บทความวิจัย

การศึกษาอัลกอริทึมสำหรับระบบแนะนำรายการอาหารและเครื่องดื่มบนชุดข้อมูลภาษาไทย

อนันต์ยศ แก้วละมุล สถิตย์โชค โพธิ์สอาด* ธรรมศักดิ์ เธียรนิเวศน์ และ ศุภกฤษฎ์ นิวัฒนากูล
สำนักวิชาศาสตร์และศิลป์ดิจิทัล มหาวิทยาลัยเทคโนโลยีสุรนารี

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 0 4422 3789 อีเมล: s@sut.ac.th
DOI: 10.14416/j.kmutnb.2024.08.008
รับเมื่อ 3 ตุลาคม 2565 แก้ไขเมื่อ 3 กุมภาพันธ์ 2566 ตอบรับเมื่อ 23 กุมภาพันธ์ 2566 เผยแพร่ออนไลน์ 19 สิงหาคม 2567

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาอัลกอริทึมที่เหมาะสมสำหรับการสร้างระบบแนะนำรายการอาหารตามรสนิยมความชื่นชอบของผู้บริโภคจากชุดข้อมูลภาษาไทยเนื่องจากการได้รับอาหารที่เหมาะสมจะส่งผลให้ผู้บริโภคมีความสุข เกิดความรู้สึกเชิงบวกและมีสุขภาพที่ดีนอกจากนั้นในมิติของธุรกิจร้านอาหารระบบแนะนำยังสามารถช่วยเสริมสร้างความพึงพอใจ ความจงรักภักดีของลูกค้าและลดระยะเวลาในการตัดสินใจสั่งซื้อ ซึ่งในงานนี้ผู้วิจัยทำการศึกษาเชิงสำรวจจากอัลกอริทึมการแนะนำอาหารที่มีความโดดเด่น 2 ประเภท ได้แก่ เทคนิคการกรองแบบอิงเนื้อหาและเทคนิคการกรองแบบมีส่วนร่วมซึ่งผลของการศึกษาจะแสดงให้เห็นลักษณะเฉพาะของแต่ละวิธีการ สำหรับเทคนิคการกรองแบบอิงเนื้อหาผู้วิจัยได้ประยุกต์เทคนิค Bag-of-Words (BoW) และ TF-IDF ร่วมกับการหาความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) และมากไปกว่านั้นงานวิจัยนี้พบว่า BoW จะสามารถแนะนำรายการอาหารโดยทั่วไปได้มากกว่า TF-IDF เล็กน้อย สำหรับเทคนิคการกรองแบบมีส่วนร่วมผู้วิจัยได้ดำเนินการผ่านอัลกอริทึม KNN Basic, SVD และ SVD++ เพื่อสร้างแบบจำลองการแนะนำและประเมินประสิทธิของแบบจำลองด้วยวิธีการของ MAE และ $RMSE$ ซึ่งพบว่า SVD++ ให้ผลลัพธ์ที่ดีกว่าด้วยค่าผิดพลาด MAE และ $RMSE$ น้อยที่สุด คือ 0.885 1.105 ตามลำดับ

คำสำคัญ: ระบบแนะนำอาหาร เทคนิคการกรองแบบอิงเนื้อหา เทคนิคการกรองแบบมีส่วนร่วม การประมวลผลภาษาธรรมชาติของไทย

การอ้างอิงบทความ: อนันต์ยศ แก้วละมุล, สถิตย์โชค โพธิ์สอาด, ธรรมศักดิ์ เธียรนิเวศน์ และ ศุภกฤษฎ์ นิวัฒนากูล, "การศึกษาอัลกอริทึมสำหรับระบบแนะนำรายการอาหารและเครื่องดื่มบนชุดข้อมูลภาษาไทย," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 34, ฉบับที่ 4, หน้า 1–12, เลขที่บทความ 244-206388, ต.ค.–ธ.ค. 2567.

วารสารวิชาการพระจอมเกล้าพระนครเหนือ ปีที่ 34 ฉบับที่ 4 ต.ค.–ธ.ค. 2567
The Journal of KMUTNB., Vol. 34, No. 4, Oct.–Dec. 2024

Research Article

The Study of Algorithms for Food and Beverage Recommendation System on Thai Language Dataset

Ananyot Keawlamoon, Satidchoke Phosaard*, Thammasak Thianniwet and Suphakit Niwattanakul
School of Digital Technology, Institute of Digital Arts and Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand

* Corresponding Author, Tel. 0 4422 3789, E–mail: s@sut.ac.th DOI: 10.14416/j.kmutnb.2024.08.008
Received 3 October 2022; Revised 3 February 2023; Accepted 23 February 2023; Published online: 19 August 2024

Abstract

This research aims to investigate suitable algorithms for food recommendation systems on a Thai language dataset based on user preferences. Having preferred and appropriate meals can make consumers feel positive, happy, and healthy. For restaurants, this system can help improve customers' satisfaction and loyalty, and reduce the decision time. In this work, an exploratory study was conducted on two prominent food recommendation algorithms: the content-based filtering and collaborative filtering methods. The results of the study revealed the characteristics of each method. For the content-based filtering method, Bag-of-Words (BoW) and TF-IDF technique were applied together with cosine similarity metrics. The BoW technique recommended food with more general ingredients. For the collaborative filtering method, KNN Basic, SVD, and SVD++ were employed to construct the recommendation models. The algorithms were evaluated using MAE and $RMSE$. The SVD++ provided a better result with the lowest MAE (0.885) and $RMSE$ (1.105) values.

Keywords: Food Recommendation, Collaborative Filtering, Content-Based Filtering, Thai Natural Language Processing

Please cite this article as: A. Keawlamoon, S. Phosaard, T. Thianniwet, and S. Niwattanakul , "The study of algorithms for food and beverage recommendation system on Thai language dataset ," *The Journal of KMUTNB*, vol. 34, no. 4, pp. 1–12, ID. 244-206388, Oct.–Dec. 2024 (in Thai).

