บทความวิจัย (Research Article)

การกรองแบบร่วมกันโดยพิจารณาสินค้าเป็นหลักเพื่อแนะนำบอร์ดเกมที่มีประสิทธิภาพ

Item-Based Collaborative Filtering for Effective Board Game Recommendation

สุคชาธัม เชียวศิริถาวร ทิวัตถ์ ทรัพย์รัตนกุล และ กฤษณะ ไวยมัย*

ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยเกษตรศาสตร์
* ผู้รับผิดชอบบทความ
fengknw@ku.ac.th

Received: 1 Aug 2024
Revised: 4 Dec 2024
Accepted: 6 Dec 2024

บทคัดย่อ

ผู้เล่นบอร์ดเกมมักเผชิญกับความยากลำบากในการค้นหาเกมที่เหมาะสม การหาผู้ที่จะเล่นบอร์ดเกมด้วยกัน และการค้นหาสถานที่ในการเล่นบอร์ดเกม งานวิจัยนี้นำเสนอเว็บไซต์ใหม่ที่จัดการกับปัญหาเหล่านี้ โดยผสมผสานระบบแนะนำบอร์ดเกมและแพลตฟอร์มชุมชนสำหรับผู้เล่นบอร์ดเกม ผู้วิจัยได้พัฒนาและประเมินโมเดลคำแนะนำต่าง ๆ ได้แก่การกรองตามเนื้อหา (content-based filtering), การกรองการทำงานร่วมกันตามผู้ใช้ (user-based collaborative filtering), การกรองการทำงานร่วมกันตามสินค้า (item-based collaborative filtering) และการเรียนรู้เชิงลึก (deep learning) ในขณะที่โมเดลการเรียนรู้เชิงลึกแสดงประสิทธิภาพสูงสุด การกรองการทำงานร่วมกันตามสินค้ากลับเป็นโมเดลที่ใช้งานได้จริงที่สุด เนื่องจากมีประสิทธิผลและความเป็นไปได้ในการดำเนินการงานวิจัยนี้แสดงให้เห็นถึงประสิทธิภาพของโมเดลการกรองการทำงานร่วมกัน โดยพิจารณาสินค้าเป็นหลักในการแนะนำบอร์ดเกมที่เหมาะสมให้กับผู้ใช้ โดยเว็บไซต์นี้จะมีระบบแนะนำบอร์ดเกม และพื้นที่ชุมชนสำหรับผู้เล่นบอร์ดเกม ซึ่งจะช่วยเพิ่มประสบการณ์ที่ดีในการเล่นบอร์ดเกมสำหรับผู้เล่น

คำสำคัญ: ระบบแนะนำบอร์ดเกม แพลตฟอร์มชุมชนสำหรับผู้เล่นบอร์ดเกม การกรองข้อมูลตามเนื้อหา การกรองแบบร่วมกันโดยพิจารณาสินค้าเป็นหลัก การเรียนรู้เชิงลึก

Abstract

Board game players often face difficulties in discovering suitable games, finding players, and locating places to play board games. This research presents a novel website addressing these issues by incorporating a board game recommendation system and a dedicated community platform for board game players. We developed and evaluated various recommendation models, including content-based filtering, user-based collaborative filtering, item-based collaborative filtering, and deep learning-based approaches. While the deep learning model exhibited the highest performance, the item-based collaborative filtering emerged as the most practical solution due to its effectiveness and implementation feasibility. Our experiments demonstrate the effectiveness of the item-based collaborative filtering model in recommending suitable board games to users. This website, equipped with this recommendation system and a community space, holds significant potential to enhance the overall board game experience for players.

Keywords: board game recommendation system, community platform for board game players, content-based filtering, item-based collaborative filtering, deep learning

1. บทนำ

เกมกระดานหรือบอร์ดเกมคือเกมที่ต้องใช้ชิ้นส่วนหรือตัวหมากวางไว้บนพื้นที่เล่นที่เปรียบได้กับกระดาน ซึ่งจะมีพื้นผิวหรือรูปภาพเฉพาะสำหรับเกมนั้น ๆ [1] โดยบอร์ดเกมเริ่มพัฒนาอย่างต่อเนื่องเมื่อต้นศตวรรษที่ 20 ตลาดที่ใหญ่ที่สุดสองแห่งคือเยอรมนีและสหรัฐอเมริกา [2]

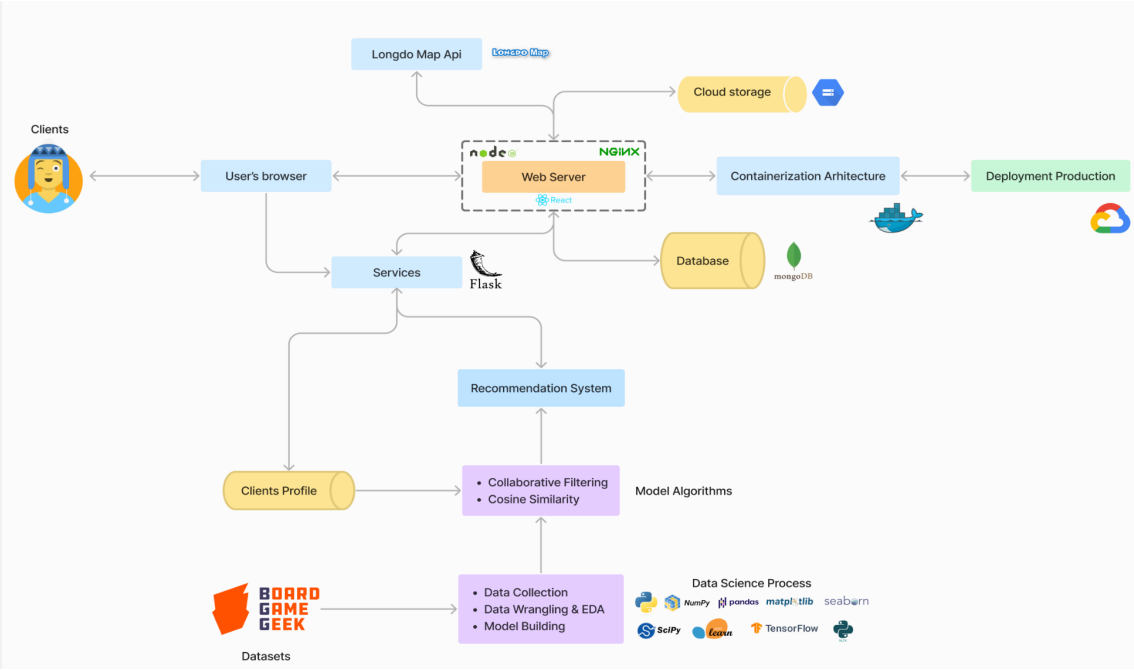
ตลาดบอร์ดเกมมีการเติบโตอย่างเนื่อง โดยขนาดตลาดบอร์ดเกมทั่วโลกคาดว่าจะเติบโตที่ CAGR 7.55% และเพิ่มขึ้น 3,423.69 ล้านเหรียญสหรัฐ ระหว่างปี 2022 ถึง 2027 [3]

