Evaluation of Recommendation System Models

Summary of Evaluation Metrics Used

The following metrics are used to evaluate the performance of the recommendation system models:

- **Root Mean Squared Error (RMSE):** Measures the average difference between predicted and actual ratings. Lower RMSE indicates better accuracy. It's calculated as:

$$RMSE = N1\sum i = 1N(rui - r\hat{}ui)2$$

where (r_{ui}) is the actual rating, (\hat{r}_{ui}) is the predicted rating, and N is the number of ratings.

- **Precision@K:** The proportion of recommended items (top K) that are relevant to the user. A higher precision score indicates that the system is good at recommending relevant items.

$$Precision@K = | Relevant\ Items \cap Recommended\ Items | \div | Recommended\ Items |$$

Recall@K: The proportion of relevant items that are found in the top K recommendations. A higher recall score indicates that the system is good at finding most of the relevant items.

$$Recall@K = | Relevant\ Items \cap Recommended\ Items | \div | Relevant\ Items |$$

- **Diversity:** Measures how dissimilar the recommended items are from each other. High diversity means the recommendations cover a broader range of items, which can be important for user satisfaction. Diversity is calculated as the average dissimilarity between all pairs of recommended items.

Table of Model Results

Model	RMSE	Precision@10	Recall@10
Collaborative Filtering	0.8732	0.1563	0.0985
Content-Based Filtering	N/A	0.1333	0.0889

Comparison of Models Based on Performance

- **Collaborative Filtering:**
    - The Collaborative Filtering model achieved an RMSE of 0.8732, indicating a reasonable level of accuracy in predicting user ratings.
    - Precision@10 was 0.1563, meaning that about 15.63% of the top 10 recommendations were relevant to the users.
    - Recall@10 was 0.0985, indicating that the model captured about 9.85% of the total relevant items in the top 10 recommendations.
- **Content-Based Filtering:**
    - RMSE is not applicable for Content-Based Filtering because it does not predict ratings.
    - Precision@10 was 0.1333, meaning that about 13.33% of the top 10 recommendations were relevant.

    - Recall@10 was 0.0889, indicating that the model captured about 8.89% of the total relevant items in the top 10 recommendations.

Discussion on Performance

- The **Collaborative Filtering** model shows slightly better performance in terms of both precision and recall compared to the Content-Based Filtering model. This suggests that Collaborative Filtering is more effective at recommending relevant books to users based on past user-item interactions. The lower RMSE also confirms that it predicts ratings more accurately.
- **Content-Based Filtering** relies solely on the features of the books (title, author, publisher). Its performance is decent, but it may suffer from not using any user interaction data.