# การพัฒนาระบบวิเคราะห์ความคิดเห็นของนักท่องเที่ยวบนแพลตฟอร์ม X: กรณีศึกษาจังหวัดภูเก็ต

# Development of a Tourist Opinion Analysis System on Platform X: A Case Study of Phuket Province

**ธิดารัตน์ ศิวะนาถนุสรณ์[1] พิทา จารุพูนผล[2] และ ณสิทธิ์ เหล่าเส็น[3*]**

Thidarat Siwanartnusorn[1], Pita Jarupunphol[2] and Nasith Laosen[3*]

[1]สาขาวิชาเทคโนโลยีดิจิทัล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏภูเก็ต 21 ม.6 ถ.เทพกระษัตรี ต.รัษฎา อ.เมือง จ.ภูเก็ต 83000

[1]Department of Digital Technology, Faculty of Science & Technology, Phuket Rajabhat University 21 Moo 6, Ratsada Subdistrict, Mueang District, Phuket 83000

*Corresponding author E-mail: nasith.l@pkru.ac.th

**บทคัดย่อ**

บทความนี้นำเสนอการพัฒนาระบบวิเคราะห์ความคิดเห็นของนักท่องเที่ยวบนแพลตฟอร์ม X (ทวิตเตอร์) โดยใช้กรณีศึกษาจังหวัดภูเก็ต เพื่อสนับสนุนการวางแผนและพัฒนาการท่องเที่ยวให้ตอบสนองต่อความต้องการของนักท่องเที่ยวได้อย่างมีประสิทธิภาพ ระบบนี้จะรวบรวมข้อมูลทวีตจากแพลตฟอร์ม X และนำเทคนิคการเรียนรู้ของเครื่องและการประมวลผลภาษาธรรมชาติมาประยุกต์ใช้ในการสกัดข้อมูล เช่น การระบุประเทศ การดึงแฮชแท็ก การวิเคราะห์หัวข้อ และการวิเคราะห์ขั้วอารมณ์ พร้อมทั้งแสดงผลข้อมูลในรูปแบบแดชบอร์ดปฏิสัมพันธ์ที่ใช้งานง่าย ผลการวิจัยแสดงให้เห็นว่าระบบสามารถทำงานได้ตรงตามข้อกำหนดของซอฟต์แวร์ โดยมีความถูกต้องในกระบวนการดึงข้อมูล การสกัดประเทศ และการสกัดแฮชแท็กในทุกกรณี การวิเคราะห์หัวข้อสามารถจัดกลุ่มข้อความได้อย่างถูกต้องในอัตราความแม่นยำร้อยละ 86.25 ขณะที่การวิเคราะห์ขั้วอารมณ์สามารถจำแนกข้อความเป็นเชิงบวก เชิงลบ หรือเป็นกลางได้ในอัตราความแม่นยำร้อยละ 76.25

**คำสำคัญ**: การวิเคราะห์ความคิดเห็นของนักท่องเที่ยว การวิเคราะห์ข้อมูลจากสื่อสังคมออนไลน์ การวิเคราะห์ขั้วอารมณ์ การวิเคราะห์หัวข้อ

**Abstract**

This article presents the development of a tourist opinion analysis system on Platform X (Twitter), using Phuket Province as a case study. The system aims to support the planning and development of tourism to better meet the needs of tourists effectively. The system collects tweet data from Platform X and applies Machine Learning (ML) and Natural Language Processing (NLP) techniques to extract and analyze information such as country, hashtag, topic, and sentiment. The processed data is then presented in an interactive dashboard format. The research findings demonstrate that the system operates in accordance with the software specification. The system achieved high accuracy in data retrieval, country identification, and hashtag extraction across all test cases. Topic modeling successfully grouped tweets with an accuracy rate of 86.25%, while sentiment analysis classified tweets as positive, negative, or neutral with an accuracy rate of 76.25%.

**Keywords**: Tourist Opinion Analysis, Social Media Analytics, Sentiment Analysis, Topic Modeling

**บทนำ**

การท่องเที่ยวเป็นหนึ่งในอุตสาหกรรมหลักที่ขับเคลื่อนเศรษฐกิจของประเทศไทย โดยมีบทบาทสำคัญในการสร้างงานและรายได้ให้กับประชาชนในภาคส่วนต่าง ๆ เช่น ธุรกิจโรงแรม บริการนำเที่ยว ร้านอาหาร และอื่น ๆ จังหวัดภูเก็ตถือเป็นเมืองท่องเที่ยวสำคัญที่มีชื่อเสียงระดับโลก ด้วยทรัพยากรธรรมชาติที่สวยงามและวัฒนธรรมอันเป็นเอกลักษณ์ ทำให้มีนักท่องเที่ยวจากทั่วโลกเดินทางมาเยี่ยมชมอย่างต่อเนื่อง

