

A Movie Recommender System: Using Machine Learning

Mansi Patil¹ and Riya Mishra²

Research Scholars, Department of MCA^{1,2}

Thakur Institute of Management Studies, Career Development & Research (TIMSCDR), Mumbai, India

mansipatil2504@gmail.com and riya443450@gmail.com

Abstract: A movie recommendation system has changed the style of searching the thing base on our interest. This is information filtering approach that is used to predict the preference of the user. The recommender System has become popular in very search base system like book, news, articles, music, movie etc. We have proposed a movie recommendation System in this paper. It is based on hybrid system using both content base filtering and Collaborative filtering. There are so many movies available now that no one can see them all. and even those movies are not interesting so the person might get confused which movie to watch. Watch find your movie filters out movies according to horror, romantic, animated, comedy or historic and even based on language. place so that there is no need of finding or downloading movies from different sites that might increase the chance of virus or slow down the device..

Keywords: K-Means, Recommendation System, Data Mining, Clustering Movie, Collaborative Filtering, Content Based Filtering

I. INTRODUCTION

Earlier people use to buy DVDs to watch movies which would probably not have good quality, later on when internet took over the world it became easier to download movies from different sites but still downloading movies were a bit difficult because of the slow speed of internet. Soon after, the hybrid system emerged, in which individuals could simply search for their favourite film and view it without difficulty. At a time, person could only buy 2 to 3 DVDs or download movies because it would cost more and would even occupy some space, so the person could only watch that particular movie present in DVDs or the one which is download in system. To avoid this problem this system contains all types of movies that one can watch without waiting for the movie to be downloaded, where earlier if the movie was to be downloaded one had to wait till the time movie is being downloaded and later on watch that movie, this system analysis the requirements of customer and accordingly recommends the movie to user. If the person usually watches particular type of movie like if he/she just prefer to watch comedy movie than it will also suggest some latest movie that are present at the nearby cinemas with the distance and the show time for particular movie. We can classify the recommendersystems in two broad categories:

1. Collaborative filtering approach
2. Content-based filtering approach

1.1 Collaborative Filtering

The collaborative system recommends based on similarity measures between users and/or items. The system recommends those items that are preferred by similar kind of users.

Collaborative filtering has many advantages

- It is content-independent i.e., it relies on connections only
- Since in CF people makes explicit ratings so real quality assessment of items are done.

It provides recommendation based on user's similarity rather than item's similarity

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

1.2 Content-based Filtering

Content-based (cognitive filtering) recommendation system is to recommend an item based on a comparison between the content of the items and a user profile. In simple words, I may get recommendation for a movie based on the description of other movies. The idea underlying collaborative filtering and how it may be used with user or movie cooperation. Taking an example, there are two users A and B, user A likes movie P, Q, R, S and user B likes movies Q, R, S, T. Since movies Q, R and S are similar to both users, therefore, movie P will be recommended to user B and movie T will be recommended to user A.

II. RELATED WORK

The authors implement two methods important for movie recommendation: movie swarm mining that mines a set of movies suited for producers for new movie planning and new item suggestion, popular and intriguing movie mining that can be utilized to solve the problem of new users. The effectiveness of our proposed methods is demonstrated using Movie lens Data Sets.

The authors develop specialized recommendation models by different types of user's choice patterns. We initially categorized the data to better understand the features of the various types of movie data and to identify groupings of genres with comparable user tendencies. We then develop a separate fabricated model (NR engine) by each type of user choice and predefine datasets (called, ml) (genre, cast, director, rating, year) that uses as inputs to learning data for a calculated class corresponding with user choice. This makes our work more effective by the way of relationship between user interest and traditional data pattern.

III. RESEARCH METHODOLOGY

3.1 The Basic K-means Algorithm

In K-means clustering technique we choose K initial centroids, where K is the desired number of clusters. Each point is then assigned to the cluster with nearest mean i.e., the centroid of the cluster. The centroid of each cluster is then updated depending on the points allocated to the cluster. We continue the operation till the cluster centre does not change (centroid). Finally, the goal of this technique is to minimize an objective function, which in this case is a squared error function. The objective function. The algorithm is composed of the following steps:

1. Select K points as initial centroids.
2. Repeat
3. From k clusters by assigning each point to its closest centroid.
4. Re-compute the centroid of each cluster.
5. Until Centroid do not change.