1. บทนำ

การบริโภคอาหารเครื่องดื่มที่ตรงตามรสนิยมความชื่นชอบหรือเงื่อนไขจะช่วยทำให้ผู้บริโภคมีความสุข เกิดความรู้สึกเชิงบวกต่อร่างกายตนเองและอาจนำไปสู่ การปรับเปลี่ยนพฤติกรรมการบริโภคโดยเฉพาะอย่างยิ่งโรคที่เกิดจากลักษณะการบริโภคอาหารที่ไม่เหมาะสม เช่น กรณีของโรคเบาหวานซึ่งหากผู้ป่วยบริโภคอาหารที่ชื่นชอบและมีโภชนาการสอดคล้องต่อความต้องการของร่างกายช่วยให้สามารถควบคุมระดับน้ำตาลในเลือดให้เป็นปกติได้โดยไม่รู้สึกเบื่ออาหารและทำให้ผู้ป่วยใส่ใจการบริโภคอาหารเพื่อสุขภาพมากขึ้น สำหรับสถานประกอบการธุรกิจอาหารหากนำเสนอเมนูให้ตรงตามรสนิยมหรือเงื่อนไขของลูกค้าจะทำให้เกิดผลกระทบเชิงบวกต่อความพึงพอใจ ความจงรักภักดีของลูกค้า [1] และลดระยะเวลาในการตัดสินใจสั่งซื้อแต่การทราบว่าเมนูใดที่สอดคล้องต่อความต้องการของผู้บริโภคต้องใช้เวลานานในการพิจารณารายการอาหารซึ่งมีความหลากหลาย ณ ปัจจุบันระบบแนะนำเข้ามามีบทบาทสำคัญต่อปัญหาดังกล่าวผ่านการกรองข้อมูลขนาดใหญ่ให้มีขนาดที่เล็กลงเหลือเพียงแค่ข้อมูลที่คิดว่าผู้ใช้สนใจและตรงตามปัจเจกของผู้ใช้แต่ละคน สำหรับแนวทางหลักในการพัฒนาระบบแนะนำสามารถแบ่งออกเป็น 3 ประเภท ได้แก่ ประการแรก คือ เทคนิคการกรองแบบอิงเนื้อหา (Content-based Filtering) หรือ CBF จะแนะนำรายการ (Items) ด้วยการหาความคล้ายคลึง (Similarity) จากคุณลักษณะ (Attributes) [2] หรือคำอธิบายของรายการในระบบกับรายการก่อนหน้าที่ผู้ใช้เคยปฏิสัมพันธ์ด้วยโดยไม่มีความเกี่ยวข้องกับข้อมูลรายการของผู้ใช้คนอื่น ๆ ทำให้ง่ายต่อการดำเนินการแต่ส่งผลให้การแนะนำถูกจำกัดแค่ช่วงข้อมูลของผู้ใช้คนเดียวเท่านั้น [3] ไม่เปิดโอกาสให้ผู้ใช้เข้าถึงรายการหรือเนื้อหาที่หลากหลาย [2] และอีกปัจจัยที่ส่งผลกระทบต่อประสิทธิภาพของ CBF คือ การเตรียมข้อมูล (Data Preprocessing) สำหรับสร้างระบบแนะนำซึ่งโดยปกติแล้วคุณลักษณะหรือคำอธิบายจะเป็นข้อความและจากการสืบค้นพบว่า การเตรียมข้อมูลมีขั้นตอนที่แตกต่างกันตามภาษาของชุดข้อมูลที่ใช้เนื่องจากภาษาจะมีลักษณะเฉพาะทั้งโครงสร้างของคำและไวยากรณ์ [4] ซึ่งสำหรับการแนะนำรายการอาหารงานวิจัยจะนิยมใช้ชื่อรายการอาหารเป็นภาษาอังกฤษ [5], [6] และสำหรับขั้นตอนการสกัดข้อมูลเพื่อนำไปหาค่าความคล้ายคลึงระหว่างรายการอาหารจะประกอบด้วย 1) การสกัดคำสำคัญหรือวลีด้วยมือ [7] ซึ่งจะทำให้ได้ข้อมูลที่ถูกต้องแต่ไม่เหมาะสมกับข้อมูลที่มีขนาดใหญ่ 2) สกัดด้วยวิธี BoW [8] เป็นการหาความถี่ของคำที่ปรากฏในเอกสารและ 3) สกัดด้วย TF-IDF [5], [9] เป็นการประเมินค่าทางสถิติเพื่อสะท้อนความสำคัญของคำที่เกิดขึ้นภายในเอกสารที่อยู่ในคลังเอกสาร หลังจากนั้นทำการหาค่าความคล้ายคลึงระหว่างรายการอาหารด้วยวิธีของความคล้ายคลึงแบบโคไซน์ [5], [8], [9] เพื่อแนะนำเมนูอาหารแก่ผู้ใช้แต่ทว่ายังคงพบกับคำถามว่ารายการอาหารที่ระบบแนะนำนั้นเป็นรายการอาหารทั่วไปที่สามารถหาร้านรับประทานได้ง่ายหรือไม่ ซึ่งงานวิจัยนี้จะวิเคราะห์รายการอาหารทั่วไปสำหรับระบบแนะนำรายการอาหารแบบอิงเนื้อหาด้วย

ประการที่สอง คือ เทคนิคการกรองแบบมีส่วนร่วม (Collaborative Filtering) หรือ CF จะแนะนำรายการโดยหาความคล้ายคลึงระหว่างผู้ใช้หรือรายการ [10] ในระบบจากค่าคะแนน (Rating) ซึ่งบอกระดับความชอบ [11]ทำให้รายการที่แนะนำแก่ผู้ใช้เป้าหมายมีความแปลกใหม่และรายการไม่ซ้ำเดิมซึ่ง CF ได้มีการประยุกต์ใช้งาน กับขอบเขตข้อมูลหลากหลายและยังมีอัลกอริทึมสำหรับ การพัฒนาหลายวิธีซึ่งหนึ่งในนั้น คือ Singular Value Decomposition (SVD) โดย SVD สามารถทำงานได้ดีกับขอบเขตข้อมูลที่หลายประเภท [12]–[14] จึงเป็นที่นิยมใช้และนอกจากนั้นได้มีการพัฒนาอัลกอริทึมใหม่จากแนวคิดของ SVD เรียกว่า Singular Value Decomposition Plus Plus (SVDpp) และได้รับความนิยมเช่นเดียวกันเนื่องด้วยความสามารถในการค้นหาค่าคะแนนปริยาย (Implicit Rating) เพื่ออธิบายการให้ค่าคะแนนจริงของผู้ใช้ทำให้ผลการทำนายมีประสิทธิภาพมากขึ้นแต่กระนั้นจากการทบทวนงานศึกษาก่อนหน้ายังพบว่าบางขอบเขต SVD จะให้ค่าความผิดพลาดน้อยกว่า SVDpp [15], [16] แต่กับอีกขอบเขต SVDpp จะให้ค่าความพลาดน้อยกว่า SVD [17] ซึ่งเป็นเหตุผลที่ก่อให้เกิดความลังเลใน

การตัดสินใจเลือกอัลกอริทึมที่เหมาะกับงาน นอกจากนั้น K-Nearest Neighbor (KNN) ยังเป็นอีกหนึ่งอัลกอริทึมที่ใช้พัฒนา CF ได้เช่นเดียวกับ CBF เนื่องจากพัฒนาได้ค่อนข้างง่าย รวดเร็วและนิยม [18] อัลกอริทึมสำหรับพัฒนาระบบแนะนำนั้นมีหลากหลายประเภทจึงมีงานวิจัยที่มุ่งเปรียบเทียบประสิทธิภาพของระบบ [5], [15]–[18] และปรับค่าพารามิเตอร์ให้เหมาะสมกับขนาดของชุดข้อมูลที่ใช้พัฒนา [17] และเพื่อให้ทราบว่าแบบจำลองใดที่เหมาะสมกับงานในขอบเขตที่ตนสนใจอย่างไรก็ตาม CF ยังประสบกับปัญหาการทำนายค่าคะแนนหากไม่มีข้อมูลคะแนนของผู้ใช้คนใหม่หรือรายการใหม่ (Cold Start Problem) [10] และปัญหาความบางเบาของข้อมูล (Sparsity Problem) จะเกิดขึ้นเมื่อข้อมูลค่าคะแนนที่บรรจุในเมทริกซ์ไม่เพียงพอทำให้ระบบทำการแนะนำรายการได้ไม่สมเหตุสมผล และแนวทางพัฒนาประการสุดท้าย คือ เทคนิคแนะนำแบบผสมผสาน ซึ่งจะนำเอาแต่ละเทคนิคของระบบแนะนำมาทำงานร่วมกันเพื่อนำข้อดีของแต่ละเทคนิคมาทำให้ผลการแนะนำดีขึ้น

