



การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น

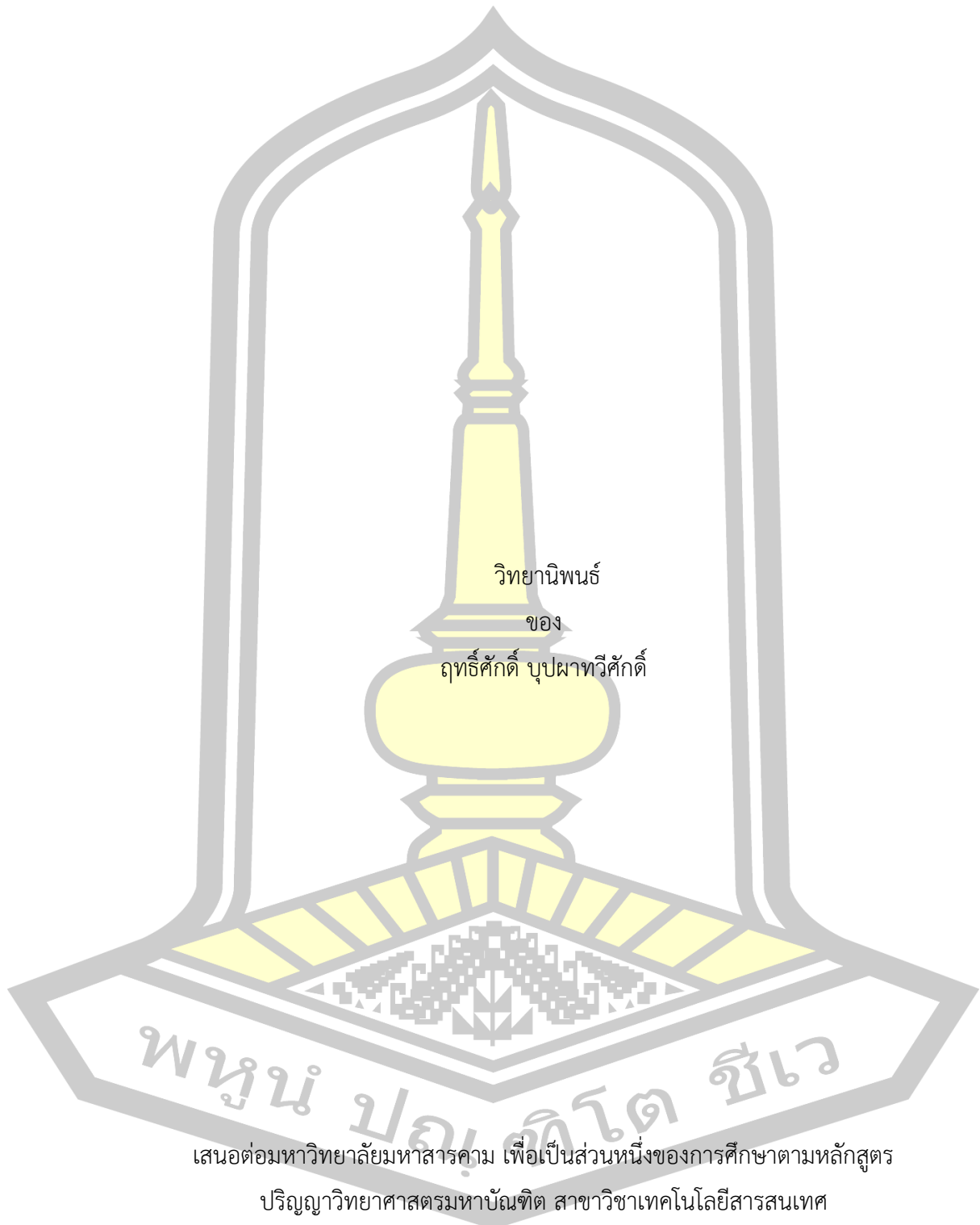
วิทยานิพนธ์
ของ
ฤทธิ์ศักดิ์ บุปผาทวีศักดิ์

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

กรกฎาคม 2567

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น



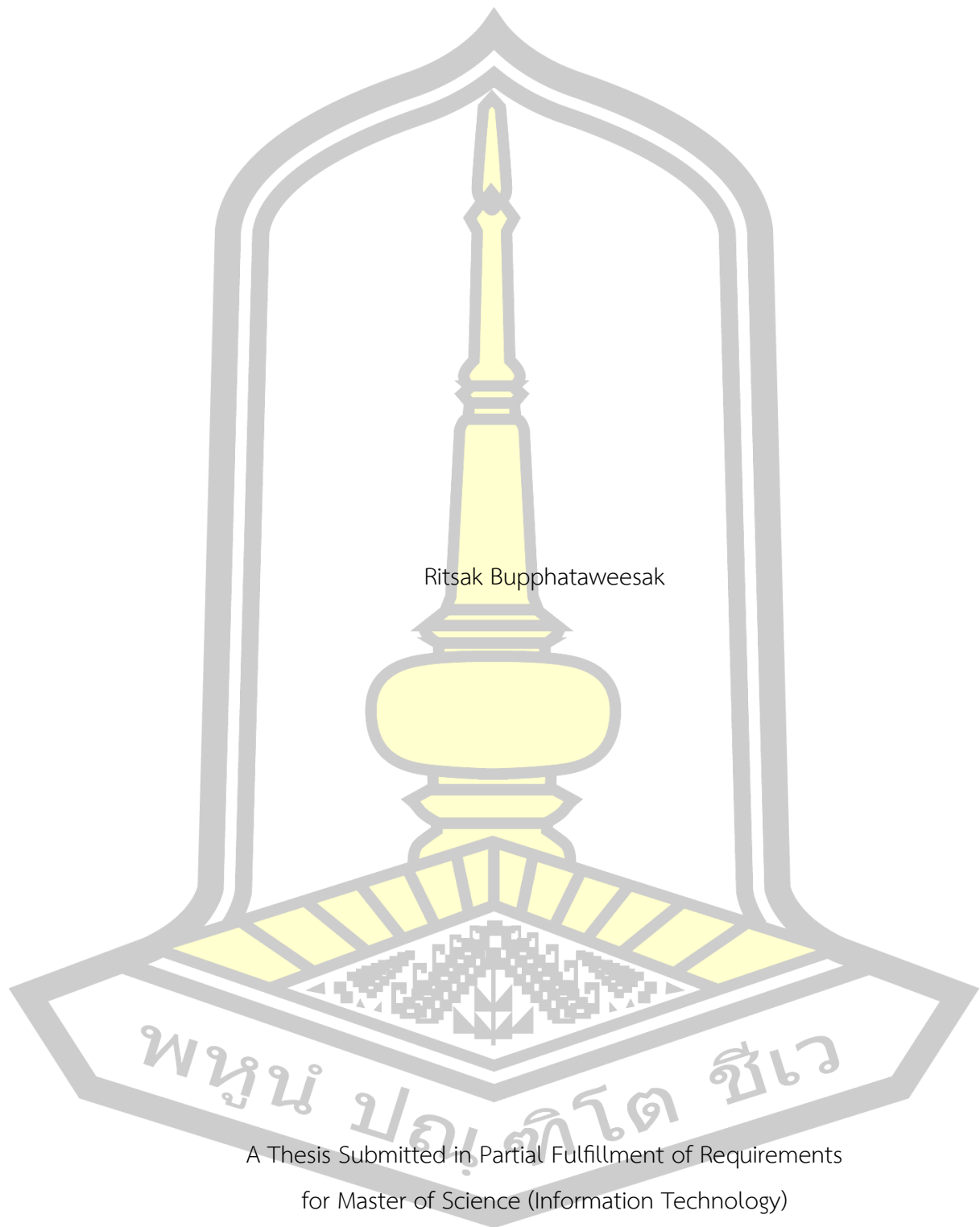
เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

กรกฎาคม 2567

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Opinion analysis of alternative-energy vehicle users utilizing the opinion mining



Ritsak Bupphataweesak

A Thesis Submitted in Partial Fulfillment of Requirements
for Master of Science (Information Technology)

July 2024

Copyright of Mahasarakham University



คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายฤทธิศักดิ์ บุปผา
ทวีศักดิ์ แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ผศ. ดร. ธาดา จันทะคุณ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผศ. ดร. จารีย์ ทองคำ)

.....กรรมการ

(ผศ. ดร. แกมกาญจน์ สมประเสริฐศรี)

.....กรรมการผู้ทรงคุณวุฒิภายนอก

(ผศ. ดร. ธาดา จันทะคุณ)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

.....
(รศ. ดร. จันทิมา พลพินิจ)

คณบดีคณะวิทยาการสารสนเทศ

.....
(รศ. ดร. กริสน์ ชัยมูล)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น		
ผู้วิจัย	ฤทธิ์ศักดิ์ บุปผาพิทักษ์ดี		
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. จาริ ทองคำ		
ปริญญา	วิทยาศาสตรมหาบัณฑิต	สาขาวิชา	เทคโนโลยีสารสนเทศ
มหาวิทยาลัย	มหาวิทยาลัยมหาสารคาม	ปีที่พิมพ์	2567

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยฉบับนี้เพื่อคัดเลือกเทคนิคในการเรียนรู้ของเครื่อง รวมถึงการเรียนรู้เชิงลึก ที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือก เทคนิคที่นำมาใช้ในการสร้างแบบจำลอง ประกอบด้วย เทคนิคต้นไม้ตัดสินใจ ซี 4.5 เทคนิคเนอรัลเน็ต เทคนิคการเรียนรู้ของเครื่อง และเทคนิคเพอร์เซปตรอนหลายชั้น ผู้วิจัยได้ทำการรวบรวมข้อมูลจากเว็บ Facebook, YouTube, Pantip และ TikTok จำนวน 2,265 ความคิดเห็น จากนั้นได้นำกระบวนการเหมืองความคิดเห็นมาใช้ในการจำแนกความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกเป็นความคิดเห็นเชิงบวกและเชิงลบ ในการวัดประสิทธิภาพของแบบจำลองผู้วิจัยได้ใช้หลักการไขว้ทับแบบ 10 ส่วน เพื่อแบ่งข้อมูลเป็นชุดเรียนรู้ และชุดทดสอบ ค่าความถ่วง (F-Measure) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ได้ถูกนำมาใช้ในการแสดงประสิทธิภาพของแบบจำลอง จากการทดลองเพื่อคัดเลือกเทคนิคที่สามารถสร้างแบบจำลองที่มีประสิทธิภาพสูงสุด พบว่าเทคนิคเพอร์เซปตรอนหลายชั้น สร้างสามารถสร้างแบบจำลองที่มีประสิทธิภาพสูงสุดที่ค่าความถ่วงที่ร้อยละ 95.30 ค่าความแม่นยำที่ร้อยละ 95.30 และค่าความระลึกที่ร้อยละ 95.30

คำสำคัญ : เหมืองความคิดเห็น, การวิเคราะห์ความคิดเห็น, ยานยนต์พลังงานทางเลือก, เหมืองข้อมูล

TITLE Opinion analysis of alternative-energy vehicle users utilizing the opinion mining

AUTHOR Ritsak Bupphataweesak

ADVISORS Assistant Professor Jaree Thongkam , Ph.D.

DEGREE Master of Science **MAJOR** Information Technology

UNIVERSITY Mahasarakham **YEAR** 2024
University

ABSTRACT

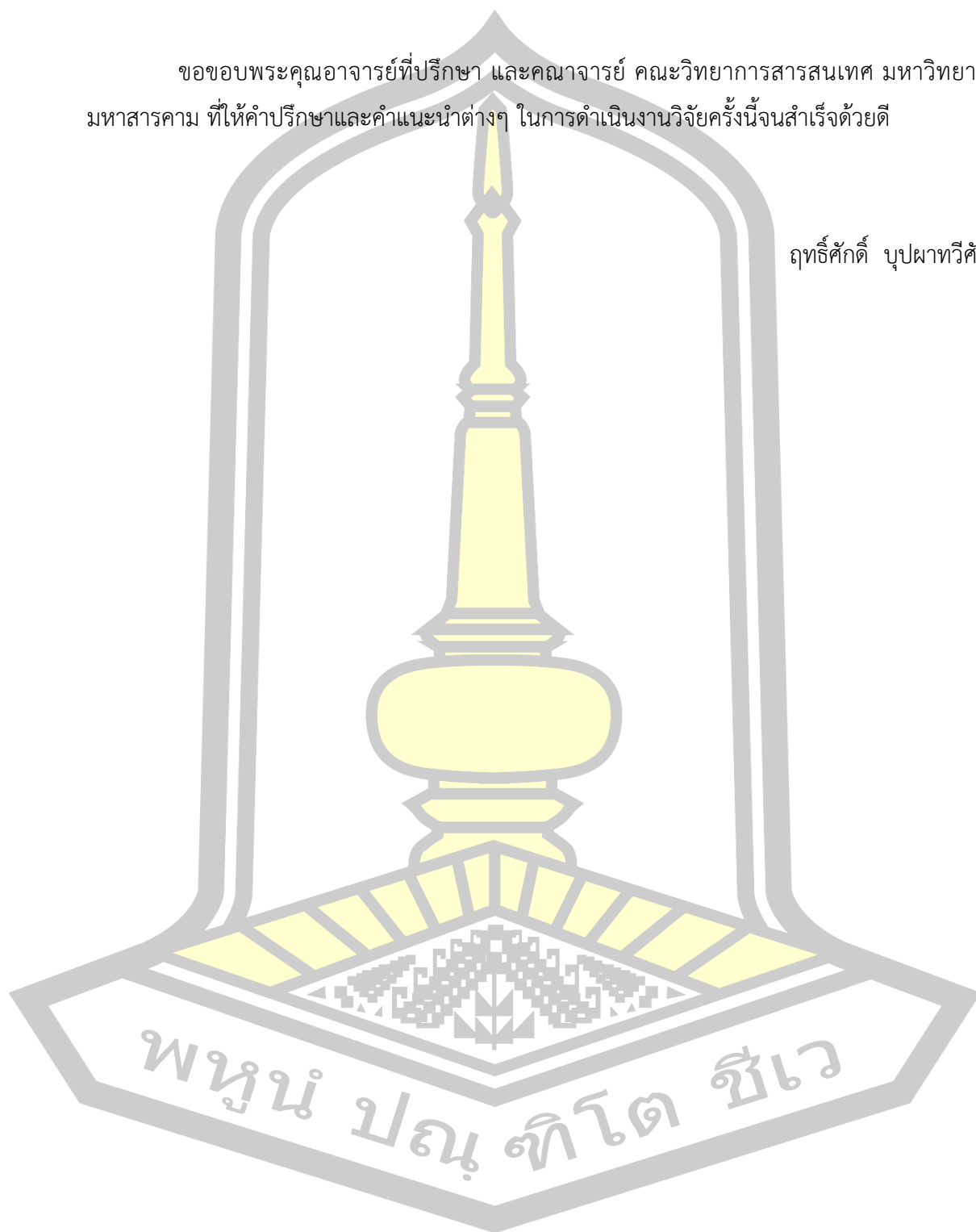
The objective of this research is to select the most efficient machine learning techniques, including deep learning, for classifying user opinions on alternative energy vehicles. The techniques considered for model creation include decision tree C4.5, Naive Bayes, machine learning, and multilayer perceptron. The researchers collected data from Facebook, YouTube, Pantip, and TikTok, comprising 2,265 comments. Opinion mining was employed to categorize the comments as positive or negative. In order to evaluate model performance, the researchers utilized the 10-fold cross-validation method to split the data into training and testing sets. The F-Measure, Precision, and Recall metrics were employed to demonstrate the effectiveness of the models. Experimental results aimed at selecting the technique that could generate the most efficient model revealed that the multilayer perceptron technique achieved the highest performance, with F-Measure, Precision, and Recall values all at 95.30 percent.

Keyword : Opinion Mining, Opinion Analysis, Alternative Energy Vehicles, Data Mining

กิตติกรรมประกาศ

ขอขอบพระคุณอาจารย์ที่ปรึกษา และคณาจารย์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม ที่ให้คำปรึกษาและคำแนะนำต่างๆ ในการดำเนินงานวิจัยครั้งนี้จนสำเร็จด้วยดี

ฤทธิ์ศักดิ์ บุปผาวิศักดิ์

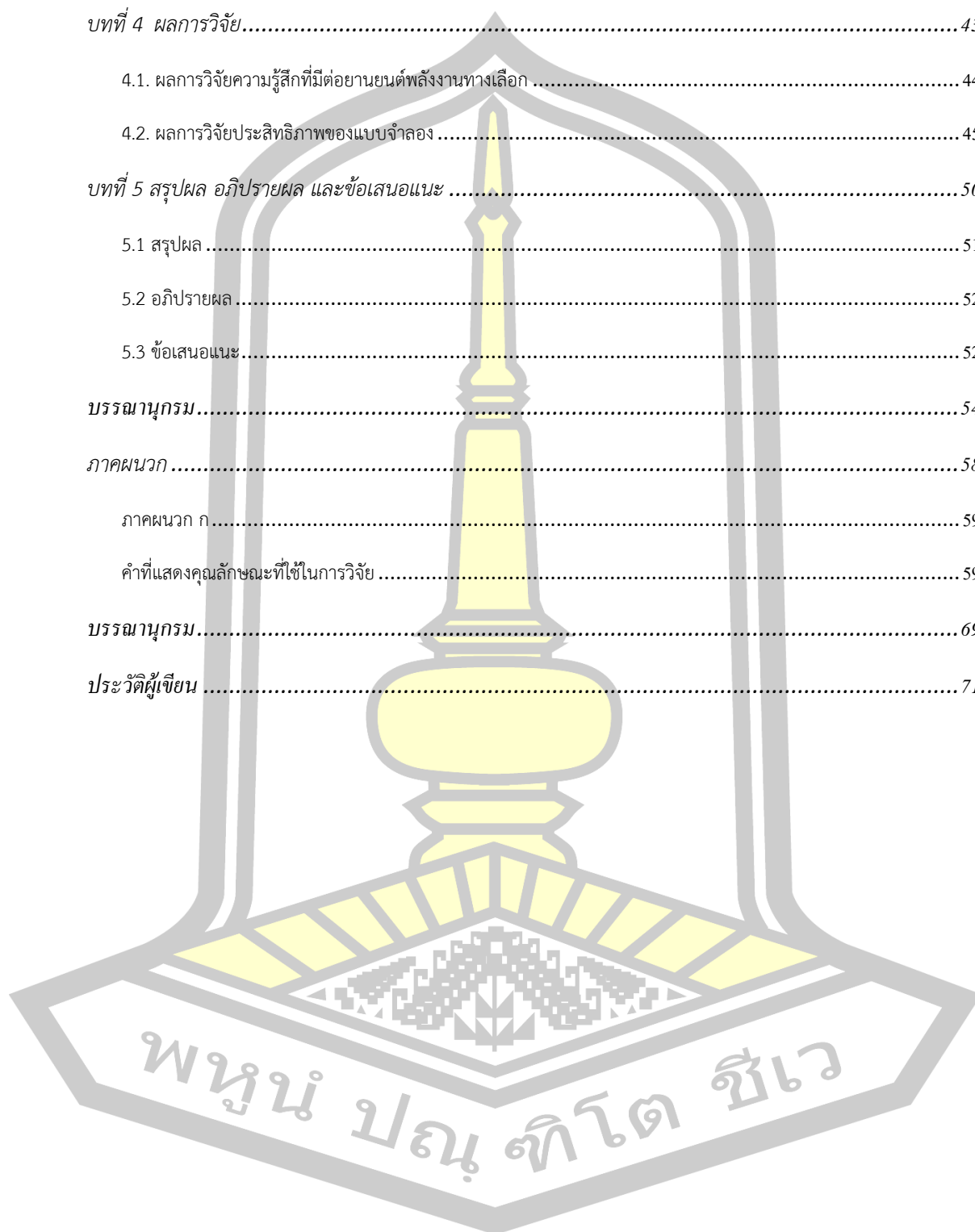


สารบัญ

หน้า

บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
บทที่ 1 บทนำ	1
1.1 ความเป็นมาของการวิจัย.....	1
1.2 ความมุ่งหมายของการวิจัย.....	2
1.3 ขอบเขตของการวิจัย.....	2
1.4 ผลที่คาดว่าจะได้รับการวิจัย	2
1.5 นิยามศัพท์เฉพาะ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	6
ในการวิจัยครั้งนี้มีทฤษฎีที่เกี่ยวข้องได้แก่ ยานยนต์พลังงานทางเลือกทั้ง 5 ชนิด การทำเหมืองข้อมูล กระบวนการการทำเหมือง ความคิดเห็น เทคนิคที่ใช้ในการวิจัย และงานวิจัยที่เกี่ยวข้อง	6
2.1 ทฤษฎีที่เกี่ยวข้อง	6
2.1.1 ยานยนต์พลังงานทางเลือก (Alternative Energy Vehicle)	6
2.1.2 การทำเหมืองข้อมูล.....	7
2.1.3 กระบวนการการทำเหมืองความคิดเห็น.....	8
2.1.4 เทคนิคที่ใช้ในการวิจัย	9
2.2 งานวิจัยที่เกี่ยวข้อง	12
บทที่ 3 วิธีดำเนินการวิจัย.....	16
3.1 การเตรียมข้อมูล	17
3.2 ขั้นตอนก่อนการประมวลผลข้อความ	28
3.3 การสร้างคุณลักษณะ	29
3.4 การสร้างแบบจำลองเพื่อการจำแนก	34

3.5 การแปลผลและการประเมินผล	41
บทที่ 4 ผลการวิจัย.....	43
4.1. ผลการวิจัยความรู้สึกที่มีต่อยานยนต์พลังงานทางเลือก	44
4.2. ผลการวิจัยประสิทธิภาพของแบบจำลอง	45
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ	50
5.1 สรุปผล	51
5.2 อภิปรายผล	52
5.3 ข้อเสนอแนะ	52
บรรณานุกรม.....	54
ภาคผนวก.....	58
ภาคผนวก ก.....	59
คำที่แสดงคุณลักษณะที่ใช้ในการวิจัย	59
บรรณานุกรม.....	69
ประวัติผู้เขียน	71



สารบัญภาพ

ภาพประกอบที่ 2.1 กระบวนการในการทำเหมืองความคิดเห็นงานวิจัย ของ Han	9
ภาพประกอบที่ 2.2 เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5.....	10
ภาพประกอบที่ 2.3 เทคนิคเอนีฟเบย์.....	11
ภาพประกอบที่ 2.4 เทคนิค Support Vector Machine	11
ภาพประกอบที่ 2.5 เทคนิคเพอร์เซปตรอนหลายชั้น	12
ภาพประกอบที่ 3.1 ผังงานการดำเนินการวิจัย.....	16
ภาพประกอบที่ 3.2 ตัวอย่างความคิดเห็นจาก Youtube	17
ภาพประกอบที่ 3.3 ตัวอย่างความคิดเห็นจาก Pantip.....	17
ภาพประกอบที่ 3.4 เมนูส่วนขยายของ Chrome.....	18
ภาพประกอบที่ 3.5 หน้าต่างจัดการส่วนขยาย	19
ภาพประกอบที่ 3.6 ค้นหาโปรแกรม Web Scraper สำหรับติดตั้ง	19
ภาพประกอบที่ 3.7 หน้าต่างติดตั้งโปรแกรม Web Scraper.....	20
ภาพประกอบที่ 3.8 ยืนยันการติดตั้งโปรแกรม	20
ภาพประกอบที่ 3.9 ตรวจสอบหลังจากการติดตั้ง.....	21
ภาพประกอบที่ 3.10 การใช้งานโปรแกรม Web Scraper กับเว็บไซต์ Pantip.....	21
ภาพประกอบที่ 3.11 การใช้งานโปรแกรม Web Scraper กับเว็บไซต์ Tiktok.....	21
ภาพประกอบที่ 3.12 เปิดใช้งาน Web Scraper.....	22
ภาพประกอบที่ 3.13 การสร้างผังของเว็บ	23
ภาพประกอบที่ 3.14 ข้อมูลสำหรับสร้างผังของเว็บ.....	23
ภาพประกอบที่ 3.15 เตรียมเลือกข้อมูลความคิดเห็น	23
ภาพประกอบที่ 3.16 ตั้งค่าเลือกความคิดเห็นพร้อมกันหลาย ๆ ความคิดเห็น.....	24
ภาพประกอบที่ 3.17 หน้าต่างก่อนการเลือกความคิดเห็น	24
ภาพประกอบที่ 3.18 หน้าต่างหลังจากเลือกความคิดเห็น	25

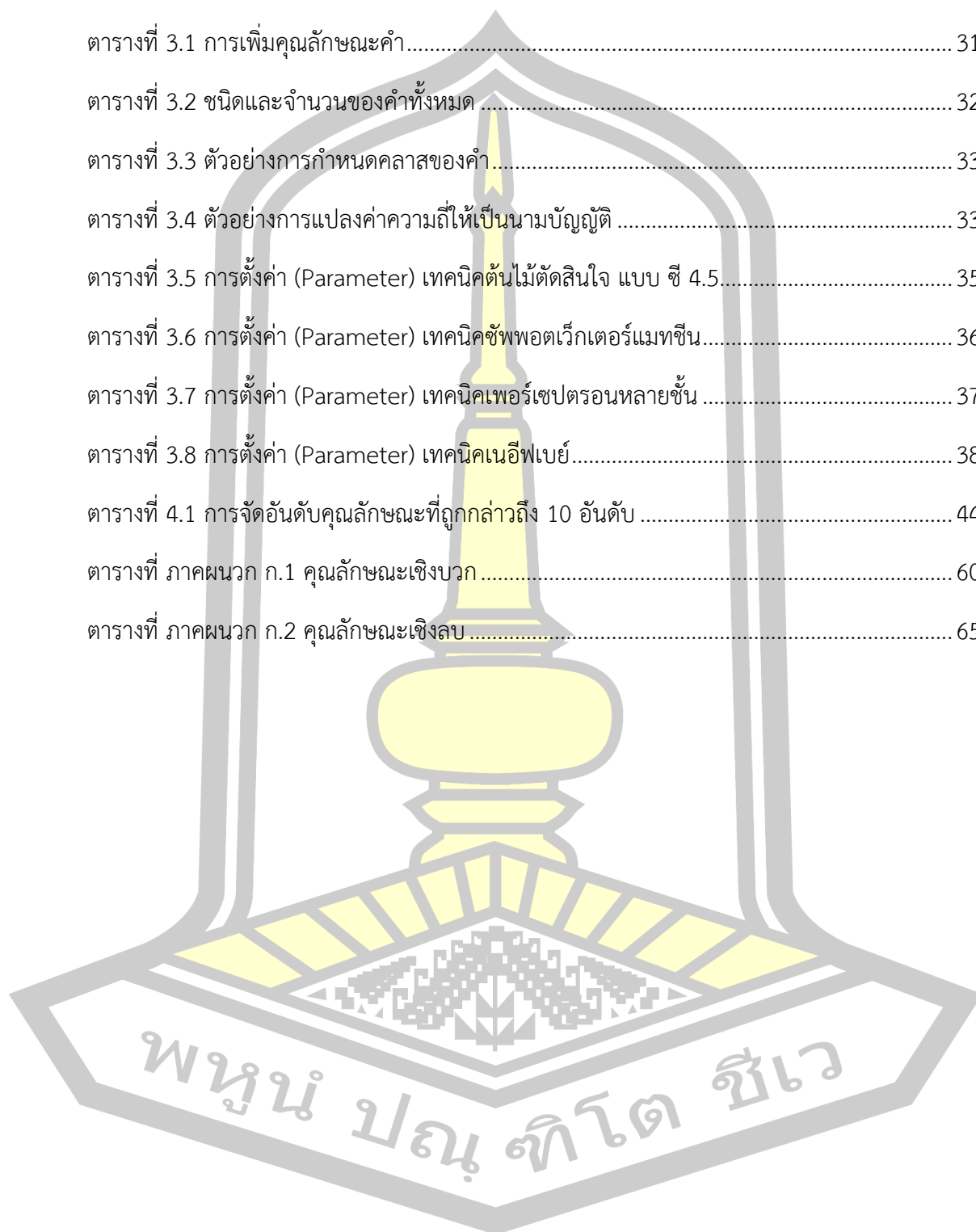
สารบัญภาพ (ต่อ)

ภาพประกอบที่ 3.19 ID ของความคิดเห็นที่ถูกเลือก	26
ภาพประกอบที่ 3.20 ความคิดเห็นที่เลือกทั้งหมด	26
ภาพประกอบที่ 3.21 ภาษา Python ใช้คัดเลือกความคิดเห็น.....	27
ภาพประกอบที่ 3.22 ตัวอย่างการแสดงความเห็นที่มีการเพิ่ม Sticker	28
ภาพประกอบที่ 3.23 ตัวอย่างการแสดงความเห็นที่มีการเพิ่ม Emoticon.....	28
ภาพประกอบที่ 3.24 ความคิดเห็นก่อนและหลังการทำความสะอาดคำ	29
ภาพประกอบที่ 3.25 ความคิดเห็นหลังจากการทำความสะอาดและแก้ไขคำ.....	29
ภาพประกอบที่ 3.26 การแยกคำเพื่อแยกคุณลักษณะ.....	30
ภาพประกอบที่ 3.27 ข้อมูลหลังจากการตัดคำ.....	30
ภาพประกอบที่ 3.28 ชนิดของคำ (Part of Speech)	32
ภาพประกอบที่ 4.1 การเปรียบเทียบความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก	44
ภาพประกอบที่ 4.2 การเปรียบเทียบค่าความแม่นยำ.....	46
ภาพประกอบที่ 4.3 การเปรียบเทียบค่าความระลึก	47
ภาพประกอบที่ 4.4 การเปรียบเทียบค่าความถ่วง	48



สารบัญตาราง

ตารางที่ 3.1 การเพิ่มคุณลักษณะคำ.....	31
ตารางที่ 3.2 ชนิดและจำนวนของคำทั้งหมด.....	32
ตารางที่ 3.3 ตัวอย่างการกำหนดคลาสของคำ.....	33
ตารางที่ 3.4 ตัวอย่างการแปลงค่าความถี่ให้เป็นนามบัญญัติ.....	33
ตารางที่ 3.5 การตั้งค่า (Parameter) เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5.....	35
ตารางที่ 3.6 การตั้งค่า (Parameter) เทคนิคซัพพอร์ตเวกเตอร์แมทชีน.....	36
ตารางที่ 3.7 การตั้งค่า (Parameter) เทคนิคเพอร์เซปตรอนหลายชั้น.....	37
ตารางที่ 3.8 การตั้งค่า (Parameter) เทคนิคเนอรัลเน็ต.....	38
ตารางที่ 4.1 การจัดอันดับคุณลักษณะที่ถูกกล่าวถึง 10 อันดับ.....	44
ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก.....	60
ตารางที่ ภาคผนวก ก.2 คุณลักษณะเชิงลบ.....	65



บทที่ 1

บทนำ

1.1 ความเป็นมาของการวิจัย

จากภาวะวิกฤตพลังงานเชื้อเพลิงโลกมีราคาผันผวน [1] ส่งผลกระทบต่อราคาน้ำมันในเมืองไทย ทำให้มีราคาที่สูงขึ้น จึงทำให้ผู้ใช้งานยานยนต์มีความสนใจในยานยนต์พลังงานทางเลือกเพิ่มมากขึ้น [2] เพื่อเป็นทางเลือกหนึ่งในการประหยัดค่าใช้จ่ายประจำวัน [3] ยานยนต์พลังงานทางเลือก รถยนต์ไฮบริด รถยนต์ปลั๊กอินไฮบริด รถยนต์ไฟฟ้า รถยนต์ไฟฟ้าแบบเซลล์เชื้อเพลิง เป็นต้น แต่อย่างไรก็ตามยานยนต์พลังงานทางเลือกยังไม่แพร่หลายมากนักเนื่องจากข้อจำกัดต่าง ๆ ในการเดินทางรวมถึงความไม่มั่นใจการใช้งานในช่วงฤดูฝนที่มีน้ำท่วมขัง [4] จึงทำให้กลุ่มผู้ใช้งานจริงจำนวนมากที่ได้แสดงความคิดเห็นผ่านทางสื่อสังคมออนไลน์ต่าง ๆ เช่น เว็บไซต์ Facebook, YouTube, Pantip และ TikTok ในรูปแบบของการแสดงความคิดเห็นเชิงบวก และเชิงลบ ทำให้ไม่สามารถแยกได้ว่าความคิดเห็นนั้นเป็นเชิงบวกหรือเชิงลบ ดังนั้นในงานวิจัยนี้ได้ทำการศึกษาความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือก โดยใช้กระบวนการเหมืองความคิดเห็น

