



Movie Recommendation System using ML

¹Krishna Gudi ²Monica S, ³Aruna P, ⁴Dhruv Jyoti Shukla, ⁵Dhanalakshmi B

¹Assistant Professor , ^{2, 3, 4, 5} Student

¹Computer Science Engineering,

¹K S Institute of Technology, Bangalore, India

Abstract : In digital era, we have wide-range of resources such as items, music, books, movies and so on, finding the data based on user's interest becomes unerring, thus recommender systems will play a major role. A recommendation engine is all about suggesting a similar type of content or items by predicting user preferences. Eg: In case of OTT platforms the movies to watch, in the case of e-commerce sites the items to purchase will be suggested using recommender systems. This can be done by using Machine learning (ML) algorithms such as Content-based filtering (CBF), Collaborative filtering (CF), and hybrid filtering approach (HF). CF is classified into model-based approach and memory based approach. HF is a technique which is a combination of both content based and model based CF approach. XG Boost is used for improving the performance of recommendation engine.

IndexTerms - Machine learning (ML), user choice's, Hybrid Filtering (HF), Collaborating filtering (CF), Recommendation, XG Boost

I. INTRODUCTION

Recommender Systems are powerful filtering methods/techniques in the field of business. Google, Aha, YouTube, Netflix are all large business organizations that will use effective Recommender engines in order to grow their business. So there are mostly three types of recommender systems exists. They are content based approach (CBF), collaborative-based filtering (CF), and hybrid filtering technique (HF).

Recommender Engine is one of the facilities given to user. Recommender engines are the most instantaneously identifiable ML technique in today's technology world. We notice many services or platforms which will recommend movies that are based on previous interaction of user with the platform. They attempt to get likeliness and inclinations by ratings given by users and to find unknown movies or data that are not realized but will fascinate or attract user. Due to improvement in recommendation engines all users round-the-clock anticipate for good recommendations, this has generated a large significance in organizations for enhancing their recommender systems. However, the difficulty of the issue seems more compound than it looks. Each user has various inclinations, in addition every single user will have a different preference based upon number of factors, like time, mood of the person or depending upon the user's current work. Building a system like this is quite challenging..

II. LITERATURE SURVEY :

Collaborative filtering (CF) is classified using various approaches like matrix factorization, recommendations based on users and items. By using algorithms such as K-nearest, SVD, Alternating Least Squares (ALS). Movie recommendation systems can be improved by using the Pytorch library, in which the model would be trained to find the latent factors.[1]

Traditional approaches majorly consists of (CF) collaborative filtering, CBF Content based Filtering, The common CF techniques are neighborhood based CF, Hidden/latent factor model based CF, and Graph based CF. KNN model and XGBoost model are combined to form a hybrid collaborative filtering recommendation.[2]

Filtering techniques are rating predictions, ranking fashion supports implicit feedback. Methodology followed are User item sparse data matrix, User-to-user similarity matrix, item-to-item similarity matrix. Major problem identified is the cold start problem.[3]

Different methods used are CF, CBF, Multi - criteria recommender Systems, Risk - aware recommender, Mobile and hybrid recommender system. These are model-based approaches and uses Kernal - mapping and KNN approach. Few hybridization techniques are Weighted, Switching, Feature combination, Feature Augmentation.[4]

Collaborative filtering is based up on the behavioral learning matrix, behavioral characteristics and calculation of weights, user's similarity-based behavior, the divine response and recommender outputs. RMSE is used for evaluation of the performance. Three algorithms described are HCF, T_UCF, T_ICF. The problem identified is matrix sparsity [5].

This paper discusses about three primary techniques or approaches used in recommendation engines such as CF approach, CBF approach, HF approach. There are many difficulties and limitations for recommender systems such as cold-start problem, data sparseness, scalability, multi-objective minimization, privacy and trust of the system.[6]

In this paper, they have discussed about building a related items table which is built by item based CF algorithm, so this makes the recommender engine quick and faster. They analyze customer's present interest and past interest to provide recommendation. They mention importance of time which enhance the quality of recommender system.[7]

This majorly discuss about solving the limitation of unique recommendation that arises from disregard of data for a particular user. They propose a engine named as CBCF-content boosted Collaborative filtering (CF) is to analyze the strenuous level of every case of every trainee.[8]

III. PROBLEM STATEMENT :

“Eradication of collection of new user's information, movies with few ratings, one account with multiple users”

The cold start problem is faced when the movies have to be recommended for new users, whose preferences are not known. The movies with very few ratings and the user who is not rating most of the movies watched becomes slight difficult to be recommended as the users' preferences will not be known to the recommendation system. The multiple users will use the same account, and this is where the personalized recommendation becomes inappropriate.

