

การวิเคราะห์ความรู้สึกในรีวิวและคะแนนผลิตภัณฑ์สำหรับแอป Lazada

SENTIMENT ANALYSIS ON PRODUCT REVIEWS AND RATINGS FOR THE  
LAZADA APP



การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์  
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง  
คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง  
พ.ศ. 2567

KMITL-2024-SC-M-017-018

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SENTIMENT ANALYSIS ON PRODUCT REVIEWS AND RATINGS FOR THE  
LAZADA APP



AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE AND ANALYTICS  
KMUTL DIGITAL ANALYTICS AND INTELLIGENCE CENTER SCHOOL OF SCIENCE  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2024

KMITL-2024-SC-M-017-018

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2024**

**SCHOOL OF SCIENCE**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อการค้นคว้าอิสระ	การวิเคราะห์ความรู้สึกในรีวิวและคะแนนผลิตภัณฑ์สำหรับแอป Lazada
ชื่อนักศึกษา	วณดา ตั้งถาวรกิจ
รหัสประจำตัว	65056076
ปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการข้อมูลและการวิเคราะห์) ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
พ.ศ.	2567
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ	รองศาสตราจารย์ ดร.ณัฐไชย ลีนาวงศ์

### บทคัดย่อ

จากข้อมูลทางสถิติ กระทรวงพาณิชย์ พบว่า ในปี พ.ศ. 2564 มูลค่าการซื้อสินค้าออนไลน์เพิ่มขึ้นถึงร้อยละ 45.05 เมื่อเทียบกับปี พ.ศ.2563 อย่างไรก็ตาม การซื้อสินค้าผ่านช่องทางออนไลน์ยังพบปัญหาที่ทำให้ผู้บริโภคเกิดความกังวลในการใช้บริการ เนื่องจากการซื้อสินค้าผ่านช่องทางออนไลน์ไม่เหมือนการซื้อสินค้าผ่านหน้าร้านค้าปลีกเนื่องจากไม่ได้เห็นหรือสัมผัสสินค้าโดยตรง ซึ่งจากข้อมูลทางสถิติพบว่าผู้บริโภคตัดสินใจเลือกซื้อสินค้าได้จากคะแนนที่มีต่อสินค้า และความคิดเห็นที่ผู้บริโภครายอื่นได้ให้ไว้ก่อนหน้า โดยการแสดงความคิดเห็นของผู้บริโภคที่มีต่อสินค้านั้นมีอิทธิพลอย่างมากต่อการตัดสินใจซื้อสินค้าของผู้บริโภคอื่นเนื่องจากก่อนการซื้อสินค้า ร้อยละ 90 ของผู้บริโภคมักจะหาข้อมูลความคิดเห็นต่อสินค้าเพื่อประกอบการตัดสินใจซื้อสินค้า

ดังนั้นการศึกษานี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบความรู้สึกของความคิดเห็นกับคะแนนรีวิวหรือดาวต่อสินค้าในแพลตฟอร์ม Lazada โดยใช้วิธีการตัวแบบอิงตามกฎ (Rule-based) และทำการจัดกลุ่มคะแนนรีวิวหรือหัวใจใหม่ ผลการศึกษาพบว่าการจัดกลุ่มคะแนนรีวิวหรือหัวใจใหม่สะท้อนความคิดเห็นที่มีต่อสินค้าน่ามากยิ่งขึ้นได้ ซึ่งสามารถนำไปประกอบการตัดสินใจซื้อสินค้าผ่านแพลตฟอร์มได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น และทำให้ผู้ประกอบการร้านค้าบนแพลตฟอร์มทราบถึงกลุ่มคะแนนรีวิวที่มีต่อสินค้าและสามารถปรับปรุงสินค้าและบริการให้ดีขึ้น นอกจากนี้แพลตฟอร์มยังสามารถนำไปประยุกต์ใช้เพื่อปรับปรุงบริการ เช่นสร้างระบบการจัดกลุ่มคะแนนรีวิวให้เป็นคะแนนหัวใจที่สะท้อนความคิดเห็นที่มีต่อสินค้าน่ามากยิ่งขึ้นได้

<b>Independent Study Title</b>	Sentiment Analysis On Product Reviews And Ratings For The Lazada App
<b>Student</b>	Ms. Warunada Thangthawornkit
<b>Student ID</b>	65056076
<b>Degree</b>	Master of Science (Data Science and Analytics) KMITL Digital Analytics and Intelligence Center
<b>Year</b>	2024
<b>Independent Study Advisor</b>	Assoc. Prof. Dr. Chartchai Leenawong

### Abstract

According to statistics from the Ministry of Commerce, it was found that in the year 2021 (B.E. 2564), the value of online purchases increased by 45.05% compared to the year 2020 (B.E. 2563). However, purchasing products through online channels still presents challenges that cause consumers to hesitate in using the service. This is because online shopping differs from retail shopping as consumers cannot directly see or touch the products. From statistical data, it was found that consumers make purchase decisions based on the scores and reviews provided by other consumers. Before making a purchase, 90% of consumers typically seek product reviews to aid their decision-making process.

Therefore, the objective of this study is to compare consumer sentiments with review scores or stars for products on the Lazada platform. This is done using a rule-based model to categorize new review scores or stars. The study found that categorizing review scores or stars helps consumers understand more useful review scores, enabling them to make purchase decisions quickly and more efficiently through the platform. It also allows business owners on the platform to be aware of the review score groups related to their products, enabling them to improve their products and services. Furthermore, the platform can be utilized to enhance services, such as creating tools to automatically categorize product review scores based on consumer feedback, making them more effective.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## กิตติกรรมประกาศ

การค้นคว้าอิสระฉบับนี้ สำเร็จลุล่วงได้อย่างสมบูรณ์เนื่องจากได้รับความอนุเคราะห์และเอาใจใส่ให้คำปรึกษาเป็นอย่างดียิ่งจาก รองศาสตราจารย์ ดร.ฉัฐไชย สีนาวงศ์ อาจารย์ที่ปรึกษา ที่ได้ให้คำปรึกษา แนะนำแก้ไขปรับปรุงข้อบกพร่อง รวมทั้งให้คำแนะนำองค์ความรู้แนวทางในการศึกษาค้นคว้ามาโดยตลอด ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

สุดท้ายนี้ขอกราบขอบพระคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ซึ่งเป็นสถานศึกษาที่ให้ความรู้แก่นักศึกษาทุกคน เพื่อให้นักศึกษาได้นำความรู้ไปใช้ประโยชน์ต่อตนเองและผู้อื่น รวมทั้งได้รับความรู้เพื่อที่จะนำไปประยุกต์ใช้ในการประกอบอาชีพสืบไป

วฤนดา ตั้งถาวรกิจ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ

	หน้า
บทคัดย่อ.....	ก
Abstract.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ฉ
สารบัญรูป.....	ช
<b>บทที่ 1 บทนำ.....</b>	<b>1</b>
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	2
1.3 ขอบเขตของงานวิจัย.....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	2
<b>บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....</b>	<b>3</b>
2.1 แนวคิดที่เกี่ยวข้อง.....	3
2.1.1 แนวคิดเกี่ยวกับความรู้สึก.....	3
2.1.2 การวิเคราะห์ความรู้สึก (Sentiment Analysis).....	3
2.2 งานวิจัยที่เกี่ยวข้อง.....	6
<b>บทที่ 3 วิธีการดำเนินงานวิจัย.....</b>	<b>9</b>
3.1 ขอบเขตการศึกษา.....	9
3.2 ข้อมูลที่ใช้และการเก็บข้อมูล.....	9
3.3 กรอบแนวคิดในการศึกษา.....	12
3.4 วิธีการศึกษา.....	12
3.4.1 การเตรียมข้อมูล.....	12
3.4.2 สร้างตัวแบบ Rule-based Model.....	18
3.4.3 หาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าเชิงบวกและค่าเชิงลบ.....	19
3.4.4 จัดอันดับคะแนนรีวิวกใหม่ให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น.....	20
3.4.5 เปรียบเทียบคะแนนหัวใจใหม่กับคะแนนรีวิวดาวจริงแบบถ่วงน้ำหนัก.....	22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

บทที่ 4 ผลการวิจัยและการอภิปรายผล.....	23
4.1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าเชิงบวกและค่าเชิงลบ.....	23
4.2 ผลการจัดอันดับคะแนนรีวิว โดยการกำหนดคะแนนหัวใจใหม่.....	24
4.2.1 การกำหนดคะแนนหัวใจหลัก.....	24
4.2.2 การปรับคะแนนหัวใจเสริม.....	26
4.2.3 ผลการกำหนดคะแนนหัวใจใหม่.....	28
4.3 เปรียบเทียบประสิทธิภาพของการกำหนดคะแนนหัวใจกับคะแนนรีวิวดาวจริง.....	28
4.4 ผลการจัดกลุ่มคะแนนหัวใจใหม่.....	29
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	31
5.1 สรุปผลการวิจัย.....	31
5.2 ข้อเสนอแนะ.....	32
5.3 ข้อจำกัดในการศึกษา.....	32
เอกสารอ้างอิง.....	33
ประวัติผู้เขียน.....	34

## สารบัญตาราง

ตารางที่	หน้า
3.1 แสดงจำนวนข้อมูลแยกตามดาวหรือคะแนนที่มีต่อสินค้า.....	11
3.2 แสดงจำนวนข้อมูลแยกตามร้านค้า และคะแนนหรือดาวที่มีต่อสินค้า.....	11
4.1 ผลการกำหนดคะแนนหัวใจหลักใหม่ เทียบกับ mean positive ของแต่ละร้านค้า.....	25
4.2 ผลการปรับคะแนนหัวใจเสริม เทียบกับ mean negative ของแต่ละร้านค้า.....	27
4.3 ตารางคะแนนหัวใจใหม่ของแต่ละร้านค้า.....	28
4.4 ตารางเปรียบเทียบการกำหนดคะแนนหัวใจใหม่ กับ คะแนนดาวจริงแบบถ่วงน้ำหนัก.....	29
4.5 ตารางแสดงผลการจัดกลุ่มคะแนนหัวใจใหม่.....	30



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป

รูปที่	หน้า
2.1 การนำ SenticNet ไปใช้ในการวิเคราะห์ความรู้สึก.....	5
3.1 แสดงขั้นตอนการดึงข้อมูลจากเว็บไซต์.....	9
3.2 ตัวอย่างความคิดเห็นในรูปแบบ CSV จากการดึงข้อมูลผ่านเว็บไซต์.....	10
3.3 กรอบการศึกษาการเปรียบเทียบความรู้สึกของความคิดเห็นกับคะแนนหรือดาวต่อสินค้า.....	12
3.4 ตัวอย่างความคิดเห็นในรูปแบบของตาราง.....	13
3.5 ตัวอย่างคำฟุ่มเฟือย (Stop Word).....	13
3.6 ตัวอย่างความคิดเห็นจากการตัดคำ Stop word ในภาษาไทย.....	14
3.7 ตัวอย่างความคิดเห็นจากการกรองคำภาษาไทย.....	15
3.8 ตัวอย่างความคิดเห็นจากการตัดแบ่งคำ.....	16
3.9 ตัวอย่างความคิดเห็นจากการเตรียมข้อมูล.....	17
3.10 ตัวอย่างคำในกระเป๋าคำเชิงบวก.....	18
3.11 ตัวอย่างคำในกระเป๋าคำเชิงลบ.....	18
3.12 ตัวอย่างการปรับค่าคะแนนหัวใจหลัก.....	21
4.1 ค่าเฉลี่ย (mean) และ ส่วนเบี่ยงเบนมาตรฐาน (sd) ระดับร้านค้า.....	23
4.2 mean positive ของแต่ละร้านค้า เทียบกับ overall mean positive.....	24
4.3 การกำหนดคะแนนหัวใจหลักใหม่ของแต่ละร้านค้า.....	25
4.4 mean negative ของแต่ละร้านค้า เทียบกับ overall mean negative.....	26
4.5 การกำหนดคะแนนหัวใจเสริมของแต่ละร้านค้า.....	27

