[2025]

MOVIE RECOMMENDATION SYSTEM

BUSINESS PROBLEM

In today's digital world, users are overwhelmed with vast amounts of content, whether it's movies, products, music, or news articles.

Businesses struggle to keep users engaged by providing personalized recommendations. Without an effective recommender system, customers may churn, engagement may decline, and businesses may lose revenue opportunities.

For example, in an online movie streaming platform, users need relevant and personalized movie recommendations based on their viewing history and preferences.

A poor recommendation system may result in users struggling to find interesting content, leading to lower customer satisfaction and reduced subscription retention.

GOALS AND OBJECTIVES

Objectives

By implementing an effective recommender system, businesses can:

1.Increase user engagement and retention.

2. Improve customer satisfaction by offering relevant recommendations.

3.Enhance revenue opportunities through personalized marketing.

Goals

The goal of this study is to develop a personalized recommendation system that improves user experience and engagement by suggesting relevant content based on past interactions. This will be achieved using:

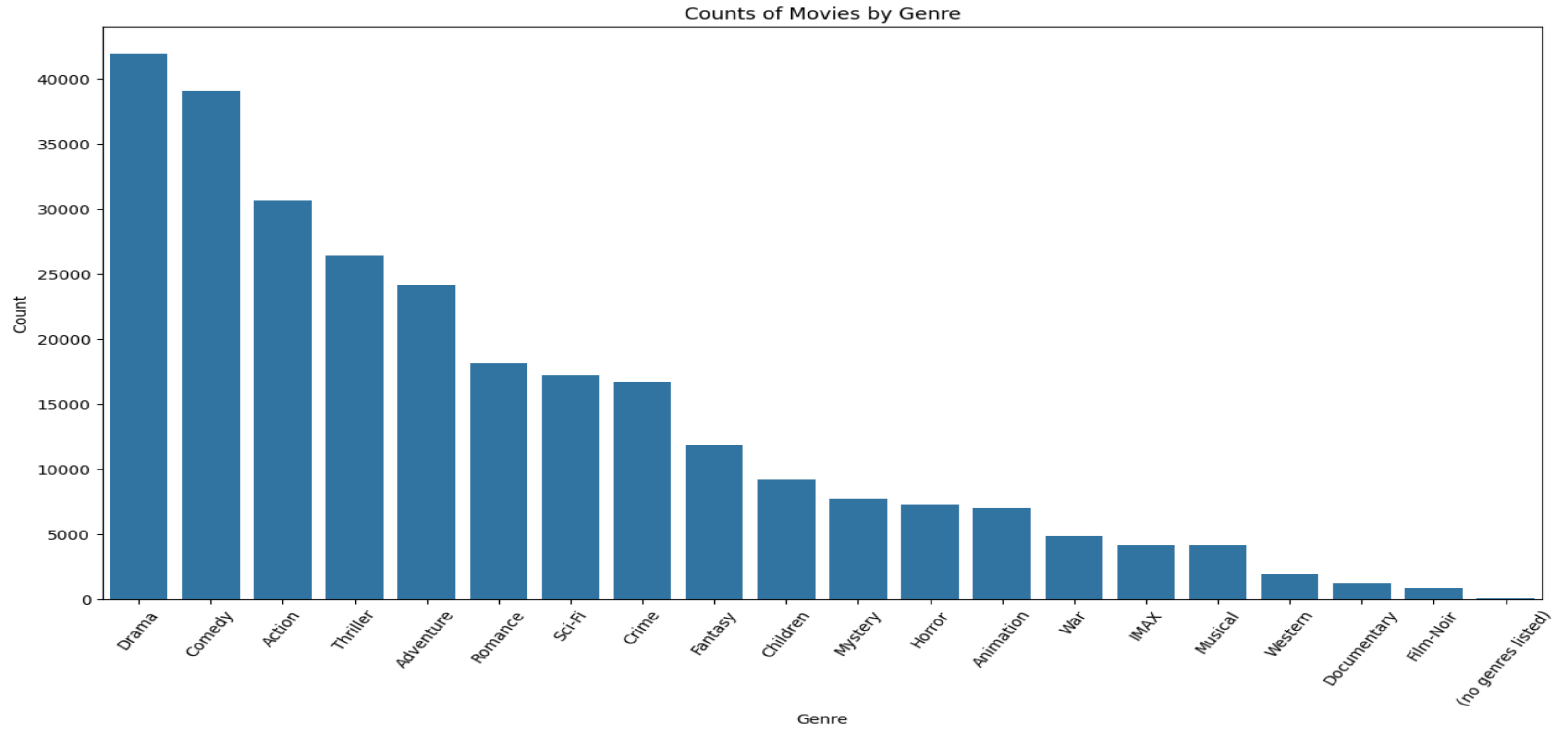
1.Collaborative Filtering: Predict user preferences based on similar users.

