

การวิเคราะห์ทัศนคติสาธารณะต่อรถยนต์ไฟฟ้าในไทย: จากกระทู้สนทนาบน Pantip ด้วยวิธีการเรียนรู้ของเครื่อง

Public Sentiment Toward Electric Vehicles in Thailand: A Machine Learning Analysis of Pantip Discussions

ชนกานต์ กิ่งแก้ว^{1*}, พรสิริ ขาติปริชา² และ บัณฑิต อินทรีย์มีศักดิ์³

^{1,3} คณะวิศวกรรมศาสตร์และเทคโนโลยี สถาบันการจัดการปัญญาภิวัฒน์ 85/1 หมู่ 2 ถ.แจ้งวัฒนะ ต.บางตลาด อ.ปากเกร็ด จ.นนทบุรี 11120

² คณะอุตสาหกรรมเกษตร มหาวิทยาลัยเกษตรศาสตร์ 50 ถนนงามวงศ์วาน แขวงลาดยาว เขตจตุจักร กรุงเทพมหานคร 10900

*ติดต่อ: chanakarnkh@pim.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้า (Electric Vehicles หรือ EVs) ในประเทศไทย โดยใช้ข้อมูลจากการสนทนาในเว็บไซต์ Pantip ซึ่งเป็นแหล่งข้อมูลความคิดเห็นของผู้บริโภคที่มีความหลากหลาย ข้อมูลที่ใช้ในการศึกษามีจำนวนมากกว่า 10,000 ข้อความ ครอบคลุมช่วงปี พ.ศ. 2560–2568 กระบวนการวิจัยประกอบด้วย การรวบรวมข้อมูล การทำความสะอาดข้อความ การตัดคำภาษาไทย และการแปลงข้อความเป็นคุณลักษณะด้วยเทคนิค TF-IDF จากนั้นนำไปพัฒนาแบบจำลองการเรียนรู้ของเครื่อง ได้แก่ Naïve Bayes (NB), Logistic Regression (LR) และ Support Vector Machine (SVM) เพื่อจำแนกความรู้สึกออกเป็น 3 ประเภท ได้แก่ เชิงบวก เชิงลบ และเป็นกลาง ผลการทดลองพบว่า Logistic Regression ให้ประสิทธิภาพสูงสุด โดยมีค่า Accuracy เท่ากับ 0.7425 และค่า F1-score เฉลี่ยเท่ากับ 0.7215 รองลงมาคือ SVM และ Naïve Bayes ตามลำดับ นอกจากนี้ยังพบว่าแบบจำลองภาษาขนาดใหญ่ (LLM) แบบ zero-shot มีประสิทธิภาพต่ำกว่าแบบจำลองดั้งเดิมอย่างมีนัยสำคัญ ผลการวิจัยสามารถนำไปใช้เป็นข้อมูลสนับสนุนการตัดสินใจเชิงนโยบายและธุรกิจเกี่ยวกับการส่งเสริมการใช้ยานยนต์ไฟฟ้าในประเทศไทย

คำหลัก: รถยนต์ไฟฟ้า (EV), การวิเคราะห์ทัศนคติ, การเรียนรู้ของเครื่อง, การประมวลผลภาษาธรรมชาติ

Abstract

This study aims to analyze public sentiment toward electric vehicles (EVs) in Thailand using discussions collected from the Pantip online platform. The dataset consists of over 10,000 Thai-language comments spanning from 2017 to 2025. The research process includes data collection, text cleaning, Thai tokenization, and feature extraction using TF-IDF. Three machine learning models—Naïve Bayes (NB), Logistic Regression (LR), and Support Vector Machine (SVM)—were developed to classify sentiment into three categories: positive, negative, and neutral. The experimental results show that Logistic Regression achieved the best performance with an accuracy of 0.7425 and a macro F1-score of 0.7215, followed by SVM and Naïve Bayes. Additionally, a large language model (LLM) using a zero-shot approach performed

significantly lower than traditional machine learning models. The findings provide valuable insights for policymakers and businesses in promoting EV adoption in Thailand.

Keywords: Electric Vehicles (EV), Sentiment Analysis, Machine Learning, Thai Natural Language Processing (NLP)

1. บทนำ

ท่ามกลางวิกฤตการณ์ราคาน้ำมันที่ปรับตัวสูงขึ้นอย่างต่อเนื่องและความผันผวนด้านการจัดหาพลังงาน งานวิจัยฉบับนี้จึงมุ่งศึกษาและสำรวจทัศนคติรวมถึงความต้องการของประชาชนในสังคมไทยที่มีต่อรถยนต์ไฟฟ้า (Electric Vehicles หรือ EVs) โดยมุ่งเน้นที่ปัจจัยด้านความคุ้มค่า ความเชื่อมั่นในเทคโนโลยี และข้อจำกัดในการใช้งานจริง เพื่อสะท้อนให้เห็นถึงแนวโน้มการปรับตัวและมุมมองของผู้บริโภคไทยต่อการเปลี่ยนผ่านสู่การใช้พลังงานทางเลือกในยุคปัจจุบัน

