปแบบมาตรฐานสามารถคำนวณได้ดังนี้

$$tf_{ij} = \frac{\text{ความถี่ของคำที่ } i \text{ ในเอกสารที่ } d_j}{\text{จำนวนคำทั้งหมดในเอกสารที่ } d_j} \quad (3)$$

การใช้ tf อย่างเดียวอาจจะส่งผลให้พลังในการจำแนกไม่เพียงพอเนื่องจากคำที่มีความถี่สูง ๆ ก็จะทำให้มีค่าน้ำหนักของคำในเอกสารมากตามไปด้วย แต่คำที่มีอำนาจการจำแนกสูง ๆ ก็ควรที่จะปรากฏในเอกสารจำนวนไม่มาก ซึ่งในที่นี้ก็คือการคำนวณหาค่าความถี่เอกสารผกผัน (idf) ซึ่งสามารถคำนวณได้ดังนี้

$$idf_i = log\left(\frac{N}{Num_doc(t_i, D)}\right) \quad (4)$$

โดยที่ $Num_doc(t_i, D)$ คือ จำนวนเอกสารในคลังเอกสารที่ประกอบด้วยคำ t_i และในงานวิจัยนี้ในการคำนวณจะใช้ลอการิทึมฐาน 2 ในการคำนวณ ในขณะที่การคำนวณหาค่าน้ำหนักของคำที่ t_i ในเอกสาร d_j ด้วย if.idf สามารถคำนวณได้ดังนี้

$$w_{ij} = tf_{ij} \cdot idf_i \quad (5)$$

การปรับค่าน้ำหนักคำด้วย tf.idf เป็นหลักการที่มาจากการสืบค้นเอกสาร (Information retrieval) และในช่วงหลายสิบปีที่ผ่านมามีงานวิจัยจำนวนมากได้นำหลักการนี้มาประยุกต์ใช้ในการจำแนกประเภทเอกสาร บางงานวิจัยก็นำเทคนิคหลักการอื่น ๆ มาประยุกต์ใช้กับ tf.idf เพื่อปรับปรุงประสิทธิภาพในการจัดหมวดหมู่เอกสาร นอกจากนี้หลักการปรับค่าน้ำหนักคำด้วย tf.idf นั้นเป็นหลักการของการปรับค่าน้ำหนักคำแบบไม่มีผู้สอนโดยจะให้ความสำคัญอย่างมากกับความถี่ของคำที่ปรากฏในเอกสารและจำนวนของเอกสารที่ปรากฏคำนั้น อย่างไรก็ตามการนำหลักการของ tf.idf มาคำนวณหาค่าน้ำหนักคำในการวิเคราะห์ความพึงพอใจ อาจจะมีประสิทธิภาพไม่มากพอในการจำแนกประเภทความพึงพอใจว่าเป็นด้านบวก หรือ ลบ เนื่องจากบทวิจารณ์นั้นเป็นประโยคที่มีความยาวจำนวนคำไม่มากและมีแนวโน้มที่ในประโยคจะถูกเน้นด้วยคำหลักบางคำและความรู้สึกโดยรวมอาจไม่ถูกเน้นผ่านการใช้คำเดิม ๆ ซ้ำ ๆ

นอกจากนี้ยังมีวิธีการปรับค่าน้ำหนักคำแบบ rtf-sisf [18] เป็นอีกวิธีการหนึ่งที่มีการปรับเปลี่ยนพัฒนามาจาก tf.idf โดยการพิจารณาความถี่สัมพัทธ์ของคำ t_i ในเอกสารที่มีการรวบรวมและมีการบวก 1 เพิ่มเข้าไปในการคำนวณ idf เพื่อป้องกันกรณีที่ค่าอินพุตของลอการิทึมเป็น 1 ซึ่งการคำนวณน้ำหนักคำแบบ rtf-sisf สามารถคำนวณได้ดังนี้

$$w_{ij} = (tf_{ij}/MaxFreq_i) \cdot A \quad (6)$$

โดยที่ $A = \log(N/(1 + Num_doc(t_i, D)))$ และ $MaxFreq_i$ คือ จำนวนความถี่สูงสุดของคำ t_i

การปรับค่าน้ำหนักคำที่อธิบายข้างต้นนั้นเป็นรูปแบบหนึ่งของการปรับค่าน้ำหนักแบบไม่มีผู้สอนซึ่งเป็นหลักการที่สนใจหรือให้ความสำคัญกับความถี่ของคำที่เกิดขึ้นในเอกสาร โดยไม่ได้สนใจว่าคำนั้นมีชนิดของคลาสของเอกสารเป็นคลาสอะไร ซึ่งในหัวข้อถัดไปจะกล่าวถึงการปรับค่าน้ำหนักคำแบบมีผู้สอนที่

จะมีการพิจารณาชนิดของคลาสของเอกสารของคำใด ๆ เพื่อช่วยเพิ่มประสิทธิภาพการจำแนกเอกสารให้ดีขึ้น

2.3) การปรับค่าน้ำหนักคำแบบมีผู้สอน (Supervised Term Weighting)

การปรับค่าน้ำหนักคำแบบมีผู้สอนนั้นจะใช้ชนิดของคลาสของเอกสารที่ใช้สำหรับการสอนในการปรับค่าน้ำหนักคำ t_i ทั้งนี้ขึ้นอยู่ว่าแต่ละคำสามารถจำแนกเอกสารได้ถูกต้องมากน้อยแค่ไหนซึ่งจะส่งผลต่อค่าน้ำหนักคำที่จะปรับเพื่อให้คำนั้นมีพลังอำนาจในการจำแนกที่สูงในการทำนายหมวดหมู่ของเอกสาร ซึ่งในงานวิจัยนี้จะอธิบายหลักการที่นำมาเปรียบเทียบประสิทธิภาพกับวิธีการที่นำเสนอดังนี้

2.3.1) การแจกแจงเทอม (Term Distributions) ในส่วนนี้เป็นการอธิบายการปรับค่าน้ำหนักคำโดยใช้การแจกแจงเทอมหรือคำ มีหลายงานวิจัยที่นำหลักการนี้มาใช้ในการจำแนกประเภทเอกสาร [9], [10], [19], [20] เพื่อปรับปรุงประสิทธิภาพของการปรับค่าน้ำหนัก tf.idf การแจกแจงเทอมที่ใช้กันมีสามชนิด ได้แก่ ค่าเบี่ยงเบนมาตรฐานระหว่างคลาส (icsd), ค่าเบี่ยงเบนมาตรฐานของคลาส (csd) และค่าเบี่ยงเบนมาตรฐาน (sd) ถ้ากำหนดให้ tf_{ijq} เป็นค่าความถี่ของคำ t_i ในเอกสาร d_j ซึ่งอยู่ในคลาส c_q โดยสามารถนิยามค่า icsd ของคำ t_i, csd ของคำ t_i ในคลาส q และ sd ของคำ t_i ตามลำดับได้ดังนี้

$$icsd_i = \sqrt{\frac{\sum_q \left[\overline{tf_{iq}} - \frac{\sum_q \overline{tf_{iq}}}{|C|}\right]}{|C|}} \tag{7}$$

$$csd_{iq} = \sqrt{\frac{\sum_{d_j \in C_q} \left[tf_{ijq} - \overline{tf_{iq}}\right]^2}{|C_q|}} \tag{8}$$

$$sd_i = \sqrt{\frac{\sum_q \sum_{d_j \in C_q} \left[tf_{ijq} - \frac{\sum_q \sum_{d_j \in q} tf_{ijq}}{\sum_q |C_q|}\right]^2}{\sum_q |C_q|}} \tag{9}$$

โดยที่ $\overline{tf_{iq}} = \frac{\sum_{d_j \in C_q} tf_{ijq}}{|C_q|}$ เป็นค่าเฉลี่ยความถี่ของเทอม t_i ในเอกสารทั้งหมดที่อยู่ภายในคลาส c_q, $|C|$ คือจำนวนคลาส และ $|C_q|$ คือ จำนวนเอกสารในคลาส c_q

การคำนวณการปรับค่าน้ำหนักคำด้วยวิธี icsd ของคำใด ๆ จะไม่ขึ้นกับคลาส คำที่มีค่า icsd สูงจะทำให้เกิดพลังในการจำแนกความแตกต่างกันระหว่างคลาส และทำให้มีอำนาจในการแยกแยะการจำแนกประเภทเอกสารที่สูงกว่าคำอื่น ๆ ดังนั้นค่า icsd จะใช้ในการส่งเสริมคำที่มีอยู่ในเกือบทุกคลาสแต่มีความถี่ของคำสำหรับคลาสเหล่านั้นค่อนข้างแตกต่างกันมาก ในขณะที่คำซึ่งมีค่า csd สูงนั้นแสดงว่าคำนี้จะปรากฏอยู่ในเอกสารส่วนใหญ่ของคลาสด้วยความถี่ที่ต่างกันค่อนข้างมาก และไม่ใช่ตัวแทนที่ดีของคลาส ในขณะที่ค่า csd ของคำนั้นมีค่าน้อยแสดงว่าอาจเกิดขึ้นจากหนึ่งในสองสาเหตุคือ การเกิดขึ้นของคำนั้นเกือบจะเท่ากันสำหรับทุกเอกสารในคลาส หรืออาจจะเป็นคำที่ไม่ค่อยเกิดขึ้นในคลาสก็ได้ สำหรับคำที่มีค่า sd สูงแสดงว่าคำนั้นจะปรากฏในเอกสารส่วนใหญ่ที่มีการรวบรวมด้วยความถี่ที่ต่างกันอย่างมาก ในทางกลับกันคำที่มีค่า sd ต่ำอาจจะเกิดจากสาเหตุสองประการ คือ การเกิดขึ้นของคำนั้นเกือบจะเท่ากันสำหรับทุกเอกสาร หรือเป็นคำที่ไม่ค่อยเกิดขึ้นในเอกสารที่มีการรวบรวม อย่างไรก็ตามการปรับค่าน้ำหนักคำด้วยการแจกแจงเทอมทั้ง 3 หลักการนี้ icsd จะใช้สำหรับในการส่งเสริมคำ ในขณะที่ csd และ sd จะใช้สำหรับการลดการส่งเสริมคำ โดยสามารถคำนวณหาค่าน้ำหนัก w_{ij} ของคำ t_i ของเอกสาร d_j ได้ดังนี้

$$w_{ij} = tf_{ij} \cdot idf_i \cdot icsd_i^p \tag{10}$$

$$w_{ij} = (tf_{ij} \cdot idf_i)/csd_{iq}^p \tag{11}$$

$$w_{ij} = (tf_{ij} \cdot idf_i)/sd_i^p \tag{12}$$

ซึ่งในการทดลองงานวิจัยนี้จะใช้ค่า p เป็น 0.5 และ 1 โดยเป็นค่าที่ได้จากผลการทดลองของงานวิจัย [10] ในการเปรียบเทียบเทียบประสิทธิภาพการปรับค่าน้ำหนักคำกับวิธีการอื่น ๆ ในการจำแนกประเภทความพึงพอใจ

2.3.2) ความถี่ของคำ-ความถี่ของเอกสารที่ตรงกับความต้องการ (Term Frequency-Relevance Frequency: tf.rf)

การปรับค่าน้ำหนักคำแบบมีผู้สอนด้วยวิธีการ tf.rf [21] เป็นหลักการปรับค่าน้ำหนักคำที่จะพิจารณาสัดส่วนของจำนวนเอกสารที่มีคำ t_i ประกอบอยู่และอยู่ในคลาสบวก (positive class) ต่อจำนวนเอกสารที่มีคำ t_i ประกอบอยู่และอยู่ในคลาสลบ (negative class) โดยสามารถนิยามสูตรการปรับค่าน้ำหนักคำ t_i ของเอกสาร d_j ได้ดังสมการต่อไปนี้

$$w_{ij} = tf_{ij} \cdot \log\left(2 + \frac{D_{pos}(t_i)}{\max(1, D_{neg}(t_i))}\right) \tag{13}$$

โดยที่ $D_{pos}(t_i)$ คือ จำนวนเอกสารที่ประกอบด้วยคำ t_i และอยู่ในคลาสบวก

$D_{neg}(t_i)$ คือ จำนวนเอกสารที่ประกอบด้วยคำ t_i และอยู่ในคลาสลบ

ในงานวิจัยนี้จะนำเสนอวิธีการปรับค่าน้ำหนักคำทั้งแบบมีผู้สอนโดยใช้หลักการคลาสมิวชวลอินฟอร์เมชันที่แก้ข้อเสียของวิธีการ tf.idf ที่เป็นวิธีการปรับค่าน้ำหนักคำที่ให้ความสำคัญกับ

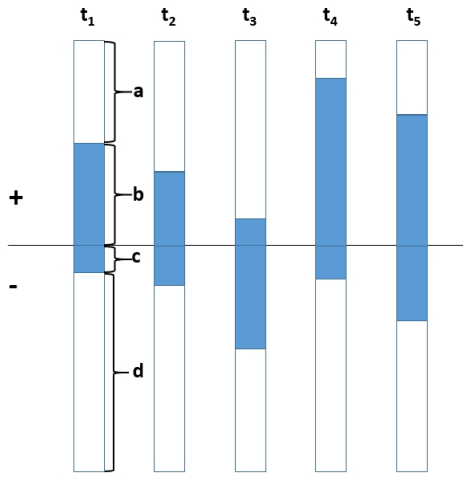
ความถี่ของคำที่เกิดขึ้นโดยไม่ได้มีการพิจารณาถึงความถี่ของคำนั้นมาจากคลาสใด ยิ่งเมื่อนำหลักการ tf.idf มาใช้ในการปรับค่าน้ำหนักคำสำหรับการจำแนกความพึงพอใจก็อาจจะทำให้ประสิทธิภาพความแม่นยำในการจัดหมวดหมู่บทวิจารณ์ไม่สูงมากนัก เนื่องจากบทวิจารณ์ออนไลน์ของผู้ใช้โดยทั่วไปแล้วจะเป็นโพสต์ข้อความสั้น ๆ ไม่ยาวมาก และความรู้สึกที่โพสต์ลงไปโดยรวมแล้วอาจไม่เน้นการใช้คำเดิม ๆ ซ้ำ ๆ เหมือนกับในเอกสารเช่น หนังสือ หรือ หนังสือพิมพ์ ซึ่งรายละเอียดแนวคิดของวิธีการที่นำเสนอและแนวทางในการแก้ข้อเสียของ tf.idf จะอธิบายในหัวข้อถัดไป

3) การปรับค่าน้ำหนักคำด้วยคลาสมิวชวลอินฟอร์เมชัน

3.1) ปัญหาการปรับค่าน้ำหนักคำด้วยวิธีการ idf ในการจำแนกประเภทความพึงพอใจ

โดยทั่วไปแล้วการวิเคราะห์ความพึงพอใจก็จะเกี่ยวข้องกับวิเคราะห์ทิศทางความรู้สึกบนหลักพิจารณาของข้อความ การพิจารณาดูข้อความนั้นว่าเป็นแบบมุมมองความคิดเห็นส่วนตัว หรือ มุมมองที่อยู่บนเนื้อหาข้อเท็จจริง และถ้าเป็นแบบมุมมองความคิดเห็นส่วนตัวจึงจะวิเคราะห์ต่อไปว่าข้อความนั้นเป็นความพึงพอใจที่เป็นเชิงบวก หรือ เชิงลบ ซึ่งในงานวิจัยโดยทั่วไปก็จะมองว่าปัญหาการจำแนกประเภทความพึงพอใจจัดว่าเป็นรูปแบบหนึ่งของปัญหาการจำแนกประเภทเอกสารที่เกี่ยวกับหัวข้อที่สนใจ ดังนั้นกระบวนการขั้นตอนต่าง ๆ ในการจำแนกประเภทความพึงพอใจก็จะเหมือนกับการจัดหมวดหมู่ของเอกสาร รวมไปถึงการแสดงถึงรูปแบบการแสดงของเอกสารจะอยู่ในรูปแบบเวกเตอร์น้ำหนักของคำซึ่งค่าน้ำหนักของคำที่นิยมใช้ในการศึกษาทดลองวิจัยส่วนใหญ่จะเป็น idf [22], [23] อย่างไรก็ตามรูปแบบการปรับค่าน้ำหนักด้วย idf อาจจะขาดประสิทธิภาพในบางสถานการณ์ซึ่งจะนำไปสู่ประสิทธิภาพความแม่นยำในการทำนายด้วยตัวจำแนกประเภทและตัวชี้วัดอื่น ๆ ไม่สูงมากนัก [22], [23]

ความถี่ของคำ tf จะแสดงถึงความสัมพันธ์ที่ใกล้ชิดกันระหว่างคำและเนื้อหาในเอกสารที่ประกอบด้วยคำนั้น เป็นที่สังเกตได้ว่าถ้าความถี่ของคำมีค่ามากและแพร่กระจายไปยังจำนวนเอกสารเป็นจำนวนมากก็จะทำให้ไม่สามารถสืบค้นเอกสารให้ตรงกับที่ต้องการจากเอกสารทั้งหมดได้ เพื่อแก้ปัญหานี้จึงมีการนำ idf มาช่วยปรับปรุงเพื่อเพิ่มพลังในการจำแนกของคำในการสืบค้นเอกสาร แต่อย่างไรก็ตามการนำหลักการนี้มาใช้ในการจำแนกประเภทของความพึงพอใจก็อาจจะพบเจอปัญหาบางประการในการวิเคราะห์ความพึงพอใจดังในรูปที่ 2 โดยการกระจายของเอกสารจะประกอบด้วย 5 คำ คือ t_1, t_2, t_3, t_4 และ t_5 แต่ละคอลัมน์จะพิจารณาว่าเป็นการกระจายของเอกสารในคลังเอกสารสำหรับแต่ละคำ ความสูงของแต่ละแท่งก็จะแทนจำนวนเอกสารที่อยู่ในคลังเอกสารและเส้นแนวนอนจะใช้แบ่งเอกสารออกเป็น 2 กลุ่มคือ คลาสบวก (Positive) และ คลาสลบ (Negative) ความสูงแต่ละคอลัมน์ที่อยู่เหนือเส้นแนวนอนและใต้เส้นแนวนอนก็จะแทนจำนวนเอกสารที่อยู่ในคลาสบวกและคลาสลบ ตามลำดับ ความสูงของพื้นที่แรงเงาของแต่ละคอลัมน์ก็จะแทนจำนวนเอกสารที่อยู่ในคลาสบวกและลบ เพื่อให้การอธิบายเข้าใจได้ง่ายก็จะใช้ a, b, c และ d แทนจำนวนเอกสารที่แตกต่างกันดังรายการต่อไปนี้



รูปที่ 2 : ตัวอย่างการกระจายที่แตกต่างกันของเอกสารที่ประกอบด้วย 6 เทอมในเอกสารที่รวบรวมทั้งหมด

โดย a แทนจำนวนเอกสารที่อยู่ในคลาสบวกที่ไม่ได้ประกอบด้วยคำนี้

b แทนจำนวนเอกสารที่อยู่ในคลาสบวกที่ประกอบด้วยคำนี้

c แทนจำนวนเอกสารที่อยู่ในคลาสลบที่ประกอบด้วยคำนี้

d แทนจำนวนเอกสารที่อยู่ในคลาสลบที่ไม่ได้ประกอบด้วยคำนี้

ดังนั้น N แทนจำนวนเอกสารทั้งหมดที่รวบรวมโดยเป็นผลรวมของ a, b, c และ d

ในที่นี้สมมุติว่าทั้ง 5 คำนี้มีค่าความถี่ของคำ (tf) เท่ากัน โดยที่สามคำแรก t_1, t_2, t_3 มีค่า idf เท่ากันซึ่งจะแทนด้วยค่า idf1 ส่วนสองคำสุดท้าย t_4 และ t_5 ก็จะมีค่า idf เท่ากันซึ่งแทนด้วยค่า

idf2 ซึ่งจากรูปจะเห็นได้ว่าค่า idf1 > idf2 ดังนั้นแสดงว่า 3 คำแรกก็จะมีค่าน้ำหนักของคำมากกว่าเนื่องจากมีความถี่ของเอกสารในคลังเอกสารน้อยกว่า 2 คำสุดท้าย จะเห็นได้ชัดว่าทั้ง 5 คำนั้นจะสนับสนุนการจำแนกประเภทความพึงพอใจได้แตกต่างกันซึ่งจะให้ความสำคัญอย่างมากกับเทอมหรือคำที่ทำให้มีจำนวนคลาสบวกมากกว่าคลาสลบซึ่งจะใช้ในการคัดแยกตัวอย่างที่เป็นคลาสบวกออกจากตัวอย่างที่เป็นคลาสลบ ถ้าพิจารณาจากรูปที่ 2 จะพบว่า t_1 จะสนับสนุนมากกว่า t_2 และ t_3 ในการคัดแยกตัวอย่างที่เป็นคลาสบวกออกจากตัวอย่างที่เป็นคลาสลบในทำนองเดียวกัน t_4 จะสนับสนุนการคัดแยกได้มากกว่า t_5 เช่นกัน

อย่างไรก็ตามถ้ามีการกำหนดค่าน้ำหนักคำด้วย idf จะเห็นได้ว่าไม่สามารถเห็นความแตกต่างระหว่าง 3 คำแรกและเช่นเดียวกันก็ไม่สามารถแยกความแตกต่างระหว่าง 2 คำสุดท้ายได้เช่นกัน ถ้าพิจารณาจากรูปที่ 2 จะเห็นได้ว่า t_1 จะสนับสนุนพลังในการจำแนกข้อความบทวิจารณ์ที่เป็นคลาสบวกออกจากข้อความบทวิจารณ์ที่เป็นคลาสลบได้ดีกว่า t_2 และ t_3 ดังนั้น t_1

ควรจะมีค่าน้ำหนักมากกว่า t_2 และ t_3 อย่างไรก็ตามการกำหนดค่าน้ำหนักด้วย idf ทั้ง t_1, t_2 และ t_3 จะให้พลังในการจำแนกประเภทความพึงพอใจได้เท่า ๆ กัน จึงเห็นได้ว่ากรณีนี้การกำหนดค่าน้ำหนักด้วย idf จะล้มเหลวในการกำหนดค่าที่เหมาะสมให้กับคำเมื่อพิจารณาจากการสนับสนุนที่แตกต่างกันในงานการจำแนกประเภทข้อความบทวิจารณ์

