



การใช้เมทริกซ์ในระบบแนะนำ Recommendation System Applying Matrix Techniques in Recommendation System

เกศธิญา สุทธิพัฒน์อนันต์¹ เปรมยาดา รัตนบุรี¹

อนัน จันทรรัตน์^{1*}

¹โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย สตูล, ตำบลฉลุง, อำเภอเมือง, จังหวัดสตูล 91140, ประเทศไทย

Email: ketsiya.s@pccst.ac.th

บทคัดย่อ

ปัจจุบันระบบแนะนำ (Recommendation System) เป็นเทคโนโลยีสำคัญที่ถูกนำมาใช้ในแพลตฟอร์มออนไลน์อย่างแพร่หลาย เพื่อช่วยคัดกรองและนำเสนอเนื้อหาที่ตรงกับความสนใจของผู้ใช้แต่ละราย โดยเฉพาะในด้านภาพยนตร์และสื่อดิจิทัล ซึ่งมีข้อมูลจำนวนมากและมีความหลากหลายสูง การทำความเข้าใจพฤติกรรมผู้ใช้และแนะนำเนื้อหาให้เหมาะสมจะช่วยเพิ่มประสิทธิภาพของระบบแนะนำเนื้อหาได้อย่างแม่นยำและมีประสิทธิภาพมากยิ่งขึ้น

โครงงานนี้จัดทำขึ้นโดยมีวัตถุประสงค์เพื่อศึกษาหลักการการใช้เมทริกซ์ในระบบแนะนำ และสร้างแบบจำลองระบบแนะนำภาพยนตร์ (Recommendation System) อย่างง่าย โดยนำข้อมูลการให้คะแนนภาพยนตร์ของผู้ใช้มาจัดเก็บในรูปแบบเมทริกซ์ผู้ใช้-ภาพยนตร์ (User-Movie Rating Matrix) จากนั้นนำหลักการทางพีชคณิตเชิงเส้นและสถิติ ได้แก่ การวิเคราะห์ค่าความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) และการการคำนวณค่าเฉลี่ยถ่วงน้ำหนัก มาใช้ในการทำนายความสนใจหรือพยากรณ์คะแนนที่หายไปของผู้ใช้ที่มีต่อภาพยนตร์เรื่องต่าง ๆ ซึ่งเป็นการบูรณาการและประยุกต์ใช้ความรู้ทางคณิตศาสตร์ร่วมกับเทคโนโลยีสารสนเทศ ผลการศึกษพบว่า แบบจำลองระบบแนะนำอย่างง่ายที่พัฒนาขึ้นสามารถวิเคราะห์ความสัมพันธ์ของความชอบระหว่างผู้ใช้ได้อย่างเป็นระบบ และสามารถทำนายแนวโน้มความสนใจของคะแนนที่หายไปได้อย่างสมเหตุสมผล โดยมีค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ (MAE) เท่ากับ 1.50 คะแนน และรากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ประมาณ 1.54 คะแนน ซึ่งอยู่ในระดับที่น่าพอใจและยอมรับได้ แสดงให้เห็นว่าแนวคิดเชิงเมทริกซ์และเวกเตอร์สามารถนำมาเป็นต้นแบบในการพัฒนาระบบแนะนำรายการได้อย่างมีประสิทธิภาพและสามารถนำไปประยุกต์ใช้ได้จริง

คำสำคัญ: เมทริกซ์, Recommendation System, ค่าความคล้ายคลึง, ค่าเฉลี่ยถ่วงน้ำหนัก

1. บทนำ

ในปัจจุบันเทคโนโลยีสารสนเทศและอินเทอร์เน็ต ก่อให้เกิดข้อมูลจำนวนมากมหาศาล (Big Data) จากแพลตฟอร์มต่าง ๆ เช่น เว็บไซต์ดูภาพยนตร์ ร้านค้าออนไลน์ และสื่อสังคมออนไลน์ ส่งผลให้ผู้ใช้ประสบปัญหา “Information Overload” หรือภาวะข้อมูลล้นเกิน จึงจำเป็นต้องมีระบบที่ช่วยคัดกรองและแนะนำเนื้อหาที่

เหมาะสม ซึ่งระบบแนะนำ (Recommendation System) ถูกพัฒนาขึ้นเพื่อตอบโจทย์นี้

หนึ่งในเทคนิคสำคัญคือ Collaborative Filtering ที่อาศัยแนวคิดพีชคณิตเชิงเส้นและเมทริกซ์ในการจัดเก็บข้อมูลคะแนนผู้ใช้ในรูปแบบ Rating Matrix และวิเคราะห์ความสัมพันธ์ด้วยค่าความคล้าย เช่น Cosine Similarity และ Pearson Correlation เพื่อนำไป



