

การพัฒนาการวิเคราะห์ลักษณะเฉพาะทางความรู้สึกลักษณะทางการภาพของรูปภาพ ในสื่อเครือข่ายสังคม

ประสิทธิ์ จุฬานุกรกุล¹ และ วรพล พงษ์เพชร²

สาขาวิชาการวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ (NIDA)

148 ถนนเสรีไทย แขวงคลองจั่น เขตบางกะปิ กรุงเทพมหานคร 10240

Received: 18 June 2021; Revised: 21 December 2021; Accepted: 3 May 2022

บทคัดย่อ

การพิจารณาเลือกดัชนีถ้อยคำ (Hashtag) หรือการเลือกรูปภาพบนสื่อเครือข่ายสังคมทางอินเทอร์เน็ตให้เหมาะสมเป็นส่วนหนึ่งของการทำการตลาดออนไลน์ที่ดี แต่การพิจารณาว่าดัชนีถ้อยคำ (Hashtag) ใดมีความเหมาะสมนั้น มีความยุ่งยากอยู่พอสมควร เนื่องจากการเลือกดัชนีถ้อยคำ (Hashtag) มักใช้ประสบการณ์ของผู้ทำการตลาดออนไลน์เพียงอย่างเดียว โดยไม่ได้คำนึงถึงรูปภาพที่อยู่ภายใต้ดัชนีถ้อยคำ (Hashtag) งานวิจัยชิ้นนี้เป็นการสร้างเครื่องมือในการวิเคราะห์รูปภาพที่อยู่ภายใต้ดัชนีถ้อยคำ (Hashtag) หรือรูปภาพบนสื่อเครือข่ายสังคมของผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ต เพื่อใช้ในการตัดสินใจเลือกดัชนีถ้อยคำ (Hashtag) หรือสามารถประยุกต์ใช้เพื่อการหาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ตให้เหมาะสม และมีการนำระบบที่พัฒนาได้มาลองใช้เป็นกรณีศึกษาเพื่อใช้ในการตัดสินใจเลือกหาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ต

คำสำคัญ: การเรียนรู้เชิงลึก, การวิเคราะห์ภาพ, เครือข่ายสังคมออนไลน์, แฮชแท็ก

* Corresponding author. E-mail: prasit.chul@stu.nida.ac.th

¹ นักศึกษาปริญญาโท สาขาวิชาการวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

² อาจารย์ สาขาวิชาการวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

The Development of the Sentiment and Visual Characterization Analyzer for the Social Media Images

Prasit Chulanutrakul¹ and Worapol Pongpech²

Business Analytics and Data Science, BADS Graduate School of Applied Statistics

National Institute of Development Administration (NIDA)

148 Serithai Road, Klong-Chan, Bangkok, Thailand, 10240

Received: 18 June 2021; Revised: 21 December 2021; Accepted: 3 May 2022

Abstract

The consideration of Hashtag or choosing it appropriately to pictures on social media is one of a good strategy for online marketing. However, it is quite complicated to consider for a proper hashtag because it only requires the experience of online marketers regardless of the pictures under the hashtag. This research is made to generate a tool for analysis of pictures under a searched hashtag or pictures of influencers on social media. Moreover, it is used for making decision on the hashtag and be able to integrate in searching for the suitable influencers on social media. Moreover, the developed system is demonstrated as a case study for decision making in selecting influencers in social media networks on the Internet.

Keywords: deep learning, image analytics, social network, hashtag

* Corresponding author. E-mail: prasit.chul@stu.nida.ac.th

¹ Master's degree in Graduate School of Applied Statistics (GSAS), National Institute of Development Administration (NIDA)

² Lecturer in Graduate School of Applied Statistics (GSAS), National Institute of Development Administration (NIDA)

1. ที่มาและความสำคัญ

ในปัจจุบันการตลาดออนไลน์เข้ามามีบทบาทในการทำรายได้ของบริษัทต่างๆ ทั้งส่วนที่เป็นการโฆษณาสินค้าของบริษัทและส่วนของการสร้างการรับรู้ของยี่ห้อสินค้า และการสร้างความน่าเชื่อถือให้กับสินค้า การโฆษณาในสื่อสังคมออนไลน์มีหลายปัจจัยที่เป็นปัจจัยส่งเสริมการขายสินค้า การเลือกดัชนีถ้อยคำ (Hashtag) เป็นหนึ่งในปัญหาหลักของงานวิจัยนี้ การพิจารณาเลือกดัชนีถ้อยคำ (Hashtag) ให้มีความเหมาะสมกับรูปภาพโฆษณาบนสื่อเครือข่ายสังคมทางอินเทอร์เน็ต เป็นส่วนที่มีความสำคัญอย่างหนึ่ง โดยปกติการพิจารณาเลือกนั้นมักมาจากประสบการณ์ของผู้ทำการตลาดออนไลน์เพียงอย่างเดียวไม่ได้มาจากการตัดสินใจบนพื้นฐานของข้อมูล เนื่องจากการที่จะให้ได้มาซึ่งข้อมูลที่จะนำมาใช้ในการตัดสินใจนั้นทำได้ยากและอาจจะต้องใช้มนุษย์เพื่อทำการวิเคราะห์รูปภาพทีละรูป ทำให้ใช้ทั้งกำลังคนจำนวนมากและใช้เวลานาน ด้วยเหตุนี้จึงอาจทำให้เกิดการตัดสินใจผิดพลาดได้ งานวิจัยนี้ยังสามารถประยุกต์ใช้กับการเลือกหาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ตเพื่อเป็นผู้ส่งเสริมการขายสินค้าได้อีกด้วยจาก ระบบที่สร้างขึ้น ดังนั้นจึงนำเสนอทั้ง การพิจารณาเลือกดัชนีถ้อยคำ (Hashtag) และ การเลือกหาผู้มีอิทธิพล (Influencer)

การวิเคราะห์รูปภาพที่อยู่ภายใต้ดัชนีถ้อยคำ (Hashtag) ยังสามารถนำไปประยุกต์ใช้ในแง่มุมต่างๆ ได้อีกมาก เช่นในบางหมวดหมู่ของดัชนีถ้อยคำ (Hashtag) ที่เป็นยี่ห้อของสินค้า รูปภาพภายใต้ดัชนีถ้อยคำ (Hashtag) เหล่านั้นจะเป็นการกล่าวถึงสินค้าของยี่ห้อเหล่านั้น การวิเคราะห์ในรายละเอียดจะเปิดเผยให้เห็นถึงเสียงของลูกค้า (Voice of Customers) ที่แท้จริง ผ่านลูกค้าที่ซื้อสินค้าหรือบริการ และนำมาลงในสื่อเครือข่ายสังคมทางอินเทอร์เน็ต ซึ่งสามารถนำมาใช้กับการวิเคราะห์ในช่วงเวลาใดเวลาหนึ่งเพื่อวิเคราะห์ถึงกระแสที่เกิดขึ้นในช่วงนั้นได้อีกด้วย โดยปกติแล้วกระแสของสังคมเป็นสิ่งที่เปลี่ยนแปลงตามเวลา ทำให้ต้องทำการวิเคราะห์อย่างสม่ำเสมอ อย่างไรก็ตามเครื่องมือที่นำเสนอในงานวิจัยนี้จะสามารถนำเข้ามาช่วยเพื่อให้การวิเคราะห์เป็นไปอย่างอัตโนมัติ

ในงานวิจัยนี้จะเป็นการสร้างเครื่องมือวิเคราะห์เพื่อวิเคราะห์รูปภาพจำนวนมากเพื่อหาลักษณะทางกายภาพของภาพเบื้องต้น คือ ขนาด, การวางแนวภาพ และอัตราส่วนภาพ กับการวิเคราะห์องค์ประกอบของรูปภาพเชิงลึก เช่นลักษณะของรูปภาพ (Image Style) , ความรู้สึกในรูปภาพ (Image Sentiment), ฉากหลังของรูปภาพ (Scene) และ วัตถุที่อยู่ในรูปภาพ (Object) พร้อมทั้งสรุปผลเป็นคะแนนให้พร้อมสำหรับการตัดสินใจ

