การประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่มแบบไฮบริดร่วมกับเทคนิคการเรียนรู้ของเครื่องเพื่อแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม

APPLIED HYBRID SWARM INTELLIGENCE ALGORITHM WITH MACHINE LEARNING TECHNIQUES FOR APPROPRIATE NUTRITIOUS THAI FOOD RECOMMENDATIONS

ธนากร ปุรารัมย์

TANAKORN PURARAM

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2566

KMITL-2023-SC-M-002-048

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

APPLIED HYBRID SWARM INTELLIGENCE ALGORITHM WITH MACHINE LEARNING TECHNIQUES FOR APPROPRIATE NUTRITIOUS THAI FOOD RECOMMENDATIONS

TANAKORN PURARAM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR
THE DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE
DEPARTMENT OF COMPUTER SCIENCE, SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRANANG
2023

KMITL-2023-SC-M-002-048

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2023
SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่มแบบไฮบริดร่วมกับเทคนิคการเรียนรู้ของเครื่องเพื่อแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม
ชื่อนักศึกษา	นายธนากร ปุรารัมย์
รหัสประจำตัว	61605045
ปริญญา	วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
ภาควิชา	วิทยาการคอมพิวเตอร์
พ.ศ.	2566
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.วรางคณา กิ้มปาน
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	ดร.สุพร พงษ์นุ่มกุล

บทคัดย่อ

ผลสำรวจพฤติกรรมการบริโภคอาหารของคนไทยพบว่า ส่วนใหญ่เลือกรับประทานโดยคำนึงถึงความชอบมากที่สุด ตามด้วยรสชาติ ความสะอาดและคุณค่าทางโภชนาการตามลำดับ ทำให้อาจจะไม่ได้รับคุณค่าทางโภชนาการอย่างเหมาะสมกับช่วงวัยและสุขภาพ ปัจจุบันเทคโนโลยีทางปัญญาประดิษฐ์ได้รับความนิยมในวงการอาหารและสุขภาพเป็นวงกว้าง ซึ่งระบบแนะนำอัจฉริยะสามารถสร้างมูลค่าให้กับธุรกิจและอุตสาหกรรมการแพทย์ จากการสำรวจพบว่าระบบแนะนำอาหารไทยยังไม่แพร่หลายและไม่ได้พัฒนาด้านโภชนาการอาหารส่วนบุคคล ผู้วิจัยจึงนำเสนอการพัฒนาระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล โดยศึกษาและวิจัยการประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่มแบบไฮบริด Hybrid-Swarm-Km ประกอบด้วย Particle Swarm Optimization และ Cuckoo Search Algorithm ร่วมกับการจัดกลุ่มพฤติกรรมการรับประทานอาหารเชิงลึกด้วย K-Means และเทคนิคการแนะนำด้วย User-based Collaborative Filtering ซึ่งผลการศึกษาพบว่าขั้นตอนวิธีแบบไฮบริด คือ WCCSA-KM, WCPSO-KM และ PSO-CSA-KM สามารถเพิ่มประสิทธิภาพการจัดกลุ่มและแนะนำอาหารไทยได้ดีกว่าขั้นตอนวิธีอื่นโดยแบบจำลอง WCPSO-KM มีความถูกต้องด้วยค่า NDCG สูงที่สุด ทั้งนี้งานวิจัยได้สำรวจความพึงพอใจชุดอาหารที่แนะนำตามโภชนาการรายบุคคล พบว่า ผู้ใช้งานมีความพึงพอใจต่อชุดรายการอาหารในระดับพึงพอใจมากที่สุด

คำสำคัญ : การแนะนำอาหารไทย คุณค่าทางโภชนาการ ขั้นตอนวิธีความฉลาดแบบกลุ่มแบบไฮบริด ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Applied Hybrid Swarm Intelligence Algorithm with Machine Learning Techniques for Appropriate Nutritious Thai Food Recommendations
Student Name	Mr. Tanakorn Puraram
Student ID	61605045
Degree	Master of Science (Computer Science)
Department	Computer Science
Year	2023
Thesis Advisor	Asst.Prof.Dr.Warangkhana Kimpan
Thesis Co-advisor	Dr.Suporn Pongnumkul

Abstract

A recent survey on the food consumption behavior of Thai people revealed that their primary preference is taste, followed by cleanliness and nutritional value. However, many people may not consider the nutritional value suitable for their age and overall health. Artificial intelligence technology is widely popular in the food and health industry, and intelligent recommendation systems can provide significant benefits to businesses and the medical industry. Despite this, the study found that the Thai food recommendation system was not widespread and needed to develop personalized food nutrition. Therefore, this study proposed a Thai food recommendation system with appropriate nutritional value. The proposed method is the Hybrid-Swarm-KM algorithm, which includes Particle Swarm Optimization and Cuckoo Search Algorithm, deep dietary behavior clustering with K-Means, and recommendation techniques with User-based Collaborative Filtering. The experimental results show that the Hybrid algorithm: WCCSA-KM, WCPSO-KM, and PSO-CSA-KM recommended Thai food better than other algorithms. Among these, the WCPSO-KM algorithm was the most accurate with an NDCG score. Additionally, a satisfaction survey was conducted on food sets recommended by personal nutrition, which revealed that users were highly satisfied with the recommended food sets.

Keywords : Thai Food Recommendation, Nutrition Value, Hybrid Swarm Intelligence, Particle Swarm Optimization, Cuckoo Search Algorithm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ โดยได้รับคำแนะนำที่ดีทั้งความรู้ในการวิจัยขั้นตอนวิธีการทำวิทยานิพนธ์ การเรียบเรียงงานเขียนวิจัยภาษาอังกฤษจากอาจารย์ที่ปรึกษาทั้ง 3 คน ได้แก่ อาจารย์ที่ปรึกษาหลัก ผศ.ดร.วรางคณา กิ้มปาน ภาควิชาวิทยาการคอมพิวเตอร์ สจล. และอาจารย์ที่ปรึกษาร่วม ดร.สุพร พงษ์นุ่มกุล และดร.พิมพ์วดี เชาวลิต อาหวาด นักวิจัยอาวุโส ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

ขอขอบคุณ รศ.ดร.ไกรศักดิ์ เกษร อาจารย์ผู้ทรงคุณวุฒิจากภายนอกสถาบันฯ ประธานกรรมการสอบ และ ดร.รุ่งรัตน์ เวียงศรีพนาวัลย์ อาจารย์บัณฑิตประจำภาควิชา กรรมการสอบวิทยานิพนธ์ที่ได้ให้คำแนะนำและชี้แนะจนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วง

ขอขอบคุณ ดร. สุปิยา เจริญสิริวัฒน์ หัวหน้าทีม คุณอาภาธา ปี่ทอง คุณพงษ์ศักดิ์ ติยานันทิ ทีมวิจัยการวิเคราะห์พฤติกรรมมนุษย์ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ที่ได้สนับสนุนชุดข้อมูลพฤติกรรมการรับประทานอาหาร ข้อมูลอาหาร ข้อมูลโภชนาการ จากแอปพลิเคชัน FoodiEat และได้ร่วมพัฒนางานวิจัยเพื่อเป็นส่วนหนึ่งในการทำวิทยานิพนธ์ฉบับนี้

ขอขอบคุณ นางสาวสุธินี แป้นเงิน นักโภชนาการอาหาร โรงพยาบาลมหาวิทยาลัยเทคโนโลยีสุรนารี จังหวัดนครราชสีมา ที่ให้ความรู้และคำแนะนำเกี่ยวกับการแนะนำอาหารและคุณค่าทางโภชนาการให้กับคนทั่วไป ทำให้วิทยานิพนธ์ฉบับนี้สมบูรณ์และถูกต้องความหลักวิชาการ

ขอขอบคุณเพื่อนร่วมงานในการทำงานทุกคนไม่ว่าจะเป็นทีมวิจัยการวิเคราะห์พฤติกรรมมนุษย์ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ทีม Data Science – DNA บริษัท จีเอเบิล จำกัด และงานพัฒนาระบบงานข้อมูลสถาบันการเงิน ฝ่ายเทคโนโลยีสารสนเทศ ธนาคารแห่งประเทศไทย ที่ให้คำแนะนำและคอยสนับสนุนความรู้ในการพัฒนางานวิจัยของวิทยานิพนธ์ฉบับนี้

ขอขอบพระคุณคุณแม่ พี่ชาย แฟนและครอบครัวที่ได้ให้การสนับสนุนและให้กำลังใจในการเรียนและการทำวิจัยจนสำเร็จการศึกษาในระดับชั้นปริญญาโท และขอบคุณ 4 สาววง Blackpink ที่เป็นแรงบันดาลใจให้มีพลังในการทำวิจัยตลอด 4 ปีที่ผ่านมา ทำให้รู้ว่าความสำเร็จต้องแลกด้วยความพยายาม ความอดทนและไม่ย่อท้อต่ออุปสรรค ทำให้ก้าวผ่านช่วงเวลาที่ยากลำบากจนสำเร็จในวันนี้

งานวิจัยฉบับนี้ได้รับทุนสนับสนุนจาก “สถาบันบัณฑิตวิทยาศาสตร์และเทคโนโลยีไทย (TGIST) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ สัญญารับทุนเลขที่ SCA-CO-2562-9838-TH” และขอขอบคุณโครงการที่ให้การสนับสนุนทุนการศึกษา

ธนากร ปุรารัมย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ค
สารบัญ	ง
สารบัญตาราง	ฉ
สารบัญรูป	ซ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของงานวิจัย	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	3
1.4 ขั้นตอนการศึกษาและดำเนินงานวิจัย	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 คุณค่าทางโภชนาการอาหารไทย	4
2.2 การจัดกลุ่มข้อมูลเคมีน	6
2.3 ระบบแนะนำ	7
2.3.1 ชนิดของข้อมูลในระแบบแนะนำ	8
2.3.2 เทคนิคการพัฒนาระบบแนะนำ	8
2.3.3 การประเมินผลระบบแนะนำ	9
2.4 ขั้นตอนวิธีความฉลาดแบบกลุ่ม	11
2.4.1 ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค	12
2.4.2 ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า	15
2.5 การโปรแกรมเชิงเส้น	18
2.6 งานวิจัยที่เกี่ยวข้อง	18
2.6.1 การพัฒนาระบบแนะนำและเทคนิคระบบแนะนำรายการอาหาร	18
2.6.2 การปรับปรุงประสิทธิภาพของระบบแนะนำด้วย ขั้นตอนวิธีความฉลาดแบบกลุ่ม	20
บทที่ 3 การวิจัยและวิธีการดำเนินงาน	21
3.1 การรวบรวม จัดเตรียมและประมวลผลข้อมูล	23
3.1.1 ชุดข้อมูลที่ใช้ในการวิจัย	23
3.1.2 การเตรียมและประมวลผลข้อมูล	25
3.2 การจัดกลุ่มผู้ใช้งาน	28
3.2.1 การเตรียมข้อมูลสำหรับการจัดกลุ่มพฤติกรรมการรับประทานอาหาร	28
3.2.2 การลดมิติข้อมูลด้วยการวิเคราะห์องค์ประกอบหลัก	29
3.2.3 การเลือกจำนวนกลุ่มในการจัดกลุ่มข้อมูลพฤติกรรมการรับประทานอาหาร	29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
3.3 การทำนายและแนะนำอาหาร	30
3.3.1 รูปแบบข้อมูลสำหรับการพัฒนาแบบจำลองระบบแนะนำ	30
3.3.2 การเลือกชุดข้อมูลเรียนรู้และทดสอบแบบจำลอง	31
3.3.3 การเลือกขั้นตอนวิธีสำหรับการวิจัยแบบจำลองระบบแนะนำ	31
3.3.4 การทำนายค่าคะแนนความชื่นชอบ	31
3.3.5 การจัดลำดับรายการแนะนำ	33
3.4 การประเมินประสิทธิภาพของแบบจำลองการแนะนำรายการอาหาร	33
3.5 การพัฒนาและปรับปรุงแบบจำลอง	33
3.5.1 การพัฒนาแบบจำลองสำหรับการแนะนำรายการอาหาร	33
3.5.2 การทดลองการทำงานแบบจำลองโครงสร้างไฮบริด ด้วยพารามิเตอร์ที่ดีที่สุด	37
3.6 การคำนวณหาชุดรายการอาหารที่เหมาะสม	38
3.7 การแนะนำรายการอาหาร	40
3.8 การประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล	40
3.9 เครื่องมือสำหรับงานวิจัย	41
บทที่ 4 ผลการวิจัยและการอภิปรายผล	42
4.1 ผลการวิจัย	42
4.1.1 การประเมินประสิทธิภาพแบบจำลองแนะนำรายการอาหาร	42
4.1.2 การประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล	53
4.2 อภิปรายผล	54
4.2.1 การประเมินประสิทธิภาพแบบจำลองแนะนำรายการอาหาร	54
4.2.2 การประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล	55
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	57
5.1 สรุปผลการวิจัย	57
5.2 ข้อเสนอแนะ	57
เอกสารอ้างอิง	59
ภาคผนวก	63
ภาคผนวก ก งานวิจัยที่เผยแพร่	64
ภาคผนวก ข ผลการทดลองประสิทธิภาพแบบจำลอง Hybrid-Swarm-KM ตามจำนวนประชากรรายครั้ง	72
ภาคผนวก ค แบบสำรวจความพึงพอใจระบบแนะนำชุดรายการอาหารไทย	80
ประวัติผู้เขียน	83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 สัดส่วนการกระจายพลังงานของสารอาหารหลัก	6
3.1 ตัวอย่างข้อมูลประวัติการบันทึกการรับประทานอาหาร (Dietary Logs)	23
3.2 ตัวอย่างข้อมูลผู้ใช้งาน (Users)	23
3.3 ตัวอย่างข้อมูลรายการอาหาร (Foods)	24
3.4 ตัวอย่างข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs)	24
3.5 ตัวอย่างข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs)	25
3.6 ตัวอย่างข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity logs)	25
3.7 ช่วงคะแนนและค่า Rating ของพฤติกรรมการรับประทานอาหาร	27
3.8 รายละเอียดชุดข้อมูลสำหรับการจัดกลุ่มพฤติกรรมการรับประทานอาหาร	28
3.9 รายละเอียดชุดข้อมูลสำหรับการทำนายคะแนนระบบแนะนำ	30
3.10 รายการแบบจำลองการแนะนำรายการอาหาร	33
3.11 ค่าพารามิเตอร์เริ่มต้นของขั้นตอนวิธี Cuckoo Search Algorithm	37
3.12 ค่าพารามิเตอร์เริ่มต้นของขั้นตอนวิธี Particle Swarm Optimization	37
3.13 ค่าพารามิเตอร์ของการทดลองขั้นตอนวิธีความฉลาดแบบกลุ่ม	38
3.14 คะแนนความพึงพอใจเฉลี่ยของแบบสอบถามความพึงพอใจชุดรายการอาหารแนะนำ	41
4.1 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร	43
4.2 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร	44
4.3 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร	45
4.4 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร	46
4.5 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร	47
4.6 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร	49
4.7 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM	50
4.8 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM	51
4.9 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM	52
4.10 ผลการประเมินความพึงพอใจชุดรายการอาหารแนะนำ	53
4.11 การเปรียบเทียบประสิทธิภาพของแบบจำลองกลุ่ม Hybrid-Swarm-KM	55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่ **หน้า**

4.12 คะแนนความพึงพอใจเฉลี่ยต่อชุดรายการอาหารแนะนำ 56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2.1 การทำงานของการจัดกลุ่ม K-Means	6
2.2 โครงสร้างพื้นฐานในการพัฒนาระบบแนะนำ	8
2.3 การทำงานของเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม	9
2.4 การทำงานของการวัดประสิทธิภาพแบบจำลองด้วยวิธี Cross Validation	10
2.5 ขั้นตอนการทำงานของขั้นตอนวิธี Particle Swarm Optimization	13
2.6 การเคลื่อนที่ของนกในรูปแบบ Levy-Flights	16
2.7 Pseudo code การทำงานของขั้นตอนวิธีแบบนกดุเหว่า (Cuckoo Search)	17
3.1 โครงสร้างการทำงานระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม	22
3.2 การเตรียมและประมวลผลข้อมูลสำหรับการจัดกลุ่มและแนะนำรายการอาหาร	25
3.3 เดนโดแกรมของการจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งาน	30
3.4 ขั้นตอนการทำงานสำหรับการคำนวณทำนายค่าคะแนนความชื่นชอบ (Rating)	32
3.5 ตัวอย่างรายการแนะนำอาหารจำนวน 10 รายการ	33
3.6 โครงสร้างการทำงานการจัดกลุ่มข้อมูลของแบบจำลอง Hybrid-Swarm-KM	35
3.7 ตัวอย่างข้อมูลรายบุคคลและรายการแนะนำอาหารชุดที่เหมาะสม	39
4.1 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร	42
4.2 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร	43
4.3 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร	44
4.4 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร	46
4.5 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร	47
4.6 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร	48
4.7 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid	50
4.8 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid	51
4.9 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid	52
4.10 กราฟเปรียบเทียบประสิทธิภาพ แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid-Swarm-KM	55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของงานวิจัย

พฤติกรรมและวัฒนธรรมการทานอาหารของคนไทยในปัจจุบันมีความเปลี่ยนแปลงไปจากเดิม โดยการบริโภคของมนุษย์ในอดีตนั้นมีทางเลือกในการตัดสินใจบริโภคอาหารอย่างจำกัด ทำให้ไม่สามารถเลือกตามความต้องการหรือตามหลักโภชนาการได้มากนัก เมื่อวิวัฒนาการที่เปลี่ยนแปลงไปผนวกกับเทคโนโลยีที่เข้ามามีบทบาทในชีวิตประจำวันมากขึ้นนั้น ทำให้การคัดสรรและการเลือกรับประทานอาหารของผู้บริโภค รวมทั้งช่องทางที่สามารถเข้าถึงการบริโภคอาหารนั้นมีความหลากหลาย แปลกใหม่ สะดวกสบายและรวดเร็ว อาทิเช่น อาหารสำเร็จรูปในร้านสะดวกซื้อ การทานอาหาร fast food การทานอาหารที่มีการประกอบอาหารผ่านกระบวนการทางวิทยาศาสตร์หรือเทคโนโลยีขั้นสูง เป็นต้น ก่อให้เกิดพฤติกรรมในการบริโภคอาหารที่หลากหลายและแตกต่างกัน จากผลการสำรวจพฤติกรรมการบริโภคอาหารของประชากรไทย สำนักงานสถิติแห่งชาติ [1] พบว่า สิ่งที่คนส่วนใหญ่คำนึงถึงเป็นอันดับแรกก่อนเลือกซื้ออาหารคือ ความชอบมากที่สุด ร้อยละ 22.1 รองลงมาคือ รสชาติ ความอยากทาน และความสะอาด ร้อยละ 18.5 18.2 และ 17.8 คุณค่าทางโภชนาการ ร้อยละ 12.9 ความสะดวกรวดเร็วและราคา ร้อยละ 10.5 สำหรับกลุ่มอายุพบว่า ทุกกลุ่มอายุส่วนใหญ่ก่อนซื้ออาหารคำนึงถึงความชอบเป็นอันดับแรก ยกเว้นกลุ่มวัยสูงอายุ (60 ปีขึ้นไป) ที่คำนึงถึงคุณค่าทางโภชนาการเป็นอันดับแรก รองลงมาคือกลุ่มวัยเด็ก (6-14 ปี) จากข้อมูลข้างต้นพบว่า พฤติกรรมการบริโภคอาหารของคนส่วนใหญ่ยังคำนึงถึงอาหารที่มีคุณค่าทางโภชนาการเป็นอันดับรองลงมา ซึ่งถือว่าจำเป็นอย่างมากที่ร่างกายจะต้องได้รับคุณค่าสารอาหารที่มีประโยชน์ต่อสุขภาพและการเจริญเติบโตของร่างกาย

ระบบแนะนำ (Recommendation System) ได้รับการนำมาใช้งานอย่างแพร่หลายและสร้างมูลค่าให้กับธุรกิจ อาทิเช่น Amazon.com เว็บไซต์บริการขายสินค้าออนไลน์ แนะนำสินค้าที่ใกล้เคียงในรายการแนะนำจากการวิเคราะห์สินค้าที่ผู้ใช้กำลังสนใจ เคยซื้อหรือเคยค้นหา Netflix.com เว็บไซต์สำหรับชมภาพยนต์ออนไลน์ แนะนำภาพยนตร์ที่มีความคล้ายกับภาพยนตร์ที่ผู้ใช้งานแต่ละบุคคลชื่นชอบ เพื่อนำเสนอเนื้อหาแนะนำที่ปรับแต่งตามความชอบส่วนบุคคล (Personalize) ซึ่งทำให้มีจำนวนการเข้าชมภาพยนตร์เพิ่มขึ้นและสร้างความประทับใจให้แก่ผู้ใช้งานระบบที่สามารถเข้าชมภาพยนตร์ตรงตามความสนใจของตนเองมากขึ้น เป็นต้น สำหรับงานวิจัยและแอปพลิเคชันทางด้านการแนะนำเกี่ยวกับสุขภาพและโภชนาการอาหารนั้นมีการนำมาใช้งานอย่างแพร่หลายทั้งในภาคธุรกิจและทางการแพทย์ เช่น ระบบแนะนำเมนูอาหารสำหรับผู้ป่วยโรคต่าง ๆ เพื่อให้ผู้ป่วยได้รับสารอาหารที่เหมาะสมกับโรค สามารถแนะนำได้ว่าผู้ป่วยควรหรือไม่ควรบริโภคอาหารประเภทใด การแนะนำปริมาณอาหารและช่วงเวลาในการทานอาหารที่เหมาะสม การแนะนำเมนูอาหารและร้านค้าที่ได้รับความนิยมหรือมีลักษณะคล้ายคลึงกับสิ่งที่ผู้ใช้กำลังสนใจ เป็นต้น โดยระบบในการแนะนำรายการอาหารจะสามารถช่วยให้ผู้ใช้งานได้รับอาหารที่ชื่นชอบและมีคุณค่าทางโภชนาการอย่างเหมาะสม ตรงต่อความต้องการ และเป็นเครื่องมือที่ช่วยให้ผู้ใช้งานสามารถนำไปวางแผนในการดูแลสุขภาพได้ อีกทั้งเพื่อเป็นทางเลือกให้ผู้ใช้งานในการเลือกรับประทานอาหารที่มีคุณค่าทางโภชนาการเหมาะสมกับพฤติกรรมของแต่ละบุคคล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เทคนิคในการพัฒนาระบบแนะนำที่ได้รับความนิยม ได้แก่ เทคนิคการคัดกรองแบบร่วมมือ (Collaborative Filtering) เทคนิคการพิจารณาจากเนื้อหา (Content Based) และเทคนิคผสมผสาน (Hybrid) [2] ซึ่งเป็นเทคนิคที่นำมาใช้งานอย่างยาวนานและมีการปรับปรุงเพื่อให้สามารถทำงานร่วมกับเทคโนโลยีในปัจจุบันได้ดียิ่งขึ้น สำหรับการพัฒนาระบบแนะนำนั้น การแนะนำสินค้า (item) ให้ตรงต่อความสนใจของผู้ใช้งานระบบมากที่สุดถือว่าเป็นเรื่องยากและมีความท้าทายอย่างยิ่ง จึงต้องการขั้นตอนวิธีที่มีความแม่นยำและประสิทธิภาพในการทำงานที่สูงขึ้น ซึ่งการพัฒนาระบบแนะนำสามารถประยุกต์เข้าร่วมกับขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) เพื่อปรับปรุงประสิทธิภาพของระบบแนะนำ ตัวอย่างเช่น ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) เป็นต้น โดยข้อดีของเทคนิคความฉลาดแบบกลุ่มนั้นสามารถคำนวณหาค่าข้อมูลที่มีความเหมาะสมที่สุดในกลุ่มข้อมูลนั้นได้ดี ผนวกกับการทำงานที่รวดเร็ว ส่งผลให้ระบบแนะนำมีประสิทธิภาพการทำงานเพิ่มมากขึ้น

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าการพัฒนาระบบงานแนะนำอาหารไทยยังไม่ได้นำข้อมูลโภชนาการอาหารไทย [3] และขั้นตอนวิธีความฉลาดแบบกลุ่ม [4] มาประยุกต์ใช้ในการพัฒนาเพื่อเพิ่มประสิทธิภาพการแนะนำอาหารให้ตรงต่อพฤติกรรมรายบุคคลได้ ดังนั้น งานวิจัยฉบับนี้ผู้วิจัยจึงนำเสนอระบบแนะนำโดยพัฒนาจากเทคนิคปัญญาประดิษฐ์เพื่อแนะนำอาหารที่เหมาะสมกับพฤติกรรม ความชื่นชอบ และมีคุณค่าทางโภชนาการแก่ผู้ใช้งาน ประยุกต์เข้ากับขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) ซึ่งได้รับแรงจูงใจจากพฤติกรรมทางธรรมชาติของสัตว์สังคม ประกอบด้วย ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) และขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) โดยขั้นตอนวิธีดังกล่าวสามารถนำมาประยุกต์ใช้ในการปรับปรุงประสิทธิภาพของระบบแนะนำ เพื่อให้ได้ผลลัพธ์ที่เหมาะสมที่สุดในการแนะนำอาหารให้ตรงตามหลักโภชนาการและมีคุณค่าทางสารอาหารมากที่สุดแก่ผู้ใช้งานรายบุคคลได้ สำหรับข้อมูลที่นำมาใช้ในงานวิจัยเพื่อวิเคราะห์หาความสัมพันธ์เชิงลึกของผู้ใช้งานรายบุคคลและพัฒนาขั้นตอนวิธีในการแนะนำอาหาร ได้แก่ ข้อมูลพฤติกรรมการบริโภคอาหารและข้อมูลสุขภาพของผู้ใช้งานในแอปพลิเคชัน FoodiEat ของทีมวิจัยวิเคราะห์พฤติกรรมมนุษย์ (Human Behavior Analytics) ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) ข้อมูลอาหารและโภชนาการสำหรับทุกช่วงวัยของคณะกรรมการอาหารแห่งชาติ และข้อมูลอาหารอื่น ๆ จากทีมวิจัยที่ร่วมมือกับสถาบันโภชนาการ มหาวิทยาลัยมหิดล สำนักงานคณะกรรมการอาหารและยา และกรมอนามัย กระทรวงสาธารณสุข ซึ่งจะช่วยเพิ่มความถูกต้องและความน่าเชื่อถือตามหลักโภชนาการมากยิ่งขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

1) เพื่อประยุกต์ใช้งานเทคนิคความฉลาดแบบกลุ่ม (Swarm Intelligence) แบบไฮบริดสำหรับการปรับปรุงประสิทธิภาพการทำงานของระบบแนะนำ

2) เพื่อพัฒนาเทคนิคในการพัฒนาระบบแนะนำอาหารอย่างเหมาะสมต่อผู้ใช้แต่ละบุคคลโดยใช้วิธีการเรียนรู้ของเครื่อง (Machine Learning) ทางด้านระบบแนะนำและเพิ่มประสิทธิภาพการทำงานด้วย

3) เพื่อวิเคราะห์ข้อมูลพฤติกรรมการบริโภคอาหารและข้อมูลสุขภาพจากการบันทึกข้อมูลของผู้ใช้งานแต่ละบุคคลผ่านแอปพลิเคชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตของงานวิจัย

1) งานวิจัยนี้ทำการศึกษาจากชุดข้อมูลการบันทึกข้อมูลการรับประทานอาหาร การออกกำลังกาย การบันทึกน้ำหนัก ส่วนสูงของผู้ใช้งานแอปพลิเคชัน FoodiEat ช่วงอายุ 15-70 ปี ทั้งเพศชายและหญิง จำนวน 300,000 ครั้ง

2) งานวิจัยนี้ใช้ข้อมูลคุณค่าโภชนาการอาหารไทย (Nutritive Values of Meals) ที่ได้รับความร่วมมือจากทีมวิจัยการวิเคราะห์พฤติกรรมมนุษย์ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ ร่วมมือกับสถาบันโภชนาการ มหาวิทยาลัยมหิดล สำนักงานคณะกรรมการอาหารและยา และกรมอนามัย กระทรวงสาธารณสุข ในการคำนวณคุณค่าทางโภชนาการอาหาร

3) งานวิจัยนี้มุ่งเน้นการประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) แบบไฮบริด โดยเป็นการผสมผสานขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) ร่วมกับการแบ่งกลุ่มข้อมูลแบบเคมีน (K-Means) เพื่อการจัดกลุ่มพฤติกรรมการกินอาหาร และประยุกต์ใช้เทคนิคระบบแนะนำแบบการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) ในการแนะนำเมนูอาหาร

1.4 ขั้นตอนการศึกษาและดำเนินงานวิจัย

1) ศึกษาทฤษฎีทางด้านการวิเคราะห์ข้อมูลพฤติกรรมผู้ใช้งาน เทคนิคระบบแนะนำ (Recommendation System Techniques) เทคนิคความฉลาดแบบกลุ่ม (Swarm Intelligence) และศึกษาข้อมูลทางด้านโภชนาการอาหาร (Nutritive Values of Meals)

2) รวบรวม วิเคราะห์ และประมวลผลข้อมูลการบันทึกการทานอาหารและข้อมูลสุขภาพรวมทั้งข้อมูลคุณค่าโภชนาการอาหาร (Nutritive Values of Meals)

3) ออกแบบงานวิจัยสำหรับการทำงานระบบแนะนำโดยใช้เทคนิคความฉลาดแบบกลุ่ม ในการประยุกต์ใช้ร่วมกับเทคนิคระบบแนะนำ

4) ทดลองงานวิจัย

5) วิเคราะห์และประเมินผลการทดลอง

6) สำรวจความพึงพอใจของผู้ใช้งานระบบแนะนำ

7) สรุปผลทดลองพร้อมเสนอแนวทางการพัฒนางานวิจัย

8) เรียบเรียงวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1) ได้ขั้นตอนวิธีที่มีประสิทธิภาพเพิ่มขึ้นเพื่อสามารถนำพัฒนาต่อยอดเป็นฟังก์ชันแนะนำอาหารผ่านแอปพลิเคชันได้

2) ได้พัฒนาและปรับปรุงประสิทธิภาพของขั้นตอนวิธีระบบแนะนำอาหารไทย

3) ได้ระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการตรงต่อพฤติกรรมและช่วงวัยผู้ใช้งานรายบุคคล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาและพัฒนาระบบแนะนำอาหารไทยในปัจจุบันยังไม่ได้นำข้อมูลโภชนาการอาหารไทยและเทคนิคกลุ่มขั้นตอนวิธีความฉลาดแบบกลุ่มมาประยุกต์ใช้เพื่อปรับปรุงประสิทธิภาพการแนะนำรายการอาหารที่มีคุณค่าโภชนาการอย่างเหมาะสมรายบุคคล ดังนั้น เนื้อหาในบทนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องกับการวิจัยและพัฒนาเทคนิคระบบแนะนำอาหารไทย ได้แก่ คุณค่าทางโภชนาการอาหารไทยและพฤติกรรมการรับประทานอาหารของคนไทย การจัดกลุ่มข้อมูล K-Means เทคนิคระบบแนะนำ ขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization : PSO) ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) การโปรแกรมเชิงเส้น (Linear Programming) และงานวิจัยที่เกี่ยวข้อง สำหรับการศึกษาและพัฒนาระบบแนะนำอาหาไทย

2.1 คุณค่าทางโภชนาการอาหารไทย

คุณค่าทางโภชนาการอาหาร เป็นสิ่งที่จำเป็นและมีประโยชน์ต่อร่างกาย ซึ่งในแต่ละช่วงวัยจำเป็นต้องเลือกบริโภคอาหารที่จะให้ได้รับคุณค่าทางโภชนาการอย่างเหมาะสม เพราะแต่ละช่วงวัยนั้นจะต้องการสารอาหารในปริมาณและความหลากหลายที่ต่างกัน หากได้รับสารอาหารน้อยหรือมากเกินไป เป็นผลให้เกิดปัญหาทางโภชนาการ อาทิ ภาวะเตี้ย ผอม อ้วน เป็นต้น คณะกรรมการอาหารแห่งชาติ (2559) [5] ได้แบ่งกลุ่มตามช่วงอายุจำนวน 5 ช่วงอายุ ได้แก่ 1) อาหารและโภชนาการสำหรับทารกแรกเกิดถึงอายุ 5 ปี 2) อาหารและโภชนาการสำหรับเด็กระดับประถมศึกษา (6-12 ปี) 3) อาหารและโภชนาการสำหรับวัยรุ่น 4) อาหารและโภชนาการสำหรับผู้บริโภควัยทำงานและหญิงเจริญพันธุ์ 5) อาหารและโภชนาการสำหรับผู้สูงอายุ (> 60 ปี) นอกจากการแบ่งกลุ่มตามช่วงอายุยังสามารถแบ่งกลุ่มตามอาการหรือโรคของผู้ป่วย อาทิเช่น อาหารสำหรับผู้ที่เป็นโรคไตเรื้อรัง อาหารสำหรับผู้ป่วยโรคไขมันในเลือดสูง อาหารสำหรับผู้ป่วยโรคความดันโลหิตสูง หัวใจ และอัมพาต อาหารสำหรับผู้ที่เป็นโรคอ้วน โรคเบาหวาน รวมทั้งการบริโภคอาหารของหญิงตั้งครรภ์ ดังนั้น ปริมาณสารอาหารที่เหมาะสมกับความต้องการในแต่ละวันของทุกเพศและวัย จึงมีความสำคัญต่อระบบการทำงานของร่างกาย

กลุ่มงานวิเคราะห์อาหารและโภชนาการ กองโภชนาการ ดำเนินการวิเคราะห์คุณค่าทางโภชนาการของอาหารไทย โดยทำการจัดกลุ่มอาหาร (Food groups) ทั้งหมด 14 กลุ่ม [6] ได้ดังนี้

1. ธัญพืชและผลิตภัณฑ์ (Cereals and products)
2. ราก หัวของพืช และผลิตภัณฑ์ (Starchy roots, tubers and products)
3. ผลไม้เปลือกแข็ง พืชเมล็ด ถั่วเมล็ดแห้ง และผลิตภัณฑ์ (Pulses, nuts, seeds and products)
4. ผักและผลิตภัณฑ์ (Vegetables and products)
5. ผลไม้ และผลิตภัณฑ์ (Fruits and products)
6. เนื้อสัตว์ เป็ด ไก่ และผลิตภัณฑ์ (Meat, poultry and products)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. ปลา กุ้ง สัตว์น้ำอื่น ๆ และผลิตภัณฑ์จากสัตว์น้ำ (Finfish, shellfish, other aquatic animals and products)
8. ไข่ (Eggs)
9. นมและผลิตภัณฑ์ (Milk and products)
10. เครื่องเทศและเครื่องปรุงรส (Spices and condiments)
11. อาหารปรุงสำเร็จและอาหารจานเดียว (Local mixed food dishes and one plate dishes)
12. ขนมหวาน และขนมขบเคี้ยว (Desserts and snacks)
13. แมลง (Insects)
14. อื่น ๆ (Miscellaneous)

การกำหนดความต้องการพลังงานที่ควรได้รับประจำวันเป็นไปตามอัตราการครองธาตุพื้นฐาน (Basal Metabolic Rate : BMR) ซึ่งแตกต่างกันขึ้นกับเพศ อายุ น้ำหนัก ส่วนสูงและระดับของกิจกรรม ร่างกายต้องการพลังงานจากอาหารที่บริโภค เพื่อการทำงานของเซลล์ในระบบต่าง ๆ การประกอบกิจกรรมในชีวิตประจำวัน การรักษาอุณหภูมิของร่างกาย การเจริญเติบโต ถ้าได้รับพลังงานน้อยเกินไปจะมีผลต่อการชะลอ การเจริญเติบโตและระบบการทำของร่างกายพร่องได้ แต่ถ้าได้รับพลังงานมากเกินไป จะมีการสะสมพลังงาน โดยเปลี่ยนเป็นไขมันสะสมตามส่วนต่าง ๆ ของร่างกาย