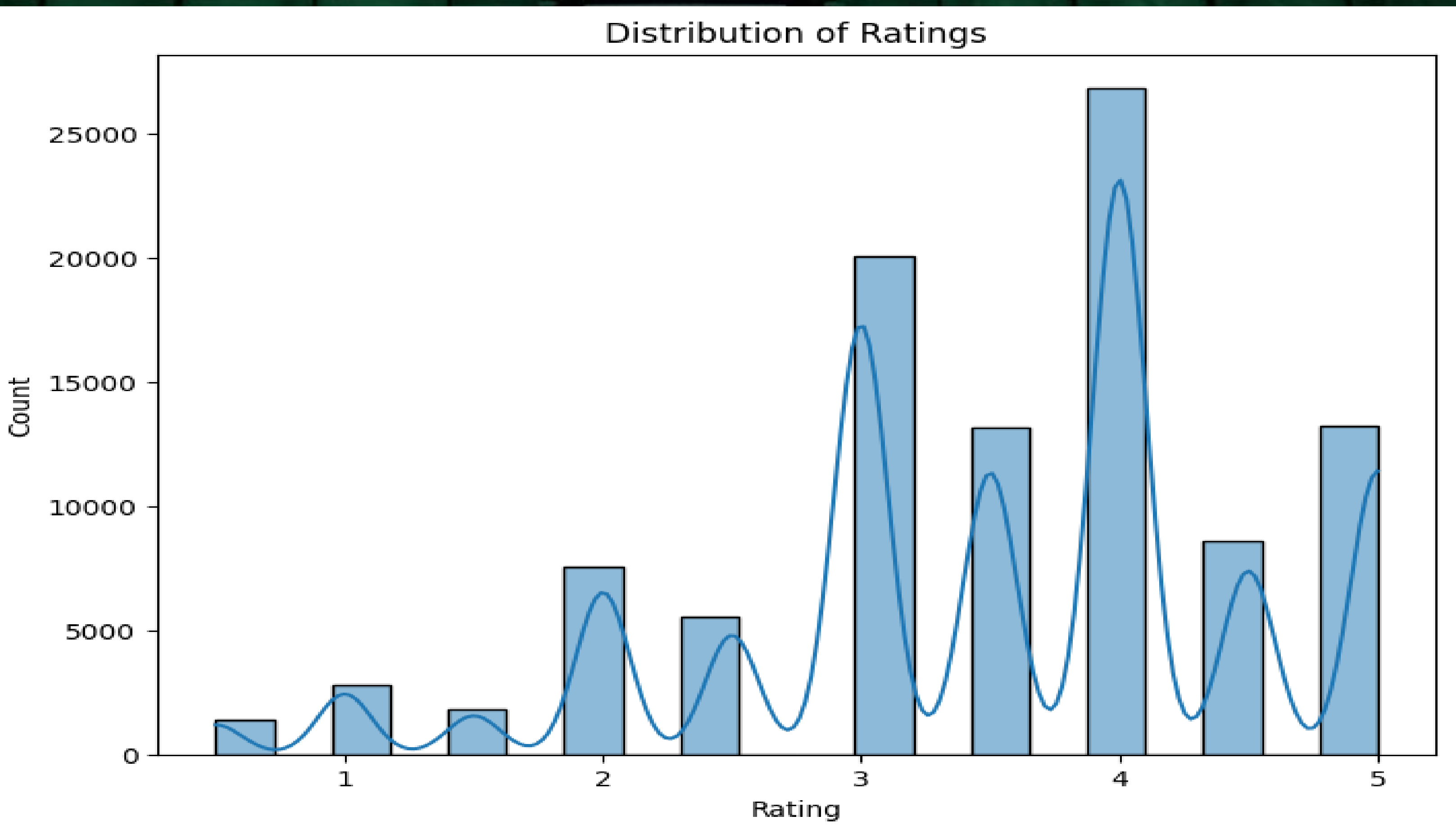
2.Content-Based Filtering: Recommend items similar to what a user has liked before.