พยากรณ์คะแนนที่ผู้ชมมีแนวโน้มจะให้กับรายการที่ยังไม่เคยเลือก แสดงให้เห็นถึงการประยุกต์ใช้เมทริกซ์ เวกเตอร์ และสถิติในการแก้ปัญหาจริง

โครงการนี้จึงมีความสำคัญทั้งด้านทฤษฎีและการประยุกต์ใช้สามารถนำแนวคิดไปพัฒนาระบบแนะนำภาพยนตร์ เพลง หรือสินค้าออนไลน์ได้อย่างมีประสิทธิภาพ พร้อมทั้งส่งเสริมทักษะการคิดวิเคราะห์เชิงคณิตศาสตร์ในยุคดิจิทัล

2. วิธีดำเนินการ

2.1 ศึกษาทฤษฎีและกำหนดกรอบ

ศึกษาหลักการของระบบแนะนำ (Recommendation System) และวิธี Collaborative Filtering แบบอิงผู้ใช้ ซึ่งตั้งอยู่บนแนวคิดที่ว่าผู้ใช้ที่มีพฤติกรรมทำให้คะแนนคล้ายกันในอดีตจะมีแนวโน้มสนใจสิ่งใกล้เคียงกันในอนาคต ศึกษาการจัดเก็บข้อมูลในรูปแบบเมทริกซ์ คะแนน (Rating Matrix) และสูตรทางคณิตศาสตร์ที่เกี่ยวข้อง เช่น Cosine Similarity และสูตรพยากรณ์แบบถ่วงน้ำหนัก เพื่อกำหนดแนวทางการทดลองอย่างชัดเจน

2.2 การจัดเตรียมข้อมูลทดลอง

กำหนดผู้ใช้ 10 คน และภาพยนตร์ 30 เรื่อง ให้คะแนนในช่วง 0-10 คะแนน ข้อมูลที่ได้มีลักษณะไม่ครบถ้วน (Missing Data) เนื่องจากผู้ใช้ไม่ได้ให้คะแนนทุกรายการ จากนั้นจัดเก็บข้อมูลในรูปตาราง และพัฒนาเว็บไซต์ต้นแบบเพื่ออำนวยความสะดวกในการบันทึกข้อมูล

2.3 การสร้างเมทริกซ์คะแนน

นำข้อมูลที่รวบรวมได้มาเขียนในรูปเมทริกซ์

$$R = [r_{u,j}]$$

โดยที่

$u = 1, 2, \dots, m$ แทนผู้ใช้

$j = 1, 2, \dots, n$ แทนรายการ

$r_{u,j}$ แทนคะแนนที่ผู้ใช้ u ให้รายการ j

ค่าที่ไม่มีคะแนนถือเป็นค่าที่ต้องพยากรณ์ เมทริกซ์นี้มีลักษณะเป็น เมทริกซ์เบาบาง (Sparse Matrix)

2.4 การแทนคะแนนในรูปเวกเตอร์

แทนคะแนนของผู้ใช้แต่ละคนในรูปเวกเตอร์ในปริภูมิ n มิติ เพื่อใช้แนวคิดผลคูณเชิงจุด (Dot Product) และความยาวเวกเตอร์ (Vector Norm) ในการวัดความคล้ายระหว่างผู้ใช้

2.5 การคำนวณค่าความคล้าย

คำนวณค่าความคล้ายระหว่างผู้ใช้โดยใช้สูตร Cosine Similarity

$$\text{sim}(u, v) = \frac{R_u \cdot R_v}{\|R_u\| \|R_v\|}$$

โดยที่

$\text{sim}(u, v)$ แทนค่าความคล้ายคลึง ซึ่งมีค่าสูงสุดเท่ากับ 1

R_u แทนเวกเตอร์คะแนนของผู้ใช้ u

R_v แทนเวกเตอร์คะแนนของผู้ใช้ v

หากค่าใกล้ 1 แสดงว่ามีความสนใจคล้ายกัน จากนั้นจัดเรียงค่าเพื่อเลือกกลุ่มผู้ใช้ที่มีความคล้ายสูงสุด