ถ้าพิจารณา t_1 และ t_4 จะเห็นได้ว่าการปรับค่าน้ำหนักด้วย idf จะทำให้ idf1(t_1) > idf2(t_4) ซึ่งก็คือ t_1 มีพลังในการจำแนกได้ดีกว่า t_4 แต่จากรูปที่ 2 ถ้าพิจารณาแล้วจะเห็นได้ชัดเจนว่า t_4 ควรมีพลังในการจำแนกได้ดีกว่า t_1 ดังนั้น t_4 ควรมีค่าน้ำหนักมากกว่า t_1 ซึ่งในทำนองเดียวกันค่า idf ของ t_4 และ t_5 เท่ากันแสดงว่าทั้ง 2 คำจะมีพลังในการจำแนกข้อความบทวิจารณ์คลาสบวกออกจากคลาสลบได้เท่า ๆ แต่ในความเป็นจริงแล้ว t_4 ควรจะมีพลังในการจำแนกคลาสบวกออกจากคลาสลบได้มากกว่า t_5 ดังนั้น t_4 ควรจะต้องมีค่าน้ำหนักมากกว่า t_5 เช่นกัน

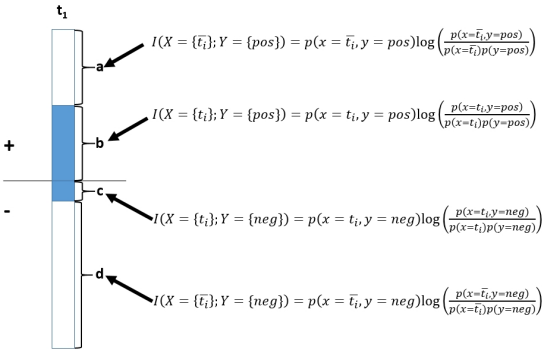
จากที่กล่าวมาข้างต้นเป็นความล้มเหลวของการกำหนดค่าน้ำหนักคำแบบ idf ที่ไม่สามารถแก้ปัญหาที่กล่าวมาข้างต้นได้จึงทำให้การกำหนดค่าน้ำหนักคำด้วยวิธีการ tf.idf ไม่ทำให้ประสิทธิภาพในงานการจำแนกประเภทความพึงพอใจสูงขึ้นได้ ดังนั้นในงานวิจัยนี้จึงนำเสนอวิธีการกำหนดค่าน้ำหนักคำด้วยคลาสมิวชวลอินฟอร์เมชันที่จะแก้จุดด้อยจากปัญหาข้างต้นของวิธีการ idf และช่วยเพิ่มประสิทธิภาพการจำแนกประเภทความพึงพอใจ รวมถึงตัวชี้วัดต่าง ๆ ที่สำคัญ ให้สูงขึ้นกว่าวิธีการ idf ซึ่งจะได้อธิบายรายละเอียดให้หัวข้อถัดไป

3.2) แนวคิดของวิธีการที่นำเสนอ

จากปัญหาข้างต้นที่เกิดขึ้นกับวิธีการกำหนดค่าน้ำหนักคำด้วย idf ซึ่งเป็นการปรับค่าน้ำหนักคำแบบไม่มีผู้สอน เนื่องจากวิธีนี้ให้ความสำคัญกับค่าความถี่ของคำที่เกิดขึ้นเท่านั้น โดยไม่ได้พิจารณาว่าข้อความบทวิจารณ์มาจากคลาสบวกหรือคลาสลบ ในงานวิจัยนี้จะแก้ปัญหาของวิธีการ idf ด้วยการปรับค่าน้ำหนักคำ tf.idf ด้วยคลาสมิวชวลอินฟอร์เมชัน (Class Mutual Information: cmi) โดยวิธีการ cmi จัดว่าเป็นการปรับค่าน้ำหนักคำแบบมีผู้สอน ซึ่งน้ำหนักคำที่ปรับจะขึ้นอยู่กับพลังอำนาจในการสนับสนุนการจำแนกคลาสบวกออกจากคลาสลบของคำนั้น

ในการคำนวณ cmi นั้นมีหลักการตรงไปตรงมา ซึ่งจะให้ความสำคัญกับการคำนวณหาข่าวสารร่วมของ $\boldsymbol{t_i}$ และคลาสบวก (คำนวณบริเวณส่วนของ b ในรูปที่ 2) ซึ่งจะเห็นได้ว่าถ้าส่วนที่แรงเงาส่วน b มีค่ามากก็จะทำให้มีพลังอำนาจมากในการสนับสนุนการจำแนกเอกสารข้อความคลาสบวกออกจากคลาสลบ ซึ่งถ้าพิจารณาในส่วนของ 3 คำแรก คือ t_1, t_2 และ t_3 จะเห็นได้ว่าค่า idf เท่ากันทุกคำซึ่งไม่สะท้อนความเป็นจริงและล้มเหลวในการกำหนดค่าที่เหมาะสมให้กับคำ แต่เมื่อพิจารณาจากส่วนแรงเงา b ซึ่งเป็นข่าวสารร่วมระหว่างคำ $\boldsymbol{t_i}$ ที่ปรากฏอยู่ในเอกสารข้อความและเอกสารข้อความที่อยู่ในคลาสบวกจะพบว่าพลังอำนาจในการสนับสนุนการของคำ t_1 มากกว่า t_2 และ t_2 มากกว่า t_3 นั่นคือ cmi(t_1) > cmi(t_2) > cmi(t_3) ซึ่งสามารถดูตัวอย่างวิธีการคำนวณส่วนต่างของคำได้จากรูปที่ 3

ในทำนองเดียวกันเมื่อพิจารณา t_1 และ t_4 จะเห็นได้ว่าการปรับค่าน้ำหนักด้วย cmi จะทำให้ cmi(t_4) > cmi(t_1) ซึ่งก็คือ t_4 จะมีพลังในการจำแนกเอกสารข้อความได้ดีกว่า t_1 ดังนั้น t_4 ก็จะมีค่าน้ำหนักมากกว่า t_1 นอกจากนี้ถ้าพิจารณาจากค่า cmi ของ t_4 และ t_5 พบว่าการกำหนดค่าน้ำหนักของคำ t_4 มากกว่า t_5 ซึ่งสะท้อนถึงความเป็นจริงที่จำนวนเอกสารข้อความในคลาสบวกของเทอม t_4 มากกว่า รวมไปถึงการสนับสนุนในการคัดแยกตัวอย่างของคลาสบวกออกจากคลาสลบของเทอม t_4 และ t_5 จะมากกว่า t_1, t_2 และ t_3 ซึ่งการกำหนดค่าน้ำหนักคำของ t_4 และ t_5 ควรจะต้องมากกว่า t_1, t_2 และ t_3 ด้วยเช่นกัน โดยการคำนวณคลาสมิวชวลอินฟอร์เมชันของแต่ละส่วนสามารถพิจารณาได้จากรูปที่ 3



รูปที่ 3 : ตัวอย่างการคำนวณคลาสมิวชวลอินฟอร์เมชัน สำหรับพื้นที่แต่ละส่วนของคำ

จากรูปที่ 3 สามารถเขียนสมการในการคำนวณค่าน้ำหนักคำ t_i ด้วยวิธีการ idf ใหม่ได้ดังนี้

$$idf_i = log\left(\frac{a+b+c+d}{b+c}\right) \quad (14)$$

กำหนดให้ t_i แทนคำหรือเทอมที่ปรากฏอยู่ในเอกสาร, $\overline{t_i}$ แทนคำหรือเทอมที่ไม่ปรากฏอยู่ในเอกสาร และกำหนดให้ pos แทนเอกสารที่อยู่ในคลาสบวก และ neg แทนเอกสารที่อยู่ในคลาสลบ ดังนั้นจากสมการ (2) การคำนวณหาค่าข่าวสารร่วมของคลาสและเอกสารที่ประกอบด้วยคำ t_i สามารถเขียนสมการได้เป็น

$$I(X;Y) = \sum_{x\epsilon\Psi}\sum_{y\epsilon\Omega} p(x,y)\log\frac{p(x,y)}{p(x)\cdot p(y)} \quad (15)$$

โดยที่ $\Psi = \{t_i, \overline{t_i}\}, \Omega = \{pos, neg\}$

เนื่องจากแนวคิดในงานวิจัยนี้จะพิจารณากำหนดค่าน้ำหนักคำโดยให้ความสนใจเฉพาะขนาดพื้นที่ส่วนแรงเงาของ b ที่จะสามารถสนับสนุนอำนาจในการจำแนกเอกสารข้อความที่เป็นคลาสบวกออกจากคลาสลบ ได้ดีกว่าส่วนอื่น ๆ ซึ่งถ้าปริมาณเอกสารในส่วนของ b มีค่าปริมาณข่าวสารมาก ๆ ย่อมส่งผลดีต่อการกำหนดค่าน้ำหนักของคำในการเพิ่มอำนาจในการจำแนกเอกสารข้อความบทวิจารณ์คลาสบวกออกจากคลาสลบ ดังนั้นการคำนวณหาข่าวสารของจำนวนเอกสารข้อความบทวิจารณ์ที่ประกอบด้วยคำ t_i และเอกสารเหล่านั้นอยู่ในคลาสบวก (pos) สามารถคำนวณได้จากสมการต่อไปนี้

$$I(X=\{t_i\};Y=\{pos\}) = Mlog(S) \quad (16)$$

โดยที่ $M = p(x=t_i, y=pos)$, $S = \frac{p(x=t_i,y=pos)}{p(x=t_i)p(y=pos)}$

จากการพิจารณารูปที่ 2 และ 3 และ จากสมการ (16) ข้างต้นแทนจะมีการแทนค่า $p(t_i, pos)$ ด้วยค่า b/N แทนค่า $p(t_i)$ และ $p(pos)$ ด้วยค่า (b+c)/N และ (a+b)/N ตามลำดับ ดังนั้นการแทนค่าจะได้เป็น

$$I(X=\{t_i\};Y=\{pos\}) = \left(\frac{b}{N}\right) log\left(\frac{\frac{b}{N}}{\left(\frac{b+c}{N}\right)\left(\frac{a+b}{N}\right)}\right) \quad (17)$$

เพื่อความสะดวกในการนำสมการที่ (17) ไปประยุกต์ใช้และเปรียบเทียบกับวิธีการอื่น ๆ ในการคำนวณหาข่าวสารร่วมระหว่างเอกสารที่ประกอบด้วยคำ t_i และเอกสารที่อยู่ในคลาส C_q จะมีการสร้างตัวแปร cmi_{iq} ที่แสดงความสัมพันธ์กับคลาสมิวชวลอินฟอร์เมชัน สามารถเขียนแสดงความสัมพันธ์ได้ดังนี้ คือ

$$cmi_{iq} = I\left(X=\{t_i\};Y=C_q\right) \quad (18)$$

โดยที่ $C_q = \{pos\}$

สำหรับวิธีการปรับค่าน้ำหนักคำที่นำเสนอนี้สามารถคำนวณหาค่าน้ำหนัก w_{ij} ของคำ t_i และเอกสาร d_j ได้ดังนี้

$$w_{ij} = tf_{ij}\cdot idf_i\cdot cmi_{iq} \quad (19)$$

ซึ่งการปรับค่าน้ำหนักคำด้วยวิธีการที่นำเสนอนี้จะนำค่าน้ำหนักคำ tf.idf ไปคูณกับ cmi_{iq} ตามสมการที่ (19) เพื่อทำการปรับค่าน้ำหนักคำให้เหมาะสมกับศักยภาพของคำต่าง ๆ ที่สนับสนุนพลังอำนาจในการจำแนกเอกสารข้อความคลาสบวกออกจากคลาสลบ ซึ่งในหัวข้อถัดไปจะได้อธิบายถึงขั้นตอนของการนำวิธีการปรับค่าน้ำหนักคำที่นำเสนอไปประยุกต์ใช้ในการจำแนกประเภทความพึงพอใจข้อความบทวิจารณ์ที่มีกระบวนการขั้นตอนต่าง ๆ ที่จะได้อธิบายต่อไป

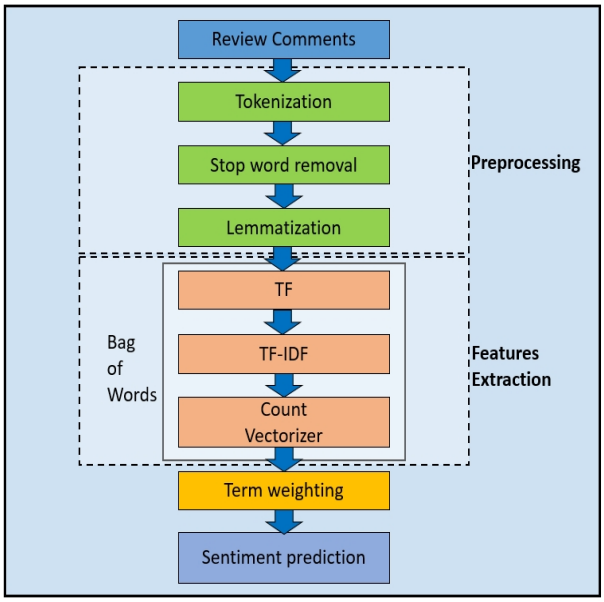
3.3) การปรับค่าน้ำหนักคำด้วยคลาสมิวชวลอินฟอร์เมชัน

การจำแนกประเภทความพึงพอใจของข้อความบทวิจารณ์นั้นเป็นกระบวนการหนึ่งในการพิจารณาบทวิจารณ์ เช่น บทวิจารณ์ผลิตภัณฑ์/บริการ แบบสำรวจออนไลน์ เป็นต้น ว่ามีทัศนคติเป็นเชิงบวก หรือ เชิงลบ ซึ่งจะมีผลดีต่อผู้ผลิตสินค้าหรือ ผู้ให้บริการที่จะเข้าใจความชื่นชอบของผู้บริโภคที่มีต่อสินค้าและบริการและสามารถรวบรวมความคิดเห็นของผู้บริโภคมาปรับปรุงการให้บริการ [24] ต่าง ๆ ได้ ขั้นตอนการประมวลผลการจำแนกประเภทความพึงพอใจของความคิดเห็นแสดงดังรูปที่ 4 มีขั้นตอนดังนี้

3.3.1) การเตรียมข้อมูลก่อนการประมวลผล (Preprocessing) ทำการแตกประโยคออกเป็นคำย่อย จากนั้นจะเป็นการลบคำหยุด คำเว้นวรรค จุด รวมไปถึง ‘a’, ‘an’, ‘the’ เป็นต้น แล้วจากนั้นจึงไปหารากศัพท์ของคำ และในขั้นตอนสุดท้ายจะเป็นกระบวนการแปลงคำด้วยรายการคำศัพท์ในภาษาเพื่อลดคำที่ผันอยู่ในรูปแบบต่าง ๆ

3.3.2) การสกัดคุณลักษณะของคำ (Feature Extraction) ขั้นตอนนี้จะทำการแปลงคำให้อยู่ในรูปตัวเลขโดยมองแต่ละแถวเป็นเอกสารและคอลัมน์เป็นคำ ขั้นตอนนี้เริ่มต้นด้วยการคำนวณหาค่าน้ำหนักคำด้วยความถี่ของคำ (tf) และคำนวณหาความถี่เอกสารผกผันและสร้างโทเค็นของการรวบรวมเอกสารข้อความและสร้างคำศัพท์สำหรับคำที่สร้างขึ้นให้สามารถที่จะเข้ารหัสเอกสารใหม่ได้โดยใช้คำศัพท์นั้น

3.3.3) การจำแนกประเภทความพึงพอใจ (Sentiment Classification) เป็นการจำแนกประเภทเอกสารข้อความบทวิจารณ์ของผลิตภัณฑ์และบริการทำได้โดยใช้แนวทางการเรียนรู้ของเครื่อง เช่น เบย์อย่างง่าย (Naïve Bayes) และ ป่าของต้นไม้ตัดสินใจแบบสุ่ม (Random Forest) เป็นต้น ในการสร้างโมเดลการจำแนกจากชุดข้อมูลที่ใช้สอน เพื่อนำมาใช้ในการจำแนกข้อมูลทดสอบเพื่อการทำนาย



รูปที่ 4 : ขั้นตอนการจำแนกประเภทความพึงพอใจเอกสารข้อความบทวิจารณ์

ขั้นตอนการจำแนกประเภทเอกสารข้อความบทวิจารณ์จะเริ่มต้นด้วยการอ่านข้อความบทวิจารณ์เข้ามาเพื่อทำขั้นตอนการเตรียมข้อมูลก่อนการประมวลผล แล้วเข้าสู่กระบวนการการสกัดคุณลักษณะของคำ หลังจากนั้นก็จะคำนวณหาค่า tf.idf สำหรับนำไปใช้คำนวณร่วมกับวิธีการที่นำเสนอ โดยวิธีการ cmi_{iq} จะมีการคำนวณโดยพิจารณาจากค่าของความถี่ของคำ tf_{ij} ที่แปลงให้อยู่ในรูปแบบมาตรฐานที่เรียกว่า ถุงของคำ (Bag of words) โดยแถวแทนเอกสารทั้งหมดในคลังและคอลัมน์แทนค่าน้ำหนักของคำดังรูปที่ 5 โดยแต่ละแถวนั้นก็จะแทนด้วยเวกเตอร์ของเอกสาร ที่แสดงในรูปแบบ $d_j = (tf_{1j}, tf_{2j}, tf_{3j}, \dots, tf_{ij})$ ซึ่งก่อนที่จะมีการคำนวณหาค่าคลาสมิวชวลอินฟอร์เมชันของเทอม t_i ของแต่ละคลาส เพื่อความสะดวกในการคำนวณหาค่าความน่าจะเป็นในการคำนวณนั้นจะต้องมีการแปลงตารางที่อยู่ในรูปที่ 5 ไปเป็นตารางที่แสดงค่าความถี่ของคำในรูปแบบไบนารี่ก่อน ดังรูปที่ 6

จากนั้นก็จะทำการคำนวณหาค่าน้ำหนักของคำด้วยคลาสมิวชวลอินฟอร์เมชัน หรือ คำนวณหาข่าวสารร่วมระหว่างเทอม t_i ที่สัมพันธ์ร่วมกับคลาสที่ q สามารถคำนวณด้วยสมการที่ (17) เมื่อคำนวณ cmi_{iq} ของแต่ละคลาสแล้ว ก็จะมีการคำนวณค่าน้ำหนักเทอม t_i สำหรับแต่ละข้อความบทวิจารณ์ที่อยู่ในคลาส q ด้วยการใช้สมการที่ (19)

Bag of Words

	tf_1	tf_2	tf_3	. . .	tf_i	class
d_1	0	0.4	0.33		0	pos
d_2	0.8	0.5	0		0.3	pos
d_3	0.4	0.1	0.1		0.2	neg
	⋮	⋮	⋮	⋮	⋮	⋮
d_N	0.7	0	0		0.6	pos

รูปที่ 5 : ถุงของคำที่แสดงค่าความถี่ของคำในรูปแบบมาตรฐาน

Binary Bag of Words

	tf_1	tf_2	tf_3	. . .	tf_i	class
d_1	0	1	1		0	pos
d_2	1	1	0		1	pos
d_3	1	1	1		1	neg
	⋮	⋮	⋮	⋮	⋮	⋮
d_N	1	0	0		1	pos

รูปที่ 6 : ถุงของคำในรูปแบบทวิภาค

ในงานวิจัยนี้ได้นำเสนอการเพิ่มประสิทธิภาพการปรับค่าน้ำหนักของคำ tf.idf ในการแก้ข้อด้อยที่เกิดขึ้นของวิธีการ idf เนื่องจากวิธีการนี้พิจารณาความถี่ของคำเป็นสำคัญแต่เพิกเฉยต่อศักยภาพอำนาจในการจำแนกตัวอย่างคลาสบวกออกจากคลาสลบของเทอมนั้น ดังนั้นในงานวิจัยนี้จะนำเสนอวิธีการที่จะช่วยเพิ่มประสิทธิภาพในการจำแนกประเภทความพึงพอใจข้อความบท

วิจารณ์ด้วยการคำนวณ tf.idf ร่วมกับคลาสมิวชวลอินฟอร์เมชัน นอกจากนี้ในหัวข้อถัดไปจะได้มีการเปรียบเทียบประสิทธิภาพวิธีการที่นำเสนอกับวิธีการ tf.idf และยังทำการเปรียบเทียบประสิทธิภาพกับวิธีการปรับค่าน้ำหนักเทอมด้วยวิธีการการแจกแจงเทอม และวิธีการอื่น ๆ ซึ่งผลการทดลองแสดงให้เห็นถึงประสิทธิภาพของวิธีการที่นำเสนอเมื่อพิจารณาทั้ง ค่าความถูกต้อง, ค่าความแม่นยำ, ค่าความระลึก และ ค่าความถ่วงดุล

4) ผลการทดลอง

ในส่วนนี้จะแสดงผลการทดลองระหว่างวิธีที่นำเสนอและเทคนิคอื่นที่ได้รับความนิยมในการปรับค่าน้ำหนักคำร่วมกับข้อมูลบทวิจารณ์ที่นิยมนำมาใช้ในการทดลองและขนาดของข้อมูลก็ไม่ใหญ่มากที่จะเกินข้อจำกัดของเครื่องคอมพิวเตอร์ที่ใช้ในการทดลอง โดยในส่วนนี้จะมีการเปรียบเทียบประสิทธิภาพวิธีการปรับค่าน้ำหนักคำกับตัวชี้วัดต่าง ๆ ที่นิยมใช้กันในงานด้านการหมวดหมู่ประเภทเอกสาร ซึ่งตัวชี้วัดที่ใช้ประกอบด้วย ค่าความถูกต้องการทำนาย (Accuracy), ค่าความแม่นยำ (Precision), ค่าความระลึก (Recall) และค่าความถ่วงดุล (F1-Measure) โดยในการทดลองงานครั้งนี้จะแบ่งชุดข้อมูลใช้สอนและชุดข้อมูลใช้ทดสอบออกเป็น 5 ชุด (5 Fold Cross Validation) และทำการทดสอบบนตัวจำแนกประเภทข้อมูลที่นิยมนำมาใช้กันในงานวิจัย ซึ่ง ได้แก่ เบย์อย่างง่าย (Naïve Bayes), ป่าของต้นไม้ตัดสินใจแบบสุ่ม (Random forest) และ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ (K-NN) เพื่อใช้เป็นตัววัดประสิทธิภาพของตัวชี้วัดต่าง ๆ ของเทคนิคการกำหนดค่าน้ำหนักคำที่นำมาเปรียบเทียบประสิทธิภาพ โดยวิธีการจำแนกประเภทด้วยเพื่อนบ้านที่ใกล้ที่สุด K อันดับ ในการทดลองนี้จะใช้ K = 3