แม้แต่ในประเทศไทยก็มีการนำบอร์ดเกมมาแปลเป็นภาษาไทยและมีบอร์ดเกมคาเฟ่ (สถานที่ที่ทุกคนสามารถมาเล่นบอร์ดเกมต่าง ๆ โดยจะคิดค่าบริการเป็นรายชั่วโมงรวมถึงมีอาหารและเครื่องดื่มไว้จำหน่ายและให้บริการ) [4] เปิดเป็นจำนวนมากทั้งในห้างสรรพสินค้าและบริเวณสถานีรถไฟฟ้า

เนื่องจากชุมชนของผู้เล่นมีขนาดใหญ่ วิธีในการขอคำแนะนำบอร์ดเกมมักจะมาจาก 1) ใช้ Forums เช่น boardgamegeek.com และ reddit.com 2) ตรวจสอบจากความเห็นบน YouTube.com 3) เว็บไซต์ต่าง ๆ ที่มีการรีวิวบอร์ดเกม ข้อเสียของวิธีการเหล่านี้คือคำแนะนำมักเป็นเกม "กระแสหลัก" เดิม ๆ ซึ่งแทบไม่มีคำแนะนำสำหรับเกมที่ไม่ค่อยมีคนรู้จักหรือเป็นที่นิยมในเฉพาะกลุ่ม โดยทั่วไปผู้เล่นจะต้องพิจารณาความคิดเห็นหรือการให้คะแนนเหล่านี้หากไม่เคยเล่นเกมนั้น ๆ มาก่อน

จากการวิเคราะห์ปัญหาดังกล่าวผู้วิจัยจึงจัดทำงานวิจัยนี้ขึ้น โดยพัฒนาเว็บไซต์ที่มีระบบแนะนำบอร์ดเกม (Board Game Recommender System) สำหรับผู้ใช้ที่ต้องการหาบอร์ดเกมเล่นและเป็นชุมชนสำหรับผู้เล่นบอร์ดเกม (Board Game Community) โดยจะมีระบบเพิ่มเติมได้แก่ระบบการสร้างปาร์ตี้สำหรับนัดเล่นบอร์ดเกม การให้ข้อมูลต่าง ๆ เช่น บอร์ดเกมยอดที่เป็นที่นิยมในช่วงเวลานั้น ๆ และช่องทางการติดต่อร้านบอร์ดเกมที่ใกล้เคียง

ในอนาคตเมื่อมีข้อมูลการเล่นบอร์ดเกมของคนไทยมากขึ้น สามารถนำโมเดลและเว็บไซต์นี้ไปต่อยอดให้เหมาะสมกับคนไทยโดยเฉพาะ ซึ่งผู้จัดทำหวังเป็นอย่างยิ่งว่างานวิจัยนี้จะสามารถช่วยอำนวยความสะดวกให้ผู้ใช้และมีส่วนร่วมในการผลักดันวงการบอร์ดเกมในประเทศไทยได้



รูปที่ 1 กระบวนการพัฒนาและส่งมอบเว็บไซต์ให้กับผู้ใช้งาน

2. ขั้นตอนการดำเนินงานวิจัย

จากรูปที่ 1 แสดงภาพรวมสถาปัตยกรรมของระบบและการทำงานของแต่ละส่วน โดยมีรายละเอียดการดำเนินงานวิจัยในแต่ละขั้นตอน ดังนี้

2.1 แหล่งข้อมูล

ในงานวิจัยนี้ใช้การบูรณาการข้อมูลจากหลายแหล่งข้อมูล ได้แก่ 1) ข้อมูลบอร์ดเกม 2) ข้อมูลประเภทบอร์ดเกม 3) ข้อมูลรีวิวบอร์ดเกม 4) ข้อมูลบอร์ดเกมยอดนิยม 5) ข้อมูลร้านบอร์ดเกมคาเฟ่

2.1.1 BoardGameGeek

เว็บไซต์ BoardGameGeek (BGG) [5] เป็น Forum ออนไลน์สำหรับผู้ที่ชื่นชอบการเล่นบอร์ดเกมและเป็นฐานข้อมูลเกมที่รวบรวมข้อมูลบอร์ดเกมกว่า 125,600 เกม นอกจากนี้เว็บไซต์ยังให้ผู้ใช้สามารถให้คะแนนเกมต่าง ๆ ได้ [6]

- BoardGameGeek API เว็บไซต์นี้มี API (BGG XML API2) [7] สำหรับเข้าถึงข้อมูลต่าง ๆ ของเว็บไซต์ BoardGameGeek
- BoardGameGeek Wiki [8] เป็นพื้นที่สำหรับค้นหาข้อมูลต่าง ๆ ที่เว็บไซต์ BoardGameGeek

2.1.2 Wizards of Learning

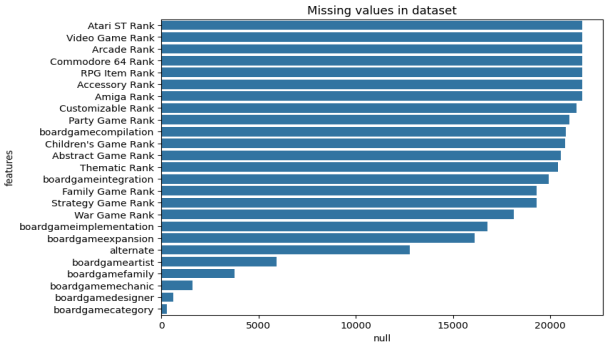
Wizards of Learning เป็นเว็บไซต์ที่ให้บริการออกแบบบอร์ดเกม และมีการรวบรวมสถานที่เล่นบอร์ดเกมไว้ [9]

2.2 การเตรียมข้อมูล

ในขั้นตอนการเตรียมข้อมูลนี้แบ่งการเตรียมข้อมูลออกเป็น 4 ขั้นตอนย่อย ดังต่อไปนี้

2.2.1 การสำรวจข้อมูล Exploratory Data Analysis)

- สำรวจว่ามี Feature ไหนบ้างที่มีข้อมูลสูญหาย (Missing values) ปริมาณมาก ซึ่งหากมีข้อมูลสูญหายมากเกินไปก็อาจจะไม่สามารถนำ Feature นั้นไปใช้ประโยชน์ได้ ดังรูปที่ 2
- สำรวจการการกระจายตัวของแต่ละ Feature เพื่อนำไปใช้แบ่ง Filter ของระบบแนะนำบอร์ดเกม
- สำรวจจำนวนบอร์ดเกมในแต่ละประเภท (Category) เพื่อนำไปใช้เป็น Filter ของระบบแนะนำบอร์ดเกม
- สำรวจอันดับ (Rank) ของบอร์ดเกม ว่าเกมส่วนใหญ่ที่หาเล่นได้ง่ายในประเทศไทย อยู่อันดับประมาณเท่าไหร่ เผื่อกรองให้ระบบแนะนำเฉพาะเกมที่สามารถหาเล่นได้ไง



