



**INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH
TECHNOLOGY**

**BOOK RECOMMENDATION USING K-MEAN CLUSTERING AND COLLABORATIVE
FILTERING**

Ritu Rani*, Rahul Sahu

M.TECH. Student, Dept. of CSE, LNCT, Bhopal

Professor, Dept. of CSE LNCT, BHOPAL

DOI: 10.5281/zenodo.1086613

ABSTRACT

With the increase in demand of items amongst customer enhances the growth in information technology and e-commerce websites. This demand is increased due to the availability of web services Personalized preferences and decision making are generated in an application called Recommendation system using an information filtering technique. Relevant features and related items are the characteristic on which this technique works. Suggestion of items, according to user preferences is most important, so suggestion according to similarities provides suitable recommendation. The working of recommendation system for administration has been researched in recent years

KEYWORDS: Book recommendation; Collaborative filtering; K-mean clustering

I. INTRODUCTION

We have lot of choices in the real world and selecting one out of many is the biggest problem. We always take help from some or the other person in choosing the one amongst many. Our family members, friends who have same preference like ours suggest us. But if the same thing happens in virtual world then there selection of interested item is based on recommender system, which is very helpful in selecting the one amongst the many. Through recommender system people can share preferences and most preferred items are offered among them to user from which he can select the interested one. This technique deals with the information overload problem. Useful suggestions are created for customers to get the interesting items.

1.1 USE OF CLOUD COMPUTING

Cloud computing is the most widely used technology, with a huge storage in which bulk amount of data can be stored and retrieved. This technology servers their customers with different services and also provides data in a secure manner. It is also called as storing center where sharing and storing can be done at anytime from any place. Many websites have large amount of data to be stored and this large data is to be stored at a single place with accordance, this type of website are e-commerce website, which have bulk data and arrangement of this data is also essential. On demand services are provided to users over an Internet. By the mean of Internet users are delivered with the services of storing, deploying, sharing, configuring etc. In this approach cloud is used for storing large data of websites like emails, networking site etc. Data of these sites are stored in cloud.

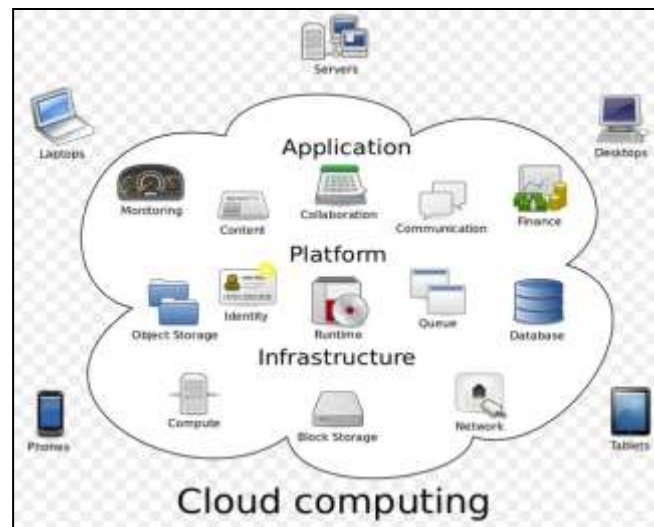


Figure 1. Cloud computing and its services

1.2 BOOK RECOMMENDATION SYSTEM

On the basis of price, quality, publisher and author, recommendation is done. Academician and students are more preferred by the book recommendation system. The purpose is to work on the profile of students and according to the interest based on the store profile recommend items at the time of searching, based on the search and interest of user. Variety of books are offered by book recommendation system, it display the results based on the search of user.

1.3 NEED OF RECOMMENDATION SYSTEM

1. It is used in many e-commerce website because recommendation system is an online approach.
2. Recommendation system helps in finding user interest items from large data.
3. Stores student profile and work according to the profile interest.
4. As the website offers many alternatives, so it evaluates it and presents it to user.
5. It works on the basis of rating and ranking and getting preferences by filtering information.
6. It suggest user to search for the interested items.
7. List of items is displayed on the basis of search.
8. It is a tool which is used over Internet, to navigate and retrieve information according to the preferences.
9. It is a direct way for searching any product, depending on the requirement of user.

II. RELATED WORK

T. Zuva et al. In[1] described about the recommendation system where prediction is done according to the user interest. Several technologies like collaborative filtering and content based filtering which are based on the records of past transaction etc. are used to predict the interest of user.

