Recommender Systems

Vatsal Raicha

Projects

[Hybrid Recommender System for Product Recommendation using LightFM]

[Movie Recommender using Collaborative Filtering]

Hybrid Product Recommender System Using LightFM

This project demonstrates the development of a hybrid recommender system designed to suggest relevant products to customers. It leverages the LightFM library to combine the strengths of collaborative filtering (learning from user-item interactions) and content-based filtering (utilizing product features). The model incorporates data from an Excel file containing order, customer, and product information.

Introduction

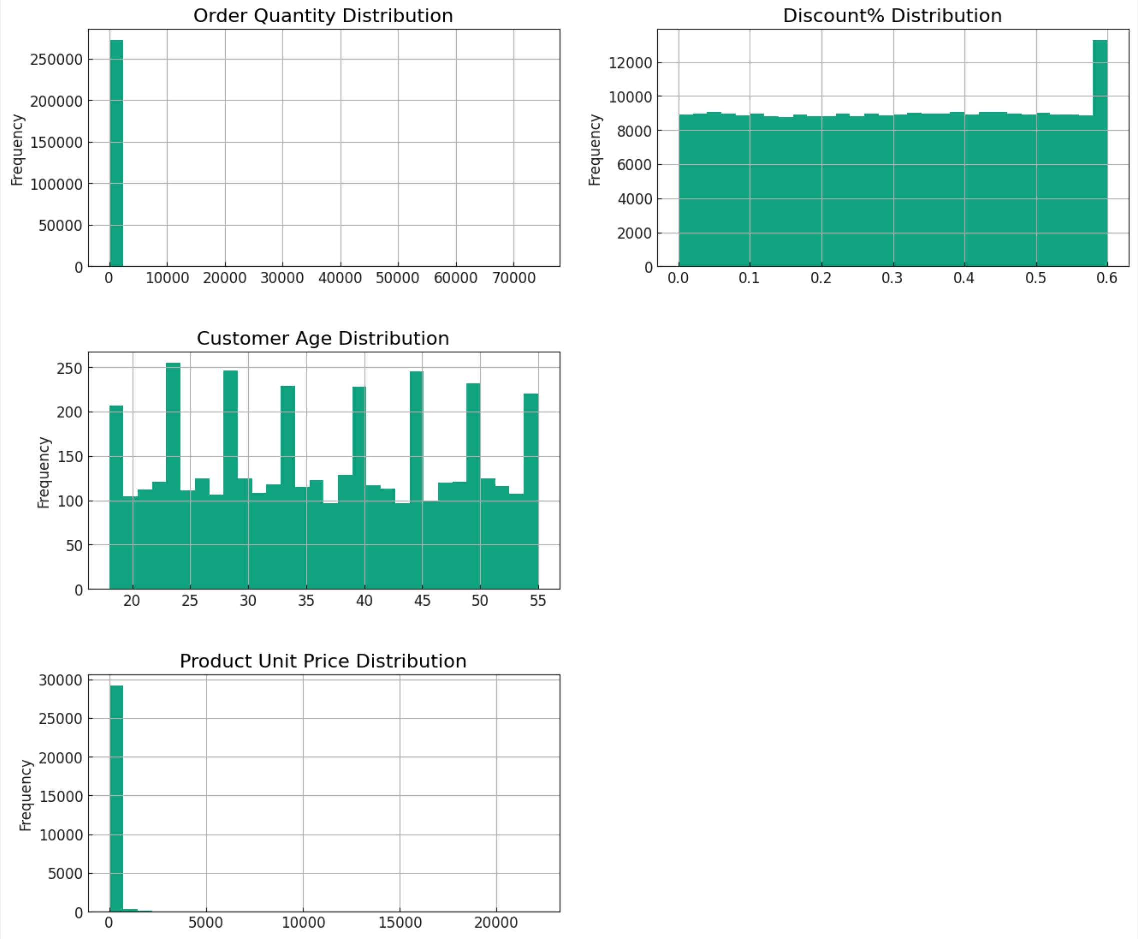
Recommender systems are known as essential tools in e-commerce, helping to personalize the shopping experience and drive sales. This project explores the following types of recommender systems:

- Collaborative Filtering: Relies on past user-item interactions to identify patterns and recommend products that similar users have liked.
- Content-Based Filtering: Recommends items based on their similarity to products a user has previously shown interest in, as defined by product features.
- Hybrid Recommender: Blends the strengths of collaborative and content-based approaches, providing more robust and adaptable recommendations.