4.1) ข้อมูลบทวิจารณ์สินค้าที่ใช้ในการทดลอง

เพื่อแสดงให้เห็นถึงประสิทธิผลของวิธีการที่นำเสนอสำหรับการจำแนกประเภทความพึงพอใจจะขออธิบายรายละเอียดถึงชุดข้อมูลที่ใช้ในการวิเคราะห์ความพึงพอใจจากหลาย ๆ เว็บไซต์ที่นิยมใช้กันในการศึกษางานวิจัย ข้อมูลแรกคือ Amazon เป็นเอกสารข้อความที่สกัดมาจากบทวิจารณ์สินค้าต่าง ๆ ของลูกค้าจากเว็บไซต์ amazon.com ประกอบด้วยข้อความ 1000 บทวิจารณ์ ซึ่งประกอบด้วยข้อความที่เป็นความรู้สึกด้านบวกต่อสินค้า 500 บทความคิดเห็นและความรู้สึกที่เป็นด้านลบต่อสินค้า 500 บทความคิดเห็นเช่นกัน ในขณะที่ข้อมูล Yelp เป็นเอกสารข้อความที่สกัดมาจากบทความคิดเห็นร้านอาหารของลูกค้าจากเว็บไซต์ yelp.com ประกอบด้วยข้อความบทความคิดเห็น 1000 บทความคิดเห็น ซึ่งประกอบด้วยความรู้สึกที่เป็นด้านบวกต่อร้านอาหาร 500 บทความคิดเห็นและความรู้สึกที่เป็นด้านลบ 500 บทความคิดเห็น ส่วนชุดข้อมูลถัดไป คือ IMDb ก็จะสกัดมาจากบทความคิดเห็นภาพยนต์ของลูกค้าจากเว็บไซต์ imdb.com ที่ประกอบด้วยข้อความบทความคิดเห็น 1000 บทความคิดเห็น โดยแบ่งเป็นความรู้สึกที่เป็นด้านบวก 500 บทความคิดเห็นและความรู้สึกที่เป็นด้านลบ 500 บทความคิดเห็นเช่นกัน ทั้งข้อมูล Amazon, Yelp และ IMDb เป็นข้อมูลที่แต่ละเอกสารประกอบด้วยข้อความบทความคิดเห็นและตามด้วยชนิดของคลาสของความรู้สึก โดยทั้ง 3 ข้อมูลนี้เป็นข้อมูลที่อ้างอิงมาจากงานวิจัย [25] ซึ่งนักวิจัยสามารถที่จะดาวน์โหลดข้อมูลเหล่านี้ได้ได้จากเว็บไซต์ https://archive.ics.uci.edu/ml/m-achine-learning-databases/00331/

ตารางที่ 1 : สรุปรายละเอียดชุดข้อมูลที่ใช้ในการทดลอง

ลำดับ	ชุดข้อมูล	คำอธิบาย
1	Amazon	บทวิจารณ์สำหรับผลิตภัณฑ์ที่ขายบนเว็บไซต์ amazon.com ในหมวดโทรศัพท์มือถือและอุปกรณ์เสริม
2	Yelp	บทวิจารณ์จากชุดข้อมูล Yelp challenge2 ซึ่งได้รวบรวมมาจากบทความคิดเห็นร้านอาหารบนเว็บไซต์ yelp.com
3	IMDb	ข้อมูลความคิดเห็นบทวิจารณ์ภาพยนตร์ซึ่งเป็นรายการบทวิจารณ์ภาพยนตร์ที่โพสต์บนเว็บไซต์ imdb.com
4	Eco-Pre	ข้อมูลบทความคิดเห็นผลิตภัณฑ์ EcoDot สำหรับ Gen3 ของ amazon จากเว็บไซต์ amazon.in
5	Magazine	ข้อมูลบทวิจารณ์ความคิดเห็นนิตยสารของสมาชิกโดยได้รวบรวมมาจากเว็บไซต์ amazon.com

ในขณะที่ข้อมูลบทความคิดเห็นสินค้า Eco-Pre เป็นข้อมูลความคิดเห็นผลิตภัณฑ์ EcoDot ซึ่งเป็นรุ่น Gen3 ของ amazon ที่มีการรวบรวมความคิดเห็นบทวิจารณ์จากผู้ใช้งานมาจากเว็บไซต์ amazon.in โดยข้อมูลบทความคิดเห็นนี้จะมีทั้งสิ้น 4,084 เอกสารข้อความและแต่ละข้อความก็จะมีการระบุถึงความรู้สึกที่มีต่อสินค้าซึ่งเป็นความรู้สึกด้านบวกจำนวน 3,066

เอกสารข้อความ ความรู้สึกด้านลบจำนวน 482 เอกสารข้อความ และความรู้สึกเป็นกลางจำนวน 536 เอกสารข้อความ โดยข้อมูล Eco-Pre สามารถที่จะทำการดาวน์โหลดได้ข้อมูลนี้จากเว็บไซต์ https://www.kaggle.com/datasets/pradeeshprabhakar/preprocessed-dataset-sentiment-analysis และสำหรับข้อมูลบทความคิดเห็นนิตยสาร Magazine เป็นข้อมูลบทวิจารณ์นิตยสารของสมาชิกโดยได้รวบรวมมาจากเว็บไซต์ amazon.com โดยข้อมูลประกอบด้วยข้อความบทวิจารณ์ 1,571 ข้อความซึ่งเป็นความรู้สึกด้านบวกจำนวน 1,454 ข้อความและความรู้สึกด้านลบจำนวน 117 ข้อความ โดยข้อมูลบทวิจารณ์ความคิดเห็นนิตยสารนี้สามารถดาวน์โหลดได้จากเว็บไซต์ https://nijianmo.github.io/amazon/index.html โดยข้อมูลทั้งหมดที่ใช้ในการทดลองในงานวิจัยครั้งนี้สามารถพิจารณาดูรายละเอียดได้จากตารางที่ 1 และ ตารางที่ 2

ตารางที่ 2 : คุณลักษณะของข้อมูลที่ใช้ในการจำแนกประเภทความพึงพอใจ

ลำดับที่	ชุดข้อมูล	#คอลัมน์	#แถว	คลาส
1	Amazon	422	1,000	บวก/ลบ
2	Yelp	464	1,000	บวก/ลบ
3	IMDb	748	1,000	บวก/ลบ
4	Eco-Pre	1,330	4,084	บวก/ลบ/กลาง
5	Magazine	1,213	1,571	บวก/ลบ

ในงานวิจัยนี้จะแสดงให้เห็นถึงประสิทธิภาพที่ดีกว่าของวิธีการที่นำเสนอ cmi ซึ่งสามารถแก้ปัญหาข้อจำกัดต่าง ๆ เมื่อเปรียบเทียบกับวิธีการ tf.idf โดยในการเปรียบเทียบจะพิจารณาค่าเฉลี่ยของตัวชี้วัดทั้งหมด 5 ครั้ง ทั้งในแง่ของค่าความถูกต้องในการทำนายเอกสารข้อความ (Accuracy: Acc) ค่าความแม่นยำ (Precision: Pre) เพื่อแสดงถึงสัดส่วนของการทำนายคลาสนั้นที่ถูกต้องจากการทำนายคลาสนั้นทั้งหมด ค่าความระลึก (Recall: Rec) ที่แสดงถึงสัดส่วนของการทำนายคลาสนั้นที่ถูกต้องจากคลาสจริงนั้นทั้งหมด และค่าความถ่วงดุล (F1) นอกจากนี้วิธีการที่นำเสนอยังได้เปรียบเทียบประสิทธิภาพกับเทคนิคการปรับค่าน้ำหนักคำด้วยวิธีการอื่น ๆ ได้แก่ การแจกแจงเทอม (icsd, csd และ sd), วิธีการปรับค่าน้ำหนักด้วยมิวชวลอินฟอร์เมชัน (MI), วิธีการปรับค่าน้ำหนักคำด้วยเทคนิค rtf-sisf [18] และวิธีการ tf.rf [21] โดยรายละเอียดของแต่ละวิธีการปรับค่าน้ำหนักคำที่นำมาเปรียบเทียบประสิทธิภาพในงานวิจัยนี้แสดงรายละเอียดดังในตารางที่ 3

ตารางที่ 3 : สรุปวิธีการปรับค่าน้ำหนักคำที่ใช้ในการเปรียบเทียบผลการทดลอง

วิธีการ	สัญลักษณ์ที่ใช้	คำอธิบาย
การปรับค่าน้ำหนักคำแบบไม่มีผู้สอน	tf.idf	ใช้สมการ (5)
	rtf-sisf	ใช้สมการ (6)
การปรับค่าน้ำหนักคำแบบมีผู้สอน	tf.idf.icsd	ใช้สมการ (10)
	tf.idf.csd	ใช้สมการ (11)
	tf.idf.sd	ใช้สมการ (12)
	tf.rf	ใช้สมการ (13)
	tf.idf.MI	ใช้สมการ (15)
	tf.idf.cmi	วิธีการที่นำเสนอ

3.2) ผลการทดลองเมื่อเปรียบเทียบประสิทธิภาพบนตัวชี้วัด

ดังที่ได้อธิบายไปแล้วข้างต้นว่าในงานวิจัยนี้จะแบ่งชุดข้อมูลใช้สอนและชุดข้อมูลทดสอบออกเป็น 5 ชุด (5 Fold Cross Validation) โดยได้ทำการทดสอบประสิทธิภาพบนตัวจำแนกประเภทข้อมูล เบย์อย่างง่าย, ป่าของต้นไม้ตัดสินใจแบบสุ่ม และเพื่อนบ้านที่ใกล้ที่สุด K อันดับ โดยค่าความถูกต้องการทำนาย (Acc) เป็นการคำนวณจากค่าเฉลี่ยจากการทำนาย 5 ครั้ง พร้อมทั้งหาค่าส่วนเบี่ยงเบนมาตรฐาน และสำหรับตัวชี้วัดอื่น ๆ ก็จะมีการคำนวณหาค่าเฉลี่ยด้วยเช่นกัน

ตารางที่ 4 ได้แสดงการเปรียบเทียบประสิทธิภาพของตัวชี้วัดต่าง ๆ บนชุดข้อมูล Amazon ซึ่งพบว่าวิธีการที่นำเสนอ cmi สามารถปรับค่าน้ำหนักคำได้ดีกว่าวิธีการ tf.idf, rtf-sisf, MI และการแจกแจงเทอมทั้ง 3 แบบ ในทุกตัวจำแนกประเภทข้อมูล และให้ค่าตัวชี้วัด Acc, Pre, Rec และ F1 สูงที่สุดทุกค่าเมื่อเปรียบเทียบกับวิธีการอื่น ๆ บนตัวจำแนกประเภทเบย์อย่างง่ายและป่าของต้นไม้ตัดสินใจแบบสุ่ม ในขณะที่บนตัวจำแนกประเภทเพื่อนบ้านที่ใกล้ที่สุด K อันดับ วิธีการ tf.rf ให้ผลลัพธ์ของตัวชี้วัดที่สูงกว่า จึงกล่าวได้ว่าบนข้อมูลบทวิจารณ์ Amazon วิธีการปรับค่าน้ำหนักคำที่นำเสนอมีประสิทธิภาพที่ดีกว่าวิธีการอื่น ๆ บนตัวจำแนกประเภทข้อมูลทั้ง เบย์อย่างง่าย และป่าของต้นไม้ตัดสินใจแบบสุ่ม ในขณะที่บนตัวจำแนกประเภทข้อมูล เพื่อนบ้านที่ใกล้ที่สุด K อันดับ วิธีการที่นำเสนอให้ผลที่ดีกว่าวิธีการการปรับค่าน้ำหนักแบบไม่มีผู้สอน rtf-sisf, tf.idf และการปรับค่าน้ำหนักแบบมีผู้สอน ได้แก่ การแจกแจงเทอมทั้ง 3 แบบ และวิธีการ MI แสดงว่าวิธีการ cmi สามารถปรับค่าน้ำหนักคำได้มีประสิทธิภาพที่ดีกว่าวิธีการเหล่านี้

ตารางที่ 4 : ผลการทดลองเปรียบเทียบประสิทธิภาพตัวชี้วัดของวิธีการปรับค่าน้ำหนักบนชุดข้อมูล Amazon

วิธีการปรับค่าน้ำหนัก	เบย์อย่างง่าย				ป่าของต้นไม้ตัดสินใจแบบสุ่ม				เพื่อนบ้านที่ใกล้ที่สุด K อันดับ			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
rtf-sisf	0.689±0.02	0.705	0.693	0.683	0.753±0.02	0.756	0.757	0.752	0.685±0.04	0.692	0.688	0.681
tf×idf	0.713±0.02	0.730	0.714	0.708	0.786±0.03	0.788	0.791	0.785	0.650±0.08	0.686	0.648	0.624
tf×idf×icsd	0.749±0.02	0.760	0.750	0.746	0.779±0.02	0.781	0.784	0.778	0.715±0.03	0.716	0.716	0.713
tf×idf×sqrt(icsd)	0.745±0.02	0.760	0.746	0.741	0.779±0.03	0.781	0.784	0.782	0.740±0.04	0.747	0.742	0.737
tf×idf/csd	0.627±0.03	0.637	0.630	0.621	0.810±0.04	0.813	0.813	0.809	0.511±0.03	0.52	0.521	0.508
tf×idf/sqrt(csd)	0.617±0.03	0.630	0.621	0.610	0.775±0.03	0.777	0.780	0.774	0.534±0.02	0.545	0.543	0.531
tf×idf/sd	0.693±0.01	0.702	0.695	0.689	0.786±0.02	0.788	0.790	0.785	0.643±0.02	0.649	0.648	0.641
tf×idf/sqrt(sd)	0.723±0.03	0.731	0.726	0.720	0.775±0.02	0.776	0.779	0.775	0.683±0.04	0.689	0.69	0.683
tf×idf×cmi	**0.825±0.03**	**0.830**	**0.826**	**0.824**	**0.867±0.02**	**0.866**	**0.868**	**0.866**	0.753±0.03	0.755	0.752	0.750
tf×idf×sqrt(Miik)	0.799±0.03	0.808	0.800	0.797	0.833±0.02	0.834	0.836	0.832	0.710±0.04	0.718	0.71	0.706
tf×idf×MI	0.728±0.02	0.743	0.729	0.724	0.776±0.02	0.78	0.782	0.776	0.661±0.05	0.663	0.659	0.657
tf×idf×sqrt(MI)	0.725±0.02	0.740	0.726	0.721	0.782±0.02	0.784	0.787	0.782	0.666±0.05	0.668	0.664	0.662
tf×rf	0.795±0.02	0.799	0.798	0.794	0.795±0.04	0.8	0.8	0.794	**0.770±0.04**	**0.787**	**0.772**	**0.766**

ตารางที่ 5 : ผลการทดลองเปรียบเทียบประสิทธิภาพตัวชี้วัดของวิธีการปรับค่าน้ำหนักบนชุดข้อมูล Yelp

วิธีการปรับค่าน้ำหนัก	เบย์อย่างง่าย				ป่าของต้นไม้ตัดสินใจแบบสุ่ม				เพื่อนบ้านที่ใกล้ที่สุด K อันดับ			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
rtf-sisf	0.706±0.02	0.729	0.707	0.699	0.750±0.01	0.666	0.664	0.663	0.664±0.03	0.752	0.75	0.749
tf×idf	0.713±0.02	0.73	0.714	0.707	0.757±0.00	0.66	0.625	0.605	0.629±0.02	0.762	0.757	0.756
tf×idf×icsd	0.749±0.02	0.762	0.749	0.745	0.751±0.01	0.726	0.724	0.723	0.725±0.0	0.759	0.752	0.749
tf×idf×sqrt(icsd)	0.745±0.02	0.759	0.745	0.741	0.763±0.02	0.735	0.729	0.729	**0.731±0.01**	0.768	0.763	0.761
tf×idf/csd	0.647±0.03	0.664	0.648	0.638	0.775±0.01	0.521	0.517	0.511	0.519±0.04	0.778	0.774	0.773
tf×idf/sqrt(csd)	0.645±0.03	0.662	0.646	0.636	0.751±0.01	0.546	0.544	0.538	0.544±0.02	0.753	0.75	0.75
tf×idf/sd	0.700±0.02	0.717	0.701	0.694	0.761±0.01	0.653	0.648	0.644	0.647±0.01	0.766	0.761	0.759
tf×idf/sqrt(sd)	0.729±0.02	0.744	0.73	0.725	0.764±0.02	0.667	0.655	0.65	0.657±0.02	0.77	0.764	0.762
tf×idf×cmi	**0.825±0.03**	**0.829**	**0.825**	**0.824**	**0.858±0.01**	0.725	0.72	0.72	0.722±0.02	**0.867**	**0.858**	**0.857**
tf×idf×sqrt(Miik)	0.799±0.03	0.808	0.799	0.797	0.823±0.02	0.691	0.685	0.684	0.688±0.02	0.831	0.823	0.822
tf×idf×MI	0.728±0.02	0.742	0.729	0.724	0.764±0.02	0.631	0.63	0.629	0.632±0.03	0.768	0.764	0.763
tf×idf×sqrt(MI)	0.725±0.02	0.74	0.725	0.72	0.755±0.01	0.631	0.624	0.621	0.627±0.03	0.76	0.755	0.753
tf×rf	0.802±0.03	0.81	0.803	0.801	0.792±0.01	**0.809**	**0.792**	**0.789**	0.716±0.05	0.719	0.714	0.714

เมื่อพิจารณาการการจำแนกประเภทความพึงพอใจที่ทดสอบกับข้อมูล Yelp ดังตารางที่ 5 พบว่าวิธีการ cmi ให้ค่า Acc สูงที่สุดในตัวจำแนกประเภทข้อมูล เบย์อย่างง่าย และป่าของต้นไม้ตัดสินใจแบบสุ่ม และให้ค่าตัวชี้วัด Pre, Rec และ F1 สูงกว่าวิธีการ tf.idf, rtf-sisf, MI, การแจกแจงเทอมทั้ง 3 แบบและ tf.rf บนตัวจำแนกประเภท เบย์อย่างง่าย และ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ ในขณะที่การทดสอบบนป่าของต้นไม้ตัดสินใจแบบสุ่มนั้นวิธีการ tf.rf ให้ค่าที่ดีกว่าบนตัวชี้วัด Pre, Rec และ F

ในขณะที่วิธีการ icsd ซึ่งเป็นการแจกแจงเทอมรูปแบบหนึ่งให้ค่าความถูกต้องในการทำนายสูงที่สุดบนเพื่อนบ้านที่ใกล้ที่สุด K อันดับ เมื่อพิจารณาประสิทธิภาพโดยรวมบนข้อมูลบทวิจารณ์ร้านอาหาร Yelp การปรับค่าน้ำหนักคำด้วยวิธีการ cmi ให้ผลลัพธ์ที่ดีกว่า rtf-sisf และ tf.idf อย่างชัดเจน เนื่องจากวิธีการเหล่านี้เป็นการปรับค่าน้ำหนักเทอมแบบไม่มีผู้สอนและพิจารณาความถี่เป็นสำคัญ ซึ่งย่อมให้ผลที่ไม่ดีกว่าวิธีการแบบมีผู้สอน นอกจากนี้วิธีการ cmi ยังให้ผลการทดลองที่ดีกว่าการแจกแจงเทอม (csd และ sd) และวิธีการ MI อย่างเห็นได้ชัดและบางครั้งให้ผลที่ดีกว่า tf.rf ในกรณีส่วนใหญ่

ตารางที่ 6 : ผลการทดลองเปรียบเทียบประสิทธิภาพตัวชี้วัดของวิธีการปรับค่าน้ำหนักบนชุดข้อมูล IMDb

วิธีการปรับค่าน้ำหนัก	เบย์อย่างง่าย				ป่าของต้นไม้ตัดสินใจแบบสุ่ม				เพื่อนบ้านที่ใกล้ที่สุด K อันดับ			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
rtf-sisf	0.674±0.02	0.683	0.679	0.673	0.731±0.02	0.739	0.737	0.731	0.618±0.04	0.651	0.625	0.603
tf×idf	0.683±0.04	0.687	0.686	0.682	0.786±0.03	0.788	0.791	0.785	0.520±0.05	0.626	0.535	0.424
tf×idf×icsd	0.731±0.03	0.734	0.734	0.731	0.779±0.02	0.781	0.784	0.778	0.699±0.02	0.702	0.702	0.698
tf×idf×sqrt(icsd)	0.722±0.03	0.724	0.725	0.721	0.779±0.03	0.781	0.784	0.738	0.623±0.03	0.644	0.63	0.614
tf×idf/csd	0.494±0.07	0.494	0.494	0.491	0.750±0.02	0.753	0.754	0.75	0.433±0.03	0.421	0.442	0.403
tf×idf/sqrt(csd)	0.535±0.03	0.538	0.538	0.533	0.723±0.01	0.727	0.727	0.723	0.461±0.05	0.449	0.469	0.422
tf×idf/sd	0.670±0.03	0.674	0.673	0.669	0.754±0.02	0.757	0.758	0.754	0.547±0.04	0.575	0.559	0.518
tf×idf/sqrt(sd)	0.686±0.03	0.689	0.689	0.685	0.719±0.02	0.724	0.724	0.719	0.554±0.03	0.578	0.564	0.536
tf×idf×cmi	**0.852±0.03**	**0.851**	**0.853**	**0.851**	**0.867±0.02**	**0.866**	**0.868**	**0.866**	**0.753±0.03**	**0.754**	**0.751**	**0.75**
tf×idf×sqrt(Miik)	0.826±0.03	0.826	0.827	0.826	0.833±0.02	0.834	0.836	0.832	0.71±0.04	0.717	0.71	0.706
tf×idf×MI	0.687±0.04	0.689	0.69	0.686	0.776±0.02	0.78	0.782	0.718	0.66±0.05	0.662	0.659	0.657
tf×idf×sqrt(MI)	0.682±0.04	0.684	0.685	0.681	0.782±0.02	0.784	0.787	0.782	0.665±0.05	0.668	0.663	0.661
tf×rf	0.789±0.03	0.79	0.791	0.788	0.757±0.05	0.771	0.764	0.755	0.670±0.04	0.737	0.679	0.649

ตารางที่ 7 : ผลการทดลองเปรียบเทียบประสิทธิภาพตัวชี้วัดของวิธีการปรับค่าน้ำหนักบนชุดข้อมูล Magazine