IV. METHODOLOGY :

The three basic filtering approaches used in recommendation engines:

1. CBF-Content based filtering approach
2. CF-Collaborative filtering approach
3. HF-Hybrid filtering approach

1. CBF-Content based filtering:

CBF it classifies the movie based on content. The main idea behind the content based recommendation is to recommend the movie to user-X similar to previously watched movie and rated highly by X. In this approach we are going to start with the user and find out set of movies the user likes using both explicit and implicit data, here there is no need for data of other user to make recommendation to a specific user. In this approach it is able to recommend to users with unique taste. When new movie is added there is no need of any rating from others user, but it depends entirely on the features of the movie.

2. CF-Collaborative filtering:

CF is technique which is used to filter the data, its function is to find out the similar pair item by gaining the rates/likes given by the main user and make an automatic prediction which the user is looking for. The main idea behind the collaborative filtering is very simple suppose we have user-X to whom you want to make recommendation. What we are going to do we are going to find group of other users who like and dislike are similar user to X. It is based on two approaches, user to user and item to item. In user to user approach given a user we try to find other users that are similar to that user and use rating of those users to predict the ratings. In item to item approach instead of starting out with user and finding similar user we are going to start out with an item X and find similar item to item X.

3. HF-Hybrid filtering:

HF is a combination of both CF-collaborative filtering and CBF-context based filtering. This model is proposed to overcome the limitation of using CF and CBF individually, which improves the performance of system. When the data is not rated by any user, or the data with very few ratings can be filtered out by using classification (KNN) and clustering algorithms (K-means).

The proposed model in below fig, is based on hybrid approach. Its function is to approach the user by recommending various movies which the user is looking for. The user first will request for the movies which he/she is looking for. The movie which is being provided to the user after requesting for it, based on the quality and content of the movie, the user gives a proper rate according to it.

