International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

SvedbergOpen
DISSEMINATION OF KNOWLEDGE
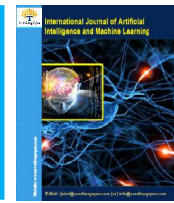
Research Paper

Open Access

Recommendation system using machine learning

Suhasini Parvatikar[1] and Deepa Parasar[2*]

[1] Assistant Professor, Department of Computer Engineering, SCOE, Kharghar, Navi Mumbai, India. E-mail: er.suha@gmail.com
[2] Associate Professor, CSE, ASET, Amity University, Mumbai, Panvel, Maharashtra, India. E-mail: dparasar@mum.amity.edu

Abstract

Due to increase of data on internet, there is an increased dependency on internet by people Thus, recommendation systems help people by suggesting products where is overload of information on ecommerce websites. There are various methods for recommendation. This paper study about various techniques used in designing of recommendation system with machine learning algorithm.

Keywords: *Content filtering, Collaborative Filtering, Naïve Bayes, KNN clustering*

1. Introduction

The use of internet and availability of data has increased enormously due to which people are relying more on internet nowadays. Most used service of the internet (Noraswaliza *et al.*, 2013) is online shopping, where many companies sell their products over the internet by using e-commerce applications like Amazon, Myntra. An e-commerce website provides large database of specified items to an user to choose, from which leads to overload problem of information.

In this scene, the users find difficulty to make decision about which product to choose with vast amount of information available to them. Recommender systems (Noraswaliza *et al.*, 2013) have come-up in response to this problem by learning from users behavior and recommending items that are similar to their interest. Thus, the recommender system helps user to find perfect choice of item required on e-commerce sites.

There are various reasons to use recommendation system, but the reason behind why companies tend to use recommender systems are:

1. To increase sell of their items/products; and
2. To understand users behavior for product purchase.

* *Corresponding author: Deepa Parasar, Associate Professor. E-mail: dparasar@mum.amity.edu*

Recommendation systems provides filtering of information by cutting down large amount of database and recommend items according to their interest. Recommendation systems generally use one of methods as shown in Figure 1 (Yagnesh and Vishal, 2015) for recommendation.

A. Personalized recommendation

Personalized recommendation (Yagnesh and Vishal, 2015) depends on users behavior for selecting item/product on e-commerce website. Generally, personalized Recommender System (RS) are of 5 types:
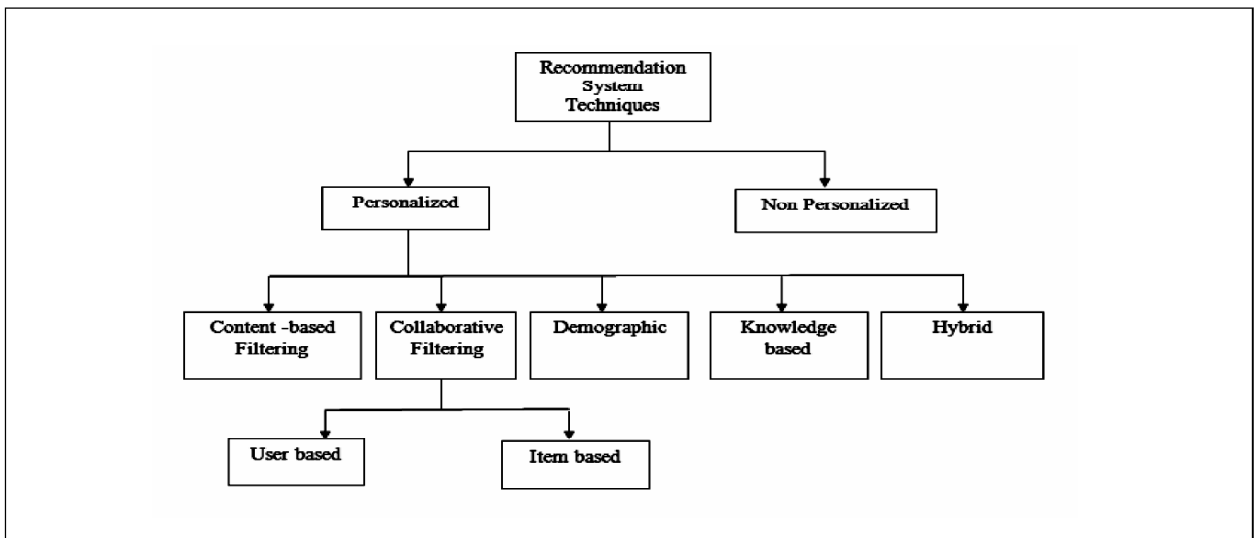

