Harsh Mishra, 2021, 9:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Movie Recommendation System

Harsh Mishra, Suraj Kumar, Yash Khare, Hradesh Kumar
School of Computer Science and Engineering,
Galgotias University,Greater Noida, India.
harshx2701@gmail.com, surajkumarnishad95@gmail.com, yashkhare189@gmail.com,
hradesh.kumar@galgotiasuniversity.edu.in

Abstract- Recommender systems are systems based on Machine learning algorithms that help users discover new products and services. Recommender systems are very essential in this era of internet where services are mostly handled on web rather than on a person to person basis. Since the number of choices present are overwhelming, recommendation systems are there to filter and prioritize the choices so as to present a subset of data which is supposedly the taste of that particular user. This increases the chances of their involvement and hence creates an atmosphere for happy customers in general. Since the amount of information present on the net is humongous, Recommender systems search through huge volume of information to provide user with personalized products or services as recommendations.

Keywords:- Machine learning algorithms, filter and prioritize Etc.

I. INTRODUCTION

With the tremendous growth in data present over the internet, along with the number of users created a big challenge to keep track of relevant data. Hence due to huge data lot of times it would hinder the actual information thus making it very hard or time taking process to access required services and hence was introduced the concept of personalization and prioritization of data. This concept is implemented in the form of Recommendation or Recommender system.

"Recommender systems are information filtering systems that deal with the problem of information overload by filtering vital information fragment out of large amount of dynamically generated information according to user's preferences, interest, or observed behaviour about item.

Recommender system has the ability to predict whether a particular user would prefer an item or not based on the user's profile." Recommender systems are a great utility for both consumer as well as the service provider. Recommendation systems are known to improve decision making process and quality. These also help users to find better content that would be closer to their taste rather than a random suggestion. Hence improving user engagement.

II. RELATED WORK

Movie recommendation systems using different types of technologies have been in use for quite some time now. These are based on various types of algorithms such as Content based filtering, Collaborative filtering, ALS and such. In majority of the cases, dataset used is the one provided by Movielens.

1. Collaborative Filtering:
In this type of filtering, data is filtered using the interactions and data collected by the system from other users. The basic idea here is "collaboration ". So if users have agreed to some extent in their evaluation, they are supposed to like the same type of content. Thus the input in such a system is the historical data from past interactions of users with the target items. This is often stored in form of matrix. Most collaborative filtering systems apply the so-called similarity index-based technique. In the neighborhood-based approach, a number of users are selected based on their similarity to the active user. Inference for the active user is made by

© 2021 Harsh Mishra. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

International Journal of Science, Engineering and Technology

An Open Access Journal

calculating a weighted average of the ratings of the selected users.

Collaborative-filtering systems focus on the relationship between users and items. The similarity of items is determined by the similarity of the ratings of those items by the users who have rated both items.

Collaborative filtering has the ability to recommend content to user that without the content being in the user's profile. But at the same time, it does have cons such as Cold start problem, data sparsity problem and scalability.

Collaborative Filtering is further sub divided into two forms;

1.1 User-Based: According to this logic, the recommendation system recommends items by finding similar users to the active user. Approach used here is "Nearest neighbor ", often known as KNN.

1.2 Item-Based: Instead of focusing on his friends, we could focus on what items from all the options are more similar to what we know he enjoys. This new focus is known as Item-Based Collaborative Filtering (IB-CF).

We could divide IB-CF in two sub tasks:
Calculate similarity among the items:

- Cosine-Based Similarity
- Correlation-Based Similarity
- Adjusted Cosine Similarity
- 1-Jaccard distance

Calculation of Prediction:

- Weighted Sum
- Regression

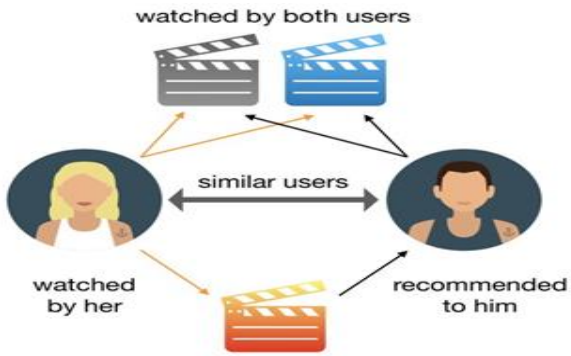

Fig 1. Collaborative Filtering.

