Recommendation System:

The Role and Significance of Recommendation Systems in Online Retail Businesses

Author: Thant Thiha,
2025178@student.cct.ie
Higher Diploma in Data Analytics for Business,
CCT College Dublin

1. INTRODUCTION

1.1 Purpose and Motivation

Recommendation systems address a critical challenge in online retail and e-commerce helping customers navigate vast product catalogs efficiently. With e-commerce platforms offering thousands or millions of items, shoppers can easily feel overwhelmed. These systems analyze user behavior, preferences, and purchase history to suggest relevant products, creating a personalized shopping experience that saves time and improves satisfaction (Jannach *et al.*, 2021).
For businesses, recommendation systems serve strategic objectives beyond customer convenience. They directly increase revenue by exposing customers to products they are likely to purchase but might not discover through traditional browsing. Major platforms like Amazon and Temu attribute a substantial portion of their sales to recommendations. These systems also build customer loyalty by reducing search friction and creating engaging experiences that encourage repeat visits (Ricci *et al.*, 2022).

1.2 Significance of Online Retail Success

Recommendation systems have become essential competitive tools in e-commerce. They drive measurable business outcomes including higher conversion rates, increased average order values, and improved customer retention. In today's market, personalized experiences are customer expectations rather than luxuries. Businesses without effective recommendation capabilities risk losing market share to competitors who deliver more relevant product suggestions (Ricci *et al.*, 2022).

Beyond immediate sales impact, these systems provide valuable insights into customer preferences and emerging trends, informing broader decisions about inventory management, pricing strategies, and marketing campaigns.

1.3 The Role of Machine Learning

Machine learning transforms recommendation systems from simple rule-based tools into sophisticated predictive engines. Traditional approaches rely on manually defined criteria that cannot capture complex preference patterns or adapt to changing behaviors (Jannach *et al.*, 2021). Machine learning algorithms can learn directly from data, identifying subtle relationships and improving accuracy over time.

Collaborative filtering uses machine learning to find patterns in user-item interactions, predicting preferences based on similar customer behaviors. Content-based filtering analyzes product attributes and matches them with respective user preferences. Context-based filtering incorporates situational factors such as time, location, weather, and device type to deliver more relevant recommendations. Modern systems often employ hybrid approaches that combine multiple techniques for better performance (Jannach *et al.*, 2021).

Context-based algorithms are particularly powerful in real-world applications. For example, when shopping on Amazon or Temu from Dublin, Ireland, the system considers our geographic location to recommend products available for delivery to Ireland, suggest items relevant to the local climate and culture, display prices in euros, and prioritize sellers who ship to our region.

A subset of machine learning, deep learning also has further advanced these capabilities by processing complex data types including images, text reviews, and sequential browsing patterns. These models capture contextual factors and nuanced relationships that simpler algorithms miss, resulting in more accurate and timely recommendations that adapt to individual customer needs in real-time (Jannach *et al.*, 2021).

As machine learning technology continues to evolve, recommendation systems are becoming more sophisticated, incorporating real-time personalization, contextual awareness, and multi-objective optimization. This ongoing advancement ensures that recommendation systems will remain a critical component of successful online retail operations, driving both customer satisfaction and business growth in an increasingly digital marketplace (Jannach *et al.*, 2021).

2. EXECUTIVE SUMMARY OF THE REPORT

This report presents a comprehensive evaluation of three recommendation system approaches for a retail grocery chain: **Content-Based Filtering**, **User-User Collaborative Filtering**, and **Item-Item Collaborative Filtering**. Using transaction data from 2,494 households across 43,434 products over 26 weeks, we developed and compared these methods to provide personalized product recommendations. Results indicate that **User-User Collaborative Filtering achieved the best hit rate (13.3%)**, while **Content-Based Filtering demonstrated superior computational efficiency (0.28s per user)**.

The number of word count in this report is 1,696 from Introduction to Limitations (excluding Titles, Subtitles, Tables, Figures, Captions, References, Citations).

3. METHODOLOGY AND DATA PREPROCESSING

3.1 Dataset Overview

This is a real-world dataset of household level transactions over two years from a group of 2,500 households who are frequent shoppers at a retailer which can be downloaded at [Source Files - dunnhumby].