A.S. Tewari et al. In[2] explained about the combining features of collaborative filtering, content-based filtering and association rule mining for book recommendation system. These techniques have there own advantages and disadvantages. The idea of combining and comparing leads to generate many advantages and disadvantages, which can be overcomes by the other.

Zhao kai et al.In[3] proposed about an improved collaborative filtering approach. The approach of improved collaborative filtering is a user similarity combination based approach in which on the basis of similarity of user combination is continued. Collaborative filtering works as recommending similar items based on user choice and here author proposed improved collaborative filtering which makes combination of similarity among user who has similar choices.

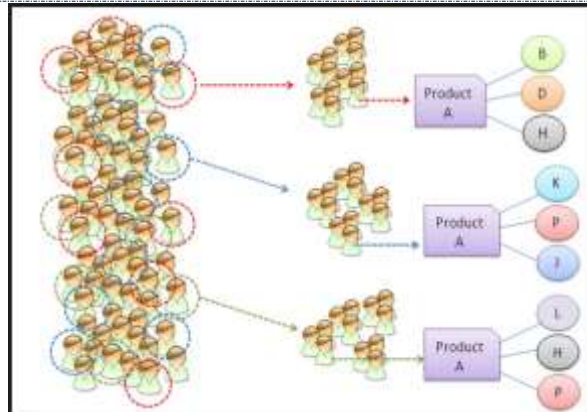


Figure 2. Collaborative Filtering

Pijitra Jomsri et al. In[4] explained about Digital Library for Book Recommendation. This digital library works on Association rule on the basis of user profile. Digital library where user can purchase and read books digitally depending on their choice of interest. Association rule is applied on it where if user is an student then there is more chances to purchase academic books.

Anand Shanker et al. In[5] described about recommendation of book for college students which is based on association rule and collaborative filtering. Author uses both the association and collaborative approach to justify his work and makes a better recommendation for students so that they can get relevant book or books based on particular author for specific subjects. Subjects based on recommendation are also the most beneficial idea.

Nirav M. Khetra et al. In[6] explained about a web personalized recommendation system using a collaborative approach. Web personalization is a trend and is beneficial for both customer and vendor. It benefits customer by providing easy search of the interested items from the huge datasets available by the vendor. It benefits customer in providing or recommending interested items to the user who is in search for the required item. Recommending right product to right customer increases the worth of vendors and will be advantageous for its business.

TABLE 1.COMPARATIVE TABLE

AUTHOR	TITLE	PROPOSED WORK
T. Zuva [1]	A Survey of Recommender System Techniques	Recommendation system where prediction is done according to the user interest.
A.S. Tewari [2]	Based on combine features of content based filtering, collaborative filtering and association rule mining	The idea of combining and comparing leads to generate many advantages and disadvantages,
Zhao kai [3]	Improved Collaborative filtering based on User similarity Combination	collaborative filtering which makes combination of similarity
Pijitra Jomsri [4]	Book recommendation for Digital Library	Digital library in which user can purchase and read

		based on User profile using Association rule	books digitally depending on their choice of interest.
Anand Shanker [5]		Book Recommendation system based on Collaborative filtering and association rule mining for college students	To justify his work and makes a better recommendation for students so that they can get relevant book or books based on particular author for specific subjects.
Nirav M. Khetra [6]		A Collaborative Approach for Web Personalized Recommendation System	Web personalization is a trend and is beneficial for both customer and vendor.

III. PROBLEM DOMAIN

Collaborative filtering depends on the scores and rating of the items, data sparsity occurs in rating matrix of user-items. Inconsistency occurs in searching books from large datasets. Then using k-mean, K-mean is used in our work, which calculates the distance from cluster center to user using a distance formula but it is also not appropriate. After it, k-mean clustering which only calculates those users who scores for items is used. Rating scales for items are different in collaborative filtering for different user. Most of the people gives low scores many gives high. For the favorite books people gives low scores and for the normal books many gives high scores. Similarity calculation for the factors is not considered but can be adjustable to overcome the defects. The approach of using it can provide with the solution balancing the average score and adjust the similarity using method of K-mean clustering. User is assigned with the most similar cluster of his search, while computation of similarity in user and cluster.

The mean of all objects in K-mean clustering is the center calculation in cluster. Moreover, there is zero rating for the user-item which refers to no rating of item, which is the cause of data sparsity. When calculating mean of the cluster center, zero is not calculated and if calculating zero then result comes will be inaccurate. It means if any item is not scored does not mean that customers are not interested in that item. So at the time of mean calculation, the item which is not scored is not calculated and only the user scored items are considered.