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

จากข้อมูลทางสถิติ กระทรวงพาณิชย์ พบว่า ในปี พ.ศ. 2564 มูลค่าการซื้อสินค้าออนไลน์เพิ่มขึ้นถึงร้อยละ 45.05 เมื่อเทียบกับปี พ.ศ. 2563 เป็นผลจากพฤติกรรมในช่วงสถานการณ์โควิด-19 และมาตรการสนับสนุนการใช้จ่ายของภาครัฐรวมทั้งแนวโน้มพฤติกรรมผู้บริโภคในยุค New Normal สำหรับสินค้าที่นิยมซื้อออนไลน์ 3 อันดับแรก ได้แก่ สินค้ากลุ่มเครื่องแต่งกายและเครื่องประดับ คิดเป็นร้อยละ 14.37 ของการซื้อสินค้าออนไลน์ทั้งหมด และในปี พ.ศ. 2565 จำนวนผู้ใช้งาน e-Commerce ทั่วโลกเติบโตขึ้นถึงร้อยละ 10 เมื่อเทียบกับปี พ.ศ. 2564 ส่วนหนึ่งที่ธุรกิจ e-Commerce มีการเติบโตเป็นอย่างมาก มาจากปัจจัยเกี่ยวกับโควิด-19 ซึ่งการปิดตัวของหน้าร้านค้าปลีก ทำให้พฤติกรรมกรรมการซื้อสินค้าถูกปรับเปลี่ยนมาใช้ช่องทาง e-Commerce ในการซื้อสินค้าออนไลน์เพิ่มขึ้น หากเปรียบเทียบการซื้อสินค้าผ่านทางออนไลน์กับผ่านทางหน้าร้านค้าปลีกแล้วนั้น การซื้อสินค้าออนไลน์มีความสะดวกและประหยัดเวลาได้มากกว่าการเดินทางไปซื้อสินค้าผ่านหน้าร้านค้าปลีก และการซื้อสินค้าผ่านแพลตฟอร์ม e-Commerce ส่วนใหญ่จะมีโปรโมชั่นบ่อยครั้งมากกว่าการซื้อผ่านหน้าร้านค้าปลีก อีกทั้งยังสามารถเปรียบเทียบคุณภาพและราคาสินค้าได้อย่างรวดเร็วผ่านข้อมูลสินค้าและรีวิวที่สินค้านั้นได้รับ และจากบทความของธนาคารแห่งประเทศไทย พบว่า ในปี พ.ศ. 2563 ที่ผ่านมามี Shopee และ Lazada คือ e-Marketplace ที่ถูกค้นหามากที่สุด คิดเป็นร้อยละ 90 จากการค้นหา e-Marketplace ทั้งหมดในไทย

อย่างไรก็ตามการซื้อสินค้าผ่านช่องทางออนไลน์ยังพบปัญหาที่ทำให้ผู้บริโภคเกิดความกังวลในการใช้บริการ เนื่องจากการซื้อสินค้าผ่านช่องทางออนไลน์ไม่เหมือนการซื้อสินค้าผ่านหน้าร้านค้าปลีก เนื่องจากไม่ได้เห็นหรือสัมผัสสินค้าโดยตรง ผู้บริโภคจะตัดสินใจซื้อสินค้าได้จากคุณสมบัติของสินค้าที่ทางผู้ประกอบการเป็นคนให้รายละเอียดไว้ และนอกจากคุณสมบัติของสินค้าแล้ว ยังตัดสินใจเลือกซื้อสินค้าได้จากคะแนนที่มีต่อสินค้าและความคิดเห็นที่ผู้บริโภครายอื่นได้ให้ไว้ก่อนหน้า

โดยการแสดงความคิดเห็นของผู้บริโภคที่มีต่อสินค้านั้น มีอิทธิพลอย่างมากต่อการตัดสินใจซื้อสินค้าของผู้บริโภคอื่น เนื่องจากก่อนการซื้อสินค้านั้น ร้อยละ 90 ของผู้บริโภคมักจะหาข้อมูลความคิดเห็นต่อสินค้าเพื่อประกอบการตัดสินใจซื้อสินค้า (Invesp, 2560)

เมื่อมีผู้บริโภคแสดงความคิดเห็นต่อสินค้าจำนวนมากและหลากหลายทิศทาง ทำให้การตัดสินใจซื้อสินค้าต้องใช้เวลาอันยาวนานมากขึ้น ถึงแม้ปัจจุบันแพลตฟอร์มออนไลน์ เช่น SHOPEE และ LAZADA จะมีการให้คะแนนรีวิวสินค้าเป็นจำนวนดาว แต่ในหลายครั้งคะแนนรีวิวสินค้าที่ได้รับสวนทางกับความคิดเห็นต่อสินค้า ผู้วิจัยคาดว่าหากมีตัวช่วยในการวิเคราะห์ความรู้สึกของความคิดเห็นต่อสินค้า และนำมาใช้เป็นเครื่องมือในการประเมินคะแนนรีวิวให้เป็นระดับคะแนนใหม่

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ การใช้งานหรือการเผยแพร่โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่แตกต่างจากดาวที่คิดจากแพลตฟอร์ม จะช่วยให้ผู้บริโภคตัดสินใจเลือกซื้อสินค้าได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น

ดังนั้น การศึกษาครั้งนี้ ผู้วิจัยจึงได้จัดทำ (1) การวิเคราะห์ความรู้สึกของความคิดเห็นต่อสินค้า ซึ่งเป็นประโยชน์สำหรับผู้บริโภคและผู้ประกอบการธุรกิจบนแพลตฟอร์ม e-Commerce ทางด้าน ผู้บริโภคสามารถนำไปประกอบการตัดสินใจซื้อสินค้าผ่านแพลตฟอร์มได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น ส่วนทางด้านผู้ประกอบการสามารถนำไปปรับปรุงสินค้าและบริการให้ดียิ่งขึ้น และ (2) เปรียบเทียบความรู้สึกของความคิดเห็นต่อคะแนนรีวิวว่ามีความสอดคล้องกันหรือไม่ เพื่อปรับปรุง การให้บริการแพลตฟอร์มสำหรับการแสดงความคิดเห็นต่อคะแนนรีวิวสินค้าให้มีประสิทธิภาพมากขึ้น (3) สร้างระบบการจัดกลุ่มคะแนนรีวิวให้เป็นคะแนนหัวใจที่สะท้อนความคิดเห็นที่มีต่อสินค้า

## 1.2 วัตถุประสงค์ของงานวิจัย

- 1) เปรียบเทียบความรู้สึกของความคิดเห็นต่อคะแนนรีวิวที่ผลิตภัณฑ์นั้นได้รับ
- 2) สร้างระบบการจัดอันดับคะแนนรีวิวใหม่ให้เป็นคะแนนหัวใจที่สะท้อนความคิดเห็นที่มีต่อสินค้า

## 1.3 ขอบเขตของงานวิจัย

- 1) วิเคราะห์ความรู้สึกของความคิดเห็นและคะแนนรีวิวสำหรับแอปพลิเคชัน LAZADA
- 2) วิเคราะห์การแสดงความคิดเห็นที่เป็นภาษาไทยเท่านั้น
- 3) ประเภทของสินค้าที่นำมาวิเคราะห์จะเลือกประเภทสินค้าที่เป็นที่นิยมสำหรับการซื้อออนไลน์ นั่นคือสินค้าหมวดหมู่เสื้อผ้าสตรี
- 4) การเลือกข้อมูลมาวิเคราะห์จะเลือกจากการสุ่มร้านค้าหมวดหมู่เสื้อผ้าสตรี 10 ร้านค้า และเรียงลำดับรายการสินค้าในร้านค้านั้น ที่มีจำนวนความคิดเห็นมากกว่า 100 รายการขึ้นไป

## 1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1) เพื่อปรับปรุงการให้บริการแพลตฟอร์มสำหรับการแสดงความคิดเห็นต่อคะแนนรีวิวสินค้าให้มี ประสิทธิภาพมากขึ้น โดยอาจพัฒนาระบบการจัดอันดับคะแนนรีวิวใหม่ให้เป็นคะแนนหัวใจที่ สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น
- 2) เพื่อปรับปรุงการจัดกลุ่มคะแนนรีวิวหรือดาวให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น
- 3) เพื่อให้ผู้บริโภคได้รับข้อมูลที่มีประโยชน์มากขึ้นซึ่งสามารถนำไปประกอบการตัดสินใจซื้อ สินค้าผ่านแพลตฟอร์มได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น
- 4) เพื่อให้ผู้ประกอบการร้านค้าบนแพลตฟอร์มสามารถปรับปรุงสินค้าและบริการให้ดียิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

# ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยครั้งนี้ผู้วิจัยได้ศึกษาทฤษฎีแนวคิดและงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการดำเนินงานวิจัย โดยมีเนื้อหาที่รวบรวมไว้เพื่อสนับสนุนงานวิจัยดังต่อไปนี้

### 2.1 แนวคิดที่เกี่ยวข้อง

#### 2.1.1 แนวคิดเกี่ยวกับความรู้สึก

การใช้ภาษาโดยทั่วไปนั้นไม่สามารถแสดงความรู้สึกได้ในทุกๆข้อความ เนื่องจากมนุษย์มักใช้ภาษาเพื่อรายงานข้อมูลเท่านั้น แต่การแสดงความรู้สึกจากข้อความเป็นสิ่งสำคัญในการวิเคราะห์ข้อมูลและทำนายพฤติกรรมของบุคคลหรือกลุ่มบุคคลที่เขียนข้อความนั้น การวิเคราะห์ความรู้สึกจึงถูกแบ่งเป็นงานวิจัยแขนงย่อยต่าง ๆ เริ่มจากการจำแนกข้อความแสดงอัตวิสัย (Subjectivity Classification) เป็นการแบ่งประเภทของข้อความออกเป็นข้อความที่แสดงอัตวิสัย (Subjective) และ ข้อความที่บอกข้อเท็จจริง (Objective) จากนั้นจึงนำข้อความแสดงอัตวิสัยมาวิเคราะห์ความรู้สึกต่อไป การวิเคราะห์ความรู้สึกของข้อความ สามารถทำได้ 2 แนวทาง คือ การวิเคราะห์แบบอิงคลังศัพท์ (Lexicon-Based) และการวิเคราะห์โดยอาศัยการเรียนรู้ด้วยเครื่อง (Machine Learning-Based) การวิเคราะห์ความรู้สึกแบบอิงคลังศัพท์อาศัยการระบุค่าที่บ่งบอกความรู้สึกในประโยคโดยอ้างอิงจากคลังศัพท์บอกความรู้สึก ซึ่งรวบรวมค่าบอกความรู้สึกชั่วคราวและชั่วคราวทั้งหมดเอาไว้ ซึ่ง การวิเคราะห์แบบอิงคลังศัพท์ (Lexicon-Based) มีข้อเสียคืออาจได้คลังศัพท์ที่ไม่ครบถ้วน ขึ้นอยู่กับธรรมชาติของภาษาที่พบในคลังข้อมูลนั้นๆ ซึ่งงานวิจัยในภาษาไทยที่พบขั้นตอนการรวบรวมค่าบอกความรู้สึกเพื่อสร้างคลังศัพท์ ได้แก่ การทดลองรวบรวมค่าบอกความรู้สึกและค่าบอกลักษณะสินค้าในบทวิจารณ์โรงแรมภาษาไทยจากเว็บ

#### 2.1.2 การวิเคราะห์ความรู้สึก (Sentiment Analysis)

การวิเคราะห์ความรู้สึกคือ กระบวนการเชิงคำนวณเพื่อกำกับข้อความหรือส่วนของข้อความที่แสดงความคิดเห็นด้วยทัศนคติของผู้เขียนข้อความนั้นๆ ซึ่งมักจะเป็นทัศนคติเชิงบวก เชิงลบ หรือเป็นกลาง เพื่อบ่งบอกความรู้สึกของผู้เขียนที่มีต่อบางสิ่งบางอย่าง ซึ่งการวิเคราะห์ความรู้สึกจะเกี่ยวข้องกับข้อความหรือประโยคที่แสดงถึงความรู้สึก ลักษณะของความรู้สึกว่า “ดี” หรือ “ไม่ดี” และความสัมพันธ์ของกัน ระหว่างความรู้สึกกับส่วนที่เป็นประธานของประโยค ซึ่งทั้งหมดจะมีความสัมพันธ์เกี่ยวข้องกัน ระหว่างความรู้สึกกับส่วนที่เป็นประธานของประโยค ซึ่งทั้งหมดจะมีความสัมพันธ์เกี่ยวข้องกัน ตัวอย่างของประโยค เช่น “สินค้า ก ดีกว่า สินค้า ข” จะเห็นได้ว่าคำว่า ดีกว่า เป็นตัวบ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บอกถึงความรู้สึกเชิงบวก (ดี) ที่มีต่อสินค้า ก และยังส่งผลบ่งบอกถึงความรู้สึกเชิงลบ (ไม่ดี) ไปยังสินค้า ข ด้วย (นริศร์, 2550) (Tetsuya and Jeonghee, 2003)

โดยการวิเคราะห์ความรู้สึกในประโยคจะใช้หลักการของการประมวลผลภาษาธรรมชาติ (NLP) ด้วยการทำให้เป็นรูปประโยคอย่างง่ายและวิเคราะห์จากชนิดของคำ (Part of Speech) โดยวิเคราะห์ในส่วน of คำที่อยู่ด้วยกันเป็นรูปแบบของประโยค โดยใช้กระบวนการแยกแยะข้อมูล (Information Extraction) เพื่อวิเคราะห์ส่วนที่แสดงถึงความสัมพันธ์ที่เกี่ยวข้องกัน

