Research Institute for DIGITAL INNOVATION IN LEARNING

Faculty Research in Partnership with RIDIL

Research and Implementation of Movie Recommendation System Based on the MovieLens Dataset

Research Author Bios:

Hongfang Zhang, Ph.D. Student, Electrical and Computer Engineering, Old Dominion University

Yi He, Assistant Professor, Data Science, College of William & Mary

Jian Wu, Assistant Professor, Computer Science, Old Dominion University

White Paper Author:

Hongfang Zhang, Yi He, Jian Wu, & Helen Crompton

Overview

Recommender systems are a pivotal research area in modern information retrieval and artificial intelligence, aiming to assist users in efficiently identifying contents that align with their preferences. With the exponential growth of users and online information, machine learning (ML)-driven recommender systems have emerged as a critical solution to this challenge. This study utilizes the MovieLens 1M [1], which is a standard benchmark dataset commonly used for training, evaluating, and comparing the performance of recommender systems. However, existing studies exhibit significant variation in utilizing this dataset, primarily due to differences in preprocessing methods, splitting strategies, and evaluation metrics. Such inconsistencies hinder the direct comparability of results across studies. This study aims to evaluate and compare the state-of-the-art recommender systems based on the MovieLens dataset, with the goal of understanding their core models and performances. In addition to the rating matrix, we design and implement methods incorporating user demographic features (Gender, Age, Occupation, zip code) and movie features (Title, Genres) as inputs. Adding user and movie profiles allows us to develop models that recommend movies to *new* users based on the similarity scores between user and movie features, thereby overcoming the known and lasting “cold-start” challenge in the domain. This study demonstrates the potential and wide applicability of RS models in real-world recommendation settings.

Purpose of the Research

This study aims to systematically compare the performance of state-of-the-art algorithms in movie recommender systems based on standard evaluation metrics. In particular, we compare preprocessing and machine learning models and found that preprocessing methods play a key role toward in training an effective recommendation model. We also design and implement a method incorporating user and movie profiles to address the cold-start problem, in which the system recommends movies to *new* users with no or very limited historical data. The key research questions explored in this report include:

1. How do state-of-the-art recommender systems perform under a uniform setting?
2. How does preprocessing influence the performance of different recommendation models?

Method

This study evaluates three movie recommendation models using the MovieLens 1M dataset, which contains over one million user ratings for nearly 3,900 movies from 6,040 users. The first model applies neural networks within a bandit framework to balance exploration, which introduces users to new recommendations, and exploitation, which refines predictions based on known preferences. The second model leverages word embeddings to encode user and movie profiles, allowing it to predict user preferences based on feature similarities. The third model employs a deep learning approach, integrating user and movie embeddings within a multi-layer perceptron to capture complex interactions between users and movies.

To assess model performance, the study uses three key metrics: ROC-AUC to measure ranking accuracy, overall accuracy to evaluate the correctness of recommendations, and regret rate to quantify incorrect predictions. The evaluation focuses on determining the most effective model for personalized recommendations while addressing challenges such as the "cold-start" problem, where new users have little or no prior data.

Results and Findings

The evaluation showed that the deep learning model performed best, making the most accurate recommendations while minimizing incorrect suggestions. This indicates that, on average, for every 10 recommendations, approximately 8–9 recommended movies are likely to be adopted, with only 1–2 misaligned with the user's preferences. The third model outperforms the other two because of its ability to capture complex interactions between user and movie features. In addition, we found that combining user demographic features, such as age, and occupation, and semantic features, such as movie titles and genres, yielded the best results across all three models.

Model/Metric	**ROC-AUC (higher is better)**	**Accuracy (higher is better)**	**Regret rate (lower is better)**
Model 1	0.7128	0.6935	0.3800
Model 2	0.7932	0.7634	0.3100
Model 3	**0.8613**	**0.8428**	**0.1572**

Conclusions

The study evaluated and compared the performance of three recommender system models based on the MovieLens 1M dataset. The results demonstrated that the deep neural network model, which integrates user and movie embeddings within a multi-layer perceptron, achieved the best performance. Overall, this study provides new insights into improving recommender system performance by effectively leveraging user and movie features. It offers potential applications in recommendations in an educational context.

Recommendation systems play a crucial role in education by tailoring content to individual students, optimizing instructional resources, and enhancing engagement. By analyzing learning behaviors, these systems can suggest materials that align with a student's pace, preferences, and needs, creating a more personalized learning experience. This adaptability ensures that learners receive the right content at the right time, supporting both reinforcement and challenge as needed.

For educators, recommendation systems provide valuable support in selecting and allocating instructional resources. By analyzing student performance and curriculum requirements, these systems can suggest relevant textbooks, videos, exercises, and supplementary materials, streamlining lesson planning and ensuring that teaching is both efficient and impactful. This helps instructors focus on facilitation rather than content curation, allowing for more meaningful interactions with students.

References

[1]. Harper, F. M., & Konstan, J. A. (2015). The movielens datasets: History and context. Acm transactions on interactive intelligent systems (tiis), 5(4), 1-19.

[2]. Ban, Y., Yan, Y., Banerjee, A., & He, J. (2021). EE-Net: Exploitation-exploration neural networks in contextual bandits. arXiv preprint arXiv:2110.03177.

[3]. Auer, P. (2002). Finite-time Analysis of the Multiarmed Bandit Problem.

[4]. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017, April). Neural collaborative filtering. In Proceedings of the 26th international conference on world wide web (pp. 173-182).