การหาตำแหน่งตราสินค้านมโคพาสเจอร์ไรส์จากความเห็นผู้บริโภคบนสื่อสังคมออนไลน์

พัชรพล อ่วมโอพาร [1*] และธนชาตย์ ฤทธิ์บำรุง [2]

Received: 19 July 2021; Revised: 20 November 2021; Accepted: 1 December 2021

บทคัดย่อ

ในปัจจุบัน การแข่งขันทางธุรกิจที่เข้มข้น และเทคโนโลยีที่พัฒนาอย่างรวดเร็ว ทำให้ทุกอุตสาหกรรมพิจารณาการสร้างสินค้าใหม่ให้แตกต่างจากคู่แข่ง เพื่อให้สินค้านั้นมีความโดดเด่นกว่าสินค้าอื่นในตลาด และทำให้ผู้บริโภคสามารถจดจำสินค้านั้นได้ เมื่อองค์กรต้องการผลิตและขายสินค้าใหม่ องค์กรมักเริ่มจากการแบ่งกลุ่มตลาด การกำหนดเป้าหมาย และการกำหนดตำแหน่งของตลาด เพื่อให้สินค้านั้นสามารถครองใจผู้บริโภคและอยู่รอดในตลาดได้ การศึกษานี้มีวัตถุประสงค์ในการหาตำแหน่งตราสินค้าของนมโคพร้อมดื่มพาสเจอร์ไรส์ในประเทศไทยจำนวน 7 ตราสินค้าที่วางขายในตลาดด้วยข้อมูลความเห็นผู้บริโภคบนสื่อสังคมออนไลน์ Pantip.com นำความเห็นมาตัดคำ คัดเลือกคำสำคัญ 500 คำ โดยใช้เทคนิคการประมวลผลภาษาธรรมชาติ ก่อนจัดตำแหน่งตราสินค้าโดยใช้เทคนิคการแบ่งสเกลหลายมิติตามมุมมองปัจจัยที่ผู้บริโภคพิจารณาในการบริโภคนมโคพาสเจอร์ไรส์ 8 ด้าน พร้อมเทียบผลกับการรับรู้ตำแหน่งตราสินค้าจากแบบสอบถามที่สร้างขึ้นจากมุมมองปัจจัย 8 ด้านเช่นกัน จากนั้นจึงสร้างกราฟโครงข่ายเชิงความหมายเพื่อพิจารณาความเชื่อมโยงระหว่างคำสำคัญ รวมถึงค้นหาข้อมูลเชิงลึกโดยใช้เทคนิคการตรวจหาเครือข่ายชุมชน พบว่าการจัดตำแหน่งตราสินค้าด้วยการแบ่งสเกลหลายมิติ เมื่อเปรียบเทียบความสัมพันธ์ระหว่างข้อมูลจากแบบสอบถามและข้อมูลจาก Pantip.com มีค่าสหสัมพันธ์ QAP เท่ากับ 0.56 และ pseudo p-value เท่ากับ 0.01

คำสำคัญ: เทคนิคการแบ่งสเกลหลายมิติ, การประมวลผลภาษาธรรมชาติ , การวิเคราะห์โครงข่ายเชิงความหมาย, การตรวจหาเครือข่ายชุมชน

** Corresponding E-mail: pacharapol.oum@stu.nida.ac.th*

[1,2]สาขาการวิเคราะห์ธุรกิจและวิทยาการข้อมูล คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

Pasteurized milk brand positioning from online social media reviews

Pacharapol Oumolarn [1*] and Thanachart Ritbumroong [2]

Received: 19 July 2021; Revised: 20 November 2021; Accepted: 1 December 2021

Abstract

The technological advancement has brought a new level of intense market competition encouraging most players in any industry to differentiate themselves from competititors in order to capture consumers' attention. As a results, various new products and services are being offered to a market. Segmentation, targeting, and positioning is a marketing technique generally utilized to sell new products and services. With the differentiation strategy employed to compete in this intense competition, choosing a unique position for new products and services is pivotal for business. This research explored the market positions of seven pasteurized cow's milk brands in Thailand. Data were collected from a famous social network forum in Thailand, Pantip.com. Comments from the forum were tokenized. The results of tokenization revealed 500 keywords. These keywords were categorized into eight influential factors affecint milk consumption. Brand perceptual maps by the influential factors were constructed using Multidimensional Scaling. Sementic Network and Community Detection techniques were performed on the dataset. From a discussion with industry experts, the results revealed new insights. To validate the results of brand perceptual maps, self-administered questionnaire survey was conducted asking respondents about brand position. There were no significant discrepancies between the results between Multidimensional Scaling and survey methods with QAP Correlation = 0.56 with pseudo p-value = 0.01.

Keywords: Multidimensional Scaling, Natural Language Processing, Semantic Network Analysis, Community Detection

** Corresponding E-mail: pacharapol.oum@stu.nida.ac.th*

[1,2]Business Analytics and Data Science, Graduate School of Applied Statistics, National Institute of Development and Administration

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบัน เมื่อองค์กรต้องการผลิตและวางขายสินค้าใหม่ จะต้องมีการวิเคราะห์ตลาดเสียก่อน ด้วยการแบ่งกลุ่มตลาด การกำหนดเป้าหมาย และกำหนดตำแหน่งทางการตลาด เพื่อให้ผู้บริโภครับรู้การมีอยู่ของสินค้า และตัวสินค้าสามารถอยู่รอดในตลาดได้ (รังสรรค์ เลิศในสัตย์, 2549)

การหาตำแหน่งทางการตลาดตราสินค้าและผลิตภัณฑ์ในประเทศไทยในช่วงที่ผ่านมา ใช้วิธีสำรวจ สัมภาษณ์ก่อนนำคำตอบมาสร้างแผนภาพการรับรู้ของผู้บริโภค (Perceptual map) โดยใช้เทคนิคการแบ่งสเกลหลายมิติ หรือ Multidimensional Scaling - MDS (ชื่นสุมล และ บุษราภรณ์, 2557) วิธีนี้แม้เป็นวิธีที่มีประสิทธิภาพ แต่ก็เสียเวลาและงบประมาณในการทำแบบสอบถามและการสัมภาษณ์ อีกทั้งอาจไม่ได้ความเห็นที่แท้จริงของผู้บริโภคเนื่องจากอคติของผู้ถูกสัมภาษณ์ (ชูชัย สมิทธิไกร, 2562) หากผู้บริโภคไม่มีความสนใจในตราสินค้านั้น หรือมีประสบการณ์ไม่ดีเกี่ยวกับผลิตภัณฑ์ จะไม่เต็มใจให้ความเห็นกับผู้สัมภาษณ์ (Malhotra, 2019)

ทุกวันนี้ การพัฒนาแพลตฟอร์มออนไลน์ในประเทศไทยได้เติบโตแบบก้าวกระโดด ผู้คนเลือกรับข้อมูลข่าวสารผ่านสื่อใหม่ (New Media) เช่น สื่อสังคม (Social Media) และแพลตฟอร์มออนไลน์อื่นๆ มากขึ้น โดยคนไทยใช้งานสื่อสังคมคิดเป็น 91.2 % ของผู้ใช้อินเทอร์เน็ตในประเทศ (สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, 2563, น. 39) ปัจจุบันแพลตฟอร์มออนไลน์นั้นเต็มไปด้วยข้อมูลที่เกี่ยวกับการบริโภคผลิตภัณฑ์ เกือบทั้งหมดอยู่ในรูปแบบข้อมูลไม่มีโครงสร้าง และด้วยปริมาณที่มากสามารถมองให้เป็นข้อมูลจากการสัมภาษณ์กลุ่มแบบเจาะจงขนาดใหญ่ได้ (Netzer, Feldman, Goldenberg & Fresko, 2012) หากสามารถนำข้อมูลดังกล่าวมาประมวลให้อยู่ในรูปแบบที่ประยุกต์ใช้งานได้ องค์กรย่อมสามารถวางแผน ออกแบบ กำหนดคุณลักษณะผลิตภัณฑ์และกลุ่มผู้บริโภค เพื่อให้ผู้บริโภคสามารถจดจำผลิตภัณฑ์ได้ และตัดสินใจเลือกซื้อต่อเนื่องอย่างเต็มใจ

1.2 วัตถุประสงค์ของงานวิจัย

งานวิจัยนี้มีแนวคิดในการระบุตำแหน่งทางการตลาดตราสินค้านมโคพร้อมดื่มพาสเจอร์ไรส์ จำนวน 7 ตราสินค้าที่วางขายในประเทศไทย ประกอบด้วย เมจิ ดัชมิลล์ ไทยเดนมาร์ค โฟร์โมสต์ แดรี่โฮม เอ็มมิลค์ และโชคชัย โดยใช้เทคนิคการแบ่งสเกลหลายมิติ (MDS) พิจารณาตำแหน่งทางการตลาดตราสินค้าในมิติปัจจัยที่มีผลต่อการบริโภคนมแตกต่างกัน 8 ด้าน และใช้เทคนิคการวิเคราะห์โครงข่ายเชิงความหมาย (Semantic Network Analysis - SNA) พิจารณาองค์ประกอบที่เกี่ยวข้องกับตราสินค้า ทั้งหมดใช้ข้อมูลความเห็นของการบริโภคสินค้าจากกระทู้ในสื่อสังคมออนไลน์ Pantip.com เสมือนการสำรวจตลาด เพื่อนำเสนอข้อมูลเชิงลึก ลดการทำแบบสอบถามหรือลงพื้นที่สำรวจความเห็น ก่อนนำข้อมูลไปใช้วางแผนในการจัดกลุ่มตลาดหรือวางสินค้าใหม่ต่อไป

1.3 ประโยชน์ที่คาดว่าจะได้รับ

การใช้สื่อสังคมออนไลน์ทำให้ทั้งองค์ประกอบสำคัญที่เคยนำมาเป็นข้อคำถามในแบบสอบถามและองค์ประกอบที่ไม่เคยทราบมาก่อน ถูกนำมาพิจารณาในการกำหนดตำแหน่งตราสินค้า อีกทั้งการสร้างภาพนิทัศน์ความเชื่อมโยงระหว่างองค์ประกอบ ทำให้ผู้บริหารหรือผู้ใช้งานทั่วไปสามารถมองเห็นข้อมูลเชิงลึกที่ไม่เคยถูกนำมาเป็นข้อคำถามในแบบสอบถาม หรือเป็นปัจจัยใหม่ที่มีผลต่อการซื้อสินค้าของผู้บริโภคที่องค์กรไม่เคยรับรู้มาก่อน

ดังที่กล่าวมาทั้งหมดนี้ องค์กรสามารถระบุความแตกต่างระหว่างตราสินค้าในมิติต่างๆ ก่อนนำไปสู่การคิดค้นสินค้าหรือบริการที่โดดเด่นจากคู่แข่งได้ ทั้งยังตรงความต้องการของผู้บริโภคมากขึ้น กำหนดกลยุทธ์การทำตลาดได้เหมาะสม ลดเวลาและงบประมาณในการสัมภาษณ์ เนื่องจากหากต้องใช้ข้อคำถามและจำนวนตัวอย่างจำนวนมากเพื่อให้มีข้อมูลมากพอจนสะท้อนภาพรวมประชากรในสื่อสังคมออนไลน์ ย่อมหมายถึงจำนวนทรัพยากรที่ต้องใช้มากขึ้นและข้อคำถามที่มีจำนวนมากนั้นสร้างอคติต่อแบบสอบถามเพิ่มมากขึ้นเช่นกัน