สำหรับในช่วงหลายปีที่ผ่านมา ยานยนต์ไฟฟ้าได้รับความสนใจเพิ่มขึ้นมาอย่างต่อเนื่องทั่วโลก รวมถึงในประเทศไทย เนื่องจากเป็นทางเลือกสำคัญในการลดการปล่อยก๊าซเรือนกระจก ลดมลพิษทางอากาศ สนับสนุนการใช้พลังงานอย่างยั่งยืน รวมถึงได้รับการสนับสนุนด้านสิทธิพิเศษทางภาษี [11] โดยมีงานวิจัยจำนวนมากที่ศึกษาปัจจัยที่มีผลต่อการยอมรับยานยนต์ไฟฟ้าของผู้บริโภค เช่น ราคา โครงสร้างพื้นฐาน และประสบการณ์การใช้งาน [6], [7] ในยุคดิจิทัล ข้อมูลความคิดเห็นของผู้บริโภคสามารถเข้าถึงได้จากแพลตฟอร์มออนไลน์ เช่น เว็บไซต์ Pantip ซึ่งเป็นแหล่งข้อมูลที่สะท้อนมุมมองของผู้ใช้งานจริง การวิเคราะห์ความรู้สึก (Sentiment Analysis) จากข้อความดังกล่าวจึงเป็นเครื่องมือสำคัญในการทำความเข้าใจความคิดเห็นของสังคม โดยมีการนำเทคนิค Machine Learning และ Deep Learning มาใช้กันอย่างแพร่หลาย [5]

อย่างไรก็ตาม การวิเคราะห์ข้อความภาษาไทยยังมีความท้าทาย เนื่องจากลักษณะเฉพาะของภาษา เช่น การไม่มีการเว้นวรรคระหว่างคำ และความกำกวมของบริบท ซึ่งทำให้ต้องใช้เครื่องมือและโมเดลที่เหมาะสมกับ

ภาษาไทยโดยเฉพาะ [2] นอกจากนี้ โมเดลภาษาไทย เช่น WangchanBERTa ได้แสดงศักยภาพสูงในการประมวลผลภาษาธรรมชาติ [1]

ในระยะหลัง Large Language Models หรือ LLMs ได้ถูกนำมาใช้ในการวิเคราะห์ข้อความแบบ Zero-Shot มากขึ้น แต่ยังคงพบข้อจำกัดในด้านความแม่นยำเมื่อเทียบกับโมเดลที่ฝึกเฉพาะงาน โดยเฉพาะในบริบทหลายภาษา [10] ดังนั้น งานวิจัยนี้จึงมุ่งวิเคราะห์ความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้าในประเทศไทย โดยใช้ข้อมูลจากเว็บไซต์ Pantip และเปรียบเทียบประสิทธิภาพของโมเดล Machine Learning ได้แก่ Naïve Bayes, Support Vector Machine และ Logistic Regression ร่วมกับ Large Language Model เพื่อให้ได้ข้อสรุปเชิงเปรียบเทียบที่เหมาะสมกับบริบทภาษาไทย

2. วัตถุประสงค์ของการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาความคิดเห็นและความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้าในประเทศไทย โดยอาศัยข้อมูลจากการสนทนาในสื่อสังคมออนไลน์ ซึ่งเป็นแหล่งข้อมูลสำคัญในการสะท้อนมุมมองของผู้บริโภคในยุคดิจิทัล [6], [7] โดยมีรายละเอียดดังนี้

1) เพื่อศึกษาความคิดเห็นและความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้าในประเทศไทยจากข้อมูลการสนทนาในเว็บไซต์ Pantip ซึ่งเป็นแหล่งข้อมูลที่สะท้อนพฤติกรรมและทัศนคติของผู้บริโภคในสังคมไทย [4]

2) เพื่อพัฒนาและประยุกต์ใช้กระบวนการวิเคราะห์ความรู้สึก (Sentiment Analysis) สำหรับข้อความภาษาไทย โดยใช้เทคนิคการประมวลผลภาษาธรรมชาติ (Natural Language Processing หรือ

NLP) ซึ่งได้รับการพัฒนาอย่างต่อเนื่องสำหรับภาษาไทย [2], [3]

3) เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง ได้แก่ Naïve Bayes (NB), Logistic Regression (LR) และ Support Vector Machine (SVM) ซึ่งเป็นโมเดลที่นิยมใช้ในการจำแนกข้อความ และมีประสิทธิภาพดีในงาน sentiment analysis [5]

4) เพื่อประเมินและเปรียบเทียบผลลัพธ์กับการใช้ Large Language Model (LLM) ในการวิเคราะห์ความรู้สึกแบบ Zero-Shot ซึ่งกำลังได้รับความสนใจในงานวิจัยด้าน NLP [10]

5) เพื่อวิเคราะห์แนวโน้มความคิดเห็นของผู้บริโภคต่อยานยนต์ไฟฟ้า เพื่อนำไปใช้เป็นข้อมูลสนับสนุนการตัดสินใจเชิงนโยบายและธุรกิจ [8], [9]

6) เพื่อสกัดองค์ความรู้เชิงลึกเกี่ยวกับปัจจัยที่มีอิทธิพลต่อทัศนคติของผู้บริโภคต่อการใช้งานยานยนต์ไฟฟ้าในประเทศไทย โดยอาศัยข้อมูลจากการวิเคราะห์ข้อความขนาดใหญ่ [6]