เหมืองความคิดเห็น (Opinion Mining) คือ กระบวนการวิเคราะห์ความคิดเห็นที่เป็นภาษาธรรมชาติ ในการจำแนกความคิดเห็นออกเป็นเชิงบวก เชิงลบ และเป็นกลาง [5] นักวิจัยจำนวนมากได้นำเหมืองความคิดเห็นมาใช้ในการจำแนกความคิดเห็น ซึ่งเทคนิคที่นำมาใช้ในการสร้างแบบจำลองโดยส่วนใหญ่เป็นเทคนิคจากเหมืองข้อมูล (Data Mining) และการเรียนรู้ของเครื่อง (Machine Learning) เช่น ปรามอยซ์ นามวงศ์ และคณะ [6] ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคนาอิวเบย์ เทคนิคเคเนียร์เสนเบอร์ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 และเทคนิคความจำระยะสั้น-ระยะยาว ในการทดลองจำแนกความคิดเห็นบนเครือข่ายสังคมออนไลน์เฟสบุ๊กของผู้ใช้บริการร้านอาหารในจังหวัดอุบลราชธานี พบว่า เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 มีประสิทธิภาพสูงสุดในการทดลองมีค่าความถูกต้องเท่ากับร้อยละ 89.00 ส่วนงานวิจัยของ พิศิษฐ์ บวรเลิศสุธี และคณะ [7] ใช้เทคนิคเคเนียร์เสนเบอร์ เทคนิคป่าไม้แบบสุ่ม เทคนิคการลดการเคลื่อนย้าย และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ในการจำแนกความคิดเห็นของผู้ใช้สินค้า พบว่า เทคนิคซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพในการจำแนกสูงที่สุดที่ค่าความแม่นยำที่ร้อยละ 79.00 ยิ่งไปกว่านั้น Adebayo Abayomi และคณะ [8] ได้ใช้เทคนิคพจนานุกรมเล็กชิตรอน เทคนิคซัพพอร์ตเวกเตอร์แมชชีน และเทคนิคการเรียนรู้เชิงลึก ในการจำแนกความคิดเห็นของผู้ใช้งานโปรแกรม Yahoo จาก Twitter พบว่า เทคนิคการเรียนรู้เชิงลึก มีประสิทธิภาพดีที่สุดที่ค่าความแม่นยำถึงร้อยละ 80.38

ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะพัฒนาแบบจำลองที่จะจำแนกความคิดเห็นของคนไทยต่อผู้ใช้นายยนต์พลังงานทางเลือก เพื่อให้ได้ความคิดเห็นของคนไทยที่มีแต่นายยนต์พลังงานทางเลือกกว่าอยู่ในเชิงบวกหรือเชิงลบ ด้วยการทำให้เหมือนความคิดเห็น โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree: C4.5) เทคนิคเนออีฟเบย์ (Naïve Bayes: NB) เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine: SVM) และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron: MLP) การวัดประสิทธิภาพแบบจำลองผู้วิจัยใช้วิธีการ 10-Fold Cross Validation เพื่อแบ่งกลุ่มข้อมูลออกเป็นชุดข้อมูล 2 ชุด คือ ชุดข้อมูลสำหรับเรียนรู้และชุดข้อมูลสำหรับทดสอบ โดยใช้ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) ในการเปรียบเทียบประสิทธิภาพของแบบจำลอง เพื่อหาว่าเทคนิคใดมีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของคนไทยที่มีต่อนายยนต์พลังงานทางเลือก

1.2 ความมุ่งหมายของการวิจัย

- เพื่อศึกษาความคิดเห็นของคนไทยที่มีต่อนายยนต์พลังงานทางเลือก
- เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกความคิดเห็นของคนไทยที่มีต่อนายยนต์พลังงานทางเลือกบนสื่อสังคมออนไลน์ด้วยเทคนิคเหมือนความคิดเห็น

1.3 ขอบเขตของการวิจัย

- เปรียบเทียบประสิทธิภาพของแบบจำลองในการทำเหมืองข้อมูลจำแนกรู้สึกของคนไทย ที่มีต่อนายยนต์พลังงานทางเลือกบนสื่อสังคมออนไลน์ 4 แบบ ได้แก่

1. Decision Tree C4.5: C4.5
2. Naïve-Bayes: NB
3. Support Vector Machine: SVM
4. Multi-layer Perceptron: MLP

- ศึกษาและวิเคราะห์ความคิดเห็นของคนไทยจากข้อมูลการแสดงความคิดเห็นที่เกี่ยวกับนายยนต์พลังงานทางเลือกจาก 4 เว็บไซต์ ได้แก่ Twitter, Facebook, Pantip และ Tiktok จำนวน 7,680 ความคิดเห็น

1.4 ผลที่คาดว่าจะได้รับการวิจัย

- ได้ความคิดเห็นของคนไทยที่มีต่อนายยนต์พลังงานทางเลือกจากการทำเหมืองความคิดเห็น

- ได้แบบจำลองที่มีประสิทธิภาพในการจำแนกความคิดเห็นของผู้ใช้รถยนต์พลังงานทางเลือกที่เป็นแบบ Hybrid, Plugin Hybrid และ Electric Vehicle จากเว็บ Twitter, Facebook, Pantip และ Tiktok

1.5 นิยามศัพท์เฉพาะ

1.5.1 ยานยนต์พลังงานทางเลือก (Alternative-Energy Vehicle) คือ รถยนต์ ที่ใช้พลังงานมากกว่า 1 อย่าง ในการขับเคลื่อนหรือจะเป็นรถยนต์ที่ไม่มีการสันดาปในเครื่องยนต์ซึ่งในปัจจุบันถูกแบ่งออกเป็น 4 กลุ่มหลักๆ ได้แก่ รถยนต์ไฟฟ้าไฮบริด (Hybrid Electric Vehicle: HEV) รถยนต์ไฟฟ้าปลั๊กอินไฮบริด (Plug-in Hybrid Electric Vehicle: PHEV) รถยนต์ไฟฟ้าแบตเตอรี่ (Battery Electric Vehicle: BEV) และ รถยนต์ไฟฟ้าเซลล์เชื้อเพลิง (Fuel Cell Electric Vehicle: FCEV) ซึ่งทั้ง 4 กลุ่มนี้จะมีลักษณะพิเศษอย่างหนึ่งที่ตรงกันคือนอกจากจะใช้พลังงานไฟฟ้าเป็นส่วนหนึ่งในการขับเคลื่อนแล้ว รถยนต์ทั้ง 4 กลุ่มได้มีการแข่งขันด้านการตลาดในประเทศไทยที่เห็นได้ชัดเจนก็คือตามสื่อสังคมออนไลน์ต่าง ๆ ได้มีการก่อตั้งชุมชน (Community) ของคนที่สนใจรถยนต์ในกลุ่มนั้น ๆ ขึ้นมามากมาย จึงทำให้ง่ายต่อการหาข้อมูล (Data) เพื่อใช้ในการวิจัย

1.5.2 รถยนต์ไฟฟ้าไฮบริด (Hybrid Electric Vehicle: HEV) คือ รถยนต์ไฮบริด คือรถยนต์ที่ประกอบด้วยเครื่องยนต์สันดาปภายใน พร้อมมีมอเตอร์ไฟฟ้าเสริมกำลังขับเคลื่อน และสนับสนุนการเบรกแบบผันกลับ (Regenerative Braking) เพื่อเก็บสำรองพลังงานในรูปแบบพลังงานไฟฟ้าเป็นการผสมผสานระหว่างเครื่องยนต์กับมอเตอร์ไฟฟ้า ซึ่งทำให้ระบบการขับเคลื่อนมีประสิทธิภาพสูงขึ้น ล้นเปลืองเชื้อเพลิงต่ำกว่ารถยนต์ปกติ

1.5.3 รถยนต์ไฟฟ้าปลั๊กอินไฮบริด (Plug-in Hybrid Electric Vehicle: PHEV) คือ รถยนต์ไฟฟ้าที่พัฒนาต่อมาจากรถยนต์ไฟฟ้าไฮบริด โดยสามารถประจุพลังงานไฟฟ้าได้จากแหล่งภายนอก (Plug-in) ทำให้รถยนต์สามารถใช้พลังงานพร้อมกันจาก 2 แหล่ง มีแบตเตอรี่ขนาดใหญ่ขึ้นจากแบบไฮบริดธรรมดา จึงสามารถวิ่งในระยะทางที่ใช้พลังงานจากไฟฟ้าโดยตรงเพิ่มขึ้น รถยนต์ไฟฟ้าแบบ PHEV มีการออกแบบอยู่ 2 รูปแบบ ได้แก่ รูปแบบ Extended Range EV หรือ (E-REV) ที่เน้นการทำงานโดยใช้พลังงานไฟฟ้าเป็นหลักก่อนและรูปแบบ Blended PHEV ที่มีการทำงานผสมผสานระหว่างเครื่องยนต์และไฟฟ้า ดังนั้น รถยนต์ไฟฟ้าแบบ E-REV สามารถวิ่งด้วยพลังงานไฟฟ้าอย่างเดียวมากกว่าแบบ Blended PHEV

1.5.4 รถยนต์ไฟฟ้าแบตเตอรี่ (Battery Electric Vehicle: BEV) คือ เกิดจากการพัฒนาทางเทคโนโลยี จึงทำให้รถยนต์ชนิดนี้ใช้เพียงพลังงานไฟฟ้าในการขับเคลื่อน โดยวิธีการทำงานของรถยนต์พลังงานไฟฟ้าแบบ 100% นั้น ไม่ได้มีความละเอียดและซับซ้อนเหมือนรถยนต์ที่ใช้น้ำมันเชื้อเพลิง มีองค์ประกอบหลักในการขับเคลื่อนเพียง 3 ส่วนเท่านั้น แต่ทั้งสามส่วนนี้ได้ก่อให้เกิดการขับเคลื่อนที่เต็มประสิทธิภาพ ซึ่งทั้งสามส่วนที่วานี้ได้แก่ แบตเตอรี่ อุปกรณ์แปลงกระแสไฟฟ้า และ มอเตอร์ไฟฟ้า

1.5.5 รถยนต์ไฟฟ้าเซลล์เชื้อเพลิง (Fuel Cell Electric Vehicle: FCEV) คือ เป็นยานยนต์ไฟฟ้าที่มีเซลล์เชื้อเพลิง (Fuel cell) ที่สามารถผลิตพลังงานไฟฟ้าได้โดยตรง รถยนต์ไฟฟ้าเซลล์เชื้อเพลิงมีข้อดีหลายๆ ประการ ข้อดีที่สำคัญที่สุดคือ ประสิทธิภาพของเซลล์เชื้อเพลิงมีค่าสูงถึง 60% และความจุพลังงานจำเพาะที่สูงกว่าแบตเตอรี่ที่มีอยู่ในปัจจุบัน รถยนต์ไฟฟ้าเซลล์เชื้อเพลิงจึงเป็นเทคโนโลยีที่บริษัทรถยนต์เชื่อว่าเป็นคำตอบที่แท้จริงของพลังงานสะอาดในอนาคต อย่างไรก็ตามก็ดียังมีข้อจำกัดในเรื่องการผลิตไฮโดรเจนและโครงสร้างพื้นฐาน

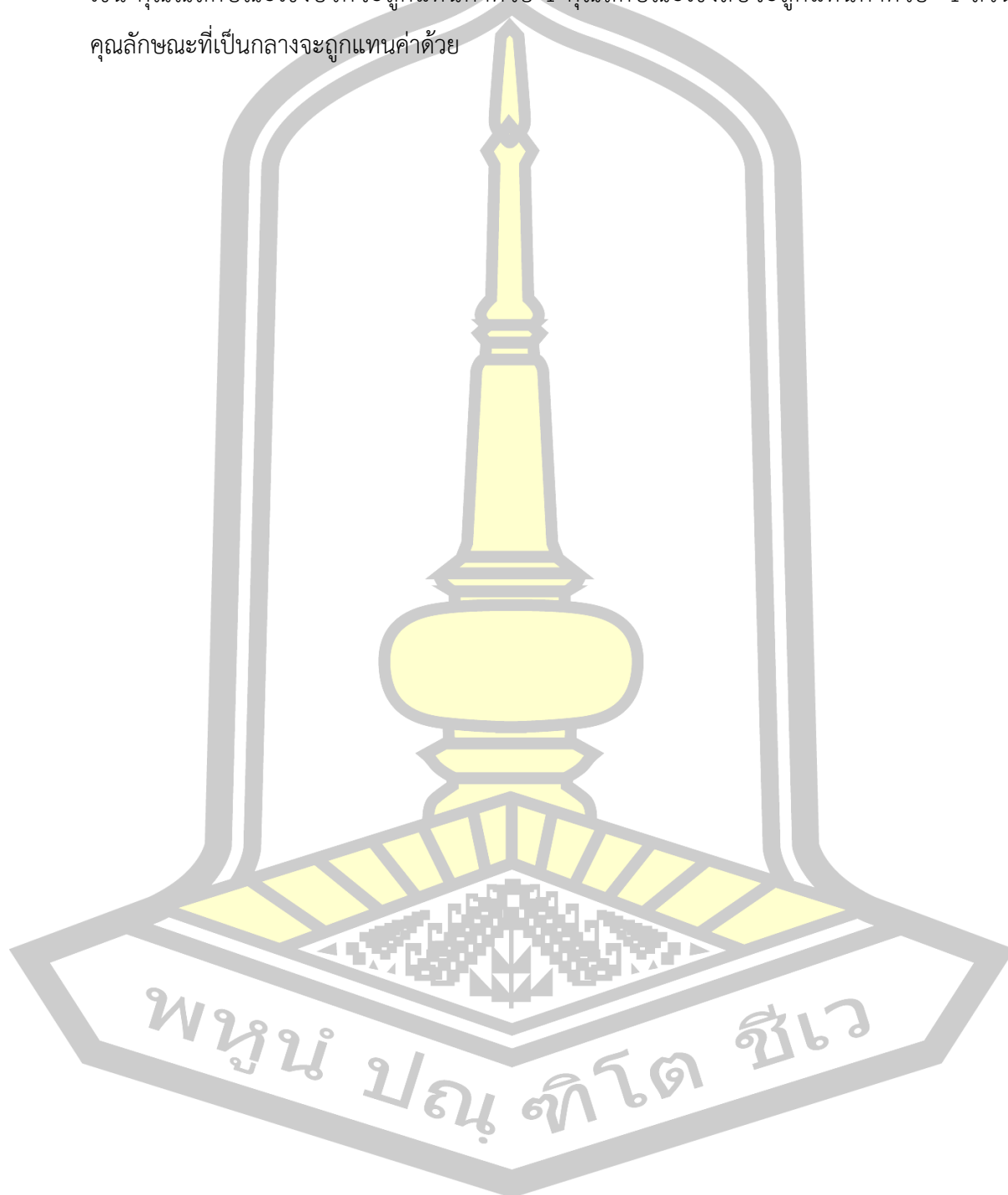
1.5.6 สื่อสังคมออนไลน์ (Social Media) คือ เว็บไซต์ที่มีกลุ่มประชากรใช้งานเพื่อสนทนาและแสดงความคิดเห็นต่าง ๆ ด้วยการ โพสต์ (Post) และ คอมเมนต์ (Comment) ซึ่งสามารถแสดงความคิดเห็นได้หลากหลายรูปแบบ เช่น ตัวอักษร รูปภาพ เสียง วิดีโอ ซึ่งในปัจจุบันมีหลากหลายผู้ให้บริการ เช่น Facebook, Twitter, Pantip, Tiktok, Youtube เป็นต้น

1.5.7 การวิเคราะห์ความคิดเห็น (Opinion Analysis) คือ การวิเคราะห์ข้อความหรือประโยคที่ได้มีการแสดงความคิดเห็นที่สามารถแสดงอารมณ์หรือความรู้สึกของคนไทยที่มีต่อ ยานยนต์พลังงานทางเลือก โดยจะแบ่งออกเป็นความคิดเห็นเชิงบวก และความคิดเห็นเชิงลบ

1.5.8 กระบวนการทำเหมืองความคิดเห็น (Opinion mining process) คือ เป็นกระบวนการต่าง ๆ ที่ใช้ในการสกัดเอาคุณลักษณะหรือความรู้สึกที่อยู่ภายในความคิดเห็นออกมา โดยจะมีทั้งหมด 5 ขั้นตอน ได้แก่ 1) การเตรียมข้อมูล 2) ขั้นตอนก่อนการประมวลผลข้อความ 3) การสร้างคุณลักษณะ 4) การสร้างแบบจำลองเพื่อจำแนกความคิดเห็น 5) การประมวลผล ซึ่งทั้ง 5 ขั้นตอนนี้ถ้าจะมีการเปลี่ยนแปลงได้ตามความเหมาะสมและวัตถุประสงค์ของนักวิจัย

1.5.9 แบบจำลอง (Model) คือ ต้นแบบจากการใช้เทคนิคต่าง ๆ ในวิเคราะห์ความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกซึ่งในงานวิจัยชิ้นนี้ได้ใช้เทคนิคในการวิจัยทั้งหมด 4 เทคนิค ได้แก่ Decision Tree C4.5, Naïve-Bayes, Support Vector Machine และ Multi-layer Perceptron

1.5.10 คุณลักษณะ (Feature) คือ คำที่ผ่านการพิจารณาแล้ว สามารถระบุได้ว่าเป็นคุณลักษณะเชิงบวก เชิงลบ หรือ เป็นกลาง โดยคุณลักษณะจะถูกแทนค่าในผ่านการพิจารณาแล้ว เช่น คุณลักษณะเชิงบวกจะถูกแทนค่าด้วย 1 คุณลักษณะเชิงลบจะถูกแทนค่าด้วย -1 ส่วนคุณลักษณะที่เป็นกลางจะถูกแทนค่าด้วย 0



บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้มีทฤษฎีที่เกี่ยวข้องได้แก่ ยานยนต์พลังงานทางเลือกทั้ง 5 ชนิด การทำเหมืองข้อมูล กระบวนการการทำเหมืองความคิดเห็น เทคนิคที่ใช้ในการวิจัย และงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ยานยนต์พลังงานทางเลือก (Alternative Energy Vehicle)

ยานยนต์พลังงานทางเลือก (Alternative Energy Vehicle: AEV) คือ รถยนต์หรือยานพาหนะที่ใช้การขับเคลื่อนโดยการใช้เชื้อเพลิงมากกว่า 1 อย่างหรือขับเคลื่อนโดยไม่ใช้การเผาไหม้และการสันดาป โดยจะมีคุณสมบัติพิเศษที่เครื่องยนต์สันดาปไม่มีนั่นคือลดการปล่อยก๊าซคาร์บอนไดออกไซด์ เป็นมิตรต่อสิ่งแวดล้อม โดยทาง Nationwide [9] ได้ระบุว่ายานยนต์พลังงานทางเลือกที่ได้รับการรับรองจากรัฐบาลสหรัฐอเมริกานั้นมีด้วยกันทั้งหมดอยู่ 5 แบบ ได้แก่

1) Hybrid electric vehicles รถยนต์ที่ใช้เชื้อเพลิงทางเลือกมากกว่าหนึ่งอย่าง คือ รถไฮบริดไฟฟ้าและแก๊ส (HEV) การผสมผสานระหว่างระบบขับเคลื่อนด้วยแก๊สและไฟฟ้า ทำให้รถไฮบริดมีจำหน่ายอย่างแพร่หลายในสหรัฐอเมริกามานานกว่า 20 ปี โดยที่ Toyota Prius เป็นรถยนต์ไฮบริดรุ่นแรกที่เกิดขึ้นเป็นจำนวนมากในปี 1997 รถยนต์ไฟฟ้าไฮบริดเหล่านี้ได้รับความนิยมเพิ่มขึ้นอย่างต่อเนื่อง โดยปัจจุบันมีรถจำหน่ายอยู่ประมาณ 50 รุ่น แบตเตอรี่ในรถไฮบริด-ไฟฟ้าแบบใช้แก๊สจะชาร์จจากเครื่องยนต์และผ่านการเบรก ซึ่งช่วยให้แบตเตอรี่สามารถขับเคลื่อนรถด้วยความเร็วต่ำและระหว่างการหยุดและสตาร์ทรถ

2) Plug-in hybrid electric vehicles Plug-in Hybrid (PHEV) คล้ายกับระบบไฮบริดที่ใช้แก๊ส-ไฟฟ้า มีแบตเตอรี่ขนาดใหญ่กว่าที่สามารถขับเคลื่อนรถยนต์โดยใช้ไฟฟ้าเพียงอย่างเดียวในระยะทางที่จำกัดโดยไม่มีการปล่อยมลพิษ เจ้าของรถสามารถเติมน้ำมันรถด้วยแก๊สปกติที่ปั๊มแล้วเสียบปลั๊กเข้ากับแหล่งพลังงานไฟฟ้าเพื่อชาร์จแบตเตอรี่ได้

3) Electric vehicles รถยนต์ไฟฟ้า หรือ รถ EV ใช้พลังงานแบตเตอรี่เพียงอย่างเดียวในการขับเคลื่อน รถ EV มีระยะทางเฉพาะที่สามารถเดินทางได้ ก่อนที่จะต้อง "ชาร์จพลังงาน" โดยเสียบปลั๊กเข้ากับแหล่งพลังงานไฟฟ้า ซึ่งทำให้ผู้ใช้งานมี "ความกังวลเรื่องระยะทาง" สำหรับผู้ขับขี่รถ EV มีจำหน่ายตั้งแต่ปี 2010 ได้รับความนิยมเพิ่มขึ้นเนื่องจากเทคโนโลยีแบตเตอรี่ได้เพิ่มช่วงการทำงาน และจำนวนสถานีชาร์จก็เพิ่มขึ้นทั้งในเมืองและตามทางหลวง

4) Natural gas vehicles การเติบโตของก๊าซธรรมชาติในฐานะแหล่งพลังงานยังช่วยกระตุ้นการเพิ่มขึ้นของยานพาหนะที่ใช้ก๊าซธรรมชาติอีกด้วย เช่นเดียวกับรถยนต์ที่ขับเคลื่อนด้วยแก๊ส ยานพาหนะที่ใช้ก๊าซธรรมชาตินั้นใช้ก๊าซธรรมชาติที่ถูกบีบอัดหรือทำให้เป็นของเหลวและมีการปล่อยไอเสียที่สะอาดกว่า

5) Fuel cell electric vehicles เป็นเวลาหลายปีแล้วที่บริษัทรถยนต์ได้ทำการวิจัยและพัฒนาเครื่องยนต์เซลล์เชื้อเพลิงที่ใช้พลังงานไฮโดรเจน (FCEV) ซึ่งมาพร้อมกับข้อดี คือ การปล่อยมลพิษเป็นศูนย์และศักยภาพในการทำงานสูงกว่ารถยนต์ที่ขับเคลื่อนด้วยน้ำมันถึงสามเท่า โดยจะทำงานด้วยการแปลงไฮโดรเจนและออกซิเจนให้เป็นไฟฟ้า ซึ่งจะให้พลังงานแก่มอเตอร์ไฟฟ้า เช่น Clarity Fuel Cell ของ Honda , Nexo Fuel Cell SUV ของ Hyundai และ Mirai ของ Toyota การใช้เทคโนโลยีเซลล์เชื้อเพลิงจากไฮโดรเจนได้กลายเป็นทางเลือกที่ใช้งานได้จริงแทนรถยนต์ที่ใช้้ำมันและแก๊ส อันที่จริง ปัจจุบันมีรถยนต์ 9 รุ่น ที่สามารถวิ่งบนถนนได้จริงจากการใช้เทคโนโลยีเซลล์เชื้อเพลิงนี้ ปัจจัยที่ขัดขวางการนำรถยนต์เซลล์เชื้อเพลิงไปใช้ในวงกว้าง ได้แก่ ต้นทุนของรถยนต์ ตลอดจนเครือข่ายสถานีเติมเชื้อเพลิงไฮโดรเจนที่จำกัด ซึ่งยานยนต์ที่อยู่ใน 5 กลุ่มนี้จะได้รับการลดอัตราภาษีจากรัฐบาลสหรัฐอเมริกาในการใช้งานบนท้องถนน

2.1.2 การทำเหมืองข้อมูล

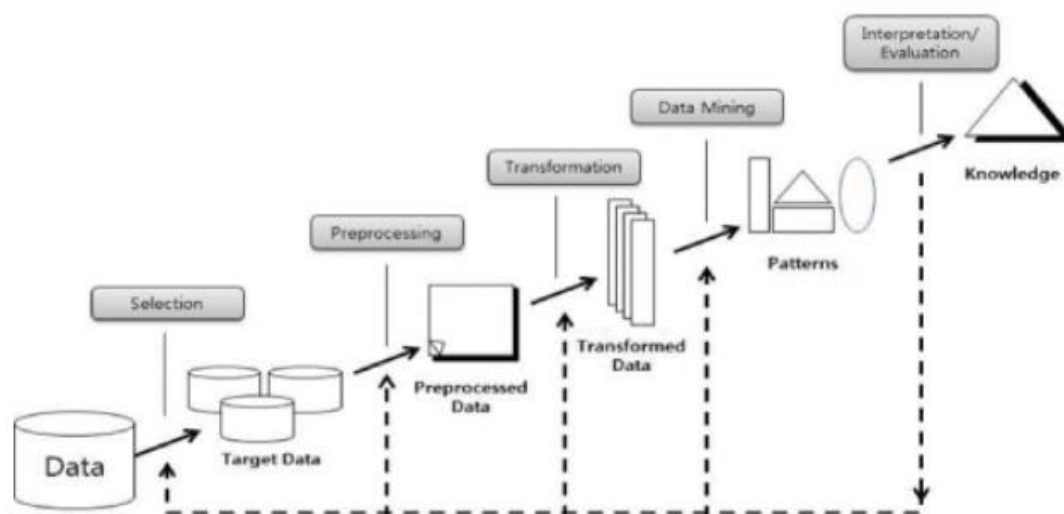
การทำเหมืองข้อมูล (Data Mining) นูชนาฏและคณะ [10] ได้กล่าวถึงการทำเหมืองข้อมูลในการวิจัยการจำแนกความคิดเห็นของคนไทยเกี่ยวกับสื่อออนไลน์โดยการทำเหมืองข้อมูลว่าเป็นกระบวนการสกัดความรู้จากภาษาธรรมชาติที่มีลักษณะข้อมูลที่ไม่มีโครงสร้างนอกจากนี้ไม่มีรูปแบบที่กำหนดไว้ล่วงหน้า 2 ผลลัพธ์ของข้อมูลที่ได้รับดังนั้นจึงไม่มีความแน่นอนในเหตุผลนั้นซ่อนอยู่ในกลุ่มของข้อความที่มีลักษณะภาษาธรรมชาติ การทำเหมืองข้อมูล จำนวนมากจะเป็นลักษณะของการนำเสนอรูปแบบสำหรับการขุดข้อความที่ใช้ประโยชน์จากเอกสารและใช้ความรู้ที่ได้จากการคัดแยกข้อความเหล่านั้นมาปรับปรุงรูปแบบและปัญหาเกี่ยวกับการค้นพบรูปแบบที่ 3 ยังคงเป็นปัญหาอยู่เนื่องจากวิธีการทำเหมืองข้อความส่วนใหญ่มักจะประสบปัญหาปัญหาของหลักภาษาศาสตร์และคำพ้องความหมายที่มีลักษณะซับซ้อนออกมาโดยมีสมมติฐานเกี่ยวกับกลุ่มคำหรือวลีที่ว่าวิธีนี้น่าจะได้ผลดีกว่า ในการวิเคราะห์ข้อความที่ไม่มีโครงสร้างแบบอัตโนมัติจึงเป็นสิ่งที่ท้าทายการประยุกต์ใช้เทคนิคการทำเหมืองข้อความความคิดเห็น โดยนำความคิดเห็นมาวิเคราะห์เพื่อจะรู้ถึงความพอใจกับมันและสกัดออกมาคำที่มีคุณลักษณะต่างกันอาจเป็นประโยชน์ต่อการพิจารณา

วิเคราะห์ที่หลากหลายและแม่นยำยิ่งขึ้น ซึ่งนักวิจัยหลายคนใช้เทคนิคการจำแนกประเภทในการทำเหมืองข้อมูลและแมชชีนเลิร์นนิงเพื่อดึงความรู้ไปใช้เป็นต้นแบบจำลองการจำแนกประเภท

2.1.3 กระบวนการการทำเหมืองความคิดเห็น

กระบวนการการทำเหมืองความคิดเห็น (Opinion Mining Process) คือ การจัดการกับข้อมูลขนาดใหญ่หรือจำนวนที่มีปริมาณมากเป็นพิเศษโดยการนำไปผ่านขั้นตอนต่างๆ อย่างเป็นกระบวนการด้วยส่วนมากจะมีตั้งแต่ 3 ถึง 5 วิธีขึ้นไป จากงานวิจัยของ Igor Mishkovski และคณะ [11] ได้เสนอ 3 ขั้นตอนในกระบวนการการทำเหมืองความคิดเห็น คือ 1) การรวบรวมข้อมูล 2) การเตรียมและแยกข้อมูลด้วยอัลกอริทึม 3) การประเมินคุณภาพข้อมูลที่ได้ ซึ่งใช้ในการวิจัย การทำเหมืองข้อมูลเพื่อการศึกษา: กรณีศึกษาการทำนายการลาออกของนักเรียนในระดับอุดมศึกษา ส่วน George Harrison และคณะ [12] ได้เสนอกระบวนการการทำเหมืองความคิดเห็นไว้ทั้งหมด 4 ขั้นตอน คือ 1) การเลือกข้อมูล 2) การเตรียมข้อมูล 3) การทำดัชนีและเหมืองความคิดเห็น 4) การประเมินผลและวัด ประสิทธิภาพ ซึ่งอยู่ในงานวิจัย การปรับปรุงกระบวนการการทำเหมืองความคิดเห็น จากงานวิจัยของทั้งสองท่านถึงแม้กระบวนการจะมีความแตกต่างกันแต่จะขาดขั้นตอนการวัดผลและประเมินผลไปไม่ได้ ซึ่งสำหรับการวิจัย การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น ในครั้งนี้ผู้วิจัยได้เลือกกระบวนการการทำเหมืองความคิดเห็นเอาไว้ 5 ขั้นตอนซึ่งเหมือนกับงานวิจัยของ ศราวุฒิ เกิดถาวร ที่ได้ใช้กระบวนการในการทำเหมืองความคิดเห็นเพื่อวิจัยของ Han [13] การคัดเลือกคุณลักษณะความรู้สึกของคนไทยต่อโรคโควิด 19 บนสื่อสังคมออนไลน์ด้วยเหมืองข้อความ โดยมีขั้นตอนดังต่อไปนี้ 1) การเลือกข้อมูล 2) การเตรียมข้อมูล 3) การทำดัชนีข้อมูล 4) การทำเหมืองความคิดเห็น 5) การแปลผลและการประเมินผล



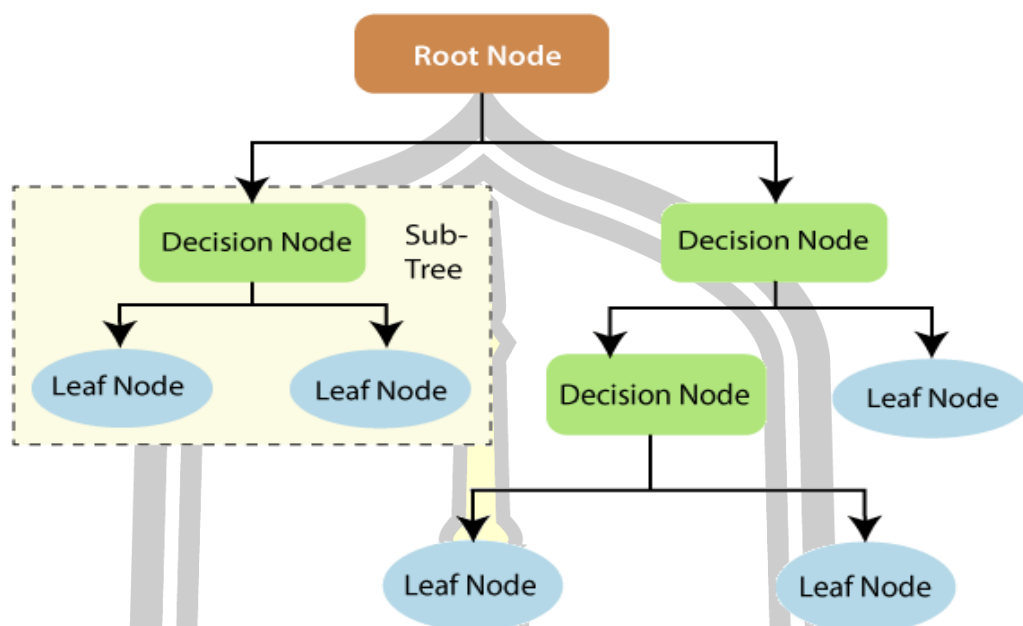


ภาพประกอบที่ 2.1 กระบวนการในการทำเหมืองความคิดเห็นงานวิจัย ของ Han

2.1.4 เทคนิคที่ใช้ในการวิจัย

สำหรับการวิจัย การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น ผู้วิจัยได้ใช้แบบจำลองในการทำเหมืองข้อความเพื่อวิเคราะห์ความคิดเห็นทั้งหมด 4 เทคนิค คือ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (C4.5) เทคนิคเนิร์ฟเบย์ (Naïve-Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และ เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) จากนั้นใช้หลักการของ 10-Fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดเพื่อเรียนรู้และชุดข้อมูลเพื่อทดสอบ โดยจะวัดประสิทธิภาพของแบบจำลองด้วย ค่าความกว้าง (F-Measure) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall)