งานนี้จะทำการศึกษาเชิงสำรวจอัลกอริทึมที่เหมาะสมสำหรับการสร้างระบบแนะนำรายการอาหารตามรสนิยมความชื่นชอบของผู้บริโภคจากชุดข้อมูลภาษาไทยจากอัลกอริทึมการแนะนำอาหารที่มีความโดดเด่น 2 ประเภท ได้แก่ เทคนิคการกรองแบบอิงเนื้อหา (CBF) และเทคนิคการกรองแบบมีส่วนร่วม (CF) ซึ่งผลของการศึกษาจะแสดงให้เห็นถึงขั้นตอนและเทคนิคที่เหมาะสมในการสร้างแบบจำลองแนะนำอาหารบนชุดข้อมูลภาษาไทยตลอดจนรูปแบบของผลลัพธ์รายการอาหารที่แนะนำจากเทคนิคและอัลกอริทึมที่ใช้ สำหรับเทคนิคการกรองแบบอิงเนื้อหาผู้วิจัยได้ประยุกต์เทคนิค Bag-Of-Words (BoW) และ TF-IDF ร่วมกับการหาความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) และวิเคราะห์รายการอาหารทั่วไป (Common Menu)

2. วัสดุ อุปกรณ์และวิธีการวิจัย

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ระบบแนะนำ (Recommendation Systems) คือ การกรองสารสนเทศให้เหลือเพียงรายการสารสนเทศที่เหมาะสมกับความชอบของผู้ใช้แต่ละบุคคล [2] ด้วยการวิเคราะห์ข้อมูลในอดีตของผู้ใช้งาน [19] และประเภทหลัก ๆ ของระบบแนะนำสามารถแบ่งออกได้เป็น 3 ประเภท คือ 1) CBF จะทำการแนะนำรายการด้วยการหาความคล้ายคลึงจากคุณลักษณะ [2] หรือคำอธิบายของรายการในระบบกับรายการก่อนหน้าที่ผู้ใช้เคยปฏิสัมพันธ์ด้วยโดยไม่มีความเกี่ยวข้องกับข้อมูลรายการผู้ใช้คนอื่น ๆ ทำให้ง่ายต่อการดำเนินการแต่จะส่งผลให้การแนะนำถูกจำกัดแค่ช่วงข้อมูลรสนิยมของผู้ใช้เท่านั้น [3] ส่งผลให้ได้รายการเดิมซ้ำ ๆ ขาดความหลากหลายของข้อมูล 2) CF จะแนะนำรายการจากการให้ค่าคะแนนความชอบของผู้ใช้คนอื่น ๆ ที่มีความคล้ายคลึงกับผู้ใช้เป้าหมาย (Target User) การหาความคล้ายคลึงกันระหว่างผู้ใช้ [10] หาได้จากลักษณะการให้ค่าคะแนนความชอบ (Rating) ของผู้ใช้อื่นในระบบซึ่งจะบ่งชี้ถึงระดับความชอบของแต่ละรายการ [11] สำหรับเทคนิคนี้แบ่งออกได้ 2 ประเภท คือ 2.1) การกรองแบบมีส่วนร่วมบนพื้นฐานความจำ โดยข้อมูลจะถูกเก็บอยู่ในรูปแบบของเมทริกซ์คะแนน (Rating Matrix) ประกอบด้วย การกรองบนพื้นฐานผู้ใช้ (User-based) และการกรองบนพื้นฐานรายการ (Item-based) [11] และ 2.2) การกรองแบบร่วมมือบนพื้นฐานแบบจำลอง เป็นการสร้างแบบจำลองแนะนำจากข้อมูล [11] ด้วยวิธีการเรียนรู้ของเครื่อง (Machine Learning) เพื่อทำการค้นหาความคล้ายคลึงของผู้ใช้หรือรายการเพื่อใช้ในการทำนาย [20] และ 3) เทคนิคแบบผสมผสานจะนำเทคนิคของระบบแนะนำประเภทต่าง ๆ มาทำงานร่วมกันเป็นการประยุกต์ข้อดีของแต่ละเทคนิคเพื่อให้ได้ผลลัพธ์ของการแนะนำที่ดีกว่าเดิม

2.1.2 การเตรียมข้อมูล (Data Preprocessing)

การแนะนำรายการอาหารและเครื่องดื่มจากข้อมูลที่อยู่ในรูปแบบข้อความภาษาไทยจะเตรียมข้อมูลด้วยกระบวนการของการประมวลผลทางภาษาธรรมชาติ (NLP) แต่ขั้นตอนการเตรียมข้อมูลของแต่ละภาษาจะมีความแตกต่างกันเนื่องด้วยโครงสร้างคำและไวยากรณ์ [4] มีเอกลักษณ์ของตนเอง สำหรับการจัดการกับชุดข้อมูลภาษาไทยจากแนวคิดของณัฐกิจ [21] มีอยู่ทั้งหมด

4 ขั้นตอน คือ 1) การปรับข้อความให้เป็นมาตรฐานเป็นการลบสัญลักษณ์พิเศษและตัวเลข 2) การตัดคำหรือวลีจากประโยคและแบ่งคำหรือวลีด้วยเครื่องหมาย | 3) การกำจัดคำหยุดทำการลบคำที่ไม่สื่อความหมายหรือไม่มีนัยสำคัญใด ๆ และ 4) การสกัดคุณลักษณะเป็นขั้นตอนที่มีความสำคัญอย่างยิ่งสำหรับการค้นหาความคล้ายคลึงระหว่างรายการอาหารและเครื่องดื่มซึ่งเป็นการเปลี่ยนรูปแบบข้อความให้อยู่ในรูปแบบของเวกเตอร์ตัวเลขซึ่งดำเนินการผ่าน 2 วิธีด้วยกัน คือ Bag of Word (BoW) และ Term Frequency Inverse Document Frequency (TF-IDF) โดยมีรายละเอียดดังนี้

Bag of Word (BoW) โดยข้อความภายในเอกสารจะถูกทำให้เป็นเวกเตอร์ตัวเลขของจำนวนความถี่การเกิดคำในเอกสารนั้น ๆ

Term frequency inverse document frequency (TF-IDF) คือ จะประเมินคำที่เกิดขึ้นภายในเอกสารที่อยู่ในคลังเอกสารนั้นมีความสำคัญมาน้อยเพียงใด และ TF-IDF เกิดจากการผสมผสานการทำงานของ 2 เทคนิคดังนี้ Term Frequency (TF) และ Inverse Document Frequency (IDF)