3. ขอบเขตของการวิจัย

1) ขอบเขตด้านข้อมูล (Data Scope) สำหรับข้อมูลที่ใช้ในการศึกษาได้มาจากการรวบรวมความคิดเห็นของผู้ใช้งานจากเว็บไซต์ Pantip โดยเลือกเฉพาะกระทู้ที่เกี่ยวข้องกับยานยนต์ไฟฟ้าในระหว่างปี พ.ศ. 2560-2568

2) ขอบเขตด้านเนื้อหา (Content Scope) ทางที่มวิจัยมุ่งเน้นการวิเคราะห์ความรู้สึก (Sentiment Analysis) ของข้อความภาษาไทย โดยจำแนกออกเป็น 3 กลุ่ม ได้แก่ ความรู้สึกเชิงบวก (Positive), เชิงลบ (Negative) และเป็นกลาง (Neutral)

3) ขอบเขตด้านเทคนิคและวิธีการ (Technical and Methodological Scope) ทางที่มวิจัยใช้กระบวนการประมวลผลข้อความภาษาไทย ประกอบด้วย การทำความสะอาดข้อความ (Text Cleaning), การตัดคำ (Tokenization) และการแปลงข้อมูลเป็นคุณลักษณะด้วยเทคนิค TF-IDF

4) ขอบเขตด้านแบบจำลอง (Model Scope) สำหรับการศึกษาที่ใช้แบบจำลองการเรียนรู้ของเครื่องจำนวน 3 วิธี ได้แก่ Naïve Bayes (NB), Logistic Regression (LR) และ Support Vector Machine (SVM) เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกความรู้สึก

5) ขอบเขตด้านการประเมินผล (Evaluation Scope) สำหรับการประเมินประสิทธิภาพของแบบจำลองโดยใช้ตัวชี้วัด ได้แก่ Accuracy, Precision, Recall และ F1-score

4. วิธีการดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Analysis) โดยประยุกต์ใช้เทคนิคการประมวลผลภาษาธรรมชาติ และการเรียนรู้ของเครื่อง เพื่อวิเคราะห์ความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้า ในประเทศไทย จากข้อมูลการสนทนาในเว็บไซต์ Pantip ซึ่งเป็นแหล่งข้อมูลที่สะท้อนความคิดเห็นของผู้บริโภคในบริบทจริง [4], [6] กระบวนการวิจัยถูกออกแบบเป็นลำดับขั้น (Pipeline) ตั้งแต่การรวบรวมข้อมูล การเตรียมข้อมูล การสร้างคุณลักษณะ การพัฒนาแบบจำลอง และการประเมินผล เพื่อให้ได้ผลลัพธ์ที่มีความน่าเชื่อถือและสามารถนำไปใช้ประโยชน์ได้จริง แสดงขั้นตอนการดำเนินการวิจัยดังรูปที่ 1



รูปที่ 1 แสดงขั้นตอนการดำเนินการวิจัย

จากรูปที่ 1 แสดงขั้นตอนการดำเนินการวิจัย ซึ่งประกอบด้วยขั้นตอนหลักดังนี้

1. การรวบรวมข้อมูล (Data Collection)
2. การเตรียมข้อมูล (Data Preprocessing)

3. การแปลงข้อความเป็นคุณลักษณะ (Feature Extraction)
4. การสร้างและฝึกแบบจำลอง (Model Training)
5. การประเมินผล (Evaluation)

4.1 การรวบรวมข้อมูล (Data Collection)

สำหรับกลุ่มเป้าหมายในการวิจัยครั้งนี้ คือ ข้อมูลความคิดเห็น (User-generated content) ของผู้ใช้งานเว็บไซต์ Pantip ที่เกี่ยวข้องกับยานยนต์ไฟฟ้าในประเทศไทย โดยมีรายละเอียด ดังนี้

- เลือกกระทู้และความคิดเห็นที่เกี่ยวข้องกับหัวข้อ EV เช่น รถยนต์ไฟฟ้า แบตเตอรี่ การชาร์จไฟ และนโยบายภาครัฐ
- ข้อมูลครอบคลุมตั้งแต่ช่วงปี พ.ศ. 2560 – 2568 เพื่อสะท้อนแนวโน้มความคิดเห็นในระยะยาว
- ในการพัฒนาโปรแกรมสำหรับดึงข้อมูลจากเว็บไซต์ Pantip ผู้วิจัยได้ใช้เครื่องมือปัญญาประดิษฐ์ (AI-assisted coding) ในบางขั้นตอนเพื่อช่วยแนะนำและปรับปรุงโค้ดภาษา Python ร่วมกับการพัฒนาโดยผู้วิจัยเอง โดยใช้ไลบรารี เช่น Selenium และ BeautifulSoup สำหรับการดึงข้อมูลเว็บแบบอัตโนมัติ ทั้งนี้ ผู้วิจัยได้ทำการออกแบบโครงสร้างโปรแกรม ตรวจสอบความถูกต้องของโค้ด ปรับแต่งเงื่อนไขการดึงข้อมูล และทดสอบผลลัพธ์ด้วยตนเอง เพื่อให้มั่นใจว่าข้อมูลที่ได้มีความครบถ้วน ถูกต้อง และเหมาะสมสำหรับการวิเคราะห์
- โดยนำข้อมูลดึงจากเว็บ Pantip ที่ได้มาใส่ไว้ในรูปแบบตาราง CSV (Comma-Separated Values) โดยแยกตามปี ดังตารางที่ 1 ซึ่งประกอบไปด้วย 1. topic_url หรือ ที่อยู่ของเว็บไซต์ 2. title หรือ ชื่อหัวข้อของกระทู้ 3. type หรือ ประเภทของข้อความในกระทู้ว่าเป็นหัวข้อหลัก (topic) หรือ ความคิดเห็น (comment) 4. comment_no หรือ ตำแหน่งของความคิดเห็น และ 5. text หรือ ข้อความที่ได้จากการดึงข้อมูลดิบ