Figure 1: Recommendation system techniques

a) Collaborative based filtering

Collaborative filtering systems generally works by collecting users rating for items in a specified field and then finds likeliness in rating among many users to recommend an item. There are two ways of Collaborative filtering:

1) User based CBF: User-based CBF gives suggestion based on behavior of users who are having similar kind of taste in options. It correlates users rating given to the items as shown in Figure 2 (Anand *et al*., 2014; Yagnesh and Vishal, 2015).

2) Item based CBF: Item-based CBF (Yagnesh and Vishal, 2015) gives suggestion based on the items rated similarly by user as shown in Figure 3 (Anand and Priyanka 2014; Yagnesh and Vishal, 2015).

b) Content based filtering

This type of Recommender System as shown in Figure 4 is dependent on inputs provided by user in past. Examples of content based recommendations are Google, Wikipedia etc. (Francesco *et al.*, 2011).
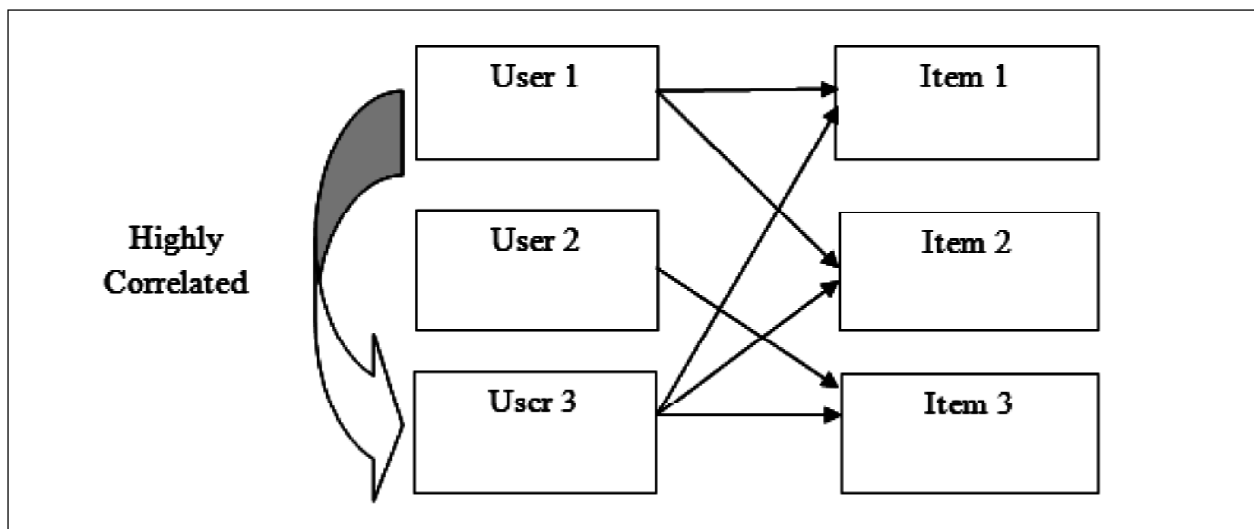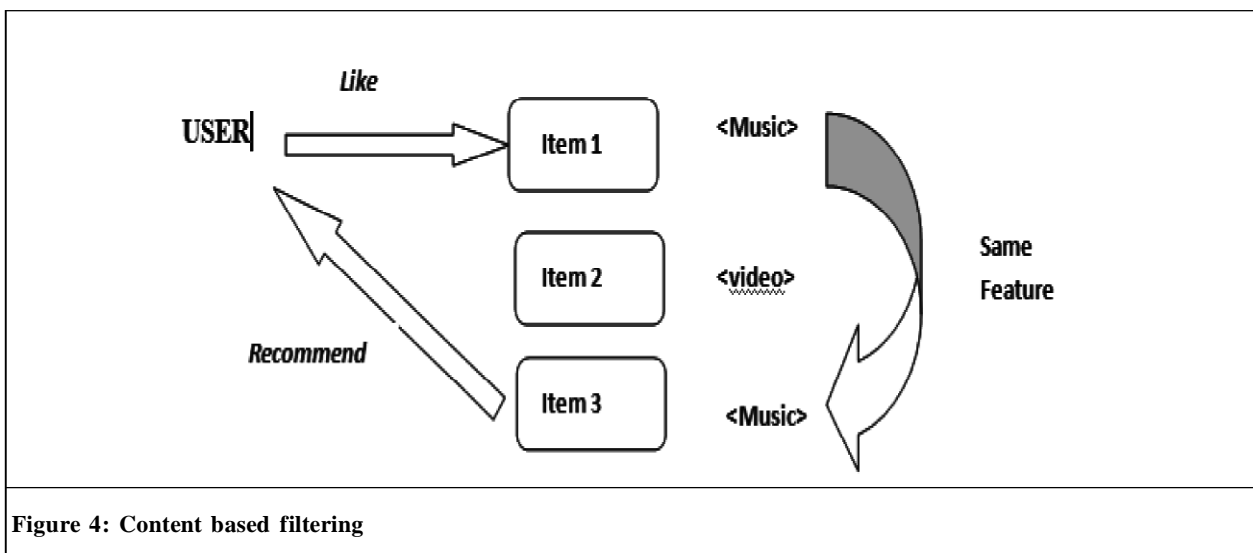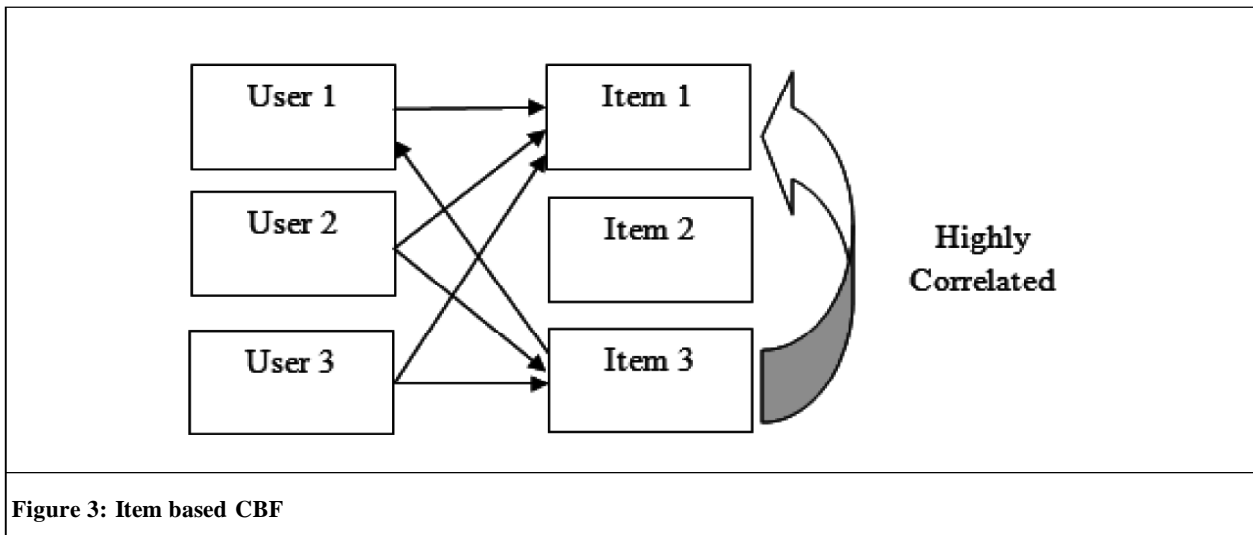

Figure 2: User based CBF

Figure 3: Item based CBF

Figure 4: Content based filtering

c) Demographic based filtering

Demographic based filtering uses user profile information like for e.g., gender, marital status, age, postal-code (Monali *et al.*, 2014). But, this method is time consuming process as it requires users information.

d) Knowledge based filtering

These types of recommendation system ask user for preferences. No interaction of data is needed.

e) Hybrid filtering

Hybrid approach includes the combination of all above methods (Nana Yaw Asabere, 2012).

B. Non personalized recommendation

These RS systems don't rely on user's behavior; they just suggest items randomly to customers based on average feedback about product. Therefore, all customers get the same suggestions on product.

