ระบบจำแนกทัศนคติการท่องเที่ยวโดยเทคนิคนาอีพเบย์

TOURISM ATTITUDE CLASSIFICATION SYSTEM USING NAIVE BAYES TECHNIQUE

ธีระพล ลิ้มศรัทธา [1] ทรงพล นคเรศเรืองศักดิ์ [2*] สุทธิลักษณ์ ชุนประวัติ [3]

Theerapol Limsatta [1] Songpon Nakharacruangsak [2*] Suttilug Choonprawat [3]

[1, 2] คณะวิทยาศาสตร์และเทคโนโลยี วิทยาลัยเซาธ์อีสท์บางกอก

[3] คณวิศวกรรมศาสตร์และเทคโนโลยี วิทยาลัยเทคโนโลยีสยาม

*Corresponding author, E-mail: songpon@southeast.ac.th

Received : 23 April 2020
Revised : 20 Jun 2020
Accepted : 28 Jun 2020

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ พัฒนาโปรแกรมระบบจำแนกทัศนคติการท่องเที่ยวในขั้วบวกและลบ โดยใช้เทคนิคนาอีฟเบย์ และประเมินประสิทธิภาพของระบบทัศนคติการท่องเที่ยวโดยใช้เมตริกของความแม่นยำ การระลึกได้ ค่าทดสอบเอฟ (F-measure) และค่าความถูกต้อง โดยใช้ข้อมูลวิจารณ์การท่องเที่ยวในประเทศจาก Google Map และจากเว็บไซต์ทริปแอดไวเซอร์ ของนักท่องเที่ยวไทย ผลการพัฒนาโปรแกรมระบบจำแนกทัศนคติการท่องเที่ยวพบว่าโปรแกรมสามารถวิเคราะห์เพื่อจำแนกความรู้สึกจากข้อความจากข้อความทั่วไปได้ แต่อาจจะมีปัญหากับประโยคที่มีความหมายกำกวม ไม่แน่ชัด เช่น ไม่ถึงกับไม่ดี ไม่ถึงกับไม่แย่ เป็นปัญหาต่อการตัดคำและความเข้าใจของโปรแกรม ทำให้การให้คะแนนผิดพลาดไป ผลการประเมินประสิทธิภาพของระบบจำแนกทัศนคติการท่องเที่ยวด้วยเอกสารทดสอบที่แตกต่างเอกสารที่ใช้ในการสร้างโมเดล จำนวน 120 เอกสาร ให้ค่าการระลึก ค่าความแม่นยำ และค่าทดสอบเอฟ มีค่า 80, 100, และ 120 ตามลำดับ ได้ผลดังเฉลี่ยจากทั้ง 3 โมเดล ดังนี้การระลึก มีค่าเฉลี่ยที่ 0.86 ความแม่นยำมีค่าเฉลี่ยที่ 0.80 ค่าทดสอบเอฟ มีค่าเฉลี่ยที่ 0.80 และมีค่าความถูกต้อง 0.80

คำสำคัญ : ทัศนคติการท่องเที่ยวม, นาอีพเบย์

Abstract

The purposes of this research study are to develop the program for distinguishing the positive and negative tourism attitudes by using the Naive Bayes technique and to evaluate the efficiency of the tourism attitude system by using accuracy matrices, recognitions, F-measure and correctness. The data were Thai tourists' reviews about domestic tourism from Google Map and TripAdvisor website. It was found that the program could analyze and distinguish the feelings from general texts. However, there might be problems about sentences with unclear sentences and phrases such as not really bad. These texts were problematic for cutting words and interpreting by the program. Consequently, scores are incorrect. By evaluating the efficiency of the system with 120 different documents for creating models; the recognition, accuracy and F-measure values are 80, 100 and 120, respectively. The averages from the three models are as follows. The average recognition is 0.86. The average accuracy is 0.80. The average F-measure is 0.80. The average correctness is 0.80.