รูปที่ 2 Missing values ของแต่ละ feature ใน dataset

2.2.2 การทำความสะอาดข้อมูล (Data Cleansing)

- ตัด Features ที่ไม่จำเป็น, ซ้ำซ้อน นำไปวิเคราะห์ยาก หรือ Missing values เยอะจนนำไปใช้ไม่ได้
- คัดเฉพาะบอร์ดเกมที่มีที่มีผู้ใช้ให้คะแนน >= 300 เนื่องจากข้อมูลการรีวิวที่น้อยเกินไป อาจทำให้เกิด Bias
- คัดเฉพาะบอร์ดเกมที่ Rank <= 6,000 และ Publication Year >= 1995 เนื่องจากเกมที่เก่าเกินไปหรืออันดับไม่ดีมักจะเป็นเกมที่หามาเล่นได้ยาก โดยเฉพาะในประเทศไทย
- แทนที่รีวิวที่ผู้ใช้คนเดิมรีวิวเกมเดิมหลายครั้งด้วยค่าเฉลี่ย เนื่องจากการรีวิวซ้ำเกิดจากการที่ความรู้สึกของผู้ใช้ที่มีต่อเกมนั้น ๆ เปลี่ยนไป และข้อมูลที่รวบรวมมาไม่ได้ระบุไว้ว่ารีวิวไหนเป็นรีวิวล่าสุด จึงแทนที่ด้วยค่าเฉลี่ยซึ่งจะเป็นการเฉลี่ยความรู้สึกของผู้ใช้ที่มีต่อเกม
- ตัดรีวิวที่ Rating น้อยกว่า 1 (Rating มี 1-10)
- ตัดผู้ใช้ที่รีวิวน้อยกว่า 450 เกม (ข้อจำกัดเรื่องทรัพยากรและการประมวลผล จึงเลือกเฉพาะผู้ใช้ที่มีข้อมูลเยอะ)

2.2.3 การแปลงข้อมูล (Data Transformation)

- One-Hot Encoding แปลงข้อมูลจัดกลุ่ม (Categorical data) ให้อยู่ในรูปแบบของค่า 0 และ 1 (Binary values)

- Data Normalization ทำข้อมูลให้เป็นมาตรฐาน เป็นเทคนิค Feature scaling ที่จะทำให้โมเดลสามารถอ่านข้อมูลได้ง่ายขึ้นและไม่มี Bias

2.2.4 การรวมข้อมูล (Data Integration)

ทำการรวมข้อมูลที่ผ่านกระบวนการตามขั้นตอนที่ 2.2.1 ถึง 2.2.3 เพื่อสร้างเป็นชุดข้อมูลตั้งต้นที่เตรียมจะนำไปสร้างโมเดลต่อไป จะได้ข้อมูลดังต่อไปนี้

- ข้อมูลบอร์ดเกม 4,582 เกม
- ข้อมูลรีวิว 2,018,715 รีวิว
- ข้อมูลประเภทของบอร์ดเกม 20 ประเภท
- ข้อมูลบอร์ดเกมที่เป็นที่นิยม 50 เกม ณ ช่วงเวลานั้น ๆ

2.3 โมเดลแนะนำบอร์ดเกม

โมเดลในการทำระบบแนะนำ (Recommendation system) มีหลายวิธีการ โดยในงานวิจัยนี้ได้ทดลองโมเดล Content-based filtering และ Collaborative filtering [10, 11]. Content-based filtering ใช้คุณสมบัติของบอร์ดเกมและความชอบของผู้เล่น แต่มีข้อจำกัดเรื่องความหลากหลาย (over-specialization) ส่วน Collaborative filtering เช่น Item-based collaborative filtering [12] ใช้ข้อมูลการโต้ตอบระหว่างผู้เล่นและเกมในอดีต ซึ่งให้คำแนะนำที่หลากหลายกว่า แต่มีปัญหาเรื่องข้อมูลกระจัดกระจาย (data sparsity) ส่วน Collaborative filtering เช่น Model-based จะสามารถวิเคราะห์ความสัมพันธ์ที่ซับซ้อนได้ และมีความสามารถในการปรับขนาด (scalability) แต่ต้องใช้ทรัพยากรมาก

2.3.1 Content-based filtering

Content-based filtering เป็นการสร้างโปรไฟล์ของผู้เล่นแต่ละคนแยกกันไป แล้วแนะนำบอร์ดเกม ตามโปรไฟล์ของผู้เล่นนั้นเป็นคน ๆ ไป [13]

name	Pandemic	Carcassonne	Catan	7 Wonders	Dominion	Ticket to Ride	Codenames	Terraforming Mars	7 Wonders Duel	Agricola	...	R
Pandemic	1.000000	0.236573	0.113901	0.385922	0.202567	0.248093	0.241890	0.177954	0.314378	0.300188	...	0.000
Carcassonne	0.236573	1.000000	0.138159	0.139320	0.364013	0.167183	0.177822	0.179878	0.132407	0.252861	...	0.081
Catan	0.113901	0.138159	1.000000	0.029514	0.144589	0.088542	0.094176	0.095265	0.000000	0.107134	...	0.000

รูปที่ 3 ผลลัพธ์ของโมเดล Content-based filtering

โดยในงานวิจัยนี้ ได้ทำการทดลองสร้างโมเดล Content-based filtering โดยใช้ Cosine similarity ในการหาความคล้ายของเกม ซึ่ง Features ที่ใช้จะมีทั้งการใช้ Numerical data และ Categorical data

จากรูปที่ 3 เมื่อผ่านโมเดล Content-based filtering จะได้ผลลัพธ์เป็นเมทริกซ์ (Matrix) ความคล้ายของแต่ละเกม ซึ่งสามารถนำไปใช้ต่อในการแนะนำบอร์ดเกมได้

2.3.2 Collaborative filtering

หลักการของ Collaborative filtering คือ การนำข้อมูลจากคนหมู่มากหลาย ๆ คนมาช่วยกันวิเคราะห์ว่าผู้ใช้คนนี้จะชอบอะไร ซึ่งก็สามารถทำได้หลายแบบเช่นกัน [13] โดยในงานวิจัยนี้จะทดลองใช้โมเดลประเภท Memory-based และ Model-based