On the basis of user interest recommendation is provided by the system. For such systems challenges occurred is like personalized recommendation for the new users, this is a cold start problem. Because the system does not have much detailed information about user so it becomes difficult to system to recommend items. And system does not burden any customer by asking choices and interest of the customers. Getting of information of new users is quite difficult. K-mean algorithm is used to group the users who are distributed into a cluster and the result of it depends on the value of initial cluster center.

In the used approach we have studied about use of data mining so that clustering approach can be observed, this is beneficial in the social networking approach to group the similar users and by grouping of similarities recommendation effort can be reduced. K-mean approach is used for the purpose of recommendation and grouping relevant elements.

IV. ALGORITHM USED

The authorize user can login and read the information and scores of the item. Our system search information in cluster by using the scores of item. Preferred item is scanned in the cluster. It suggests for the items having highest probability of preference in category. Then the system creates top- n number of list for the highest preferences of items and recommend item for purchase ability.

System compares with the past transaction, so as to remove duplicate recommendation. The proposed system algorithm describes personalized recommendation system using K-mean and collaborative filtering.

Step1. When user login, his/her information is created and managed like age, gender, income, occupation and nature etc. of user through social data.

Step2. The login user can read the information of other user through the social data. After it clusters are classified using demographic variables and codes.

Step3. After this, system searches for the preferences of category of items in the cluster of data.

Step4. Highest preferences of items are selected by system based on the data of purchased item. The ordered recommendation is created by descending the preferences.

Step5. Top-n number of list of recommendation is created with highest preferences of item, with the efficiency of purchase ability.

Step6. Comparison is done with the transaction of history data to avoid duplicate recommendation.

V. METHODOLOGY

5.1 COLLABORATIVE FILTERING

This method matches the people of similar taste and then on the basis of personalized recommendation, recommend the user. This algorithm is classified into two entities, the user entity and the item entity. The user entity works on the basis of rating, they rate the item according to their opinion about that item. Recommender system mainly uses collaborative filtering or the combination of it with other algorithm. It mainly focuses on user with same preference and taste and suggest items to them on the basis of selection of items by those users.



Figure 3. Collaborative Filtering for Book Recommendation

5.2 K-MEAN CLUSTERING

K-mean algorithm is proposed by Macqueen, in the year 1967. This algorithm is a simple and learning algorithm. The clustering algorithm used in data mining is k-mean clustering algorithm, it is popularly used in data mining for the clustering of large data. This algorithm works as selection of k objects which is the initial cluster center. Distance between each cluster center and object are calculated and is assign to nearest cluster. Average of all cluster is updated and the process repeats until starts functioning.

K-mean clustering is described as:-

Input data:

database, $N = \{ x_1, x_2, x_3, \dots, x_n \}$

n data objects and k number of cluster,

Output:-

1. from N datasets, k objects are randomly selected and cluster center ($m_1, m_2, m_3, \dots, m_k$),
2. distance between each object and cluster center is calculated and then each object is assigned to the nearest cluster.

VI. PROPOSED WORK

6.1 Data Collection and Preparation

The proposed work implies that to maintain quality and authenticity of any system, information and data source are required. Collection of this data source is as important, data collection can be directly done from user or from any existing system. Primary data collection is done by survey and secondary data collection is from existing system or datasets. Primary data collection is a purpose of primary source.

6.2 Clustering Approach

Similar elements are grouped on the basis of similarities, this approach of grouping a similar elements are called Clustering Approach. The similarity is based on the maximum size of cluster and relevant value. Different elements which have different values exist in another group. Here, K-mean clustering algorithm is used for grouping similarities of books based on similar user. In it work becomes easy for mining and classifying

[Rani* *et al.*, 6(12): December, 2017]
 ICTM Value: 3.00

because it is a simple learning algorithm, which checks for similar elements using K-mean. Distance calculation is done by calculating distance between single cluster center and user. Quantification technique is also used in proposed architecture, which is used to convert relevant values into a matching score.