ตารางที่ 1 แสดงรูปแบบการเก็บข้อมูลที่ได้จากการดึงข้อมูลดิบจากเว็บ Pantip ในปี 2565

คอลัมน์	ค่าของข้อมูล
topic_url	https://pantip.com/topic/41626508
title	BYD เตรียมผลิตรถยนต์ไฟฟ้าในระยอง หวังเป็นแหล่งผลิตในอาเซียน รัฐบาลไทยหนุน - Pantip
type	topic
comment_no	0
text	มาไทยแล้ว! BYD เตรียมผลิตรถยนต์ไฟฟ้าในระยอง หวังเป็นแหล่งผลิตในอาเซียน รัฐบาลไทยหนุน . คณะกรรมการส่งเสริมการลงทุนแห่งประเทศไทย (BoI) ระบุว่า บริษัทผลิตรถยนต์ไฟฟ้าสัญชาติจีน BYD จะเข้ามาลงทุนเพื่อผลิตรถยนต์ไฟฟ้าในจังหวัดระยองด้วยเงินมูลค่า 17.9 พันล้านบาท โดยซื้อที่ดิน 700 ไร่จากนิคมอุตสาหกรรม WHA ...
period	2022
ตัวอย่างความเห็นที่ 1 ของกระทู้ที่แสดง:	ความคิดเห็นที่ 1 เบอร์ 2 ของโลกมาแล้ว รอเบอร์ 1 อีกเจ้าครับ ตอบกลับ 0 0 ถูกใจให้พอยต์ สมาชิกหมายเลข 724783 10 กันยายน 2565 เวลา 00:31 น.

จะเห็นว่าข้อมูลดิบที่ได้ ยังมีข้อมูลอื่น ๆ ที่ไม่ต้องการ เช่น ลำดับความคิดเห็น วัน เดือน ปี หมายเลขสมาชิก และอื่น ๆ จึงจะได้เข้าสู่กระบวนการเตรียมข้อมูลต่อไป

4.2 การเตรียมข้อมูล

ข้อมูลดิบประเภทข้อความ (Text) ที่ได้จาก Pantip ถูกนำมาผ่านกระบวนการเตรียมข้อมูล (Text preprocessing) เพื่อเพิ่มคุณภาพของข้อมูลก่อนนำไปวิเคราะห์ โดยมีขั้นตอน ดังนี้:

- ลบข้อมูลซ้ำ (Duplicate Removal)
- ลบ URL สัญลักษณ์พิเศษ และตัวอักษรที่ไม่จำเป็น
- แปลงข้อความเป็นรูปแบบมาตรฐาน (Normalization)
- ตัดคำภาษาไทย (Tokenization) โดยใช้เครื่องมือที่รองรับภาษาไทย [2]
- ลบ stop words และคำที่ไม่มีความหมายซึ่งกระบวนการดังกล่าวช่วยลด Noise และเพิ่มประสิทธิภาพของโมเดลในการเรียนรู้ข้อมูล [3]

จากนั้นเมื่อได้ข้อความที่สะอาดแล้ว ทำการกำหนดป้ายกำกับข้อมูล (Data Labeling) สำหรับข้อมูลถูกกำหนดป้ายกำกับความรู้สึกออกเป็น 3 กลุ่ม ได้แก่ เชิงบวก (Positive) เชิงลบ (Negative) และเป็นกลาง (Neutral) โดยใช้วิธีการกำหนดป้ายกำกับแบบกึ่งอัตโนมัติ (Semi-automatic labeling) ร่วมกับการตรวจสอบโดยผู้วิจัย เพื่อเพิ่มความถูกต้องและไม่นำเชื่อถือของข้อมูล

4.3 การแปลงคุณลักษณะ (Feature Extraction)

งานวิจัยนี้ใช้เทคนิค TF-IDF (Term Frequency-Inverse Document Frequency) ในการแปลงข้อความเป็นเวกเตอร์เชิงตัวเลข ซึ่งเป็นวิธีที่ได้รับความนิยมและมีประสิทธิภาพในงานจำแนกข้อความ [5] โดยการใช้ทั้ง word-level และ n-gram ช่วยให้โมเดลสามารถเรียนรู้รูปแบบของภาษาไทยได้ดีขึ้น

4.4 การสร้างและฝึกแบบจำลอง (Model Training)