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ปัจจัยที่มีผลต่อการบริโภคนม

ผู้วิจัยได้ทบทวนวรรณกรรมเกี่ยวกับปัจจัยที่มีผลต่อการบริโภคนมทั้งในประเทศไทยและต่างประเทศ สามารถสรุปได้ 8 กลุ่ม ดังต่อไปนี้

1. แม่และเด็ก Kurajdova, Petrovicova & Kascakova (2015) ได้ระบุว่าเป็นปัจจัยหนึ่งที่มีผลต่อการบริโภคจากผลการศึกษา และ Hsu & Lin (2006) ยังได้ระบุว่า เด็กมักจะชอบนมที่มีรสอร่อย จึงทำให้เป็นที่นิยม รวมถึงในประเทศไทย ปิยฉัตร ช่างเหล็ก, ยอดมนี เทพานนท์ และณัฐพล พันธุ์ภักดี (2561) ได้มีการกล่าวถึงประเด็นเกี่ยวกับนมโรงเรียน และความสำคัญในประเด็นเกี่ยวกับเด็ก

2. การส่งเสริมการขาย หรือการจัดโปรโมชัน De Alwis, Edirisinghe, & Athauda (2011) และสุภชาติ ชัยณรงค์สิงห์ (2539) ได้ระบุตรงกันว่าเป็นปัจจัยที่มีความสำคัญในการพิจารณาบริโภคนม เช่น การลดราคา ของแถม และการจัดโปรโมชันต่างๆ

3. คุณภาพ ความปลอดภัย ความสะอาด และมาตรฐานการผลิต จากการศึกษาของ Yin et al. (2016) และ Hsu & Lin (2006) ได้กล่าวถึงอย่างชัดเจนจากกรณีศึกษาในประเทศจีน ผู้บริโภคยอมจ่ายมากขึ้นเพื่อซื้อนมที่ผ่านการรับรองด้านคุณภาพจากทางการ เพื่อหลีกเลี่ยงสารปรุงแต่งที่ไม่มีประโยชน์หรือเป็นโทษต่อสุขภาพ

4. ความยากง่ายในการหาซื้อ De Alwis et al. (2011) และสุภชาติ ชัยณรงค์สิงห์ (2539) ได้ระบุตรงกันว่าเป็นปัจจัยที่มีความสำคัญในการพิจารณาบริโภคนม เช่น ประเภทห้างสรรพสินค้า ช่องทางการขายที่แตกต่างกัน

5. การทำกาแฟ ทำอาหาร และขนม Kurajdova et al. (2015) ได้ระบุว่าเป็นปัจจัยที่มีผลต่อการจูงใจให้บริโภคนม โดยผู้บริโภคนำนมมาผสมกับกาแฟ ขนม หรืออาหาร Kurajdova et al. ได้กล่าวเพิ่มเติมว่าบางวัฒนธรรมถือว่านมเป็นส่วนหนึ่งของอาหารที่ต้องมีในการบริโภคแต่ละมื้อ

6. ปัจจัยด้านสุขภาพ (Kurajdova et al., 2015; De Alwis et al., 2011; สุภชาติ ชัยณรงค์สิงห์, 2539) สำหรับปัจจัยในข้อนี้ มีการกล่าวถึงเหมือนกันในหลายการศึกษา ถือเป็นปัจจัยหลักในการพิจารณาบริโภคนม เช่น ความสูง การขับถ่าย คุณค่าทางอาหาร โรคภัย และการออกกำลังกาย

7. ส่วนผสม และสารอาหาร Hsu & Lin (2006) และ De Alwis et al. (2011) ได้ระบุตรงกันว่าเป็นปัจจัยที่มีความสำคัญเชิงบวกในการพิจารณาบริโภคนม เช่น ไขมัน แคลเซียม โดย Hsu & Lin กล่าวเพิ่มเติมอีกว่าผู้บริโภคยอมจ่ายมากขึ้นเพื่อสารอาหารที่มีคุณประโยชน์มากกว่า

8. ราคาและความคุ้มค่า ปิยฉัตร ช่างเหล็ก และคณะ (2561) และ De Alwis et al. (2011) จากผลการศึกษาได้ระบุว่าปัจจัยทางเศรษฐกิจและราคามีผลต่อการบริโภค เช่น รายได้ การตั้งราคา ความคุ้มค่า

2.2 การวิเคราะห์โครงข่ายเชิงความหมาย (Semantic Networks Analysis)

โครงข่ายเชิงความหมาย หรือ Semantic Networks เป็นเทคนิคหนึ่งในการแสดงความสัมพันธ์และความเชื่อมโยงระหว่างคู่ของวัตถุ การรับรู้ หรือพฤติกรรม โดยใช้กราฟโครงข่าย (Doerfel, 1998) เช่น องค์ประกอบทั่วไปของตราสินค้า การมีส่วนร่วมของผู้บริโภคต่อสินค้า และภาพลักษณ์ในสังคม (Henderson, Iacobucci & Calder, 1998)

กราฟโครงข่ายนั้นสามารถสร้างขึ้นได้โดยใช้ Adjacency Matrix ที่มีลักษณะเป็น Matrix จัตุรัส ในบางการศึกษา การสร้าง Semantic Networks ของคำสำคัญจากแหล่งที่สนใจ สามารถทำได้โดยนำคำเหล่านั้นนับความถี่การปรากฏร่วมกันระหว่างคำ 2 คำในเอกสารเดียวกัน นำมาสร้าง Co-occurrence Matrix หรือ Co-word Analysis ซึ่งมีลักษณะคล้ายกับ Adjacency Matrix ก่อนแปลงองค์ประกอบคำสำคัญระหว่างแถวและคอลัมน์ใน Matrix นั้นให้กลายเป็น Node หรือ Vertex (จุดยอด) และเส้นเชื่อมความสัมพันธ์ระหว่างกัน โดยแปลงค่าความถี่ให้กลายเป็นน้ำหนักของเส้น เช่น ความถี่ของการติดต่อสื่อสารผ่านอุปกรณ์ระหว่าง Actor 2 คน (Tabassum, Pereira, Fernandes & Gama, 2018) หรืออาจสร้างเป็น Distance Matrix เพื่อเชื่อมโยงความสัมพันธ์ระหว่างแถวและคอลัมน์ด้วยค่า

Interpersonal Tie ซึ่งสามารถแปลงให้เป็นความยาวของเส้นได้ (Vemprala, Xiong, Liu & Choo, 2019)

สำหรับผลลัพธ์เชิงประจักษ์ของการเชื่อมโยงความสัมพันธ์ระหว่างตราสินค้าและองค์ประกอบสินค้า Aaker (1996, น.94), Henderson et al. (1998) และ Peter & Olson (2010, น.55) ได้แสดงกราฟโครงข่ายซึ่งประกอบด้วยตราสินค้าร่วมกับองค์ประกอบต่างๆ Netzer et al. (2012) ได้ให้ความเห็นที่เกี่ยวข้องกับการเชื่อมโยงความสัมพันธ์ไว้ว่า “ตราสินค้า 2 ตัวที่เชื่อมโยงใกล้กันมีแนวโน้มว่าเกิดจากความจำระยะยาวที่ผู้คนนึกถึงบ่อย ทำนองเดียวกันองค์ประกอบของสินค้าใดๆ ที่เชื่อมโยงใกล้กับตราสินค้า มักปรากฏขึ้นในประโยคร่วมกันกับชื่อตราสินค้าบ่อยครั้ง”

2.3 Similarity Measure

ในการศึกษาเกี่ยวกับ Semantic Network Analysis ของตราสินค้า มีการใช้ Similarity Measure หรือมาตรวัดค่าความคล้ายคลึงสำหรับการ Normalize ความสัมพันธ์ระหว่างองค์ประกอบหรือคำสำคัญของตราสินค้าใน Co-occurrence Matrix เพื่อป้องกันผลกระทบจากอคติที่เกิดจากขนาดตลาด (Market size bias) เนื่องจากแต่ละตราสินค้ามีจำนวนการกล่าวถึงอย่างไม่สมดุลกัน (Malhotra & Bhattacharyya, 2019) จากการทบทวนวรรณกรรม งานวิจัยมีการใช้ Similarity Measure สำหรับการสร้าง Semantic Network ดังต่อไปนี้

2.3.1 Jaccard Index

เป็น Similarity Measure บนพื้นฐานของเซต โดยพิจารณา Cardinality ของเซตองค์ประกอบหรือคำสำคัญ 2 ตัวที่นำมาเทียบกัน ดังแสดงในสมการที่ (1)

$$J(A,B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \qquad (1)$$

จากสมการที่ (1) *A* และ *B* คือ องค์ประกอบหรือคำสำคัญ A และ B ที่ปรากฏในทุกเอกสาร ค่า $A \cap B$ คือ องค์ประกอบหรือคำสำคัญ A และ B ปรากฏพร้อมกันในเอกสารเดียวกัน

2.3.2 Lift

เป็น Similarity Measure บนพื้นฐานของความน่าจะเป็น ซึ่งมีสมการแตกต่างจาก Lift Ratio ที่ใช้ในทฤษฎี Association Rules ดังแสดงในสมการที่ (2)

$$Lift\ (A,B) = \frac{P(A \cap B)}{P(A)\ P(B)'} \qquad (2)$$

โดย $P(A \cap B)$ หมายถึงความน่าจะเป็นที่องค์ประกอบหรือคำสำคัญ A และ B ปรากฏพร้อมกันในเอกสารเดียวกัน $P(A)$ และ $P(B)$ คือความน่าจะเป็นที่องค์ประกอบหรือคำสำคัญ A และ B ปรากฏขึ้นจากทุกเอกสาร และค่า $P(B)'$ เท่ากับ 1-$P(B)$ ถ้าหากว่าตัวส่วน $P(A)$ เป็น 0 ค่า Similarity Measure ที่ได้จะเป็นค่าอนันต์ ซึ่งการเกิด 0 นั้นสามารถเกิดจากกรณีองค์ประกอบ A ไม่มีปรากฏขึ้นจริงในทุกรายการเอกสาร

2.3.3 Inclusion Index

เป็น Similarity Measure ที่คิดค้นโดย Qin He ในปี 1999 เพื่อนำมาทำ Co-word Analysis ระหว่างคำสำคัญ ดังแสดงในสมการที่ (3)

$$I\ (A,B) = \frac{C_{AB}}{min(C_A, C_B)} \qquad (3)$$

C_{AB} = จำนวนเอกสารที่องค์ประกอบหรือคำสำคัญ A และ B ปรากฏพร้อมกันในเอกสารเดียวกัน และ $min(C_A, C_B)$ เป็นฟังก์ชันเลือกค่าต่ำสุดของจำนวนเอกสารที่องค์ประกอบหรือคำสำคัญ A และ B ปรากฏทั้งหมด อย่างไรก็ตาม Similarity Measure นี้ยังคงมีปัญหาการเกิดค่าอนันต์ลักษณะคล้ายกันกับ Lift ถ้าค่าในฟังก์ชัน min = 0

2.3.4 Cosine Similarity

เป็น Similarity Measure ที่พิจารณาองศาบน Inner Product Space ระหว่างองค์ประกอบหรือคำสำคัญ A และ B นิยมประยุกต์ใช้กับงานหลายประเภท เช่น ตัวแบบช่วยแนะนำสินค้า (Product Recommendation) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ดังแสดงในสมการที่ (4)

$$cos(\theta) = \frac{AB}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i}\sqrt{\sum_{i=1}^{n} B_i}} \quad (4)$$