หน่วยของพลังงาน กล่าวคือ พลังงานมีหน่วยเป็นแคลอรี (Calorie) 1 แคลอรี หมายถึง ปริมาณความร้อนที่ทำให้น้ำบริสุทธิ์ 1 มิลลิลิตรที่อุณหภูมิ 15 องศาเซลเซียส มีอุณหภูมิสูงขึ้น 1 องศาเซลเซียส แคลอรีเป็นหน่วยที่ใช้กันตั้งแต่เริ่มต้นเนื่องจากเป็นการวัดปริมาณความร้อน ต่อมาปี พ.ศ. 2513 มีการปรับเป็นการวัดในรูปของพลังงานมีหน่วยเป็นจูล (joule) ซึ่งเป็นหน่วยวัดการทำงาน (mechanical work) พลังงาน 1 จูล หมายถึง พลังงานที่เคลื่อนวัตถุด้วยแรง 1 นิวตัน ได้ระยะทาง 1 เมตร ดังนั้น 1 กิโลแคลอรี มีค่าเท่ากับ 4.18 กิโลจูล หรือ 1 กิโลจูล มีค่าเท่ากับ 0.24 กิโลแคลอรี การนำสารอาหาร ได้แก่ คาร์โบไฮเดรต โปรตีน ไขมัน ไปทำการเผาไหม้โดยใช้เครื่อง Bomb calorimeter จากนั้นคิดปัจจัยของการย่อยและการดูดซึม ในการปรับเปลี่ยนเพื่อให้ได้ค่าพลังงานที่แท้จริงในร่างกาย จากการคำนวณโดยใช้ Atwater factor จะได้ค่าพลังงานจากสารอาหาร [7] ดังนี้

คาร์โบไฮเดรต	1 กรัม	ให้พลังงาน 4 กิโลแคลอรี (17 กิโลจูล)
โปรตีน	1 กรัม	ให้พลังงาน 4 กิโลแคลอรี (17 กิโลจูล)
ไขมัน	1 กรัม	ให้พลังงาน 9 กิโลแคลอรี (38 กิโลจูล)

การกำหนดความต้องการพลังงาน เป็นไปตามการใช้พลังงานของร่างกายทั้งหมด (Total Energy Expenditure : TEE) ซึ่งประกอบด้วยพลังงานการครองธาตุพื้นฐาน (Basal Energy Expenditure : BEE) บางครั้งเรียกว่าอัตราการครองธาตุพื้นฐาน (Basal Metabolic Rate : BMR) เป็นพลังงานจำเป็นต่อการดำรงชีวิต เพื่อการทำงานของเซลล์และอวัยวะต่าง ๆ เช่น หัวใจ ปอด หลอดเลือด ทางเดินอาหาร ระบบประสาท ซึ่งใกล้เคียงกับการใช้พลังงานขณะพักผ่อน พลังงานที่ใช้ในกิจกรรมทางกาย และพลังงานความร้อนจากการบริโภคอาหาร นอกจากนี้สำหรับทารกและเด็กยังต้องการพลังงานเพิ่มขึ้นสำหรับการเจริญเติบโตด้วย รวมทั้งหญิงตั้งครรภ์และหญิงให้นมบุตรก็ต้องการพลังงานเพิ่มขึ้นสำหรับการสร้างเนื้อเยื่อของทารก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แนวทางการบริโภคอาหารเพื่อให้ได้รับปริมาณพลังงานที่เหมาะสม

การเลือกบริโภคอาหารเพื่อให้ได้รับปริมาณพลังงานที่เพียงพอกับความต้องการของร่างกายนั้น พิจารณา จากแหล่งสารอาหารหลักที่ให้พลังงาน ได้แก่ คาร์โบไฮเดรต ไขมัน และโปรตีน เป็นหลัก [7]

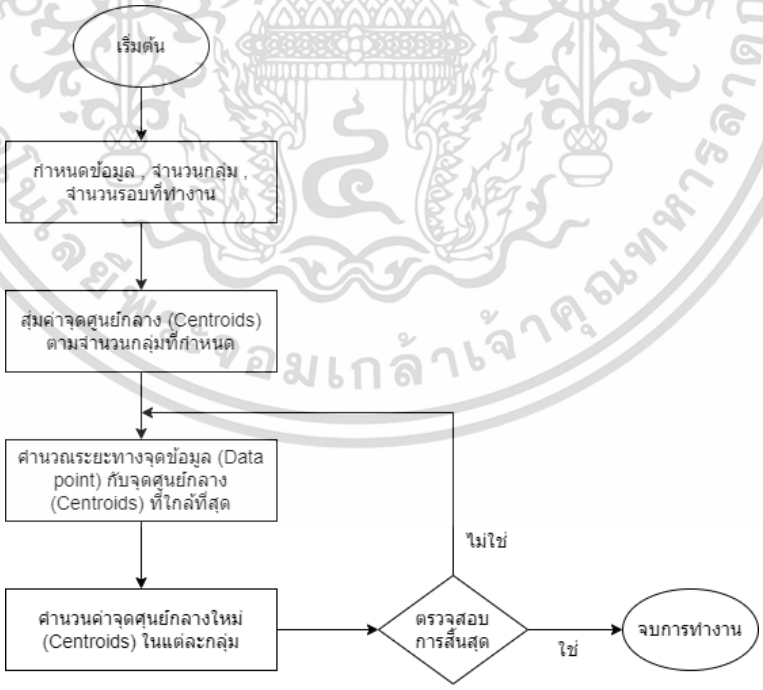
ตารางที่ 2.1 ตารางสัดส่วนการกระจายพลังงานของสารอาหารหลัก

สารอาหาร	สัดส่วนการกระจายพลังงาน
คาร์โบไฮเดรต	45-65
โปรตีน	10-15
ไขมัน	20-35

หมายเหตุ สัดส่วนการกระจายพลังงานของความต้องการพลังงานที่ร่างกายควรได้รับใน 1 วัน

2.2 การจัดกลุ่มข้อมูล K-Means

การจัดกลุ่มข้อมูล K-Means หรือเรียกอีกอย่างหนึ่งว่า การวิเคราะห์กลุ่มแบบไม่เป็นขั้นตอน (Nonhierarchical Cluster Analytics) เป็นเทคนิคการเรียนรู้ของเครื่องประเภทหนึ่งแบบไม่มีผู้สอน (Unsupervised Learning) ซึ่งเป็นวิธีการจัดกลุ่มที่ต้องทราบจำนวนกลุ่มก่อนการทำงานและเป็นการจัดกลุ่มข้อมูลแบบแบ่งส่วน (Partition) กล่าวคือข้อมูลที่มีลักษณะใกล้เคียงกันจะได้รับการแบ่งกลุ่มให้อยู่ในกลุ่ม (Cluster) เดียวเท่านั้น โดยใช้ค่าจุดศูนย์กลาง (Centroids) เป็นตัวกำหนดระยะห่างของข้อมูลที่ใกล้ที่สุดในกลุ่มเดียวกัน [8] มีขั้นตอนการทำงานดังรูปที่ 2.1



รูปที่ 2.1 การทำงานของการจัดกลุ่ม K-Means

จากรูปที่ 2.1 แสดงขั้นตอนการทำงานของการจัดกลุ่ม K-Means โดยมีขั้นตอนการทำงานทั้งหมด 5 ขั้นตอน ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. กำหนดข้อมูล จำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูล และจำนวนรอบในการคำนวณหาจุดศูนย์กลาง (Centroids) และปรับกลุ่มข้อมูล
2. เริ่มต้นการทำงานด้วยการสุ่มจุดศูนย์กลาง (Centroids) ตามจำนวนกลุ่มที่กำหนดไว้
3. ดำเนินการคำนวณระยะทางของข้อมูล (Data point) ไปยังจุดศูนย์กลางที่ใกล้ที่สุด โดยใช้สูตรระยะทางแบบยุคลิด (Euclidean Distance)

$$\mathrm{d}(\mathrm{x},\mathrm{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

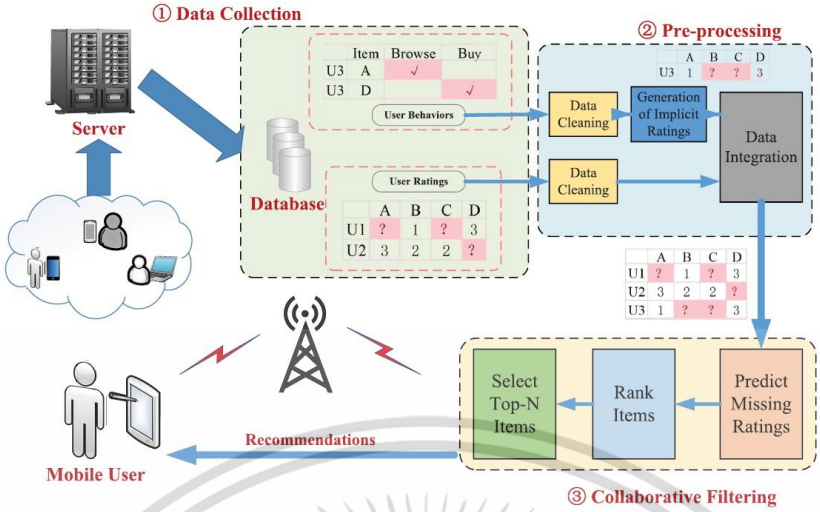
4. ดำเนินการคำนวณค่าจุดศูนย์กลางใหม่ โดยใช้ค่าระยะทางเฉลี่ยจากทุกจุดข้อมูลภายในกลุ่ม และกำหนดเป็นจุดศูนย์กลางใหม่
5. ดำเนินการวนการทำงานตามข้อ 3-5 จนกว่าจะจบการทำงานตามเงื่อนไขที่กำหนด เช่น ครบจำนวนรอบที่กำหนดการทำงาน หรือค่าจุดศูนย์กลางของกลุ่ม (Centroids) ไม่มีการเปลี่ยนแปลงแล้วทุกกลุ่ม

2.3 ระบบแนะนำ (Recommendation System)

ระบบแนะนำ (Recommendation System) คือ ระบบสนับสนุนการตัดสินใจที่ให้การแนะนำสินค้าหรือบริการ ซึ่งสามารถแทนด้วย วัตถุ (item) ที่มีความเหมาะสมกับรูปแบบและพฤติกรรมของลูกค้าแต่ละคน แทนด้วย ผู้ใช้ (user) โดยอาศัยข้อมูลของผู้ใช้งานร่วมกับข้อมูลประกอบภายนอกมาใช้ในการวิเคราะห์คัดกรองให้ได้สิ่งที่มีความเหมาะสมต่อผู้ใช้งานมากที่สุด การทำงานของระบบผู้แนะนำพยายามวิเคราะห์และทำนายว่าผู้ใช้จะให้คะแนนความชอบหรือมีพฤติกรรมต่อสิ่งที่ถูกแนะนำนั้น ๆ อย่างไร ซึ่งกระบวนการทำงานดังกล่าวได้ถูกจำแนกออกเป็นสองขั้นตอน คือขั้นตอนการทำนาย (prediction phase) และขั้นตอนการแนะนำ (recommendation phase) สำหรับการสร้างระบบผู้แนะนำนั้นสามารถสร้างได้จากหลายวิธีการ [9] ซึ่งจำแนกออกเป็น 3 ชนิดหลัก ได้แก่ 1) Content-based, 2) Collaborative Filtering, และ 3) Hybrid recommendation ในการพัฒนาระบบแนะนำ พบว่ามีปัญหาหลัก 2 ปัญหา [10] ได้แก่ Cold-start problem คือ ปัญหาที่ระบบไม่สามารถดึงข้อสรุปใด ๆ สำหรับผู้ใช้ใหม่หรือสินค้า (item) รายการที่ยังไม่ได้รวบรวมข้อมูลเพียงพอ และปัญหา Sparsity data คือการมีข้อมูลในปริมาณที่น้อย ทำให้พบการขาดหายข้อมูลเป็นจำนวนมากเมื่อเปรียบเทียบสัดส่วนต่อจำนวนของผู้ใช้งานกับจำนวนวัตถุ

โครงสร้างพื้นฐานในการพัฒนาระบบแนะนำ ยกตัวอย่างโดยการใช้เทคนิค Collaborative Filtering ดังรูปที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 โครงสร้างพื้นฐานในการพัฒนาระบบแนะนำ

2.3.1 ชนิดของข้อมูลในระบบแนะนำ (Recommendation System)

ชนิดของข้อมูลในระบบแนะนำสามารถแบ่งได้เป็น 2 ประเภท ดังต่อไปนี้

1. Explicit Feedback (User Rating) ข้อมูลที่ได้จากผู้ใช้งานโดยตรง [11] เช่น การให้คะแนนภาพยนตร์ (Rating) จากคะแนน 1 ถึง 5 ซึ่งถือว่าข้อมูลดังกล่าวเป็นข้อมูลที่ผู้ใช้งานยินยอมในการให้คะแนนและสามารถบ่งบอกถึงความรู้สึกต่อวัตถุนั้นได้อย่างดี โดยสามารถคัดแยกความสนใจของผู้ใช้เป็นเชิงบวกและเชิงลบได้ เช่น หากผู้ใช้งานให้คะแนนภาพยนตร์ 4-5 คะแนน หมายถึงผู้ใช้งานชื่นชอบภาพยนตร์เรื่องนั้นมาก ให้ความสนใจในเชิงบวก แต่หากผู้ใช้งานให้คะแนนภาพยนตร์เพียง 1 คะแนน หมายถึงผู้ใช้งานอาจจะไม่ชอบภาพยนตร์เรื่องนี้ เป็นต้น

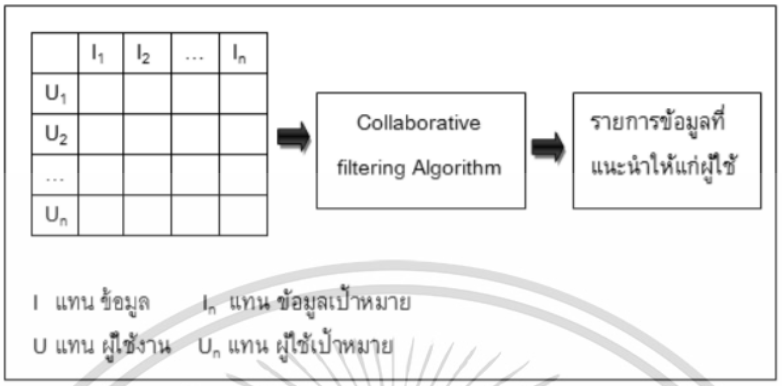
2. Implicit Feedback (User Behavior) ข้อมูลที่ไม่ได้มาจากผู้ใช้งานโดยตรง [7] ซึ่งตรงข้ามกับข้อมูล Explicit Feedback กล่าวคือ ข้อมูลที่มาจากการพฤติกรรมของผู้ใช้งาน เช่น การเข้าชมหนังสือที่สนใจ ระยะเวลาในการเข้าชมภาพยนตร์ การเลือกรับประทานอาหารในแต่ละมื้อ เป็นต้น โดยสามารถรวบรวมพฤติกรรมของผู้ใช้งานเป็นเชิงปริมาณ เช่น จำนวนครั้งที่รับประทานอาหาร เวลาทั้งหมดในการเข้าชมภาพยนตร์เรื่องนั้น ข้อมูล Implicit Feedback นั้นอาจไม่เหมาะในการคัดแยกความสนใจของผู้ใช้เป็นเชิงบวกและเชิงลบตามปริมาณในการกระทำต่อวัตถุได้ เช่น ผู้ใช้มีพฤติกรรมต่อวัตถุหนึ่งมากเกิน 5 ครั้ง ไม่ได้หมายถึงว่าผู้ใช้คนนั้นมีความสนใจที่แท้จริงต่อวัตถุ หรือผู้ใช้ไม่มีพฤติกรรมใดเลยต่อวัตถุหนึ่งทำให้ข้อมูลเชิงปริมาณเป็น 0 ก็ไม่ได้หมายความว่าผู้ใช้คนนั้นจะมีความสนใจเชิงลบต่อวัตถุ อาจเป็นเพราะผู้ใช้ไม่เคยได้รับรู้ว่ามีวัตถุนั้นอยู่ก็เป็นได้ ดังนั้น ข้อมูลชนิด Implicit Feedback จึงไม่สามารถคัดแยกความสนใจของผู้ใช้เป็นเชิงบวกหรือเชิงลบได้อย่างชัดเจนเหมือนกับ Explicit Feedback

2.3.2 เทคนิคการพัฒนาระบบแนะนำ (Recommendation System Techniques)

เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) เป็นเทคนิคการกรองข้อมูลที่ได้รับความนิยมในการนำไปใช้แนะนำข้อมูล ซึ่งแนวความคิดพื้นฐานของการแนะนำข้อมูลของวิธีนี้ขึ้นอยู่กับความคิดเห็นของผู้ใช้ในระบบ ซึ่งระบบทำการค้นหากลุ่มสมาชิกใกล้เคียง (Neighbors)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่มีความชอบหรือพฤติกรรมเหมือนกันกับผู้ใช้งานเป้าหมาย (Active User) เช่น ความชอบภาพยนตร์ หนังสือ การรับประทานอาหาร เป็นต้น โดยข้อมูลจากกลุ่มสมาชิกใกล้เคียงนี้ได้นำมาวิเคราะห์ในระบบเพื่อแนะนำข้อมูลไปยังผู้ใช้เป้าหมายต่อไป ดังรูปที่ 2.3



รูปที่ 2.3 การทำงานของเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) มีขั้นตอนการทำงาน 4 ขั้นตอน ได้แก่

ขั้นตอนที่ 1) การคำนวณความคล้ายคลึง (Similarity Computation) ระหว่างผู้ใช้ 2 คน โดยมีการรวบรวมข้อมูลในรูปแบบตาราง (Matrix) ซึ่งมีวิธีการที่ได้รับความนิยม 2 วิธี คือ การคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Correlation Coefficient) และ Cosine-based

ขั้นตอนที่ 2) การเลือกสมาชิกใกล้เคียง (Neighbor Selection) เป็นวิธีการที่มีความสำคัญ โดยทำการเลือกจากผู้ใช้ที่มีอยู่ภายในระบบทั้งหมดเพื่อใช้ในการทำนาย ซึ่งเทคนิคหลักในการเลือกสมาชิกใกล้เคียงมี 2 วิธีคือ Similarity Threshold และ Best K-Neighbor

ขั้นตอนที่ 3) การทำนาย (Prediction) เป็นการพยากรณ์ค่าความพึงพอใจของผู้ใช้ต่อรายการ (item) ใดรายการหนึ่ง โดยพิจารณาจากความพึงพอใจและความคล้ายคลึงกันระหว่างรายการนั้นกับรายการอื่น ซึ่งจะนำกลุ่มของสมาชิกข้างเคียงที่ได้เลือกข้อมูลแล้ว นำมาคำนวณเพื่อสร้างรายการแนะนำต่อไป

ขั้นตอนที่ 4) การสร้างรายการแนะนำ (Recommendation) โดยนำค่าที่ได้จากการพยากรณ์ในแต่ละรายการ (item) มาเรียงตามลำดับ เริ่มตั้งแต่รายการที่มีค่าการพยากรณ์มากที่สุดจนถึงรายการที่มีค่าการพยากรณ์น้อยที่สุด การเลือกจำนวนรายการแนะนำมาแสดงสามารถเลือกจำนวนได้ตามความต้องการของผู้ใช้ว่าจะให้มีการแสดงผลการแนะนำจำนวนเท่าใด เช่น การแนะนำจำนวน 10 รายการไปยังผู้ใช้งาน (Top-10 Recommendation)

เทคนิคที่ใช้ในการกรองสารสนเทศแบบพึ่งพาผู้ใช้มีการนำหลายเทคนิคมาใช้ เช่น การหาสมาชิกที่ใกล้ที่สุด (nearest neighbor) การจัดกลุ่มข้อมูล (clustering) โครงข่ายประสาทเทียม (artificial neural networks) โดยสามารถแบ่งเทคนิคเป็น 2 กลุ่ม ได้แก่ 1) Memory based ประกอบด้วย User-based และ Item-based ทั้ง 2 เทคนิคนี้ในกลุ่มนี้ใช้การหาค่าความคล้ายของข้อมูล (Similarity Measure) เหมือนกัน แตกต่างกันเพียงแค่ใช้ข้อมูล User หรือ item

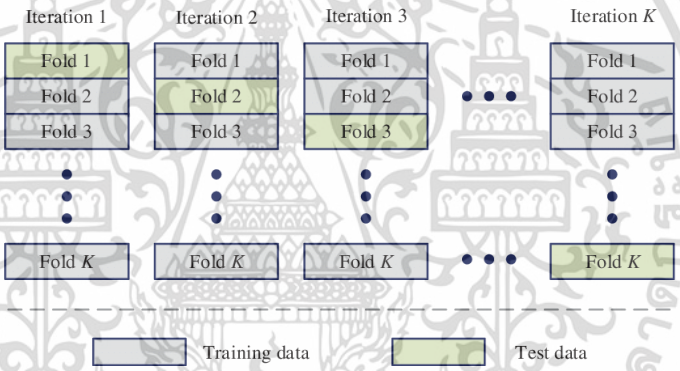
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการคำนวณอย่างใดอย่างหนึ่งเท่านั้น และ 2) Model based เช่น Clustering techniques, Neural Network เป็นต้น นอกจาก 2 เทคนิคที่ผู้วิจัยกล่าวข้างต้นแล้ว ยังมีเทคนิคอื่นๆ ที่ได้รับความนิยมในการพัฒนาระบบแนะนำ ได้แก่ Matrix Factorization, Content-based, Association Rules, Deep Learning เช่น Auto Encoder, Restricted Boltzmann Machine (RBM) เป็นต้น

2.3.3 การประเมินผลระบบแนะนำ (Recommendation System Evaluation)

ในการพัฒนาระบบแนะนำจำเป็นต้องมีการประเมินผลระบบ เพื่อให้มั่นใจว่ารายการที่แนะนำมีความถูกต้อง แม่นยำ และสามารถแนะนำวัตถุไปยังผู้ใช้งานได้อย่างเหมาะสม วิธีการที่ได้รับความนิยมในการประเมินผลระบบแนะนำได้แก่

2.3.3.1 Train/Test and Cross Validation เป็นการแบ่งข้อมูลออกเป็น 2 ชุดข้อมูล ได้แก่ ชุดข้อมูลฝึกสอน (Training data) และ ชุดข้อมูลทดสอบ (Test data) และใช้เทคนิค Cross Validation ที่มีการแบ่งข้อมูลออกเป็น k กลุ่ม หรือ เรียกว่า fold ในการเลือกชุดข้อมูลแต่ละกลุ่มเป็นชุดทดสอบ สำหรับชุดข้อมูลที่เหลือจะเป็นชุดฝึกสอน และมีการวนซ้ำเลือกใช้ครบทุกกลุ่มข้อมูลวัดประสิทธิภาพของแบบจำลอง ดังรูปที่ 2.4



รูปที่ 2.4 การทำงานของการวัดประสิทธิภาพแบบจำลองด้วยวิธี Cross Validation

2.3.3.2 Accuracy Metrics (RMSE, MAE) เป็นการวัดผลการทำนายค่าคะแนน (Predicted Rating) ที่เป็นผลลัพธ์ของกับแบบจำลองเปรียบเทียบกับค่าคะแนนจริง (Actual Rating) ดังสูตร 2.1

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(predicted_i - actual_i)^2}{N}} \qquad (2.1)$$

2.3.3.3 Precision@K, MAP@K, NDCG@K สำหรับตัววัดประเมินผลระบบแนะนำในกลุ่มนี้เป็นการวัดผลจากรายการแนะนำจากระบบ โดยเปรียบเทียบวัตถุ (items) ในรายการแนะนำกับรายการวัตถุที่ควรอยู่ในรายการแนะนำจริง ซึ่งจะคำนึงถึงอันดับด้วยหรือไม่ขึ้นอยู่กับวัตถุประสงค์ของตัววัดนั้น ๆ แสดงรายละเอียด ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Precision@K เป็นการประเมินสัดส่วนของรายการวัตถุที่อยู่ในรายการแนะนำจำนวน K รายการตามที่กำหนด โดยจะเป็นการตรวจสอบว่ารายการวัตถุที่เราสนใจอยู่ในรายการแนะนำจำนวนเท่าใด

$$P@K = \frac{1}{\min(n, K)} \sum_{k=1}^{K} rel(k) \tag{2.2}$$

2. MAP@K หรือ Mean Average Precision at K เป็นตัววัดจากการต่อยอดตัววัด Average Precision (AP) ของแต่ละรายการแนะนำตามจำนวน K รายการตามที่กำหนด โดยทำการคำนวณหาค่าเฉลี่ยจากค่า AP ของทุกรายการแนะนำ [12]

$$MAP@K = \frac{1}{|U|} \sum_{u=1}^{I} U \frac{1}{\min(n, K)} \sum_{k=1}^{K} P_u(k) \cdot rel_u(k) \tag{2.3}$$

3. NDCG หรือ Normalized Discounted Cumulative Gain เป็นตัววัดประสิทธิภาพของผลลัพธ์รายการแนะนำแบบเรียงลำดับ โดยใช้คะแนนค่าชื่นชอบ (Rating) เป็นเกณฑ์ให้คะแนนกับรายการที่เกี่ยวข้องและให้ความสำคัญกับรายการวัตถุที่มีคะแนนค่าชื่นชอบมากอยู่ในลำดับต้น ๆ [12] ค่าคะแนน NDCG ยิ่งมีค่ามากหมายถึงเรียงลำดับของรายการวัตถุที่มีค่าชื่นชอบมากอยู่ในลำดับต้น

$$NDCG@K = \frac{1}{\sum_{i=1}^{rel@K} \frac{2^{rel_i - 1}}{log_2(i+1)}} \sum_{i=1}^{K} \frac{2^{rel_i - 1}}{log_2(i+1)} \tag{2.4}$$

2.4 ขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence)

ขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) นำเสนอโดย Gerardo Beni และ Jing Wang ในปี ค.ศ. 1989 เป็นศาสตร์แขนงหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) และเป็นศาสตร์แขนงในสาขา Biologically-inspired computing โดยเป็นการศึกษาและเลียนแบบพฤติกรรมของสิ่งมีชีวิตในธรรมชาติที่มีการทำงานแบบกลุ่ม ประกอบด้วยตัวแทนหรือเอเจนต์ ซึ่งตัวแทนภายในกลุ่มจะมีปฏิสัมพันธ์สามารถสื่อสารตอบโต้กับตัวแทนอื่นหรือสภาพแวดล้อมอื่นเพื่อช่วยกันแก้ไขปัญหา โดยสามารถนำไปสู่คำตอบที่เหมาะสมที่สุดหรือวัตถุประสงค์แบบวงกว้าง (Global) ได้อย่างมีประสิทธิภาพ [13] ระบบภายในของขั้นตอนวิธีความฉลาดแบบกลุ่มมีพฤติกรรม 2 รูปแบบคือ พฤติกรรมส่วนตัวของแต่ละตัวแทน ซึ่งมีลักษณะที่ไม่ซับซ้อน ไม่มีโครงสร้างหรือหน้าที่ที่ได้รับการกำหนดจากส่วนกลางอย่างชัดเจน จึงสามารถจัดการตัวเอง (Self-organize) และปรับปรุงการทำงานของตนเองเพื่อนำไปสู่วัตถุประสงค์แบบท้องถิ่น (Local) และพฤติกรรมโดยรวมของกลุ่ม ซึ่งเป็นตัวกำหนดเงื่อนไขที่ตัวแทนแต่ละตัวจะต้องกระทำต่อไป การทำงานในรูปแบบนี้จึงทำให้เกิดพฤติกรรมของระบบที่ซับซ้อนขึ้นได้ ตัวอย่างของกลุ่มสิ่งมีชีวิต อาทิตเช่น การบินเป็นฝูงของนกหรือการว่ายน้ำเป็นฝูงของปลา การทำงานในอาณาเขตและการหาอาหารของมด เป็นต้น ข้อดีของขั้นตอนวิธีความฉลาดแบบกลุ่มนั้นคือ ขั้นตอนไม่ซับซ้อนแต่สามารถทำงานได้อย่าง

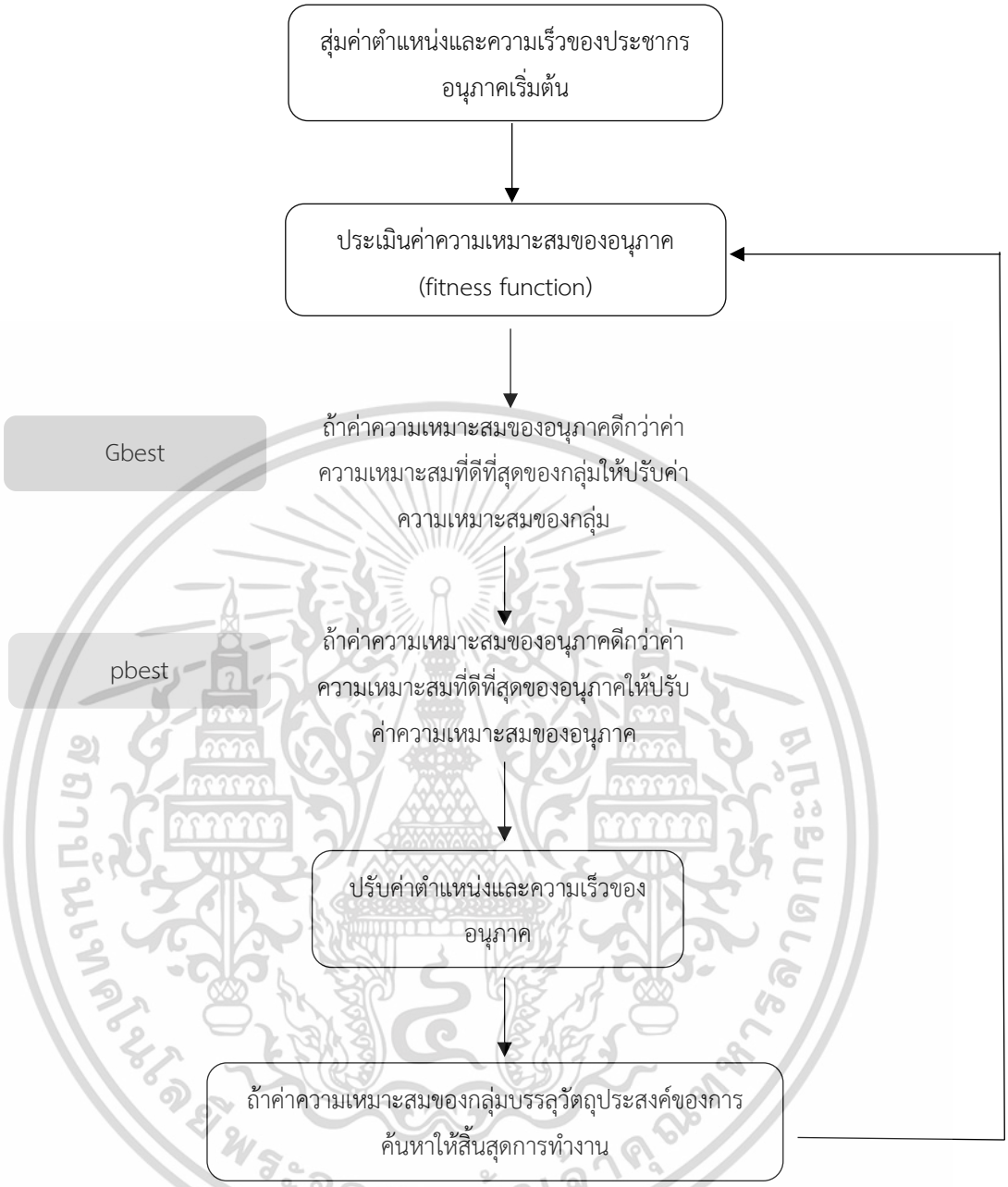
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีประสิทธิภาพ สามารถนำไปใช้ในปัญหาการหาค่าเหมาะสมที่สุดหลาย ๆ อย่างได้ตัวอย่างของความฉลาดแบบกลุ่มที่มาจากธรรมชาติได้แก่ ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization), ขั้นตอนวิธีหาค่าเหมาะสมที่สุดด้วยระบบอาณาจักรมด (Ant Colony Optimization), ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบค้างคาว (Bat Algorithm), ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo Search), ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบหิ่งห้อย (Firefly Algorithm), ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบดอกไม้ไฟ (Fireworks Algorithm) เป็นต้น ซึ่งขั้นตอนวิธีข้างต้นสามารถนำมาประยุกต์ใช้ในการพัฒนาระบบแนะนำ (Recommendation System) เพื่อให้มีประสิทธิภาพมากขึ้นได้ การประยุกต์ใช้งานของ Swarm Intelligence เช่นการคัดเลือกคุณลักษณะ (Feature Selection) [14]

2.4.1 ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization)

ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization : PSO) เป็นวิธีการที่ได้รับแรงบันดาลใจจากการเคลื่อนที่เป็นฝูงของปลาและนก นำเสนอโดย James Kennedy และ Russell Eberhart ในปี ค.ศ. 1995 PSO มีคุณลักษณะที่คล้ายกันกับขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA) โดยมีการสุ่มประชากรเริ่มต้นใช้ในการหาคำตอบที่เหมาะสมที่สุด และทำการปรับค่าของประชากรแต่ละตัวในทุกรอบการคำนวณเหมือนกัน สิ่งที่ทำให้ PSO และ GA ต่างกันคือ PSO ไม่มีการกลายพันธุ์หรือการปรับเปลี่ยนและสร้างข้อมูลใหม่ เช่น การทำครอสโอเวอร์หรือมิวเทชัน เช่นเดียวกันกับ GA แต่คำตอบของ PSO จะถูกแทนด้วยอนุภาค (particle) ซึ่งจะเคลื่อนที่ไปในปริภูมิค้นหา (Search Space) โดยมีทิศทางไปตามอนุภาคที่ใกล้เคียงกับคำตอบที่เหมาะที่สุดในช่วงเวลานั้น การเคลื่อนที่ของกลุ่มประกอบไปด้วย การเคลื่อนที่ของกลุ่มอนุภาคโดยที่แต่ละอนุภาคเป็นตัวแทนของคำตอบของระบบที่ต้องการค้นหา ตำแหน่งของแต่ละอนุภาคจะเปลี่ยนไปตามข้อมูลของสภาะแวดล้อมที่เคลื่อนที่ผ่านมาและข้อมูลจากอนุภาคภายในกลุ่มด้วยกันเอง ตัวอย่างการประยุกต์ใช้ PSO เช่น การหาค่าเหมาะที่สุดของฟังก์ชัน การฝึกสอนเครือข่ายประสาทเทียม เป็นต้น [13] รายละเอียดขั้นตอนของอัลกอริทึม PSO สามรถสรุปได้ดังรูปที่ 2.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5 ขั้นตอนการทำงานของขั้นตอนวิธี Particle Swarm Optimization

รายละเอียดขั้นตอนการทำงานของ Particle Swarm Optimization : PSO

กำหนดให้ $\overrightarrow{p_i}(t)$ แทนเวกเตอร์ตำแหน่ง (position vector) ของอนุภาค $\mathrm{p_i}$ ณ เวลา t

1. สุ่มค่าตำแหน่งและความเร็วของประชากรอนุภาคเริ่มต้น : ทำการสุ่มค่าเวกเตอร์ตำแหน่งและความเร็วของแต่ละอนุภาคในกลุ่มประชากร เวกเตอร์ตำแหน่งของอนุภาคมีมิติเท่ากับ N ซึ่งเป็นขนาดของตัวแปรในปัญหาที่ต้องการค้นหา ดังนั้นเราสามารถมองตำแหน่งของนุภาคในรูปเวกเตอร์ $\overrightarrow{p_i}(t)$ ขนาด N ได้ และสุ่มความเร็วของแต่ละอนุภาคจะมีขนาดเท่ากับ N ด้วย โดยเรียกว่าเวกเตอร์ความเร็ว (Velocity vector) กำหนดให้ $\overrightarrow{v_i}(t)$ แทนเวกเตอร์ความเร็วดังกล่าว ณ เวลา t ซึ่งแต่ละองค์ประกอบของเวกเตอร์ความเร็วจะเป็นค่าความเร็วของแต่ละตัวแปรในอนุภาคนั่นเอง ดังนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตำแหน่งของอนุภาค p_i จะเปลี่ยนแปลงไปด้วยการบวดเวกเตอร์ตำแหน่งเข้ากับเวกเตอร์ความเร็วดังสมการ 2.6

$$\overrightarrow{p_i}(t) = \overrightarrow{p_i}(t-1) + \overrightarrow{v_i}(t) \quad (2.6)$$