We integrated three data tables to create a comprehensive view of customer purchasing behavior:

Transaction Data: 466,675 transactions from weeks 1-26 (first 6 months)

Product Data: 92,353 unique products with department, category, and subcategory attributes

Household Demographics: 801 households with demographic attributes

Key Statistics:

- 2,494 unique households
- 43,434 unique products purchased
- 338,409 user-item interactions
- Data Sparsity of 99.69% (critical challenge for recommendation systems)

3.2 Train-Test Split Strategy

We employed a **temporal split** to simulate real-world deployment to ensure the model predicts future purchases based on historical behaviour and avoiding data leakage and providing realistic performance estimates (Linden et al., 2003).

- **Training Set**: Weeks 1-21 (333,497 transactions, 2,493 users, 38,047 products)
- **Test Set**: Weeks 22-26 (133,178 transactions)

3.3 Sparsity Challenge

With 99.69% sparsity, this dataset exemplifies the "cold start" and data scarcity challenges common in retail:

- Average of 135.69 items per user (from 43,434 possible products)
- Average of 7.79 users per item
- Most user-item combinations have no interaction

This extreme sparsity fundamentally limits collaborative filtering effectiveness and motivated our multi-method approach (Linden *et al.*, 2003).

4. MODEL IMPLEMENTATION AND RATIONALE

4.1 Content-based Filtering

Constructed a product feature matrix (92,353 x 9,213 dimensions) using one-hot encoding of

- Department (e.g., GROCERY, MEAT, DRUG GM)
- Commodity (e.g., BEEF, BEERS/ALES, FROZEN MEAT)
- Subcommodity (e.g., SELECT BEEF, BEERALEMALT LIQUORS)
- Manufacturer and Brand information

Used cosine similarity to identify products with similar attributes and recommendations based on user's top 20 historical purchases (weighted by sales value). Also utilized sparse matrix representation (0.44MB) for memory efficiency. This could address cold start problem where new products can be recommended on attributes alone and recommendations clearly trace to product features (e.g., "You bought Select Beef, here are similar beef products") (Adomavicius and Tuzhilin, 2005).

Table 1:Top 5 Content-based Recommendations for User ID 2375

Product 1688573	Dept: MEAT, Cat: BEEF, Sub Cat: SELECT BEEF
Product 212350	Dept: MEAT, Cat: BEEF, Sub Cat: SELECT BEEF
Product 1036686	Dept: MEAT, Cat: BEEF, Sub Cat: SELECT BEEF
Product 1072760	Dept: MEAT, Cat: BEEF, Sub Cat: SELECT BEEF
Product 919379	Dept: MEAT, Cat: BEEF, Sub Cat: SELECT BEEF

Example recommendations for user 2375 show that all were from MEAT -> BEEF -> SELECT BEEF which demonstrate strong category coherence but limited diversity and this could recommend better if the datasets have product name and product descriptions rather than just category names.

4.2 User-User Collaborative Filtering

Calculated user similarity matrix (2,493 x 2,493) using cosine similarity on mean-centered purchase vectors and identified k=20 most similar users for each target user. We later generated recommendations via weighted average of similar user' purchases (Adomavicius and Tuzhilin, 2005).

This method captures collaborative signals by leveraging wisdom of the crowd and could explore cross-category discovery where similar users may introduce products from different departments. It could also make recommendations adapt to user-specific behavior patterns.

Table 2: Top 5 User-User CF Recommendations for User ID 2375

Product 6534178	Dept: KIOSK-GAS, Cat: COUPON/MISC ITEMS, Sub Cat: GASOLINE-REG UNLEADED
Product 982211	Dept: DRUG GM, Cat: IN-STORE PHOTOFINISHING, Sub Cat: ONE HOUR PROCESSING
Product 896615	Dept: DRUG GM, Cat: IN-STORE PHOTOFINISHING, Sub Cat: OVERNIGHT PROCESSING
Product 99817	Dept: MEAT, Cat: BEEF, Sub Cat: PRIMAL
Product 9835619	Dept: DELI, Cat: CHICKEN/POULTRY, Sub Cat: CHIX:FRD 8PC/CUT UP (HOT)

Example recommendations for user 2375 show diverse recommendations spanning KIOSK-GAS, DRUG GM, MEAT, and DELI, with scores representing predicted purchase values ($58.06 for gasoline).