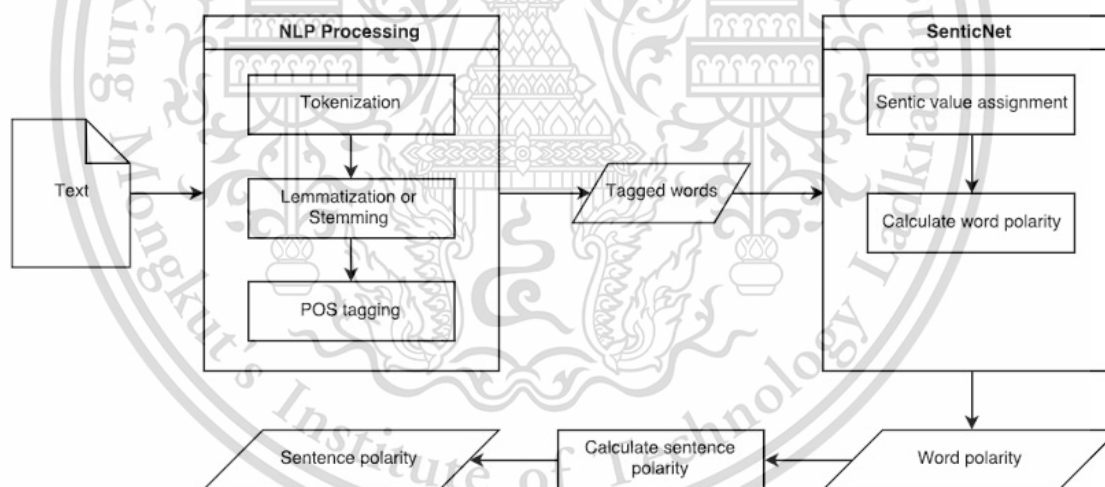
การวิเคราะห์หาความสัมพันธ์กันขึ้นอยู่กับขอบเขตที่ระบุไว้ด้วย ตัวอย่างเช่นประโยคที่มีประธานและส่วนที่แสดงความรู้สึกนั้นอยู่ในประโยคเดียวกันและยังอยู่ใกล้กัน แต่มีความรู้สึกที่แสดงออกมาทั้งสองส่วนกลับขัดแย้งกัน ตัวอย่างเช่น “ฉันชอบรองเท้าคู่นี้ ใส่สบาย แต่สีไม่สวย” จะเห็นได้ว่าภายในประโยคมีสองส่วนที่ขัดแย้งกัน คือ “ชอบรองเท้า เพราะ ใส่สบาย” แต่กลับไม่ชอบ “สีของรองเท้า” ซึ่งทำให้เป็นต้นเหตุสำคัญที่ทำให้การวิเคราะห์ความสัมพันธ์กันระหว่างประธานและส่วนที่บอกถึงความรู้สึกได้ผลลัพธ์ออกมาไม่ดีเท่าที่ควร ซึ่งอาจเกิดจากโปรแกรมคอมพิวเตอร์ที่ใช้ในการประมวลผลนั่นเอง โดยโปรแกรมส่วนใหญ่จะเน้นไปที่การพยายามวิเคราะห์ค่าของความพึงพอใจที่เป็นเชิงบวกหรือเป็นเชิงลบที่มีต่อสิ่งนั้นออกมาทั้งหมดในคราวเดียว ไม่ว่าจะสามารถวิเคราะห์ได้อย่างชัดเจนหรือไม่ชัดเจนก็ตาม ตัวอย่างเช่นการวิเคราะห์การรีวิวร้านอาหารว่าร้านนี้ดีหรือไม่ดี คืออาจจะให้ค่าความรู้สึกที่เป็นลบสำหรับพนักงานร้าน เช่น “พนักงานร้านนี้บริการได้แย่มาก” แต่ว่าในส่วน of อาหารทั้งหมดกลับให้ค่าเป็นบวก เช่น “อาหารร้านนี้ถือว่าอร่อยมาก” เป็นต้น ทำให้เกิดความสับสนในการตัดสินใจได้ ในทางตรงกันข้าม ถ้าใช้การวิเคราะห์โดยใช้ความสัมพันธ์ระหว่างประธานและส่วนที่บอกถึงความรู้สึก สามารถวิเคราะห์ให้ละเอียดลงไปกว่านั้นได้ว่าอะไรที่ดีและอะไรที่ไม่ดี

การกำหนดรูปแบบที่แสดงถึงความพึงพอใจนอกจากคำที่ขยายคำนามแล้ว คำอื่น ๆ ก็สามารถนำมาใช้เพื่อบอกถึงความรู้สึกชอบหรือไม่ชอบได้ เช่น คำนาม คำกริยา และคำขยายที่บอกถึงความรู้สึก เช่นคำว่า “ดี” เป็นคำที่บอกความรู้สึกที่มีต่อคำนาม เช่น “สินค้าดี” และในประโยคของคำนามทั้งหมดจะเป็นการบอกถึงชนิดของความรู้สึกที่มีค่าเดียวกับคำขยาย (“ดี” ในกรณีนี้มีค่าเป็นบวก) ในทำนองเดียวกันการบอกความพึงพอใจโดยใช้คำขยายกริยา เช่น คำว่า “มาก” เป็นตัวที่บอกถึงความพึงพอใจของคำกริยา เช่น “ทำงานเร็วมาก” โดยชนิดของความพึงพอใจจะได้มาจากคำกริยา

ดังนั้นการบ่งบอกถึงความพึงพอใจโดยใช้คำขยายคำนาม คำขยายกริยา และคำนาม สามารถบอกได้ว่าความพึงพอใจเป็นเชิงบวกหรือเป็นเชิงลบ ในทางกลับกันดังตัวอย่างที่ผ่านมา เช่น “สินค้า ก ดีกว่า สินค้า ข” ชนิดของความพึงพอใจที่ได้ออกมาจากคำกริยา ก จะต้องขึ้นอยู่กับความสัมพันธ์กันด้วย ในกรณีนี้ค่าความพึงพอใจที่เป็นเชิงบวกจะชี้ไปที่ประธาน (สินค้า ก) และค่าความพึงพอใจที่เป็นเชิงลบจะชี้ไปที่ กรรม (สินค้า ข) ยิ่งไปกว่านั้นคำกริยาบางคำไม่สามารถที่จะบอกความพึงพอใจ

ได้ด้วยตัวเอง แต่เป็นเพียงตัวที่ส่งผ่านค่าความพึงพอใจไปยังส่วนอื่นที่เกี่ยวข้อง เช่น คำว่า “มี” จะเป็นตัวส่งผ่านความพึงพอใจไปยังประธาน ตัวอย่างเช่น “สินค้ามีมาก” ซึ่งคำว่า “มาก” มีค่าความพึงพอใจเป็นเชิงบวก ซึ่งจะส่งผลไปถึงประธานทำให้รูปประโยคมีค่าความพึงพอใจเป็นเชิงบวกไปด้วย (นริศร์, 2550)

อย่างไรก็ตามในการกำกับข้อความระดับคำสามารถใช้ทรัพยากร SenticNet ในการแทนค่าอารมณ์ Sentic Value 4 มิติ คือ ความพึงพอใจ (Pleasantness) ความสนใจ (Attention) ความอ่อนไหว (Sensitivity) และ ความฉลาด (Aptitude) และเมื่อต้องการวิเคราะห์ความรู้สึกต้องใช้ SenticNet ในการวิเคราะห์ SenticNet จะต้องนำข้อความที่ต้องการวิเคราะห์ผ่านกระบวนการ NLP (Natural Language Processing) โดยต้องทำ Tokenization เพื่อแยกข้อความออกเป็นคำ ซึ่งในภาษาไทยต้องมีการตัดคำ (Word Segmentation) หลังจากนั้นนำมาทำ Lemmatization หรือ Stemming ซึ่งโดยปกติกระบวนการ NLP จะเลือกใช้ Stemming เพราะสามารถทำงานได้รวดเร็วกว่า จากนั้นนำมากำกับชนิดของคำจากข้อความ (Part Of Speech Tagging) อาทิเช่น คำนาม คำกริยา เป็นต้น หลังจากนั้นนำคำที่ผ่านกระบวนการ NLP แล้วมาหาค่าความรู้สึกและขั้วอารมณ์จากกระบวนการทำ Senticnet ต่อไป (สันติ,2561) ดังรูปที่ 1.1



รูปที่ 2.1 การนำ SenticNet ไปใช้ในการวิเคราะห์ความรู้สึก

ที่มา : แนวคิดในการสร้างและการใช้ประโยชน์ของ Sentiment Corpus

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ ธนัญชัย เพ็งพรหม และ วรารัตน์ สงฆ์แป้น (2562) ศึกษาการทำแบบจำลองแบบผสมเพื่อสร้างกฎ ถาม-ตอบ โดยใช้เทคนิคการทำเหมืองข้อความ พบว่าด้วยเทคนิคการหาเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor) และผสมกับแบบสร้างกฎความสัมพันธ์โดยใช้อัลกอริทึมแอฟริอริ ได้ผลความถูกต้องโดยเฉลี่ยของกฎถึงร้อยละ 91.43 และแบบจำลองสามารถตอบคำถามได้ถูกต้องเป็นที่น่าพอใจ

งานวิจัยของ นิภาพร เทศศรีเมือง (2559) ศึกษาการสรุปความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์เครื่องสำอาง โดยมีขั้นตอนได้แก่การใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) ในการจัดกลุ่มเพื่อจำแนกประเภทของเครื่องสำอาง จากนั้นสร้างกลุ่มคำที่มีความคิดเห็นเชิงบวกและลบเพื่อใช้วิเคราะห์ความรู้สึกของประโยค และหาความคล้ายคลึงของเอกสาร (Cosine similarity) จากนั้นใช้อัลกอริทึม PageRank, อัลกอริทึม Hopfield networks, และการหาความถี่ของเอกสาร (Term Frequency Inverted Document Frequency) เพื่อหาประโยคที่มีความสำคัญที่สุด พบว่าการวัดประสิทธิภาพของโปรแกรมโดยใช้ ROUGE มีความแม่นยำสูงที่สุดที่ 90.462% โดยเป็นผลลัพธ์ที่ได้จากอัลกอริทึม Hopfield networks

งานวิจัยของ พนิดา กาศกลางดอน (2563) ศึกษาการวิเคราะห์ประสบการณ์จากการใช้บริการโรงพยาบาลในประเทศไทยจากความคิดเห็นของผู้ใช้บริการ โดยใช้การทดสอบการจำลองโดยใช้เทคนิคอัลกอริทึม 3 เทคนิค คือ นาร์อิวเบย์ (Naive Bayes), การสุ่มป่าไม้ (Random Forest) และ โครงข่ายประสาทเทียมแบบหน่วยความจำระยะสั้นแบบยาว (Long-Short-Term Memory) จากผลการทดสอบพบว่าอัลกอริทึม Random Forest มีประสิทธิภาพสูงสุดในการทำนายความคิดเห็นเชิงลบโดยมีค่าการเรียกคืน (Recall) ที่ 0.71

งานวิจัยของ พิศิษฐ์ บวรเลิศสุธี และ วรภัทร ไพรีเกรง (2565) ศึกษาตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์โดยใช้เทคนิค Machine Learning ร่วมกับเทคนิคการตัดคำ (Word Tokenization) และการสร้างคลังคำศัพท์ (Bag of Words) จากนั้นนำเข้ากระบวนการจำแนกประเภทผลการวิเคราะห์ 4 เทคนิค ได้แก่ LSTM, SGD, Logistic Regression และ Support Vector Machines จากผลการทดลองสร้างตัวแบบการวิเคราะห์ระดับความรู้สึกทางอารมณ์ 3 ระดับ คือความรู้สึกเชิงบวก (Positive) เป็นกลาง (Neutral) และเชิงลบ (Negative) โดยแต่ละตัวแบบให้ค่าความถูกต้องในการทำนายผล ดังนี้ LSTM 81.27%, Logistic Regression 69%, SGD 66%, และ Support Vector Machines 65% สรุปได้ว่าการสร้างตัวแบบโดยใช้เทคนิคการจำแนกประเภทแบบ LSTM ให้คะแนนความถูกต้องสูงที่สุด จึงเหมาะสมแก่การนำไปใช้สร้างตัวแบบในการวิเคราะห์ระดับความรู้สึกกับข้อความที่เป็นภาษาไทยโดยพิจารณาจากค่า F1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิจัยของ วิชาญดา เทศเมือง และ นิเวศ จิระวิจิตรชัย (2560) ศึกษาการวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์โดยใช้ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมทซิง งานวิจัยนี้ได้นำเสนอวิธีการจำแนกความคิดเห็นโดยการวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์ด้านการบริการห้องพักโรงแรมรีสอร์ตจาก Agoda Thailand และ Twitter Thailand ที่จดทะเบียนหลักทรัพย์ โดยใช้เทคนิคเหมือนข้อความวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์และสร้างแบบจำลองด้วยอัลกอริทึม 4 วิธีได้แก่ ซัพพอร์ตเวกเตอร์แมทซิง ต้นไม้ตัดสินใจ นาอีฟเบย์ และเคเนียร์สเนเบอร์เพื่อเปรียบเทียบประสิทธิภาพ จากการทดลองพบว่าคุณลักษณะที่ดีที่สุดคือซัพพอร์ตเวกเตอร์แมทซิง รองลงมาเป็น นาอีฟเบย์ , ต้นไม้ตัดสินใจ และเคเนียร์สเนเบอร์ ตามลำดับ