Keyword : tourism attitudes, naive bayes

1. บทนำ

อินเทอร์เน็ตได้กลายเป็นช่องทางสำหรับการติดต่อสื่อสารสำคัญสำหรับสังคมยุคใหม่ ทำให้ส่วนใหญ่คุ้นเคยกับการใช้ชีวิตในสังคมออนไลน์ในรูปแบบต่าง ๆ เช่น เฟสบุ๊ค (Facebook) ยูทูป (YouTube) เป็นต้น เว็บไซต์ต่าง ๆ เป็นแหล่งแลกเปลี่ยนและแสดงความคิดเห็น ที่สำคัญมากในปัจจุบัน รวมทั้งยังเป็นแหล่งสืบค้นข้อมูลที่มีผู้ใช้จำนวนมาก และผู้ที่ใช้อินเทอร์เน็ตเป็นผู้ที่สร้างข้อมูลในหลายรูปแบบไม่ว่าจะเป็น ข้อความ หรือรูปภาพ ต่างทำให้มีข้อมูลอย่างมากมายมหาศาลบนโลกอินเทอร์เน็ต ที่สามารถนำมาใช้ในการวิเคราะห์และประมวลผลโดยคอมพิวเตอร์ได้ ซึ่งการวิเคราะห์ข้อมูลขนาดใหญ่จะเรียกว่า บิ๊กดาต้า ซึ่งข้อมูลเหล่านี้สามารถนำไปใช้ประโยชน์ในด้านการวิเคราะห์ความรู้สึก (sentiment analysis) โดยอาศัยเทคนิคต่าง ๆ ของการประมวลภาษาธรรมชาติ (natural language processing) ในการระบุขั้วความรู้สึก (sentiment polarity) ของข้อความซึ่งมักจำแนกออกเป็น 2 ขั้ว คือขั้วบวก (positive) และขั้วลบ (negative)

งานวิจัยนี้สนใจศึกษาเกี่ยวกับการวิเคราะห์ความรู้สึกเนื่องจากการวิเคราะห์ความรู้สึกนี้ เป็นประโยชน์มากต่อการประกอบการในยุคปัจจุบัน ซึ่งสินค้าหรือบริการหนึ่งๆ มักมีผู้ใช้จำนวนมาก ทำให้มีปริมาณข้อมูลการวิจารณ์เพิ่มมากขึ้นบนเว็บไซต์ต่าง ๆ ตามไปด้วย การวิเคราะห์ความรู้สึกนี้ จึงเป็นเครื่องมือที่สามารถใช้สรุปความคิดเห็นที่นักท่องเที่ยวมีต่อสถานที่นั้น ๆ ได้โดยอัตโนมัติ ซึ่งจะเป็นประโยชน์ต่อทั้งนักท่องเที่ยวและผู้ประกอบการนั้น ๆ กล่าวคือ หากนักท่องเที่ยวอยากรู้ว่าสถานที่ที่ตนสนใจมีคุณภาพดีมากแค่ไหนก่อนจะตัดสินใจไปท่องเที่ยว ก็สามารถค้นหาข้อมูลเกี่ยวกับสถานที่นั้น ๆ บนอินเทอร์เน็ตได้ และเมื่อไปท่องเที่ยวแล้ว ก็อาจเขียนคำวิจารณ์ของสถานที่นั้นไว้สำหรับผู้ที่ต้องการท่องเที่ยวคนอื่น ๆ ส่วนสถานประกอบการก็สามารถนำบทวิจารณ์ด้านลบ กลับมาปรับปรุงสถานที่ของตนเองให้ดีขึ้นได้ หรืออาจนำบทวิจารณ์ด้านบวกมาใช้ปรับเปลี่ยนแนวทางการทำการตลาดต่อไปได้ในอนาคต

2. วัตถุประสงค์การวิจัย

1. เพื่อจำแนกความรู้สึกจากความเห็นเกี่ยวกับสถานที่ท่องเที่ยวที่อยู่ในบทวิจารณ์ออนไลน์ด้วยเทคนิคนาอีฟเบย์

