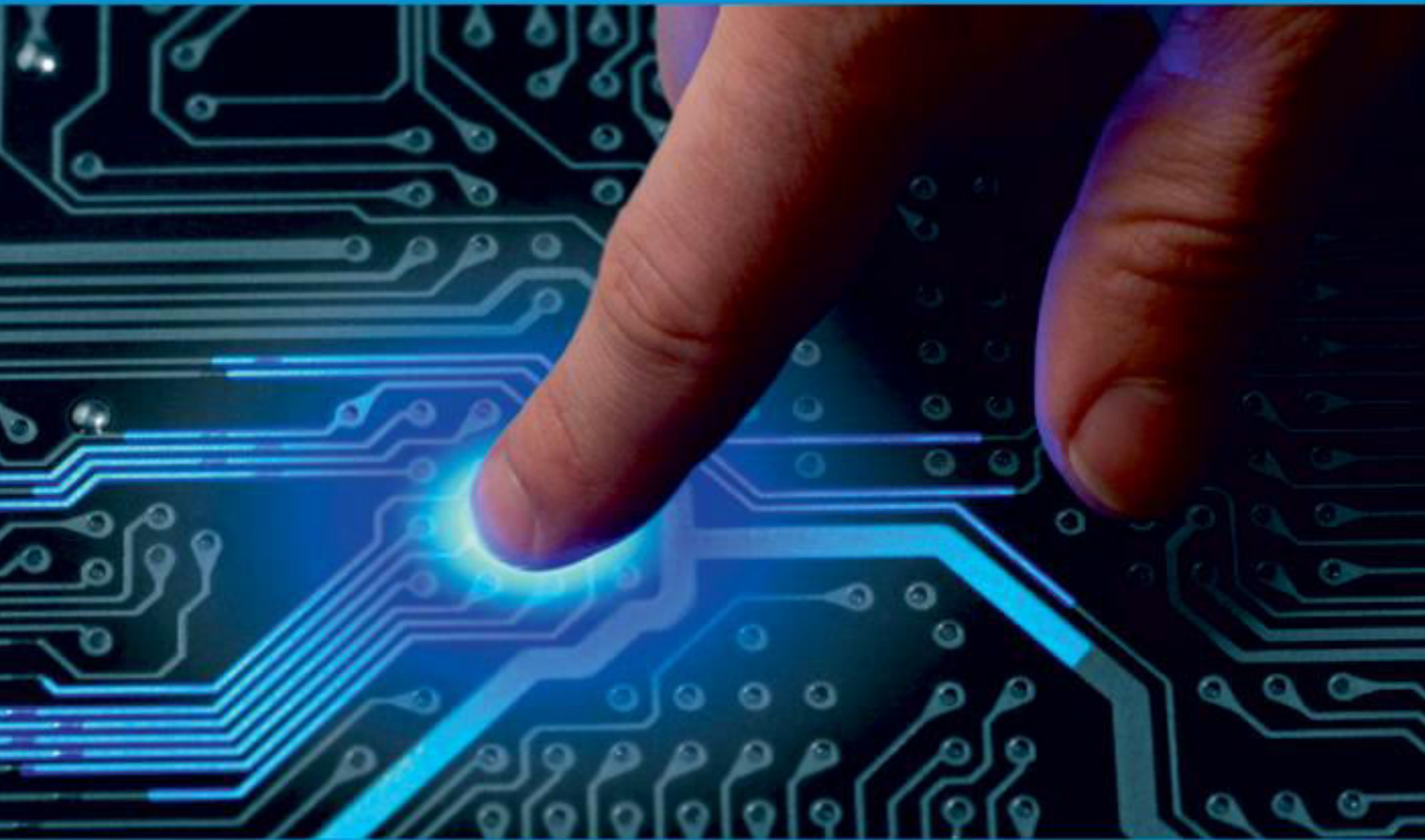
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 11, November 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011009 |

Movie Recommendation Engine using Collaborative Filtering

Kishor Sonawane, Vedantika Wadkar, Khushi Tiwari, Rohit Prasad, Prof. Varsha Gosavi