2. ประเมินค่าความเหมาะสมของอนุภาค (fitness function) : ทำการประเมินค่าความเหมาะสม $fitness$ ของแต่ละอนุภาค ซึ่งการประเมินค่าดังกล่าวจะขึ้นอยู่กับแต่ละปัญหา ค่าความเหมาะสมที่ได้จากการประเมินจะถูกพิจารณาในสองขั้นตอน ดังนี้

ถ้าค่าความเหมาะสมของอนุภาค p_i มีค่าดีกว่าค่าความเหมาะสมที่ดีทุ่ดของทั้งกลุ่มอนุภาค ให้ทำการบันทึกเวกเตอร์ตำแหน่งของอนุภาคนี้ไว้ โดยเรียกว่าค่าความเหมาะสมที่ดีที่สุดแบบวงกว้าง (global best fitness) หรือ $gbest$ กล่าวคือถ้า $fitness(\vec{p}_i(t)) < gbest$ (กรณีหาค่าที่น้อยที่สุด) ให้ทำการบันทึกค่าความเหมาะสมของระบบและค่าเวกเตอร์ตำแหน่งของอนุภาคนั้น ๆ ไว้ดังสมการ 2.7 และ 2.8

$$gbest_i = fitness(\vec{p}_i(t)) \quad (2.7)$$
$$\vec{p}_{gbest} = \vec{p}_i(t) \quad (2.8)$$

ถ้าค่าความเหมาะสมของอนุภาค มีค่าดีกว่าค่าความเหมาะสมที่ดีที่สุดของอนุภาคนั้น ๆ ซึ่งเรียกว่า $pbest_i$ (p มาจาก personal) กล่าวคือ $fitness(\vec{p}_i(t)) < pbest_i$ ให้ทำการบันทึกค่าความเหมาะสมและเวกเตอร์ตำแหน่งของอนุภาคไว้ใน ดังสมการ 2.9 และ 2.10

$$pbest_i = fitness(\vec{p}_i(t)) \quad (2.9)$$
$$\vec{p}_{pbest_i} = \vec{p}_i(t) \quad (2.10)$$

3. ปรับค่าตำแหน่งและความเร็วของอนุภาค : ทำการปรับค่าความเร็วของอนุภาค p_i ดัง 2.11

$$\overrightarrow{v_i}(t) = \overrightarrow{v_i}(t-1) + \rho_p\left[\vec{p}_{pbest_i} - \vec{p}_i(t)\right] + \rho_g\left[\vec{p}_{gbest} - \vec{p}_i(t)\right] \quad (2.11)$$

โดยที่ ρ_p และ ρ_g เป็นตัวแปรสุ่ม เทอมที่สองของสมการข้างต้นเรียกว่า องค์ประกอบเชิงปริชาน (cognitive component) และเทอมสุดท้ายเรียกว่า องค์ประกอบทางสังคม (social component)

4. ทำการปรับค่าเวกเตอร์ตำแหน่งของอนุภาค p_i ดังสมการ 2.12

$$\overrightarrow{p_i}(t) = \overrightarrow{p_i}(t-1) + \overrightarrow{v_i}(t) \quad (2.12)$$

5. ทำการปรับค่าตัวแปรเวลา $t = t + 1$ และดำเนินขั้นตอนทั้งหมดกับอนุภาคถัดไปในกลุ่มประชากรจนครบทุกอนุภาค

6. วนรอบการทำงานทั้งหมดจนกระทั่งมีการลู่เข้าคำตอบของอนุภาคที่ดีที่สุดในกลุ่ม หรือตามเงื่อนไขหยุดการทำงานที่กำหนดไว้

โดยปกติทำการกำหนดจำนวนรอบในการค้นหาไว้หรือกำหนดการค้นหาสามารถสิ้นสุดลงได้เมื่อค่า gbest บรรลุวัตถุประสงค์ของการค้นหา แต่ละองค์ประกอบของเวกเตอร์อนุภาคที่ได้คือตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

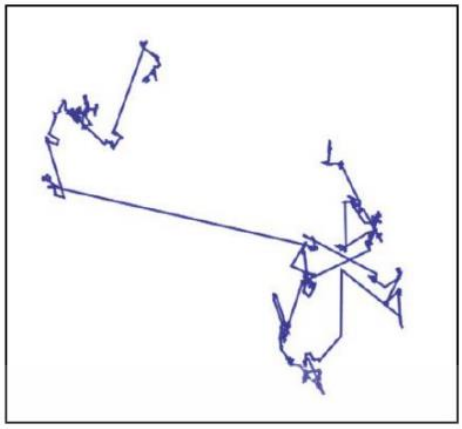
แปรของปัญหาที่ต้องค้นหานั่นเอง องค์ประกอบเชิงปริชาน (cognitive component) แทนประสบการณ์ค้นหาคำตอบของแต่ละอนุภาค ในขณะที่องค์ประกอบเชิงสังคม (social component) แทนประสบการณ์ค้นหาของกลุ่มอนุภาค ถ้าหากไม่พิจารณาใช้องค์ประกอบเชิงสังคมในการปรับค่าเวกเตอร์ความเร็ว ขั้นตอนวิธี PSO นั้นจะเรียกว่าเป็นแบบเฉพาะที่ (local best) ค่าของ ρ_p เป็นพารามิเตอร์ที่มีค่าสุ่มและกำหนดขนาดโดยผู้ใช้ พารามิเตอร์ดังกล่าวใช้สำหรับขจำกัดขนาดเส้นทางเดินของอนุภาคออกจากตำแหน่งที่เหมาะสมที่สุด ค่า ρ_p ที่มาก จะทำให้เส้นทางเดินของอนุภาคมีการกวัดแกว่ง (oscillate) ในขณะที่ค่า ρ_p มีค่าน้อยจะทำให้อนุภาคเคลื่อนที่อย่างเรียบ (smooth) ในทำนองเดียวกัน ค่า ρ_g เป็นพารามิเตอร์ที่จำกัดเส้นทางเดินของกลุ่ม ทั้งค่า ρ_p และ ρ_g กำหนดได้เป็น $\rho_p = r_p c_p$ และ $\rho_g = r_g c_g$ โดยที่ $r_p, r_g \in U(0,1)$ เป็นค่าสุ่มการแจกแจงแบบสม่ำเสมอระหว่าง ถึง และ c_p และ c_g เป้นค่าคงที่ความเร่ง (acceleration constant) ซึ่งมีค่าเป็นบวกทั้งคู่ จากการศึกษาของ Kennedy ค่าของ ρ_p และ ρ_g ในเส้นทางการเคลื่อนที่ของอนุภาคและได้เสนอผลไว้ว่า $c_p + c_g \leq 4$ (Kennedy, 1998) ถ้า $c_p + c_g > 4$ จะส่งผลให้ตำแหน่งและความเร็วมีค่าเป็นอนันต์ [13]

2.4.2 ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search)

ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) เป็นหนึ่งในขั้นตอนวิธี nature-inspired metaheuristic ได้รับการพัฒนาขึ้นจาก Xin-She Yang จากมหาวิทยาลัยแคมบริดจ์ และ Suash Deb จากวิทยาลัยวิศวกรรมรามานในปี 2009 ซึ่งได้รับแรงบันดาลใจจากพฤติกรรมการวางไข่ของนกดุเหว่าที่จะวางไข่ในรังของนกสายพันธุ์อื่น โดยนกดุเหว่าจะไม่มีการสร้างรังของตัวเองและอาศัยในรังสายพันธุ์อื่นเพื่อเจ้าของรังฟักและเลี้ยงดูลูกนกของตนเองจนเติบโต นกดุเหว่าบางชนิดมีไหวพริบในการใช้รังร่วมกับนกชนิดอื่น เพื่อสร้างสมาชิกใหม่ขึ้นมา ซึ่งพฤติกรรมของนกดุเหว่าเช่นนี้จะเรียกว่า “การเบียนฟังไข่” แม่นกดุเหว่ามีความสามารถในการเลียนแบบไข่ของนกเจ้าของรังอย่างแนบเนียน ซึ่งถือว่าเป็นความมหัศจรรย์ของธรรมชาติ หากไม่ถูกโยนไข่ทิ้งและนกสายพันธุ์นั้นทิ้งรังแล้วไปสร้างรังใหม่แทน ลูกนกดุเหว่าจะฟักออกมาและมีสัญชาตญาณในการขับไล่ไข่ของรังนกนั้น ถือเป็นการขยายถิ่นอาศัยและเพิ่มการหาอาหารของนกดุเหว่า จุดเด่นของขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่าคือการค้นหาคำตอบที่รวดเร็ว จึงได้รับความนิยมนำมาใช้แก้ปัญหา Optimization เพื่อหา global optimum ที่สามารถหาผลลัพธ์เป็นคำตอบที่ดีที่สุด ขั้นตอนวิธีแบบนกดุเหว่าใช้การปรับปรุงวิธีการทำงานด้วย Lévy-flights และจากการศึกษาพบว่ามีประสิทธิภาพสูงกว่าขั้นตอนขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm) และขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) การทำงานของการค้นหาแบบนกดุเหว่า (Cukcoo Search) มีแนวคิดพื้นฐาน 3 ข้อ [15] คือ 1) นกดุเหว่าทำการวางไข่ครั้งละ 1 ฟอง และทิ้งไข่ของตนเองไว้ในรังของนกสายพันธุ์อื่นที่เลือกสุ่มไว้ 2) รังไข่ที่ไม่ถูกนกสายพันธุ์อื่นโยนทิ้ง/ทิ้งรังและปล่อยให้ไข่ลูกนกดุเหว่าที่มีคุณภาพสูงฟักตัว จะเป็นการสร้างนกดุเหว่ารุ่นถัดไป และ 3) กำหนดจำนวนรังของนกสายพันธุ์อื่นที่ใช้ได้ (n) และโอกาสที่ไข่ของลูกนกดุเหว่าถูกค้นพบโดยนกสายพันธุ์อื่นให้มีค่าความน่าจะเป็นระหว่าง 0 - 1

ในการศึกษาพบว่า การบินของสัตว์หลาย ๆ สายพันธุ์มีลักษณะที่คล้ายคลึงกัน รวมถึงการบินของนกดุเหว่าเพื่อค้นหารังของนกสายพันธุ์ในการวางไข่ของลูกตนเองนั้น [16] สามารถอธิบายได้ในรูปแบบ Lévy-flights ดังรูปที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 การเคลื่อนที่ของนกในรูปแบบ Lévy-flights

วิธีการสร้างแนวทางการแก้ปัญหาใหม่ของการค้นหาคำตอบของปัญหา ดังสมการ 2.13

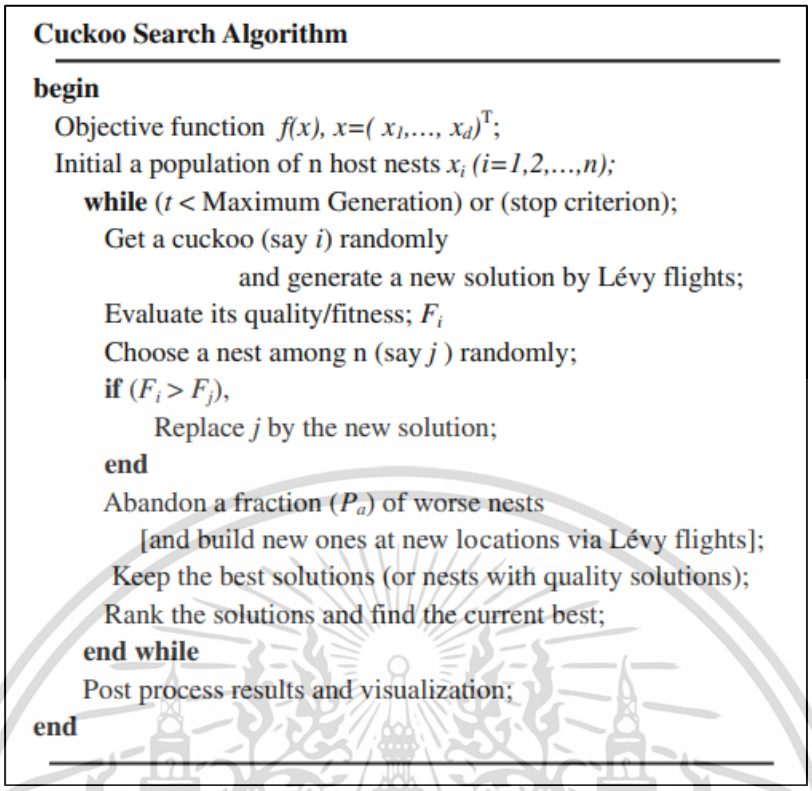
$$x_i^{(t+1)} = x_i^t + \alpha \oplus \mathbf{L\acute{e}vy}(\lambda) \tag{2.13}$$

โดย α คือขนาดขั้นตอนที่เกี่ยวข้องกับมาตราส่วนของปัญหาที่กำลังสนใจแก้ไข สมการถัดไปเป็นสมการสำหรับการสุ่มทางเดิน ซึ่งโดยปกติจะเป็นการสุ่มทางเดินแบบ Markove Chain ที่มีสถานะหรือตำแหน่งถัดไปขึ้นกับตำแหน่งปัจจุบันและความน่าจะเป็นของการเปลี่ยนแปลง เครื่องหมาย $\oplus$ ในสมการ คือ Entrywise multiplications ซึ่งจะมีลักษณะคล้ายกับการทำงานที่ถูกใช้ในขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) แต่มีประสิทธิภาพในการสำรวจพื้นที่มากขึ้นสำหรับการสุ่มการเดินของ Lévy-flights [16] ดังสมการ 2.14

$$\mathbf{L\acute{e}vy}(\lambda) \backsim \mathcal{U} = \mathrm{t} - \lambda, 1 < \lambda \leq 3 \tag{2.14}$$

โดยในการสุ่มการเคลื่อนที่ของ Lévy-flights จะมีการนำ Mantegna's Algorithm มาช่วยในการสร้างตัวเลขสุ่ม เนื่องจากขั้นตอนวิธีดังกล่าวสามารถสร้างตัวเลขตามการกระจายที่มีเสถียรภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Cuckoo Search Algorithm

begin
Objective function $f(x)$, $x=(x_1,\ldots,x_d)^T$;
Initial a population of n host nests x_i $(i=1,2,\ldots,n)$;
while (t < Maximum Generation) or (stop criterion);
Get a cuckoo (say i) randomly
and generate a new solution by Lévy flights;
Evaluate its quality/fitness; F_i
Choose a nest among n (say j) randomly;
if ($F_i > F_j$),
Replace j by the new solution;
end
Abandon a fraction (P_a) of worse nests
[and build new ones at new locations via Lévy flights];
Keep the best solutions (or nests with quality solutions);
Rank the solutions and find the current best;
end while
Post process results and visualization;
end

รูปที่ 2.7 Peudo code การทำงานของขั้นตอนวิธีแบบนกดุเหว่า (Cukcoo Search)

จากรูปที่ 2.7 แสดงการทำงานของขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) สามารถสรุปขั้นตอนได้ดังนี้ [15]

1. กำหนดฟังก์ชันวัตถุประสงค์ (Objective Function) $f(x), x = (x_1, x_2, \ldots, x_d)^T$
2. สร้างประชากรเริ่มต้นเป็นจำนวนรังนก (n) โดย $x_i = (i = 1,2, \ldots, n)$
3. สร้างไข่ลูกนกดุเหว่าโดยการสุ่มการเคลื่อนที่ด้วย Lévy-flights
4. ประเมินคุณภาพของรังจากฟังก์ชันวัตถุประสงค์
5. เลือกรังใหม่จากจำนวนรังทั้งหมด (n) โดยการสุ่มและประเมินคุณภาพของรังด้วยฟังก์ชันวัตถุประสงค์
6. เปรียบเทียบคุณภาพของรังปัจจุบันและรังใหม่จากการประเมินคุณภาพในขั้นตอน (5) หากคุณภาพของรังใหม่มีผลที่ดีกว่า ให้ทำการแทนที่รังใหม่เป็นคำตอบใหม่
7. ในการค้นพบรังใหม่นั้นจะต้องมีค่าความน่าจะเป็นมากว่าค่าที่กำหนดไว้เท่านั้น หากไม่เป็นตามเงื่อนไขให้ทิ้งรังดังกล่าวและสุ่มเลือกรังใหม่ด้วย Lévy-flights ตามข้อ (3)
8. จัดเก็บรังที่มีคุณภาพหรือมีคำตอบที่ดีที่สุด
9. เรียงลำดับของผลลัพธ์ที่ได้ และตรวจสอบการสิ้นสุดการทำงาน โดยตรวจสอบจากจำนวนรอบที่กำหนด หรือหยุดตามเงื่อนไขที่กำหนดไว้ หากไม่ใช่ให้ดำเนินการขั้นตอนที่ 4 – 9 หากใช่ให้ถือว่าสิ้นสุดการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 การโปรแกรมเชิงเส้น (Linear Programming)

การโปรแกรมเชิงเส้น เป็นฟังก์ชันเชิงเส้นทางคณิตศาสตร์ที่สามารถการแก้ปัญหาการจัดสรรทรัพยากรที่มีอยู่อย่างจำกัด โดยการหาคำตอบที่ดีที่สุดในปัญหาที่ดำเนินการจะเป็นตัวแลขแสดงจำนวนของแต่ละตัวแปรที่สนใจและอยู่ภายในขอบเขต ยกตัวอย่างเช่น กำไรสูงที่สุด หรือต้นทุนต่ำที่สุด เป็นต้น ซึ่งมีองค์ประกอบและวิธีการกำหนดรูปแบบสำหรับการแก้ไขปัญหาของการโปรแกรมเชิงเส้น [17] ดังนี้

1. กำหนดตัวแปรที่ต้องตัดสินใจ (Decision Variable) เป็นตัวแปรที่ต้องพิจารณาเพื่อหาคำตอบที่ดีที่สุด

2. ฟังก์ชันวัตถุประสงค์ของปัญหา (Objective Function) เป็นฟังก์ชันที่บอกเป้าหมายว่าต้องการค่าสูงสุด (Maximum) หรือต่ำสุด (Minimum)

3. เงื่อนไขแสดงขอบเขต (Constraints) เป็นข้อจำกัดของคำตอบหรือค่าที่ตัวแปรตัดสินใจจะเป็นได้

4. บริเวณที่เป็นไปได้ของคำตอบ (Feasible Region) เป็นกลุ่มของคำตอบที่อยู่ภายในเงื่อนไขขอบเขตทั้งหมดของปัญหาที่จะแก้ไข ซึ่งจะพิจารณาค่าทุกค่าที่ค่าบวกและศูนย์เท่านั้น

2.6 งานวิจัยที่เกี่ยวข้อง

ในการพัฒนาระบบแนะนำ (Recommendation System) มีการพัฒนาจากหลายเทคนิคตามวัตถุประสงค์ของธุรกิจและลักษณะของชุดข้อมูลที่นำมาใช้ ผู้วิจัยได้แบ่งกลุ่มงานวิจัยเป็น 2 กลุ่ม ได้แก่ 1) กลุ่มการพัฒนาระบบแนะนำรายการอาหาร 2) กลุ่มการปรับปรุงประสิทธิภาพระบบแนะนำโดยการประยุกต์ใช้เทคนิคทาง Swarm Intelligence ตัวอย่างงานวิจัยแสดงรายละเอียด ดังนี้

2.6.1 กลุ่มของการพัฒนาระบบแนะนำและเทคนิคระบบแนะนำรายการอาหาร

การศึกษาและพัฒนาระบบแนะนำมีขั้นตอนการดำเนินงานและเทคนิคที่หลากหลาย กว้างขวางขึ้นอยู่กับวัตถุประสงค์ของการนำรายการแนะนำไปใช้กับกลุ่มบุคคลใด เช่น บุคคลทั่วไป กลุ่มเด็กเล็ก กลุ่มควบคุมอาหาร กลุ่มกิจกรรมและกีฬา กลุ่มผู้ป่วย เป็นต้น รวมถึงความต้องการในการแนะนำเพื่อปรับเปลี่ยนพฤติกรรมการบริโภคอาหารและสนับสนุนด้านการดูแลสุขภาพ โดยมีงานวิจัยของ M. Sedlak (2021) [18] เรื่อง Content-based recommender System for Food Recipes ศึกษาวิธีการค้นหาพฤติกรรมการรับประทานอาหารของผู้ใช้งานด้วย similarity metric กับชุดข้อมูลการบันทึกประวัติการบริโภคอาหารเพื่อแนะนำรายการอาหารที่ผู้ใช้งานไม่เคยรู้จักหรือบริโภคมาก่อน C. N. Sanchez และคณะ (2023) [19] ทำการศึกษาพฤติกรรมการสั่งรายการอาหารบนแอปพลิเคชันของผู้ใช้งานและพัฒนาการแนะนำรายการของผู้ใช้งานที่มีพฤติกรรมคล้ายคลึงกันด้วยเทคนิคเพื่อนบ้านที่ใกล้เคียงที่สุด (K-Nearest Neighbor) Abhari S., Safdari R. (2019) [1] และคณะได้ศึกษางานวิจัยการพัฒนาเทคนิคระบบแนะนำอาหารที่มีคุณค่าทางโภชนาการ พบว่า Hybrid recommender system ได้รับความนิยมในการพัฒนามากกว่า 40% ตามด้วย Knowledge-based, rule-based และ ontology ซึ่งในการพัฒนาระบบแนะนำประเภทนี้จำเป็นต้องมีผู้เชี่ยวชาญเฉพาะด้านเข้าร่วมการออกแบบและพัฒนา งานวิจัยเรื่อง Healthy Food Recommendation Using a Time Aware Community Detection Approach and Reliability

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Measurement ของ S. Ahmadian และคณะ (2022) [20] เสนอระบบแนะนำอาหารโดยอิงจากการวัดความน่าเชื่อถือด้านสุขภาพและพัฒนาวิธีการ time-aware community detection ด้วยการค้นหาเพื่อนบ้านที่ใกล้ที่สุดและคาดคะเนความชื่นชอบรายการอาหาร ซึ่งผลการทดลองพบว่ามีค่า NDCG เท่ากับ 16.17 งานวิจัยของ T. Mokdara [2] และคณะนำเสนอแบบจำลอง deep neural network กับชุดข้อมูลการรับประทานอาหารไทย ซึ่งจะดำเนินการแยกส่วนประกอบของรายการอาหารแต่ละจานและสร้างแบบจำลองหาความสัมพันธ์ คาดคะเนและแนะนำรายการอาหารไทยไปยังผู้ใช้งาน ผลการทดลองมีความแม่นยำ 90% งานวิจัยเรื่อง What-To-Taste: A Food Recommendation System ของ K. Ahuja, M. Goel (2020) [21] และคณะได้นำเสนอระบบแนะนำรูปแบบเว็บแอปพลิเคชันที่พัฒนาด้วยข้อมูลส่วนบุคคล รสนิยม ความสนใจ และประวัติการรับประทานอาหารของผู้บริโภคด้วยเทคนิค Content Based และ Collaborative Filtering ซึ่งทางผู้วิจัยได้เสนอแนะการเพิ่มปัจจัยด้านโภชนาการในการแนะนำรายการอาหารไปยังผู้ใช้งานเพื่อให้มีประสิทธิภาพและส่งผลดีต่อสุขภาพมากยิ่งขึ้น R. Y. Toledo, A. A. Alzahrani และ L. Martinez (2019) [22] ได้ศึกษาระบบแนะนำอาหารพบว่าระบบทั่วไปมีวิธีการดำเนินงานที่หลากหลาย เช่น preference-aware, semantic-based หรือ optimization based กับพฤติกรรมการรับประทานอาหารแต่ยังไม่สนใจโภชนาการอาหารมากนัก จึงศึกษากรอบการพัฒนาระบบแนะนำอาหารโดยสร้างแผนการบริโภคอาหารส่วนบุคคลภายใต้คุณค่าโภชนาการอาหารที่จำเป็นและความชื่นชอบอาหารในอดีตของผู้ใช้งาน ได้แก่ โปรตีน โซเดียม คลอเรสเตอรอล และไขมัน เป็นต้น ทั้งนี้มีงานวิจัยจำนวนไม่น้อยที่ดำเนินการศึกษาข้อมูลโภชนาการอาหารเพื่อให้ผลการแนะนำมีประสิทธิภาพและตอบโจทย์การใช้งานมากยิ่งขึ้น ได้แก่ M. Showafah และคณะ (2021) [23] ศึกษาและพัฒนาระบบแนะนำอาหารเสริมสำหรับคุณแม่และทารกแรกเกิด 6 -24 เดือน ด้วยการพิจารณาโภชนาการอาหาร ได้แก่ คาร์โบไฮเดรต โปรตีนและไขมัน ด้วยวิธีการ ontology Knowledge-based สำหรับโภชนาการอาหาร Naïve Bays สำหรับความสนใจหรือพฤติกรรมของผู้ใช้งาน และ TOPSIS สำหรับการแนะนำรายการอาหารที่เหมาะสมกับโภชนาการอาหารและพฤติกรรมของผู้ใช้งาน ซึ่งผลลัพธ์อยู่ในระดับที่น่าพึงพอใจ การแนะนำชุดรายการอาหารไทย P. Lertkrai และคณะ (2018) [4] ได้ศึกษาเรื่อง Developing ontology of food and nutrition for Thai preschool ดำเนินการพัฒนา ontology สำหรับข้อมูลอาหารและโภชนาการ เพื่อแนะนำอาหารและพลังงานแก่เด็กในโรงเรียนให้สอดคล้องกับระดับค่า BMI รายบุคคล M. N. Jasim และ A. B. Hamid (2022) [24] ได้ศึกษาเรื่อง Food recommendation system based on nutritional needs of human beingsand user preferences ซึ่งพัฒนาระบบแนะนำโดยอาศัยข้อมูลโภชนาการอาหารผนวกกับความสนใจของผู้ใช้งานด้วยวิธีการจัดกลุ่มข้อมูล K-means ซึ่งสารอาหารประกอบด้วย คาร์โบไฮเดรต แมงกานีส โปรตีน โซเดียม สังกะสี รวมถึงวิตามิน A B6 B12 C E K เป็นต้น การศึกษาของ B. Houthami และ M. Gangappa ได้นำเสนองานวิจัยเรื่อง Nutrition Diet Recommendation System using User's interest (2020) [25] ซึ่งพัฒนาระบบแนะนำอาหารที่ได้รับคุณค่าทางโภชนาการตามการบริโภคของผู้ใช้งานเพื่อการวางแผนการบริโภคอาหารใน 1 วัน โดยใช้ข้อมูลจำนวน 2 ชุด ได้แก่ USDA nutrient dataset เป็นชุดข้อมูลอาหารและโภชนาการ และ Grocery dataset ชุดข้อมูลการบริโภคอาหารตามความสนใจของผู้ใช้งาน พัฒนาแบบจำลองด้วยเทคนิคผสมผสานระหว่าง collaborative filtering กับ content-based filtering และคำนวณร่วมกับค่า BMI เพื่อแนะนำรายการอาหารสำหรับเติมเต็มสารอาหารต่าง ๆ ในแต่ละวันเพื่อให้ได้รับสารอาหารที่ครบถ้วน อาทิเช่น โปรตีน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คาร์โบไฮเดรต ไอมัน แคลเซียม เป็นต้น อีกทั้ง รายการอาหารที่แนะนำครอบคลุมถึงอาหารกินเล่น มื้อกลางวัน และมื้อเย็น

2.6.2 การปรับปรุงประสิทธิภาพของระบบด้วย Swarm Intelligence

Hui Jing (2022) ได้ศึกษาเรื่อง Application of Improved K-Means Algorithm in Collaborative Recommendation System [26] โดยนำเสนอการเพิ่มประสิทธิภาพระบบแนะนำของเทคนิค Collaborative Filtering ด้วยการประยุกต์ Bee Colony Alogrithm ปรับปรุงจุด centroid ของการจัดกลุ่ม K-means จากการทดลองพบว่าเทคนิคที่นำเสนอมีประสิทธิภาพในการแนะนำดีว่าเทคนิค user recommendation-based , user-based clustering และ ICCFRA ด้วยค่า MAE เท่ากับ 0.767 ในขณะที่ A. Al-Asadi และ M. N. Jasim (2022) [27] ได้นำเสนอประยุกต์ใช้ Artificial Fish Swarm Algorithm (AFSA) ในการปรับปรุงจุด centroids ของเทคนิค K-means และพัฒนาระบบแนะนำด้วย user-based collaborative filtering แต่ละกลุ่มข้อมูล งานวิจัยเรื่อง Improving the performance of video collaborative filtering recommender system using optimization algorithm ของ N. Tohidi และ C. Dadkhah (2020) [28] ได้ศึกษา Accelerated Particle Swarm Optimization (APSO) และ Forest Optimization Algorithm (FOA) มาประยุกต์ใช้เพิ่มประสิทธิภาพของการจัดกลุ่ม K-Means เช่นเดียวกัน R. Katarya และ O. P. Verma (2017) [29] ได้ศึกษาเรื่อง An effective collaborative movie recommender system with cuckoo search ซึ่งนำเสนอการปรับปรุงการจัดกลุ่ม K-means ด้วยเทคนิค cuckoo search เพื่อหาค่า centroid ที่เหมาะสมที่สุดของแต่ละกลุ่มข้อมูล และพัฒนาระบบแนะนำภายใต้กลุ่มของผู้ใช้งานนั้น ๆ จากการศึกษาบนชุดข้อมูล MovieLens พบว่าเทคนิคที่นำเสนอมีประสิทธิภาพด้วยค่า MAE ดีกว่าเทคนิคอื่นที่จำนวนกลุ่ม เท่ากับ 64 เช่น K-mean , SOM , GAKM และข้อมูล PCA ในเทคนิคเดียวกัน งานวิจัยของ K. Jabbar (2021) เรื่อง Performance Optimization of Swarm Intelligence-based Clustering Algorithms [30] ได้ศึกษาการประยุกต์ใช้ Cuckoo Search Algorithm (CSA) และ Firefly Algorithm (FA) ในการจัดกลุ่ม K-Means เพื่อปรับปรุงปัญหาของ K-Means ที่ไม่สามารถลด distance cost ในการค้นหาจุด centroids และลดระยะทางของกลุ่มข้อมูลได้ จากการศึกษากับชุดข้อมูล BioColour พบว่าเทคนิค CSA มีข้อดีเรื่องการทำงานและระยะเวลาในการประมวลผลดีกว่า FA งานวิจัยเรื่อง A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city ของ R. Logesh และคณะ (2018) [31] ทำการศึกษาการปรับปรุงประสิทธิภาพของการจัดกลุ่มและระบบแนะนำ user based collaborative filtering โดยนำเสนอ 2 เทคนิค ชื่อว่า QPSO และ QICE ซึ่งเป็นเทคนิค ensemble ที่รวบรวมผลการทำงานของแบบจำลอง 3 ตัวคือ QPSO, K-PSO และ K-MWO และเปรียบเทียบกับเทคนิคอื่น เช่น K-Means , C-Means PSO , MWO และ MSA จากการวิจัยโดยใช้ข้อมูลชุด Yelp และ TripAdvisor และประเมินผลด้วยค่า Hit-Rate, Precision , Recall และ F-Measure พบว่าเทคนิค ensemble QICE มีประสิทธิภาพการทำงานและสามารถแนะนำได้ดีที่สุด โดยมี QPSO, MSA , MWO , PSO รองลงมาตามลำดับ งานวิจัยของ N. Y. Gelvez-Garcia และคณะ (2023) [32] ได้ศึกษากลุ่ม bio-inspired algorithms โดยเฉพาะ particle swarm optimization (PSO) ที่มีการค้นหาคำตอบรวดเร็วกว่าเทคนิคอื่นในกลุ่มเดียวกัน เพื่อปรับปรุงและเพิ่มประสิทธิภาพการทำงานของเทคนิค Collaborative Filtering ซึ่งเป็นการเรียนรู้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

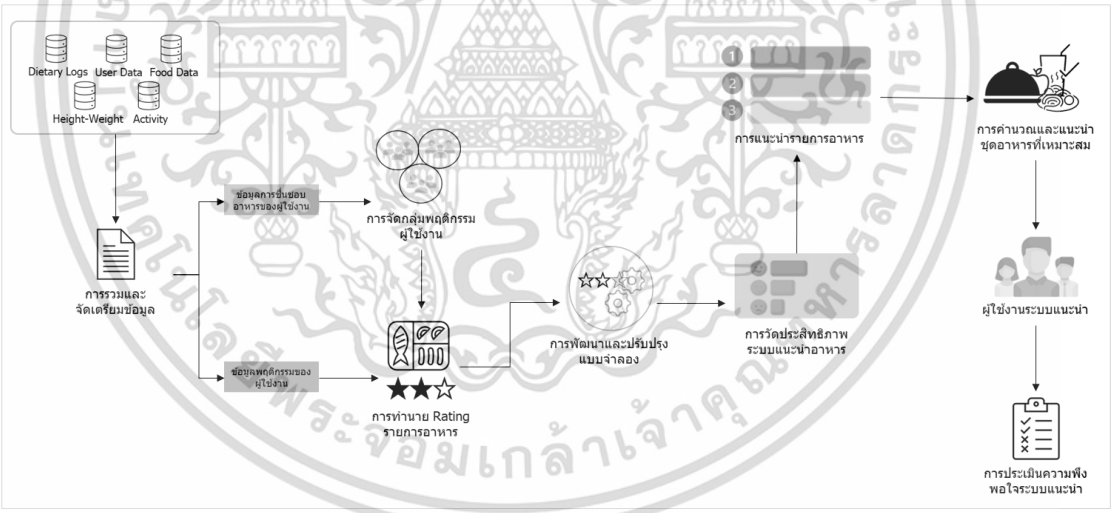
ความชอบส่วนตัวของผู้ใช้งาน ทั้งนี้ได้นำเสนอเทคนิค Cognitive Information Fitering System (CIFS) เพื่อเรียนรู้ feedback ของผู้ใช้งาน และงานวิจัยของ Fanny, L. A. Wulandhari และ S. M. Isa พัฒนาเรื่อง Optimum Nutrition Intake from Daily Dietary Recommendation for Indonesian Children Using Binary Particle Swarm Optimization Algorithm (2019) [33] ทำการวิจัยและพัฒนาแบบจำลองสำหรับการแนะนำอาหารอินโดนีเซียด้วยขั้นตอนวิธี Binary Particle Swarm Optimization เพื่อช่วยแนะนำผู้ใช้งานทั่วไปให้ได้รับสารอาหารที่ครบถ้วนตามคุณค่าทางโภชนาการอย่างเหมาะสมในการรับประทานอาหารแต่ละวันและสอดคล้องกับความต้องการพลังงานโดยรวม ประกอบด้วย เมนูหลัก ผักเครื่องเคียง เนื้อ ผักและผลไม้ ซึ่งพัฒนาบนชุดข้อมูลอาหารอินโดนีเซียที่ชื่อว่า Table Komposisi Pangan Indonesia จำนวนมากกว่า 1600 รายการ และข้อมูลส่วนบุคคล เช่น อายุ น้ำหนัก เพศ และระดับกิจกรรมการอออกกำลังกาย จากการทดลองพบว่า แบบจำลองที่นำเสนอ BPSO มีค่าความถูกต้อง 99.14% โดยมากกว่า Genetic Algorithm ที่ความถูกต้อง 97.87%

จากการศึกษางานวิจัยที่เกี่ยวข้องในกลุ่มของเทคนิคและการพัฒนาระบบแนะนำอาหารและการปรับปรุงประสิทธิภาพของระบบด้วยขั้นตอนวิธี Swarm Intelligence พบว่าการพัฒนาระบบแนะนำอาหารด้วยเทคนิค Hybrid Recommendation System ได้รับความนิยมอย่างแพร่หลายและช่วยเพิ่มประสิทธิภาพในการทำงานของระบบ การพิจารณาสารอาหารและคุณค่าทางโภชนาการสามารถตอบโจทย์และสร้างความพึงพอใจมากขึ้นสำหรับการแนะนำรายการอาหาร อย่างเหมาะสมตามพฤติกรรมรายบุคคลได้ ทั้งนี้ มีการศึกษาและพัฒนาระบบแนะนำรายการอาหารไทยในหลายงานวิจัยแต่ยังไม่ได้พัฒนาบนชุดข้อมูลโภชนาการอาหารไทยมากนัก การประยุกต์ขั้นตอนวิธี Swarm Intelligence ในการพัฒนาระบบแนะนำสามารถช่วยเพิ่มความถูกต้องและประมวลผลเชิงลึกเพื่อปรับปรุงประสิทธิภาพการทำงานของระบบ เช่น การประยุกต์ใช้ Particle Swarm Optimization หรือ Cuckoo Search Algorithm ร่วมกับ K-Means และ Collaborative Filtering Technique เพื่อเรียนรู้พฤติกรรม ความชื่นชอบเชิงลึกส่วนบุคคล และสามารถแนะนำรายการอาหารได้อย่างถูกต้องมากยิ่งขึ้น ดังนั้น ผู้วิจัยจึงนำเสนอการศึกษาและพัฒนาระบบแนะนำอาหารไทยแบบ Hybrid Recommender System ที่พัฒนาบนชุดข้อมูลโภชนาการอาหารไทยเพื่อแนะนำรายการอาหารตามพฤติกรรมเชิงลึกรายบุคคลอย่างเหมาะสม รวมทั้งการประยุกต์ใช้ขั้นตอนวิธี Particle Swarm Optimization และ Cuckoo Search Algorithm สำหรับการศึกษากระบวนการทำงาน ทดลองและพัฒนาแบบจำลองร่วมกับเทคนิค K-Means และ Collaborative Filtering เพื่อประสิทธิภาพสูงสุดในการแนะนำรายการอาหารตามโภชนาการรายบุคคล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การวิจัยและวิธีการดำเนินงาน