Data Exploration and Preparation

- **Data Import**: Data is imported into Pandas DataFrames for analysis and manipulation.
- **Data Merging**: The DataFrames are merged to create a comprehensive dataset encompassing order history, customer demographics, and product information.
- **Data Preprocessing**:
  - Customer age is binned into categories for richer insights.
  - Unique users, items, and features are extracted to prepare interaction matrices.
  - Data is split into training and testing sets for model evaluation.
- **Data Visualization**: Histograms are used to analyze the distributions of order quantity, discount percentages, customer age, and product unit prices. These visualizations reveal valuable insights about customer behavior and product patterns.

Hybrid Product Recommender System Using LightFM



Order Quantity Distribution: The majority of orders contain a relatively small quantity of items, as evidenced by the sharp peak at the lower end of the scale. There are very few large quantity orders, which indicates that bulk purchases are rare.

Discount% Distribution: The distribution of discounts is fairly uniform across the range, with a slight increase in frequency as the discount percentage increases. This could suggest that larger discounts are more commonly offered or that orders with larger discounts are more frequent.

Customer Age Distribution: The age of customers appears to be fairly evenly distributed, with slight increases around the ages of 30 and 50. There are no particularly dominant age groups, which suggests a diverse customer base in terms of age.

Product Unit Price Distribution: There is a large spike in the number of products with a very low unit price, indicating that most products are priced on the lower end. The frequency drops significantly for higher-priced items, showing that expensive products are much less common.

Hybrid Product Recommender System Using LightFM

Model Development

1. Feature Engineering:
   - Customer segments, age groups, and gender are used as additional features to enrich recommendations.
   - Mappings are created to convert user IDs, item names, and features into numerical indices for compatibility with the LightFM library.
2. Interaction Matrices:
   - User-item interactions (e.g., product purchases) are represented as a sparse matrix.
   - Item-feature interactions are constructed for customer segment, age group, and gender, enabling better content-based recommendations.
3. Hybrid Model with LightFM:
   - The LightFM library is used to create a hybrid recommender model.
   - The model is trained using:
     - User-item interaction matrix
     - Item-feature interaction matrices
     - Loss functions (WARP, Logistic, BPR) for optimization

Evaluation:

- The model's performance is measured using the Area Under the ROC Curve (AUC) metric.
- Various loss functions and hyperparameters are experimented with to optimize performance.
- Logistic loss function resulted in the best model with an AUC score of 0.89

Recommendations

- Generating Recommendations:
  - The trained model is used to predict user preferences for unseen items.
  - Scores are normalized for better interpretability.
  - Top recommendations are presented to the user, excluding items they have already purchased.
- Incorporating User Features:
  - Utilizing user features (demographics, behavior) in the LightFM model can further personalize recommendations.

Hybrid Product Recommender System Using LightFM

Results and Discussion

The hybrid recommender system demonstrates the effectiveness of combining collaborative and content-based filtering techniques.

Evaluation metrics (e.g., AUC) provide insights into model performance.

Findings on the importance of different loss functions and hyperparameter optimization are discussed.

Conclusion

This project successfully implemented a hybrid product recommendation system using LightFM. Further improvements and considerations include:

Exploring alternative feature engineering techniques.

Experimenting with additional datasets for broader applicability.

