



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 12, December 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Recommendation System for Movies and Books

Kulkarni A.A.¹, Dudhe V.A.², Chinchkar O.P.³, Lokhande S.A.⁴, Dr. Nikita Kulkarni⁵

Graduation Students, Department of Computer Engineering, KJEE'S Trinity Academy of Engineering, Pune, Maharashtra, India^{1,2,3,4}

Associate Professor, Department of Computer Engineering, KJEE'S Trinity Academy of Engineering, Pune, Maharashtra, India⁵

ABSTRACT: In this paper we are making a recommendation system which recommends movies and books to the users as per their preferences. Now while creating a recommendation system; we should take note of the fact that how unique are the choices and the preferences of each individual user in terms of movie or book preferences. When we take all that into an account, it's very difficult to generalize a movie or a book and say that everyone would like it. But in spite of all that, still it is seen that similar movies or books are liked by a specific part of the society, so the system will make the group of users with similar taste. For this, various methods such as content-based filtering and collaborative filtering are used. Each method has several advantages over another method so we are using both of these techniques to make recommendations better.

KEYWORDS: Recommendation system, Book recommendation system, content-based filtering, collaborative filtering, Hybrid method.

I. INTRODUCTION

As we all know in today's world, the usage of the internet has increased which directly leads to an increase in data consumption. So finding the exact content that the user wants is getting harder. This is why personalized recommendation systems have become a necessity in many areas such as entertainment, e-commerce, social media, etc. Recommendation Systems are designed to make the user experience much better. This system will help users to save their time and increase the probability of user's satisfaction as a result; users will purchase more content, so the service provider will also earn more. But on the other hand, it may look relatively simple but behind the scenes, a complex statistical algorithm gets executed in order to predict these recommendations.

Such systems are called Recommender Systems, Recommendation Systems, or Recommendation Engines. One of the most famous applications of data science and machine learning is the recommendation system.

As a result, we're developing a movie recommendation system that will suggest movies or books to consumers based on their tastes. Now, when constructing a recommendation system, we need to consider how distinctive each person's movie interests are. Talking about the movies some people prefer genre films, such as thrillers, romances, or science fiction, while others are more interested in the leading stars, directors, and production companies. This system also recommends books based on the user's movie taste.[9] With all of this in mind, it's tough to generalize a film and declare that everyone will enjoy it. Even though it is clear that identical films are enjoyed by a specific segment of the population, the system will create a group of users with similar tastes [2]. For example, if one user from one group likes a horror movie there is a large possibility that other members from the same group will like the same horror movie as well. For this, various methods such as content-based filtering, collaborative filtering are used. Each method has several advantages over another method so we will use both of these techniques to make recommendations better.

Content-Based recommender system attempts to guess the features or behavior of the item's features given by a user, if a user reacts positively to author used features such as movie plots, ratings, countries of production and release year to find similarity between movies [8]. Collaborative Filtering is used to predict the rating of items. For every user Collaborative filtering method uses a user-item-rate matrix. The user-item-rate matrix indicates a specific user and each column indicates a specific item. Also, ratings are the elements of a matrix. Collaborative filtering methods are divided



into two main groups: model and memory-based methods [4][5]. We are going to use Kaggle MovieLens movies Dataset for movies and Kaggle Goodreads Books Dataset for books in this system

II. RELATED WORK

In this recommendation system, data will be filtered according to the user's choice. To recommend movies, this model is used, basically research on recommender systems started back in the 90's and it increased eventually. In these researches, so many different powerful methods have been designed and implemented. These methods can be divided into three groups: content-based filtering, collaborative filtering, and hybrid methods. In addition to the methods mentioned above, in recent years, deep learning and reinforcement learning methods have also been used [2]. Content-based filtering methods use the features of items and the user's history of actions to learn the preferences of the user. So in this method, we train the model by defining features of each item. Content based systems work on the concept of similar content that the user has liked in the past. The basic idea behind such a system is to recommend items or similar products which are already liked by users in the past. based on similar features, similarity between two or more items can be calculated. To understand better, continuing with the same example discussed above, whenever you watch any video on Instagram like the ones posted on Animal lover's page, when you are done watching, you get similar videos/reels on your home page. Also, when you like some pages, you get suggestions of pages similar to them. So what actually is behind all this is Content based filtering i.e. if you like any one item from some particular category, it is likely that you might like any other item similar from the same category or some category similar to it [8].