ในยุคดิจิทัลที่สื่อสังคมออนไลน์เข้ามามีบทบาทสำคัญ นักท่องเที่ยวสามารถแบ่งปันประสบการณ์ แสดงความคิดเห็น และแลกเปลี่ยนข้อมูลได้อย่างรวดเร็ว ข้อมูลเหล่านี้สะท้อนถึงพฤติกรรม ความต้องการ และความคิดเห็นของนักท่องเที่ยว ซึ่งเป็นประโยชน์ต่อการพัฒนาบริการและการวางกลยุทธ์ทางการตลาด การวิเคราะห์ข้อมูลเหล่านี้จึงมีความสำคัญอย่างยิ่งเพื่อให้สามารถตอบสนองต่อความต้องการของตลาดและกลุ่มเป้าหมายได้อย่างมีประสิทธิภาพ

ปัจจุบันมีงานวิจัยหลายงานที่นำเทคนิคทางวิทยาการข้อมูล (Data Science) และปัญญาประดิษฐ์ (Artificial Intelligence) เช่น เครื่องมือทางสถิติ (Statistics) การเรียนรู้ของเครื่อง (Machine Learning) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) มาประยุกต์ใช้เพื่อสกัดข้อมูลเชิงลึกจาก

สื่อสังคมออนไลน์ ตัวอย่างงานวิจัย ได้แก่ Huahai Yang และ Yunyao Li (2013) ได้พัฒนาแบบจำลองการวิเคราะห์ความต้องการของผู้ใช้งานสื่อสังคมออนไลน์ โดยใช้ NLP ร่วมกับเครื่องมือวัดทางจิตวิทยา Amrutha Benny และ Mintu Philip (2015) ใช้ NLP ในการสกัดคำสำคัญและตรวจสอบความสัมพันธ์ของข้อความบนทวิตเตอร์ เพื่อวิเคราะห์และสรุปหัวข้อ พิเชษฎ์ จุลรอด และ ดินสอ บุญมา (2566) วิเคราะห์ขั้วอารมณ์ (Sentiment) จากรีวิวร้านอาหารและร้านค้าบนเว็บไซต์ Wongnai ด้วยการเรียนรู้ของเครื่อง และแสดงผลในรูปแบบกราฟและ Word Cloud ศิวกร ดีชูศร และคณะ (2566) พัฒนาเว็บไซต์รวบรวมข้อมูลสัตว์เลี้ยงที่ต้องการการอุปถัมภ์ โดยดึงข้อมูลจากทวิตเตอร์และนำเสนอในรูปแบบแดชบอร์ด และปภัสสนันท์ ชมจันทึก (2561) ศึกษาพฤติกรรมการเปิดรับทัศนคติและกระบวนการตัดสินใจท่องเที่ยวผ่านแฮชแท็กบนทวิตเตอร์ พบว่าทวิตเตอร์เป็นแหล่งข้อมูลสำคัญที่สามารถชี้ให้เห็นถึงพฤติกรรมและการตัดสินใจของนักท่องเที่ยว จากตัวอย่างงานวิจัยข้างต้น สื่อสังคมออนไลน์ เช่น ทวิตเตอร์ ได้แสดงให้เห็นถึงศักยภาพในฐานะแหล่งข้อมูลสำคัญที่สามารถนำมาใช้วิเคราะห์ความคิดเห็น ความต้องการ และพฤติกรรมของผู้ใช้งานได้อย่างมีประสิทธิภาพ

บทความนี้นำเสนอการพัฒนาระบบเพื่อรวบรวมและวิเคราะห์ความคิดเห็นของนักท่องเที่ยวบนแพลตฟอร์ม X (ทวิตเตอร์) ที่เกี่ยวข้องกับจังหวัดภูเก็ต โดยใช้เทคนิคการเรียนรู้ของเครื่องและการประมวลผลภาษาธรรมชาติในการสกัดข้อมูล เช่น แฮชแท็ก หัวข้อ และขั้วอารมณ์ และนำเสนอข้อมูลในรูปแบบแดชบอร์ดปฏิสัมพันธ์ที่เข้าใจง่าย ระบบนี้มีเป้าหมายเพื่อสนับสนุนหน่วยงานรัฐและผู้ประกอบการในจังหวัดภูเก็ตในการวางแผนและพัฒนาบริการให้ตอบสนองความต้องการของนักท่องเที่ยวได้ดียิ่งขึ้น ตลอดจนช่วยในการวางกลยุทธ์ทางการตลาดที่ตรงกลุ่มเป้าหมาย และเพิ่มขีดความสามารถในการตัดสินใจเชิงข้อมูล