The most common recommender systems applications (Manithva and Sunitha Reddy, 2014; Jaimeel and Lokesh, 2015) are:

- Recommendation system in music domain;
- Recommendation system in e-commerce; and
- Recommendation system in social network.

2. Literature survey

Sunitha *et al.* (2014) proposed hybrid recommendation system which combines clustering and association mining. The main focus of authors was to find solution on CBF challenge, i.e., cold start problem. The proposed algorithm works as

follows: In the first step, user clusters are formed. These clusters are formed on the basis of similarity calculated by cosine similarity method, thus using collaborative filtering. In the second step, each cluster is converted to transactional database. This transactional database is used to find frequent item set by using extended FP tree as shown in Figure 5 (Sunitha *et al.,* 2014). From the frequent item sets strong association rules are formed by traversing the tree when it is required. This method improves the quality of recommendation.

Pijitra (2014), designed a system which uses association rule technique. This method identifies relationship between book items (Mohd Abdul Hameed *et al.,* 2012). The motive of this project is to help user for searching books. This project proposed various steps for book recommendation. In the first step, to collect dataset, apply association rule to recommend books. The results are evaluated by using precision and recall.

Figure 5: Extended FP-tree
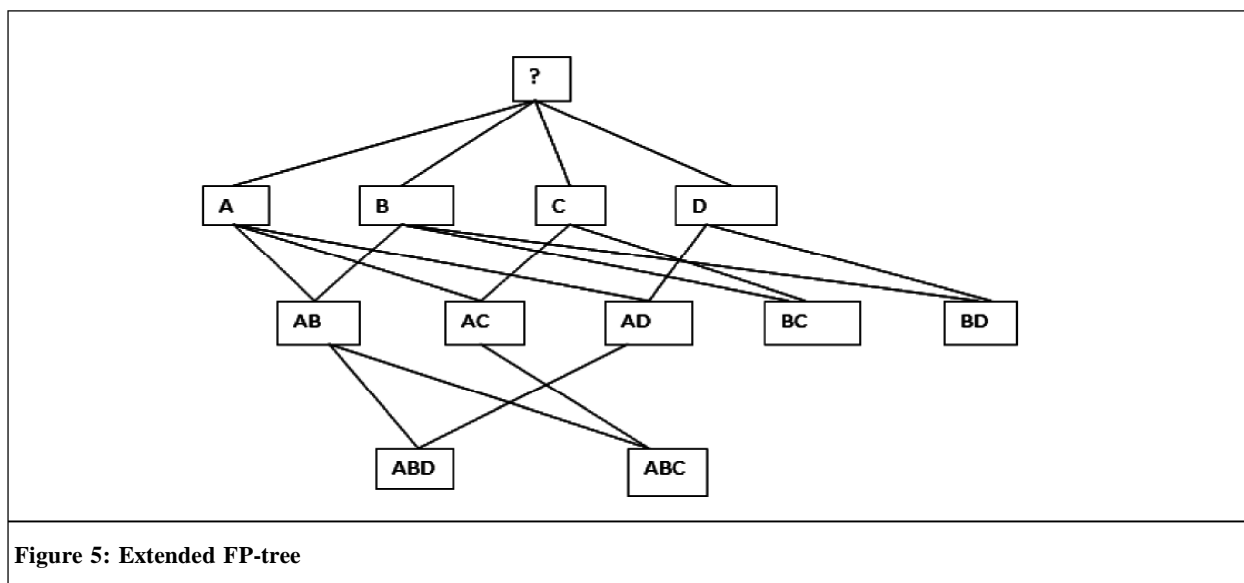
Monali *et al.* (2014), discusses about the reason behind using hybridization of algorithms. The authors described the different methods of recommendation. Association mining alone cannot be effective when the data set is not sufficiently available. The proposed system divides its modules of work into 4 phases. In the first phase, database preselection is done then, association rule mining is used and at last collaborative filtering is done. Finally, top N items are recommended to users.

Hongwu (2011) proposed RS self-organizing maps using combination of collaborative filtering, association mining and to remove data sparsity problem in recommender system.