UG Student, Dept. of C.S., SKNCOE, Savitribai Phule Pune University, Pune, India

UG Student, Dept. of C.S., SKNCOE, Savitribai Phule Pune University, Pune, India

UG Student, Dept. of C.S., SKNCOE, Savitribai Phule Pune University, Pune, India

UG Student, Dept. of C.S., SKNCOE, Savitribai Phule Pune University, Pune, India

Professor, Dept. of C.S., SKNCOE, Savitribai Phule Pune University, Pune, India

ABSTRACT: A recommender system is a software system that aims to make a prediction that quantifies how strong a user's interest in an object is, to recommend the user exactly those objects from the set of all existing objects, which he is probably most interested in. Typical objects of a recommendation service are, for example, products from a web shop, pieces of music or artists or films. A recommendation service is intended to help deal with the information overload by recommending a subset to the user from a confusing set of objects. To determine the appropriate recommendations, a recommendation service uses methods of machine learning and information retrieval.

Collaborative recommendation services (also known as collaborative filtering) recommend the objects that users with similar rating behavior (similar users) are most interested in. No further knowledge about the object itself is required for this. The recommendation service includes the user-related and element-based algorithm, as well as the model and memory-based method

This report discusses a wide variety of the choices available and their implications, aiming to provide both practitioners and researchers with an introduction to the important issues underlying recommenders and current best practices for addressing these issues.

KEYWORDS: Recommendation Engine; Movie Recommendation Engine; Collaborative Filtering; Hybrid Filtering; Content-based filtering

I. INTRODUCTION

Collaborative Filtering is the most popular recommender systems approach. It recommends items based on the user's past behavior as well as similar decisions made by other users. There are three major filtering algorithms to make recommendations.

Content-based filtering is to learn what kinds of contents a user likes and then match the contents of a current article with a "content prototype" that we believe describes well what the user likes. Collaborative filtering (user-based filtering) assumes that if users who are like the current user like some items, the current user might also like it. A hybrid recommender combines the two, probably also involving knowledge-based and demographic techniques.

Collaborative filtering is mostly used for very large amounts of data. Collaborative filtering is used for a wide variety of areas such as B. in the financial services sector for the integration of financial sources or in applications in eCommerce and Web 2.0. This article deals with collaborative filtering for user data, although some methods and approaches can be transferred to other areas.

The aim of the method is an automatic prediction (filtering) of user interests. For this purpose, information about the behavior and preferences of as many users as possible is collected. The underlying assumption of collaborative filtering is that if two people have the same preferences about similar products, they agree about other products as well. Hence the term collaboration: If you want to know the opinion of user A on an article, you look at the opinion of other users on this article. Whereby one only considers users whose opinion on as many articles as possible agrees with the opinion of user A. The other users work together to solve the question of what the opinion of user A is.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011009 |

The goal of our recommendation system is to better assist users in finding the movie that they are looking for without having to enter a very concise query to the information retrieval system. For this project, we will be reading in the Database of the movies containing the user ratings for various movies. The dataset is sourced from Movie Lens.

II. RELATED WORK

Mukesh Kumar Kharita [1] presented a model where a prediction of unknown movie ratings can be done using a similarity model which can in turn help in the recommendation of the movies to the given user according to his past ratings and likings. This makes it easier to work with the data where the underlying meta-data of the data is unknown.

Dharmendra Pathak [2] proposed a hybrid model in which it brought best of both worlds. That is, it consisted of the collaborative filtering technique as well as the content based filtering too. It helped to overcome the challenges of both the algorithms which made the hybrid model very much effective. This came with one drawback though. It was that the model needed to be provided with some semi-labelled data to start with. Contrary to this, Ching-Seh (Mike) Wu [3] displayed that a recommendation model can be trained without the use of any co-relation equation and using the similarity matrix. This is very helpful for the low-end systems. As the matrix can be quickly traversed in constant time, it is faster than most of the other models.

