การทำเหมืองความคิดเห็นของผู้ใช้รถพลังงานทางเลือก
Opinion Mining of Alternative-Energy Vehicle Users

ฤทธิ์ศักดิ์ บุปผาทวีศักดิ์* และ จารี ทองคำ
Ritsak Bupphataweesak* and Jaree Thongkam

สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
Information Technology Department, Faculty of Informatics, Mahasarakham University
*Email: myp4nda@gmail.com
Received: September 17, 2024; Revised: November 07, 2024; Accepted: November 15, 2024

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อคัดเลือกเทคนิคการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึกที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือก โดยแบ่งออกเป็น 3 เป้าหมายหลัก ได้แก่ 1) การคัดเลือกเทคนิคที่ดีที่สุดสำหรับการจำแนกความคิดเห็นทั่วไปของผู้ใช้ 2) การคัดเลือกเทคนิคที่ดีที่สุดสำหรับการจำแนกความคิดเห็นที่พิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป และ 3) การศึกษาความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกจากความคิดเห็น 3,000 ข้อความที่รวบรวมจาก Facebook, YouTube, Pantip และ TikTok โดยใช้กระบวนการเหมืองความคิดเห็นและการประเมินประสิทธิภาพของแบบจำลองด้วยวิธีไขว้ทับแบบ 10 ส่วน (10-fold cross-validation) เพื่อวัดค่าความถ่วง (F-Measure), ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ผลการวิจัยพบว่าเทคนิคเพอร์เซปตรอนหลายชั้น มีประสิทธิภาพสูงสุดทั้งในการจำแนกความคิดเห็นทั่วไปและความคิดเห็นตามคุณลักษณะโดยพิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป โดยมีค่าความถ่วง ค่าความแม่นยำ และค่าความระลึกสูงถึง 95.30% และ 97.20% ตามลำดับ ในงานวิจัยนี้พบว่าคนไทยมีความคิดเห็นเชิงบวกต่อยานยนต์พลังงานทางเลือกร้อยละ 69.97 เชิงลบร้อยละ 28.67 และเป็นกลางร้อยละ 1.36

คำสำคัญ : เหมืองความคิดเห็น, การวิเคราะห์ความคิดเห็น, ยานยนต์พลังงานทางเลือก, เหมืองข้อมูล

ABSTRACT

This research aims to identify the most effective machine learning and deep learning techniques for classifying user opinions on alternative energy vehicles. The study is divided into three main objectives: 1) to select the best technique for classifying general user opinions, 2) to select the best technique for classifying opinions based on features mentioned more than 10 times, and 3) to analyze Thai users' opinions on alternative energy vehicles. A total of 3,000 comments were collected from Facebook, YouTube, Pantip, and TikTok, with sentiment analysis techniques used to classify the comments into positive and negative opinions. The performance of the models was evaluated using 10-fold cross-validation, measuring F-Measure, Precision, and Recall. The findings indicate that the Multilayer Perceptron technique achieved the highest performance in both general opinion

classification and feature-based classification, with F-Measure, Precision, and Recall values of 95.30% and 97.20%, respectively. Additionally, it was found that 69.97% of Thai users expressed positive opinions toward alternative energy vehicles, 28.67% expressed negative opinions, and 1.36% were neutral.

Keywords: Opinion Mining, Opinion Analysis, Alternative Energy Vehicles, Data Mining.

1. บทนำ

จากภาวะวิกฤตพลังงานและความผันผวนของราคาน้ำมันทั่วโลก ส่งผลกระทบต่อราคาน้ำมันในประเทศไทย ทำให้ราคาน้ำมันเพิ่มสูงขึ้น [1] ซึ่งกระตุ้นให้ผู้ใช้ยานยนต์หันมาสนใจในยานยนต์พลังงานทางเลือกมากขึ้น [2] เพื่อเป็นทางเลือกในการลดค่าใช้จ่ายประจำวัน ยานยนต์พลังงานทางเลือกที่ได้รับความสนใจ ได้แก่ รถยนต์ไฮบริด รถยนต์ปลั๊กอินไฮบริด รถยนต์ไฟฟ้า และรถยนต์ไฟฟ้าแบบเซลล์เชื้อเพลิง เป็นต้น อย่างไรก็ตาม ยานยนต์พลังงานทางเลือกยังคงมีข้อจำกัดหลายประการ เช่น ความไม่มั่นใจในการใช้งานในช่วงฤดูฝนที่มีน้ำท่วมขัง [3] ส่งผลให้ความคิดเห็นของผู้ใช้ที่เผยแพร่ผ่านสื่อสังคมออนไลน์ เช่น Facebook, YouTube, Pantip และ TikTok มีลักษณะทั้งเชิงบวกและเชิงลบ ซึ่งทำให้การแยกแยะความคิดเห็นเหล่านี้เป็นเรื่องที่ท้าทายและน่าสนใจ ดังนั้น งานวิจัยนี้จึงมุ่งศึกษาความคิดเห็นของผู้ใช้ยานยนต์พลังงานทางเลือกโดยใช้กระบวนการเหมืองความคิดเห็นเพื่อการวิเคราะห์และเข้าใจความคิดเห็นของผู้ใช้และผู้ที่สนใจในยานยนต์พลังงานทางเลือกได้อย่างแม่นยำ

เหมืองความคิดเห็น (Opinion Mining) คือ กระบวนการวิเคราะห์ความคิดเห็นที่เป็นภาษาธรรมชาติ ในการจำแนกความคิดเห็นออกเป็นเชิงบวก เชิงลบ และเป็นกลาง [4] นักวิจัยจำนวนมากได้นำเหมืองความคิดเห็นมาใช้ในการจำแนกความคิดเห็น ซึ่งเทคนิคที่นำมาใช้ในการสร้างแบบจำลองโดยส่วนใหญ่เป็นเทคนิคจากเหมืองข้อมูล (Data Mining) และการเรียนรู้ของเครื่อง (Machine Learning) เช่น ปราโมชย์ นามวงศ์ และคณะ [5] ใช้เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคนาอีฟเบย์ เทคนิคเคเนียเรสเนเบอร์ เทคนิคต้นไม้ตัดสินใจ ซี4.5 และเทคนิคความจำระยะสั่น-ระยะยาว ในการทดลองจำแนกความคิดเห็นบนเครือข่ายสังคมออนไลน์เฟสบุ๊คของผู้ใช้บริการร้านอาหารในจังหวัดอุบลราชธานี พบว่า เทคนิคต้นไม้ตัดสินใจ มีประสิทธิภาพสูงสุดในการทดลองมีค่าความถูกต้องเท่ากับร้อยละ 89.00 ส่วนงานวิจัยของ พิศิษฐ์ บวรเลิศสุธี และคณะ [6] ใช้เทคนิคเคเนียเรสเนเบอร์ เทคนิคป่าไม้แบบสุ่ม เทคนิคการลดการเคลื่อนย้าย และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ในการจำแนกความคิดเห็นของผู้ใช้สินค้า พบว่า เทคนิคซัพพอร์ตเวกเตอร์แมชชีนมีประสิทธิภาพในการจำแนกสูงที่สุดที่ค่าความแม่นยำที่ร้อยละ 79.00 ยิ่งไปกว่านั้น Adebayo Abayomi และคณะ [7] ได้ใช้เทคนิคพจนานุกรมเล็กซิตรอน เทคนิคซัพพอร์ตเวกเตอร์แมชชีน และเทคนิคการเรียนรู้เชิงลึก ในการจำแนกความคิดเห็นของผู้ใช้งานโปรแกรม Yahoo จาก Twitter พบว่า เทคนิคการเรียนรู้เชิงลึกมีประสิทธิภาพดีที่สุดที่ความแม่นยำถึงร้อยละ 80.38