Term frequency (TF) คือ จำนวนครั้งของคำ (Word), วลี (Phrase) หรือคำเฉพาะ (Term) ที่ปรากฏในเอกสาร (Document) [13] สำหรับ Scikit-learn Library สามารถประเมิน $tf\,(w,d)$ ได้จากสมการที่ 1 โดย w คือ คำ วลีหรือคำเฉพาะ d คือ เอกสาร และ f คือ ความถี่ของ w ใน d

$$tf(w,d) = f_{w,d} \tag{1}$$

Inverse document frequency (IDF) คือ การลดค่าน้ำหนักความสำคัญของคำหรือวลีที่ปรากฏบ่อยในเอกสาร [14] สามารถคำนวณ $idf(w)$ ได้จากสมการที่ 2 โดย n คือ จำนวนของเอกสารทั้งหมด, w คือ คำ วลีหรือคำเฉพาะ และ $df(w)$ คือ จำนวนเอกสารที่มี w ปรากฏ

$$idf\left(w\right) = \log\left(\frac{1+n}{1+df\left(w\right)}\right) \tag{2}$$

หลังจากคำนวณค่าจากทั้ง 2 สมการ ดังกล่าวข้างต้น นำผลลัพธ์ที่ได้มาคูณดังสมการที่ 3 สำหรับการหาค่า TF-IDF ด้วย Scikit-learn โดยค่าเริ่มต้นหลังจากได้ผลลัพธ์ของ TF-IDF ผลลัพธ์จะถูกนำไปคำนวณด้วย Euclidean norm

$$tf\text{-}idf\left(w,d\right) = tf\left(w,d\right) \times idf\left(w\right) \tag{3}$$

2.1.3 ความคล้ายคลึงแบบโคไซน์ (Cosine Similarity)

ความคล้ายคลึงแบบโคไซน์เป็นการวัดความคล้ายคลึงระหว่าง 2 เวกเตอร์ด้วยวิธีการหาโคไซน์ของมุมซึ่งค่าโคไซน์อยู่ระหว่าง [0,1] และสามารถคำนวณ $sim(x,y)$ ได้จากสมการที่ 4 โดย x คือ เวกเตอร์ x, y คือ เวกเตอร์ y, $||x||$ คือ ความยาวของเวกเตอร์ x และ $||y||$ คือ ความยาวของเวกเตอร์ y

$$sim\left(x,y\right) = \frac{x \cdot y}{\| x \| \| y \|} \tag{4}$$

ตัวอย่างเช่น เอกสาร x = โตเกียวไส้เค็มผสมไส้หวาน เวกเตอร์ $x = [1,2,1,1,1]$ และ เอกสาร y = โตเกียวไส้เค็ม เวกเตอร์ $y = [1,1,1,0,0]$ จากนั้น $x \cdot y = 4$ และหาความยาวของเวกเตอร์ $||x|| = 2.830$ และ $||y|| = 1.730$ ดังนั้น $sim(x,y) = 4 / (2.830 \times 1.730) = 0.820$

2.1.4 การหาเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor; KNN)

KNN คือ อัลกอริทึมสำหรับการจัดหมวดหมู่ข้อมูลจากลักษณะที่คล้ายคลึงหรือมีความใกล้ชิดกันเพื่อทำการแบ่งข้อมูลเป็นกลุ่ม ๆ KNN เป็นอัลกอริทึมประเภทการเรียนรู้แบบมีผู้สอนโดยจะมองหาเพื่อนบ้านที่ใกล้เคียงกับตัวเองมากที่สุด (Nearest Neighbor) จากนั้นจะถูกจัดเข้ากลุ่ม

2.1.5 การแยกส่วนประกอบค่าเอกพจน์ (Singular Value Decomposition; SVD)

SVD เทคนิคสำหรับการแยกตัวประกอบของเมทริกซ์ผ่านการใช้เมทริกซ์สี่เหลี่ยม [22] และเมทริกซ์ดังกล่าวจะถูกแบ่งแยกออกเป็น 3 เมทริกซ์ย่อยที่มีขนาดเล็กกว่าเมทริกซ์ตั้งต้นและ SVD สามารถแก้ไขปัญหาความเบาบาง

ของข้อมูลในเมทริกซ์ได้ [23] ซึ่งสามารถอธิบายจากสมการที่ 5 โดยกำหนดให้ m แทนจำนวนแถวและ n แทนจำนวนคอลัมน์ดังนั้น M คือ เมทริกซ์ตั้งต้น $m \times n$, U คือ เมทริกซ์มุมฉากขนาด $m \times m$ เรียกว่า เวกเตอร์เอกพจน์ซ้าย (Left Singular Vectors), Σ คือ เมทริกซ์มุมฉากเอกพจน์ที่มีองค์ประกอบไม่ติดลบขนาด $m \times n$ เรียกว่า ค่าเอกพจน์ของ A (Singular of A) และ V คือ เมทริกซ์มุมฉากขนาด $n \times n$ เรียกว่า เวกเตอร์เอกพจน์ขวา (Right Singular Vectors)

$$M = U \cdot \Sigma \cdot V^T \quad (5)$$

สำหรับการทำนายค่าคะแนนความชอบของ SVD จากไลบรารี่ของ Surprise Python [24] โดยมีการคำนวณค่าอคติร่วมเนื่องจากการให้คะแนนของผู้ใช้บางคนในระบบจะสูงกว่าผู้ใช้คนอื่น ๆ หรือรายการบางรายการได้คะแนนความชอบมากกว่ารายการอื่น ๆ ในระบบเช่นเดียวกัน [25]

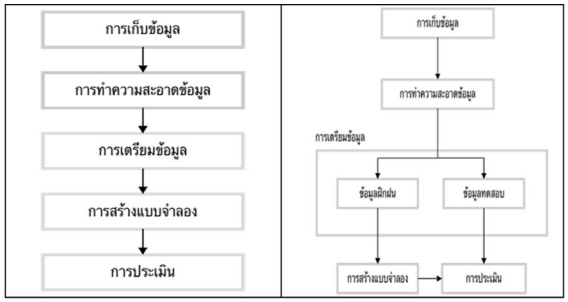
2.1.6 การแยกส่วนประกอบค่าเอกพจน์พลัสพลัส (Singular Value Decomposition Plus Plus; SVD++)

SVD++ คือ เทคนิคที่ถูกพัฒนาต่อยอดจากเทคนิค SVD ซึ่งจะพิจารณาการให้คะแนนปริยายและการให้คะแนนฯ ดังกล่าวจะทำการอธิบายข้อเท็จจริงการให้คะแนนของผู้ใช้ต่อรายการต่าง ๆ ในระบบ

2.3 วิธีดำเนินการวิจัย

2.3.1 ผู้ตอบแบบสอบถาม

งานวิจัยนี้ดำเนินการเก็บข้อมูลรายการอาหารและเครื่องดื่มจากบริษัทธุรกิจอาหารเดลิเวอรี่ซึ่งเป็นแพลตฟอร์มออนไลน์จึงดำเนินการกำหนดกลุ่มผู้ตอบแบบสอบถามสำหรับการเก็บข้อมูลค่าคะแนนความชอบ (Rating) ต่อรายการอาหารและเครื่องดื่มที่มีช่วงอายุระหว่าง 15–24 ปี ซึ่งเป็นช่วงอายุการใช้งานอินเทอร์เน็ตมากที่สุดคิดเป็นร้อยละ 98.4 จากประชากรประเทศไทยจำนวนทั้งหมด 66.17 ล้านคน คิดเป็นร้อยละ 12.73 และเมื่อคำนวณกลุ่มผู้ตอบแบบสอบถามด้วยสมการของทาโร ยามาเนซึ่งกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้ที่ 0.1 ทำให้ได้กลุ่มผู้ตอบแบบสอบถามทั้งหมด 100 คน