2. สร้างส่วนปฏิสัมพันธ์ให้สามารถใช้งานเพื่อการจำแนกความรู้สึกได้สะดวก

3. เอกสารและงานวิจัยที่เกี่ยวข้อง

[1] ได้ทำวิจัยเรื่องการวิเคราะห์ความรู้สึกแบบอัตโนมัติจากข้อความแสดงความคิดเห็น โดยมีวัตถุประสงค์เพื่อวิเคราะห์ความรู้สึกของนักท่องเที่ยวที่มีการเขียนในรูปแบบของข้อความบนโซเชียลมีเดียเป็น 2 กลุ่มความรู้สึก คือ นักท่องเที่ยวความรู้สึกที่ดี และนักท่องเที่ยวที่ความรู้สึกที่ไม่ดีต่อสถานที่ท่องเที่ยวที่ได้เข้าชม โดยงานวิจัยได้พัฒนาโมเดลเพื่อการวิเคราะห์ ความรู้สึก โดยทำการจัดกลุ่มเอกสารแบบนาอีฟเบย์ และทำการทดสอบด้วยการวัดค่าเอฟ ผลลัพธ์ที่ได้มานั้นได้แสดงว่าวิธีที่ที่ใช้ในการนำเสนอ สามารถวิเคราะห์เนื้อหา จากความถูกต้องอย่างมีประสิทธิภาพ ตัวอย่างในแต่ละจุดซ้ำ 2 ครั้ง เมื่อเก็บตัวอย่างครบตามกำหนดเวลา หลอดเก็บตัวอย่างจะถูกบรรจุอยู่ในถุงซิปล๊อคและเก็บรักษาอุณหภูมิตามที่วิธีมาตรฐานกำหนดจนมาถึงห้องปฏิบัติการเพื่อทำการวิเคราะห์

[2] ได้ทำวิจัยเรื่อง การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์โดยใช้ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมทชีน โดยมีวัตถุประสงค์เพื่อนำเสนอวิธีการจำแนกความ คิดเห็น โดยการวิเคราะห์ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์ ด้านการบริการห้องพัก โรงแรม รีสอร์ท จาก Agoda Thailand และ Twitter Thailand ที่จดทะเบียนหลักทรัพย์ โดยเก็บข้อมูลจำนวน 2,890 ข้อมูล โดยใช้เทคนิคเหมืองข้อความ วิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิว สินค้าออนไลน์ และสร้างแบบจำลองด้วยอัลกอริทึม 4 วิธี ได้แก่ ซัพพอร์ตเวกเตอร์แมทชีน ต้นไม้ตัดสินใจ นาอีฟเบย์ และเคเนียเรสเนเบอร์ เพื่อเปรียบเทียบ ประสิทธิ์ภาพการวิเคราะห์ความคิดเห็น

ภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์ จากการทดลอง พบว่า คุณลักษณะที่ดีสุด คือ ซัพพอร์ตเวกเตอร์แมทชีน ระดับรองลงมาเป็น นาอีฟเบย์ , ต้นไม้ตัดสินใจ และเคเนีย เรสเนเบอร์ ตามลำดับ

[3] ได้ทำวิจัยเรื่อง ระบบวิเคราะห์ข้อความแสดงความคิดเห็นสำหรับโรงแรม โดยมีวัตถุประสงค์เพื่อวิเคราะห์ข้อความแสดงความคิดเห็นสำหรับโรงแรม เป็นระบบที่สามารถวิเคราะห์ข้อมูลที่อยู่ ในรูปแบบการแสดงความคิดเห็นด้วยภาษาธรรมชาติโดยอาศัยหลักการการวิเคราะห์และสรุปผลจาก ทัศนคติในระดับคุณลักษณะทำให้ข้อมูลการแสดงความคิดเห็นอยู่ในรูปแบบที่เป็นข้อมูลสรุปให้เข้าใจได้ง่ายในรูปแบบกราฟิก โดยสามารถเลือกกลับไปดูข้อความแสดงความคิดเห็นที่ถูกนำมาวิเคราะห์ในหัวข้อของมุมมองที่สนใจได้

[4] ได้นำเสนอเรื่องการทำการวิเคราะห์ความรู้สึก จากวิเคราะห์อารมณ์และความรู้สึกจากข้อความ เพื่อบ่งบอกความรู้สึกของผู้คนที่มีต่อบางสิ่งบางอย่าง เช่น ความรู้สึกดี หรือ ความรู้สึกที่ไม่ดี ตัวอย่างเช่น เขาเป็นคนดี ช่วยแมวที่กำลังจะตกจากอาคารสูง น่าชื่นชมนะ อันนี้คือ ความรู้สึกดีเขาทำงานบกพร่องจนทำให้มีคนเดือดร้อน อันนี้คือความรู้สึกที่ไม่ดี