เนื้อหาในบทนี้อธิบายถึงขั้นตอนการศึกษาและพัฒนาระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการรายบุคคล ซึ่งสร้างโดยใช้เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) ร่วมกับเทคนิคการแบ่งกลุ่มข้อมูล K-Means และประยุตก์ใช้ขั้นตอนวิธี Swarm Intelligence แบบไฮบริดประกอบด้วยขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization) และขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) ในการปรับปรุงประสิทธิภาพของการสร้างแบบจำลอง (Model) โดยมีการทำงานทั้งหมด 8 ขั้นตอนหลักได้แก่ 1) การรวบรวม จัดเตรียมและประมวลผลข้อมูล (Data Integration and Preparation) 2) การจัดกลุ่มผู้ใช้งาน (User's Behavior Clustering) 3) การทำนายและแนะนำอาหาร (Score prediction & Top-N Ranking) 4) การวัดประสิทธิภาพของการแนะนำรายการอาหาร (Model Evaluation) 5) การพัฒนาและปรับปรุงแบบจำลอง (Model implement and tunning) 6) การคำนวณหาชุดรายการอาหารที่เหมาะสม (Food Optimization) 7) การแนะนำชุดรายการอาหาร (Food Recommendation) และ 8) การประเมินความพึงพอใจชุดรายการอาหารแนะนำ แสดงโครงสร้างการทำงานในการพัฒนาระบบดังรูปที่ 3.1



รูปที่ 3.1 โครงสร้างการทำงานระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1 การรวบรวม จัดเตรียมและประมวลผลข้อมูล

3.1.1 ชุดข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่ใช้ในงานวิจัยนี้ เป็นข้อมูลการบันทึกการทานอาหารบนแอพพลิเคชัน FoodiEat ของห้องวิจัยการวิเคราะห์พฤติกรรมมนุษย์ ศูนย์เทคโนโลยีอิเล็กทรอกนิกส์และคอมพิวเตอร์แห่งชาติ จำนวน 6 ชุดข้อมูล ประกอบด้วย ข้อมูลประวัติการบันทึการรับประทานอาหาร (Dietary logs) ข้อมูลผู้ใช้งาน (Users) ข้อมูลรายการอาหาร (foods) ข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs) ข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs) และข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity Log) ซึ่งเก็บข้อมูลตั้งแต่ปี 2014-2019 จำนวน 307,102 ข้อมูล ตัวอย่างข้อมูลที่ใช้ในการวิจัย แสดงดังตารางที่ 3.1 – 3.6

ตารางที่ 3.1 ตัวอย่างข้อมูลประวัติการบันทึกการรับประทานอาหาร (Dietary logs)

food_id	user_id	date_eat	interval_type	portion	cal
5221	11	4/25/2557	2	1	115
2440	18	4/25/2557	2	1	357.2
1508	11	4/25/2557	2	1	656.1
5137	11	4/26/2557	3	1	320
1414	11	4/26/2557	2	1	261

จากตารางที่ 3.1 แสดงข้อมูลประวัติการบันทึกการรับประทานอาหารรายครั้งของผู้ใช้งานแอปพลิเคชัน โดยจะบันทึกเมนูอาหาร วันที่รับประทาน และจำนวน portion โดยแอปพลิเคชันจะคำนวณและบันทึกปริมาณแคลอรี่ให้อัตโนมัติจากเมนูอาหารและจำนวน portion

ตารางที่ 3.2 ตัวอย่างข้อมูลผู้ใช้งาน (Users)

user_id	gender	Date of birth
1	0	10/25/1977
3	1	10/17/1980
4	1	10/17/1978

จากตารางที่ 3.2 แสดงข้อมูลผู้ใช้งาน จัดเก็บข้อมูลเพศ และวัน/เดือน/ปีที่เกิดสำหรับคำนวณอายุของผู้ใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ตัวอย่างข้อมูลรายการอาหาร (foods)

id	food_name	food_type	cal	protein	fat	carb
1201	ก๋วยจั๊บ	อาหารจานเดียว	465	24	25	36
1202	กระเพาะปลา	อาหารจานเดียว	240	19	10	18
1203	ก๋วยเตี๋ยวน้ำน่องไก่ตุ๋น	อาหารจานเดียว	428	23	13	55
1204	ก๋วยเตี๋ยวเนื้อสับ	อาหารจานเดียว	417	13	15	56
1205	ก๋วยเตี๋ยวเป็ด	อาหารจานเดียว	330	18	9	44

จากตารางที่ 3.3 แสดงข้อมูลรายการอาหาร แสดงชื่อเมนู ประเภทรายการอาหาร ปริมาณสารอาหารได้แก่ แคลอรี่ โปรตีน ไขมัน และคาร์โบไฮเดรต ซึ่งได้รับจากฐานข้อมูลจากสถาบันโภชนาการ มหาวิทยาลัยมหิดล

ตารางที่ 3.4 ตัวอย่างข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs)

wid	ref_mid	weight	create_date
1	2	52	3/13/2014 16:40
2	3	65	3/13/2014 16:47
3	4	76	3/13/2014 17:05
4	5	59	3/14/2014 16:47
5	6	100	3/14/2014 16:59

จากตารางที่ 3.4 แสดงข้อมูลประวัติการบันทึกการชั่งน้ำหนักของผู้ใช้งานแอปพลิเคชันรายบุคคล โดยจัดเก็บข้อมูลปริมาณน้ำหนักที่ชั่งวัดและวันที่ เวลา ณ เวลาที่บันทึกในแต่ละครั้ง

ตารางที่ 3.5 ตัวอย่างข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs)

hid	ref_mid	height	create_date
1	2	151	3/13/2014 16:40
2	3	172	3/13/2014 16:47
3	4	170	3/13/2014 17:05
4	5	147	3/14/2014 16:47
5	6	183	3/14/2014 16:59

จากตารางที่ 3.5 แสดงข้อมูลประวัติการบันทึกการวัดส่วนสูงของผู้ใช้งานแอปพลิเคชันรายบุคคล โดยจัดเก็บข้อมูลส่วนสูงที่วัดและวันที่ เวลา ณ เวลาที่บันทึกได้ในแต่ละครั้ง

ตารางที่ 3.6 ตัวอย่างข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity Log)

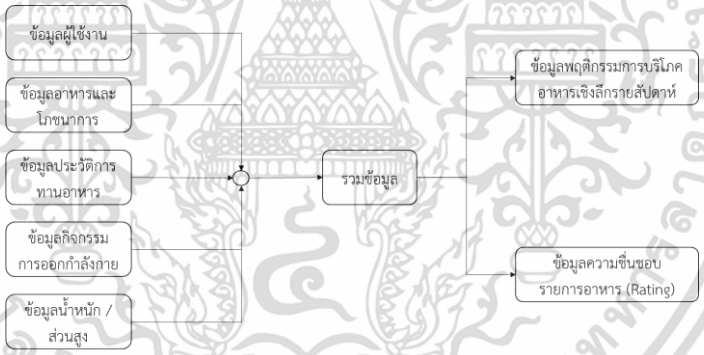
act_id	user_id	desc	date_act	duration	cal	interval_type
2	3	เดิน	3/25/2014	180	100	1
3	3	เดิน	3/26/2557	200	400	1
6	3	ขี่จักรยาน	3/26/2557	120	960	2
25	11	การเดินทาง	3/28/2557	5	15	3
32	14	วิ่ง	3/29/2557	45	180	3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 3.6 แสดงข้อมูลประวัติการบันทึกการบันทึกการออกกำลังกายของผู้ใช้งานแอปพลิเคชันรายบุคคล โดยจัดเก็บประเภทการออกกำลังกาย วันที่ออก ระยะเวลาที่ออกกำลังกาย โดยแอปพลิเคชันคำนวณปริมาณแคลอรี่ที่เผาผลาญให้อัตโนมัติ

3.1.2 การเตรียมและประมวลผลข้อมูล

สำหรับการพัฒนาระบบแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม จะดำเนินการวิเคราะห์พฤติกรรมการรับประทานอาหารของผู้ใช้งานด้วยการจัดกลุ่มผู้ใช้ที่มีลักษณะการใช้ชีวิตและการเลือกบริโภคอาหารอยู่ภายใต้กลุ่มเดียวกัน จึงมีการจัดเตรียมและเปลี่ยนแปลงหรือคำนวณข้อมูลอื่นเพื่อให้ได้ข้อมูลเชิงลึกอันเป็นประโยชน์ในการใช้งานต่อไป ตัวอย่าง ข้อมูลสุขภาพและการออกกำลังกายสำหรับวิเคราะห์หาข้อมูลความต้องการขั้นพื้นฐานต่อวันรายบุคคล ซึ่งขึ้นอยู่กับข้อมูลพฤติกรรมและข้อมูลสุขภาพของผู้ใช้งานรายบุคคล เช่น อัตราความต้องการเผาผลาญของร่างกายในชีวิตประจำวัน หรือ Basal Metabolic Rate (BMR) ปริมาณการใช้พลังงานรายบุคคลต่อวัน หรือ Total Daily Energy Expenditure (TDEE) เป็นต้น ทั้งนี้ เนื่องด้วยข้อมูลมีที่มาจากหลายแหล่งและเป็นข้อมูลที่ได้จากการบันทึกข้อมูลจากผู้ใช้งานแอปพลิเคชันเป็นหลัก จึงต้องมีการเตรียมและประมวลผลข้อมูลก่อนนำไปพัฒนาระบบแนะนำในขั้นตอนต่อไป โดยแบ่งเป็น 3 ขั้นตอนหลักตามรูปที่ 3.2 ได้แก่ 1) ขั้นตอนการทำความสะอาดข้อมูล 2) ขั้นตอนการรวมข้อมูลเข้าด้วยกัน และ 3) ขั้นตอนการแปลงข้อมูล



รูปที่ 3.2 การเตรียมและประมวลผลข้อมูลสำหรับการจัดกลุ่มและแนะนำรายการอาหาร

จากรูปที่ 3.2 แสดงรายละเอียดการเตรียมและจัดการข้อมูล ได้ดังนี้

3.1.2.1 การทำความสะอาดข้อมูล (Data Cleansing)

ขั้นตอนการทำความสะอาดข้อมูล (Data Cleansing) เป็นขั้นตอนแรกในการเตรียมข้อมูล โดยทำการคัดเลือกและปรับแต่งข้อมูลให้สอดคล้องกัน และเพื่อให้ง่ายต่อการวิเคราะห์ข้อมูล มีรายละเอียดการทำงานดังนี้

1. ข้อมูลผู้ใช้งาน (users) ทำการแปลงข้อมูลวัน/เดือน/ปีเกิด (dob) จากปีพุทธศักราช (B.C) เป็นปีคริสต์ศักราช (A.D) ข้อมูลชื่อ (name) แปลงข้อมูลจากอีเมลเต็มเหลือเพียงชื่อผู้ใช้งานเท่านั้น

2. ข้อมูลรายการอาหาร (foods) เลือกข้อมูลประเภทอาหารจำนวน 9 ประเภท ได้แก่ 'เครื่องดื่ม', 'อาหารจานเดียว', 'ขนมหวาน', 'กับข้าว', 'ขนมกรุบกรอบ', 'ผลไม้', 'อาหาร brand',

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และ 'อาหารแช่แข็ง' และลบข้อมูลรายการอาหารที่ค่าสารอาหาร ได้แก่ แคลอรี่ (cal), โปรตีน (protein), ไขมัน (fat) และคาร์โบไฮเดรต (carb) มีค่าเป็น ค่าว่าง หรือ NaN

3. ข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs) เลือกข้อมูลน้ำหนักโดยอ้างอิงจากการสำรวจรูปร่างทั่วประเทศ โครงการ SizeThailand ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (เนคเทค) ซึ่งเพศชายมีน้ำหนักเฉลี่ย 68.83 กิโลกรัม และเพศหญิง 57.40 กิโลกรัม ทั้งนี้เพื่อให้ครอบคลุมกลุ่มตัวอย่างข้อมูล ผู้วิจัยจึงเลือกช่วงน้ำหนักเพศชาย ระหว่าง 40 – 120 กิโลกรัม และช่วงน้ำหนักเพศหญิง ระหว่าง 35 – 110 กิโลกรัม ซึ่งห่างจากค่าเฉลี่ยประมาณ 20-50 กิโลกรัม และทำการเลือกข้อมูลน้ำหนักที่มีประวัติการบันทึกล่าสุดจากผู้ใช้งาน เพื่อใช้ในการคำนวณค่า BMI ประจำตัวต่อไป

4. ข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs) ทำการเลือกข้อมูลส่วนสูง โดยอ้างอิงจากการสำรวจรูปร่างทั่วประเทศ โครงการ SizeThailand เช่นกัน โดยข้อมูลเพศชายมีส่วนสูงเฉลี่ย 169.46 เซนติเมตร และเพศหญิง 157 เซนติเมตร ทั้งนี้เพื่อให้ครอบคลุมกลุ่มตัวอย่างข้อมูล ดังนั้น ผู้วิจัยจึงเลือกช่วงส่วนสูงเพศชายระหว่าง 150-200 เซนติเมตร และเพศหญิงระหว่าง 135-160 เซนติเมตร ซึ่งห่างจากค่าเฉลี่ยประมาณ 20-30 เซนติเมตร และทำการเลือกข้อมูลส่วนสูงที่มีประวัติการบันทึกล่าสุดจากผู้ใช้งาน เพื่อใช้ในการคำนวณค่า BMI ประจำตัวต่อไป

5. ข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity Log) ทำการเลือกข้อมูลออกกำลังกายของผู้ใช้งานระหว่างปี 2013 – 2019 เท่านั้น

6. ข้อมูลประวัติการบันทึกการรับประทานอาหาร (Dietary logs) ทำการเลือกข้อมูลที่มีข้อมูลผู้ใช้งานและข้อมูลรายการอาหารภายในระบบเท่านั้น

3.1.2.2 การรวมข้อมูลเข้าด้วยกัน (Data Integration)

ขั้นตอนการรวมข้อมูลเข้าด้วยกัน (Data Integration) เป็นการรวมข้อมูลจำนวน 6 ชุดข้อมูล ประกอบด้วย ข้อมูลประวัติการบันทึกการรับประทานอาหาร (Dietary logs) ข้อมูลผู้ใช้งาน (Users) ข้อมูลรายการอาหาร (foods) ข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs) ข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs) และข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity Log) โดยมีรายละเอียดการรวมข้อมูล ดังนี้

1. รวมข้อมูลระหว่างข้อมูลประวัติการบันทึการรับประทานอาหาร (Dietary logs) และข้อมูลรายการอาหาร (foods) ด้วยข้อมูลรหัสผู้ใช้งาน (food_id)

2. รวมข้อมูลระหว่างข้อมูลประวัติการบันทึกการชั่งน้ำหนัก (Weight logs) และข้อมูลผู้ใช้งาน (users) ด้วยข้อมูลรหัสผู้ใช้งาน (user_id)

3. รวมข้อมูลระหว่างข้อมูลประวัติการบันทึกการวัดส่วนสูง (Height logs) และข้อมูลผู้ใช้งาน (users) ด้วยข้อมูลรหัสผู้ใช้งาน (user_id)

4. รวมข้อมูลระหว่างข้อมูลประวัติการบันทึกการออกกำลังกาย (Activity Log) และข้อมูลผู้ใช้งาน (users) ด้วยข้อมูลรหัสผู้ใช้งาน (user_id)

5. รวมข้อมูลระหว่างข้อมูลประวัติการบันทึการรับประทานอาหาร (Dietary logs) และข้อมูลผู้ใช้งาน (Users) ด้วยข้อมูลรหัสผู้ใช้งาน (user_id)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.3 การแปลงข้อมูล (Data Transformation)

ขั้นตอนการแปลงข้อมูล (Data Transformation) เป็นการจัดทำหรือดัดแปลงข้อมูลที่มีอยู่เดิมให้เกิดข้อมูลหรือสารสนเทศใหม่ที่จำเป็นต่อการวิเคราะห์ข้อมูล เช่นการคำนวณอายุจากวันเดือนปีเกิด การหาค่าร้อยละจากการกินอาหารประเภทต่าง ๆ เป็นต้น โดยมีรายละเอียดดังนี้

1. คำนวณอายุจากวัน/เดือน/ปีเกิดของข้อมูลผู้ใช้งาน และเลือกข้อมูลผู้ใช้งานที่มีอายุระหว่าง 15-70 ปีเท่านั้น และดำเนินการแบ่งช่วงข้อมูลตามข้อมูลการสำรวจพฤติกรรมการบริโภคอาหารของประชากร ปี 2560 [1] ซึ่งแบ่งเป็น 3 ช่วง ได้แก่ 15-24 ปี, 25-59 ปี และ 60 ปีขึ้นไป เนื่องจากผู้วิจัยเล็งเห็นว่าช่วงอายุ 25-59 ปี มีช่วงที่กว้างมากกว่ากลุ่มอื่นๆ จึงดำเนินการแบ่งกลุ่มตามจำนวนข้อมูลในวิจัยเพื่อจำนวนที่ใกล้เคียงกัน ได้จำนวน 2 กลุ่ม ได้แก่ 25-39 ปี และ 40-59 ปี
2. คำนวณค่าเฉลี่ยจำนวนหน่วยบริโภคอาหาร (Portion) ของผู้ใช้งานจากการกินอาหารทั้งหมด
3. คำนวณค่าดัชนีมวลกาย (BMI) จากน้ำหนัก-ส่วนสูงของผู้ใช้งานที่มีการบันทึกล่าสุดตามเกณฑ์ดัชนีมวลกายของกรมควบคุมโรค กระทรวงสาธารณสุข
4. คำนวณค่าอัตราความต้องการเผาผลาญของร่างกายในชีวิตประจำวัน (BMR)
5. คำนวณค่าปริมาณการใช้พลังงานต่อวันของผู้ใช้งาน (TDEE)
6. คำนวณค่าเฉลี่ยแคลอรี่รวมต่อวันจากการออกกำลังกายของผู้ใช้งาน
7. คำนวณค่าเฉลี่ยของจำนวนการออกกำลังกายต่อสัปดาห์ของผู้ใช้งาน
8. คำนวณค่าร้อยละของจำนวนการกินอาหารแต่ละประเภทของผู้ใช้งาน
9. คำนวณค่าร้อยละของช่วงเวลาการกินอาหารของผู้ใช้งาน
10. คำนวณค่าเฉลี่ยแคลอรี่ของจำนวนการกินอาหารแต่ละประเภท
11. คำนวณค่าเฉลี่ยแคลอรี่ของการกินอาหารแต่ละช่วงเวลา
12. คำนวณค่าคะแนนต่อรายการอาหารของผู้ใช้งานในแต่ละสัปดาห์ โดยได้ทำการออกแบบสูตรคำนวณขึ้นมาใหม่ [34] เพื่อใช้สำหรับเป็นตัวแทนของข้อมูลความนิยมที่ผู้รับประทานมีต่อเมนูอาหารแต่ละอย่างโดยใช้สมการ 3.1

$$food_score_{(u,f)} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{m} portion_{(f,j)}}{m_i}\right) \quad (3.1)$$

ซึ่ง $portion_{(f,j)}$ คือ จำนวน portion ในการกินอาหารแต่ละเมนู

$m_{(i)}$ คือ จำนวนสัปดาห์ในการกินอาหาร

N คือ จำนวนสัปดาห์ทั้งหมดของการกินอาหารแต่ละเมนู

13. คำนวณค่าความชื่นชอบ (rating) ต่อรายการอาหารของผู้ใช้งาน โดยใช้หลักเกณฑ์การแบ่งช่วงคะแนน ดังตารางที่ 3.7

ตารางที่ 3.7 ช่วงคะแนนและค่า Rating ของพฤติกรรมการกินอาหาร

ช่วงคะแนน	Rating	ช่วงคะแนน	Rating
< 0.5	1	> 1.0 - < 2.5	4
>= 0.5 - < 1.0	2	>= 2.5	5
1.0	3		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การจัดกลุ่มผู้ใช้งาน (User's Behavior Clustering)

3.2.1 การเตรียมข้อมูลสำหรับการจัดกลุ่มพฤติกรรมการรับประทานอาหาร

การเตรียมข้อมูลที่ใช้ในการจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งานที่ผ่านการรวบรวม จัดเตรียมและประมวลผลข้อมูลตาม 3.1 แล้วนั้น จะได้จำนวน 5,957 ผู้ใช้งาน พร้อมกับประวัติและพฤติกรรมการรับประทานอาหารภายในแอปพลิเคชัน FoodiEat และข้อมูลที่ใช้ในการจัดกลุ่มนั้น จะถูกประมวลผลและดัดแปลงให้อยู่ในรูปแบบที่เหมาะสมสำหรับขั้นตอนวิธีการจัดกลุ่มตามพฤติกรรมการรับประทานอาการของผู้ใช้งาน ได้เป็น 27 รายการ โดยมีรายละเอียด ดังตารางที่ 3.8

ตารางที่ 3.8 รายละเอียดชุดข้อมูลสำหรับการจัดกลุ่มพฤติกรรมการรับประทานอาหาร

ลำดับ	ชื่อรายการข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
1	age_group_1	อายุของผู้ใช้งานอยู่ระหว่าง 15-24 ปี	0
2	age_group_2	อายุของผู้ใช้งานอยู่ระหว่าง 25-39 ปี	1
3	age_group_3	อายุของผู้ใช้งานอยู่ระหว่าง 40-59 ปี	0
4	age_group_4	อายุของผู้ใช้งานมากกว่า 60 ปี	0
5	bmi_group_1	BMI มีค่าน้อยกว่า 18.5	0
6	bmi_group_2	BMI อยู่ในช่วง 18.5 – 22.9	0
7	bmi_group_3	BMI อยู่ในช่วง 23.0 – 24.9	1
8	bmi_group_4	BMI อยู่ในช่วง 25.0 – 29.90	0
9	bmi_group_5	BMI มีค่าเท่ากับหรือมากกว่า 30	0
10	gender_0	เพศชาย	1
11	gender_1	เพศหญิง	0
12	calories	ค่าเฉลี่ยของแคลอรี่ที่ผู้ใช้งานบันทึกการรับประทานอาหารทั้งหมด	416.44
13	protein	ค่าเฉลี่ยของโปรตีนที่ผู้ใช้งานบันทึกการรับประทานอาหารทั้งหมด	20.86
14	carbohydrates	ค่าเฉลี่ยของคาร์โบไฮเดรตที่ผู้ใช้งานบันทึกการรับประทานอาหารทั้งหมด	49.23
15	fats	ค่าเฉลี่ยของไขมันที่ผู้ใช้งานบันทึกการรับประทานอาหารทั้งหมด	14.96
16	food_weekly_mean	ค่าเฉลี่ยของจำนวนเมนูอาหารที่ทานรายสัปดาห์	11
17	ratio_seafoods	อัตราส่วนอาหารประเภททะเลที่รับประทาน	2.33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.8 รายละเอียดชุดข้อมูลสำหรับการจัดกลุ่มพฤติกรรมการรับประทานอาหาร (ต่อ)

ลำดับ	ชื่อรายการข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
18	ratio_meatfoods	อัตราส่วนอาหารประเภทเนื้อที่รับประทาน	32.56
19	ratio_fishfoods	อัตราส่วนอาหารประเภทปลาที่รับประทาน	13.95
20	ratio_eggfoods	อัตราส่วนอาหารประเภทไข่ที่รับประทาน	2.33
21	ratio_otherfoods	อัตราส่วนอาหารประเภทอื่นที่รับประทาน	48.84
22	Interval_bf_rank	อันดับของมื้ออาหารเช้าที่นิยมรับประทาน	1
23	Interval_ln_rank	อันดับของมื้ออาหารกลางวันที่นิยมรับประทาน	3
24	Interval_dn_rank	อันดับของมื้ออาหารเย็นที่นิยมรับประทาน	2
25	Interval_sp_rank	อันดับของมื้ออาหารดึกที่นิยมรับประทาน	4
26	side_dish	ค่าเฉลี่ยของกับข้าวที่รับประทานรายสัปดาห์	9.5
27	single_dish	ค่าเฉลี่ยของอาหารจานเดียวที่รับประทานรายสัปดาห์	10

3.2.2 การลดมิติข้อมูลด้วย Principal Component Analysis (PCA)

การลดมิติข้อมูล (Dimensionality Reduction) เป็นวิธีการเรียนรู้ของเครื่อง โดยทำการการลดรายการข้อมูล (Feature) ตั้งต้นที่มีรายการข้อมูลจำนวนมากหรืออาจไม่มีประโยชน์ต่อการนำมาวิเคราะห์ข้อมูลให้มีขนาดเล็กลงโดยสูญเสียลักษณะสำคัญของข้อมูลและความถูกต้องของผลลัพธ์น้อยที่สุด และสามารถเลือกข้อมูลที่มีความสำคัญและเป็นตัวแทนของชุดข้อมูลส่วนใหญ่ได้ โดยทำการคำนวณส่วนประกอบหลักของข้อมูลโดยการค้นหาเวกเตอร์ลักษณะเฉพาะ ของความแปรปรวนของข้อมูล จำนวนส่วนประกอบหลักที่เหมาะสมคือจำนวนที่สามารถครอบคลุมตัวแปรที่มีความสัมพันธ์ ซึ่งจับความแปรปรวนจำนวนมากที่สุดในข้อมูลและจะมีจำนวนน้อยกว่าหรือเท่ากับจำนวนตัวแปรรายการข้อมูลเดิม ดังนั้น PCA สามารถจับข้อมูลที่จำเป็นภายในคุณสมบัติของข้อมูลและเปลี่ยนเป็นส่วนประกอบหลัก ซึ่งจำนวนมิตินั้นต่ำกว่าข้อมูลต้นฉบับ [35]

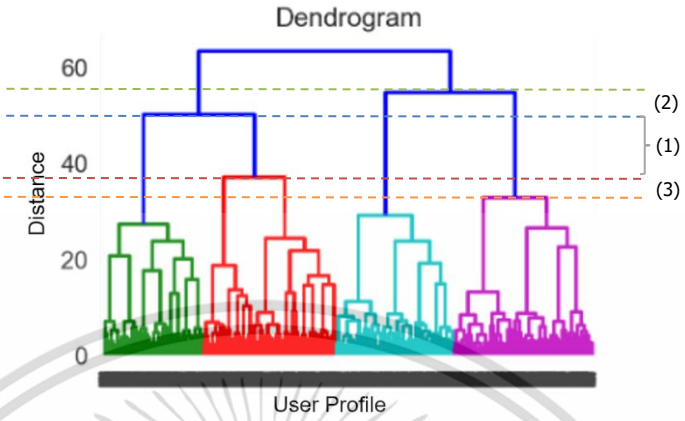
การคัดเลือกจำนวนมิติของข้อมูล (dimension) โดยยอมรับจำนวนมิติน้อยที่สุดที่สามารถอธิบายความแปรปรวนครอบคลุมข้อมูลทั้งหมด (Total variation explained) อย่างน้อย 80% จึงเลือกจำนวนมิติข้อมูลที่ 8 dimension ซึ่งสามารถอธิบายความแปรปรวนของข้อมูลได้ที่ 80.37%

3.3.3 การเลือกจำนวนกลุ่มในการจัดกลุ่มข้อมูลพฤติกรรมการรับประทานอาหาร

ในงานวิจัยนี้ ผู้วิจัยได้เลือกวิธีการสร้างแผนภาพเดนโดรแกรม (Dendrogram) เพื่อหาจำนวนกลุ่มของพฤติกรรมการรับประทานอาหารของผู้ใช้งานที่เหมาะสม โดยคำนึงถึงระยะความห่างอย่างชัดเจนของแต่ละกลุ่ม การทำงานของการสร้างแผนภาพ มีการคำนวณความคล้ายคลึงหรือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระยะทางระหว่างข้อมูล เลือกหลักการการรวมกลุ่ม โดยงานวิจัยนี้ ผู้วิจัยเลือกหลักการรวมกลุ่ม คือ Ward's Method โดยพิจารณาจากค่า sum of the squared within-cluster distance และพิจารณาจำนวนกลุ่มขึ้นอยู่กับที่มีระยะทางห่างกับจำนวนกลุ่มอื่นอย่างชัดเจน



รูปที่ 3.3 ภาพเดนโดรแกรมของการจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งาน

จากรูปที่ 3.3 จะพบว่าภาพเดนโดรแกรมของกลุ่มข้อมูลในเส้นประสีน้ำเงินที่มีระยะทางห่างกันจากกลุ่มอื่นในช่วง 40-50 อย่างชัดเจน ซึ่งสามารถแบ่งข้อมูลออกได้เป็น 4 กลุ่ม หากพิจารณาการแบ่งเส้นออกเป็นจำนวนกลุ่มอื่น ๆ หมายเลข (2) และ (3) พบว่าระยะทางจะแคบกว่าในช่วง 40-50 หมายเลข (1) ยกตัวอย่างเช่น การแบ่งกลุ่มเส้นสีน้ำเงินในช่วง 50-55 (เส้นประสีเขียว) แบ่งออกเป็น 3 กลุ่ม หรือการแบ่งกลุ่มเส้นสีแดงช่วง 30-35 (เส้นประสีแดง) แบ่งออกเป็น 5 กลุ่ม และการแบ่งกลุ่มเส้นสีม่วงช่วง 20-25 (เส้นประสีส้ม) แบ่งออกเป็น 6 กลุ่ม ดังนั้น ผู้วิจัยจึงได้เลือกจำนวนการแบ่งกลุ่มข้อมูลออกเป็น 4 กลุ่มสำหรับการวิจัยในงานนี้ เพื่อประสิทธิภาพในการจัดกลุ่มที่หาความสัมพันธ์ที่คล้ายกันในแต่ละกัน และมีความแตกต่างกันจากกลุ่มอื่น

3.3 การทำนายและแนะนำอาหาร (Score prediction & Top-N Ranking)

3.3.1 รูปแบบข้อมูลสำหรับการพัฒนาแบบจำลองระบบแนะนำ

ชุดข้อมูลที่ใช้ในขั้นตอนนี้ เป็นข้อมูลที่คำนวณหาความชื่นชอบที่ผู้ใช้งานมีต่อ 1 รายการอาหาร โดยใช้สูตร $food\ score_{(u,f)} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{m} portion_{(f,j)}}{m_i}\right)$ และได้ดัดแปลงข้อมูลจาก food score สู่คะแนนความชื่นชอบหรือ Rating ให้อยู่ในช่วง 1-5 จำนวน 59,255 รายการ โดยมีรูปแบบข้อมูลตามตารางที่ 3.9

ตารางที่ 3.9 รายละเอียดชุดข้อมูลสำหรับการทำนายคะแนนระบบแนะนำ

ลำดับ	ชื่อรายการข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
1	User_id	รหัสของผู้ใช้งาน	U001
2	Food_id	รหัสของเมนูอาหาร	F005
3	Rating	คะแนนความชื่นชอบ	3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 การเลือกชุดข้อมูลเรียนรู้และทดสอบแบบจำลอง

การเลือกข้อมูลจากชุดข้อมูลสำหรับพัฒนาแบบจำลองระบบแนะนำ จำนวน 59,255 รายการ โดยแบ่งชุดสอนแบบจำลอง (Training) และชุดทดสอบ (Testing) อย่างละ 70% และ 30% ตามลำดับ เพื่อให้ในการฝึกสอนแบบจำลองและทดสอบประสิทธิภาพของแบบจำลอง

3.3.3 การเลือกขั้นตอนวิธีสำหรับการวิจัยแบบจำลองระบบแนะนำ

การคัดเลือกขั้นตอนวิธีเพื่อนำมาศึกษาและทดลองในงานวิจัยนี้ ผู้วิจัยได้ทำการคัดเลือกและพัฒนาแบบจำลองด้วยเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) ในรูปแบบ User-User หรือเรียกอีกอย่างหนึ่งว่า User-User Collaborative Filtering Technique เนื่องจากเป็นขั้นตอนวิธีที่ยังถูกใช้ในระบบแนะนำอย่างแพร่หลาย และนำมาพัฒนาต่อยอดร่วมกับขั้นตอนวิธีอื่นได้โดยง่ายและเพื่อเพิ่มประสิทธิภาพการทำงานให้เทียบเท่าขั้นตอนวิธีที่การทำงานมีความซับซ้อนได้ แต่ยังคงเอกลักษณ์ของความเข้าใจง่ายในหลักการทำงาน และเลือกใช้ข้อมูลที่ได้จากการจัดกลุ่มตาม 3.2 ที่มีประสิทธิภาพของการจัดกลุ่มพฤติกรรมการรับประทานอาหารเหมาะสมที่สุด เพื่อให้แบบจำลองระบบแนะนำ สามารถแนะนำรายการอาหารไทยได้อย่างถูกต้อง

3.3.4 การทำนายค่าคะแนนความชื่นชอบ (Rating)

ในงานวิจัยนี้ได้คัดเลือกขั้นตอนวิธีสำหรับพัฒนาแบบจำลองระบบแนะนำด้วยวิธี User-User Collaborative Filtering ซึ่งเป็นการหาความสัมพันธ์ของผู้ใช้งาน (user) ที่มีความสนใจหรือพฤติกรรมที่ใกล้เคียงกัน โดยดูจากค่าคะแนนความชื่นชอบ (Rating) ที่ผู้ใช้งานมีต่ออาหาร (food) ในระบบตามความสนใจรายบุคคล ทำการคำนวณหาความคล้ายคลึงกันของพฤติกรรมของผู้ใช้งานทั้งหมด จากนั้นจึงเลือกผู้ใช้งานที่มีพฤติกรรมใกล้เคียงกันมากที่สุดตามจำนวนที่สนใจ และคำนวณทำนายค่าคะแนนความชื่นชอบให้กับรายการอาหารที่ผู้ใช้งานคนนั้น ๆ ยังไม่ได้มีความสนใจเกิดขึ้น

สมการที่ใช้ในการคำนวณทำนายค่าคะแนนความชื่นชอบรายการอาหาร แสดงดังด้านล่าง

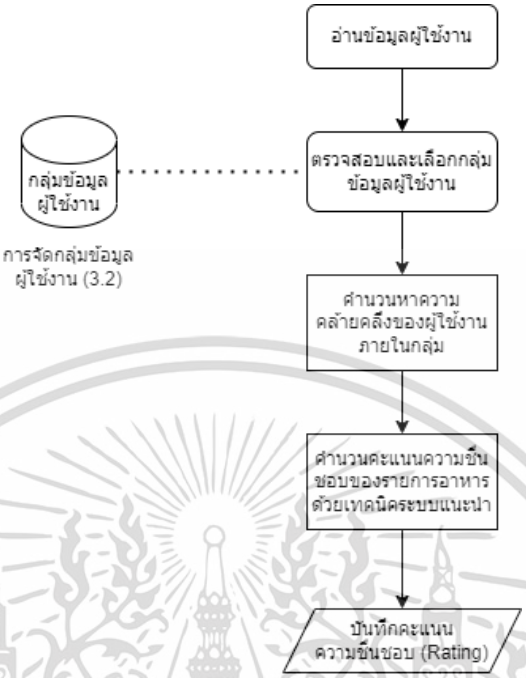
$$r_{ij} = \bar{r}_i + \frac{\sum_k similarity(i,u)(r_{uj} - \bar{r}_u)}{\sum_k |similarity(i,u)|} \tag{3.2}$$