3.Create a model and carry out a model evaluation

4.Performance Evaluation: Assess the effectiveness of different models using evaluation metrics such as RMSE (Root Mean Squared Error) and Cosine Similarities

VISUALIZATIONS





EXPLORATORY DATA ANALYSIS (EDA)

- The dataset shows a skewed rating distribution, with most movies receiving higher ratings, particularly around 4 stars.
- Drama is the most watched genre, followed by Comedy and Action, while Thriller, Adventure, and Romance have moderate popularity.
- The most-rated movies, such as *Hollywood Chainsaw Hookers* and *Calcium Kid*, significantly impact collaborative filtering recommendations. Highly rated genres include Film-Noir, War, Documentary, Crime, and Drama, meaning users who favor these genres are likely to receive similar recommendations.
- Conversely, Horror, Comedy, Children, Action, and Sci-Fi tend to have lower average ratings.

BUILDING A RECOMMENDERSYSTEM

Collaborative Filtering

The recommendation model uses both item-based and user-based collaborative filtering to suggest movies based on similarities. It incorporates SVD for collaborative filtering, TF-IDF for genre-based recommendations, and RMSE for evaluation. In item-based filtering, movies similar to those a user has liked (e.g., *Hollywood Chainsaw Hookers*) are recommended.

In user-based filtering, recommendations are based on users with similar tastes—if a user liked *Hollywood Chainsaw Hookers*, they might be suggested *Calcium Kid* or *Chinese Puzzle* based on other users' preferences. This approach enhances personalized recommendations by leveraging both item and user similarities.

Content Based Filtering

The Content-Based Movie Recommender System suggests movies based on their features, such as genres and descriptions, using TF-IDF vectorization to convert text into numerical data and cosine similarity to measure movie similarity.

It focuses on recommending movies with similar characteristics to a given input movie, allowing users to receive customized recommendations based on content rather than user ratings.

MODEL EVALUATION AND PREDICTION

Model Evaluation

•The model achieved an RMSE of 0.8748, meaning the predicted ratings deviate from the actual ratings by approximately 0.87 on average.

•Since RMSE is a measure of error, a lower value indicates better prediction accuracy.

•However, an RMSE close to 1 suggests that while the model performs reasonably well, there is still room for improvement through hyperparameter tuning, incorporating additional features, or using a more advanced recommendation technique.

Model Prediction

•The performance is evaluated using Root Mean Squared Error (RMSE), which measures the difference between predicted and actual ratings.

•The resulting RMSE is 0.9734, indicating that, on average, the model's predictions deviate from actual ratings by approximately 0.97 rating points.

CONCLUSION

The analysis revealed several important findings:

- Rating distribution shows a positive skew, with most movies receiving 4-5 star ratings
- Both collaborative and content-based filtering methods demonstrated effectiveness in generating recommendations
- The combination of both approaches provides a more robust recommendation system
- The models show promise in capturing user preferences and suggesting relevant content

RECOMMENDATION

1.Hybrid System Implementation

•Combine collaborative and content-based filtering

•Leverage the strengths of both methods for more accurate recommendations Model Optimization

2.Implement continuous monitoring of model performance

•Regular updates to adapt to changing user preferences

•Consider implementing A/B testing for different recommendation strategies

3.Data Enhancement

•Regular updates to adapt to changing user preferences

•Expand the dataset with additional features:

- User demographics
- Movie reviews
- Social media interactions

•This will improve recommendation accuracy and personalization

4.User Engagement Strategy

•Use personalized recommendations to increase platform engagement

•Implement features to encourage content exploration

•Track and analyze user interaction with recommendations

5.Technical Improvements

•Regular system performance monitoring

•Optimization of recommendation algorithms

•Implementation of real-time recommendation updates