ดังนั้นเพื่อพัฒนาแบบจำลองที่สามารถจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกได้อย่างมีประสิทธิภาพ ผู้วิจัยจึงได้ประยุกต์เทคนิคต่าง ๆ ในการทำเหมืองความรู้สึก ได้แก่ เทคนิคต้นไม้ตัดสินใจ, เทคนิคเนอีฟเบย์, เทคนิคซัพพอร์ตเวกเตอร์แมชชีน และเทคนิคเพอร์เซปตรอนหลายชั้น ในการประเมินประสิทธิภาพของแบบจำลอง ได้ใช้วิธีการ 10-Fold Cross Validation เพื่อแบ่งข้อมูลออกเป็นสองชุด คือ ชุดข้อมูลสำหรับการเรียนรู้และชุดข้อมูลสำหรับการทดสอบ โดยใช้ค่าความแม่นยำ (Precision), ค่าความระลึก (Recall), และค่าความถ่วง (F-Measure) เป็นตัวชี้วัดเพื่อเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง

2. วัตถุประสงค์ของงานวิจัย

2.1. เพื่อศึกษาความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกในกลุ่มยานยนต์พลังไฟฟ้า (Electric Vehicle: EV) และยานยนต์กึ่งไฟฟ้า (Hybrid Electric Vehicle: HEV)

2.2. เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกในกลุ่มยานยนต์พลังไฟฟ้า (EV) และยานยนต์กึ่งไฟฟ้า (HEV)

2.3 เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกในกลุ่มยานยนต์พลังไฟฟ้า (EV) และยานยนต์กึ่งไฟฟ้า (HEV) โดยพิจารณาคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป

3. วิธีดำเนินการวิจัย

3.1. เครื่องมือการวิจัย

1. คอมพิวเตอร์ Core i5-11400 RAM 32GB (16* 2Gb) bus 3200
2. โปรแกรม Weka 3.8.6
3. เทคนิคในการสร้างแบบจำลอง
   - 3.1 เทคนิคต้นไม้ตัดสินใจ แบบ ซี 4.5 : C4.5
   - 3.2 เทคนิคเนอีฟเบย์ : NB
   - 3.3 เทคนิคซัพพอตเว็กเตอร์แมทชีน : SVM
   - 3.4 เทคนิคเพอร์เซปตรอนหลายชั้น : MLP

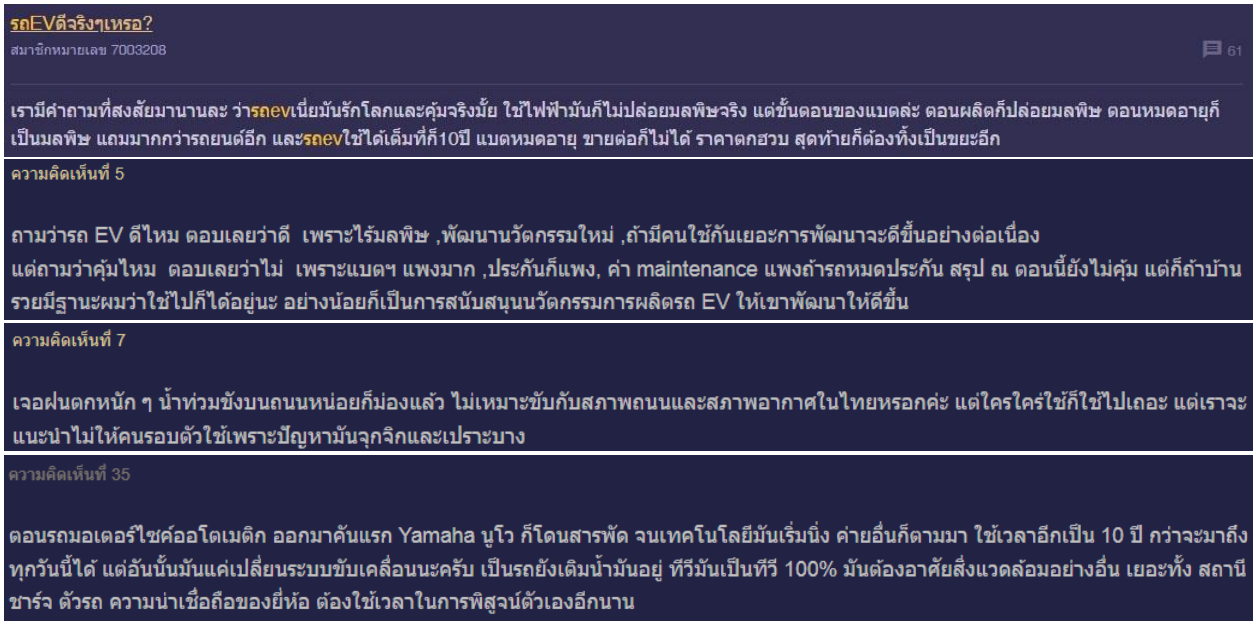
3.2. กลุ่มเป้าหมาย

3.2.1 ความคิดเห็นของคนไทยจากเว็บไซต์ Facebook, YouTube, Pantip และ TikTok จำนวน 3,000 ความคิดเห็น ที่เกี่ยวข้องกับรถยนต์พลังงานทางเลือกในกลุ่ม EV และ HEV

3.3. ขั้นตอนในการดำเนินการวิจัย ผู้วิจัยได้แบ่งการดำเนินการเป็น 5 ขั้นตอน ดังนี้

3.3.1 การเลือกข้อมูล

การเลือกข้อมูลในขั้นตอนนี้ผู้วิจัยได้ทำการค้นหาและเก็บรวมรวมความคิดเห็นจาก Facebook, YouTube, Pantip และ TikTok โดยการเข้าไปในช่องหรือกระทู้ต่าง ๆ ที่มีส่วนเกี่ยวของกับยานยนต์พลังงานทางเลือก ดังปรากฏในรูปที่ 1



รูปที่ 1 ตัวอย่างกระทู้ในชุมชนออนไลน์เกี่ยวกับยานยนต์พลังงานทางเลือกจากเว็บ Pantip

จากรูปที่ 1 แสดงตัวอย่างความคิดเห็นที่ถูกเลือกมาใช้เป็นข้อมูลในการวิจัยจากเว็บไซต์ Pantip ซึ่งไม่ว่าจะเป็นสื่อสังคมออนไลน์ใดก็ตามมักจักมีชุมชน (Community) ไว้ให้ผู้ใช้งานที่มีความคิดเห็นตรงกันหรือสนใจในสิ่งเดียวกันได้แสดงความคิดเห็นต่าง ๆ ร่วมกันโดยส่วนมากมักจะใช้ # (Hashtag) เพื่อแยกหรือติดตามสิ่งที่ตนเองสนใจ เช่น #รถไฟฟ้า, #รถEV เป็นต้น จุดนี้เองทำให้ง่ายต่อการรวบรวมข้อมูล ผู้วิจัยได้ใช้ภาษา Python ในการเลือกความคิดเห็นที่มีเนื้อความมากกว่า 15 คำขึ้นไปมาใช้ เพื่อให้มีเนื้อหาความคิดเห็นที่ละเอียดมากขึ้น จากนั้นนำความคิดเห็นทั้งหมดเก็บในโปรแกรม Microsoft Excel จำนวนทั้งสิ้น 3,000 ความคิดเห็น