[5] ได้ทำวิจัยเรื่องการสร้างคลังศัพท์บอกความรู้สึกในภาษาไทยจากบทวิจารณ์ออนไลน์ โดยมีวัตถุประสงค์เพื่อสร้างคลังศัพท์บอกความรู้สึกจากบทวิจารณ์สินค้าและบริการออนไลน์ โดยวิธีการประมวลภาษาทางธรรมชาติ และเลือกแหล่งข้อมูลจากบทวิจารณ์ของโรงแรม บทวิจารณ์ภาพยนตร์ของโรงภาพยนต์ และบทวิจารณ์แอปพลิเคชันโทรศัพท์มือถือ โดยวิเคราะห์และให้คะแนนจากคำบอกเล่า วิเคราะห์ความรู้สึกจากข้อมูลของบทวิจารณ์ ซึ่งความรู้สึกจะเกิดขึ้นจากคำอธิบายลักษณะของสินค้า มีทั้งบวกและลบ ซึ่งดูได้จากะดูจากค่า tf-idf เชิงบวกและเชิงลบซึ่งคำนวณจากความถี่การปรากฏในกลุ่มข้อมูลบทวิจารณ์ที่มีการให้คะแนนเชิงบวกและเชิงลบตามลำดับ

[6] ได้นำเสนอและอธิบาย 10 Metrics พื้นฐานสำหรับวัดผลโมเดล Machine Learning ไว้ซึ่งเป็นประโยชน์สำหรับการใช้วัดผลความแม่นยำของโมเดลที่สร้างขึ้น 10 Metrics อันประกอบไปด้วย Regression – MAE, MSE, RMSE, R2 Classification – Accuracy, Precision, Recall, F1-Score, F-Beta Score, AUC

4. วิธีดำเนินการวิจัย

1. เครื่องมือการวิจัย

    1.1 ภาษาที่ใช้ในการพัฒนา ได้แก่ Python

    1.2 โปรแกรมที่ใช้พัฒนา ได้แก่ Pycharm

    1.3 ชุดโปรแกรมที่ใช้พัฒนา ได้แก่ Natural Language Toolkit, Natural Language Process, scikit-learn

2. ขั้นตอนการดำเนินการวิจัย

    2.1 รวบรวมข้อมูลการให้ความเห็นเกี่ยวกับสถานที่ท่องเที่ยว จากแหล่งข้อมูล 2 แหล่งคือ จากข้อความวิจารณ์ของ Google Map (https://www.google.co.th/maps) และเว็บไซต์ทริปแอดไวเซอร์ (https://th.tripadvisor.com) ของนักท่องเที่ยวไทย ที่เป็นภาษาไทย จำนวน 300 ตัวอย่าง และเพื่อการทดสอบ 200 ตัวอย่าง ในช่วงปี 2560 - 2561

    2.2 จัดระเบียบข้อมูลในรูปแบบที่พร้อมจะนำไปประมวลผลได้ง่าย ในรูปแบบไฟล์ CSV

    2.3 พัฒนาระบบ และออกแบบระบบ

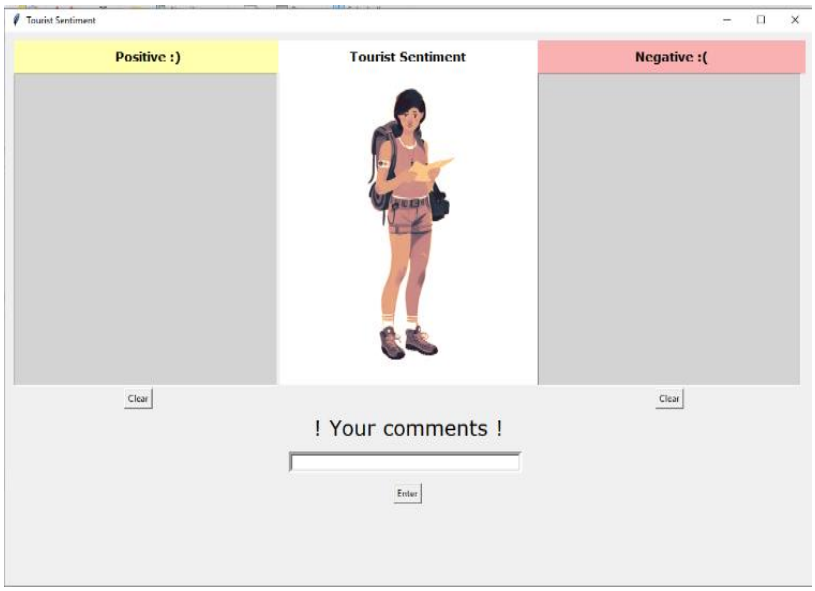
    2.4 ทดลองใช้ระบบ และประเมินประสิทธิภาพ