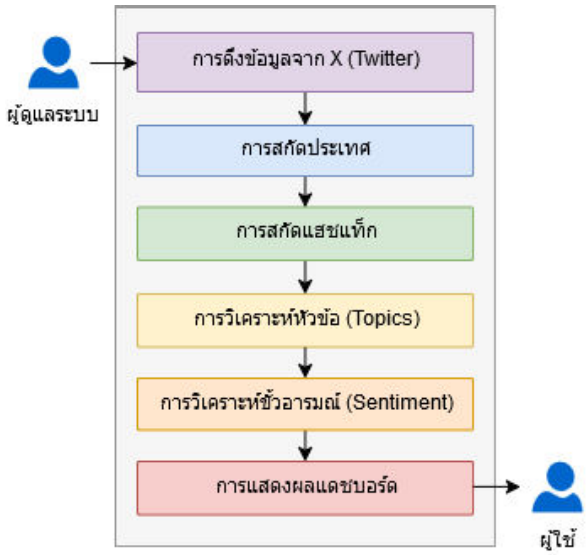
## วิธีการดำเนินการวิจัย

ผู้วิจัยใช้กระบวนการวิศวกรรมซอฟต์แวร์ในการทำวิจัย ซึ่งมีขั้นตอนหลัก 3 ขั้นตอน ได้แก่ การวิเคราะห์และออกแบบระบบ การพัฒนาระบบ และการทดสอบระบบ แต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

### 1. การวิเคราะห์และออกแบบระบบ

ผู้วิจัยเริ่มต้นการพัฒนาระบบด้วยการวิเคราะห์องค์ประกอบและการทำงานของระบบ โดยแบ่งขั้นตอนการทำงานออกเป็น 6 ส่วนตามภาพที่ 1 ส่วนแรกเป็นการดึงข้อมูลทวีตผ่าน API ของแพลตฟอร์ม X เพื่อให้ได้ข้อความที่เกี่ยวข้อง จากนั้นจึงสกัดข้อมูลประเทศของเจ้าของทวีต เนื่องจากข้อมูลดังกล่าวจะช่วยให้เข้าใจธรรมชาติและความต้องการของนักท่องเที่ยวแต่ละประเทศได้ดียิ่งขึ้น ถัดมาจึงทำการสกัดแฮชแท็ก (Hashtag) ซึ่งสะท้อนถึงสิ่งที่ผู้โพสต์ให้ความสนใจหรือเกี่ยวข้องกับโพสต์นั้น ๆ ขั้นตอนถัดไปเป็นการวิเคราะห์ว่าแต่ละทวีตกล่าวถึงหัวข้อใด (Topic Modeling) หลังจากนั้นจึงวิเคราะห์ขั้วอารมณ์ (Sentiment) ของทวีตเพื่อประเมิน

แนวโน้มเชิงบวก เชิงลบ หรือเป็นกลาง สุดท้ายจะนำข้อมูลทั้งหมดที่ได้มาสร้างเป็นแดชบอร์ด (Dashboard) เพื่อให้ใช้งานและทำความเข้าใจได้อย่างสะดวกยิ่งขึ้น




**ภาพที่ 1** ขั้นตอนการทำงานของระบบ

**2. การพัฒนาระบบ**

ผู้วิจัยใช้ภาษา Python บน Google Colab ในการเขียนโปรแกรมและใช้ Google Looker Studio ในการสร้างแดชบอร์ด โดยมีรายละเอียดดังต่อไปนี้

**2.1 การดึงข้อมูลทวีต**

การดึงข้อมูลเริ่มจากการสมัครบัญชีผู้พัฒนา (Developer Account) เพื่อขอรับ API Key และ Access Token สำหรับเข้าถึง API ของแพลตฟอร์ม X จากนั้นจึงเขียนโปรแกรมดึงข้อมูลโดยใช้ชุดคำค้น (Keywords) ที่เกี่ยวข้องกับการท่องเที่ยวของประเทศไทย (ภาษาอังกฤษ) เช่น “Phuket”, “Thailand”, “Travel”, “Tourism”, และ “Taxi” ในการดึงข้อมูลทวีต ข้อมูลที่ได้ประกอบด้วยวันเวลาที่เผยแพร่ทวีต เนื้อหาทวีต ลิงก์ทวีต และข้อมูลของผู้โพสต์ (User Profile) สุดท้ายทำการลบข้อมูลที่ซ้ำกัน (Duplicate Data) เพื่อเตรียมความพร้อมสำหรับขั้นตอนการวิเคราะห์ต่อไป

**2.2 การสกัดประเทศ**

ผู้วิจัยใช้ข้อมูลประเทศของผู้โพสต์ทวีตจากตำแหน่งผู้ใช้ (User Location) ซึ่งเป็นส่วนหนึ่งของข้อมูลผู้โพสต์ที่ดึงมา อย่างไรก็ตาม ผู้ใช้แพลตฟอร์ม X หลายรายมักระบุข้อมูลที่อยู่ไม่ชัดเจน โดยอาจใส่เพียงระดับเมือง หรือใช้ตัวย่อของประเทศ เช่น “Delhi”, “Colorado, USA” หรือ “JPN” ทำให้ไม่สามารถระบุชื่อประเทศได้

โดยตรง เพื่อแก้ไขปัญหานี้ ผู้วิจัยจึงนำฐานข้อมูล World Cities Database (https://simplemaps.com/data/world-cities) ซึ่งเป็นไฟล์ CSV ที่ประกอบด้วยชื่อเมือง ชื่อประเทศ และตัวย่อของประเทศทั้งแบบ 2 และ 3 ตัวอักษร มาประยุกต์ใช้ในการวิเคราะห์ โดยใช้ไลบรารี NLTK (https://www.nltk.org/) ช่วยในกระบวนการตัดคำ (Tokenization) ขั้นตอนการทำงานสามารถอธิบายเป็นรหัสเทียม (Pseudocode) ได้ดังภาพที่ 2

**Function:** `extract_country(user_location)`

**Input:**
- `user_location` (String): The location string provided by the user.

**Output:**
- `country_name` (String): The corresponding country name, if found; otherwise, returns "Other".

**Algorithm:**
**1. Load Data:**
`world_cities = read("world-cities.csv")`
(`world_cities` is a table with columns: `country`, `country2char`, `country3char` and `city`.)