ในการศึกษานี้ใช้แบบจำลองทั้งหมด 4 รูปแบบ ได้แก่

- 1) Naïve Bayes (NB)
- 2) Logistic Regression (LR)
- 3) Support Vector Machine (SVM) และ
- 4) Large Language Model (LLM) (zero-shot)

โดยโมเดล Machine Learning ถูกฝึกด้วยข้อมูล train set ขณะที่ LLM ถูกใช้เป็น baseline โดยไม่ต้องฝึกเพิ่มเติม [10]

4.5 การประเมินผล (Evaluation)

ในการประเมินผลแบบจำลองใช้หลักในการแบ่งข้อมูลสำหรับฝึก และ สำหรับทดสอบ โดยคัดเลือกข้อมูลแบบสุ่มจำนวน 2,000 ข้อความลำดับตามความใหม่ของข้อมูล สำหรับข้อมูลถูกแบ่งออกเป็น ชุดข้อมูลสำหรับฝึก (Training Set) 80% และชุดข้อมูลสำหรับทดสอบ (Testing Set) 20% โดยใช้วิธีการสุ่มเลือกแบบ stratified sampling เพื่อรักษาสัดส่วนของแต่ละ class ให้สมดุล

สำหรับตัวชี้วัดในด้านประสิทธิภาพของโมเดลถูกประเมินโดยใช้ตัวชี้วัด ได้แก่ ความถูกต้องโดยรวม (Accuracy), ความแม่นยำในการทำนายเชิงบวก (Precision), ความสามารถในการตรวจจับข้อมูลเชิงบวก (Recall) และค่าความสมดุลระหว่าง Precision และ Recall (F1-score) นอกจากนี้ยังใช้ Confusion Matrix เพื่อวิเคราะห์ความสามารถในการจำแนกแต่ละคลาส

5. ผลการดำเนินการวิจัย

5.1 การวิเคราะห์ข้อมูล

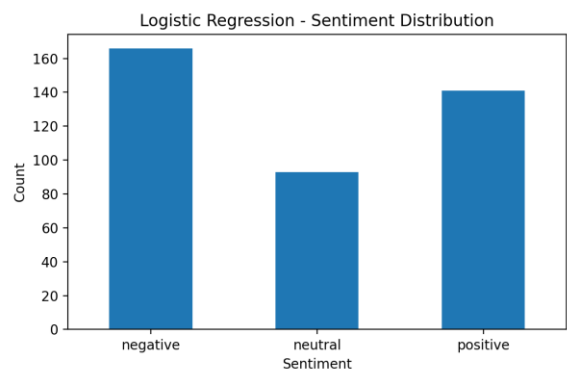
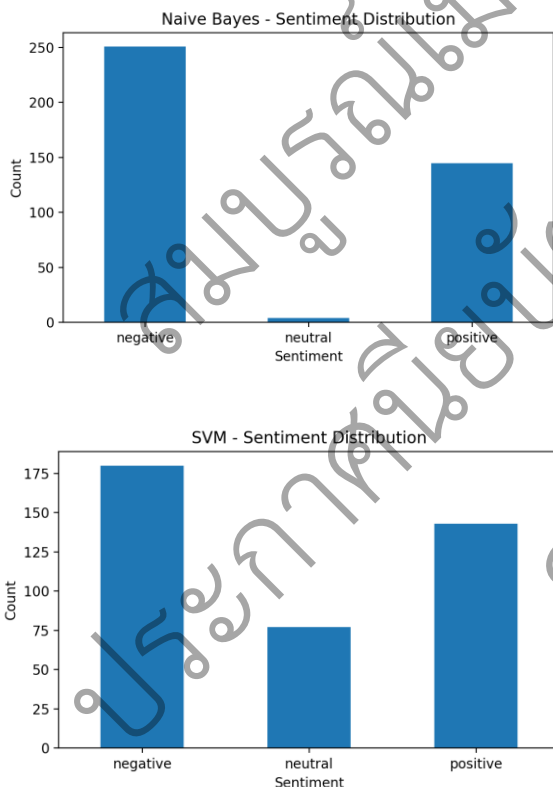
ผลการวิเคราะห์พบว่าความคิดเห็นเชิงลบมีสัดส่วนสูงกว่ากลุ่มอื่นเล็กน้อย โดยเมื่อพิจารณาเชิงเนื้อหา สามารถจำแนกประเด็นสำคัญ (Key Barriers) ที่ส่งผลต่อทัศนคติของผู้บริโภคได้ดังนี้

(1) ด้านราคา (Price Concern) เช่น ความกังวลเกี่ยวกับราคาจำหน่ายและมูลค่าการขายต่อของยานยนต์ไฟฟ้า

(2) โครงสร้างพื้นฐานการชาร์จ (Charging Infrastructure) เช่น ความไม่สะดวกในการเข้าถึงสถานีชาร์จและระยะเวลาในการชาร์จ

(3) ระยะทางการใช้งาน (Driving Range) เช่น ความกังวลเกี่ยวกับระยะทางต่อการชาร์จหนึ่งครั้ง

ผลลัพธ์ดังกล่าวสะท้อนให้เห็นถึงปัจจัยสำคัญ (Key Drivers and Barriers) ที่มีอิทธิพลต่อการยอมรับยานยนต์ไฟฟ้าในประเทศไทย และสามารถนำไปใช้เป็นข้อมูลสนับสนุนในการกำหนดนโยบายและกลยุทธ์ทางธุรกิจได้ สามารถแสดงการกระจายของข้อมูลความรู้สึกจริง (Sentiment Distribution) ดังรูปที่ 2