3.3.2 การเตรียมข้อมูล

การเตรียมข้อมูลสำหรับการจำแนกความคิดเห็นเริ่มจากการทำความสะอาดข้อความ โดยแก้คำสะกดผิด ลบคำซ้ำ ตัวเลข สัญลักษณ์ (! @ $ % *) และภาพ ตลอดจนการเว้นวรรค ย่อหน้า และคำที่ไม่สื่อความหมาย (เช่น ค่ะ, ครับ) โดยตัวอย่างความคิดเห็นก่อนและหลังการทำความสะอาดแสดงในรูปที่ 2 และรูปที่ 3 ตามลำดับ

1	คันนี้อ้อฟชั่นน้อย พอใช้ แต่ ผมเน้นขับ แนวนี้ คือโดนใจอ่ะ กับการทรงตัวของคันนี้ คือตอบโจทย์มาก กับผม ตอนแรกแอบมา BYD ตอนนี้ คือ เท มาทางนี้เลย
2	ที่เท้าแขน แอร์หลัง มือจับ เอาไปทำไมไม่เห็นจำเป็น ถ้าขับเองนะ คนนั่งหลังก็เรื่องของมัน แค่นั่งสบายก็พอที่เหลือของจุกจิกไม่จำเป็น
3	1. ถ้าไม่ใช่ blade battery จะเป็นไรไหมครับ? ด้อยกว่าแค่ไหน 2. เห็นว่ารถ EV ตปท.ขายดีเลยสงสัย ตปท.แต่ละรัฐห่างกันมาก ขับข้ามกันเป็นวัน รถ EV ที่ต้องชาร์จทุก 350km. เขาบริหารความเซ็งกันยังไงน้อ
4	ปกติ เป็นคนขับรถน้ำมัน ไม่ชอบค่ายนี้ เห็นเสียบ่อย อะไหล่ไม่ได้คุณภาพ วันนี้ไปลองรถไฟฟ้าครั้งแรก mg4x ขับดีจริง เปิดใจเลย
5	ออฟชั่นน้อยไป,เช่นมือจับหลังคา,แอร์หลัง,ที่เท้าแขน,ซันรูปก็ไม่มี,ที่ปัดน้ำฝนหลัง,ยางอะไหล่ไม่มีให้,คนรีวิวพูดเวอร์ไป,ช่วงล่างมันดิสิงดีกว่ารถราคา2ล้านขึ้นไป,เขาระบบมันดิสิงทั้งนั้น
6	สวัสดี ครับ น้อง นิธิ มีFC ฝากให้ผมบอก ขอให้น้อง นิธิ เวลา แนะนำรถ ต่าง จะมีการแจงรายละเอียด คำว่า (สมมุติ)แรงบิด ที่ 1950 นิวตั้นฯที่ 3,500รอบ FC บางคนยังไม่กระจ่างครับ นิวตั้น มันคือแรงม้าหรือไม่ นการปีนป่าย เนิน รีที่สูงชัน ใช่หรือไม่ครับ

รูปที่ 2 ความคิดเห็นก่อนการทำความสะอาดคำ

```
['คันนี้อ้อฟชั่นน้อยพอใช้แต่ผมเน้นขับแนวนี้คือโดนใจอ่ะกับการทรงตัวของคันนี้คือตอบโจทย์มากกับผมตอนแรกแอบมาตอนนี้คือเทมาทางนี้เลย',
 'ที่เท้าแขนแอร์หลังมือจับเอาไปทำไม่ไม่เห็นจำเป็นถ้าขับเองนะคนนั่งหลังก็เรื่องของมันแค่นั่งสบายก็พอที่เหลือของจุกจิกไม่จำเป็น',
 'ถ้าไม่ใช่จะเป็นไรไหมครับด้อยกว่าแค่ไหนเห็นว่ารถตปทขายดีเลยสงสัยตปทแต่ละรัฐห่างกันมากขับข้ามกันเป็นวันรถที่ต้องชาร์จทุกเขาบริหารคว
 'ปกติเป็นคนขับรถน้ำมันไม่ชอบค่ายนี้เห็นเสียบ่อยอะไหล่ไม่ได้คุณภาพวันนี้ไปลองรถไฟฟ้าครั้งแรกขับดีจริงเปิดใจเลย',
 'ออฟชั่นน้อยไปเช่นมือจับหลังคาแอร์หลังที่เท้าแขนซันรูปก็ไม่มีที่ปัดน้ำฝนหลังยางอะไหล่ไม่มีให้คนรีวิวพูดเวอร์ไปช่วงล่างมันดิสิงดีกว่ารถราคา
 'สวัสดีครับน้องนิธิมีฝากให้ผมบอกขอให้น้องนิธิเวลาแนะนำรถต่างจะมีการแจงรายละเอียดคำว่าสมมุติแรงบิดที่นิวตั้นฯที่รอบบางคนยังไม่กระจ่างค
ตามกำหนดจึงจะได้แรงบิดกำลังม้าที่รถจะใช้กำลังในการปีนป่ายเนินรีที่สูงชันใช่หรือไม่ครับ',
 'เวลาขับรีวิวนิธิถ้าไม่ดีจะบอกว่าไม่ดีแต่ถ้าดีนิธิจะบอกว่าดีนะเห็นรีวิวหลายยี่ห้อเชื่อครับเคยลองรุ่นก่อนคือก็ยังไม่ดีคือไฟจอแสดงผลไม่ค่อยสว่าง
 'ช่วงล่างสู่บีวายดีไม่ได้เอ็มจีออฟชั่นก็น้อยเช่นไม่มีแอร์มือจับที่เท้าแขนปัดน้ำฝนยางอะไหล่ไม่มีสักอย่าง',
```