(ก) CBF (ข) CF

รูปที่ 1 กระบวนการพัฒนาระบบแนะนำอาหารและเครื่องดื่มโดย (ก) สำหรับ CFB และ (ข) สำหรับ CF

2.3.2 รายการอาหารและเครื่องดื่มยอดนิยม

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลรายการอาหารและเครื่องดื่มยอดนิยมเพื่อใช้เป็นข้อมูลสำหรับการพัฒนาแบบจำลองด้วย CF โดยจะเก็บจากบริษัทธุรกิจอาหารเดลิเวอรี่ในประเทศไทยและจากการรายงานของ Economic Intelligence Center (EIC) พบบริษัทธุรกิจอาหารเดลิเวอรี่ 5 บริษัท ประกอบด้วย Foodpanda, LINEMAN, Grab, Robinhood, และ Airasia โดยมีจำนวนพื้นที่ให้บริการ คือ 77 59 52 6 และ 5 จังหวัด ตามลำดับ งานวิจัยนี้จะเลือกรายการอาหารจากบริษัทที่มีพื้นที่ให้บริการมากที่สุด 3 อันดับแรก โดยเก็บจากการรายงานอาหารยอดนิยมของบริษัทดังกล่าวระหว่าง พ.ศ. 2562–2564 เพื่อนำรายการอาหารและเครื่องดื่มทั้งหมดที่ได้ไปหาค่าคะแนนความชอบจากผู้ตอบแบบสอบถามโดยค่าคะแนนแต่ระดับอยู่ระหว่าง 1 ถึง 5 ระดับ มีความหมายดังนี้ 1 คือ ชอบน้อยที่สุด 2 คือ ชอบน้อย 3 คือ ชอบปานกลาง 4 คือ ชอบมาก และ 5 คือ ชอบมากที่สุด

2.3.3 กระบวนการพัฒนาระบบแนะนำ

งานวิจัยนี้แบ่งกระบวนการพัฒนาระบบแนะนำได้ 2 แบบ ตามเทคนิคที่ใช้ซึ่งอธิบายดังรูปที่ 1

กระบวนการพัฒนาระบบแนะนำด้วยเทคนิคการกรองแบบอิงเนื้อหา (CBF) ตามรูปที่ 1 (ก)

1) การเก็บรวบรวมข้อมูล จะเก็บข้อมูลชื่อและวัตถุดิบ

ของอาหารและเครื่องดื่มจากเว็บไซต์ที่เปิดให้บริการบนอินเทอร์เน็ตด้วยเทคนิค Web Scraping

2) ทำความสะอาดข้อมูล โดยใช้เทคนิคการประมวลผลทางภาษาธรรมชาติ (NLP) เพื่อประมวลผลข้อความเช่น การปรับข้อความให้เป็นมาตรฐาน (Text Normalization) นอกจากมีการลบเครื่องหมายต่าง ๆ แล้วงานวิจัยนี้จะดำเนินการปรับรูปคำหรือวลีที่สื่อความหมายถึงสิ่งเดียวกันให้ใช้คำหรือวลีที่เป็นมาตราฐานเดียวกัน เช่น น้ำเชื่อมและไซรัป โดยให้เลือกใช้น้ำเชื่อมหรือไซรัปอย่างใดอย่างหนึ่งซึ่งทั้งสองคำมีความหมายเดียวกันเนื่องจากวิธีที่ใช้ในการวิเคราะห์ยังไม่ได้วิเคราะห์ถึงรากศัพท์หรือความหมายของคำซึ่งจะวิเคราะห์ตามข้อความที่ปรากฏเท่านั้น ขั้นตอนถัดไปเป็นการลบคำหยุด (Stop Words) จากข้อความ และขั้นสุดท้ายเป็นการตัดคำ (Word Tokenization) ซึ่งพบว่า หากคำหรือวลีในคลังคำที่มีอยู่ไม่ครอบคุมขอบเขตข้อมูลที่ศึกษาจะส่งผลให้การตัดคำนั้นผิดพลาดผู้วิจัยจึงดำเนินการแบ่งคำโดยใช้เครื่องหมาย | คั้นระหว่างคำหรือวลี

3) การเตรียมข้อมูล เป็นการแปลงข้อความ คำหรือวลีให้อยู่ในรูปแบบของเวกเตอร์ตัวเลขโดยใช้ Count Vectorizer สำหรับ BoW และ TF-IDF Vectorizer สำหรับ TF-IDF จากไลบรารี่ Scikit-learn

4) การสร้างแบบจำลอง นำเวกเตอร์ตัวเลขที่ได้จาก BoW และ TF-IDF มาหาค่าความคล้ายคลึงระหว่างรายการอาหารและเครื่องดื่มด้วยวิธีการหาความคล้ายแบบโคไซน์ (Cosine Similarity)

5) การประเมิน การประเมินระบบแนะนำด้วยค่าเฉลี่ยของค่าความคล้ายคลึงของรายการแนะนำสูงสุด N รายการหรือ Top N [5]

กระบวนการพัฒนาระบบแนะนำด้วยเทคนิคการกรองแบบมีส่วนร่วม (CF) ตามรูปที่ 1 (ข)

1) การเก็บรวบรวมข้อมูล จะเก็บรวบรวมข้อมูลค่าคะแนนความชอบต่อรายการอาหารและเครื่องดื่มจากกลุ่มผู้ตอบแบบสอบถามด้วย Google Form

2) การทำความสะอาดข้อมูล เป็นการลบข้อมูลที่ไม่มีประสิทธิภาพ หลังจากนั้นทำการกำหนดรหัสให้กับเมนูและผู้ตอบแบบสอบถาม

3) การเตรียมข้อมูล เป็นการสร้างชุดข้อมูลซึ่งประกอบด้วย ข้อมูลรหัสผู้ใช้ (User ID) ข้อมูลรหัสเมนู (Menu ID) และข้อมูลค่าคะแนนความชอบ (Rating) ในรูปแบบ User-Item Matrix สุดท้ายแบ่งข้อมูลสำหรับการเรียนรู้แบบจำลอง 90% และอีก 10% สำหรับการประเมิน

4) การสร้างแบบจำลอง ดำเนินการสร้างแบบจำลองด้วยอัลกอริทึม KNNBasic, SVD และ SVD++ โดยที่ KNNBasic และ SVD ใช้ค่าพารามิเตอร์ตามค่าตั้งต้น (Default) SVD++ อ้างอิงค่าพารามิเตอร์จากงานของ Shaikh [17] ซึ่งเป็นค่าพารามิเตอร์ที่ให้ค่าความผิดพลาดน้อยที่สุด โดยรายละเอียดของพารามิเตอร์เป็นไปดังตารางที่ 1–3

ตารางที่ 1 การตั้งค่าอัลกอริทึม KNNBasic

พารามิเตอร์	ค่า
k	40
min_k	1

ตารางที่ 2 การตั้งค่าอัลกอริทึม SVD

พารามิเตอร์	ค่า
n_factors	100
n_epochs	20
biased	true
init_mean	0
init_std_dev	1
lr_all	0.005
reg_all	0.02