เนื่องด้วยการวิเคราะห์ลักษณะของรูปภาพ (Image Style) ความรู้สึกในรูปภาพ (Image Sentiment), ฉากหลังของรูปภาพ (Scene) และ วัตถุที่อยู่ในรูปภาพ (Object) ใช้เทคนิคการวิเคราะห์เพื่อการจำแนกประเภทรูปภาพ (Image Classification) เป็นพื้นฐาน ซึ่งการวิเคราะห์เพื่อการจำแนกประเภทรูปภาพ (Image Classification) นั้นเป็นการวิเคราะห์ข้อมูลที่ไม่มีโครงสร้าง (unstructured data) การวิเคราะห์ในรูปแบบนี้จำเป็นต้องมีสองกระบวนการคือ กระบวนการสกัดตัวแทนที่มีลักษณะเฉพาะ (Feature Extraction) และกระบวนการจำแนกประเภทรูปภาพ (Image Classification) ในแนวทางดั้งเดิมปัญหาหนึ่งในการวิเคราะห์คือกระบวนการสกัดตัวแทนที่มีลักษณะเฉพาะ (Feature Extraction) แต่เดิมการจัดการกับรูปภาพ และการสร้างตัวแทนที่มีลักษณะเฉพาะมักเป็นการสร้างจากเทคนิคทำมือ (Hand-Crafted Feature Extraction) เพื่อเลือกหาวิธีการสกัดลักษณะเฉพาะ เช่น Haar-like feature, Histogram of Oriented Gradients (HOG), bag of visual words (BOVW), Scale-Invariant Feature Transform (SIFT) และ Speeded Up Robust Feature (SURF) กระบวนการเหล่านี้ใช้เวลาอย่างมาก และต้องพึ่งพาผู้เชี่ยวชาญเชิงลึกในศาสตร์สาขาใดสาขาหนึ่ง (Domain Expert) ในประเภทของภาพที่ทำการแยก และความสามารถในการสกัดนำเอาตัวแทนที่มีลักษณะเฉพาะออกมาได้ดีหรือไม่นั้นจะขึ้นกับความรู้เชิงลึก (Domain Knowledge) ของผู้เชี่ยวชาญในศาสตร์สาขาใดสาขาหนึ่ง (Domain Expert) ของชนิดรูปภาพนั้นๆ ซึ่งผลของการทำนายด้วยตัวแบบการเรียนรู้ของเครื่องแบบดั้งเดิม (Traditional machine learning) เช่น SVM จะขึ้นกับขั้นตอนการสกัดตัวแทนที่มีลักษณะเฉพาะ (Feature Extraction) เป็นหลักด้วย

ต่อมาได้มีการพัฒนาเทคนิคที่ใช้กระบวนการเรียนรู้เชิงลึก (Deep Learning) เข้ามาเพื่อสกัดตัวแทนที่มีลักษณะเฉพาะ (Feature Extraction) ได้อย่างรวดเร็ว โดยไม่จำเป็นต้องใช้ผู้เชี่ยวชาญเชิงลึก (Domain Expert) มาสกัดตัวแทนที่มีลักษณะเฉพาะ (feature) โดยเทคนิคที่นิยมใช้คือ Convolutional Neural Networks (CNN) ที่สามารถสกัดเอาตัวแทนที่ยังคงรักษาเอกลักษณ์เฉพาะของรูปได้เป็นอย่างดีหรือสามารถรักษาความเหมือนกัน (Preserve Similarity) ของรูปแม้จะถูกแปลง (Map) จากปริภูมิรูปที่ป้อนเข้า (Input Space) ไปอยู่บนโดเมนของปริภูมิมิติสูง (High Dimensional Space) ของปริภูมิตัวแทนที่มีลักษณะเฉพาะ (Feature Space) แล้วก็ตาม ทำให้ในงานวิจัยนี้จึงจะนำเอาตัวแบบประเภท Convolutional Neural Networks (CNN) มาใช้งานเพื่อให้ตอบโจทย์ของการวิเคราะห์ดังกล่าว

2. ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

โดยปกติงานวิจัยที่ทำการวิเคราะห์ดัชนีถ้อยคำ (Hashtag) ส่วนมากจะเป็นงานวิจัยที่ทำกับเครือข่ายสังคมประเภทที่เป็นข้อความ เช่น Twitter เป็นหลัก ส่วนงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์เนื้อหาของดัชนีถ้อยคำ (Hashtag) ของรูปภาพ หรือรูปภาพที่ผู้มีอิทธิพล (Influencer) นำมาลงเครือข่ายสังคม (Social Network) มีจำนวนไม่มาก หนึ่งในนั้นคืองานวิจัยของ Hu และ คณะ (2014) [1] ได้ศึกษาเกี่ยวกับการหาประเภทของรูปภาพที่ผู้ใช้นิยมนำไปยัง Instagram และจัดกลุ่มของผู้ใช้บริการ Instagram จากรูปที่นำไปยัง Instagram ซึ่งการทำความเข้าใจกลไกของของเครือข่ายสังคม (Social Network) เหล่านี้หมายถึงการทำงานกับพลวัตของมนุษย์ วัฒนธรรมสังคม และสิ่งแวดล้อมซึ่งจะทำให้เห็นถึงสถานการณ์ต่างๆ จากมุมมองทางธุรกิจ การสื่อสาร และการตลาด [2]

ปัจจัยสำคัญอย่างหนึ่งของสื่อสังคมออนไลน์ที่เน้นการให้ผู้ใช้งานลงรูปภาพคือ ดัชนีถ้อยคำ (Hashtag) ที่ประกอบไปด้วย “#” และข้อความ เช่น “#อร่อย” เพื่อจัดหมวดหมู่รูปภาพในแอปพลิเคชันสื่อสังคมออนไลน์ เช่น Instagram [3] ซึ่งดัชนีถ้อยคำนี้มีหน้าที่เป็นตัวบอกลักษณะ

ของรูปภาพ หรือมีความเกี่ยวข้องกับคำอธิบายรูปภาพ (Caption) เพื่ออธิบายเนื้อหาให้ชัดเจนยิ่งขึ้น ดังนั้นดัชนีถ้อยคำ (Hashtag) จึงเป็นหนึ่งในวิธีการที่ใช้เพื่อจัดการข้อมูลสารสนเทศ [4] ในทางเดียวกัน Ibba และ คณะ (2015) [2] ได้กล่าวว่า ดัชนีถ้อยคำ (Hashtag) คือ ระบบที่คั่นหนังสือทางสังคม (Social Bookmarking System) ซึ่งเป็นระบบที่ผู้ใช้สร้างขึ้นเองในการจำแนกและจัดระเบียบเนื้อหาออนไลน์เป็นหมวดหมู่ต่างๆ โดยใช้ข้อมูลอภิพันธุ์ (Metadata) เช่นป้ายอิเล็กทรอนิกส์ (Electronic Tag) การใช้ดัชนีถ้อยคำ บางประเภทยังสามารถทำให้ชุมชนบางประเภทก่อตัวขึ้น เช่นกลุ่มสาธารณะเฉพาะกิจที่สร้างขึ้นอย่างรวดเร็ว เพื่อตอบสนองกับเหตุการณ์เฉพาะหรือปัญหาเฉพาะหน้าได้ [5] ทำให้สามารถใช้ดัชนีถ้อยคำ (Hashtag) เพื่อติดตามกระแสสังคมได้ ดังนั้นการหา ดัชนีถ้อยคำ (Hashtag) ที่เหมาะสมจะทำให้ผู้ที่เข้ามาใช้งานสามารถเข้าถึง สินค้าและบริการ ได้อย่างรวดเร็วเนื่องจากถูกจัดหมวดหมู่อย่างเหมาะสมแล้ว

การใช้ผู้มีอิทธิพล (Influencer) อย่างมีประสิทธิภาพต้องทำความเข้าใจผู้ชม และการเข้าใจอินฟลูเอนเซอร์ด้วยจำนวนผู้ติดตามที่แท้จริงของผู้มีอิทธิพลไม่จำเป็นต้องแปลว่าบุคคลนั้นเหมาะสมสำหรับแบรนด์ใดแบรนด์หนึ่งเนื่องจากผู้มีอิทธิพล (Influencer) บน Social Media ต่างจากโฆษณาแบบดั้งเดิม หรือการใช้คนดังมาพูดแทนแบรนด์คือ Influencer จะโต้ตอบและพูดคุยกับผู้ติดตามอยู่เป็นประจำ ดังนั้นจะเป็นการเชื่อมโยงกับกลุ่มเป้าหมาย (Relatability) และความเป็นเอกลักษณ์ (Authenticity) การเลือกหา ผู้มีอิทธิพล (Influencer) บน Social Media จึงจำเป็นต้องหาผู้ที่มีความเหมาะสมครบถ้วน

การวิเคราะห์ลักษณะเฉพาะทางความรู้สึกและลักษณะทางกายภาพของดัชนีถ้อยคำของรูปภาพในสื่อเครือข่ายสังคมมีความจำเป็นอย่างยิ่งที่ต้องใช้เทคนิคต่างๆ มาประกอบกัน การจำแนกความรู้สึกรูปภาพ (Image Sentiment Classification) และการจำแนกประเภทลักษณะรูปภาพ (Image Style Classification) จึงเป็นส่วนหนึ่งเพื่อการวิเคราะห์นี้ จากงานวิจัยที่ผ่านมา Satoh (2016) [6] ได้สร้างตัวแบบสำหรับจำแนกความรู้สึก