Incorporating real-time updates for more dynamic recommendations.

```
(hybridrecommenderlightfm-py3.9) vatsal@VRALIEN:~/LCode/HybridRecommenderLightFM/src$ python engine.py
Config Path -  /home/vatsal/LCode/HybridRecommenderLightFM/data/config.ini
order.shape=(272404, 9), customer.shape=(4372, 6), product.shape=(29912, 6)
time taken for fitting = 3.33 seconds
User index = 2888
User 17017
     Known positives:
                  Ganma Superheroes Ordinary Life Case For Samsung Galaxy Note 5 Hard Case Cover
                  MightySkins Skin Decal Wrap Compatible with Nintendo Sticker Protective Cover 100's of Color Options
                  Mediven Sheer and Soft 15-20 mmHg Thigh w/ Lace Silicone Top Band CT Wheat II - Ankle 8-8.75 inches
                  MightySkins Skin Decal Wrap Compatible with OtterBox Sticker Protective Cover 100's of Color Options
                  MightySkins Skin Decal Wrap Compatible with DJI Sticker Protective Cover 100's of Color Options
                  MightySkins Skin Decal Wrap Compatible with Lenovo Sticker Protective Cover 100's of Color Options
                  Ebe Reading Glasses Mens Womens Tortoise Bold Rectangular Full Frame Anti Glare grade ckbdp9088
                  Window Tint Film Chevy (back doors) DIY
                  Union 3" Female Ports Stainless Steel Pipe Fitting
                  Ebe Women Reading Glasses Reader Cheaters Anti Reflective Lenses TR90 ry2209
     Recommended:
                  Owlpack Clear Poly Bags with Open End, 1.5 Mil, Perfect for Products, Merchandise, Goody Bags, Party Favors (4x4 inches)
                  MightySkins Skin Decal Wrap Compatible with DJI Sticker Protective Cover 100's of Color Options
                  MightySkins Skin Decal Wrap Compatible with Apple Sticker Protective Cover 100's of Color Options
                  Mediven Sheer and Soft 15-20 mmHg Thigh w/ Lace Silicone Top Band CT Wheat II - Ankle 8-8.75 inches
                  3 1/2"W x 20"D x 20"H Funston Craftsman Smooth Bracket, Douglas Fir
                  MightySkins Skin Decal Wrap Compatible with HP Sticker Protective Cover 100's of Color Options
                  Ebe Women Reading Glasses Reader Cheaters Anti Reflective Lenses TR90 ry2209
                  Handcrafted Ercolano Music Box Featuring "Luncheon of the Boating Party" by Renoir, Pierre Auguste - New YorkNew York
                  MightySkins Skin Decal Wrap Compatible with Smok Sticker Protective Cover 100's of Color Options
                  Eye Buy Express Kids Childrens Reading Glasses Violet Rectangular Full Frame Style Anti Glare grade d5343
```

Movie Recommender using memory-based Collaborative Filtering

This project showcases the development of a collaborative filtering recommendation system using user similarity to suggest movies. It incorporates a Streamlit web application for an interactive user experience. This system is built on Python, utilizing libraries like Pandas, NumPy, and Scikit-learn, with the goal of providing personalized movie recommendations.

Introduction

The aim of this project was to implement a recommendation system that could predict and suggest movies to users based on their historical preferences and similarities with other users. By applying collaborative filtering, the system analyzes patterns in user behavior to recommend movies liked by similar users.

Data Exploration and Preparation

The project used the IMDB movie dataset, requiring an initial phase of data exploration and preparation. This involved loading the dataset, cleaning the data, and understanding its structure to ensure it was suitable for the recommendation engine. Tools like Pandas and NumPy were crucial for this stage, enabling efficient data manipulation and analysis.

Interaction Analysis

Creating an interaction matrix was a key step in the project. This matrix mapped user interactions with movies, serving as the foundation for identifying patterns and similarities among users. Analyzing this matrix helped to understand the relationships between users and movies, guiding the recommendation logic.

Recommendation System Development

The development of the recommendation system was the core of the project. This involved creating a similarity matrix to identify users with similar tastes and recommending movies based on this similarity. Scikit-learn played a vital role here, facilitating the creation of models to calculate user similarities and generate recommendations.

Evaluation

Evaluating the system's effectiveness was crucial to ensure accurate and relevant recommendations. This involved testing the system with various user profiles to assess the accuracy of its predictions and tweaking the model to improve its performance based on feedback and results.