ตารางที่ 3 การตั้งค่าอัลกอริทึม SVD++

พารามิเตอร์	ค่า
n_factors	40
n_epochs	40
init_mean	0
init_std_dev	0.1
lr_all	0.008
reg_all	0.1

5) การประเมิน จะประเมินค่าผิดพลาดของการทำนายค่าคะแนนความชอบของแบบจำลองด้วยค่า MAE และ $RMSE$ หากแบบจำลองมีค่าทั้งสองต่ำแสดงให้เห็นว่าแบบจำลองนั้น ทำนายค่าคะแนนผิดพลาดน้อยและมีประสิทธิภาพที่ดี [5], [17], [18]

Mean Absolute Error (MAE) การคำนวณขนาดค่าเฉลี่ยของค่าผิดพลาดของค่าทำนายกับข้อมูลทดสอบแต่ละรายการซึ่ง MAE สามารถคำนวณได้จากสมการที่ 6 โดย r_i คือ ค่าคะแนนจริง, p_i คือ ค่าคะแนนที่ได้จากการทำนาย และ n คือ จำนวนของคะแนนทั้งหมด

$$MAE = \frac{\sum_{i=1}^{n} | r_i - p_i |}{n} \quad (6)$$

Root Mean Square Error ($RMSE$) ให้ค่าน้ำหนักเพิ่มเติมจากค่าผิดพลาดจริงทำให้เห็นความแตกต่างของค่าผิดพลาดมากกว่า MAE ซึ่ง $RMSE$ คำนวณได้ดังสมการที่ (7) โดย r_i คือ ค่าคะแนนจริง และ p_i คือ ค่าคะแนนที่ได้จากการทำนาย

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - p_i)^2} \quad (7)$$

3. ผลการทดลอง

3.1 การเก็บรวบรวมข้อมูล

3.1.1 ผลการเก็บรายการอาหารและเครื่องดื่มยอดนิยม

การเก็บรวบรวมข้อมูลจากรายการอาหารยอดนิยมของของทั้ง 3 บริษัทซึ่งได้ดำเนินการลบข้อมูลที่ซ้ำออกและสามารถรวบรวมได้ทั้งหมด 38 รายการ ประกอบด้วย อาหาร 25 รายการและเครื่องดื่ม 13 รายการ

3.1.2 ผู้ตอบแบบสอบถาม

ข้อมูลสำหรับการสร้างแบบจำลองด้วย CF จากการคำนวณกลุ่มผู้ตอบแบบสอบถาม 100 คน มีผู้ตอบแบบสอบถามจริงทั้งสิ้นจำนวน 106 คน กำจัดข้อมูลที่ผิดพลาดของการตอบแบบสอบถามจำนวน 3 คน ทำให้เหลือ 103 คน และสามารถแบ่งข้อมูลผู้ตอบแบบสอบถามตามเพศสภาพได้ดังนี้ เพศหญิง 73 คน เพศชาย 22 คน ไม่ต้องการระบุเพศ 6 และเพศทางเลือก 2 คน คิดเป็นร้อยละ 70.9 21.4 5.8 และ 1.9 ตามลำดับ

3.2 ผลการเตรียมข้อมูล

ผลของการเตรียมข้อมูลสามารถแบ่งได้เป็น 2 ประเภทตามเทคนิคพัฒนาแบบจำลองดังนี้

3.2.1 เทคนิคการกรองแบบอิงเนื้อหา

ข้อมูลรายการอาหารและเครื่องดื่มสำหรับการพัฒนาแบบจำลองนี้จะประกอบด้วย ชื่อเมนูและวัตถุดิบดังตารางที่ 4 ซึ่งมีการแบ่งคำด้วยเครื่องหมาย | มีจำนวนข้อมูลทั้งสิ้น 1,167 ระเบียนที่ได้จากการเก็บผ่านอินเทอร์เน็ต

ตารางที่ 4 ตัวอย่างข้อมูลสำหรับการพัฒนา CBF

ชื่อเมนู	วัตถุดิบ
ชาเขียว|นม|เย็น	ชาเขียว|น้ำร้อน|นมข้นหวาน|น้ำตาลทราย|นมสดจืด|น้ำแข็ง
กาแฟ|ดำ|เย็น	กาแฟ|น้ำเปล่า|น้ำแข็ง|น้ำเชื่อม

3.2.2 เทคนิคการกรองแบบมีส่วนร่วม

ข้อมูลรายการอาหารและเครื่องดื่มสำหรับการพัฒนาแบบจำลองนี้จะประกอบด้วย รหัสผู้ตอบแบบสอบถาม (UserId) รหัสเมนู (Menuld) และค่าคะแนนความชอบ (Rating) ดังตารางที่ 5 และมีจำนวนข้อมูลทั้งหมด 3,811 ระเบียนที่ได้จากผู้ตอบแบบสอบถาม

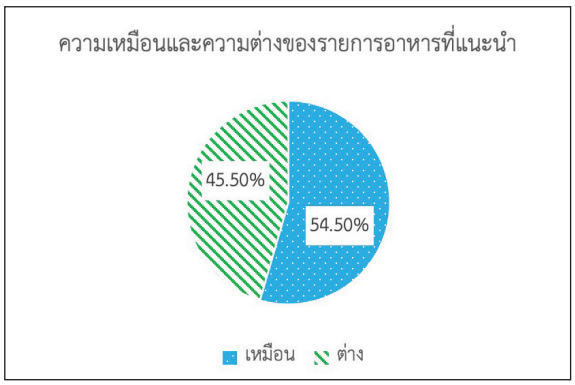
ตารางที่ 5 ตัวอย่างข้อมูลสำหรับการพัฒนา CF

รหัสผู้ใช้	รหัสเมนู	ค่าคะแนนความชอบ
U0001	M0014	5
U0004	M0014	2

3.4 ผลการพัฒนาแบบจำลอง

3.4.1 เทคนิคการกรองแบบอิงเนื้อหา

ทดลองแนะนำอาหารจากแบบจำลองโดยใช้การแนะนำแบบสูงสุด 1 รายการ (Top 1) ของทุกรายการอาหารในชุดข้อมูลเพื่อเปรียบเทียบรายการอาหารที่แนะนำนั้นเหมือนหรือต่างกัน



รูปที่ 5 การเปรียบเทียบความและความต่างของรายการอาหารแนะนำ Top 1 จากทุกรายการอาหารในชุดข้อมูล