(Sentiment Analysis) โดยใช้รูปภาพในโซเชียลมีเดียจากเว็บไซต์ Flickr และ Instagram ด้วยเทคนิค “ถุงของภาพคำศัพท์ (Bag of visual words)” นอกจากนี้ ในงานวิจัยที่ผ่านมาได้มีการใช้ตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) ประเภท Convolutional Neural Network (CNN) เพื่อการจำแนกประเภทรูปภาพ (Image Classification) ที่สามารถนำมาวิเคราะห์รูปภาพบนสื่อเครือข่ายสังคมได้

Gajarla และ Gupta (2011) [7] ได้ทำการใช้ตัวแบบโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks, CNN) เพื่อการจำแนกประเภท (Classification) สองตัวแบบเพื่อการตรวจจับรูปภาพตามอารมณ์ (Emotional Detection) และ การวิเคราะห์ความรู้สึกในรูปภาพ (Sentiment Analysis) โดยเก็บรวบรวมรูปภาพในโซเชียลมีเดียจากเว็บไซต์ Flickr นอกจากนี้ Karayev และ คณะ (2014) [8] ได้ใช้รูปภาพจาก Flickr และ Wikipaintings เพื่อสร้างตัวแบบประเภท Convolutional Neural Network (CNN) เพื่อจำแนกตามตามสไตล์ของรูปภาพ (Recognizing Image Style) จากนั้น Zhou และ คณะ (2017) [9] ได้นำเสนอการจำแนกพื้นหลังรูปภาพ (Image Scene Classification) ที่สามารถนำมาใช้กับการวิเคราะห์พื้นหลังของภาพ

การจำแนกรูปภาพด้วยตัวแบบโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks, CNN) นั้น มีผู้พัฒนาออกมาหลายตัว แต่ EfficientNet ที่ถูกพัฒนาโดย Tan และ Le (2020) [10] เป็นตัวแบบ state-of-the-art ที่ให้ความแม่นยำถึง 84.3% สูงเป็นอันดับหนึ่งบนชุดข้อมูล ImageNet ซึ่งมีขนาดเล็กกว่า 8.4 เท่า และเร็วกว่า 6.1 เท่า เมื่อเทียบกับตัวแบบคอนโวลูชันที่ดีที่สุด จึงเป็นตัวแบบที่เหมาะสมต่อการนำมาผ่านเทคนิค กระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) เพื่อการจำแนกรูปภาพได้เป็นอย่างดี ซึ่งกระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) คือ เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยตัวแบบได้รับการฝึกอบรมและพัฒนาสำหรับงานหนึ่งงานแล้วสามารถนำกลับมาใช้ใหม่ในงานที่เกี่ยวข้องได้ [11]

การตรวจจับวัตถุ (Object Detection) เป็นอีกส่วนสำคัญที่สามารถนำมาใช้กับการตรวจจับวัตถุที่ต้องการรวมถึงหาจำนวนวัตถุที่เกิดขึ้นในภาพ ซึ่งมีผู้พัฒนาออกมาหลายตัว แต่ Redmon และ คณะ (2016) [12] ได้นำเสนอตัวแบบ YOLO ที่สามารถตรวจจับวัตถุได้อย่างรวดเร็วและใช้ทรัพยากรการประมวลผลที่น้อย โดย YOLO ถูกสอนบนชุดข้อมูล COCO [13] เทคนิคที่กล่าวมาข้างต้นสามารถนำมาประยุกต์ให้รวมกันเพื่อวิเคราะห์ ลักษณะเฉพาะของดัชนีถ้อยคำ (Hashtag) ได้

3. ระเบียบวิธีวิจัย

การวิเคราะห์ลักษณะเฉพาะทางความรู้สึกและลักษณะทางกายภาพในงานวิจัยนี้สามารถแบ่งได้เป็นสองส่วนใหญ่ คือ ส่วนแรก ลักษณะทางกายภาพของภาพเบื้องต้น ได้แก่ ขนาด การวางแนวภาพ และ อัตราส่วนภาพที่จะได้ผลลัพธ์เป็นค่าสถิติเชิงพรรณนาเพื่อการอธิบายลักษณะทางกายภาพเบื้องต้นของรูปภาพ

ส่วนที่สอง การวิเคราะห์องค์ประกอบของรูปภาพเชิงลึก เช่น ลักษณะของรูปภาพ (Image Style) ความรู้สึกในรูปภาพ (Image Sentiment) ฉากหลังของรูปภาพ (Scene) และ วัตถุที่อยู่ในรูปภาพ (Object) ซึ่งการวิเคราะห์องค์ประกอบของรูปภาพเชิงลึกจะใช้เทคนิคการประยุกต์ตัวแบบสองรูปแบบคือ “ตัวแบบสำเร็จรูป” ซึ่งหมายถึง ตัวแบบที่ผ่านการสอนแล้ว (Pre-trained model) ที่ไม่ผ่านกระบวนการปรับปรุงตัวแบบใดๆ เนื่องจากมีความเหมาะสมที่จะนำมาใช้ “ตัวแบบใหม่” ที่เกิดจากการนำเอาตัวแบบที่ผ่านการสอนแล้ว (Pretrained-Model) มาผ่านกระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) ให้ได้ตัวแบบที่มีความสามารถเฉพาะ เพื่อใช้ในการวิเคราะห์ที่เฉพาะเจาะจงยิ่งขึ้น แล้วนำผลลัพธ์ของแต่ละตัวแบบมาคิดเป็นค่าสถิติเชิงพรรณนา และคะแนนความเหมาะสมของดัชนีถ้อยคำ (Hashtag) ที่จะนำมาใช้ หรือ ผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ต เพื่อใช้สำหรับการตัดสินใจ

ทุกเทคนิคทางการประมวลผลภาพและปัญหาดังที่กล่าวมาจำถูกนำมาประยุกต์ใช้โดยการนำมาสร้างเป็นระบบ

ที่ใช้ในการวิเคราะห์ดัชนีถ้อยคำ (Hashtag) หรือ การเลือกหาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ตที่เหมาะสม

ตัวแบบสำเร็จรูป

“ตัวแบบสำเร็จรูป” ที่นำมาใช้งาน มีสองตัวแบบคือ ตัวแบบจำแนกประเภทฉากหลัง (Scene Classification Model) และ ตัวแบบการตรวจจับวัตถุ (Object Detection) ซึ่งมีรายละเอียดดังตารางที่ 1 รายละเอียดตัวแบบสำเร็จรูป

ตารางที่ 1 รายละเอียดตัวแบบสำเร็จรูป

ชื่อ	ประเภท	ชนิด	ข้อมูลที่ใช้สอน	จุดประสงค์ตัวแบบ
ตัวแบบจำแนกประเภทฉากหลัง	จำแนกประเภท	VGG16	Place365	จำแนกประเภทฉากหลังรูป
ตัวแบบตรวจจับวัตถุ	ตรวจจับวัตถุ	YOLO V4	COCO	ตรวจจับวัตถุทั่วไป

ตัวแบบจำแนกประเภทฉากหลัง (Scene Classification Model) เป็นตัวแบบที่มีความสามารถในการทำนายฉากหลังของภาพได้ 365 ชนิด โดยสามารถดูรายละเอียดเพิ่มเติมเกี่ยวกับชนิดพื้นหลังได้ที่เว็บไซต์ <http://places2.csail.mit.edu/explore.html> ตัวแบบนี้ยังสามารถทำนายออกมาเป็นภาพภายนอก หรือภายในอาคารสถานที่ได้อีกด้วย ตัวแบบนี้จะนำมาใช้วิเคราะห์หาว่าภาพถูกถ่ายขึ้นที่ไหน และอยู่ภายนอก หรือภายในอาคารสถานที่

ตัวแบบตรวจจับวัตถุ (Object Detection) สามารถตรวจจับวัตถุได้ 80 วัตถุซึ่งจะนำมาใช้เพื่อตรวจสอบวัตถุที่สนใจในรูปภาพว่ามีปรากฏอยู่เท่าใดและยังสามารถนำมาหา Rule Set ว่าวัตถุใดเกิดขึ้นพร้อมกับวัตถุใดได้ด้วย

ตัวแบบใหม่

การพัฒนา “ตัวแบบใหม่” ทั้งสองตัวแบบ จะใช้ตัวแบบแกน (Backbone) คือ EfficientNet B0 นำมาผ่านกระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) เพื่อให้ได้ตัวแบบสำหรับ

1. ตัวแบบจำแนกประเภทความรู้สึกรูปภาพ (Image Sentiment Classification)

2. ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification)

3.1 แหล่งที่มาของข้อมูล

เนื่องจากในงานวิจัยนี้ต้องการสร้างตัวแบบการจำแนกประเภทความรู้สึกของรูปภาพ (Image Sentiment Classification) และ ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification) ดังนั้นจึงมีข้อมูลรูปภาพสองส่วนหลักที่ต้องใช้เพื่อการสอนตัวแบบ โดยเป็นข้อมูลสำหรับการสอนตัวแบบมาจากจากเว็บไซต์ ซึ่งเป็นชุดข้อมูลที่อยู่ของรูปภาพ (URL) จากเว็บไซต์ Flickr ที่ถูกติดสลากไว้แล้ว ใช้สำหรับการสอนตัวแบบจำแนกประเภทความรู้สึกรูปภาพ (Image Sentiment Classification) ซึ่งมีรายละเอียดดังนี้

1) ชุดข้อมูลความรู้สึกรูปภาพ (Image Sentiment Dataset)

ข้อมูลสำหรับการสอนตัวแบบมาจากจากเว็บไซต์ <https://mm.doshisha.ac.jp/senti/CrossSentiment.html>

2) ชุดข้อมูลลักษณะรูปภาพ (Image Style Dataset)

ข้อมูลสำหรับการสอนตัวแบบมาจากจากเว็บไซต์ <http://vislab.berkeleyvision.org/datasets.html>

3.2 การเตรียมข้อมูล

3.2.1 การดึงรูปภาพจากอยู่ของรูปภาพ (URL)

ข้อมูลทั้งสองชุดข้อมูลเป็นรายการที่อยู่ของรูปภาพ (URL) ดังนั้นจึงใช้วิธีการเขียนโปรแกรมเพื่อทำการดึงข้อมูลรูปภาพออกมา ซึ่งที่อยู่ของรูปภาพ (URL) บางส่วนมีความเสียหายไม่สามารถเข้าถึงข้อมูลได้ โดยชุดข้อมูลความรู้สึกรูปภาพ (Image Sentiment Dataset) มีจำนวนรูปภาพที่สามารถเข้าถึงได้ รวมรูปภาพทั้งหมด 66,839 รูป ส่วนชุดข้อมูลลักษณะรูปภาพ (Image Style Dataset) มีจำนวนรูปภาพที่สามารถเข้าถึงได้ รวมรูปภาพทั้งหมด 63,989 รูป

3.2.2 การให้ฉลากข้อมูล

1) ชุดข้อมูลความรู้สึกรูปภาพ (Image Sentiment Dataset)

สำหรับชุดข้อมูลความรู้สึกรูปภาพ (Image Sentiment Dataset) จะมีการให้ฉลากข้อมูล (Data Labelling) โดยให้ผู้ติดฉลาก (labeller) จำนวน 3 คน โดยแต่ละรูปภาพผู้ติดฉลากแต่ละคนจะระบุความรู้สึกให้รูปภาพนั้นได้เพียงความรู้สึกเดียวซึ่งประกอบด้วย แง่ลบ (Negative), กลาง (Neutral) และ แง่บวก (Positive) เท่านั้น แต่เพื่อให้การระดับของความรู้สึกมีความละเอียดยิ่งขึ้นในงานวิจัยนี้จะให้วิธีแบ่งความรู้สึกออกเป็น 5 ระดับ คือ แง่ลบอย่างยิ่ง (Highly Negative), แง่ลบ (Negative), กลาง (Neutral), แง่บวก (Positive) และ แง่บวกอย่างยิ่ง (Highly Positive) โดยมีเกณฑ์การแปลงคะแนนจากคะแนนที่ผู้ติดฉลากให้ดังนี้

- แง่ลบอย่างยิ่ง (Highly Negative) ก็ต่อเมื่อ ผู้ติดฉลาก (labeller) ทั้ง 3 คนให้รูปเป็น (Negative)
- แง่ลบ (Negative) ก็ต่อเมื่อ ผู้ติดฉลาก (labeller) 2 คนให้รูปเป็น ลบ (Negative) และ 1 คนให้เป็น กลาง (Neutral)
- กลาง (Neutral) ก็ต่อเมื่อ ผู้ติดฉลาก (labeller) ทั้ง 3 คนให้รูปเป็น กลาง (Neutral)
- แง่บวก (Positive) ก็ต่อเมื่อ ผู้ติดฉลาก (labeller) 2 คนให้รูปเป็น บวก (Positive) และ 1 คนให้เป็น กลาง (Neutral)
- แง่บวกอย่างยิ่ง (Highly Positive) ก็ต่อเมื่อ ผู้ติดฉลาก (labeller) ทั้ง 3 คนให้รูปเป็น (Positive)

สามารถสรุปได้ตามตารางที่ 2 การแปลงคะแนนความรู้สึก

ตารางที่ 2 การแปลงคะแนนความรู้สึก

	แง่ลบ (Negative)	กลาง (Neutral)	แง่บวก (Positive)
แง่ลบอย่างยิ่ง (Highly Negative)	3	0	0
แง่ลบ (Negative)	2	1	0
กลาง (Neutral)	1	2	0
กลาง (Neutral)	0	3	0
กลาง (Neutral)	0	2	1
แง่บวก (Positive)	0	1	2
แง่บวกอย่างยิ่ง (Highly Positive)	0	0	3

ภายหลังจากการแปลงคะแนน จะมีจำนวนรูปภาพในแต่ละประเภทฉลากดังตารางที่ 3 จำนวนรูปภาพในชุดข้อมูลความรู้สึกรูปภาพแยกตามความรู้สึก

ตารางที่ 3 จำนวนรูปภาพในชุดข้อมูลความรู้สึกรูปภาพแยกตามความรู้สึก

ประเภท	จำนวน (รูป)
Highly Negative	7,122
Negative	7,005
Neutral	9,838
Positive	13,697
Highly Positive	29,177
รวมรูปภาพทั้งหมด	66,839 รูป

2) ชุดข้อมูลลักษณะรูปภาพ (Image Style Dataset)

สำหรับข้อมูลสำหรับสอนตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification) ฉลากข้อมูลที่ให้มามีความสมบูรณ์พร้อมใช้งาน ไม่จำเป็นต้องให้ฉลากข้อมูลเพิ่มเติม ซึ่งมีจำนวนรูปภาพในแต่ละประเภทฉลากดังตารางที่ 4 จำนวนรูปภาพในชุดข้อมูลลักษณะรูปภาพแยกตามลักษณะรูปภาพ

ตารางที่ 4 จำนวนรูปภาพในชุดข้อมูลลักษณะรูปภาพแยกตามลักษณะรูปภาพ

ประเภท	จำนวน (รูป)	ประเภท	จำนวน (รูป)
Bokeh	3,262	Macro	3,322
Bright	3,208	Melancholy	2,753
Depth of Field	3,102	Minimal	3,195
Detailed	2,743	Noir	3,570
Ethereal	3,220	Pastel	2,912
Geometric Composition	3,150	Romantic	2,957
Hazy	3,385	Serene	3,625
HDR	3,283	Sunny	3,210
Horror	3,511	Texture	3,306
Long Exposure	3,097	Vintage	3,178
รวมรูปภาพทั้งหมด	63,989 รูป		

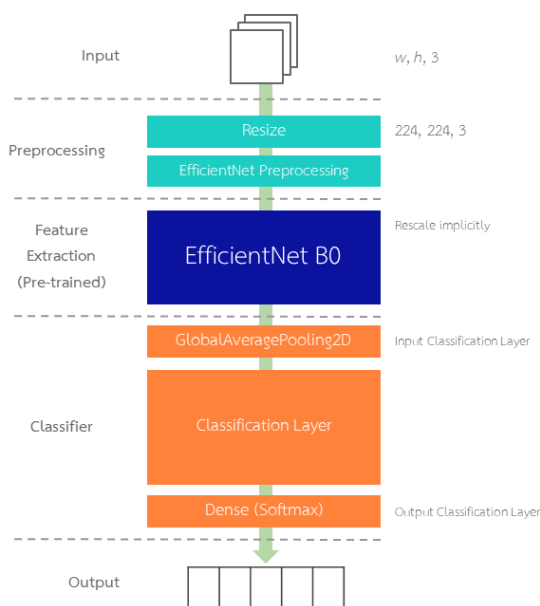
3.3 การสร้างตัวแบบใหม่

ทั้งสองตัวแบบ เกิดจากการนำเอาตัวแบบที่ผ่านการสอนแล้ว (Pretrained-Model) มาผ่านกระบวนการถ่ายทอดการเรียนรู้ (Transfer Learning) โดยใช้ ส่วนการสกัดลักษณะเฉพาะ (Feature Extraction) เป็นโครงข่าย

ประสาท (Neural Network) ประเภทคอนโวลูชัน (Convolutional Neural Network) ชนิด EfficientNet B0 ที่ผ่านการสอนแล้วบนชุดข้อมูล ImageNet (Pre-trained on ImageNet Dataset) โดยนำชั้นตัวแบบจำแนกประเภท (Classification Layer) ออก แล้วสร้างชั้นตัวแบบจำแนกประเภทใหม่ จากนั้นทำการป้องกันไม่ให้น้ำหนักของชั้นการสกัดลักษณะเฉพาะ (Feature Extraction Layer) เปลี่ยนแปลง แล้วทำการสอนตัวแบบทั้งหมดด้วยชุดข้อมูลสอน