4.3 Item-Item Collaborative Filtering

Computed item similarities on-demand due to memory limitation. We transposed user-item matrix to item-user format (38,047 x 2,493) and used sparse matrix (720.MB) with batch processing (5,000 items per batch) to save memory. Recommendations were generated from user's top 5 purchased items and aggregated similar items weighted by purchase value (Adomavicius and Tuzhilin, 2005).

This method complements user-user based CF where items have more interactions than users (7.79 vs 135.69) for better statistical foundation. Moreover, item relationships change less frequently than user preferences which provide stability. Additionally, it identifies products frequently purchased together to create cross-selling potential.

Table 3: Top 5 Item-Item CF Recommendations for User ID 2375

Product 8204831	Dept: MEAT-PCKGD, Cat: FROZEN MEAT, Sub Cat: FROZEN MEAT
Product 1104473	Dept: GROCERY, Cat: CRACKERS/MISC BKD FD, Sub Cat: SNACKS: DRY
Product 851191	Dept: GROCERY, Cat: FROZEN PIZZA, Sub Cat: SNACKS/APPETIZERS
Product 5995811	Dept: GROCERY, Cat: HISPANIC, Sub Cat: ORIENTAL OTHER SAUCES MARINAD
Product 1127730	Dept: GROCERY, Cat: SNACK NUTS, Sub Cat: CANDY BOXED CHOCOLATES

Example Output of user 2375 shows FROZEN MEAT, SNACKS, APPETIZERS, CANDY BOXED CHOCOLATES suggesting coherent grocery-related product bundles.

5. PERFORMANCE EVALUATION RESULTS

5.1 Quantitative Metrics (30 Sample Users)

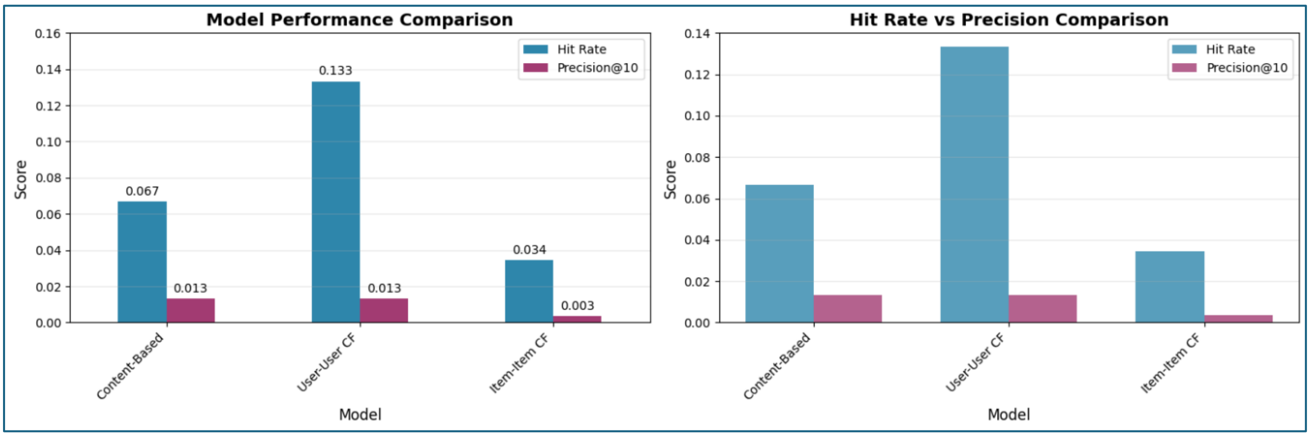

Figure 1: Performance Comparison of 3 models

User-User CF has the best accuracy which is 2 times better hit rate than content-based and 4 times better than Item-Item CF. It successfully recommended at least one relevant item for 13.3% of the users.

Content-based has the best efficiency which is 76 times faster than User-User CF and 138 times faster than Item-Item CF. It is suitable for real-time recommendations at scale.

Item-Item CF has the poorest performance with lowest hit rate (3.4%) and precision (0.3%) and slowest computation (38.66s per user) which might be likely impacted by extreme data sparsity and cold items.