Bagher Rahimpour Cami [4] et al: gathered the movies information from IMDB2 and employ them into a Bayesian nonparametric framework that presented in to capture the temporal preferences of a user. This had some performance benefits which stood apart from the others.

Noor Ifada [5] et al: analyzes and compares the performance of the Collaborative Filtering and Hybrid based approaches in generating movie recommendations. The paper presents two variations of movie recommendation methods built by implementing the Collaborative Filtering and Hybrid approaches. The former's development requires the learning process to get the movie's rating pattern given by users. Jochen Nessel [6] outlines an approach to recommend movies according to the genres, runtime, artists involved and other factors that need to be hard coded for each movie. Such filtering is called as content based filtering and is pretty accurate. But as discussed earlier, the method needs labelled data which is not very efficient to start with. Yajie Hu, Ziqi Wang [7] et al: aims at finding the stars or films which are close to the queried stars or films and recommending them to users. We define “close” as similar performance style, genre and film styles. When the user submits a query, the system recommends the relevant movies and stars similar to the query in genre and style.

Sajal Halder [8] suggests a system that can suggest a set of movies to users based on their interest, or the popularities of the movies. Although, a set of movie recommendation systems have been proposed, most of these either cannot recommend a movie to the existing users efficiently or to a new user by any means.

III. APPROACH

Matrix factorization is used to resolve the sparsity problem in collaborative filtering. It is one of the models of matrix factorization that is considered as the values from the user’s item list. Selecting several parameters to perform analytics on the ALS algorithm helps in building an efficient final recommendation system. This method computes the statistical correlation (also known as Pearson’s Coefficient) between the common ratings of two users to determine the similarity.Firstly, we can calculate the mean rate of a user using the formula given below.

$$w(a,u) = \frac{\sum v_{ai}v_{ui} - \frac{\sum v_{ai} \sum v_{ui}}{n}}{\sqrt{\sum v_{ai}^2 - \frac{\sum v_{ai}^2}{n}}\sqrt{\sum v_{ui}^2 - \frac{\sum v_{ui}^2}{n}}}$$

To explain how these methods works we are going to use the following notations. Let U be a set of N users and I a set of M items. Vui denotes the rating of user u U on item i I, and S I stands for the set of items that user u has rated The approximation is better as the value can be found in one pass thus reducing the computational time.

Collaborative filtering usually takes place in two steps.

1. Search for users who have the same behavioral pattern as the active user. (= the user for whom the prediction is made)
2. Use of behavior patterns to make a prediction for the active user.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011009 |

Alternatively, there is the article-based collaborative filtering, which became known through Amazon.com ("You might also be interested in"). and was first presented by Vucetic and Obradovic in 2000.

1. Create a similarity matrix to determine relationships between articles.
2. The preferences of the active user are derived from the matrix.

Other forms of collaborative filtering can be based on implicit observation of user behavior. These forms of filtering compare the behaviour of each user with the behaviour of all other users (What music did they listen to? What products did they buy?). This data is used to be able to predict the future behavior of the user. It does not make sense to offer a user a specific piece of music if he has made it clear through his behavior that he already owns it. Likewise, it does not make sense to offer a user other Paris travel guides if he already owns a travel guide for this city.

In today's information age, these and similar technologies turn out to be extremely helpful for product selection, especially when certain product groups (e.g. B. Music, movies, books, news, websites) have become so big that individuals cannot see the entire offer

IV. Conclusion and Future Work

It is undisputed that personalized content will play an increasingly important role in the future. Providers and customers will "work out" the best form together. This means that customers will show to what extent personalization is desired and which areas it should affect. This also determines which methods are the most effective and which methods remain under discussion with certain content and are therefore oversized.

Recommender systems have become ubiquitous. People use them to find books, music, news, smart phones, vacation trips, and romantic partners. Nearly every product, service, or type of information has recommenders to help people select from among the myriad alternatives the few they would most appreciate.

