



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 11, Issue 5, May 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Review Based Recommendation System Using Data Mining

Shubhangi Ragit, Shilpa Kar, Pragati Dahake, Rutuja Raut, Prof. Deepak Dhanwani

Department of Computer Science and Engineering, P. R. Pote Patil College of Engineering and Management,
Amravati, Maharashtra, India

Department of Computer Science and Engineering, P. R. Pote Patil College of Engineering and Management,
Amravati, Maharashtra, India

Department of Computer Science and Engineering, P. R. Pote Patil College of Engineering and Management,
Amravati, Maharashtra, India

Department of Computer Science and Engineering, P. R. Pote Patil College of Engineering and Management,
Amravati, Maharashtra, India

Department of Computer Science and Engineering, P. R. Pote Patil College of Engineering and Management,
Amravati, Maharashtra, India

ABSTRACT: Recommender system has been proven to be significantly crucial in many fields and is widely used by various domains. Most of the conventional recommender systems rely on the numeric rating given by a user to reflect his opinion about a consumed item; however, these ratings are not available in many domains. As a result, a new source of information represented by the user-generated reviews is incorporated in the recommendation process to compensate for the lack of these ratings. The reviews contain prosperous and numerous information related to the whole item or a specific feature that can be extracted using the sentiment analysis field. This paper gives a comprehensive overview to help researchers who aim to work with recommender system and sentiment analysis. It includes a background of the recommender system concept, including phases, approaches, and performance metrics used in recommender systems. Then, it discusses the sentiment analysis concept and highlights the main points in the sentiment analysis, including level, approaches, and focuses on aspect-based sentiment analysis.

KEYWORDS :recommender system, ratings, user reviews, sentiment analysis, aspects.

I. INTRODUCTION

Recently, there has been a vast flow of information on the Web, and it continues to grow exponentially while providing users/customers with various resources about services such as products, hotels, and restaurants. Despite such data's benefits, the vast flow of information causes challenges for users to deal with and choose from a vast number of available options. This causes an information overload problem (Liu et al. 2011) and complicates the decision-making process. In this case, it is essential to filter the information to a limited amount based on the current user/customer preferences to assist them in making the correct decision (Hdioud et al. 2013). Such a filtering process is typically done by recommender systems (RSs), which are developed to solve the information overload problem by providing personalized suggestions of services (i.e. items) to specific customers according to their preferences (Adomavicius&Tuzhilin 2005; Ebadi&Krzyzak 2016). Since RS emerged more than a few decades ago, the field has been of considerable importance in academia, business, and industry. They are widely used in various domains such as shopping (Amazon), music (Pandora), movies (Netflix), travel (TripAdvisor), restaurant (Yelp), people (Facebook), and articles (TED). With the recent advancement of e-commerce websites, it has been shown that RSs have a significant benefit in assisting users/customers to discover the relevant items that suit their needs and, most likely, to their preferences. For example, according to Amazon statistics in 2015, 35% of sales/revenues come from recommended items to users. There are three main recommendation approaches which are content-based, collaborative-based, and hybrid. The content-based (CB) approach mines the appropriate recommendations for a user



based on his recent behaviours according to what the user liked, bought, or watched (Pazzani&Billsus 2007). While the collaborative filtering approach (CF) generates the recommendation for a user based on the similarities among users who have similar preferences/interests to him in the past (Aciar et al. 2007). The last approach is the hybrid which integrates two or more recommendation components or algorithms implementations into a single recommendation system (Danilova&Ponomarev 2016). However, these classical RS approaches rely on a single-criterion rating (overall rating) as a primary source for the recommendation process. A single-criterion rating for a recommendation is insufficient to give an accurate recommendation because the overall ratings cannot express fine-grained analysis behind the users' behaviours. This made the RS a broad research topic that encouraged more research works to find practical solutions for improving the RS's performance.

II. RELATED WORK

There are three techniques of recommendation system: Collaborative Filtering, Content-Based Filtering and Hybrid Filtering. In Content Based recommender system, user provides data either explicitly (rating) or implicitly (by clicking on a link). The system captures this data and generates user profile for every user. By making use of user profile, recommendation is generated. In content based filtering, recommendation is given by only watching single user's profile. System tries to recommend item similar to that item based on users past activity. Unlike content based, collaborative filtering finds those users whose likings are similar to a given user. It then recommends item or any product, by considering that the given user will also like the item which other users like because their taste are similar . Both these technique have their own strength and weakness so to overcome this, hybrid technique came into picture, which is a combination of both these techniques. Hybrid filtering can be used in various types. We can use content based filtering first and then pass those results to collaborative recommender (and vice-versa) or by integrating both the filter into one model to generate the result. These kinds of modifications are also uses to cope up with cold start, data sparsity and scalability problem.