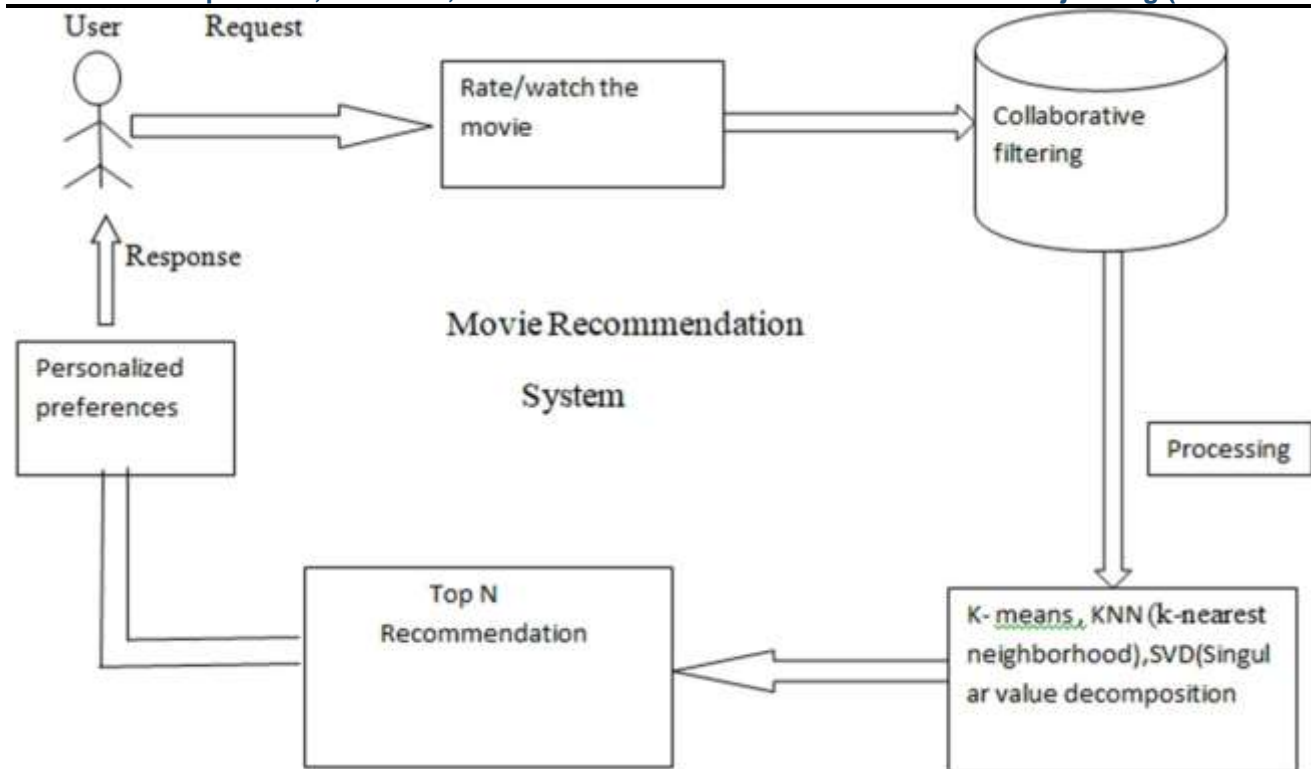


Fig : A flow chart of proposed movie recommendation system

The request which is being obtained from the user it is being sent to collaborative filtering. its function is to find out the similar pair item by gaining the rates /likes given by the main user and make an automatic prediction which the user is looking for, it is based on two approach user to user and item to item

Now the information which is being sent in form of similar item is based on the matrix factorization. It is used to reduce the dimension or size of the matrix and also decompose the user to user and item to item matrix into product of lower rectangular matrix.

By using the collaborative filtering algorithms such as k-mean and KNN. K-mean provides a partition or the division of different dataset into many cluster variables k represents the total number of group created in the groups. This algorithm help to form a new data point and this formation is based on closed centre of mass. It assigns the nearest centroid distance. It later provides similar and effective recommendations based up on the users interest.

V. CONCLUSION :

In this paper we briefly review how recommendation system help in recommending movies that are not recognized by the user but are similar to the user's interest. This proposed system consist of various methodology which provides more appropriate results and personalized recommendation. It has its own limitations.

The recommendation engine can be implemented on a large datasets, and improve its limitations by using ML algorithms and improve its accuracy.

REFERENCES

- [1] Nirav Raval, Vijayshri Khedar, (December 2019), International Journal of Scientific & Technology Research, ISSN2277-8616, Volume 8.
- [2] Wang Juan, Lan Yue-Xin, Wu Chun-Ying, Survey of Recommendation based on Collaborative Filtering, ICEMCE – 2019, JOP Conf.Series 1314.
- [3] M.Chenna Keshava, S.Srinivasulu, P.Narendra Reddy, B.Dinesh Naik, Machine Learning Model for Movie Recommendation System, IJERT (April – 2020), volume 9, Issue no. 04, ISSN: 2278-0181.

[4] Ashritha Kashyap, Sunitha B, Sneh Srivastava, A Movie Recommender System: MOVREC using Machine Learning Techniques, IJESC – 2020, Volume 10, Issue No.06, ISSN: 2321-3361.

[5] Xuejian Huang, Gensheng Wang, Learning recommendation based on hybrid collaborative filtering algorithm, ICAMLDS – 2020, JOP: Conference series. 1629(012008).

[6] Sindhu J, Roopa G M, Approaches and challenges for CF Recommendation System (2018), International Journal of Scientific Research in Computer Science, (IJSRCSEIT), ISSN:2456-3307, Volume 3.

[7] Brent Smith, Greg Linden, Two decades of Recommender Systems - Amazon, IEEE Internet Computing article(2017).

[8] Poonam Sharm, Lokesh Yadav, Movie Recommendation System Using Item-Based Collaborative Filtering(2020),International Journal of Scientific Research in Computer Science, (IJSRCSEIT), ISSN:2347-5552, Volume 8.