วิธีการปรับค่าน้ำหนัก	เบย์อย่างง่าย				ป่าของต้นไม้ตัดสินใจแบบสุ่ม				เพื่อนบ้านที่ใกล้ที่สุด K อันดับ			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
rtf-sisf	0.627±0.01	0.514	0.549	0.455	0.938±0.01	0.877	0.618	0.666	0.943±0.01	0.895	0.877	0.7
tf×idf	0.625±0.01	0.514	0.548	0.455	0.940±0.02	0.927	0.612	0.663	0.938±0.02	0.903	0.596	0.637
tf×idf×icsd	0.700±0.01	0.531	0.593	0.501	0.941±0.01	0.944	0.612	0.663	0.942±0.02	0.844	0.666	0.713
tf×idf×sqrt(icsd)	0.637±0.01	0.515	0.549	0.46	0.941±0.02	0.944	0.618	0.671	0.939±0.02	0.852	0.621	0.669
tf×idf/csd	0.603±0.01	0.506	0.52	0.438	0.945±0.02	0.944	0.642	0.7	0.840±0.01	0.513	0.511	0.509
tf×idf/sqrt(csd)	0.610±0.01	0.509	0.532	0.444	0.940±0.02	0.891	0.617	0.668	0.908±0.01	0.616	0.559	0.572
tf×idf/sd	0.619±0.01	0.516	0.555	0.454	0.941±0.01	0.944	0.612	0.663	0.938±0.01	0.918	0.595	0.639
tf×idf/sqrt(sd)	0.635±0.01	0.511	0.538	0.455	0.942±0.02	0.941	0.623	0.676	0.932±0.01	0.783	0.617	0.648
tf×idf×cmi	0.729±0.01	0.542	0.623	0.523	0.971±0.01	**0.978**	**0.825**	**0.882**	**0.965±0.01**	**0.974**	**0.773**	**0.839**
tf×idf×sqrt(Miik)	0.682±0.01	0.526	0.582	0.489	0.958±0.02	0.978	0.73	0.797	0.943±0.02	0.96	0.621	0.669
tf×idf×MI	0.680±0.02	0.523	0.575	0.486	0.940±0.01	0.940	0.606	0.656	0.936±0.01	0.815	0.603	0.645
tf×idf×sqrt(MI)	0.630±0.01	0.515	0.551	0.457	0.942±0.02	0.944	0.623	0.677	0.936±0.02	0.906	0.582	0.619
tf×rf	**0.904±0.02**	**0.662**	**0.69**	**0.67**	**0.986±0.00**	0.798	0.707	0.736	0.783±0.03	0.558	0.636	0.555

อย่างไรก็ดีเมื่อพิจารณาเมื่อพิจารณาดูประสิทธิภาพของวิธีการต่าง ๆ ในการปรับค่าน้ำหนักคำบนข้อมูลบทวิจารณ์ IMDb ดังในตารางที่ 6 จะพบว่าวิธีการที่นำเสนอให้ผลลัพธ์ที่ดีที่สุดในทุกตัวชี้วัดสำหรับการวัดประสิทธิภาพการสืบค้นเอกสารข้อความบทวิจารณ์และมีประสิทธิภาพสูงกว่าทุกวิธีที่มีการเปรียบเทียบและดีกว่าอย่างน่าประทับใจ โดยเฉพาะอย่างยิ่งเมื่อเปรียบเทียบกับวิธีการ rtf-sisf, tf.idf, การแจกแจงเทอมทั้ง 3 วิธี, MI และ tf.rf ซึ่งเป็นเครื่องบ่งชี้ได้ว่าหลักการที่นำเสนอนี้สามารถปรับปรุงการปรับค่าน้ำหนักคำ tf.idf ให้มีพลังอำนาจในการจำแนกเอกสารข้อความบทวิจารณ์ของคลาสบวกออกจากเอกสารข้อความคลาสลบได้อย่างมีประสิทธิภาพ

ในตารางที่ 7 เป็นการพิจารณาผลการทดลองบนชุดข้อมูลบทวิจารณ์ Magazine ซึ่งพบว่าโดยรวมแล้ววิธีการที่นำเสนอก็ยังคงให้ประสิทธิภาพของตัวชี้วัดดีกว่าวิธีการอื่น ๆ บนตัวจำแนกประเภทเอกสารข้อความ ป่าของต้นไม้ตัดสินใจแบบสุ่ม และ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ ในขณะที่วิธี tf.rf ให้ผลของประสิทธิภาพของตัวชี้วัดดีกว่าบนตัวจำแนกประเภทเอกสารข้อความเบย์อย่างง่ายและให้ Acc ดีที่สุดบนป่าของต้นไม้ตัดสินใจแบบสุ่ม

ตารางที่ 8 : ผลการทดลองเปรียบเทียบประสิทธิภาพตัวชี้วัดของวิธีการปรับค่าน้ำหนักบนชุดข้อมูล Eco-Pre

วิธีการปรับค่าน้ำหนัก	เบย์อย่างง่าย				ป่าของต้นไม้ตัดสินใจแบบสุ่ม				เพื่อนบ้านที่ใกล้ที่สุด K อันดับ			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
rtf-sisf	0.502±0.03	0.456	0.52	0.437	0.713±0.00	0.769	0.726	0.726	0.854±0.00	0.571	0.59	0.55
tf×idf	0.529±0.03	0.476	0.537	0.461	0.462±0.01	0.797	0.695	0.724	0.858±0.00	0.548	0.48	0.379
tf×idf×icsd	0.649±0.02	0.552	0.618	0.555	0.803±0.01	0.804	0.697	0.727	0.859±0.01	0.656	0.664	0.657
tf×idf×sqrt(icsd)	0.573±0.03	0.501	0.563	0.494	0.771±0.01	0.79	0.687	0.716	0.853±0.00	0.645	0.644	0.621
tf×idf/csd	0.485±0.03	0.446	0.507	0.424	0.535±0.01	0.817	0.746	0.76	0.877±0.01	0.392	0.462	0.379
tf×idf/sqrt(csd)	0.488±0.03	0.447	0.508	0.425	0.480±0.02	0.797	0.721	0.738	0.862±0.00	0.441	0.464	0.359
tf×idf/sd	0.517±0.03	0.468	0.529	0.451	0.620±0.01	0.804	0.693	0.723	0.858±0.00	0.529	0.481	0.441
tf×idf/sqrt(sd)	0.550±0.03	0.483	0.544	0.474	0.645±0.01	0.801	0.693	0.723	0.857±0.00	0.587	0.577	0.492
tf×idf×cmi	**0.832±0.01**	**0.758**	**0.753**	**0.738**	0.815±0.01	**0.881**	**0.828**	**0.842**	**0.917±0.00**	0.708	0.676	0.661
tf×idf×sqrt(Miik)	0.690±0.01	0.627	0.642	0.605	0.761±0.01	0.858	0.77	0.797	0.893±0.00	0.689	0.638	0.606
tf×idf×MI	0.610±0.02	0.517	0.579	0.517	0.746±0.01	0.798	0.699	0.726	0.859±0.00	0.541	0.53	0.531
tf×idf×sqrt(MI)	0.557±0.03	0.489	0.552	0.48	0.698±0.02	0.798	0.699	0.728	0.857±0.01	0.517	0.527	0.507
tf×rf	0.744±0.014	0.624	0.64	0.621	**0.906±0.01**	0.829	0.818	0.82	0.893±0.013	**0.825**	**0.82**	**0.812**

จากการทดลองพบว่าวิธีการแจกแจงเทอม icsd นั้นจะให้ผลที่ดีกว่า rtf-sisf, tf.idf, csd และ sd ซึ่งก็สอดคล้องกับงานวิจัยก่อนหน้านี้ที่ได้ศึกษากันมา [9], [10], [19], [20] โดยปกติแล้วหลักการแจกแจงเทอมจะพยายามเพิ่มค่าน้ำหนักของคำที่มาจากคลาสเดียวกันและลดค่าน้ำหนักของคำที่อยู่คนละคลาสกันโดยใช้หลักการทางสถิติโดยการคำนวณหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคำสำหรับแต่ละเอกสาร แต่อย่างไรก็ดีเมื่อเปรียบเทียบวิธีการ icsd กับ cmi แล้วจะพบว่าวิธีการ cmi ให้ผลการทดลองที่ดีกว่า icsd อย่างเห็นได้ชัดในทุกตัวชี้วัด ซึ่งเป็นสิ่งที่ยืนยันได้ว่าวิธีการปรับค่าน้ำหนักคำ rtf-sisf, tf.idf, icsd, csd และ sd ไม่เหมาะที่จะนำมาประยุกต์ใช้ในการจำแนกประเภทความพึงพอใจของเอกสารบทวิจารณ์ โดยวิธีการ cmi และ tf.rf จะมีความเหมาะสมกว่าตามลำดับ

เมื่อพิจารณาตารางที่ 8 เป็นผลการทดลองการเปรียบเทียบตัวชี้วัดบนการทดสอบกับข้อมูลบทวิจารณ์ Eco-Pre ซึ่งเป็นข้อมูลที่ประกอบด้วย 3 คลาส คือ เชิงบวก เชิงลบ และเป็นกลาง โดยหลักการแล้วการจำแนกประเภทความพึงพอใจงานวิจัยโดยส่วนใหญ่จะพิจารณาเฉพาะข้อมูลที่มี 2 คลาสเท่านั้น คือ เชิงบวกและเชิงลบ เนื่องจากคลาสเป็นกลางไม่ได้สะท้อนถึงความรู้สึกที่แท้จริงของผู้บริโภคที่มีต่อสินค้าในการปรับปรุงและพัฒนาผลิตภัณฑ์และบริการ แต่ในการทดลองนี้จะนำข้อมูลนี้เพื่อมาเปรียบเทียบเพื่อดูประสิทธิภาพของตัวชี้วัดต่าง ๆ เมื่อข้อมูลประกอบด้วย 3 คลาส ว่าส่งผลอย่างไรต่อวิธีการกำหนดค่าน้ำหนักคำของวิธีการต่าง ๆ ทั้งในเรื่องของประสิทธิภาพของตัวชี้วัดการสืบค้นเอกสารข้อความและในเรื่องของเวลาที่ใช้ในการคำนวณ ซึ่งวิธีการปรับค่าน้ำหนักคำแบบมีผู้สอน cmi และ tf.rf ให้ค่าตัวชี้วัดสูงกว่าและมีประสิทธิภาพดีกว่าวิธีการอื่น ๆ ซึ่งเมื่อเปรียบเทียบกับวิธีการ rtf-sisf, tf.idf, การแจกแจงเทอมทั้ง 3 วิธี และ MI ซึ่งเป็นสิ่งที่ยืนยันได้ว่าการพิจารณาค่าความถี่ของคำและความถี่เอกสารผกผัน (rtf-sisf และ tf.idf) จะมีประสิทธิภาพไม่เพียงพอในการจำแนกประเภทเอกสารข้อความบทวิจารณ์ เนื่องจากข้อความบทวิจารณ์นั้นจะมีประโยคที่ยาวไม่มากและคำที่มีพลังอำนาจในการแยกแยะคลาสเชิงบวกและเชิงลบก็ไม่ได้เกิดขึ้นบ่อยและไม่ค่อยเกิดขึ้นซ้ำ ๆ ในประโยคเหมือนกับในเอกสารอื่น ๆ เช่น หนังสือ หรือ หนังสือพิมพ์

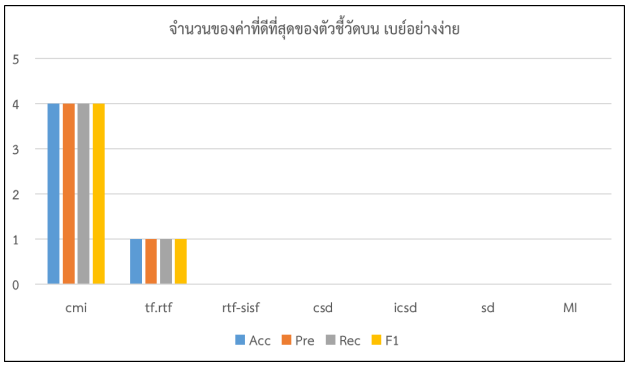
ในขณะที่วิธีการ tf.rf เมื่อพิจารณาจากการทดลองบนข้อมูล Eco-Pre พบว่าให้ผลลัพธ์ที่ดีที่รองจากวิธีการที่นำเสนอเนื่องจากให้ค่าความถูกต้องจากการทำนาย Acc สูงที่สุดบน ป่าของต้นไม้ตัดสินใจแบบสุ่ม และบนตัวจำแนกประเภทเอกสารข้อความเพื่อนบ้านที่ใกล้ที่สุด K อันดับ ก็ให้ค่าตัวชี้วัด Pre, Rec และ F1 ดีที่สุด ในขณะที่วิธีการ cmi ให้ค่า Acc ดีที่สุดบน เบย์อย่างง่าย และ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ และสำหรับค่าตัวชี้วัด Pre, Rec และ F1 ก็ให้ผลลัพธ์ที่ดีที่สุดบนตัวจำแนกประเภทเอกสารข้อความบน เบย์อย่างง่าย และ ป่าของต้นไม้ตัดสินใจแบบสุ่ม

จากการทดลองบนชุดข้อมูล Eco-Pre ในเรื่องของเวลาในการคำนวณสามารถพิจารณาได้จากตารางที่ 9 ซึ่งพบว่าวิธีการแจกแจงเทอมจะใช้เวลาในการคำนวณนานที่สุด พบว่าวิธีการแจกแจงเทอมทั้ง 3 แบบใช้เวลาคำนวณนานกว่าวิธีการที่นำเสนอ โดย

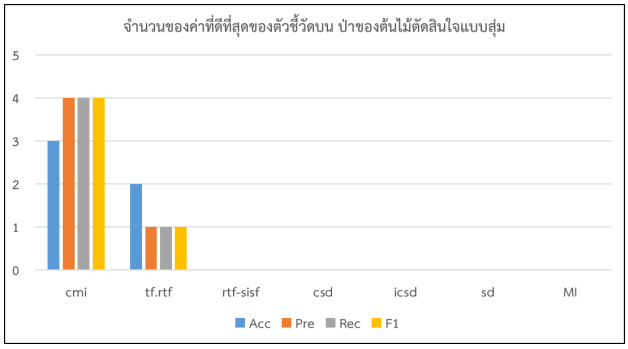
ใช้เวลานานกว่าประมาณ 800 เท่าสำหรับวิธีการ icsd ในขณะที่ csd และ sd จะใช้เวลานานกว่า 1500 เท่า ทั้งนี้เวลาในการคำนวณนั้นก็จะเพิ่มขึ้นอย่างทวีคูณตามจำนวนเอกสาร จำนวนคำ และจำนวนคลาสของเอกสาร ดังนั้นวิธีการปรับค่าน้ำหนักด้วยการแจกแจงเทอมจึงไม่เหมาะที่จะนำไปใช้กับข้อมูลเอกสารขนาดใหญ่และยิ่งมีจำนวนคลาสมากก็ยิ่งช้าเป็นทวีคูณ

ตารางที่ 9 : การเปรียบเทียบเวลาในการคำนวณการปรับค่าน้ำหนัก

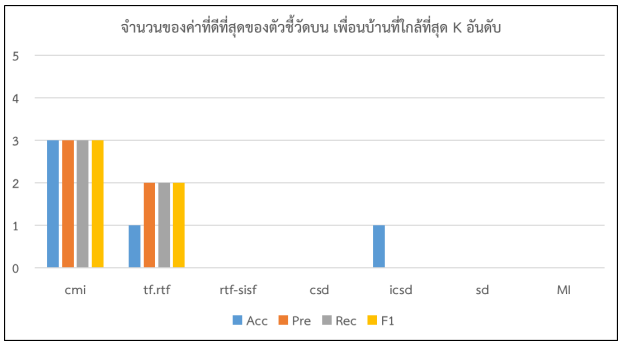
วิธีการปรับค่าน้ำหนัก	เวลาที่ใช้ประมวลผล (วินาที)
csd	26,020
sd	26,020
icsd	14,241
cmi	17
tf.idf, tf.rt, rtf-sisf	< 17



รูปที่ 7 : จำนวนของค่าที่ดีที่สุดของตัวชี้วัดประสิทธิภาพการสืบค้นเอกสารข้อความบนเบย์อย่างง่าย



รูปที่ 8 : จำนวนของค่าที่ดีที่สุดของตัวชี้วัดประสิทธิภาพการสืบค้นเอกสารข้อความบนป่าของต้นไม้ตัดสินใจแบบสุ่ม



รูปที่ 9 : จำนวนของค่าที่ดีที่สุดของตัวชี้วัดประสิทธิภาพการสืบค้นเอกสารข้อความบนเพื่อนบ้านที่ใกล้ที่สุด K อันดับ

รูปที่ 7–9 ได้สรุปถึงประสิทธิภาพตัวชี้วัดของการสืบค้นเอกสาร โดยแสดงค่าที่สูงที่สุดของแต่ละวิธีที่ได้รับบนตัวจำแนกประเภทเอกสารข้อความ เบย์อย่างง่าย, ป่าของต้นไม้ตัดสินใจแบบสุ่ม และ เพื่อนบ้านที่ใกล้ที่สุด K อันดับ ที่ได้มีการทดสอบกับชุดข้อมูลบทวิจารณ์ทั้ง 5 ชุด เมื่อพิจารณาจากกราฟที่แสดงนั้นจะเห็นได้ว่าการปรับค่าน้ำหนัก ด้วยคลาสมิวชวลอินฟอร์เมชัน cmi จะมีประสิทธิภาพที่ดีกว่าวิธีการปรับค่าน้ำหนักวิธีการอื่น ๆ โดยเฉพาะอย่างยิ่งเมื่อเปรียบเทียบกับวิธีการปรับค่าน้ำหนักคำแบบดั้งเดิมที่นิยมใช้กัน tf.idf และ rtf-sisf ที่พิจารณาใช้ความถี่ของคำและความถี่เอกสารผกผัน รวมถึงการแจกแจงเทอมทั้ง 3 แบบ (icsd, csd, และ sd) ที่ใช้กับหลักการทางสถิติซึ่งเป็นสิ่งที่ยืนยันถึงประสิทธิภาพของวิธีการปรับค่าน้ำหนักคำที่นำเสนอนี้ได้เป็นอย่างดี

5) สรุปผลการทดลอง

งานวิจัยนี้มีแนวคิดโดยใช้มิวชวลอินฟอร์เมชันในการหาความสัมพันธ์ร่วมระหว่างคำและคลาสของประเภทความพึงพอใจ การปรับค่าน้ำหนักคำ tf.idf ด้วยวิธีการที่นำเสนอนี้ช่วยเพิ่มประสิทธิภาพในการสืบค้นเอกสารข้อความบทวิจารณ์ อย่างไรก็ตามวิธีการนี้จะต้องมีการปรับค่าความถี่คำให้เป็นแบบไบนารีก่อนเพื่อคำนวณค่าความน่าจะเป็นจึงทำให้ต้องเสียเวลาในขั้นตอนนี้ รวมถึงอาจจะทำให้ข่าวสารบางอย่างหายไป ซึ่งงานวิจัยในอนาคตจะคำนวณค่ามิวชวลอินฟอร์เมชันโดยตรงจากค่าความถี่ของคำด้วยเทคนิคอื่น ๆ เช่น ฟัซซี่เซต (Fuzzy sets) หรือ การให้เหตุผลบนหลักการกรณีศึกษา (Case based reasoning) เป็นต้น

เทคนิคที่นำเสนอนี้ให้ความสำคัญกับจำนวนเอกสารที่อยู่ในคลาสบวก ถ้าคำใดให้ค่าข่าวสารร่วมกับคลาสบวกมากก็จะถูก

ปรับค่าน้ำหนักคำเพื่อเพิ่มอำนาจการจำแนกให้มากขึ้น ซึ่งต่างจากวิธี tf.rf ที่พิจารณาสัดส่วนจำนวนของเอกสารคลาสบวกต่อคลาสลบ ซึ่งอาจจะประสบปัญหาในกรณีที่คำต่าง ๆ มีสัดส่วนของคลาสบวกต่อคลาสลบเท่ากัน ซึ่งก็จะมีการปรับค่าน้ำหนักคำให้เท่ากัน ในขณะที่วิธีการ cmi อาจจะประสบปัญหาในกรณีที่คำต่าง ๆ มีจำนวนเอกสารในคลาสบวกเท่ากันแต่ในคลาสลบแตกต่างกันก็จะมีการปรับค่าน้ำหนักคำเหล่านั้นเท่ากัน ซึ่งถ้ามีการพิจารณาคลาสลบด้วยจะสามารถช่วยปรับค่าน้ำหนักคำต่าง ๆ ให้เหมาะสมได้

การทดลองการจำแนกประเภทความพึงพอใจด้วยเบย์อย่างง่ายพบว่าวิธีการที่นำเสนอให้ค่าความถูกต้องเฉลี่ยและค่าเฉลี่ยของตัวชี้วัดอื่น ๆ สูงกว่าวิธีการต่าง ๆ แสดงว่าวิธีการที่นำเสนอสามารถทำให้คำในถุงของคำมีความเป็นอิสระกันมากกว่าวิธีการอื่น ๆ ในขณะที่การวัดประสิทธิภาพบนป่าของต้นไม้ตัดสินใจแบบสุ่ม ก็ให้ผลใกล้เคียงกับเบย์อย่างง่าย เนื่องจากตัวจำแนกประเภทชนิดนี้มีการพิจารณาความสำคัญของคำเช่นเดียวกับเทคนิคที่นำเสนอ และเมื่อพิจารณาจากเพื่อนบ้านที่ใกล้ที่สุด K อันดับพบว่าวิธีการ tf.rf และ icsd ให้ผลลัพธ์ซึ่งบางครั้งสูงกว่าวิธีการที่นำเสนอ เนื่องจากวิธีการ tf.rf นี้ช่วยปรับให้เอกสารที่อยู่ในคลาสเดียวกันอยู่ใกล้กัน ในขณะที่วิธีการ icsd จะพยายามทำให้กลุ่มของคำที่มาจากคลาสเดียวกันมีความกระชับและอยู่ใกล้กัน ดังนั้นจากผลการทดลองสามารถกล่าวได้ว่า วิธีการแจกแจงคำทั้ง 3 แบบ รวมถึง tf.idf และ rtf-sisf จึงไม่เหมาะที่จะนำมาใช้ในการจำแนกประเภทความพึงพอใจเนื่องจากแต่ละบทวิจารณ์เป็นประโยคไม่ยาว โดยวิธีการที่นำเสนอนี้มีความเหมาะสมกว่าทั้งในแง่ประสิทธิภาพและเวลาที่ใช้ จึงเหมาะกับการวิเคราะห์ความพึงพอใจแบบประมวลผลทันทีเพื่อการปรับปรุงการให้บริการและคุณภาพของสินค้าให้ตรงกับความต้องการของลูกค้าและทันสถานการณ์มากที่สุด

REFERENCES

[1] O. Gokalp, E. Tasci, and A. Ugur, “A novel wrapper feature selection algorithm based on iterated greedy metaheuristic for sentiment classification,” *Expert Syst. Appl.*, vol. 146, May 2020, Art. no. 113176.

[2] W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” *Ain Shams Eng. J.,* vol. 5, no. 4, pp. 1093–1113, Dec. 2014.

[3] D. M. E.-D. M. Hussein, “A survey on sentiment analysis challenges,” *J. King Saud Univ. – Eng. Sci.*, vol. 30, no.4, pp. 330–338, Oct. 2018.

[4] G. Wang, J. Sun, J. Ma, K. Xu, and J. Gu, “Sentiment classification: The contribution of ensemble learning,” *Decis. Support Syst.*, vol. 57, pp. 77–93, Jan. 2014.