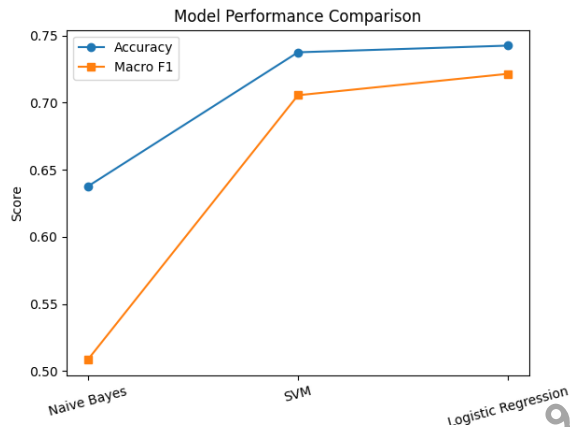
รูปที่ 2 การกระจายของข้อมูลความรู้สึกจริง (Sentiment Distribution)

จากรูปที่ 2 แบ่งออกเป็น 3 กลุ่ม ได้แก่ เชิงลบ เชิงกลาง และเชิงบวก โดยพบว่าความคิดเห็นเชิงลบมีสัดส่วนสูงกว่ากลุ่มอื่นเล็กน้อย สำหรับการวิเคราะห์การกระจายของผลการทำนาย (Prediction Distribution Analysis) มีดังนี้ 1) Naïve Bayes (NB) ผลการทดลองพบว่าโมเดล Naïve Bayes มีแนวโน้มในการทำนายค่าความรู้สึกไม่สมดุล โดยมีการกระจายตัวในบางคลาสมากกว่าคลาสอื่น แสดงให้เห็นถึงข้อจำกัดของโมเดลในการเรียนรู้บริบทของข้อความภาษาไทย 2) Support Vector Machine (SVM) โมเดล SVM มีความสามารถในการกระจายผลการทำนายได้ดีกว่า Naïve Bayes โดยสามารถจำแนกข้อมูลในแต่ละคลาสได้ชัดเจนขึ้น อย่างไรก็ตามยังพบความคลาดเคลื่อนในกลุ่ม Neutral ซึ่งมีลักษณะข้อความที่กำกวม และ 3) Logistic Regression (LR) โมเดล Logistic Regression แสดงผลลัพธ์ที่มีความสมดุลมากที่สุด โดยสามารถสะท้อน distribution ของข้อมูลจริงได้ใกล้เคียง และมีความสามารถในการแยกคลาสได้ดีในทั้งสามกลุ่ม

5.2 ผลการเปรียบเทียบโมเดล (Model Performance Comparison)

จากรูปที่ 3 กราฟแสดงการเปรียบเทียบประสิทธิภาพของโมเดล Machine Learning ได้แก่ Naïve Bayes, SVM และ Logistic Regression ในด้าน Accuracy และ Macro F1-score รวมถึงการเปรียบเทียบการกระจาย

ของผลการทำนายในแต่ละโมเดล Machine Learning ทั้ง 3 แบบ ได้แก่ Naïve Bayes, SVM และ Logistic Regression

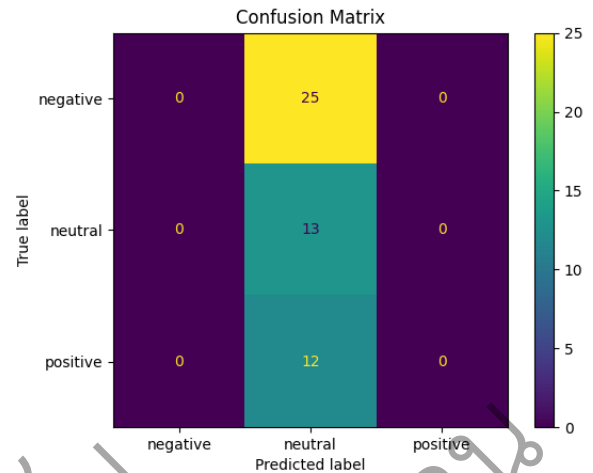


รูปที่ 3 กราฟเชิงเส้นแสดงค่าความถูกต้องโดยรวม และ F1-Score

พบว่า Logistic Regression ให้ผลลัพธ์ที่ดีที่สุดในด้าน Accuracy และ F1-score ส่วน SVM มีประสิทธิภาพรองลงมา และสามารถจัดการข้อมูลที่มีมิติสูงได้ดี และ Naïve Bayes มีความเร็วในการประมวลผลสูง แต่มีข้อจำกัดด้านความแม่นยำผลการทดลองสอดคล้องกับงานวิจัยก่อนหน้านี้ที่ระบุว่าโมเดลเชิงเส้น เช่น Logistic Regression มีประสิทธิภาพสูงในงาน text classification ที่ใช้ TF-IDF เป็น feature

5.3 ผลการเปรียบเทียบกับ Large Language Model (LLM)

จากรูปที่ 4 Confusion Matrix และ Normalized Confusion Matrix ของโมเดล Large Language Model (LLM) แบบ zero-shot