2.1.4.1 เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (C4.5) เป็นเทคนิคเทคนิคใช้หลักการสร้างโหนดราก (root node) จากการเลือกคุณลักษณะด้วย grain ratio แล้วทำการเลือกโหนดลูกตามลำดับจนถึงโหนดใบ (leaf node) จุดเด่นของเทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 เป็นเทคนิคที่มีขั้นตอนการทำงานที่ไม่ซับซ้อนและมีความรวดเร็วในการประมวลผล สามารถใช้งานกับข้อมูลที่มีความซับซ้อน สามารถแสดงความสัมพันธ์ระหว่างตัวแปรได้เป็นอย่างดี และแบบจำลองที่ได้จากการสร้างด้วยเทคนิคนี้สามารถแปลผลได้ง่าย แต่อย่างไรก็ตามเทคนิคนี้ยังมีจุดด้อยเกี่ยวกับปัญหา Overfitting ได้ [14]



ภาพประกอบที่ 2.2 เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5

2.1.4.2 เทคนิคเนอ์ฟเบย์ เป็นเทคนิคที่ใช้ทฤษฎีทางด้านสถิติโดยนำความน่าจะเป็นมาประเมินความไม่แน่นอนของให้เป็นตัวเลข ได้กล่าวถึงความน่าจะเป็นของเหตุการณ์ที่เกิดขึ้น (A) ถ้ามี เหตุการณ์อีกเหตุการณ์หนึ่งเกิดมาแล้ว (B) สามารถ เขียนให้อยู่ในรูปอย่างง่าย ซึ่งข้อดีของเทคนิคเนอ์ฟเบย์ คือ เทคนิคเข้าใจง่าย มีความรวดเร็วในการประมวล และมีความแม่นยำในการประมวลผลกับชุดข้อมูลขนาดใหญ่ โดยเทคนิคนี้เหมาะกับการวิจัยที่ต้องการจำแนกประเภทข้อมูลหลาย ๆ ชั้น [15] ดังสมการที่ 1

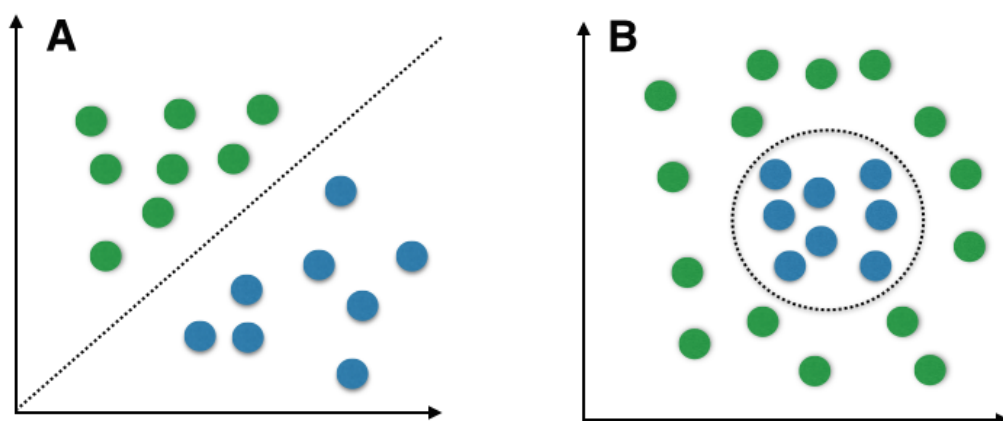
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

$P(A|B)$ คือ ความน่าจะเป็นที่เหตุการณ์ A จะเกิดขึ้น ถ้าเหตุการณ์ B เกิดขึ้นแล้ว

$P(B|A)$ คือ ความน่าจะเป็นที่เหตุการณ์ B จะเกิดขึ้น ถ้าเหตุการณ์ A เกิดขึ้นแล้ว

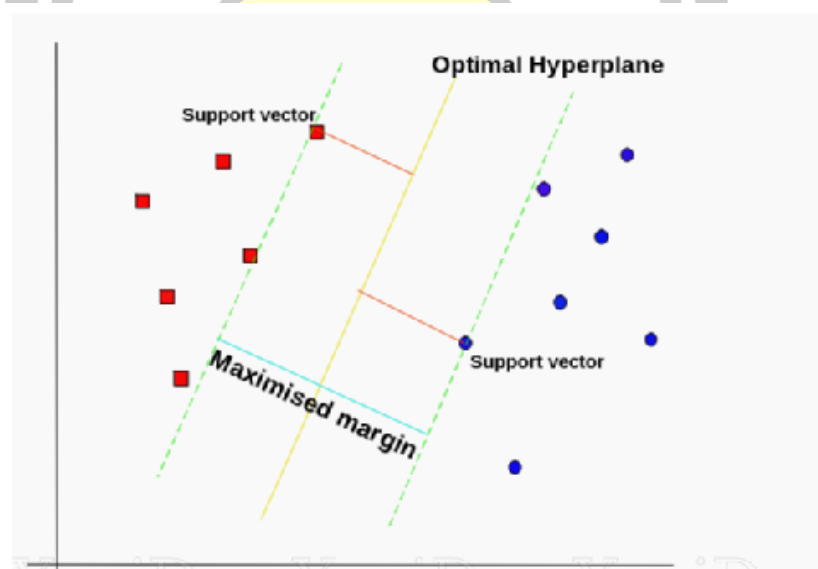
$P(A)$ คือ ความน่าจะเป็นที่จะเกิด เหตุการณ์ A

$P(B)$ คือ ความน่าจะเป็นที่จะเกิด เหตุการณ์ B



ภาพประกอบที่ 2.3 เทคนิคเนอรัลเน็ต

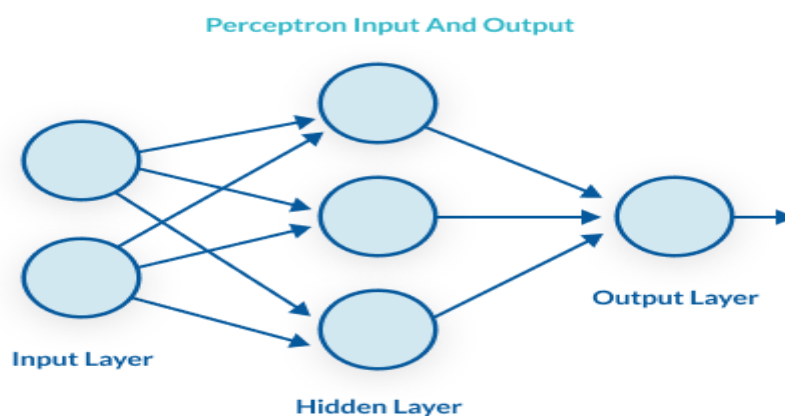
2.1.4.3 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เป็นเทคนิคการเรียนรู้ของเครื่องแบบจำแนกประเภทที่มีประสิทธิภาพสูง SVM ทำงานโดยการสร้างเส้นแบ่ง (hyperplane) ระหว่างข้อมูลสองคลาส โดยเส้นแบ่งนี้จะทำให้ข้อมูลของทั้งสองคลาสอยู่ห่างจากเส้นแบ่งมากที่สุด SVM มีข้อดีคือ มีประสิทธิภาพสูงสำหรับชุดข้อมูลที่มีขนาดใหญ่ สามารถจัดการกับข้อมูลส่วนเกินหรือข้อมูลที่ไม่เกี่ยวข้องการประมวลผลได้ดี และสามารถนำไปใช้กับข้อมูลที่มีหลายมิติได้ จึงเหมาะกับการทำเหมืองข้อมูล [16]



ภาพประกอบที่ 2.4 เทคนิค Support Vector Machine

2.1.4.4 เทคนิคเพอร์เซปตรอนหลายชั้น เป็นโมเดลการเรียนรู้ของเครื่องแบบจำแนกประเภทที่ใช้โครงข่ายประสาทเทียมแบบฟีดฟอร์เวิร์ด (feedforward neural network) MLP

ประกอบด้วยชั้นอินพุต (input layer) ชั้นซ่อน (hidden layer) และชั้นเอาต์พุต (output layer) โดยชั้นอินพุตจะรับข้อมูลเข้า ชั้นซ่อนจะทำการประมวลผลข้อมูล และชั้นเอาต์พุตจะส่งออกผลลัพธ์ MLP ทำงานโดยการคำนวณผลรวมของค่าน้ำหนัก (weight) ที่เชื่อมต่อกับเซลล์ประสาท (neuron) ในแต่ละชั้น จากนั้นนำค่ารวมมาผ่านฟังก์ชันการเปิดใช้งาน (activation function) เพื่อหาผลลัพธ์ ข้อดีของ MLP คือ สามารถเรียนรู้ความสัมพันธ์ที่ซับซ้อนของข้อมูลได้ สามารถทำงานกับข้อมูลที่มีมิติสูงได้ และสามารถปรับแต่งให้เหมาะกับข้อมูลเฉพาะได้ [17]



ภาพประกอบที่ 2.5 เทคนิคเพอร์เซปตรอนหลายชั้น

2.2 งานวิจัยที่เกี่ยวข้อง

ในการการวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็นครั้งนี้ ผู้วิจัยได้ทำการ การทบทวนวรรณกรรม (Literature Review) ของงานนักวิจัยหลายๆท่านโดยพบว่า มีนักวิจัยหลายๆท่านได้เสนอการวิจัยที่ใช้เทคนิคแบบเดียวกัน และได้ผลการวิจัยออกมาเป็นที่น่าสนใจ ธีรยุทธ และคณะ[18] ได้เสนอ กระบวนการคัดเลือกคุณลักษณะสำหรับเพิ่มประสิทธิภาพในการจำแนกความคิดเห็นของลูกค้า เกี่ยวกับร้านอาหารข้อมูลรวบรวมจากเว็บไซต์ wongnai.com จำนวน 4,487 ข้อความ ได้ใช้เทคนิค 3 ประการในการเลือกคุณสมบัติข้อความ ได้แก่ Chi-Square, Information Gain และ Information Gain Ratio เพื่อวัดประสิทธิภาพของเทคนิคการเลือกคุณสมบัติและใช้ Naive Bayes, Support Vector Machine, K-Nearest Neighbor และ C4.5 สำหรับการจัดประเภทให้ออนบวกรนอกจากนี้ยังมีการใช้ Cross Validation 10 เท่าเพื่อแบ่งข้อมูลออกเป็นชุดการเรียนรู้และวัดความถูกต้อง (Accuracy) ความแม่นยำ (Precision) และการเรียกคืน (Recall) จากการทดลองพบว่าเทคนิคการเลือกฟีเจอร์ Information Gain ร่วมมือกับเทคนิค

Naive Bayes ให้ผลลัพธ์ที่ดีที่สุดในการจำแนกประเภทของความคิดเห็นโดยความถูกต้องคือ 89.08% ความแม่นยำเท่ากับ 89.12% และการเรียกคืนเท่ากับ 89.10%

งานวิจัยการจัดหมวดหมู่ภาษาไทย ของ ราชวิทย์ และคณะ[19] ได้เสนอ การจำแนกกลุ่มคำถามอัตโนมัติบนกระดานสนทนาโดยใช้เทคนิคเหมืองข้อความเพื่อศึกษาและเปรียบเทียบประสิทธิภาพการจำแนกข้อความ 3 เทคนิควิธีได้แก่ การหาเพื่อนบ้านใกล้สุด ต้นไม้ตัดสินใจ การเรียนรู้แบบอย่างง่าย ผลการทดลองพบว่าเทคนิคการหาเพื่อนบ้านใกล้สุด มีประสิทธิภาพในการจำแนกกลุ่มคำถามได้ดีที่สุด มีค่าความถูกต้องเท่ากับ 89% ค่าความเที่ยงเท่ากับ 90% และค่าความระลึก 89% ส่วนงานวิจัยของนิเวศและคณะ[20] ได้เสนอการพัฒนาประสิทธิภาพการจัดหมวดหมู่เอกสารภาษาไทยแบบอัตโนมัติ ซึ่งได้ลดคุณลักษณะข้อมูลโดยใช้ค่าเกนความรู้ (Information Gain) และเทคนิคชัพพอตเว็กเตอร์แมทซิน ผลการทดลอง พบว่า อัลกอริทึมชัพพอร์ทเวกเตอร์แมทซิน ให้ประสิทธิภาพสูงสุด คือ 94.3% ด้านการลดขนาดคุณลักษณะสามารถลดลงได้ 90% โดยไม่ส่งผลต่อประสิทธิภาพการจัดหมวดหมู่

สฤติโชคและคณะ [21] ได้เสนอการจำแนกพฤติกรรมการขับซึ่รถโดยสารสาธารณะโดยใช้วิธีการสกัดข้อความและเทคนิคชัพพอตเว็กเตอร์แมทซิน เพื่อจำแนกประเภทข้อร้องเรียนแบบอัตโนมัติที่เกี่ยวกับพฤติกรรมการขับซึ่และคุณภาพการบริการ จำนวน 4 คลาส ผลการทดลองพบว่าตัวแบบโครงข่ายประสาทเทียมสามารถจำแนกข้อความพฤติกรรมการขับรถโดยสารสาธารณะได้ค่าความแม่นยำสูงสุดซึ่งมีค่าเท่ากับ 90.23% ค่าความแม่นยำ เท่ากับ 91.9% ค่าความระลึก เท่ากับ 90.2% และค่าความถ่วงเท่ากับ 90.5% งานวิจัยนี้สามารถนำไปประยุกต์ใช้สร้างระบบจำแนกหมวดหมู่เกี่ยวกับพฤติกรรมการขับรถโดยสารสาธารณะและระบบจำแนกหมวดหมู่เอกสารอัตโนมัติได้

Cambria และคณะ [22] ได้นำเสนอการทำเหมืองข้อมูลความรู้สึกของผู้ใช้สมาร์ทโฟนบนระบบปฏิบัติการ Android และ iOS โดยใช้กระบวนการในการตัดคำและเลือกเฉพาะคำที่แสดงถึงเจตนาในการใช้งานเช่น ชอบ ไม่ชอบ ดี ไม่ดี โดยพิจารณาจากภาษาธรรมชาติ (NLP) เท่านั้น และใช้เทคนิคแบบ Machine Learning การเรียนรู้ของเครื่อง และเทคนิค Support Vector Machine (SVM) มาจำแนกรู้สึกของผู้ใช้งานสมาร์ทโฟนทั้ง 2 ระบบปฏิบัติการโดยพบว่า มีค่าความถูกต้องที่ 75% ความแม่นยำเท่ากับ 79.32% และการเรียกคืนเท่ากับ 82.10% สอดคล้องกับงานวิจัยของ Bing Liu [23] ในการ การสำรวจความคิดเห็นและการวิเคราะห์ความคิดเห็น ว่าถ้าใช้เทคนิคในการแยกคำโดยใช้ภาษาธรรมชาติ (NLP) ที่เจาะจงเลือกคำที่เป็นคำเฉพาะในปริมาณที่น้อย จะทำให้

ประสิทธิภาพของเทคนิคที่นำมาใช้ในการสร้างเหมือนความคิดเห็นนั้นค่อนข้างลดลง ซึ่งในงานวิจัยสำรวจความคิดเห็นของผู้ใช้งานโทรศัพท์ iPhone จึงได้มีการเพิ่มส่วนในการพิจารณาให้กับภาษาธรรมชาติ 2 ส่วน โดยการ นิยามความคิดเห็นและการสรุปความคิดเห็นใหม่ เพื่อให้เป็นภาษาธรรมชาติที่สามารถนำไปใช้งานได้จริงในขั้นตอนการทำเหมือนความคิดเห็น

Alexander และคณะ [24] ได้เสนอกระบวนการทำเหมือนความคิดเห็นในงานวิจัย การวิเคราะห์ความรู้สึกและเหมืองข้อมูลจากผู้ใช้งาน Twitter ด้วยการแยกคำออกเป็น 2 กลุ่ม คือ กลุ่มที่มีความหมายและกลุ่มที่ไม่มีความหมาย โดยการใช้ภาษาธรรมชาติแบบ NPL และ NPS จากนั้นใช้เทคนิค Support Vector machine (SVM) และ Naive Bayes ในการทำเหมือนความคิดเห็น ผลที่ได้คือมีความถูกต้องถึง 89.99% ความแม่นยำเท่ากับ 88.89% และค่าความถ่วง (F0.5) 0.60 โดยเฉพาะเมื่อทดลองกับชุดข้อมูลที่มีขนาดใหญ่พิเศษ ระดับค่าความถูกต้องยังสามารถเพิ่มขึ้นได้อีก ซึ่งในการวิจัยครั้งนี้ได้ใช้ข้อมูลทั้งหมด 300,000 ความคิดเห็นโดยเป็นการเลือกแบบสุ่มจากผู้ใช้งานบน Twitter

มาสวี่ร์ มาศดิศรโชติ [25] ได้สรุปกระบวนการในการทำเหมือนความคิดเห็น 3 ระดับ ดังนี้ ระดับเอกสาร ระดับประโยค และระดับคุณลักษณะของวัตถุ ซึ่งทั้ง 3 ระดับจะต้องมีองค์ประกอบที่เป็นกระบวนการลำดับ และจะต้องวัดประสิทธิภาพด้วยการหาค่าความถ่วง (F-Measure) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ไม่ว่าจะใช้เทคนิคใดในการทำเหมือนความคิดเห็นก็ตาม โดยเฉพาะภาษาไทยที่เป็นภาษาธรรมชาติ จะต้องมีการวิเคราะห์เชิงองค์ประกอบละเอียดมาก เช่น คำว่า การประปา สามารถแยกออกเป็น คำว่า การ กับ คำว่า ประปา ต่อมาต้องวิเคราะห์ระดับโครงสร้างทางภาษาหรือวากยสัมพันธ์ คือ การวิเคราะห์คำตามหน้าที่ของคำเพื่อดูว่าประโยคที่ประกอบด้วยหลายๆคำมีโครงสร้างทางภาษาเป็นอย่างไร คำใดเป็นประธาน กริยา กรรม และเมื่อผ่านขั้นตอนนี้ต้องมีการวิเคราะห์ระดับความหมายอีกรอบเพื่อหาความหมายของประโยครวมถึงต้องวิเคราะห์ประโยคข้างเคียงกันด้วย และขั้นตอนสุดท้ายคือการวิเคราะห์ทางปฏิบัติเพื่อแปลความตั้งใจของผู้แสดงความคิดเห็นว่าต้องการจะสื่อความหมายอะไร

Magdalini และคณะ [26] ได้สรุปให้เห็นถึงความสำคัญของคุณลักษณะในการทำเหมือนความคิดเห็น (Feature opinion mining) ในแบบต่าง ๆ ไว้ในงานวิจัย การขุดและการจัดอันดับความคิดเห็นตามคุณสมบัติ โดยให้ความสำคัญกับกระบวนการและเครื่องมือในการคุณลักษณะที่สำคัญของความรู้สึกจากข้อมูลที่มีจำนวนมาก โดยการกำหนดต้องการไว้ล่วงหน้า เช่น ลูกค้าน่าจะมีความรู้สึกอย่างไรหลังจากได้ใช้ผลิตภัณฑ์ หรือ ประชาชนจะรู้สึกอย่างไรถ้ามีประธานาธิบดีที่ตนไม่ได้

เลือกเข้ามา โดยเสนออัลกอริทึม 3 ตัว ได้แก่ อัลกอริทึมการนับคำคุณศัพท์สูง อัลกอริทึมคะแนนความคิดเห็นสูงสุด และส่วนขยายสำหรับสื่อสังคมออนไลน์อื่นๆ ซึ่งใช้งานคู่กับการหาค่าส่วนกลับเอกสาร (TF-IDF) โดยผลการวิจัยพบว่าอัลกอริทึมที่นำมาใช้มีความถูกต้องถึง 87%

งานวิจัยของ Shaik Thanveer และคณะ [27] งานวิจัยนี้เริ่มต้นด้วยการให้นิยามการวิเคราะห์ความรู้สึก (sentiment analysis) และการทำเหมืองความคิดเห็น (opinion mining) ว่าเป็นกระบวนการในการระบุและวัดอารมณ์หรือทัศนคติของผู้เขียนในข้อความ การวิเคราะห์ความรู้สึกและการทำเหมืองความคิดเห็น โดยการใช้เทคนิคในการวิจัย 5 เทคนิคได้แก่ พจนานุกรมคำศัพท์ที่มีค่าความรู้สึก เทคนิคชัพพอดเว็กเตอร์แมทซิน เทคนิคตัวกรองคำศัพท์ เทคนิคการเรียนรู้เชิงลึก และเทคนิคต้นไม้ตัดสินใจ ในการวิจัยการวิเคราะห์ความรู้สึกและการขุดความคิดเห็นเกี่ยวกับข้อมูลการศึกษาโดยใช้แบบสำรวจ โดยผลการวิจัยพบว่า เทคนิคการเรียนรู้เชิงลึก ให้ผลลัพธ์ที่ดีที่สุดโดยมีค่าความแม่นยำที่ร้อยละ 98.87 ค่าความระลึกที่ร้อยละ 97.71 และค่าความถ่วงที่ร้อยละ 97.32

งานวิจัยของ Khafaga และคณะ [28] ได้ใช้ เทคนิคเทคนิคต้นไม้ตัดสินใจ เทคนิคป่าไม้แบบสุ่ม และเทคนิคเพอร์เซปตรอนหลายชั้น ในการวิเคราะห์และพยากรณ์สภาวะการไหลเวียนในช่องท้องของผู้ป่วยโรคถุงลมโป่งพอง พบว่า เทคนิคเพอร์เซปตรอนหลายชั้นให้ผลลัพธ์ที่ดีที่สุดโดยมีค่าความแม่นยำที่ร้อยละ 78.03

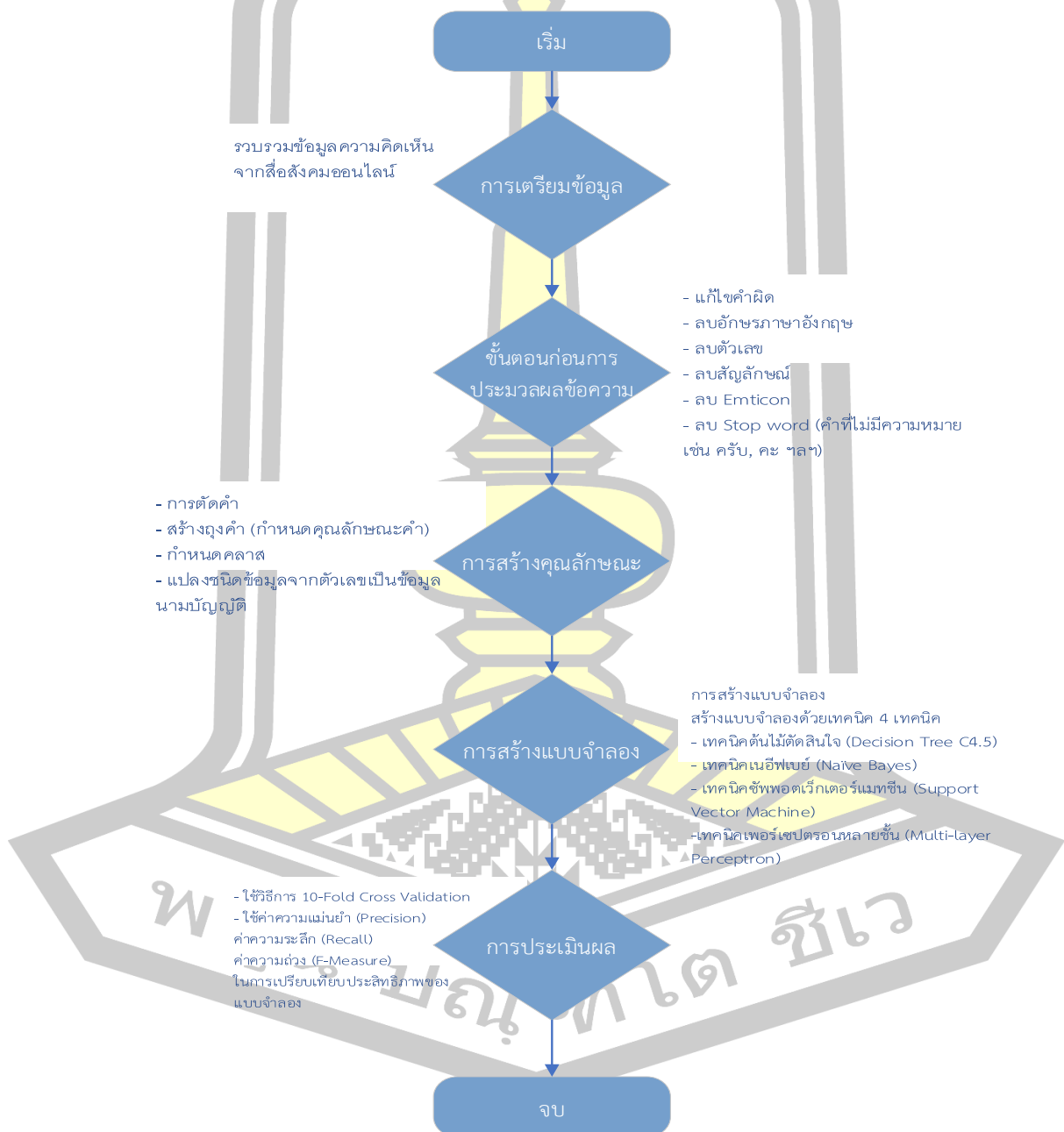
งานวิจัยของ Davis และคณะ [29] ได้สร้างเหมืองข้อความโดยการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ด้วยการสร้างแบบจำลองเพื่อวิเคราะห์ข้อความ 2 ตัวได้แก่ แบบจำลองการวิเคราะห์ความหมายแฝง Latent Semantic Analysis (LSA) และแบบจำลองการเร่งสถานการณ์ผ่านการสร้างแบบจำลองอัตโนมัติ Scenario Acceleration through Automated Modelling (SAAM) เพื่อจำแนกความคิดเห็นและจัดกลุ่มของข้อมูล โดยการนำเอาเทคนิคต้นไม้ตัดสินใจ มาใช้ในการทดสอบแบบจำลองทั้ง 2 ตัว ผลการวิจัยพบว่า แบบจำลองการวิเคราะห์ความหมายแฝง มีค่าความแม่นยำที่ร้อยละ 90.00% ในการจัดกลุ่มของข้อมูล ส่วนแบบจำลองการเร่งสถานการณ์ผ่านการสร้างแบบจำลองอัตโนมัติ มีค่าความแม่นยำที่ร้อยละ 89.13

พูน ปณ จิต ชเว

บทที่ 3

วิธีดำเนินการวิจัย

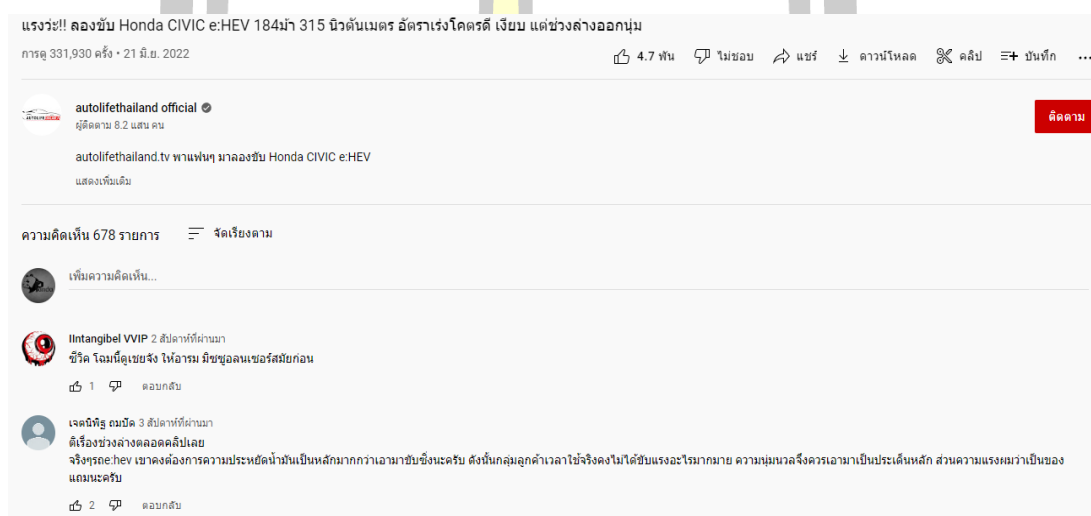
ในการวิจัยการวิเคราะห์ความคิดเห็นของผู้ใช้งานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น ครั้งนี้ ผู้วิจัยได้กำหนดวิธีดำเนินการวิจัยออกเป็น 5 ขั้นตอน ดังภาพที่ 3.1



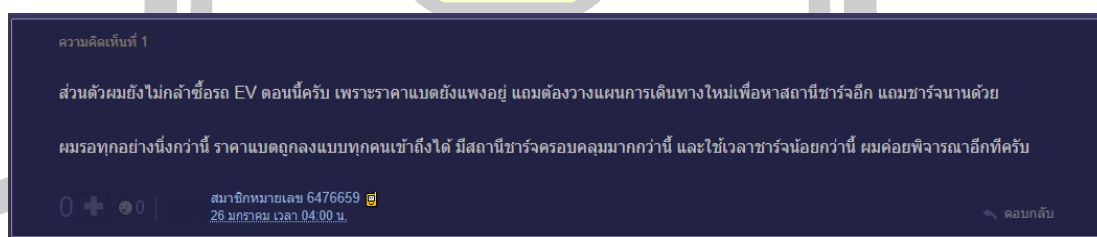
ภาพประกอบที่ 3.1 ผังงานการดำเนินการวิจัย

3.1 การเตรียมข้อมูล

ในขั้นตอนนี้เป็นการเลือกข้อมูลที่จะนำมาทำการวิจัยการวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมือนความคิดเห็นโดยการใช้ความคิดเห็นของคนสื่อสังคมออนไลน์หรือสื่อสังคมออนไลน์ต่าง ๆ เช่น Twitter, Facebook, Pantip และ Tiktok โดยการเข้าไปในช่องหรือกระทู้ต่าง ๆ ที่มีส่วนเกี่ยวข้องกับยานยนต์พลังงานทางเลือก และเลือกเอาสิ่งที่ผู้คนแสดงความคิดเห็น (Comment) ที่ไว้ โดยใช้หลักการและวิธีการดังต่อไปนี้



ภาพประกอบที่ 3.2 ตัวอย่างความคิดเห็นจาก Youtube



ภาพประกอบที่ 3.3 ตัวอย่างความคิดเห็นจาก Pantip

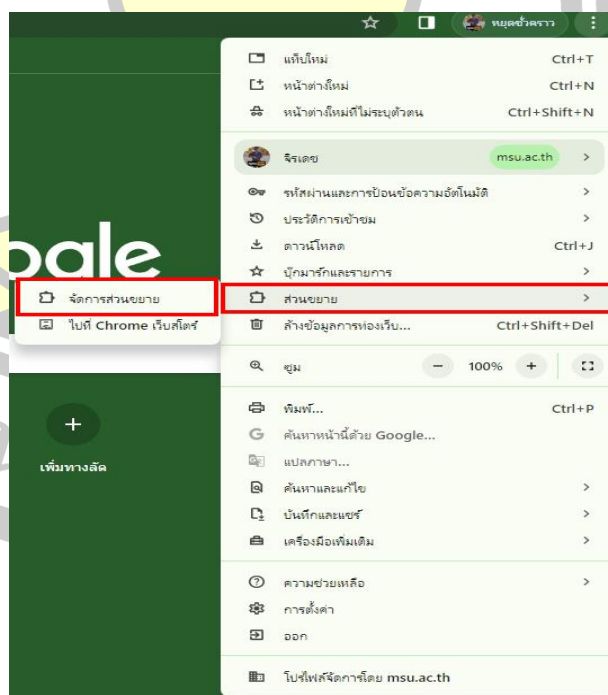
ผู้วิจัยได้ให้ความสำคัญกับขั้นตอนการเลือกข้อมูลเป็นอย่างมาก เพื่อให้การวิจัยครั้งนี้เป็นไปตามวัตถุประสงค์และได้ความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกที่เป็นข้อมูลจริง โดยการศึกษางานวิจัยจากหลาย ๆ แหล่ง เช่น สมศักดิ์ ศรีสุวรรณ และสมัย ศรีสวย [30] ได้คัดเลือกความคิดเห็น 1,056 ความคิดเห็น สำหรับการวิเคราะห์เหมือนความคิดเห็นโดยใช้เทคนิคการสกัดคำหลังจากที่เก็บรวบรวมความคิดเห็นทั้งหมด 13,564 ความคิดเห็น โดยพิจารณาความคิดเห็นที่มี