[5] J. Chen, J. Yu, S. Zhao, and Y. Zhang, “User’s review habits enhanced hierarchical neural network for document-level sentiment classification,” *Neural Process. Lett.*, vol. 53, pp. 2095–2111, Apr. 2021.

[6] G. Wang, Z. Zhang, J. Sun, S. Yang, and C. A. Larson, “POS-RS: A Random Subspace method for sentiment classification based on part-of-speech analysis,” *Inf. Process. & Manage.*, vol. 51, no. 4, pp. 458–479, Jul. 2015.

[7] C. Yang, X. Chen, L. Liu, and P. Sweetser, “Leveraging semantic features for recommendation: Sentence-level emotion analysis,” *Inf. Process. Manage*., vol. 58, no. 3, May 2021, Art. no. 102543.

[8] R. K. Yadav, L. Jiao, O.-C. Granmo, and M. Goodwin, “Human-level interpretable learning for aspect-based sentiment analysis,” in *Proc. 35th AAAI Conf. Artif. Intell. (AAAI-21)*, Palo Alto, CA, USA, Feb. 2021, pp. 14023–14212.

[9] V. Lertnattee and T. Theeramunkong, “Effect of term distributions on centroid-based text categorization,” *Inf. Sci.*, vol. 158, pp. 89–115, Jan. 2004.

[10] U. Buatoom, W. Kongprawechnon, and T. Theeramunkong, “Improving seeded k-means clustering with deviation- and entropy-based term weightings,” *IEICE Trans. Inf. Syst.*, vol. E103–D, no. 4, pp. 748–758, Apr. 2020.

[11] H. Liu, X. Chen, and X. Liu, “A study of the application of weight distributing method combining sentiment dictionary and TF-IDF for text sentiment analysis,” *IEEE Access*, vol. 10, pp. 32280–32289, 2022, doi: 10.1109/ACCESS.2022.3160172.

[12] I. Almalis, E. Kouloumpris, and I. Vlahavas, “Sector-level sentiment analysis with deep learning,” *Knowledge-Based Syst*., vol. 258, 2022, Art. no. 109954.

[13] Y. Dang, Y. Zhang, and H. Chen, “A lexicon-enhanced method for sentiment classification: An experiment on online product reviews,” *IEEE Intell. Syst.*, vol. 25, no. 4, pp. 46–53, 2010.

[14] S. Foithong, O. Pinngern, and B. Attachoo, “Feature subset selection wrapper based on mutual information and rough sets,” *Expert Syst. Appl.*, vol. 39, no. 1, pp. 574–584, 2012.

[15] A. K. Paul and P. C. Shill, "Sentiment mining from Bangla data using mutual information," in *Proc. 2nd Int. Conf. Elect., Comput. & Telecommun. Eng. (ICECTE)*, Rajshahi, Bangladesh, Dec. 2016, pp. 1–4.

[16] X. -Y. Jiang and J. Shui, "An improved mutual information-based feature selection algorithm for text classification," in *Proc. 5th Int. Conf. Intell. Human-Machine Syst. and Cybernetics*, Hangzhou, China, Aug. 2013, pp. 126–129.

[17] A. Bagheri, M. Saraee, and F. de Jong, "Sentiment classification in Persian: Introducing a mutual information-based method for feature selection," in *Proc. 21st Iranian Conf. Elect. Eng. (ICEE)*, Mashhad, Iran, May 2013, pp. 1–6.

[18] J. M. Sanchez-Gomez, M. A. Vega-Rodríguez, and C. J. Pérez, "The impact of term-weighting schemes and similarity measures on extractive multi-document text summarization," *Expert Syst. Appl.*, vol. 169, May 2021, Art. no. 114510.

[19] V. Lertnattee and T. Theeramunkong, "Multidimensional text classification for drug information," *IEEE Trans. Inf. Technol. Biomed.,* vol. 8, no. 3, pp. 306–312, Sep. 2004, doi: 10.1109/TITB.2004.832542.

[20] U. Buatoom, W. Kongprawechnon, and T. Theeramunkong, "Document clustering using k-means with term weighting as similarity-based constraints," *Symmetry*, vol. 12, no. 6, p. 967, 2020.

[21] M. Lan, C. L. Tan, J. Su, and Y. Lu, "Supervised and traditional term weighting methods for automatic text categorization," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 31, no. 4, pp. 721–735, Apr. 2009.

[22] Z. Feng, H. Zhou, Z. Zhu, and K. Mao, " Tailored text augmentation for sentiment analysis," *Expert Syst. Appl.*, vol. 205, Nov. 2022, Art. no. 117605.

[23] H. Zhao, Z. Liu, X. Yao, and Q. Yang, "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach," *Infor. Process. Manage.*, vol. 58, Sep. 2021, Art. no. 102656.

[24] I. Chaturvedi, E. Cambria, R. E. Welsch, and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges", *Inf. Fusion*, vol. 44, pp. 65–77, 2018.

[25] D. Kotzias, M. Denil, N. D. Freitas, and P. Smyth, "From group to individual labels using deep features," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery and Data Mining*, Sydney, Australia, Aug. 2015, pp. 597–606.

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

แบบจำลองการพยากรณ์ปริมาณฝุ่นละออง $PM_{2.5}$ ของจังหวัดพะเยา

คมกฤต เมฆสกุล[1*] ทรงวุฒิ ประกายวิเชียร[2] อธิคม บุญซื่อ[3]

[1*,2,3]*สาขาวิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยพะเยา, พะเยา, ประเทศไทย*

*ผู้ประพันธ์บรรณกิจ อีเมล : kit_9m@hotmail.com

รับต้นฉบับ : 3 กุมภาพันธ์ 2566; รับบทความฉบับแก้ไข : 25 เมษายน 2566; ตอบรับบทความ : 26 เมษายน 2566
เผยแพร่ออนไลน์ : 27 ธันวาคม 2566

บทคัดย่อ

ฝุ่นละอองในบรรยากาศขนาดเล็กไม่เกิน 2.5 ไมครอน ($PM_{2.5}$) เป็นปัญหามลพิษทางอากาศที่สำคัญในพื้นที่ภาคเหนือตอนบนของประเทศไทยซึ่งส่งผลกระทบต่อสุขภาพของประชาชนทั้งทางตรงและทางอ้อม โดยจังหวัดพะเยามีมลภาวะนี้จัดอยู่ 3 ลำดับแรกของภาคเหนือ การพยากรณ์ระดับความเข้มข้นของปริมาณฝุ่นละอองในบรรยากาศขนาดเล็กไม่เกิน 2.5 ไมครอน โดยใช้แบบจำลองทางคณิตศาสตร์ เป็นทางเลือกหนึ่งที่สามารถนำไปใช้ประกอบเป็นแนวทางในการวางแผนการแก้ไขหรือป้องกันปัญหามลพิษอากาศ การสร้างแบบจำลองการพยากรณ์ด้วยการประยุกต์ใช้วิธีการวิเคราะห์การถดถอยโลจิสติก โดยใช้ชุดข้อมูลจากกรมควบคุมมลพิษอากาศและกรมอุตุนิยมวิทยาที่ตรวจวัดสภาพอากาศในจังหวัดพะเยา ซึ่งนำข้อมูลสถิติย้อนหลัง 5 ปี ในช่วงปี พ.ศ. 2558–2564 ไปสร้างแบบจำลองการพยากรณ์ด้วยโปรแกรมภาษาไพทอน ผลการทดสอบ พบว่า ตัวแปรอิสระที่นำไปสร้างสมการจำลองทางคณิตศาสตร์สำหรับการพยากรณ์ค่า $PM_{2.5}$ ของงานวิจัยนี้ ได้แก่ ค่า PM_{10}, NO_2, O_3 และ CO เมื่อนำมาตรวจสอบความถูกต้องของแบบจำลองการพยากรณ์ โดยให้ค่าความไวและค่าความจำเพาะเท่ากับ 0.96 และ 0.72 ตามลำดับ โดยมีค่าพื้นที่ใต้เส้นโค้งเท่ากับ 0.98

คำสำคัญ : มลพิษอากาศ การถดถอยโลจิสติก สมการพยากรณ์

Journal of Engineering and Digital Technology (JEDT)
Vol.11 No.2 July - December 2023

The Prediction Model of $PM_{2.5}$ in Phayao Province

Khomkit Meksagul[1*] Songwut Prakaiwichien[2] Athikom Boonsue[3]

[1*,2,3]*Industrial Engineering Program, School of Engineering, University of Phayao, Phayao, Thailand*

*Corresponding Author. E-mail address: kit_9m@hotmail.com

Received: 3 February 2023; Revised: 25 April 2023; Accepted: 26 April 2023
Published online: 27 December 2023

Abstract

The dust in the atmosphere (Particulate Matter with a diameter less than or equal to 2.5 micrometers; $PM_{2.5}$) is the most significant air pollution in the northern region of Thailand, directly and indirectly affecting the health of the people. This pollution in Phayao province has consistently ranked in the top 3 of Thailand's northern regions. Forecasting the concentration of $PM_{2.5}$ in the atmosphere using a mathematical model is an alternative that can be employed as a guideline for planning solutions or preventing air pollution problems. Logistic regression analysis and data from the Thai Meteorological Department and Pollution Control Department were applied to this prediction model. The 5 years of data between 2015–2021 were utilized for this model and created using Python programming. The results demonstrate that the independent variables of this model are PM_{10}, NO_2, O_3 and CO. The sensitivity and specificity are 0.96 and 0.72, respectively. The area under the curve (AUC) is 0.98.

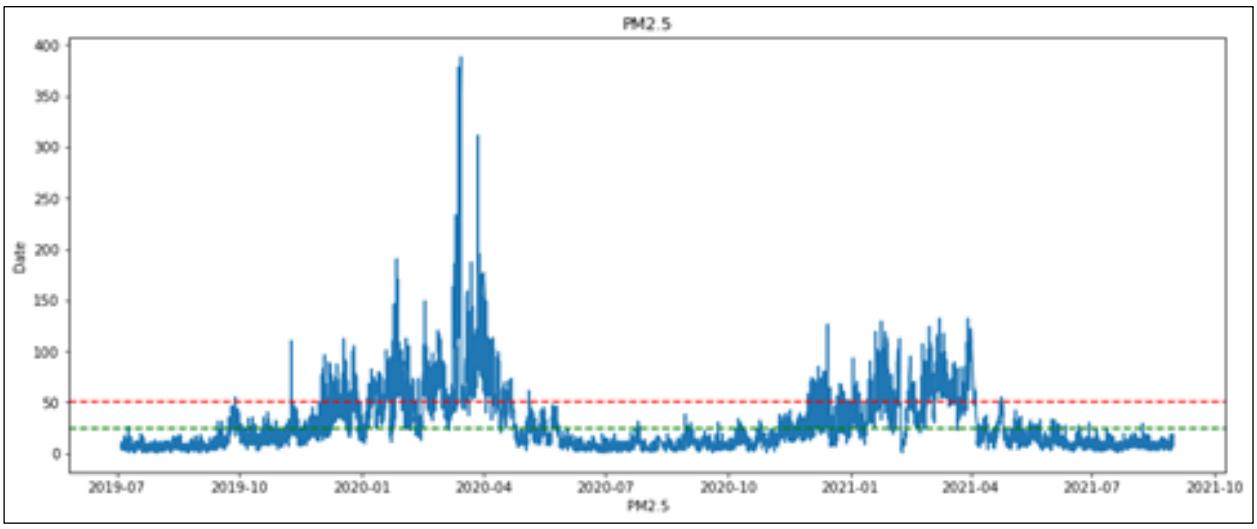
***Keywords*:** Air pollution, Logistic regression, Predict model

1) บทนำ

ปัญหามลพิษทางอากาศเป็นปัญหาที่สำคัญที่ส่งผลกระทบต่อสุขภาพของประชาชนทั้งทางตรงและทางอ้อม ซึ่งเกิดจากฝุ่นละอองในบรรยากาศ เป็นอนุภาคของแข็งขนาดเล็กที่ล่องลอยอยู่ในอากาศ ซึ่งเกิดจากวัตถุที่ถูกทุบ ตี บด กระแทก จนแตกเป็นชิ้นส่วนเล็ก ๆ หรือปฏิกิริยาทางเคมี เมื่อถูกกระแสลมพัด ก็จะปลิวกระจายตัวอยู่ในอากาศ และอนุภาคบางส่วนตกลงสู่พื้น [1] ซึ่งปัญหานี้ส่งผลกระทบต่อความเป็นอยู่ของประชาชนในพื้นที่ โดยเฉพาะผลกระทบต่อสุขภาพ ระบบทางเดินหายใจ สภาพเศรษฐกิจและการท่องเที่ยว การเดินทางและการขนส่ง [2] จากการศึกษาผลกระทบจากฝุ่นละอองขนาดเล็กในบรรยากาศ กับจำนวนผู้ป่วยที่มีอาการของโรคที่เกิดจากการได้รับปริมาณฝุ่นเข้าสู่ร่างกาย มีความสอดคล้องกันตามลำดับเวลา เช่น ช่วงเดือนกุมภาพันธ์ถึงกลางเดือนมีนาคม ซึ่งมีระดับฝุ่นละอองขนาดเล็กปริมาณสูง ก็พบว่า มีปริมาณจำนวนผู้ป่วยเพิ่มมากขึ้น และช่วงเดือนเมษายนถึงเดือนพฤษภาคม ระดับฝุ่นละอองขนาดเล็กมีปริมาณลดต่ำลง จำนวนผู้ป่วยก็มีแนวโน้มลดลง ซึ่งมีสัดส่วนในการเจ็บป่วยในกลุ่มคอหอยอักเสบเฉียบพลัน เจ็บคอมากที่สุด รองลงมาคือ ไข้หวัด เยื่อจมูกอักเสบเฉียบพลัน เยื่อตาอักเสบเฉียบพลัน ต้อเนื้อ และการติดเชื้อทางเดินอาหาร [3]

จังหวัดพะเยาเป็นจังหวัดที่มีฝุ่น $PM_{2.5}$ สูงติด 1 ใน 3 ของภาคเหนือ และสาเหตุหลักของฝุ่นในพะเยาเกิดจากการเผาชีวมวลในที่โล่ง [4] โดยข้อมูลของฝุ่นละออง $PM_{2.5}$ ระยะเวลาตั้งแต่กรกฎาคม พ.ศ. 2562 ถึง สิงหาคม พ.ศ. 2564 จากกรมควบคุมมลพิษของจังหวัดพะเยา แสดงให้เห็นว่าจังหวัดพะเยามีค่า $PM_{2.5}$ เกินค่ามาตรฐาน ซึ่งในบางวันสูงเกือบ 400 ไมโครกรัม ($\mu g/m^3$) ดังรูปที่ 1 โดยค่า $PM_{2.5}$ ในบรรยากาศทั่วไป เฉลี่ย 24 ชั่วโมง ตามมาตรฐานของประเทศไทย อยู่ที่ 50 $\mu g/m^3$ [5]

ดังนั้น ปัญหาฝุ่นละอองในพื้นที่จึงมีความจำเป็นต้องได้รับการแก้ไขอย่างทันท่วงที ซึ่งการนำข้อมูลด้านคุณภาพอากาศที่ตรวจวัดได้ มาทำนายหาค่าระดับ $PM_{2.5}$ ในอนาคต เพื่อบ่งบอกสถานการณ์คุณภาพอากาศ และเตรียมการล่วงหน้าได้มีประสิทธิภาพมากยิ่งขึ้น [6] จึงมีความจำเป็นต้องใช้คอมพิวเตอร์เป็นเครื่องมือช่วยในการหาคำตอบที่ต้องการในรูปแบบของเครื่องจักรการเรียนรู้ (machine learning) ที่เกิดจากการพัฒนาโปรแกรมทางคอมพิวเตอร์ ซึ่งสามารถเรียนรู้ได้จากชุดข้อมูลที่จัดเตรียมไว้ให้ [6] โดยใช้การวิเคราะห์สมการถดถอยโลจิสติก (logistic regression) ซึ่งจะไม่มีข้อกำหนดที่เกี่ยวกับตัวแปรตาม (dependent variable) และความความคลาดเคลื่อนที่จะต้องมีการแจกแจงแบบปกติ (normal distribution) รวมไปถึงความสัมพันธ์เชิงเส้นตรง (linear function) ระหว่างตัวแปรอิสระและตัวแปรตาม นอกจากนี้ การถดถอยโลจิสติกยังสามารถใช้วิเคราะห์ข้อมูลได้ทุกระดับการวัดของตัวแปรอีกด้วย ซึ่งถือว่าเป็นจุดแข็งของเทคนิคทางสถิตินี้ [7]



รูปที่ 1 : ผลการตรวจวัดฝุ่นละออง $PM_{2.5}$ ของ จ.พะเยา

2) ทบทวนวรรณกรรม

2.1) ดัชนีคุณภาพอากาศ (Air Quality Index หรือ AQI)

เป็นดัชนีข้อมูลคุณภาพอากาศในรูปแบบที่ง่ายต่อความเข้าใจ เป็นเผยแพร่ให้ได้รับทราบถึงสถานการณ์มลพิษทางอากาศ ซึ่งดัชนีคุณภาพอากาศ 1 ค่า ใช้เป็นตัวแทนค่าความเข้มข้นของสารมลพิษทางอากาศ 6 ชนิด ได้แก่ ก๊าซโอโซน (O_3) ก๊าซคาร์บอนมอนอกไซด์ (CO) ก๊าซไนโตรเจนไดออกไซด์ (NO_2) ก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน (PM_{10}) และฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน ($PM_{2.5}$)

โดยการคำนวณค่าความเข้มข้นของสารมลพิษทางอากาศ ที่ได้มาจากข้อมูลผลการตรวจวัดคุณภาพอากาศ โดยการคำนวณดัชนีคุณภาพอากาศภายในช่วงระดับเป็นสมการเส้นตรง ดังนี้

$$I = \frac{I_j - I_i}{x_j - x_i}(x - x_i) + I_i \qquad (1)$$

โดย I คือ ค่าดัชนีย่อยคุณภาพอากาศ

x คือ ความเข้มข้นของมลพิษทางอากาศจากการตรวจวัด

x_i, x_j คือ ค่าต่ำสุด, ค่าสูงสุด ของช่วงความเข้มข้นมลพิษ

I_i, I_j คือ ค่าต่ำสุด, ค่าสูงสุด ของช่วงดัชนีคุณภาพอากาศที่ตรงกับช่วงความเข้มข้นของ x

2.2) การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression)

เป็นกระบวนการทางสถิติ เพื่อให้ได้สมการถดถอย สำหรับทำนาย ปรากฏการณ์ต่าง ๆ โดยสิ่งที่ถูกทำนายก็จะถูกเรียกว่า ตัวแปรตาม ส่วนตัวแปรทำนายก็คือ ตัวแปรอิสระ ในกระบวนการทำนายนี้ ตัวแปรตามจะมีเพียงตัวเดียวเท่านั้น ส่วนตัวแปรอิสระอาจจะมีหลายตัวแปร โดยแบ่งออกเป็น

2.2.1) การวิเคราะห์การถดถอยอย่างง่าย (Simple Linear Regression) สมการถดถอยอย่างง่ายเป็นความสัมพันธ์ที่มีตัวแปรตาม และตัวแปรอิสระ อย่างละ 1 ตัว เท่านั้น โดยมีสมการการถดถอยอย่างง่าย ดังนี้

$$z = a + b_0 x_0 + e \qquad (2)$$

2.2.2) การวิเคราะห์การถดถอยพหุคูณ (Multiple Linear Regression) สมการถดถอยพหุคูณ ความสัมพันธ์ที่มีตัวแปรตามเพียง 1 ตัว และมีตัวแปรอิสระ ตั้งแต่ 2 ตัวขึ้นไป ซึ่งสมการถดถอยพหุคูณสามารถแสดงได้ในสมการ (3) ดังต่อไปนี้

$$z = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n + e \qquad (3)$$

โดยที่ z คือ ค่าของตัวแปรตาม

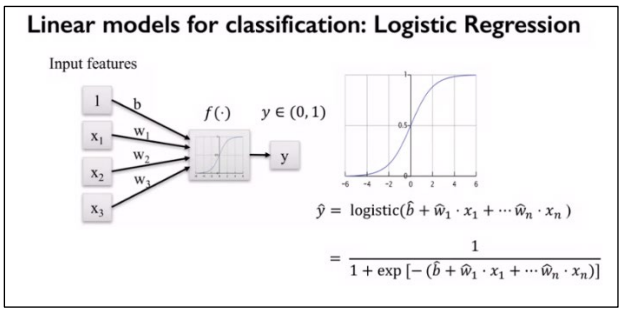
e คือ ค่าความคลาดเคลื่อน (Error or Residual)

a คือ ค่าคงที่ (Constant) ของสมการถดถอย

$x_0, x_1, x_2, \ldots, x_n$ คือ ค่าของตัวแปรอิสระ

$b_0, b_1, b_2, \ldots, b_n$ คือ ค่าสัมประสิทธิ์ของสมการถดถอย

2.2.3) การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression) เป็นการวิเคราะห์ข้อมูลเพื่อทำนาย โอกาสการเกิดเหตุการณ์หนึ่ง ๆ หรือโอกาสเกิดเหตุการณ์มากหรือน้อย โดยมีตัวแปรตัวหนึ่ง หรือหลายตัวที่ส่งผลต่อการเกิดเหตุการณ์นั้น ๆ โดยตัวแปรอิสระจะต้องเป็นตัวแปรข้อมูลในระดับอันตรภาค (interval scale) เป็นอย่างต่ำ กรณีที่เป็นข้อมูลเชิงกลุ่มให้แปลงเป็นตัวแปรหุ่น (dichotomous variable) โดยที่มีค่าเป็น 0 กับ 1 เท่านั้น ส่วนตัวแปรตามจะเป็นการถดถอยโลจิสติกแบบทวิ (Binary Logistic Regression) ซึ่งกำหนดเพียง 2 ค่าเท่านั้น โดยอาจเป็นคำว่า ใช่ (Yes) หรือ ไม่ใช่ (No) ซึ่งจะแทนด้วยค่าเป็น 1 กับ 0 ตามลำดับ ซึ่งสามารถแสดงความสัมพันธ์จากสมการ (3) ได้ดังรูป



รูปที่ 2 : ความสัมพันธ์ของการถดถอยเชิงเส้นพหุคูณ และการถดถอยโลจิสติก [8]

ดังนั้น การถดถอยโลจิสติกจึงสามารถเขียนเป็นสมการ (4) ได้ดังนี้

$$Prob(event) = \frac{1}{1+\exp[-(\hat{a}+\hat{b}_1 x_1+\cdots+\hat{b}_n x_n)]} \qquad (4)$$

2.3) สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation)