- Memory-based เป็นการใช้ Nearest neighbor คือการที่เราคาดเดา Rating จากการดูจากผู้ใช้ (User-based) หรือบอร์ดเกม (Item-based) ที่ใกล้เคียงกัน [13] โดยในงานวิจัยนี้ ได้ทำการทดลองสร้างโมเดล Memory-based โดยใช้ Cosine similarity ในการหาความคล้ายของผู้ใช้และความคล้ายของเกม ดังรูปที่ 4 และ 5

user	1 Family Meeple	1000rpm	28green	42amu	549sd	AHforever	AJBrandon	ASSaali	Aarkas	AaronBE85	...	zhiwiller
1 Family Meeple	1.000001	0.351328	0.441230	0.463583	0.431945	0.327051	0.307018	0.348769	0.365398	0.422336	...	0.413991
1000rpm	0.351328	1.000003	0.384807	0.314870	0.343502	0.153420	0.327590	0.341947	0.277185	0.328638	...	0.287583
28green	0.441230	0.384807	0.999999	0.414176	0.423904	0.287099	0.382935	0.345915	0.311133	0.390536	...	0.454744

รูปที่ 4 โมเดล User-based Collaborative Filtering

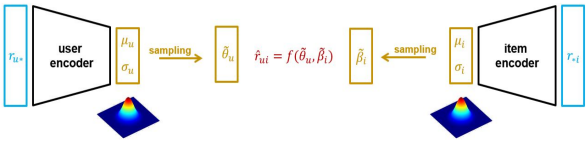
name	...and then, we held hands.	...und tschüss!	10 Days in Africa	10 Days in Asia	10 Days in Europe	10 Days in the Americas	10 Days in the USA	10 Minute Heist: The Wizard's Tower	10' to Kill	1000 Blank White Cards	...	Zooloretto Junior
...and then, we held hands.	1.000000	0.080115	0.158781	0.108188	0.140938	0.108727	0.127127	0.147714	0.141141	0.061741	...	0.075312
...und tschüss!	0.080115	1.000000	0.164283	0.135121	0.228677	0.106158	0.131682	0.073456	0.054965	0.035974	...	0.121544

รูปที่ 5 โมเดล Item-based Collaborative Filtering

- Model-based เป็นการแก้ปัญหาจาก Memory-based ที่มีข้อจำกัดของ ในเรื่องของความกระจัดกระจาย (Sparse) และความสามารถในการขยายขนาด (Scalability) ที่จะต้อง

เก็บเมทริกซ์ Rating ตามจำนวนของผู้ใช้และบอร์ดเกมทั้งหมด ในส่วนของ Model-based จะเป็นการใช้โมเดลต่าง ๆ ในการหาเวกเตอร์ที่มีขนาดเล็กที่จะมาใช้เป็นตัวแทนของผู้ใช้และสินค้า [13] โดยในงานวิจัยนี้ ได้ทำการทดลองสร้างโมเดล Deep Learning โดยใช้โมเดล Bilateral Variational Autoencoder (BiVAE) [14]

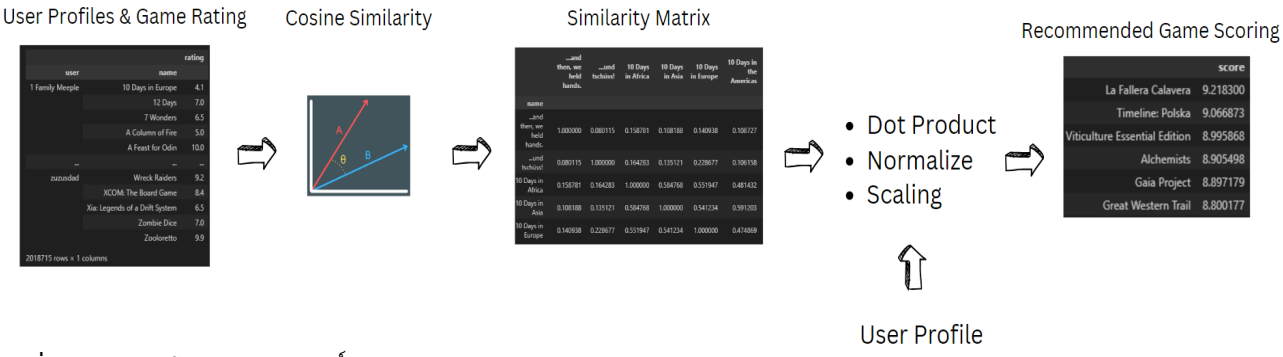
$$r_{ui} \sim \text{EXPFAM}(r_{ui}; \eta(\theta_u; \beta_i; \omega))$$



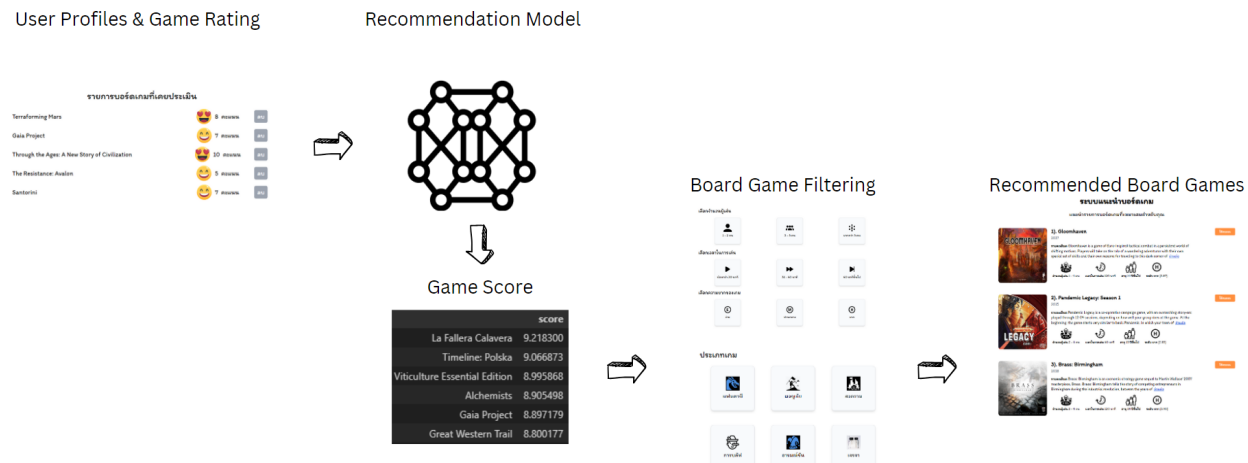
รูปที่ 6 สถาปัตยกรรมของ BiVAE

โดยทั่วไป Collaborative filtering (CF) จะประกอบด้วยเมทริกซ์การโต้ตอบระหว่างผู้ใช้และสินค้า โดยเป็นรูปแบบหนึ่งของข้อมูลไดอะดิก (Dyadic data) อย่างไรก็ตาม โมเดลประเภทนี้เป็นที่ทราบกันว่ามีความสามารถที่จำกัด เนื่องจากสามารถจับเฉพาะรูปแบบเชิงเส้นทั้งในข้อมูลและพื้นที่แฝง (Latent spaces) เท่านั้น จึงเกิดโมเดล Variational Autoencoder (VAE) ที่ถูกนำมาใช้กับ CF โดย VAE ไม่ได้พยายามเรียนรู้การกระจายตัวแทน จึงสามารถอธิบายความไม่แน่นอนในพื้นที่แฝงได้