**2. Tokenize User Location:**
Split `user_location` into individual tokens using `word_tokenize(user_location)`.

**3. Iterate Through Tokens:**
For each token `t` in the tokenized `user_location`:
- If `t` matches any value in the `country` column of `world_cities` then return `t`.
- If `t` matches any value in the `city` column of `world_cities` then return the corresponding value from the `country` column.
- If `t` matches any value in the `country2char`column of `world_cities` then return the corresponding value from the `country` column.
- If `t` matches any value in the `country3char`column of `world_cities` then return the corresponding value from the `country` column.

Return "Other".

**ภาพที่ 2** รหัสเทียมการสกัดประเทศ

### 2.3 การสกัดแฮชแท็ก

ผู้วิจัยได้ใช้ Regular Expression (regex) เพื่อสกัดแฮชแท็ก (Hashtags) จากข้อความทวีต เนื่องจาก regex เป็นเครื่องมือที่มีประสิทธิภาพและเหมาะสมสำหรับการจัดการกับข้อความที่มีรูปแบบเฉพาะ เช่น แฮชแท็ก ซึ่งมีลักษณะเริ่มต้นด้วยเครื่องหมาย # และตามด้วยตัวอักษร ตัวเลข หรือเครื่องหมายขีดล่าง _ โดยไม่มีการเว้นวรรคระหว่างข้อความ การใช้ regex ในการสกัดแฮชแท็กสามารถจัดการได้ง่าย รวดเร็ว และรองรับการประมวลผลข้อมูลจำนวนมาก รหัสเทียมสำหรับกระบวนการนี้แสดงดังภาพที่ 3

**Function:** `extract_hashtags(tweet)`

**Input:**
- `tweet` (String): The input tweet text.

**Output:**
- `hashtag_list` (List): A list of hashtags extracted from the tweet.

**Algorithm:**

**1. Initialize an Empty List:**
Create an empty list `hashtag_list` to store extracted hashtags.

**2. Tokenize the Tweet:**
Split `tweet` into individual tokens using `word_tokenize(tweet)`.

**3. Iterate Through Tokens:**
For each token `t` in the tokenized `tweet`:
- If `t` matches the regular expression pattern for a hashtag (i.e., #\w+) then append `t` to `hashtag_list`.

**4. Return the result:**
Return `hashtag_list`.

**ภาพที่ 3** รหัสเทียมการสกัดแฮชแท็ก

### 2.4 การวิเคราะห์หัวข้อ

การวิเคราะห์หัวข้อ (Topic Modeling) ของทวีตเริ่มต้นด้วยการทำความสะอาดข้อมูล โดยใช้ Regular Expression เพื่อกำจัดข้อความที่ไม่จำเป็น เช่น URL, ที่อยู่อีเมล, ข้อความที่อ้างอิงถึงผู้ใช้อื่น (ข้อความที่มีเครื่องหมาย @), และ Emojicons ออกจากข้อความ ทั้งนี้ผู้วิจัยเลือกคงแฮชแท็กไว้ เนื่องจากแฮชแท็กมีความสำคัญต่อการวิเคราะห์หัวข้อ จากนั้น ผู้วิจัยนำทวีตที่ผ่านการทำความสะอาดแล้วมาวิเคราะห์หัวข้อ (Topic Modeling) โดยใช้ไลบรารี BERTopic (https://maartengr.github.io/BERTopic/index.html) โดยผู้วิจัยตั้งค่า BERTopic ให้ใช้ Sentence Embedding Model แบบ all-MiniLM-L6-v2 ในการแปลงข้อความแต่ละทวีตเป็นเวกเตอร์ (Vectorization) ซึ่งเป็นการเตรียมข้อมูลสำหรับการจัดกลุ่มทวีต (Clustering) หลังจากจัดกลุ่มทวีตสำเร็จ BERTopic จะสกัดคำสำคัญของแต่ละกลุ่มโดยใช้วิธี Term Frequency–Inverse Document Frequency (TF-IDF) เพื่อระบุคำที่มีความสำคัญในบริบทของแต่ละกลุ่ม จากนั้นผู้วิจัยกำหนดให้ BERTopic เรียกใช้ ChatGPT (ผ่าน API) เพื่อสรุปหัวข้อคำสำคัญของแต่ละกลุ่มในรูปแบบที่ชัดเจนและเข้าใจง่าย รหัสเทียมสำหรับกระบวนการนี้แสดงดังภาพที่ 4

**Function:** `topic_modeling(tweets)`

**Input:**
- `tweets` (List): A list containing all tweet texts..

**Output:**
- `topics` (List): A list of identified topics corresponding to the input tweets.

**Algorithm:**
**1. Clean Data:**
For each `tweet` in `tweets`
- Remove URLs, email addresses, user mentions, and emojicons from `tweet`

**2. Generate Topics:**
- Use the BERTopic model to analyze the topics for the cleaned `tweets`.
- Store the resulting list of topics in the variable `topics`.