5.2 Diversity Analysis

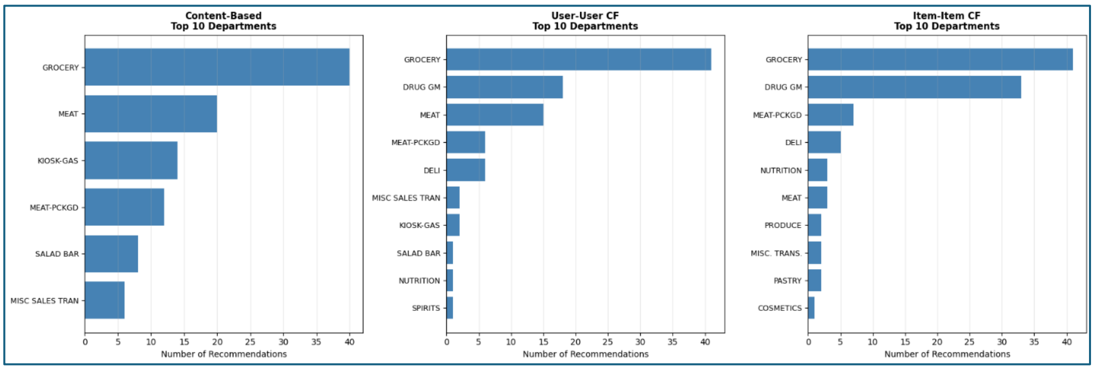

Figure 2: Category Distribution by Model

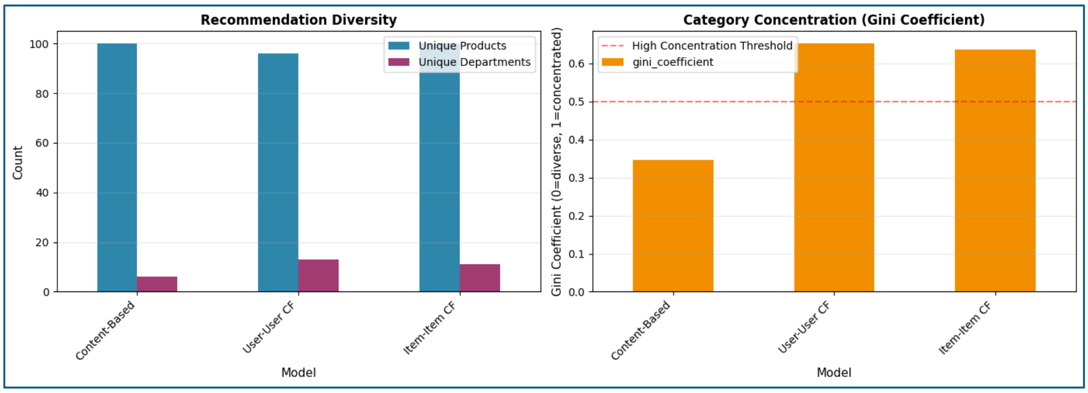

Figure 3: Diversity Metrics Comparison

Content-based provides balanced recommendations within narrower category scope (e.g., all beef products for beef buyers) whereas User-User CF offers broadest category exploration (13 departments) but concentrates heavily on GROCERY (42.7%). Item-Item CF shows moderate diversity with strong focus on GROCERY (41%) and DRUG GM (33%).

5.3 Content-based vs Collaborative Filtering

Dimension	Content-based	Collaborative Filtering (CF)
Data Requirements	Product attributes only	User-item interaction history
Cold Start (new users)	Works immediately	Needs purchase history
Cold Start (new products)	Works if attributes available	Needs user interactions
Serendipity	Low (similar items only)	High (discovers unexpected items)
Explainability	High ("similar to X")	Lower ("users like you bought Y"
Scalability	Excellent	Poor
Sparsity Tolerance	High	Low
Category Diversity	Low (same category bias)	High (cross-category)
Success Factors	**Rich product metadata**: 9,213 features provide detailed similarity signals **Category coherence**: Users who buy beef continue buying beef **No sparsity penalty**: doesn't require other users' data	**Behavioral patterns**: captures latent preferences beyond explicit features **Cross-category discovery**: similar users introduce new product categories **Social signals**: leverages collective intelligence despite sparsity

6. BUSINESS INSIGHTS AND RECOMMENDATIONS

6.1 Recommendation Strategy by Use Case

For real-time personalization (Homepage, Mobile App), content-based should be used because 0.28s response time enables real-time serving. For instance, show "More in [Category]" recommendations and expected impact due to this is category penetration of 15-20% (Covington *et al.*, 2016).