Collaborative Filtering is used to predict the rating of items. For every user Collaborative filtering methods use a user-item-rate matrix. Each column of the user-item-rate matrix indicates a specific item and each row is a specific user. Also, the elements of the matrix are the ratings. Collaborative filtering methods are divided into two main groups: model and memory-based methods. In the model-based methods, the model will learn and recommend. The matrix factorization is the most popular method to reduce the number of unknown parameters, it factorizes the user-item-rate matrix into two low-rank matrices. Also, to find similar users/items memory-based methods use the KNN algorithms.[4][5] KNN k nearest neighbour (KNN) is a popular classification method in data mining and statistics because of its simple implementation and significant classification performance [6]. The drawbacks of collaborative filtering methods are Cold start problem and data sparsity. Cold start problem is a situation where it is very difficult to make any correct recommendation if the user is new to the system or a given item or product is added to the system for the first time. So there are two types of cold-start problems: new user cold start problems and new item cold start problems. The other problem is data sparsity problem, it implies the irregular, insufficient or highly varying user ratings.

One of the major problems encountered by recommender systems is the cold start problem. The major reason behind sparsity is that most of the users do not provide ratings and the ones available are usually too scattered or sparse; for example rating of any mobile app. Mostly we ignore when asked to rate the app and even if we do rate, then some of us either rate it positively or negatively, as a result the ratings available are quite scattered. This is a major issue for Collaborative filtering where the recommendations are based on the user ratings[3]. It is the combination of two methods i.e. content-based and collaborative filtering. The idea is to combine the features of two methods in such a way that it gives the best of both worlds. The rate estimation of an item and a user can be achieved in this method by the rating pattern of the similar users to the same item. For our recommendation system we are gonna be using this hybrid method [1]. The most important aspect of the recommendation system is its accuracy. It can be evaluated through two main measures: Mean Absolute Error and Root Mean Squared Error is a measure of how spread out these residuals are. It tells you; how concentrated the data is around the line of best fit. Mean Absolute Error is an average of the absolute difference between the predicted rating and the actual rating. They are negatively-oriented scores(0-∞), Lower values better the prediction [7].

III. METHODOLOGY

The most important aspect of the recommendation system is its accuracy. It can be evaluated through two main measures: Mean Absolute Error and Root Mean Squared Error is a measure of how spread out these residuals are. It tells you; how concentrated the data is around the line of best fit. They are negatively-oriented scores(0-∞), Lower values better the prediction.

Mean Absolute Error (MAE): Mean Absolute Error is an average of the absolute difference between the predicted rating and the actual rating



$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

Root Mean Square Error(RMSE): Root Mean Square Error(RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}}$$

IV. SYSTEM ARCHITECTURE

At the start, the new user gets registered and logs into the system. The user gets recommendations according to the latest, trending or popular movies and books. Whichever movies/books user watches, gets saved in the database and after that, the system recommends similar kinds of movies and books based on the user's watch history. For these recommendations to work properly and effectively; Content-Based Filtering and Collaborative Filtering plays an important role.

With the help of Content-Based Filtering, the system considers two major factors. One is, it considers similar movies and books based on the user's watch history and it also takes explicit data from the user for a better recommendation.

A collaborative filtering system makes a group of people with similar taste when one user likes movies/books, that movie/book gets recommended to every other user who belongs to the same group.

Both these techniques i.e. Collaborative Filtering and Content-Based Filtering creates a hybrid recommender system that leverages both and continues to recommend movies and books which the user will like. New users get popular and latest movies as well as Books as a recommendation.

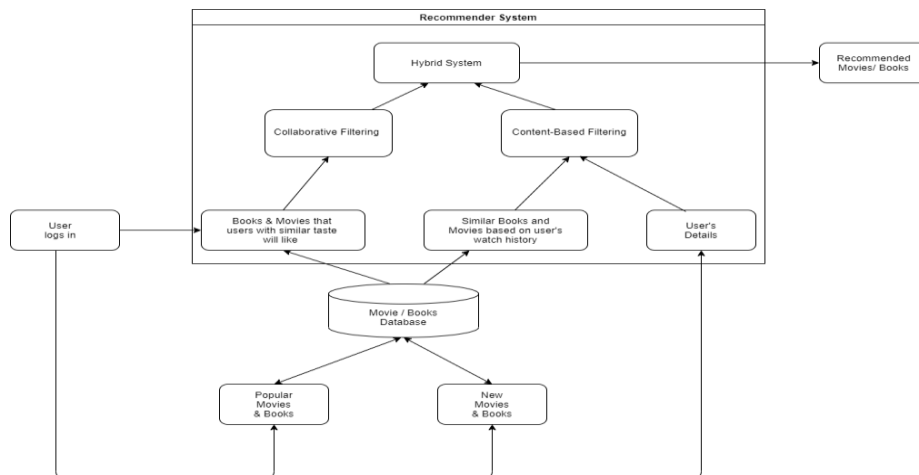


Fig.1 System architecture for movies and book recommendation system.