2.3.5 Association Strength

เป็น Similarity Measure อีกรูปแบบที่คิดค้นและปรับปรุงโดย Van Eck & Waltman (2010) ใช้กับการแสดงผลในเทคนิค Visualization of Similarity (VOS) สามารถใช้จำนวนการปรากฏขององค์ประกอบหรือคำสำคัญ i และ j ในเอกสารทั้งหมด หรือ ผลรวมจำนวนการปรากฏร่วมกันระหว่างองค์ประกอบหรือคำสำคัญที่ได้คำนวณลงใน Matrix มาแล้ว ดังแสดงในสมการที่ (5)

$$S_{ij} = \frac{C_{ij}}{w_i, w_j} \quad (5)$$

โดยกำหนดให้ S_{ij} คือ Similarity Measure, C_{ij} คือจำนวนการปรากฏพร้อมกันขององค์ประกอบหรือคำสำคัญ i กับ j, w_i คือ จำนวนการปรากฏองค์ประกอบหรือคำสำคัญ i และ w_j คือจำนวนการปรากฏองค์ประกอบหรือคำสำคัญ j

2.4 Community Detection แบบ Modularity Base

เป็นวิธีการแบ่งกลุ่มขององค์ประกอบในกราฟโครงข่ายที่คิดค้นโดย Newman (2004) เพื่อเป็นวิธีทางเลือกนอกเหนือจาก Community Detection แบบ Girvan-Newman ที่ Newman เองได้ร่วมคิดค้นขึ้นในปี 2002 ต่อมาได้มีนักวิจัยหลายคนได้นำแนวคิดนี้ไปประยุกต์เป็นอัลกอริทึมใหม่ เช่น Louvain Community Detection (Blondel et al., 2008) หรือ Visualization of Similarity Clustering (Waltman, Van Eck & Noyons, 2010)

Modularity เป็นกระบวนการหาค่าเหมาะสมที่สุดอย่างหนึ่ง ตัว Modularity เป็นค่า Scalar ที่มีค่าระหว่าง -1 ถึง 1 หาโดยวัดความหนาแน่นของจำนวนจุดเชื่อมต่อภายใน Community แล้วนำไปเปรียบเทียบกับจำนวนจุดเชื่อมต่อภายในของ Community อื่น เพื่อสรุปผลการแบ่งกลุ่มดังสมการที่ (6)

$$Q = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j) \quad (6)$$

กำหนดให้ Q คือ Modularity, A_{ij} คือน้ำหนักของเส้นระหว่าง Vertex i และ j, k_i และ k_j คือผลรวมของน้ำหนักเส้นที่เชื่อมต่อกับ Vertex i, j ในกราฟ, (c_i, c_j) คือ Community ที่กำหนดไว้ที่ Vertex i และ j, δ เป็นฟังก์ชัน δ (u, v) ที่มีค่าเท่ากับ 1 ถ้า u = v แต่ถ้าไม่ใช่ δ มีค่าเท่ากับ 0 และ m มีค่าดังสมการที่ (7)

$$m = \frac{1}{2}\sum_{j} A_{ij} \quad (7)$$

เมื่อหาค่า Modularity ได้แล้ว ค่า Q จะถูกใช้เป็นตัวเทียบคุณภาพของ Community Detection โดยใช้เทคนิคการหาค่าเหมาะสมที่สุด เพื่อให้ได้ค่า Q ที่ต่ำที่สุด วนรอบคำนวณค่า Q ซ้ำจนครบทุก Vertex ซึ่งวิธีการประเมินค่า Q นั้นแตกต่างกันไปตามวิธีของนักวิจัยที่คิดค้นอัลกอริทึม อาจถือได้ว่า Modularity เป็นฟังก์ชันวัตถุประสงค์ (Objective function) ในเทคนิคการหาค่าเหมาะสมที่สุดเพื่อใช้หา Community

2.5 Visualization of Similarity (VOS)

Van Eck & Waltman (2006) ได้คิดค้นวิธีการนำเสนอกราฟโครงข่ายขนาดใหญ่ ชื่อว่า Visualization of Similarities (VOS) มีวัตถุประสงค์ในการแก้ไขปัญหาการซ้อนทับกันขององค์ประกอบในกราฟโครงข่ายที่มีจำนวนมาก แต่แรกเดิมนั้น VOS ใช้สำหรับสร้างกราฟโครงข่ายความสัมพันธ์การอ้างอิงบรรณานุกรมงานวิจัย ซึ่งมีข้อมูลจำนวนมหาศาล โดยผู้วิจัยแต่ละท่าน มีจำนวนงานวิจัยและจำนวนการอ้างอิงไม่เท่ากัน ทั้งนี้ VOS ยังคงใช้ Co-occurrence Matrix เป็นข้อมูลตั้งต้นในการสร้างกราฟโครงข่าย เหมือนกับการสร้างกราฟโครงข่ายทั่วไป แต่จะใช้ Similarity Measure แบบ Association Strength ดังที่ได้กล่าวไปในข้อที่ 2.3.5 แล้ว

การสร้างกราฟโครงข่ายจาก Co-occurrence Matrix มี 3 ขั้นตอน

1. สร้าง Co-occurrence Matrix คำนวณ Similarity Measure โดยใช้ Association Strength

2. Mapping แต่ละ Vertex องค์ประกอบ โดยให้ค่า Similarity สูงอยู่ใกล้กัน และ Similarity น้อยอยู่ห่างกัน ดำเนินการหาค่าเหมาะสมที่สุดโดยให้ผลรวมของ Square Euclidean Distance มีค่าน้อยที่สุดระหว่างคู่องค์ประกอบหรือคำสำคัญ

3. ดำเนินการ Translation, Rotation และ Reflection

เมื่อสร้างกราฟโครงข่ายเชิงความหมายเรียบร้อยแล้ว VOS จะทำ Community Detection องค์ประกอบภายในต่อ โดยใช้วิธีแบบ Modularity Base ดังที่ได้กล่าวไปใน 2.4 เพียงแต่มีการปรับเล็กน้อยตามการศึกษาของ Waltman et al. (2010) โดยมีการใช้พารามิเตอร์พิเศษเพิ่มเติม ด้วยการเพิ่มส่วนสำหรับหาค่า Resolution γ ที่ได้จากส่วนกลับของค่า Distance ระหว่าง Vertex 2 จุด และนำค่า γ ไปแทนค่าในสมการวัตถุประสงค์ ดังสมการที่ (8)

$$V = \frac{1}{2m}\sum_{i<j}\delta(x,x_j)w_{ij}\left[c_{ij}-\gamma\frac{c_i c_j}{2m}\right] \qquad (8)$$

โดย $w_{ij}=\frac{2m}{c_i c_j}$ และ $distance_{ij}=\begin{cases}0 & if\ x_i = x_j\\ \frac{1}{\gamma} & if\ x_i \neq x_j\end{cases}$

ค่า Resolution ของสมการวัตถุประสงค์นี้ มีไว้เพื่อป้องกันปัญหาการเกิด Community ที่เล็กเกินไปจนทำให้ไม่สื่อความหมาย ซึ่ง Community ที่มีขนาดเล็กนั้นจะถูกรวบเข้าไปใน Community ที่ใหญ่กว่า

2.6 การแบ่งสเกลหลายมิติ (Multidimensional Scaling)

การแบ่งสเกลหลายมิติ หรือ Multidimensional Scaling (MDS) เป็นกระบวนการนำเสนอการรับรู้ ความชอบ ความสัมพันธ์ทางจิตวิทยาระหว่างสิ่งเร้า (Stimuli) ด้วยกราฟพิกัดเรขาคณิต การคำนวณตำแหน่งวัตถุที่สนใจนั้นใช้ตาราง Proximity Matrix สรุป Dissimilarity Measure ระหว่างองค์ประกอบ หรือใช้เป็นค่า Distance ระหว่างองค์ประกอบที่สนใจ ค่าดังกล่าวสามารถแปลงค่ามาจาก Similarity Measure ที่กล่าวถึงในข้อที่ 2.3 ได้ จากนั้นนำเสนอข้อมูลผ่านกราฟพิกัดเรขาคณิตที่เรียกว่า Spatial map หรือ Perceptual map

Malhotra (2019) ได้ระบุว่าปัจจุบันการแบ่งสเกลหลายมิตินั้นนิยมนำมาสร้างตำแหน่งทางการตลาดระหว่างวัตถุที่สนใจในมิติเฉพาะด้านสำหรับการวิเคราะห์ตลาด เช่น การวัดระดับภาพลักษณ์องค์กร การแบ่งกลุ่มตลาด การประเมินประสิทธิผลของการโฆษณา การเลือกช่องทางการขาย และการวิเคราะห์ราคา ตำแหน่งทางการตลาดสามารถช่วยระบุความแตกต่างของการกำหนดราคาในตลาดของตราสินค้าคู่แข่งได้ และทราบว่าสินค้าของตนเองหากเข้าตลาดไปแล้วจะต้องแข่งขันกับใครในมิติดังกล่าว

กระบวนการสร้างตัวแบบ MDS มีการประเมิน Goodness-of-fit ของตัวแบบ โดยใช้ค่า Stress ในส่วนนี้นิยมใช้ Kruscal's Stress ดังสมการที่ (9)

$$Stress = \sqrt{\frac{\sum\left(d_{ij} - \hat{d}_{ij}\right)^2}{\sum d_{ij}^{\ 2}}} \qquad (9)$$

โดย d_{ij} คือค่า Distance ระหว่างสิ่งเร้า 2 ตัวใน Proximity Matrix และ $\hat{d}_{ij}$ คือค่า Distance ระหว่างสิ่งเร้า 2 ตัว ที่ตัวแบบ MDS คำนวณตำแหน่งออกมา เกณฑ์การประเมิน Goodness-of-fit นั้น ถ้าค่ามากกว่า 0.2 ขึ้นไป = ไม่ดี, 0.1-0.2 = พอใช้, 0.05-0.099 = ดี, 0.025 = ดีมาก และ 0.000 สมบูรณ์แบบ

ในด้านจำนวนตัวอย่างกับการแบ่งสเกลหลายมิติ ไม่ได้มีการระบุจำนวนตัวอย่าง (n) ที่เหมาะสมสำหรับการทำแบบสอบถามการรับรู้ แต่เมื่อทบทวนวรรณกรรมแล้ว พบว่ามีงานวิจัยของ Rodger (1991) ได้ทำการทดลองหาว่าควรใช้ n เท่าใดจะเหมาะสมกับการทำ MDS โดยประเมิน Goodness-of-fit ด้วยค่า Stress ดำเนินการสร้างแบบจำลอง Monte-Carlo เพื่อหา n ที่มีผล Stress ดีที่สุด พบว่า n ระหว่าง 6-15 คน จะให้ผล Stress ได้อยู่ในระดับพอใช้จนถึงดี n ที่มากกว่านั้นไม่ได้ทำให้ Stress ลดลง หรือ เพิ่มขึ้น แต่กลับมีแนวโน้มคงที่ นอกจากนี้ Howard (1998) ได้อ้างอิงผลการศึกษาของ Rodger เช่นกัน แต่ระบุเพิ่มเติมว่าควรเลือกตัวอย่างให้มีความหลากหลายครอบคลุมทุกมิติ