Movie Recommender using memory-based Collaborative Filtering

Streamlit Application

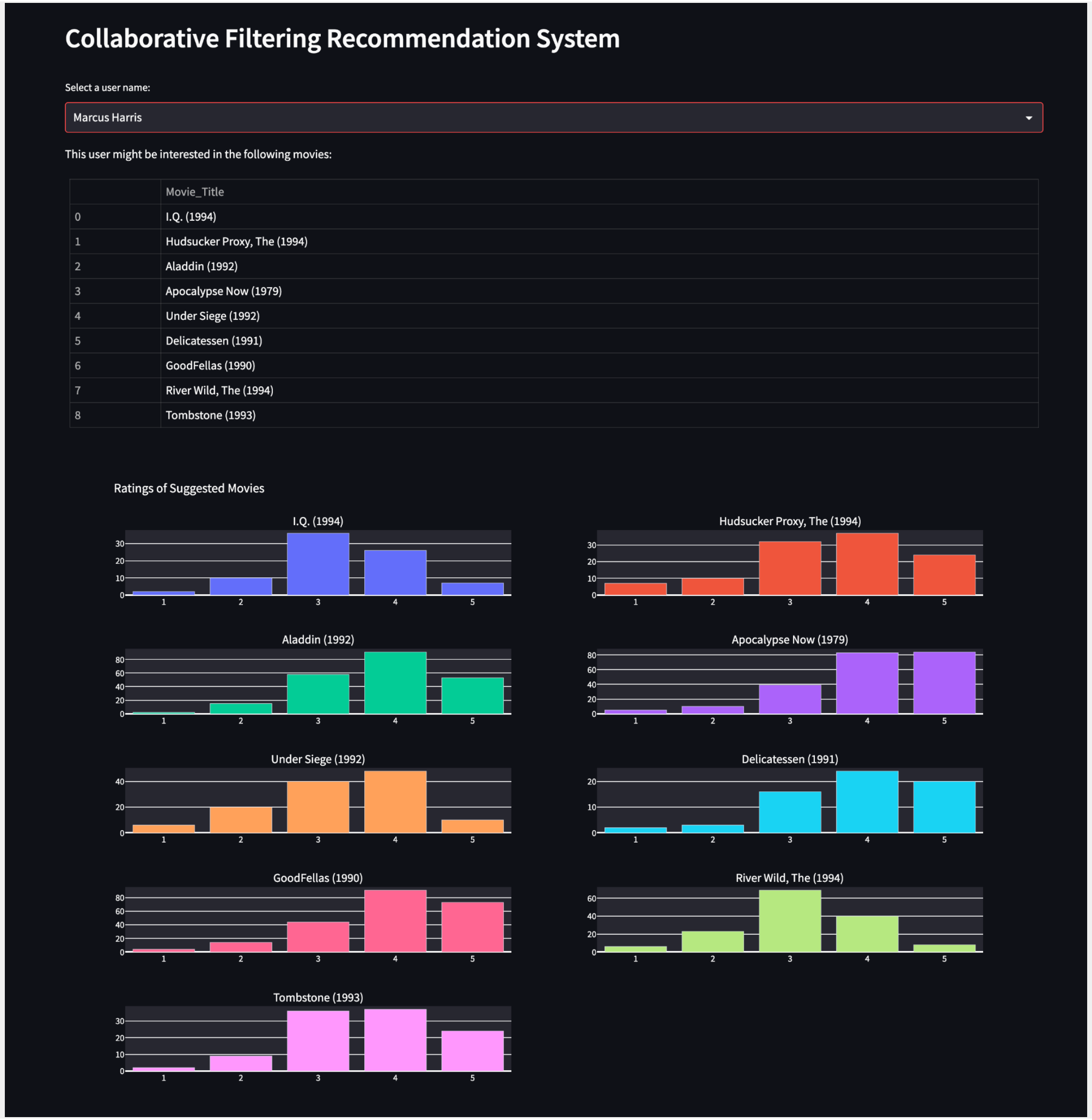
The Streamlit library was used to develop a web application, making the recommendation system accessible and interactive. This application allowed users to view movie recommendations personalized to their tastes, showcasing the system's capabilities in a real-world scenario.

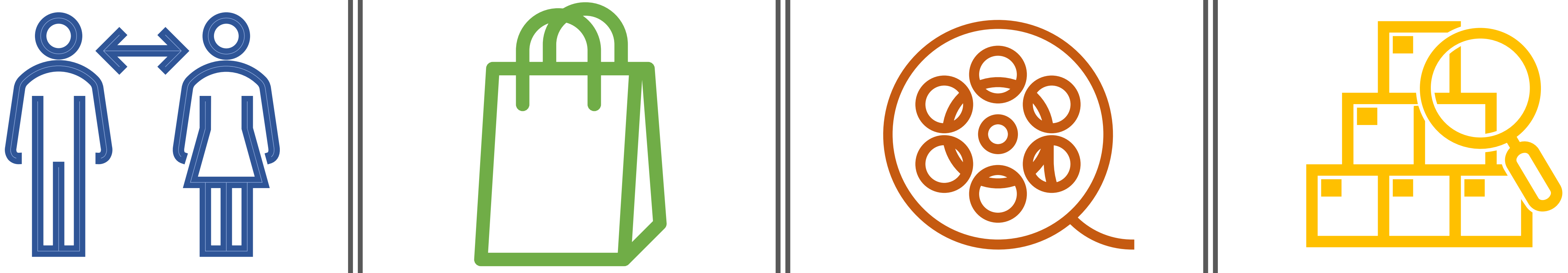
Results

The project successfully developed a recommendation system that could accurately predict user preferences and suggest movies. The Streamlit application provided a user-friendly interface for interaction with the system, demonstrating its practicality and effectiveness.

Conclusion

This project highlighted the application of collaborative filtering in creating personalized experiences. Through careful data preparation, analysis, and model development, the system achieved its goal of recommending movies to users based on their preferences and similarities with others. The Streamlit application further emphasized the project's success, offering an engaging platform for users to discover movie recommendations.





Recommender System Projects

Vatsal Raicha