

Movie Recommendation System Using Machine Learning

Aryan Raj
School Of Computing Science And
Engineering
Galgotias University Uttar Pradesh,
India
aryan_raj.scsebtch@galgotiasuniversit
y.edu.in

Harsh Panjwani
School Of Computing Science And
Engineering
Galgotias University
Uttar Pradesh, India
harshpanjwani116@gmail.com

Soumya Ranjan
School Of Computing Science And
Engineering at Galgotias University
Uttar Pradesh, India
soumya_ranjan.scsebtch@galgotiasuni
versity.edu.in

Abstract— Recommender system as the name suggests recommends something whether it is a movie or any other item. The research paper mainly focuses on collaborative filtering based recommender system. collaborative filtering is a method in which a recommendation is given to the user based on his past search and based on the profile of the others who have similar choices. Recommender System nowadays has become a crucial need in our life because if there is no recommendation system then each time we visit a movie site or youtube, it will show the items which are of no use to us. It saves our time from unnecessary search and provides an exact choice based content. There is user to user, user to an item, and item to the item-based recommender system. Facebook is an example of the user to user based recommender system. But mainly the three most popular recommender systems are content-based, collaborative filtering, and hybrid-based recommender system. Various movie datasets are used to recommend movies to users. Nowadays mostly the recommender systems are based on the combination of hybrid and collaborative filtering models. In this project, we will mainly focusing on the collaborative filtering model of the recommendation system.

Keywords-Collaborative filtering, hybrid-based, content based, recommender system.

1. Introduction

A Recommendation system is a type of filtering system which is used to predict the choices of users Based on their likes and dislikes. There are various applications of recommender systems in our day-to-day life starting from friends on FB page to items on e-commerce site, to movies on Netflix or any other similar platforms. A recommender system applies different algorithms to filter data and then recommends the most suitable items to the user. Based on the previous behavior of the user it recommends items and if there is no history of the user or there is a new customer who has a sign up for the first time, that we have a different proposed solution or different methods of a recommender system. There are three main types of recommendation systems. The first one is a Demographic-based filtering system that gives recommendations to the users based on the popularity or current high trending topics. Since every user has a different choice of view and various other perspectives this algorithm is quite simple. The next one is the content based [12] recommender system in which we try to profile the users' choices using the information collected from

similar profiles and recommend the items to the user based on that profile and the last one is collaborative filtering [13] based recommender system

1.1 KNN Algorithm

Collaborative filtering [8] system is widely used and it is more suitable than another filtering system. Its of two types first One is item-based and the second one is user-based It is generally preferred to use item based [9] collaborative filtering method because the user based method is difficult to implement. After all, the choice of users changes from time to time whereas items generally do not change very often. Item-based filtering can also be done without the use of any software and there is also no need to use any algorithm for training the data set.

To implement item-based collaborative filtering, generally, KNN is a good approach. KNN is also known as the K-nearest neighbor algorithm [14]. It is based on clustering [6] where it uses a database where there is a collection of data points in the form of clusters and the sample experimental data is compared to find the most suitable nearest data points.

KNN algorithm [11] does not make any extraordinary The assumption of the existing data distribution but mainly focuses on the item feature similarity. when the KNN algorithm observes the similarity between the different clustered data points then it calculates the distance between the target movie and all other movies in the database and then finally it returns the best movies or top k-Nearest neighbor movies and returns them as the most recommended movies.

In calculating the distance between several other movies in the database it uses Euclidean distance but if the data is of high dimensionality it fails and at that position, it uses cosine similarity for finding the nearest distance between different data points.

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$

$$\|\vec{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2}$$

$$\|\vec{b}\| = \sqrt{b_1^2 + b_2^2 + b_3^2 + \dots + b_n^2}$$

Figure:- cosine similarity

There are other methods like locality sensitive hashing[10] based on the hash map algorithm to calculate the distance between K-nearest neighbor which we will avoid in this discussion because of its vast nature.