3. Proposed methodology

The proposed framework consists of following modules: (1) Database; (2) ML Algorithm; and (3) Item based CBF as shown in Figure 6. The database maintains content-based information, i.e., it uses the log details of books which are rated by user positively in the past (Rohini and Kavita, 2013). Our system computes clustering/classification using ML algorithm and then finds similarity between items and gives recommendation. When the user searches for any item from database, the database runs these algorithms and recommends item with user's interest.

A) ML algorithm

We can use unsupervised method, i.e., clustering or supervised method, i.e., Naïve bayes algorithm for classification. Clustering is one of machine learning technique to find hidden patterns in a given dataset. It is an unsupervised learning. In this step the proposed system forms the user clusters based on accessing the frequent book itemset.

This clustering technique represents each user with partial participation in several clusters. If any new user, access any of these items we will be added to corresponding cluster. The use of this type of clustering is just to predict the user's interest in future.

B) Collaborative filtering

Collaborative Filtering (CF) (Heng and Hong, 2009) is most common technique used for recommendation. The main motive of CBF is to find similar users. The users having similar taste, i.e., almost similarly rated items, form a group. For
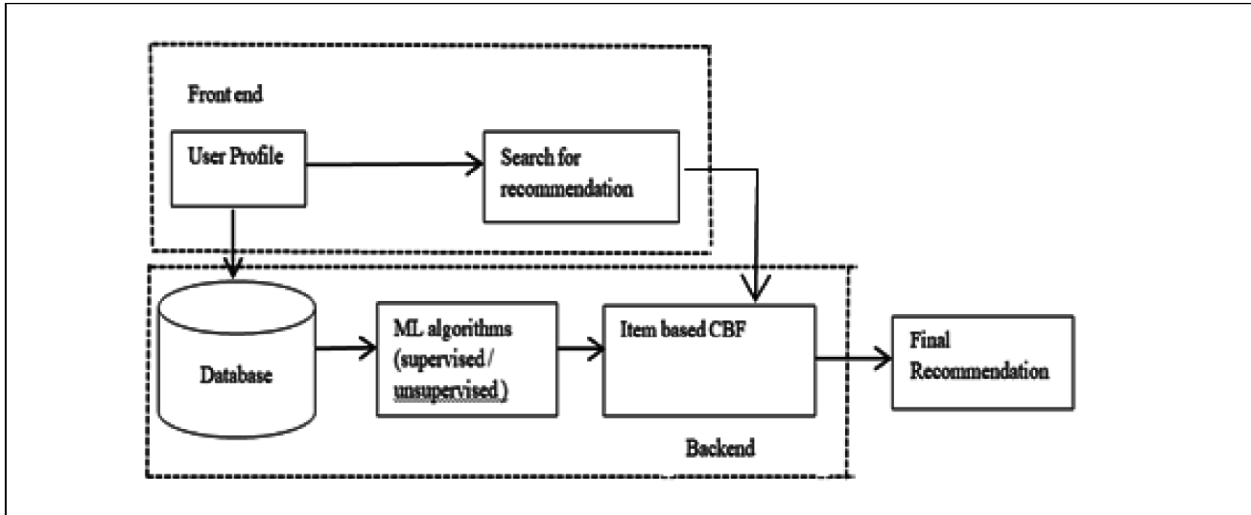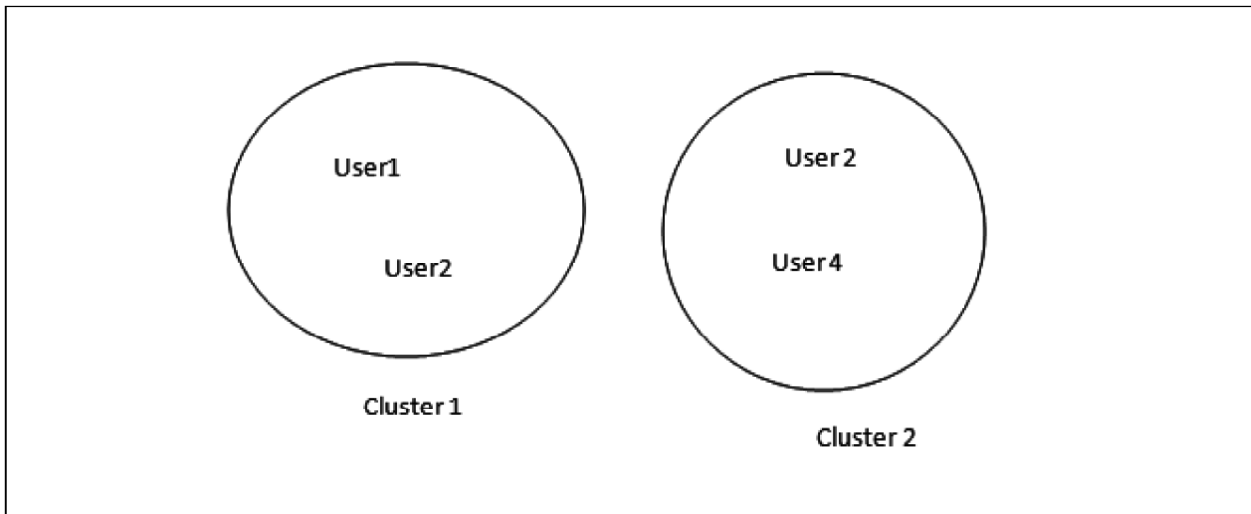
Figure 6: Recommendation architecture