ตารางที่ 6 ผลการแนะนำจากรายการกาแฟดำเย็น

TF-IDF	SIM	BoW	SIM
เอสเพรสโซ่เย็น	0.628	เอสเพรสโซ่เย็น	0.630

จากตารางที่ 6 รายการอาหารแนะนำที่มีความคล้ายคลึงกับรายการกาแฟดำเย็นพบว่า จากวิธีการของ BoW และ TF-IDF แนะนำรายการที่เหมือนกันแม้ว่าค่าความคล้ายคลึง (SIM) ของ TF-IDF จะต่ำกว่า BoW แต่อย่างไรก็ตามแบบจำลองจากวิธีการของ BoW และ TF-IDF พบว่า มีการแนะนำรายการที่แตกต่างกัน ผู้วิจัยจึงพิจารณารายการอาหารแนะนำที่แตกต่างกันนั้นซึ่งเมื่อพิจารณาผลการแนะนำ Top 1 ที่แตกต่างกัน โดย “ข้าวมันไก่ต้ม” เป็นรายการสืบค้น พบว่า BoW จะแนะนำรายการ “หมูสับต้มเกี๋ยมบ๊วย” (SIM = 0.442) และ TF-IDF แนะนำ “ข้าวมันไก่หม้อหุงข้าว” (SIM = 0.401) จากรายการอาหารแนะนำที่ได้พบว่า ค่า SIM ของ BoW จะมีค่าที่มากกว่า SIM ของ TF-IDF อย่างไรก็ตามความสอดคล้องของรายการที่แนะนำนั้นพบว่า TF-IDF แนะนำได้สอดคล้องกว่า BoW โดยจากผลการทดสอบค่า SIM อาจจะไม่สามารถบ่งชี้ถึงเรื่องประสิทธิของระบบแนะนำที่พัฒนาด้วยเทคนิค CBF และจากรูปที่ 5 เป็นการเปรียบเทียบรายการอาหารที่แนะนำสูงสุด Top 1 รายการ เพื่อดูว่ารายการอาหารที่แนะนำจากระบบที่พัฒนาด้วย BoW และ TF-IDF นั้นแตกต่างกันอยู่เท่าไรซึ่งพบว่าแนะนำรายการต่างกัน 45.50% และเหมือนกัน 54.50% เพื่อให้ทราบว่ารายการที่ต่างกันนั้นมีลักษณะเป็นอย่างไรผู้วิจัยจึงได้พัฒนาสมการที่ 8 ขึ้นซึ่งใช้แนวคิดจากสมการของ IDF เพื่อตอบคำถามว่า BoW และ TF-IDF สามารถให้ผลลัพธ์ของรายการอาหารและเครื่องดื่มที่ถูกแนะนำแตกต่างกันอย่างไร

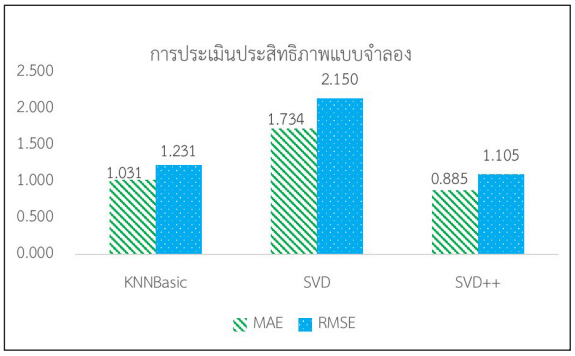
$$\psi = \frac{\sum_{i=1}^{n} df(w_i)}{(n \cdot N)} \tag{8}$$

โดย ψ หรือ Phi คือ ค่าสำหรับบ่งชี้รายการอาหารเพื่อบอกว่าเป็นรายการทั่วไป (Common Menu) อยู่เท่าไรเมื่อเทียบกับรายการอาหารทั้งหมดในชุดข้อมูลหรือฐานข้อมูล ซึ่งค่าจะอยู่ระหว่าง [0,1] หากค่าเข้าใกล้ 1 หมายความว่าเป็นรายการอาหารทั่วไปสูงและหากเข้าใกล้ 0 หมายความว่าเป็นรายการอาหารทั่วไปต่ำ โดยกำหนดให้ w_i คือ คำหรือวลีของรายการอาหารตำแหน่งที่ i ซึ่ง w_i จะมีลักษณะเป็นอาเรย์ของคำหรือวลี, $df(w_i)$ คือ ความถี่ (Frequency) ของ w_i ที่ปรากฏในเอกสาร (Documents) ในคลังเอกสารโดยหาก w_i ปรากฏในเอกสารใด ๆ มากกว่า 1 ครั้งจะถือว่าปรากฏแค่ 1 ครั้งเท่านั้น, n คือ จำนวนคำหรือวลีในรายการอาหารและวัตถุดิบ, N คือ จำนวนรายการอาหารทั้งหมด และผลลัพธ์คำนวณหาค่ารายการอาหารโดยทั่วไป (ψ) พบว่า ลักษณะของรายการอาหารแนะนำที่แตกต่างอยู่ 45.50% นั้น BoW ($\psi = 0.135$) มีลักษณะเป็นรายการอาหารโดยทั่วไปได้มากกว่ารายการอาหารที่แนะนำจาก TF-IDF ($\psi = 0.125$) เมื่อเทียบกับรายการอาหารทั้งหมดในชุดข้อมูล แต่อย่างไรก็ตามไม่ได้หมายความว่า BoW หรือ TF-IDF จะมีประสิทธิภาพที่ดีกว่าซึ่งกันและกัน

3.4.2 เทคนิคการกรองแบบมีส่วนร่วม

สำหรับเทคนิคนี้ดำเนินการเปรียบค่าผิดพลาดของ *MAE* และ *RMSE*

รูปที่ 6 พบว่า แบบจำลองที่ให้ค่า *MAE* และ *RMSE* น้อยที่สุดคือ SVD++ (*MAE* = 0.885, *RMSE* = 1.105) และตามด้วย KNNBasic (*MAE* = 1.031, *RMSE* =1.231) และ SVD (*MAE* = 1.734, *RMSE* = 2.150)



รูปที่ 6 การประเมินประสิทธิภาพแบบจำลองด้วยค่า MAE และ $RMSE$

4. สรุปและอภิปรายผล

งานวิจัยนี้จะทำการศึกษาเชิงสำรวจอัลกอริทึม 2 ประเภท ได้แก่ เทคนิคการกรองแบบอิงเนื้อหา (CBF) และเทคนิคการกรองแบบมีส่วนร่วม (CF)

CBF มีขั้นตอนเทคนิคในการสร้างแบบจำลองประกอบด้วย 5 ขั้นตอน คือ การเก็บข้อมูล การทำความสะอาดข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง และการประเมิน จากผลการศึกษาพบว่า เมื่อนำระบบที่พัฒนาด้วย BoW และ TF-IDF ร่วมกับการหาค่าความคล้ายคลึงแบบโคไซน์มาทำการทดลองแนะนำรายการอาหารสูงสุด Top 1 รายการด้วยรายการอาหารสืบค้นที่เหมือนกัน พบว่า BoW และ TF-IDF สามารถแนะนำรายการอาหารหรือเครื่องดื่มที่เหมือนกันคิดเป็น 54.50% แต่ยังพบว่า แตกต่างกันคิดเป็น 45.50% ผู้วิจัยจึงพิจารณาถึงความสอดคล้องของรายการอาหารที่แนะนำแตกต่างกันกับรายการอาหารที่ใช้สืบค้น พบว่า TF-IDF จะให้รายการอาหารที่มีความสอดคล้องมากกว่า BoW แม้ว่าค่า SIM ของ BoW จะสูงกว่า TF-IDF ซึ่งจากผลการศึกษาดังกล่าวทำให้ทราบว่าค่า SIM อาจไม่สามารถบ่งชี้ถือประสิทธิภาพของระบบแนะนำที่พัฒนาด้วย CBF ได้ และมากไปกว่านั้นผู้วิจัยได้ดำเนินการพัฒนาสมการเพื่อหาว่ารายการอาหารที่แนะนำที่แตกต่างกันจาก BoW และ TF-IDF นั้นมีลักษณะเป็นรายการอาหารโดยทั่วไปมากน้อยเพียงใดและพบว่า รายการอาหารที่ BoW แนะนำนั้นเป็นคุณลักษณะโดยทั่วไปมากกว่า แต่อย่างไรก็ตามค่า ψ ที่ได้ยังแตกต่างกันไม่มากนัก