รูปที่ 4 Confusion Matrix และ Normalized Confusion Matrix ของโมเดล LLM

แสดงให้เห็นว่าโมเดลมีแนวโน้มทำนายเป็น Neutral เกือบทั้งหมด ส่งผลให้ไม่สามารถจำแนกคลาสได้อย่างมีประสิทธิภาพเมื่อเปรียบเทียบกับ LLM ซึ่งใช้วิธี zero-shot classification พบว่า LLM มีแนวโน้มทำนายเป็น Neutral ในสัดส่วนที่สูงส่งผลให้ค่า F1-score โดยรวมต่ำกว่าโมเดล Machine Learning สะท้อนว่า LLM ในรูปแบบ zero-shot ยังไม่เหมาะสมกับงานที่ต้องการความแม่นยำสูงในบริบทภาษาไทย อีกทั้งตาม LLM มีข้อดีในด้านความสามารถในการเข้าใจบริบท และสามารถพัฒนาเพิ่มเติมได้ด้วยเทคนิค few-shot learning

6. สรุปผลและอภิปราย

การวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ความรู้สึกของประชาชนต่อยานยนต์ไฟฟ้าในประเทศไทย โดยใช้ข้อมูลจากเว็บไซต์ Pantip และเปรียบเทียบประสิทธิภาพของแบบจำลอง Machine Learning ได้แก่ Naive Bayes (NB), Support Vector Machine (SVM), Logistic Regression (LR) และ Large Language Model (LLM)

ผลการทดลองพบว่าแบบจำลอง Machine Learning สามารถจำแนกความรู้สึกได้อย่างมีประสิทธิภาพ โดยเฉพาะ Logistic Regression ซึ่งให้

ผลลัพธ์ที่ดีที่สุด รองลงมาคือ SVM และ Naïve Bayes ตามลำดับ โดยมีการกระจายผลการทำนายที่สอดคล้องกับข้อมูลจริงมากกว่า ในขณะที่ LLM แบบ zero-shot มีแนวโน้มทำนายเป็น Neutral ในสัดส่วนสูง ส่งผลให้ไม่สามารถสะท้อนความแตกต่างของความคิดเห็นได้อย่างชัดเจน

ในด้านเทคนิค ผลการศึกษาชี้ให้เห็นว่าโมเดล Machine Learning ยังคงเหมาะสมสำหรับงานวิเคราะห์ข้อความภาษาไทย โดยเฉพาะเมื่อใช้ร่วมกับเทคนิค TF-IDF ซึ่งสามารถจัดการข้อมูลที่มีมิติสูงได้ดี อย่างไรก็ตาม วิธีการดังกล่าวยังมีข้อจำกัดในการสะท้อนความสัมพันธ์เชิงความหมายของคำ ดังนั้น งานวิจัยในอนาคตอาจพิจารณาการใช้เทคนิค word embedding เช่น Word2Vec หรือ FastText เพื่อเพิ่มความสามารถในการจับบริบทของภาษาไทยได้ดียิ่งขึ้น

นอกจากนี้ แม้ว่าโมเดลแบบดั้งเดิมจะมีประสิทธิภาพสูง แต่ยังมีข้อจำกัดในการทำความเข้าใจบริบทเชิงลึกของภาษา โดยเฉพาะข้อความที่มีความกำกวม งานวิจัยในอนาคตจึงอาจพิจารณาการนำโมเดลภาษาขั้นสูงที่ได้รับการฝึกฝนล่วงหน้าสำหรับภาษาไทย เช่น WangchanBERTa มาใช้ เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์ความรู้สึก ซึ่งสอดคล้องกับแนวโน้มของงานวิจัยด้าน Natural Language Processing (NLP) ในปัจจุบัน

ในเชิงการนำไปใช้ ผลการวิจัยสะท้อนให้เห็นว่าผู้บริโภคยังมีความกังวลเกี่ยวกับราคายานยนต์ไฟฟ้า โครงสร้างพื้นฐานการชาร์จ และระยะทางการใช้งาน ซึ่งข้อมูลดังกล่าวสามารถนำไปใช้สนับสนุนการกำหนดนโยบายของภาครัฐ และการพัฒนากลยุทธ์ทางธุรกิจของภาคเอกชนได้อย่างมีประสิทธิภาพ

6.1 ข้อจำกัดของงานวิจัย (Limitations)

งานวิจัยนี้มีข้อจำกัดในด้านแหล่งข้อมูล เนื่องจากใช้ข้อมูลจากแพลตฟอร์ม Pantip เพียงแหล่งเดียว ซึ่งอาจไม่สามารถสะท้อนความคิดเห็นของผู้ใช้งานจากแพลตฟอร์มสื่อสังคมออนไลน์อื่น เช่น Facebook,

Twitter หรือ YouTube ได้อย่างครอบคลุม นอกจากนี้ ข้อมูลยังอาจมีอคติจากลักษณะของผู้ใช้งานในแพลตฟอร์มดังกล่าว

6.2 แนวทางการวิจัยในอนาคต (Future Work)