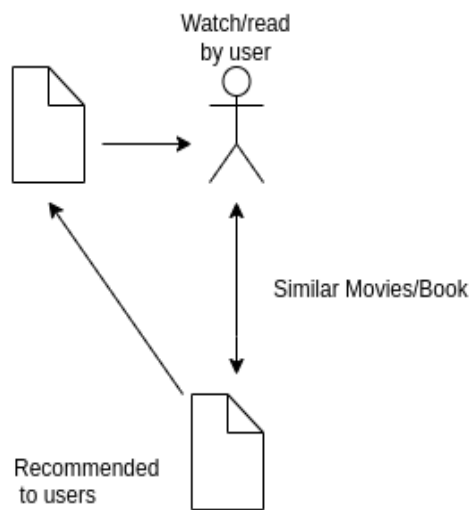


Fig.2 Content-based filtering

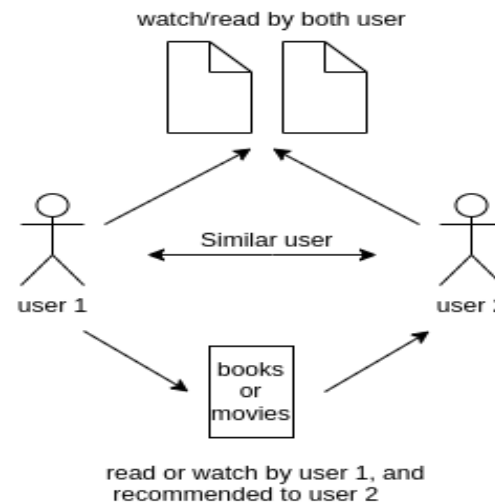


Fig.3 Collaborative filtering

V. CONCLUSION

The content based filtering and collaborative filtering are combined together to minimize the errors and improve the performance as a hybrid approach. Hybrid systems are found to be more effective and provide more accurate recommendations. Our system considers the user ratings to recommend movies and books. Features such as the genre of the movie/book, the directors, the actors, author and so on are considered as well to provide suggestions. This approach gives lower RMSE & MAE values than other existing systems. Other more effective techniques to compare accuracy of a recommendation system like Hit rate, Average reciprocal hit rate (ARHR), cumulative hit rate gives promising results. Other important aspects like coverage, diversity, and the novelty. of the unique list of movies and books generated for every individual user recommendation gets improved.

REFERENCES

- [1] Nguyen N.T., Rakowski M., Rusin M., Sobocki J., Jain L.C. (2007) Hybrid Filtering Methods Applied in Web-Based Movie Recommendation System. In: Apolloni B., Howlett R.J., Jain L. (eds) Knowledge-Based Intelligent Information and Engineering Systems. KES 2007. Lecture Notes in Computer Science, vol 4692. Springer, Berlin, Heidelberg.
- [2] Rahul Kataria, Om Prakash Verma. An effective collaborative movie recommender system with cuckoo Issue 2, July 2017, Pages 105-112.
- [3] Mishra N., Chaturvedi S., Mishra V., Srivastava R., Bargah P. (2017) Solving Sparsity Problem in Rating Based Movie Recommendation System. In: Behera H., Mohapatra D. (eds) Computational Intelligence in Data.
- [4] Ching-Seh (Mike) Wu, Deepti Garg, Unnathi Bhandary, Movie Recommendation System Using Collaborative Filtering. Dept. of Computer Science San Jose State University San Jose, CA, USA
- [5] M. K. Kharita, A. Kumar and P. Singh, "Item-Based Collaborative Filtering in Movie Recommendation in Real time," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018, pp. 340-342, doi: 10.1109/ICSCCC.2018.8703362
- [6] S. Zhang, X. Li, M. Zong, X. Zhu and R. Wang, "Efficient KNN Classification With Different Numbers of Nearest Neighbors," in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 5, pp. 1774-1785, May 2018. doi: 10.1109/TNNLS.2017.2673241.
- [7] Chai, Tianfeng & Draxler, R.R.. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development. 7. 1247-1250. 10.5194/gmd-7-1247-2014.
- [8] R. Singla, S. Gupta, A. Gupta and D. K. Vishwakarma, "FLEX: A Content Based Movie Recommender," 2020 International Conference for Emerging Technol (INCET), 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154163
- [9] M. Kommineni, P. Alekhya, T. M. Vyshnavi, V. Aparna, K. Swetha and V. Mounika, "Machine Learning based Efficient Recommendation System for Book Selection using User based Collaborative Filtering Algorithm," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), 2020, pp. 66-71, doi: 10.1109/ICISC47916.2020.9171222.



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details