**3. Return the result:**
Return `topics`.

**ภาพที่ 4** รหัสเทียมการวิเคราะห์หัวข้อ

**Function:** `sentiment_analysis(tweet)`

**Input:**
- `tweet` (String): The input tweet text.

**Output:**
- `sentiment` (String): The sentiment classification (`Positive`, `Neutral`, `Negative`) of the input tweet.

**Algorithm:**
**1. Clean Data:**
Remove URLs, email addresses, user mentions, and hashtags from `tweet`

**2. Initialize Sentiment Analysis Model:**
- Download `movie_review` corpus from the NLTK library.
- Create a `pattern_analyzer` model using the loaded corpus.

**3. Classify Sentiment:**
- Use `pattern_analyzer` to analyze the sentiment of the cleaned `tweet`.
- Store the result in the variable `sentiment`.

**4. Return the result:**
Return `sentiment`.

**ภาพที่ 5** รหัสเทียมการวิเคราะห์ขั้วอารมณ์

**2.5 การวิเคราะห์ขั้วอารมณ์**

ขั้นตอนนี้เริ่มด้วยการทำความสะอาดข้อมูลทวีต โดยใช้ Regular Expression เพื่อกำจัดข้อความที่ไม่จำเป็น เช่น URL, ที่อยู่อีเมล, ข้อความที่อ้างอิงถึงผู้ใช้อื่น, และแฮชแท็กออกไป ในขณะที่คง Emojicons ซึ่งเป็นไอคอนแสดงความรู้สึกไว้ เนื่องจากมีความสำคัญต่อการวิเคราะห์อารมณ์ หลังจากการทำความสะอาดข้อมูล ผู้วิจัยได้นำทวีตที่ผ่านกระบวนการดังกล่าวมาวิเคราะห์ขั้วอารมณ์ (Sentiment Analysis) โดยใช้ไลบรารี TextBlob (https://textblob.readthedocs.io/en/dev/) ซึ่งทำงานร่วมกับ Movie Review Corpus ของไลบรารี NLTK ในการสอนแบบจำลองวิเคราะห์อารมณ์แบบ PatternAnalyzer โดย TexBlob จะแบ่งขั้วอารมณ์ออกเป็น 3 แบบ คือ Positive, Neutral และ Negative รหัสเทียมสำหรับกระบวนการนี้แสดงดังภาพที่ 5

**2.6 การแสดงผลแดชบอร์ด**

การทำ Data Visualization เป็นกระบวนการที่มีความสำคัญอย่างยิ่ง เนื่องจากช่วยแปลงข้อมูลที่ผ่านการจัดเตรียมมาแล้วให้เป็นรูปแบบที่เข้าใจง่าย สร้างความดึงดูดใจ และสนับสนุนการวิเคราะห์ข้อมูลเพิ่มเติมได้อย่างมีประสิทธิภาพและรวดเร็ว สำหรับงานวิจัยนี้ ผู้วิจัยได้นำข้อมูลที่สกัดและวิเคราะห์จากขั้นตอนก่อนหน้ามาบันทึกในรูปแบบไฟล์ CSV บน Google Drive จากนั้นจึงใช้ Google Looker Studio สร้างและแสดงผลข้อมูลในรูปแบบแดชบอร์ดแบบปฏิสัมพันธ์ ซึ่งช่วยให้ผู้ใช้งานสามารถมองเห็นภาพรวมของข้อมูลและเชื่อมโยงความสัมพันธ์ต่าง ๆ ได้อย่างสะดวก

**3. การทดสอบระบบ**

ผู้วิจัยได้ออกแบบการทดสอบเพื่อประเมินความถูกต้องและประสิทธิภาพของระบบในแต่ละส่วน โดยประกอบด้วยการทดสอบดังต่อไปนี้ (ผลการทดสอบระบบในแต่ละส่วนจะถูกสรุปและนำเสนอในหัวข้อถัดไป)

1. *ทดสอบการดึงข้อมูล*: ตรวจสอบว่าระบบสามารถดึงข้อมูลจากแหล่งที่มาโดยใช้คำสำคัญที่กำหนดได้อย่างถูกต้อง
2. *ทดสอบการสกัดประเทศ*: ตรวจสอบความสามารถของระบบในการระบุประเทศจากข้อมูลตำแหน่งผู้ใช้งาน (User Location) ตามขั้นตอนที่ออกแบบไว้
3. *ทดสอบการสกัดแฮชแท็ก*: ตรวจสอบความถูกต้องของระบบในการแยกและดึงแฮชแท็กจากข้อความทวีต
4. *ทดสอบการวิเคราะห์หัวข้อ*: ประเมินความถูกต้องของระบบในการจัดกลุ่มและสรุปหัวข้อ
5. *ทดสอบการวิเคราะห์ขั้วอารมณ์*: ตรวจสอบความแม่นยำของระบบในการจัดหมวดหมู่อารมณ์ของข้อความ (Positive, Neutral, Negative)