โดยชุดข้อมูลจะถูกแบ่งออกเป็นชุดสอน (Train Set) และ ชุดทดสอบ (Test Set) ในอัตราส่วน 80:20 เหมือนกัน ทั้งตัวแบบการจำแนกประเภทความรู้สึกของรูปภาพ (Image Sentiment Classification) และ ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification)

รูปที่ 1 แสดงถึงโครงสร้างโดยคร่าว ของตัวแบบใหม่ โดยตัวแบบใหม่จะนำเรารูปภาพป้อนเข้า (Input) ผ่านกระบวนการประมวลผล (Preprocessing) เพื่อปรับรูปภาพให้มีความเหมาะสมกับตัวแบบ โดยขั้นตอนนี้จะมีสองส่วนคือการปรับขนาดรูปภาพ (Resize) และ การประมวลผลที่จำเพาะเฉพาะของตัวแบบ EfficientNet (EfficientNet Preprocessing) หลังจากนั้นภาพที่ผ่านกระบวนการประมวลผลแล้วจะถูกผ่านชั้นการสกัดลักษณะเฉพาะ และ ผ่านชั้นตัวแบบจำแนกประเภทแล้ว จะได้ออกมาเป็นผลลัพธ์การทำนายของตัวแบบใหม่



รูปที่ 1 โครงสร้างโดยคร่าวของตัวแบบใหม่

รายละเอียดของประเภท ขนาดตัวแบบจำแนกประเภทของแต่ละตัวแบบแสดงดังรายละเอียดใน ตารางที่ 5 รายละเอียดของตัวแบบใหม่ที่สร้างขึ้น

ตารางที่ 5 รายละเอียดของตัวแบบใหม่ที่สร้างขึ้น

รายการ	ตัวแบบการจำแนกประเภท	
	ความรู้สึกของรูปภาพ	ลักษณะรูปภาพ
ประเภท	Multiclass Classification	Multiclass Classification
ขนาดรูปภาพป้อนเข้า	224 x 224 (3 Channels)	224 x 224 (3 Channels)
Mini Batch Size	32	32
Optimizer	Adam	Adam
Loss Function	Categorical Cross Entropy	Categorical Cross Entropy

สำหรับรายละเอียดของชั้นตัวแบบจำแนกประเภท (Classification Layer) ของตัวแบบ ตัวแบบการจำแนกประเภทความรู้สึกของรูปภาพ (Image Sentiment Classification) และ ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification) แสดงในตารางที่ 6 รายละเอียดชั้นการจำแนกประเภทของตัวแบบการจำแนกประเภทความรู้สึกของรูปภาพและ ตารางที่ 7 รายละเอียดชั้นการจำแนกประเภทของตัวแบบจำแนกประเภทลักษณะรูปภาพการคำนวณคะแนน ตามลำดับ

ตารางที่ 6 รายละเอียดชั้นการจำแนกประเภทของตัวแบบการจำแนกประเภทความรู้สึกของรูปภาพ

Layer Type	Dimension	Activation Function	Drop out	Batch Normalization
Global Average Pooling 2D	-	-	-	-
Fully Connected	1024	ReLU	30%	Yes
Fully Connected	1024	ReLU	30%	Yes
Fully Connected	1024	ReLU	30%	Yes
Fully Connected	512	ReLU	30%	Yes
Fully Connected	256	ReLU	30%	Yes
Fully Connected	128	ReLU	30%	Yes
Fully Connected	64	ReLU	30%	Yes
Fully Connected (Output)	5	SoftMax	-	-

ตารางที่ 7 รายละเอียดขั้นตอนการจำแนกประเภทของตัวแบบ
จำแนกประเภทลักษณะรูปภาพการคำนวณคะแนน

Layer Type	Dimension	Activation Function	Drop out	Batch Normalization
Global Average Pooling 2D	-	-	20%	Yes
Fully Connected	512	ReLU	30%	Yes
Fully Connected	256	ReLU	30%	Yes
Fully Connected	128	ReLU	30%	Yes
Fully Connected	64	ReLU	30%	Yes
Fully Connected (Output)	20	SoftMax	-	-

3.4 การคำนวณคะแนนความเหมาะสม

3.4.1 การคำนวณคะแนนความเหมาะสมสำหรับ ความรู้สึกรูปภาพ

$$S(P, E) = W(P) \times F(E)^T \quad (1)$$

เมื่อ

$S(P, E)$ คือ คะแนน มีค่าอยู่ในช่วง $[0, 1]$

P คือ เมทริก ขนาด $m \times n$

p_{ij} คือสมาชิกของ P ซึ่งเป็นความน่าจะเป็นของรูป i ที่จะเป็นประเภท j

- i จะมีค่าตั้งแต่ $1 - n$ โดย n คือจำนวนรูปที่มีทั้งหมด
- j จะมีค่าตั้งแต่ $1 - 5$ (มี 5 ความรู้สึก เรียงตามลำดับคือแง่บวกอย่างยิ่ง (Highly Positive), แ่บวก (Positive), กลาง (Neutral), ลบ (Negative) และ ลบอย่างยิ่ง (Highly Negative))

โดยที่ p_{ij} จะมีคุณสมบัติคือ

$$\sum_{j=1}^m p_{ij} = 1; \text{ เมื่อ } i = \{1, 2, \dots, n\} \quad (2)$$

$$\sum_{i=1}^n \sum_{j=1}^m p_{ij} = n \quad (3)$$

E คือ ความรู้สึกที่สนใจซึ่งมีค่าได้ คือ แ่บวก (Positive), กลาง (Neutral) และ ลบ (Negative)

$W(P)$ คือ เมทริกผลรวมความน่าจะเป็นของแต่ละความรู้สึก
ทุกรูปภาพสามารถคำนวณได้จาก

$$W(P) = \left[\sum_{i=0}^n p_{i,1}, \sum_{i=0}^n p_{i,2}, \sum_{i=0}^n p_{i,3}, \sum_{i=0}^n p_{i,4}, \sum_{i=0}^n p_{i,5} \right] \quad (4)$$

$F(E)$ คือ เมทริกตัวคูณปัจจัยสามารถคำนวณได้จาก

$$F(E) = [w_{\text{highly positive}}, w_{\text{positive}}, w_{\text{neutral}}, w_{\text{negative}}, w_{\text{highly negative}}] \quad (5)$$

โดยที่ w น้ำหนักคะแนนในแต่ละความรู้สึกที่สนใจ

กรณี $E = \text{Positive}$

$$w_{\text{highly positive}} = 1$$

$$w_{\text{positive}} = 0.75$$

$$w_{\text{neutral}} = 0.5$$

$$w_{\text{negative}} = 0.25$$

$$w_{\text{highly negative}} = 0$$

กรณี $E = \text{Neutral}$

$$w_{\text{highly positive}} = 0$$

$$w_{\text{positive}} = 0.5$$

$$w_{\text{neutral}} = 1$$

$$w_{\text{negative}} = 0.5$$

$$w_{\text{highly negative}} = 0$$

กรณี $E = \text{Negative}$

$$w_{\text{highly positive}} = 0$$

$$w_{\text{positive}} = 0.25$$

$$w_{\text{neutral}} = 0.5$$

$$w_{\text{negative}} = 0.75$$

$$w_{\text{highly negative}} = 1$$

3.4.2 การคำนวณคะแนนความเหมาะสมสำหรับ ฉากหลังและลักษณะรูปภาพ

$$S(P, E) = \sum_{k=1}^d e_k \quad (6)$$

เมื่อ

$S(P, E)$ คือ คะแนน มีค่าอยู่ในช่วง $[0, 1]$

P คือ เมทริก ขนาด $m \times n$

p_{ij} คือสมาชิกของ P ซึ่งเป็นความน่าจะเป็นของรูป i ที่จะเป็นประเภท j

- i จะมีค่าตั้งแต่ 1 - n โดย n คือจำนวนรูปที่มีทั้งหมด
- j จะมีค่าตั้งแต่ 1 - m โดย m คือจำนวนประเภทของภาพที่มีทั้งหมด

โดยที่ p_{ij} จะมีคุณสมบัติคือ

$$\sum_{j=1}^m p_{ij} = 1; \text{ เมื่อ } i = \{1, 2, \dots, n\} \quad (7)$$

$$\sum_{i=1}^n \sum_{j=1}^m p_{ij} = n \quad (8)$$

d คือ จำนวนประเภทของรูปที่สนใจ เป็นจำนวนเต็ม ในช่วง $[1, m]$ เช่นหากมีประเภทของภาพที่ตัวแบบสามารถทำนายได้ทั้งหมด 10 ประเภท และ เราสนใจประเภทของภาพ 2 ประเภท ที่เป็นซับเซตของภาพที่ตัวแบบสามารถทำนายได้ ดังนั้น $d = 2$

e_k คือคะแนนของแต่ละประเภทของภาพที่สนใจ โดยที่ k มีค่าตั้งแต่ 1 ถึง d สามารถคำนวณได้จาก

$$e_k = \begin{cases} \frac{1}{n} \sum_{i=0}^n p_{ij}, & \frac{1}{n} \sum_{i=0}^n p_{ij} < \frac{1}{d} \\ \frac{1}{d}, & \frac{1}{n} \sum_{i=0}^n p_{ij} \geq \frac{1}{d} \end{cases} \quad (9)$$