2.7 Quadratic Assignment Procedure (QAP)

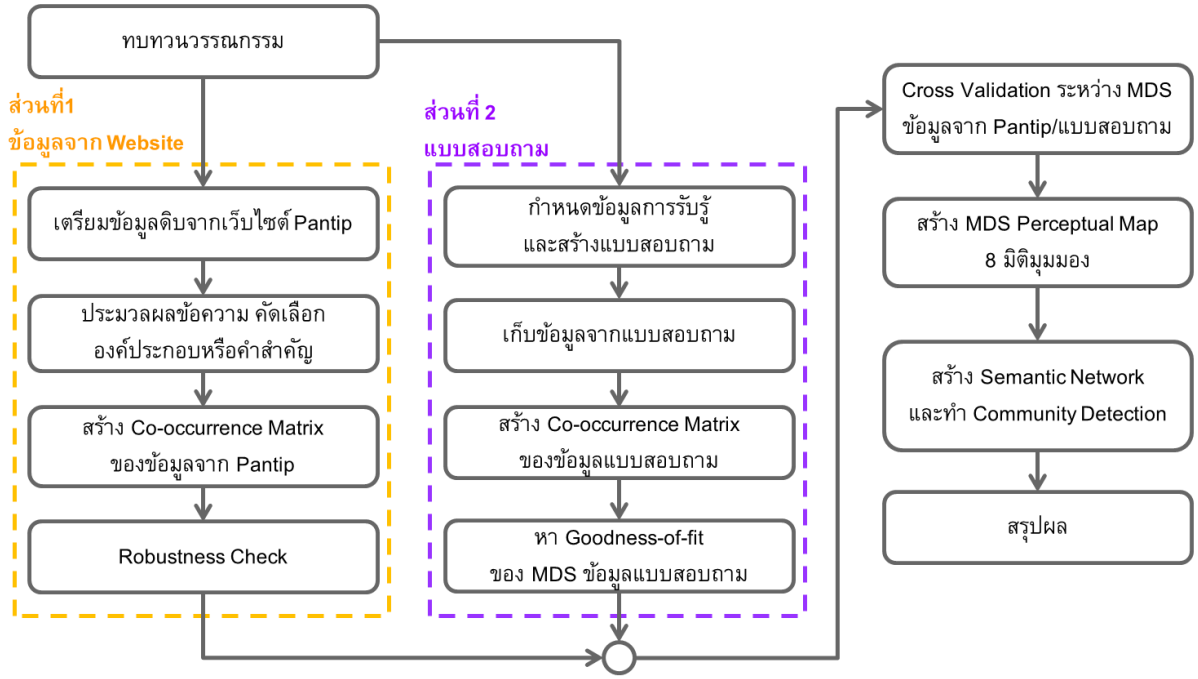
เป็นวิธีทดสอบสหสัมพันธ์แบบหนึ่งที่ใช้ใน Semantic Network Analysis และ Social Network Analysis Simpson (2001) ได้นิยามวิธีการนี้ไว้ว่า Quadratic Assignment Procedure (QAP) ใช้เพื่อวิเคราะห์ความสัมพันธ์ระหว่างแถวและคอลัมน์ของตารางข้อมูลแบบ Dyadic (ตารางข้อมูลที่แสดงความสัมพันธ์คู่ต่อคู่) เช่น Proximity Matrix ด้วยการใช้วิธีจำลองการเรียงสับเปลี่ยนข้อมูลในตาราง (QAP Permutations) เพื่อทดสอบค่าสหสัมพันธ์ระหว่างแถวและคอลัมน์ของตาราง Dyadic อย่างไรก็ตาม ตัวองค์ประกอบในแถวกับคอลัมน์นั้นอาจจะไม่มีความสัมพันธ์ระหว่างกันจริง จึงเป็นข้อที่ต้องระวังในการตีความ การทดสอบสหสัมพันธ์ ใช้ค่า Pseudo p-value ในการทดสอบสมมติฐานค่าสหสัมพันธ์ สามารถกำหนดสมมติฐานได้ดังนี้

H_0: No relationship between 2 matrices

H_1: Otherwise

3. ขั้นตอนวิธีดำเนินการวิจัย

การศึกษานี้อิงขั้นตอนวิธีดำเนินการวิจัยที่ใช้ร่วมกันโดยหลักจาก Netzer et al. (2012), ชื่นสุมล และ บุษราภรณ์ (2557), Ringel & Skiera (2016), Vemprala et al. (2019) และ Malhotra & Bhattacharyya (2019) เน้นการปรับปรุงและต่อยอดการศึกษาของ ชื่นสุมล และ บุษราภรณ์ และ Netzer et al. เป็นหลัก ผู้วิจัยเห็นว่าแนวทางการสร้าง Percetual Map ของ ชื่นสุมล และ บุษราภรณ์ มีความเหมาะสมเนื่องจากสามารถเลือกสรุปการรับรู้ตามกลุ่มปัจจัยหรือสิ่งเร้าที่สนใจได้ แต่จำนวนสิ่งเร้านั้นมีไม่มากเนื่องจากสร้าง Percetual Map จากแบบสอบถาม ผู้วิจัยจึงประยุกต์ใช้ข้อมูลสิ่งเร้าจากคำสำคัญที่เลือกมาจาก Pantip.com แทน และแสดงความสัมพันธ์ระหว่างตราสินค้าและองค์ประกอบหรือสิ่งเร้าเพิ่มเติมนอกเหนือการใช้ Percetual Map โดยผสมผสานขั้นตอนดำเนินการและจุดแข็งของงานวิจัยทั้ง 5 งานที่ได้กล่าวถึงในข้างต้น ทั้งนี้ เพื่อให้เข้าใจภาพรวมได้ง่าย ผู้วิจัยได้นำเสนอ Flow Chart ขั้นตอนวิธีดำเนินการวิจัยโดยรวม ดังในรูปที่ 1



รูปที่ 1 Flow Chart ภาพรวมของขั้นตอนวิธีดำเนินการวิจัย

3.1 ขั้นตอนการเตรียมข้อมูลดิบ

ผู้วิจัยได้สร้างโปรแกรมสำหรับการเก็บ URL ของกระทู้ที่มีการกล่าวถึงตราสินค้าด้วยโปรแกรมภาษา Python โดยใช้ URL ของการค้นหากระทู้และข้อความของ Pantip ที่มี Prefix เป็น https://pantip.com/search?q=xxx โดย xxx คือคำค้นที่ต้องการค้นหา เช่น https://pantip.com/search?q=นมเมจิ ดำเนินการลักษณะนี้ให้ครบทั้ง 7 ตราสินค้า โดยสามารถปรับคำค้นให้เหมาะสมได้กับบริบทของชื่อตราสินค้านั้น หากการค้นหาในแต่ละครั้งพบ URL กระทู้ที่เหมือนกันให้ลบ URL กระทู้ที่ปรากฏซ้ำออก เมื่อได้ URL กระทู้ที่ต้องการทั้งหมดแล้ว ทำการ Scraping ทีละ URL โดยเก็บเนื้อหาทั้งหัวกระทู้ และความเห็นแยกกัน ไม่ได้เก็บข้อมูลผู้เขียนข้อความ จากนั้นบันทึกข้อมูลทั้งหมดลงฐานข้อมูล MongoDB เมื่อทำการสำรวจข้อมูลพบว่า มีจำนวนกระทู้รวม 4,995 กระทู้ และจำนวนความเห็นทั้งหมด 225,360 ความเห็น ตั้งแต่วันที่ 3 มีนาคม 2556 ถึงกระทู้ใหม่ที่สุด 20 มีนาคม 2564

3.2 ขั้นตอนการประมวลผลข้อความ และคัดเลือกองค์ประกอบสำคัญ

ผู้วิจัยทำการลบกระทู้และความเห็นที่ถูกระงับการแสดงผล หรือ ถูกลบออกเพื่อรักษาบรรยากาศการสนทนาด้วยโปรแกรมภาษา Python เนื่องจาก Pantip ไม่ลบ URL ให้หายไปแต่ใช้ข้อความระบุว่ากระทู้หรือความเห็นนั้นถูกลบตามด้วยเหตุผลในการลบ รวมถึงลบความเห็นที่ถูกเผยแพร่ซ้ำโดยไม่ได้ตั้งใจ และความเห็นที่ไม่มีผู้ใช้เข้าไปกรอกความเห็น

เมื่อลบกระทู้และความเห็นเสร็จ ใช้ PyThaiNLP (Phatthiyaphaibun et al., 2016) ในการตัดคำ (Tokenized) โดยใช้อัลกอริทึม Newmm เพื่อให้เหมาะสมกับบริบทของ Pantip และใช้ระยะเวลาในการตัดข้อความที่ไม่มากจนเกินไป การประมวลผลใช้คำภาษาไทยเป็นหลัก ถ้ามีภาษาอังกฤษให้แปลงเป็นภาษาไทยให้มากที่สุด ยกเว้นชื่อห้างสรรพสินค้าและชื่อสารอาหารบางตัว ร่วมกับการสร้าง Customized Dictionary เฉพาะ Domain ของอุตสาหกรรมนม ประกอบด้วย ชื่อตราสินค้า ชื่อผลิตภัณฑ์ และองค์ประกอบที่สำคัญอื่น เมื่อตัดคำแล้ว ดำเนินการสร้าง Label ตราสินค้า และ Domain ของข้อความจากข้อความที่ถูกตัด เพื่อระบุว่ากระทู้หรือความเห็นนั้นกล่าวถึงชนิดของนมและชื่อตราสินค้าใดบ้าง ดัดแปลงตามวิธีของ Lee & Bradlow (2011) เพื่อนำไปใช้เป็นตัวกรองเอกสารที่ไม่ต้องการออก ทั้งนี้การระบุ Label เป็นการระบุโดยตรงจากคำสำคัญที่ปรากฏแล้วตรงตามเงื่อนไขแบบตรงไปตรงมา พิจารณาจากข้อความในหัวกระทู้ หรือในความเห็น เมื่อกรองเอกสารตามเงื่อนไขแล้ว เหลือกระทู้เพียง 1,077 กระทู้ และ 12,546

ความเห็น ด้วยเงื่อนไขของ Domain นมพาสเจอร์ไรส์เท่านั้น

หลังจากกรองเอกสารออก สร้าง Bag of Words ของคำทั้งหมดพร้อมสร้าง LDA Model ด้วย LDAvis (Sievert & Shirley, 2014) เพื่อช่วยคัดเลือกชุดคำที่จะนำมาเป็นองค์ประกอบสำหรับการสร้าง Co-occurrence Matrix โดยไม่ได้มีเป้าหมายในการสร้างตัวแบบหัวข้อ (Topic Modeling) แต่เพื่อลดจำนวนคำที่ถูกตัดให้เหลือแต่คำที่มีความน่าจะเป็นของการปรากฏในทุกเอกสารไม่มากและไม่น้อยเกินไป ลดโอกาสการเกิดค่าอนันต์ของ Similarity Measure ระหว่างการสร้าง Co-occurrence Matrix ก่อนนำคำทั้งหมดที่ถูกเลือกมาจัดหัวข้อของคำด้วยมือ พิจารณาตามปัจจัยที่มีผลต่อการบริโภคนม 8 กลุ่มปัจจัย หากความหมายกำกวมหรือไม่ทราบ ให้จัดเข้ากลุ่มที่ 9 คือกลุ่มที่ความหมายไม่เข้าพวก เพื่อนำไปหาข้อมูลเชิงลึกในขั้นตอนการทำ Community Detection ต่อไป สำหรับคำสำคัญทั้งหมดที่ได้คัดเลือกมา มีทั้งหมด 500 คำ ใน 9 กลุ่มหัวข้อ เหตุผลที่จัดหัวข้อของคำด้วยมือ เนื่องจากต้องการให้ข้อมูลบางส่วนของการหาตำแหน่งตราสินค้าอยู่ในรูปแบบเปรียบเทียบกันได้กับสิ่งเร้าในแบบสอบถามที่จำเป็นต้องกำหนดกลุ่มของคำถามการรับรู้ไว้ล่วงหน้าก่อน (A Priori) ในขณะที่ข้อความที่มาจาก Pantip เป็นประสบการณ์ของตนหรือคนอื่นที่ไม่ได้กำหนดหัวข้อตายตัว (A posteriori)