Content-Based Filtering : Content-Based Filtering are also known as cognitive filtering [1]. This filtering recommends item to the user based on his past experience. For example, if a user likes only action movies then the system predicts him only action movies similar to it which he has highly rated. The broader explanation could be suppose the user likes only politics related content so the system suggests the websites, blogs or the news similar to that content. Unlike collaborative filtering, content based filtering do not face new user problem. It does not have other user interaction in it. It only deals with particular user's interest. Content based filtering first checks the user preference and then suggest him with the movies or any other product to him. It only focus on single user's ideas, thoughts and give prediction based on his interest. So if we talk about movies, then the content based filtering technique checks the rating given by the user. The approach checks which movies are given high ratings by the user by checking the genre categories in the user profile. After analysing user profile, the technique recommends movies to user according to his taste. The figure 2 shows us how Content-Based Filtering works.

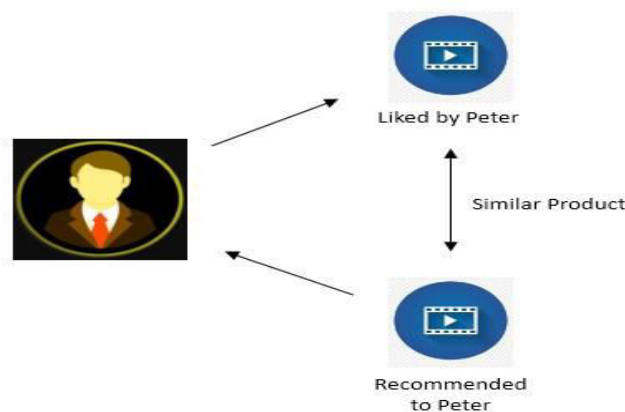


Figure2.1: Content-based filtering

Collaborative filtering: The concept of collaborative filtering was first introduced in 1991 by Goldberg et al. [5]. The Tapestry system applies only to smaller user groups (e.g. a single unit), and has too many demands on the user. As a proto-type of collaborative filtering recommendation system, Tapestry presents a new recommendation, but there are many technical deficiencies. Since then, there has been a scoring based collaborative filtering recommendation system, such as Grouplens, which recommends news and films. At present many ecommerce sites have been using the recommendation system such as Amazon, CDNow, Drugstor and Moviefinder etc. There is massive amount of data available. As we all know that today in this busy life no one has time to search hundreds of thousands of item and select the one which is similar to their taste. So collaborative filtering is one of the ways to filter the data and provide the relevant information in which the user is interested in. Collaborative Filtering is one of the most well known techniques for recommending items. This technique suggests relevant item to the user based on neighbour’s choice. It first finds out the similarity between the user and his neighbour and then predicts the items. There can be n number of users. This technique finds the similar user from the list of user’s. But the similarity between users is found out based the ratings which the users have given to the particular item. This way the approach continues and the desired result is generated. This strategy takes ratings given by user for any item from the large catalog of item catalog of ratings given by the user.

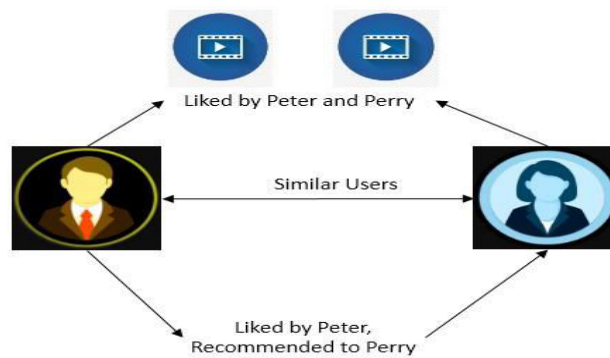


Figure 2.2 :Collaborative filtering

Memory Based Method:

Different colours are used to distinguish movies. By conducting the experiments, it is seen that for highly recommended movies, node adds yellow edge and the thickness of the edge represent film recommendation. It is observed that as the radius of the nodes increases and as the edges become thick, the score of the movies is increased. For implementation, py1neo is used which is a working library with Neo4j and movielens100k dataset is used.

This approach is also known as item -to- item filtering. It is used to recommend any item based on related item’s rating. By analyzing the rating , the users whose rating are similar for different item, only those items are available for recommendation [6]. It has been widely used by all Web giants including Netflix, YouTube etc.

Model Based Method:

In Model-based method it develops a user model using ratings of each user to evaluate the expected value of unrated items. This method generally uses machine learning or data mining algorithm to create a model. The model is developed using the utility matrix which is build using rating given by user for any item. The model is trained by getting the information from the utility matrix. Now this model is trained using the given data to generate prediction for the users .The model based approach is further classified into various categories. They are as Association rule mining, Decision Tree, Clustering, Artificial Neural Network, Regression etc. There are various examples working on model based approach. Some of them are Latent Semantic methods like Latent Semantic Analysis and Latent



Semantic Indexing and Dimensionality Reduction techniques like Singular Value Decomposition (SVD), Matrix Factorization etc. Model based techniques are use to solve sparsity problem occurring in recommendation system.