3. สถิติที่ใช้ในการวิจัย ได้แก่ นาอีฟเบย์ การระลึกได้ ความแม่นยำ ค่าทดสอบเอฟ และ ความถูกต้อง

5. ผลการวิจัย

1. ผลการพัฒนาระบบการจำแนกทัศนคติการท่องเที่ยว

การใช้งานโปรแกรมจะทำการดึงข้อมูลที่เตรียมไว้จาก pos.txt และ neg.txt มาไว้ใน list จากนั้นก็จะทำการติดแท็กที่แต่ละประโยคว่าเป็น positive หรือ negative จากนั้นโปรแกรมก็จะนำประโยคที่ได้มาตัดคำออกจากประโยคโดยใช้ Pythainlp แล้วหาเวกเตอร์ของคำแล้วประมวลผลด้วยอัลกอริทึม Naïve-Bayes เพื่อคำนวณหาผลลัพธ์ ทำงานโดยการนำประโยคไปใส่ไว้ในไฟล์ pos.txt และ neg.txt และโปรแกรมจะนำเข้ามาเก็บไว้ใน list Pos = สถานที่นี้สวยงาม Neg = สถานที่นี้แย่มาก แล้วจาก list ทำการแท็กทีละประโยคว่าเป็น pos หรือ neg ในรูปแบบ [(สถานที่นี้สวยงาม’,’pos’), (สถานที่นี้แย่มาก’,’neg’)] จากนั้นนำมาตัดคำออกจากประโยค {'นี้', 'สถานที่', 'สวยงาม', 'แย่มาก'} จากนั้นนำคำที่ผ่านการตัดแล้วไปวนลูปในแต่ละประโยคเพื่อหาความถี่ของคำ เมื่อเจอคำในประโยคจะคืนค่า True หากไม่เจอจะคืนค่า False ({'แย่มาก': False,'สวยงาม': True, 'นี้': True, 'สถานที่': True}, 'positive') ({'แย่มาก': True, 'สวยงาม': False, 'นี้': True, 'สถานที่': True}, 'negative') จากการทดลองนี้ได้ทำการเพิ่มประโยคไปอีกอย่างละ 1 ประโยคเพื่อให้สามารถให้น้ำหนักคำได้ แย่มาก = False pos : neg = 1.7 : 1.0 สวยงาม = False neg : pos = 1.7 : 1.0 ดี = False neg : pos = 1.7 : 1.0 ไม่ดี = False pos : neg = 1.7 : 1.0 นี้ = True neg : pos = 1.0 : 1.0 สถานที่ = True neg : pos = 1.0 : 1.0 ผลจากการนำไปประมวลผลผ่านอัลกอริทึม Naïve-Bayes ในที่นี้ True คือเมื่อพบข้อความ False คือเมื่อไม่พบข้อความ กรณีตัวอย่าง เช่น แย่มาก = False คือ ถ้าหากไม่พบข้อความว่า “แย่มาก” ก็จะมีน้ำหนักไปทางฝั่ง positive = 1.7



รูปที่ 1 หน้าแรกของระบบจำแนกทัศนคิการท่องเที่ยว

จากรูปที่ 1 เมื่อเข้ามาในโปรแกรมแล้วจะมีชื่อของโปรแกรม Tourist Sentiment จะสังเกตุได้ว่าจะมีช่องสำหรับกรอกข้อความอยู่ข้างล่าง ตัวหนังสือที่เขียนว่า Your comments ให้ใส่ประโยคที่เราต้องการวิเคราะห์ลงไปในช่องว่าง จากนั้นกด Enter และจะมีช่องสำหรับผลลัพธ์ทางฝั่ง Positive และ Negative อยู่ทางด้านซ้ายและขวาและ มีปุ่ม clear ไว้สำหรับลบข้อความที่แสดงผล

