

Movie Recommendation Using Vectorisation

Darshit Madhukar Gharat

M.E Data Science, Zeal College of Engineering and Research, Pune, Maharashtra, India

ABSTRACT

In today's busy world, entertainment has become a necessity for everyone to refresh our energy and mood. Entertainment refreshes our mind and makes us confident for work so that we can perform more enthusiastically. For revitalising ourselves, we can either listen to music or can watch movies of our choice. For this, one can make use of Recommendation Systems which are more reliable, since in searching we require more and more time which one cannot afford to waste. Here I have designed a system for Movie Recommendation. Now, to improve the accuracy of a movie recommender, a Hybrid approach by combining collaborative filtering and content based filtering using Support Vector Machine as a classifier is presented in the proposed methodology and comparative results have been shown which depicts that the proposed approach shows an improvement in the accuracy, quality and scalability of the movie recommendation system than the pure approaches in three different datasets. Hybrid approach helps to get the advantages from both the approaches as well as tries to eliminate the drawbacks of both methods.

Keywords: Recommendation System, Movie Recommendation, SVM, Entertainment.

Article Info

Volume 9, Issue 2

Page Number : 524-528

Publication Issue

March-April-2022

Article History

Accepted : 03 April 2022

Published : 20 April 2022

I. INTRODUCTION

A recommendation system is a model that is commonly used for filtering information. Here it tries to predict the users preferences and provide the optimal suggestions based on their preferences. Now a days, these systems have become popular and are widely used today in areas such as movies, books, videos, clothing, music, restaurants, food, places and many other utilities. These systems collect information according to user's choice or preferences, and then use this information to improve their suggestions in the future. Movies are now become an

important part and parcel in everyone's life. There are different types of movies or one can say different genres where most of them are for entertainment, some for educational purposes, animated movies for children while some are horror movies or action films. Movies can be easily differentiated through their genres like drama, comedy, action, thriller, animation etc. Other way to distinguish among movies, they can be either by star-cast, releasing year, director, language etc. While searching movies online, there are millions of movies to search for and finding a movie according to our mood and liking is a very difficult task unless we know the name of

movie. But again we avoid watching same movie again and again. Movie Recommendation Systems helps user to search his preferred movies among all of these different types of movies and hence reduce the time for searching.

So it is expected that the movie recommendation system should be very accurate in prediction and should provide us with the recommendation of movies which are almost same or should match closely with our preferences. Now a days, N number of companies are making use of recommendation systems to increase the user interaction and enhance ones experience towards that company or product. Movie Recommendation System is very powerful and important system and has various benefits.

II. METHODS AND MATERIAL

The hybrid approach that is proposed in an integrative method by merging SVM as classifier method and genetic algorithm based weighted similarity measure to construct a movie recommendation system. The proposed movie recommendation system gives finer similarity metrics and quality than the existing Movie recommendation system but the computation time which is taken by the proposed recommendation system is more than the existing recommendation system. This problem can be fixed by taking the clustered data points as an input dataset

The proposed approach is for improving the scalability and quality of the movie recommendation system. We use a Hybrid approach, by unifying Content-Based Filtering and Collaborative Filtering, so that the approaches can be profited from each other. For computing similarity between the different movies in the given dataset efficiently and in least time and to reduce computation time of the movie recommender engine we used cosine similarity measure.

Agile Methodology:

1. **Collecting the data sets:** I have collected all the required data set from Kaggle web site. In this project we require `tmdb_5000_movie.csv`, `tmdb_5000_credits.csv` files.
2. **Data Analysis:** This is to make sure that the collected data sets are correct and analysing the data in the csv files that is after checking whether all the column fields are present or not in the data sets.
3. **Algorithms:** In my project, I have only two algorithms one is cosine similarity and other is single valued decomposition are used to build the machine learning recommendation model.
4. **Training and testing the model:** After implementing algorithm, we will have to train the model to get the desired result of the user. I have tested it several times if the model recommends different set of movies to different users.
5. **Improvements in the project:** In the later stage we can implement different algorithms and methods for better recommendation also we can improve the interface.

III. RESULTS AND DISCUSSION

Since this project is movie recommendation system, one can develop a movie recommendation system by using either content based or collaborative filtering or combining both. In our project we have developed a hybrid approach i.e combination of both content and collaborative filtering .Both the approaches have advantages and dis-advantages.in content based filtering the it based on the user ratings or user likes only such kind of movie will recommended to the user. Advantages: it is easy to design and it takes less time to compute Dis-advantages: the model can only make recommendations based on existing interests of the user. other words, the model has limited ability to

expand on the users' existing interests. In Collaborative filtering the recommendation is comparison of similar users.

Advantages:

No need domain knowledge because the embeddings are automatically learned. The model can help users discover new interests. In isolation, the ML system may not know the user is interested in a given item, but the model might still recommend it because similar users are interested in that item.

Dis-advantages:

The prediction of the model for a given (user, item) pair is the dot product of the corresponding embeddings. So, if an item is not seen during training, the system can't create an embedding for it and can't query the model with this item. This issue is often called the cold-start problem. The hybrid approach will resolve all these limitations by combining both content and collaborative filtering

IV. LITERATURE SURVEY

Over the years, many recommendation systems have been developed using either collaborative, content based or hybrid filtering methods. These systems have been implemented using various big data and machine learning algorithms.