Hybrid Filtering:

This filtering is an information filtering system that takes ratings of the movies as input from the users and then apply the collaborative and content based filtering and generate recommendation list [13]. It is a combination of the two technique i.e. collaborative filtering and content based filtering. When only the single method i.e. the collaborative filtering or content based filtering alone cannot solve the problem then hybrid filtering concept comes into picture. By using hybrid filtering many problems of collaborative filtering and content based filtering can be resolved. The problem like cold start problem in collaborative filtering is a major challenge in it. So if we apply content based filtering and then use collaborative filtering can be a solution to it. So making it hybrid can resolve the problem.

S.M.Ali et al. proposed hybrid model ,which uses genomic tags of movies with content based filtering concept [14]. It uses Principal Component Analysis with Pearson Correlation technique to reduce the tags which are redundant, so that computation complexity can be reduced. Here Movie lens dataset is used which was released on October 18, 2016.

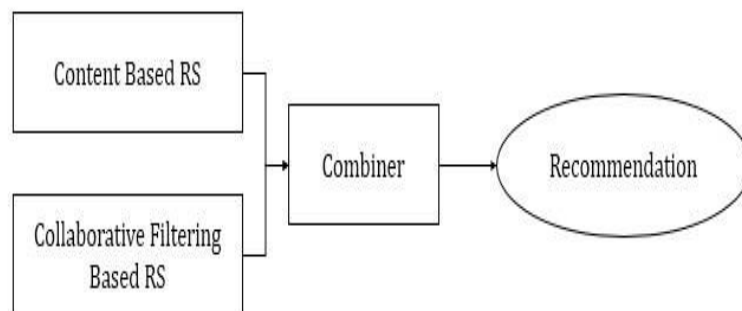


Figure2.3 :Hybrid filtering

III. PROPOSED METHODOLOGY

The proposed approach is the combination of the Content Recommendation System and the Memory-based and Model-based algorithms. The analysis of the parameters and their outcome is performed. In this approach, the user history which is stored in database is used to find the rating predictions. After using the Domain recommendation if further algorithms are used then the time required to find the rating predictions gets reduced. Thus the users will get predictions very quickly. These predictions include the user history and the opinions of like-minded users.

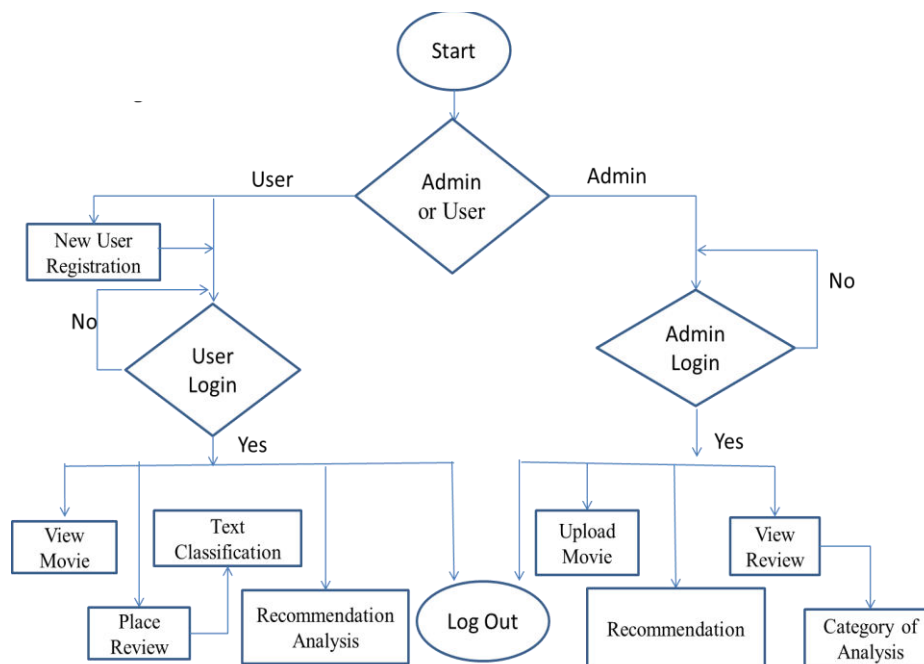


Figure 2.4: Flowchart of User or Admin Login

Recommender systems apply data mining methods and prediction algorithms to predict users' interest on information, products and services among the marvelous amount of available items. The central component of all recommender systems is the user prototypical that contains information about the individual preferences which control his or her behavior in a complex environment of web-based systems.