3.4.3 การคำนวณคะแนนความเหมาะสมสำหรับวัตถุที่สนใจ

$$S = \frac{I(E)}{n} \quad (10)$$

เมื่อ

S คือ คะแนน มีค่าอยู่ในช่วง $[0, 1]$

$I(E)$ คือ จำนวนภาพที่มี วัตถุที่สนใจ E ปรากฏอยู่ในภาพ

n คือจำนวนรูปที่มีทั้งหมด

3.4.4 การคำนวณคะแนนความเหมาะสมรวม

$$S_{Overall} = \frac{\sum S}{n} \quad (11)$$

เมื่อ

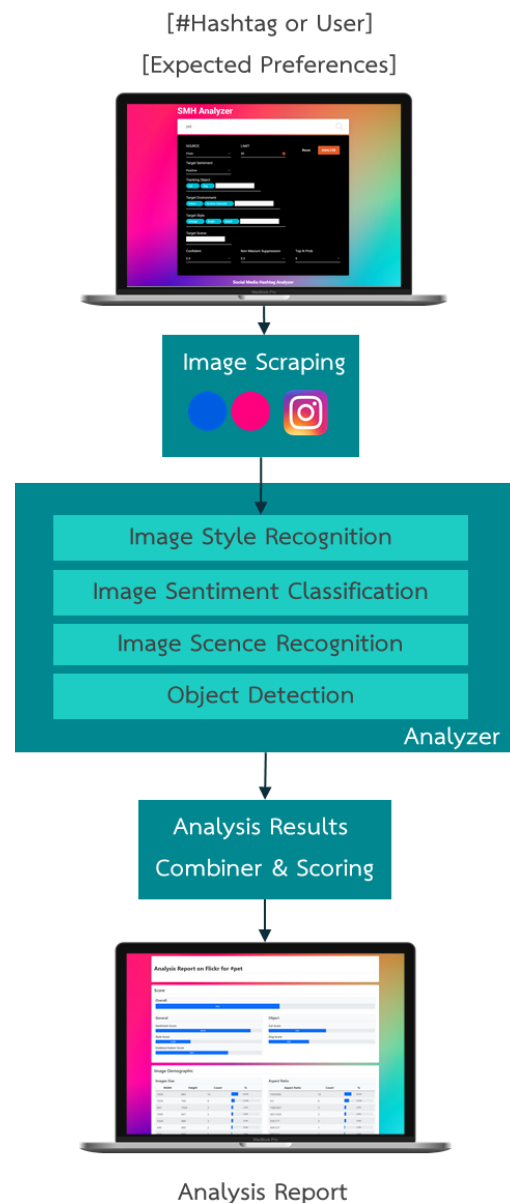
$S_{Overall}$ คือ คะแนนรวม มีค่าอยู่ในช่วง $[0, 1]$

S คือ คะแนน มีค่าอยู่ในช่วง $[0, 1]$

n คือจำนวนคะแนนที่นำมาใช้ทั้งหมด

3.5 การวิเคราะห์และการแสดงผลของระบบ

เพื่อแสดงให้เห็นว่าตัววิเคราะห์ที่สร้างขึ้นสามารถใช้ได้จริง ตัววิเคราะห์นี้จึงถูกครอบด้วยส่วนต่อประสานกับผู้ใช้ (User Interface) ที่พัฒนาขึ้นด้วย ภาษา Python ร่วมกับ Flask web framework และ Bootstrap โดยระบบสามารถวิเคราะห์ได้สองรูปแบบคือ การวิเคราะห์หัตถ์ฉลาก (Hashtag) บน Instagram และ Flickr และการวิเคราะห์ผู้ใช้หรือผู้มีอิทธิพล (Influencer) บน Instagram ได้อย่างอัตโนมัติโดยมีผังการทำงานดัง รูปที่ 2 ผังการทำงานของระบบ



รูปที่ 2 ผังการทำงานของระบบ

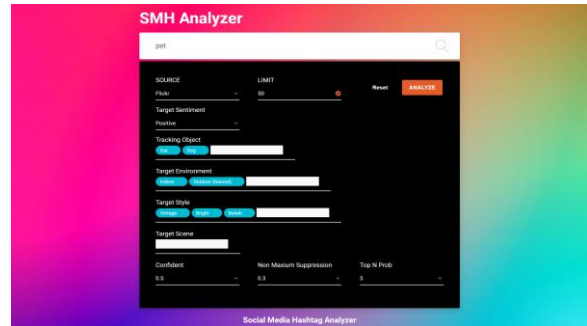
โดยระบบจะเริ่มจากการรับเอาข้อมูลป้อนเข้า (Input) โดยผ่านทางหน้าเว็บดังในรูปที่ 3 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าสำหรับรับข้อมูลเพื่อทำการวิเคราะห์ซึ่งประกอบไปด้วยสองส่วนคือ ส่วนเป้าหมายที่เป็นดัชนีถ้อยคำ (Hashtag) หรือชื่อบัญชีผู้ใช้ (Username) ที่เปิดบัญชีเป็นสาธารณะ โดยระบบสามารถรับดัชนีถ้อยคำ (Hashtag) ของแอปพลิเคชัน Instagram และ Flickr ได้ แต่สำหรับบัญชีผู้ใช้จะสามารถรับได้เพียงแค่ของแอปพลิเคชัน Instagram เท่านั้น พร้อมทั้งต้องระบุจำนวนภาพที่จะนำมาวิเคราะห์ด้วย ส่วนที่สองที่รับคือ ค่าความคาดหวังในด้านต่างๆของชุดรูปภาพที่ต้องการซึ่งมีดังนี้

- ความรู้สึกต่อรูปภาพที่ต้องการ
- วัตถุที่ต้องการให้ปรากฏบนรูปภาพ
- พื้นหลังที่ต้องการ
- บรรยากาศที่ต้องการ (ภายในหรือภายนอกสถานที่)

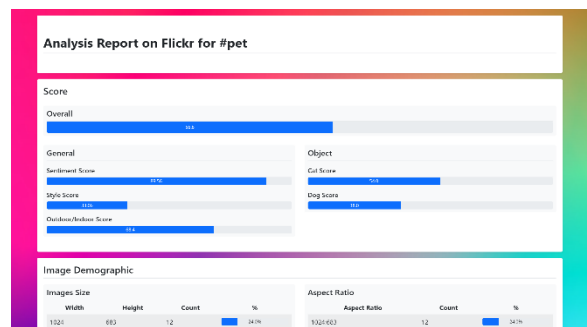
ต่อมาระบบทำการดาวน์โหลดรูปภาพล่าสุดภายใต้ดัชนีถ้อยคำ (Hashtag) หรือ ชื่อบัญชีผู้ใช้ (Username) นั้นๆ โดยอัตโนมัติเพื่อรอการวิเคราะห์ จากนั้นรูปภาพทั้งหมดจะผ่านตัวแบบทั้งหมดที่ประกอบไปด้วย

- 1) ตัวแบบจำแนกประเภทฉากหลัง (Scene Classification Model)
- 2) ตัวแบบการตรวจจับวัตถุ (Object Detection)
- 3) ตัวแบบจำแนกประเภทความรู้สึกรูปภาพ (Image Sentiment Classification)
- 4) ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification)

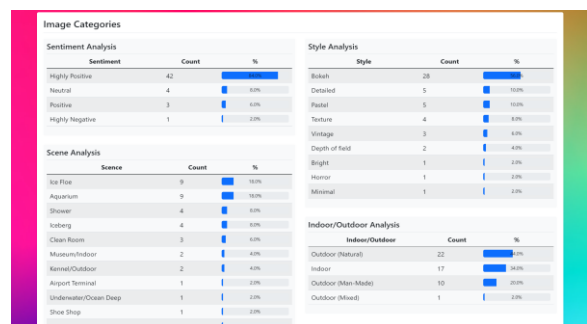
และส่งผลการทำนายทั้งหมดไปยังตัวรวมผลการวิเคราะห์ และคิดค่าคะแนนความเหมาะสม พร้อมทั้งออกรายงานผ่านทางหน้าเว็บเพื่อให้ผู้ใช้งานใช้เพื่อการตัดสินใจ ดังรูปที่ 4 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าแสดงรายงานการวิเคราะห์ (1) และ รูปที่ 5 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าแสดงรายงานการวิเคราะห์ (2)