มากกว่า 20 คำ ขึ้นไป ส่วนงานวิจัยของ วีรศักดิ์ ฟองเงิน และสุรพงษ์ เพ็ชรหาญ [31] คัดเลือกความคิดเห็น 1,056 ความคิดเห็น จากความคิดเห็นที่เก็บรวบรวมมาทั้งหมด 18,790 ความคิดเห็น โดยพิจารณาจากความคิดเห็นที่มีมากกว่า 15 คำ ขึ้นไป ในการวิจัย การวิเคราะห์เหมือนความคิดเห็นเครือข่ายเฟสบุ๊กโดยใช้เทคนิคเหมืองข้อมูล และงานวิจัยของ วสวัตดี อินทร์แปลง และคณะ [32] ได้คัดเลือกความคิดเห็น 3,123 ความคิดเห็น ในการวิจัยการวิเคราะห์ความคิดเห็นต่อเกมมือถือนักบ๊วด้วยเหมืองข้อความ โดยพิจารณาความคิดเห็นที่มี 10 คำ ขึ้นไป

ดังนั้นเพื่อให้งานวิจัยครั้งนี้เป็นไปตามวัตถุประสงค์ผู้วิจัยจึงรวบรวมเอาความคิดเห็นทั้งหมด 7,680 ความคิดเห็น โดยใช้เครื่องมือโปรแกรม Web Scraper ซึ่งเป็นโปรแกรมฟรี ที่สามารถใช้งานบน Google Chrome ในการรวบรวมความคิดเห็นทั้งหมด และใช้ภาษา Python ในการพิจารณา คำที่อยู่ในความคิดเห็น โดยพิจารณาเลือกเอาเฉพาะความคิดเห็นที่มีคำมากกว่า 15 คำ ขึ้นไป เมื่อผ่านขั้นตอนนี้ จึงได้ความคิดเห็นที่ใช้ในการวิจัยครั้งนี้ 2,265 ความคิดเห็น โดยมีขั้นตอนดังนี้

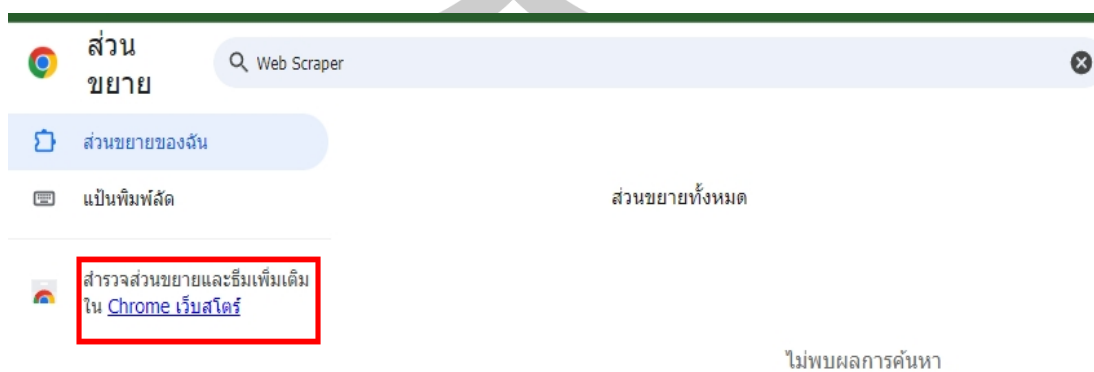
3.1.1 การติดตั้งโปรแกรม

การติดตั้งโปรแกรม Web Scraper สามารถทำได้โดยการเพิ่มส่วนขยายของโปรแกรม Google Chrome โดยเริ่มจาก เข้าไปตรวจสอบ ส่วนขยาย และเข้าไปที่เมนู จัดการส่วนขยาย เพื่อเข้าไปเพิ่มโปรแกรม Web Scraper ดังภาพที่ 3.4



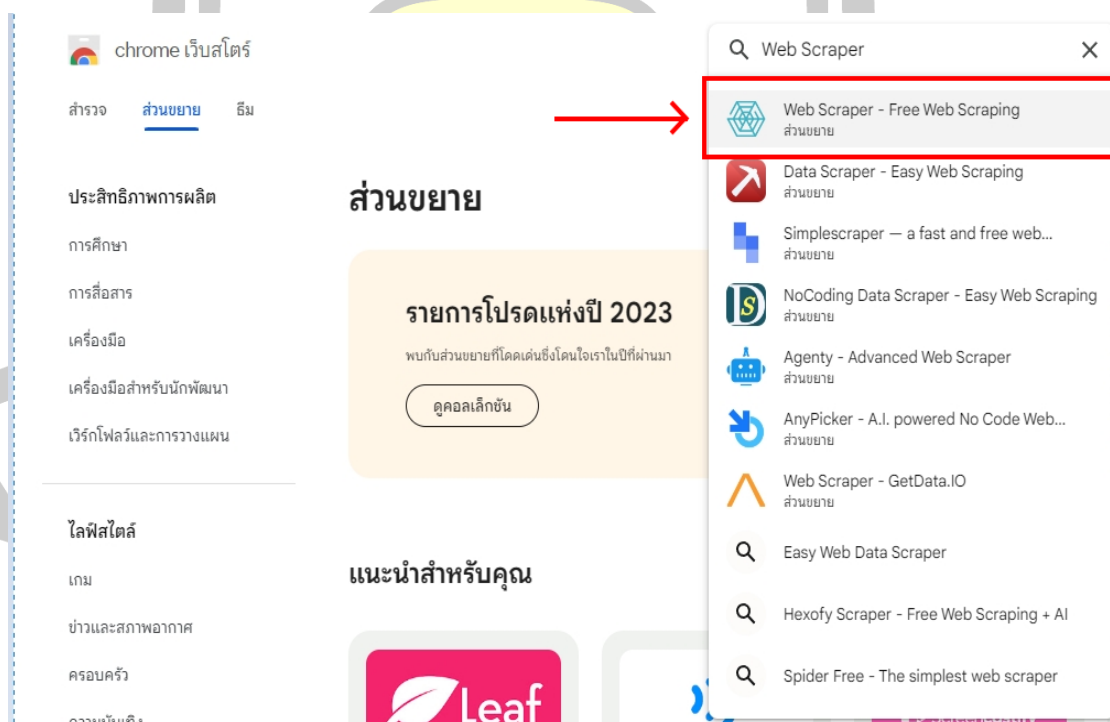
ภาพประกอบที่ 3.4 เมนูส่วนขยายของ Chrome

เมื่อเข้าสู่หน้าเพิ่มส่วนขยายแล้ว ให้กดเข้าไปที่เมนู Chrome เว็บสโตร์ เพื่อค้นหาโปรแกรม Web Scraper สำหรับติดตั้ง ดังภาพที่ 3.5



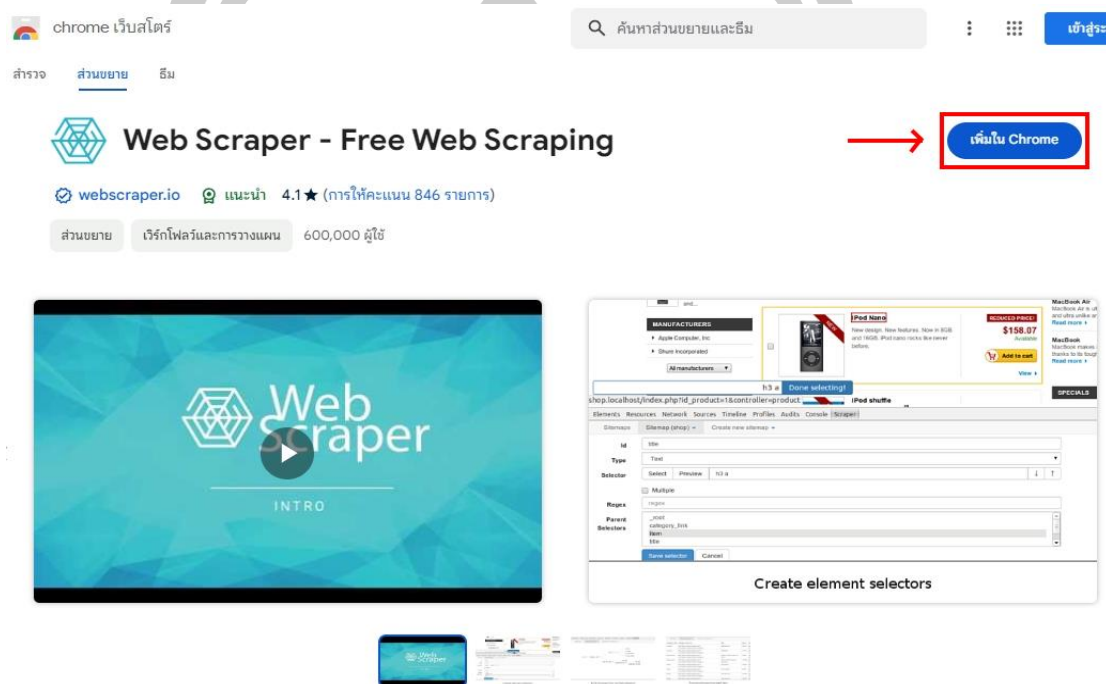
ภาพประกอบที่ 3.5 หน้าต่างจัดการส่วนขยาย

เมื่อเข้ามาหน้าต่าง Chrome เว็บสโตร์ ให้ใช้ ช่องค้นหาส่วนขยายและธีม โดยการพิมพ์ชื่อโปรแกรม Web Scraper เข้าไปเพื่อค้นหาได้โปรแกรมที่ใช้ในการติดตั้ง ซึ่งในขั้นตอนนี้จะพบโปรแกรมหลายโปรแกรมที่มีชื่อหรือความสามารถในการทำงานที่คล้ายกันให้เลือกโปรแกรมชื่อว่า Web Scraper – Free Web scraping ดังภาพประกอบที่ 3.6

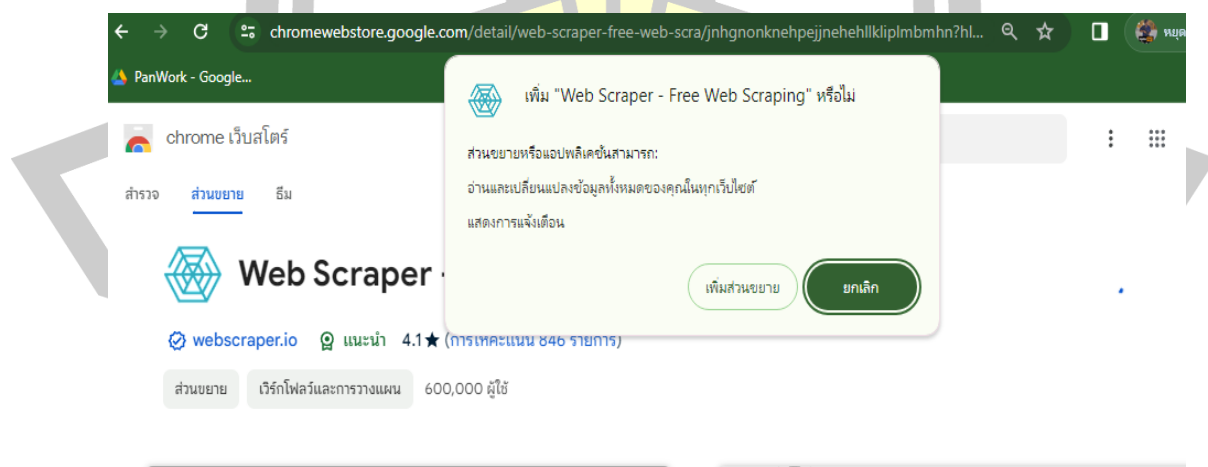


ภาพประกอบที่ 3.6 ค้นหาโปรแกรม Web Scraper สำหรับติดตั้ง

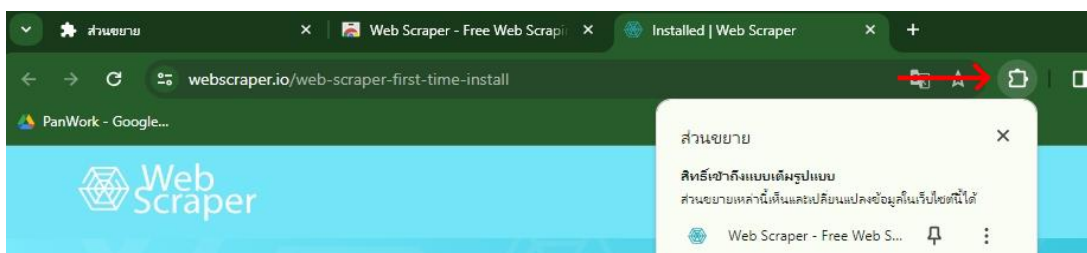
เมื่อเลือกโปรแกรม Web Scraper – Free Web scraping จะเข้าสู่หน้าแนะนำการใช้งาน ให้กดที่ปุ่ม เพิ่มใน Chrome เพื่อติดตั้งเป็นส่วนขยาย โดยผู้ใช้งานจะต้องยืนยันการเพิ่มส่วนขยายก่อน จากขั้นตอนนี้อาจจะใช้เวลาประมาณ 3-5 นาที ในการติดตั้ง หลังจากนั้นก็สามารถตรวจสอบว่าได้ทำการติดตั้งโปรแกรมหรือยังได้ที่ เมนูส่วนขยาย หลังช่อง Address bar ดังภาพที่ 3.7 ภาพที่ 3.8 และภาพที่ 3.9



ภาพประกอบที่ 3.7 หน้าต่างติดตั้งโปรแกรม Web Scraper



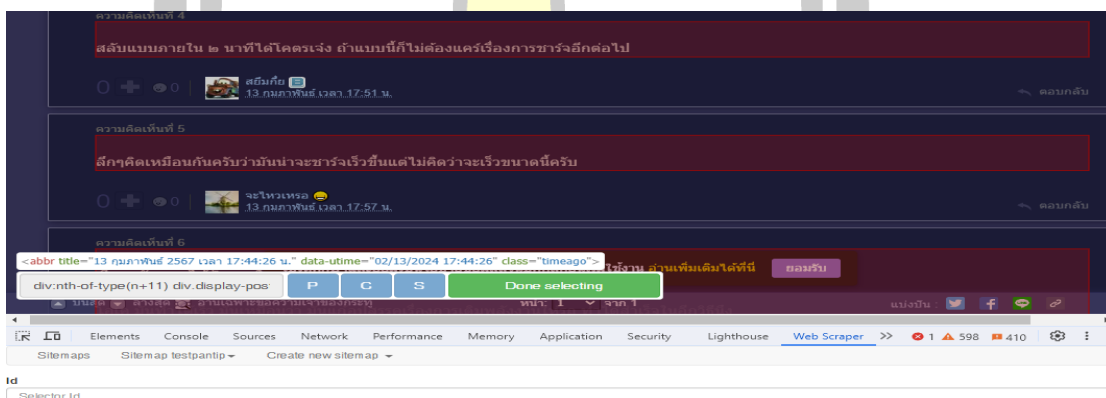
ภาพประกอบที่ 3.8 ยืนยันการติดตั้งโปรแกรม



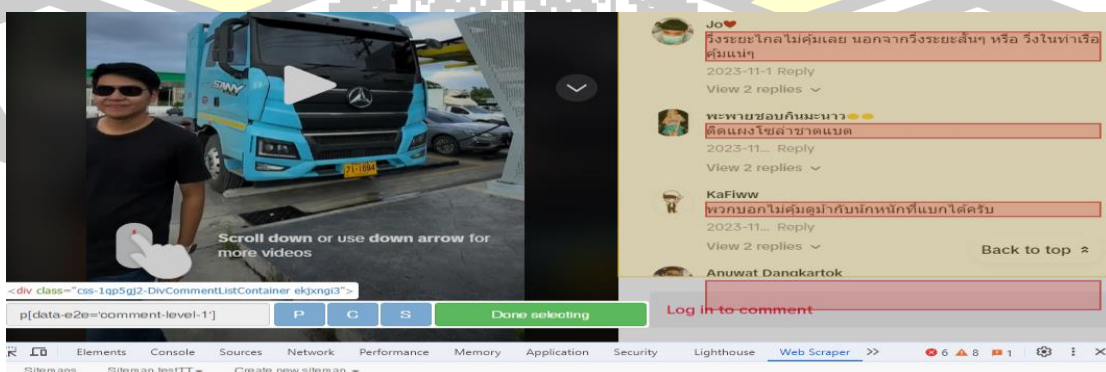
ภาพประกอบที่ 3.9 ตรวจสอบหลังจากการติดตั้ง

3.1.2 การใช้งานโปรแกรมในการเก็บรวบรวมความคิดเห็น

การใช้งานโปรแกรม Web Scraper เพื่อรวบรวมความคิดเห็นจากสื่อสังคมออนไลน์ โดยการเข้าไปยังหน้าเว็บที่มีความคิดเห็นที่ได้แสดงไว้จากนั้นเปิดหน้าต่างเครื่องมือ เพื่อใช้งานโปรแกรมด้วยการกดปุ่ม F12 แล้วเลือกแท็บ Web Scraper ดังภาพที่ 3.12 โดยข้อดีของโปรแกรม Web Scraper คือ สามารถรวบรวมความคิดเห็นที่ถูกแสดงไว้ได้ครั้งละมากและรวดเร็ว โดยใช้วิธีการเดียวกันกับทุกเว็บไซต์ ดังภาพประกอบที่ 3.10 และภาพประกอบที่ 3.11



ภาพประกอบที่ 3.10 การใช้งานโปรแกรม Web Scraper กับเว็บไซต์ Pantip



ภาพประกอบที่ 3.11 การใช้งานโปรแกรม Web Scraper กับเว็บไซต์ Tiktok

[spin9] รีวิว ORA Good Cat เหมือนไฟฟ้า ขับไม่ยาก แต่ใช้ไม่ง่าย! รุ่นท็อป 1.199 ล้านบาท

spin9 8.61 แสน คน

1.1 หมื่น

การดู 9.8 แสน ครั้ง 2 ปีที่แล้ว

รีวิวใหม่นี้ พบกับ ORA Good Cat รถไฟฟ้า 100% แบตเตอรี่ใหม่ ที่รถตลาดเมืองไทยด้วยดีไซน์สุดน่ารัก สไตล์ Retro Futuristic พร้อมสีสันทันใจเลือกถึง 7 สี จัดเต็มด้วยอุปกรณ์ช่วยเหลือการขับขี่ มีหลังคา Panoramic Sunroof ในรุ่นท็อป (500ULTRA) สามารถวิ่งได้ไกลสูงสุด 500 กิโลเมตรต่อการชาร์จหนึ่งครั้ง (ตาม ...เพิ่มเติม)

ความคิดเห็น 907 รายการ

เพิ่มความคิดเห็น...

@tokgun4768 2 ปีที่แล้ว (แก้ไขแล้ว)

ถ้าวิ่งในเมือง มีบ้าน มีที่จอดรถ และเป็นรถคันที่สอง รังในเมืองๆไม่เกิน 300km ไปกลับที่ทำงานทุกวัน ผมก็ว่าพอสำหรับชีวิตประจำวัน น่าจะเป็นจุดเริ่มต้นที่ดีนะครับ อย่างน้อยคนที่ใช้ก็สบายใจได้ด้วยครับ ส่วนตัว

21

ตอบกลับ

@teamthailand5209 2 ปีที่แล้ว

เป็นการรีวิวที่ดีที่สุดในเวลาการ Ora good cat ทั้งวันเต็มๆ ได้เห็นข้อดี และข้อเสียอย่างละเอียดครับ

51

ตอบกลับ

@AngkutRoongsangjan 2 ปีที่แล้ว

คุณรีวิวได้ละเอียดดีมากครับ ชอบ ตั้งแต่กระป๋องท้าย กระเป๋าหลัง เบาะหลัง User interface ซึ่งหาคนรีวิวละเอียดแบบนี้ได้ยาก

37

ตอบกลับ

@timelaps_hunter4865 2 ปีที่แล้ว

ส่วนตัวคิดว่า ราคาจะขายได้ยาก ไม่เหมาะจะกับบ้านที่ต้องการมีรถคันเดียว หรือ มีรถคันแรก ในราคา 1m น่าจะเหมาะจะเป็นรถคันที่สองเป็นเจ้าบ้าน หรือ รถอวดๆเป็น Hybrid DHEV ดีกว่าจะมาดาวน์คัน

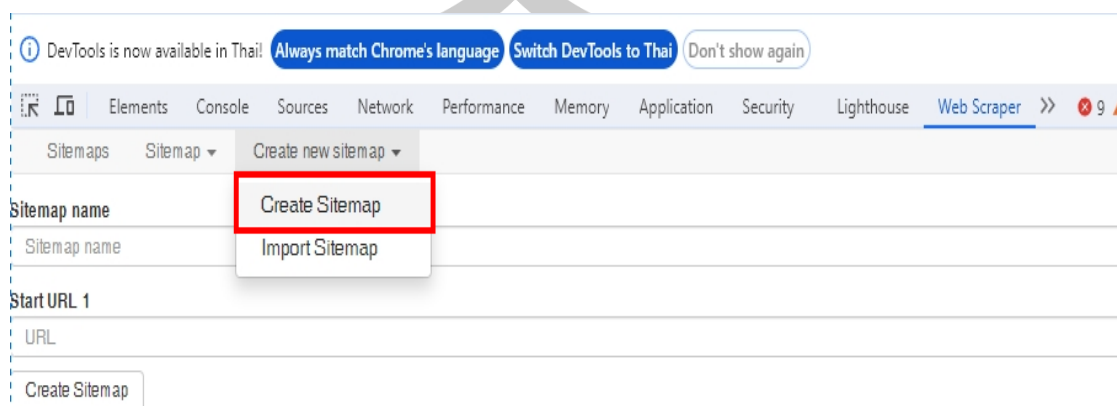
DevTools is now available in Thai! Always match Chrome's language Switch DevTools to Thai Don't show again

Web Scraper

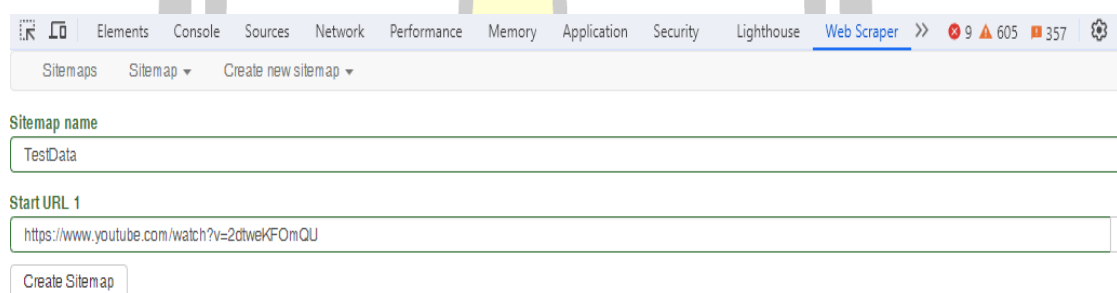
ภาพประกอบที่ 3.12 เปิดใช้งาน Web Scraper

จากนั้นทำการสร้าง ผังของเว็บ เพื่อใช้เป็นกรอบในจำกัดพื้นที่บนหน้าเว็บสำหรับการเก็บข้อมูลความคิดเห็นโดยกดที่เมนู Create Sitemap จากนั้นนำเอา URL ของหน้าที่จะเก็บเอาความคิดเห็นมาใส่ในช่อง Start URL 1 และทำการตั้งชื่อผังของเว็บในช่อง Sitemap name และกดปุ่ม Create Sitemap ดังภาพประกอบที่ 3.13 และ 3.14 ข้อสังเกตในช่องกรอกข้อมูลทั้ง 2 ช่อง ถ้าสามารถใช้งานได้ชื่อช่องข้อมูลและช่องข้อมูลจะเป็นสีเขียว ซึ่งหลังจากสร้างผังเพื่อจำกัดพื้นที่ในการ

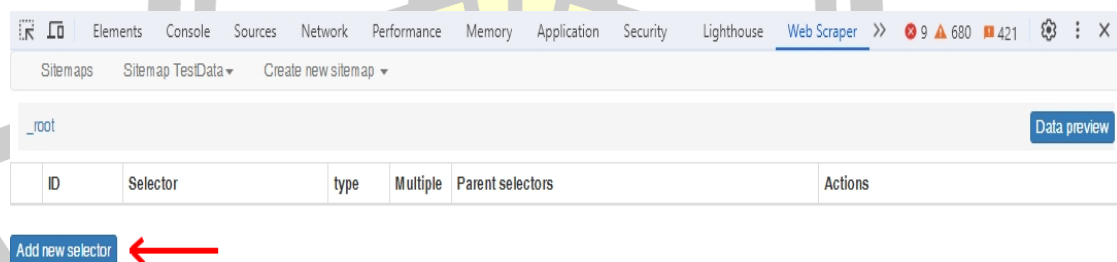
เลือกข้อมูลแล้วจะเข้าสู่ขั้นตอนการเลือกข้อมูล สังเกตจากหน้าเมนูของ Web Scraper จะเปลี่ยนไป จะปรากฏเมนู Add new selector ขึ้นมา ดังภาพประกอบที่ 3.15



ภาพประกอบที่ 3.13 การสร้างผังของเว็บ

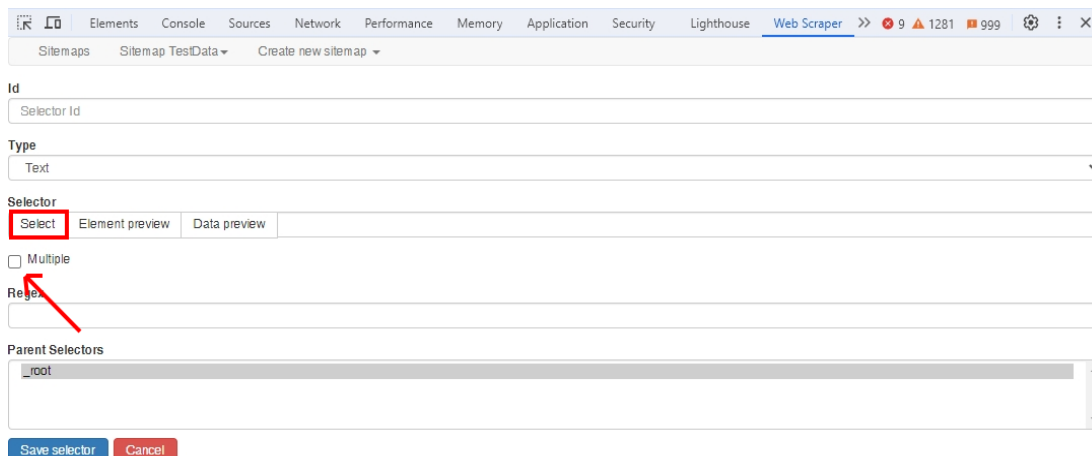


ภาพประกอบที่ 3.14 ข้อมูลสำหรับสร้างผังของเว็บ



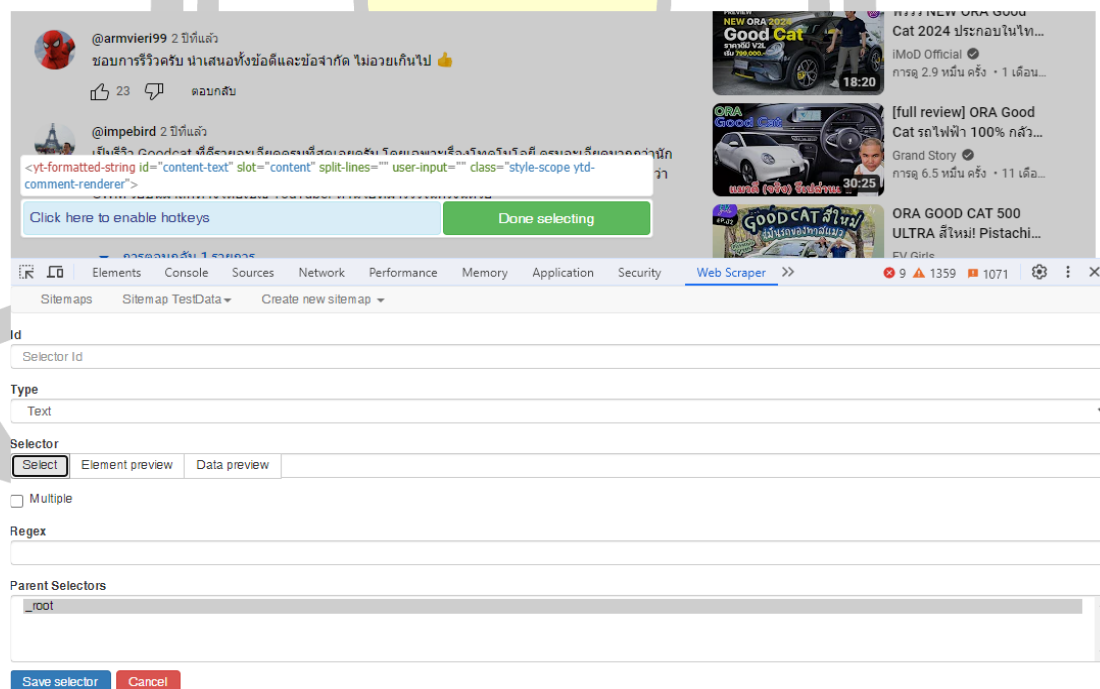
ภาพประกอบที่ 3.15 เตรียมเลือกข้อมูลความคิดเห็น

หลังจากเลือกเมนู Add new selector ก็จะได้ปรากฏหน้าสำหรับเลือกความคิดเห็น เราได้สร้างผังเอาไว้ก่อนหน้านี้ ซึ่งก่อนการเลือกข้อมูลให้เลือกเช็คบล็อกร Multiple ก่อนเพื่อเป็นการเลือกหลายๆความคิดเห็นพร้อมกัน จากนั้นกดที่เมนู Select เพื่อทำการเลือกข้อมูลความคิดเห็นที่เราต้องการ ดังภาพประกอบที่ 3.16



ภาพประกอบที่ 3.16 ตั้งค่าเลือกความคิดเห็นพร้อมกันหลาย ๆ ความคิดเห็น

จากนั้นทำการเลือกความคิดเห็นภายในหน้าเว็บที่เราได้สร้างผังไว้ ข้อควรระวัง ใน 1 หน้าเว็บสามารถสร้างผังเพื่อจำกัดการเก็บความคิดเห็นได้แค่ 1 ผัง เท่านั้น เมื่อต้องการความคิดเห็นจากหน้าเว็บอื่นต้องทำการสร้างผังหน้าเว็บใหม่ทุกครั้ง ในส่วนของการเลือกข้อมูลให้นำเมาส์คลิกเลือกที่ความคิดเห็นที่อย่างน้อย 2-3 ความคิดเห็น โปรแกรมจะทำการเลือกความคิดเห็นทั้งหมดให้โดยอัตโนมัติ ดังปรากฏในภาพที่ 3.17 ข้อสังเกต ความคิดเห็นที่ถูกเลือกจะปรากฏกรอบสีแดงขึ้นมาดังภาพที่ 3.18



ภาพประกอบที่ 3.17 หน้าต่างก่อนการเลือกความคิดเห็น

The screenshot shows a YouTube page for the video "ora good cat". The page displays several comments from users, many of which are highlighted with red boxes. The comments discuss the car's features, performance, and availability. Below the comments, a web scraper interface is visible, showing the "Selector" field with the value "yt-formatted-string#content-text". The interface includes buttons for "Save selector" and "Cancel".