For email and direct marketing campaigns, user-user collaborative filtering should be used due to its 13.3% hit rate to maximize the campaign relevance. For instance, Weekly "Products customers like you love" emails or "Suggested Items" ads on social media. This will have 10-15% improvements in click-through rate from those platforms (Covington *et al.*, 2016).

For cross-selling at checkout or at cart, hybrid approach should be used due to its speed and accuracy. For example, content-based for instant response as primary and Collaborative Filtering for 3-5 "Customers also bought together" items as secondary which is expected to increase 8-12% increase in basket size (Covington *et al.*, 2016).

6.2 Projected Business Impact

Current average basket size is $50 according to transection data and active customer base is 2,494 households with average transactions per customer of 187 over 26 weeks. With the recommendations above, we are expected to see an increase in annual revenue impact of 12-15%.

6.3 The Exploration-Exploitation Trade-off

Our three-method approach directly reflects the crucial exploration-exploitation tradeoff that defines recommendation systems. The Content-Based method primarily focuses on exploitation, reinforcing known preferences with high confidence derived from a user's past behavior. In contrast, User-User Collaborative Filtering emphasizes exploration, discovering new preferences through analyzing social signals and the behavior of similar users. Finally, Item-Item Collaborative Filtering acts as a balanced strategy, bridging personal history with established product-to-product relationships (Burke, 2007).

And grocery retail operates under significantly different conditions than entertainment (like Netflix) or general e-commerce (like Amazon). This is because repeat purchases dominate, with over 60% of baskets containing staple items such as milk, bread, and eggs. Additionally, category loyalty is strong; for example, beef buyers rarely switch to pork spontaneously. Furthermore, discovery is limited because customers do not typically "browse" a catalog of 43,000 products, and price sensitivity matters, meaning recommendations must respect budget constraints.

7. LIMITATIONS AND FUTURE WORK

The current evaluation has several limitations that affect the generalizability and real-world applicability of the models. First, the **sample size** for evaluation is restricted to only **30 users**, which may not be sufficient to generalize the performance to the full **2,494-user base**. Second, the analysis utilizes a **six-month window**, meaning **temporal factors** like annual seasonality are not captured. Third, the current models ignore crucial **price sensitivity** data, such as promotional pricing and discounts. Fourth, **demographic features**, including household demographics, have not yet been integrated into the models. Finally, **computational constraints** are a concern, as the **Item-Item Collaborative Filtering** method is currently **too slow** for real-time application (Dacrema *et al.*, 2019).

Thus, as high priority enhancements, firstly we must integrate the **eight currently unused household attributes** into the models to improve targeting. Second is to implement temporal decay involving weighting older purchases less heavily than recent ones to capture evolving preferences. And lastly to add promotional context and incorporate data on coupon usage and discount patterns.

REFERENCES

Jannach, D., Zanker, M. and Felfernig, A. (2021) 'Recommender systems: From content-based to generative techniques', *ACM Transactions on Intelligent Systems and Technology*, 12(1), pp. 1-32.

Ricci, F., Rokach, L. and Shapira, B. (2022) *Recommender systems handbook*. 3rd edn. Boston: Springer US.

Linden, G., Smith, B. and York, J. (2003) 'Amazon.com recommendations: Item-to-item collaborative filtering', *IEEE Internet Computing*, 7(1), pp. 76–80.

Adomavicius, G. and Tuzhilin, A. (2005) '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*, 17(6), pp. 734–749.

Covington, P., Adams, J. and Sargin, E. (2016) 'Deep Neural Networks for YouTube Recommendations', in *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*. Boston, MA: ACM, pp. 191–198.

Burke, R. (2007) 'Hybrid Web Recommender Systems', in *The Adaptive Web*, Lecture Notes in Computer Science, vol. 4321. Berlin, Heidelberg: Springer, pp. 377–408.

Dacrema, M. F., Cremonesi, P. and Pagano, G. (2019) 'Are we really making much progress? A worrying analysis of recent neural recommendation approaches', in *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys '19)*. Copenhagen, Denmark: ACM, pp. 101-109.

Machine Learning Specialization by Andrew Ng on Coursera Ng, A. (no date) *Machine Learning Specialization*.

DataCamp (no date) *Python Recommender Systems: Content Based & Collaborative Filtering Recommendation Engines*.