รูปที่ 3 ข้อมูลหลังจากการทำความสะอาดคำ

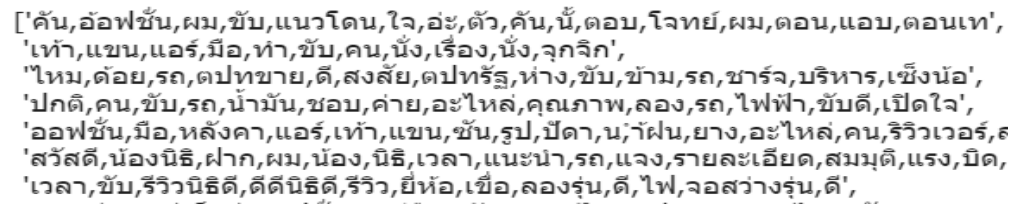
3.3.3 การทำถุงคำและการกำหนดคลาส

1. การสร้างถุงคำ (Bag of Words) เป็นขั้นตอนสำคัญในการแยกและนับคำที่ปรากฏในข้อความเพื่อใช้ในการวิเคราะห์ข้อมูลความคิดเห็น ในการวิจัยครั้งนี้ ผู้วิจัยได้เลือกใช้ภาษา Python ร่วมกับ Library PyThaiNLP [8] เพื่อให้การแยกคำเป็นไปอย่างมีประสิทธิภาพและสะดวกในการนับจำนวนคำที่เกิดขึ้นในแต่ละข้อความ นอกจากนี้ ยังมีการตรวจสอบความเหมาะสมของข้อมูลหลังจากการแยกคำ เพื่อให้มั่นใจว่าคำที่ได้สามารถนำไปใช้ในการวิเคราะห์คุณลักษณะในขั้นตอนถัดไปได้อย่างแม่นยำ ตัวอย่างผลลัพธ์ของการแยกคำแสดงให้เห็นใน รูปที่ 4

```
1  from pythainlp.tokenize import word_tokenize
2
3  text = "คันนี้อ้อฟชั่นน้อยพอใช้แต่ผมเน้นขับแนวนี้คือโดนใจอ่ะกับการทรงตัวของคันนี้คือตอบโจทย์มากกับผม ตอนแรกแอบมาตอนนี้เทมาทางนี้เลย"
4  tokens = word_tokenize(text, engine="newmm")
5  print(tokens)
6
7
8  ['คันนี้', 'อ้อฟชั่น', 'น้อย', 'พอใช้', 'แต่', 'ผม', 'เน้น', 'ขับ', 'แนว', 'นี้', 'คือ', 'โดนใจ', 'อ่ะ', 'กับ', 'การ',
9  'ทรงตัว', 'ของ', 'คันนี้', 'คือ', 'ตอบโจทย์', 'มาก', 'กับ', 'ผม', 'ตอนแรก', 'แอบ', 'มา', 'ตอนนี้', 'เท'
10 , 'มา', 'ทาง', 'นี้', 'เลย']
```

รูปที่ 4 การแยกคำเพื่อแยกคุณลักษณะ

จากรูปที่ 4 แสดงการแยกคำเพื่อให้ง่ายต่อการนับคำและระบุชนิดของคำ ซึ่งขั้นตอนนี้ผู้วิจัยต้องตรวจสอบความถูกต้องหลังจากการแยกคำ หากพบข้อผิดพลาดต้องทำการแก้ไขทันที โดยในขั้นตอนการแยกคำอาจจะเกิดข้อผิดพลาดได้ เช่น การแยกตัวสะกดผิดหรือคำบางคำไม่ถูกแยก เป็นต้น เมื่อตรวจสอบความถูกต้องแล้วนำคำทั้งหมดมาเรียงใหม่ดังรูปที่ 5



รูปที่ 5 ข้อมูลหลังจากการแยกคำ

จากรูปที่ 5 แสดงคำที่ถูกเรียงใหม่หลังจากการแยกคำซึ่งมีจำนวนคำทั้งหมด 3,044 คำ จากนั้นนำคำไปตรวจสอบและระบุชนิดของคำ (Part of Speech) ด้วยภาษา Python และเลือกเอาคำกริยา คำวิเศษณ์ คำคุณศัพท์และคำกริยาวิเศษณ์ [4] ในการกำหนดคลาสของคำในขั้นตอนต่อไปดังตัวอย่างใน ตารางที่ 1

ตารางที่ 1 การระบุชนิดคำและกำหนดคุณลักษณะจาก

ลำดับ	คำ	ชนิดคำ	จำนวนที่พบ	คุณลักษณะ
1	ค้น	คำกริยา	17	0
2	ผม	คำนาม	77	0
3	แย่	คำวิเศษณ์	51	-1
4	ดี	คำวิเศษณ์	70	1
5	ชอบ	คำกริยา	50	1
6	กลัว	คำกริยา	21	-1

จากตารางที่ 1 แสดงการกำหนดคุณลักษณะให้กับคำหลังจากการแยกคำและระบุชนิดของคำแล้ว โดยคุณลักษณะเชิงบวกแทนค่าด้วย 1 และคุณลักษณะเชิงลบจแทนค่าด้วย -1 ส่วนคำที่มีคุณลักษณะเป็นกลางแทนค่าด้วย 0 ขั้นตอนนี้ผู้วิจัยจะประเมินคุณลักษณะร่วมกับผู้เชี่ยวชาญด้านภาษาไทยจากโรงเรียนเทศบาล ๑ อำเภอซุมแพ จังหวัดขอนแก่น เพื่อให้ได้คุณลักษณะทางความรู้สึกที่ถูกต้องที่สุด

2. การกำหนดคลาส เป็นการนับจำนวนครั้งของคุณลักษณะที่ปรากฏในความคิดเห็นของแต่ละความคิดเห็นโดยไม่ต้องคำนึงถึงหลักไวยากรณ์ ความถี่ของคำ และลำดับคำในแต่ละประโยค จากนั้นทำการกำหนดคลาสของคำออกมา โดยกำหนดให้เป็นคุณลักษณะที่เป็นเชิงบวก (Positive) คือ คลาส P ส่วนคุณลักษณะที่เป็นเชิงลบ (Negative) คือ คลาส N และคุณลักษณะที่เป็นกลางคือ 0 ความถี่ของคุณลักษณะนับจากการนำเอา “คำ” ที่เป็นคุณลักษณะเชิงบวกหรือเชิงลบ ไปเทียมกับความคิดเห็นในแต่ละความคิดเห็น เช่น ความคิดเห็นที่ 2 มีการกล่าวถึงคำว่า “ไม่” จำนวน 2 ครั้ง ความถี่ของคุณลักษณะจึงเป็น 2 และขั้นตอนนี้เป็นการนับเอาคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้ง ขึ้นไป ดังแสดงใน ตารางที่ 2

ตารางที่ 2 การกำหนดคลาส

ลำดับ	คุณลักษณะเชิงบวก		คุณลักษณะเชิงลบ		ความถี่ของคุณลักษณะเชิงบวก	ความถี่ของคุณลักษณะเชิงลบ	คลาส
	ชอบ	ได้	ไม่	แย่			
1	1	0	0	0	1	0	P
2	0	1	2	0	1	2	N
3	1	1	0	0	2	0	P
4	0	1	0	0	1	0	P
5	0	0	0	0	0	0	0
6	1	1	1	0	2	1	P
7	1	1	1	1	2	2	0
8	0	0	0	1	0	1	N

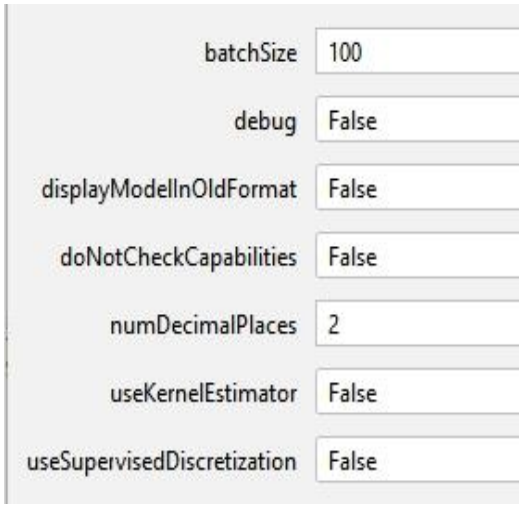
3. การแปลงค่าความถี่ให้เป็นนามบัญญัติ เป็นการเปลี่ยนค่าความถี่ของคุณลักษณะที่พบในความคิดเห็นของแต่ละความคิดเห็น เช่น ความคิดเห็นที่ 2 มีการกล่าวถึงคำว่า “ไม่” จำนวน 2 ครั้ง ดังนั้นจึงต้องแปลงจาก 2 เป็น 1 ซึ่งหมายความว่า คำว่า “ไม่” ได้มีการกล่าวถึง หากความถี่เป็น 0 หมายถึง คุณลักษณะนี้ไม่ถูกกล่าวถึง ดังตารางที่ 3