เป็นการศึกษาความสัมพันธ์เชิงเส้น (linear relationship) ระหว่างตัวแปร 2 ตัวหรือข้อมูล 2 ชุด โดยมีข้อตกลงเบื้องต้น คือ ตัวแปรทั้งสองจะต้องเป็นตัวแปรข้อมูลระดับอันตรภาค (interval scale) ขึ้นไป และมีการแจกแจงปกติสองตัวแปร (bivariate normal distribution) โดยค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน จะมีค่าอยู่ระหว่าง -1 ถึง +1 ถ้า เข้าใกล้ ±1

หมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์กันอยู่ในระดับมาก ถ้าสัมประสิทธิ์สหสัมพันธ์มีค่าเข้าใกล้ 0 หมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์กันในระดับน้อยหรือไม่สัมพันธ์กันเลย เครื่องหมายบวกลบ (±) จะแสดงทิศทางของความสัมพันธ์ว่ามีทิศทางเป็นแบบใด โดยถ้าสัมประสิทธิ์สหสัมพันธ์มีเครื่องหมายบวก (+) หมายความว่าตัวแปรทั้งสองมีความสัมพันธ์ไปทิศทางเดียวกัน หากมีเครื่องหมายลบ (-) หมายความว่าตัวแปร ทั้งสองมีความสัมพันธ์ไม่ไปทิศทางเดียวกันหรือตรงกันข้าม สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันสามารถคำนวณจากสมการ (5) ดังนี้ [9]

$$r_{xy} = \frac{n(\Sigma xy)-(\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2)-(\Sigma x)^2][n(\Sigma y^2)-(\Sigma y)^2]}} \quad (5)$$

โดยให้ r_{xy} คือ สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน
Σx คือ ผลรวมข้อมูลที่วัดจากตัวแปร x
Σy คือ ผลรวมข้อมูลที่วัดจากตัวแปร y
Σxy คือ ผลรวมผลคูณของตัวแปร x และ ตัวแปร y
Σx^2 คือ ผลรวมกำลังสองข้อมูลที่วัดจากตัวแปร x
Σy^2 คือ ผลรวมกำลังสองข้อมูลที่วัดจากตัวแปร y
n คือ ขนาดตัวอย่าง

2.4) การแปลผลการวิเคราะห์ข้อมูลสหสัมพันธ์และการถดถอย

เป็นการแปลผลการวิเคราะห์ความสัมพันธ์ของข้อมูลสหสัมพันธ์และสมการถดถอย เมื่อพบว่ามีนัยสำคัญทางสถิติที่ .05 หรือ .01 ให้แปลว่า มีความสัมพันธ์กัน โดยทั่วไป การแปลความหมายของขนาดความสัมพันธ์จะพิจารณาจาก ค่าสัมประสิทธิ์สหสัมพันธ์ (ใช้สัญลักษณ์ r) ซึ่งมีค่าอยู่ระหว่าง -1 ถึง +1 ดังนี้ [10]

ถ้า ค่า r มีค่าเท่ากับ 0.00 แปลว่า ไม่มีความสัมพันธ์กัน
ค่า r มีค่าต่ำกว่า 0.40 แปลว่า มีความสัมพันธ์น้อย
ค่า r มีค่า 0.40 ถึง 0.60 แปลว่า มีความสัมพันธ์ปานกลาง
ค่า r มีค่า 0.60 ขึ้นไป แปลว่า มีความสัมพันธ์มาก

การแปลความหมายทิศทางของความสัมพันธ์ คือ หากค่าที่คำนวณได้เป็น บวก หมายถึง มีสัมพันธ์ในทิศทางเดียวกัน แต่ถ้าค่าที่คำนวณได้มีเครื่องหมายเป็น ลบ หมายถึง มีสัมพันธ์ในทิศทางตรงกันข้าม

สำหรับการแปลความหมายของสมการถดถอย จะพิจารณาจากค่าสัมประสิทธิ์การถดถอย ซึ่งใช้สัญลักษณ์ b (หรือ Beta) จากสมการ (2) และ (3) ซึ่งแปลความหมายเช่นเดียวกับค่าสหสัมพันธ์ แต่เปลี่ยนเป็นการพยากรณ์ หรือเป็นการทำนายค่าตัวแปรตาม ส่วนการแปลผลการถดถอยพหุคูณ ซึ่งจะต้องแปลงเป็น ค่า R-Square หรือ R^2 [9] ซึ่งหมายถึง ความสามารถของตัวแปรอิสระในการทำนายตัวแปรตาม

3) ผลการจำลอง

การสร้างสมการรูปแบบการพยากรณ์ได้ถูกพัฒนาขึ้นบนโปรแกรมไพทอน (Python) โดยใช้คอมพิวเตอร์ที่มีหน่วยประมวลผลกลาง Intel® Core™ i5 Gen.11 2.40 GHz และมีหน่วยความจำขนาด 8 GB ซึ่งรวบรวมข้อมูลผ่านโปรแกรม Microsoft Excel

3.1) การเก็บรวบรวมข้อมูลทางอากาศจากสถานีตรวจวัดอากาศแบบอัตโนมัติบริเวณพื้นที่ทั่วไป

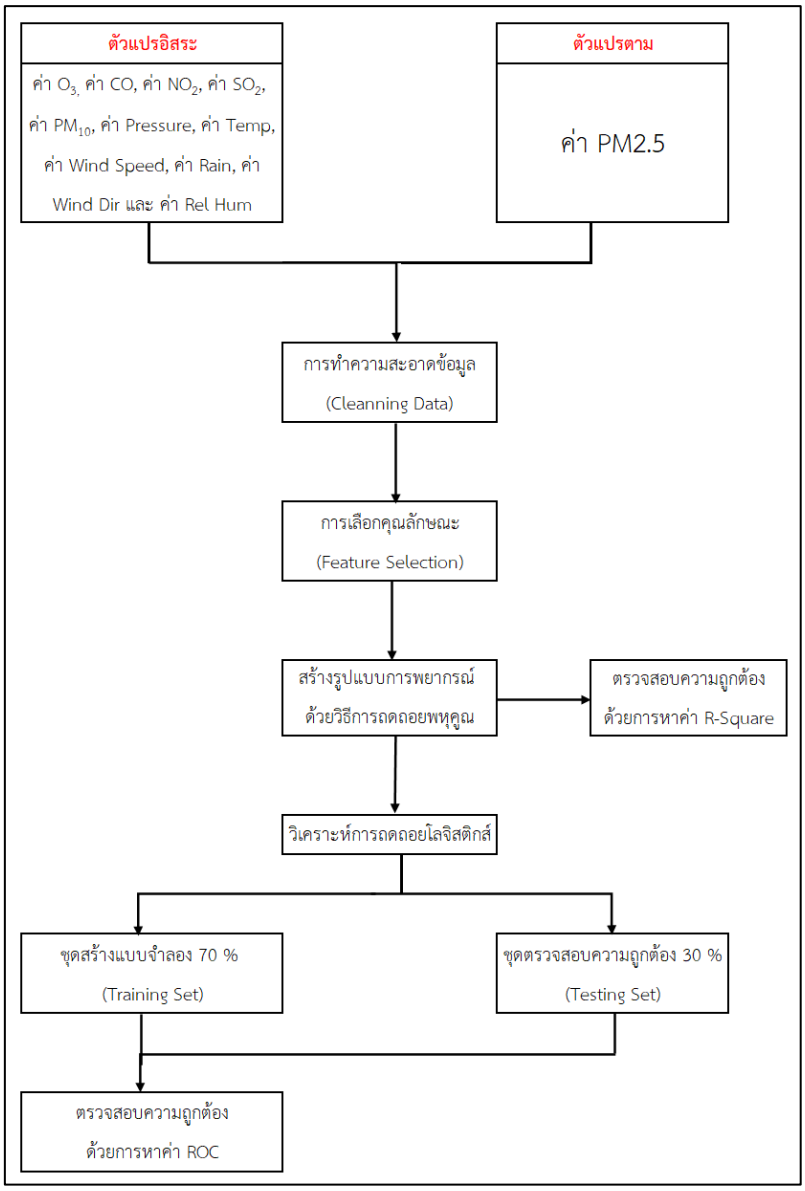
เป็นการข้อมูลย้อนหลังเป็นรายชั่วโมง ตั้งแต่วันที่ 1 มกราคม พ.ศ. 2558 จนถึง วันที่ 31 สิงหาคม พ.ศ. 2564 รวม 5 ปี จากสถานีตรวจวัดอากาศของจังหวัดพะเยา ทั้งสิ้น 12 ชุดข้อมูล โดยชุดข้อมูลที่ถูกวิเคราะห์ คือ ฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน มีหน่วยเป็น μg/m^3 แทนด้วย “PM2.5” โดยกำหนดให้เป็นตัวแปรตาม ส่วนการกำหนดตัวแปรอิสระ คือ ชุดข้อมูลที่เหลือทั้ง 11 ชุดข้อมูล อันได้แก่ 1) ก๊าซโอโซน มีหน่วยเป็น ppm แทนด้วย “O3” 2) ก๊าซคาร์บอนมอนอกไซด์ มีหน่วยเป็น ppm แทนด้วย “CO” 3) ก๊าซไนโตรเจนไดออกไซด์ มีหน่วยเป็น ppb แทนด้วย “NO2” 4) ก๊าซซัลเฟอร์ไดออกไซด์ มีหน่วยเป็น ppb แทนด้วย “SO2” 5) ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน มีหน่วยเป็น μg/m^3 แทนด้วย “PM10” 6) อุณหภูมิ มีหน่วยเป็น °C แทนด้วย “Temp” 7) ปริมาณน้ำฝน มีหน่วยเป็น mm แทนด้วย “Rain” 8) ความเร็วลม มีหน่วยเป็น km/h แทนด้วย “Wind Speed” 9) ทิศทางลม มีหน่วยเป็น องศา แทนด้วย “Wind Dir” 10) ความชื้นสัมพัทธ์ มีหน่วยเป็น % แทนด้วย “Rel Hum” และ 11) ความกดอากาศ มีหน่วยเป็น mb แทนด้วย “Pressure” โดยนำค่าตัวแปรต่าง ๆ มาสร้างแบบจำลองทางคณิตศาสตร์ ด้วยสมการ (3) เป็นสมการรูปแบบการพยากรณ์ โดยนำข้อมูลจากโปรแกรม Microsoft Excel ไปทำการแปลงด้วยโปรแกรมไพทอน (Python) ซึ่งสามารถแสดงผล ดังรูปที่ 3

3.2) การสร้างสมการรูปแบบการพยากรณ์

เป็นการนำชุดข้อมูลที่แปลได้จากโปรแกรมไพทอน มาสร้างแบบจำลองการพยากรณ์ (Prediction Model) โดยมีขั้นตอนและกระบวนการ ดังรูปที่ 4

	Year_Month_Day	Hour	CO	NO2	SO2	O3	PM10	Wind speed	Wind dir	Temp	Rel hum	Pressure	Rain	PM2.5
0	150101	100	1.50	12.6	1.0	11.0	36.0	0.1	106.0	20.9	81.4	725.9	0.0	NaN
1	150101	200	1.40	12.7	1.0	11.0	50.0	0.2	68.0	20.1	81.5	725.8	0.0	NaN
2	150101	300	0.60	10.3	1.0	11.0	46.0	0.1	79.0	19.1	86.0	725.7	0.0	NaN
3	150101	400	0.50	9.1	1.0	9.0	42.0	0.2	360.0	18.2	92.9	725.7	0.0	NaN
4	150101	500	0.50	7.2	0.0	9.0	41.0	0.2	314.0	18.0	97.6	725.9	0.0	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
56251	210831	2000	0.49	5.0	0.0	8.0	16.0	0.5	136.0	25.3	84.0	721.0	0.0	11.0
56252	210831	2100	0.57	4.0	0.0	4.0	20.0	0.2	90.0	24.6	85.0	721.0	0.0	10.0
56253	210831	2200	0.52	4.0	0.0	4.0	19.0	1.2	65.0	24.5	86.0	721.0	0.0	7.0
56254	210831	2300	0.40	2.0	0.0	5.0	17.0	1.3	67.0	24.4	86.0	722.0	0.0	9.0
56255	210831	2400	0.44	2.0	0.0	3.0	18.0	0.8	69.0	24.2	87.0	722.0	0.0	9.0

รูปที่ 3 : การแปลงข้อมูลที่ได้จากการเก็บรวบรวมด้วยโปรแกรมไพทอน (Python)



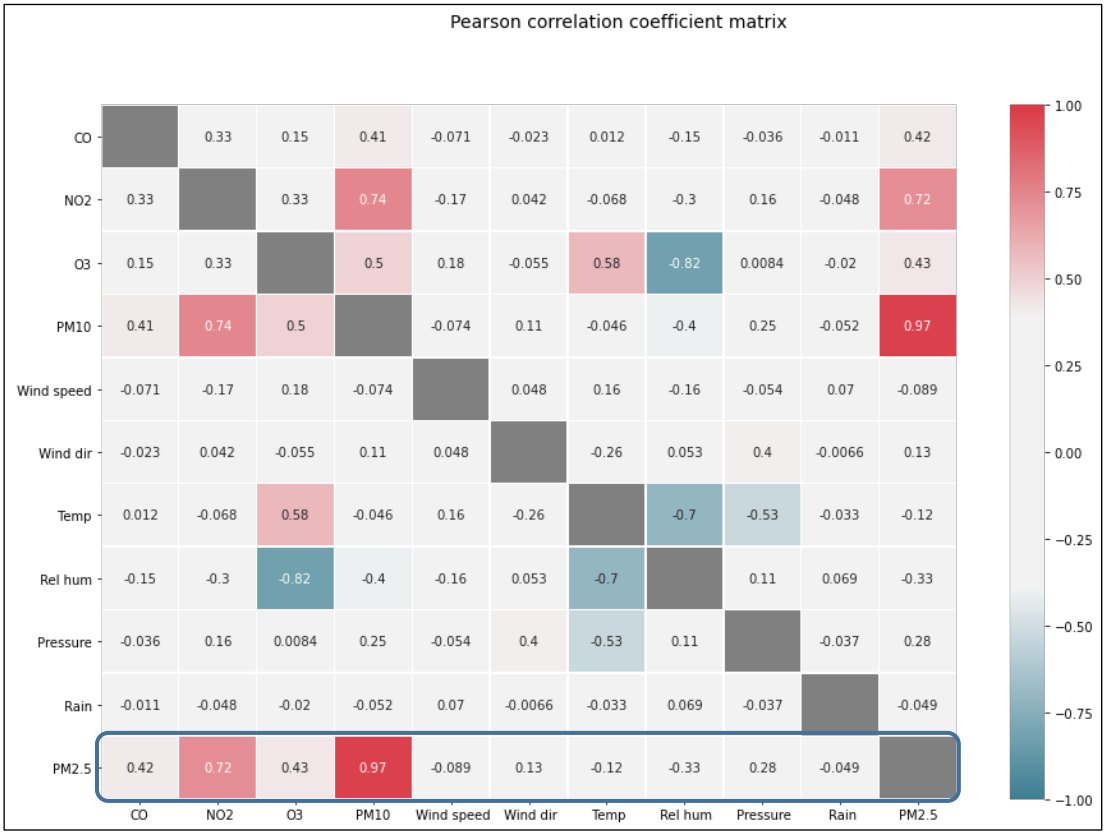
รูปที่ 4 : ขั้นตอนและกระบวนการสร้างแบบจำลองการพยากรณ์

3.2.1) การทำความสะอาดข้อมูล (Cleaning Data) เป็นขั้นตอนกระบวนการตรวจสอบ และแก้ไข (หรือลบ) รายการชุดข้อมูลที่ไม่ถูกต้องหรือไม่สมบูรณ์ ออกไปจากชุดข้อมูลในการกำหนดตัวแปรอิสระเพื่อนำสร้างแบบจำลองการพยากรณ์ เนื่องจาก ชุดข้อมูลอาจจะมีลักษณะของข้อมูลที่ไม่สมบูรณ์ (incomplete) มีข้อมูลรบกวน (noise) หรือมีข้อมูลที่ไม่สอดคล้อง (inconsistent) จึงต้องมีการประมวลผลชุดข้อมูลก่อน (pre-processing) ซึ่งใช้วิธีการทำความสะอาด แต่เนื่องจากข้อมูลไม่ครบถ้วนสมบูรณ์ ข้อมูลที่ขาดหายไปอย่างต่อเนื่องติดต่อกันนานจะถูกตัดทิ้ง และทดแทนค่าสูญหายกรณีอื่นด้วยค่าเฉลี่ย (Mean) ซึ่งพบว่าข้อมูลสถานีตรวจวัดอากาศสนามกีฬา จ.พะเยา ไม่มีการตรวจจับค่าก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) ในช่วงวันที่ 5 กรกฎาคม พ.ศ. 2562 ถึง วันที่ 31 สิงหาคม พ.ศ. 2564 มีค่าที่ขาดหายไปกว่า 25% จึงมีความจำเป็นต้องนำชุดข้อมูลนี้ ออกจากการสร้างแบบจำลองการพยากรณ์ [11] ดังรูปที่ 5

3.2.2) การเลือกคุณลักษณะ (Feature Selection) เป็นวิธีการลดมิติ โดยจะเลือกคุณลักษณะ หรือข้อมูลที่เหมาะสม เนื่องจาก ข้อมูลที่นำมาสร้างรูปแบบการพยากรณ์มีจำนวนมากและมีลักษณะข้อมูลที่กระจัดกระจาย และบางข้อมูลไม่มีประโยชน์ในการวิเคราะห์ ซึ่งจะส่งผลกระทบต่อความถูกต้องของการประมวลผล จึงต้องมีกระบวนการเลือกคุณลักษณะ ซึ่งในกรณีนี้ใช้วิธีการฟิลเตอร์ (filter) ข้อดีของวิธีเลือกคุณลักษณะแบบฟิลเตอร์นี้ คือ เป็นกระบวนการประเมินประสิทธิภาพของคุณลักษณะของข้อมูลแต่ละตัวว่ามีความเหมาะสมกับการวิเคราะห์ข้อมูลมากน้อยเพียงใด โดยไม่ขึ้นกับขั้นตอนวิธีของการเรียนรู้แบบใดแบบหนึ่ง การเลือกคุณลักษณะแบบนี้ จะทำการจัดลำดับ (Ranking) ตามความสำคัญของคุณลักษณะแต่ละตัว และเลือกคุณลักษณะที่มีระดับความสำคัญสูงสุดตามจำนวนที่ผู้ใช้ระบุ ซึ่งเป็นเทคนิคที่คำนวณได้ง่าย รวดเร็ว และหลีกเลี่ยงการเกิด Overfitting [12] ตัวแปรที่ได้ทั้งหมดที่นำมาทำการวิเคราะห์สหสัมพันธ์แบบเพียร์สัน โดยจะคัดเลือกจากตัวแปรที่มีความสัมพันธ์กันมากกว่า 0.40 ซึ่งได้รับการทดสอบนัยสำคัญของสหสัมพันธ์ มีแนวโน้มควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ให้อยู่ภายใต้ระดับนัยสำคัญที่ .05 [13] ซึ่งขนาดของตัวแปรที่มากขึ้นส่งผลให้การวิจัยและการวิเคราะห์ข้อมูลมีความคลาดเคลื่อนลดลงด้วย [14] และได้ตารางเมทริกซ์ค่าสหสัมพันธ์ ดังรูปที่ 6

Date	CO	NO2	O3	PM10	Wind speed	Wind dir	Temp	Rel hum	Pressure	Rain	PM2.5
2019-07-05	0.37	5.0	20.0	23.0	1.3	134.0	26.1	82.0	717.0	0.0	9.0
2019-07-05	0.35	5.0	20.0	20.0	1.4	136.0	25.6	81.0	717.0	0.0	8.0
2019-07-05	0.38	5.0	19.0	10.0	1.4	136.0	25.2	81.0	717.0	0.0	7.0
2019-07-05	0.41	6.0	18.0	9.0	1.5	136.0	25.0	81.0	717.0	0.0	6.0
2019-07-05	0.43	7.0	15.0	10.0	1.4	135.0	24.8	82.0	717.0	0.0	7.0
...	...	...	...	...	...	...	...	...	...	...	...
2021-08-31	0.49	5.0	8.0	16.0	0.5	136.0	25.3	84.0	721.0	0.0	11.0
2021-08-31	0.57	4.0	4.0	20.0	0.2	90.0	24.6	85.0	721.0	0.0	10.0
2021-08-31	0.52	4.0	4.0	19.0	1.2	65.0	24.5	86.0	721.0	0.0	7.0
2021-08-31	0.40	2.0	5.0	17.0	1.3	67.0	24.4	86.0	722.0	0.0	9.0
2021-08-31	0.44	2.0	3.0	18.0	0.8	69.0	24.2	87.0	722.0	0.0	9.0

รูปที่ 5 : การแปลผลหลังจากการทำความสะอาดข้อมูล



รูปที่ 6 : ตารางเมทริกซ์สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

3.2.3) การสร้างสมการพยากรณ์ด้วยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยการคัดเลือกตัวแปรที่มีความสัมพันธ์กับค่า $PM_{2.5}$ ที่มีความเหมาะสมที่ได้จากการเลือกคุณลักษณะ เพื่อกำหนดเป็นตัวแปรอิสระ โดยในการวิเคราะห์เพื่อสร้างสมการจำลองทางคณิตศาสตร์ สำหรับการพยากรณ์ค่า $PM_{2.5}$ อันได้แก่ ค่า PM_{10}, NO_2, O_3 และ CO ตามรูปแบบสมการ (3) ได้เป็นสมการพยากรณ์ของแบบจำลอง (6) ดังนี้

$$Z = -3.412+1.1819X_1-0.1392X_2-0.1149X_3+0.8260X_4 \quad (6)$$

โดยที่ Z คือ ฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน ($PM_{2.5}$)

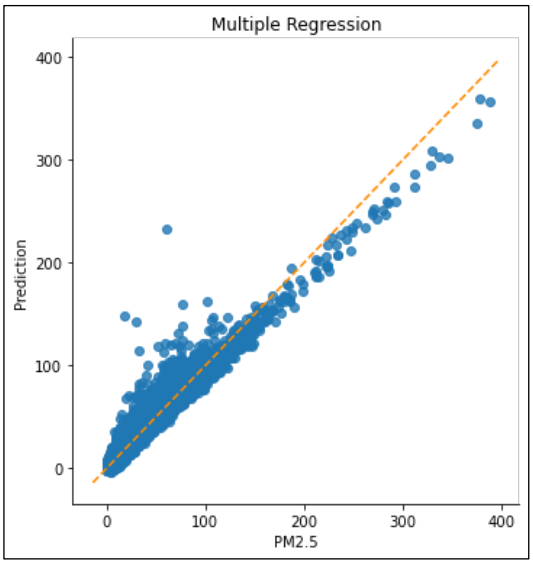
X_1 คือ ก๊าซคาร์บอนมอนอกไซด์ (CO)