งานวิจัยของ วนัสวรรณ มีประเสริฐ และ เอกรัฐ รัชฎาญจน์ (2564) ศึกษาการวิเคราะห์ความคิดเห็นจากทวีตเตอร์ของลูกค้าบริษัทซ้อปปีประเทศไทย โดยแยกแต่ละประเด็นสำคัญ ซึ่งแบ่งออกเป็น 5 ประเด็นคือ (1) บริษัท (2) ขนส่ง (3) โปรโมชั่น (4) ร้านค้า (5) ระบบ โดยผู้วิจัยได้สร้าง Model สำหรับการแยกหัวข้อประเด็นสำคัญคือ Random Forest และใช้ Model ที่มีผู้พัฒนาไว้แล้วในการแยกความรู้สึกแต่ละประเด็นสำคัญคือ WangchanBERTa ซึ่งให้ประสิทธิภาพที่ดีที่สุดสำหรับข้อความภาษาไทยและจากการหาภาพรวมของการวิเคราะห์ความรู้สึกแต่ละประเด็นสำคัญของลูกค้าบริษัทซ้อปปีสิ่งที่พบในข้อมูลเชิงลึกของลูกค้าคือทางบริษัทซ้อปปีควรให้ความสำคัญกับการปรับปรุงและพัฒนากลยุทธ์ในเรื่องของขนส่ง (Logistics) และระบบ (System) เป็นอันดับแรกเพื่อสร้างความพึงพอใจให้แก่ลูกค้าและเป็นผู้นำในด้านการให้บริการ E-Commerce

งานวิจัยของ สมัคร ชัยสงวน (2561) ศึกษาการพัฒนากระบวนการวิเคราะห์ความรู้สึกแบบเรียลไทม์ของนักศึกษานานาชาติโดยใช้ตัวจำแนกข้อมูลนาอีฟเบย์ (NaNe Bayes Classifier) สำหรับภาษาไทยพบว่าจากหลักการพื้นฐานของการประมวลผลภาษาธรรมชาติด้วยวิธีการตัดคำภาษาไทยและประยุกต์ตัวจำแนกข้อมูลนาอีฟเบย์ โดยผลลัพธ์ทั้งหมดสามารถแสดงผลแบบเรียลไทม์กับระบบจัดการคำคลังคำศัพท์สรุปผลการวิเคราะห์ความรู้สึก เครื่องมือทดสอบประโยคและระบบแจ้งเตือนข้อความผ่านทางไลน์ โดยได้ผลลัพธ์ค่าความแม่นยำร้อยละ 80.03 ค่าเรียกคืนร้อยละ 90.52 ค่าความถูกต้องร้อยละ 89.63 และการวัดประสิทธิภาพโดยรวมร้อยละ 84.81 ตามลำดับ

งานวิจัยของ อธิศักดิ์ ศรีดำ (2565) ศึกษาการจัดหมวดหมู่ข้อความความคิดเห็นและข้อเสนอแนะของประชาชนที่มีต่อโครงการของรัฐโดยวิธีปัญญาประดิษฐ์ โดยใช้การหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความถูกต้อง (Accuracy) และค่าเอฟเมเชอร์ (F-measure) เป็นตัวชี้วัดพบว่าวิธีปัญญาประดิษฐ์ด้วยการเรียนรู้ของเครื่องด้วยกรรมวิธีต้นไม้ตัดสินใจ (Decision Tree) สำหรับการจำแนกหมวดหมู่ตามคุณลักษณะของข้อมูลทั้ง 3 ลักษณะ ได้แก่หมวดหมู่ข้อความความคิดเห็นและข้อเสนอแนะประเภทโครงการ หน่วยงานหรือจังหวัดที่รับผิดชอบมีความเหมาะสมต่อการนำไปใช้งานระดับสูง โดยการจำแนกคุณลักษณะข้อความความคิดเห็นและข้อเสนอแนะมีค่าความแม่นยำเท่ากับ 0.88 ค่าความระลึกเท่ากับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

0.86 ค่าความถูกต้องเท่ากับ 86.69 และค่าเอฟเมเชอร์เท่ากับ 0.87 ส่วนการจำแนกคุณลักษณะด้านประเภทโครงการพบว่าค่าความแม่นยำเท่ากับ 0.87 ค่าความระลึกลับเท่ากับ 0.85 ค่าความถูกต้องเท่ากับ 85.23 และค่าเอฟเมเชอร์เท่ากับ 0.85 การจำแนกคุณลักษณะด้านหน่วยงานหรือจังหวัดที่รับผิดชอบมีค่าความแม่นยำเท่ากับ 0.86 ค่าความระลึกลับเท่ากับ 0.83 ค่าความถูกต้องเท่ากับ 83.94 และค่าเอฟเมเชอร์เท่ากับ 0.84



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 3

### วิธีการดำเนินงานวิจัย

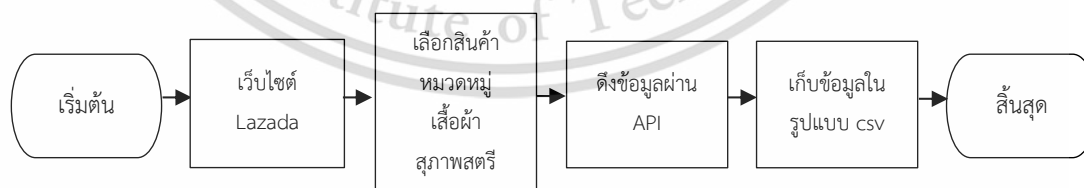
การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อทำการเปรียบเทียบความรู้สึกของความคิดเห็นกับคะแนนรีวิวหรือดาวต่อสินค้าในแพลตฟอร์ม Lazada โดยใช้วิธีการตัวแบบอิงตามกฎ และทำการจัดกลุ่มคะแนนรีวิวหรือหัวใจใหม่ โดยมีขั้นตอนในการศึกษาและดำเนินการศึกษา ดังนี้

#### 3.1 ขอบเขตการศึกษา

- 1) แพลตฟอร์มที่ใช้ในการศึกษาความรู้สึกของความคิดเห็นและคะแนนคือ LAZADA
- 2) เลือกวิเคราะห์ความคิดเห็นที่เป็นภาษาไทยเท่านั้น
- 3) ประเภทของสินค้าที่นำมาวิเคราะห์จะเลือกประเภทสินค้าที่เป็นที่นิยมสำหรับการซื้อออนไลน์ นั่นคือ สินค้าหมวดหมู่ เสื้อผ้าสุภาพสตรี
- 4) การเลือกข้อมูลมาวิเคราะห์จะเลือกจากการสุ่มร้านค้าหมวดหมู่เสื้อผ้าสุภาพสตรี 10 ร้านค้า และเรียงลำดับรายการสินค้าในร้านค้า ที่มีจำนวนความคิดเห็นมากกว่า 100 รายการขึ้นไป

#### 3.2 ข้อมูลที่ใช้และการเก็บข้อมูล

งานวิจัยนี้ทำการเก็บรวบรวมข้อมูลจากแพลตฟอร์ม Lazada โดยทำการสุ่มเก็บข้อมูลการแสดงความคิดเห็นจากสินค้าหมวดหมู่เสื้อผ้าสุภาพสตรี 10 ร้านค้า และเก็บข้อมูลจากการเรียงลำดับรายการสินค้าในร้านค้านั้น ที่มีจำนวนความคิดเห็นมากกว่า 100 รายการขึ้นไป ทำการเก็บข้อมูลในเดือนกุมภาพันธ์ พ.ศ. 2567 โดยใช้วิธีดึงข้อมูลผ่านเว็บไซต์ด้วยภาษา Python ผ่าน API ของแพลตฟอร์ม Lazada และเก็บข้อมูลจากการดึงผ่าน API ในรูปแบบของ CSV ดังรูปที่ 3.1



รูปที่ 3.1 แสดงขั้นตอนการดึงข้อมูลจากเว็บไซต์

สามารถเก็บข้อมูลได้ทั้งหมด 14,877 ความคิดเห็น โดยข้อมูลที่ได้อยู่ในรูปแบบของ CSV ดังรูปที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ข้อมูลความคิดเห็นทั้งหมด 14,877 สามารถแยกตามคะแนนหรือดาวที่มีต่อสินค้าได้ ดังตารางที่ 3.1

ตารางที่ 3.1 แสดงจำนวนข้อมูลแยกตามดาวหรือคะแนนที่มีต่อสินค้า

ดาว	จำนวนข้อมูล
5	14,506
4	248
3	76
2	23
1	24

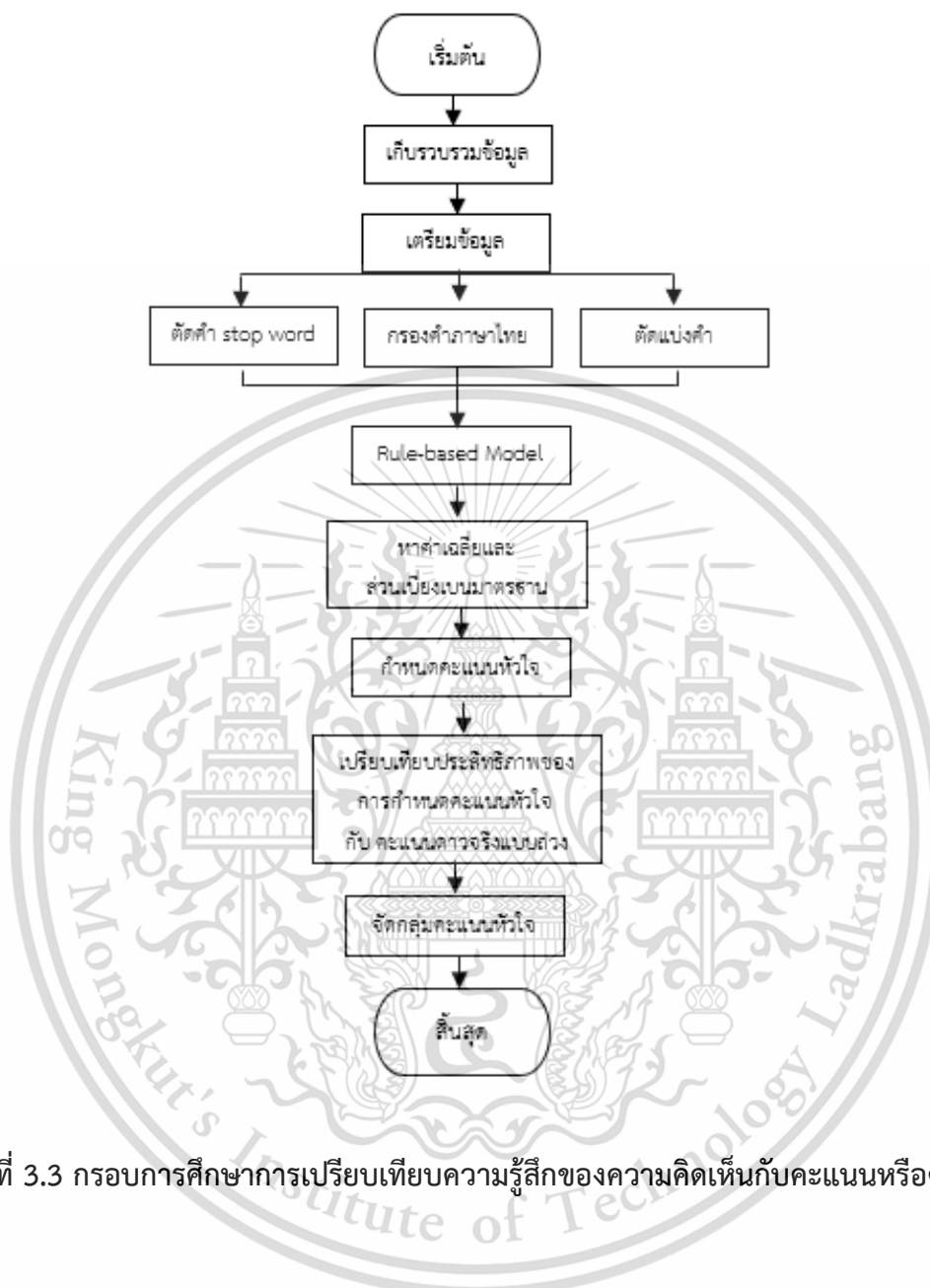
ข้อมูลความคิดเห็นทั้งหมด 14,877 สามารถแยกตามร้านค้า และคะแนนหรือดาวที่มีต่อสินค้าได้ ดังตารางที่ 3.2

