Shopping Recommender System Report

1. Introduction

In the rapidly evolving e-commerce industry, personalized product recommendations have become a critical component for enhancing user experience, increasing customer engagement, and boosting sales. As consumers are bombarded with choices, an effective shopping recommender system can help users discover products tailored to their preferences and browsing history. This project aims to develop a shopping recommender system that leverages customer data to suggest relevant products, thereby increasing the likelihood of purchases and overall customer retention.

2. Problem Statement

The main objective of this project is to develop an effective shopping recommender system that:

- Recommends products based on user preferences.
- Enhances the shopping experience with personalized suggestions.
- Increases customer engagement and satisfaction.
- Leverages various recommendation techniques to deliver optimal solutions.

3. Project Scope

The project focuses on the following aspects:

- **Data Collection**: Gathering data related to customer profiles, purchase history, and product information.
- **Data Preprocessing**: Cleaning and transforming data for modelling.
- **Recommendation Techniques**: Implementing collaborative filtering, content-based filtering, and hybrid models.
- **Evaluation Metrics**: Assessing the model's performance using precision, recall, F1-score, and mean average precision.

4. Methodology

4.1 Data Collection

Data was collected from three primary sources:

- **Customer Data**: Information about customers, including demographics and profiles. (Includes features like customer ID, date of birth (DOB), gender, and city code)
- **Transaction Data**: Historical purchase records that detail what products customers have bought. (Includes transaction ID, product category codes, subcategory codes, quantities purchased, rates, taxes, and total amounts)

- **Product Category Information**: Details about product categories and subcategories to aid in recommendations.

.**4.2 Data Preprocessing**

Cleaning

Data cleaning involved removing duplicate transaction IDs, handling missing values, and converting negative quantities and rates to their absolute values. We ensured the datasets were free from inconsistencies before merging.

Transformation

The data was transformed by:

- Converting date columns to a datetime format.
- Creating a new feature for 'age' derived from the DOB.
- One-hot encoding categorical variables to prepare them for the recommendation models.

4.3 Recommendation Techniques

4.3.1 Collaborative Filtering

Collaborative filtering utilizes historical interactions between users and products. Collaborative filtering was implemented using Singular Value Decomposition (SVD) to uncover latent features that could explain user-item interactions. This method analyses the user-item matrix and provides recommendations based on similar users' behaviours.

4.3.2 Content-Based Filtering

The content-based filtering approach involved creating user profiles based on their purchase history. By analysing product categories and subcategories, we employed cosine similarity to recommend products similar to those previously purchased by the user.

4.3.3 Hybrid Model

The hybrid model combines both collaborative and content-based filtering techniques to enhance recommendation accuracy. By assigning weights to both models (collaborative: 0.6, content: 0.4), the system generated a more robust recommendation list, leveraging the strengths of both methodologies.

5. Evaluation Metrics

The performance of the recommender system was assessed using the following metrics:

- **Precision**: The ratio of relevant items recommended to the total items recommended.
- **Recall**: The ratio of relevant items recommended to the total relevant items available.
- **F1-Score**: The harmonic mean of precision and recall, providing a single metric to evaluate the system’s accuracy.

The system achieved a precision of 0.67, recall of 1.0, and an F1-score of 0.8, indicating a robust recommendation capability.

6. Results

6.1 Collaborative Recommendations

The top 5 collaborative product recommendations for user 266783 include:

Electronics	Computers	1.368258
Clothing	Kids	1.212041
Home and kitchen	Furnishing	0.825129
	Bath	0.778332
	Tools	0.734618

6.2 Content-Based Recommendations

The content-based filtering approach produced the following top recommendations:

prod_cat_Bags	Recommended
prod_cat_Books	Recommended
prod_cat_Clothing	Not Recommended
prod_cat_Electronics	Not Recommended
prod_cat_Footwear	Recommended

6.3 Hybrid Recommendations

The hybrid model provided recommendations as follows:

Home and kitchen	Tools	971
Electronics	Mobiles	947
Footwear	Women	943
Books	Fiction	939
	Children	933

7. Conclusion

The shopping recommender system developed in this project effectively enhances user experience by providing personalized and relevant product suggestions. By employing collaborative filtering, content-based filtering, and hybrid models, the system demonstrates a high level of accuracy and user satisfaction.

Implementing a recommender system can significantly enhance customer experience by providing tailored suggestions that meet individual preferences. This personalization increases the chances of repeat purchases and customer retention. For the e-commerce organization, such a system can lead to higher conversion rates and improved customer satisfaction metrics.

8. Future Work

Future enhancements could include:

- Incorporating user feedback to refine recommendations further.
- Implementing deep learning techniques for more sophisticated recommendations.
- Exploring additional metrics for a more nuanced evaluation of recommendations.

9. Appendices

- **Source Code**: The complete source code for the recommender system is included.
- **User Guide**: Instructions for installing and running the recommender system are provided.