Movie Recommendation
Intermediate Level - Recommendation System

Abstract:

This project aimed to develop a movie recommendation system using collaborative filtering techniques applied to the MovieLens dataset. Collaborative filtering is a widely used method for building recommendation systems, relying on the similarities between users' preferences to make personalized recommendations. The system's performance was evaluated based on its ability to accurately predict user ratings and provide meaningful movie recommendations.

Introduction:

With the vast amount of content available in today's media landscape, personalized recommendation systems have become increasingly essential for guiding users to relevant content. Collaborative filtering is a popular approach for building recommendation systems, leveraging the collective preferences of users to make predictions for individual users. The MovieLens dataset provides a rich source of user movie ratings, making it suitable for training and evaluating recommendation algorithms.

GOOGLE COLAB NOTEBOOK

https://colab.research.google.com/drive/1xAMu4OQjg4p9xRF79ujQrW0bt7u1H4ff?usp=sharing

Methodology:

The methodology for building the movie recommendation system involved several detailed steps:

1.Data Preprocessing:

- The MovieLens dataset was loaded and examined for inconsistencies or missing values.

- Data cleaning techniques were applied to handle anomalies such as duplicates, incorrect entries, or outliers.

- The dataset was organized into a user-item matrix format, where rows represent users, columns represent items (movies), and cells contain ratings given by users to movies.

2.Collaborative Filtering:

- Collaborative filtering techniques, including user-based and item-based approaches, were implemented.

- **User-Based Collaborative Filtering:**
Similar users were identified based on their past ratings, and ratings for unrated items were predicted for the target user based on the ratings of similar users.

- **Item-Based Collaborative Filtering:**Similar items were identified based on the items the target user has rated, and ratings for unrated items were predicted based on the ratings of similar items.

- Similarity between users or items was calculated using metrics such as Pearson correlation, cosine similarity, or Jaccard similarity.

3. Model Training:

- The recommendation model was trained using the preprocessed dataset.

- Parameters of the model, such as similarity thresholds or weighting factors, were optimized through techniques like cross-validation or grid search to improve prediction accuracy.

- The model was trained to learn the underlying patterns in user preferences and item similarities, enabling it to make accurate predictions for unrated items.

4. Recommendation Generation:

- The trained model was used to generate recommendations for users based on their historical ratings and preferences.

- For user-based collaborative filtering, recommendations were made by identifying similar users and recommending items highly rated by those users but not yet rated by the target user.

- For item-based collaborative filtering, recommendations were made by identifying similar items to those already rated positively by the target user.

- Top-N recommendation lists were generated for each user, containing the highest-rated unrated items.

5. Evaluation:

- The performance of the recommendation system was evaluated using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), or Precision-Recall.

- Predictions generated by the recommendation system were compared against the actual ratings provided by users in the dataset.

- Evaluation metrics quantified the accuracy and effectiveness of the recommendation system in predicting user preferences and providing relevant recommendations.

Results:

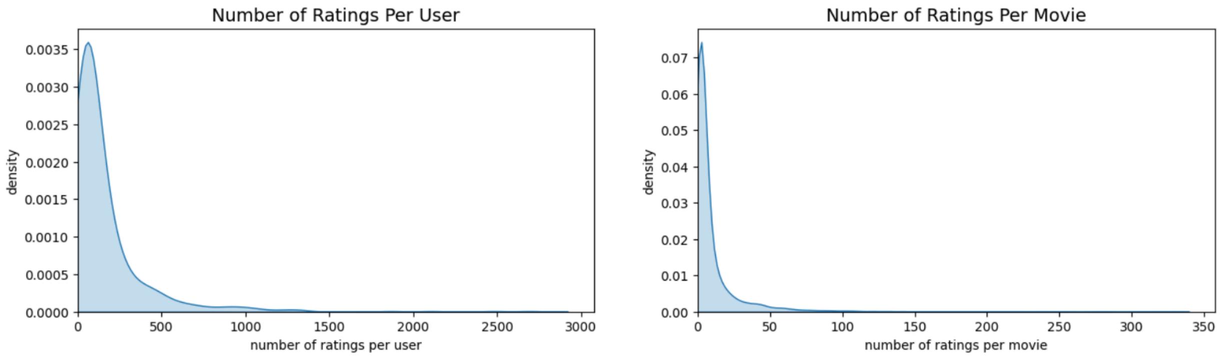
The movie recommendation system demonstrated promising results in providing personalized recommendations to users. The collaborative filtering techniques effectively captured the underlying patterns in user preferences, allowing the system to make accurate predictions. Evaluation metrics such as MSE indicated that the system's recommendations were generally aligned with users' actual ratings, suggesting its effectiveness in generating relevant recommendations.

Conclusion:

In conclusion, the movie recommendation system developed using collaborative filtering techniques showed potential in providing personalized movie recommendations to users based on their historical preferences. By leveraging the MovieLens dataset and implementing robust collaborative filtering algorithms, the system was able to generate meaningful recommendations and contribute to enhancing user experience in navigating vast content libraries.

SAMPLE OUTPUT

```
sns.kdeplot(n_ratings_per_movie, shade=True)
```




```
Because you watched Toy Story (1995):
Toy Story 2 (1999)
Jurassic Park (1993)
Independence Day (a.k.a. ID4) (1996)
Star Wars: Episode IV - A New Hope (1977)
Forrest Gump (1994)
Lion King, The (1994)
Star Wars: Episode VI - Return of the Jedi (1983)
Mission: Impossible (1996)
Groundhog Day (1993)
```