ตารางที่ 3.2 แสดงจำนวนข้อมูลแยกตามร้านค้า และคะแนนหรือดาวที่มีต่อสินค้า

ร้านค้า	คะแนนดาว					รวมทั้งหมด
	1	2	3	4	5	
Acaseyshop	6	10	17	24	2,110	2,167
BEMING.BKK		1	4	14	710	729
BabarryJeans	1	1	11	58	3,403	3,474
CHIC.th	11	3	9	22	1,161	1,206
Esolo		1	3	18	854	876
Giordano		1	4	33	1,877	1,915
Nessaj	3	2	9	22	693	729
SPINNYHOUSE	3	3	11	23	1,335	1,375
Yuedpao				2	913	915
wipcloset		1	8	32	1,450	1,491
รวมทั้งรวม	24	23	76	248	14,506	14,877

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3 กรอบแนวคิดในการศึกษา



รูปที่ 3.3 กรอบการศึกษาการเปรียบเทียบความรู้สึกของความคิดเห็นกับคะแนนหรือดาวต่อสินค้า

### 3.4 วิธีการศึกษา

#### 3.4.1 การเตรียมข้อมูล

จากการดึงข้อมูลความคิดเห็นผ่านเว็บไซต์ (Web Scraping) ซึ่งได้ข้อมูลมาในรูปแบบของ CSV โดยผู้วิจัยจึงจัดรูปแบบข้อมูลใหม่ให้อยู่ในรูปแบบของตาราง เพื่อความสะดวกต่อการนำไปประมวลผลข้อความต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





### 3.4.1.2 กรองคำภาษาไทย

ทำการตัดสัญลักษณ์แสดงความรู้สึกหรืออีโมจิ (Emoji) ในความคิดเห็น และทำการตัดตัวเลข และอักขรพิเศษออก ดังนี้

Comment	Token Thai Only
เป็นกระโปรงทรงA แบบที่อยากได้ ผ้าเย็นส์สวย ใส่แล้วดูเพรียวดีค่ะ ทางร้านจัดส่งรวดเร็วดีค่ะ ส่วนไซส์ของเราดูตามตารางไซส์แล้ว +1ค่ะ กระโปรงจริงไซส์3xl เอว 34 ค่ะ สะโพก 42 ใส่พอดีเป๊ะ!	กระโปรงทรง แบบที่อยากได้ ผ้าเย็นส์สวย ใส่ดูเพรียวดี ทางร้านจัดส่งรวดเร็วดี ส่วนไซส์ดูตามตารางไซส์แล้ว กระโปรงจริงไซส์ เอว สะโพก ใส่พอดี
เนื้อผ้าเย็นส์ดีมากเลย ตัดเย็บดี ทรงสวย สีสวย ใส่แล้วเดินสะดวกสบาย ขนาดมาตรฐานอยู่แล้วแต่เหมือนสั่งไซส์เอวหลวมไปนิดนึงเอง แต่ไม่เป็นปัญหา เมื่อเราอ้วนได้อยู่	เนื้อผ้าเย็นส์ดีมาก ตัดเย็บดี ทรงสวย สีสวย ใส่แล้วเดินสะดวกสบาย ขนาดมาตรฐานอยู่แล้วเหมือนสั่งไซส์เอวหลวมไปนิดนึง ไม่เป็นปัญหา เมื่อเราอ้วนได้
ได้รับกางเกงแล้ว ผ่านิม ยืด ใส่ได้ แต่เอวไม่ได้สูงนะทุกคน เอวแค่นี้คือ เราสูง 164 กางเกงยาวคลุมตามม แต่ส่วนตัวรู้สึกว่ใส่กางเกงนี้แล้วขาใหญ่อะ อาจจะอยู่ที่ทรงด้วย สีด้วยมั้ง ส่วนตัวยังไม่ค่อยปลื้มเท่าไรแล้วเหมือนเวลาเดินตรงเป้าจะเป็นสามเหลี่ยมด้วย ไม่มั่นใจเลยน่าจะต้องใส่เสื้อยาวๆ 🙄🙄	ได้รับกางเกงแล้ว ผ่านิม ยืด ใส่ได้ เอวไม่สูงทุกคน เอวแค่นี้คือ เราสูง กางเกงยาวคลุมตามม ส่วนตัวรู้สึกว่ใส่กางเกงนี้แล้วขาใหญ่ อาจจะอยู่ที่ทรง สี ส่วนตัวยังไม่ค่อยปลื้มเท่าไรแล้วเหมือนเวลาเดินตรงเป้าเป็นสามเหลี่ยม ไม่มั่นใจน่าจะต้องใส่เสื้อยาว
ได้รับสินค้าตามที่สั่งแล้ว กระโปรงสีเย็นส์สีเข้มสวยดี ร้านค้าบริการดี พนักงานขนส่งบริการดี	ได้รับสินค้าตามที่สั่งแล้ว กระโปรงสีเย็นส์สีเข้มสวยดี ร้านค้าบริการดี พนักงานขนส่งบริการดี
ของจริงๆสีสวยมากกกกกก ❤️❤️❤️ ใส่ได้พอดี ของส่งไวดี 🛒🛒🛒🛒🛒 ชอบบบบบ 😄😄😄😄😄😄😄😄😄	ของจริงๆสีสวยมาก ใส่ได้พอดี ของส่งไวดี ชอบ

รูปที่ 3.7 ตัวอย่างความคิดเห็นจากการกรองคำภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



จากการเตรียมข้อมูลทั้ง 3 ขั้นตอนข้างต้นแล้ว จะได้ผลของการเตรียมข้อมูลความคิดเห็น ดังนี้

Comment	Token
เป็นกระโปรงทรงA แบบที่อยากได้ ผ้าเย็นส์สวย ใส่แล้วดูเพรียวดีคะ ทางร้านจัดส่งรวดเร็วดีคะ ส่วนไซส์ของเราดูตามตารางไซส์แล้ว +1คะ กระโปรงจริงไซส์3xl เอว 34 ค่ะ สะโพก 42 ใส่พอดีเป๊ะ!	[กระโปรง,ทรง,แบบ,ที่,อยากได้,ผ้า,เย็นส์,สวย,ใส่,แล้ว,ดู,เพรียว,ดี,ทาง,ร้าน,จัดส่ง,รวดเร็ว,ดี,ส่วน,ไซส์,ดู,ตาม,ตาราง,ไซส์,แล้ว,กระโปรง,จริง,ไซส์,เอว,สะโพก,ใส่,พอดี]
เนื้อผ้าเย็นส์ดีมากเลย ตัดเย็บดี ทรงสวย สีสวย ใส่แล้วเดินสะดวกสบาย ขนาดมาตรฐานอยู่แล้วแต่เหมือนสั่งไซส์เอวหลวมไปนิดนึงเอง แต่ไม่เป็นปัญหา เพื่อเราอ้วนได้อยู่	[เนื้อผ้า,เย็นส์,ดีมาก,ตัดเย็บ,ดี,ทรง,สวย,สี,สวย,ใส่,แล้ว,เดิน,สะดวก,สบาย,ขนาด,มาตรฐาน,อยู่แล้ว,แต่,เหมือน,สั่ง,ไซส์,เอว,หลวม,ไป,นิด,นึง,ไม่,เป็น,ปัญหา,เพื่อ,เรา,อ้วน,ได้]
ได้รับกางเกงแล้ว ผ้านุ่ม ยืด ใส่ได้ แต่เอวไม่ได้สูงนะทุกคน เอวแค่นี้คือ เราสูง 164 กางเกงยาวคลุมตาตุ่ม แต่ส่วนตัวรู้สึกว่ใส่กางเกงนี้แล้วขาใหญ่อะ อาจจะอยู่ที่ทรงด้วย สีด้วยมั้ง ส่วนตัวยังไม่ค่อยปลื้มเท่าไรแล้วเหมือนเวลาเดินตรงเป้าจะเป็นสามเหลี่ยมด้วย ไม่มั่นใจเลยน่าจะต้องใส่เสื้อยาวๆ 🙄🙄	[ได้รับ,กางเกง,แล้ว,ผ้า,นุ่ม,ยืด,ใส่,ได้,เอว,ไม่,สูง,ทุกคน,เอว,แค่นี้,คือ,เรา,สูง,164,กางเกง,ยาว,คลุม,ตาตุ่ม,แต่,ส่วน,ตัว,รู้สึก,ว่า,ใส่,กางเกง,นี้,แล้ว,ขา,ใหญ่,อะ,อาจ,จะ,อยู่,ที่,ทรง,ด้วย,สี,ด้วย,มั้ง,ส่วน,ตัว,ยัง,ไม่,ค่อย,ปลื้ม,เท่า,ไร,แล้ว,เหมือน,เวลา,เดิน,ตรง,เป้า,จะ,เป็น,สาม,เหลี่ยม,ด้วย,ไม่,มั่นใจ,เลย,น่า,จะ,ต้อง,ใส่,เสื้อ,ยาว,ๆ]
ได้รับสินค้าตามที่สั่งแล้ว กระโปรงสีเย็นส์สีเข้มสวยดี ร้านค้าบริการดี พนักงานขนส่งบริการดี	[ได้รับ,สินค้า,ตาม,ที่,สั่ง,แล้ว,กระโปรง,สี,เย็นส์,สี,เข้ม,สวย,ดี,ร้าน,ค้า,บริการ,ดี,พนักงาน,ขนส่ง,บริการ,ดี]
ของจริงๆสีสวยมากกกกกก ❤️❤️❤️ ใส่ได้พอดี ของส่งไวดี 🛒🛒🛒🛒🛒 ชอบบบบบ 🍷🍷🍷🍷🍷🍷🍷🍷🍷	[ของ,จริง,สี,สวย,มาก,ใส่,ได้,พอดี,ของ,ส่ง,ไว,ดี,ชอบ,บบบบ]

รูปที่ 3.9 ตัวอย่างความคิดเห็นจากการเตรียมข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.4.2 สร้างตัวแบบการวิเคราะห์ความรู้สึกของความคิดเห็นโดยหลักอิงตามกฎ (Rule-based Model)

ในการสร้างตัวแบบโดยหลักอิงตามกฎ จะนำความคิดเห็นที่ผ่านการเตรียมข้อมูลมาแล้วนั้น มาเทียบว่าแต่ละคำในความคิดเห็นมีปรากฏอยู่ในกระเป๋าของคำในเชิงบวก (Bag of Positive word) และกระเป๋าของคำในเชิงลบ (Bag of Negative word) มากน้อยแค่ไหน โดยที่กระเป๋าของคำในเชิงบวกและเชิงลบ อ้างอิงตามข้อมูลสาธารณะจากเว็บไซต์ <https://github.com/PyThaiNLP/lexicon-thai> เข้าถึงเมื่อวันที่ 25 กุมภาพันธ์ 2567 โดยกระเป๋าของคำในเชิงบวกมีคำทั้งหมด 509 คำ และกระเป๋าของคำในเชิงลบมีคำทั้งหมด 1,218 คำ ดังนี้

ความเอือออาทร, ความเอือเอื้อเพื่อ, พอใจ, น่าพอใจ, น่าประทับใจ, น่าพึงพอใจ, น่าฟัง, น่าภาคภูมิใจ, น่าน่าหัตถ์จรรยา, น่ายกย่อง น่ายินดี, สนุกสนาน, สวยมาก, คุ่ม, คุ่มมาก, คุ่มค่า, ไม่มีปัญหา, สด, สดใส, สนใจไยดี, เกียรติ, สนิท, สนิทสนม, สบาย, สมจริง, สมบูรณ์, มั่นคง, ยอดเยี่ยม, เยี่ยม, เยี่ยมยอด

รูปที่ 3.10 ตัวอย่างคำในกระเป๋าคำเชิงบวก

ทรมาน, ทะลุ, ทุเรศ, เทียม, นักเลง, ลบ, เนื้อ, ปรัมปรา, เท็จ, ต่อแหล, ตาขาว, ฉุนเฉียว, เณโก, ฉื้อย, ชะงัด, ทรหด, ทรมาน ทะลุ คือ, ทุเรศ, เท็จ, เทียม, น้อย, นักเลง, ช่วย, บิน, ดื้อ, เบียง, เปื้อ, ผิดหวัง, รุ่ยร่าย, ล้ม, ลัก, ล้าสมัย, ลำบาก, แคว, โลเนล, ยาก, เหม็น ไม่เชื่อ