งานวิจัยในอนาคตควรขยายแหล่งข้อมูลไปยังหลายแพลตฟอร์ม เพื่อเพิ่มความครอบคลุมของข้อมูล และความสามารถในการสรุปผลโดยทั่วไป (generalizability) รวมถึงเพิ่มความแข็งแกร่งของแบบจำลองในการใช้งานจริง

นอกจากนี้ ควรพัฒนาการวิเคราะห์ในระดับเชิงลึก เช่น การวิเคราะห์เชิงหัวข้อ (Topic Analysis) และการวิเคราะห์เชิงเหตุผล (Thematic Analysis) เพื่อทำความเข้าใจปัจจัยที่มีอิทธิพลต่อทัศนคติของผู้บริโภค โดยสามารถจำแนกเป็นปัจจัยผลักดัน (Drivers) และอุปสรรค (Barriers) ได้อย่างชัดเจน

อีกทั้ง การประยุกต์ใช้เทคนิคขั้นสูง เช่น deep learning และ pretrained language models จะช่วยยกระดับความแม่นยำของระบบวิเคราะห์ความรู้สึก และเพิ่มศักยภาพในการประยุกต์ใช้งานในบริบทจริงได้มากยิ่งขึ้น

7. เอกสารอ้างอิง

- [1] Lowphansirikul, L., Polpanumas, C., Jantrakulchai, N. and Nutanong, S. (2021). WangchanBERTa: Pretraining transformer-based Thai language models, *arXiv preprint arXiv:2101.09635*.
- [2] Arreerard, R., Mander, S. and Piao, S. (2022). Survey on Thai NLP Language Resources and Tools, *Proceedings of the Thirteenth Language Resources and Evaluation Conference (LREC)*, pp. 6495–6505.
- [3] Pasupa, K. and Seneewong Na Ayutthaya, T. (2022). Hybrid Deep Learning Models for Thai Sentiment Analysis, *Cognitive Computation*, Vol. 14, pp. 167–193.
- [4] Jitboonyapinit, C., Maneerat, P. and Chirawichitchai, N. (2023). Sentiment Analysis

on Thai Social Media Using Convolutional Neural Networks and Long Short-Term Memory, *International Scientific Journal of Engineering and Technology*, Vol. 7 (1), pp. 74–80.

- [5] Khamphakdee, N. and Seresangtakul, P. (2023). An Efficient Deep Learning for Thai Sentiment Analysis, *Data*, Vol. 8 (5).
- [6] Gong, B., Liu, R., Zhang, X., Chang, C. T. and Liu, Z. (2023). Sentiment analysis of online reviews for electric vehicles using the SMAA-2 method and interval type-2 fuzzy sets, *Applied Soft Computing*, Vol. 147.
- [7] Qin, Q., Zhou, Z., Zhou, J., Huang, Z., Zeng, X. and Fan, B. (2024). Sentiment and attention of the Chinese public toward electric vehicles: A big data analytics approach, *Engineering Applications of Artificial Intelligence*, Vol. 127.
- [8] Ashby, C., Weir, D. and Fussey, P. (2025). Understanding public views on electric vehicle charging: A thematic analysis, *Transportation Research Interdisciplinary Perspectives*, Vol. 29.
- [9] Ruan, T. and Lv, Q. (2024). Exploring equity perception of electric vehicles from a social media perspective, *Transportation Research Interdisciplinary Perspectives*, Vol. 25.
- [10] Wu, C., Ma, B., Zhang, Z., Deng, N., He, Y. and Xue, Y. (2025). Evaluating zero-shot multilingual aspect-based sentiment analysis with large language models, *International Journal of Machine Learning and Cybernetics*, Vol. 16 (10), pp. 8079–8101.
- [11] Chonsalasin, D., Champahom, T., and Jomnonkwao, S. (2024). Exploring the Influence of Thai Government Policy Perceptions on Electric Vehicle Adoption: A

Measurement Model and Empirical Analysis, *Smart Cities*, 7(4), pp. 2258–2282.

ประวัติผู้เขียนบทความ

ชนกานต์ กิ่งแก้ว

หัวหน้าสาขา เทคโนโลยีสารสนเทศ

คณะวิศวกรรมศาสตร์และเทคโนโลยี

สถาบันการจัดการปัญญาภิวัฒน์

งานวิจัยที่สนใจ : การวิเคราะห์ข้อมูล และ IT Software Analyst



ผู้ช่วยศาสตราจารย์ พรสิริ ชาติปรีชา

อาจารย์ประจำ เทคโนโลยีอุตสาหกรรม

คณะอุตสาหกรรมเกษตร

มหาวิทยาลัยเกษตรศาสตร์

งานวิจัยที่สนใจ: วิศวกรรมซอฟต์แวร์ และการเขียน

โปรแกรมคอมพิวเตอร์



บัณฑิต อินทรีย์มีศักดิ์

อาจารย์ประจำ คณะวิศวกรรมศาสตร์

และเทคโนโลยี สาขาวิศวกรรมการผลิต

ยานยนต์ สถาบันการจัดการปัญญาภิวัฒน์

งานวิจัยที่สนใจ : ออกแบบชิ้นส่วนยานยนต์

ซ่อมบำรุงยานยนต์สันดาปภายในและยานยนต์ไฟฟ้า

การปรับปรุงผิวมวล วัสดุผสมจากเส้นใยธรรมชาติ