Figure 7: Example of User based Clustering

Table 1: Example of Collaborative filtering

Books User	Java	C++	C#	ASP.Net
A	8	7	9	–
B	9	7	9	9
C	9	–	9	8

example, Table 1 shows that 3 users rated the books positively with similar rating, i.e., they have common taste and thus they form group. The user A has not rated the book ASP.Net, might be he has not watched it yet. As other users have positively rated that, user A will get this item recommended.

The main functions (Noraswaliza *et al.,* 2013) of collaborative filtering recommendations system are:

1. Recommendation: It displays items to a user. The items are listed in the order of usefulness to the user.
2. Prediction: In prediction a calculation of predicted rating is made for a particular item. User based CBF (Figure 7) systems perform poorly when they had many items and comparatively few ratings are present.

So item based algorithm is usually recommended. The Item based CBF algorithm can be summarized in the following steps:

Step 1: Compute similarity between items from above cluster by using Equation (1) as given below.

Step 2: Select *n* active items that have the highest similarity.

Step 3: Compute a prediction from a weighted combination.

a) Item similarity computation

This is most important step in Item based CBF (Sarwar *et al.*, 2011). It finds similarity between items and selects the most similar item. There are different kinds of ways to find the similarity between items using formula:

1. Pearson Correlation: It finds the linear correlation between two vectors of ratings the target item *t* and the remaining item *r*.

$$sim(t,r) = \frac{\sum_{i=1}^{m}(R_{it} - A_t)(R_{ir} - A_r)}{\sqrt{\sum_{i=1}^{m}(R_{it} - A_t)^2 \sum_{i=1}^{m}(R_{ir} - A_r)^2}} \qquad ...(1)$$

where, R_{it} is the rating of the target item *t* by user *i*, R_{ir} is the rating of the remaining item *r* by user *i*, A_t is the average rating of the target item t for all the co-rated users, A_r is the average rating of the remaining item *r* for all the co-rated users, and *m* is the number of all rating users to the item *t* and item *r*.

2. Cosine Vector Similarity: It is the angle between two vectors of ratings as target item *t* and the remaining item *r*.

$$sim(t,r) = \frac{\sum_{i=1}^{m} R_{it} R_{ir}}{\sqrt{\sum_{i=1}^{m}(R_{it})^2 \sum_{i=1}^{m}(R_{ir})^2}} \qquad ...(2)$$

where R_{it} is the rating of the target item *t* by user *i*, R_{ir} is the rating of the remaining item *r* by user *i*, and *m* is the number of all rating users to the item *t* and item *r*.

b) Prediction using item based collaborative filtering

The most important step (Binge and Xin, 2009) during a collaborative filtering system is to get prediction. Since we have similar items supported calculation of similarity measures, subsequent step to urge prediction of rating of target user. The rating (Ujwala and Sheetal, 2013) of the target user *u* to the target item *t* is as follows:

$$P_{u,t} = \frac{\sum_{i=1}^{c} R_{ui} \times sim(t,i)}{\sqrt{\sum_{i=1}^{c} sim(t,i)}} \qquad ...(3)$$

where, R_{ui} is the rating of the target user *u* to the neighbor item *i*, *sim(t, i)* is the similarity of the target item *t* and the neighbor it user *i* for all the co-rated items, and *m* is the number of all rating users to the item *t* and item *r*.