รูปที่ 3 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าสำหรับรับข้อมูลเพื่อทำการวิเคราะห์



รูปที่ 4 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าแสดงรายงานการวิเคราะห์ (1)



รูปที่ 5 ตัวอย่างส่วนต่อประสานกับผู้ใช้งาน: หน้าแสดงรายงานการวิเคราะห์ (2)

4. ผลการทดลอง

4.1 ผลของตัวแบบใหม่

ในงานวิจัยนี้มีการสร้างตัวแบบเพื่อทำการวิเคราะห์รูปภาพขึ้นสองตัวแบบ ซึ่งรายละเอียดความแม่นยำของตัวแบบเป็นดังนี้

4.1.1 ตัวแบบจำแนกประเภทความรู้สึกรูปภาพ (Image Sentiment Classification)

ผลของความแม่นยำของตัวแบบเมื่อประเมินด้วยชุดข้อมูลสอนคือ 49.36% แต่เมื่อประเมินด้วยชุดข้อมูลทดสอบจะได้ ความแม่นยำ 47.21% สำหรับรายละเอียดความแม่นยำที่วัดด้วยมาตรวัดอื่นอยู่ใน ตารางที่ 8

แม้ว่าจะได้คะแนนความแม่นยำจะไม่สูงนักแต่หากพิจารณาว่า แบ่งความรู้สึกออกเป็น 5 ระดับคือ แง่ลบอย่างยิ่ง (Highly Negative), แ่ลบ(Negative), กลาง(Neutral), แ่บวก (Positive) และ แ่บวกอย่างยิ่ง (Highly Positive) สามารถเปลี่ยนเป็นตัวเลขตั้งแต่ 0 ที่แทนแง่ลบอย่างยิ่ง ถึง 4 ที่แทนแง่บวกอย่างยิ่ง หากนำเอามาคำนวณเพื่อหาค่า MAE และ MAPE จะได้ 1.05 และ 49.98% ตามลำดับ แสดงถึงค่าที่ทำนายได้จะผิดพลาดโดยเฉลี่ยประมาณหนึ่งลำดับและเมื่อพิจารณา ค่าMSE และ MSPE จะได้ 2.68 และ 147.35% ที่สามารถบอกได้ว่าค่าที่ทำนายมีความต่างจากค่าจริงแบบแตกต่างกันมากมีจำนวนน้อย แสดงให้เห็นว่าในการใช้ตัวแบบนี้ให้ ประเมินผลรูปภาพยังคงให้ผลที่สะท้อนความเป็นจริงและมีความน่าเชื่อถือ

ตารางที่ 8 ตารางแสดงค่าความแม่นยำของตัวแบบจำแนกประเภทความรู้สึกรูปภาพโดยประเมินจากชุดข้อมูลทดสอบ

CLASS	MACRO AVG	WEIGHTED AVG
PRECISION	27%	35%
RECALL	30%	47%
F1-SCORE	24%	35%
SUPPORT	13,366	13,366
ACCURACY	47.21%	

4.1.2 ตัวแบบจำแนกประเภทลักษณะรูปภาพ (Image Style Classification)

ผลของความแม่นยำของตัวแบบเมื่อประเมินด้วยชุดข้อมูลสอนคือ 49.16% แต่เมื่อประเมินด้วยชุดข้อมูลทดสอบจะได้ ความแม่นยำ 43.50% สำหรับรายละเอียดความแม่นยำที่วัดด้วยมาตรวัดอื่นแสดงในตารางที่ 9

Error! Reference source not found. ตารางแสดงค่าความแม่นยำของตัวแบบจำแนกประเภทลักษณะรูปภาพโดยประเมินจากชุดข้อมูลทดสอบ

CLASS	MACRO AVG	WEIGHTED AVG
-------	-----------	--------------

PRECISION	41%	41%
RECALL	43%	44%
F1-SCORE	40%	41%
SUPPORT	12,792	12,797
ACCURACY	43.50%	

เนื่องจากภาพที่ใช้ในการสอนตัวแบบนั้นมีความกำกวมเนื่องจาก ภาพหนึ่งๆนั้น สามารถมีได้หลายลักษณะในหนึ่งภาพทำให้ความแม่นยำของตัวแบบไม่สูงนัก แต่เมื่อพิจารณาด้วยตัวบุคคลจะพบว่าภาพส่วนมากมักเป็นลักษณะผสมและผลความน่าจะเป็นที่สูงเป็นสามอันดับแรกเป็นลักษณะที่ผสมกัน หากเราทำการประเมินตัวแบบใหม่โดย หากประเภทของรูปภาพที่ตัวแบบทำนาย ในความน่าจะเป็นสูงสุด 3 อันดับแรกของ มีหนึ่งในสามเป็นค่าที่ตรงกับค่าลักษณะที่ให้ไว้ จะถือว่าตัวแบบทำนายถูก จะได้ ค่าความแม่นยำสูงขึ้นมาก โดยได้ค่าความแม่นยำที่ 68.83% รายละเอียดความแม่นยำที่วัดด้วยมาตรวัดอื่นอยู่ใน ตารางที่ 9

ตารางที่ 9 ตารางแสดงค่าความแม่นยำของตัวแบบจำแนกประเภทลักษณะรูปภาพโดยประเมินจากชุดข้อมูลทดสอบ คิดจาก ความน่าจะเป็นสามอันดับสูงสุด

CLASS	MACRO AVG	WEIGHTED AVG
PRECISION	70%	70%
RECALL	68%	69%
F1-SCORE	68%	68%
SUPPORT	12,792	12,797
ACCURACY	68.83%	

4.2 กรณีศึกษากรณีศึกษาการวิเคราะห์หาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ต ที่เหมาะสม

เพื่อแสดงให้เห็นว่าตัววิเคราะห์ที่สร้างขึ้นสามารถทำการวิเคราะห์หาผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมทางอินเทอร์เน็ตที่เหมาะสมในการเป็นผู้นำเสนอสินค้า (Presenter) ได้ จึงได้ทำการทดลองทดสอบหาผู้มีอิทธิพลใน

สื่อเครือข่ายสังคมทางอินเทอร์เน็ตที่เหมาะสมที่สุดในการเป็นตัวแทนสินค้าสเก็ตบอร์ดที่กำลังเป็นกระแสในปัจุบัน ในกรณีนี้จะวิเคราะห์ภาพที่ถูกลำขึ้นไปยังบัญชี Instagram ของตัวอย่าง 7 ราย ตั้งแต่วันที่ 01/01/2021 ถึง วันที่ 01/06/2021 โดยสองราย errorbxx และ weir____things เป็นผู้ใช้ทั่วไปที่ชื่นชอบสเก็ตบอร์ดมีการนำภาพที่มีสเก็ตบอร์ดขึ้นไปยังบัญชีของตนจำนวนมากอีก 5 รายเป็นดารานักแสดงโดยมีนักแสดงที่มีความชื่นชอบในการเล่นสเก็ตบอร์ดจำนวน 4 รายได้แก่ joeybangkokboy, paloyh, apitsada และ toeyjarinporn โดยอีกหนึ่งราย คือ sirinissirin ที่ไม่ได้เล่นสเก็ตบอร์ด รายละเอียดอื่นๆ เกี่ยวกับบัญชี Instagram จะแสดงในตาราง ตารางที่ 10

ตารางที่ 10 ตารางแสดงรายละเอียดของบัญชี Instagram ตัวอย่าง

ชื่อผู้ใช้	ประเภทผู้ใช้	จำนวนโพสต์ทั้งหมด	จำนวนผู้ติดตาม	จำนวนรูปที่นำมาวิเคราะห์
errorbxx	ทั่วไป	1,215	1,294	54
weir____things	ทั่วไป	2,941.00	4,624	397
joeybangkokboy	ดารา	1,626.00	1.2 ล้าน	32
sirinissirin	ดารา	997.00	1.6 ล้าน	146
paloyh	ดารา	9,953.00	1.2 ล้าน	761
apitsada	ดารา	8,619.00	3.7 ล้าน	374
toeyjarinporn	ดารา	8,967.00	7.1 ล้าน	347

หากกำหนดความต้องการให้ ผู้ที่จะมาเป็นผู้นำเสนอสินค้า (Presenter) จะต้องมีความรู้สีกุภาพเป็นบวก, ลักษณะรูปภาพ (Style) เป็นแบบ Bright และ Noir, สภาพแวดล้อมเป็นแบบกลางแจ้ง ที่มนุษย์สร้างขึ้น มีวัตถุที่สำคัญคือ บุคคลและสเก็ตบอร์ด เพื่อให้เข้ากับแบรนด์ จะได้ผลลัพธ์การวิเคราะห์ ดังตารางที่ 12Error! Reference source not found. ผลการวิเคราะห์ผู้ใช้งานเหมาะสมในการเป็นผู้นำเสนอสินค้า (Presenter)