6. *ทดสอบการแสดงผลบนแดชบอร์ด*: ทดสอบการอัปเดตและแสดงผลข้อมูลบนแดชบอร์ดแบบอัตโนมัติ รวมถึงตรวจสอบความสอดคล้องของข้อมูลที่แสดง

**ผลการวิจัย**

ผู้วิจัยได้พัฒนาระบบจนเสร็จสมบูรณ์ และดำเนินการทดสอบระบบตามขั้นตอนที่ออกแบบไว้ในหัวข้อก่อนหน้า โดยผลการวิจัยมีรายละเอียดดังนี้

ในการทดสอบการดึงข้อมูล ผู้วิจัยทำการทดลองใช้คำค้นที่เกี่ยวข้องกับการท่องเที่ยวที่หลากหลายเพื่อทดสอบการดึงข้อมูลจากแพลตฟอร์ม X ผลการทดสอบพบว่าระบบสามารถดึงข้อมูลได้ตรงตามเงื่อนไขที่กำหนด อย่างไรก็ตาม กระบวนการดึงข้อมูลต้องมีข้อจำกัด คือ ระบบสามารถดึงข้อมูลได้ครั้งละจำนวนจำกัด ซึ่งเป็นข้อกำหนดที่เกิดจากการใช้งาน API ฟรี หากต้องการดึงข้อมูลในปริมาณมากขึ้นจำเป็นต้องเสียค่าใช้จ่ายมากขึ้น

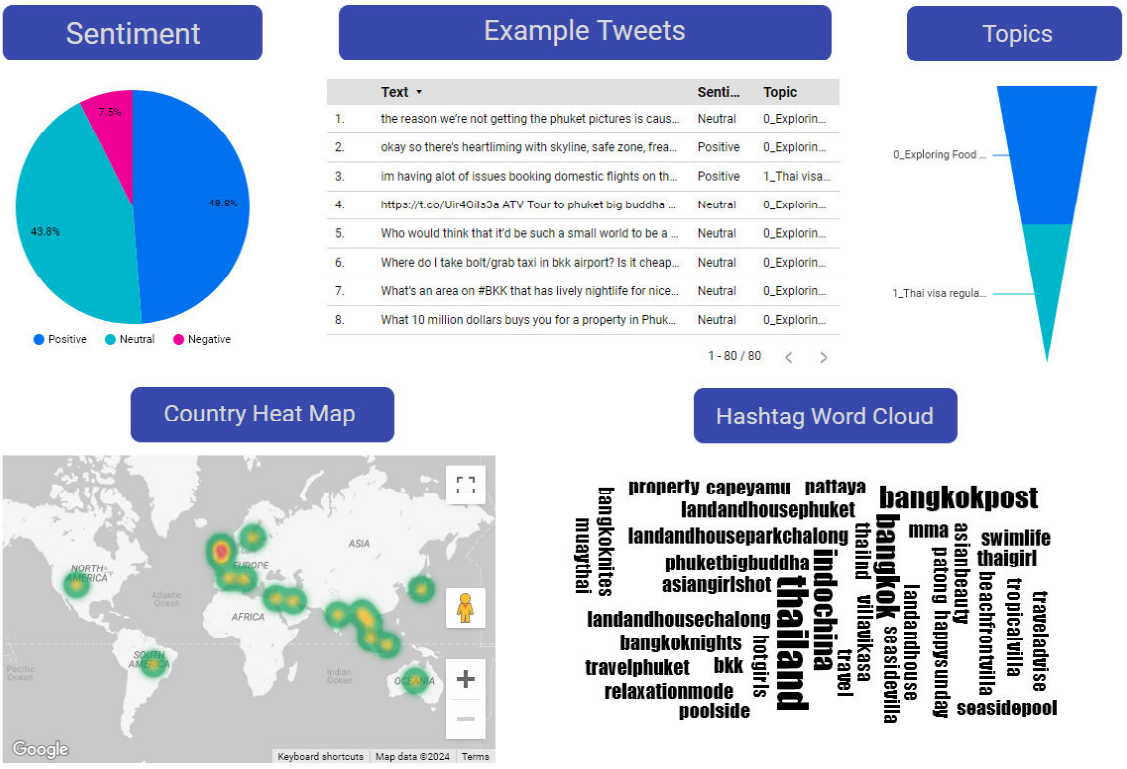
สำหรับการทดสอบการสกัดประเทศ ผู้วิจัยได้สุ่มเลือกทวีตตัวอย่างจำนวน 80 ทวีต เพื่อประเมินความถูกต้องของการทำงานของโปรแกรม จากนั้นผู้วิจัยตรวจสอบตำแหน่งผู้ใช้ (User Location) ของทั้ง 80 ทวีต เทียบกับผลลัพธ์ที่โปรแกรมสกัดได้ ผลการทดสอบพบว่าโปรแกรมสามารถระบุประเทศได้อย่างถูกต้องตรงกับข้อมูลต้นฉบับในทุกกรณี