โดยการคำนวณหาค่าความคล้ายคลึงของผู้ใช้งานนั้น สามารถประยุกต์ใช้สูตรคำนวนความคล้ายคลึงที่ได้รับความนิยมในการพัฒนาระบบแนะนำตามความต้องการ เช่น Euclidean distance similarity, Pairwise similarity, Pearson’s correlation, Spearman’s correlation หรือ Cosine similarity เป็นต้น ซึ่งในงานวิจัยนี้ ได้เลือกตัววัด Cosine similarity ในการคำนวณหาค่าความคล้ายคลึงของผู้ใช้งาน มีสมการในการคำนวณ ดังนี้

$$similarity_{(A,B)} = \cos(\theta) = \frac{A \cdot B}{||A||\,||B||} = \frac{\sum_{i=1}^{N} A_i B_i}{\sqrt{\sum_{i=1}^{N} A_i^2}\sqrt{\sum_{i=1}^{N} B_i^2}} \tag{3.3}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการคำนวณทำนายค่าคะแนนความชื่นชอบ จะได้ค่าคะแนนความชื่นชอบ (r_{ij}) ของรายการอาหารที่ผู้ใช้งานคนนั้นๆ ที่ยังไม่เคยบันทึกในแอปพลิเคชัน



รูปที่ 3.4 ขั้นตอนการทำงานสำหรับการคำนวณทำนายค่าคะแนนความชื่นชอบ (Rating)

จากรูปที่ 3.4 แสดงขั้นตอนการทำงานสำหรับการคำนวณทำนายค่าคะแนนความชื่นชอบ (Rating) โดยใช้เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมแบบ User-User Collaborative filtering ซึ่งจะใช้ข้อมูลผู้ใช้งานจากการจัดกลุ่มตามพฤติกรรมการรับประทานอาหาร ที่ได้จากขั้นตอน 3.2 มีรายละเอียดการทำงานดังนี้

1. ดำเนินการอ่านข้อมูลผู้ใช้งานที่จัดเตรียมไว้สำหรับการนำเข้าการคำนวณทำนายค่าคะแนนตามตาราง 3.10
2. ดำเนินการตรวจสอบและเลือกกลุ่มผู้ใช้งานตามกลุ่มที่ได้รับการจัดกลุ่มในขั้นตอน 3.2 การจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งาน
3. คำนวณหาความคล้ายคลึงความชื่นชอบต่อรายการอาหารของผู้ใช้งานเป้าหมายที่ต้องการแนะนำกับผู้ใช้งานอื่นภายในกลุ่มเท่านั้น ด้วยสูตรการคำนวณหาความคล้ายผู้ใช้งาน Cosine Similarity
4. คำนวณหาค่าคะแนนความชื่นชอบของผู้ใช้งานเป้าหมายที่มีต่อรายการอาหารอื่นทั้งหมดภายในระบบที่ผู้ใช้เป้าหมายยังไม่เคยสนใจหรือบันทึกผลการรับประทานอาหารด้วยเทคนิค User-User Collaborative Filtering
5. บันทึกผลการคำนวณคะแนนความชื่นชอบที่ได้จากข้อ 4 เพื่อใช้สำหรับการจัดลำดับและแนะนำรายการอาหารต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.5 การจัดอันดับรายการแนะนำ

หลังจากการคำนวณทำนายค่าคะแนนความชื่นชอบ (r_{ij}) ของผู้ใช้งานต่อรายการเมนูอาหารทั้งหมดในระบบแล้ว จึงนำค่าคะแนนความชื่นชอบมาเรียงลำดับจากมากไปน้อย และคัดเลือกจำนวนรายการแนะนำตามจำนวนที่กำหนด โดยในงานวิจัยนี้ทำการคัดเลือกจำนวน 4 ชุดรายการ เพื่อใช้ในการทดสอบประสิทธิภาพการพัฒนาแบบจำลอง ได้แก่ Top-5 , Top-10 , Top-15 และ Top-20 ดังรูปที่ 3.5

รายการแนะนำอาหาร (Top-10)

ลำดับ	เมนูอาหาร	ปริมาณแคลอรี่
1	หมูทอด(สะโพก,สันนอก)	315.0
2	เกาเหลาหมู	163.0
3	ข้าวเหนียว	222.2
4	ข้าวราดผัดคะน้าหมูกรอบ	516.0
5	ข้าวหมูกรอบ	490.0
6	ก๋วยเตี๋ยวเส้นเล็กหมู น้ำตก	463.0
7	เกาเหลาหมูตุ๋น	202.6
8	ข้าวไข่เจียวทรงเครื่อง (หมู)	740.8
9	ข้าวราดแกงเขียวหวานไก่	483.0
10	แพนงหมู	183.7

รูปที่ 3.5 ตัวอย่างรายการแนะนำอาหาร จำนวน 10 รายการ

3.4 การประเมินประสิทธิภาพของแบบจำลองการแนะนำรายการอาหาร (Model Evaluation)

การประเมินผลที่ใช้ในงานวิจัยนี้ เป็นการวัดประสิทธิภาพของแบบจำลองการแนะนำเมนูอาหาร โดยมีตัววัดการทำงานของจำนวน 3 ตัววัด ได้แก่ Normalized Discount Cumulative Gain (NDCG) , MAP (Mean Average Precision) และ Precision at K ซึ่งดำเนินการตรวจสอบผลลัพธ์ตามที่ได้รับรายการแนะนำจากขั้นตอนที่ 3.3.5

3.5 การพัฒนาและปรับปรุงแบบจำลอง (Model implement and tunning)

3.5.1 การพัฒนาแบบจำลองสำหรับการแนะนำรายการอาหาร งานวิจัยนี้ได้นำเสนอการพัฒนาแบบจำลองด้วยโครงสร้างแบบไฮบริด ชื่อว่า Hybrid-Swarm-KM และพัฒนาแบบจำลองสำหรับประเมินการเปรียบเทียบประสิทธิภาพการทำงานจำนวนทั้งหมด 9 แบบจำลอง ดังตารางที่ 3.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.10 รายการแบบจำลองการแนะนำรายการอาหาร

ลำดับ	ชื่อแบบจำลอง	คำอธิบาย
1	K-MEANS	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means บนชุดข้อมูลปกติ
2	KM-PCA	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means บนชุดข้อมูล PCA
3	KM-PSO	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids ที่เหมาะสมที่สุดด้วย PSO บนชุดข้อมูลปกติ
4	KM-PCA-PSO	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids ที่เหมาะสมที่สุดด้วย PSO บนชุดข้อมูล PCA
5	KM-CSA	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids ที่เหมาะสมที่สุดด้วย CSA บนชุดข้อมูลปกติ
6	KM-PCA-CSA	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids ที่เหมาะสมที่สุดด้วย PSO บนชุดข้อมูล PCA
7	WCPSO-KM	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids และค่าน้ำหนักของข้อมูลที่เหมาะสมที่สุดด้วย PSO บนชุดข้อมูลปกติ
8	WCCSA-KM	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids และค่าน้ำหนักของข้อมูลที่เหมาะสมที่สุดด้วย CSA บนชุดข้อมูลปกติ
9	PSO-CSA-KM	แบบจำลองที่พัฒนาจากขั้นตอนวิธี K-means และหาค่า centroids และค่าน้ำหนักของข้อมูลที่เหมาะสมที่สุดด้วย PSO และ CSA บนชุดข้อมูลปกติ

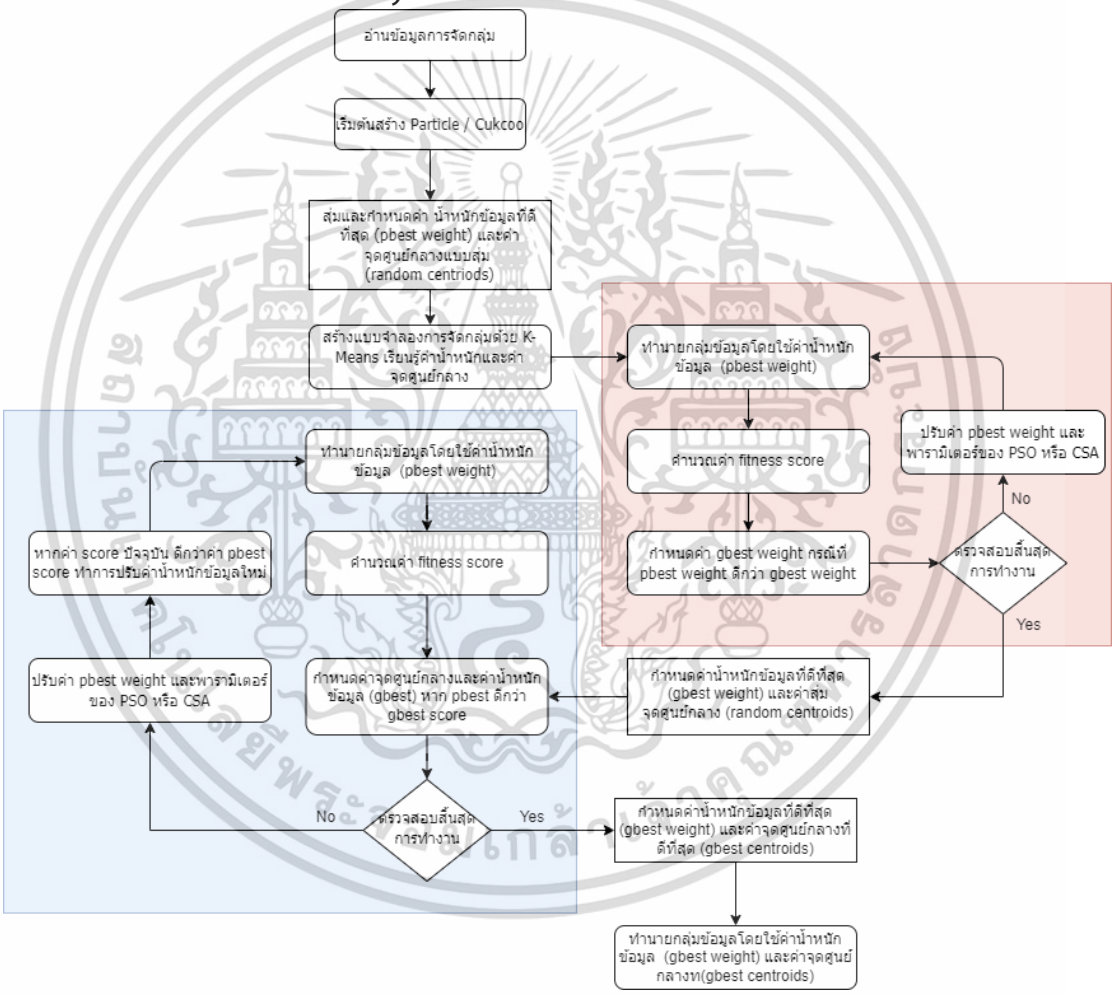
จากตารางที่ 3.10 แสดงรายละเอียดของแบบจำลองที่ใช้พัฒนาการจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งานและแนะนำรายการอาหารทั้ง 9 แบบจำลองเพื่อเปรียบเทียบประสิทธิภาพการทำงาน โดยผู้วิจัยมีความสนใจและศึกษาในการนำขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) จำนวน 2 ขั้นตอนวิธี ได้แก่ ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization : PSO) ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) มาประยุกต์ใช้ร่วมกับขั้นตอนวิธีการจัดกลุ่ม K-Means และการทำงานกับขั้นตอนวิธี Collaborative Filtering เพื่อเพิ่มประสิทธิภาพในการทำงานให้ดีขึ้น ซึ่งในงานวิจัยนี้ได้สนใจและนำเสนอการพัฒนาโครงสร้างขั้นตอนวิธีแบบผสม (Hybrid) โดยดำเนินการออกแบบโครงสร้างการทำงาน 2 ส่วนหลัก ได้แก่ การค้นหาค่าจุดศูนย์กลาง (Centroids) และค่าน้ำหนักข้อมูล (Weight) ที่เหมาะสมที่สุดในการคำนวณหากลุ่มของข้อมูล เรียกว่า Hybrid-Swarm-KM แสดงรายละเอียดดังรูป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ 3.4 โดยโครงสร้างนี้ได้ดำเนินไปประยุกต์ใช้กับขั้นตอนวิธีฉลาดแบบกลุ่ม PSO และ CSA ได้จำนวน 3 แบบจำลอง ได้แก่ แบบจำลองที่ใช้ความฉลาดแบบกลุ่มเพียง 1 ขั้นตอนวิธีร่วมกับเทคนิค K-Means คือ WCPSO-KM และ WCCSA-KM และแบบจำลองที่ผสมผสานค่าพารามิเตอร์ของแบบจำลอง PSO และ CSA ที่ดีที่สุด จำนวน 1 แบบจำลอง คือ PSO-CSA-KM

ทั้งนี้ 6 แบบจำลองที่เหลือดังตารางที่ 3.9 ได้แก่ แบบจำลอง K-Means , KM-PCA , KM-PSO, KM-PCA-PSO , KM-CSA และ KM-PCA-CSA ผู้วิจัยได้พัฒนาขึ้นเพื่อต้องการแบบจำลอง Baseline ในการเปรียบเทียบผลการทำงานและพิสูจน์สมมติฐานของแบบจำลอง Hybrid จะมีการทำงานที่ดีกว่าขั้นตอนวิธีแบบปกติ

ภาพการทำงานของแบบจำลอง Hybrid-Swarm-KM



รูปที่ 3.6 โครงสร้างการทำงานการจัดกลุ่มข้อมูลของแบบจำลอง Hybrid-Swarm-KM

จากรูปที่ 3.6 แสดงโครงสร้างการทำงานของแบบจำลอง Hybrid-Swarm-KM ซึ่งเป็นแบบจำลองแบบไฮบริดที่ประยุกต์ขั้นตอนวิธีจำนวน 3 ขั้นตอนวิธี ได้แก่ เทคนิคการจัดกลุ่ม K-Means ขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization : PSO) และขั้นตอนวิธีหาค่าเหมาะสมที่สุดแบบนกดุเหว่า (Cuckoo search) ในการจัดกลุ่มข้อมูลพฤติกรรมการรับประทานอาหารของผู้ใช้งาน โดยมีขั้นตอนการทำงาน ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ดำเนินการอ่านข้อมูลที่จัดเตรียมดังตาราง 3.8 สำหรับการจัดกลุ่มโดยใช้ข้อมูลแบบปกติ จำนวน 27 รายการ และข้อมูลที่มีการลดมิติข้อมูลด้วย PCA จำนวน 8 รายการ

2. ดำเนินการเริ่มต้นสร้างอนุภาค (Particle) สำหรับแบบจำลองที่ใช้ PSO หรือสร้างรังไข่นกกดเหว่า (Cukcoo) สำหรับแบบจำลองที่ใช้ CSA

3. ดำเนินการสุ่มข้อมูลและกำหนดเป็นค่าน้ำหนักข้อมูลที่ดีที่สุด (Pbest weight) และค่าจุดศูนย์กลาง (random centroids) เพื่อต้องการหาค่าน้ำหนักข้อมูลที่ดีที่สุดก่อน จึงจะหาค่าจุดศูนย์กลางที่ดีที่สุดต่อไป

4. ดำเนินการสร้างแบบจำลองการจัดกลุ่มด้วยเทคนิค K-Means เพื่อเรียนรู้และปรับค่าน้ำหนักข้อมูลและค่าจุดศูนย์กลาง และเป็นแบบจำลองในการปรับจูนในขั้นตอนต่อไป

5. ดำเนินการทำนายกลุ่มของข้อมูลโดยใช้ค่าน้ำหนักข้อมูลและค่าจุดศูนย์กลางจากข้อ 4

6. หลังจากได้ผลลัพธ์การจัดกลุ่มข้อมูลแล้ว จึงดำเนินการคำนวณค่า Fitness score ผ่าน Fitness function โดยงานวิจัยนี้ได้เลือกวิธีการคำนวณค่าเฉลี่ยของระยะห่างแต่ละจุดภายในกลุ่มเดียวกัน (Mean within-cluster distance) ดังสูตร

$$fitness\ score = \frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{n}|x_i - \mu_j|^2 \qquad (3.4)$$

7. ดำเนินการเปรียบเทียบค่า fitness score และกำหนดค่าน้ำหนักข้อมูลที่ดีที่สุดใหม่ (gbest weight) หากค่า score ในรอบปัจจุบันมีผลดีกว่าค่า score ในรอบก่อนหน้า

8. ตรวจสอบเงื่อนไขสำหรับจบการทำงานการหาค่าน้ำหนักข้อมูลที่ดีที่สุด โดยในงานวิจัยนี้กำหนดเงื่อนไขเป็นจำนวนรอบการทำงาน

9. ดำเนินการปรับค่าน้ำหนักข้อมูลในรอบปัจจุบันและค่าพารามิเตอร์ที่กำหนดในขั้นตอนวิธี PSO และ CSA เพื่อหาโอกาสในการเพิ่มประสิทธิภาพการทำงาน หลังจากนั้นจะย้อนกลับไปทำในขั้นตอนที่ 5 – 8 ก่อนจะจบการทำงานหากพบว่าเงื่อนไขเป็นไปตามที่กำหนดไว้ ภายหลังจบการทำงานจะได้ค่าน้ำหนักข้อมูลที่ดีที่สุด (Best Weight) แล้ว และจะดำเนินการหาค่าจุดศูนย์กลางที่ดีที่สุดในขั้นตอนต่อไป

10.ดำเนินการกำหนดค่าจุดศูนย์กลางและค่าน้ำหนักข้อมูลที่ดีที่สุดใหม่ หาก score รอบปัจจุบันมีผลดีกว่าทุกรอบก่อนหน้า และตรวจสอบเงื่อนไขจบการทำงาน

11.ดำเนินการปรับค่าจุดศูนย์กลางในรอบปัจจุบันและค่าพารามิเตอร์ที่กำหนดในขั้นตอนวิธี PSO และ CSA เพื่อหาโอกาสในการเพิ่มประสิทธิภาพการทำงาน

12.หากค่า score ปัจจุบันได้ผลลัพธ์ที่ต่ำกว่า score ทุกรอบที่ผ่านมา ดำเนินการปรับค่าน้ำหนักข้อมูล

13.ดำเนินการทำนายกลุ่มของข้อมูลโดยใช้ค่าน้ำหนักข้อมูลและค่าจุดศูนย์กลางจากข้อ 11 และ 12 (หากมีการปรับ)

14.หลังจากได้ผลลัพธ์การจัดกลุ่มข้อมูลแล้ว จึงดำเนินการคำนวณค่า Fitness score ผ่าน Fitness function จากนั้นเปรียบเทียบค่า fitness score และกำหนดค่าน้ำหนักข้อมูลที่ดีที่สุดใหม่ (Gbest weight) หากค่า score ในรอบปัจจุบันมีผลดีกว่าค่า score ในรอบก่อนหน้า หาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตรวจสอบเงื่อนไขจบการทำงานยังไม่สิ้นสุด จะย้อนกลับไปทำในขั้นตอนที่ 10-14 จะได้ค่าจุดศูนย์กลาง (Best centroids) น้ำหนักข้อมูลที่เหมาะสมที่สุด (Best Weight) กับชุดข้อมูลนั้น

15.ดำเนินการทำนายกลุ่มของข้อมูลโดยใช้ค่าน้ำหนักข้อมูลและค่าจุดศูนย์กลางที่ดีที่สุดจากการค้นหาโดยใช้แบบจำลอง Hybrid-PSO-CSA-KM

ในการพัฒนาแบบจำลองการแนะนำและประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่มร่วมกันนั้น งานวิจัยดำเนินการกำหนดค่าพารามิเตอร์ของขั้นตอนวิธี CSA และ PSO โดยอ้างอิงจากงานวิจัยที่เกี่ยวข้อง ซึ่งมีรายละเอียดของขั้นตอนวิธี Cuckoo Search Algorithm (K. G. Dhal, et al., 2018) ดังตารางที่ 3.11 และขั้นตอนวิธี Particle Swarm Optimization (A. L. Ballardini, 2018 และ S.U. Mane, et al., 2014) ดังตารางที่ 3.12

ตารางที่ 3.11 ค่าพารามิเตอร์เริ่มต้นของขั้นตอนวิธี Cuckoo Search Algorithm

พารามิเตอร์	ค่าพารามิเตอร์	
	Centroids	Weight
Population Size	{10, 15, 20, ..., 50}	{10, 15, 20, ..., 50}
Dimension	(4, 27)	27
Lambda	1.5	1.5
Probability (Pa)	0.2	0.2
Step Size	0.01	0.01
Min-Max bound	[0, 1]	[0, 1]
Iteration	300	300

ตารางที่ 3.12 ค่าพารามิเตอร์เริ่มต้นของขั้นตอนวิธี Particle Swarm Optimization

พารามิเตอร์	ค่าพารามิเตอร์	
	Centroids	Weight
Population Size	{3, 6, 9, ..., 30}	{3, 6, 9, ..., 30}
Dimension	(4, 27)	27
Inertial Weight	0.72	0.72
C1	1.49	1.49
C2	1.49	1.49
Iteration	600	600

3.5.2 การทดลองการทำงานแบบจำลอง Hybrid-Swarm-KM ด้วยพารามิเตอร์ที่ดีที่สุด

ในขั้นตอนนี้เป็นการศึกษาและพัฒนาแบบจำลองด้วยพารามิเตอร์ที่แตกต่างกันของแต่ละขั้นตอนวิธีความฉลาดแบบกลุ่ม เพื่อหาค่าพารามิเตอร์ที่สามารถให้ผลลัพธ์การทำงานที่ดีที่สุด และสามารถนำค่าพารามิเตอร์ดังกล่าว ไปประยุกต์ใช้กับแบบจำลองผสมหรือ PSO-CSA-KM เพื่อทดสอบว่าสามารถให้ผลลัพธ์ที่ดีที่สุดได้หรือไม่ ในงานวิจัยนี้สนใจพารามิเตอร์จำนวนประชากร หรือ population size ว่าส่งผลต่อการค้นหาคำตอบหรือจัดกลุ่มและแนะนำรายการอาหารได้ถูกต้องตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มเป้าหมายหรือไม่ ค่าพารามิเตอร์ของขั้นตอนวิธีความฉลาดแบบกลุ่มทั้ง 2 ขั้นตอนวิธี แสดงดังตารางที่ 3.13

ตารางที่ 3.13 ค่าพารามิเตอร์ของการทดลองขั้นตอนวิธีความฉลาดแบบกลุ่ม

Param / ขั้นตอนวิธี	1	2	3	4	5	6	7	8	9	10
PSO	3	6	9	12	15	18	21	24	27	30
CSA	10	15	20	25	30	35	40	45	50	-

3.6 การคำนวณหาชุดรายการอาหารที่เหมาะสม (Food Optimization)

ในการคำนวณหาชุดรายการอาหารที่เหมาะสมรายบุคคล งานวิจัยนี้ดำเนินการคำนวณตามค่าของพลังงานที่จำเป็นต้องใช้ทั้งหมดในแต่ละวัน (Total Daily Energy Expenditure : TDEE) ซึ่งคำนวณจากการใช้พลังงานพื้นฐานที่จำเป็นของร่างกาย ภายใน 24 ชม. (BMR) และการออกกำลังหายต่อสัปดาห์ โดยมีรายละเอียดการคำนวณ ดังนี้

- **BMR** (Basal Metabolic Rate) : การใช้พลังงานพื้นฐานที่จำเป็นของร่างกายภายใน 24 ชม.

  สำหรับผู้ชาย

  BMR = 66 + (13.7 x น้ำหนักตัวเป็น กก.) + (5 x ส่วนสูงเป็น ซม.) – (6.8 x อายุ)

  สำหรับผู้หญิง

  BMR = 665 + (9.6 x น้ำหนักตัวเป็น กก.) + (1.8 x ส่วนสูงเป็น ซม.) – (4.7 x อายุ)

- **TDEE** (Total Daily Energy Expenditure) : ค่าของพลังงานที่ใช้ทั้งหมดในแต่ละวัน เมื่อมีการทำกิจกรรมต่างๆ

  ออกกำลังกายน้อยมากหรือไม่ออกเลย : TDEE = 1.2 x BMR

  ออกกำลังกาย 1-3 ครั้งต่อสัปดาห์ : TDEE = 1.375 x BMR

  ออกกำลังกาย 4-5 ครั้งต่อสัปดาห์ : TDEE = 1.55 x BMR

  ออกกำลังกาย 6-7 ครั้งต่อสัปดาห์ : TDEE = 1.7 x BMR

  ออกกำลังกายวันละ 2 ครั้งขึ้นไป : TDEE = 1.9 x BMR

เมื่อคำนวณค่าของพลังงานที่ใช้ทั้งหมดในแต่ละวันแล้ว จึงนำค่าดังกล่าวมาคำนวณสัดส่วนพลังงาน/สารอาหารขั้นต่ำใน 1 วัน ของคาร์โบไฮเดรต, โปรตีน และไขมัน เป็น 50-20-30 ตามลำดับ ด้วยสูตรด้านล่าง เพื่อนำไปคำนวณหารายการอาหารที่มีคุณค่าทางโภชนาการขั้นต่ำรายบุคคลต่อไป

คาร์โบไฮเดรต : (TDEE * 0.5) / 4 = Carb

โปรตีน : (TDEE * 0.2) / 4 = Protein

ไขมัน : (TDEE * 0.3) / 9 = Fat *หน่วยเป็นกรัม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการคำนวณหาชุดรายการอาหารที่เหมาะสม เป็นขั้นตอนที่ใช้ในการจัดรายการที่มีความเหมาะสมกับความต้องการพลังงานขั้นพื้นฐานของผู้ใช้งานรายบุคคล (TDEE) ขึ้นอยู่กับข้อมูลอายุ น้ำหนัก ส่วนสูง และข้อมูลกิจกรรมการออกกำลังกายส่วนบุคคล ในงานวิจัยนี้ใช้เทคนิคการโปรแกรมเชิงเส้น (Linear Programming) ในการหาค่าพลังงานของรายการอาหารรวมมากที่สุดที่ต้องประกอบด้วยคุณค่าทางโภชนาการขั้นต่ำตามการคำนวณ ได้แก่ คาร์โบไฮเดรต โปรตีน และไขมัน ซึ่งถือเป็นการหาค่ามากที่สุด (Maximize Objective) โดยอยู่ภายใต้เงื่อนไข (Constrains) จำนวน 5 ค่า ได้แก่

1. ค่าผลรวมพลังงานจากรายการอาหารทั้งหมดต้องไม่มากกว่าค่าของพลังงานที่ต้องการใช้ทั้งหมดในแต่ละวัน (TDEE) รายบุคคล
2. ค่าโปรตีนต้องไม่น้อยกว่าค่าปริมาณโปรตีนขั้นต่ำที่กำหนด
3. ค่าไขมันต้องไม่มากกว่าค่าปริมาณไขมันที่กำหนด
4. ค่าคาร์โบไฮเดรตต้องไม่มากกว่าค่าปริมาณคาร์โบไฮเดรตที่กำหนด
5. รายการอาหารที่แนะนำในชุดอาหารต้องไม่เกิน 4 รายการอาหาร

สูตรการใช้การโปรแกรมเชิงเส้น (Linear Programming)

Decision variable : [TDEE (Calories), Carb (x), Protein (y), Fat (z), Menu (n), Menu selected (d)]

Objective function : “*Maximize*” .... sum(calories) <= TDEE

Constrains : $sum_n(x)$ <= Carb, $sum_n(z)$ <= Fat , $sum_n(d)$ <= $Menu\ selected_3$

: $sum_n(y)$ >= Protein

: sum(n) = $Menu_4$

รูปที่ 3.7 ตัวอย่างข้อมูลรายบุคคลและรายการแนะนำอาหารชุดที่เหมาะสม

จากภาพที่ 3.7 แสดงตัวอย่างการคำนวณหาชุดรายการอาหารที่เหมาะสมรายบุคคล โดยใช้เทคนิคการโปรแกรมเชิงเส้น (Linear Programming) ในการหาชุดรายการอาหารที่ 1 ชุดรายการอาหารประกอบไปด้วยรายการอาหารจำนวนไม่เกิน 4 รายการ จากรายการแนะนำอาหารจำนวน 10 รายการที่ได้จากทำงานในขั้นตอน 3.3.5 การจัดอันดับรายการแนะนำที่มีผลรวมของค่าพลังงานมากที่สุดและไม่เกินความต้องการพลังงานขั้นพื้นฐานของผู้ใช้งานรายบุคคล (TDEE) โดยรายการอาหาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต้องมีผลรวมประกอบด้วยคุณค่าทางโภชนาการได้แก่ คาร์โบไฮเดรต โปรตีน และไขมัน ไม่ขั้นต่ำกว่าเกณฑ์ที่ต้องการรายบุคคล

3.7 การแนะนำรายการอาหาร (Food Recommendation)

ในงานวิจัยนี้ ขั้นตอนการแนะนำชุดรายการอาหารไทย ดำเนินการแนะนำชุดรายการอาหารจากรายการอาหารแนะนำตามผู้ใช้งาน จำนวน 5 ชุด และในแต่ละชุดรายการอาหารต้องมีรายการอาหารที่ซ้ำกันไม่เกิน 3 รายการอาหาร โดยเรียงข้อมูลจากชุดที่มีผลรวมของพลังงานมากที่สุด ยกตัวอย่างเช่น

ข้อมูลผู้ใช้งาน นายประชา ใจดี เพศชาย อายุ 45 ปี ส่วนสูง 161 เซนติเมตร น้ำหนัก 57 กิโลกรัม ค่าดัชนีมวลกาย (BMI) 21.99 ค่า ค่า TDEE เท่ากับ 1,350.15

ดังนั้น ต้องการพลังงานแคลอรี่ไม่เกิน 1350.15 กิโลแคลอรี่/วัน โปรตีนไม่น้อยกว่า 67 กรัม ไขมัน ไม่มากกว่า 45 กรัม และคาร์โบไฮเรตร ไม่มากกว่า 168 กรัม

ผลการแนะนำชุดรายการอาหารไทย (ยกตัวอย่างจำนวน 2 ชุดรายการอาหาร)

ชุดรายการอาหารชุดที่ 1 :

รายการอาหาร ได้แก่ ข้าวเหนียว (222.2), เกาเหลาหมูตุ๋น (202.6), เกาเหลาลูกชิ้น (155.5) และ ข้าวไก่ย่าง (759.6)

รวมทั้งสิ้น 1339.9 กิโลแคลอรี่/วัน (โปรตีน 71 ก. ไขมัน 45 ก. และคาร์โบไฮเดรต 162 ก.)

ชุดรายการอาหารชุดที่ 2 :

รายการอาหาร ได้แก่ หมูทอด(สะโพก,สันนอก) (315), ข้าวเหนียว (222.2), เกาเหลาหมูตุ๋น (202.6) และข้าวมันไก่ทอด (599.9)

รวมทั้งสิ้น 1339.7 กิโลแคลอรี่/วัน (โปรตีน 94 ก. ไขมัน 41 ก. และคาร์โบไฮเดรต 149 ก.) เป็นต้น และมีการดำเนินการแนะนำชุดรายการอาหารอีก 3 ชุดด้วยวิธีการเช่นเดียวกันไปยังผู้ใช้งาน

3.8 การประเมินความพึงพอใจต่อชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

งานวิจัยนี้ดำเนินการศึกษาและพัฒนาระบบแนะนำรายการอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล โดยจัดเก็บและประมวลผลข้อมูลส่วนบุคคล สุขภาพร่างกายและพฤติกรรมการทานอาหารและการออกกำลังราย ซึ่งจะนำรายการอาหารแนะนำ Top15 มาจัดทำชุดรายการอาหารจำนวน 5 ชุดต่อบุคคลด้วยวิธีการโปรแกรมเชิงเส้น หรือ Linear Programming ตามค่าของพลังงานที่จำเป็นต้องใช้ทั้งหมดในแต่ละวัน (Total Daily Energy Expenditure : TDEE) และคุณค่าทางโภชนาการขั้นต่ำตาม 3.6 และ 3.7 จากนั้นดำเนินการสำรวจความคิดเห็นและความพึงพอใจต่อชุดรายการอาหารแนะนำดังกล่าวด้วยแบบสอบถาม ผู้ตอบแบบสอบถามจะประเมินชุดรายการอาหารแต่ละชุดและให้คะแนนความพึงพอใจรายชุด ซึ่งงานวิจัยนี้จะนำคะแนนทั้งหมดคำนวณหาค่าเฉลี่ยความพึงพอใจและเทียบความพึงพอใจภาพรวม โดยมีเกณฑ์การวัดผลดังตารางที่ 3.14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.14 คะแนนความพึงพอใจเฉลี่ยของแบบสอบถามความพึงพอใจชุดรายการอาหารแนะนำ

คะแนนความพึงพอใจเฉลี่ย	ความพึงพอใจ
4.51 – 5.00	ความพึงพอใจมากที่สุด
4.01 – 4.50	ความพึงพอใจมาก
3.01 – 4.00	ความพึงพอใจปานกลาง
2.51 – 3.00	ความพึงพอใจเล็กน้อย
0.00 – 2.50	ไม่พึงพอใจ

3.9 เครื่องมือสำหรับงานวิจัย

การทดลองในงานวิจัยนี้ได้นำข้อมูลจากแอปพลิเคชัน FoodiEat ของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ โดยมีการจัดเก็บไว้ในฐานข้อมูล MySQL และนำข้อมูลมาทดลองแบ่งออกเป็น 3 ขั้นตอนได้แก่ ขั้นตอนการทำความสะอาดข้อมูล ขั้นตอนการรวมข้อมูล และขั้นตอนการดัดแปลงข้อมูล สำหรับการพัฒนาแบบจำลองทั้ง 3 ส่วน คือ การจัดกลุ่มพฤติกรรมการรับประทานอาหารของผู้ใช้งาน การทำนายค่าคะแนนและการจัดลำดับ Top-N การจัดรายการอาหารตามความเหมาะสมส่วนบุคคล จะถูกพัฒนาจากการเขียนโปรแกรมโดยภาษา Python บนเครื่องมือ Jupyter Notebook โดยผลลัพธ์ในการทำงานจะถูกจัดเก็บเป็นไฟล์รูปแบบ csv และ text

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการวิจัยและการอภิปรายผล

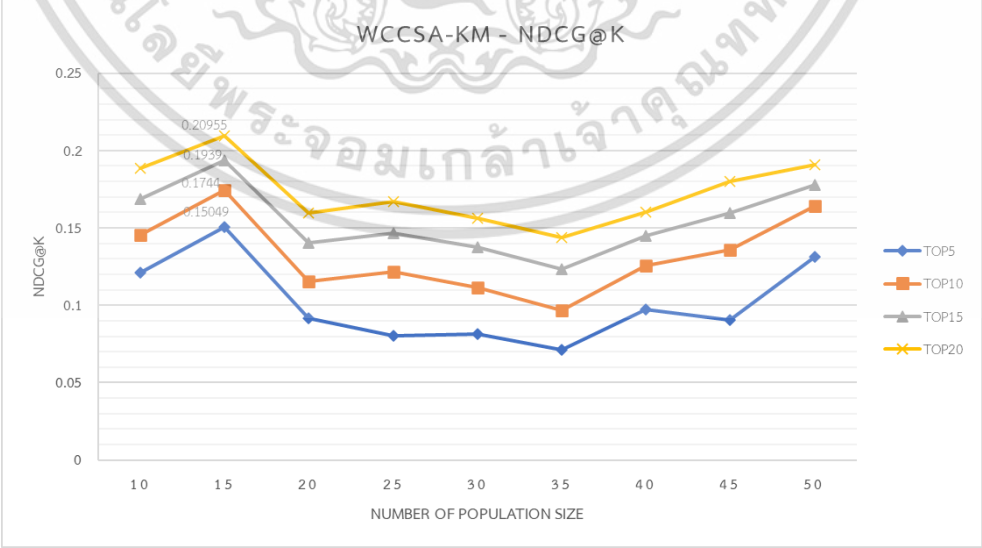
งานวิจัยนี้ได้ดำเนินการทดลองของการผสมผสานการเรียนรู้ของเครื่องทางด้านระบบแนะนำและขั้นตอนวิธีความฉลาดแบบกลุ่มเพื่อแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมและสำรวจความพึงพอใจของชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนากาการอย่างเหมาะสมรายบุคคล โดยมีผลการวิจัยและการอภิปรายดังนี้