2.6 การพยากรณ์คะแนน

คำนวณค่าเฉลี่ยถ่วงน้ำหนักจากกลุ่มผู้ใช้ที่มีความคล้ายสูง เพื่อประมาณคะแนนที่ผู้ใช้เป้าหมายยังไม่ได้ให้ โดยให้น้ำหนักตามระดับของค่าความคล้ายคลึงจากสูตร

$$r_{u,j} = \frac{\sum_{v \in N_k(u)} \text{sim}(u, v) \cdot r_{v,j}}{\sum_{v \in N_k(u)} |\text{sim}(u, v)|}$$

โดยที่

$r_{u,j}$ แทนคะแนนพยากรณ์ที่ผู้ใช้ u จะให้กับ

รายการ j

$\text{sim}(u, v)$ แทนค่าความคล้ายคลึง

$r_{v,j}$ แทนคะแนนจริงที่ผู้ใช้ v ให้กับรายการ j

$N_k(u)$ แทนกลุ่มของผู้ใช้จำนวน k คนที่มีความคล้ายคลึงสูงสุดกับผู้ใช้ u

2.7 การประเมินและสรุปผล

ประเมินความแม่นยำของแบบจำลองด้วยค่า MAE และ RMSE โดยเปรียบเทียบค่าพยากรณ์กับค่าจริงค่าที่ต่ำแสดง ถึงความแม่นยำสูง จากนั้นจัดเรียงผลลัพธ์



เพื่อแสดงรายการแนะนำ และสรุปผลการทดลองพร้อม
วิเคราะห์ข้อดีข้อจำกัด

3. ผลการทดลอง

3.1 ผลการจัดรูปข้อมูลในลักษณะเมทริกซ์

ข้อมูลคะแนนที่เก็บรวบรวมจากกลุ่มตัวอย่าง
สามารถ จัดให้อยู่ในรูปเมทริกซ์คะแนน $R = [r_{uj}]$ โดยที่
 r_{uj} แทนคะแนนของผู้ใช้ลำดับที่ u ต่อรายการลำดับที่ j
เมทริกซ์ที่ได้มีลักษณะเป็นเมทริกซ์แบบเบาบาง (Sparse
Matrix) เนื่องจากผู้ใช้แต่ละคนไม่ได้ให้คะแนนครบทุก
รายการ ส่งผลให้มีตำแหน่งของข้อมูลที่หายไป ซึ่งเป็นค่าที่
ต้องการพยากรณ์

3.2 ผลการคำนวณค่าความคล้ายระหว่างผู้ใช้

เมื่อแทนคะแนนของผู้ใช้แต่ละคนในรูปเวกเตอร์
ในปริภูมิหลายมิติ แล้วนำมาคำนวณค่าความคล้ายด้วย
สูตร Cosine Similarity พบว่าผู้ใช้บางกลุ่มมีค่าความ
คล้ายกันในระดับสูง แสดงถึงรูปแบบการให้คะแนนที่
ใกล้เคียงกัน ค่าความคล้ายที่ได้มีค่าอยู่ในช่วง 0 ถึง 1 โดย
ค่าที่เข้าใกล้ 1 หมายถึงมีความคล้ายสูง ผลการจัดอันดับ
ความคล้ายช่วยให้สามารถเลือกกลุ่มผู้ใช้ที่เหมาะสม
สำหรับนำไปใช้ในการพยากรณ์คะแนน

3.3 ผลการพยากรณ์คะแนนที่หายไป

เมื่อเลือกกลุ่มผู้ใช้ที่มีความคล้ายสูงแล้วได้นำ
สูตร ค่าเฉลี่ยถ่วงน้ำหนักมาใช้ในการพยากรณ์คะแนนใน
ตำแหน่งที่ยังไม่มีข้อมูล

ผลการพยากรณ์แสดงให้เห็นว่าสามารถ
ประมาณค่าคะแนนที่หายไปได้อย่างสมเหตุสมผลโดยค่าที่
ได้อยู่ในช่วงคะแนนที่กำหนดไว้ และสะท้อนแนวโน้ม
พฤติกรรมของผู้ใช้เป้าหมายได้อย่างเหมาะสม จากการ
คำนวณครบทุกตำแหน่งที่ต้องการพยากรณ์ ระบบสามารถ
จัดลำดับรายการตามค่าคะแนนที่คาดการณ์ได้ และเลือก
รายการที่มีคะแนนสูงสุดเป็นรายการแนะนำ