รูปที่ 2 หน้าแรกของระบบ Classification ทัศนติการท่องเที่ยว โดย Naïve-Bayes

จากรูปที่ 2 เมื่อกรอกข้อความเสร็จแล้วทำการกดปุ่ม Enter โปรแกรมก็จะทำการประมวลผลจากนั้นจะได้ผลลัพธ์ออกมาทางด้านซ้ายและขวา ถ้าเป็นขั้วบวก ก็จะแสดงที่ด้านซ้าย ถ้าเป็น ขั้วลบก็จะแสดงทางด้านขวา

2. ผลการประเมินความแม่นยำของระบบจำแนกทัศนคติการท่องเที่ยว

การประเมินประสิทธิผลของการค้นค้นสารสนเทศ ได้ทำการทดสอบเพื่อวัดประสิทธิผลด้วยค่าความแม่นยำ (Recall) และค่าแม่นยำ (Precision) เป็นค่าการวัดประสิทธิภาพพื้นฐานในการจัดกลุ่มเอกสาร เพื่อประเมินค่ารวมของค่าทดสอบเอฟ (F) และความถูกต้อง (Accuracy) เมื่อทำการทดสอบแบบจำลองการจัดกลุ่มเอกสาร เพื่อประยุกต์ใช้ในการวิเคราะห์ความรู้สึก ด้วยเอกสารทดสอบที่แตกต่างเอกสารที่ใช้ในการสร้างแบบจำลอง จำนวน 120 เอกสาร ด้วยค่าความระลึก ค่าความแม่นยำ ให้ค่าทดสอบเอฟ และความถูกต้อง โดยพิจารณาจากจำนวนเอกสาร ทีละกลุ่มด้วยจำนวนที่แตกต่างกัน คือ 80, 100, และ 120 ตามลำดับ สามารถแสดงผลของการทดสอบได้ดังตารางที่ 1

ตารางที่ 1 ผลการประเมินความแม่นยำของระบบจำแนกทัศนคติการท่องเที่ยว

แบบจำลอง	Recall	Precision	F-measure	Accuracy
แบบจำลองที่สร้างจากข้อความ ข้อความ 80	0.73	0.84	0.78	0.78
แบบจำลองที่สร้างจากข้อความ ข้อความ 100	0.93	0.73	0.81	0.80
แบบจำลองที่สร้างจากข้อความ ข้อความ 120	0.93	0.73	0.83	0.83
เฉลี่ย	**0.86**	**0.80**	**0.80**	**0.80**

จากตารางที่ 1 การทดสอบแบบจำลองการจัดกลุ่มเอกสารที่ประยุกต์ใช้ในการวิเคราะห์ความรู้สึก ให้ผลการทดสอบอย่างน่าพึงพอใจ เพราะเมื่อทดสอบด้วยค่าความถูกต้องมีค่าเฉลี่ยที่ 0.80

5. อภิปรายผลการวิจัย

งานวิจัยนี้เป็นการวิเคราะห์ข้อความที่แสดงความรู้สึกแบบอัตโนมัติของนักท่องเที่ยวที่มีต่อสถานที่ ท่องเที่ยวในประเทศไทยที่มีการแสดงออกไว้ในข้อความแสดงความรู้สึกผ่านสื่อสังคมออนไลน์ต่าง ๆ โดยในงานวิจัยนี้ทำการวิเคราะห์ความรู้สึกที่ดี และความรู้สึกที่ไม่ดีต่อสถานที่ท่องเที่ยว โดยใช้เทคนิคนาอีฟเบย์ ซึ่งในขั้นตอนของการวัดประสิทธิภาพการจัดขั้ว จะมีการสร้างแบบจำลองสำหรับวิเคราะห์ความรู้สึกโดยใช้ค่าความแม่นยำ ค่าความระลึก ค่าทดสอบเอฟ และค่าความถูกต้องเมื่อทำการทดสอบวัดประสิทธิภาพของการสร้างจำลองในการค้นคืน พบว่าค่าความถูกต้อง และค่าทดสอบเอฟ เฉลี่ยได้เท่ากับ 0.80 แม้จากผลการทดสอบจะให้ ผลลัพธ์ที่น่าพอใจ แต่ก็ยังพบความผิดพลาดอันเนื่องมาจากการรวบรวมเอกสารจากสื่อสังคมออนไลน์นั้น มีเอกสารหลายฉบับที่ใช้ประโยคกำกวม ไม่ชัดเจนว่าจะแสดงความรู้สึกทางด้านไหนมากกว่ากัน ส่งผลให้แบบจำลองที่สร้างมา เพื่อวิเคราะห์ความรู้สึกมีความผิดพลาด ส่งผลทำให้ประสิทธิภาพของการวิเคราะห์ลดต่ำลงได้

