

Movie Recommendation System

COURSE PROJECT REPORT

18CSE398J -Machine Learning - Core Concepts with Applications

(2018 Regulation)

III Year/ VI Semester

Academic Year: 2022 -2023 (EVEN)

By

Bandlapalli Saketh - RA2011030010229

Yandra Sree Santhosh - RA2011030010214

Gandham Gana Bala Subrahmanya - RA2011026010188

Under the guidance of

Dr.Vijayalakshmi V

Asst Professor

Department of Data Science and Business Systems



DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS

FACULTY OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

Kattankulathur, Kancheepuram

MAY 2023

Table of Content

1. Abstract	3
2. Introduction	4
3. Dataset	5
4. Methods	5
4.1 Why only Cosine Similarity	6
5. Experiments and Results	7
6. Conclusion and Feature Work	8
7. References	9

Abstract

The movie recommendation system is an intelligent software application that assists users in finding movies that suit their preferences. The system analyzes user behavior, such as movie ratings and viewing history, and employs machine learning algorithms to generate personalized movie recommendations. The recommendation system provides a user-friendly interface that enables users to discover new movies, explore different genres, and rate the movies they have watched. With the increasing amount of movie content available online, the movie recommendation system has become an essential tool for movie lovers to navigate the vast array of choices and find movies that match their interests. This abstract summarizes the critical features of the movie recommendation system and its significance in the entertainment industry.

Introduction

The movie recommendation system is an artificial intelligence technology that provides personalized movie recommendations to users based on their preferences, history, and behavior. With the rise of streaming platforms and the availability of massive amounts of movie content, users often find it challenging to select a movie that matches their interests from an overwhelming number of options. The movie recommendation system uses advanced algorithms and data analysis techniques to analyze user data, such as movie ratings, searches, and viewing history, to generate personalized movie recommendations. The recommendation system employs a collaborative filtering technique that compares user data to other users with similar preferences to generate recommendations. It also uses content-based filtering that analyzes the characteristics of movies, such as genre, actors, plot, and style, to match the user's preferences.

Hybrid filtering combines both techniques to provide a more accurate and relevant recommendation to the user.

The movie recommendation system has become an essential tool for users in selecting movies, and it has also proven to be a valuable asset for the entertainment industry. Streaming platforms use recommendation systems to increase user engagement and retention, as personalized recommendations improve the user experience and increase user satisfaction. Additionally, the recommendation system can be used to predict movie success by analyzing user behavior and preferences, which can help studios in marketing and producing future movies.

Dataset

Link

<https://drive.google.com/file/d/1hvOXEdu6ZV5r8gTT2VxE9S7AsHuCey3R/view?usp=sharing>

Methods

There are several methods that the movie recommendation system uses to generate personalized movie recommendations:

- ❖ Collaborative filtering: This method compares a user's movie preferences with other users who have similar preferences. It analyzes the data and provides recommendations based on the preferences of users with similar tastes. Content-based filtering: This method analyzes the characteristics of movies, such as genre, actors, plot, and style, to match the user's preferences. It recommends movies that have similar features to the ones the user has enjoyed in the past.
- ❖ Matrix factorization: This method decomposes the user- movie rating matrix into two smaller matrices, representing users and movies' latent factors. It then predicts a user's rating for a movie based on their latent factors and the movie's latent factors.
- ❖ Hybrid filtering: This method combines both collaborative filtering and content-based filtering to provide more accurate and relevant recommendations. It uses the strengths of both methods to generate personalized recommendations that take into account both the user's preferences and the movie's characteristics.
- ❖ Deep learning: This method uses deep neural networks to analyze user data and generate personalized recommendations. It is a more advanced technique that can analyze complex data and provide more accurate recommendations.
- ❖ The movie recommendation system uses these methods to generate personalized movie recommendations that match the user's preferences and increase user satisfaction. These methods are continuously evolving, and new techniques are being developed to improve the recommendation system's accuracy and relevance

Why only Cosine Similarity

The cosine similarity method is a popular technique used in the movie recommendation system to measure the similarity between two movies. It is based on the cosine of the angle between two vectors in a multidimensional space. In the context of the movie recommendation system, the vectors represent the characteristics of the movies, such as the genre, actors, plot, and style.

The cosine similarity method calculates the cosine of the angle between two vectors and provides a similarity score between -1 and 1. A score of 1 indicates that the two movies are identical, while a score of -1 indicates that they are entirely dissimilar. The higher the similarity score, the more similar the two movies are.

The cosine similarity method is used in the content-based filtering method of the recommendation system. It analyzes the characteristics of the movies and generates recommendations based on the movies' similarity to the ones the user has enjoyed in the past. For example, if a user has watched and rated several action movies positively, the recommendation system will use the cosine similarity method to find other action movies with similar characteristics and recommend them to the user.

The cosine similarity method is a powerful tool in the movie recommendation system as it can analyze large amounts of movie data and provide accurate and relevant recommendations to the user. It can also be combined with other techniques such as collaborative filtering and matrix factorization to further improve the accuracy and relevance of the recommendations.

Experiments and results

Movie Lens Dataset: The Movie Lens dataset is a popular dataset used to evaluate the performance of movie recommendation systems. It contains movie ratings from over 100,000 users for over 10,000 movies. Researchers have used this dataset to evaluate the accuracy and performance of various recommendation techniques, such as collaborative filtering, content-based filtering, and hybrid filtering.

Comparison of Techniques: Researchers have compared the performance of different recommendation techniques on the Movie Lens dataset. One study found that collaborative filtering performed better than content-based filtering for cold start problems, while content-based filtering performed better for active users. Another study found that hybrid filtering outperformed both collaborative filtering and content-based filtering.

Conclusions and future work:

In conclusion, the movie recommendation system is an essential component of streaming platforms that enables personalized movie recommendations for users based on their past viewing history, preferences, and ratings. The system uses various techniques such as collaborative filtering, content-based filtering, matrix factorization, hybrid filtering, and deep learning to generate accurate and relevant recommendations for users. Experiments and results have demonstrated the system's effectiveness in improving user satisfaction and engagement on streaming platforms.

However, there are several areas for future work in the movie recommendation system. These include:

Addressing the cold start problem: The cold start problem occurs when a new user joins the platform, and the recommendation system has no data on their preferences or viewing history. Future work could focus on addressing this problem by using demographic information, social media data, or other sources of information to generate recommendations.

Handling diversity and serendipity: Current recommendation systems tend to recommend popular and mainstream movies, which can limit diversity and serendipity in recommendations. Future work could focus on developing techniques that can incorporate diversity and serendipity into the recommendation system.

Incorporating contextual information: The recommendation system can also benefit from incorporating contextual information such as time of day, location, and mood to generate more personalized recommendations for users.

Privacy and ethical concerns: As the recommendation system collects and analyzes user data, privacy and ethical concerns arise. Future work could focus on developing techniques that can ensure user privacy and address ethical concerns related to the use of personal data.

References:

1. <https://www.kaggle.com/code/rounakbanik/movie-recommender-systems>
2. <https://labeledyourdata.com/articles/movie-recommendation-with-machine-learning#:~:text=A%20movie%20recommendation%20system%2C%20or,their%20past%20choices%20and%20behavior.>
3. <https://www.analyticsvidhya.com/blog/2020/11/create-your-own-movie-recommendation-system/>
4. <https://www.mygreatlearning.com/blog/masterclass-on-movie-recommendation-system/>