4.1 ผลการวิจัย

ในงานวิจัยนี้ได้ดำเนินการทดลองและพัฒนาแบบจำลองสำหรับการแนะนำรายการอาหารตามที่กล่าวมาเพื่อค้นหาค่าพารามิเตอร์ของขั้นตอนวิธีความฉลาดแบบกลุ่มและเปรียบเทียบประสิทธิภาพของแบบจำลองบนชุดข้อมูลพฤติกรรมการรับประทานอาหาร โดยทดลองแบบจำลองในกลุ่ม Hybrid-Swarm-KM ได้แก่ WCCSA-KM, WCPSO-KM และ PSO-CSA-KM กับข้อมูลพฤติกรรมเชิงลึกของการรับประทานอาหารและการออกกำลังกายรายบุคคล ซึ่งข้อมูลอยู่ในรูปแบบ implicit data การดำเนินงานประกอบด้วยการประมวลผลแบบจำลองทั้งหมด 3 ครั้ง และแสดงผลประสิทธิภาพแบบจำลองด้วยค่าเฉลี่ย ผลการทดลองค่าพารามิเตอร์ของขั้นตอนวิธีและการเปรียบเทียบแบบจำลองมีรายละเอียดดังต่อไปนี้

4.1.1 การประเมินประสิทธิภาพแบบจำลองแนะนำรายการอาหาร

4.1.1.1 การเปรียบเทียบประสิทธิภาพของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ของพารามิเตอร์จำนวนประชากร



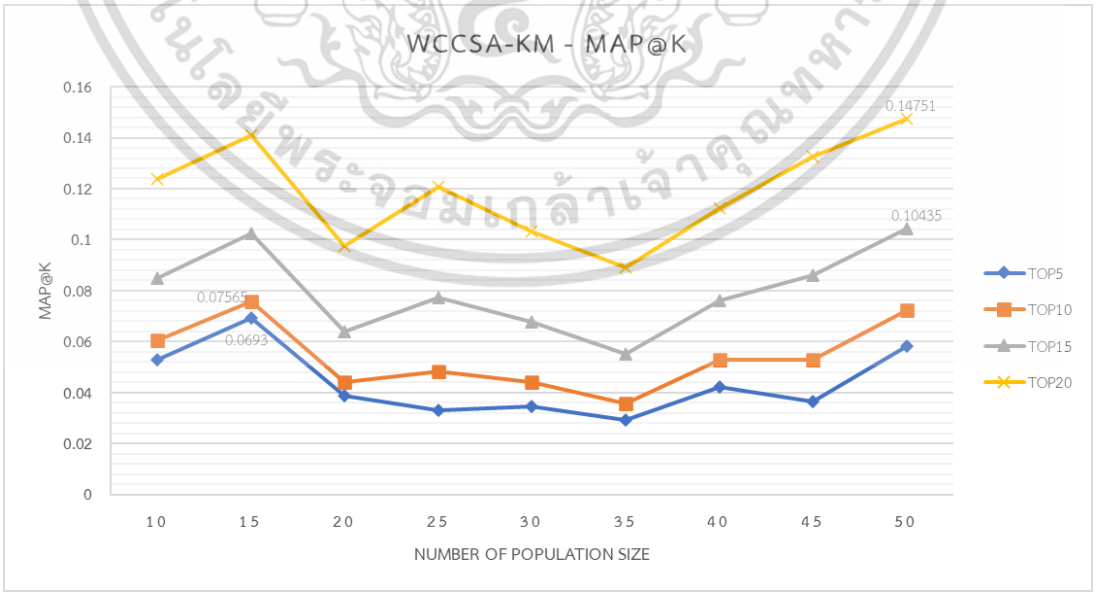
รูปที่ 4.1 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร

Population size	NDCG@K			
	Top5	Top10	Top15	Top20
10	0.12133	0.14555	0.16853	0.18850
15	0.15049	0.17440	0.19390	0.20955
20	0.09152	0.11521	0.14019	0.15971
25	0.08056	0.12166	0.14662	0.16674
30	0.08159	0.11120	0.13753	0.15634
35	0.07107	0.09684	0.12320	0.14353
40	0.09714	0.12557	0.14488	0.16041
45	0.09033	0.13588	0.15974	0.18005
50	0.13122	0.16431	0.17765	0.19077

จากรูปที่ 4.1 และตารางที่ 4.1 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี CSA ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCCSA-KM ซึ่งทดลองจำนวนประชากรทั้งหมด 9 ชุด คือ 10, 15, 20, 25, 30, 35, 40, 45 และ 50 ด้วยวิธีการประเมินผลแบบ NDCG@K พบว่า การกำหนดจำนวน population size หรือจำนวนรังที่วางไข่นกกุเหว่า เท่ากับ 15 จะแสดงผลการแนะนำรายการอาหารได้ถูกต้องตามความสนใจมากที่สุด ครอบคลุมทั้ง Top5, Top10, Top15 และ Top20 โดยห่างจากจำนวนประชากรที่ให้ผลลัพธ์ถูกต้องอันดับที่ 2 อยู่ 0.02-0.03 ซึ่งมีจำนวนประชากรเท่ากับ 50 และ 10 ให้ประสิทธิภาพดีตามลำดับสอดคล้องกันทุก TopN



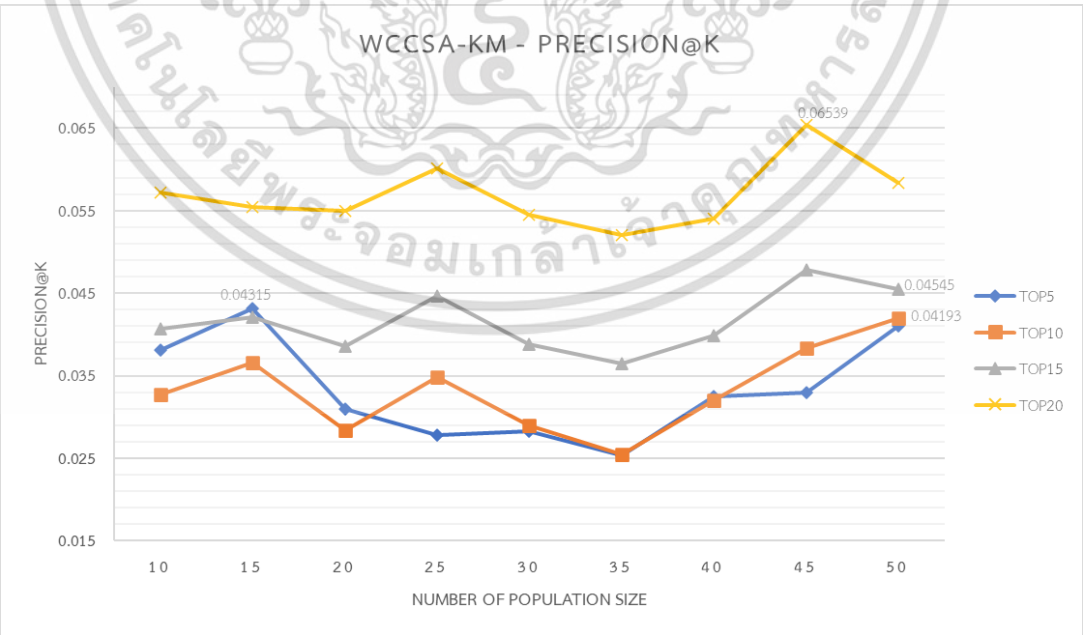
รูปที่ 4.2 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร

Population size	MAP@K			
	Top5	Top10	Top15	Top20
10	0.05300	0.06033	0.08488	0.12370
15	0.06930	0.07565	0.10232	0.14114
20	0.03864	0.04408	0.06389	0.09765
25	0.03300	0.04825	0.07735	0.12065
30	0.03465	0.04419	0.06760	0.10309
35	0.02938	0.03573	0.05522	0.08903
40	0.04228	0.05271	0.07599	0.11234
45	0.03629	0.05305	0.08603	0.13240
50	0.05812	0.07241	0.10435	0.14751

จากรูปที่ 4.2 และตารางที่ 4.2 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี CSA ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCCSA-KM จากวิธีการประเมินผลแบบ MAP@K พบว่าจำนวนประชากรเท่ากับ 15 ให้ผลลัพธ์ถูกต้องมากที่สุดใน Top5 และ Top10 เท่านั้น สำหรับ Top15 และ Top20 จำนวนประชากรเท่ากับ 50 มีผลลัพธ์ดีที่สุด ตามด้วยจำนวนประชากรเท่ากับ 15 ดังนั้น กรณีมีจำนวนรายการแนะนำน้อย จำนวนประชากร 15 สามารถให้ผลลัพธ์ที่ถูกจัดอันดับอย่างถูกต้องกว่าจำนวนประชากรอื่น ๆ โดยที่จำนวนประชากร 50 แสดงผลลัพธ์ที่ดีที่สุดกรณีต้องการจำนวนรายการแนะนำมาก



รูปที่ 4.3 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี CSA ตามจำนวนประชากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA ตามพารามิเตอร์จำนวนประชากร

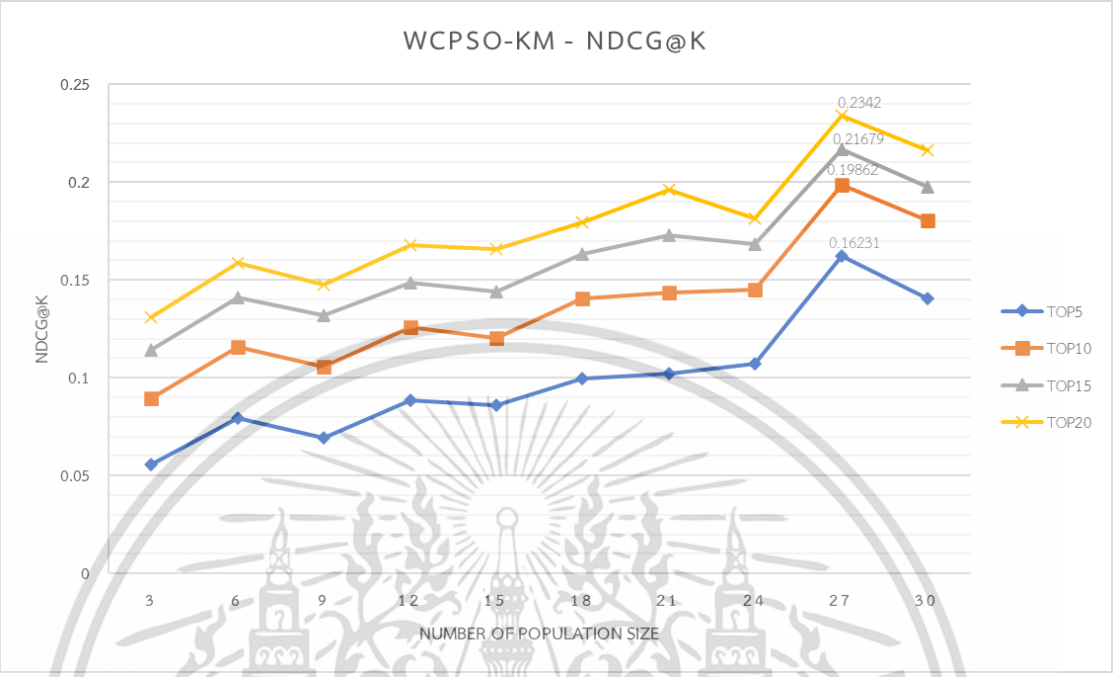
Population size	Precision@K			
	Top5	Top10	Top15	Top20
10	0.03815	0.03267	0.04073	0.05717
15	0.04315	0.03654	0.04209	0.05541
20	0.03095	0.02838	0.03861	0.05491
25	0.02778	0.03479	0.04467	0.06017
30	0.02829	0.02902	0.03877	0.05454
35	0.02536	0.02545	0.03642	0.05204
40	0.03253	0.03206	0.03983	0.05398
45	0.03297	0.03828	0.04780	0.06539
50	0.04107	0.04193	0.04545	0.05836

จากรูปที่ 4.3 และตารางที่ 4.3 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี CSA ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCCSA-KM วิธีการประเมินผลแบบ Precision@K พบว่ากรณีมีจำนวนรายการแนะนำน้อย จำนวนประชากรเท่ากับ 15 สามารถรวบรวมรายการอาหารที่ถูกต้องได้มากที่สุด กรณีแนะนำรายการอาหาร Top10, Top15 จำนวนประชากรเท่ากับ 50 สามารถรวมรวบและแนะนำรายการอาหารได้ดี และจำนวนประชากรเท่ากับ 45 จะแสดงผลลัพธ์ได้มากที่สุดกรณีแนะนำรายการอาหาร 20 รายการ

จากการทดลองเปรียบเทียบประสิทธิภาพของแบบจำลองที่พัฒนาด้วยขั้นตอนวิธี CSA ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCCSA-KM พบว่าจำนวนประชากรเท่ากับ 15 สามารถแสดงผลลัพธ์ที่ตรงต่อความสนใจและถูกต้องมากที่สุดในทุก TopN ด้วยค่า NDCG@K ซึ่งมีค่ามากที่สุดเท่ากับ 0.20955 ใน Top20 ตามด้วย 0.19390, 0.17440, 0.15049 ของ Top15, Top10 และ Top 5 ตามลำดับ นอกจากนี้สามารถจัดลำดับและรวบรวมรายการแนะนำกรณีจำนวน รายการน้อยได้ดี จำนวนประชากรเท่ากับ 15 มีประสิทธิภาพในการแนะนำจำนวนมากได้ดีรองลงมาจากจำนวนประชากรเท่ากับ 50 ซึ่งมีค่า MAP@15 น้อยกว่าเพียง 0.002 และ MAP@20 น้อยกว่า 0.006 ดังนั้นจำนวนประชากรเท่ากับ 15 สามารถแสดงรายการแนะนำได้ดีทุกรายการ TopN ทั้งนี้ เนื่องจากการกำหนดพารามิเตอร์จำนวนประชากรที่แตกต่างกันมีผลต่อเวลาในการประมวลผลการทดลองเพิ่มจำนวนประชากรผลลัพธ์มีแนวโน้มประสิทธิภาพที่ลดลงและเพิ่มขึ้น แต่ขณะเดียวกันการประมวลผลมีเวลาที่มากขึ้นเช่นกัน ดังนั้นผู้วิจัยเล็งเห็นว่าสิ้นสุดการทดลองจำนวนประชากรเท่ากับ 50 ซึ่งได้ผลลัพธ์จำนวนประชากรเท่ากับ 15 ที่มีผลลัพธ์ดีที่สุดแล้วในช่วงจำนวนประชากรตั้งแต่ 10 ถึง 50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.2 การเปรียบเทียบประสิทธิภาพของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ของพารามิเตอร์จำนวนประชากร



รูปที่ 4.4 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร

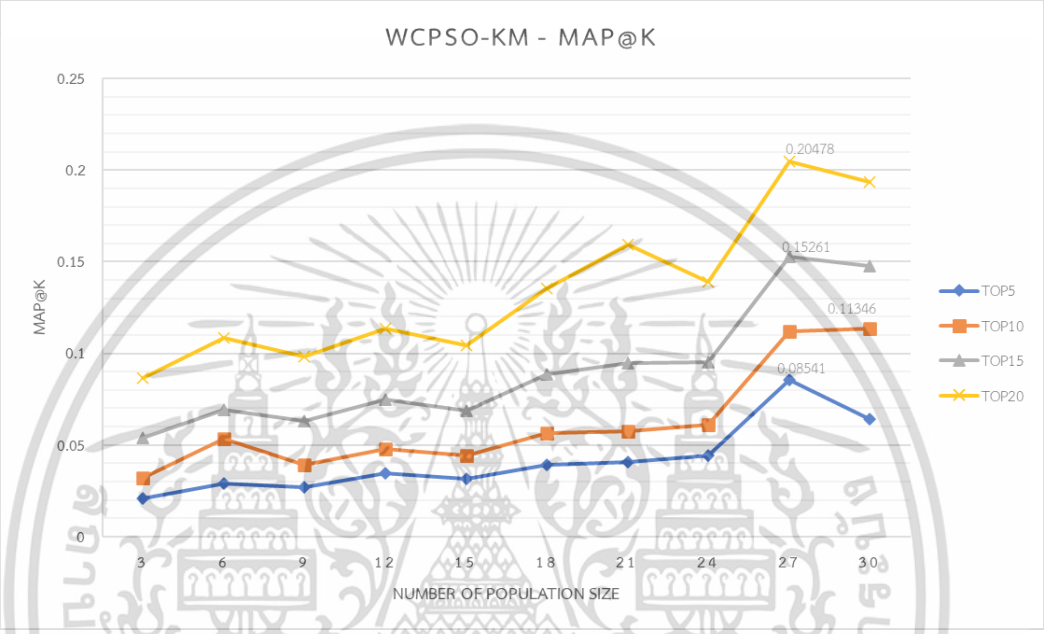
ตารางที่ 4.4 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร

Population size	NDCG@K			
	Top5	Top10	Top15	Top20
3	0.05555	0.08931	0.11440	0.13114
6	0.07951	0.11580	0.14081	0.15892
9	0.06904	0.10584	0.13195	0.14757
12	0.08871	0.12576	0.14871	0.16767
15	0.08606	0.12043	0.14416	0.16590
18	0.09978	0.14029	0.16342	0.17944
21	0.10205	0.14343	0.17267	0.19607
24	0.10700	0.14512	0.16826	0.18141
27	0.16231	0.19862	0.21679	0.23420
30	0.14033	0.18065	0.19749	0.21637

จากรูปที่ 4.4 และตารางที่ 4.4 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี PSO ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCPSO-KM ดำเนินการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทดลองจำนวนประชากรทั้งหมด 10 ชุด คือ 3, 6, 9, 12, 15, 18, 21, 24, 27 และ 30 การแสดงผลด้วยค่าเฉลี่ย NCDG@K พบว่าจำนวนประชากรหรือจำนวน particle เท่ากับ 27 แสดงผลลัพธ์รายการแนะนำได้ถูกต้องและตรงต่อความสนใจมากที่สุดในทุก TopN ได้แก่ NDCG@5 = 0.16231, NDCG@10 = 0.19862, NDCG@15 = 0.21679 และ NDCG@20 = 0.23420 โดยมีผลประเมินมากกว่าอันดับ 2 อยู่ระหว่าง 0.01-0.02 ซึ่งมีจำนวนประชากรเท่ากับ 30, 21, 24 และ 18 แสดงผลลัพธ์ออกมาได้ดีตามลำดับ



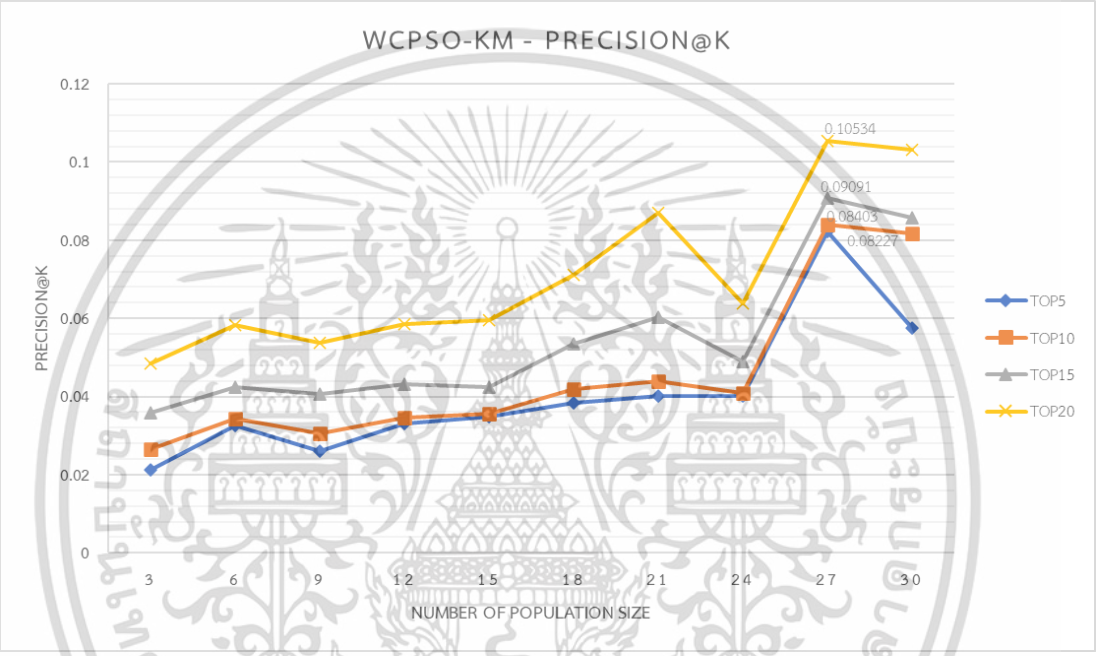
รูปที่ 4.5 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร

ตารางที่ 4.5 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร

Population size	MAP@K			
	Top5	Top10	Top15	Top20
3	0.02075	0.03180	0.05387	0.08673
6	0.02890	0.05341	0.06909	0.10858
9	0.02670	0.03929	0.06316	0.09818
12	0.03464	0.04777	0.07469	0.11365
15	0.03130	0.04431	0.06875	0.10458
18	0.03908	0.05639	0.08877	0.13566
21	0.04047	0.05776	0.09480	0.15930
24	0.04432	0.06128	0.09501	0.13894
27	0.08541	0.11227	0.15261	0.20478
30	0.06415	0.11346	0.14754	0.19337

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.5 และตารางที่ 4.5 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี PSO ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCPSO-KM วิธีการประเมินผลแบบ MAP@K พบว่าจำนวนประชากรเท่ากับ 27 แสดงผลลัพธ์ภาพรวมการจัดลำดับรายการแนะนำได้ถูกต้องมากที่สุดใน Top5, Top15 และ Top20 เท่ากับ 0.08541, 0.15261 และ 0.20478 ตามลำดับ สำหรับ Top10 จำนวนประชากรเท่ากับ 27 และ 30 แสดงผลลัพธ์ได้อย่างถูกต้องใกล้เคียงกัน โดยมีค่า MAP@10 มากกว่าเพียง 0.001 เท่านั้น การประเมินผลแบบจัดลำดับรายการแนะนำ จำนวนประชากรเท่ากับ 30, 27, 21, 24 และ 18 สามารถจัดลำดับและแสดงผลลัพธ์ได้ถูกต้องมากที่สุด ตามลำดับ



รูปที่ 4.6 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี PSO ตามจำนวนประชากร

จากรูปที่ 4.6 และตารางที่ 4.6 เป็นการเปรียบเทียบจำนวนประชากร (population size) ของขั้นตอนวิธี PSO ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCPSO-KM วิธีการประเมินผลแบบ Precision@K พบว่าจำนวนประชากรเท่ากับ 27 สามารถรวบรวมข้อมูลและแสดงผลลัพธ์ได้มากที่สุดในทุกจำนวนรายการแนะนำ ได้แก่ Top5 Precision@K = 0.08227 , Top10 Precision@k = 0.08403 , Top15 Precision@K = 0.09091 และ Top20 Precision@K = 0.10534 โดยจำนวนประกรเท่ากับ 30 , 21 , 18 และ 24 สามารถรวบรวมและแสดงผลลัพธ์ได้ดีตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี PSO ตามพารามิเตอร์จำนวนประชากร

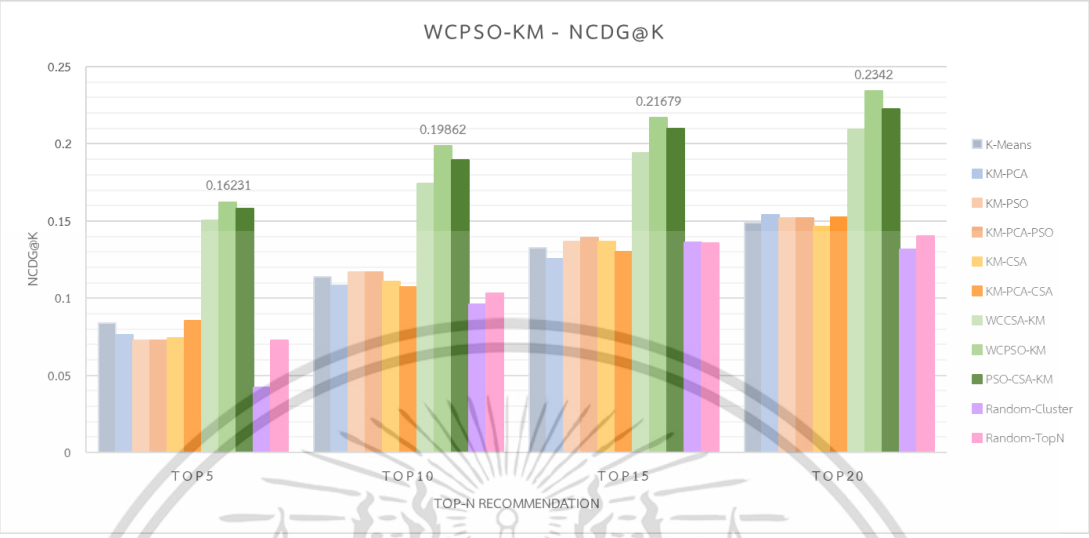
Population size	Precision@K			
	Top5	Top10	Top15	Top20
3	0.02133	0.02660	0.03588	0.04835
6	0.03249	0.03424	0.04231	0.05829
9	0.02596	0.03068	0.04076	0.05386
12	0.03302	0.03472	0.04320	0.05860
15	0.03489	0.03563	0.04232	0.05964
18	0.03831	0.04186	0.05362	0.07117
21	0.04012	0.04387	0.06023	0.08705
24	0.04008	0.04084	0.04898	0.06386
27	0.08227	0.08403	0.09091	0.10534
30	0.05741	0.08163	0.08569	0.10323

จากการทดลองเปรียบเทียบประสิทธิภาพของแบบจำลองที่พัฒนาด้วยขั้นตอนวิธี PSO ด้วยโครงสร้าง Hybrid-Swarm-KM เป็นแบบจำลอง WCPSO-KM พบว่าจำนวนประชากรเท่ากับ 27 สามารถแสดงผลลัพธ์ที่ตรงต่อความสนใจและถูกต้องมากที่สุดในทุก TopN ด้วยค่า NDCG@5 = 0.16231 , NDCG@10 = 0.19862 , NDCG@15 = 0.21679 และ NDCG@20 = 0.23420 และมีประสิทธิภาพที่ดีที่สุดสอดคล้องกันในวิธีการประเมินแบบอื่น ได้แก่ MAP@K และ Precision@K ทั้งนี้จำนวนประชากรเท่ากับ 30 , 21 , 24 และ 18 แสดงผลลัพธ์ที่ดีตามลำดับเช่นเดียวกันในทุกแบบวิธีการประเมิน ทั้งนี้ เนื่องจากการกำหนดพารามิเตอร์จำนวนประชากรที่แตกต่างกันมีผลต่อเวลาในการประมวลผล การทดลองเพิ่มจำนวนประชากรผลลัพธ์มีแนวโน้มประสิทธิภาพที่เพิ่มขึ้นและลดลง แต่ขณะเดียวกันการประมวลผลมีเวลาที่มากขึ้นเช่นกัน ซึ่งหลังจากจำนวนประชากรเท่ากับ 30 ประสิทธิภาพลดลงเรื่อย ๆ ดังนั้นผู้วิจัยเล็งเห็นว่าสิ้นสุดการทดลองและแสดงผลจำนวนประชากรเท่ากับ 30 โดยจำนวนประชากรเท่ากับ 27 มีผลลัพธ์ที่ดีที่สุดในช่วงจำนวนประชากรตั้งแต่ 6 ถึง 30

การทดลองค่าพารามิเตอร์ของขั้นตอนวิธี CSA และ PSO ด้วยโครงสร้าง Hybrid-Swarm-KM พบว่าขั้นตอนวิธี CSA มีจำนวนประชากรเท่ากับ 15 สามารถแสดงผลลัพธ์ได้มีประสิทธิภาพเหมาะสมที่สุด และขั้นตอนวิธี PSO มีจำนวนประชากรเท่ากับ 27 แสดงรายการแนะนำได้อย่างถูกต้องมากที่สุด ดังนั้นจึงนำค่าพารามิเตอร์จำนวนประชากรหรือ population size ของขั้นตอนวิธี CSA และ PSO ไปประยุกต์ใช้ในการพัฒนาแบบจำลอง PSO-CSA-KM เพื่อประเมินประสิทธิภาพต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1.3 การเปรียบเทียบประสิทธิภาพของแบบจำลองแนะนำรายการอาหารด้วยขั้นตอนวิธี CSA และ PSO ของพารามิเตอร์จำนวนประชากร



รูปที่ 4.7 กราฟเปรียบเทียบประสิทธิภาพ NDCG@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid

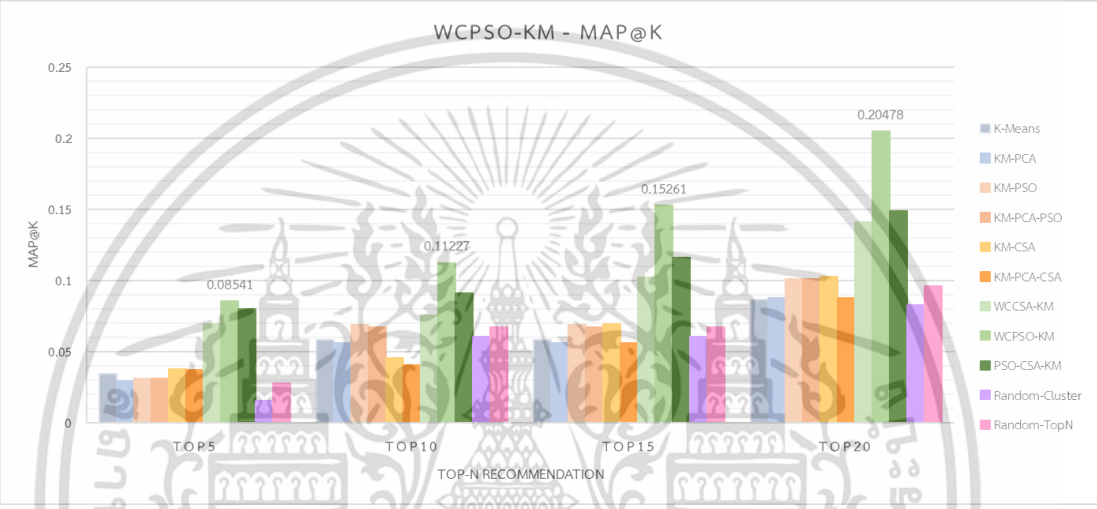
ตารางที่ 4.7 ผลการเปรียบเทียบค่าประสิทธิภาพ NDCG@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM

แบบจำลอง	NDCG@K			
	Top5	Top10	Top15	Top20
K-Means	0.08389	0.11369	0.13265	0.14880
KM-PCA	0.07652	0.10827	0.12568	0.15400
KM-PSO	0.07280	0.11688	0.13668	0.15219
KM-PCA-PSO	0.07280	0.11688	0.13914	0.15219
KM-CSA	0.07421	0.11101	0.13659	0.14620
KM-PCA-CSA	0.08539	0.10722	0.12992	0.15255
WCCSA-KM	0.15049	0.17440	0.19390	0.20955
WCPSO-KM	0.16231	0.19862	0.21679	0.23420
PSO-CSA-KM	0.15803	0.18968	0.21005	0.22246
Random-Cluster	0.04230	0.09626	0.13610	0.13147
Random-TopN	0.07282	0.10314	0.13586	0.14044

จากรูปที่ 4.7 และตารางที่ 4.7 เป็นการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งหมด 11 แบบจำลอง ประกอบด้วย กลุ่ม baseline algorithm ได้แก่ K-Means, KM-PCA กลุ่มประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่ม ได้แก่ KM-PSO, KM-PCA-PSO, KM-CSA และ KM-PCA-CSA กลุ่มสุ่มข้อมูลเพื่อเปรียบเทียบประสิทธิภาพกับการสุ่มข้อมูล ได้แก่ Random-Cluster เป็นการสุ่มกลุ่มผู้ใช้งานให้กับผู้ใช้งาน 1 คน และ Random-TopN เป็นการสุ่มรายการแนะนำตามจำนวน TopN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และกลุ่มการนำเสนอในงานวิจัยนี้ กลุ่มผสมผสานการประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่มมากกว่า 1 ส่วน ได้แก่ โครงสร้าง Hybrid-Swarm-KM คือ WCCSA-KM, WCPSO-KM และ PSO-CSA-KM จากการประเมินผลด้วย NDCG@K พบว่าแบบจำลอง WCPSO-KM ที่อยู่ในกลุ่ม Hybrid-Swarm-KM สามารถแสดงผลลัพธ์ในรายการแนะนำได้อย่างถูกต้องและตรงต่อความสนใจมากกว่าแบบจำลองอื่น ซึ่งมีประสิทธิภาพสูงที่สุดในทุกรูปแบบการแนะนำ ได้แก่ Top5 ด้วยค่า NDCG เท่ากับ 0.16231, Top10 ค่า NDCG เท่ากับ 0.19862, Top15 ค่า NDCG เท่ากับ 0.21679 และ Top20 ค่า NDCG เท่ากับ 0.23420 โดยแบบจำลองที่อยู่ในกลุ่ม Hybrid-Swarm-KM ได้แก่ PSO-CSA-KM, WCCSA-KM ให้ประสิทธิภาพในการแนะนำรองลงมาตามลำดับ



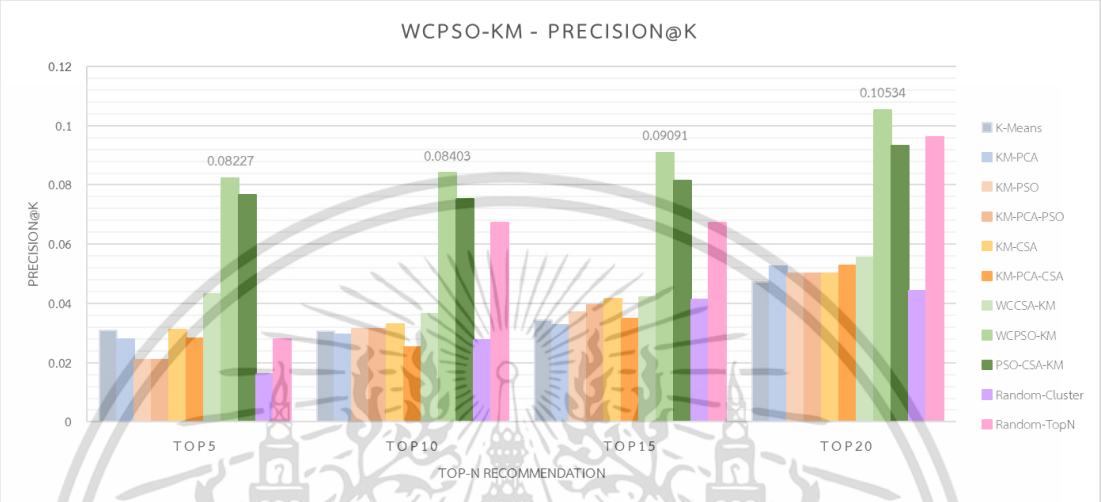
รูปที่ 4.8 กราฟเปรียบเทียบประสิทธิภาพ MAP@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid

ตารางที่ 4.8 ผลการเปรียบเทียบค่าประสิทธิภาพ MAP@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM

แบบจำลอง	MAP@K			
	Top5	Top10	Top15	Top20
K-Means	0.03456	0.05843	0.05843	0.08677
KM-PCA	0.02944	0.05623	0.05623	0.08760
KM-PSO	0.03130	0.06891	0.06891	0.10087
KM-PCA-PSO	0.03130	0.06702	0.06702	0.10087
KM-CSA	0.03762	0.04541	0.06956	0.10258
KM-PCA-CSA	0.03704	0.04058	0.05626	0.08748
WCCSA-KM	0.06930	0.07565	0.10232	0.14114
WCPSO-KM	0.08541	0.11227	0.15261	0.20478
PSO-CSA-KM	0.07978	0.09118	0.11605	0.14895
Random-Cluster	0.01560	0.06056	0.06056	0.08279
Random-TopN	0.02793	0.06728	0.06728	0.09617