Apriority Algorithm:The recommendation system using apriori algorithms is a classic algorithm for learning association rules. Apriori is designed to function on databases containing transactions. Additional algorithms are calculated for finding association rules in data having no relations. It is collective in association rule mining, given a set of element sets, the algorithm attempts to find subsets which are common to at least a lowest number C of the item groups. Apriority is a bottom up methodology, wherecollective subsets are extended one item at a time and sets of candidates are verified against the data. The algorithm trimmings when no further successful extensions are found

Collaborative Filtering:The aim of a collaborative filtering algorithm is to propose new items or to calculate the utility of a certain item for a particular user based on the users prevision likings and the opinions of other like-minded users.

Item-Based Collaborative Filtering: A predictable a different method in the area of filtering algorithms, that was proposed newly is based on item relations and not on user relations, as in typical Collaborative Filtering. In Item-based Collaborative Clarifying process, we look into the group of items, that the dynamic user, has rated, compute how like they are to the goal item and then choice the k most similar items $\{i_1, i_2, \dots, i_k\}$, based on their parallel similarities $\{s_{i1}, s_{i2}, \dots, s_{ik}\}$. The calculations can then be calculated by taking a weighted average of the dynamic users scores on these associated items. The first step in this new approach is the Representation. Its resolve is the related as with the classic Collaborative Filtering procedure: represent the data in an ordered manner.

Content-Based Collaborative Filtering: The basic idea behind Content-Boosted Collaborative Filtering is to use a content-based predictor to enhance current user data, communicated via the user-item matrix, R, and then deliver personalized suggestions through collaborative filtering. The content-based analyst is practical on each row from the



first user-item matrix, corresponding to every separate user, and gradually makes a pseudo user-item matrix, PR. At the end, each row, i , of the pseudo user-item matrix PR consists of the scores providing by user u_i , when available, and those grades predicted by the content-based predictor.

IV. CONCLUSION AND FUTURE SCOPE

The proposed scheme consists of three algorithms that are Book Sensitive Recommendation algorithm (DsRec), Memory-based algorithm and Model-based algorithm. The proposed method provides the efficient approach to find the appropriate rating predictions. In other words provides the better recommendations to the users of same interest. This approach calculates the ratings of the items searched by the users. After that the top-N items are searched and recommended to the other users having similar area of interest.

The analysis of the products history is carried out to show which product is good and getting better response. The reviews on the product help in the analysis phase. The use of few recommendations optimize the less time required for calculation of ratings using all recommendations. It makes the analysis phase easy. It also takes less time to find quality recommendations.

In a future work, recommendation engines and other representations of the contextual variables will also be used. It would be interesting to study how item-based works with other interesting methods or techniques.

REFERENCES

- [1] Patel, A. Thakkar, N. Bhatt, and P. Prajapati, *Survey and Evolution Study Focusing Comparative Analysis and Future Research Direction in the Field of Recommendation System Specific to Collaborative Filtering Approach*, Springer, Singapore, pp.155-163, 2019.
- [2] J. Son and S. B. Kim, "Content-based filtering for recommendation systems using multiattribute networks", *Expert Systems with Applications*, pp.89:404-412, 2017.
- [3] S. Khatwani and D. M. B. Chandak, "Building Personalized and Non Personalized Recommendation Systems", *IEEE International Conference on Automatic Control and Dynamic Optimization Techniques*, pp.623-628, 2016.
- [4] M. Nilashi, O. Ibrahim, and K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques", *Expert Systems with Applications* pp.92:507-520, 2018.
- [5] T. Zhou, L. Chen, and J. Shen, "Movie Recommendation System Employing the User-Based CF in Cloud Computing", *IEEE International Conference on Computational Science and Engineering and IEEE International Conference on Embedded and Ubiquitous Computing*, pp. 2:46-50, 2017.
- [6] H. Koochi and K. Kiani, "User based Collaborative Filtering using fuzzy C-means", *Measurement: Journal of the International Measurement Confederation* pp.91:134-139, 2016.
- [7] N. Yi, C. Li, M. Shi, and X. Feng, "Design and Implementation of Movie Recommender System Based on Graph Database", *IEEE 14th Web Information Systems and Applications Conference*, pp.132-135, 2017.
- [8] D. Bokde, S. Girase, and D. Mukhopadhyay, "Matrix Factorization model in Collaborative Filtering algorithms: A survey", *Procedia Computer Science* 2015.
- [9] K. N. Jain, V. Kumar, P. Kumar, and T. Choudhury, *Movie recommendation system: Hybrid information Filtering System*, Springer, Singapore, pp.677-686, 2018.
- [10] S. M. Ali, G. K. Nayak, R. K. Lenka, and R. K. Barik, *Movie Recommendation System Using Genome Tags and Content-Based Filtering*, Springer, Singapore, 2018.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



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