Comments (บางส่วน):

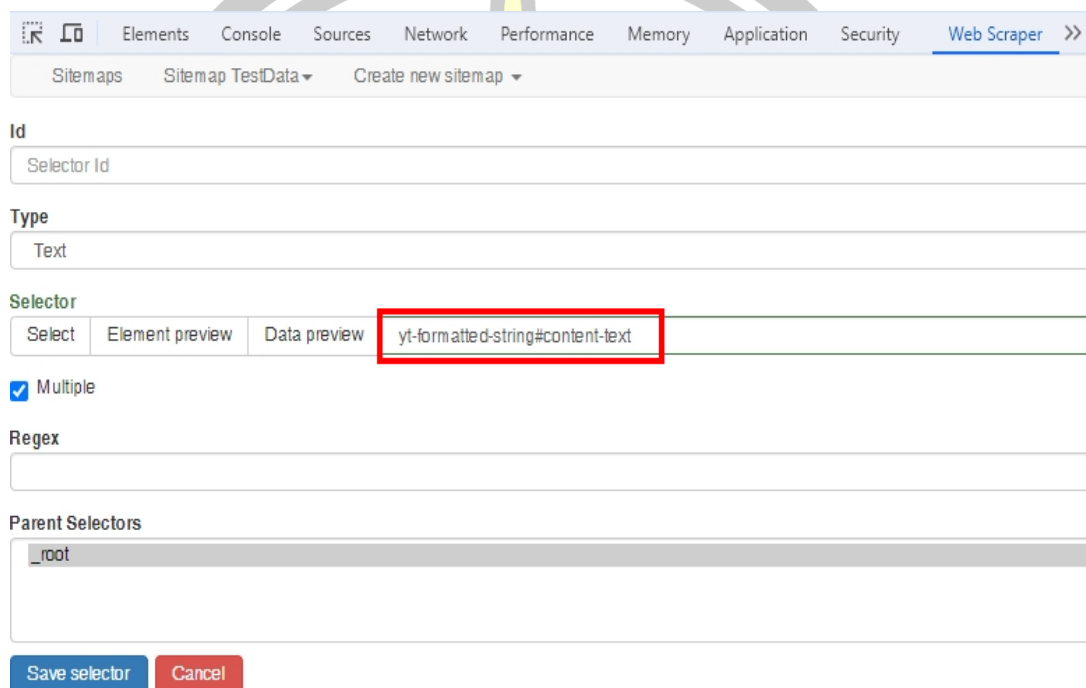
- @tokgun4768 2 ปีที่แล้ว (แก้ไขแล้ว): ถ้าวิ่งในเมือง มีบ้าน มีที่จอดรถ และเป็นรถคันที่สอง ริ่งในเมืองๆไม่เกิน 300km ไปกลับที่ทำงานทุกวัน ผมถือว่าพอสำหรับชีวิตประจำวัน น่าจะเป็นจุดเริ่มต้นที่ดีนะครับ อย่างน้อยคนที่มียานยนต์ก็จับต้องได้ด้วยครับ ส่วนตัว
- @teamthailand5209 2 ปีที่แล้ว: เป็นการรีวิวที่ดีที่สุด ไขว่คว้ากับรถ Ora good cat ทั้งวันเต็มๆ ได้เห็นข้อดี และข้อเสียอย่างละเอียดครับ
- @AngkutRoongsangjan 2 ปีที่แล้ว: คุณดูรีวิวดูละเอียดดีมากครับ ชอบ ตั้งแต่กระโปรงท้าย กระຈกขานหลัง เบาะหลัง User interface ซึ่งหาตามรีวิวละเอียดแบบนี้ได้ยาก
- @timelaps_hunter4865 2 ปีที่แล้ว: ส่วนตัวคิดว่า ราคาจะขายได้ยาก ไม่เหมาะกับบ้านที่ต้องการมีรถคันเดียว หรือ มีรถคันแรก ในราคา 1m บวกवल ควรไปซื้อรถที่ใหญ่กว่าที่เป็นน้ำมัน หรือ รถตลาดที่เป็น Hybrid PHEV อื่นๆ จะคุ้มกว่ามากครับ ถ้าต้องการรถไฟฟ้าจริงๆ MG ZS EV ยังคุ้มกว่าครับในราคาที่ใกล้เคียงกัน
- @ssann8778 2 ปีที่แล้ว: ดูหลายๆคนรีวิว อูรีวิวดีสุดละ รู่เรื่อง ชัดเจน ครบ
- @thanakornthipmanee7473 2 ปีที่แล้ว: ขอมไม่ดูช่องอื่นเพื่อรอดูช่องนี้เลย สม่กับการรอคอยจริงๆครับ
- @armvieri99 2 ปีที่แล้ว: ขอจบการรีวิวครับ น่าสนใจทั้งข้อดีและข้อจำกัด ไม่อวยเกินไป
- @impebird 2 ปีที่แล้ว: เป็นรีวิว Goodcat ที่ดีรายละเอียดครบที่สุดเลยครับ โดยเฉพาะเรื่องเทคโนโลยี ครับละเอียดมากกว่านัก

Web Scraper Interface:

- Id: Selector Id
- Type: Text
- Selector: Select Element preview Data preview
- Multiple: ☒
- Regex:
- Parent Selectors: _root
- Buttons: Save selector, Cancel

ภาพประกอบที่ 3.18 หน้าต่างหลังจากเลือกความคิดเห็น

เมื่อเลือกความคิดเห็นทั้งหมดแล้วกดปุ่ม Done selecting (ปุ่มสีเขียว) เพื่อจบการเลือกความคิดเห็น หน้าต่างของโปรแกรม Web Scraper ในเมนู Selector จะแสดง ID ของความคิดเห็นที่ถูกเลือก ดังที่ปรากฏในภาพ 3.19 และสามารถตรวจสอบข้อมูลความคิดเห็นที่ถูกเลือกทั้งโดยการเลือกเมนู Data preview ดังภาพที่ 3.20



Id

Selector Id

Type

Text

Selector

Select Element preview Data preview **yt-formatted-string#content-text**

☒ Multiple

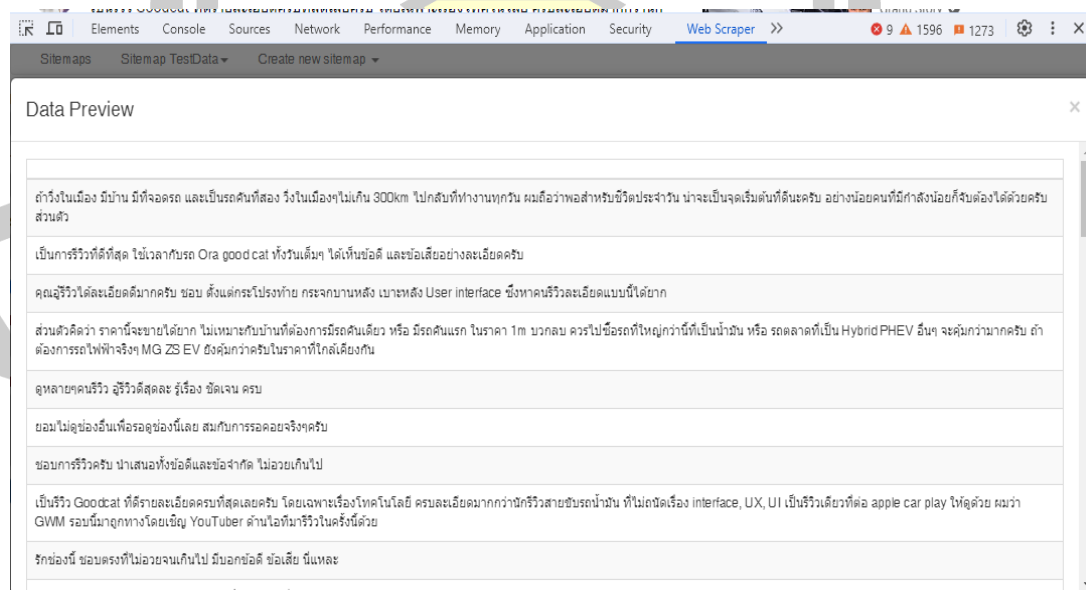
Regex

Parent Selectors

_root

Save selector Cancel

ภาพประกอบที่ 3.19 ID ของความคิดเห็นที่ถูกเลือก



Data Preview

ถ้าวิ่งในเมือง มีบ้าน มีที่จอดรถ และเป็นรถคันที่สอง รังในเมืองไม่เกิน 300km ไปกลับที่ทำงานทุกวัน ผมถือว่าพอสำหรับชีวิตประจำวัน น่าจะเป็นจุดเริ่มต้นที่ดีครับ อย่างน้อยคนที่กำลังมองหาซื้อด้วยครับ ส่วนตัว
เป็นการรีวิวที่ดีสุด ไม่เวลาขับรถ Ora good cat ทั่ววันเต็มๆ ได้เห็นข้อดี และข้อเสียอย่างละเอียดครับ
คุณผู้รีวิวได้ละเอียดดีมากครับ ชอบ ตั้งแต่กระป๋องท้าย กระเป๋าหลัง เบาะหลัง User interface ซึ่งหาคนรีวิวละเอียดแบบนี้ได้ยาก
ส่วนตัวคิดว่า ราคาจะขายได้ยาก ไม่เหมาะกับบ้านที่ต้องการมีรถคันเดียว หรือ มีรถคันแรก ในราคา 1m บวกลบ ควรไปซื้อรถที่ใหญ่กว่านี้ที่เป็นน้ำมัน หรือ รถผลัดที่เป็น Hybrid PHEV อื่นๆ จะคุ้มกว่ามากครับ ถ้าต้องการรถไฟฟ้าจริงๆ MG ZS EV ยังคุ้มกว่าครับในราคาที่ใกล้เคียงกัน
ดูหลายๆคนรีวิว ผู้รีวิวดีสุดละ รู้เรื่อง ชัดเจน ครับ
ยอมไม่ดูช่องอื่นเพื่อรอช่องนี้เลย สนุกกับการรอคอยจริงๆครับ
ชอบการรีวิวครับ น่าสนใจทั้งข้อดีและข้อจำกัด ไม่อวยเกินไป
เป็นรีวิว Goodcat ที่รายละเอียดครบที่สุดเลยครับ โดยเฉพาะเรื่องเทคโนโลยี ครับละเอียดมากกว่านักรีวิวสายซบถน้ำมัน ที่ไม่ถนัดเรื่อง interface, UX, UI เป็นรีวิวเดียวที่ต่อ apple car play ไปด้วย ผมว่า GWM รอบนี้มากถูกทางโดยเชิญ YouTuber ด้านไอทีมารีวิวในครั้งนี้นี้ด้วย
รักช่องนี้ ชอบตรงที่ไม่อวยจนเกินไป มีบอกข้อดี ข้อเสีย นี้แหละ

ภาพประกอบที่ 3.20 ความคิดเห็นที่เลือกทั้งหมด

เนื่องจากการใช้งานโปรแกรมฟรีจึงไม่สามารถส่งออกไฟล์ในรูปแบบของโปรแกรม Microsoft Excel ได้ ซึ่งสามารถแก้ปัญหาได้โดยการ เลือกข้อมูลทั้งหมด (Ctrl+A) จากนั้นสำเนาไปวางในโปรแกรม Microsoft Excel ได้ ข้อควรระวังในการใช้งานโปรแกรม Web Scraper อาจจะมีบางครั้งที่โปรแกรมนำเอาความคิดเห็นที่มี Emoticon, Sticker อยู่ในความคิดเห็นที่ถูกเลือก ซึ่งหลังจากที่ได้ความคิดเห็นทั้งหมดที่รวบรวมมาแล้วจะเข้าสู่กระบวนการคัดเลือกความคิดเห็นด้วยภาษา Python โดยพิจารณาจากความคิดเห็นที่มีคำในความคิดเห็นมากกว่า 15 คำ ขึ้นไป ดังภาพประกอบที่ 3.21

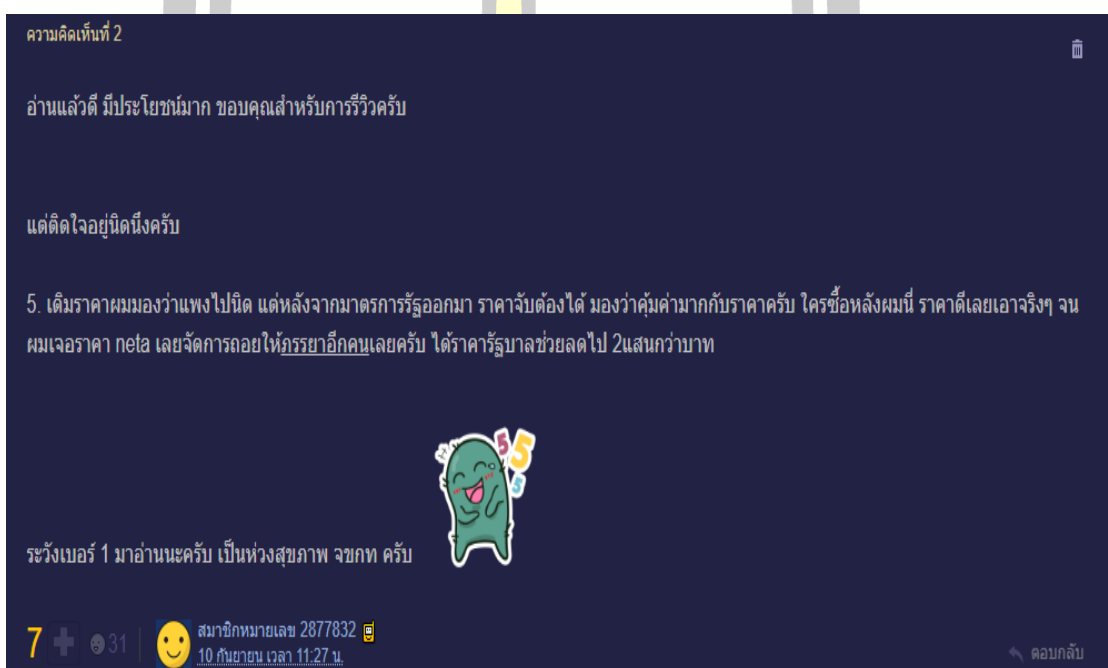
```

1 from google.colab import drive
2 import pandas as pd
3
4 # ดึงไฟล์ .xls จาก Google Drive
5 drive.mount('/content/drive')
6 file_path = '/content/drive/My Drive/select_data.xls'
7
8 df = pd.read_excel(file_path)
9
10 comments = df['comment_column']
11
12 # ฟังก์ชันตัดคำ
13 def split_words(text):
14     return text.split()
15
16 # ฟังก์ชันนับคำ
17 def count_words(words):
18     return len(words)
19
20 # เก็บผลลัพธ์
21 results = []
22
23 # วนลูปแต่ละความคิดเห็น
24 for comment in comments:
25     # ตัดคำ
26     words = split_words(comment)
27     # นับคำ
28     word_count = count_words(words)
29     # เก็บผลลัพธ์
30     if word_count >= 15:
31         results.append({
32             'comment': comment,
33             'word_count': word_count
34         })
35
36 # แปลง results เป็น Pandas DataFrame
37 df_results = pd.DataFrame(results)
38
39 # เขียน df_results ไปยัง Google Sheet
40 sheet_name = 'your_sheet_name'
41 df_results.to_csv(f'/content/drive/My Drive/{sheet_name}.csv', index=False)
42
43 # แจ้งเตือน
44 print('Done!')
```

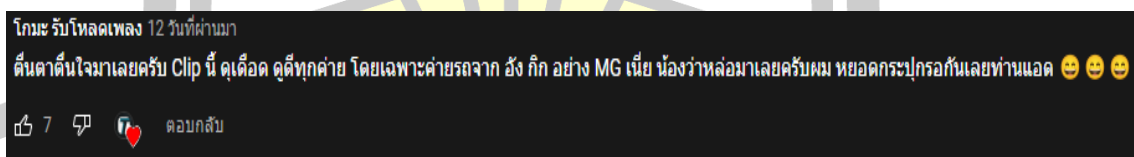
ภาพประกอบที่ 3.21 ภาษา Python ใช้คัดเลือกความคิดเห็น

3.2 ขั้นตอนก่อนการประมวลผลข้อความ

ขั้นตอนก่อนการประมวลผลข้อความ เป็นขั้นตอนที่จะนำเอาความคิดเห็นของคนไทยที่ใช้สื่อสังคมออนไลน์ จำนวน 2,265 ความคิดเห็น ซึ่งเก็บข้อมูลอยู่ในรูปของไฟล์ประเภท .xls ด้วยโปรแกรม Microsoft Excel ซึ่งขั้นตอนนี้ความคิดเห็นที่ถูกเลือกจากขั้นตอนที่ 3.1 จะถูกเปลี่ยนให้เป็นประโยคและตัดส่วนที่ไม่ได้ใช้งานในการประมวลผลออกไป เช่น อักษรภาษาอังกฤษ, ตัวเลข, Emoticon, Sticker และสัญลักษณ์ต่างๆ ด้วยเครื่องมือภาษา Python



ภาพประกอบที่ 3.22 ตัวอย่างการแสดงความเห็นที่มีการเพิ่ม Sticker



ภาพประกอบที่ 3.23 ตัวอย่างการแสดงความเห็นที่มีการเพิ่ม Emoticon

	text	cleaning
0	อ่านแล้วดี มีประโยชน์มาก ขอบคุนสำหรับการรวิวค...	อ่านแล้วดีมีประโยชน์มากขอบคุณสำหรับการรวิวค...
1	เรื่องไม่มีล่ออะไรมากมาให้มีมานานแล้วครับไม่เฉ...	เรื่องไม่มีล่ออะไรมากมาให้มีมานานแล้วครับไม่เฉ...
2	ยากให้ราคาไม่แพงเกินไปอ้อฟชั่นเยอะๆครับ	ยากให้ราคาไม่แพงเกินไปอ้อฟชั่นเยอะๆครับ
3	รอ i3 เลยครับ สวยมาก อยากได้รถไฟฟ้าแบบเก้ง	รอเลยครับสวยมากอยากได้รถไฟฟ้าแบบเก้ง
4	อะไรก็มาเถอะครับ แต่จะขอเลือกกรรยี่ห้อผลิตในประ...	อะไรก็มาเถอะครับแต่จะขอเลือกกรรยี่ห้อผลิตในประ...
5	Mine กระบะไทย หัวใจไฟฟ้า มาอุดหนุนแน คับ เชิญ...	กระบะไทยหัวใจไฟฟ้ามาอุดหนุนแนคับเชิญชวนคนไทย...
6	ดินดาดีใจมาเลยครับ Clip นี้ ดูเด็ด ดูดีทุก...	ดินดาดีใจมาเลยครับนี้ดูเด็ดดูดีทุกค่ายโดย...
7	ชอบ Smart อ่าสวยงามรัก เห็นมีแค่ Suv, B-suv ค...	ชอบอ่าสวยงามรักเห็นมีแค่คันใหญ่ๆแบบนี้น่ารัก...
8	CHALIS LL 12 วันที่ผ่านมา Series 5 วันที่เห็นวิ...	วันที่ผ่านมาวันนี้เห็นวิ่งบนถนนแล้วครับสวยมาก...
9	ต้องขออุดหนุนบริการก่อนครับ ว่ามีมากน้อยแค่ไหน...	ต้องขออุดหนุนบริการก่อนครับว่ามีมากน้อยแค่ไหน...

ภาพประกอบที่ 3.24 ความคิดเห็นก่อนและหลังการทำความสะอาดคำ

ซึ่งหลังจากทำความสะอาดแล้ว ผู้วิจัยจะทำการตรวจสอบประโยคอีกครั้งและทำการแก้ไขหากพบจุดผิดพลาด เช่น พิมพ์ผิด พิมพ์ตก เพื่อให้ได้ คำ ที่สมบูรณ์ที่สุดในการค้นหาคุณลักษณะของคำ ดังภาพที่ 3.25

['อ่านแล้วดีมีประโยชน์มากขอบคุณสำหรับการรวิวคครับแต่ติดใจอยู่นิดนึงครับเดิมราคาผมมองว่าแพงไปนิดแต่หลังจากมาตรวจได้ราคาถูกลงช่วยลดไปแสนกว่าบาทะวังเบอร์มาอ่านนะครับเป็นห่วงสุขภาพขกทครับ', 'เรื่องไม่มีล่ออะไรมากมาให้มีมานานแล้วครับไม่เฉพาจะรถจีนรถยุโรปนี้แหละครับ', 'ยากให้ราคาไม่แพงเกินไปอ้อฟชั่นเยอะๆครับ', 'รอเลยครับสวยมากอยากได้รถไฟฟ้าแบบเก้ง', 'อะไรก็มาเถอะครับแต่จะขอเลือกกรรยี่ห้อผลิตในประเทศก่อนก็ขอให้ผลิตในบ้านเราด้วยเถิดแถมแถมเดียวขับหลังก็เหลือก็', 'กระบะไทยหัวใจไฟฟ้ามาอุดหนุนแนคับเชิญชวนคนไทยอุดหนุนยี่ห้อไทยด้วยกันให้ก้าวไประดับโลกมีเทคโนโลยีขาร์จแบด', 'ดินดาดีใจมาเลยครับนี้ดูเด็ดดูดีทุกค่ายโดยเฉพาะค่ายรถจากอังกฤษก็อย่างเนี่ยน้องว่าล่อมาเลยครับผมหยอดกระป๋องกัน', 'ชอบอ่าสวยงามรักเห็นมีแค่คันใหญ่ๆแบบนี้น่ารักเลยแต่รถเล็กคงทำใจเรื่องระยะทางแต่เอาไปแต่ละค่ายไม่เน้นไปทำเรื่องแแต่เนิสสันบอกสวยหรือไม๊ขาย', 'วันที่ผ่านมาวันนี้เห็นวิ่งบนถนนแล้วครับสวยมากๆไม่'ได้ถ่ายมาที่แรกไม่รู้จักเห็นข้างหลังเขียนเลยลองดูเพราะรถสวยจริงๆ', 'ต้องขออุดหนุนบริการก่อนครับว่ามีมากน้อยแค่ไหนค่อยตัดสินใจอีกทีครับ']

ภาพประกอบที่ 3.25 ความคิดเห็นหลังจากการทำความสะอาดและแก้ไขคำ

3.3 การสร้างคุณลักษณะ

การสร้างคุณลักษณะ ในขั้นตอนนี้ผู้วิจัยได้นำเอาข้อมูลทำความสะอาดแล้วมาทำการตัดคำสร้างถูงคำ และกำหนดคลาส โดยมีขั้นตอนดังต่อไปนี้

การตัดคำ (Tokenization) ในขั้นตอนนี้ผู้วิจัยได้ใช้ ภาษา python และ Library PyThaiNLP [33] โดยนำข้อมูลทำความสะอาดแล้วมาแยกเป็นคำ เพื่อให้ง่ายต่อการนับคำ ซึ่งขั้นตอนนี้จะพิจารณาเพิ่มเติมว่าการตัดคำมีความเหมาะสมหรือไม่ เพื่อนำไปแยกคุณลักษณะในขั้นตอน

ต่อไป ดังปรากฏในภาพประกอบที่ 3.26 จากนั้นนำข้อมูลไปเรียงใหม่และใช้ ภาษา python ในตรวจสอบและระบุชนิดของคำ ดังภาพประกอบที่ 3.27

```
1 from pythainlp import word_tokenize
2
3 text = "รถสวยราคาไม่แพง มีออฟชั่นครบน่าใช้มากๆเลยครับ"
4
5 print("default (newmm):")
6 print(word_tokenize(text))
7 print("\nnewmm and keep_whitespace=False:")
8 print(word_tokenize(text, keep_whitespace=False))
```

default (newmm):

['รถ', 'สวย', 'ราคา', 'ไม่', 'แพง', ' ', 'มี', 'ออฟ', 'ชั่น', 'ครบ', 'น่า', 'ใช้', 'มาก', 'ๆ', 'เลย', 'ครับ']

newmm and keep_whitespace=False:

['รถ', 'สวย', 'ราคา', 'ไม่', 'แพง', 'มี', 'ออฟ', 'ชั่น', 'ครบ', 'น่า', 'ใช้', 'มาก', 'ๆ', 'เลย', 'ครับ']

ภาพประกอบที่ 3.26 การแยกคำเพื่อแยกคุณลักษณะ

['คัน, ออฟชั่น, ผม, ขับ, แนวโดน, ใจ, อะ, ตัว, คัน, นี้, ดอบ, โจทย์, ผม, ดอน, แอบ, ดอนเท',
'เท้า, แขน, แอร์, มือ, ทำ, ขับ, คน, นั่ง, เรือง, นั่ง, จุกจิก',
'ใหม่, ด้อย, รถ, ดปทขาย, ดี, สงสัย, ดปทรู้, ห่าง, ขับ, ข้าม, รถ, ชาร์จ, บริหาร, แข็งน้อ',
'ปกติ, คน, ขับ, รถ, น้ำมัน, ชอบ, ค่าย, อะไหล่, คุณภาพ, ลอง, รถ, ไฟฟ้า, ขับดี, เปิดใจ',
'ออฟชั่น, มือ, หลังคา, แอร์, เท้า, แขน, ชัน, รูป, บิดา, น, ำฝน, ยาง, อะไหล่, คน, รีวิวเวอร์, 6',
'สวัสดี, น้องนิธิ, ฝาก, ผม, น้อง, นิธิ, เวลา, แนะนำ, รถ, แฉง, รายละเอียด, สมมติ, แรง, บิด',
'เวลา, ขับ, รีวิวนิธิ, ดีดีนิธิ, รีวิว, ยี่ห้อ, เชื้อ, ลองรุ่น, ดี, ไฟ, จอสว่างรุ่น, ดี',

ภาพประกอบที่ 3.27 ข้อมูลหลังจากการตัดคำ

การทำถุงคำ (Bag of Word) [34] คือการแปลงคำแต่ละคำที่ไม่ซ้ำกัน ให้เป็น ID ของคำนั้น ๆ และทำการนับจำนวนครั้งที่แต่ละคำปรากฏ โดยไม่ต้องคำนึงถึงหลักไวยากรณ์ ความถี่ของคำ และลำดับคำในแต่ละประโยค ขั้นตอนนี้ผู้วิจัยจะทำการกำหนดคุณลักษณะของคำ โดยเลือกใช้ชนิดของคำ คำกริยา คำวิเศษณ์ และคำคุณศัพท์ ซึ่งสามารถแสดงถึงคุณลักษณะของคำได้ และแทนค่าลักษณะคำ ที่เป็นความคิดเห็นเชิงบวกด้วย 1 และความคิดเห็นเชิงลบด้วย -1 ส่วนคำที่มีความคิดเห็นเป็นกลางแทนค่าด้วย 0 ดังแสดงใน ตารางที่ 3.1

ตารางที่ 3.1 การเพิ่มคุณลักษณะคำ

ลำดับ	คำ	ชนิดคำ	จำนวนที่พบ	คุณลักษณะ
1	ฉัน	คำกริยา	70	0
2	ออฟชั่น	คำวิเศษณ์	14	1
3	ผม	คำนาม	20	0
4	ขับ	คำกริยา	74	0
5	แนว	คำนาม	5	0
6	โดนใจ	คำคุณศัพท์	1	1
7	ตอบโจทย์	คำคุณศัพท์	2	1
8	จุกจิก	คำกริยา	1	-1
9	ดี	คำคุณศัพท์	696	1
10	ปัญหา	คำคุณศัพท์	202	-1

ในขั้นตอนการแยกคำผู้วิจัยได้ใช้ภาษา Python ในการแยกคำและกำหนดชนิดของคำ (Part of speech) โดยใช้ Library PyThaiNLP ที่มีความสามารถในการแยกคำภาษาไทยและกำหนดชนิดของคำได้อย่างแม่นยำ ซึ่งมีความสามารถในการจำแนกและระบุชนิดคำได้ทั้งหมด 15 ชนิด ซึ่งขั้นตอนนี้เป็นขั้นตอนที่สำคัญอีกหนึ่งขั้นตอนเพราะภาษาไทยมีรูปแบบการเขียนติดกันไม่มีการเว้นวรรคเหมือนภาษาอังกฤษหากการแยกคำออกมาเพื่อกำหนดหรือระบุชนิดคำจึงต้องมีการตรวจสอบความถูกต้องซ้ำหลายรอบ เพราะหากแยกคำผิดพลาดอาจจะส่งผลถึงผลลัพธ์ของงานวิจัยในครั้งนี้ และทำให้ผิดวัตถุประสงค์ออกไป Library PyThaiNLP สามารถแยกคำภาษาไทยและกำหนดชนิดของคำได้ ดังภาพประกอบที่ 3.28 และสามารถรวบรวมคำที่จะใช้ในการสร้างดัชนีข้อมูลได้ ดังตารางที่ 3.2

พูน ปณ จิต ชเว

Abbreviation	Part-of-Speech tag	Examples
ADJ	Adjective	ใหม่, พิเศษ , ก่อน, มาก, สูง
ADP	Adposition	แม้, ว่า, เมื่อ, ของ, สำหรับ
ADV	Adverb	ก่อน, ก็, เล็กน้อย, เลย, สุด
AUX	Auxiliary	เป็น, ใช้อยู่, คือ, คล้าย
CCONJ	Coordinating conjunction	แต่, และ, หรือ
DET	Determiner	ที่, นี้, ซึ่ง, ทั้ง, ทุก, หลาย
INTJ	Interjection	อ้อ, โอ๊ย
NOUN	Noun	กำมือ, พวก, สนาม, กีฬา, บัญชี
NUM	Numerical	5,000, 103.7, 2004, หนึ่ง, ร้อย
PART	Particle	มา ขึ้น ไม่ ได้ เข้า
PRON	Pronoun	เรา, เขา, ตัวเอง, ใคร, เธอ
PROPN	Proper noun	โอบามา, แคปิตอลฮิล, จีโอพี, ไมเคิล
PUNCT	Punctuation	(,), “ ” :
SCONJ	Subordinating conjunction	หาก
VERB	Verb	เปิด, ให้, ไข, เผชิญ, อ่าน

ภาพประกอบที่ 3.28 ชนิดของคำ (Part of Speech)

ตารางที่ 3.2 ชนิดและจำนวนของคำทั้งหมด

ชนิดคำ	ADJ	ADV	DET	NOUN	NUM	PRON	PROPN	SCONJ	VERB
จำนวน	16	39	1	2471	1	2	301	29	184
รวม	3,044								

จากการศึกษาจากงานวิจัยของผู้วิจัยท่านอื่น ๆ ที่ได้ทำเหมือนความคิดเห็นพบว่า การเลือกเอา คำกริยา (VERB) คำคุณศัพท์ (ADJ) และคำวิเศษณ์ (ADV) มาทำเป็นถ้อยคำ จะมีประสิทธิภาพสูงสุดในการจำแนกความรู้สึก คำกริยา สามารถแสดงออกด้านความรู้สึกได้อย่างชัดเจนว่าอยู่ในเชิงบวกหรือเชิงลบ คำวิเศษณ์ มักจะขยายความรู้สึกให้ชัดเจนมากยิ่งขึ้น เช่น สวยดี, ใหญ่ดี, ดุดี และ คำคุณศัพท์ มีความใกล้เคียงกันโดยจะสามารถบอกความรู้สึกได้ในระดับหนึ่ง เช่น เกลียด, ไม่ชอบ, แย่ เป็นต้น ผู้วิจัยจะนำเอาคำเหล่านี้ไปใช้ในการกำหนดคลาสของคำ ดังตารางที่ 3.3

ตารางที่ 3.3 ตัวอย่างการกำหนดคลาสของคำ

ลำดับ	คุณลักษณะเชิงบวก		คุณลักษณะเชิงลบ		ความถี่ของลักษณะเชิงบวก	ความถี่ของลักษณะเชิงลบ	คลาส
	ดี	ชอบ	แย่	แพง			
1	1	0	0	0	1	0	P
2	0	0	0	1	0	1	N
3	1	0	0	1	1	1	O
4	2	1	0	0	3	0	P
5	0	0	0	0	0	2	N

จากตารางที่ 3.3 แสดงการกำหนดคลาสของคำ โดยการนำเอาค่าที่ได้ระบุชนิดคำและคุณลักษณะแล้วมานับความถี่โดย การนำเอา “คำ” ที่มีคุณลักษณะเชิงบวกหรือเชิงลบ ไปเทียบกับความคิดเห็นในแต่ละความคิดเห็น เช่น ความคิดเห็นลำดับที่ 4 มีการกล่าวถึงคำว่า “ดี” จำนวน 2 ครั้ง และกล่าวถึงคำว่า “ชอบ” จำนวน 1 ครั้ง ความถี่ของคุณลักษณะเชิงบวกจึงเป็น 3 เมื่อได้ความถี่ครบถ้วนทั้งหมดแล้ว นำเอาค่าความถี่ของทั้ง 2 คุณลักษณะลบกันก็จะได้รูปแบบคลาสของคำออกมา 3 รูปแบบ ดังนี้

- รูปแบบที่ 1 เป็นความคิดเห็นเชิงบวก (Positive) แทนค่าของคลาสด้วย P
- รูปแบบที่ 2 เป็นความคิดเห็นเชิงลบ (Negative) แทนค่าของคลาสด้วย N
- รูปแบบที่ 3 เป็นความคิดเห็นเป็นกลาง (Neutral) แทนค่าของคลาสด้วย 0 ซึ่งรูปแบบที่มีความคิดเห็นเป็นกลางนี้จะนับรวมความคิดเห็นที่ถูกตัดออกจากชุดข้อมูลด้วย

หลังจากกำหนดคลาสเสร็จแล้วผู้วิจัยจะตรวจสอบความถูกต้องอีกครั้งก่อนจะทำการแปลงค่าของความถี่คุณลักษณะเพื่อนำไปสร้างแบบจำลองจากเทคนิคต่าง ๆ ได้ ดังตารางที่ 3.4