Bilateral Variational Autoencoder (BiVAE) เป็นโมเดลที่ประกอบด้วย Generative model ของ User-item based, Uesr-based และ Item-based ที่กำหนดพารามิเตอร์โดยใช้โครงข่ายประสาทเทียมหลายชั้น ซึ่งรวมทั้งหมดเข้าด้วยกันสำหรับเข้ารหัสข้อมูลไดอะดิกอัตโนมัติ (Auto-encode dyadic data) โดย BiVAE สามารถจับความไม่แน่นอนของข้อมูลไดอะดิกทั้งสองด้าน (มุมมองจาก users และ items) ซึ่งจะเพิ่มประสิทธิภาพของข้อมูลแบบกระจัดกระจาย เมื่อเปรียบเทียบกับ VAE [14] โดยโมเดล BiVAE มีสถาปัตยกรรม (architecture) ดังรูปที่ 6



รูปที่ 7 Pipeline โมเดลแนะนำบอร์ดเกม



รูปที่ 8 Pipeline ของระบบแนะนำบอร์ดเกม

2.4 การพัฒนาเว็บไซต์ (Web Development)

การพัฒนาเว็บไซต์เพื่อนำไปใช้งานกับโมเดล สำหรับกระบวนการพัฒนาจะแบ่งเป็น 3 ส่วนหลัก คือ

2.4.1 FrontEnd Development

FrontEnd เป็นการพัฒนาในส่วนของหน้าตาเว็บไซต์ จะเป็นส่วนที่ผู้ใช้งานเห็นและสามารถโต้ตอบกับผู้ใช้งานได้

ในการพัฒนานั้นจะเป็นการรูปแบบ Single Page Application ที่ซึ่งเว็บไซต์จะถูกแทนที่เนื้อหาแบบไดนามิกด้วยข้อมูลที่ได้จาก Server โดยไม่ต้องมีการ Reload เว็บเพจ นอกจากนี้ยังมีส่วนที่ต้องเชื่อมต่อกับ Service อื่นเพื่อใช้งานข้อมูล เช่น แสดงแผนที่ (Longdo Map API) และการใช้บัญชี Social Media ในการเข้าสู่ระบบเพื่อใช้งาน

2.4.2 BackEnd Development

BackEnd เป็นการพัฒนาในส่วนของเบื้องหลังที่ซึ่ง FrontEnd สามารถเชื่อมต่อผ่าน API (Application Programming Interface) ในการติดต่อสื่อสารกัน มีหน้าที่สำคัญหลายอย่าง ไม่ว่าจะเป็นการประมวลผลข้อมูลจำนวนมาก มีการตรวจสอบและยืนยันตัวของผู้ใช้งาน (Authentication) มีการจำกัดการเข้าถึงในข้อมูลของผู้อื่นโดยที่ตนเองไม่ได้มีสิทธิ (Authorization) มีการเชื่อมต่อกับ API หรือ Service อื่นๆ โดยนำข้อมูลมาคัดกรองพร้อมกับส่งไปแสดงผลที่ส่วนของ FrontEnd และเป็นตัวกลางในการติดต่อกับฐานข้อมูล (Database) เพื่อใช้ในการเก็บข้อมูลของผู้ใช้งาน

2.4.3 Model Boardgame Service Development

เนื่องจากโมเดลแนะนำบอร์ดเกมพัฒนาโดยใช้ภาษาโปรแกรม Python จึงต้องมีการแยกเป็น API Service ให้ส่วนของ Backend ใช้งานอีกที เพื่อความเข้ากันได้ของภาษาโปรแกรมมิ่งและความเรียบง่ายในการจัดการ Package

2.5 ภาพรวมการทำงานของระบบแนะนำบอร์ดเกม

เมื่อนำข้อมูลผู้ใช้จากเว็บไซต์ในหัวข้อ 2.4 มาเข้าโมเดล แนะนำบอร์ดเกมในหัวข้อ 2.3 จะได้ระบบแนะนำบอร์ดเกมโดยมีภาพรวมดังรูปที่ 7 และ 8

3. ผลการดำเนินงานวิจัย

3.1 การประมินผล

3.1.1 Precision และ Recall

Precision (เรียกอีกอย่างว่า positive predictive value) ใช้สำหรับวัดผลว่าเกมที่โมเดลแนะนำไปแล้วผู้ใช้ชอบคิดเป็นอัตราส่วนเท่าไหร่จากเกมที่โมเดลแนะนำไปทั้งหมด

Recall (เรียกอีกอย่างว่า sensitivity) ใช้สำหรับวัดผลว่าเกมที่โมเดลแนะนำไปแล้วผู้ใช้ชอบคิดเป็นอัตราส่วนเท่าไหร่เมื่อเทียบกับเกมที่ผู้ใช้ชอบทั้งหมด โดยมีสูตรดังรูปที่ 9

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

รูปที่ 9 สูตรของ Precision และ Recall [15]

โดยที่

- True Positive (tp): ค่าจริงเป็น True และทำนายได้ True
- False Positive (fp): ค่าจริงเป็น False แต่ทำนายได้ True
- True Negative (tn): ค่าจริงเป็น False และทำนายได้ False
- False Negative (fn): ค่าจริงเป็น True แต่ทำนายได้ False

3.1.2 Normalized Discounted Cumulative Gain (NDCG)

NDCG เป็นตัวชี้วัดคุณภาพการจัดอันดับ (Ranking quality) โดยในงานวิจัยนี้ โมเดลแนะนำบอร์ดเกมที่ใช้นอกจากจะแนะนำเกมที่ผู้ใช้ชอบแล้ว จะต้องมีการเรียงลำดับเกมให้ถูกต้องตามความชอบของผู้ใช้ เพื่อให้ NDCG มีค่าสูง โดยมีสูตรดังรูปที่ 10

$$NDCG = \frac{DCG}{IDCG}$$

รูปที่ 10 สูตรของ NDCG [16]

ซึ่ง NDCG เป็นอัตราส่วนของ Discounted Cumulative Gain (DCG) ของลำดับที่แนะนำต่อ DCG ของลำดับที่แท้จริง (iDCG) โดย DCG มีสูตรดังรูปที่ 11

$$\mathrm{IDCG_k} = \sum_{i=1}^{k} \frac{\mathrm{rel}_i}{\log_2(i+1)}$$

รูปที่ 11 สูตรของ DCG [16]

โดยที่

- k: จำนวนของรายการที่เราพิจารณาในลำดับ (Top-k)
- rel_i: ค่าความเกี่ยวข้อง (relevance) ของรายการในลำดับที่ i โดยค่าจะเป็นตัวเลขที่บ่งบอกถึงความสำคัญของรายการนั้น

3.1.3 เปรียบเทียบประสิทธิภาพของแต่ละโมเดล

งานวิจัยนี้ได้นำเมทริกซ์วัดผลในหัวข้อที่ 3.1.1 และ 3.1.2 มาใช้ในการเปรียบเทียบประสิทธิภาพของแต่ละโมเดลในหัวข้อที่ 2.3 โดยมีขั้นตอนดังนี้

