Deepak Shinde (Vyorius Test)

Performance Report: Personalized Movie Recommendation System

Overview

This report evaluates a **personalized movie recommendation system** that uses **collaborative filtering** to predict ratings for movies based on a user's past preferences and the similarities to other users. The system then generates a ranked list of recommended movies. The performance of the system is assessed using key metrics, including **Root Mean Squared Error (RMSE)**, **Mean Absolute Error (MAE)**, **Precision@N**, **Recall@N**, and **Coverage** to gauge both its predictive accuracy and the quality of the recommendations provided.

1. Root Mean Squared Error (RMSE)

RMSE measures the average difference between predicted ratings and actual ratings, offering an indication of prediction accuracy.

- **Performance**: The system achieved an **RMSE of 1.2748**, meaning that, on average, the predicted ratings deviate from actual ratings by 1.27 points. This indicates a moderate error in prediction, suggesting that the system is not perfectly accurate in predicting user ratings.

Interpretation: While this RMSE value indicates that the system is somewhat accurate, there is significant room for improvement. A lower RMSE would reflect more precise predictions and improve user experience.

2. Mean Absolute Error (MAE)

MAE measures the average absolute difference between predicted and actual ratings, providing a simpler understanding of the prediction errors. /

- **Performance**: The system achieved an **MAE of 0.6250**, meaning that, on average, the predicted ratings deviate from the actual ratings by 0.625 points.

Interpretation: The MAE value suggests that the system's predictions are fairly close to the actual ratings, but there is still room to reduce errors. Lowering the MAE would improve the accuracy and reliability of the predictions.

3. Precision@N

Precision@N measures how many of the top-N recommended movies are relevant to the user. It is an important metric to evaluate the recommendation quality.

- **Performance**: The system achieved a **Precision@3 of 0.3333**, meaning that 33.33% of the top-3 recommended movies were relevant to the user.

Interpretation: This suggests that only about one-third of the recommended movies are relevant to the user's preferences. The system could benefit from better filtering to ensure that more of the top-N recommendations are aligned with user interests.

4. Recall@N

Recall@N measures how many of the relevant movies for the user appear in the top-N recommendations, indicating how well the system captures the user's interests.

- **Performance**: The system achieved a **Recall@3 of 0.3333**, indicating that 33.33% of the relevant movies for the user were included in the top-3 recommendations.

Interpretation: This implies that the system misses a significant portion of relevant movies. While it recommends some relevant movies, it does not capture all the relevant items, meaning improvements are needed to increase recall and ensure that more relevant items are recommended.

Summary of System Performance

- **Prediction Accuracy**: The system achieved an **RMSE of 1.2748** and an **MAE of 0.6250**, indicating that the system's predictions are fairly accurate, but there is

room for improvement in reducing prediction errors. The model could benefit from better similarity calculations or more sophisticated predictive algorithms.

- **Recommendation Quality**: The **Precision@3 of 0.3333** and **Recall@3 of 0.3333** suggest that only one-third of the top-3 recommendations are relevant, and the system misses a significant portion of relevant items. This indicates the need to improve both precision and recall by better capturing user preferences and recommending more relevant items.

Conclusion

The movie recommendation system demonstrates reasonable performance but has several areas for improvement:

- **Prediction accuracy** (RMSE of 1.2748 and MAE of 0.6250) can be enhanced by refining the similarity measures or implementing more advanced algorithms.
- **Recommendation quality** (Precision@3 and Recall@3 both at 0.3333) suggests that the system needs better mechanisms for selecting top-N recommendations and capturing all relevant items.
- **Diversity and coverage** (Coverage of 0.60) indicate that the system provides a reasonably varied set of recommendations, but greater diversity could improve user experience.