ผู้วิจัยได้ทำการทดสอบความถูกต้องของโปรแกรมสกัดแฮชแท็ก โดยเปรียบเทียบแฮชแท็กที่ปรากฏอยู่จริงในทวีตตัวอย่าง 80 ทวีตที่สุ่มเลือกมาก่อนหน้านี้กับผลลัพธ์ที่โปรแกรมสกัดได้ ผลการทดสอบพบว่าโปรแกรมสามารถสกัดแฮชแท็กได้อย่างถูกต้องและครบถ้วนทุกกรณี

ในการทดสอบความสามารถของโปรแกรมในการวิเคราะห์หัวข้อ ผู้วิจัยได้ทำการทดลองโดยให้โปรแกรมวิเคราะห์หัวข้อของทวีตตัวอย่าง 80 ทวีต โดยผลลัพธ์ที่ได้คือโปรแกรมสามารถสกัดหัวข้อและจัดกลุ่มทวีตออกเป็น 2 กลุ่ม คือ กลุ่มที่ 1 Exploring_Food_Markets_and_Beach_Life_in_Phuket และ กลุ่มที่ 2 Thai_Visa_Regulations_and_Travel_Inquiries หลังจากนั้น ผู้วิจัยได้ตรวจสอบเนื้อหาของแต่ละทวีตเพื่อประเมินความถูกต้องของการจัดกลุ่ม ผลการตรวจสอบพบว่ามีทวีตจำนวน 69 ทวีตที่ถูกจัดกลุ่มอย่างถูกต้อง คิดเป็นร้อยละ 86.25

ในทำนองเดียวกันกับการทดสอบการวิเคราะห์หัวข้อ ผู้วิจัยให้โปรแกรมวิเคราะห์ขั้วอารมณ์ของทวีตตัวอย่างทั้ง 80 ทวีต จากนั้นจึงตรวจสอบเนื้อหาของแต่ละทวีตเทียบกับผลลัพธ์ที่ได้จากโปรแกรมเพื่อประเมินความถูกต้องของการวิเคราะห์ ผลการตรวจสอบพบว่าโปรแกรมวิเคราะห์ถูกต้องจำนวน 61 ทวีต คิดเป็นร้อยละ 76.25

จากการทดสอบระบบแสดงผลแดชบอร์ด พบว่าระบบสามารถแสดงข้อมูลที่ได้จากการสกัดและการวิเคราะห์ในรูปแบบที่ถูกต้อง ครบถ้วน รวมทั้งทำงานแบบปฏิสัมพันธ์และอัตโนมัติ นอกจากนี้ข้อมูลที่แสดงในแดชบอร์ดช่วยให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้อย่างสะดวกและเข้าใจง่าย ตัวอย่างการแสดงผลของแดชบอร์ดสามารถดูได้ดัง ภาพที่ 6




**ภาพที่ 6** ตัวอย่างแดชบอร์ดนำเสนอข้อมูลทวีต

## วิจารณ์ผล

จากผลการวิจัยพบว่า ระบบที่พัฒนาขึ้นสามารถทำงานได้ตรงตามข้อกำหนดซอฟต์แวร์ (Software Specification) ที่กำหนดไว้ โดยทุกกระบวนการตั้งแต่การดึงข้อมูล การสกัดประเทศ การสกัดแฮชแท็ก และการแสดงผลแดชบอร์ดสามารถทำงานได้อย่างถูกต้องและสมบูรณ์

ในส่วนของการวิเคราะห์หัวข้อ พบว่าโปรแกรมสามารถจัดกลุ่มหัวข้อของทวีตตัวอย่างได้ถูกต้องในอัตราความถูกต้อง ร้อยละ 86.25 แต่พบว่ากลุ่มที่ 1 (Exploring_Food_Markets_and_Beach_Life_in_Phuket) มีอัตราความผิดพลาดสูงกว่ากลุ่มที่ 2 (Thai_Visa_Regulations_and_Travel_Inquiries) ซึ่งอาจเกิดจากลักษณะของข้อมูลในกลุ่มที่ 1 ที่มีความหลากหลายและคลุมเครือ ขณะที่กลุ่มที่ 2 มีหัวข้อที่เฉพาะเจาะจงมากกว่า ทำให้

การวิเคราะห์ง่ายและแม่นยำขึ้น ดังนั้นการเพิ่มจำนวนกลุ่มเพื่อให้เนื้อหาในแต่ละกลุ่มมีความชัดเจนและเฉพาะเจาะจงมากขึ้น อาจช่วยลดอัตราความผิดพลาดได้