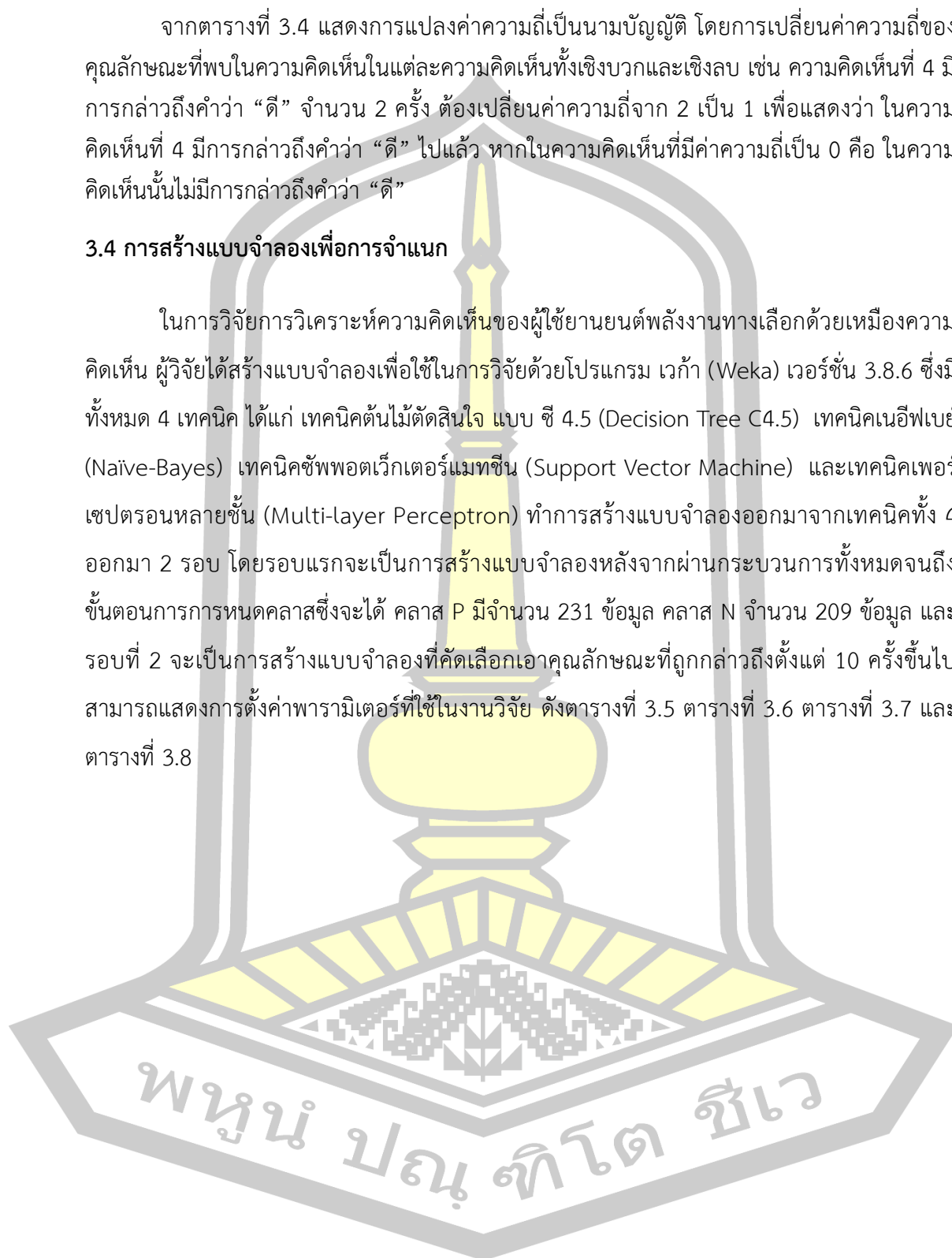
ตารางที่ 3.4 ตัวอย่างการแปลงค่าความถี่ให้เป็นนามบัญญัติ

ลำดับ	คุณลักษณะเชิงบวก		คุณลักษณะเชิงลบ		คลาส
	ดี	ชอบ	แย่	แพง	
1	1	0	0	0	P
2	0	0	0	1	N
3	1	0	0	1	O
4	1	1	0	0	P
5	0	0	0	0	N

จากตารางที่ 3.4 แสดงการแปลงค่าความถี่เป็นนามบัญญัติ โดยการเปลี่ยนค่าความถี่ของ คุณลักษณะที่พบในความคิดเห็นในแต่ละความคิดเห็นทั้งเชิงบวกและเชิงลบ เช่น ความคิดเห็นที่ 4 มีการกล่าวถึงคำว่า “ดี” จำนวน 2 ครั้ง ต้องเปลี่ยนค่าความถี่จาก 2 เป็น 1 เพื่อแสดงว่า ในความคิดเห็นที่ 4 มีการกล่าวถึงคำว่า “ดี” ไปแล้ว หากในความคิดเห็นที่มีค่าความถี่เป็น 0 คือ ในความคิดเห็นนั้นไม่มีการกล่าวถึงคำว่า “ดี”

3.4 การสร้างแบบจำลองเพื่อการจำแนก

ในการวิจัยการวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น ผู้วิจัยได้สร้างแบบจำลองเพื่อใช้ในการวิจัยด้วยโปรแกรม เวก้า (Weka) เวอร์ชัน 3.8.6 ซึ่งมีทั้งหมด 4 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5) เทคนิคเนอ์ฟเบย์ (Naïve-Bayes) เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ทำการสร้างแบบจำลองออกมาจากเทคนิคทั้ง 4 ออกมา 2 รอบ โดยรอบแรกจะเป็นการสร้างแบบจำลองหลังจากผ่านกระบวนการทั้งหมดจนถึงขั้นตอนการการหนดคลาสซึ่งจะได้ คลาส P มีจำนวน 231 ข้อมูล คลาส N จำนวน 209 ข้อมูล และรอบที่ 2 จะเป็นการสร้างแบบจำลองที่คัดเลือกเอาคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป สามารถแสดงการตั้งค่าพารามิเตอร์ที่ใช้ในงานวิจัย ดังตารางที่ 3.5 ตารางที่ 3.6 ตารางที่ 3.7 และ ตารางที่ 3.8



ตารางที่ 3.5 การตั้งค่า (Parameter) เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5

Weka.classifierstrees.J48	
หัวข้อ	การตั้งค่า
batchSize	100
binatySplits	false
CollapseTree	True
confidenceFactor	0.25
Debug	False
doNotcheckCapabilities	False
doNotMakeSplitPonitActualValue	False
minNumObj	2
numDecimalPlaces	2
numFolds	3
ReducedErrorPruning	False
saveInstanceData	False
Seed	1
subtreeRaising	True
Unpruned	False
useLaplace	False
useMDLcorrection	True

พหุ ประทีป ชีวะ

ตารางที่ 3.6 การตั้งค่า (Parameter) เทคนิคซ์ฟพอดเว็กเตอร์แมทชีน

Weka.classifiers.functions.LibSVM	
หัวข้อ	การตั้งค่า
SVMTType	C-SVC (classification)
batchSize	100
cacheSize	40.0
Coefo	0.0
Cost	1.0
Debug	False
Degree	3
doNotCheckCapabilities	Flase
doNotReplaceMissingValues	False
Eps	0.001
Gamma	0.0
doNotCheckCapabilities	False
doNotReplaceMMissingValues	False
Eps	0.001
kernalType	Radial basis function: $\exp(-\gamma \ u-v\ ^2)$
Loss	0.1
modelFile	Ad
Normalize	False
Nu	0.5
numDecimalPlaces	2
probabilityEstimates	False
Seed	1
Shrinking	True
weight	

ตารางที่ 3.7 การตั้งค่า (Parameter) เทคนิคเพอร์เซปตรอนหลายชั้น

Weka.classifiers.functions.MutiLayerPerceptron	
หัวข้อ	การตั้งค่า
GUI	False
autoBuild	True
batchSize	100
Debug	False
Decay	False
doNotCheckCapabilities	False
hiddenLayers	A
learningRate	0.3
Momentum	0.2
nominalToBinaryFilter	True
normalizeAttributes	True
normalizeNumericClass	True
numdecimalPlaces	2
Reset	True
Resume	False
Seed	0
trainingTime	500
validationSetSize	0
validationThreshold	20

ตารางที่ 3.8 การตั้งค่า (Parameter) เทคนิคเนอ์ฟเบย์

Weka.classifiers.bayes.NaiveBayes	
หัวข้อ	การตั้งค่า
batchSize	100
Debug	False
displayModelInOldFormat	False
doNotCheckCapabilities	False
numDecimalPlaces	2
useKernelEstimator	False
useSupervisedDiscretization	False

3.4.1 เทคนิคต้นไม้ตัดสินใจ แบบ ซี4.5 (C4.5) [35] เป็นการเรียนรู้โดยการจำแนก (Classification) ชุดข้อมูล ออกเป็นกลุ่ม (Class) ต่าง ๆ โดยใช้คุณสมบัติ (Attribute) ของข้อมูลในการแยกแยะต้นไม้ตัดสินใจที่ได้จากการเรียนรู้ ทำให้ทราบคุณสมบัติใดของ ข้อมูลที่เป็นตัวกำหนดการจำแนก และคุณสมบัติแต่ละตัวของข้อมูลนั้นมีความสำคัญมากน้อยอย่างไร ซึ่งเป็นประโยชน์ที่จะช่วยให้ ผู้ใช้สามารถวิเคราะห์ข้อมูลและตัดสินใจได้ถูกต้องยิ่งขึ้น ผลลัพธ์ของการเรียนรู้ต้นไม้ตัดสินใจประกอบด้วย 1) โหนดภายใน (Internal Node) 2) กิ่ง (Branch, Link) คือ ค่าคุณสมบัติของคุณสมบัติในโหนดภายในที่แตกกิ่งนี้ออกมา ซึ่งโหนดภายในจะแตกกิ่ง เป็นจำนวนเท่ากับจำนวนค่าคุณสมบัติของโหนดภายในนั้น 3) โหนดใบ (Leaf Node) คือ กลุ่ม (Class) ต่าง ๆ ซึ่งเป็นผลลัพธ์ในการ จำแนกข้อมูล ซึ่งเป็นขั้นตอนวิธีหนึ่งที่ใช้ในการสร้างต้นไม้ตัดสินใจจะใช้ค่าที่วัดที่เรียกว่าค่า เกน (Gain) เป็นตัวตัดสินใจว่าจะใช้แอตทริบิวต์ ใดในการแบ่งข้อมูลเพื่อใช้ในการตัดสินใจ โดยวิธีการกำหนดโครงสร้างต้นไม้ตัดสินใจจะเป็นตามลำดับของค่าตัวชี้วัดหรือค่า เกนของแอตทริบิวต์ ที่มีค่าสูงที่สุด ซึ่งเริ่มหาค่าสารสนเทศ (Information) แต่ละแอตทริบิวต์ ซึ่งบ่งบอกความสามารถของแอตทริบิวต์ ในการแยกแต่ละคลาส (Class) และหาค่าเอนโทรปี (Entropy) เป็นค่าของผลรวมแต่ละแอตทริบิวต์ ตามสมการที่ 2 และ 3

$$I(S_1, S_2, \dots, S_n) = - \sum_i^n \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (2)$$

โดย S คือ จำนวนข้อมูลทั้งหมด ซึ่งมีจำนวนข้อมูล s_i ของ $CLASS_i$

n คือ จำนวน Class ทั้งหมด

$$E(A) = \sum_{j=1}^m \frac{s_{1j} + \dots + s_{1j}}{s} I(S_{1f}, \dots, S_{2j}) \quad (3)$$

ค่าเกน (Gain) ของแอตทริบิวท์ที่ถูกเลือกหาได้จากการแยกประเภทข้อมูล สารสนเทศทั้งหมดของแอตทริบิวท์นั้นดังสมการที่ 4

$$\text{Gain}(A) = I(S_1, S_2, \dots, S_n) - E(A) \quad (4)$$

3.4.2 เทคนิคเนอ์ฟเบย์ (Naïve-Bayes) [36] มีพื้นฐานมาจากกฎของเบย์ เป็นทฤษฎีทางด้านสถิติโดยนำความน่าจะเป็นมาใช้ประเมินความไม่แน่นอนให้เป็นตัวเลข ได้กล่าวถึงความน่าจะเป็นของเหตุการณ์ที่เกิดขึ้น (A) ถ้ามี เหตุการณ์อีกเหตุการณ์หนึ่งเกิดมาแล้ว (B) สามารถ เขียนให้ อยู่ในรูปอย่างง่าย ดังในสมการที่ 5

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5)$$

$P(A|B)$ คือ ความน่าจะเป็นที่เหตุการณ์ A จะเกิดขึ้น ถ้าเหตุการณ์ B เกิดขึ้นแล้ว

$P(B|A)$ คือ ความน่าจะเป็นที่เหตุการณ์ B จะเกิดขึ้น ถ้าเหตุการณ์ A เกิดขึ้นแล้ว

$P(A)$ คือ ความน่าจะเป็นที่จะเกิดเหตุการณ์ A

$P(B)$ คือ ความน่าจะเป็นที่จะเกิด เหตุการณ์ B

วิธีการจำแนกหมวดหมู่โดยใช้หลักการความน่าจะเป็น เป็นการแก้ปัญหาแบบ Classification สามารถคาดการณ์ผลลัพธ์ และสามารถอธิบายได้ ทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร เพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละ ความสัมพันธ์ เป็นวิธีการจำแนกประเภทข้อมูลที่มี ประสิทธิภาพวิธีหนึ่ง โดยใช้ในการจำแนกหมวดหมู่เอกสารข้อความ (Text Classification) ได้ดี การทำงานไม่ซับซ้อนเหมาะกับกรณีของเซตตัวอย่างมีจำนวนมากและคุณสมบัติ (Attribute) ไม่ขึ้นต่อกัน โดย กำหนดให้ความน่าจะเป็นของข้อมูลที่จะเป็น

3.4.3 เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (SVM) [37] คือ ขั้นตอนวิธีการที่มีความรวดเร็วและเป็นอัลกอริทึมที่มีความสามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มของข้อมูลที่ถูกป้อนเข้าสู่กระบวนการทดสอบแบบให้ระบบเรียนรู้โดยเน้นไปที่เส้นแบ่งแยกกลุ่มของข้อมูลได้ดีที่สุด แนวความคิดของเทคนิควิธี SVM นั้นเกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางในฟีเจอร์สเปซ (Feature Space) ในลักษณะเชิงเส้นแบ่ง (Hyperplane) ที่เป็นเส้นตรงขึ้นมา เพื่อให้ทราบว่เส้นตรงที่แบ่งกลุ่มสองกลุ่มออกจากกันนั้นเส้นใดเป็นเส้นที่ดีที่สุด โดยการทำงานแบบ binary classification จะใช้สมการที่ 6 ในการสร้างแบบจำลอง

$$(\max w^T)x - b \quad (6)$$

w คือ เวกเตอร์ของค่าสัมประสิทธิ์

x คือ เวกเตอร์ของข้อมูล

b คือ ค่า *bias*

3.4.4 เทคนิคเพอร์เซปตรอนหลายชั้น (MLP) [38] เป็นโมเดลการเรียนรู้ของเครื่องประเภทหนึ่งที่จัดอยู่ในกลุ่มโครงข่ายประสาทเทียม (Neural Network) ประกอบไปด้วยชั้นของหน่วยประมวลผล (Processing Unit) เรียกว่า "เพอร์เซปตรอน" เรียงซ้อนกันหลายชั้น โมเดลนี้ได้รับแรงบันดาลใจจากโครงสร้างของระบบประสาทในสมองมนุษย์ โดยการทำงาน 3 ชั้น คือ ชั้นข้อมูลเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นข้อมูลออก (Output Layer) ซึ่งการทำงานจะเป็นการปรับค่าน้ำหนัก (Weight) ของการเชื่อมต่อระหว่างเพอร์เซปตรอนในแต่ละชั้น เพื่อลดค่าความผิดพลาด (Error) ของโมเดล โดยการใช้ สมการที่ 7 ในชั้นซ่อน และ สมการที่ 8 ในชั้นข้อมูลออก

$$\frac{z}{j} = + \sqrt{(i = 1)^n \frac{w}{ji} \left(\frac{x}{i} + \frac{b}{j} \right)} \quad (7)$$

$$\frac{y}{k} = \sum \left(\frac{y}{k} \right) \quad (8)$$

$\frac{z}{j}$ คือ ค่าอินพุตของเพอร์เซปตรอนตัวที่ j ในชั้นซ่อน

$\frac{w}{ji}$ คือ ค่าน้ำหนักของการเชื่อมต่อระหว่างเพอร์เซปตรอนตัวที่ i ในชั้นก่อนหน้ากับเพอร์เซปตรอนตัวที่ j ในชั้นปัจจุบัน

$\frac{x}{i}$ คือ ค่าอินพุตของเพอร์เซปตรอนตัวที่ i ในชั้นก่อนหน้า

$\frac{b}{j}$ คือ ค่าไบแอส (Bias) ของเพอร์เซปตรอนตัวที่ j ในชั้นซ่อน

$\frac{y}{k}$ คือ ค่าเอาต์พุตของเพอร์เซปตรอนตัวที่ k ในชั้นข้อมูลออก

Σ คือ ฟังก์ชันการกระตุ้น (Activation Function)

3.5 การแปลผลและการประเมินผล

การแปลผลและการประเมินผล เป็นขั้นตอนการแปลความหมาย การตีความและการประเมินผลลัพธ์ว่ามีความเหมาะสมหรือตรงกับวัตถุประสงค์ที่ต้องการหรือไม่ ซึ่งควรนำเสนอผลการวิเคราะห์ในรูปแบบที่ผู้ใช้งานสามารถเข้าใจได้ง่าย คือ วัดประสิทธิภาพของแบบจำลองที่สร้างขึ้น โดยใช้เทคนิคการวัดประสิทธิภาพแบบ 10-Fold Cross Validation [39] ซึ่งเป็นวิธีการวัดประสิทธิภาพที่ได้รับความนิยม โดยการแบ่งข้อมูลออกเป็น 10 กลุ่ม แต่ละกลุ่มมีปริมาณข้อมูลที่เท่ากัน โดยในแต่ละรอบการทดสอบจะใช้ข้อมูล 1 ชุด สำหรับการเรียนรู้ และใช้ข้อมูล 9 ชุดสำหรับทดสอบ ซึ่งการวัดประสิทธิภาพของแบบจำลองครั้งนี้จะทำการทดสอบ 10 รอบ ดังภาพที่ 3.29



ภาพประกอบที่ 3.29 การทดสอบแบบ 10-Fold cross validation

ส่วนค่าที่ใช้ในการวัดประสิทธิภาพของแบบจำลองที่สร้างขึ้นมาในครั้งนี้คือ ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure)

$$\text{Precision} = \frac{TP}{TP+FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

$$F - \text{Measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (11)$$

โดยที่ TP แทนจำนวนข้อมูลที่แบบจำลองทำนายถูกต้องในคลาส P

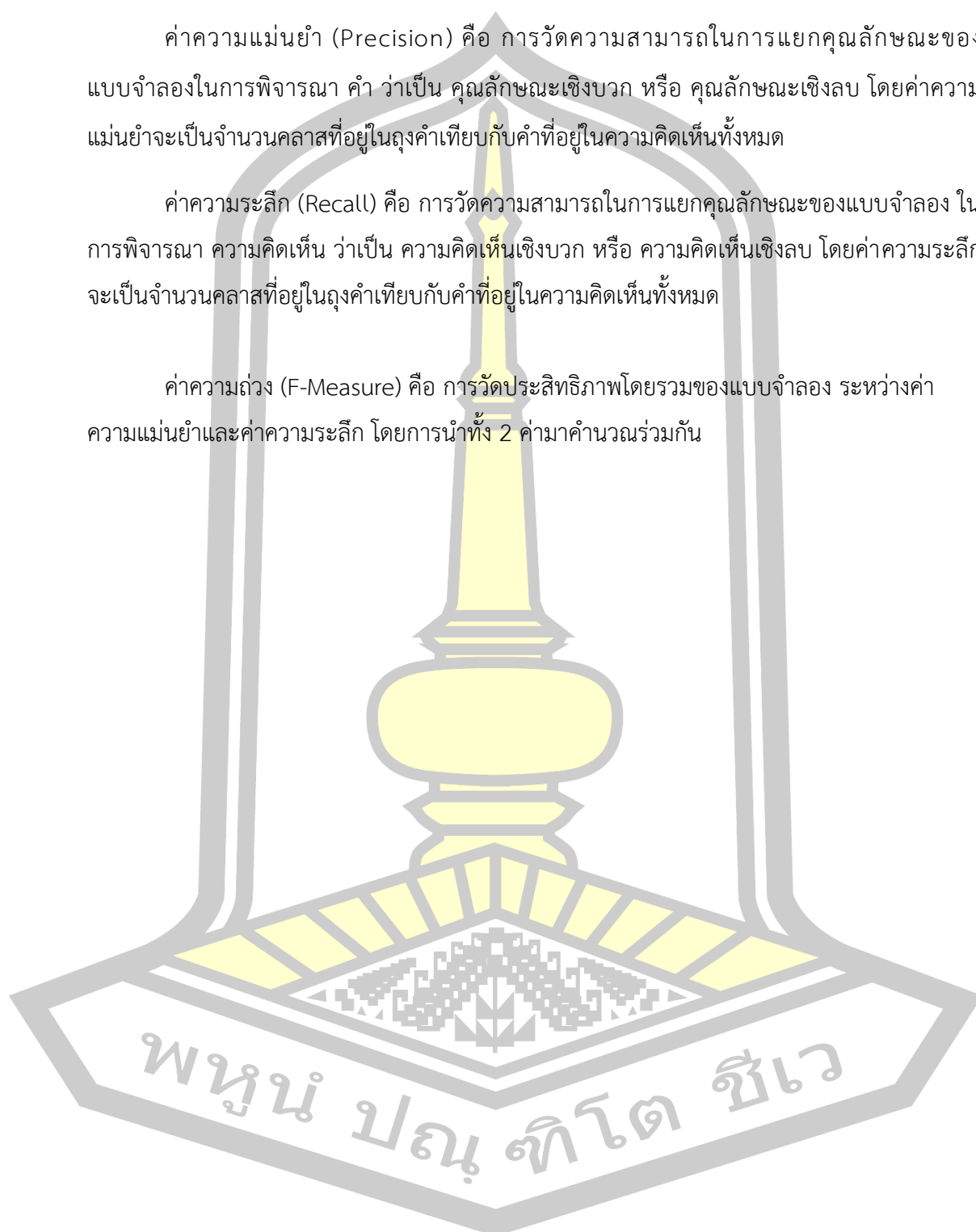
FP แทนจำนวนข้อมูลที่แบบจำลองทำนายผิดในคลาส P

FN แทนจำนวนข้อมูลที่แบบจำลองทำนายผิดในคลาส N

ค่าความแม่นยำ (Precision) คือ การวัดความสามารถในการแยกคุณลักษณะของแบบจำลองในการพิจารณา คำ ว่าเป็น คุณลักษณะเชิงบวก หรือ คุณลักษณะเชิงลบ โดยค่าความแม่นยำจะเป็นจำนวนคลาสที่อยู่ในกลุ่มคำเทียบกับคำที่อยู่ในความคิดเห็นทั้งหมด

ค่าความระลึก (Recall) คือ การวัดความสามารถในการแยกคุณลักษณะของแบบจำลอง ในการพิจารณา ความคิดเห็น ว่าเป็น ความคิดเห็นเชิงบวก หรือ ความคิดเห็นเชิงลบ โดยค่าความระลึกจะเป็นจำนวนคลาสที่อยู่ในกลุ่มคำเทียบกับคำที่อยู่ในความคิดเห็นทั้งหมด

ค่าความถ่วง (F-Measure) คือ การวัดประสิทธิภาพโดยรวมของแบบจำลอง ระหว่างค่าความแม่นยำและค่าความระลึก โดยการนำทั้ง 2 ค่ามาคำนวณร่วมกัน



บทที่ 4

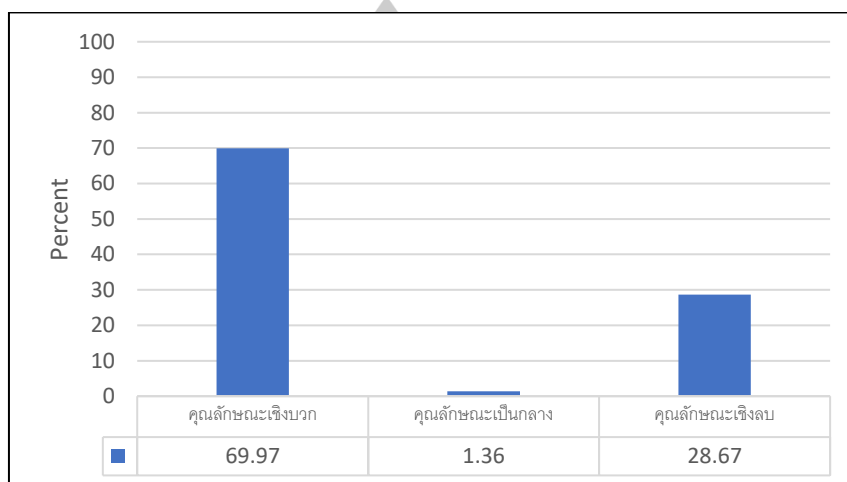
ผลการวิจัย

ในบทนี้เป็นการแสดงผลการวิจัยจากการศึกษาความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกและการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก โดยใช้หลักการของเหมืองความคิดเห็นที่มี 5 ขั้นตอน ซึ่งการรวบรวมข้อมูลความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกบนสื่อสังคมออนไลน์ต่าง ๆ ได้แก่ Twitter, Facebook, Pantip และ Tiktok ตั้งแต่วันที่ 1 มีนาคม ถึง 31 สิงหาคม พ.ศ. 2565 จำนวนทั้งหมด 2,265 ความคิดเห็น นำมาผ่านกระบวนการต่าง ๆ เช่น การทำความสะอาดคำ การทำดัชนี การแยกประเภทคำที่อยู่ในแต่ละประโยค การสร้างถ่วงคำ และกระบวนการอื่นๆ คงเหลืออยู่ทั้งหมด 3,045 คำ โดยการใช้พจนานุกรมภาษาไทยเพื่อแยกชนิดของคำ จากนั้นนำเอาคำที่สามารถบ่งบอกได้ถึงความรู้สึกที่เป็นความหมายในเชิงบวกและเชิงลบ แทนที่ด้วยตัวเลขโดยหากเป็นคำที่มีคุณลักษณะความรู้สึกเป็นเชิงบวกจะแทนค่าด้วย 1 และคำที่มีคุณลักษณะความรู้สึกเป็นเชิงลบจะแทนค่าด้วย -1 ส่วนคำอื่น ๆ ที่มีคุณลักษณะเป็นกลางหรือไม่สามารถบ่งบอกถึงความรู้สึกได้จะแทนค่าด้วย 0

การสร้างแบบจำลอง เพื่อจำแนกความรู้สึกผู้วิจัยได้นำเอาเทคนิคต่างๆที่ใช้กันอย่างแพร่หลาย โดยเลือกเอาทั้งหมด 4 เทคนิคได้แก่ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5), เทคนิคเนออีฟเบย์ (Naïve-Bayes), เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) นำมาสร้างเป็นแบบจำลองเพื่อจำแนกความรู้สึกของคนไทยเพื่อใช้ในการวิเคราะห์ข้อมูล และวัดประสิทธิภาพของแบบจำลองด้วยวิธี 10-Fold Cross Validation ซึ่งวัดประสิทธิภาพจากผลของการจำแนกกลุ่มข้อมูลและหาค่า ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) ซึ่งสามารถแสดงผลการวิจัยได้ ดังนี้

พูน ปณ จิตโต ชีเว

4.1. ผลการวิจัยความรู้สึกที่มีต่อยานยนต์พลังงานทางเลือก



ภาพประกอบที่ 4.1 การเปรียบเทียบความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก

จากภาพประกอบที่ 4.1 การวิจัยการวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมือนความคิดเห็น มีวัตถุประสงค์เพื่อศึกษาความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือก โดยการนำเอาความรู้สึกของคนไทยต่อยานยนต์พลังงานทางเลือกบนสื่อสังคมออนไลน์ ว่ามีความคิดเห็นไปในทิศทางใดซึ่งจะแบ่งออกเป็น 2 กลุ่ม ได้แก่ความคิดเห็นเชิงบวกและความคิดเห็นเชิงลบ โดยการเปรียบเทียบเทคนิคที่ใช้ในการสร้างแบบจำลองเพื่อจำแนกความคิดเห็นจากการทดลอง พบว่าคนไทยมีความรู้สึกเป็นเชิงบวกร้อยละ 69.97 ส่วนความรู้สึกเชิงลบอยู่ที่ร้อยละ 28.67 และความรู้สึกเป็นกลางร้อยละ 1.36

ตารางที่ 4.1 การจัดอันดับคุณลักษณะที่ถูกกล่าวถึง 10 อันดับ

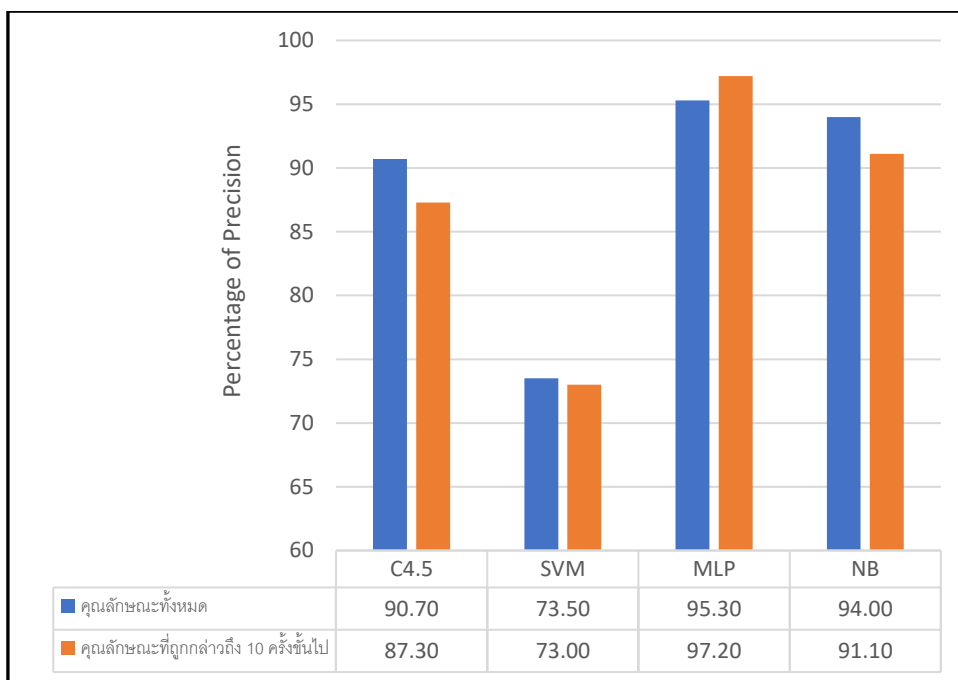
ลำดับ	คุณลักษณะเชิงบวก	จำนวน(ครั้ง)	คุณลักษณะเชิงลบ	จำนวน(ครั้ง)
1	ดี	696	ปัญหา	202
2	สวย	510	แพง	195
3	ชอบ	350	ง่าย	78
4	ขอบคุณ	112	ซ่อม	51
5	สนใจ	101	ร้อน	50
6	สนุก	93	ต่ำ	48
7	คุ้ม	88	ภาษี	46
8	คุ้มค่า	87	พัง	45
9	พัฒนา	84	ตาย	36
10	ประหยัด	76	กลัว	35

จากตารางที่ 4.1 แสดงการเปรียบเทียบค่าที่มีคุณลักษณะเชิงบวกและเชิงลบที่ถูกกล่าวถึงมากที่สุด 10 อันดับ โดยคุณลักษณะเชิงบวกที่ถูกกล่าวมากที่สุดคือ “ดี” ทั้งหมด 696 ครั้ง และค่าที่มีคุณลักษณะเชิงลบที่ถูกกล่าวถึงมากที่สุดคือ “ปัญหา” จำนวน 202 ครั้ง