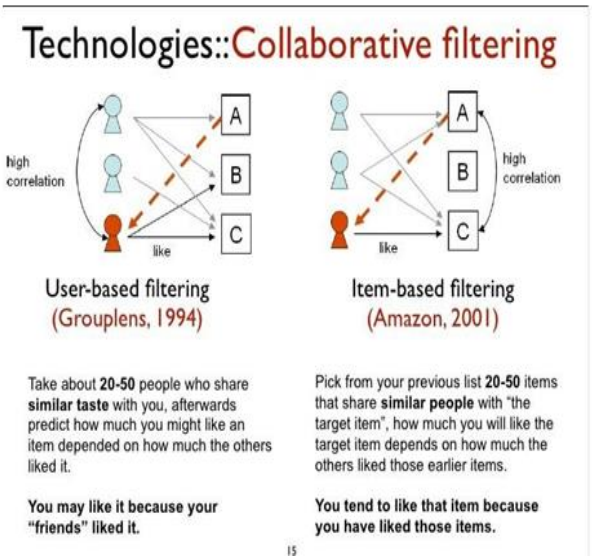

Fig 2. User based vs Item based filtering.

2. Content Based filtering:

Content based filtering is more of a domain dependent algorithm, hence can't just be applied to every single section it emphasizes more on the analysis of the attributes of items in order to generate predictions.

When documents such as web pages, publications and news are to be recommended, content-based filtering technique is the most successful. In content based filtering, recommendations are made on the past behavior of the user.

Items that are mostly related to the positively rated items are recommended to the user. Content based filtering does not rely on the evaluation of other users as it won't affect the outcome.

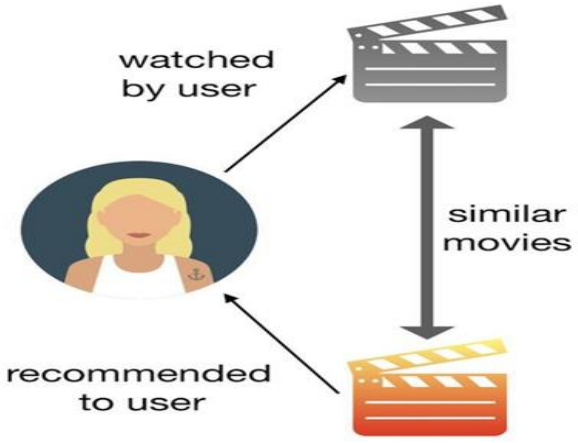

Fig 3. Content Based Filtering.

There are certain pros and cons that come in with content based filtering. These systems have ability to

International Journal of Science, Engineering and Technology

An Open Access Journal

recommend to the user even if they have not rated anything. Also, with changing user preferences, it has ability to adapt in short period of time.

So basically, users still get recommendations without sharing their profiles, hence the system ensure privacy. But, content based filtering does depend on how descriptive data is, hence this factor directly affects the accuracy of the system. Content specialization tends to be another con of Content based filtering.

III. REQUIRED TOOLS

- **Language:** Python Libraries: Pandas, SciPy
- **Editor:** Google Colab
- **Dataset:** Movielens by Grouplens

1. About Dataset:

For implementing the Movie recommendation system, we have used the smaller subset of famous Movielens dataset by Grouplens.

The dataset comprises of 100,000 ratings and 3,600 tag applications applied to 9,000 movies by 600 users.

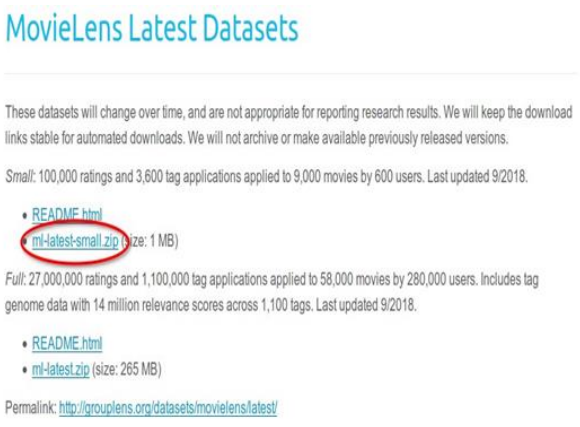

Fig 4. Movielens Dataset

IV. IMPLEMENTATION FLOWCHART

Recommendation systems are great utility for providing personalized services and content to the users. It also tries to reduce the problem of information overload. This paper includes a summary and literature studies related to Movie recommendation system based on collaborative filtering. Other approach of Content based filtering is also discussed.