4. Conclusion

Recommendation system is most important part of life. Due to recommendation system, online shopping has gained much popularity. Adequate services are being provided by recommender system. Thus, recommender system has proved a gift for the society which helps customer to decide among their choices. The recommender system uses several approaches for recommendation.

References

Bech, Morten., and Garratt, Rodney. (2017). Central Bank Cryptocurrencies. *BIS Quarterly Review.*

Anand, Shanker, Tewari., Abhay, Kumar., and Asim, Gopal, Barman. (2014). Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining, IEEE International Advance Computing Conference, 500-503.

Anand, Shanker, Tewari., and Kumari, Priyanka. (2014). Book recommendation system based on collaborative filtering and association rule mining for college students. *IEEE International Conference on Contemporary Computing and Informatics*, 135-138.

Binge, Cui., and Xin, Chen. (2009). An online book recommendation system based on web service. IEEE 6th International Conference on Fuzzy Systems and Knowledge Discovery, 520-524.

Francesco, Ricci., Lior, Rokach., and Bracha, Shapira. (2011). *Introduction to Recommender Systems Handbook*, Springer.

Heng, Song, Tan., and Hong, Wu, Ye. (2009). A collaborative filtering recommendation algorithm based on item classification. *IEEE Pacific Asia Conference on Circuits , Communication and System*, 694-697.

Hongwu. Ye. (2011). A personalized collaborative filtering recommendation using association rule mining and self organizing map. *Journal of Software*, 6(4), 732-73.

Jaimeel, Shah., and Lokesh, Sahu. (2015). Recommendation based on clustering and association rules. *International Journal of Advance Research and Innovative Ideas in Education,* 1(2), 49-52.

Manithva V., and Sunitha, Reddy, M. (2014). Music recommend-ation system using association rule mining and clustering technique to address cold start problem. *International Journalof Engineering and Computer Science*, 3(7), 6855-6858.

Mohd, Abdul, Hameed., Omar, Al Jadaan., and Ramachandram, S. (2012). Collaborative filtering based recommendation system: A survey. *International Journal on Computer Science and Engineering*, 5(5), 859-876.

Monali, Gandhi., Khushali, Mistry., and Mukesh, Patel. (2014). A modified approach towards tourism recommendation system with collaborative filtering and association rule mining. *International Journal of Computer Application*, 91(6), 17-21.

Nana, Yaw, Asabere. (2012). Towards a perspective of hybrid approaches and methodologies in recommender system. *International Journal of Computer Science and Telecommunication*, 3(11), 23-32.

Noraswaliza, Abdullah., Yue, Xu., and Shlomo, Geva. (2013). Integrating collaborative filtering and matching-based search for product recommendations. *Journal of Theoretical and Applied Electronic Commerce Research*, 8(2), 34-48.

Pijitra, Jomsri. (2014, August). Book recommendation system for digital library based on user profiles by using association rule. IEEE 4th International Conference on Innovative Computing Technology, 130-134. London. UK.

Rohini, Nair., and Kavita, Kelkar. (2013). Implementation of item and content based collaborative filtering techniques based on ratings average for recommender systems. *International Journal of Computer Applications*, 65(24), 1-4.

Sarwar, B., Karypis, G., Konstan, J.A., and Reidl, J. (2011). Item-based collaborative filtering recommendation algorithms. Proceedings of the 10th international conference on World Wide Web. 285-295. Hong Kong. ACM Press. http://www.ibm.com/developerworks/library/os-recommender1/www.google.com

Sunitha, Reddy, M., Adilakshmi, T., and Swathi, V. (2014). A novel association rule miming and clustering based hybrid method for music recommendation system. *International Journal of Research in Engineering and Technology,* 3(5), 55-59.

Ujwala, Wanaskar., Sheetal. R. Vij., and Debajyoti, Mukhopadhyay. (2013). A hybrid web recommendation system based on the improved association rule mining algorithm.*Journal of Software Engineering and Applications*, 6(8), 396-404.

Yagnesh, Patel., and Vishal, Patel. (2015). A survey on various techniques of recommendation system in web mining. *International Journal of Engineering Development and Research*, 3(4), 696-700.