Figure 1. K-means Algorithm

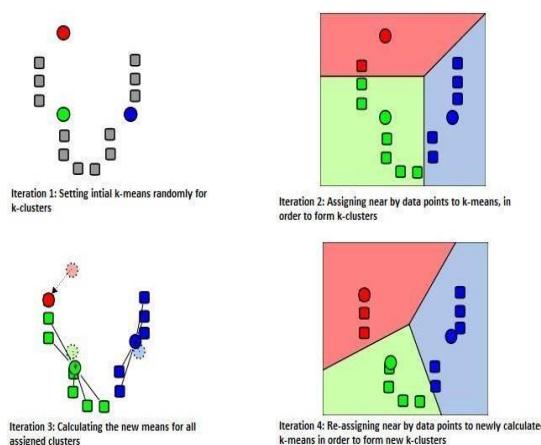


Figure 2: The 5 steps of the K-Means algorithm

3.2 Data Description

In proposed model we use a pre filter before applying K- means algorithm. The attributes used to calculate distance of each point from centroid are

1. Genre
2. Actor
3. Director
4. Year
5. Rating

Different attributes have different weights. In our research we have found that the appropriate recommendations that can be generated should be based on the as a result, we've given the rating attribute greater weight than the other attributes when it comes to movies. attributes. These ratings have been taken from movie data lens because perhaps it has the largest collection of movies alongbased on the ratings provided to these films by a vast number of individuals from all around the world. The total number of votes collected by a single film is another key parameter in our suggested model. We separated the votes into threecategories: less than or equal to 1000 votes, more than 1000 but less than or equal to 10,000 votes, and more than 10,000 votes.

Weightage and matching of attributes

A. Actor (Wa)

Wa= No. of movies of Actor(a) in data set Total no. of movies in data set Director (Wd)

Wd= No. of movies of Director(d) in data set Total no. of movies in data set

B. Rating

Rating	Weight		
	If number of votes <=1000	If number of 1000 < votes <=10000	If number of votes >10000
10	10	20	30
9	9	18	27
8	8	16	24
7	7	14	21
6	6	12	18
5	5	10	15
1-4.9	1	2	3

C. Genre (Wg)

Wg= No. of movies of Genre(g) in data set total no. of movies in data set

D. Year (Wy)

Wy= No. of movies in Year (Y) in data set Total no. of movies in data set

Total weight of a particular movie m is given by

$$W_m = W_r + W_a + W_d + W_g + W_y$$



IV. CHALLENGES FACED

1. Lack of data-biggest issue faced in recommendation system is that there needs to be a lot of data to make suggestions with effectiveness to overcome this user must rate at least 20 movies.
2. Changing user preferences- User preferences change day-by-day so it becomes difficult for the system to recommend the preferred movie.
3. Collecting information related to movies.
4. Changing data- Updating the movie collection list regularly.
5. Storage of user similarity data is a problem because amount of user data is huge.

V. CONCLUSION

In our project, collaborative filtering algorithm and content-based filtering algo have been used to predict movie recommendations based on users' ratings.

By overcoming the limitations of existing recommendation systems, it enhances performance. Traditional recommendation system was not much secure.

The primary purpose of our system is to provide movie suggestions to clients based on their viewing history and ratings. Collaborative filtering and content-based filtering are the prime approaches to provide recommendation to users. In this project, we've used collaborative filtering to predict user's movie rating.

A movie recommendation system gives a degree of convenience and customization that aids in the selection of films. The user interacts better with the system and watch movies that cater to his needs.

According to this our system has pretty good prediction performance.

REFERENCES

- [1]. Sajal halder, A.M. Jehad Sarkar and Young-Koo Lee "Movie Recommendation System Based on Movie Swarm" 2012 Second International Conference on Cloud and Green Computing.
- [2]. Shreya Agrawal and Pooja Jain "An Improved Approach for Movie Recommendation System" 2017 International Conference on I-SMAC.
- [3]. Sajal Halder and Jehad Sarkar "Mining Technique for Movie Recommendation System" 2018 International Conference on Electrical and Computer Engineering.
- [4]. Ching-Seh, Deepti Garg and Unnathi Bhandary "Movie Recommendation System Using Collaborative Filtering" 2018 4th International Conference on Electrical Engineering and Information & Communication Technology.
- [5]. Sajal halder, A.M. Jehad Sarkar and Young-Koo Lee "Movie Recommendation System Based on Movie Swarm" 2012 Second International Conference on Cloud and Green Computing.