รูปที่ 3.11 ตัวอย่างคำในกระเป๋าคำเชิงลบ

หลังจากทำการตรวจสอบคำของแต่ละความคิดเห็นแล้วว่ามียู่ในกลุ่มคำเชิงบวกและเชิงลบ เป็นจำนวนกี่คำ จะนำมาคิดสัดส่วนของคำเชิงบวกและสัดส่วนของคำเชิงลบ เพื่อหาสัดส่วนของคำเชิงบวกและคำเชิงลบ ดังสมการต่อไปนี้

$$\text{Positive Ratio} = \frac{N_{\text{positive}}}{N} \times 100$$

โดยที่  $N_{\text{positive}}$  คือ จำนวนค่าที่พบที่กระเปาะของค่าเชิงบวกในแต่ละความคิดเห็น  
 $N$  คือ จำนวนค่าทั้งหมดในแต่ละความคิดเห็น

$$\text{Negative Ratio} = \frac{N_{\text{negative}}}{N} \times 100$$

โดยที่  $N_{\text{negative}}$  คือ จำนวนค่าที่พบที่กระเปาะของค่าเชิงลบในแต่ละความคิดเห็น  
 $N$  คือ จำนวนค่าทั้งหมดในแต่ละความคิดเห็น

### 3.4.3 หาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าเชิงบวกและค่าเชิงลบ

เมื่อได้ค่า positive ratio และ negative ratio ของแต่ละความคิดเห็นมาแล้ว จากนั้นนำไปหาค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (sd) ของค่าเชิงบวกและค่าเชิงลบของความคิดเห็นแต่ละร้านค้า ดังนี้

$$\text{Mean} = \frac{\sum X}{N}$$

โดยที่  $X$  คือ จำนวนค่าที่เป็น positive หรือ negative ของแต่ละความคิดเห็นของร้านค้าหนึ่ง  
 $N$  คือ จำนวนค่าทั้งหมดในทุกความคิดเห็นของร้านค้า

$$\text{S.D.} = \sqrt{\frac{\sum_{i=1}^n (x - \mu)^2}{N}}$$

โดยที่  $X$  คือ ค่าในชุดข้อมูลแต่ละตัว  
 $\mu$  คือ ค่าเฉลี่ยของชุดข้อมูล  
 $N$  คือ จำนวนข้อมูลทั้งหมด

### 3.4.4 จัดอันดับคะแนนรีวิวใหม่ให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น

เมื่อทำการวิเคราะห์ความรู้สึกของความคิดเห็นต่อสินค้าแล้ว จากนั้นจะทำการจัดกลุ่มคะแนนดาวให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น โดยทำการกำหนดคะแนนหัวใจใหม่ ดังนี้

ในการกำหนดคะแนนหัวใจหลัก ตั้งแต่ 4.00 ถึง 5.00 หัวใจ จะใช้ข้อมูลความคิดเห็นที่เป็น positive ratio ในการกำหนดเป็นคะแนนหัวใจหลัก

โดยที่ค่ากลางจะเป็น ค่าเฉลี่ย positive ของทุกร้านค้า และ ขอบบนจะเป็นค่า +3 sd positive ขอบล่างจะเป็นค่า -3 sd positive ของทุกร้านค้า

และสำหรับการปรับคะแนนหัวใจเสริม ตั้งแต่ 0.00 ถึง 0.50 หัวใจ จะใช้ข้อมูลความคิดเห็นที่เป็น negative ratio ในการปรับคะแนนหัวใจเสริม

โดยที่ค่ากลางจะเป็น ค่าเฉลี่ย negative ของทุกร้านค้า และ ขอบบนจะเป็นค่า +3 sd negative ขอบล่างจะเป็นค่า -3 sd negative ของทุกร้านค้า

#### 3.4.4.1 วิธีการปรับค่าคะแนนรีวิวดาวจริง ให้เป็นคะแนนหัวใจใหม่

วิธีการแปลงค่าของคะแนนดาวจริงให้เป็นคะแนนหัวใจที่กำหนดใหม่ จะใช้วิธีการแปลงค่าข้อมูลเดิมไปยังช่วงใหม่ที่กำหนดขึ้น โดยใช้เทคนิคการปรับค่าแกน ซึ่งมีวิธีการดังนี้

1. คำนวณอัตราส่วนระหว่างค่า  $y$  แต่ละค่ากับช่วงข้อมูลเดิม

$$\text{Ratio} = \frac{(y - y_{\min})}{(y_{\max} - y_{\min})}$$

โดยที่	$y$	คือ	ค่าข้อมูลเดิม
	$y_{\min}$	คือ	ค่าต่ำสุดของข้อมูลเดิม
	$y_{\max}$	คือ	ค่าสูงสุดของข้อมูลเดิม

ratio จะอยู่ระหว่าง 0 ถึง 1 แสดงถึงสัดส่วนของค่า  $y$  แต่ละค่า

2. แปลงค่าข้อมูล  $y$  เดิมไปยังช่วงใหม่ที่กำหนด

$$y_{scaled} = ratio * (y_{max} - y_{min}) + y_{min}$$

โดยที่  $y_{scaled}$  คือ ค่าข้อมูลที่ปรับขนาดแล้ว  
 $y_{max}$  คือ ค่าสูงสุดที่ต้องการสำหรับข้อมูล  $y$  ที่ปรับขนาดแล้ว  
 $y_{min}$  คือ ค่าต่ำสุดที่ต้องการสำหรับข้อมูล  $y$  ที่ปรับขนาดแล้ว

ตัวอย่างการเขียน Python สำหรับการกำหนดคะแนนหัวใจหลัก ตั้งแต่ 4.00 ถึง 5.00 หัวใจ โดยจะใช้ข้อมูลความคิดเห็นที่เป็น positive ratio ในการกำหนดเป็นคะแนนหัวใจหลัก

```
import numpy as np

# ข้อมูล y
y = [43.200, 44.970, 29.484, 47.878, 41.063, 40.820, 16.712, 57.076, 50.791, 39.604]

# กำหนดค่าช่วงใหม่
new_ymin = 4.0
new_ymax = 5.0

# คำนวณค่าต่ำสุดและสูงสุดของข้อมูลดั้งเดิม
y_min = np.min(y)
y_max = np.max(y)

# ปรับขนาดค่าข้อมูล
y_scaled = []
for y_value in y:
    y_scaled_value = (y_value - y_min) / (y_max - y_min) * (new_ymax - new_ymin) + new_ymin
    y_scaled.append(y_scaled_value)

# แสดงผล
print(y_scaled)
```

รูปที่ 3.12 ตัวอย่างการปรับค่าคะแนนหัวใจหลัก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.4.5 เปรียบเทียบคะแนนหัวใจใหม่กับคะแนนรีวิวดาวจริงแบบถ่วงน้ำหนัก

เมื่อทำการกำหนดคะแนนหัวใจใหม่แล้ว จากนั้นจะนำมาเปรียบเทียบกับคะแนนรีวิวดาวจริงแบบถ่วงน้ำหนักของแต่ละร้านค้า โดยทำการหาคะแนนรีวิวจริงแบบถ่วงน้ำหนัก ดังนี้

$$\text{คะแนนรีวิวของร้านค้าแบบถ่วงน้ำหนัก} = \frac{\sum_{i=1}^n W_i X_i}{N}$$

โดยที่  $X_i$  คือ คะแนนดาว  $i$   
 $W_i$  คือ จำนวนความคิดเห็นที่ได้รับคะแนนดาว  $i$   
 $N$  คือ จำนวนความคิดเห็นทั้งหมด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

### ผลการวิจัยและการอภิปรายผล

#### 4.1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคำเชิงบวกและคำเชิงลบ

ค่าเฉลี่ย (mean) และ ส่วนเบี่ยงเบนมาตรฐาน (sd) ของแต่ละร้านค้า ได้ผลลัพธ์ดังนี้

ร้านค้า	mean_pos	sd_pos	mean_neg	sd_neg
Acaseyshop	43.20	6.20	2.76	2.75
BEMING.BKK	29.48	5.54	2.24	2.36
BabarryJeans	44.97	6.38	1.75	2.47
CHIC.th	47.88	6.47	1.83	2.51
Esolo	41.06	6.41	2.35	2.81
Giordano	40.82	6.13	1.62	2.25
Nessaj	16.71	5.57	1.57	2.50
SPINNYHOUSE	57.08	6.62	1.85	2.55
Yuedpao	39.60	6.40	2.18	2.34
wipcloset	50.79	6.83	2.89	2.95

รูปที่ 4.1 ค่าเฉลี่ย (mean) และ ส่วนเบี่ยงเบนมาตรฐาน (sd) ระดับร้านค้า

จากการสร้างตัวแบบการวิเคราะห์ความรู้สึกของความคิดเห็นโดยหลักอิงตามกฎ เพื่อหาค่า positive ratio และ negative ratio ของแต่ละความคิดเห็น จากนั้นนำไปหาค่าเฉลี่ย (mean) และ ส่วนเบี่ยงเบนมาตรฐาน (sd) ของคำเชิงบวกและคำเชิงลบของความคิดเห็นแต่ละร้านค้า พบว่าค่าเฉลี่ย positive ของแต่ละร้านค้าอยู่ระหว่าง 16.71 ถึง 57.08 และส่วนเบี่ยงเบนมาตรฐาน positive ของแต่ละร้านค้าอยู่ระหว่าง 5.54 ถึง 6.83

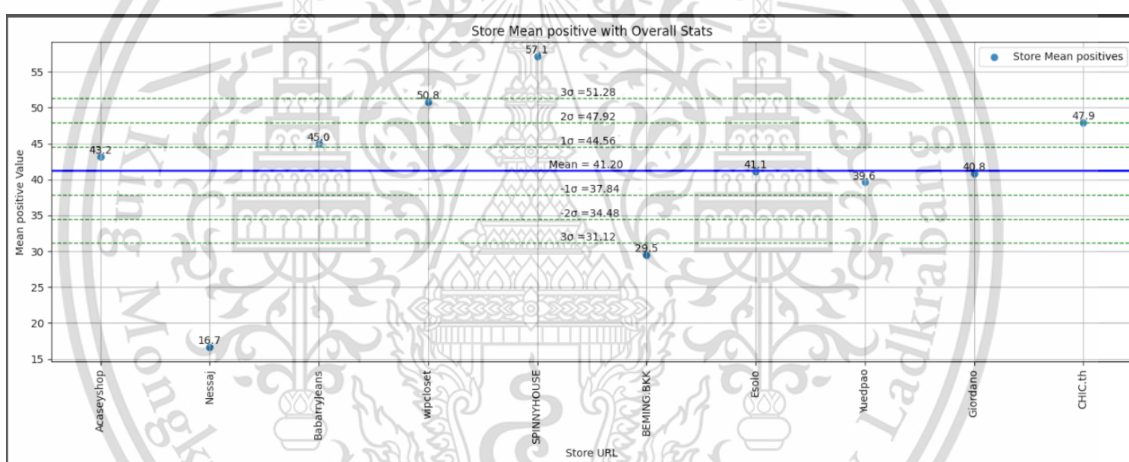
และค่าเฉลี่ย negative ของแต่ละร้านค้าอยู่ระหว่าง 1.57 ถึง 2.89 และส่วนเบี่ยงเบนมาตรฐาน negative ของแต่ละร้านค้าอยู่ระหว่าง 2.25 ถึง 2.95

