Hybrid Approach for Movie Recommendation System

Ishita Gabhane, Amanda Tsui, Shatakshi Ranjan
Rutgers University
img50,alt165,sr1647

ABSTRACT

This research investigates the implementation and evaluation of a hybrid recommendation system that combines collaborative filtering with user-specific averages. Using the MovieLens dataset, the system predicts user ratings and generates top-10 recommendation lists. Performance metrics such as MAE, RMSE, Precision, Recall, F-measure, and NDCG are used to evaluate the effectiveness of the recommendations. Results indicate improved accuracy in rating prediction and demonstrate the potential of hybrid methods in addressing the limitations of traditional collaborative filtering.

KEYWORDS

Recommendation System, Collaborative Filtering, Hybrid Approach, Precision, Recall, NDCG, MovieLens Dataset, MAE, RMSE

1 INTRODUCTION

Recommender systems have been widely studied in both academia and industry. Collaborative filtering, particularly item-based and user-based approaches, is among the most widely used techniques for recommendation systems. Matrix factorization methods, such as Singular Value Decomposition (SVD), have been employed for better accuracy in rating prediction [1]. However, challenges like data sparsity and cold-start problems remain significant obstacles. This project builds on these methods by focusing on item similarity, transparency, and fairness, and evaluating their effectiveness in movie recommendation.

2 RELATED WORK

Recommender systems have been widely studied in both academia and industry. Collaborative filtering, particularly item-based and user-based approaches, is among the most widely used techniques for recommendation systems. Matrix factorization methods, such as Singular Value Decomposition (SVD), have been employed for better accuracy in rating prediction [1]. However, challenges like data sparsity and cold-start problems remain significant obstacles. This project builds on these methods by focusing on item similarity, transparency, and fairness, and evaluating their effectiveness in movie recommendation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference'17, Washington, DC, USA
© 2016 ACM. 978-x-xxxx-xxxx-x/YY/MM…$15.00
DOI: 10.1145/nnnnnnn.nnnnnnn

3 PROBLEM FORMALIZATION

Given a set of users U, a set of items (movies) I, and a set of known ratings $R_{u,i}$ for user u and item i, the goal is to predict the unknown ratings $\hat{R}_{u,i}$ for each user-item pair and generate a top-N recommendation list for each user. The performance is evaluated using the following metrics:

- **MAE:**
$$\text{MAE} = \frac{1}{|T|} \sum_{(u,i) \in T} |R_{u,i} - \hat{R}_{u,i}|$$
- **RMSE:**
$$\text{RMSE} = \sqrt{\frac{1}{|T|} \sum_{(u,i) \in T} (R_{u,i} - \hat{R}_{u,i})^2}$$
- **Precision and Recall:** Evaluated on the top-10 recommendations generated for each user.

4 THE PROPOSED MODEL

The system uses collaborative filtering with item-item cosine similarity and a hybrid approach combining collaborative filtering and user-specific averages. Transparency and controllability features were added to enhance user trust and satisfaction. Fairness was addressed by analyzing genre diversity in recommendations. Robustness and privacy protection were implemented through anomaly detection and data anonymization.

5 EXPERIMENTS

The system was implemented using Python, leveraging the MovieLens dataset. The training dataset was used to compute item similarity and predict ratings for the testing dataset. The following results were obtained:

- **MAE:** 0.7396
- **RMSE:** 0.9522
- **Precision:** 0.1448
- **Recall:** 0.0670
- **F-measure:** 0.0916
- **NDCG:** 0.1475

The results demonstrate an improvement in ranking quality (NDCG) and relevance (Precision) compared to a baseline collaborative filtering approach. While the hybrid approach improves overall prediction accuracy, the system still struggles to retrieve a significant portion of relevant items (low Recall).

6 OPTIONAL TASKS

6.1 Transparency and Explainability

Problem Statement: Users often lack trust in recommendation systems because they cannot understand why certain items (e.g., movies) are recommended. This lack of transparency reduces user engagement and satisfaction, especially when recommendations feel irrelevant or arbitrary.

Solution: The system provides explanations for recommendations by:

- Identifying the genres of the recommended movie.
- Analyzing the user's previous ratings, focusing on movies they rated highly (e.g., 4 stars or more).
- Calculating similarity scores between the recommended movie and the user's highly-rated movies using the item similarity matrix.
- Generating a natural-language explanation that highlights shared genres and mentions similar movies the user liked.