ตารางที่ 3 ตัวอย่างการแปลงค่าความถี่ให้เป็นนามบัญญัติ

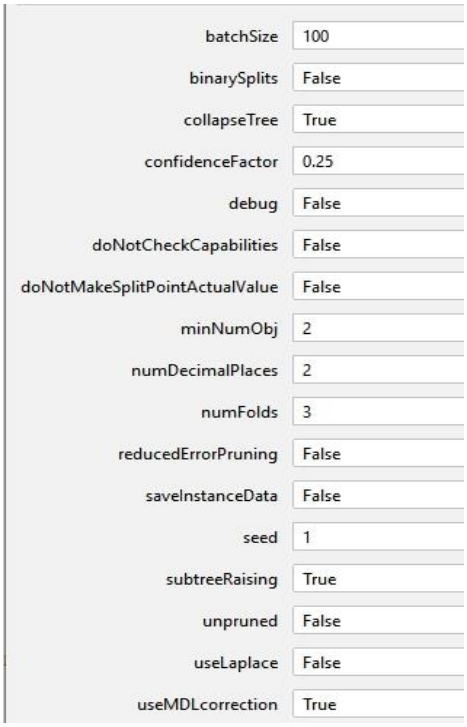
ลำดับ	คุณลักษณะเชิงบวก		คุณลักษณะเชิงลบ		คลาส
	ชอบ	ได้	ไม่	แย่	
1	1	0	0	0	P
2	0	1	1	0	N
3	1	1	0	0	P
4	0	0	0	0	0

3.3.4 การสร้างแบบจำลองเพื่อการจำแนก

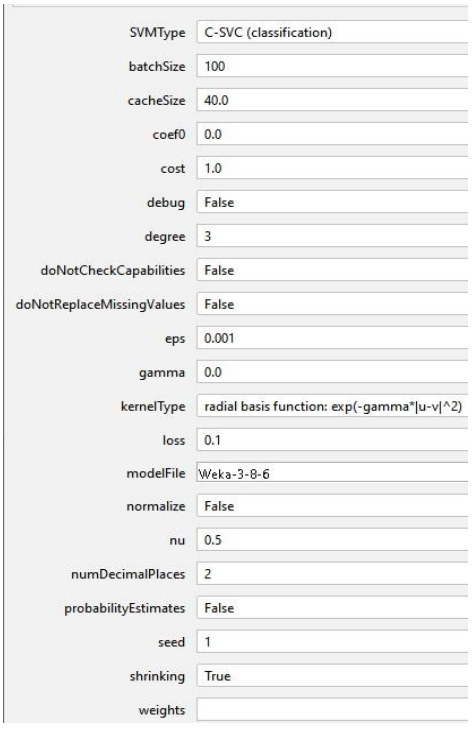
ในการวิจัยการทำเหมืองความคิดเห็นของผู้ใช้รถพลังงานทางเลือก ผู้วิจัยได้สร้างแบบจำลองสำหรับการวิเคราะห์โดยใช้โปรแกรม Weka เวอร์ชัน 3.8.6 ซึ่งประกอบด้วยเทคนิคหลัก 4 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ (Decision Tree C4.5), เทคนิคเนอีฟเบย์ (Naïve-Bayes), เทคนิคซัพพอร์ตเวกเตอร์ แมชชีน (Support Vector Machine), และเทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ซึ่งเป็นเทคนิคที่ได้รับความนิยมในการสร้างเหมืองความคิดเห็น โดยเทคนิคเทคนิคต้นไม้ตัดสินใจ แบบซี 4.5 ทำงานโดยการแบ่งข้อมูลออกเป็นกลุ่มต่าง ๆ ตามลักษณะของข้อมูล โดยเลือกใช้คุณลักษณะที่ช่วยแบ่งข้อมูลได้ดีที่สุด การแบ่งจะเกิดขึ้นเรื่อย ๆ จนกว่าจะสร้างโครงสร้างต้นไม้ครบ ซึ่งใช้เกณฑ์ในการเลือกคุณลักษณะโดยอิงจากข้อมูลความไม่แน่นอน (entropy) และการเพิ่มขึ้นของข้อมูล (information gain) มีข้อดีคือ สามารถใช้งานกับข้อมูลเชิงประเภท (categorical) และเชิงตัวเลข (numerical) ได้ดีส่วนข้อเสียคือ มีโอกาสเกิดการเรียนรู้ที่มากเกินไป (overfitting) ได้ง่าย เทคนิคเนอีฟเบย์ เป็นเทคนิคการจำแนกประเภทที่อิงจากกฎเบย์ (Bayes' Theorem) ซึ่งสมมติให้ทุกคุณลักษณะเป็นอิสระต่อกัน กล่าวคือจะพิจารณาโอกาสของข้อมูลแต่ละชุดตามความน่าจะเป็นโดยไม่คำนึงถึงการเชื่อมโยงกัน การทำงานนี้จึงเหมาะกับการประมวลผลข้อความที่มีการกระจายตัวของคำ ข้อดีของเทคนิคเทคนิคเนอีฟเบย์คือทำงานได้เร็ว และเหมาะสมกับข้อมูลขนาดใหญ่และทำงานได้ดีแม้จะมีข้อมูลที่ไม่มีการเชื่อมโยง (independent features) ส่วนข้อเสียคือ ข้อสมมติที่ว่าคุณลักษณะเป็นอิสระต่อกัน อาจไม่สอดคล้องกับความจริง ทำให้มีความแม่นยำลดลง เทคนิคซัพพอร์ตเวกเตอร์แมชชีน ทำงานโดยการหาขอบเขตที่เหมาะสมที่สุดในการแบ่งข้อมูลให้มีระยะห่างระหว่างกลุ่มข้อมูลที่ต่างกันให้มากที่สุด เรียกว่าขอบเขตการแบ่งแบบไฮเปอร์เพลน (hyperplane) ซึ่งเทคนิคนี้ จะพิจารณาจุดที่เรียกว่าเวกเตอร์สนับสนุน (support vectors) ที่ช่วยกำหนดขอบเขตเหล่านี้ ข้อดีคือ มีความแม่นยำสูง และเหมาะกับข้อมูลที่มีการจำแนกอย่างชัดเจน ยากต่อการเกิดการเรียนรู้ที่มากเกินไป (overfitting) และสามารถปรับ kernel trick เพื่อจัดการกับข้อมูลที่ไม่เป็นเส้นตรงได้ ข้อเสียคือ ใช้เวลาคำนวณสูงเมื่อข้อมูลมีจำนวนมาก การปรับค่าพารามิเตอร์ (parameter tuning) แม้เพียงเล็กน้อยจะมีผลต่อการคำนวณ และเทคนิคเพอร์เซปตรอนหลายชั้น เป็นโครงข่ายประสาทเทียมที่มีการเชื่อมต่อระหว่างชั้นข้อมูลหลายชั้น ทำงานโดยการเรียนรู้ผ่านกระบวนการถ่ายทอดข้อมูลจากชั้นอินพุต ผ่านชั้นซ่อน (hidden layer) และให้ผลลัพธ์ที่ชั้นเอาต์พุต เทคนิคนี้มีการปรับค่าเวท (weight) ของการเชื่อมต่อผ่านการคำนวณย้อนกลับ (backpropagation) ข้อดีของเทคนิคนี้คือ มีความยืดหยุ่นสูง และสามารถปรับแต่งโครงสร้างเครือข่ายเพื่อเพิ่มความแม่นยำ และข้อเสียคือ ต้องการข้อมูลที่มีปริมาณมากและคุณภาพดีเพื่อให้เกิดการเรียนรู้ที่แม่นยำ สำหรับการวิจัยในครั้งนี้ผู้วิจัยได้ใช้การตั้งค่าในโปรแกรม Weka ในบางกรณีเท่านั้น เนื่องจากทดลองปรับแต่งค่าต่าง ๆ ในการทดลองหลาย ๆ ครั้ง ผลที่ได้มักจะเกิดข้อผิดพลาด (Error) หรือ เกิดการเรียนรู้ที่มากเกินไป (Overfitting) ดังแสดงในรูปที่ 6 – 9