## 4.2 ผลการจัดอันดับคะแนนรีวิว โดยการกำหนดคะแนนหัวใจใหม่

### 4.2.1 การกำหนดคะแนนหัวใจหลัก ตั้งแต่ 4.00 ถึง 5.00 หัวใจโดยใช้ข้อมูลความคิดเห็นที่เป็น positive ratio

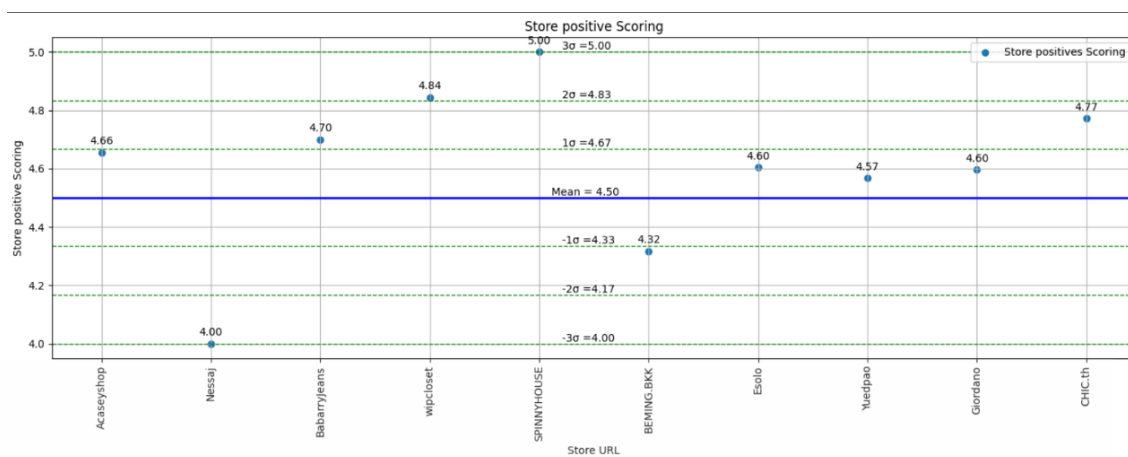
ในการกำหนดคะแนนหัวใจหลักตั้งแต่ 4.00 ถึง 5.00 หัวใจ จะใช้ข้อมูลความคิดเห็นที่เป็น positive ratio ในการกำหนดเป็นคะแนนหัวใจหลัก

โดยที่ค่ากลางจะเป็น ค่าเฉลี่ย positive ของทุกร้านค้า และ ขอบบนจะเป็นค่า +3 sd positive ขอบล่างจะเป็นค่า -3 sd positive ของทุกร้านค้าได้ผลลัพธ์ ดังนี้



รูปที่ 4.2 mean positive ของแต่ละร้านค้า เทียบกับ overall mean positive

และจากนั้นทำการแปลงค่าของคะแนนดาวจริงให้เป็นคะแนนหัวใจที่กำหนดใหม่ตั้งแต่ 4.00 ถึง 5.00 หัวใจ โดยใช้วิธีการแปลงค่าข้อมูลเดิมไปยังช่วงใหม่ที่กำหนดขึ้น โดยใช้เทคนิคการปรับค่าแกนได้ผลลัพธ์ ดังนี้



รูปที่ 4.3 การกำหนดคะแนนหัวใจหลักใหม่ของแต่ละร้านค้า

จากนั้นทำการสรุปผลการกำหนดคะแนนหัวใจหลักใหม่ตั้งแต่ 4.00 ถึง 5.00 หัวใจ เทียบกับ mean positive ที่ยังไม่ได้รับการปรับของแต่ละร้านค้า ได้ผลลัพธ์ดังตารางต่อไปนี้

ตารางที่ 4.1 ผลการกำหนดคะแนนหัวใจหลักใหม่ เทียบกับ mean positive ของแต่ละร้านค้า

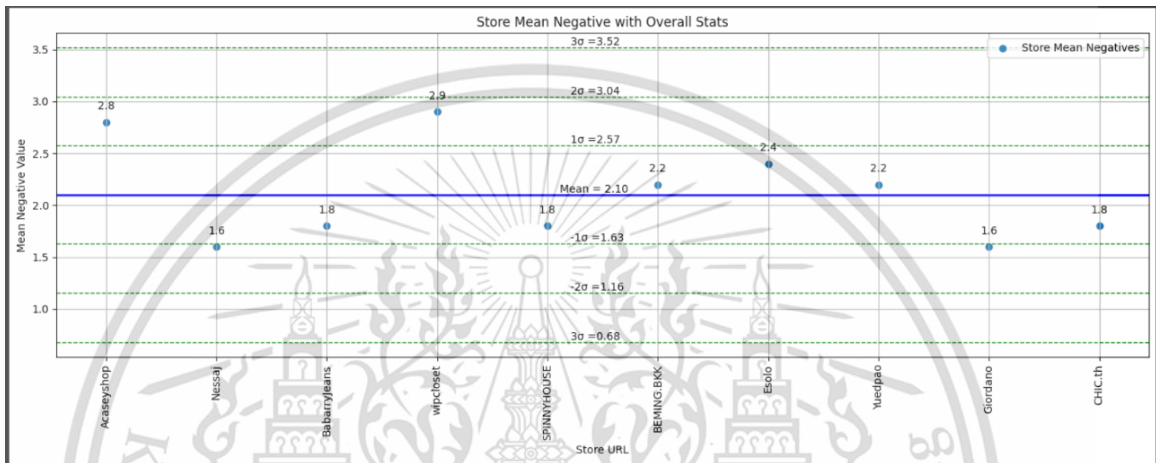
ร้านค้า	Mean Positive	คะแนนหัวใจหลักใหม่
Acaseyshop	43.20	4.66
BEMING.BKK	29.50	4.32
BabarryJeans	45.00	4.70
CHIC.th	47.90	4.77
Esolo	41.10	4.60
Giordano	40.80	4.60
Nessaj	16.70	4.00
SPINNYHOUSE	57.10	5.00
Yuedpao	39.60	4.57
wipcloset	50.80	4.84

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.2.2 การปรับคะแนนหัวใจเสริม ตั้งแต่ 0.00 ถึง 0.50 หัวใจโดยใช้ข้อมูลความคิดเห็นที่เป็น negative ratio

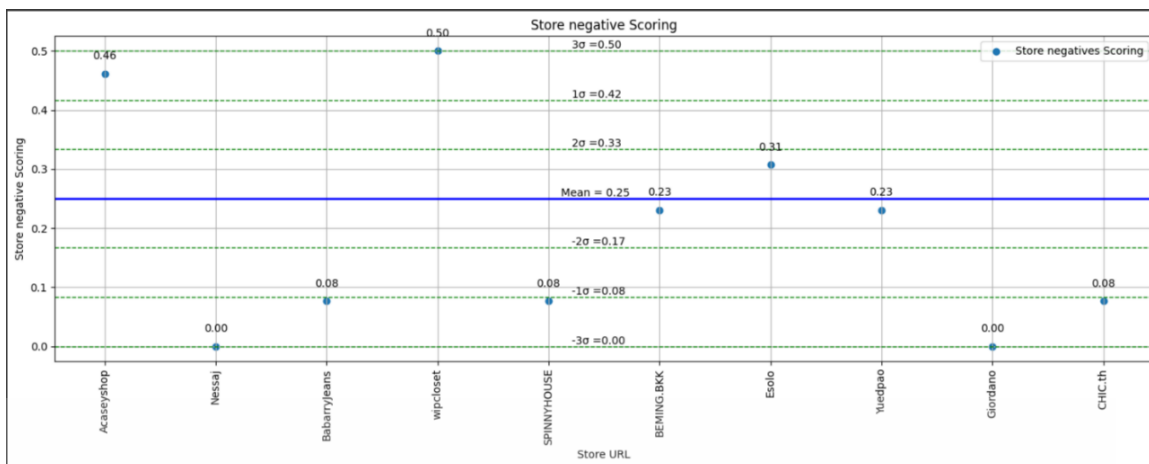
ในการปรับคะแนนหัวใจเสริมตั้งแต่ 0.00 ถึง 0.50 หัวใจ จะใช้ข้อมูลความคิดเห็นที่เป็น negative ratio ในการกำหนดเป็นคะแนนหัวใจเสริม

โดยที่ค่ากลางจะเป็น ค่าเฉลี่ย negative ของทุกร้านค้า และ ขอบบนจะเป็นค่า +3 sd negative ขอบล่างจะเป็นค่า -3 sd negative ของทุกร้านค้าได้ผลลัพธ์ ดังนี้



รูปที่ 4.4 mean negative ของแต่ละร้านค้า เทียบกับ overall mean negative

และจากนั้นทำการแปลงค่าของคะแนนดาวจริงให้เป็นคะแนนหัวใจเสริมที่กำหนดใหม่ตั้งแต่ 0.00 ถึง 0.50 หัวใจ โดยใช้วิธีการแปลงค่าข้อมูลเดิมไปยังช่วงใหม่ที่กำหนดขึ้น โดยใช้เทคนิคการปรับค่าแทนได้ผลลัพธ์ ดังนี้



#### รูปที่ 4.5 การกำหนดคะแนนหัวใจเสริมของแต่ละร้านค้า

จากนั้นทำการสรุปผลการกำหนดคะแนนหัวใจเสริมตั้งแต่ 0.00 ถึง 0.50 หัวใจ เทียบกับ mean negative ที่ยังไม่ได้รับการปรับของแต่ละร้านค้า ได้ผลลัพธ์ดังตารางต่อไปนี้

#### ตารางที่ 4.2 ผลการปรับคะแนนหัวใจเสริม เทียบกับ mean negative ของแต่ละร้านค้า

ร้านค้า	Mean Negative	คะแนนหัวใจเสริม
Acaseyshop	2.80	0.46
BEMING.BKK	2.20	0.23
BabarryJeans	1.80	0.08
CHIC.th	1.80	0.08
Esolo	2.40	0.31
Giordano	1.60	0.00
Nessaj	1.60	0.00
SPINNYHOUSE	1.80	0.08
Yuedpao	2.20	0.23
wipcloset	2.90	0.50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 4.2.3 ผลการกำหนดคะแนนหัวใจใหม่

การกำหนดคะแนนหัวใจใหม่ จะใช้การกำหนดคะแนนหัวใจหลักตั้งแต่ 4.00 ถึง 5.00 หัวใจของข้อมูลความคิดเห็นที่เป็น positive ratio เป็นคะแนนหลัก จากนั้นนำมาหักลบกับการปรับคะแนนหัวใจเสริม ตั้งแต่ 0.00 ถึง 0.50 หัวใจ โดยใช้ข้อมูลความคิดเห็นที่เป็น negative ratio ได้ผลลัพธ์ของคะแนนหัวใจใหม่ของแต่ละร้านค้าแสดงดังตารางต่อไปนี้

ตารางที่ 4.3 ตารางคะแนนหัวใจใหม่ของแต่ละร้านค้า

ร้านค้า	คะแนนหัวใจหลัก	คะแนนหัวใจเสริม	คะแนนหัวใจใหม่
Acaseyshop	4.66	0.46	4.20
BEMING.BKK	4.32	0.23	4.09
BabarryJeans	4.70	0.08	4.62
CHIC.th	4.77	0.08	4.69
Esolo	4.60	0.31	4.29
Giordano	4.60	0.00	4.60
Nessaj	4.00	0.00	4.00
SPINNYHOUSE	5.00	0.08	4.92
Yuedpao	4.57	0.23	4.34
wipcloset	4.84	0.50	4.34

### 4.3 เปรียบเทียบคะแนนหัวใจใหม่กับคะแนนรีวิวดาวจริงแบบถ่วงน้ำหนัก

ผลลัพธ์ของการกำหนดคะแนนหัวใจใหม่เปรียบเทียบกับคะแนนรีวิวจริงแบบถ่วงน้ำหนักของระดับร้านค้า ได้ผลดังตารางต่อไปนี้

#### ตารางที่ 4.4 ตารางเปรียบเทียบการกำหนดคะแนนหัวใจใหม่ กับ คะแนนดาวจริงแบบถ่วงน้ำหนัก

ร้านค้า	คะแนนดาวจริง จากการถ่วงน้ำหนัก	คะแนนหัวใจใหม่	ระยะห่างระหว่าง คะแนนรีวิวจริง และคะแนนหัวใจใหม่
Acaseyshop	4.95	4.20	0.75
BEMING.BKK	4.97	4.09	0.88
BabarryJeans	4.98	4.62	0.36
CHIC.th	4.92	4.69	0.23
Esolo	4.97	4.29	0.68
Giordano	4.98	4.60	0.38
Nessaj	4.92	4.00	0.92
SPINNYHOUSE	4.95	4.92	0.03
Yuedpao	5.00	4.34	0.66
wipcloset	4.97	4.34	0.63

#### 4.4 ผลการจัดกลุ่มคะแนนหัวใจใหม่ ให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น

เมื่อได้ผลลัพธ์ของการกำหนดคะแนนหัวใจใหม่มาแล้ว จากนั้นทำการจัดกลุ่มคะแนนหัวใจให้สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น โดยจะทำการกำหนดกลุ่มคะแนนหัวใจใหม่ ดังนี้

ร้านค้าที่มีคะแนนหัวใจใหม่ตั้งแต่ 3.50 ถึง 4.00 คะแนน จะจัดให้อยู่ในกลุ่มคะแนนรีวิวดี

ร้านค้าที่มีคะแนนหัวใจใหม่ตั้งแต่ 4.01 ถึง 4.50 คะแนน จะจัดให้อยู่ในกลุ่มคะแนนรีวิวดีมาก




ร้านค้าที่มีคะแนนหัวใจใหม่ตั้งแต่ 4.51 ถึง 5.00 คะแนน จะจัดให้อยู่ในกลุ่มคะแนนรีวิวดีที่สุด