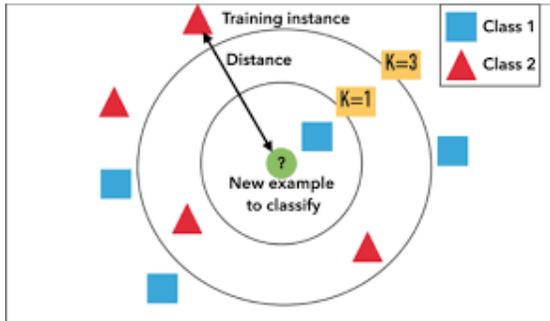


Figure: KNN ALGORITHM

1.2 Literature Review

A film recommendation framework focused on collaborative suggestions[11], was suggested by Kumar et. Filtering method Collaborative filtering takes all users data and is focused on that information to Produce recommendation. H.KVirk.[3] also proposed a hybrid framework. In this method, the system integrates both a collective and a content-based approach. De Campos et al.[2] made, to an Analysis in all the conventional methods for the suggestion. Since both of these approaches have certain limitations He recommended another framework that is a mixture of the Bayesian network with some losses. Clustering was suggested by Kuzelewska[6] as an alternative to working with therules. Two clustering approaches were analyzed first one is Centroid-based andthe other one Method Focused on Memory. The outcome was the production of reliable advice. Sharma and Mann[7] in their research paperhave used all three methods that are content, collaborative, and hybrid.

1.3 . Existing System

The explanation behind this move is the success gained by Companies such as Netflix, whose primary focus is customer loyalty. Individuals can physically pick movies to watch from movie libraries until the recommendation system operates. The users to get the most popular movie recommended have to read the reviews by other customers or have to look at the star cast or look at the budget or the ratings given by the users. As there is a wide range of people with different reviews, rating about a particular movie or the movie which you like has not even been rated or rated worst by someoneso this technique is not feasible and optimistic for better search results.

1.4 . Study of the Research Work

We use a pre-channel in the proposed model before implementing the k-means clustering algorithm. The credits used to practice centroid detachment from each stage are:

- A. Genre
- B. Rating

The different range of movies have different variety of aspects and characteristics .so to get the best recommendation of a particular movie it should be rated currently or the reviews should be current, so far the rating characteristic has become more relevant than different properties. The user should rate at least 7 to 8 moviesbased on the genre so that he must be recommended based on that. let us suppose a new user visited the movie site then to get recommended he must search a movie in the search bar and rate some movies based on genres and the recommender system will recommend movies based on his liking and the users having a similar profile and then will the films be proposed to him/her.

1.5 .Problem Description

This suggestion scheme suggests to users numerous movies. As this method is based on a collaborative approach, in comparison to current frameworks which are basedon a content-based approach, there would be increasingly transparent effects. Content-based recommendation systems[4] are limited to individuals, Items are not recommended out of the box by these systems. These programs run on the scores of individual users, thus restricting the ability to explore further. Although our interactive approach based framework measures the relation between different customers and relies on their scores, it recommends films to those users which are havingof choices similar to the required user, enabling the customers to have a wider range of view anda better approach to finding movies of their choice. Our project is a python flask based application where the user is required to rate movies based on the genres he likes and based on using collaborative filtering movies are recommended to the user.

1.6 Solution Methodologies

Collaborative filtering[15] works are based on users that have similar tastes. In table 1. as we can see that the user A and B have given similar ratings to the Sci-fi and horror genre .so when next time user B wants the recommendation for movies then he will recommend movies similar to user A And also according to the history.

Genres/Us ers	comedy	horror	Sci-fi	Action	Rom ance
A	1	5	5	4	3
B	2	5	4		
C	4	1		3	5
D	2	3	4	4	2

Figure: Ratings based on collaborative filtering

1.7. Flowchart of the proposed system

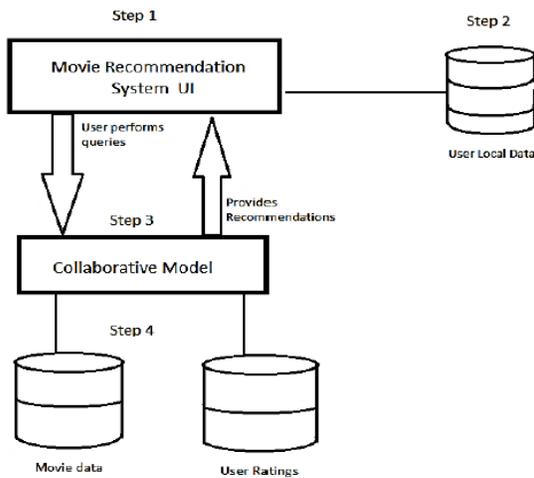


Figure 2. The architecture of movie recommendation system.