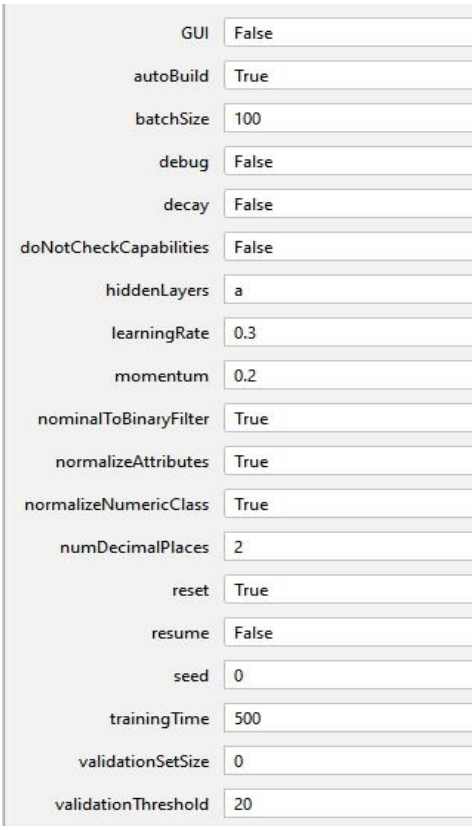
รูปที่ 6 แสดงการปรับพารามิเตอร์ของเทคนิคเนอีฟเบย์



รูปที่ 7 แสดงการปรับพารามิเตอร์ของเทคนิคต้นไม้ตัดสินใจ



รูปที่ 8 แสดงการปรับพารามิเตอร์ของเทคนิคซัพพอร์ตเวกเตอร์แมชชีน



รูปที่ 9 แสดงการปรับพารามิเตอร์ของเทคนิคเพอร์เซปตรอนหลายชั้น

3.3.5 การแปลผลและการประเมินผล

การประมวลผลของแบบจำลองในงานวิจัยนี้ผู้วิจัยได้เรียกใช้เทคนิคการวัดประสิทธิภาพของแบบจำลองแบบ 10 - Fold Cross Validation [9] ซึ่งแบ่งข้อมูลทั้งหมดออกเป็น 10 ชุด โดยใช้ 1 ชุด สำหรับเป็นชุดการเรียนรู้ (Training Data) และอีก 9 ชุด ใช้เป็นชุดทดสอบ (Test Data) ซึ่งหากคิดเป็นอัตราข้อมูลในการทดสอบออกเป็นร้อยละที่ 10:90 โดยมีค่าที่ใช้ในการวัดประสิทธิภาพของแบบจำลองทั้งหมด 3 ค่าได้แก่ ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วง (F-Measure)

สถิติที่ใช้ในการวิจัย

ในงานทดลองครั้งนี้ ใช้สถิติสามตัวหลักในการประเมินประสิทธิภาพของแบบจำลอง ได้แก่ ค่าความแม่นยำ (Precision), ค่าความระลึก (Recall), และค่าความถ่วง (F-Measure) แบบจำลองที่มีค่าความแม่นยำ ค่าความระลึกและค่าความถ่วงในอัตราร้อยละสูงสุดจะถือเป็นแบบจำลองที่มีประสิทธิภาพดีที่สุดในการวิจัยนี้ การคำนวณค่าดังกล่าวดำเนินการตามสมการทั้ง 3 สมการ ดังต่อไปนี้

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (2)$$

$$\text{F} - \text{Measure} = 2 * \frac{\text{precision}*\text{recall}}{\text{precision}+\text{recall}} \quad (3)$$

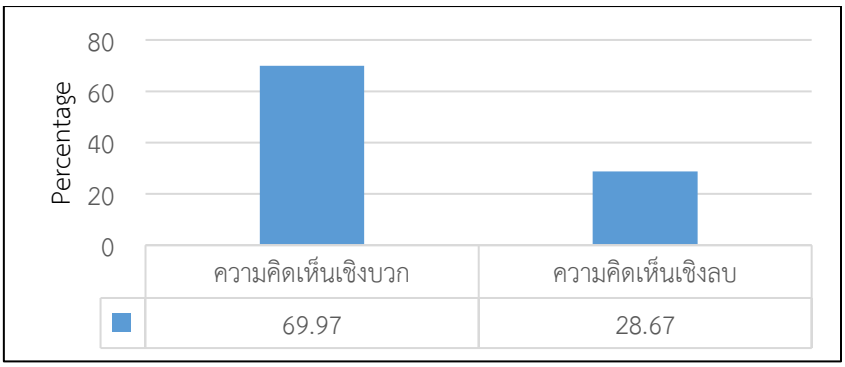
โดยที่ TP แทนจำนวนการทำนายที่ตรงกับข้อมูลจริงในคลาสที่กำลังพิจารณา

FP แทนจำนวนการทำนายที่ผิดในคลาสที่กำลังพิจารณา

FN แทนจำนวนการทำนายที่ผิดในคลาสที่ไม่ได้พิจารณา

4. ผลการวิจัย

4.1. ผลศึกษาและจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก



รูปที่ 10 กราฟการเปรียบเทียบความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทดแทน

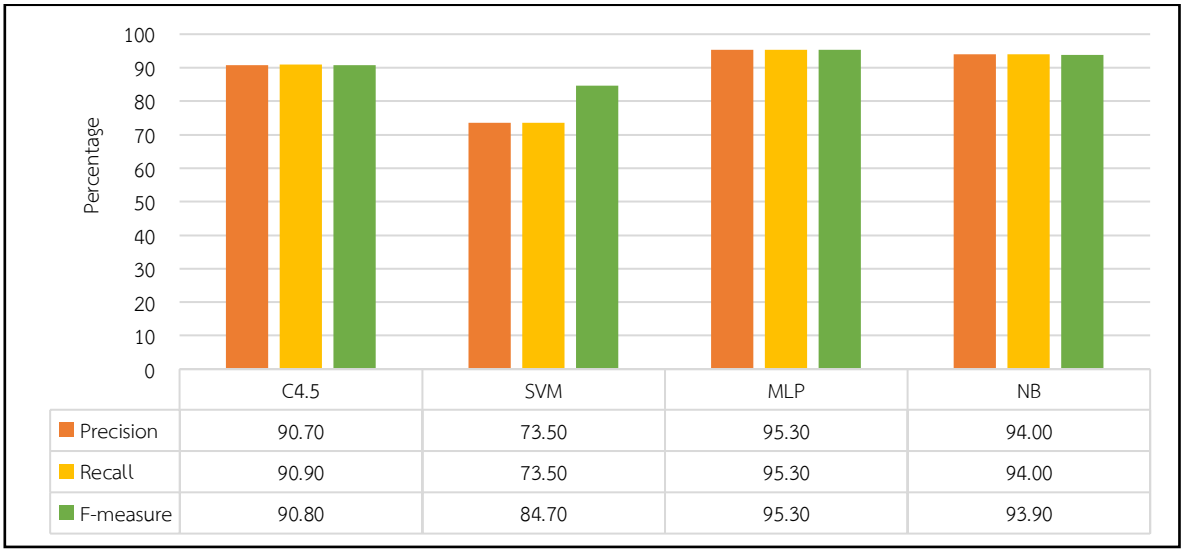
จากรูปที่ 10 แสดงการเปรียบเทียบการจำแนกความคิดเห็นของคนไทยต่อยานยนต์พลังงานทางเลือก ตามขั้นตอนการวิจัยทั้งหมด ใช้ข้อมูลความคิดเห็น 3,000 ชุด โดยแยกคุณลักษณะในความคิดเห็นและให้น้ำหนักความคิดเห็นเชิงบวกด้วยค่า 1, ความคิดเห็นเชิงลบด้วยค่า -1 และ 0 สำหรับความคิดเห็นที่เป็นกลางหรือถูกตัดออก ผลการทดลองพบว่าคนไทยมีความรู้สึกเชิงบวกอยู่ที่ร้อยละ 69.97 และความรู้สึกเชิงลบที่ร้อยละ 28.67 คำที่มีคุณลักษณะเชิงบวกที่ถูกกล่าวถึงมากที่สุดคือ "ดี" จำนวน 696 ครั้ง และคำที่มีคุณลักษณะเชิงลบที่ถูกกล่าวถึงมากที่สุดคือ "ปัญหา" จำนวน 202 ครั้ง ซึ่งสามารถแสดงคุณลักษณะที่ถูกกล่าวถึง 10 อันดับทั้งในเชิงบวกและเชิงลบได้ดังปรากฏในตารางที่ 4