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.8 และตารางที่ 4.8 เป็นการเปรียบเทียบแบบจำลองด้วย MAP@K พบว่าแบบจำลองในกลุ่ม Hybrid-Swarm-KM ให้ผลลัพธ์สูงกว่าแบบจำลองกลุ่มอื่น ๆ ซึ่งแบบจำลอง WCPSO-KM มีประสิทธิภาพการแนะนำในทุก TopN สูงที่สุด ได้แก่ Top5 ด้วยค่า MAP เท่ากับ 0.08541, Top10 ค่า MAP เท่ากับ 0.11227, Top15 ค่า NDCG เท่ากับ 0.15261 และ Top20 ค่า NDCG เท่ากับ 0.20478 ตามด้วยแบบจำลอง PSO-CSA-KM และ WCCSA-KM ตามลำดับ



รูปที่ 4.9 กราฟเปรียบเทียบประสิทธิภาพ Precision@K แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid

ตารางที่ 4.9 ผลการเปรียบเทียบค่าประสิทธิภาพ Precision@K ของแบบจำลองแนะนำรายการอาหารด้วย Hybrid-Swarm-KM

แบบจำลอง	Precision@K			
	Top5	Top10	Top15	Top20
K-Means	0.03093	0.03053	0.03449	0.047143
KM-PCA	0.02805	0.02954	0.03287	0.052569
KM-PSO	0.02102	0.03130	0.03702	0.050039
KM-PCA-PSO	0.02102	0.03130	0.03950	0.050039
KM-CSA	0.03126	0.03311	0.04172	0.050047
KM-PCA-CSA	0.02826	0.02516	0.03480	0.052837
WCCSA-KM	0.04315	0.03654	0.04209	0.05541
WCPSO-KM	0.08227	0.08403	0.09091	0.10534
PSO-CSA-KM	0.07670	0.07523	0.08149	0.09316
Random-Cluster	0.01614	0.02774	0.04139	0.04439
Random-TopN	0.02702	0.02838	0.03805	0.05065

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.9 และตารางที่ 4.9 เป็นการเปรียบเทียบแบบจำลองด้วย Precision@K พบว่าแบบจำลอง WCPSO-KM มีประสิทธิภาพในการแนะนำดีที่สุดด้วย ค่า Precision@5 = 0.08227, Precision@10 = 0.08403, Precision@15 = 0.09091 และ Precision@20 = 0.10534 และพบว่าแบบจำลองในกลุ่ม Hybrid-Swarm-KM ให้ประสิทธิภาพดีกว่าแบบจำลองกลุ่มอื่น ๆ อย่างชัดเจน ได้แก่ PSO-CSA-KM และ WCCSA-KM โดยมีค่า Precision มากกว่าแบบจำลองกลุ่มอื่นอยู่ระหว่าง 0.01-0.06

4.1.2 การประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

งานวิจัยนี้ได้ศึกษาและพัฒนาระบบแนะนำรายการอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสม ซึ่งได้ดำเนินการพัฒนาชุดรายการอาหารที่เหมาะสมตามคุณลักษณะและพฤติกรรมการทานอาหารและการออกกำลังรายบุคคล โดยนำรายการอาหารแนะนำ Top15 มาจัดทำชุดรายการอาหารจำนวน 5 ชุดต่อบุคคลด้วยวิธีการ Linear Programming และดำเนินการสำรวจความคิดเห็นและความพึงพอใจต่อชุดรายการอาหารแนะนำดังกล่าวด้วยแบบสอบถาม งานวิจัยนี้สำรวจความพึงพอใจจำนวน 5 บุคคล และมีผลสำรวจดังตารางที่ 4.10

ตารางที่ 4.10 ผลการประเมินความพึงพอใจชุดรายการอาหารแนะนำ

บุคคล	**ชุดรายการอาหารแนะนำ**					**คะแนนความพึงพอใจเฉลี่ย**	**ความพึงพอใจ**
	ชุด 1	ชุด 2	ชุด 3	ชุด 4	ชุด 5		
T001	5	4	5	5	5	4.8	มากที่สุด
T002	5	5	5	4	4	4.6	มากที่สุด
T003	4	4	4	5	5	4.4	มาก
T004	4	5	5	4	5	4.6	มากที่สุด
T005	4	4	5	5	5	4.6	มากที่สุด
เฉลี่ย	4.4	4.4	4.8	4.6	4.8	**4.6**	**มากที่สุด**

ตารางที่ 4.10 ผลสำรวจความคิดเห็นและความพึงพอใจของชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล จำนวน 5 บุคคล พบว่า ผู้ตอบสอบถามมีความพึงพอใจต่อชุดรายการอาหารอยู่ในค่าคะแนน 4-5 และมีคะแนนความพึงพอใจเฉลี่ยทั้ง 5 ชุดรายการอาหารอยู่ระหว่าง 4.4 (ความพึงพอใจมาก) ถึง 4.8 (ความพึงพอใจมากที่สุด) โดยมีคะแนนความพึงพอใจเฉลี่ยต่อชุดรายการอาหารแต่ละชุดอยู่ระหว่าง 4.4 (ความพึงพอใจมาก) ถึง 4.8 (ความพึงพอใจมากที่สุด) เช่นเดียวคะแนนความพึงพอใจเฉลี่ยรายบุคคล จากผลการสำรวจความพึงพอใจทั้งหมด สรุปได้ว่าชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการรายบุคคล มีคะแนนความพึงพอใจเฉลี่ยเท่ากับ 4.6 คือ ความพึงพอใจมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 อภิปรายผล

ในการอภิปรายผลการวิจัยการแนะนำอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมด้วยการผสมผสานการเรียนรู้ของเครื่องทางด้านระบบแนะนำและขั้นตอนวิธีความฉลาดแบบกลุ่มจากผลการวิจัยข้อ 4.1 ซึ่งอภิปรายได้ดังนี้

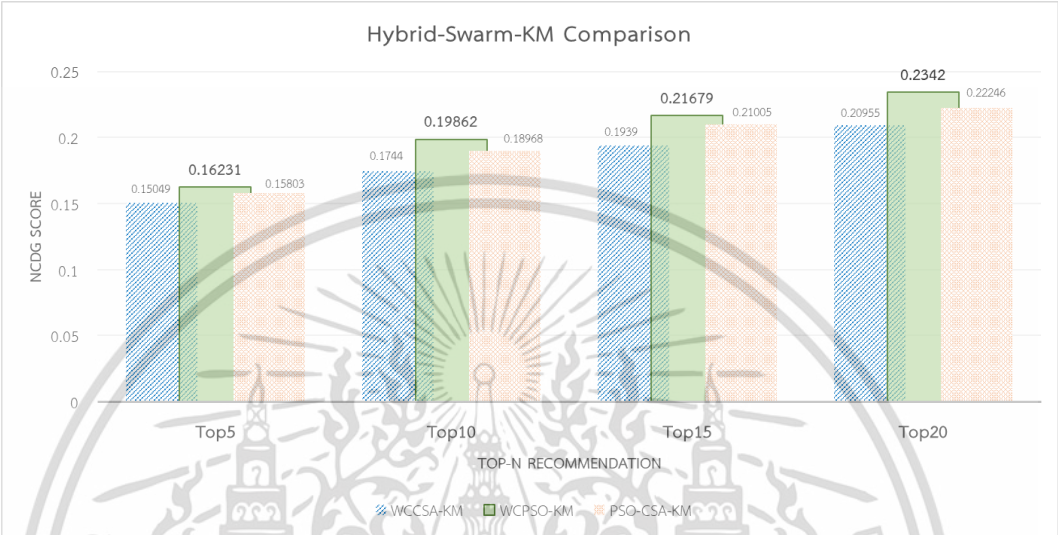
4.2.1 การประเมินประสิทธิภาพแบบจำลองรายการอาหารแนะนำ

การศึกษาครั้งนี้เป็นการวิจัยเพื่อศึกษาพารามิเตอร์และนำเสนอขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence) แบบไฮบริดเพื่อเพิ่มประสิทธิภาพการจัดกลุ่มข้อมูลพฤติกรรมการรับประทานอาหารเชิงลึกและแนะนำชุดรายการอาหารที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล โดยพัฒนาโครงสร้างแบบจำลองรูปแบบ Hybrid-Swarm-KM ประกอบด้วย 3 แบบจำลอง ได้แก่ WCCSA-KM, WCPSO-KM และ PSO-CSA-KM ทำงานกับชุดข้อมูลแบบ implicit feedback หรือข้อมูลพฤติกรรมของผู้ใช้งาน และมีการดัดแปลงข้อมูลเพื่อวิเคราะห์หาข้อมูลพฤติกรรมเชิงลึก การศึกษาพารามิเตอร์ของขั้นตอนวิธีแต่ละชนิดนั้น ได้ดำเนินการศึกษาจำนวนประชากร (population size) เพื่อค้นหาจำนวนประชากรที่ช่วยเพิ่มประสิทธิภาพการทำงานของแต่ละขั้นตอนวิธี เช่น CSA ดำเนินการศึกษาจำนวนประชากรทั้งหมด 9 แบบ เริ่มต้นด้วย 10 และได้ทดลองเพิ่มครั้งละ 5 ได้แก่ 15, 20, 25, 30, 35, 40, 45 และ 50 จากการวิจัยพบว่าจำนวนประชากรเท่ากับ 15 มีประสิทธิภาพในการแนะนำรายการอาหารที่ตรงใจผู้ใช้งานได้มากที่สุดด้วย NDCG score การตรวจสอบความถูกต้องของรายการแนะนำ ดำเนินการเปรียบเทียบข้อมูลชุดทดสอบซึ่งทำการแบ่งข้อมูลในสัดส่วน 70: 30 สำหรับ PSO ดำเนินการศึกษาจำนวนประชากรทั้งหมด 10 แบบ เริ่มต้นด้วย 3 และได้ทดลองเพิ่มครั้งละ 3 ได้แก่ 6, 9, 12, 15, 18, 21, 24, 27 และ 30 จากการวิจัยพบว่าจำนวนประชากรเท่ากับ 27 มีประสิทธิภาพในการแนะนำรายการอาหารได้ดีที่สุดในทุก TopN และตัววัดประเมินผล คือ NDCG@K, MAP@K และ Precision@K ดังนั้นจึงนำจำนวนประชากรที่ดีที่สุดของทั้ง 2 แบบจำลองมาพัฒนาแบบจำลอง PSO-CSA-KM โดยใช้ค่าจำนวนประชากรดังกล่าวในแต่ละส่วนตามขั้นตอนวิธีของแบบจำลอง จากการวิจัยพบว่า แบบจำลอง PSO-CSA-KM มีประสิทธิภาพในการแนะนำอยู่ในระดับที่ 2 ซึ่งมีประสิทธิภาพน้อยกว่าแบบจำลอง WCPSO-KM ที่ประกอบด้วยขั้นตอนวิธี PSO ทั้งหมดด้วยค่า NDCG@K ระหว่าง 0.004-0.011 ดังรูปที่ 4.10 และตารางที่ 4.11

จากการประเมินผลแบบจำลองที่นำเสนอในงานวิจัยนี้ Hybrid-Swarm-KM เปรียบเทียบกับกลุ่มแบบจำลอง เช่น กลุ่ม baseline algorithm ได้แก่ K-Means , KM-PCA กลุ่มประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่ม ได้แก่ KM-PSO, KM-PCA-PSO, KM-CSA และ KM-PCA-CSA กลุ่มสุ่มข้อมูลเพื่อเปรียบเทียบประสิทธิภาพกับการสุ่มข้อมูล ได้แก่ Random-Cluster และ Random-TopN พบว่ากลุ่ม Hybrid สามารถแนะนำรายการอาหารได้ถูกต้องกว่ากลุ่มอื่นๆ ในทุก Top-N ตามรูปที่ 4.7 โดยมีค่าความต่างของค่า NDCG@K อยู่ระหว่าง 0.07- 0.10 ซึ่งถือว่ามีความต่างอย่างชัดเจนในผลการทดลองของงานวิจัยนี้ ทั้งนี้ ในการประมวลผลเพื่อการเรียนรู้แบบจำลอง กลุ่ม Hybrid จะใช้เวลาในการประมวลผลที่นานกว่าแบบจำลองกลุ่มอื่น เนื่องจากมีขั้นตอนในการทำงานที่มากกว่าและซับซ้อน ได้แก่ ขั้นตอนในการหาค่าน้ำหนักของข้อมูล (Weight) ภายในเทคนิคการจัดกลุ่มเคมีน โดยใช้ขั้นตอนวิธี PSO หรือ CSA ซึ่งมีเวลาในการประมวลผลที่นานตามจำนวนประชากรที่กำหนด และขั้นตอนในการหาค่าจุดกลางของกลุ่ม (Centroids) ที่มีการประยุกต์ใช้ขั้นตอนวิธี PSO หรือ CSA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่นเดียวกัน เมื่อผนวกการทำงานของทั้งสองขั้นตอนทำให้มี time complexity ในการทำงานของแบบจำลองที่นานกว่าแบบจำลองกลุ่มอื่น ๆ สำหรับขั้นตอนการทำนายกลุ่มข้อมูลของข้อมูลใหม่พบว่า เวลาในการทำงานไม่แตกต่างกัน เนื่องจากแบบจำลองกลุ่ม hybrid สามารถนำผลลัพธ์การหาค่าน้ำหนักและจุดกลางของกลุ่มมาประยุกต์ใช้กับเทคนิคการจัดกลุ่มเคมีนได้เลย ทำให้เทคนิคเคมีนสามารถทำนายกลุ่มของชุดข้อมูลใหม่ได้ทันที



รูปที่ 4.10 กราฟเปรียบเทียบประสิทธิภาพ แบบจำลองแนะนำด้วยขั้นตอนวิธี Hybrid-Swarm-KM

ตารางที่ 4.11 การเปรียบเทียบประสิทธิภาพของแบบจำลองกลุ่ม Hybrid-Swarm-KM

แบบจำลอง	Top5	Top10	Top15	Top20
WCCSA-KM	0.15049	0.17440	0.19390	0.20955
WCPSO-KM	**0.16231**	**0.19862**	**0.21679**	**0.23420**
PSO-CSA-KM	0.15803	0.18968	0.21005	0.22246

จากรูปที่ 4.10 และตารางที่ 4.11 จากการทดลองเปรียบเทียบประสิทธิภาพของแบบจำลองจำนวน 3 แบบจำลองในกลุ่ม Hybrid ด้วยโครงสร้าง Hybrid-Swarm-KM ที่นำเสนอ พบว่าแบบจำลองที่มีส่วนประกอบของขั้นตอนวิธี PSO ได้แก่ WCPSO-KM และ PSO-CSA-KM มีประสิทธิภาพที่สูงกว่าแบบจำลองที่มีส่วนประกอบของขั้นตอนวิธี CSA คือ WCCSA-KM ซึ่งอธิปรายได้ว่า ขั้นตอนวิธี PSO มีการเลียนแบบพฤติกรรมสัตว์ในการบินหาอาหารของฝูงนก เนื่องจากการทำงานจะมีการเรียนรู้จากพฤติกรรมของเพื่อนนกหรือข้อมูลที่ใกล้เคียงกันในกลุ่มข้อมูลนั้น ๆ ทำให้เมื่อเรียนรู้กับข้อมูลพฤติกรรมรับประทานอาหารของผู้ใช้งานและต้องทำงานร่วมกับการจัดกลุ่มเคมีนเพื่อจัดกลุ่มพฤติกรรมที่มีความคล้ายคลึงหรือใกล้เคียงกัน สามารถเรียนรู้และจัดกลุ่มข้อมูลได้ดีกว่าขั้นตอนวิธี CSA ซึ่งเลียนแบบพฤติกรรมการวางไข่ของนกดุเหว่าที่เลือกสุ่มการวางไข่ตามจุดต่าง ๆ โดยอาศัยความน่าจะเป็นของรังไข่ในการค้นหาคำตอบรุ่นถัด ๆ ไป

ดังนั้น สรุปได้ว่าแบบจำลองกลุ่ม Hybrid-Swarm-KM ที่ประกอบด้วยขั้นตอนวิธี PSO ทั้งสองส่วนจะมีประสิทธิภาพในการแนะนำมากกว่าขั้นตอนวิธี CSA เป็นส่วนประกอบ โดยประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบจำลอง WCPSO-KM , PSO-CSA-KM และ WCCSA-KM มีผลลัพธ์ที่ดีกว่าตามลำดับกับข้อมูลประเภท implicit feedback ของพฤติกรรมการรับประทานอาหารด้วยค่า NCDG@K, MAP@K และ Precision@K

4.2.2 การประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

งานวิจัยนี้ได้สำรวจความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคลด้วยแบบสอบถามจำนวน 5 คน โดยสำรวจรายการอาหารที่รับประทานใน 2 สัปดาห์และประวัติการออกกำลังกาย จากนั้นนำเสนอชุดรายการอาหารจำนวน 5 ชุดต่อคน จากการสำรวจและประเมินผลพบว่า คะแนนเฉลี่ยความพึงพอใจของชุดรายการอาหารแนะนำอยู่ระหว่าง 4.4 (พึงพอใจมาก) ถึง 4.8 (พึงพอใจมากที่สุด) และคะแนนความพึงพอใจเฉลี่ยทุกชุดเท่ากับ 4.6 คือ พึงพอใจมากที่สุด ดังตารางที่ 4.12 ดังนั้นสรุปได้ว่าชุดรายการอาหารที่แนะนำไปยังผู้ใช้งานแต่ละบุคคลสามารถสร้างความพึงพอใจได้มากที่สุด

ตารางที่ 4.12 คะแนนความพึงพอใจเฉลี่ยต่อชุดรายการอาหารแนะนำ

	ชุดรายการอาหารแนะนำ					**คะแนนความพึงพอใจเฉลี่ย**	**ความพึงพอใจ**
	ชุด 1	ชุด 2	ชุด 3	ชุด 4	ชุด 5		
เฉลี่ย	4.4	4.4	4.8	4.6	4.8	**4.6**	**มากที่สุด**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้ได้ทำการพัฒนาแบบจำลองที่ผสมผสานระหว่างการเรียนรู้ของเครื่องทางด้านระบบแนะนำและขั้นตอนวิธีความฉลาดแบบกลุ่ม ได้แก่ ขั้นตอนวิธี K-Mean, ขั้นตอนวิธีความฉลาดแบบกลุ่ม (Swarm Intelligence), เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (User-User Collaborative Filtering) และโปรแกรมเชิงเส้น (Linear Programming) กับชุดข้อมูลการบันทึกประวัติการรับประทานอาหารและออกกำลังกายแบบ implicit feedback โดยมีขั้นตอนการดำเนินงาน 7 ขั้นตอน ประกอบด้วย การรวบรวม จัดเตรียมและประมวลผลข้อมูล, การจัดกลุ่มผู้ใช้งาน, การทำนายและแนะนำอาหาร, การวัดประสิทธิภาพของการแนะนำรายการอาหาร, การพัฒนาและปรับปรุงแบบจำลอง, การคำนวณหาชุดรายการอาหารที่เหมาะสม, การแนะนำชุดรายการอาหาร และการประเมินความพึงพอใจชุดรายการอาหารแนะนำ

ในการพัฒนาและปรับปรุงแบบจำลอง ได้ดำเนินการทดลองค่าจำนวนประชากรของขั้นตอนวิธี CSA และ PSO ด้วยโครงสร้างแบบจำลอง Hybrid-Swarm-KM พบว่าจำนวนประชากรเท่ากับ 15 ของขั้นตอนวิธี CSA และจำนวนประชากรเท่ากับ 27 ของขั้นตอนวิธี PSO จะมีประสิทธิภาพในการแนะนำรายการอาหารได้ดีที่สุด จากนั้นจึงนำจำนวนประชากรของทั้งสองขั้นตอนวิธีไปประยุกต์ใช้กับแบบจำลอง PSO-Swarm-KM พบว่าประสิทธิภาพของแบบจำลองอยู่ในลำดับที่ 2 ซึ่งน้อยกว่าแบบจำลอง WCPSO-KM เพียงเล็กน้อย การประเมินประสิทธิภาพของแบบจำลอง WCPSO-KM ด้วยค่า NDCG@5 เท่ากับ 0.16231, NDCG@10 เท่ากับ 0.19862 , NDCG@15 เท่ากับ 0.21679 และ NDCG@20 เท่ากับ 0.23420 ทั้งนี้ งานวิจัยดำเนินการเปรียบเทียบประสิทธิภาพของแบบจำลองในกลุ่มอื่น ๆ จำนวน 3 กลุ่ม 8 แบบจำลอง ได้แก่ กลุ่ม baseline algorithm คือ K-Means และ KM-PCA กลุ่มประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่ม ได้แก่ KM-PSO, KM-PCA-PSO, KM-CSA และ KM-PCA-CSA กลุ่มสุ่มข้อมูล คือ Random-Cluster และ Random-TopN พบว่ากลุ่ม Hybrid-Swarm-KM มีประสิทธิภาพมากที่สุดอย่างชัดเจน ตามด้วย กลุ่มประยุกต์ใช้ขั้นตอนวิธีความฉลาดแบบกลุ่ม, กลุ่ม baseline algorithm และกลุ่มสุ่มข้อมูล ตามลำดับ

งานวิจัยนี้ได้ประเมินความพึงพอใจชุดรายการอาหารแนะนำที่มีคุณค่าทางโภชนาการรายบุคคลด้วยข้อมูลประวัติส่วนตัว ข้อมูลร่างกาย ประวัติการรับประทานอาหารและการออกกำลังกายพบว่า ผู้ใช้งานมีความพึงพอใจต่อชุดรายการอาหารแต่ละชุดที่แตกต่างกันระดับมากถึงมากที่สุด ด้วยค่าคะแนนระหว่าง 4.4 – 4.8 และมีความพึงพอใจต่อชุดรายการอาหารที่แนะนำทั้งหมดเฉลี่ยเท่ากับ 4.6 คะแนน ซึ่งมีความพึงพอใจระดับมากที่สุด

5.2 ข้อเสนอแนะ

งานวิจัยนี้ใช้ข้อมูลพฤติกรรมการรับประทานอาหารที่ผู้ใช้งานทุกช่วงวัยบันทึกผ่านแอปพลิเคชัน FoodiEat ของทีมวิจัยการวิเคราะห์พฤติกรรมมนุษย์ ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ พบว่ามีข้อมูลที่ผิดปกติจำนวนมากและมีปัญหาความเบาบางของข้อมูล (data sparsity) พบข้อมูลประวัติการบันทึกรายการอาหารที่ซ้ำซ้อนทำให้ต้องขจัดข้อมูลเพื่อให้มีความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถูกต้องมากที่สุด ในงานวิจัยนี้สนใจเฉพาะรายการอาหารประเภทอาหารจานเดียวและกับข้าว ซึ่งสามารถพัฒนาต่อยอดกับอาหารประเภทอื่น ๆ ได้ เช่น ผลไม้ ของหวาน ขนมกินเล่น เป็นต้น อีกทั้งสามารถประยุกต์ข้อมูลโรคประจำตัวหรือความสนใจเฉพาะด้าน เช่น มังสวิรัติ เข้ามาเป็นเงื่อนไขเพิ่มเติมในการแนะนำรายการอาหารได้ สำหรับแนวทางการเพิ่มประสิทธิภาพระบบแนะนำรายการอาหารไทยสามารถปรับปรุงเทคนิค User-based Collaborative Filtering ซึ่งอยู่ในเทคนิคกลุ่ม memory-based เป็นเทคนิคกลุ่ม model-based เช่น SVD, ALS หรือ Deep learning เพื่อให้แบบจำลองสามารถเรียนรู้และปรับปรุงประสิทธิภาพได้แม่นยำมากยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] สำนักงานสถิติแห่งชาติ. 2561. **การสำรวจพฤติกรรมการบริโภคอาหารของประชากร พ.ศ. 2560**. กรุงเทพฯ. กองสถิติพยากรณ์ สำนักงาสถิติแห่งชาติ.

[2] Abhari S, et. Al. 2019. “A Systematic Review of Nutrition Recommendation Systems: With Focus on Technical Aspects”. J Biomed Phys Eng. 2019 Dec 1;9(6):591-602.

[3] M. Tossawat et al. 2018. “Personalized Food Recommendation Using Deep Neural Network”. International Student Project Conference (ICT-ISPC). pp 1-4.

[4] Puriwat Lertkrai, Nattapong Kaewboonma, Jutaporn Lertkrai. 2018. “Developing ontology of food and nutrition for Thai pre-school”. AIP Conference Proceedings.

[5] คณะกรรมการอาหารแห่งชาติ. 2559. **องค์ความรู้ด้านอาหารและโภชนาการสำหรับทุกช่วงวัย**. พิมพ์ครั้งที่ 1. กรุงเทพฯ. ฝ่ายเลขานุการคณะกรรมการอาหารแห่งชาติ สำนักอาหาร สำนักงานคณะกรรมการอาหารและยา.

[6] กรมอนามัย กระทรวงสาธารณสุข. 2561. **ตารางแสดงคุณค่าทางโภชนาการของหารไทย**. พิมพ์ครั้งที่ 1. กรุงเทพฯ. ห้างหุ้นส่วนจำกัด เอ.วี. โปรเกรสซีฟ

[7] กรมอนามัย กระทรวงสาธารณสุข. 2563. **ปริมาณสารอาหารอ้างอิงที่ควรได้รับประจำวันสำหรับคนไทย พ.ศ. 2563**. พิมพ์ครั้งที่ 1. กรุงเทพฯ. ห้างหุ้นส่วนจำกัด เอ.วี. โปรเกรสซีฟ.

[8] Kanungo, Tapas, et al. 2000. "The analysis of a simple k-means clustering algorithm." Proceedings of the sixteenth annual symposium on Computational geometry. pp. 100-109.

[9] Ko, Hyeyoung, et al. "A survey of recommendation systems: recommendation models, techniques, and application fields." Electronics 11.1 (2022): 141.

[10] Guibing Guo. 2012. “Resolving Data Sparsity and Cold Start in Recommender Systems”. *International Conference on User Modeling, Adaptation, and Personalization*. P. 361-364.

[11] Jawaheer, Gawesh, Peter Weller, and Patty Kostkova. "Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback." ACM Transactions on Interactive Intelligent Systems (TiiS) 4.2 (2014): 1-26.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

[12] พิศิษฐ์ บวรเลิศสุธี. 2560. “ตัวแบบสำหรับการเรียงลำดับผลลัพธ์การค้นคืนของเอกสารบนอินทราเน็ต : กรณีศึกษา มหาวิทยาลัยราชภัฏธนบุรี”. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมเว็บ วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์.

[13] อาทิตย์ ศรีแก้ว. 2558. **ปัญญาเชิงคำนวณ (Computational Intelligence).** พิมพ์ครั้งที่ 2. นครราชสีมา. สาขาวิชาวิศวกรรมไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี

[14] Lucija, B., ztok Fister, Jr. and Vili P. 2018. “Swarm Intelligence Algorithms for Feature Selection:A Review”. *Applied Sciences*. 8(1521).

[15] ฐิติพร พงศ์กิดาการ และ ดุลพิเชษฐ์ ฤกษ์ปรีดาพงศ์. 2557. “การจ่ายโหลดอย่างประหยัดด้วยวิธีการค้นหาแบบนกกาเหว่า”. วารสารวิศวกรรมสาร มก. 27(90) : 57-66

[16] วีระ พุ่มเกิด. 2560. “การพัฒนาวิธีการค้นหาแบบนกกาเหว่าโดยการผสมผสานด้วยเจเนติอัลกอริทึมสำหรับหาค่าที่เหมาะสมของการทำแห้งมะละกอแบบออสโมซิส”. ดุษฎีนิพนธ์ปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา วิทยาวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา.

[17] ทวีพล ซื่อสัตย์. 2542. **การวิจัยดำเนินงาน (Operation Research).** คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมการวัดคุม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

[18] Matus Sedlak. 2021. “Content-based Recommender System for Food Recipes.”. Master’s Thesis, Faculty of Informatics, Masaryk University.

[19] Sánchez, Claudia N., Julieta Domínguez-Soberanes, Alejandra Arreola, and Mario Graff. 2023. "Recommendation System for a Delivery Food Application Based on Number of Orders" Applied Sciences 13, no. 4: 2299.

[20] Ahmadian, S., et al. 2022. “Healthy Food Recommendation Using a Time-Aware Community Detection Approach and Reliability Measurement”. International Journal Computer Intelligence System 15, 105 (2022).

[21] Ahuja, Kashish, et al. 2020. "What-to-taste: a food recommendation system.".

[22] R. Yera Toledo, A. A. Alzahrani and L. Martínez. 2019. "A Food Recommender System Considering Nutritional Information and User Preferences," in IEEE Access, vol. 7, pp. 96695-96711.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

[23] Showafah, M., Sihwi, S.W., & ., W. 2021. “Ontology-based Daily Menu Recommendation System for Complementary Food According to Nutritional Needs using Naïve Bayes and TOPSIS”. International Journal of Advanced Computer Science and Applications.

[24] Jasim, Mahdi Nsaif, and Ahmed Bahaddin Hamid. 2018. "Food recommendation system based on nutritional needs of human beings and user preferences."

[25] B. Gouthami and M. Gangappa. 2020. “Nutrition Diet Recommendation System using user’s interest”. International Journal of Advanced Research in Engineering and Technology. 11(12). Pp 2910-2919.

[26] Jing, H., 2022. Application of Improved K-Means Algorithm in Collaborative Recommendation System. Journal of Applied Mathematics, 2022.

[27] Al-Asadi, A. and Jasim, M.N., 2022, September. Clustering-Based Collaborative Recommender System Using a Nature-Inspired Algorithm. In 2022 Iraqi International Conference on Communication and Information Technologies (IICCIT) (pp. 95-100). IEEE.

[28] Tohidi, N. and Dadkhah, C., 2020. Improving the performance of video collaborative filtering recommender systems using optimization algorithm. International Journal of Nonlinear Analysis and Applications, 11(1), pp.483-495.

[29] Katarya, R. and Verma, O.P., 2017. An effective collaborative movie recommender system with cuckoo search. Egyptian Informatics Journal, 18(2), pp.105-112.

[30] Jabbar, Kamran. 2021. Performance Optimization of Swarm Intelligence-based Clustering Algorithms. Master thesis of Computer Science of University of Eastern Finland

[31] Logesh, R., et al., 2018. A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city. Future Generation Computer Systems, 83, pp.653-673.

[32] Gelvez-García, et al, 2023. Optimization of Recommender Systems Using Particle Swarms. Ingeniería, 28.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

[33] Wulandhari, Lili Ayu, and Sani Muhamad Isa. 2019. "Optimum nutrition intake from daily dietary recommendation for Indonesian children using binary particle swarm optimization algorithm." Procedia Computer Science 157. pp. 16-24.

[34] Puraram, T., Chaovalit, P., Peethong, A., Tiyanunti, P., Charoensiriwath, S. and Kimpan, W., 2021. Thai food recommendation system using hybrid of particle swarm optimization and K-means algorithm. In 2021 6th International Conference on Machine Learning Technologies (pp. 90-95).

[35] ภรัณยา อำมฤครัตน์ และพยุง มีสัจ. 2553. การเปรียบเทียบประสิทธิภาพการลดมิติข้อมูลและจำแนกข้อมูลโดยวิธีการทางเครือขายประสาทเทียม. การประชุมทางวิชาการการเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 11 มหาวิทยาลัยขอนแก่น. หน้า 58-65

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก
งานวิจัยที่เผยแพร่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thai Food Recommendation System using Hybrid of Particle Swarm Optimization and K-Means Algorithm

Tanakorn Puraram
King Mongkut's Institute of Technology Ladkrabang, Thailand

Pimwadee Chaovalit
National Electronics and Computer Technology Center, Thailand

Apatha Peethong
National Electronics and Computer Technology Center, Thailand

Pongsak Tiyanunti
National Electronics and Computer Technology Center, Thailand

Supiya Charoensiriwath
National Electronics and Computer Technology Center, Thailand

Warangkhana Kimpan*
King Mongkut's Institute of Technology Ladkrabang, Thailand

ABSTRACT

A food recommendation system is an information filtering tool that helps suggest appropriate food menus to users based on their dietary behavior, nutrition, health, or activity. In this paper, a hybrid method of Particle Swarm Optimization (PSO) and K-Means algorithm is proposed to improve the user's dietary behavior clustering and using Principal Component Analysis (PCA) to reduce the data dimension. Moreover, the User-Based Collaborative Filtering technique is used to predict the rating of relevant Thai food menus and recommendation. The experimental result shows the hybrid method improves the clustering performance from 3 models: Hierarchical Clustering, K-Means, and K-Means with PCA, in terms of silhouette coefficient score. In addition, the hybrid method improves the Davies-Bouldin index score by 44%, 19%, and 17% compared to those models, respectively. The rating prediction result shows the hybrid method outperforms the other methods.

CCS CONCEPTS

• **Information System**; • **Recommender systems**; • **Clustering and classification.**;

KEYWORDS

Particle Swarm Optimization, K-Means, Collaborative Filtering, Thai food Recommendation

ACM Reference Format:

Tanakorn Puraram, Pimwadee Chaovalit, Apatha Peethong, Pongsak Tiyanunti, Supiya Charoensiriwath, and Warangkhana Kimpan. 2021. Thai Food Recommendation System using Hybrid of Particle Swarm Optimization and K-Means Algorithm. In *2021 6th International Conference on Machine Learning Technologies (ICMLT 2021), April 23–25, 2021, Jeju Island, Republic of Korea.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3468891.3468904

*uthors' addresses: Tanakorn Puraram, 61605045@kmitl.ac.th, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand; Pimwadee Chaovalit, National Electronics and Computer Technology Center, Pathum Thani, Thailand, pimwadee.chaovalit@nectec.or.th; Apatha Peethong, National Electronics and Computer Technology Center, Pathum Thani, Thailand, apatha.peethong@nectec.or.th; Pongsak Tiyanunti, National Electronics and Computer Technology Center, Pathum Thani, Thailand, pongsak.tiyanunti@nectec.or.th; Supiya Charoensiriwath, National Electronics and Computer Technology Center, Pathum Thani, Thailand, supiya.charoensiriwath@nectec.or.th; Warangkhana Kimpan, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, warangkhana.ki@kmitl.ac.th.

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

ICMLT 2021, April 23–25, 2021, Jeju Island, Republic of Korea
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8940-2/21/04…$15.00
https://doi.org/10.1145/3468891.3468904

1 INTRODUCTION

The recommendation system aims to anticipate the interests of users and recommend products that are quite interesting to them. It is one of the most powerful machine learning systems used by e-commerce or online businesses to increase sales and user satisfaction. Those businesses have earned a reputation for their intelligent systems such as Amazon, Netflix, Spotify and so on. The recommendation system is popularly used in various industries for particular objectives. For example, a music recommendation platform recommends unheard of songs which are similar to a user's style. Similar concepts are applied to tourism, travel, and banking industries. A recommendation system provides personalization, leaving good impression and good satisfaction for a business's customers, for which personalization can benefit both consumers and the company involved. Businesses can increase the customer's familiarity with the brand, improve business image, increase customer loyalty, and gain a higher probability to attract future purchases from customers.

In terms of a food recommendation system, it recommends appropriate food items for each user based on dietary habits and healthy activities. The system can help users aware and change behavior in a positive way. A type of food recommendation system [1] such as balancing between user preferences and nutritional needs or either and group system which implemented by several techniques. A clustering technique is commonly used in recommendation systems because it is easy to understand and can be implemented very quickly. However, the technique has some disadvantages, such as the accuracy of the recommended items or few data for model training. Thus, to improve the model performance to get the best item recommendations is a challenge. Several studies have proposed a combination of clustering and collaborative filtering techniques [2],[3],[4], in which the results of the recommendations were better than using only either clustering or collaborative filtering.

The aim of this research is to build the hybrid method in clustering the user's behavior and recommend food using Thai food dietary behavior data of Thai people. This idea presents the data preparation to build factors to analyze the user's dietary behavior preference and food rating calculation method for collaborative filtering. The model applied Particle Swarm Optimization (PSO) algorithm in the K-Means clustering technique for finding the best

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

centroids in each cluster and use the User-based Collaborative Filtering technique for finding the Thai food item of their dietary behavior. This study is the first phase of a hybrid optimization model for Top-N recommendation which focuses on optimizing the user clustering performance. Finally, The Top-N recommendation model will be improved in future work.

The rest of the paper is organized as follows: Section 2, we introduce the algorithm and related works. In Section 3, the data preparation and the methodology is described. Section 4 illustrates experimental results. We finally conclude the work with some future directions in Section 5.

2 ALGORITHM AND RELATED WORKS

2.1 User-User Collaborative Filtering Technique

User-User Collaborative Filtering is a memory-based system that uses user rating data to compute the similarity between the users and make recommendations. It is effective and easy to implement. This technique retrieves items from users who are similar to the queried users and estimate the desired item's ratings. The distance measure used for user similarity including Pearson Correlation and Cosine Similarity. It calculates the ratings for an unseen item and prediction indicating to what degree the current user will like or dislike a certain item, then it orders the highest rating prediction result to a Top-N list of recommended items.