- สุ่มนำบอร์ดเกมที่ผู้ใช้ชอบมาก (Rating >= 9) ออก 10 เกม
- นำ user profile เข้าโมเดลแนะนำบอร์ดเกม
- นำผลลัพธ์ 10 เกมที่โมเดลให้คะแนนว่าเหมาะสมกับผู้ใช้ที่สุดไปประเมินว่าเป็นเกมที่ผู้ใช้ชอบหรือไม่ (Rating > ค่าเฉลี่ย rating ของผู้ใช้)
- วัดผลด้วย precision, recall, NDCG, execution time

ได้ผลของแต่ละโมเดลดังตารางที่ 1. เราได้นำเสนอโมเดลที่ใช้ในการทดลองทั้งหมด 6 โมเดล ซึ่งรวมถึงวิธี Content-Based Filtering (โมเดล 1-3), Collaborative Filtering ทั้งแบบ User-Based (โมเดล 4) และ Item-Based (โมเดล 5) รวมถึงโมเดล Deep Learning (โมเดล 6) เพื่อเปรียบเทียบประสิทธิภาพของวิธีการที่มีความหลากหลาย การเลือกโมเดลเหล่านี้เกิดจากการต้องการทดสอบทั้งวิธีการที่เน้นข้อมูลของผู้ใช้และสินค้า (Content-Based), การใช้ข้อมูลการโต้ตอบของผู้ใช้ (Collaborative Filtering), รวมถึงการใช้โมเดลที่ซับซ้อนเพื่อจับความสัมพันธ์ที่ไม่เป็นเชิงเส้น (Deep Learning) เพื่อให้การแนะนำมีความหลากหลายและแม่นยำมากขึ้น โมเดลที่เลือกมาในการทดลองนี้ครอบคลุมการกรองข้อมูลหลายประเภทและเทคนิคที่แตกต่างกัน เพื่อเปรียบเทียบประสิทธิภาพและความเหมาะสมในแต่ละกรณี โดยเฉพาะโมเดล Item-Based Collaborative Filtering ซึ่งเป็นโมเดลที่มีความเหมาะสมที่สุดสำหรับการใช้งานจริง เนื่องจากมีสมดุลระหว่างความแม่นยำและความสามารถในการปรับใช้ในสภาพแวดล้อมที่มีข้อมูลจำนวนมากและต้องการประมวลผลที่รวดเร็ว

ตารางที่ 1 วัดผลโมเดลแนะนำบอร์ดเกม

Models	Precision	Recall	NDCG	Time
1) Content-based filtering โดยใช้ Numerical Data	0.352	0.011	0.306	0.34
2) Content-based filtering โดยใช้ Categorical Data	0.309	0.010	0.281	0.37
3) Content-based filtering โดยใช้ Numerical Data และ Categorical Data	0.301	0.009	0.276	0.42
4) User-based Collaborative Filtering	0.566	0.018	0.493	1.28
5) Item-based Collaborative Filtering	0.708	0.023	0.680	1.26
6) Bilateral Variational Autoencoder (BiVAE)	0.765	0.050	0.779	29.17

จากตารางที่ 1 โมเดล BiVAE ให้ผลลัพธ์ที่ดีที่สุดทั้งเมทริกซ์ Precision, Recall, NDCG และ Execution Time ทรัพยากรที่ใช้ในการทำงานวิจัยนี้ด้วย Intel(R) Core(TM) i7-9750H CPU @ 2.60GHz 2.59 GHz แต่โมเดลดังกล่าวมีข้อเสียเรื่องการประมวลผลที่นานและจำเป็นต้อง Retrain ทุกครั้งที่มีผู้ใช้ใหม่เพิ่มขึ้นซึ่งนำไปใช้จริงในเว็บไซต์ได้ยาก จึงเลือกใช้โมเดล Item-

based Collaborative Filtering ซึ่งมีประสิทธิภาพที่ดีรองลงมา แต่มีความเร็วในการประมวลผลสูง โดยวัดผลประสิทธิภาพของโมเดลได้ดังรูปที่ 12

นอกจากนี้จากตารางที่ 1 สังเกตุได้ว่า Recall นั้นมีค่าที่ต่ำมากในทุกโมเดล ซึ่งสาเหตุมาจากระบบจะทำการแนะนำเกมเพียง 10 เกม แต่เกมที่ผู้ใช้ชอบนั้นมีตั้งแต่หลักร้อยจนถึงหลายพันเกม ผลลัพธ์ที่ได้คือแม้ว่าเกมที่แนะนำจะเป็นที่ผู้ใช้ชอบทั้ง 10 เกม ก็จะได้ค่า Recall ที่น้อยเสมอ

3.2 ผลการวิจัย

เมื่อรวบรวมข้อมูล โมเดลแนะนำบอร์ดเกม และเว็บไซต์เข้าด้วยกัน จะได้เว็บไซต์ Board Game RecCommu ซึ่งมีระบบต่าง ๆ ดังนี้

3.2.1 ระบบแนะนำบอร์ดเกม

งานวิจัยนี้ได้พัฒนาระบบแนะนำบอร์ดเกม 2 ประเภท คือ ระบบแนะนำสำหรับผู้ใช้ที่ไม่ได้เข้าสู่ระบบ (Guest user) ซึ่งมีไว้ให้ผู้ใช้ได้ทดลองใช้เว็บไซต์ โดยจะแนะนำบอร์ดเกมที่มีความใกล้เคียงกับบอร์ดเกมที่ผู้ใช้ค้นหา และระบบแนะนำสำหรับผู้ใช้ที่เข้าสู่ระบบแล้ว (Logged-in user) ซึ่งจะเป็นการนำ User profile มาใช้เพื่อแนะนำบอร์ดเกมที่เหมาะสมกับผู้ใช้มากขึ้น นอกจากนี้ยังสามารถกรองบอร์ดเกมตามจำนวนผู้เล่น เวลาที่ใช้ ความยาก และประเภทของเกมดังรูปที่ 13 และ 14

3.2.2 ระบบให้คะแนนบอร์ดเกม

เมื่อผู้ใช้ได้ลองเล่นบอร์ดเกมที่แนะนำไป ผู้ใช้สามารถกลับมาให้คะแนน เพื่อเป็นการสร้าง User profile ที่โมเดลสามารถนำไปวิเคราะห์เพิ่มเพื่อให้การแนะนำบอร์ดเกมครั้งต่อ ๆ ไปมีความเหมาะสมกับผู้ใช้มากขึ้น

3.2.3 ระบบปาร์ตี้

ระบบปาร์ตี้สำหรับผู้ใช้ที่ต้องการหาเพื่อนเล่นบอร์ดเกม โดยสามารถใส่บอร์ดเกมที่ต้องการเล่น เวลาอยากที่เล่น จำนวนสมาชิกต้องการ สถานที่เล่นบอร์ดเกม และข้อมูลอื่น ๆ ได้