X_2 คือ ก๊าซไนโตรเจนไดออกไซด์ (NO_2)

X_3 คือ ก๊าซโอโซน (O_3)

X_4 คือ ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน (PM_{10})

โดยเราสามารถนำสมการถดถอยพหุคูณที่ได้มาสร้างกราฟความสัมพันธ์เชิงเส้น เพื่อวิเคราะห์ความสัมพันธ์ระหว่าง ค่าฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน ($PM_{2.5}$) กับ ค่าการพยากรณ์หรือตัวแปรอิสระ ดังรูปที่ 7



รูปที่ 7 : กราฟความสัมพันธ์เชิงเส้นพหุคูณ

เมื่อนำค่าจากสมการ (6) แปลผลด้วยวิธีการหาค่า R-Square หรือ R^2 มีค่าเท่ากับ 0.9557 ซึ่งหมายความว่า ตัวแปรอิสระทั้ง 4 ชุดข้อมูล อันได้แก่ ค่า PM_{10}, NO_2, O_3 และ CO ร่วมกันทำนายค่าตัวแปรตาม หรือค่าฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน ($PM_{2.5}$) ได้ร้อยละ 95.57

3.2.4) การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression Analysis) เพื่อสร้างรูปแบบการเรียนรู้การจำลองการถดถอยเพื่อการพยากรณ์ โดยการแบ่งข้อมูลออกเป็น 2 ชุดข้อมูล คือ ชุดการสร้างแบบจำลอง (Training Set) โดยวิธีการวิเคราะห์การถดถอยโลจิสติก ซึ่งจะใช้ข้อมูล 70% จากชุดข้อมูลทั้งหมด ในการสร้างรูปแบบสมการพยากรณ์ จากนั้นนำรูปแบบสมการพยากรณ์ที่สร้างได้มาทดสอบด้วยการทำนายในชุดข้อมูลที่เหลืออีก 30% ซึ่งจะใช้สำหรับทำการตรวจสอบความถูกต้อง (Testing Set) และการวิเคราะห์จำแนกประเภท (Discriminant Analysis) คือ การศึกษาความสัมพันธ์ระหว่างตัวแปรตาม และตัวแปรอิสระ โดยกำหนดค่า $PM_{2.5}$ เกินกว่า 50 μg/m^3 ให้เป็นค่าเกินมาตรฐาน หรือ จริง (Yes) หากต่ำกว่า 50 μg/m^3 ถือว่าเป็นค่าที่ไม่เกินมาตรฐาน หรือ ไม่จริง (No) ซึ่งจะแทนด้วยค่าเป็น 1 กับ 0 ตามลำดับ แล้วทำการสร้างแบบจำลองการพยากรณ์ด้วยสมการถดถอยโลจิสติก (4) ได้ดังนี้

$$Prob\ (Y) = \frac{1}{1 + e^{-(8.6581+0.4365X_1 -0.706X_2 -0.0212X_3 +0.0598X_4)}} \quad (7)$$

โดยที่ *Prob (Y)* คือ ฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน ($PM_{2.5}$) เกินมาตรฐาน หรือไม่เกินมาตรฐานที่ 50 ไมโครกรัม

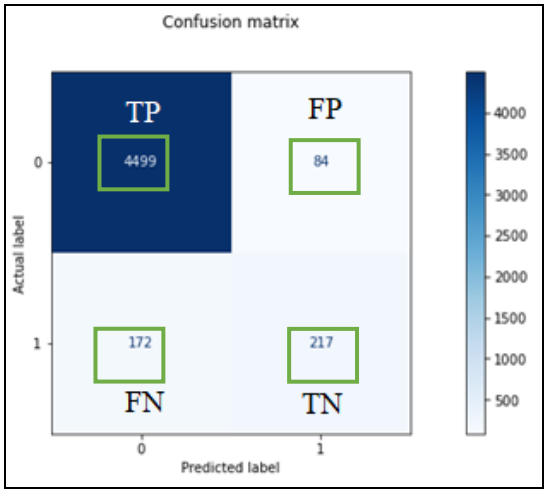
X_1 คือ ก๊าซคาร์บอนมอนอกไซด์ (CO)

X_2 คือ ก๊าซไนโตรเจนไดออกไซด์ (NO_2)

X_3 คือ ก๊าซโอโซน (O_3)

X_4 คือ ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน (PM_{10})

3.2.5) การตรวจสอบความถูกต้อง (Accuracy) โดยสร้างตารางสำคัญในการวัดความสามารถ (Confusion Matrix) จากสมการ (7) ได้ดังรูปที่ 8



รูปที่ 8 : ตารางสำคัญในการวัดความสามารถ (Confusion Matrix)

จากนั้น นำค่าที่ได้จากตารางสำคัญในการวัดความสามารถ โดยหาค่าความถูกต้อง (Accuracy) ได้ตามสมการ (8)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

โดยที่ *TP* คือ ทำนายว่า “จริง” ตรงกับสิ่งที่เกิดขึ้น “จริง”

TN คือ ทำนายว่า “ไม่จริง” ตรงกับสิ่งที่เกิดขึ้น “ไม่จริง”

FP คือ ทำนายว่า “จริง” แต่สิ่งที่เกิดขึ้น “ไม่จริง”

FN คือ ทำนายว่า “ไม่จริง” แต่สิ่งที่เกิดขึ้น “จริง”

ค่าความไว (Sensitivity) คือ สัดส่วนของผลบวกที่เป็นจริงสำหรับการทดลองนั้น ๆ เช่น การเกิดมลพิษอากาศที่เป็นจริง (true positives) จึงช่วยในการแยกการทดลองที่เป็นผลลบปลอม (false negative) เพราะว่าการทดสอบยิ่งไวเท่าไหร่ โอกาสการได้ผลลบ (เช่นพบว่า ไม่เกิดมลพิษอากาศ) ที่ไม่เป็นจริง (เช่น ผลที่เกิดจริง คือ เกิดมลพิษอากาศ) ก็น้อยลงเท่านั้น สามารถคำนวณได้จากสมการ (9) ดังนี้

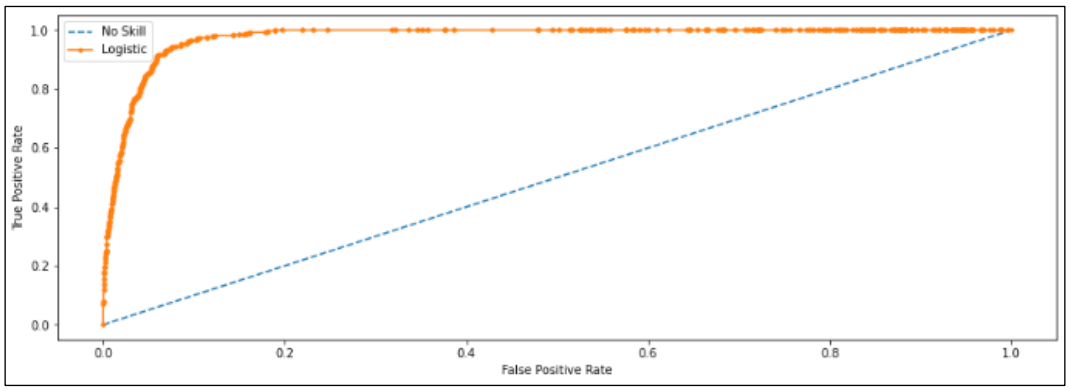
$$Sensitivity = \frac{TP}{TP+FN} \quad (9)$$

ค่าความจำเพาะ (Specificity) คือ สัดส่วนของผลลบที่เป็นจริงสำหรับการทดลองนั้น ๆ เช่น การไม่เกิดมลพิษอากาศ ซึ่งจะช่วยในการยืนยันสภาวะที่มีผลบวกปลอม (false positive) เพราะว่าการทดสอบยิ่งจำเพาะเท่าไหร่ โอกาสที่ได้ผลบวก (เช่นพบว่า เกิดมลพิษอากาศ) ที่ไม่เป็นจริง (ผลที่เกิดคือ ไม่เกิดมลพิษอากาศ) ก็น้อยลงเท่านั้น คำนวณได้จากสมการ (10) ดังนี้

$$Specificity = \frac{TN}{TN+FP} \quad (10)$$

เมื่อคำนวณค่าความไว (Sensitivity) และค่าความจำเพาะ (Specificity) จากสมการ (9) และ (10) ก็จะได้ผลการคำนวณอยู่ที่ 0.9632 และ 0.7209 ตามลำดับ หรือคิดเป็นร้อยละ 96.32 และ 72.09 โดยมีค่าความถูกต้อง (Accuracy) ซึ่งคำนวณได้จากสมการ (8) อยู่ที่ 0.9485 หรือร้อยละ 94.85

3.2.6) การทดสอบประสิทธิภาพของแบบจำลองการพยากรณ์ด้วยค่า Receiver Operating Characteristic Curve (ROC Curve) และค่าพื้นที่ใต้เส้นโค้ง หรือ Area under the curve (AUC) เป็นการหาสมการสหสัมพันธ์ระหว่างค่าจริงที่เกิดจากการวัด และค่าที่ได้จากการทำนายที่ระดับความเชื่อมั่นร้อยละ 95 เพื่อให้การทดสอบประสิทธิภาพการพยากรณ์ที่จะเกิดมลพิษทางอากาศหรือไม่เกิดมลพิษอากาศ ซึ่งเป็นวิธีการที่สามารถนำมาใช้ในการเลือกจุดตัดที่เหมาะสมได้ คือ การสร้างความสัมพันธ์ระหว่าง true positive (Sensitivity) กับ false positive rate (1-Specificity) ด้วยการแปรค่าจุดตัด (Cut-off Point)



รูปที่ 9 : กราฟ Receiver Operating Characteristic

โดยการเลือก Optimal cut-off point หรือจุดที่อยู่บนกราฟที่จะได้ค่า AUC มากที่สุด (ให้ผลดีที่สุด) คือ จุดที่อยู่ใกล้มุมซ้ายบนมากที่สุด ซึ่งจะเป็นจุดที่ Sensitivity สูง และ Specificity สูงด้วยเช่นกัน ดังนั้น หากค่า AUC ยิ่งเข้าใกล้ 1 ก็แสดงว่า แบบจำลองการพยากรณ์จะมีประสิทธิภาพมาก

เมื่อทำการทดสอบความถูกต้องจากค่าความไว (sensitivity) ที่ได้ คือ 0.9632 และค่าความจำเพาะ (specificity) ที่ได้ คือ 0.7209 มาคำนวณลงบนค่าพื้นที่ใต้เส้นโค้ง หรือค่า AUC ก็จะได้เท่ากับ 0.9757 หรือ 0.98 ดังรูปที่ 9

4) อภิปรายผลการวิจัย

การศึกษานี้เป็นการวิเคราะห์สหสัมพันธ์เพื่อคัดเลือกตัวแปรที่เหมาะสมในการสร้างสมการแบบจำลองการพยากรณ์ โดยใช้ข้อมูลสภาพอากาศย้อนหลัง 5 ปี พบว่า ตัวแปรที่มีสหสัมพันธ์กับความเข้มข้นของ $PM_{2.5}$ โดยเรียงจากตัวแปรที่มีความสัมพันธ์มากไปหาน้อย จากค่าสัมประสิทธิ์ของสมการถดถอย คือ ก๊าซคาร์บอนมอนอกไซด์ (CO), ฝุ่นละอองขนาดไม่เกิน 10 ไมครอน (PM_{10}), ก๊าซไนโตรเจนไดออกไซด์ (NO_2) และ ก๊าซโอโซน (O_3) ตามลำดับ ซึ่งพบว่า มีค่าความไว (sensitivity) อยู่ในระดับสูงที่ 96.32% แต่ค่าความจำเพาะ (specificity) มีระดับเพียง 72.09% เนื่องจาก ปัจจัยของก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) ถูกทำความสะอาด (cleaning) ออกไป ซึ่งค่าความเข้มข้นของ SO_2 เกิดมาจากการเผาไหม้เชื้อเพลิงฟอสซิลที่มีกำมะถันสูง เช่น น้ำมันเตาและถ่านหิน โดยจะค่าสูงขึ้นในเขตอุตสาหกรรม และโรงผลิตไฟฟ้า [6] สอดคล้องกับผลการวิจัยของ ศรีรักษ์ ศรีทองชัย [14] ซึ่งพบว่า ตัวแปร PM_{10}, NO_2, CO, O_3, และ SO_2 เป็นแบบจำลองที่ให้ความแม่นยำสูงสุดในการพยากรณ์ $PM_{2.5}$ และผลการศึกษาของ Zhai *et al.* [15] ก็ยังพบว่า ค่า SO_2, CO และ NO_2 มีความสัมพันธ์กับค่า $PM_{2.5}$ นอกจากนี้ ผลการศึกษาของ Lachatre *et al.* [16] ยังพบอีกว่า ค่า $PM_{2.5}$ มีความสัมพันธ์กับค่า SO_2, NO_x และ NH_3 อีกด้วย

5) สรุปผล

แบบจำลองการพยากรณ์นี้ มีประสิทธิภาพของแบบจำลองการพยากรณ์ด้วยค่า Receiver Operating Characteristic Curve (ROC Curve) และค่าพื้นที่ใต้เส้นโค้ง หรือ Area Under the Curve (AUC) อยู่ที่ 0.9757 หรือ 0.98 ซึ่งมีค่าอยู่ในระดับสูงแสดงว่า แบบจำลองสามารถพยากรณ์ได้ค่อนข้างแม่นยำ แต่ค่าความจำเพาะ (Specificity) ไม่ดีเท่าที่ควร

หากมีชุดข้อมูลตัวแปรอิสระของก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) ซึ่งถูกขจัดออกไปในขั้นตอนการทำความสะอาด (cleaning) โดยสามารถนำชุดข้อมูลนี้ มาสร้างแบบจำลองการพยากรณ์ ก็จะทำให้การพยากรณ์มีประสิทธิภาพที่สูงขึ้น และมีความแม่นยำเพิ่มมากขึ้น

6) ข้อเสนอแนะ

ปัญหาการวิจัยแบบจำลองการพยากรณ์นี้ พบว่า มีข้อมูลชุดขาดหายเป็นจำนวนมาก จึงมีข้อเสนอแนะให้ ขั้นตอนการเก็บข้อมูลต่าง ๆ ควรมีค่าที่ครบถ้วน ไม่ตกหล่น หรือสูญหาย ซึ่งจะทำให้ชุดข้อมูลถูกคัดกรองออกน้อย โดยเฉพาะชุดข้อมูลตัวแปรอิสระของก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) หรือการเก็บข้อมูลที่มีปัจจัยความสัมพันธ์เพิ่มเติม อาทิเช่น ค่าแอมโมเนีย (NH_3) ซึ่งจะทำให้สมการแบบจำลองการพยากรณ์จะมีความถูกต้องมากขึ้นและแม่นยำเพิ่มขึ้น ซึ่งจะช่วยพัฒนาอัลกอลิทึมให้ฉลาดเพิ่มมากขึ้นอีกด้วย

REFERENCES

[1] S. Pansripong, "Development of artificial neural network model and multiple regression models for prediction of PM10 concentrations in Bangkok and Nakhon Ratchasima municipality area," (in Thai), M.S. thesis, Dept. Environ. Eng., Suranaree Univ. Technol., Nakhon Ratchasima, Thailand, 2007.

[2] W. Simachaya. "Smog Pollution in the Northern: Problems and Solutions." (in Thai), PCD.go.th. https://www.pcd.go.th/wp-content/uploads/2020/06/pcdnew-2020-06-05_07-33-31_040770.pdf (accessed Feb. 1, 2023).

[3] S. Pimonsree and P. Arrin, "Situation of Particulate Matter in Ambient Air and Impacts on Healthy," (in Thai), in *Proc. Naresuan Res. Conf.: Sustain. Ways of Living Based on Sufficiency Economy,* Phitsanulok, Thailand, Jul. 2010, pp. 670–681.

[4] T. Rattanapong, "University of Phayao joins forces with RISC by MQDC to move forward with Fasai 2 to Phayao, expanding urban air purification towers for people in the region to further research towards understanding PM 2.5 dust in the northern region." (in Thai), UP.ac.th. https://www.up.ac.th/th/NewsRead.aspx?itemID=23918 (accessed Jul. 2, 2021).

[5] Greenpeace Thailand. "The different air quality measurement standards." (in Thai), GREENPEACE.org. https://www.greenpeace.org/thailandexplore/protect/cleanair/air-standard/ (accessed Nov. 23, 2022).

[6] K. Meksagul and A. Phantong, "The prediction model of climate in Phayao," (in Thai), in *Proc. Operations Res. Netw. (OR-net) Conf.*, Chiang Mai, Thailand, Feb. 2019, pp. 30–40.

[7] U. Amornnimit, "Logistic regression: Choice of risk analysis," (in Thai), *Univ. Thai Chamber of Commerce J.,* vol. 23, no. 2, pp. 21–35, 2003.

[8] V. Kovenko and V. Shevchuk. "Machine Learning with Python: Differences between classification and regression." READ THEDOCS.io. https://machine-learning-and-data-science-with-python.readthedocs.io/en/latest/assignment1_sup_ml_cls.html (accessed July. 2, 2021).

[9] K. Pearson, "Notes on the history of correlation," *Biometrika*, vol. 13, no. 1, pp. 25–45, 1920.

[10] B. Pommapun, "Techniques for data interpreting for using correlation and regression in research," (in Thai), *STOU Educ. J.*, vol. 11, no. 1, pp. 32-45, 2018.

[11] S. Srithongchai, "A PM2.5 prediction model using LSTM Neural Network in Bangkok area" (in Thai), *J. Eng. Digit. Technol. (JEDT)*, vol.10, no.1, pp. 1–9, 2022.

[12] R. Kladchuen and C. Sanrach, "An efficiency comparison of algorithms and feature selection methods to predict the learning achievement of vocational students," (in Thai), *Res. J. Rajamangala Univ. Technol. Thanyaburi*, vol. 17, no. 1, pp. 1–10, 2018.

[13] C. Pruekpramool, N. Jaroentaku, and S. Srisuttiyakorn, "Efficiency of Pearson, Spearman and Kendall's correlation coefficients when data is non-normal distributed," (in Thai), *An Online J. Educ. (OJED)*, vol. 15, no. 2, pp. 1–16, 2020.

[14] B. Srisa-ard, *Basic Research*, 10th ed. Bangkok, Thailand: SUWEERIVASARN (in Thai), 2017.

[15] S. Zhai *et al.*, "Fine particulate matter ($PM_{2.5}$) trends in China, 2013–2018: separating contributions from anthropogenic emissions and meteorology," *Atmos. Chem. Phys.*, vol. 19, no. 16, pp. 11031–11041, Aug. 2019, doi: 10.5194/acp-19-11031-2019.

[16] M. Lachatre *et al.*, "The unintended consequence of SO_2 and NO_2 regulations over China: Increase of ammonia levels and impaction $PM_{2.5}$ concentrations," *Atmos. Chem. Phys.*, vol. 19, no. 10, pp. 6701–6716, May 2019, doi: 10.5194/acp-19-6701-2019.

คำแนะนำสำหรับผู้เขียนบทความเพื่อลงตีพิมพ์

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล เป็นวารสารวิชาการทางด้านวิศวกรรมศาสตร์และเทคโนโลยี ของสถาบันเทคโนโลยีไทย-ญี่ปุ่น บทความที่นำเสนอจะต้องพิมพ์เป็นภาษาไทย หรือภาษาอังกฤษตามรูปแบบที่กำหนด และพร้อมที่จะนำไปตีพิมพ์ได้ทันที การเสนอบทความเพื่อพิจารณาตีพิมพ์ในวารสาร มีรายละเอียดดังนี้

1. หลักเกณฑ์การพิจารณาบทความเพื่อตีพิมพ์

1.1 เป็นบทความที่ไม่ได้อยู่ระหว่างการพิจารณาตีพิมพ์ หรือไม่ได้อยู่ระหว่างการพิจารณาของสื่อสิ่งพิมพ์อื่น ๆ และไม่เคยตีพิมพ์ในวารสารหรือรายงานการสืบเนื่องจากการประชุมวิชาการใดมาก่อนทั้งในประเทศและต่างประเทศ หากตรวจสอบพบว่ามีการตีพิมพ์ซ้ำซ้อน ถือเป็นความรับผิดชอบของผู้เขียนแต่เพียงผู้เดียว

1.2 เป็นบทความที่แสดงให้เห็นถึงความคิดริเริ่มสร้างสรรค์ มีคุณค่าทางวิชาการ มีความสมบูรณ์ของเนื้อหา และมีความถูกต้องตามหลักวิชาการ

1.3 บทความที่ได้รับการตีพิมพ์จะต้องผ่านการประเมินคุณภาพจากผู้ทรงคุณวุฒิ (Peer reviewer) อย่างน้อย 2 ท่านต่อบทความ ซึ่งผู้ทรงคุณวุฒิอาจให้ผู้เขียนแก้ไขเพิ่มเติมหรือปรับปรุงบทความให้เหมาะสมยิ่งขึ้น

1.4 กองบรรณาธิการขอสงวนสิทธิ์ในการตรวจแก้ไขรูปแบบบทความที่ส่งมาตีพิมพ์ตามที่เห็นสมควร

1.5 บทความ ข้อความ ภาพประกอบ และตารางประกอบ ที่ตีพิมพ์ลงวารสารเป็นความคิดเห็นส่วนตัวของผู้เขียน กองบรรณาธิการไม่มีส่วนรับผิดชอบใด ๆ ถือเป็นความรับผิดชอบของผู้เขียนแต่เพียงผู้เดียว

1.6 ต้องเป็นบทความที่ไม่ละเมิดลิขสิทธิ์ ไม่ลอกเลียน หรือตัดทอนข้อความของผู้อื่นโดยไม่ได้รับอนุญาต

1.7 หากเป็นงานแปลหรือเรียบเรียงจากภาษาต่างประเทศ ต้องมีหลักฐานการอนุญาตให้ตีพิมพ์เป็นลายลักษณ์อักษรจากเจ้าของลิขสิทธิ์

1.8 ต้องมีการอ้างอิงที่ถูกต้อง เหมาะสมก่อนการตีพิมพ์ ซึ่งเป็นความรับผิดชอบของเจ้าของผลงาน

1.9 บทความที่ส่งถึงกองบรรณาธิการ ขอสงวนสิทธิ์ที่จะไม่ส่งคืนผู้เขียน

2. รูปแบบการกลั่นกรองบทความก่อนลงตีพิมพ์ (Peer-review)

ในการประเมินบทความโดยผู้ทรงคุณวุฒิเป็นการประเมินแบบ Double-blind peer review คือ ผู้ทรงคุณวุฒิไม่ทราบชื่อและรายละเอียดของผู้เขียนบทความ และผู้เขียนบทความไม่ทราบชื่อและรายละเอียดของผู้ทรงคุณวุฒิ