ร้านค้าที่มีคะแนนหัวใจใหม่ ตั้งแต่ 3.50 ถึง 4.00 คะแนน ที่อยู่ในกลุ่มคะแนนรีวิวดี ได้แก่ร้านค้า Nessaj เพียงหนึ่งร้าน

ร้านค้าที่มีคะแนนหัวใจใหม่ ตั้งแต่ 4.01 ถึง 4.50 คะแนน ที่อยู่ในกลุ่มคะแนนรีวิวดีมาก ได้แก่ร้านค้า Acaseyshop, BEMING.BKK, Esolo, Yuedpao และ wipcloset

ร้านค้าที่มีคะแนนหัวใจใหม่ ตั้งแต่ 4.51 ถึง 5.00 คะแนน ที่อยู่ในกลุ่มคะแนนรีวิวดีที่สุด ได้แก่ร้านค้า BabarryJeans, CHIC.th, Giordano และ SPINNYHOUSE แสดงดังตารางที่ 4.5

#### ตารางที่ 4.5 ตารางแสดงผลการจัดกลุ่มคะแนนหัวใจใหม่

คะแนนรีวิวร้านค้า	ร้านค้า
 ดีที่สุด	BabarryJeans CHIC.th Giordano SPINNYHOUSE
 ดีมาก	Acaseyshop BEMING.BKK Esolo Yuedpao wipcloset
 ดี	Nessaj

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

#### 5.1 สรุปผลการวิจัย

จากการศึกษาการเปรียบเทียบความรู้สึกของความคิดเห็นกับคะแนนหรือดาวต่อสินค้า เมื่อทำการวิเคราะห์ความรู้สึกของความคิดเห็น โดยใช้ตัวแบบอิงตามกฎ (Rule-based) พบว่าค่า Positive Ratio โดยเฉลี่ยอยู่ที่ 41.20 และค่า Negative Ratio โดยเฉลี่ยอยู่ที่ 2.10 และพบว่าคะแนนดาวจริงจากการถ่วงน้ำหนักอยู่ในช่วง 4.92 ถึง 5.00 คะแนน ทำให้ก่อนการจัดอันดับคะแนนหัวใจใหม่ คะแนนร้านค้าทุกร้านค้าจะอยู่ในกลุ่มคะแนนรีวิวดีที่สุด

จากผลการศึกษาข้างต้นสรุปได้ว่า การดึงข้อมูลผ่านเว็บไซต์ (Web Scraping) โดยเลือกจากการสุ่มร้านค้าหมวดหมู่เสื้อผ้าสุภาพสตรี 10 ร้านค้า และเรียงลำดับรายการสินค้าในร้านค้านั้นที่มีจำนวนความคิดเห็นมากกว่า 100 รายการขึ้นไป ทำให้ได้ข้อมูลที่มีการให้คะแนนหรือดาว 4-5 ดาวเป็นจำนวนมาก ซึ่งคิดเป็นสัดส่วนถึงร้อยละ 99 ซึ่งเมื่อนำมาวิเคราะห์แล้วนั้น ทำให้ตัวแบบเอนเอียงไปที่การทำนายความรู้สึกว่าเป็นการให้ 5 ดาว เป็นส่วนใหญ่ และจากการสำรวจผู้วิจัยพบว่าโดยส่วนมากคนมักจะให้ความคิดเห็นต่อสินค้าน้อยกว่าร้อยละ 10 เมื่อเทียบกับปริมาณที่สินค้าสามารถขายได้ ทำให้การเลือกสินค้าที่มีความคิดเห็นเกิน 100 ข้อความนั้นเปรียบเสมือนการเลือกสินค้าที่เป็นที่น่าพอใจต่อผู้บริโภคอยู่แล้ว จึงทำให้ความคิดเห็นที่ได้จากการดึงข้อมูลมานั้นมีความโน้มเอียงไปที่คะแนน 4-5 ดาวเป็นส่วนใหญ่

ดังนั้นการศึกษาในครั้งนี้จึงทำการจัดอันดับคะแนนรีวิวดีใหม่ให้เป็นคะแนนหัวใจที่สะท้อนความคิดเห็นที่มีต่อสินค้า จากการวิเคราะห์ความรู้สึกของความคิดเห็นต่อสินค้าโดยใช้ตัวแบบอิงตามกฎ (Rule-based) ออกเป็นร้านค้าที่ได้รับคะแนนรีวิวดีที่สุด ดีมาก และ ดี และเมื่อทำการจัดอันดับคะแนนรีวิวให้เป็นคะแนนหัวใจใหม่ พบว่าทุกร้านค้าจะได้รับคะแนนรีวิวลดลงตั้งแต่ 0.03 จนถึง 0.92 คะแนน และผลการกำหนดคะแนนหัวใจใหม่มีค่าอยู่ในช่วง 4.00 ถึง 4.92 หัวใจ ทำให้การจัดอันดับคะแนนหัวใจใหม่ คะแนนรีวิวร้านค้าดีที่สุด เหลือเพียง 4 ร้านค้าเท่านั้น

การศึกษาในครั้งนี้จึงทำให้ผู้บริโภคทราบคะแนนรีวิวที่สะท้อนความคิดเห็นที่มีต่อสินค้ามากขึ้น ซึ่งสามารถนำไปประกอบการตัดสินใจซื้อสินค้าผ่านแพลตฟอร์มได้รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น และเพื่อให้ผู้ประกอบการร้านค้าบนแพลตฟอร์มทราบถึงกลุ่มคะแนนรีวิวที่มีต่อสินค้าและสามารถปรับปรุงสินค้าและบริการให้ดียิ่งขึ้น นอกจากนี้แพลตฟอร์มยังสามารถนำไปประยุกต์ใช้เพื่อปรับปรุงบริการ เช่นสร้างระบบการจัดกลุ่มคะแนนรีวิวให้เป็นคะแนนหัวใจที่สะท้อนความคิดเห็นที่มีต่อสินค้ามากยิ่งขึ้นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 5.2 ข้อเสนอแนะ

สำหรับการวิจัยต่อไปควรมีการวิเคราะห์การเพิ่มประสิทธิภาพของตัวแบบเพื่อให้ผลการทำนายมีประสิทธิภาพขึ้น โดยนำเทคนิคอื่น ๆ ร่วมในกระบวนการจัดเตรียมข้อมูล เช่น การเพิ่มจำนวนคำศัพท์ เพิ่มเทคนิคการแบ่งชนิดของคำ (Named-Entity Tagging) เพื่อช่วยให้การแยกคำมีความชัดเจนมากขึ้น รวมทั้งการเพิ่มชุดข้อมูลคำแสดงเพื่อให้ตัวแบบสามารถวิเคราะห์คำประชดประชันได้ รวมถึงเพิ่มข้อมูลของคำและประโยค ที่เป็นการคัดเลือกคำบอกความรู้สึก บวก ลบ และ เป็นกลาง (Sentiment Word level) อีกทั้งยังควรเพิ่มปริมาณข้อมูลของแต่ละคะแนนให้มีความสมดุลมากขึ้น (Balance Data) เพื่อไม่ให้เกิดการทำนายเอนเอียงไปทางประเภทใดประเภทหนึ่ง และในการวิจัยครั้งนี้ ผู้วิจัยมีการนำการแสดงความรู้สึกด้วยสัญลักษณ์หรืออีโมจิ (Emoji) ออกจากความคิดเห็น ซึ่งอีโมจินั้นมีการสื่อถึงความรู้สึกในเชิงบวกและเชิงลบอยู่ด้วย เช่น อีโมจิยิ้ม ที่สื่อถึงเชิงบวก หรือ อีโมจิร้องไห้ ที่สื่อถึงเชิงลบ ดังนั้นหากสามารถนำอีโมจิเหล่านี้มาทำการวิเคราะห์ได้ก็อาจทำให้ผลการวิเคราะห์ความรู้สึกของความคิดเห็นและคะแนนรีวิวมีประสิทธิภาพมากขึ้น

## 5.3 ข้อจำกัดในการศึกษา

- 1) ตัวแบบยังไม่สามารถวิเคราะห์คำประชดประชัน (Sarcasm) ได้เนื่องจากชุดข้อมูลที่นำมาใช้ ยังไม่มีคำประเภทดังกล่าว ซึ่งต้องอาศัยตัวแบบในการเรียนรู้แยกต่างหากและอาศัยเหมืองข้อความ (Text mining) แบบใช้กฎ (Rule-based)
- 2) กระเป๋าของคำเชิงบวก (Bag of Positive word) และกระเป๋าของคำเชิงลบ (Bag of Negative word) จาก PyThaiNLP ยังไม่ครอบคลุมคำศัพท์ในเชิงบวกและเชิงลบทั้งหมด เนื่องจากปัจจุบันในสื่อสังคมออนไลน์มีการใช้คำศัพท์ใหม่ๆ เกิดขึ้น
- 3) ตัวแบบในงานวิจัยนี้สามารถทำงานได้บนภาษาไทยเท่านั้น ซึ่งสามารถไปพัฒนาต่อยอดให้รองรับภาษาอื่นๆ ได้ โดยการใช้คลังคำศัพท์ของภาษานั้นๆ

## เอกสารอ้างอิง

- Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L., Chormai, P., & Phatthiyaphaibun, W. (2016). *Thai Natural Language Processing in Python*. Retrieved from pythainlp.github: <https://pythainlp.github.io/docs/2.2>
- ชญัญชัย เฟื่องพรหม, และ วรรัตน์ สงฆ์แป้น. (2019). แบบจำลองแบบผสมเพื่อสร้างกฎสำหรับ คำถาม-คำตอบโดยใช้เทคนิคการทำเหมืองข้อความ. *Walailak Procedia*.
- นภัสสร แยมอุทัย. (2015). *ทัศนคติของผู้บริโภคในการอ่านบทความรีวิวผ่านช่องทางออนไลน์*. นิภาพร เทศศรีเมือง. (2016). *ความคิดเห็นที่เป็นตัวแทนผลิตภัณฑ์*.
- พนิดา กาศกลางดอน. (2020). *การวิเคราะห์ประสบการณ์จากการใช้บริการโรงพยาบาลในประเทศไทยจากความคิดเห็นของผู้ใช้บริการ*.
- พิศิษฐ์ บวรเลิศสุธี, และ วรภัทร ไพรีเกรง. (2022). ตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับ จำแนกประเภทบทความแนะนำสินค้าออนไลน์. *Engineering and Digital Technology (JEDT)*.
- วิสุตา เทศเมือง, และ นิเวศ จิระวิจิตรชัย. (2017). การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับ การรีวิวสินค้าออนไลน์โดยใช้ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน. *Engineering Journal of Siam University*.
- วันसारณ มีประเสริฐ, และ เอกรัฐ รัชฎาญจน์. (2021). การวิเคราะห์ความคิดเห็นจากทวีตเตอร์ของ ลูกค้าบริษัทข้อปีประเทศไทย. *วารสารระบบสารสนเทศด้านธุรกิจ (JISB)*.
- สมัคร ชัยสงวน. (2018). การพัฒนาระบบวิเคราะห์ความรู้สึกแบบเรียลไทม์ของนักศึกษาบนเฟซบุ๊ก โดยใช้ตัวจำแนกข้อมูล นาอ็ฟ เบย์ สำหรับภาษาไทย. *The 8th STOU National Research Conference*.
- สันติ สุขเกษม. (2018). การวิเคราะห์รูปแบบความคิดเห็นของข้อมูลโรงแรมโดยใช้เทคนิคการตัดคำ แบบผสมผสาน. *วารสารวิชาการวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏ นครสวรรค์*.
- อิทธิศักดิ์ ศรีดำ. (2022). ระบบปัญญาประดิษฐ์สำหรับการจัดหมวดหมู่ข้อคิดเห็นและข้อเสนอแนะ ของประชาชนที่มี ต่อโครงการของรัฐ. *วารสารวิทยาศาสตร์และเทคโนโลยีมหาวิทยาลัย มหาสารคาม*.
- อิสราภาพ ล้อรัตนไชยยงค์. (2017). *การสร้างคลังคำศัพท์บอกความรู้สึกในภาษาไทยจากบทวิจารณ์ ออนไลน์*.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ประวัติผู้เขียน

ชื่อ	นางสาว วรณดา ตั้งถาวรกิจ
วัน เดือน ปีเกิด	20 สิงหาคม 2540
ที่อยู่ปัจจุบัน	339/1104 ถนน จตุรทิศ แขวง มักกะสัน เขต ราชเทวี กรุงเทพฯ 10400
ประวัติการศึกษา	(2562) วิทยาศาสตรบัณฑิต สาขาสถิติ เกรตเฉลี่ย 3.33 (มหาวิทยาลัยธรรมศาสตร์) (2566) วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูลและการวิเคราะห์ (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้