1.8. Implementation of our project

Firstly, the user will open the web page for the movie recommendation system then he/she will give ratings to the movies based on a different genre,s and then the internal system will apply item - based collaborative filtering on the data to give a list of similar recommended movies to the users. The implementation is described in the following figures.



Figure 1. Home page



Figure 2. Rating page

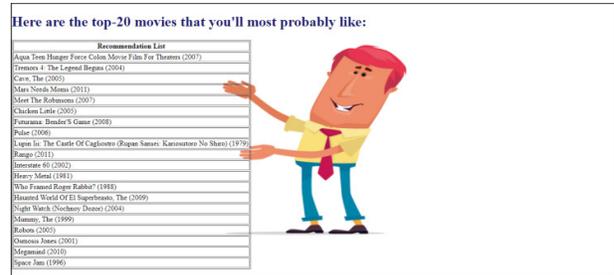


Figure 3. Recommendation page

2. DATASETS SOURCE FOR PROJECT

For recommending the movies using the movie recommendation system the data has been taken from Kaggle and the movie lens dataset. The data sets mostly rely on the data used by amazon prime movies, Netflix, and Hulu. The ratings are given similar to the IMDB website. Few references are also taken from tmdb for more analysis.

3. Conclusion

The recommendation system that we have made in this project is mostly based on item-based collaborative filtering.KNN algorithm with cosine similarity is applied to a 20 million rating movie lens dataset. When we compare our proposed work with the existing system, it is found to be more accurate and efficient. The item-based collaborative filtering is found more accurate as compared to the user based. The main purpose of this project is to improve the old recommendation engine. our proposed work is based on item-based collaborative filtering that mainly analyses the user's past search and also compares it with the users having the same profile. our project is mainly the web application that allows the user to give ratings to the movies according to the genres they like and based on that they are recommended the most relevant movies

4.REFERENCES

- [1] Recommender System based on the collaborative approach by M Kumar, D. Yadav, A.singh, and Vijay Kr. International journal of computer Application 124(4):7-11.DOI:10.5120/ijca.2015904111.
- [2] Costin-Gabriel chiru, V- Nicolae Dinu, M Macri, C. Preda IEEE, ICCP DOI:10.1109/ICCP.2015.7312611.
- [3] Er. M. Singh, Er A. Singh, Er. H.K. Virk, Design of hybrid Movie Recommender System. (IJJET).
- [4] Recommender System Challenges and solutions Survey by MH Mohamed, M. Khafagy, M.H Ibrahim.2019 ITCE.
- [5] <https://towardsdatascience.com/prototyping-a-a-recommender-system-step-by-step-by-knn-algorithm>. written by Kevin Liao.
- [6] A hybrid Recommender systemusing a clustering algorithm on movie Dataset by Kuzelewska.Studies in Logic , 37(50) 204 DOI:10.2478/slur-2014-0021.
- [7] Combining a content Based and collaborative recommendation system based on a hybrid approachby Luis

de compos, J.M. F-Luna, JF -Huete, M.A.R-Morals. International Journal of Approximate Reasoning 151(7):785-799. DOI:10.1016/j.ijar.2010.04.001.

[8] A Review Paper on Collaborative filtering Based Movie Recommender system by N.Raval, V Khedkar. International Journal of Scientific and Technology Research Volume 8, ISSUE,,2019.

[9] Movie Recommendation system using collaborative model by M.K Kharita, Atul Kumar, P Singh, NIT Hamirpur(ICSCCC) 2018.

[10] Locality-Sensitive Hashing scheme based on the heap sort algorithm by B.fang, H Huang, the 14th ICCSE 2019, Toronto, Canada.

[11] Research on K-means clustering algorithm. Third International Symposium on Intelligent Information Technology and Security Informatics.

[12] Movie Recommender System using Machine Learning. F. Furtado, A Singh IJRIC 2020.

[13] A movie Recommender system using collaborative filtering by M. Gupta, A Thakkar, Aashish, ICES2020.

[14] Movie Recommendation System using K-means clustering algorithm and k-nearest neighbor. R.ahuja, A Solanki, Anand Nayyar. IEEE 2019.

[15] Movie Recommendation System using collaborative filtering by N.kalra, D.yadav, Gourav Bathwa International Journal on Future Revolution in Computer science and communication engineering.