3.2.4 ระบบค้นหาร้านบอร์ดเกมใกล้เคียง

ผู้ใช้ที่ต้องการหาร้านบอร์ดเกมเล่น สามารถใช้ระบบค้นหาร้านบอร์ดเกมใกล้เคียงเพื่อหาสถานที่เล่น ช่องทางติดต่อร้าน และเส้นทางจากตำแหน่งปัจจุบันไปยังร้านได้ ดังรูปที่ 15

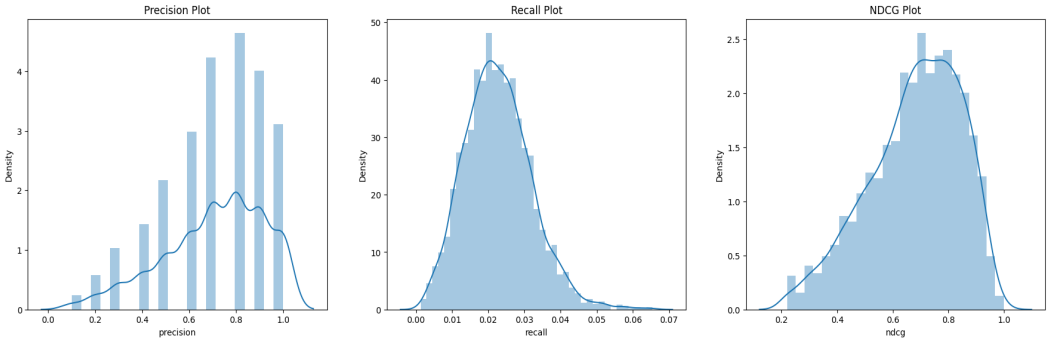
3.2.5 บอร์ดเกมยอดนิยม

ในหน้า Landing Page ของเว็บไซต์จะมีการแสดงบอร์ดเกมที่เป็นที่นิยมในปัจจุบันเพื่อให้ผู้ใช้ทราบว่ามีบอร์ดเกมใดบ้างที่กำลังเป็นกระแส ณ ปัจจุบัน ซึ่งข้อมูลดังกล่าวนำมากจาก BoardGameGeek API [7]

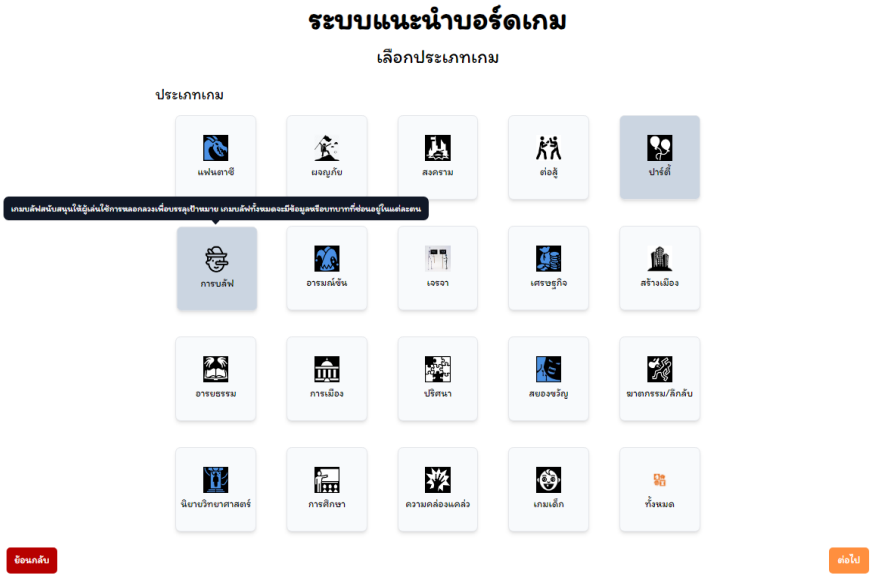
4. สรุปผลการดำเนินงานวิจัย

งานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพโมเดลต่าง ๆ สำหรับระบบแนะนำบอร์ดเกมดังหัวข้อที่ 3.1.3 พบว่า

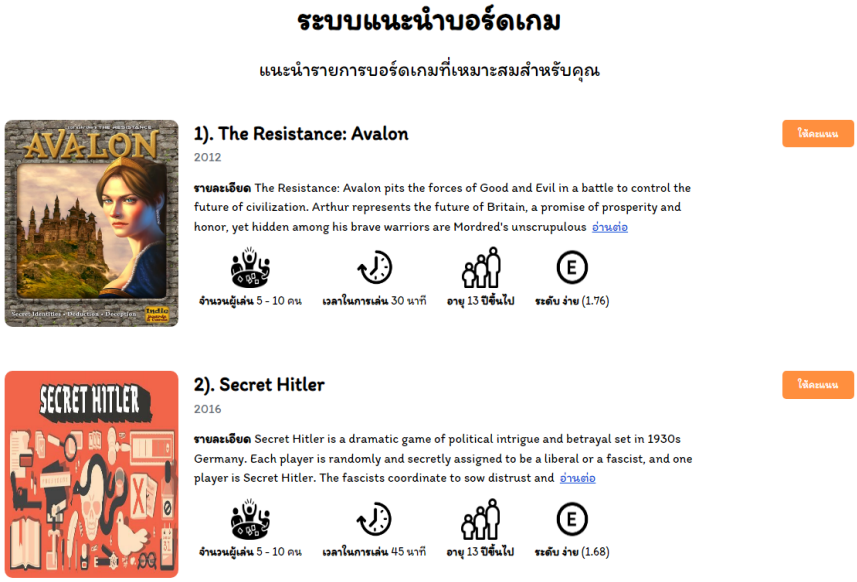
- จากการพัฒนาโมเดล Content-based filtering ทั้ง 3 แบบ พบว่าการใช้ Numerical Data ร่วมกับ Categorical Data ให้ผลลัพธ์ที่แย่ที่สุด ซึ่งการเพิ่ม Feature ให้โมเดลไปเยอะ ๆ ไม่ได้เป็นผลดีเสมอไป อาจส่งผลให้โมเดลทำการเรียนรู้สิ่งที่ไม่จำเป็นหรือมีความผิดพลาด
- โมเดล Content-based filtering มีประสิทธิภาพต่ำสุด เนื่องจากเป็นเพียงการแนะนำบอร์ดเกมที่มี Feature คล้าย ๆ กันเช่น นักออกแบบคนเดียวกัน ผู้เผยแพร่เดียวกัน ปีที่ผลิตปีเดียวกัน เป็นต้น ซึ่งผู้เล่นบอร์ดเกมแต่ละคนให้น้ำหนัก Feature เหล่านี้ไม่เหมือนกัน การใช้ Content-based filtering จึงได้ประสิทธิภาพที่ค่อนข้างต่ำ
- โมเดล Item-based Collaborative Filtering ให้ผลลัพธ์ที่ดีกว่า User-based Collaborative Filtering เนื่องจาก User-based เป็นการหาผู้ใช้ที่คล้ายกันแล้วอ้างอิงการแนะนำจากผู้ใช้คนนั้นเป็นหลัก ซึ่งในความเป็นจริงแม้จะได้ผู้ใช้ที่คล้ายกันก็มักจะมีความชอบบางอย่างที่ไม่ได้เหมือนกัน และยังมีปัญหาผู้ใช้ใหม่ (Cold start problem) ที่ยังมีข้อมูลน้อยทำให้หาผู้ใช้ที่คล้ายกันได้ยาก ซึ่งเมื่อเทียบกับ Item-based การแนะนำจะอ้างอิงจากบอร์ดเกมที่ผู้ใช้ชอบเป็นหลักซึ่งจะได้ประสิทธิภาพที่ดีและนำไปใช้จริงได้ง่ายกว่า
- โมเดล Deep learning (BiVAE) เป็นโมเดลที่มีความซับซ้อนและให้ประสิทธิภาพที่ดี แต่มีข้อจำกัดในเรื่องเวลาในการประมวลผลทำให้ไม่สามารถให้การแนะนำแบบ Real time ได้ จึงจำเป็นต้องใช้โมเดลอื่นแทน