3.3 ขั้นตอนการสร้างแบบสอบถาม และการเตรียมข้อมูลการรับรู้จากแบบสอบถาม

ถึงแม้ว่าการใช้ข้อมูลจากสื่อสังคมออนไลน์นั้นมีเป้าหมายเพื่อลดการทำแบบสอบถาม แต่ในการศึกษานี้ยังคงใช้เพื่อแสดงกระบวนการ Cross Validation ของผลลัพธ์ระหว่างตำแหน่งทางการตลาดตราสินค้าจาก Pantip.com กับแบบสอบถามให้เป็นที่ประจักษ์ โดยเริ่มจากกำหนดชุดสิ่งเร้าจากปัจจัยที่มีผลต่อการบริโภคนม 8 กลุ่ม ก่อนเริ่มสร้างคำถาม แบบสอบถามนั้นแบ่งออกเป็น 3 ส่วน ประกอบด้วย รูปภาพประกอบซึ่งเป็นรูปภาพตัวอย่างสินค้าที่วางขายในตลาด ตัวอย่างการกรอกแบบสอบถาม และส่วนสุดท้ายคือช่องสำหรับกรอกข้อมูลการรับรู้ตราสินค้าจากสิ่งเร้า โดยใช้ Scale แบบ Attribute Rating พร้อมระบุความหมายของ Scale อย่างชัดเจนในทิศทางเดียวกัน จำนวนตัวอย่างที่ใช้ในการสำรวจ ใช้เพียง 40 คน เป็นพนักงานในโรงพยาบาลรามาธิบดี เขตราชเทวี จังหวัดกรุงเทพมหานคร เหตุผลที่ใช้ตัวอย่างเพียง 40 คนนั้น เนื่องจากพบว่า Rodger (1991) สรุปไว้ว่าจำนวนตัวอย่างระหว่าง 6-15 คนจะให้ค่า Stress ซึ่งเป็นค่า Goodness-of-fit ของตัวแบบ MDS อยู่ในเกณฑ์ดี ทำให้ถึงแม้แบบสอบถามที่ตอบกลับมาไม่สมบูรณ์และถูกตัดออกถึง 50% จำนวนตัวอย่างนั้นยังคงสอดคล้องกับการศึกษาของ Rodger ผู้วิจัยได้จัดทำแบบสอบถามและดำเนินการทดสอบกับกลุ่มตัวอย่างที่ไม่ได้อยู่ในกลุ่มสำรวจ จำนวน 5 คน เพื่อตรวจสอบคำผิดและความเข้าใจของตัวแบบสอบถาม จากนั้นจึงดำเนินการสุ่มแบบตามสะดวก ร่วมกับการสุ่มแบบเจาะจง เมื่อได้ข้อมูลกลับมา ทำ Feature Scaling กับผลลัพธ์ ด้วยวิธี Min-Max Normalization

3.4 ขั้นตอนการสร้าง Co-occurrence Similarity Matrix

หลังจากรวบรวมข้อมูลจาก Pantip และแบบสอบถามแล้ว ดำเนินการสร้าง Co-occurrence Matrix จากคำสำคัญที่เลือกมาด้วยโปรแกรมภาษา Python เพื่อนำไปสร้างตัวแบบ MDS และกราฟโครงข่าย โดยใช้ Similarity Measure แบบ Jaccard Index เหตุผลที่ใช้ Jaccard นั้นเพราะต้องการให้ตัวแบบยืดหยุ่นต่อการสำรวจองค์ประกอบใหม่ในตลาด ลดโอกาสการเกิดค่าอนันต์หากคำสำคัญคำใดคำหนึ่งไม่ปรากฏจริงในเอกสาร ทั้งนี้แต่ละ Matrix มีกระบวนการในการคำนวณที่แตกต่างกัน แบ่งเป็น 2 กลุ่ม ประกอบด้วย Matrix ขนาดใหญ่ที่นำไปใช้สร้างกราฟโครงข่าย และ Matrix ขนาดเล็กที่นำไปใช้ทำ Robustness check, MDS และ Cross Validation ดังนี้

กลุ่มที่ 1 คำนวณค่า Jaccard Index เพื่อสร้าง Co-occurrence Matrix จากการเทียบคู่การปรากฏคำสำคัญพร้อมกันในเอกสาร ขนาด 500x500 ทำการ Normalized ด้วย Jaccard Index

กลุ่มที่ 2 สร้าง Co-occurrence Matrix ขนาด 7x7 ตราสินค้า ด้วย Jaccard Index ด้วยวิธีเทียบการปรากฏร่วมกัน 3 ส่วน ระหว่างตราสินค้า A คำสำคัญ X และตราสินค้า B เทียบคู่ตราสินค้ากับคำสำคัญจนครบทั้ง 500 คำ และครบทุกคู่ตราสินค้า จากนั้นหาค่าเฉลี่ยของ Jaccard Index จากคู่ตราสินค้าเทียบกับคำสำคัญ เพื่อให้เหลือค่า Jaccard

Index เฉพาะระหว่างคู่ตราสินค้า 7 ตราเท่านั้น ดัดแปลงสอดคล้องกับวิธีของ Lee (2007) ที่ได้ดำเนินการในลักษณะคล้ายกันในการหา Jaccard Index เฉลี่ย ในตาราง Collaborative Filtering ของการให้คะแนน Rating ภาพยนตร์ใน MovieLens และ Netflix ระหว่างผู้ใช้มากกว่า 1 คนที่สนใจภาพยนตร์เรื่องเดียวกัน

3.5 ขั้นตอนการทดสอบความแข็งแกร่งทางสถิติ (Robustness check) และ Cross Validation

ผู้วิจัยได้เลือกวิธีการทดสอบความแข็งแกร่งทางสถิติ (Robustness check) และ Cross Validation ตามแนวทางของ Netzer et al. (2012) และ Malhotra & Bhattacharyya (2019) แต่ไม่ได้เลือกใช้ทั้งหมดเนื่องจากบางวิธีไม่สามารถปรับใช้กับข้อมูลจาก Pantip.com ได้ เพื่อให้แน่ใจว่ากระบวนการสร้าง Co-occurrence Matrix มีความถูกต้อง โดยสร้าง Matrix เพื่อทำการทดสอบด้วยวิธีการแบบกลุ่มที่ 2 ในข้อที่ 3.4 ด้วยโปรแกรมภาษา Python ดังนี้

1. Robustness check โดยเทียบค่าสหสัมพันธ์ QAP ระหว่าง Co-occurrence Matrix ขนาดมิติ 7x7 ที่สร้างในรูปแบบทางเลือก (Alternative form) จำนวน 5 Matrix ด้วย Similarity Measure ที่แตกต่างกัน 5 แบบ ประกอบด้วย Jaccard Index, Cosine Similarity, Lift, Inclusion Index และ Association Strength

2. Robustness check โดยเทียบค่าสหสัมพันธ์ QAP ระหว่าง Co-occurrence Matrix ขนาดมิติ 7x7 จำนวน 5 Matrix ที่สร้างด้วย Jaccard Index โดย Matrix ทั้งหมดถูกสร้างจากจำนวนเอกสารตั้งต้นที่สุ่มมาจากเอกสารทั้งหมด 12,546 เอกสาร ในสัดส่วน 1/1, 1/2, 1/4, 1/8 และ 1/16 ตามลำดับ การสุ่มแต่ละครั้งเป็นแบบสุ่มแล้วใส่คืน

3. Cross Validation เทียบค่าสหสัมพันธ์ QAP ระหว่างข้อมูลการรับรู้จาก Pantip.com และข้อมูลการรับรู้จากแบบสอบถาม โดยใช้ Co-occurrence Matrix 2 ตัว สร้างด้วย Jaccard Index ขนาดมิติ 7x7 ตราสินค้า โดย Matrix ข้อมูลจาก Pantip.com สร้างจากชุดคำสำคัญที่ถูกจัดเข้ากลุ่มปัจจัย 8 กลุ่มตามวิธีในข้อที่ 3.2 โดยเอาคำในกลุ่มทั้งหมดมารวมกัน แต่ไม่รวมกลุ่มที่ความหมายไม่เข้าพวก เพื่อให้สามารถเปรียบเทียบกันได้กับข้อมูลจากแบบสอบถาม

3.6 ขั้นตอนการสร้าง MDS Perceptual map

สร้าง Co-occurrence Matrix 9 ชุด ด้วยโปรแกรมภาษา Python แบ่งตามมิติปัจจัยที่มีผลต่อการบริโภคนมจากคำสำคัญที่ความหมายสอดคล้องในแต่ละปัจจัย จำนวน 8 ชุด และ Matrix ที่สร้างจากคำสำคัญทุกคำรวมกัน 1 ชุด คำนวณค่า Co-occurrence ด้วยวิธีการแบบกลุ่มที่ 2 ในข้อที่ 3.4 แต่เปลี่ยนจาก Jaccard Index เป็น Jaccard Distance เมื่อได้ Matrix แล้ว นำเข้าข้อมูลสู่ตัวแบบ MDS เพื่อกำหนดพิกัดตราสินค้าบน Perceptual map ในแต่ละมิติปัจจัย โดยมิติของแกนฉายที่ใช้เป็นแบบ 2 มิติ ผู้วิจัยได้เพิ่มฟังก์ชันการทำ K-Mean Clustering เข้าไปด้วย ใช้วิธีหาจำนวนกลุ่มที่เหมาะสมแบบ Silhouette Score เพื่อช่วยจัดกลุ่มตราสินค้าใน Perceptual Map เพิ่มเติม สาเหตุที่ใช้ Silhouette Score เนื่องจากง่ายต่อการสร้างตัวแบบ เพราะว่าค่า Score ที่ต้องการเป็นค่าสูงสุด แตกต่างจากการใช้ Elbow Method

3.7 ขั้นตอนการสร้างโครงข่ายเชิงความหมาย (Semantic Network)

นำ Co-occurrence Similarity Matrix ขนาดมิติ 500x500 ที่ใช้วิธีคำนวณค่า Jaccard Distance ด้วยวิธีการแบบกลุ่มที่ 1 ในข้อที่ 3.4 จำนวน 1 ชุด มาเป็นข้อมูลตั้งต้นในการสร้าง Semantic Network โดยใช้โปรแกรม R Studio ร่วมกับ Package ชื่อ Bibliometrix และ VOS Viewer ที่พัฒนาขึ้นตามแนวทางของ Visualization of Similarity เหตุผลที่ผู้วิจัยเลือก VOS Viewer เพราะว่าวิธีการแสดงผลดังกล่าวสามารถแก้ไขปัญหาความหนาแน่นของ Vertex และเส้นในกราฟโครงข่ายเมื่อต้องการแสดงผลองค์ประกอบและคำสำคัญจำนวนมาก นอกจากนี้ VOS Viewer ใช้วิธีการทำ Community Detection แบบ Modularity Base ในตัว ซึ่งมีประสิทธิภาพสำหรับกราฟที่มีขนาดใหญ่

4. ผลการวิจัย

4.1 ผลการตอบแบบสอบถาม และ Goodness-of-fit ของตัวแบบ MDS

จากแบบสอบถามที่ได้รับกลับมา มีจำนวน 32 ตัวอย่าง จาก 40 ตัวอย่างที่ได้ส่งไป มีรายละเอียดของ เพศผู้ตอบแบบสอบถาม เป็นชาย 6 คน หญิง 26 คน อยู่ในช่วงอายุ 21-30 ปี มากที่สุด 19 คน ช่วงอายุ 31-40 ปี จำนวน 10 คน ช่วงอายุ 41-50 ปี 2 คน และมากกว่า 51 ปี 1 คน จากตัวอย่างทั้งหมด มีการกรอกข้อมูลในแต่ละตราสินค้าจำนวนแตกต่างกันตามการรับรู้ หากตราสินค้าเป็นที่รู้จักดีจะมีผู้กรอกข้อมูลมาก อย่างไรก็ตาม เนื่องจากข้อมูลบางส่วนไม่สมบูรณ์ จึงตัดตัวอย่างออกไป 2 ชุด เหลือเพียง 30 ชุด เพื่อใช้สร้างตัวแบบ MDS หลังจากได้ตัวแบบแล้วคำนวณค่า Goodness-of-fit ของตัวแบบ ด้วย Kruskal's Stress ได้ค่าเท่ากับ 0.14 ซึ่งอยู่ในเกณฑ์พอใช้ การทดสอบส่วนนี้ดำเนินการด้วยโปรแกรมภาษา Python

4.2 ผลการทำ Robustness check และ Cross Validation

Matrix ที่ได้สร้างขึ้นสำหรับการทำ Robustness check ทั้งหมด ถูกนำมาหาค่าสหสัมพันธ์ QAP ด้วยวิธีการที่กำหนดไว้เพื่อพิจารณา Reliability การทดสอบค่าทั้งหมดดำเนินการในโปรแกรม R Studio เริ่มจากตารางที่ 1 แสดงผลค่าสหสัมพันธ์ระหว่าง Similarity Measure แบบ Jaccard Index และอีก 4 ตัวที่แตกต่างกันเพื่อพิจารณารูปแบบทางเลือก (Alternative form) พบว่า Jaccard Index มีความสัมพันธ์กับรูปแบบทางเลือกในระดับสูงเกือบทั้งหมดที่ระดับนัยสำคัญน้อยกว่า 0.01

สำหรับในตารางที่ 2 เป็นการคำนวณค่าสหสัมพันธ์ระหว่าง Jaccard Index สร้างจากเอกสารที่ถูกสุ่มแล้วใส่คืนมาในสัดส่วนแตกต่างกัน 4 แบบ เทียบกับ Matrix ที่ถูกสร้างจากเอกสารทั้งหมด พบว่ามีความสัมพันธ์กันในระดับสูงเกือบทั้งหมดที่ระดับนัยสำคัญน้อยกว่า 0.01

ตารางที่ 1 ผลสหสัมพันธ์ QAP แบบ Alternative form ของ Similarity Measure ที่แตกต่างกัน

	Jaccard	Lift	Inclusion	Cosine	Association
Jaccard Index	1.0000				
Lift	0.5777 *	1.0000			
Inclusion Index	0.7114 *	0.6883 *	1.0000		
Cosine Similarity	0.64 (R) *	0.2989 (R) ***	0.3299 (R) ***	1.0000	
Association Strength	0.9264 *	0.6366 *	0.842 *	0.4598 (R) **	1.0000

ตารางที่ 2 ผลสหสัมพันธ์ QAP ระหว่าง Similarity Matrix ที่ถูกสร้างจากเอกสารในสัดส่วนแตกต่างกัน

	1/16 (สุ่ม)	1/8 (สุ่ม)	1/4 (สุ่ม)	1/2 (สุ่ม)
1/1 (ทุกคำสำคัญ)	0.7531 *	0.5618 *	0.7722 *	0.9642 *

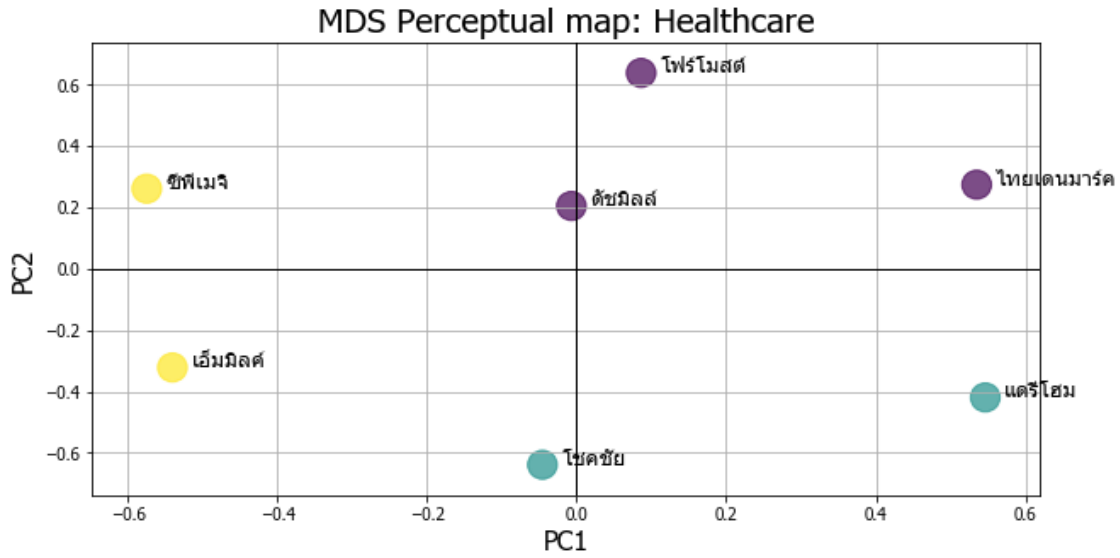
* p-value sig ที่ <= 0.01, ** p-value sig ที่ <= 0.05, *** p-value sig ที่ <= 0.10

ในส่วนการทำ Cross Validation ระหว่าง Matrix จากแบบสอบถามและ Pantip.com ที่ได้คัดเลือกชุดคำสำคัญจัดกลุ่มตามปัจจัย เพื่อให้สามารถเปรียบเทียบกันได้ ได้ค่าสหสัมพันธ์ QAP = 0.5642 ค่า p-value = 0.011 ซึ่งมีความสัมพันธ์ระหว่างกันในระดับกลางค่อนข้างสูงอย่างมีนัยสำคัญ แต่ถ้ารวมกลุ่มที่ความหมายไม่เข้าพวกเข้าไปใน Matrix ข้อที่ 3.5.3 ส่วนที่สร้างด้วยข้อมูลการรับรู้จาก Pantip.com เมื่อเทียบกับแบบสอบถามแล้ว มีผลสหสัมพันธ์ QAP สูงขึ้นอีกอย่างมีนัยสำคัญ เท่ากับ 0.6681 และค่า p-value = 0.005

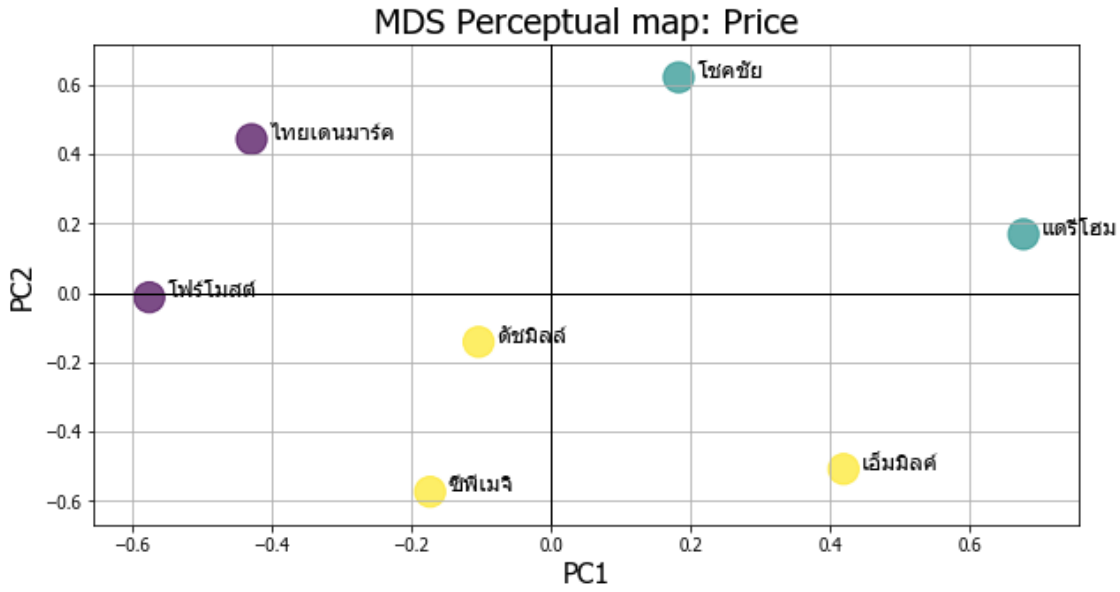
4.3 MDS Perceptual map

การสร้าง Perceptual map ตำแหน่งทางการตลาด 8 มิติจาก Dissimilarity Matrix ของข้อมูลจาก Pantip.com ร่วมกับการใช้ K-Mean Clustering ในการจัดกลุ่มตราสินค้า ผู้วิจัยใช้โปรแกรมภาษา Python สร้างตัวแบบขึ้น ได้