CF มีขั้นตอน เทคนิคในการสร้างแบบจำลองประกอบด้วย 5 ขั้นตอน การเก็บข้อมูล การทำความสะอาดข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง และการประเมิน ซึ่งในการเตรียมข้อมูลมีการแบ่งเป็น 2 ชุด คือ ข้อมูลฝึกฝนและข้อมูลทดสอบ จากผลการวิจัยพบว่า SVD++ ให้ค่าผิดพลาดของ MAE = 0.885 และ $RSME$ = 1.105 ต่ำที่สุด และสอดคล้องกับงานวิจัยของ Shaikh [17] เนื่องจากค่าพารามิเตอร์ที่ใช้พัฒนาถูกปรับค่าและประเมินประสิทธิภาพให้เข้ากับชุดข้อมูลแล้วซึ่งจะต่างจากงานศึกษาของ Pannam และ Viyanon [5] ที่พัฒนาระบบแนะนำไอศกรีมด้วยการใช้ SVD แต่พารามิเตอร์พบว่าเป็นค่าเริ่มต้นและยังไม่ได้ปรับให้เข้ากับชุดข้อมูลใด ๆ

จากผลการศึกษาสามารถนำไปต่อยอดเพื่อพัฒนาระบบแนะนำในขอบเขตข้อมูลอื่น ๆ ได้ อย่างไรก็ตามเพื่อให้ระบบแนะนำมีประสิทธิภาพผู้วิจัยแนะนำว่าควรดำเนินทดสอบค่าพารามิเตอร์ให้เหมาะสมกับปัญหา และนอกจากนั้นสามารถนำสมการที่ได้เพื่อประเมินรายการอาหารเป็นทั่วไปหรือไม่

เอกสารอ้างอิง

[1] M. Zhou, Z. Ding, J. Tang, and D. Yin, “Micro behaviors: A new perspective in E-commerce recommender systems,” in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, Marina Del Rey, CA, USA, 2018, pp. 727–735.

[2] H. Ko, S. Lee, Y. Park, and A. Choi, “A survey of recommendation systems: Recommendation models, techniques, and application fields,” *Electronics*, vol. 11, no. 1, pp. 141, 2022.

[3] P. Chavan, B. Thoms, and J. Isaacs, “A Recommender system for healthy food choices: Building a hybrid model for recipe recommendations using big data sets,” in *Proceedings of the 54th Hawaii International*

Conference on System Sciences, HICCS-54, January 2021.

[4] R. Rahutomo, F. Lubis, H. H. Muljo, and B. Pardamean, "Preprocessing methods and tools in modelling Japanese for text classification," in *Proceedings 2019 International Conference on Information Management and Technology (ICIMTech)*, 2019, vol. 1, pp. 472–476.

[5] N. Pannam and W. Viyanon, "Food Recommendation system using a hybrid recommendation method," M.S. thesis, Faculty of Science, Srinakharinwirot University, 2021 (in Thai).

[6] M. Phanich, P. Pholkul, and S. Phimoltares, "Food recommendation system using clustering analysis for diabetic patients," in *Proceedings 2010 International Conference on Information Science and Applications*, 2010, pp. 1–8.

[7] W. Sriphalang and W. Sriurai, "The development of recommender system for E-library by using collaborative filtering and user profile," *Srinakharinwirot University (Journal of Science and Technology)*, vol. 9, no. 18, pp. 150–164, 2017.

[8] N. Nilesh, M. Kumari, P. Hazarika, and V. Raman, "Recommendation of Indian cuisine recipes based on ingredients," in *Proceedings 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW)*, 2019, pp. 96–99.

[9] F. Ramadhan and A. Musdholifah, "Online learning video recommendation system based on course and sylabus using content-based filtering," *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)*, vol. 15, no. 3, pp. 265–274, 2021.

[10] K. Patel and H. B. Patel, "A state-of-the-art survey on recommendation system and prospective extensions," *Computers and Electronics in Agriculture*, vol. 178, pp. 105779, 2020.

[11] P. Valdiviezo-Diaz, F. Ortega, E. Cobos, and R. Lara-Cabrera, "A collaborative filtering approach based on Naïve Bayes classifier," *IEEE Access*, vol. 7, pp. 108581–108592, 2019.

[12] A. Starke, C. Trattner, H. Bakken, M. Johannessen, and V. Solberg, "The cholesterol factor: Balancing accuracy and health in recipe recommendation through a nutrient-specific metric," in *CEUR Workshop Proceedings*, 2021.

[13] H. S. Song and Y. A. Kim, "A dog food recommendation system based on nutrient suitability," *Expert Systems*, vol. 38, no. 2, article no. e12623, 2021.

[14] R. N. Nakano, "Predicting optimal meal kit choices: A comparison of methods," Ph.D. dissertations, Department of Mathematics and Statistics, California State University, Long Beach, 2020.

[15] A. El Majjodi, A. D. Starke, and C. Trattner, "Nudging towards health? examining the merits of nutrition labels and personalization in a recipe recommender system," in *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, 2022, pp. 48–56.

[16] M. Elahi, N. El Ioini, A. Alexander Lambrix, and M. Ge, "Exploring personalized university ranking and recommendation," in *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, 2020, pp. 6–10.

[17] M. I. Shaikh, "Top-N nearest neighbourhood based movie recommendation system using different recommendation techniques," M.S. thesis, School of Computing, Dublin, National College of Ireland, 2020.

[18] A. D. Puspita, V. A. Permadi, A. H. Anggani, and E. A. Christy, "Musical instruments recommendation system using collaborative filtering and KNN," in *UMY Grace Proceedings*, 2021, vol. 1, no. 2, pp. 1–6.

[19] R. S. Gaikwad, S. S. Udmale, and V. K. Sambhe, "E-commerce recommendation system using improved probabilistic model," in *Information and Communication Technology for Sustainable Development: Springer*, 2018, pp. 277–284.

[20] Z. Romadhon, E. Sediyono, and C. E. Widodo, "Various implementation of collaborative filtering-based approach on recommendation systems using similarity," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 5, no. 3, pp. 179–186, 2020.

[21] N. Jenkarn and M. Ketcham, "Thai-textual cyberbullying detection using support vector machines," *Science Technology and Innovation (STIJ)*, vol. 1, no. 1, pp. 24–34, 2020 (in Thai).

[22] O. N. Osmanli and İ. H. Toroslu, "Using tag similarity in svd-based recommendation systems," in *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*, 2011, pp. 1–4.

[23] A. Al-Nafjan, N. Alrashoudi, and H. Alrasheed, "Recommendation System Algorithms on Location-Based Social Networks: Comparative Study," *Information*, vol. 13, no. 4, pp. 188, 2022.

[24] P .Sudasinghe," Enhancing book recommendation with the use of Reviews," M.S. thesis, School of Computing, University of Colombo, Sri Lanka, 2022.

[25] J. Yancheng, Z. Changhua, L. Qinghua, and W. Peng, "Users' brands preference based on SVD++ in recommender systems," in *2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)*, 2014, pp. 1175–1178.