รูปที่ 12 ประสิทธิภาพของโมเดล Item-based Collaborative Filtering ที่นำไปใช้ในระบบแนะนำบอร์ดเกม

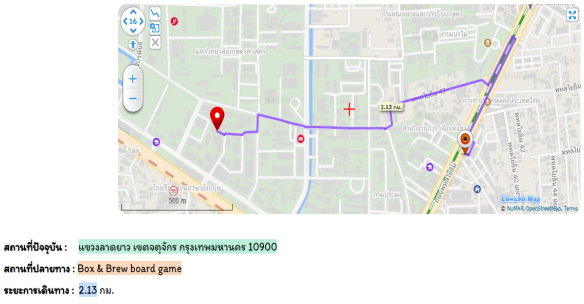


รูปที่ 13 ระบบกรองประเภทบอร์ดเกม



รูปที่ 14 ระบบแนะนำสำหรับผู้ใช้ที่เข้าสู่ระบบ

จากเหตุผลข้างต้นสรุปได้ว่าโมเดลที่มีความแม่นยำสูงที่สุดอาจจะไม่ได้เป็นโมเดลที่เหมาะสมต่อการใช้งานที่สุด ซึ่งในงานวิจัยนี้ได้เลือกใช้โมเดล Item-based Collaborative Filtering ในการทำระบบแนะนำบอร์ดเกม ที่แม้จะมีประสิทธิภาพลดลงเล็กน้อย แต่มีความเหมาะสมกับการใช้งานจริงมากกว่า และมีการประมวลผลที่รวดเร็ว



รูปที่ 15 แสดงเส้นทางจากตำแหน่งปัจจุบันไปยังร้านที่เลือก

5. เอกสารอ้างอิง

[1] Wikimedia Foundation. *(*3 ธันวาคม 2565). *เกมกระดาน,* [ระบบออนไลน์] แหล่งที่มา: https://th.wikipedia.org/wiki/เกมกระดาน

[2] Wikimedia Foundation. *(*3 ธันวาคม 2565). *ยูโรเกม,* [ระบบออนไลน์] แหล่งที่มา: https://th.wikipedia.org/wiki/ยูโรเกม

[3] technavio. *(*24 มกราคม 2567). *Board Games Market Analysis Forecast* 2023-2027, [ระบบออนไลน์] แหล่งที่มา: https://www.technavio.com/report/board-games-market-industry-analysis

[4] ปัทมา เจริญกรกิจ. (3 ธันวาคม 2565). *บอร์ดเกมคาเฟ่พื้นที่แห่งปฏิสัมพันธ์และการสร้างบทสนทนาของความเป็นมนุษย์,* [ระบบออนไลน์] แหล่งที่มา: https://www.thekommon.co/board-game-cafe-2022

[5] BoardGameGeek. *(*3 ธันวาคม 2565). *BoardGameGeek | Gaming Unplugged Since* 2000, [ระบบออนไลน์] แหล่งที่มา: https://boardgamegeek.com

[6] Wikimedia Foundation. *(*3 ธันวาคม 2565). *BoardGameGeek,* [ระบบออนไลน์] แหล่งที่มา: https://en.wikipedia.org/wiki/BoardGameGeek

[7] BoardGameGeek. *(*3 ธันวาคม 2565). *BGG XML API2,* [ระบบออนไลน์] แหล่งที่มา: https://boardgamegeek.com/wiki/page/BGG_XML_API2

[8] BoardGameGeek. *(*3 ธันวาคม 2565). *Category,* [ระบบออนไลน์] แหล่งที่มา: https://boardgamegeek.com/wiki/page/Category

[9] Wizards of Learning. *(*2 ธันวาคม 2566). *รวม 100 ที่เล่นบอร์ดเกม ร้านบอร์ดเกม ห้องสมุด ชมรม ทั่วไทย* Play *Board Games | Wizards of Learning,* [ระบบออนไลน์] แหล่งที่มา: https://wizardsoflearning.com/play-bg-thailand

[10] Ricci, F., Rokach, L., & Shapira, B. *Introduction to Recommender Systems Handbook.* Springer Boston MA, 2010. doi:10.1007/978-0-387-85820-3

[11] *Adomavicius, G., & Tuzhilin, A.* "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering.,* Vol. 17 (6), pp. 734–749, 2005. doi: *10.1109/*TKDE.2005.99

[12] Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl. *(*2001). "Item-based Collaborative Filtering Recommendation Algorithms". *Hypermedia Track of the 10th International World Wide Web Conference. 1-5 May. Hong Kong* : pp. 285-295, 2001. doi: 10.1145/371920.372071

[13] Sirinart Tangruamsub. *(*3 ธันวาคม 2565). *Recommendation System (แบบสรุปไม่ค่อยสั้น),* [ระบบออนไลน์] แหล่งที่มา: https://medium.com/@sinart.t/recommendation-system-แบบสรุปเอาเอง-*ce*6246*f*49754

[14] Jun Ki Min. *(*2 ธันวาคม 2566). *Bilateral Variational Autoencoder (BiVAE),* [ระบบออนไลน์] แหล่งที่มา: https://github.com/recommenders-team/recommenders/blob/main/examples/*02*_model_collaborative_filtering/cornac_bivae_deep_dive.ipynb

[15] Wikimedia Foundation*. (*24 มกราคม 2567). *Precision and recall,* [ระบบออนไลน์] แหล่งที่มา: https://en.wikipedia.org/wiki/Precision_and_recall

[16] Aparna Dhinakaran*. (*24 มกราคม 2567). *Demystifying NDCG,* [ระบบออนไลน์] แหล่งที่มา: https://towardsdatascience.com/demystifying-ndcg-bee*3*be*58*cfe*0*