ตัวอย่างผลลัพธ์ดังรูปที่ 2 และรูปที่ 3



รูปที่ 2 MDS Perceptual map จาก Pantip.com ในมุมมองปัจจัยด้านสุขภาพ

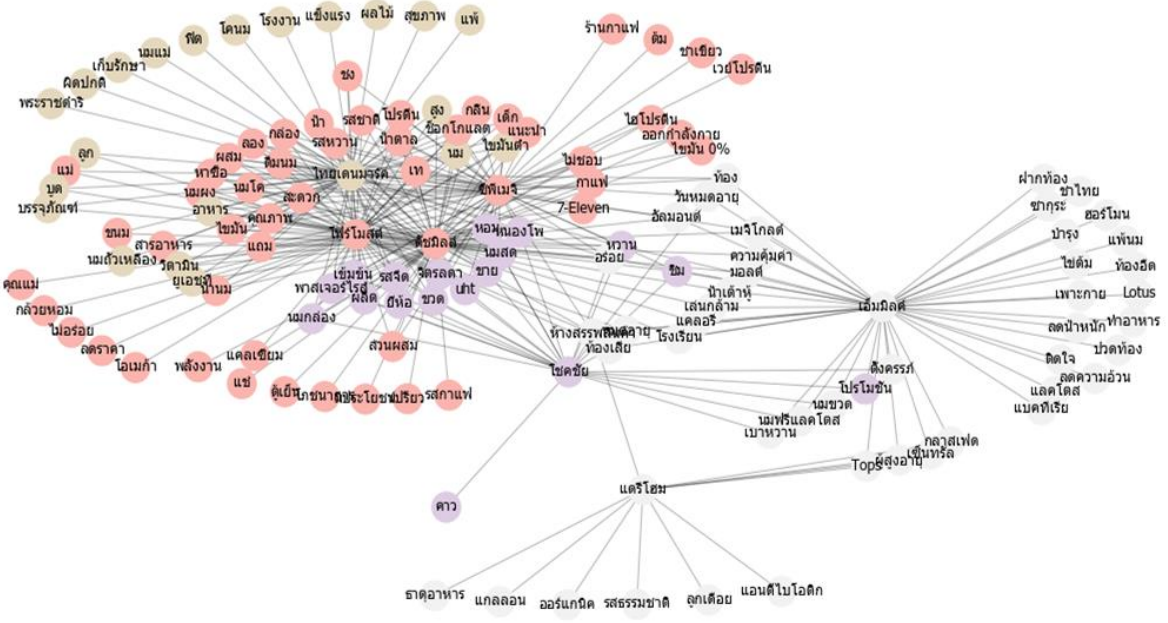


รูปที่ 3 MDS Perceptual map จาก Pantip.com ในมุมมองปัจจัยด้านราคา

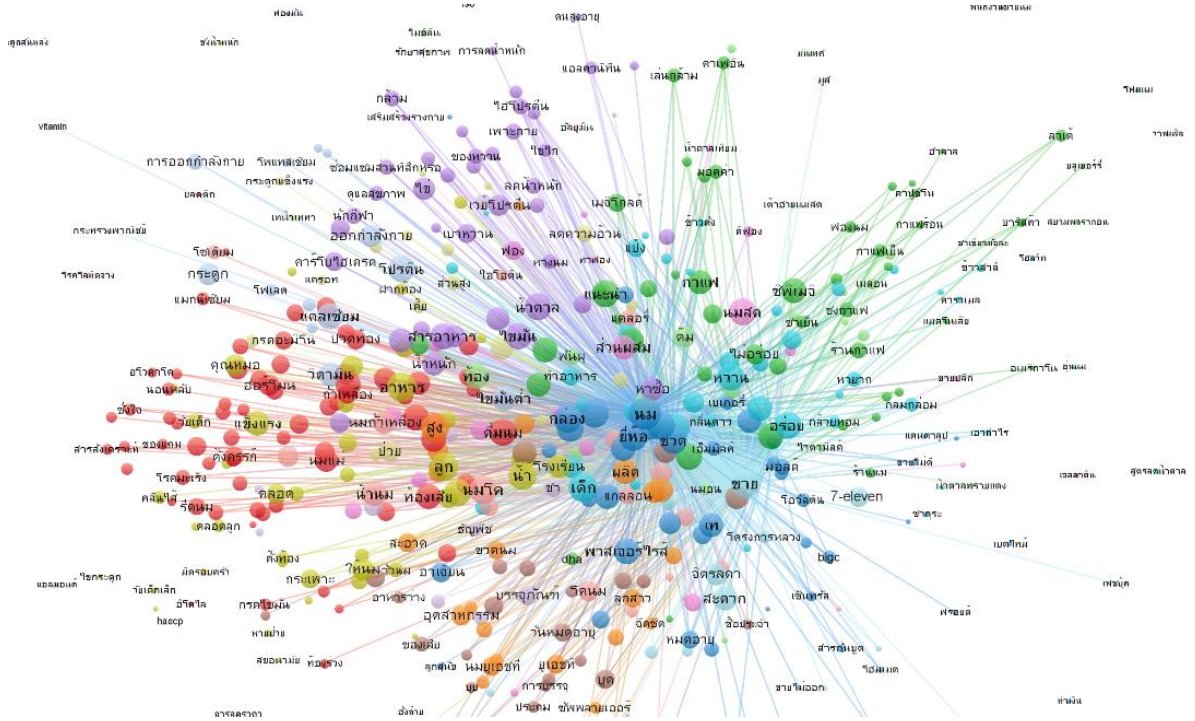
4.4 Semantic Network และ Community Detection

การสร้าง Semantic Network ของข้อมูลจาก Pantip.com ใช้ Similarity Matrix ขนาด 500x500 นำมาสร้างกราฟโครงข่าย 2 แบบ คือ แบบ Zoom-in เจาะจงเฉพาะตราสินค้า และแบบแสดงผลทุกคำสำคัญ ในแบบที่ 1 Zoom-in เจาะจงเฉพาะตราสินค้า มีผลลัพธ์ดังรูปที่ 4 ซึ่งสร้างจาก Library ชื่อว่า NetworkX ในโปรแกรมภาษา Python เพื่อแสดงผลองค์ประกอบร่วมแบบเจาะจงตราสินค้า โดยใช้ Layout แบบ Kamada-Kawai

อย่างไรก็ตาม การแสดงผลแบบ Zoom-in นั้นมีปัญหาที่สำคัญ หากต้องการแสดงผลกราฟโครงข่ายให้ครบทุกเส้นเพื่อให้ได้ความเชื่อมโยงที่สมบูรณ์ จำนวน Vertex และเส้นที่ปรากฏนั้นมีจำนวนมากซ้อนทับกันอย่างหนาแน่นจนไม่สามารถเข้าใจข้อมูลที่ปรากฏได้ ผู้วิจัยจึงได้เลือกใช้ VOS Viewer ซึ่งเป็นเครื่องมือที่สามารถแสดงผลกราฟโครงข่ายขนาดใหญ่ ดังตัวอย่างในรูปที่ 5



รูปที่ 4 Zoom-in Network Graph



รูปที่ 5 VOS Viewer Network Graph

จากรูปที่ 5 เป็นการแสดงผลบน VOS Viewer ที่มีการแบ่งสีตาม Community สามารถเลือกแสดงชนิดของเส้น (เส้นโค้ง-ตรง) หรือกรองจำนวนเส้นที่ปรากฏบน Canvas ได้ หากมองไม่ชัดว่าองค์ประกอบที่มีขนาดเล็กหมายถึงอะไร ผู้ใช้สามารถขยายมุมมองเพื่อดูได้ หรือ นำเมาส์ไปชี้ที่ Vertex ที่ต้องการให้ VOS Viewer แสดงเฉพาะเส้นและ Vertex อื่นที่เชื่อมต่อด้วย นอกเหนือจากความสัมพันธ์ในโครงข่าย ผู้วิจัยได้แจกแจงรายละเอียดองค์ประกอบที่ถูกจัดกลุ่มด้วยวิธี Community Detection เพื่อพิจารณาความสอดคล้องกับปัจจัยที่ได้จากการทบทวนวรรณกรรม และเปิดเผยข้อมูลเชิงลึกที่ซ่อนอยู่ในกลุ่มที่ความหมายไม่เข้าพวก โดยมีรายละเอียดดังตารางที่ 3

ตารางที่ 3 สรุป Community ที่ค้นพบ

ลำดับที่	ชื่อ Community	จำนวนสมาชิก
1	สุขภาพและโรคภัย	70
2	ผสมชาและกาแฟ	53
3	การขายในห้างสรรพสินค้า	52
4	แม่และเด็ก	46
5	เพาะกาย และโภชนาการของการเพาะกาย	45
6	รสชาติและส่วนผสม	43
7	นมกับการเป็นอาหารเสริมสำหรับเด็ก	30
8	คุณภาพการผลิต กับการเก็บรักษา	30
9	ราคาถูก ความสะอาด ออร์แกนิค	21
10	แหล่งโปรตีนอื่นที่นอกจากนม	21
11	ร้านค้าส่ง & ร้านของชำ	20
12	การบำรุงกระดูก	19
13	ลดความอ้วน	18
14	การส่งเสริมการขาย	15
15	ร้านสะดวกซื้อ	15

VOS Viewer ได้จัด Community จำนวน 15 กลุ่ม โดยตัดองค์ประกอบที่มีความสัมพันธ์น้อยผิดปกติ (Outlier) 2 คำสำคัญออก เหลือ 498 ชุดคำ ปัจจัยที่ได้จากการทบทวนวรรณกรรมนั้นปรากฏครบถ้วน แต่มีการแบ่งกลุ่มย่อยเพิ่มเติมอีก ซึ่งเป็นประโยชน์ในทางธุรกิจ ขอยกตัวอย่างคำสำคัญในบางกลุ่มดังต่อไปนี้

1. ตัวอย่างคำใน Community แม่และเด็ก เช่น โรงเรียน, ให้นม, คุณแม่, คุณหมอ, คลอด, ลูกชาย, กระเพาะ, เครียด, น้ำอัดลม, มีลูก, เลี้ยงลูก, ความสูง, วัยเด็ก, วัวนม, โรงพยาบาล, เด็กโต, ทำกิจกรรม, คลื่นไส้, มื้อเช้า, ตั้งท้อง, คลอดลูก, ฝากท้อง, เป็นหวัด, กลางดึก, ระบบขับถ่าย, ไข่ดาว, แพ้ท้อง, ฝากครรภ์, ตัวเตี้ย, มีครอบครัว, หายป่วย, ประกอบอาหาร, วัยเด็กเล็ก, ฟื้นฟูสภาพ, สุขอนามัย, ไขกระดูก, แม่ลูกอ่อน, เด็กอ่อน

2. ตัวอย่างคำใน Community เพาะกาย และโภชนาการ เช่น สารอาหาร, ไขมัน, โภชนาการ, พลังงาน, น้ำหนัก, ไข่, ไขมัน 0%, เวย์โปรตีน, แบคทีเรีย, คาร์โบไฮเดรต, เบาหวาน, กล้ามเนื้อ, เพาะกาย, ความหวาน, นักกีฬา, ลดน้ำหนัก, ไฮโปรตีน, กล้าม, ความคุ้มค่า, ของหวาน, ดูแลสุขภาพ, หางนม, ซ่อมแซมส่วนที่สึกหรอ, กลูโคส, ออกกำลัง, เล่นเวท, ปริมาตร, ไข่ขาว, ฟิตเนส, ไข่ไก่, แอลคานีทีน, อกไก่, เพิ่มน้ำหนัก

5. การอภิปรายผลและสรุป

งานวิจัยนี้ได้นำเสนอวิธีการสำรวจตลาดผ่านสื่อสังคมออนไลน์อย่าง Pantip.com โดยนำเสนอกราฟโครงข่ายและแผนผังการรับรู้ซึ่งแสดงตำแหน่งการตลาดของตราสินค้า สำหรับตำแหน่งทางการตลาดที่สร้างด้วย MDS ในขั้นตอน Cross Validation ระหว่างข้อมูลจากกระทู้ในเว็บไซต์กับข้อมูลจากแบบสอบถาม พบว่าอาจมีมุมมองหรือปัจจัยอื่นเพิ่มเติมที่ไม่ปรากฏในการทบทวนวรรณกรรมอีก ซึ่งอาจเกิดจากกิจกรรมตลาดที่เปลี่ยนแปลงตลอดเวลา หรือเลือกคำสำคัญไม่ครบถ้วนพอ เพราะว่าเมื่อทดลองเพิ่มชุดคำสำคัญที่ความหมายไม่เข้าพวกเข้าไปรวมกับชุดคำสำคัญ 8 กลุ่มที่มีอยู่เดิมเพื่อเป็นข้อมูลตั้งต้นในการทดสอบ ค่าสหสัมพันธ์นั้นสูงขึ้นอย่างมีนัยสำคัญ นอกจากนี้ยังพบว่าตำแหน่งทางการตลาดตราสินค้าในปัจจัยด้านส่วนผสม ด้านแม่และเด็ก และด้านสุขภาพ ตราสินค้าทั้งหมดถูกแบ่งกลุ่มได้เป็นมิติละ 3 กลุ่ม มีตราสินค้าในแต่ละกลุ่มเหมือนกันทั้งสามมิติ สามารถตีความได้ว่า สมาชิกในกลุ่มกำลังแข่งขันกัน

ภายในมิติมุมมองดังกล่าว ซึ่งสอดคล้องกับข้อมูลเชิงลึกของผู้เชี่ยวชาญจากบริษัทนม บริษัทนมต้องกำหนดกลยุทธ์อย่างระมัดระวัง เช่น การอบรมการแนะนำสินค้าให้กับพนักงานแนะนำสินค้าในมิติที่กำลังแข่งขันกันอยู่อย่างเข้มข้น การป้องกันการซื้อตัวพนักงานแนะนำสินค้าโดยบริษัทที่กำลังแข่งขันอยู่ หรือการแย่งชิงพื้นที่ขายเพื่อเข้าถึงกลุ่มเป้าหมายตามมิติตำแหน่งตราสินค้าที่ได้สรุปไว้

สำหรับ Semantic Network และกระบวนการ Community Detection หลังจากอภิปรายผลร่วมกับผู้เชี่ยวชาญ พบว่าข้อมูลเชิงลึกใหม่ที่ปรากฏ เผยให้เห็นเหตุการณ์ในตลาด กับองค์ประกอบคำสำคัญหลายคำที่บางประเด็นเป็นสิ่งที่ไม่คาดคิดว่าจะพบเห็น ทั้งหมดดำเนินการโดยไม่ต้องใช้ข้อมูลจากการทำแบบสอบถาม

เมื่อเปรียบเทียบงานวิจัยในประเทศไทยก่อนหน้านี้ งานวิจัยของ ชื่นสุมล และ บุษราภรณ์ (2557) ดำเนินการสร้างตำแหน่งทางการตลาดตราสินค้าโดยใช้ MDS ด้วยแบบสอบถาม พบว่าเมื่อใช้ข้อมูลจากสื่อสังคมออนไลน์ได้ข้อมูลที่มีจำนวนมิติมุมมองมากกว่าการทำแบบสอบถาม เพราะการใช้แบบสอบถามทำให้แนวทางการกำหนดองค์ประกอบในตลาดมีโอกาสกำหนดได้ไม่ครบถ้วน เนื่องจากแต่ละบริษัทมีการคิดค้นสินค้าใหม่หรือกิจกรรมส่งเสริมการขายใหม่อยู่เสมอ เมื่อเทียบงานของ Netzer et al. (2012) ที่ได้ดำเนินการวิจัยในลักษณะคล้ายกันกับงานนี้ โดยมีกระบวนการทำ Cross Validation จากแบบสอบถามเหมือนกัน มีค่าสหสัมพันธ์ QAP เท่ากับ 0.43 พบว่าผลลัพธ์ของงานวิจัยอยู่ในเกณฑ์ดีกว่าของ Netzer et al. ไม่มากนัก ซึ่งอาจเป็นผลมาจากการเลือกชุดคำเพื่อสร้าง Matrix เปรียบเทียบนั้นไม่ครบถ้วน แต่ว่าการศึกษานี้สามารถแก้ปัญหาการแสดงผลกราฟที่หนาแน่น ยากที่จะตีความได้ และปรับปรุงประสิทธิภาพของกระบวนการ Community Detection แบบ Girvan-Newman ในการศึกษาของ Netzer et al. ที่กินเวลาและทรัพยากรประมวลผลมากกว่าเมื่อกราฟมีขนาดใหญ่และหนาแน่นด้วยการใช้ VOS Viewer ร่วมกับกระบวนการ Community Detection แบบ Modularity Base

6. ข้อจำกัดและข้อเสนอแนะ

อย่างไรก็ตาม งานวิจัยยังมีข้อจำกัด การใช้ข้อมูลจาก Pantip.com มีความเห็นในเชิงเปรียบเทียบค่อนข้างน้อย ผู้ใช้ระบบมักไม่ค่อยใส่ชื่อตราสินค้าลงในช่องความเห็น ส่วนมากจะใส่ในหัวกระทู้ ทำให้แยกชื่อตราสินค้าได้ยาก นอกจากนี้ การแยกประเภทของ Domain นมพาสเจอร์ไรส์และตราสินค้า ทำเพียงกรองข้อความด้วยเงื่อนไข If-Else แบบตรงไปตรงมา ไม่ได้ใช้เทคนิคการเรียนรู้ของเครื่องจักร (Machine Learning) เช่นเดียวกับการคัดเลือกคำเข้าแต่ละกลุ่มปัจจัยที่ทำด้วยมือ ไม่ได้ใช้ตัวแบบหัวข้อ (Topic Modeling) รวมถึงการสุ่มตัวอย่างด้วยวิธีสุ่มแบบตามสะดวกร่วมกับการสุ่มแบบเจาะจง เนื่องจากข้อจำกัดของการเข้าถึงกลุ่มตัวอย่างเพื่อสัมภาษณ์ในพื้นที่โรงพยาบาลช่วงโรค Covid-19 ระบาด ทำให้ตัวอย่างยังมีความหลากหลายไม่เพียงพอ มีอคติ (Bias) ของการสุ่ม และการ Normalized ความสัมพันธ์ระหว่างองค์ประกอบคำสำคัญ ไม่ได้ควบคุมอคติอื่นนอกจากขนาดตลาด

สำหรับการวิจัยในอนาคต ในการหาตำแหน่งตราสินค้า สามารถพิจารณาเพิ่มช่องทางสื่อสังคมออนไลน์อื่นเพิ่มตราสินค้าให้ครอบคลุมทั้งตลาด เพิ่มมิติมุมมอง และเพิ่มประเภทสินค้านมให้ครอบคลุมถึงนมเปรี้ยว โยเกิร์ต วิปครีม นมถั่วเหลือง และนมแพะ หรืออาจพิจารณาสร้างตำแหน่งลงลึกถึงหน่วยผลิตภัณฑ์ในตลาดซึ่งต้องใช้เทคนิคการเรียนรู้ของเครื่องจักรช่วยในการจำแนกข้อความ เพื่อให้ข้อมูลมีความถูกต้องต่อการหาตำแหน่งตราสินค้าต่อไป

7. กิตติกรรมประกาศ

ขอขอบพระคุณ คุณนิธิมา พลชัย ผู้จัดการทั่วไปบริษัทซีพี-เมจิ จำกัด ที่ได้เอื้อเฟื้อข้อมูลและแนวคิดที่เกี่ยวข้อง อีกทั้งยังสนับสนุนกิจกรรมต่างๆ เป็นอย่างดีมาตลอดจนงานวิจัยชิ้นนี้สำเร็จลุล่วงไปด้วยดี

8. เอกสารอ้างอิง

ชูชัย สมิทธิไกร. (2562). **พฤติกรรมผู้บริโภค.** กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ชื่นสุมล บุนนาค, บุษราภรณ์ อารีย์. (2557). การวิเคราะห์ตำแหน่งตราสินค้าของผลิตภัณฑ์กระดาษเช็ดหน้าโดยใช้เทคนิคการแบ่งสเกลแบบหลายมิติ. **Journal of Business, Economics and Communications,** 9(2), 27-38.

ปิยฉัตร ช่างเหล็ก, ยอดมนี เทพานนท์, ณัฐพล พันธุ์ภักดี. (2561). ปัจจัยที่มีผลกับทัศนคติต่อการซื้อนมพาสเจอไรส์ ของผู้บริโภคในเขตกรุงเทพมหานคร. **การประชุมวิชาการเสนอผลงานวิจัย ระดับชาติและนานาชาติ,** 9(1), 935-946.

รังสรรค์ เลิศในสัตย์. (2549). **การตลาดเชิงกลยุทธ์เพื่อความสำเร็จสำหรับผู้บริหาร SME.** กรุงเทพฯ: สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น).

สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์. (2563). **รายงานผลสำรวจพฤติกรรมผู้ใช้อินเทอร์เน็ตในประเทศไทย ปี 2562.** กรุงเทพฯ: ผู้แต่ง.

สุภชาติ ชัยณรงค์สิงห์. (2539). **การศึกษาปัจจัยที่ส่งผลการบริโภคนมเปรี้ยวพร้อมดื่ม.** วิทยานิพนธ์ปริญญาบริหารธุรกิจมหาบัณฑิต สาขาวิชาบริหารธุรกิจ มหาวิทยาลัยเกษมบัณฑิต.

Aaker, D. A. (1996). **Building Strong Brands.** Free Press, New York.

Blondel, V., Guillaume, J. L., Lambiotte, R., Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. **Journal of Statistical Mechanics: Theory and Experiment,** 2008(10). doi:10.1088/1742-5468/2008/10/P10008

De Alwis, A., Edirisinghe, J., Athauda, A. (2011). Analysis of Factors Affecting Fresh Milk Consumption among the Mid-Country Consumers. **Tropical Agricultural Research and Extension,** 12(2), 103-109. doi:10.4038/tare.v12i2.2799

Doerfel, M. (1998). What Constitutes Semantic Network Analysis? A Comparison of Research and Methodologies. **Connections,** 21(2). 16-26.

Henderson, G. R., Iacobucci, D., Calder, B. J. (1998). Brand diagnostics: Mapping branding effects using consumer associative networks. **European Journal of Operational Research,** 111(2), 306-327.

Qin, H. (1999). Knowledge Discovery through Co-Word Analysis. **Library Trends,** 48(1). 133-59.

Howard, L. W. (1998). Validating the Competing Values Model as a Representation of Organization Cultures. **The International Journal of Organization Analysis,** 6(3), 231-250.

Hsu, J. L., & Lin, Y. T. (2006). Consumption and attribute perception of fluid milk in Taiwan. **Nutrition & Food Science,** 36(3), 177-182.

Kurajdova, K., Petrovicova, J. T., Kascakova, A. (2015). Factors Influencing Milk Consumption and Purchase Behavior - Evidence from Slovakia. **Procedia Economics and Finance,** 34, 573-580. doi:10.1016/S2212-5671(15)01670-6

Lee, S. J. (2017). Improving Jaccard Index for Measuring Similarity in Collaborative Filtering. **International Conference on Information Science and Applications,** 2017, 799-806.

Lee, T., Bradlow, E. (2011). Automated Marketing Research Using Online Customer Reviews. **Journal of Marketing Research,** 48(5), 881-894.

Malhotra, N. K. (2019). **Marketing research: An applied orientation** (7th ed.). New York, NY: Pearson.

Malhotra, P., Bhattacharyya, S. (2019). Online Brand Networks: A New Approach to Brand Positioning. **SSRN Electronic Journal,** doi:10.2139/ssrn.3484520

Netzer, O., Feldman, R., Goldenberg, J., Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance Through Text Mining. **Marketing Science,** 31(3), 521-543.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. **Physical Review E,** 69(6), 066133.

Peter, J. P., Olson, J. C., (2010). **Consumer Behavior and Marketing Strategy** (9th ed.). McGraw-Hill.

Phatthiyaphaibun, W., Chaovavanich, K., Polpanumas, C., Suriyawongkul, A., Lowphansirikul, L. & Chormai, P. (2016). PyThaiNLP: Thai Natural Language Processing in Python. **Zenodo,** doi:10.5281/zenodo. 3519354

Ringel, D., Skiera, B. (2016). Visualizing Asymmetric Competition among More Than 1,000 Products Using Big Search Data. **Marketing Science,** 35. doi:10.1287/mksc.2015.0950

Rodgers, J. L. (1991). Matrix and Stimulus Sample Sizes in the Weighted MDS Model: Empirical Metric Recovery Functions. **Applied Psychological Measurement,** 15(1), 71-77.

Sievert, C., Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. Workshop on Interactive Language Learning. **Visualization, and Interfaces,** 63-70. doi:10.13140/2.1.1394.3043.

Simpson, W. (2001). The Quadratic Assignment Procedure (QAP). **North American Stata Users Group Meetings 2001,** 1.2.

Tabassum, S., Pereira S. F. F., Fernandes S., Gama J. (2018). Social network analysis: An overview. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,** 8(5). doi:10.1002/widm.1256.

Van Eck, N. J., Waltman, L. (2006). VOS: a new method for visualizing similarities between objects. **Advances in Data Analysis: Proceedings of the 30th Annual Conference of the German Classification Society,** 299-306. doi:10.1007/978-3-540-70981-7_34

Van Eck, N. J., Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics,** 84, 523-538. doi:10.1007/s11192-009-0146-3

Vemprala, N., Xiong, R. R., Liu, C. Z., Choo, K. K. R. (2019). Where Does My Product Stand? A Social Network Perspective on Online Product Reviews. **52nd Hawaii International Conference on System Sciences,** 2345-2354.

Waltman, L., Van Eck, N. J., Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. **Journal of Informetrics,** 4(4), 629-635.

Yin, S., Chen, M., Chen, Y., Xu, Y., Zou, Z., & Wang, Y. (2016). Consumer trust in organic milk of different brands: the role of Chinese organic label. **British Food Journal,** 118(7), 1769-1782.