**GitHub link -**

**Enhancing Customer Experience with a Hybrid Recommendation System that uses both Collaborative and Content based Filtering.**

**Introduction**

In the landscape of online retail, personalized customer experiences are crucial for maintaining customer satisfaction. This project aims to develop a hybrid recommendation system by combining collaborative and content-based filtering techniques.

**Objectives.**

To develop a hybrid recommendation system that enhances the shopping experience by suggesting relevant products to customers.

To evaluate performance of collaborative and content-based filtering methods.

To integrate these methods into a single robust system for improved recommendation accuracy.

**Data Collection.**

The data set for this project was sourced from the "Online Retail" dataset available from the UCI Machine Learning Repository. It includes transactions from a UK-based online retail from December 1, 2010, to December 9, 2011.

The dataset contains several fields, including Invoice Number, Stock Code, Description, Quantity, Invoice Date, Unit Price, and Customer ID.

**Data Preprocessing**

**Cleaning:**

I removed rows with negative quantities and zero prices to ensure that only valid sales transactions are considered.

**Duplication Handling:**

I eliminated duplicate entries based on product descriptions and stock codes to maintain data integrity.

**Exploratory Data Analysis (EDA)**

I conducted preliminary analysis to understand the distribution and characteristics of the data:

**Sales Trends:**

I analyzed purchase quantities and frequency to identify popular products.

**Customer Behavior:**

I examined customer buying patterns to understand common trends and preferences.

Recommendation Techniques

**Collaborative Filtering**

**Pivot Table Creation.**

I created a pivot table to summarize how much each customer bought of each product, which serves as the foundation for collaborative filtering.

**Normalization.**

I normalized the data using standardscaler to ensure that all features contributed equally to the model, preparing it for cosine similarity calculations.

**Content-Based Filtering**

**Train-Test Split.**

I split the data into training and testing sets to validate the performance of the model on unseen data.

**TF-IDF Vectorization:**

I transformed product descriptions into a TF-IDF matrix, which then I used to calculate the similarity between products based on their descriptions.

**Hybrid Recommendation Function**

I developed a hybrid recommendation function to combine the strengths of both collaborative and content-based filtering. This function calculates a weighted average of the similarity scores from both methods to generate product recommendations.

**Conclusion**

The hybrid recommendation system successfully combines collaborative and content-based filtering methods to provide personalized product recommendations.

This system has the potential to significantly enhance the customer experience by offering more relevant products, thereby increasing customer satisfaction and potentially boosting sales.