4.2. ผลการวิจัยประสิทธิภาพของแบบจำลอง

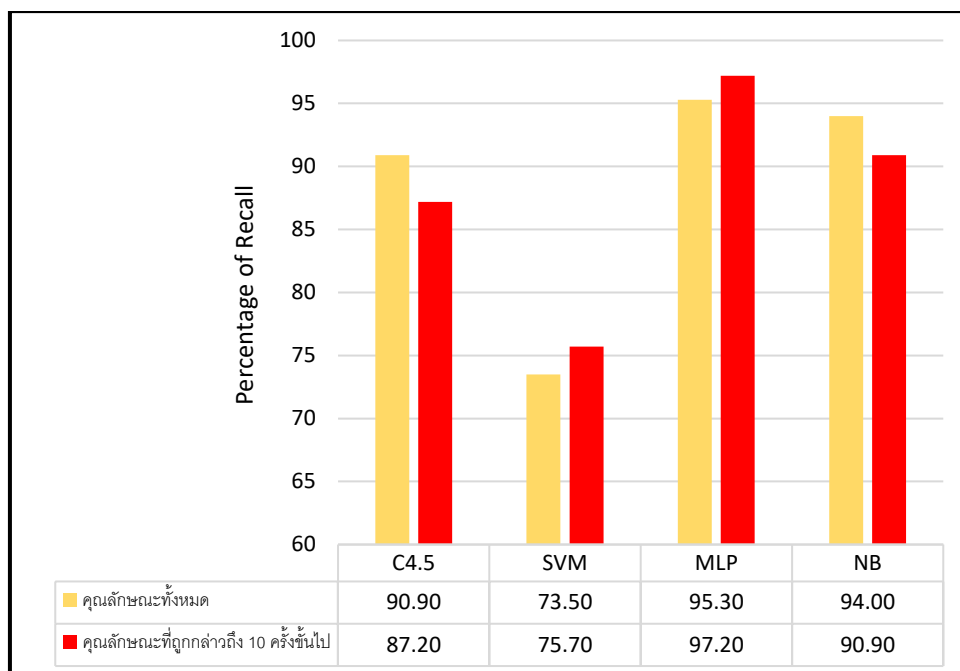
ในการวิจัยการวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเครื่องมือความคิดเห็น โดยการนำเอาเทคนิคการทำเหมืองความคิดเห็นมาใช้ในการวิเคราะห์ข้อมูลโดยใช้เทคนิคทั้งหมด 4 เทคนิคได้แก่ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5) เทคนิคเนออีฟเบย์ (Naïve-Bayes) เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ได้ใช้เครื่องมือเป็นโปรแกรม Weka เวอร์ชัน 3.8.6 ซึ่งได้รวบรวมข้อมูลจากสื่อสังคมออนไลน์ต่าง ๆ เช่น Facebook, YouTube, Pantip และ TikTok ตั้งแต่วันที่ 1 มีนาคม ถึง 31 สิงหาคม พ.ศ. 2565 จำนวนทั้งหมด 2,265 ชุดข้อมูล หลังจากผ่านกระบวนการต่าง ๆ ก่อนจะนำเข้าสู่แบบจำลองทั้ง 4 แบบ ได้มีการแบ่งชุดข้อมูลออกเป็นจำนวน 2 คลาส ได้แก่ คลาส P มีจำนวน 231 ข้อมูล คลาส N จำนวน 209 ข้อมูล รวมทั้งหมดเป็น 440 คลาส ซึ่งเมื่อคัดเลือกเอาคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้ง ขึ้นไปจะได้ คลาส P มีจำนวน 173 ข้อมูล คลาส N จำนวน 110 ข้อมูล รวม 283 คลาส ทั้งสองคลาสมีจำนวนข้อมูลที่ใกล้เคียงกัน จึงไม่ต้องปรับสมดุลหรือแก้ปัญหาคือข้อมูลที่สมดุลกัน และได้สร้างแบบจำลองด้วย เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5), เทคนิคเนออีฟเบย์ (Naïve-Bayes), เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ได้ใช้กระบวนการทดสอบประสิทธิภาพ 10 Fold Cross Validation แบ่งกลุ่มข้อมูลออกเป็น 10 ชุด โดย 9 ชุดข้อมูลเป็นชุดสำหรับทดสอบ และ 1 ชุดข้อมูลเป็นชุดสำหรับการเรียนรู้ โดยวัดประสิทธิภาพจากค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) โดยได้ผลการวิจัยจากทั้ง 2 รูปแบบ ดังภาพประกอบที่ 4.2 ภาพประกอบที่ 4.3 และ ภาพประกอบที่ 4.4

พูน ปณ จิต ชเว



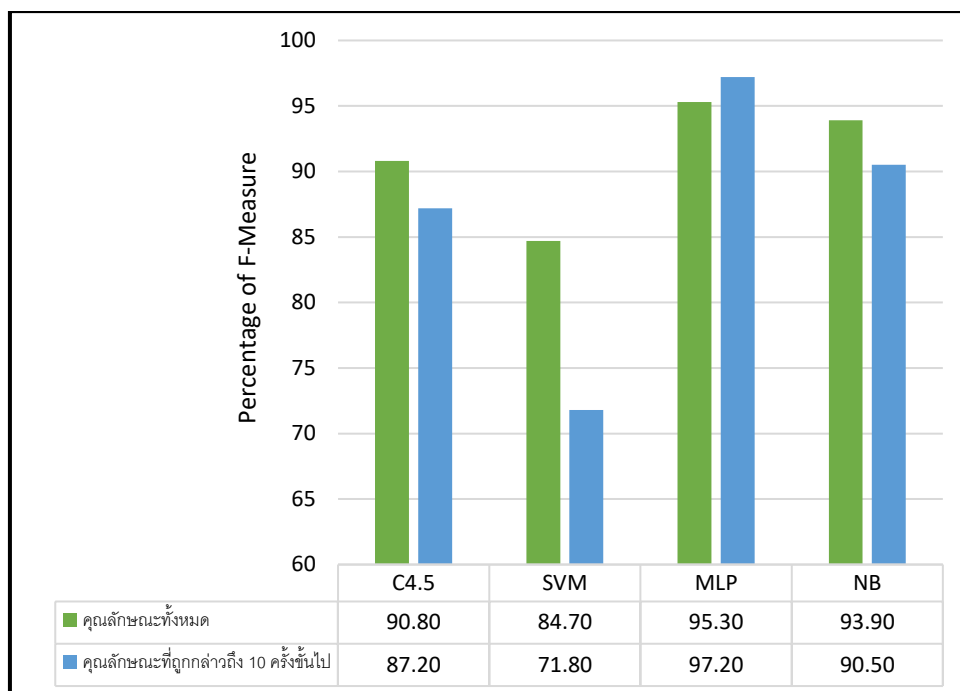
ภาพประกอบที่ 4.2 การเปรียบเทียบค่าความแม่นยำ

จากภาพประกอบที่ 4.2 แสดงค่าความแม่นยำของแบบจำลองเพื่อการจำแนกความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกด้วย เทคนิคเพอร์เซปตรอนหลายชั้น (MLP) เทคนิคเนออบี (NB) เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (C4.5) และเทคนิคซัพพอร์ตเวกเตอร์แมทชีน (SVM) การทดลองที่นำเอาคุณลักษณะทั้งหมดมาสร้างแบบจำลอง พบว่า เทคนิค MLP มีค่าความแม่นยำมากที่สุด โดยมีค่าความแม่นยำที่ร้อยละ 95.30 รองลงมาเป็น NB มีค่าความแม่นยำที่ร้อยละ 94.00 เทคนิค C4.5 มีค่าความแม่นยำที่ร้อยละ 90.70 และเทคนิค SVM มีค่าความแม่นยำน้อยที่สุดที่ร้อยละ 73.50 ส่วนการทดลองที่ใช้คุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไปมาทำการทดลอง พบว่า เทคนิค MLP มีค่าความแม่นยำมากที่สุดที่ร้อยละ 97.20 รองลงมาเป็น เทคนิค NB มีค่าความแม่นยำที่ร้อยละ 91.10 เทคนิค C4.5 มีค่าความแม่นยำที่ร้อยละ 87.30 และเทคนิค SVM มีค่าความแม่นยำน้อยที่สุดที่ร้อยละ 73.00 จากการวิเคราะห์การใช้คุณลักษณะทั้งหมดและการใช้คุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิค MLP มีค่าความแม่นยำเพิ่มขึ้นร้อยละ 1.90 เทคนิค C4.5 มีค่าความแม่นยำลดลงร้อยละ 3.40 เทคนิค NB มีค่าความแม่นยำลดลงร้อยละ 2.90 ส่วนเทคนิค SVM มีค่าความแม่นยำลดลงร้อยละ 0.50



ภาพประกอบที่ 4.3 การเปรียบเทียบค่าความระลึก

จากภาพประกอบที่ 4.3 แสดงการเปรียบเทียบค่าความระลึกของแบบจำลองเพื่อจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกด้วย เทคนิคเพอร์เซปตรอนหลายชั้น (MLP) เทคนิคเอนีฟเบย์ (NB) เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (C4.5) และเทคนิคซัพพอร์ทเวกเตอร์แมทชีน (SVM) การทดลองนำเอาคุณลักษณะทั้งหมดมาสร้างแบบจำลอง พบว่า เทคนิค MLP เป็นเทคนิคที่ให้ค่าความระลึกมากที่สุดที่ร้อยละ 95.30 รองลงมาเป็นเทคนิค NB มีค่าความระลึกที่ร้อยละ 94.00 เทคนิค C4.5 มีค่าความระลึกที่ร้อยละ 90.90 และเทคนิค SVM มีค่าความระลึกน้อยที่สุด มีค่าความระลึกที่ร้อยละ 73.50 ส่วนการทดลองสร้างแบบจำลองจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิค MLP ที่มีค่าความระลึกมากที่สุดที่ร้อยละ 97.20 เทคนิค NB รองลงมา มีค่าความระลึกที่ร้อยละ 90.90 เทคนิค C4.5 มีค่าความระลึกที่ร้อยละ 87.20 และเทคนิค SVM มีค่าความระลึกน้อยที่สุดโดยมีค่าความระลึกที่ร้อยละ 75.70 จากการวิเคราะห์การใช้คุณลักษณะทั้งหมด และการใช้คุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิค MLP มีค่าความระลึกเพิ่มขึ้นร้อยละ 1.90 เทคนิค SVM มีค่าความระลึกเพิ่มขึ้นร้อยละ 2.20 ในขณะที่ เทคนิค C4.5 และ เทคนิค NB มีค่าความระลึกลดลงร้อยละ 3.1 และร้อยละ 3.7 ตามลำดับ



ภาพประกอบที่ 4.4 การเปรียบเทียบค่าความถ่วง

จากภาพประกอบที่ 4.4 แสดงการเปรียบเทียบค่าความถ่วงของแบบจำลองเพื่อจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกด้วย เทคนิคเพอร์เซปตรอนหลายชั้น (MLP) เทคนิคเนอโฟเบย์ (NB) เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (C4.5) และเทคนิคซัพพอร์ทเวกเตอร์แมทซิน (SVM) พบว่า เทคนิค MLP ที่มีค่าความถ่วงมากที่สุดจากทั้ง 4 เทคนิค โดยมีค่าความถ่วงอยู่ที่ร้อยละ 95.30 รองลงมาคือ NB มีค่าความถ่วงที่ร้อยละ 93.90 เทคนิค C4.5 มีค่าความถ่วงที่ร้อยละ 90.80 และเทคนิค SVM มีค่าความถ่วงที่ร้อยละ 84.70 ส่วนการทดลองสร้างแบบจำลองจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิค MLP ที่มีค่าความถ่วงมากที่สุด โดยมีค่าความถ่วงอยู่ที่ร้อยละ 97.20 รองลงมาเป็น เทคนิค NB มีค่าความถ่วงที่ร้อยละ 90.50 เทคนิค C4.5 มีค่าความถ่วงที่ร้อยละ 87.20 และเทคนิคที่ให้ค่าความถ่วงน้อยที่สุดคือ เทคนิค SVM มีค่าความถ่วงที่ร้อยละ 71.80 จากการวิเคราะห์การใช้คุณลักษณะทั้งหมดและการใช้คุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิค MLP มีความระลึกเพิ่มขึ้นร้อยละ 1.90 ส่วนเทคนิคอื่น ๆ มีความถ่วงลดลงทั้งหมด โดยเทคนิค NB มีค่าความถ่วงลดลงร้อยละ 3.40 เทคนิค C4.5 มีค่าความถ่วงลดลงร้อยละ 3.60 และเทคนิค SVM มีค่าความถ่วงลดลงร้อยละ 12.90

จากการทดลองครั้งนี้ในส่วนของคุณลักษณะที่ถูกกล่าวถึงทั้งหมด พบว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) เป็นเทคนิคที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก โดยมีค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) เท่ากันที่ร้อยละ 95.30 รองลงมาได้แก่ เทคนิคเนอโฟเบย์

(Naïve-Bayes) มีค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) เท่ากันที่ร้อยละ 94.00 ค่าความถ่วง (F-Measure) ที่ร้อยละ 93.90 รองลงมาคือ เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5) มีค่าความแม่นยำ (Precision) ที่ร้อยละ 90.70 ค่าความระลึก (Recall) ที่ร้อยละ 90.90 และค่าความถ่วง (F-Measure) ที่ร้อยละ 90.80 และเทคนิคที่มีประสิทธิภาพน้อยที่สุดในการทดลองครั้งนี้คือ เทคนิคซ์พอดเว็คเตอร์แมทชีน (Support Vector Machine) ซึ่งจากการทดลองซ้ำหลายครั้ง พบว่า เทคนิคซ์พอดเว็คเตอร์แมทชีน (SVM) มีความสามารถในการจำแนกคุณลักษณะเชิงบวก หรือ คุณลักษณะเชิงลบ ได้น้อยที่สุด โดยมีค่าความแม่นยำ (Precision) ที่ร้อยละ 73.50 ค่าความระลึก (Recall) ที่ร้อยละ 73.50 และค่าความถ่วง (F-Measure) ที่ร้อยละ 84.30 ส่วนการทดลองสร้างแบบจำลองเพื่อจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก โดยใช้คุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) เป็นเทคนิคที่มีประสิทธิภาพสูงสุด โดยมีค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) เท่ากันที่ร้อยละ 97.20 รองลงมาคือ เทคนิคเนอ์ฟเบย์ (Naïve-Bayes) มีค่าความแม่นยำ (Precision) ที่ร้อยละ 91.10 ค่าความระลึก (Recall) ที่ร้อยละ 90.90 ค่าความถ่วง (F-Measure) ที่ร้อยละ 90.50 ต่อมาเป็น เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 (Decision Tree C4.5) มีค่าความแม่นยำ (Precision) ที่ร้อยละ 87.30 ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) ที่ร้อยละ 87.20 และเทคนิคซ์พอดเว็คเตอร์แมทชีน (Support Vector Machine) เป็นเทคนิคที่มีประสิทธิภาพน้อยที่สุดจากการทดลองคัดเลือกคุณลักษณะที่ถูกกล่าวถึง 10 ครั้งขึ้นไป โดยมีค่าความแม่นยำ (Precision) ที่ร้อยละ 73.00 ค่าความระลึก (Recall) ที่ร้อยละ 75.70 และค่าความถ่วง (F-Measure) ที่ร้อยละ 71.80



บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

งานวิจัยนี้ได้นำเอากระบวนการในการคัดเลือกคุณลักษณะด้วยการให้น้ำหนักค่าในการสร้างแบบจำลองที่ได้ประสิทธิภาพสำหรับจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก โดยหลักการของการทำเหมืองข้อความ เพื่อสร้างและเปรียบเทียบประสิทธิภาพของแบบจำลองความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก และเพื่อให้ได้แบบจำลองที่มีประสิทธิภาพที่สุดในการจำแนกความคิดเห็นจึงได้นำเอากระบวนการคัดเลือกคุณลักษณะโดยการให้น้ำหนักเพื่อใช้ในการสร้างแบบจำลอง

ขั้นตอนในการทำเหมืองข้อมูลมีทั้งหมด 5 ขั้นตอน โดยการรวบรวมความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกจากสื่อสังคมออนไลน์ ได้แก่ Twitter, Facebook, Pantip และ Tiktok ตั้งแต่วันที่ 1 มีนาคม ถึง 31 สิงหาคม พ.ศ. 2565 จำนวนทั้งหมด 2,265 ความคิดเห็น จากนั้นนำมาผ่านกระบวนการเตรียมข้อมูลด้วยการทำความสะอาดข้อความที่ได้จากสื่อสังคมออนไลน์

การทำดัชนีข้อมูล นำเอาข้อมูลมาทำการแยกคำและค้นหาชนิดของคำที่อยู่ในประโยค เช่น คำนาม (Noun) คำสรรพนาม (Pronoun) คำกริยา (Verb) คำวิเศษณ์ (Adverb) คำบุพบท (Prepositions) คำสันธาน (Conjunctions) เป็นต้น เนื่องจากประโยคต่าง ๆ ที่ใช้ในการพูด การสื่อสาร หรือแสดงความคิดเห็นนั้นล้วนมาจากการนำเอาคำต่าง ๆ เหล่านี้มาประกอบกัน จากนั้นทำการสร้างถุงคำและกำหนดคลาสของคำ

การสร้างถุงคำ (Bag of word) เป็นการนำคำที่แยกประเภทจากความคิดเห็นหรือข้อความที่ผ่านกระบวนการทำความสะอาดแล้วมาเพียงแค่ 1 คำ ไม่ว่าจะในกรณีที่ซ้ำกันหรือไม่ซ้ำกันก็ตาม จากทั้งหมด 7,223 คำ คงเหลือ 3,044 คำ โดยการพิจารณาเลือกใช้คำที่สมบูรณ์เท่านั้น จากนั้นทำการคัดเลือกเอา คำกริยา คำคุณศัพท์ คำวิเศษณ์ และคำวิเศษณ์วิเศษ ที่สามารถบ่งบอกถึงความรู้สึกได้ ทำการเพิ่มคุณสมบัติให้กับคำที่มีคุณลักษณะเชิงบวกและเชิงลบ โดยแทนค่าของคำที่มีคุณลักษณะเชิงบวกเป็น 1 ส่วนคำที่มีคุณลักษณะเชิงลบแทนค่าเป็น -1 และแทนค่าของคำที่มีคุณลักษณะเป็นกลางหรือไม่ปรากฏคุณลักษณะด้วย 0 โดยจะได้เป็นชุดข้อมูลใหม่ที่มี 440 คลาส แบ่งออกเป็น 2 กลุ่ม โดยกลุ่มแรก คือ คุณลักษณะที่ถูกกล่าวถึงทั้งหมด แบ่งเป็น 2 รูปแบบ ได้แก่ คุณลักษณะที่เป็นเชิงบวก 231 คลาส และคุณลักษณะที่เป็นเชิงลบ 209 คลาส กลุ่มที่ 2 คือ คุณลักษณะที่ถูกกล่าวถึง 10 ครั้งขึ้นไป จำนวน 283 คลาส แบ่งเป็น 2 รูปแบบ ได้แก่ คุณลักษณะที่เป็นเชิงบวก 173 คลาส และคุณลักษณะที่เป็นเชิงลบ 110 คลาส

การสร้างแบบจำลอง เป็นการเอาเทคนิคในการสร้างเหมืองข้อความ มาใช้ในการสร้างแบบจำลองเพื่อจำแนกรู้สึกด้วยกันทั้งหมด 4 เทคนิค จากนั้นนำเอาเทคนิคที่ใช้ในการวัดประสิทธิภาพของการคัดเลือกคุณลักษณะทำการเปรียบเทียบแบบจำลองทั้ง 4 เทคนิค ที่ใช้สร้างแบบจำลองเพื่อจำแนกรู้สึก วิเคราะห์ และวัดประสิทธิภาพของแบบจำลอง ด้วยวิธีการ 10 Fold Cross Validation โดยการหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure)

5.1 สรุปผล

วัตถุประสงค์งานวิจัยครั้งนี้เพื่อศึกษาความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกและเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกสามารถสรุปได้ ดังนี้

จากการศึกษาความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก สรุปได้ว่าคนไทยร้อยละ 69.97 มีความคิดเห็นเชิงบวกต่อยานยนต์พลังงานทางเลือก และคนไทยร้อยละ 28.67 มีความคิดเห็นในเชิงลบต่อยานยนต์พลังงานทางเลือก

จากการเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกด้วยคุณลักษณะทั้งหมด สรุปได้ว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) เป็นเทคนิคที่มีประสิทธิภาพมากที่สุดในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก มีค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) เท่ากันที่ร้อยละ 95.30 ส่วนเทคนิคที่มีประสิทธิภาพน้อยที่สุดในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก คือ เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) โดยมีค่าความแม่นยำ (Precision) ที่ร้อยละ 73.50 ค่าความระลึก (Recall) ที่ร้อยละ 73.50 และค่าความถ่วง (F-Measure) ที่ร้อยละ 84.30

จากการเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกด้วยคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป สรุปได้ว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) เป็นเทคนิคที่มีประสิทธิภาพมากที่สุดในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) เท่ากันที่ร้อยละ 97.20 ส่วนเทคนิคที่มีประสิทธิภาพน้อยที่สุดในการจำแนกรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก คือ เทคนิคซัพพอร์ทเวกเตอร์แมชชีน (Support Vector Machine) โดยมีค่าความแม่นยำ (Precision) ที่

ร้อยละ 73.00 ค่าความระลึก (Recall) ที่ร้อยละ 75.70 และค่าความถ่วง (F-Measure) ที่ร้อยละ 71.80

จากการทดลองใช้คุณลักษณะทั้งหมดและคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป สามารถสรุปได้ว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) มีค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure) เพิ่มขึ้นที่ร้อยละ 1.90 ส่วนเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) มีค่าความระลึกเพิ่มขึ้นร้อยละ 2.20 ค่าความแม่นยำลดลงร้อยละ 0.50 และค่าความถ่วงลดลงร้อยละ 12.90

5.2 อภิปรายผล

จากการวิจัย การวิเคราะห์ความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกด้วยเหมืองความคิดเห็น ในครั้งนี้ทำให้พบว่า เทคนิคที่ใช้ในการวิจัย เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ได้ให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความแม่นยำอยู่ที่ร้อยละ 95.30 รวมถึงค่าความระลึก และค่าความถ่วงที่ร้อยละ 95.30 เช่นกัน ซึ่งต่างจากงานวิจัยของ Davis และคณะ [29] ที่ได้ผลลัพธ์จากการทดลองเป็นเทคนิค ต้นไม้ตัดสินใจ ที่ให้ผลลัพธ์ที่ดีที่สุดโดยมีค่าความแม่นยำที่ 90.00% จากการวิจัย การสร้างสถานการณ์โดยอัตโนมัติจากคลังข้อความ : กรณีศึกษาเกี่ยวกับยานพาหนะไฟฟ้า ส่วนงานวิจัยของ กิตติศักดิ์ ขำจิตร และคณะ [40] ที่ได้ใช้เทคนิคเทคนิคเอนีฟเบย์ เทคนิคต้นไม้ตัดสินใจ เทคนิคป่าแบบสุ่ม และเทคนิคการเรียนรู้เชิงลึก พบว่าได้ผลลัพธ์ต่างกัน โดยเทคนิคที่มีประสิทธิภาพสูงสุด คือ เทคนิคการเรียนรู้เชิงลึก มีค่าความแม่นยำร้อยละ 95.11 ค่าความถ่วงร้อยละ 97.49 และค่าความไวร้อยละ 99.11 ในการวิจัย ประสิทธิภาพของเทคนิคเหมืองข้อมูลสำหรับพยากรณ์การเกิดโรคหลอดเลือดในสมอง และนั่นหนังสือ คู่มือการวิจัยการขุดความคิดเห็นและการวิเคราะห์ข้อความในงานวรรณกรรมและโซเชียลมีเดีย [41] ได้เสนอ เทคนิคการเรียนรู้เชิงลึก ว่าเป็นเทคนิคที่เหมาะสมในการทำเหมืองความคิดเห็นของยูทูป ซึ่งสอดคล้องกับเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ซึ่งเป็นเทคนิคต้นแบบที่พัฒนาไปสู่เทคนิคการเรียนรู้เชิงลึกจึงได้ผลลัพธ์ที่ดีที่สุดจาก 4 เทคนิคที่ใช้ในการวิจัยครั้งนี้

5.3 ข้อเสนอแนะ

ในการวิจัยครั้งนี้ผู้วิจัยได้นำเอาเทคนิคเหมืองข้อมูลเพื่อใช้ในการกระบวนการคัดเลือกคุณลักษณะด้วยการแทนค่ามาใช้ในการสร้างแบบจำลองเพื่อจำแนกคุณลักษณะความรู้สึกของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก โดยข้อมูลที่ใช้เป็นข้อมูลในช่วงที่มีการนำเข้ายานยนต์พลังงานทางเลือกในช่วงแรก ๆ ของประเทศไทย จึงทำให้มีความคิดเห็นของคนไทยถือว่าที่ยังไม่เป็นที่ทราบ

กันในวงกว้างหรือเฉพาะผู้ที่สนใจเท่านั้น แต่ในปัจจุบันยานยนต์พลังงานทางเลือกถูกนำเข้ามาและจำหน่ายภายในประเทศไทยหลากหลายยี่ห้อและมีเงื่อนไขในการใช้งานของแต่ละรุ่นแตกต่างกันไป การหาข้อมูลความคิดเห็นที่ได้จึงอาจจะมีคุณลักษณะที่มากกว่าการใช้ข้อมูลในช่วงแรกที่มีการนำเข้า ยานยนต์พลังงานทางเลือก นอกจากนี้การตั้งค่าพารามิเตอร์สำหรับการสร้างแบบจำลองจากเทคนิคต่าง ๆ ในโปรแกรม Weka ยังอยู่ในรูปแบบดั้งเดิม (default) หากมีการปรับแต่งค่าในการตั้งค่าที่มีความละเอียดมากขึ้น อาจจะได้ผลลัพธ์ที่แตกต่างออกไป

การประยุกต์ใช้เป็นแนวทางในการพัฒนาระบบสารสนเทศ เพื่อตอบสนองความต้องการที่มากขึ้นหรือค้นหาวิธีการใหม่ ๆ ที่ให้ผลลัพธ์ที่ต่างออกไปในการวิจัย สามารถศึกษาได้จากวิธีการเตรียมข้อมูล โดยใช้รูปแบบอื่น ๆ ที่ไม่ใช่ ภาษา python หรือ Library PyThaiNLP เพื่อให้ได้กระบวนการและวิธีการใหม่ ๆ ในการทำเหมืองความคิดเห็น สำหรับใช้ประกอบในการวางแผน และบริหารจัดการในด้านอื่น ๆ เช่น การประชาสัมพันธ์เพื่อให้ความรู้แก่ผู้ที่สนใจในยานยนต์พลังงานทางเลือก ใช้เป็นข้อมูลในการตัดสินใจประกอบการเลือกซื้อยานยนต์พลังงานทางเลือก หรือใช้เป็นข้อมูลเพื่อให้กลุ่มเป้าหมายรับรู้และเข้าใจในกรณีเป็นที่ปรึกษาด้านการขาย เป็นต้น

สำหรับการวิจัยครั้งต่อไป แนะนำให้เพิ่มเทคนิคอื่น ๆ หรือรูปแบบการเตรียมข้อมูลและกำหนดคุณลักษณะในแบบอื่นให้กับคำ เนื่องจาก Library PyThaiNLP ที่ผู้วิจัยเลือกใช้นั้นอยู่ช่วงระหว่างการพัฒนาและปรับปรุงจึงทำให้มีจุดบกพร่องที่ต้องแก้ไขอยู่หลายครั้ง หากในอนาคตมีเครื่องมืออื่นที่สามารถใช้งานได้เหมือนกันหรือดีกว่า Library PyThaiNLP ก็สามารถนำมาใช้ควบคู่หรือเปรียบเทียบในการในการวิจัย และข้อมูลที่ใช้ในการวิจัยครั้งต่อไปควรจะเป็นข้อมูลที่อยู่ในช่วงหลังจากมีการประชาสัมพันธ์และการสนับสนุนให้คนไทยหันมาใช้ยานยนต์พลังงานทางเลือกแล้ว เพื่อจะได้ความรู้สึกรักของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกในรูปแบบที่มีการรับรู้ของคนไทยมากขึ้นแล้ว



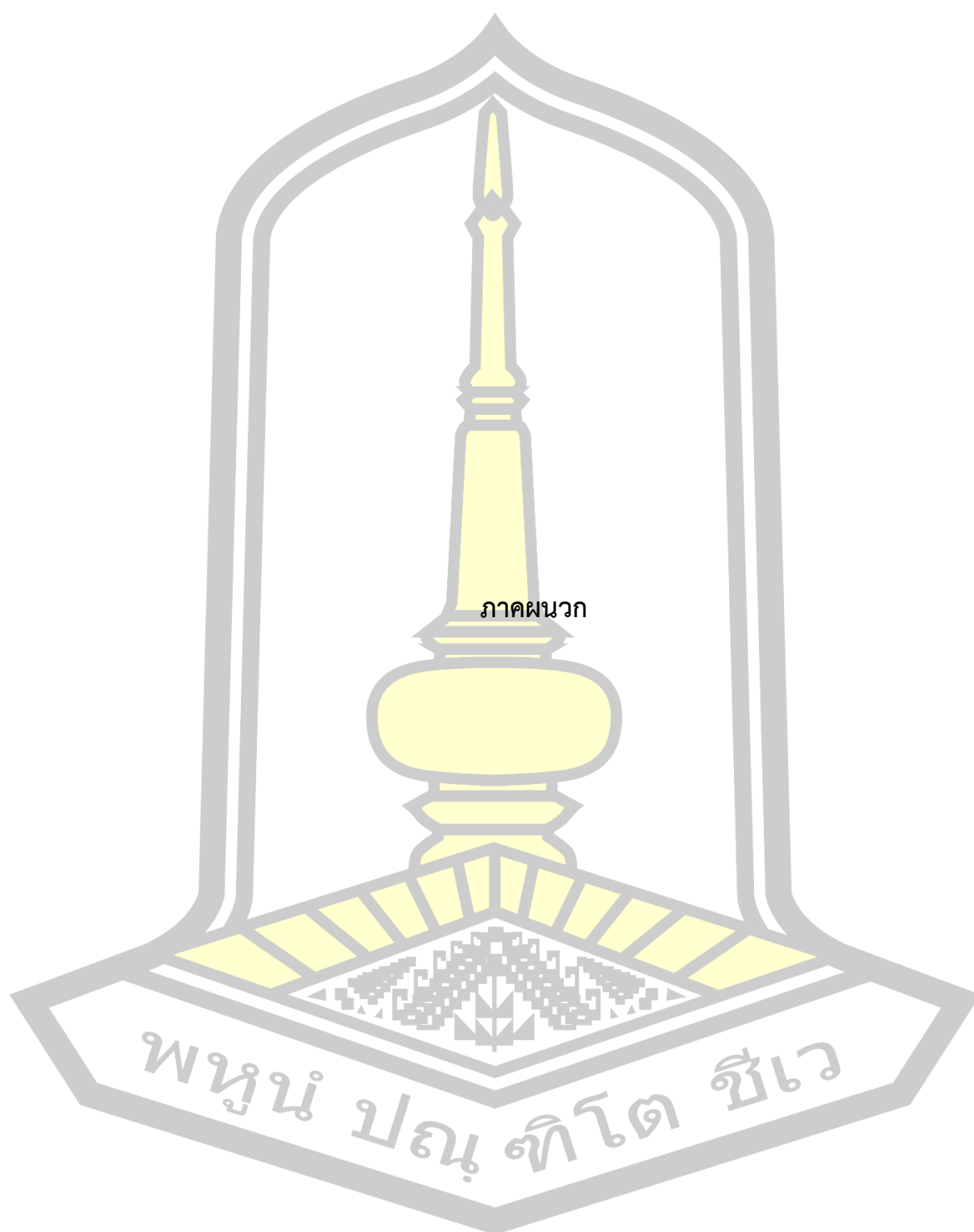
บรรณานุกรม

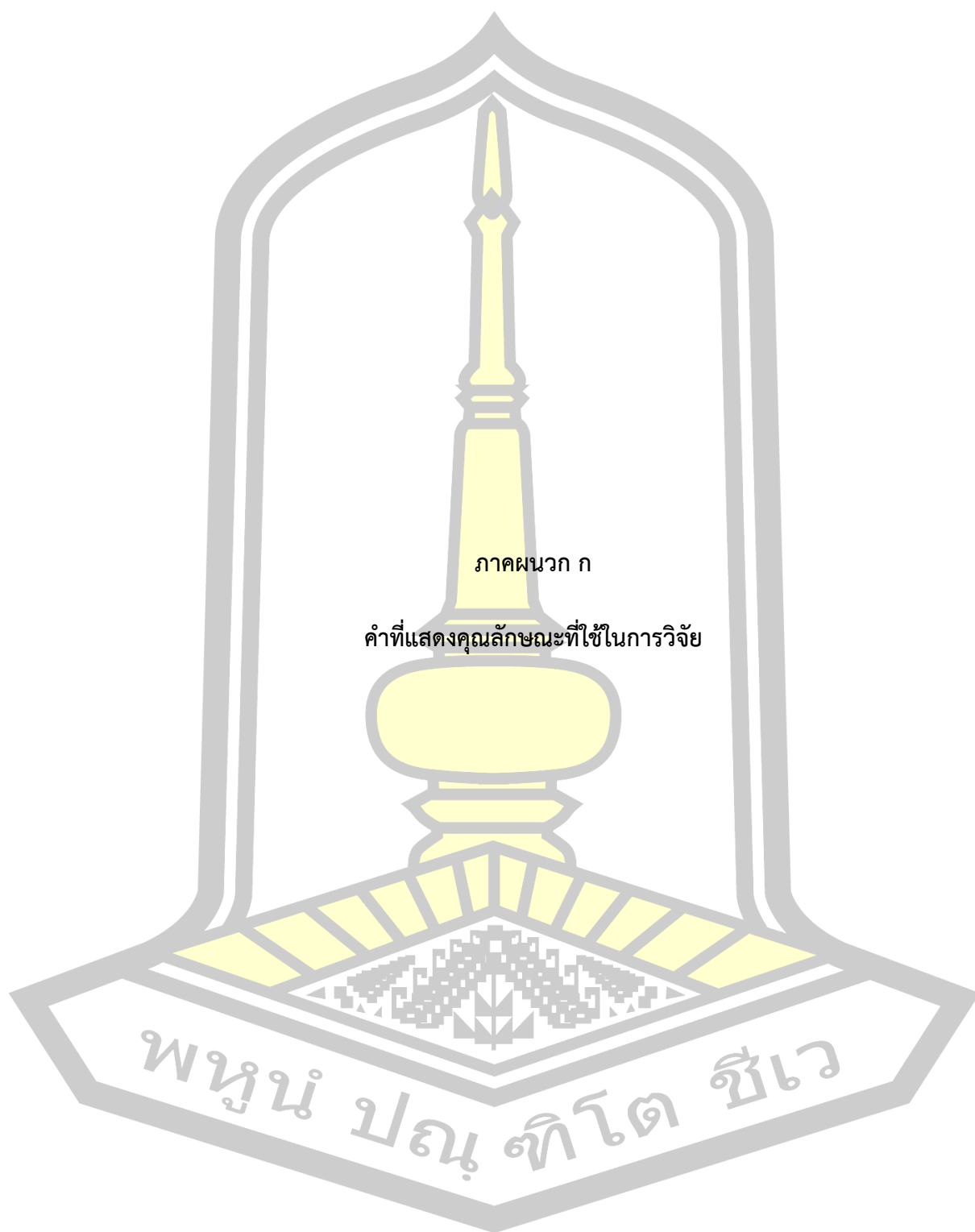
- [1] L. Schramm, "Some differences, many similarities: comparing Europe's responses to the 1973 oil crisis and the 2022 gas crisis," *European Political Science Review*, pp. 1-16, 2023.
- [2] บริษัท เนชั่น กรุ๊ป (ไทยแลนด์) จำกัด (มหาชน). (2023). เหตุผลที่คนไทยใช้ EV มากขึ้น ค่าชาร์จถูกกว่าน้ำมัน เป็นมิตรกับสิ่งแวดล้อม. Available: <https://www.posttoday.com/smart-city/704607>
- [3] M Report Co. Ltd. (2023). คนไทยเปิดรับรถยนต์ xEV มากขึ้น พร้อมก้าวสู่ยุคพลังงานทางเลือกใหม่. Available: <https://www.mreport.co.th/news/statistic-and-ranking/309-Poll-shows-Thai-people-have-more-interest-on-xEV-cars>
- [4] บริษัท บางกอก มีเดีย แอนด์ บรอดคาสติ้ง จำกัด. (2022). รถยนต์ไฟฟ้า ชั้ลยุ้นน้ำท่วม-ชาร์จไฟขณะฝนตก ปลอดภัยหรือไม่. Available: <https://www.pptvhd36.com/automotive/carcare/176850>
- [5] B. Liu, *Sentiment analysis and opinion mining*. Switzerland: Springer Nature, 2022.
- [6] Pramroj Namwong and Suphathep Sathitman, "Classifying opinions on restaurant service use on social networks in Ubon Ratchathani Province using opinion mining techniques.," *Journal of Management and Development Ubon Ratchathani Rajabhat University*, vol. 9, no. 2, pp. 1-14, 2022.
- [7] Bowornlertsutee Pisit and Paireekreng Worapat, "Sentiment Analysis Techniques of Online Product Reviews," *Rajamangala University of Technology Srivijaya Research Journal*, vol. 14, no. 3, pp. 755-769 2022.
- [8] A. Abayomi-Alli, O. Abayomi-Alli, S. Misra, and L. Fernandez-Sanz, "Study of the Yahoo-Yahoo Hash-Tag tweets using sentiment analysis and opinion mining algorithms," *Information*, vol. 13, no. 3, p. 22, 2022.
- [9] N. News, ""Types of Alternative Fuel Vehicles", " Available: <https://blog.nationwide.com/vehicle/vehicle-technology-trends/alternative-fuel-vehicles/>, pp. 1-3, 1 September 2021.