Sustaining these commercial applications is a vibrant research community, with creative interaction ideas, powerful new algorithms, and careful experiments. Still, there are many challenges for the field, especially at the interaction between research and commercial practice

Collaborative filtering evaluates behavior patterns of user groups in order to deduce the interests of individuals. This is a form of data mining that eliminates the need for explicit user input.The aim of the method is an automatic prediction (filtering) of user interests. For this purpose, information about the behaviour and preferences of as many users as possible is collected. The underlying assumption of collaborative filtering is that if two people have the same preferences for similar products, they also agree on other products. Hence the term collaboration: If you want to know what opinion a user A has about an article, look at what opinion other users have about this article. Whereby only users are considered to agree with the opinion of user A for as many articles as possible. The other users work together to solve the question of what opinion user A is.

A specific problem with collaborative filters is their latency: A new user enters the system with an empty user profile. Since his interests are not yet known, he cannot receive meaningful recommendations at the beginning. The same applies to new elements entering the system (e.g. B. products in an online shop). They have no quantifiable similarity to other elements and cannot be reasonably recommended. Collaborative filters are therefore learning systems and thus a form of artificial intelligence. So this can be an area of future work to be considered.

REFERENCES

[1] 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.

[2] D. Pathak, S. Matharia and C. N. S. Murthy, "ORBIT: Hybrid movie recommendation engine," 2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN), 2013, pp. 19-24, doi: 10.1109/ICE- CCN.2013.6528589.

[3] C. -S. M. Wu, D. Garg and U. Bhandary, "Movie Recommendation System Using Collaborative Filtering," 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018, pp. 11-15, doi: 10.1109/ICSESS.2018.8663822.

[4] B. R. Cami and M. R. M. Brujerdi, "Using Hybrid Trust Model for Handling Inaccurate Resource," 2009 International Conference on Availability, Reliability and Security, 2009, pp. 630- 634, doi: 10.1109/ARES.2009.164.

[5] N. Ifada, T. F. Rahman and M. K. Sophan, "Comparing Collaborative Filtering and Hybrid based Approaches for Movie Recommendation," 2020 6th Information Technology International Seminar (ITIS), 2020, pp. 219-223, doi: 10.1109/ITIS50118.2020.9321014.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011009 |

[6] J. Nessel and B. Cimpa, "The MovieOracle - Content Based Movie Recommendations," 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, 2011, pp. 361-364, doi: 10.1109/WI-IAT.2011.236.
[7] Y. Hu, Z. Wang, W. Wu, J. Guo and M. Zhang, "Recommendation for Movies and Stars Using YAGO and IMDB," 2010 12th International Asia-Pacific Web Conference, 2010, pp. 123-129, doi: 10.1109/APWeb.2010.51.
[8] Halder, S., Sarkar, A. M. J., & Lee, Y.-K. (2012). Movie Recommendation System Based on Movie Swarm. 2012 Second International Conference on Cloud and Green Computing. doi:10.1109/cgc.2012.121

BIOGRAPHY

Kishor Sonawane is a B.E student in the Computer Engineering Department, at Shrimati Kashibai Navale College of Engineering, SPPU. He is in the final year of his Bachelor's in Computer Engineering and will pass out in 2023. His research interests are Machine Learning, Data Science, Data Structures and Algorithms, and Web development.

Vedantika Wadkar is a B.E student in the Computer Engineering Department, at Shrimati Kashibai Navale College of Engineering, SPPU. She is in the final year of his Bachelor's in Computer Engineering and will pass out in 2023. Her research interests are Design and analysis of Algorithms, Data Analytics, DBMS, Front End Web development.

Khushi Tiwari is a B.E student in the Computer Engineering Department, at Shrimati Kashibai Navale College of Engineering, SPPU. She is in the final year of his Bachelor's in Computer Engineering and will pass out in 2023. Her research interests are Web 3.0, Blockchain Technology, DBMS, Full stack Web development, etc.

Rohit Prasad is a B.E student in the Computer Engineering Department, at Shrimati Kashibai Navale College of Engineering, SPPU. He is in the final year of his Bachelor's in Computer Engineering and will pass out in 2023. His research interests are Machine Learning, Natural Language Processing, Artificial Intelligence, Algorithms, Data science, etc.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

doi
cross ref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING