

On the Part of the Survey-Nature of the Recommender Systems and their Challenges

K.Selvarani¹ and Dr.A.Annadhason²

¹Alagappa University/Computer Applications, Karaikudi, Sivaganga (DT), India
Email: selvarani1628@gmail.com

²Govt., Arts and Science College/Computer Science, Kadaladi, Ramnad (DT), India
Email:annadhasa@gmail.com

Abstract— Regarding to applied science, mining play a significant role in the computer science world. Data mining procedure is to be enforced upon the internet data said to be called Web Mining. Based on the Web mining techniques, we can extract the user needed pattern from the websites. In web mining, Recommender System having the significant portion in the approach of machine learning. The main aim of the Recommender systems is that to construct the connection within the items (products), End users and decide to choose the most suitable product to the particular user. The Main aim of the RS is to more helpful to the end-users by giving the surface of data that most similar to them, fulfil their needs and tasks. This paper implies to provide the overview of the Recommendation Systems (RS) and also make the survey about the recommender systems with their techniques and Challenges in Recommendation Systems (RS).

Index Terms— Recommender Systems (RS), Collaborative Filtering, Content-Based Filtering, Hybrid Filtering, Demographic Filtering, Machine learning, DB (Database), SVM (Support Vector Machine), ANN (Artificial Neural Network).

I. INTRODUCTION

Increasing the consumption of the Internet and also modern web services containing more quantity of facts now accessible to everyone. Among the more difficult challenging task for End-users to refine the appropriate information what they need to access. In Recent days, The Recommender Systems are playing the most significant area in the E-Commerce Platform. Also, The Recommender System is the most significant research field in today's research by reason of it will help to upgrade the information quality to the users above the internet world. The Recommendation System (RS) main target is that to forecast the product item that would be useful to the end-user based on given information. Recommender system used to forecast the rating along the user will give to a particular item or Product. Focused on the user's past rating history similar products will be recommended to the particular user [8].

Recommender systems having four different types namely (Collaborative filtering, Content-based filtering, Demographic filtering, Hybrid filtering). These will help to upgrade the standard of the recommendation engine. Suppose if we are purchasing the item or product from the particular website, they will recommend the more items to us assumed from the what you purchased [9]. For example, suppose if we are purchasing the book from the particular website, you will be recommended with the system the same book topics or more books published

near the same authors.

Nowadays, all over the world move to build the essential techniques and software using big data to handle frequently increasing the size of the data. In this paper, we are going to focus on discussing the recommender systems and their techniques with challenges based on the user.

II. RELATED WORK

Quin Wang and Xianhu Yuan, et al [1] proposed a demographic filtering Technique to increase the system of scalability and genetic Algorithms to upgrading the perfection of the recommender system. Also they suggest the hybrid approaches model through combining the item with demographic data and finding the set of neighbouring users who have similar interests and using a genetic algorithm to resolve the weight feature of the user model.

Xeulin Zeng et al [2] proposed a model named by Parallel Latent Group Model (PLGM).they used the spark to implement the ALS Matrix Factorization to be balanced by using MLib to generate SGD Matrix Factorization.

Sajal Haldaer et al [3] propose an algorithm for the movie swarm mining concept. They tried to solve the cold start problem to suggested items to the new user and how to manage and increase viewing times for the newest and most liked movies. This paper used two pruning rules and frequent item mining.

Rahul Katarya, Om Prakash verma [4] proposes a hybrid movie recommender system based on the collaborative filtering technique .they address the scalability and accuracy of the movie recommender system they used fuzzy and c-means techniques to find the neighbourhood items to the users.

Ghazanfar and Prugel- Bennett [5] implies to use and merging all recommender system techniques and propose a new one named cascading hybrid recommender system. This applied to eliminate the redundant records problems with the recommendation systems.

Kumar, Vikas, et al [6] propose a new model to make an accurate rating matrix completion that the ordinal matrix by making the two-class structure of binary matrix factorization.

Dianping, Lakshmi et al [7] proposed discovering the relation within the different items compared to a user-item rating matrix to see the similarities and making the new recommender item list to the user named by item-based collaborative filtering techniques.

Panigrahi, et al [8] used to reduce the user-oriented collaborative filtering technique challenges such as scalability and data sparsity through proposing a new hybrid algorithm using K-means and dimensionality elimination techniques such as ALS.

III. RECOMMENDER SYSTEM TECHNIQUES

One on one greatest web service recommendation is the processing and identifying the usefulness of the services and frequently discovering the recommending services to the end-users. Effective recommendation service needs to satisfy the functional and non-functional specifications of the end- users.Recommendor system getting the details regarding the user based in the effect of contrasting methods and sources to forecast the user needed information founded on the survey process results [3].

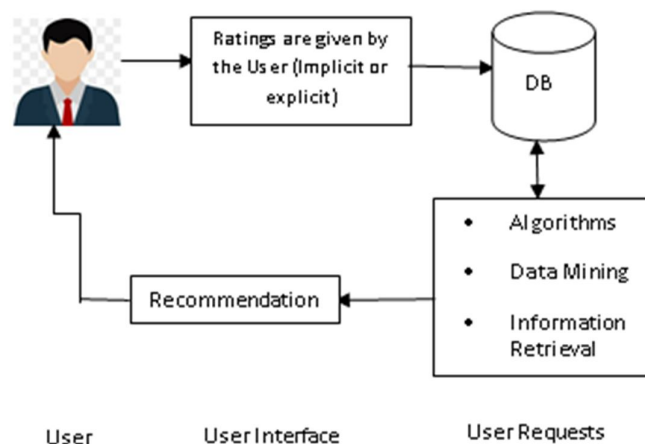


Fig 1.Design of the Recommendation System

Recommender systems are categorized into the following:

- Collaborative Filtering Recommendations
- Content-based Recommendations
- Demographic Recommendations
- Hybrid Recommendations

A. Collaborative Filtering Recommendations

Collaborative Filtering techniques[1] are constructed on the resemblance within the user's rating items or products over the particular website to recommend with other different preferences what she or he may be liked. CF Recommendations collect and analyze user's past historical behavioral details like feedback, preferences, and activities. The action of Identifying similar end-users and similar web services and also recommending the items that he or she likes is said to be collaborative filtering. The Collaborative Filtering recommending the web services to the end-user based on the user's past web service history.[11].

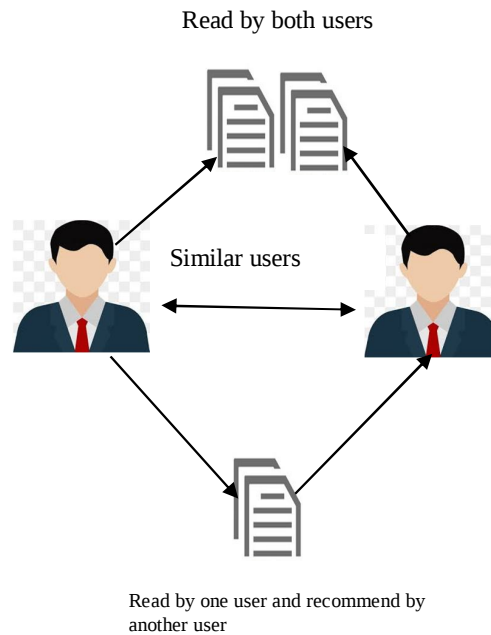


Fig 2. Collaborative Filtering Recommendations

Collaborative Filtering Techniques are two types: Memory -Based and Model-based

CF types	Strength	Weakness
Memory Based CF	Implementation is easy New data will be added easily Need not consider about the content of the items when recommended	Based on the user ratings When the data are sparse, performance will be decreased. Cannot make recommendations to the new users and new items Scalability is limited for large datasets.
Model -Based CF	Good to direct the sparsity,scalabily and other Problems. Improve the Performance of the prediction	Expensive model Have an accord within a performance prediction and scalability

B. Content-Based Filtering Recommendations

Content-based techniques depend on the items or products for the particular user that he or she most wanted in the past. It will be designed mostly to recommend text-based items. Content-based recommendations are based on the reporting of the item or product and a history of the end-user's taste behaviour. Basic operation mostly depends on the content-based techniques consists of user's basic data like age, gender, location, etc [10].

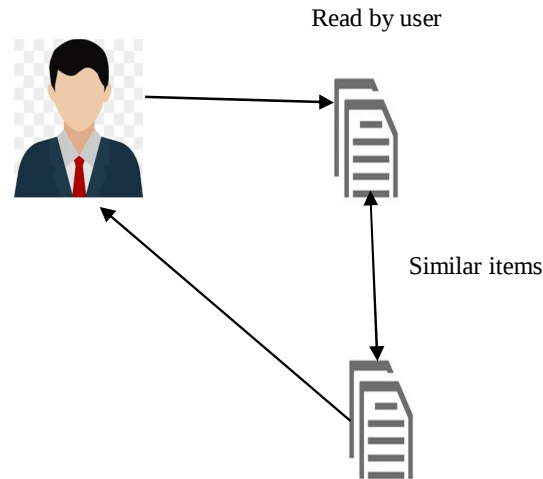


Fig 3. Content-based Filtering Recommendations

C. Demographic Filtering Recommendations

Classifies the user based on personal attributes and builds recommendations based on demographic classes i.e.) College students, Women, Businessman, etc.,. This technique depends on the user's profile information and their rating items and finding the common users have similarity rating items based on their attributes like age group and area of living [12].

D. Hybrid Filtering Recommendations

A hybrid Recommendation system combines two or more recommendation techniques to gain the best system optimization. We are merging the two recommended techniques content-based and collaborative filtering to get the better advantages and to gain the best result. Combining both different kinds of techniques getting high performance and more challenging makes by using collaborative and content-based filtering only [9].

E. Data Mining Methods used in the Recommender Systems

Now we will consider some techniques used in the data mining and Recommender Systems (RS):

- **Detection of Outlier:** We can select the values that are different than any kind of sample of data is known as Outlier.
- **Cluster Analysis:** We will build the cluster or group with the sample data which is similar in such a way with others in similar groups. These algorithms need to find the similar items in terms of specification or end-users purchasing similar items in the big data structure. Finally matching these results and suggest a newly recommended item or product list to the end-user[13].
- **Classification:** data are divided into groups using clustering methods. We will use classification to construct matching from different clusters. This will be achieved by a decision tree. We will define the main node and objects in the decision tree are items or users.
- **Association analysis:** Through this association, we can formulate the inference rule by determining the relationship between the datasets. For example, we need to clarify which kind of product is already purchased together to define the correlation.
- **Regression Analysis:** Regression methods are used to make the models find the relations with the dependent variable through multiple independent variables.

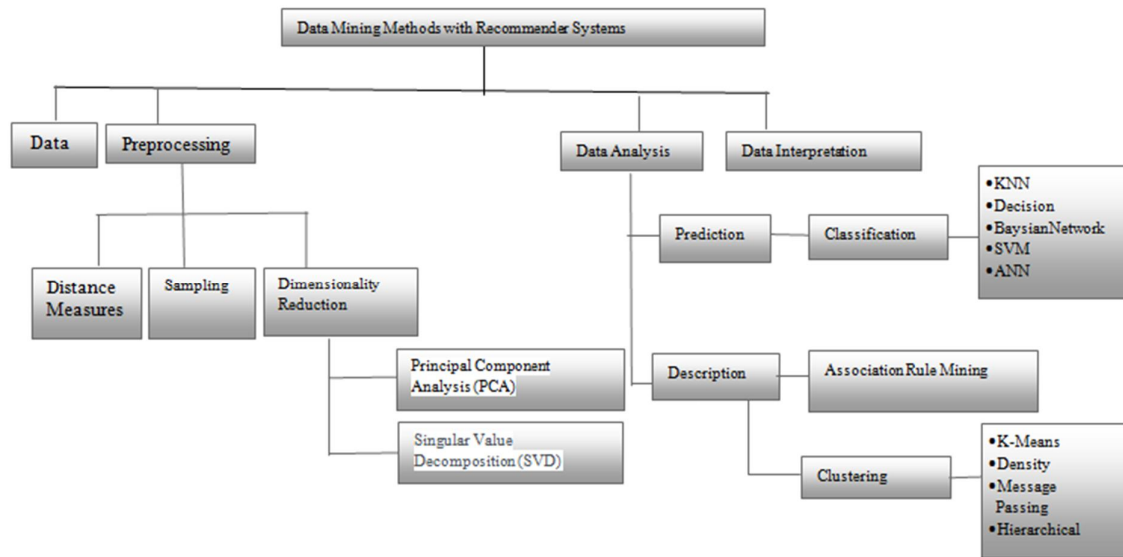


Fig 4. Data mining Methods with Recommender Systems

F. Difference between types of Recommender Systems Techniques

S.No	Techniques	Advantageousness	drawback
1.	Collaborative Filtering Recommendations	Decreasing sparsity and cold start problem on memory-based CF Upgrading scalability of memory-based CF Upgrading accuracy of model-based CF	Gray sheep problem Unscalable for large datasets
2.	Content-based Recommendations	Decreasing new item problem Reducing sparsity problem Reducing gray sheep problem Higher accuracy than memory-based CF	New user problem will arise Unscalable for large datasets The diversity of the recommendations is not higher
3.	Demographic Recommendations	Reducing the new user problem Higher accuracy than memory-based CF Improving prediction accuracy and scalability	Demographic data is difficult to obtain Unscalable for large datasets Gray sheep problem
4.	Hybrid Recommendations	Hybrid combines all the strength within the collaborative and content-based filtering. It will be based on the content description and evaluation of the user. It will solve over- specialization and also increase the satisfaction rate of the user.	Tolerate from the cold start problem. Sparsity problem.

IV. RECOMMENDER SYSTEM RESEARCH CHALLENGES

The Recommender system facing many challenges. In this section we introduce some of the challenges about Recommender Systems:

A. Cold Start Problem

Cold Start problem arising when the new user enters to the website or when adding the new items to the system. There are two problems that should be clarify:

- 1) How to recommend the item to the new user because we don't know the interests and he or she rated none of the items yet.
- 2) To whom to recommend the new item to others because none of the users didn't rate the particular new item either it is good or bad to be most likely by the end-users.[16]
- 3) Finally, the cold start problem arises when the new items are added into the list or a new user enter into the system. The main reason is that few small rated item is available in the system because the new user wanted data can be predicted before he or she purchased or new item can be rated.

B. Data Sparsity Problem

This problem is the most important issue in the recommender systems. It will happen when the user having more amount of a matrix contain purchasing items or listening music. Sparsity will occur when the user didn't rate these items. Recommender systems based on the users rating matrix users to recommend to the others [8]. Sparsity problems arise when the user rated the limited number of items. In the web, thousands of items are available but all the items are not rated or used by the user. This will lead to sparsity problem [14].

C. Scalability

Scalability measuring the efficiency of the item to work effectively with high performance when growing the information. The Recommender system is used to recommend the items to the users without any change when the number of users increased or the number of items increased too. The scalability occurs only if the number of users of items increasing hugely [17].

D. Privacy

One of the most important challenges in the recommender system is Privacy. Recommender systems to recommend items matching their interests and we should know some information about the user's data. Users need to aware, which information is needed to recommend most preferably items [7]. Recommender system need to use the personal data of the needed users but it may lead to data privacy and security issues.

E. Shilling Attacks

Shilling Attacks happen, when the malicious user enters into the system and starts giving false ratings on some of the particular items either to gaining popularity or destroying popularity [6]. Suppose if the user giving a false rating to the particular items either to decrease or increase the popularity of the items, this leads to break the trust of the user also reducing the quality and performance of the Recommendation Systems (RS) [15].

F. Diversity

The Main fact on most of the recommender system works upon the similar user is likely to have similar taste. Most of the recommender system allows us to connect with social media and also recommends the products that are purchased by their friends on social media or similar kinds of items we have viewed before. Whenever we want to view the same type of items that we have viewed before is not the necessary one. It will lead to recommendations tedious. Diversity is the problem that the user has already purchased the particular item but he doesn't want to purchase the same item again but most of the RS shows the similar items that he or she purchased once. Diversity depends on the recommender results span as much as possible our item space and does not come to all from the same cluster [5].

G. Grey Sheep Problem

Grey Sheep Problem arises when the people who do not fully agree or disagree with the group of people who rated the items. This kind of situation leads to the user choices are out of the line. Suppose sometimes one group of people may like the one season of web series but dislike the other seasons. Recommendations for this kind of customers is more difficult because they will like or dislike is not determined [16].

H. Synonym

This problem occurs when a similar item is represented with different kinds of names or similar meanings. In this situation, Recommender system cannot find easily what they needed because it will symbolize the same name or different items. Synonym problem reducing the performance of the RS

To overcome the above challenges, we need to provide better recommendations to the users.

I. Over-Specialization

Suggested item to the end-users are based on those already known or pre-defined by user-information or profile only without finding the newly items and other available options what they have.

V. CONCLUSIONS

One of the most important research fields is recommender systems in today's world. When the number of data sizes such as items and users are increasing also it will raise the big data analysis techniques like a spark, Map-Reduce, etc. [10]. recommender systems to be used to recommend the items to the particular user according to their interests and previous historical items rate list. In this paper, we discussed the four recommendation system

techniques and their challenges like cold start, scalability, privacy and, so on. To increase the performance of the recommender systems, some of the big data algorithms like spark should be improved in future.

REFERENCES

- [1] Xianhu Yuan, Quin Wang and Min sun et al. " Collaborative Filtering Recommendation Algorithm based on Hybrid user Model",FSKD,2010
- [2] Xuelin, Zeng, et al," Parallezation of Latent Group Model for Group Recommendation Algorithm". Data Science in Cyberspace (DSC), IEEE International conference on IEEE, 2016.
- [3] Halder, sajal, AM Jehad Sarkar, and Young-koo Lee,"Movie recommendation system based on movie swarm". Cloud and Green Computing (CGC), 2012, Second International Conference on IEEE 2012.
- [4] Katarya, Rahul, OM Prakash Verma et al, "A Collaborative Recommender system enhanced with particle swarm optimization technique". Multimedia tools and Applications 75.15(2016):1000-1006.
- [5] Mustansar Ali Ghazanfar, Adam Prugel-Bennett et al," A Scalable Accurate Hybrid Recommender system",2010.Third International conference on Knowledge discovery and data mining.
- [6] Kumar, Vikas, et al," Collaborative Filtering using multiple binary maximum margin matrix factorization". Information sciences 380(2017):1-11.
- [7] LakshmiTharun and Pet al," Movie recommender system using item-based collaborative filtering technique", Emerging Trends in Engineering, Technology and Science(ICETETS),International Conference on IEEE,2016.
- [8] Soanpet.Sree Lakshmi and Dr.T.Adi Lakshmi. "Recommendation Systems: Issues and Challenges". (IJCSIT) International Journal of Computer Science and Information Technologies, Vol .5(4), 2024, PP. 5771-5772.
- [9] Mohammed Fadhel Aljunid and Manjaiah D.H "A Survey on Recommendation Systems for Social media using Big Data Analytics". International Journal of Latest Trends in Engineering and Technology, Special Issue (SACAIM 2017), pp.48-58.
- [10] Radhya Sahal, Mohamed helmy Khafagy, Fatima A. Omara," Comparative Study of Multi- query Optimization Techniques Using Shared Predicate-based for Big Data", International Journal of Grid and Distributed Computing Vol. 9, No. 5 (2016), pp.229-240.
- [11] Mina Samir shnoda, Mohamed helmy khafagy, samah Ahmed senbel,"JOMR: Multi-join Optimizer Technique to Enhance Map Reduce Job", The 9th International Conference on Informatics and Systems (INFOS2014), pp.: 80-86
- [12] Hussein SH. Abdel Azez, Mohamed H. Khafagy & Fatima A. Omara ,"Optimizing Join in HIVE Star Schema Using Key/Facts Indexing", IETE Technical Review, 2017.DOI:10.1080/02564602.2016.1260498.
- [13] Shah Khuro, Zafar Ali and Irfan Ullah. "Recommender Systems: Issues, Challenges, and Research Opportunities". Information Science and Applications (ICISA) 2016, pp.1179-1189.
- [14] Dr. Sarika Jain, Anjali Grover, Praveen Singh Thakur, Sourabh Kumar Choudhary." Trends, Problems and Solutions of Recommender System". International Conference on Computing, Communication and Automation (ICCCA2015), (2015). Pp.955-958.
- [15] Yagnesh G. Patel, Vishal P.Patel. "A Survey on Various Techniques of Recommendation System in Web Mining", International Journal of Engineering Development and Research, 2015 IJEDR, Vol 3(4), 2015, pp.696-700.
- [16] Tanya Maan, Shikha Gupta, Dr. Atul Mishra." A Survey on Recommendation System", international conference on recent innovations in management, engineering, science and technology (RIMEST 2018), (2018).pp:543-549.
- [17] Hadeer Mahmoud , Abdelfatah Hegazy , Mohamed H. Khafagy," An approach for big data security based on Hadoop distributed file system", 2018 International Conference on Innovative Trends in Computer Engineering (ITCE), Aswan, 2018, pp. 109-114.