This paper also focuses on strength and weak points of the discussed approaches. Performance of the Movie recommendation system can be increased pretty much by implementing both, content based and collaborative filtering.

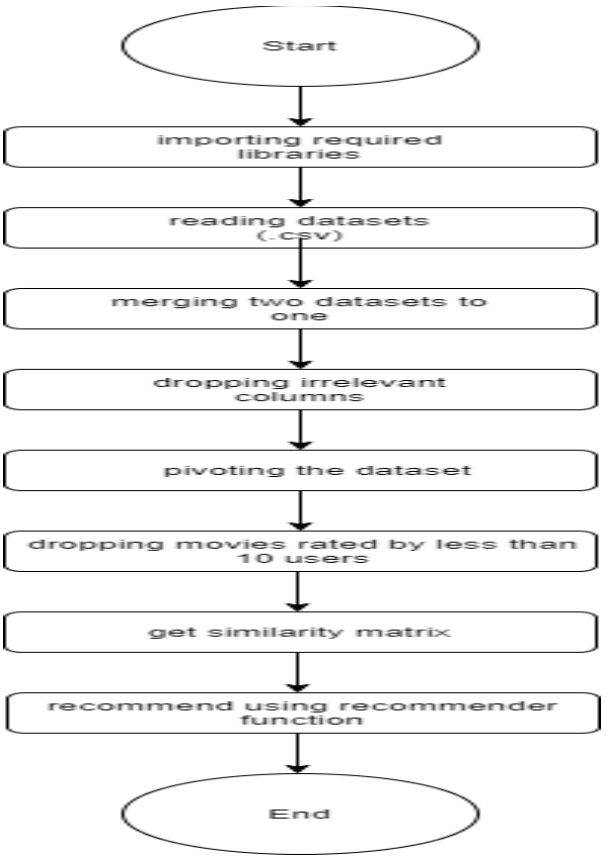

Fig 5. Basic flow of instructionsConclusion and Future scope.

V. ACKNOWLEDGEMENT

I cannot express enough thanks to our coordinator/ mentor Mr.Hradesh Kumar for his continuous support and guidance throughout the course of this research which exposed us to a pile of new knowledge.I would also like this take this opportunity to thank my team who helped a lot in finalizing this project within limited time frame.

REFERENCES

[1] J.A. Konstan, J. Riedl Recommender systems: from algorithms to user experience User Model User-Adapt Interact, 22 (2012), pp. 101-123.

[2] Miryala, Goutham & Gomes, Rahul & Dayananda, Karanam. (2017). Comparative

International Journal of Science, Engineering and Technology

An Open Access Journal

Analysis of Movie Recommendation System Using Collaborative Filtering in Spark Engine. Journal of Global Research in Computer Science. 8.10-14.

[3] Banerjee, Anurag & Basu, Tanmay. (2018). Yet Another Weighting Scheme for Collaborative Filtering Towards Effective Movie Recommendation.

[4] Zhao, Zhi-Dan & Shang, Ming Sheng.(2010). User Based Collaborative-Filtering Recommendation Algorithms on Hadoop. 3rd International Conference on Knowledge Discovery and Data Mining, WKDD 2010. 478-481.

[5] Subramaniyaswamy, V., Logesh, R., Chandra Shekhar, M., Challa, A. and Vijayakumar, V. (2017) 'A personalized movie recommendation system based on collaborative filtering,' Int. J. High Performance Computing and Networking, Vol. 10, Nos. 1/2, pp.54–63.

[6] https://www.kdnuggets.com/2019/09/mach ine-learning-recommender-systems. html#:~:text= Recommender%20systems%20are%20an%20im portant, to%20follow%20from%20example%20 code

[7] https://www.sciencedirect.com/science/arti cle/pi i/S1110866515000341.

[8] https://builtin.com/data- science/collaborative-fil tering- recommender-system.

[9] https://developers.google.com/machine-learni ng/ recommendation/collaborative/basics

[10] https://analyticsindiamag.com/how- to-build-a-content-based-movie- recommendation-system-in-python/

[11] https://dl.acm.org/doi/10.1145/371920.372 071

[12] https://medium.com/@cfpinela/recommender-systems-user-based-and-item-based-collaborati v e-filtering-5d5f375a127f