

Hybrid Movie Recommendation System using Collaborative and Content-Based Filtering

Thanapoom Phatthanaphan,¹ Chandini Rayadurgam,² Shreya Daniel Jacob³

Stevens Institute of Technology
tphattha@stevens.edu,¹ crayadur@stevens.edu,² sjacob4@stevens.edu³

Abstract

The research focuses on movie recommendation systems as its application domain, aiming to create a hybrid recommendation system that blends collaborative filtering and content-based filtering techniques. By doing so, the system can provide personalized movie recommendations to users, which is a common practice among movie streaming platforms and other entertainment providers. This research aims to enhance users' movie-watching experience by suggesting movies that match their preferences, leading to greater satisfaction and engagement.

Introduction

Movie recommendation systems have become increasingly popular in recent years, with the rise of online streaming platforms such as Netflix and Amazon Prime. These systems aim to suggest movies to users that they are likely to enjoy based on their past preferences.

However, the traditional movie recommendation systems that use either collaborative filtering (CF) or content-based filtering (CBF) techniques have several limitations that can negatively affect the quality and diversity of recommendations. CF relies on user behavior data to recommend movies, which can lead to the problem of the cold start, where new users or movies have no historical data to rely on. Additionally, this technique often leads to the problem of the popularity bias, where popular movies receive more recommendations regardless of their quality. CBF, on the other hand, recommends movies based on their features, which can lead to limited diversity in recommendations and a lack of exposure to new and different types of movies.

To address these limitations, this research proposes the development of a hybrid movie recommendation system that combines CF and CBF techniques that can address the limitations of traditional recommendation systems, evaluate the performance of the proposed hybrid system, and compare it to traditional CF and CBF techniques. The MovieLens 25M dataset from GroupLens Research will be used to train and evaluate the system. This dataset contains ratings from over 162,000 users on approximately 62,000 movies, making it a rich source of data for building a movie recommendation system.

The proposed system will utilize the strengths of both CF and CBF techniques to provide more accurate and diverse recommendations. CF will be used to identify similar users and movies based on past behavior, while CBF will be used to analyze movie attributes such as genre, director, and cast to make recommendations. By combining these two techniques, we hope to overcome some of the limitations of each method and provide a more comprehensive and personalized recommendation system.

Problem Statement

The problem statement for building a movie recommendation system is that the system may encounter challenges such as data sparsity, cold start problem, and lack of diversity. These issues can make it difficult to accurately capture user preferences and provide recommendations that are relevant and diverse.

Data sparsity

Data sparsity occurs when there is a limited amount of data available for some movies, which can make it challenging to provide accurate recommendations.

Cold start problem

The cold start problem arises when there is a lack of data available for new users, which can make it difficult to provide personalized recommendations.

Lack of diversity

The lack of diversity issue can arise when the system tends to recommend similar movies, which can limit the user's exposure to new and different content.

Artificial Intelligence Techniques

Collaborative Filtering

Collaborative filtering utilizes past user behavior and preferences to recommend movies to users with similar tastes. This technique can help address the cold start problem by

identifying similar users who have rated the same movies and providing recommendations based on their preferences. It can also address data sparsity by filling in missing data through predictions based on the preferences of similar users. Collaborative filtering can also help increase diversity by identifying less popular movies that are similar to those the user has rated positively.

Content-based Filtering

Content-based filtering, on the other hand, utilizes information about the movie's characteristics to recommend movies to users. This technique can help address the cold start problem by recommending movies based on the user's preferences for certain genres or actors. It can also address data sparsity by providing recommendations based on a movie's specific attributes, such as plot, genre, or director. Content-based filtering can also increase diversity by recommending movies that have similar attributes but may not be as popular as those the user has already seen.

Hybrid system

To address the challenges of data sparsity, cold start problem, and lack of diversity effectively, a hybrid system that combines collaborative filtering and content-based filtering can be used. This approach leverages the strengths of both techniques and can help mitigate the limitations of each. A hybrid system can provide more accurate recommendations, increase diversity, and handle data sparsity and the cold start problem more effectively by incorporating information about both user preferences and movie attributes.

References

- Hu, Y., Koren, Y., & Volinsky, C. 2008. Collaborative Filtering for Implicit Feedback Datasets. 2008 Eighth IEEE International Conference on Data Mining, 263-272.
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. 2001. Item-Based Collaborative Filtering Recommendation Algorithms. In Proceedings of the 10th International Conference on World Wide Web, 285–295. Association for Computing Machinery.
- Li, L., Zhang, Z., Zhang, S., & Jiang, Y.Z. 2021. Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation. Sci. Program., 2021.
- Kumar, M., Yadav, D., Singh, A.K., & Gupta, V.K. 2015. A Movie Recommender System: MOVREC. International Journal of Computer Applications, 124, 7-11.asd
- Khalaji, M., Dadkhah, C., & Gharibshah, J. 2020. Hybrid Movie Recommender System based on Resource Allocation. The CSI Journal on Computer Science and Engineering, 17(2).
- Geetha, G.P., Safa, M., Fancy, C., & Saranya, D. 2018. A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System. Journal of Physics: Conference Series, 1000.
- Sharma, P., & Yadav, L. 2020. Movie Recommendation System Using Item Based Collaborative Filtering. Comp-SciRN: Audio.