3.4 ผลการประเมินความแม่นยำของแบบจำลอง

เพื่อประเมินประสิทธิภาพของแบบจำลอง ได้ใช้
ตัวชี้วัดทางสถิติ ได้แก่ ค่า Mean Absolute Error (MAE)
และ Root Mean Square Error (RMSE)

ผลการประเมินพบว่าค่าความคลาดเคลื่อนจาก
การทดลองอยู่ในระดับต่ำ ซึ่งได้ค่า MAE เท่ากับ 1.50
คะแนน และค่า RMSE เท่ากับ 1.54 คะแนน แสดงให้เห็น
ว่าแบบจำลองสามารถพยากรณ์คะแนนได้ใกล้เคียงกับค่า
จริง และมีความแม่นยำในระดับที่ยอมรับได้

3.5 การวิเคราะห์ผลเชิงคณิตศาสตร์

ผลการดำเนินงานแสดงให้เห็นว่าแนวคิดเรื่อง
เวกเตอร์ ในปริภูมิหลายมิติสามารถนำมาใช้แทน
พฤติกรรมการให้คะแนนของผู้ใช้ได้อย่างมีประสิทธิภาพ
การใช้ Cosine Similarity ช่วยวัดความใกล้เคียงของ
รูปแบบข้อมูลโดยพิจารณาจากมุมระหว่างเวกเตอร์

นอกจากนี้ การใช้ค่าเฉลี่ยถ่วงน้ำหนักยังช่วยให้
คะแนนของผู้ใช้ที่มีความคล้ายสูงมีอิทธิพลมากกว่าผู้ใช้ที่มี
ความคล้ายต่ำ ซึ่งสอดคล้องกับหลักการทางพีชคณิตเชิง
เส้นและสถิติ

4. สรุปและอภิปรายผล

4.1 อภิปรายผล

จากผลการศึกษา พบว่า การแทนข้อมูลในรูป
เวกเตอร์ช่วยให้สามารถใช้แนวคิดเรื่องมุมระหว่างเวกเตอร์
มาอธิบายความคล้ายของพฤติกรรมผู้ใช้ได้อย่างชัดเจน ซึ่ง
เป็นการเชื่อมโยงแนวคิดทางคณิตศาสตร์กับการ
ประยุกต์ใช้ในชีวิตจริง อย่างไรก็ตามประสิทธิภาพของ
แบบจำลองขึ้นอยู่กับปริมาณและความสมบูรณ์ของข้อมูล
หากจำนวนข้อมูลมีน้อย หรือเมทริกซ์มีความเบาบางมาก
อาจทำให้ค่าความคล้ายที่คำนวณได้ไม่เสถียร และส่งผล
ต่อความแม่นยำของการพยากรณ์

นอกจากนี้ วิธี User-Based Collaborative
Filtering จะทำงานได้ดีเมื่อมีผู้ใช้ที่มีพฤติกรรมใกล้เคียง
กันจำนวนมาก แต่หากเป็นผู้ใช้ใหม่ที่ยังไม่มีประวัติการให้
คะแนน (ปัญหา Cold Start) ระบบจะไม่สามารถคำนวณ
ความคล้ายได้อย่างมีประสิทธิภาพ ถึงแม้จะมีข้อจำกัด
ดังกล่าวแต่ผลการทดลองในโครงการนี้แสดงให้เห็นว่า
แนวคิดทางพีชคณิตเชิงเส้นสามารถนำมาประยุกต์ใช้ได้
จริง และให้ผลลัพธ์ที่สอดคล้องกับทฤษฎีทางคณิตศาสตร์



4.2 สรุปผลการศึกษา

จากการดำเนินงาน พบว่า สามารถคำนวณค่าความคล้ายระหว่างผู้ใช้ด้วยสูตร Cosine Similarity ได้อย่างมีประสิทธิภาพ และสามารถนำค่าความคล้ายดังกล่าวไปใช้ในการพยากรณ์คะแนนที่หายไปด้วยวิธีค่าเฉลี่ยถ่วงน้ำหนักได้อย่างเหมาะสม ผลการประเมินด้วยตัวชี้วัดทางสถิติ เช่น MAE และ RMSE แสดงให้เห็นว่าค่าความคลาดเคลื่อนอยู่ในระดับต่ำ แสดงว่าแบบจำลองที่พัฒนาขึ้นมีความแม่นยำอยู่ในระดับที่น่าพอใจ และสามารถนำไปประยุกต์ใช้ในระบบแนะนำได้จริง