6. ข้อเสนอแนะ

ระบบวิเคราะห์ข้อความแสดงความคิดเห็นเกี่ยวกับการท่องเที่ยวสามารถที่จะวิเคราะห์ข้อความแสดง ความคิดเห็นเกี่ยวกับการท่องเที่ยวได้ แต่เนื่องจากข้อความแสดงความคิดเห็นภาษาไทยไม่มีรูปแบบที่แน่นอน จึงทำให้ยังมีข้อผิดพลาดอยู่ หากผู้พัฒนาอยากนำไปพัฒนาต่อ ควรพัฒนาอัลกอริทึมที่ทำการตัดประโยคสำหรับการวิเคราะห์ความรู้สึกโดยเฉพาะโดยใช้การตัดคำร่วมกับคลังคำวิเคราะห์ความรู้สึก ซึ่งหาได้จากงานวิจัยทางภาษาศาสตร์ ซึ่งอาจใช้คำหลายคำร่วมกันมากกว่าจะตัดเป็นแยกในแบบการตัดคำทั่วไป หากมีการปรับปรุงหรือแก้ไขเพิ่มเติมก็จะสามารถทำให้ระบบวิเคราะห์ ข้อความแสดงความคิดเห็นเกี่ยวกับการท่องเที่ยวสามารถวิเคราะห์ข้อความแสดงความคิดเห็นเกี่ยวกับการท่องเที่ยวได้อย่างแม่นยำยิ่งขึ้น และสามารถนำไปปรับใช้กับการวิเคราะห์ข้อความแสดงความคิดเห็นในหัวข้ออื่นที่น่าสนใจได้อย่างมีประสิทธิภาพ

7. เอกสารอ้างอิง

[1] เจตรินทร์ วงศ์ศิลป์ ณัฐกิตติ์ ศรีกาญจนเพริศ และจันทิมา พลพินิจ. (2560). การวิเคราะห์ความรู้สึกแบบอัตโนมัติจากข้อความแสดงความ*คิดเห็น. ในการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศครั้งที่ 15 (NCCIT2015),* 2560, หน้า 1-6. กรุงเทพฯ: คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

[2] รวิสุดา และ นิเวศ. (2560). *การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์ โดยใช้ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมทชีน.* สืบค้นวันที่ 5 พฤศจิกายน 2561, จาก ejsu.siam.edu/journals/PDF_34/EJSU_No.34_pp_1-12.pdf

[3] วรัญญา. (2553). ระบบวิเคราะห์ข้อความแสดงความคิดเห็นสำหรับโรงแรม. สืบค้นวันที่ 5 พฤศจิกายน 2561, จาก thailang.nectec.or.th/halloffame/images/stories/best/download/13p33c001.pdf/

[4] วรรณพงษ์. (2558). *การประมวลภาษาธรรมชาติด้วยภาษาไพทอน.* สืบค้นวันที่ 5 พฤศจิกายน 2561, จาก https://python3.wannaphong.com/2015/10/blog-post.html

[5] วรรณพงษ์. (ม.ป.ป). *คู่มือการใช้งาน PyThaiNLP 1.4.* สืบค้นวันที่ 5 พฤศจิกายน 2561, จาก https://pythainlp.readthedocs.io/en/latest/pythainlp-1-4-thai/

[6] อิสรภาพ ล้อรัตนไชยยงค์. (2560). *การสร้างคลังคำศัพท์บอกความรู้สึกในภาษาไทยจากบทวิจารณ์ออนไลน์* [อ.ม. สาขาภาษาศาสตร์]. กรุงเทพฯ: คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

[7] Kasidis. (2560). *อธิบาย 10 Metrics พื้นฐานสำหรับวัดผลโมเดล Machine Learning*. สืบค้นวันที่ 8 สิงหาคม 2562, จาก https://datarockie.com/2019/03/30/top-ten-machine-learning-metrics/