3. ประเภทของบทความที่รับพิจารณาลงตีพิมพ์

นิพนธ์ต้นฉบับต้องเป็นบทความวิจัย ประกอบด้วย บทคัดย่อ บทนำ วัตถุประสงค์ของการวิจัย วิธีดำเนินการวิจัย ผลการวิจัยและอภิปรายผล สรุปผล และเอกสารอ้างอิง*

<u>หมายเหตุ</u> : บทความภาษาไทยต้องมีบทคัดย่อภาษาไทยและภาษาอังกฤษ โดยให้หน้าบทคัดย่อภาษาไทยอยู่ก่อนหน้าบทคัดย่อภาษาอังกฤษ สำหรับบทความภาษาอังกฤษไม่ต้องมีบทคัดย่อภาษาไทย

---

* เอกสารอ้างอิง เป็นการบอกรายการแหล่งอ้างอิงตามที่มีการอ้างอิงในเนื้อหาของงานเขียน

4. จริยธรรมในการตีพิมพ์ผลงานวิจัยในวารสารวิชาการ

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล ได้คำนึงถึงจริยธรรมในการตีพิมพ์บทความ โดยจริยธรรมและบทบาทหน้าที่ของผู้เกี่ยวข้อง มีดังนี้

จริยธรรมและหน้าที่ของผู้เขียน

1. ผู้เขียนต้องเขียนบทความให้เป็นไปตามรูปแบบที่วารสารกำหนดไว้ในคำแนะนำสำหรับผู้เขียน
2. หากมีการนำข้อมูลของผู้อื่นหรือข้อมูลของผู้เขียนที่เคยตีพิมพ์ในวารสารฉบับอื่นมาใช้ ผู้เขียนต้องอ้างอิงแหล่งที่มาของข้อมูลนั้น โดยไม่ทำให้ผู้อื่นเข้าใจผิดว่าข้อมูลนั้นเป็นผลงานใหม่ของผู้เขียน
3. ผู้เขียนต้องไม่ดัดแปลงหรือบิดเบือนข้อมูล
4. ผู้เขียนต้องระบุแหล่งทุนที่สนับสนุนในการทำวิจัย (ถ้ามี)
5. ผู้เขียนต้องเปิดเผยข้อมูลเกี่ยวกับผลประโยชน์ทับซ้อนอย่างชัดเจน (ถ้ามี)

จริยธรรมและหน้าที่ของบรรณาธิการ

1. บรรณาธิการต้องดำเนินการเผยแพร่วารสารให้ตรงตามเวลา
2. บรรณาธิการต้องคัดเลือกบทความ โดยพิจารณาจากคุณภาพและความสอดคล้องของเนื้อหาบทความกับขอบเขตของวารสาร
3. บรรณาธิการต้องดำเนินการประเมินบทความอย่างเป็นธรรม ไม่ปฏิเสธการตีพิมพ์บทความโดยใช้อคติ
4. บรรณาธิการต้องไม่เปิดเผยข้อมูลของผู้เขียนบทความ และผู้ประเมินบทความแก่บุคคลอื่นที่ไม่เกี่ยวข้อง
5. บรรณาธิการต้องไม่มีผลประโยชน์ทับซ้อนกับผู้เขียน และผู้ประเมินบทความ
6. บรรณาธิการต้องใช้โปรแกรมในการตรวจสอบการคัดลอกผลงาน เพื่อป้องกันการตีพิมพ์ผลงานซึ่งคัดลอกมาจากผลงานผู้อื่น หากพบการคัดลอกผลงานของผู้อื่น บรรณาธิการต้องหยุดกระบวนการพิจารณาบทความทันที และติดต่อผู้เขียนเพื่อขอคำชี้แจง

จริยธรรมและหน้าที่ของผู้ประเมินบทความ

1. ผู้ประเมินบทความ ควรพิจารณาตอบรับการประเมินเฉพาะบทความที่สอดคล้องกับความเชี่ยวชาญของตนเองเท่านั้น เพื่อให้บทความที่ตีพิมพ์มีคุณภาพ
2. ผู้ประเมินบทความควรประเมินบทความให้แล้วเสร็จตามระยะเวลาที่กำหนด เพื่อไม่ให้บทความที่ผ่านการพิจารณาตีพิมพ์ล่าช้า
3. ผู้ประเมินควรประเมินบทความโดยให้ข้อเสนอแนะตามหลักวิชาการเท่านั้น ไม่ควรใช้ความคิดเห็นส่วนตัวที่ไม่มีเหตุผลหรือไม่มีข้อมูลรองรับ
4. ผู้ประเมินควรปฏิเสธการประเมินบทความ หากเห็นว่าตนเองอาจมีผลประโยชน์ทับซ้อนกับผู้เขียน
5. ผู้ประเมินต้องไม่เปิดเผยเนื้อหาภายในบทความให้ผู้ที่ไม่เกี่ยวข้องทราบ
6. หากผู้ประเมินเห็นว่ามีบทความอื่นที่ผู้เขียนไม่ได้กล่าวอ้างถึง แต่เป็นบทความที่สำคัญและเกี่ยวข้องกับบทความ ผู้ประเมินควรแจ้งให้ผู้เขียนกล่าวอ้างถึงบทความนั้น

5. ข้อกำหนดการจัดพิมพ์ต้นฉบับบทความ

ผู้เขียนต้องจัดพิมพ์บทความตามข้อกำหนดเพื่อให้มีรูปแบบการตีพิมพ์เป็นมาตรฐานแบบเดียวกัน ดังนี้

5.1 ขนาดของกระดาษ ให้ใช้ขนาด A4

5.2 กรอบของข้อความ ระยะห่างของขอบกระดาษ

ด้านบน 2.5 ซม. (0.98”) ด้านล่าง 2 ซม. (0.79”)

ด้านซ้าย 2 ซม. (0.79”) ด้านขวา 2 ซม. (0.79”)

5.3 ระยะห่างระหว่างบรรทัด หนึ่งช่วงบรรทัดของเครื่องคอมพิวเตอร์ (Single)

5.4 ตัวอักษร รูปแบบของตัวอักษรให้ใช้ TH Sarabun New

5.5 รายละเอียดต่าง ๆ ของบทความ กำหนดดังนี้

ชื่อเรื่อง (Title) ขนาด 22 **ตัวหนา** กำหนดกึ่งกลาง

ชื่อผู้เขียน (Author) ขนาด 16 ตัวธรรมดา กำหนดกึ่งกลาง ไม่ต้องใส่คำนำหน้า

สังกัด (Affiliation) ขนาด 14 *ตัวเอน* กำหนดกึ่งกลาง

E-mail ผู้ประพันธ์บรรณกิจ (Corresponding Author) ขนาด 12 ตัวธรรมดา กำหนดกึ่งกลาง

บทคัดย่อ (Abstract) จัดรูปแบบการพิมพ์เป็นแบบ **1 คอลัมน์** ชื่อหัวข้อ ขนาด 14 ***ตัวหนาและเอน*** กำหนดกึ่งกลาง ข้อความในบทคัดย่อ ขนาด 14 ตัวธรรมดา

คำสำคัญ (Keywords) ให้ใส่คำสำคัญ 3–5 คำ ซึ่งเกี่ยวข้องกับบทความที่นำเสนอ โดยให้จัดพิมพ์ใต้บทคัดย่อ เว้น 1 บรรทัดด้วยขนาด 12 คำสำคัญขนาด 14 ***ตัวหนาและเอน*** กำหนดชิดซ้าย ข้อความในคำสำคัญ ขนาด 14 ตัวธรรมดา

รูปแบบการพิมพ์เนื้อหาของบทความ

- **รูปแบบการพิมพ์เป็นแบบ 2 คอลัมน์** แต่ละคอลัมน์ กว้าง 8.2 ซม. (3.23”) ระยะห่างระหว่างคอลัมน์ 0.61 ซม. (0.24”)
- **หัวข้อหลัก** ประกอบด้วย บทนำ (INTRODUCTION) วัตถุประสงค์ของการวิจัย (OBJECTIVES) วิธีดำเนินการวิจัย (METHODS) ผลการวิจัยและอภิปรายผล (RESULTS AND DISCUSSION) สรุปผล (CONCLUSIONS) เอกสารอ้างอิง (REFERENCES) ขนาด 14 ตัวธรรมดา กำหนดกึ่งกลาง และมีเลขกำกับ เช่น 1) บทนำ หรือ I. INTRODUCTION เป็นต้น
- **หัวข้อรอง ระดับที่ 1** ขนาด 14 *ตัวเอน* กำหนดชิดซ้าย เมื่อขึ้นหัวข้อระดับเดียวกันหรือมากกว่าในลำดับถัดไป ให้เคาะเว้น 1 บรรทัดหลังจบเนื้อหา
- **หัวข้อรอง ระดับที่ 2** ขนาด 14 *ตัวเอน* กำหนดชิดซ้ายและเลื่อนเข้ามา 0.5 ซม.
- **เนื้อเรื่อง** ขนาด 14 ตัวธรรมดา
- **ชื่อตาราง** ขนาด 12 ตัวธรรมดา กำหนดกึ่งกลาง และใส่ชื่อเหนือตาราง
- **หัวข้อในตาราง** ขนาด 12 **ตัวหนา** กำหนดกึ่งกลาง เนื้อหาในตาราง ขนาด 12 ตัวธรรมดา
- **ชื่อภาพประกอบ** ขนาด 12 ตัวธรรมดา กำหนดกึ่งกลาง และใส่ชื่อใต้ภาพ
- **เนื้อหาในภาพประกอบ** ขนาด 12 ตัวธรรมดา

5.6 เอกสารอ้างอิง

1. การอ้างอิงในเนื้อหาบทความใช้การอ้างอิงแบบตัวเลข ตามมาตรฐานสากล โดยใช้หมายเลขในเครื่องหมายก้ามปู [ ] และเรียงลำดับการอ้างอิงตามเนื้อหา โดยมีตัวอย่างการเขียน เช่น [1] หรือ [2] หรือ [1], [2] หรือ [1], [3]–[8] หรือ [9], [10], [15], [16] หากมีการอ้างอิงซ้ำบทความเดิมให้ใช้หมายเลขเดิม ตัวอย่างเช่น by Brown [4], [5]; as mentioned earlier [2], [4]–[7], [9]; Smith [4] and Brown and Jones [5]; Wood *et al.* [7]

2. รูปแบบของชื่อหัวข้อใช้รูปแบบตัวอักษร TH Sarabun New ขนาด 14 ตัวธรรมดา ในเนื้อหาขนาด 12 ตัวธรรมดา

3. การอ้างอิงท้ายบทความ จะต้องเรียงตามลำดับบทความที่เขียนอ้างอิงในเนื้อเรื่อง และใช้การอ้างอิงตามรูปแบบการอ้างอิง IEEE ซึ่งผู้เขียนสามารถศึกษาวิธีการเขียนเอกสารอ้างอิงตามรูปแบบที่กำหนดได้ที่เว็บไซต์ https://ieeeauthorcenter.ieee.org/wp-content/uploads/IEEE-Reference-Guide.pdf โดยจะต้องเขียนเป็นภาษาอังกฤษเท่านั้น หากบทความอ้างอิงมาจากบทความภาษาไทย ต้องแปลเป็นภาษาอังกฤษให้ถูกต้อง

4. กรณีที่เอกสารที่นำมาอ้างอิงเขียนเป็นภาษาไทยให้เติมคำว่า “(in Thai)” เข้าไปในเอกสารอ้างอิง ดังเช่นตัวอย่างต่อไปนี้

ตัวอย่างรูปแบบการเขียนและการแปลเอกสารอ้างอิงภาษาไทยเป็นภาษาอังกฤษ

<u>ตัวอย่างที่ 1 การอ้างอิงจากหนังสือ(Books)</u>

Basic Format:

[Number] J. K. Author, *Title of His Published Book*, xth ed. City of Publisher, State (only U.S.), Country: Abbrev. of Publisher, year.

[Number] J. K. Author, “Title of chapter in the book,” in *Title of His Published Book*, xth ed. City of Publisher, State (only U.S.), Country: Abbrev. of Publisher, year, ch. x, sec. x, pp. xxx–xxx.

Examples:

[1] V. Rijiravanich, *Work Study: Principles and Case Studies*, 4th ed. Bangkok, Thailand: Chulalongkorn University Press (in Thai), 2005.
วันชัย ริจิรวนิช, *การศึกษาการทำงาน: หลักการและกรณีศึกษา*, พิมพ์ครั้งที่ 4, กรุงเทพฯ, ประเทศไทย: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2548.

[2] L. Edelstein-Keshet, *Mathematical Models in Biology.* New York, NY, USA: Random House, 1998.

[3] K. J. Roodbergen, “Storage assignment for order picking in multiple-block warehouses,” in *Warehousing in the Global Supply Chain: Advanced Models, Tools and Applications for Storage Systems.* London, U.K.: Springer, 2012, pp. 139–155.

ตัวอย่างที่ 2 การอ้างอิงจากวารสาร (Periodicals)

Basic Format:

[Number] J. K. Author, "Name of paper," *Abbrev. Title of Periodical*, vol. x, no. x, pp. xxx–xxx, Abbrev. Month, year.

[Number] J. K. Author, "Name of paper," *Abbrev. Title of Periodical*, vol. x, no. x, pp. xxx-xxx, Abbrev. Month, year, doi: xxx.

Examples:

[4] N. Dechampai and K. Sethanan, "An application of lean manufacturing system in the textile of lean manufacturing system in the textile and garment industry case study: Wacoal Kabinburi Co., Ltd," (in Thai), *MBA-KKU J.*, vol. 7, no. 2, pp. 13–27, 2014.
นิวัฒน์ เดชอำไพ และกาญจนา เศรษฐนันท์, "การเพิ่มประสิทธิภาพกระบวนการผลิตชุดชั้นในสตรีโดยประยุกต์ใช้แนวคิดการผลิตแบบลีน," *วารสารวิทยาลัยบัณฑิตศึกษาการจัดการมหาวิทยาลัยขอนแก่น*, ปีที่ 7, ฉบับที่ 2, หน้า 13–27, 2557.

[5] S.-F. Wang, H.-P. Chen, Y. Ku, and M.-X. Zhong, "Voltage-mode multifunction biquad filter and its application as fully-uncoupled quadrature oscillator based on current-feedback operational amplifiers," *Sensors*, vol. 20, no. 22, p. 6681, Nov. 2020, doi: 10.3390/s20226681.

[6] M. Dwiyaniti *et al.*, "Extremely high surface area of activated carbon originated from sugarcane bagasse," in *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 909, no. 1, 2020, pp. 1–8.

[7] M. Yoshiki, Y. Fujita, A. Kawamura, and H. Arai, "Instability of plates with holes (1st report)," (in Japanese), *J. Zosen Kiokai*, vol. 1967, no. 122, pp. 137–145, 1967.

[8] F. de Oliveira Santini, "Clockwise versus counterclockwise turning bias: Moderation effects of foot traffic and cognitive experience on visual attention," *J. Retail. Consum. Serv.*, vol. 67, 2022, Art. no. 102965.

ตัวอย่างที่ 3 การอ้างอิงจากวิทยานิพนธ์ (Theses and Dissertations)

Basic Format:

[Number] J. K. Author, "Title of thesis," M.S. thesis / Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State (only U.S.), year.

Examples:

[9] N. Kawasaki, "Parametric study of thermal and chemical nonequilibrium nozzle flow," M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.

[10] Y. Sornsa and P. Wanrerk, "Buckling of rectangular plates with a central square hole," (in Thai), B.S. thesis, Dept. Mech. Eng., Burapha Univ., Chonburi, Thailand, 2007.

<u>ตัวอย่างที่ 4 การอ้างอิงจากการประชุมทางวิชาการ (Conferences and Conference Proceedings)</u>

Basic Format:

[Number] J. K. Author, "Title of paper," in *Abbrev. Name of Conf.,* City, State (only U.S.), Country, Month year, pp. xxx–xxx.

Examples:

[11] N. Kriengkorakot, P. Kriengkorakot, S. Duan, P. Thung, and W. Piromsuk, "Repair work reduction in sewing process of the apparel factory," (in Thai), in *Proc. 10th Ubon Ratchathani Univ. Nat. Res. Conf.*, Ubon Ratchathani, Thailand, Jul. 2016, pp. 87–96.

นุชสรา เกรียงกรกฏ, ปรีชา เกรียงกรกฏ, สกาวเดือน พรมทุ่ง และวิจิตรา ภิรมย์สุข, "งานวิจัยการลดชิ้นส่วนงานซ่อมในขั้นตอนการเย็บของโรงงานผลิตเสื้อผ้าสำเร็จรูป," *การประชุมวิชาการระดับชาติ มอบ. วิจัยครั้งที่ 10*, อุบลราชธานี, ประเทศไทย, 7-8 กรกฎาคม, 2559, หน้า 87–96.

[12] J. Lingad, S. Karimi, and J. Yin, "Location extraction from disaster-related microblogs," in *Proc. 22nd Int. Conf. World Wide Web*, Rio de Janeiro, Brazil, May 2013, pp. 1017–1020.

<u>ตัวอย่างที่ 5 การอ้างอิงเว็บไซต์</u>

[13] B. Suttamanatwong. "Minimum Wage Rate 2022." (in Thai), MOL.go.th. https://www.mol.go.th/อัตราค่าจ้างขั้นต่ำ (accessed Dec. 31, 2022).

[14] Greenpeace Thailand. "The different air quality measurement standards." (in Thai), GREENPEACE.org. https://www.greenpeace.org/thailandexplore/protect/cleanair/air-standard/ (accessed Nov. 23, 2022).

6. วิธีการจัดส่งบทความ

ผู้เขียนส่งบทความออนไลน์ได้ที่ https://ph01.tci-thaijo.org/index.php/TNIJournal

ทุกบทความที่ส่งเข้ามาจะมีการตรวจพิจารณาเบื้องต้นโดยกองบรรณาธิการก่อน และเมื่อบทความผ่านการพิจารณาเบื้องต้นผู้เขียนจึงจะสามารถชำระเงินได้ หลังจากที่ผู้เขียนชำระเงินเรียบร้อย ทางกองบรรณาธิการจะดำเนินการส่งบทความเสนอผู้ทรงคุณวุฒิพิจารณาบทความและแจ้งผลการพิจารณาให้ผู้เขียนบทความทราบ สำหรับบทความที่ผ่านการประเมินโดยผู้ทรงคุณวุฒิแล้วจะได้รับการตีพิมพ์ลงในวารสารเพื่อเผยแพร่ต่อไป

วารสารวิศวกรรมศาสตร์และเทคโนโลยีดิจิทัล
ปีที่ 11 ฉบับที่ 2 กรกฎาคม – ธันวาคม 2566

"สร้างนักคิด ผลิตนักปฏิบัติ สร้างนักประดิษฐ์ ผลิตนักบริหาร"

คณะวิศวกรรมศาสตร์

หลักสูตรระดับปริญญาตรี

- หลักสูตรวิศวกรรมยานยนต์ (Automotive Engineering, B.Eng. : AE)
- หลักสูตรวิศวกรรมหุ่นยนต์และระบบอัตโนมัติแบบลีน (Robotics and Lean Automation Engineering, B.Eng. : RE)
- หลักสูตรวิศวกรรมคอมพิวเตอร์และปัญญาประดิษฐ์ (Computer Engineering and Artificial Intelligence, B.Eng. : CE)
- หลักสูตรวิศวกรรมอุตสาหการ (Industrial Engineering, B.Eng. : IE)
- หลักสูตรวิศวกรรมไฟฟ้า (Electrical Engineering, B.Eng. : EE)

หลักสูตรระดับปริญญาโท

- หลักสูตรเทคโนโลยีวิศวกรรม (Engineering Technology, M.Eng. : MET)

คณะเทคโนโลยีสารสนเทศ

หลักสูตรระดับปริญญาตรี

- หลักสูตรเทคโนโลยีสารสนเทศ (Information Technology, B.Sc. : IT)
- หลักสูตรเทคโนโลยีมัลติมีเดีย (Multimedia Technology, B.Sc. : MT)
- หลักสูตรเทคโนโลยีสารสนเทศทางธุรกิจ (Business Information Technology, B.Sc. : BI)
- หลักสูตรเทคโนโลยีดิจิทัลทางสื่อสารมวลชน (Digital Technology in Mass Communication, B.Sc. : DC)

หลักสูตรระดับปริญญาโท

- หลักสูตรเทคโนโลยีสารสนเทศ (Information Technology, M.Sc. : MIT)

คณะบริหารธุรกิจ

หลักสูตรระดับปริญญาตรี

- หลักสูตรการจัดการเทคโนโลยีและนวัตกรรม (Management of Technology and Innovation, B.B.A. : MI)
- หลักสูตรบริหารธุรกิจญี่ปุ่น (Japanese Business Administration, B.B.A. : BJ)
- หลักสูตรการจัดการธุรกิจระหว่างประเทศ (International Business Management, B.B.A. : IB)
- หลักสูตรการบัญชี (Accountancy, B.Acc. : AC)
- หลักสูตรการจัดการทรัพยากรมนุษย์แบบญี่ปุ่น (Japanese Human Resources Management, B.B.A. : HR)
- หลักสูตรการจัดการโลจิสติกส์และโซ่อุปทาน (Logistics and Supply Chain Management, B.B.A. : LM)
- หลักสูตรการตลาดดิจิทัล (Digital Marketing, B.B.A. : DM)
- หลักสูตรการจัดการการท่องเที่ยวและการบริการเชิงนวัตกรรม (Innovative Tourism and Hospitality Management, B.B.A. : TH)

หลักสูตรระดับปริญญาโท

- หลักสูตรบริหารธุรกิจญี่ปุ่น (Japanese Business Administration, M.B.A. : MBJ)
- หลักสูตรนวัตกรรมการจัดการธุรกิจและอุตสาหกรรม (Innovation of Business and Industrial Management, M.B.A. : MBI)
  - แขนงวิชา การวางแผนกลยุทธ์สำหรับผู้ประกอบการ (Strategic Planning and Management Entrepreneur : SPE)
  - แขนงวิชา การจัดการระบบการผลิต และโลจิสติกส์แบบลีน (Lean Manufacturing System and Logistics Management : LMS)

Thai-Nichi International College (TNIC)

Bachelor's Degree Programs

- Digital Engineering (DGE)
- Data Science and Analytics (DSA)
- Data Science and Analytics (DSA)
- Japanese for International Business (JIB)

Thai-Nichi Institute of Technology

1771/1 Pattanakarn Road, Suanluang, Bangkok 10250, Thailand

Tel: 0-2763-2600 Fax: 0-2763-2700

Website: https://ph01.tci-thaijo.org/index.php/TNIJournal E-mail : JEDT@tni.ac.th