- [10] นุชนาฏ ปิ่นเมือง และ จารีย์ ทองคำ, "การจำแนกความคิดเห็นของคนไทยเกี่ยวกับสื่อออนไลน์ โดยใช้การทำเหมืองข้อความ," วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม, vol. 37, no. 3, 2018.
- [11] Igor Mishkovski, "Educational Data Mining: Case Study for Predicting Student Dropout in Higher Education," *ResearchGate*, no. 1, p. 5, 2015.
- [12] G. H. John, *Enhancements to the data mining process*. stanford university, 1997.
- [13] K. Han, Pei, Jaiwei, Micheline, Jian,, "Data Mining : Concepts and Techniques (3rd ed.), in *Data Mining: Morgan Kaufmann*". , 2011.
- [14] R. Javed Mehedi Shamrat, Rumes, K. M. Hasib, A. Yadav, and A. H. Siddique, "Performance evaluation among id3, c4. 5, and cart decision tree algorithm," in *Pervasive Computing and Social Networking*: Springer Singapore, 2022, pp. 127-142.
- [15] J. Jefriyanto, N. Ainun, and M. A. Al Ardha, "Application of Naïve Bayes Classification to Analyze Performance Using Stopwords," *Journal of Information System, Technology and Engineering*, vol. 1, no. 2, pp. 49-53, 2023.
- [16] A. Patel, A. K. Tiwari, and S. S. Ahmad, "Fake News Detection using Support Vector Machine," Department of Computer Science & Engineering, Kamla Nehru Institute of Technology, Sultanpur, India, KL Ghurair University, Dubai, UAE, 2022.
- [17] P. S. Thilagam, "Multi-layer perceptron based fake news classification using knowledge base triples," *Applied Intelligence*, vol. 53, no. 6, pp. 6276-6287, 2023.
- [18] T. Khunsuk. and J. Thongkam., "Feature Selection Method for Improving Customer Reviews Classification,," *RMUTI JOURNAL Science and Technology*, vol. Vol. 13 No. 1, no. 1, 10 January - 07 April. 2020.
- [19] R. Tipsena, "Automatic Question Classification on Webboard Using Text Mining Techniques," *Journal of Science and Technology Mahasarakham University*, vol. 33, no. 5, pp. 493-493, 2014.

- [20] N. Chirawichitchai, P. Sa-nguansat, and P. Meesad, "Developing an effective Thai document categorization framework base on term relevance frequency weighting," 2010, pp. 19-23: IEEE.
- [21] S. P. a. P. Posawang., "Classification for Bus Driver's Behaviors Using Text Extraction and Machine Learning Technique,," ITD KMUTNB, vol. Vol. 15 No. 1 no. 1, 2019.
- [22] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent systems*, vol. 28, no. 2, pp. 15-21, 2013.
- [23] M. Hu and B. Liu, "Mining opinion features in customer reviews," vol. 4, pp. 755-760.
- [24] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining.",2022.
- [25] มาสวีร์ มาศดิศร์โชติ, "การทำเหมืองความคิดเห็นภาษาไทย," วารสารศรีปทุมปริทัศน์ ฉบับวิทยาศาสตร์และเทคโนโลยี, vol. 6, no. 1, pp. 120-128, 2014.
- [26] M. Eirinaki, S. Pisal, and J. Singh, "Feature-based opinion mining and ranking," *Journal of Computer and System Sciences*, vol. 78, no. 4, pp. 1175-1184, 2012.
- [27] T. Shaik, X. Tao, C. Dann, H. Xie, Y. Li, and L. Galligan, "Sentiment analysis and opinion mining on educational data: A survey," *Natural Language Processing Journal*, vol. 2, 2023.
- [28] D. S. Khafaga, A. Ibrahim, S. K. Towfek, and N. Khodadadi, "Data Mining Techniques in Predictive Medicine: An Application in hemodynamic prediction for abdominal aortic aneurysm disease," *Journal of Artificial Intelligence and Metaheuristics*, vol. 5, no. 1, pp. 29-37, 2023.
- [29] C. W. H. Davis, A. J. Jetter, and P. J. Giabbanelli, "Automatically Generating Scenarios from a Text Corpus: A Case Study on Electric Vehicles," *Sustainability*, vol. 14, no. 13, p. 7938, 2022.
- [30] สมศักดิ์ ศรีสวการย์ และ สมัย ศรีสว, "การวิเคราะห์เหมืองความคิดเห็นโดยใช้เทคนิคการสกัดคำ," วารสาร "การประยุกต์ใช้เทคโนโลยีสารสนเทศ", vol. 6, no. 2, pp. 95-104, 2020.

- [31] ส. วีรศักดิ์ ฟองเงิน, เพ็ชรหาญ, "การวิเคราะห์เหมืองความคิดเห็นเครือข่ายเฟสบุ๊ค โดยใช้เทคนิคเหมืองข้อมูล : *Sentiment analysis opinion mining from Facebook Comments Using Data Mining Technique*," วารสารเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏอุบลราชธานี, vol. 11, no. 1, pp. 121-132, 2021.
- [32] จ. ท. วสวัตต์ อินทร์แปลง, "การวิเคราะห์ความคิดเห็นต่อเกมมือถือผับจีด้วยเหมืองข้อความ," วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม, vol. 39, no. 5, pp. 523-531, 2020.
- [33] P. Chormai, P. Prasertsom, and A. Rutherford, "Attacut: A fast and accurate neural thai word segmenter," *arXiv preprint arXiv:1911.07056*, 2019.
- [34] L. Galke and A. Scherp, "Bag-of-words vs. graph vs. sequence in text classification: Questioning the necessity of text-graphs and the surprising strength of a wide MLP," 1, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics., Association for Computational Linguistics, 2022.
- [35] H. Chauhan and A. Chauhan, "Implementation of decision tree algorithm c4.5," *International Journal of Scientific and Research Publications*, vol. 3, no. 10, pp. 1-3, 2013.
- [36] G. I. Webb, E. Keogh, and R. Miikkulainen, "Naïve Bayes Encyclopedia of machine learning," vol. 15, pp. 713-714, 2010.
- [37] C. Gold and P. Sollich, "Model selection for support vector machine classification," *Neurocomputing*, vol. 55, no. 1-2, pp. 221-249, 2003.
- [38] I. R. Vanani and M. Amirhosseini, "Deep learning for opinion mining in Extracting knowledge from opinion mining," IGI Global, 2019, pp. 40-65.
- [39] A. Košir, A. Odić, and M. Tkalčič, "How to improve the statistical power of the 10-fold cross validation scheme in recommender systems," pp. 3-6.
- [40] K. Khamchit, D. Jai Khum Kao, W. Phumrang, a. satnago, and A. Sukprasert, "The Efficiency of Data Mining Technique for the Prognosis of Cerebrovascular Disease," *Journal of Applied Informatics and Technology*, vol. 4, no. 2, pp. 87-98, 09/22 2022.
- [41] P. Keikhosrokiani and M. Pourya Asl, *Handbook of research on opinion mining and text analytics on literary works and social media*. IGI Global, 2022.





ภาคผนวก ก

คำที่แสดงคุณลักษณะที่ใช้ในการวิจัย

พูน ปณ จิต ชีวะ

คำแสดงคุณลักษณะที่ใช้ในงานวิจัย

คำแสดงคุณลักษณะที่ใช้ในการวิจัยครั้งนี้ ทั้งหมด 440 คำ แบ่งเป็นคำที่แสดงคุณลักษณะเชิงบวก 231 คำ และคำที่แสดงคุณลักษณะเชิงลบ 209 คำ โดยเรียงลำดับตามจำนวนการถูกกล่าวถึงจากมากที่สุดไปถึน้อยที่สุดตามลำดับ

ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
1	ดี	696	23	คนนิยม	66
2	สวย	510	24	สนุกโคตร	66
3	ชอบ	350	25	คุ้มราคา	64
4	ชอบคุณ	112	26	ชม	60
5	สนใจ	101	27	สะดวก	60
6	สนุก	93	28	สุข	60
7	คุ้ม	88	29	โคตรดี	60
8	คุ้มค่า	87	30	คาดหวัง	59
9	พัฒนา	85	31	สะอาด	59
10	ประหยัด	84	32	เยี่ยม	59
11	รัก	76	33	ก็ดี	57
12	สนุกดี	74	34	ถูกใจ	57
13	สวยงาม	74	35	เจ๋ง	54
14	คุณภาพดี	70	36	ชอบใจ	51
15	งาม	70	37	ดีใจ	51
16	รวย	70	38	ปลื้ม	50
17	สนุกสนาน	70	39	คุณสมบัติดี	49
18	เงิน	70	40	ดีใจ	49
19	โคตรสวย	70	41	ปรับแต่ง	49
20	แฮปปี้	69	42	สะดวกสบาย	49
21	โดดเด่น	68	43	ประทับใจ	48
22	มาตรฐาน	67	44	บำรุงรักษา	47

ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก (ต่อ)

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
44	บำรุงรักษา	47	72	เพลีนดี	35
45	สมราคา	47	73	คุณค่า	34
46	ยอดเยี่ยม	45	74	จริงใจ	34
47	สุดยอด	45	75	ทนทาน	34
48	ชำ	44	76	ท่องเที่ยว	34
50	ดูดี	44	77	สบายใจ	34
51	น่ารัก	44	78	อัปเดต	34
52	บำรุง	44	79	เยี่ยมชอบ	34
53	ปรับปรุง	44	80	แข็งแรง	34
54	ยั่งยืน	44	81	ไวใจ	34
55	สบายดี	44	82	ผลักดัน	33
56	แนะนำ	42	83	เชื่อใจ	33
57	สาระดี	41	84	เอาใจ	33
58	เท่ห้	41	85	แจก	33
59	เพลีน	41	86	เพลิตเพลีน	32
60	โดนใจ	41	87	ดีโคตร	31
61	ก้าวหน้า	40	88	ยินดี	31
62	ร่ำรวย	40	89	เสน่ห์	31
63	พอใจ	39	90	คึกคัก	30
64	อเนกประสงค์	39	91	พัฒนาการ	30
65	แถม	38	92	ยิ้ม	30
66	ดีงาม	37	93	สบายตา	30
67	ประโยชน์	37	94	สำเร็จ	30
68	มั่นใจ	36	95	เชียร์	30
69	สมบัติ	36	96	เด่น	30
70	โด่งดัง	36	97	แบ่งปัน	30
71	รอบรู้	35	98	เด่นชัด	27

ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก (ต่อ)

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
99	ยินดีต้อนรับ	26	126	เลิศ	19
100	ยิ้มแย้ม	26	127	เอาใจช่วย	19
101	รางวัล	26	128	แง่ดี	19
102	เหมาะสม	26	129	รีวิวดี	18
103	เวิร์ค	25	130	อัปเดตเฟิร์มแวร์	18
104	สำเร็จรูป	24	131	กิมมิกดี	17
105	ส่งเสริม	24	132	ดีสุด	17
106	แจ๋ว	24	133	สะเป็ก	17
107	เสถียร	24	134	อัปเดต	17
108	ดูหรูหรา	23	135	เฉลิมฉลอง	17
109	ประหยัดพลังงาน	23	136	นับถือ	16
110	เรียบง่าย	22	137	ระเปียบ	16
111	ใสใจ	22	138	รักษา	16
112	ชัดดี	21	139	เห็นใจ	16
113	งดงาม	21	140	กำไร	15
114	ชาร์จดี	21	141	ชอบชม	15
115	ชื่นชอบ	21	142	ดีกว่า	15
116	เต็มใจ	21	143	ตื่นตา	15
117	ชนะ	20	144	นับถือใจ	15
118	ขยัน	19	145	ผ่อนคลาย	15
119	ชื่นชม	19	146	ภาคภูมิใจ	15
120	ดูแลรักษา	19	147	สมบูรณ์	15
121	บริษัทดี	19	148	อิสระ	15
122	ภูมิใจ	19	149	เป็นห่วง	15
123	ส่งงาน	19	150	คุ้มงาน	14
124	สมาร์ท	19	151	ซื้อสัตย์	14
125	หล่อ	19	152	บริการดี	14

ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก (ต่อ)

153	ปรับเปลี่ยน	14	181	แรงใจ	9
154	รับประกัน	14	182	ปัญญา	8
155	สไมล์	14	183	ขวัญใจ	7
156	เรียบง่าย	14	184	ซัพเว็ค	7
157	ไว้วางใจ	14	185	คนดี	7
158	ฉลาด	12	186	มรดก	7
159	ดีเยอะ	12	187	หลงรัก	7
160	ดูแล	12	188	อยู่ดี	7
161	พริ้วเมี่ยม	12	189	เทคโนโลยีสารสนเทศ	7
162	รุ่ง	12	190	เป็นธรรม	7
163	สาระประโยชน์	12	191	ไฮเทค	7
164	หลงไหล	12	192	การันตี	5
165	เมตตา	12	193	กะทัดรัดดี	4
166	พิชิต	11	194	การันตีดี	4
167	รุ่มร้อน	11	195	ชมเพลิน	4
168	แซร์	11	196	ชำนาญ	4
169	กำลังใจ	10	197	ตื่นตาตื่นใจ	4
170	บันดาลใจ	10	198	พัฒนาสมรรถนะ	4
171	หายห่วง	10	199	มืออาชีพ	4
172	เอาใจใส่	10	200	ละเอียดดี	4
173	ได้ใจ	10	201	ลุล่วง	4
174	ตรงใจ	9	202	ส่งเสริมพัฒนา	4
175	ที่รัก	9	203	เวอร์วังอลังการ	4
176	ปกป้อง	9	204	แพรวพราว	4
177	ไร้ไซเคิล	9	205	ซอฟต์แวร์เสถียรดี	3
178	สง่า	9	206	ดีวะ	3
179	เสียสละ	9	207	ต้อนรับ	3
180	แปลกใจ	9	208	รักดี	3

ตารางที่ ภาคผนวก ก.1 คุณลักษณะเชิงบวก (ต่อ)

208	รักดี	3
209	สังคมดี	3
210	กลืนดี	2
211	ชมรถ	2
212	ซื้อตรง	2
213	ดีเหมาะ	2
214	คูโอเค	2
215	ทรหดอดทน	2
216	ทวีปดี	2
217	บ้านดี	2
218	พัฒนาฟื้นฟู	2
219	ลึกซึ้ง	2
220	วิวดี	2
221	กล้า	1
222	กะดี	1
223	ขับเทพ	1
224	คลังรัก	1
225	คอนเทนเยียม	1
226	คูโณปการ	1
227	ขอพรสนิยม	1
228	ดียอด	1
229	ดีเวอร์	1
230	สิ่งใจ	1
231	อุดมสุข	1

ตารางที่ ภาคผนวก ก.2 คุณลักษณะเชิงลบ

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
1	ปัญหา	202	27	ช่วย	30
2	แพง	195	28	ต่ำ	30
3	จ่าย	78	29	บกพร่อง	30
4	ซ่อม	51	30	รัฐบาล	30
5	ร้อน	50	31	แจ้ง	30
6	ต่ำ	48	32	เสียใจ	30
7	ภาษีแพง	46	33	โกหก	30
8	พัง	45	34	จุดอ่อน	29
9	กวน	36	35	ถกเถียง	29
10	ตาย	36	36	ท้อ	29
11	กลัว	35	37	เสียหาย	29
12	ดราม่า	34	38	โกรธ	29
13	ผิดพลาด	34	39	ด้อย	28
14	เครียด	34	40	คาใจ	27
15	เผา	34	41	บังคับ	27
16	ขาดแคลน	33	42	ชี้เี่ยว	26
17	ขายแพง	33	43	ซึ้ง	24
18	ชน	33	44	ดูถูก	24
19	ราคาแพง	33	45	เหนียว	24
20	อคติ	33	46	ถ่ม	23
21	เสียง	33	47	บ้า	23
22	โง่	33	48	ประมาท	23
23	ทะเลาะ	32	49	มั่ว	22
24	เหี้ย	32	50	เศร้า	22
25	กาก	30	51	กวนใจ	21
26	ซึ่กลัว	30	52	กังวล	21
27	ช่วย	30	53	คนอคติ	21

ตารางที่ ภาคผนวก ก.2 คุณลักษณะเชิงลบ (ต่อ)

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
54	งก	21	81	หงุดหงิด	18
55	หลอก	21	82	พักผ่อน	16
56	เกียจ	21	83	เล็ก	15
57	โง่ง	21	84	ซื่อค	14
58	ก๊ากสเปค	20	85	ตำหนิ	14
59	รก	20	86	ห่วย	14
60	รำคาญ	20	87	อุบัติเหตุ	14
61	หนี้	20	88	เสียเปรียบ	14
62	แปลก	20	89	เหนื่อยใจ	14
63	โหดร้าย	20	90	อึดอัด	13
64	ก๊าก	19	91	เสียดสี	13
65	ขัดใจ	19	92	แซะ	13
66	ขโมย	19	93	ซู้ย	12
67	ตกเทรนด์	19	94	ดีเลย์	12
68	ทำใจ	19	95	ตกยุค	12
69	บิบบังคับ	19	96	ทะแม่ง	12
70	บ่น	19	97	ระเบิด	12
71	ประจาน	19	98	รุนแรง	12
72	พลาด	19	99	ล่าช้า	12
73	ภาระ	19	100	อิจฉา	12
74	ยุ่ง	19	101	เกลียด	12
75	ร้องเรียน	19	102	เคร่งครัด	12
76	อ้วน	19	103	เรียกร้อง	12
77	เปื้อน	19	104	ฉวย	11
78	เผาไหม้	19	105	ปากดี	11
79	เมา	19	106	ลำบาก	11
80	แบน	19	107	เหยียด	11

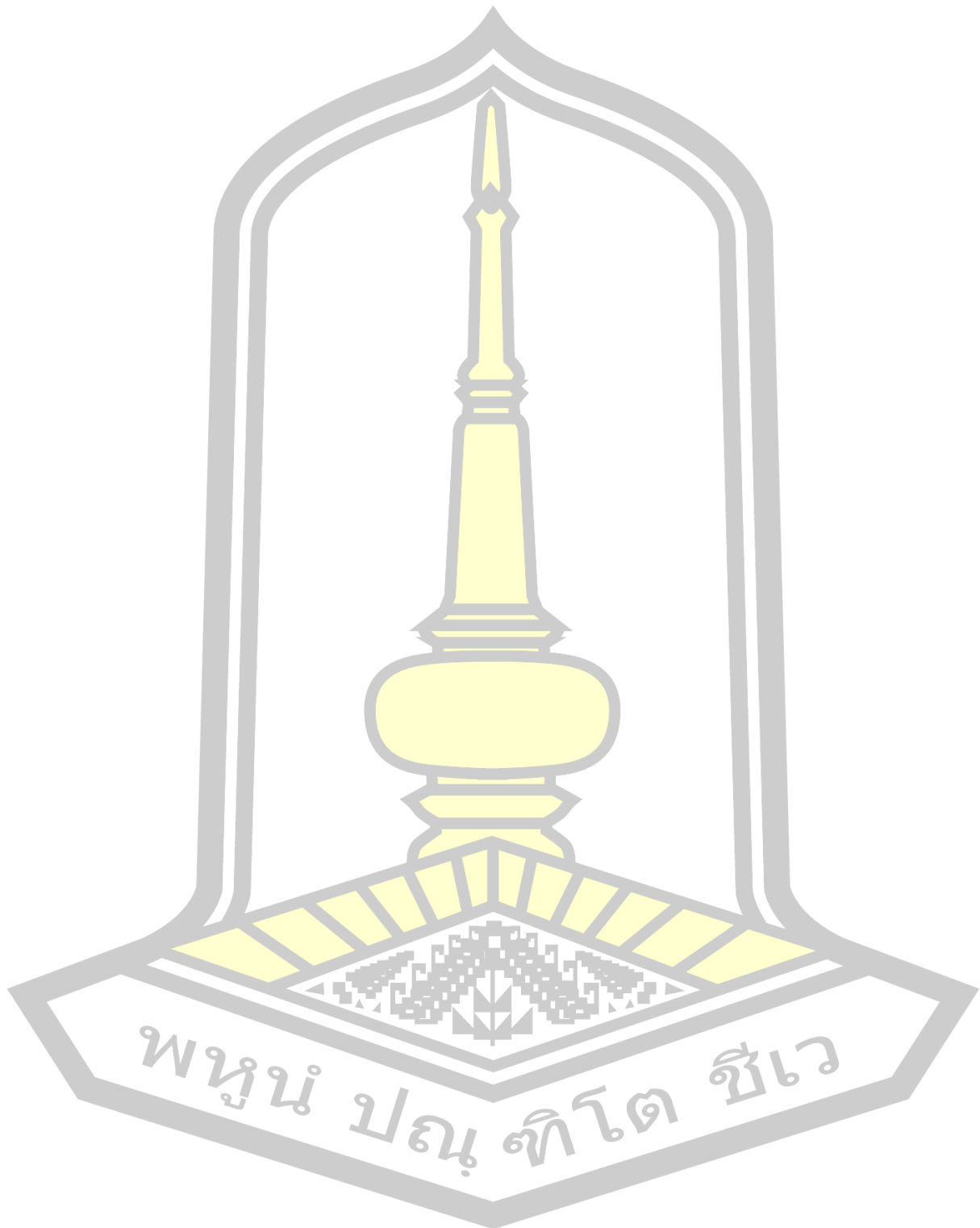
ตารางที่ ภาคผนวก ก.2 คุณลักษณะเชิงลบ (ต่อ)

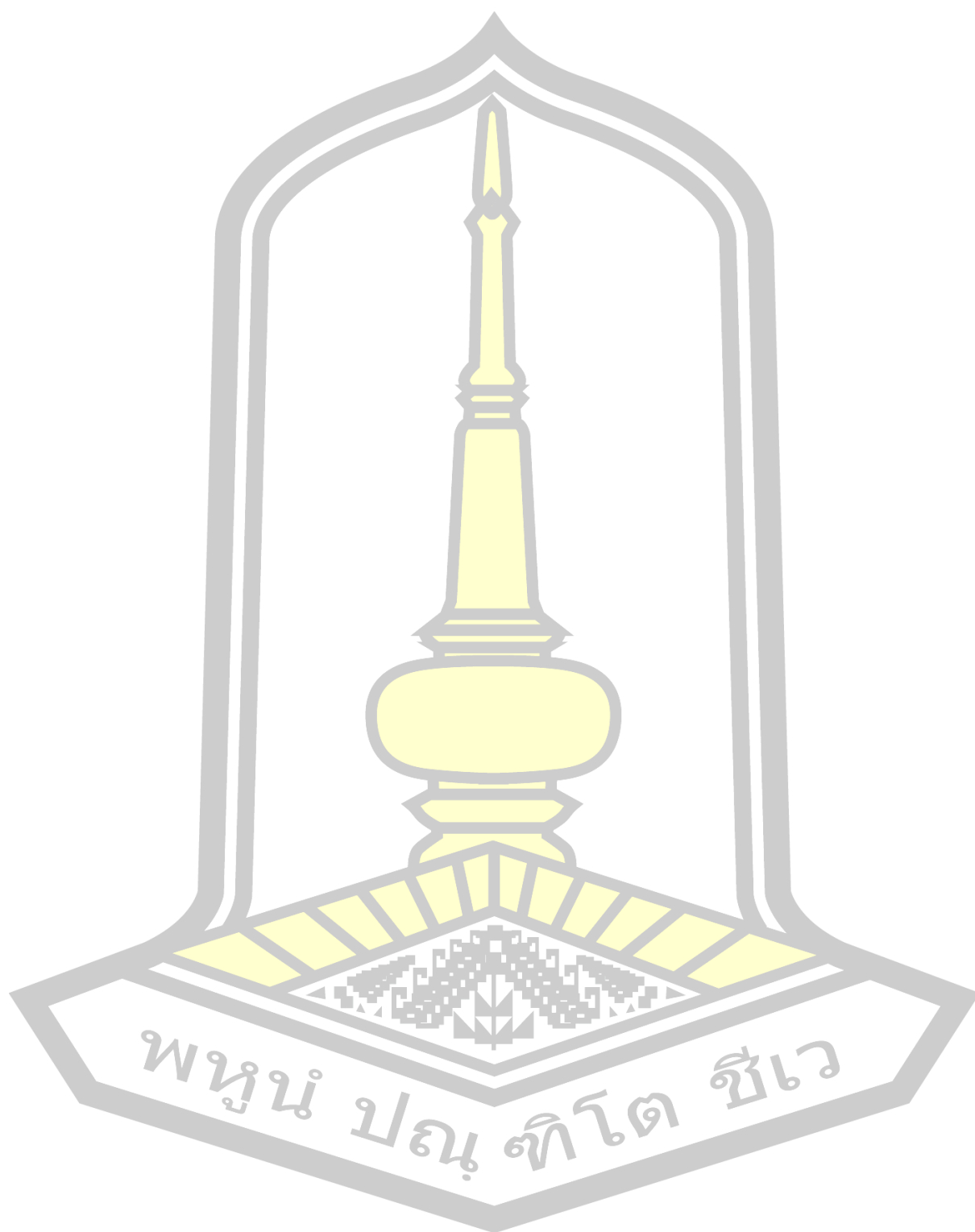
ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
108	ยุ่งยาก	10	135	ชู่	4
109	แปลกตา	10	136	คุกคาม	4
110	แพ้	10	137	จองจำ	4
111	คลัง	9	138	ชก	4
112	จุกจิก	9	139	บู๊	4
113	ชดเชย	9	140	ปะทะ	4
114	ชั่ว	9	141	ยอมใจ	4
115	พิการ	9	142	รังเกียจ	4
116	ลั้ม	9	143	ร้าว	4
117	สาป	9	144	วุ่นวาย	4
118	อัพราดา	9	145	สยอ	4
119	เลวร้าย	9	146	สูญ	4
120	พนัน	8	147	เกะกะ	4
121	ไม่ดี	8	148	เสื่อม	4
122	กักอหวัณ	7	149	เหม็น	4
123	คว่ำบาตร	7	150	เหยียดหยาม	4
124	ถลุง	7	151	กักรถ	3
125	ถ่วง	7	152	ขัดแย้ง	3
126	ผูกขาด	7	153	ชี้เหย	3
127	ยำแย	7	154	จัญไร	3
128	สูญเสี	7	155	ชักช้า	3
129	เหงา	7	156	บ้านแตก	3
130	แย่ง	7	157	เสี่ยงตาย	3
131	บ้ำจริง	6	158	พรวดพราด	3
132	ซังใจ	5	159	พาล	3
133	กู่	4	160	ยอดตก	3
134	ขัดตา	4	161	ลักทรัพย์	3

ตารางที่ ภาคผนวก ก.2 คุณลักษณะเชิงลบ (ต่อ)

ลำดับ	คำ	จำนวนครั้ง	ลำดับ	คำ	จำนวนครั้ง
162	สิ้นสุด	3	189	รา	1
163	อ้วก	3	190	ร้อนหน้า	1
164	เถียง	3	191	ลุ่มจม	1
165	แผล	3	192	สงคราม	1
166	โจร	3	193	สวะ	1
167	ไขว้เขว	3	194	สังหาร	1
168	กตดัน	2	195	สาปส่ง	1
169	กั๊น	2	196	สาหัส	1
170	งามไส้	2	197	ส่วย	1
171	มหาโหด	2	198	หัวแข็ง	1
172	ลัก	2	199	อัคคีภัย	1
173	สิ้นเปลือง	2	200	เกร	1
174	สูญสิ้น	2	201	เคลือบแคลงใจ	1
175	เจ็บป่วย	2	202	ฉะฉาน	1
176	เพี้ยน	2	203	เน่า	1
177	ใจหาย	2	204	เลื่อมล้ำ	1
178	ขาดหวัง	1	205	เอะอะ	1
179	ฆ่าตกรรม	1	206	เอะอะกวนใจ	1
180	ปักใจ	1	207	แย่ง	1
181	ป้างเจก	1	208	แฮก	1
182	ปาราสิต	1	209	โจรกรรม	1
183	พะวง	1			
184	ภัยฉุกเฉิน	1			
185	มหาแพง	1			
186	ยากจน	1			
187	ยากลำบาก	1			
188	ยาบ้า	1			

บรรณานุกรม





ประวัติผู้เขียน

ชื่อ	นายฤทธิ์ศักดิ์ บุปผาวิศักดิ์
วันเกิด	29 พฤษภาคม 2536
สถานที่เกิด	จังหวัดขอนแก่น
สถานที่อยู่ปัจจุบัน	282/1 ม.11 บ.หนองโตน ต.หนองไผ่ อ.ชุมแพ จ.ขอนแก่น 40130 โทร. 092-8242294
ตำแหน่งหน้าที่การงาน	เจ้าหน้าที่ระบบงานคอมพิวเตอร์
สถานที่ทำงานปัจจุบัน	สำนักงานขนส่งจังหวัดขอนแก่น
ประวัติการศึกษา	2556 - ปวส. สาขาเทคโนโลยีสารสนเทศ วิทยาลัยอาชีวศึกษาขอนแก่น 2559 - วท.บ. สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม ปัจจุบัน - กำลังศึกษา วท.ม. สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการ สารสนเทศ มหาวิทยาลัยมหาสารคาม
ทุนวิจัย	-
ผลงานวิจัย	ระบบจัดการตารางสอนอัตโนมัติของวิทยาลัยอาชีวศึกษาขอนแก่น Automatic Classroom Timetable Management System of Khon Kaen Vocational College

พูน ปณ ภัโต ชีเว