ตารางที่ 4 แสดงการจัดอันดับคุณลักษณะที่ถูกกล่าวถึงมากที่สุด 10 อันดับ

ลำดับ	คุณลักษณะเชิงบวก	จำนวน(ครั้ง)	คุณลักษณะเชิงลบ	จำนวน(ครั้ง)
1	ดี	696	ปัญหา	202
2	สวย	510	แพง	195
3	ชอบ	350	จ่าย	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

4.2. ผลการศึกษาและพัฒนาเทคนิคที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือก

ผู้วิจัยได้ดำเนินการศึกษาและพัฒนาเทคนิคที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของคนไทยต่อยานยนต์พลังงานทางเลือก ตามขั้นตอนการวิจัยทั้งหมด โดยใช้เทคนิคการทำเหมืองความรู้สึก 4 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคเนอีฟเบย์ เทคนิคการเรียนรู้ของเครื่อง และเทคนิคเพอร์เซปตรอนหลายชั้น ผ่านโปรแกรม Weka เวอร์ชั่น 3.8.6 การทดลองใช้ข้อมูลความคิดเห็นจำนวน 3,000 ชุด โดยแบ่งออกเป็น 2 คลาส: คลาส P ซึ่งระบุถึงความคิดเห็นเชิงบวกจำนวน 231 คลาส และคลาส N ซึ่งระบุถึงความคิดเห็นเชิงลบจำนวน 208 คลาส การทดสอบประสิทธิภาพของเทคนิคที่ใช้ในการวิจัยนี้ดำเนินการด้วยวิธี 10-Fold Cross Validation เพื่อแบ่งข้อมูลออกเป็นชุดชุดเรียนรู้และทดสอบในอัตรา 1 : 9 และเปรียบเทียบประสิทธิภาพของแบบจำลองโดยใช้ค่าความแม่นยำ (Precision), ค่าความระลึก (Recall), และค่าความถ่วง (F-Measure) ผลลัพธ์การวิจัยแสดงไว้ใน รูปที่ 11



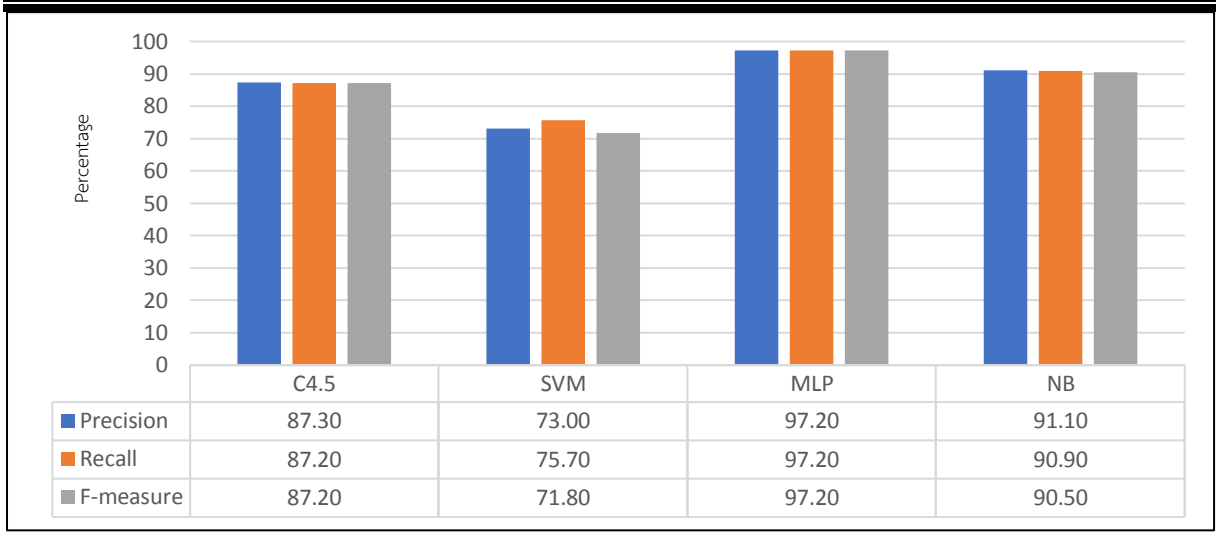
	C4.5	SVM	MLP	NB
Precision	90.70	73.50	95.30	94.00
Recall	90.90	73.50	95.30	94.00
F-measure	90.80	84.70	95.30	93.90

รูปที่ 11 กราฟการเปรียบเทียบประสิทธิภาพของแบบจำลอง

จากรูปที่ 11 แสดงการเปรียบเทียบประสิทธิภาพของแบบจำลองที่ใช้ในการวิจัยทั้ง 4 เทคนิค พบว่า เทคนิคเพอร์เซปตรอนหลายชั้น มีประสิทธิภาพสูงสุดในการวิจัยครั้งนี้ โดยมีค่าความแม่นยำ ค่าความระลึก และค่าความถ่วงที่ร้อยละ 95.30 รองลงมาได้แก่ เทคนิคเนอีฟเบย์ มีค่าความแม่นยำและค่าความระลึกที่ร้อยละ 94.00 และค่าความถ่วงที่ร้อยละ 93.90 เทคนิคต้นไม้ตัดสินใจ มีค่าความแม่นยำร้อยละ 90.70 ค่าความระลึกร้อยละ 90.90 ค่าความถ่วงร้อยละ 90.80 และเทคนิคการเรียนรู้ของเครื่องเป็นเทคนิคให้ประสิทธิภาพต่ำที่สุดในการวิจัยครั้งนี้โดยมี ค่าความแม่นยำร้อยละ 73.50 ค่าความระลึกร้อยละ 73.50 และค่าความถ่วงร้อยละ 84.70

4.3. ผลการศึกษาและพัฒนาเทคนิคที่มีประสิทธิภาพสูงสุดในการจำแนกความคิดเห็นของคนไทยที่มีต่อยานยนต์พลังงานทางเลือกโดยพิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้ง ขึ้นไป

จากการคัดเลือกคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบทั้งหมด 283 รายการ แบ่งเป็นคุณลักษณะเชิงบวก 173 รายการ และเชิงลบ 110 รายการ ในการทดสอบประสิทธิภาพของเทคนิค ใช้วิธี 10-Fold Cross Validation เพื่อแบ่งข้อมูลออกเป็นชุดชุดเรียนรู้และทดสอบในอัตรา 1 : 9 แล้วเปรียบเทียบประสิทธิภาพของแบบจำลองด้วยค่าความแม่นยำ (Precision), ความระลึก (Recall) และ F-Measure ผลลัพธ์แสดงในรูปที่ 12



	C4.5	SVM	MLP	NB
Precision	87.30	73.00	97.20	91.10
Recall	87.20	75.70	97.20	90.90
F-measure	87.20	71.80	97.20	90.50

รูปที่ 12 กราฟการเปรียบเทียบประสิทธิภาพของแบบจำลองพิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้ง ขึ้นไป

จากรูปที่ 12 แสดงการเปรียบเทียบประสิทธิภาพของแบบจำลองที่ใช้ในการวิจัยทั้ง 4 เทคนิค โดยพิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป พบว่า เทคนิคเพอร์เซปตรอนหลายชั้น ยังมีประสิทธิภาพสูงสุด โดยมีค่าความแม่นยำ ค่าความระลึก และค่าความถ่วงที่ร้อยละ 97.20 รองลงมาได้แก่ เทคนิคเนอีฟเบย์ มีค่าความแม่นยำที่ร้อยละ 91.10 ค่าความระลึกที่ร้อยละ 90.90 และค่าความถ่วงที่ร้อยละ 90.50 เทคนิคต้นไม้ตัดสินใจ มีค่าความแม่นยำร้อยละ 87.30 ค่าความระลึกและค่าความถ่วงร้อยละ 87.20 และเทคนิคการเรียนรู้ของเครื่องเป็นเทคนิคให้ประสิทธิภาพต่ำที่สุดในการวิจัยครั้งนี้ โดยมี ค่าความแม่นยำร้อยละ 73.00 ค่าความระลึกร้อยละ 75.70 และค่าความถ่วงร้อยละ 71.80

5. อภิปรายผลการวิจัย

จากการวิจัยการทำเหมืองความคิดเห็นของผู้ใช้รถพลังงานทางเลือก ในครั้งนี้ทำให้ พบว่า เทคนิคเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ได้ให้ผลลัพธ์ที่ดีที่สุดทั้งแบบทั่วไปและพิจารณาจากคุณลักษณะที่ถูกกล่าวถึงตั้งแต่ 10 ครั้งขึ้นไป โดยมีค่าความแม่นยำ ค่าความระลึกและค่าความถ่วงที่ร้อยละ 95.30 สำหรับการทดลองทั่วไป และร้อยละ 97.20 สำหรับการทดลองที่พิจารณาจากคุณลักษณะที่ถูกกล่าวถึง 10 ครั้งขึ้นไป ซึ่งต่างจากงานวิจัยของ Davis และคณะ [10] ที่ได้ผลลัพธ์จากการทดลองเป็นเทคนิค ต้นไม้ตัดสินใจ ที่ให้ผลลัพธ์ที่ดีที่สุดโดยมีค่าความแม่นยำที่ร้อยละ 90.00 จากการวิจัยการสร้างสถานการณ์โดยอัตโนมัติจากคลังข้อความ : กรณีศึกษาเกี่ยวกับยานพาหนะไฟฟ้า ส่วนงานวิจัยของกิตติศักดิ์ ขำจิตร และคณะ [11] ที่ได้ใช้เทคนิคเทคนิคเนอีฟเบย์ เทคนิคต้นไม้ตัดสินใจ เทคนิคป่าแบบสุ่ม และเทคนิคการเรียนรู้เชิงลึก พบว่าได้ผลลัพธ์ต่างกัน โดยเทคนิคที่มีประสิทธิภาพที่สุด คือ เทคนิคการเรียนรู้เชิงลึก มีค่าความแม่นยำร้อยละ 95.11 ค่าความถ่วงร้อยละ 97.49 และค่าความไวร้อยละ 99.11 ในการวิจัยประสิทธิภาพของเทคนิคเหมืองข้อมูลสำหรับพยากรณ์การเกิดโรคหลอดเลือดในสมอง และหนังสือคู่มือการวิจัยการขุดความคิดเห็นและการวิเคราะห์ข้อความในงานวรรณกรรมและโซเชียลมีเดีย [12] ได้เสนอ เทคนิคการเรียนรู้ของเครื่อง ว่าเป็นเทคนิคที่เหมาะสมในการทำเหมืองความคิดเห็นในช่วงทศวรรษนี้ ซึ่งสอดคล้องกับเทคนิคเพอร์เซปตรอนหลายชั้น ซึ่งได้พัฒนาจากเทคนิคการเรียนรู้ของเครื่องจึงได้ผลลัพธ์ที่ดีที่สุดจาก 4 เทคนิคในการวิจัยครั้งนี้ เนื่องจากชุดข้อมูลที่ต่างกันจึงทำให้ได้คุณลักษณะของคำมีความแตกต่างกันจึงส่งผลให้ผลการวิจัยแตกต่างกันแม้จะใช้เทคนิคในการวิเคราะห์ที่เหมือนกันก็ตาม ซึ่งข้อมูลที่ได้จากการวิจัยครั้งนี้ยังสามารถนำไปวิเคราะห์เพิ่มเติมได้อีก เช่น การใช้ความถี่เชิงคำ – ความถี่ผกผันเอกสาร (Term Frequency-Inverse Document Frequency: TF-IDF) เพื่อใช้วัดความสำคัญของคุณลักษณะกับชุดข้อมูลอื่น ๆ เพื่อใช้วิเคราะห์ความรู้สึกของผู้ใช้งานยานยนต์พลังงานทางเลือก

6. เอกสารอ้างอิง

[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, Aug. 2023, DOI:10.1017/S1755773923000255.

[2] Nation Group (Thailand) Public Company Limited. (2024, January 19). *The reason why Thai people*

use EV more: Charging costs are cheaper than oil, environmentally friendly [Online]. Available: https://www.posttoday.com/smart-city/704607 (in Thai)

[3] Bangkok Media and Broadcasting Co., Ltd.. (2022, September 11). *Electric cars driving through floods and charging while raining, are they safe ?* [Online]. Available: https://www.pptvhd36.com/automotive/carcare/176850 (in Thai)

[4] B. Liu, *Sentiment analysis and opinion mining*. Switzerland: Springer Nature, 2022.

[5] P. Namwong and S. 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.

[6] B. Pisit and P. Worapat, "Sentiment Analysis Techniques of Online Product Reviews," *Rajamangala University of Technology Srivijaya Research Journal,* vol. 14, no. 3, pp. 755-769, 2022.

[7] 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.

[8] GitHub, Inc. (2023, December 19). AttaCut: Fast and Reasonably Accurate Word Tokenizer for Thai [Online]. Available: https://github.com/PyThaiNLP/attacut

[9] E. Elsaeed, O. Ouda, M. M. Elmogy, A. Atwan, and E. El-Daydamony, "*Detecting fake news in social media using voting classifier," IEEE Access,* vol. 9, pp. 161909-161925, 2021.

[10] 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.

[11] 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.

[12] P. Keikhosrokiani and M. Pourya Asl, *Handbook of research on opinion mining and text analytics on literary works and social media*. IGI Global, 2022.