1. **Movie Recommendation System by K-Means Clustering AND K-Nearest Neighbour** A recommendation system collect data about the user's preferences either implicitly or explicitly on different items like movies. An implicit acquisition in the development of movie recommendation system uses the user's behaviour while watching the movies. On the other hand, a explicit acquisition in the development of movie recommendation system uses the user's previous ratings or history. The other supporting technique that are used in the development of

recommendation system is clustering. Clustering is a process to group a set of objects in such a way that objects in the same clusters are more similar to each other than to those in other clusters. K-Means Clustering along with K-Nearest Neighbour is implemented on the movie lens dataset in order to obtain the best-optimized result. In existing technique, the data is scattered which results in a high number of clusters while in the proposed technique data is gathered and results in a low number of clusters. The process of recommendation of a movie is optimized in the proposed scheme. The proposed recommender system predicts the user's preference of a movie on the basis of different parameters. The recommender system works on the concept that people are having common preference or choice. These users will influence on each other's opinions. This process optimizes the process and having lower RMSE.

2. **Movie Recommendation System using Collaborating Filtering:** By theory proposed by ChingS eh(Mike)Wu, Deepti Garg, Unnathi Bhandary, Collaborative filtering system will analyse the user's behaviour and preferences and will predict what their liking based on similarity with other users. There are two types of collaborative filtering systems; they are
 - a) user-based recommender.
 - b) item-based recommender.
3. **Use-based filtering:** User-based preferences are very common in the field of designing personalized systems. This approach is usually based on liking of the user. The process starts with users giving ratings on scale of to movies. These ratings can be implicit or explicit. Explicit ratings are when these users explicitly rates the item on some scale or indicates a thumbs-up or thumbs-down to the item. More often, explicit ratings are hard to gather as not all user are

interested in providing honest feedbacks. In such scenarios, we usually gather implicit ratings based on ones behaviour. For instance, if a user buys an item or product more than once, then it indicates that it is a positive preference. In context to movie systems, we can imply that if a user watches the entire movie, then he/she has some sort of liking to it. Note there are no clear rules in determining implicit ratings. Next, for each user, we first find predefined number of nearest neighbours. The assumption is that if two users' ratings are highly correlated, then these two users must enjoy similar items and products is used to recommend items to users.

4. Item-based filtering: Unlike the user-based filtering, item-based mainly focuses on the similarity between these item's users like instead of the user themselves. The most similar items are computed ahead of time as then for recommendation, the items which are most similar to the target items are recommended to the user.

V. SYSTEM ANALYSIS AND DESIGN

System Architecture:

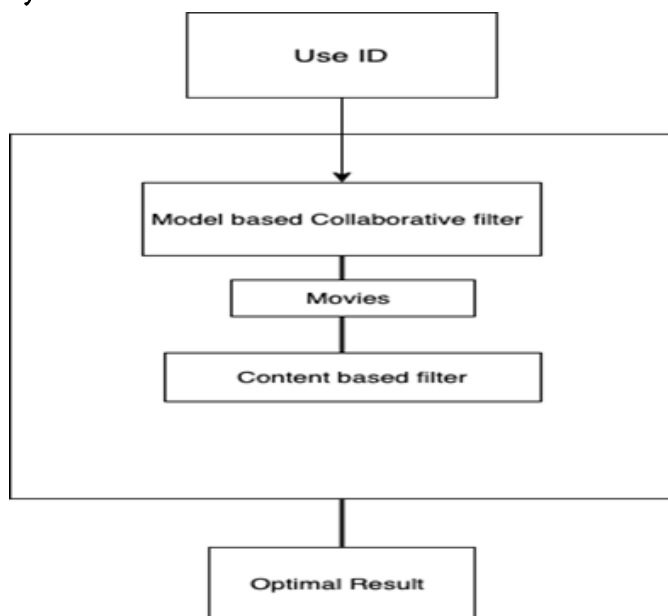


Fig:-4.1 Architecture for hybrid approach

For each different individual, different set list of movies are recommended. As user login or enters the user id, the system will recommend the set of 10 movies to the particular user by combining the both the set of movie based on the user the hybrid model will recommend the single list of movie to the user.

VI. CONCLUSION

In this project, to improve the scalability, accuracy and quality of movie recommendation system, a Hybrid approach unifying collaborative filtering and content based filtering using Vector classifier and Cosine Similarity is presented in the proposed methodology. Existing pure approaches and proposed hybrid approach is implemented on three different Movie datasets and the results are compared among them. Comparative results depicts that the proposed approach shows an improvement in the accuracy, quality and scalability of the movie recommendation system than the pure approaches. Also, computing time of the proposed approach is lesser than the other two pure approaches.

VII. REFERENCES

- [1]. Hirdesh Shivhare, Anshul Gupta and Shalki Sharma (2015), "Recommender system using fuzzy c-means clustering and genetic algorithm based weighted similarity measure", IEEE International Conference on Computer, Communication and Control.
- [2]. Manoj Kumar, D.K. Yadav, Ankur Singh and Vijay Kr. Gupta (2015), "A Movie Recommender System: MOVREC", International Journal of Computer Applications (0975 – 8887) Volume 124 – No.3.
- [3]. RyuRi Kim, Ye Jeong Kwak, HyeonJeong Mo, Mucheol Kim, Seungmin Rho, Ka Lok Man, Woon Kian Chong (2015), "Trustworthy Movie Recommender System with Correct Assessment and Emotion Evaluation", Proceedings of the

International MultiConference of Engineers and
Computer Scientists Vol II.

- [4]. Zan Wang, Xue Yu*, Nan Feng, Zhenhua Wang (2014), "An Improved Collaborative Movie Recommendation System using Computational Intelligence", Journal of Visual Languages & Computing, Volume 25, Issue 6.
- [5]. Debadrita Roy, Arnab Kundu, (2013), "Design of Movie Recommendation System by Means of Collaborative Filtering", International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 4.