สำหรับการวิเคราะห์ขั้วอารมณ์พบว่ามีความถูกต้องในอัตราร้อยละ 76.25 ซึ่งต่ำกว่าการวิเคราะห์หัวข้อ สาเหตุหนึ่งอาจมาจากลักษณะข้อความในทวีตที่มีความสั้น กระชับ และมักใช้ภาษาที่ไม่เป็นทางการ ซึ่งอาจทำให้โมเดลสับสนในการวิเคราะห์ขั้วอารมณ์ การเพิ่มข้อมูลชุดสอนที่มีลักษณะใกล้เคียงกับทวีตหรือข้อมูลจากโซเชียลมีเดียอื่น ๆ ให้มากขึ้น อาจช่วยลดข้อผิดพลาดได้

**สรุปผล**

บทความวิจัยนี้นำเสนอการพัฒนาระบบวิเคราะห์ความคิดเห็นของนักท่องเที่ยวบนแพลตฟอร์ม X โดยใช้กรณีศึกษาจังหวัดภูเก็ต ระบบถูกออกแบบและพัฒนาเพื่อรวบรวม สกัด และวิเคราะห์ข้อมูลจากข้อความทวีต เพื่อให้ได้ข้อมูลเชิงลึกที่สามารถนำไปใช้สนับสนุนการวางแผนการท่องเที่ยวและการตัดสินใจเชิงนโยบาย

ผู้วิจัยใช้ภาษา Python ในการพัฒนา และใช้ Google Looker Studio ในการแสดงผลแดชบอร์ด ผลการวิจัยแสดงให้เห็นว่าระบบสามารถทำงานได้ตรงตามข้อกำหนดซอฟต์แวร์ (Software Specification) ที่กำหนดไว้ โดยมีความถูกต้องในการดึงข้อมูล การสกัดประเทศ และการสกัดแฮชแท็ก นอกจากนี้ การวิเคราะห์หัวข้อสามารถจัดกลุ่มทวีตได้ในอัตราความถูกต้อง ร้อยละ 86.25 และการวิเคราะห์ขั้วอารมณ์มีความถูกต้อง ร้อยละ 76.25

ระบบนี้สามารถนำไปประยุกต์ใช้ในอุตสาหกรรมการท่องเที่ยว โดยช่วยให้ผู้เกี่ยวข้อง เช่น หน่วยงานรัฐ ผู้ประกอบการ และนักวิจัย สามารถเข้าใจความคิดเห็นและพฤติกรรมของนักท่องเที่ยวได้ดีขึ้น ซึ่งจะนำไปสู่การพัฒนาแผนการตลาด การบริการ และนโยบายที่ตอบสนองความต้องการของนักท่องเที่ยวได้อย่างมีประสิทธิภาพ

แนวทางการปรับปรุงระบบในอนาคต ได้แก่ การเพิ่มคำค้นให้ครอบคลุมหัวข้อที่หลากหลายยิ่งขึ้น เพื่อให้ได้ข้อมูลที่สมบูรณ์มากขึ้น การขยายการใช้งานระบบไปยัง Web Forums และ Social Media อื่น ๆ เพื่อเพิ่มแหล่งข้อมูล และการพัฒนาระบบให้สามารถวิเคราะห์ข้อความในภาษาต่าง ๆ นอกเหนือจากภาษาอังกฤษ เพื่อให้รองรับกลุ่มเป้าหมายในประเทศอื่น ๆ ได้อย่างครอบคลุม

**เอกสารอ้างอิง**

ปภัสนันท์ ชมจันทึก. 2561. พฤติกรรมการเปิดรับ ทัศนคติ และกระบวนการตัดสินใจท่องเที่ยวตามแฮชแท็กท่องเที่ยวบนทวิตเตอร์. วิทยานิพนธ์ ปริญญาวารสารศาสตรมหาบัณฑิต มหาวิทยาลัยธรรมศาสตร์. กรุงเทพฯ. https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:164800

พิเชษฎ์ จุลรอด, และดินสอ บุญมา. 2566. การวิเคราะห์ความรู้สึกของผู้บริโภคและใช้บริการร้านอาหารและร้านค้าในมหาวิทยาลัยเชียงใหม่. วารสารสารสนเทศศาสตร์ มหาวิทยาลัยเชียงใหม่. 1(1): 73-83. https://lis.human.cmu.ac.th/Journal/file/CMU_ISJ05-PRINT.pdf

ศิวกร ดีชูศร, ศิรภัสสร สุขขนาน, และ ณสิทธิ์ เหล่าเส็น. 2566. การพัฒนาเว็บไซต์เผยแพร่ทวิตสัตว์เลี้ยงที่ต้องการการอุปถัมภ์. น. 233-237. ใน: การประชุมวิชาการวิศวกรรมศาสตร์และเทคโนโลยี มทร.พระนคร ครั้งที่ 7. กรุงเทพฯ.

Amrutha, B., and Mintu, P. 2014. Keyword Based Tweet Extraction and Detection of Related. pp. 364371. In: The International Conference on Information and Communication Technologies. https://www.sciencedirect.com/science/article/pii/S1877050915000964

Huahai, Y. , and Yunyao, L. 2013. Identifying User Needs from Social Media. Available: https://www.researchgate.net/publication/264782517_Identifying_User_Needs_from_Social_Media. Accessed Feb.18, 2025.