2.2 K-Means Clustering

K-Means is one of the partitioning methods in clustering algorithms. It partitions data into K clusters using cluster centers. The cluster center or centroid of each cluster is calculated as an average of all instances in the cluster [5]. This method requires users to specify the number of clusters before training the model. There are 4 steps of the K-Means algorithm as described below:

- (1) Choose K random points as centroids.
- (2) Assign each data point (x_i) to the closest cluster by calculating its distance to each centroid using Euclidean distance.
- (3) Update new centroids by taking the average of the assigned points.
- (4) Repeat step 2 and step 3 until convergence or max iteration is reached.

2.3 Principle Component Analysis

Principle Component Analysis (PCA) is a dimensionality reduction technique for multidimensional datasets. It is particularly suitable for transforming datasets with high numbers of dimensions into datasets with a few dimensions. PCA computes the principal components by finding eigenvectors of the covariance matrix of data. The number of principal components is an optimal number that covers possibly correlated variables which capture the greatest amount of variance in the data and it has less than or equal to the number of original variables [6]. Thus, PCA can capture essential information within data features, and transform them into principal components, whose number of dimensions are much lower than its original. PCA has been widely used in both unsupervised or supervised learning techniques in machine learning for model performance improvement.

Figure 1: Particle Swarm Optimization Algorithm

2.4 Particle Swarm Optimization

Particle swarm optimization (PSO) is a population-based optimization technique inspired by the motion of bird flocking and fish schooling. The system is initialized with a population of random solutions, and the search for the optimal solution is performed by updating generations. The search can be executed by the speed of each particle. During many generations of development, only the most optimistic particles were able to transmit information to other particles [7]. The fitness function was measured for selecting the best position of particle to be the final result. The algorithm of PSO is shown in Figure 1

2.5 Related Works

Various studies related to the food recommendation system have been carried out. In 2017, D. Bianchini et al. [8] presented the PRE-Fer food recommendation system for providing personalized menus to users based on both user's short/long-term preferences and ideal user's nutrition behavior. The model used an innovative algorithm for the ranking of food menus and evaluated the model performance by response time and quality of concept-based relevance. What-To-Taste system produced by K. Ahuja et al. [9] studied food and user profiling data. Users can get recommendations for food by similar items or categories and their preferences using a collaborative filtering approach. G. Bhatia et al. [10] developed an Android application that uses content-based filtering and clustering technique to predict restaurants and food for producing recommendations. Their system constructed user profiles from the inputs and food profiles from food ingredients. S. Bundasak [11] proposed a hybrid recommender system that combines Collaborative Filtering and clustering technique with the Weighted slope one Predictor to improve the model. The objective was to recommend the food for people with the disease to be mindful with eating.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

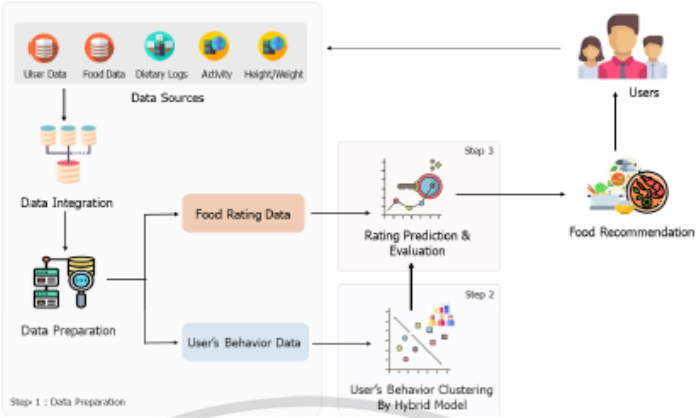


Figure 2: The proposed architecture of Thai food recommendation system

Z. Yuan and F. Luo [2] used the K-Means method to cluster the food set and user-based collaborative filtering algorithm to suggest food by user's nutritional balance in the same dietary preference cluster. The researchers used precision to measure quality of the system and the average accuracy rate is 77%. The K-Means user-based clustering was used in S. Kaushik [12] to enhance the performance in terms of RMSE and computational time or iterations and then to generate better quality and accurate recommendation for each user. The authors H. Koohi and K. Kiani [13] applied Fuzzy C-Means clustering to user-based CF then compared with K-Means, SOM, C-Means&MAX, and C-Means&COG for finding the best similar users group. The research showed the result of the proposed model could achieve the best accuracy of both average and Pearson prediction results in a recommendation. There are research studies on hybrid recommendation models with Swarm Intelligence for improvement. R.Katarya and O.P. Verma [14] used the type division method to divide user interests into different kinds of movies then applied PSO with K-Means for the optimized center of cluster calculation. The output of the previous step was used by Fuzzy C-Mean. The performance of their proposed method had the MAE of 0.7547 and it was better than that of PCA-KM, GA-KM, PCA-SOM, and other compared techniques. S. Alam et al. [15] implemented the Hierarchical Particle Swarm (HPSO) clustering to develop web usage mining based recommender systems and introduced their outlier detection technique. Moreover, the researchers developed Evolutionary PSO (EPSO) clustering method and compared model performance with other clustering algorithms such as DBSCAN and K-Means. They evaluated all models by execution time and fitness score.

3 THE METHODOLOGY

In this study, the proposed architecture of Thai food recommendation system consists of 3 main steps: 1) Data Preparation, 2) Clustering by Hybrid Model, and 3) Rating Prediction and Evaluation, as shown in Fig. 2. The first step, data preparation is the process of integrating, cleaning, and transforming data from the databases into 2 datasets for user preference clustering (User's Behavior data) and for rating prediction (Food Rating data). The second step, clustering by Hybrid Model used the user's behavior data to cluster the users to the same preference group. The last step is rating prediction and evaluation which used the user cluster data from the second step to calculate the rating of a food item by referring to user neighbor within their group. Two metrics: Silhouette Coefficient and Davies–Bouldin index (DBI) used to evaluate the clustering performance and measure the rating prediction result of recommendation system by Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Error (MAE).

3.1 Data Preparation

We acquired the data from a food log mobile application consisting of 307,102 history eating log records, 37,028 food data and nutrition records, 40,497 user profiles, 50,338 weight-height items, and 26,600 user activities. The research studied the behavior of a user's diet and health from log data and built factors for analyzing the insight behavior. We transformed all transaction data to new 27 features to cluster users with similar behaviors into the same cluster. The 27 features are shown in detail in Table. 1. The data preparation step resulted in 2 new datasets: User's Behavior data with 5,957 unique users for clustering and Food Rating data with 59,255 rating pairs to build the recommendation system.

Score and rating calculation: The score was calculated by averaging the weekly portion of eating each food item and then averaging all total weeks. The result is score of pair user with each food, as shown in Eq. 1)

$$food_score_{(u,f)} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{m} portion_{(f,j)}}{m_i}\right) \quad (1)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Table 1: Transformed 27 feature Data for User's behavior Clustering

No.	Name	Comment	No.	Name	Comment
1	age_group_1	age 15-24	15	fats	average fats
2	age_group_2	age 25-39	16	food_weekly_mean	weekly mean of food
3	age_group_3	age 39-59	17	ratio_seafoods	ratio of seafood
4	age_group_4	age 60-100	18	ratio_meatfoods	ratio of meatfood
5	bmi_group_1	bmi less than 18.5	19	ratio_fishfoods	ratio of fishfood
6	bmi_group_2	bmi 18.5-22.99	20	ratio_eggfoods	ratio of eggfood
7	bmi_group_3	bmi 23-24.99	21	ratio_otherfoods	ratio of otherfood
8	bmi_group_4	bmi 25-29.99	22	interval_bf_rank	rank of breakfast
9	bmi_group_5	bmi 30 or more	23	interval_ln_rank	rank of lunch
10	gender_0	gender male	24	interval_dn_rank	rank of dinner
11	gender_1	gender female	25	interval_sp_rank	rank of supper
12	Calories	average calories	26	side_dish	weekly mean side dish
13	Protein	average protein	27	single_dish	weekly mean single dish
14	carbohydrates	average carbohydrates			

Table 2: The conversion of food score to rating

Score Range	Rating
< 0.5	1
>= 0.5 - < 1.0	2
1.0	3
>1.0 - < 2.5	4
>= 2.5	5

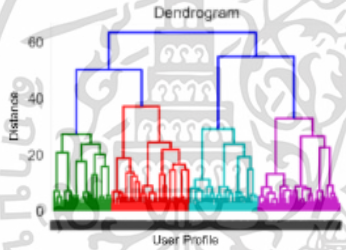


Figure 3: The Dendrogram of User's Dietary Behavior Clustering

where, $portion_{(f,j)}$ represents the food portion of user dietary for each food. m_i represents the number of weekly eating, and N represents the total weeks of user eating each food. We need to convert the food score to the rating because the collaborative filtering algorithm works well with explicit data so we transformed the transaction data or implicit data to explicit data and discretize it into 5 ranges by the criteria used only in our research, as illustrated in Table. 2.

We used the hierarchical clustering technique to build a dendrogram in the process of choosing the number of clusters (K). The dendrogram result is shown in Fig. 3. We chose K=4 to cluster the user's behavior group based on the dendrogram.

3.2 Hybrid K-Means with PCA Data and Particle Swarm Optimization Method (KM-PCA-PSO)

We presented a hybrid method called KM-PCA-PSO that applied the PSO algorithm into the K-Means technique on PCA data. The application of PSO was for improving the clustering performance of users based on the dietary behavior. KM-PCA-PSO applied the user-user collaborative filtering technique for training and predicting the food item ratings. The KM-PCA-PSO has 3 main steps to find the optimal centroids of each cluster. Firstly, it initialized particles by computing the K-Means algorithm and get the centroids of output to initial position for each particle. Secondly, it optimized the centroid position of each particle until terminated by the PSO algorithm. Finally, it assigned the best and final centroids position as the cluster centers in the K-Means algorithm and predict the data to the cluster. The algorithm of the KM-PCA-PSO method is shown in Algorithm 1.

The PSO parameter setup in the KM-PCA-PSO method refers to A. L. Ballardini [16] and S.U. Mane [17] with the following values: the number of particles is 6, inertial weight is 0.72, C_1 and C_2 values are 1.49, and max iteration value is 600. For the fitness function, we used a mean within-cluster distance for best model performance which is given by Eq. 2).

$$fitness_score = \frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{n}|x_i - \mu_j|^2 \quad (2)$$

3.3 The experiment setup

The data of recommendation evaluation was divided into two sets for training and testing by the ratio of 70% and 30%, respectively. We constructed 4 different models to cluster users and rate food items as follows:

- (1) HC : The Hierarchical Clustering which served as a baseline model to choose K (the number of clusters)
- (2) K-Means : A basic K-means algorithm to cluster the transformed 27-feature data

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Algorithm 1 KM-PCA-PSO method

Initial n particle : setup centroids position by compute K-Means clustering algorithm for assign the centroids to initial position each particle and random velocity v;
while *t iteration or criterion is satisfied*
for *loop over all n particle and dimension* do
evaluate fitness function each particle at position using equation bellow;

$$fitness_score = \frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{n}|x_i - \mu_j|^2$$

calculate new velocity;
update new centroids position;
if *pbest* > *current_pbest_score* then
pbest ← *current_pbest_score;*
centroids_position ←*updated_centroids_position;*
End
End
Find the current gbest from all pbest particle;
If *gbest* > *current_gbest* then
gbest ← *current_gbest;*
gbest_centroids_position ←*pbest_centroids_position;*
end
update time $t \leftarrow t + 1$
end
output the final result centroids position and gbest score;
assign best final centroids position result to cluster center in K-Means;
predict the data to closest cluster

- (3) KM-PCA : A K-Means algorithm which ran on PCA data (Using PCA technique to reduce data dimensions from 27 to 8 features)
- (4) KM-PCA-PSO : Applying the KM-PSO model on PCA data. The model applied PSO in a centroid initialization step for finding optimal centroids for each cluster.

4 EXPERIMENTAL RESULTS

This section describes the experimental results for clustering and prediction. We evaluated the recommendation system performance in 2 sections, which are 1) clustering performance and 2) rating prediction evaluation. For clustering evaluation, we used Silhouette Coefficient and Davies–Bouldin index (DBI) metrics to evaluate the clustering model. Silhouette Coefficient is a measure of how similar an object is to its own cluster (cohesion) compared to how similar an object is to other clusters (separation) as shown in Eq (3). The best value is 1 and the worst value is -1. Values near 0 indicate overlapping clusters.

$$SC = \frac{1}{n}\sum_{i=1}^{n}\frac{\mu_{out}^{min}(x_i)\cdot\mu_{in}(x_i)}{max\left\{\mu_{out}^{min}(x_i),\ \mu_{in}(x_i)\right\}} \quad (3)$$

Davies–Bouldin index (DBI) is a measure of the average similarity of each cluster with the closest cluster, where similarity is the ratio of the distance within the cluster to the distance between the clusters. The minimum value in Eq. 4) indicates better performance.

$$DB = \frac{1}{n}\sum_{i=1}^{n} max_{j\neq i}\left(\frac{\sigma_i + \sigma_j}{d(c_i,\ c_j)}\right) \quad (4)$$

For prediction rating, we used 3 evaluation metrics: Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Error (MAE). The RMSE formula is given by Eq. 5), where y_i is the actual value and $\hat{y}$ is the predict values which have been estimated for collaborative filtering.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \quad (5)$$

The experiments setup was a comparison of performances among 4 models: Hierarchical Clustering, K-Means, KM-PCA, and KM-PCA-PSO. The number of cluster (K) is 4.

The clustering performance comparison shown in Fig. 4 consists of Silhouette Coefficient and Davies–Bouldin Index (DBI). We consider both results and found that our method KM-PCA-PSO can partition user's behavior better than KM-PCA, K-Means, and Hierarchical clustering. KM-PCA-PSO had the best silhouette coefficient score of 0.254 and Davies-Bouldin index score of 1.537. The method improved upon baseline models (both Hierarchical clustering and K-Means technique) in silhouette coefficient score and the Davies-Bouldin index score for 44% and 19%, respectively. All K-Means techniques in this study, when applied with PSO algorithm, it could perform better.

Table 3 shows the performance of rating prediction using RMSE, MSE, and MAE measurements. All results are the mean values averaged from 4 clusters in the output of each method. The KM-PCA-PSO method has the smallest RMSE value of all models which achieved 0.538, with a 0.04 standard deviation, followed by 0.548, 0.555, and 0.558 of KM-PCA, K-Means, and Hierarchical Clustering methods, respectively. For MSE and MAE measurements, the results followed a similar trend, where KM-PCA-PSO showed the lowest error rates of 0.29 and 0.371 for MSE and MAE, respectively. Our proposed method, KM-PCA-PSO, produced more accurate prediction ratings across the board, than other methods.

5 CONCLUSION

This study presents a hybrid PSO and K-means clustering method, named KM-PCA-PSO combined with a Collaborative Filtering algorithm to user's dietary behavior data. The method applied Particle Swarm Optimization (PSO) to find the optimal centers of the clusters and then employed user-based collaborative filtering technique to predict the rating of relevant food items. KM-PCA-PSO was compared with several clustering methods such as Hierarchical Clustering, K-Means, and KM-PCA. The experimental results showed that the KM-PCA-PSO method obviously outperformed other compared methods by 2 clustering performance metrics, with the highest Silhouette Coefficient value of 0.254, and the smallest Davies–Bouldin Index value of 1.537. The rating prediction of KM-PCA-PSO recommendation system as measured by RMSE, MSE, and MAE showed the best performance among all methods, with the gap of about 0.01 between methods.

For future work, we want to improve the clustering model performance and proposed our new method by applied Particle Swarm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

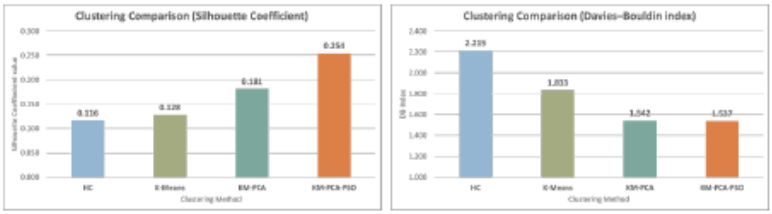


Figure 4: The Silhouette Coefficient (left) and Davies–Bouldin Index (right) of clustering evaluation

Table 3: The performance of rating prediction

Clustering	RMSE	MSE	MAE
HC	0.558 ± 0.040	0.312 ± 0.040	0.390 ± 0.030
K-Means	0.555 ± 0.030	0.309 ± 0.040	0.389 ± 0.030
KM-PCA	0.548 ± 0.020	0.301 ± 0.030	0.382 ± 0.020
KM-PCA-PSO	0.538 ± 0.040	0.290 ± 0.040	0.371 ± 0.040

Optimization (PSO) algorithm to weigh features for distance calculation measurement in the K-Means technique. The method will improve the rank of food items in the Top-N recommendation list.

ACKNOWLEDGMENTS

This work is supported by the Thailand Graduate Institute of Science and Technology Scholarship (TGIST) under the National Science and Technology Development Agency, Thailand (Contract No. SCA-CO-2563-9838-TH). The authors are very thankful to The Human Behavior Analytics Research Team of National Electronics and Computer Technology Center (NECTEC) to provide the data of FoodiEat mobile application for this study.

REFERENCES

[1] Thi Ngoc Trang Tran, Müslüm Atas, Alexander Felfernig, and Martin Stettinger. 2018. An overview of recommender systems in the healthy food domain. *Journal of Intelligent Information Systems* 50, 3 (01 Jun 2018), 501–526. https://doi.org/10.1007/s10844-017-0469-0

[2] Zhicai Yuan and Pang Luo. 2019. Personalized Diet Recommendation Based on K-means and Collaborative Filtering Algorithm. *Journal of Physics: Conference Series* 1213 (06 2019), 032013. https://doi.org/10.1088/1742-6596/1213/3/032013

[3] Urszula Kużelewska. 2019. Collaborative Filtering Recommender Systems Based on k-means Multi-clustering. In *Contemporary Complex Systems and Their Dependability*, Wojciech Zamojski, Jacek Mazurkiewicz, Jarosław Sugier, Tomasz Walkowiak, and Janusz Kacprzyk (Eds.). Springer International Publishing, Cham, 316–325.

[4] Phongsavanh Phorasim and Lasheng Yu. 2017. Movies recommendation system using collaborative filtering and k-means. *International Journal of Advanced Computer Research* 7 (02 2017), 52–59. https://doi.org/10.19101/IJACR. 2017.729004

[5] Thomas Sterling, Matthew Anderson, and Maciej Brodowicz. 2018. Chapter 19 - MapReduce. In *High Performance Computing*, Thomas Sterling, Matthew Anderson, and Maciej Brodowicz (Eds.). Morgan Kaufmann, Boston, 579 – 589. https://doi.org/10.1016/B978-0-12-420158-3.00019-8

[6] T.G. Penkova. 2017. Principal component analysis and cluster analysis for evaluating the natural and anthropogenic territory safety. *Procedia Computer Science* 112 (2017), 99–108. https://doi.org/10.1016/j.procs.2017.08.179 Knowledge Based and Intelligent Information Engineering Systems: Proceedings of the 21st International Conference, KES-20176-8 September 2017, Marseille, France.

[7] Dian Palupi Rini, S. Shamsuddin, and S. Yuhaniz. 2011. Particle Swarm Optimization: Technique, System and Challenges. *International Journal of Computer Applications* 14 (2011), 19–27.

[8] Devis Bianchini, Valeria De Antonellis, Nicola De Franceschi, and Michele Melchiori. 2017. PREFer: A prescription-based food recommender system. *Computer Standards Interfaces* 54 (2017), 64–75. https://doi.org/10.1016/j.csi.2016.10.010SI: New modeling in Big Data.

[9] Kashish Ahuja, Mukul Goel, Sunil Sikka, and Priyanka Makkar. 2020. WHAT-TO-TASTE: A FOOD RECOMMENDATION SYSTEM. *International Journal of Innovative Research in Computer Science Technology* 8 (05 2020). https://doi.org/10.21276/ijirest.2020.8.3.6

[10] Bhatia Gresha, Shedge Saloni, Sahetia Simran, Mhatre Disha, and Gangwani Sahil. 2020. FoodDicted: A Restaurant Food Recommendation System. *International Journal of Future Generation Communication and Networking* 13, 1 (2020), 284–290.

[11] S. Bundasak. 2017. A healthy food recommendation system by combining clustering technology with the Weighted slope one Predictor. In *2017 International Electrical Engineering Congress (iEECON)*. 1–5. https://doi.org/10.1109/IEECON.2017.8075820

[12] S. Kaushik. 2018. An Enhanced Recommendation System using proposed Efficient K Means User-based Clustering Algorithm. *In 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. 251–255. https://doi.org/10.1109/ICACCCN.2018.8748693

[13] Hamidreza Koohi and Kourosh Kiani. 2016. User based Collaborative Filtering using fuzzy C-means. *Measurement* 91 (2016), 134–139. https://doi.org/10.1016/j.measurement.2016.05.058

[14] Rahul Katarya and Om Verma. 2016. A collaborative recommender system enhanced with particle swarm optimization technique. *Multimedia Tools and Applications* 75 (08 2016). https://doi.org/10.1007/s11042-016-3481-4

[15] Shafiq Alam, Gillian Dobbie, and Patricia Riddle. 2012. Towards Recommender System Using Particle Swarm Optimization Based Web Usage Clustering. In *New Frontiers in Applied Data Mining*, Longbing Cao, Joshua Zhexue Huang, James Bailey, Yun Sing Koh, and Jun Luo (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 316–326.

[16] Augusto Luis Ballardini. 2018. A tutorial on Particle Swarm Optimization Clustering. arXiv:1809.01942 [cs.NE]

[17] Sandeep U. Mane and Pankaj G. Gaikwad. 2014. Article: Hybrid Particle Swarm Optimization (HPSO) for Data Clustering. *International Journal of Computer Applications* 97, 19 (July 2014), 1–5. Full text available.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข
ผลการทดลองประสิทธิภาพแบบจำลอง Hybrid-Swarm-Km ตามจำนวนประชากรรายครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Cuckoo Search Algorithm ชื่อ WCCSA-KM ครั้งที่ 1

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.08539	0.03704	0.02826	0.10722	0.04058	0.02516
15	0.17623	0.08318	0.04817	0.19846	0.08849	0.03961
20	0.10378	0.04184	0.03633	0.13118	0.05110	0.03481
25	0.08486	0.03644	0.02784	0.12464	0.04757	0.03361
30	0.08079	0.03265	0.03155	0.10534	0.04237	0.02779
35	0.04880	0.01798	0.01928	0.07938	0.02690	0.02434
40	0.11840	0.05293	0.03862	0.13919	0.06065	0.03115
45	0.08668	0.03675	0.02920	0.13537	0.05102	0.03774
50	0.14409	0.06409	0.04629	0.16737	0.07228	0.03902

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.12992	0.05626	0.03480	0.15255	0.08748	0.05284
15	0.21674	0.11508	0.04411	0.23331	0.15600	0.05799
20	0.16074	0.07914	0.04623	0.17404	0.11800	0.05905
25	0.15021	0.07368	0.04364	0.16940	0.11380	0.05794
30	0.14294	0.06965	0.04199	0.16326	0.10843	0.05965
35	0.11463	0.05351	0.04005	0.13239	0.09647	0.05397
40	0.15747	0.08269	0.03806	0.16779	0.11450	0.04915
45	0.15386	0.07970	0.04410	0.17452	0.12268	0.06179
50	0.17432	0.09563	0.03962	0.18804	0.12976	0.05303

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Cuckoo Search Algorithm ชื่อ WCCSA-KM ครั้งที่ 2

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.09637	0.03949	0.03406	0.12597	0.04987	0.03214
15	0.17887	0.08522	0.04721	0.19879	0.08859	0.03788
20	0.08539	0.03704	0.02826	0.10722	0.04058	0.02516
25	0.06515	0.02555	0.02287	0.11191	0.03950	0.03375
30	0.08539	0.03704	0.02826	0.10722	0.04058	0.02516
35	0.07903	0.03311	0.02855	0.10393	0.03971	0.02685
40	0.06709	0.02750	0.02394	0.10345	0.04246	0.03249
45	0.06777	0.02607	0.02425	0.12225	0.04795	0.03634
50	0.07072	0.02505	0.02971	0.12678	0.05637	0.04889

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.15333	0.07928	0.04330	0.17021	0.12053	0.05695
15	0.21164	0.11262	0.03888	0.22513	0.14688	0.05129
20	0.12992	0.05626	0.03480	0.15255	0.08748	0.05284
25	0.13906	0.06657	0.04129	0.16671	0.10691	0.06158
30	0.12992	0.05626	0.03480	0.15255	0.08748	0.05284
35	0.12506	0.05588	0.03442	0.14566	0.08314	0.04932
40	0.11596	0.06756	0.03830	0.13359	0.10737	0.05299
45	0.15236	0.08664	0.05023	0.17448	0.14017	0.06856
50	0.14700	0.10481	0.05783	0.15914	0.16588	0.07076

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Cuckoo Search Algorithm ชื่อ WCCSA-KM ครั้งที่ 3

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.18223	0.08248	0.05214	0.20348	0.09055	0.04070
15	0.09637	0.03949	0.03406	0.12597	0.04987	0.03214
20	0.08539	0.03704	0.02826	0.10722	0.04058	0.02516
25	0.09167	0.03701	0.03264	0.12843	0.05767	0.03701
30	0.07858	0.03427	0.02508	0.12105	0.04963	0.03410
35	0.08539	0.03704	0.02826	0.10722	0.04058	0.02516
40	0.10591	0.04642	0.03504	0.13409	0.05501	0.03253
45	0.11655	0.04606	0.04547	0.15002	0.06020	0.04075
50	0.17887	0.08522	0.04721	0.19879	0.08859	0.03788

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
10	0.22234	0.11909	0.04409	0.24275	0.16311	0.06172
15	0.15333	0.07928	0.04330	0.17021	0.12053	0.05695
20	0.12992	0.05626	0.03480	0.15255	0.08748	0.05284
25	0.15061	0.09181	0.04907	0.16411	0.14123	0.06099
30	0.13972	0.07688	0.03953	0.15321	0.11336	0.05112
35	0.12992	0.05626	0.03480	0.15255	0.08748	0.05284
40	0.16121	0.07772	0.04314	0.17984	0.11516	0.05980
45	0.17300	0.09176	0.04908	0.19115	0.13433	0.06583
50	0.21164	0.11262	0.03888	0.22513	0.14688	0.05129

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Particle Swarm Optimization ชื่อ WCPSO-KM ครั้งที่ 1

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.02193	0.01951	0.05646	0.02822	0.03241	0.09400
6	0.03661	0.03508	0.09239	0.04164	0.08405	0.13811
9	0.02100	0.02236	0.05681	0.02728	0.03380	0.09533
12	0.03745	0.02748	0.08564	0.03632	0.04376	0.12149
15	0.02844	0.02811	0.07176	0.03517	0.04117	0.11232
18	0.03536	0.04215	0.10414	0.04550	0.05769	0.14953
21	0.05311	0.05344	0.13122	0.06145	0.08194	0.18488
24	0.05639	0.05683	0.13872	0.04858	0.08163	0.17085
27	0.02067	0.02096	0.05426	0.02665	0.02935	0.08970
30	0.05213	0.04310	0.12556	0.03886	0.05942	0.14583

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.03379	0.05346	0.11325	0.04895	0.08625	0.13576
6	0.04831	0.08405	0.16057	0.07059	0.13600	0.18418
9	0.03495	0.05386	0.11618	0.04577	0.08371	0.13046
12	0.04607	0.07337	0.14935	0.06803	0.11735	0.17613
15	0.04123	0.06561	0.13537	0.06007	0.10345	0.16307
18	0.05558	0.08849	0.16952	0.07486	0.13578	0.18207
21	0.09667	0.14335	0.22705	0.14227	0.25770	0.25289
24	0.05254	0.11942	0.18580	0.06558	0.16719	0.19599
27	0.03598	0.05064	0.11457	0.04946	0.08316	0.13358
30	0.03970	0.08394	0.16083	0.04963	0.11832	0.17301

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Particle Swarm Optimization ชื่อ WCPSO-KM ครั้งที่ 2

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.02449	0.02577	0.06498	0.03000	0.03900	0.09891
6	0.02370	0.02494	0.06241	0.02580	0.03431	0.09277
9	0.03025	0.03196	0.08332	0.03778	0.04993	0.12727
12	0.03175	0.04281	0.09726	0.03628	0.05473	0.13886
15	0.03879	0.03832	0.10079	0.03540	0.04800	0.12748
18	0.03598	0.03000	0.08590	0.04338	0.05216	0.13648
21	0.03836	0.03472	0.09452	0.03730	0.04728	0.12738
24	0.02722	0.03737	0.08570	0.04022	0.04942	0.13765
27	0.14247	0.14738	0.22224	0.14845	0.15907	0.25552
30	0.03303	0.03475	0.08154	0.03508	0.04489	0.11605

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.03846	0.06467	0.12402	0.05082	0.10041	0.13922
6	0.03809	0.05555	0.12289	0.04824	0.08706	0.13543
9	0.04908	0.08277	0.15426	0.06469	0.12641	0.16855
12	0.04947	0.08756	0.16676	0.06109	0.13144	0.18067
15	0.03965	0.06726	0.14776	0.05084	0.09294	0.15852
18	0.05898	0.09123	0.16405	0.07335	0.14101	0.17635
21	0.04314	0.07357	0.15076	0.06560	0.11820	0.18056
24	0.05183	0.08549	0.16995	0.06283	0.12986	0.17937
27	0.15368	0.17756	0.27231	0.17080	0.20666	0.29666
30	0.03703	0.06386	0.13465	0.05526	0.09210	0.15987

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยขั้นตอนวิธี Particle Swarm Optimization ชื่อ WCPSO-KM ครั้งที่ 3

Population Size	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.01756	0.01697	0.04520	0.02157	0.02398	0.07502
6	0.03716	0.02667	0.08372	0.03529	0.04186	0.11652
9	0.02662	0.02578	0.06698	0.02697	0.03415	0.09492
12	0.02985	0.03365	0.08322	0.03155	0.04483	0.11694
15	0.03745	0.02748	0.08564	0.03632	0.04376	0.12149
18	0.04358	0.04508	0.10932	0.03672	0.05932	0.13486
21	0.02890	0.03326	0.08043	0.03288	0.04407	0.11803
24	0.03663	0.03875	0.09656	0.03373	0.05279	0.12684
27	0.08369	0.08788	0.21043	0.07699	0.14839	0.25063
30	0.08707	0.11459	0.21391	0.17095	0.23606	0.28007

Population Size	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
3	0.03540	0.04348	0.10593	0.04527	0.07352	0.11844
6	0.04054	0.06766	0.13896	0.05602	0.10268	0.15717
9	0.03823	0.05285	0.12540	0.05111	0.08443	0.14369
12	0.03406	0.06316	0.13003	0.04669	0.09217	0.14621
15	0.04607	0.07337	0.14935	0.06803	0.11735	0.17613
18	0.04629	0.08660	0.15668	0.06531	0.13020	0.17990
21	0.04087	0.06747	0.14019	0.05327	0.10200	0.15477
24	0.04257	0.08012	0.14905	0.06317	0.11978	0.16887
27	0.08307	0.22964	0.26350	0.09577	0.32454	0.27235
30	0.18033	0.29483	0.29698	0.20479	0.36969	0.31624

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองพารามิเตอร์ค่าจำนวนประชากร (Population Size) ของแบบจำลองระบบแนะนำรายการอาหารไทยด้วยจำนวนประชากรที่มีประสิทธิภาพสูงสุดของขั้นตอนวิธี CSA และ PSO ชื่อ PSO-CSA-KM

ครั้งที่	TOP5			TOP10		
	NDCG	MAP	Precision	NDCG	MAP	Precision
1	0.27857	0.16242	0.15027	0.29825	0.16643	0.14837
2	0.12105	0.05091	0.04798	0.15539	0.06235	0.04243
3	0.07446	0.02602	0.03185	0.11538	0.04474	0.03488

ครั้งที่	TOP15			TOP20		
	NDCG	MAP	Precision	NDCG	MAP	Precision
1	0.31059	0.18197	0.15231	0.32246	0.20467	0.16216
2	0.17109	0.08816	0.04498	0.18551	0.12224	0.05960
3	0.14847	0.07801	0.04719	0.15941	0.11994	0.05773

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค
แบบสอบถามความพึงพอใจระบบแนะนำชุดรายการอาหารไทย ที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบสอบถามความพึงพอใจระบบแนะนำชุดรายการอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

เพศ : .....
อายุ : .....
น้ำหนัก : ..... **อายุ** : .....
อาชีพ : .....

รายการแนะนำชุดอาหารไทย จำนวน 5 ชุด
โปรดให้คะแนนความพึงพอใจต่อชุดแนะนำรายการอาหารไทย โดยให้คะแนน 1 – 5 (ไม่พึงพอใจ ถึง พึงพอใจมากที่สุด)

ชุดที่ 1

รายการอาหาร	ค่าพลังงาน
เมนูที่ 1 :	
เมนูที่ 2 :	
เมนูที่ 3 :	
เมนูที่ 4 :	
พลังงานรวม	
คะแนนความพึงพอใจชุดนี้	

ชุดที่ 2

รายการอาหาร	ค่าพลังงาน
เมนูที่ 1 :	
เมนูที่ 2 :	
เมนูที่ 3 :	
เมนูที่ 4 :	
พลังงานรวม	
คะแนนความพึงพอใจชุดนี้	

ชุดที่ 3

รายการอาหาร	ค่าพลังงาน
เมนูที่ 1 :	
เมนูที่ 2 :	
เมนูที่ 3 :	
เมนูที่ 4 :	
พลังงานรวม	
คะแนนความพึงพอใจชุดนี้	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แบบสอบถามความพึงพอใจระบบแนะนำชุดรายการอาหารไทยที่มีคุณค่าทางโภชนาการอย่างเหมาะสมรายบุคคล

รายการแนะนำชุดอาหารไทย จำนวน 5 ชุด (ต่อ)

โปรดให้คะแนนความพึงพอใจต่อชุดแนะนำรายการอาหารไทย โดยให้คะแนน 1 – 5 (ไม่พึงพอใจ ถึง พึงพอใจมากที่สุด)

ชุดที่ 4

รายการอาหาร	ค่าพลังงาน
เมนูที่ 1 :	
เมนูที่ 2 :	
เมนูที่ 3 :	
เมนูที่ 4 :	
พลังงานรวม	
คะแนนความพึงพอใจชุดนี้	

ชุดที่ 5

รายการอาหาร	ค่าพลังงาน
เมนูที่ 1 :	
เมนูที่ 2 :	
เมนูที่ 3 :	
เมนูที่ 4 :	
พลังงานรวม	
คะแนนความพึงพอใจชุดนี้	

คะแนนความพึงพอใจเฉลี่ยรายบุคคล	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นายธนากร ปุรารัมย์
วัน เดือน ปีเกิด	14 เมษายน 2538
ที่อยู่ปัจจุบัน	143/34 ศุภาลัย โมด้า ปิ่นเกล้า-พุทธมณฑลสาย3 หมู่ 3 ถนนสำเร็จพัฒนา ตำบลศาลากลาง อำเภอบางกรวย จังหวัดบุรีรัมย์ 11130
ประวัติการศึกษา	(2557) รัฐศาสตรบัณฑิต เกรดเฉลี่ย 2.87 (มหาวิทยาลัยรามคำแหง) (2559) วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.95 (มหาวิทยาลัยราชภัฏบุรีรัมย์) (2566) วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.88 (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)
ทุนการศึกษาที่ได้รับ	ทุนสถาบันบัณฑิตวิทยาศาสตร์และเทคโนโลยีไทย ระดับปริญญาโท สัญญาเลขที่ SCA-CO-2562-9838-TH
ผลงานทางวิชาการ	T. Puraram, P. Chaovalit, A. Peethong, P. Tiyanunti, S. Charoensiriwath and W. Kimpan, “Thai Food Recommendation System using Hybrid of Particle Swarm Optimization and K-Means Algorithm,” ICMLT 2021: 2021 6th International Conference on Machine Learning Technologies, Jeju Island, Republic of Korea, 23-25 April 2021, pp. 90-95.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้