จากผลการทดลองเบื้องต้น สามารถสรุปได้ว่า แนวคิดทางเมทริกซ์และเวกเตอร์ในพีชคณิตเชิงเส้นสามารถนำมาประยุกต์ใช้ในการนำมาเป็นต้นแบบของระบบแนะนำรายการได้อย่างมีประสิทธิภาพ

ข้อเสนอแนะ

1. ควรเพิ่มจำนวนผู้ใช้และจำนวนรายการให้มากขึ้น เพื่อให้เมทริกซ์มีข้อมูลที่สมบูรณ์มากขึ้น และช่วยเพิ่มความแม่นยำของการพยากรณ์
2. สามารถทดลองเปรียบเทียบวิธีการวัดความคล้ายรูปแบบอื่น เช่น Pearson Correlation เพื่อตรวจสอบว่าวิธีใดให้ผลแม่นยำมากกว่า
3. อาจพัฒนาแบบจำลองต่อยอดโดยใช้เทคนิคทางพีชคณิตเชิงเส้นขั้นสูง เช่น การแยกเมทริกซ์ (Matrix Factorization) เพื่อปรับปรุงประสิทธิภาพของระบบ
4. ในอนาคตสามารถพัฒนาเว็บไซต์ต้นแบบให้รองรับผู้ใช้จริง และเก็บข้อมูลในปริมาณมากเพื่อทดสอบประสิทธิภาพในสถานการณ์จริง

กิตติกรรมประกาศ

คณะผู้จัดทำโครงการเรื่อง “การใช้เมทริกซ์ในระบบแนะนำ Recommendation System” ขอขอบพระคุณคุณครูอนัน จันทรัตน์ที่ได้ให้คำแนะนำ ข้อเสนอแนะ ตลอดจนแนวทางในการวิเคราะห์และปรับปรุงโครงการในทุกขั้นตอนอย่างต่อเนื่อง ท้ายที่สุดนี้ คณะผู้จัดทำหวังเป็นอย่างยิ่งว่า โครงการชิ้นนี้จะเป็นประโยชน์ต่อการศึกษาและสร้างความเข้าใจเกี่ยวกับการประยุกต์ใช้เมทริกซ์ในระบบแนะนำภาพยนตร์ ตลอดจน

เป็นแนวทางในการพัฒนาระบบแนะนำที่มีประสิทธิภาพ และตอบสนองความต้องการของผู้ใช้งานได้ดียิ่งขึ้นต่อไป

เอกสารอ้างอิง

- Aggarwal, C. C. (2016). Recommender systems: The textbook. - . Springer. สืบค้นเมื่อ 5 กุมภาพันธ์ 2569, จาก <https://link.springer.com/book/10.1007/978-3-319-29659-3>.
- Goodfellow, I., Bengio, Y., & Courville, A. (2015). Deep learning. - . MIT Press. สืบค้นเมื่อ 3 กุมภาพันธ์ 2569, จาก <https://www.deeplearningbook.org>.
- Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer. 42(8): 30–37. สืบค้นเมื่อ 5 กุมภาพันธ์ 2569, จาก <https://ieeexplore.ieee.org/document/5197422>.
- Lay, D. C., Lay, S. R., & McDonald, J. J. (2016). Linear algebra and its applications. ครั้งที่ 5. - . Pearson. สืบค้นเมื่อ 4 กุมภาพันธ์ 2569, จาก <https://broman.dev/download/Linear%20Algebra%20and%20its%20Applications%205th%20Edition.pdf>.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. - . Cambridge University Press. สืบค้นเมื่อ 10 กุมภาพันธ์ 2569, จาก <https://nlp.stanford.edu/IR-book/>.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. In Proceedings of the 1994 ACM conference on Computer supported cooperative work (หน้า 175-186). - . ACM. สืบค้นเมื่อ 12 กุมภาพันธ์ 2569, จาก



<https://dl.acm.org/doi/10.1145/192844>.

192905.

Ricci, F., Rokach, L., & Shapira, B. (2015).

Recommender systems handbook. ครั้งที่ 2. - .

Springer. สืบค้นเมื่อ 8 กุมภาพันธ์ 2569, จาก

<https://link.springer.com/book/10.1007/978-1-4899-7637-6>.

Su, X., & Khoshgoftaar, T. M. (2009). A survey of

collaborative filtering techniques. Advances

in artificial intelligence. 2009: 1–19. สืบค้นเมื่อ

12 กุมภาพันธ์ 2569, จาก

<https://doi.org/10.1155/2009/421425>.