ตารางที่ 12Error! Reference source not found. ผลการวิเคราะห์ผู้ใช้งานเหมาะสมในการเป็นผู้นำเสนอสินค้า (Presenter)

ชื่อผู้ใช้	คะแนนทั่วไป			คะแนนวัตถุ		รวม
	ความรู้สีก	ลักษณะ	ภายใน/นอก	บุคคล	สเก็ตบอร์ด	
joeybangkokboy	77.32	9.99	36.56	78.1	40.62	48.5
				2		2
sirinissirin	89.41	9.17	33.33	94.5	0	45.2
				2		9
apitsada	87.33	10.96	24.94	78.0	0.27	40.3
				7		1
toeyjarinporn	84.62	8.35	30.17	79.8	2.59	41.1
				3		1
paloyh	82.03	12.38	28.21	65.3	5.46	38.6
				4		8
weir____things	86.28	18.25	32.16	89.4	53.15	55.8
				2		5
errorbxx	71.6	8.72	35.5	92.5	9.26	43.5
				9		3

จากผลข้างต้นจะเห็นว่า หากต้องการเลือกดาราที่เหมาะสมมาเป็นผู้นำเสนอสินค้า (Presenter) เพียงคนเดียวจากดารา 5 ท่าน พบว่า joeybangkokboy เป็นผู้ที่เหมาะสมที่สุดเนื่องจากมีคะแนนการวิเคราะห์รวมที่สูงที่สุด

แต่หากพิจารณาเลือก 2 คน การพิจารณาต้องนำเอาคะแนนวัตถุสเก็ตบอร์ด เข้ามาพิจารณาร่วมด้วย เนื่องจากผู้ที่มีคะแนนรวมเป็นอันดับสอง คือ sirinissirin พบว่าไม่มีรูปที่มีสเก็ตบอร์ดเลย เนื่องจากการคิดคะแนนมาจากหลายส่วน ดังนั้นหากต้องการให้กรองแต่ผู้ที่มีวัตถุที่ต้องการสามารถทำได้โดยการเพิ่มน้ำหนักในการให้คะแนนของวัตถุนั้นๆในสูตรการคำนวณคะแนนได้ แต่อย่างไรก็ตามการตัดสินใจเลือกผู้มีอิทธิพล (Influencer) ในสื่อเครือข่ายสังคมอาจจะเลือกผู้ที่ยังไม่เคยลงรูปที่มีสเก็ตบอร์ดในบัญชีสื่อเครือข่ายสังคมของตนได้ เนื่องจากมีหลายส่วนที่เหมาะสมและ ในบางครั้ง การลงภาพครั้งแรก อาจจะทำให้เกิดกระแสที่แรงกว่าผู้ที่เคยลงรูปแบบดังกล่าวอยู่แล้ว แต่ในกรณีศึกษาที่ใช้เงื่อนไขการตรงกับความต้องการที่ระบุมากที่สุดจึงได้ผลลัพธ์ในแบบดังกล่าว ทั้งนี้ผลการวิเคราะห์สำหรับกลุ่มดารา ให้ผลที่ตรงกับลักษณะที่เป็นจริงของการใช้งาน Instagram จากที่ได้ทำการสังเกตพฤติกรรมการใช้งานจริง

5. สรุปผล

เนื่องจากการสร้างเครื่องมือการวิเคราะห์ที่มีความแม่นยำและให้มีความครอบคลุม (Generalization) กับข้อมูลในความเป็นจริงนั้นทำได้ยาก ข้อจำกัดของตัววิเคราะห์นี้จึงขึ้นกับตัวแบบที่นำมาใช้หากข้อมูลรูปภาพนำเข้ามีลักษณะที่คล้ายกับข้อมูลสอนตัวแบบก็จะมีความแม่นยำที่ดี แต่หากข้อมูลที่นำเข้ามีความแตกต่าง อาจส่งผลต่อประสิทธิภาพของการวิเคราะห์ ตัวแบบที่นำมาใช้มีทั้งตัวแบบสำเร็จรูป และตัวแบบใหม่ ที่มีการทำนายที่จำกัดชนิด (Class) ของผลการทำนายที่อาจจะยังไม่ครอบคลุมการวิเคราะห์ ในแง่มุมเฉพาะ หากต้องนำไปใช้จริงอาจจะต้องทำการเปลี่ยนชุดการสอนของตัวแบบให้เข้ากับลักษณะงานจริง เฉพาะด้านที่ต้องการวิเคราะห์ให้มากขึ้น และ ในระบบปัจจุบันการคิดคะแนนของระบบยังเป็นการคำนวณจากสูตรที่คิดขึ้นหากสามารถที่จะสร้างตัวแบบขึ้นอีกตัวเพื่อใช้สำหรับการคิดคะแนน จะทำให้งานวิจัยมีความน่าสนใจขึ้นอีก

ตัววิเคราะห์นี้ยังมีจุดที่นำไปพัฒนาต่อยอดได้อีก โดยสามารถที่จะเพิ่มตัวแบบในการวิเคราะห์ด้านอื่นๆ เข้าไป เช่น การวิเคราะห์ว่าภาพมีความรุนแรง การพนัน เรื่องที่เกี่ยวข้องกับการเมือง และมีภาพที่เป็นโฆษณา เป็นต้น เพื่อให้มีแง่มุมในการวิเคราะห์เพิ่มขึ้น และหากสามารถเพิ่มเติมการหาวัตถุจากภาพของวัตถุที่ป้อนเข้าไปได้ จะสามารถทำให้แบรนด์ต่างๆ หาความนิยมของกระแสสินค้าตนได้ อีกทั้งหากสามารถทำให้ระบบสามารถระบุเป้าหมายหลายเป้าหมายในคราวเดียว พร้อมทั้งทำการการเปรียบเทียบคะแนนของเป้าหมายโดยอัตโนมัติจะทำให้ผู้ใช้สามารถใช้งานได้สะดวกสบายยิ่งขึ้น อีกทั้งตัวระบบยังสามารถพัฒนาต่อยอดให้เป็นระบบช่วยตรวจสอบสถานะทางเครือข่ายสังคมออนไลน์ (Social Network Monitoring) ที่สามารถวิเคราะห์ถึงเทรนของเครือข่ายสังคมในขณะนั้นได้อีกด้วย

6. Source Code

URL สำหรับดาวน์โหลด Source Code ของงานวิจัยนี้:

<https://github.com/NickyPSCK/Instagram-Hashtag-Analyzer-App>

7. เอกสารอ้างอิง

- [1] Y. Hu, L. Manikonda and S. Kambhampati, "What We Instagram: A First Analysis of Instagram Photo Content and User Types," in *Proceedings of the 8th International Conference on Weblogs and Social Media*, The AAAI Press, pp.595-598, 2014.
- [2] S. Ibba, M. Orrù, F.E. Pani and S. Porru, "Hashtag of Instagram: From Folksonomy to Complex Network," in *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Lisbon, Portugal, pp.279-284, 2015.
- [3] I. Dorsch, "Content description on a mobile image sharing service: Hashtags on Instagram," *Journal of Information Science Theory and Practice.*, vol.6, no.2, pp.46-61, 2018.
- [4] A. Mathes, "Folksonomies - cooperative classification and communication through shared metadata," *Journal of Computer-Mediated Communication.*, vol. 47, 2004.
- [5] Bruns, Axel & Burgess and Jean, "The Use of Twitter Hashtags in The Formation of Ad Hoc Publics," in *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*, pp. 1-9, 2011.
- [6] K. S. Satoh, "Image Sentiment Analysis Using Correlations Among Visual, Textual, and Sentiment Views," in *International Conference on Acoustics, Speech, and Signal Processing IEEE*, 2016.

- [7] V. Gajarla and A. Gupta, "Emotion Detection and Sentiment Analysis of Images," Georgia Institute of Technology, Atlanta, Georgia, 2011.
- [8] S. Karayev, M. Trentacoste, H. Han, A. Agarwala, Trevor Darrell, A. Hertzmann and H. Winnemoeller, "Recognizing Image Style," in *Proceedings British Machine Vision Conference*, BMVA Press, 2014.
- [9] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva and A. Torralba, "Places: A 10 million image database for scene recognition," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1452 - 1464, 2017.
- [10] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *The International Conference on Machine Learning*, Long Beach, California, pp.6105-6114, 2019.
- [11] M. Hussain, J. J. Bird and D. R. Faria, "A Study on CNN Transfer Learning for Image Classification," in *EventUKCI'18: 18th Annual UK Workshop on Computational Intelligence - Nottingham*, United Kingdom, pp.191-202, 2018.
- [12] J. Redmon, S. Divvalay, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA. IEEE, 2016.
- [13] T. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C.L. Zitnick and P. Dollár, "Microsoft COCO: Common Objects in Context," in *European Conference on Computer Vision, Computer Vision – ECCV 2014*, pp. 740-755, 2014.