Evaluation Methods: Transparency and explainability are evaluated qualitatively by:

- Verifying that the explanations align with user preferences and highlight relevant connections (e.g., shared genres, similar movies).
- Example output:

  "Movie 333 is recommended because it shares the genres 'Comedy' with movies you rated highly (e.g., Maverick (1994), Fugitive, The (1993), Austin Powers: International Man of Mystery (1997))."
- Gathering user feedback (if available) to measure trust and satisfaction based on the clarity of explanations.

6.2 Fairness and Unbiases

Problem Statement: Recommendation systems often suffer from biases, such as favoring popular genres (e.g., "Action" or "Comedy"), which can result in a lack of diversity. This bias reduces fairness and limits exposure to less popular genres or niche content, impacting user satisfaction and inclusivity.

Solution: The system incorporates fairness by analyzing the genre diversity of recommendations:

- For each user, the system counts the number of recommended movies belonging to each genre.
- Results are stored in a file (`genre_diversity.txt`), providing a breakdown of genre representation in the recommendations.
- This ensures that users receive a balanced mix of content and that no single genre dominates.

Evaluation Methods: Fairness is evaluated quantitatively by:

- Calculating the genre diversity for each user using metrics such as the number of unique genres represented.
- Example output for User 1:

  Comedy: 6, Action: 4, Adventure: 3, Sci-Fi: 2, Crime: 2.
- Comparing the genre distribution across all users to ensure consistency and avoid favoritism toward specific genres or user groups.
- Optional metrics like Shannon Diversity Index or Gini Coefficient can be used for advanced fairness evaluation.

6.3 Controllability

Problem Statement: Users have limited control over the recommendations they receive, which reduces personalization and adaptability. For example, users might prefer content from a specific genre or want to exclude certain types of recommendations.

Solution: The system empowers users to filter recommendations based on genre preferences:

- Users can input their user ID and a genre they are interested in (e.g., "Comedy").
- The system filters the top-N recommendations for the user to include only movies from the selected genre.
- This tailored approach ensures that users can adjust recommendations to match their specific needs.

Evaluation Methods: Controllability is evaluated qualitatively and quantitatively by:

- Verifying the accuracy of filtered recommendations against user-specified genres.
- Example output:

  Filtered Recommendations for User 4 by Genre 'Comedy': Fish Called Wanda, A (1988), Blues Brothers, The (1980), Raising Arizona (1987).
- Measuring user satisfaction through feedback (if available), focusing on whether the filtered recommendations meet expectations.
- Tracking the success rate of genre filtering (e.g., the percentage of filtered recommendations that belong to the specified genre).

6.4 Privacy Protection

Problem Statement: The use of real user IDs and identifiable movie titles in datasets can compromise user privacy, making recommendation systems vulnerable to breaches of sensitive information.

Solution: The system implements data anonymization:

- User IDs are replaced with pseudonyms (e.g., User_1, User_2).
- Movie titles are replaced with their respective movie IDs (e.g., Movie_1, Movie_2).
- The anonymized data is stored in a file (`anonymized_data.txt`), enabling secure data sharing and compliance with privacy regulations like GDPR.

Evaluation Methods: Privacy protection is evaluated by:

- Ensuring that no personally identifiable information (PII) is present in the anonymized data.
- Example output from the anonymized data file:

  User ID: User_1 **Movie Title:** Movie_1 **Rating:** 5
- Confirming that the recommendation system performs equally well with anonymized data compared to original data, maintaining usability while safeguarding privacy.

7 CONCLUSIONS AND FUTURE WORK

This project demonstrates the feasibility of a hybrid approach for building a recommendation system. While the system achieves good accuracy in predicting ratings (low MAE and RMSE), the recommendation relevance (Precision, Recall, and NDCG) remains low. Future work will focus on:

(1) Incorporating content-based features like genres and tags.
(2) Using advanced algorithms like Matrix Factorization or Neural Collaborative Filtering.
(3) Addressing data sparsity and cold-start challenges.

ACKNOWLEDGEMENT

We would like to thank Professor Zhang and our TAs Wujiang Xu and Divanshi Patel for their guidance and support throughout this project. This research was inspired by the work of Professor Zhang on fairness-aware and explainable recommendation systems.

REFERENCES

[1] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 8 (2009), 30–37.
[2] Yongfeng Zhang. Explainable Recommendation: A Survey and New Perspectives. 2019.
[3] Yongfeng Zhang. Fairness in Recommendation Systems: Research Directions and Open Questions. 2020.