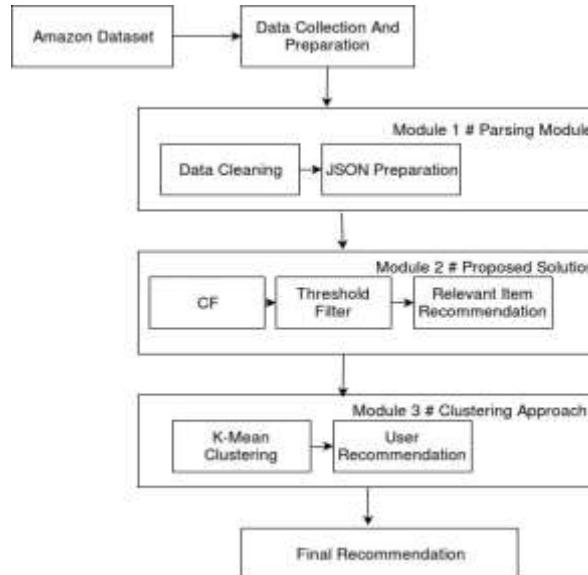


Figure 4.Cluster making

6.3 Recommendation Approach

This approach defines filtering which works on the basis of rating and scoring. Mainly scores are used for predicting the frequency, frequency of items appearing which depends on similarity and preferences. The system of recommendation is now-a-days becoming more popular because of the use of searching items, many users are there who prefer it for books, articles, newspapers, different products and items etc. A recommendation for customized algorithm with K-mean clustering approach is used in order to convey the best recommendation and similar solution. Moreover, cluster recommended as input of data source through which similar score is calculated.

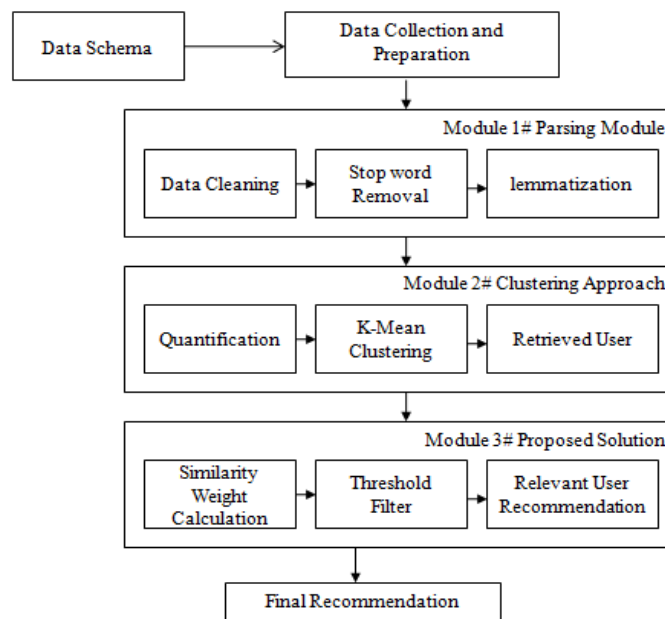


Figure 5.Proposed architecture

VI. RESULT ANALYSIS

Until the execution of any work completed, it cannot be stated as finish. The conditions on which the recorded observation work is as:

1. Accuracy
2. Precision
3. F Score

This all three condition works on the basis of Hybrid Algorithm.

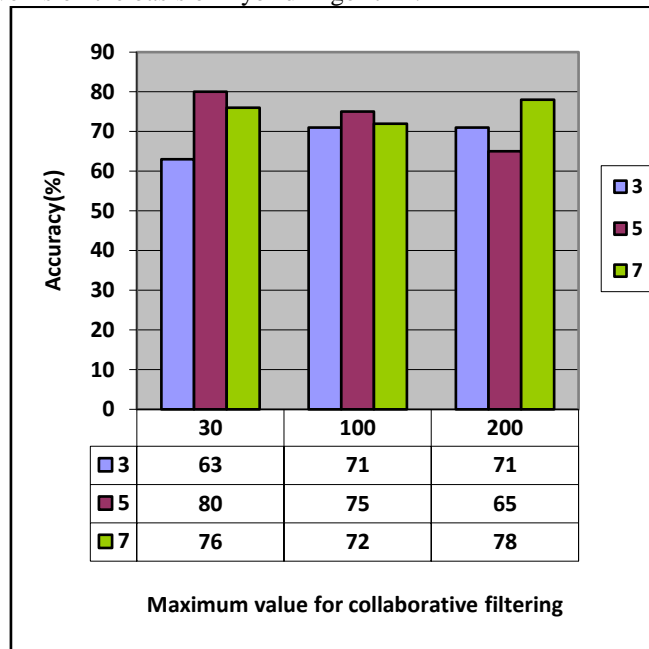


Figure 6. Graph representing accuracy of Hybrid Algorithm

The above graph represents the accuracy of hybrid algorithm showing the maximum value for collaborative filtering. It is shown at different values of 30 100 and 200 with accuracy percentage.

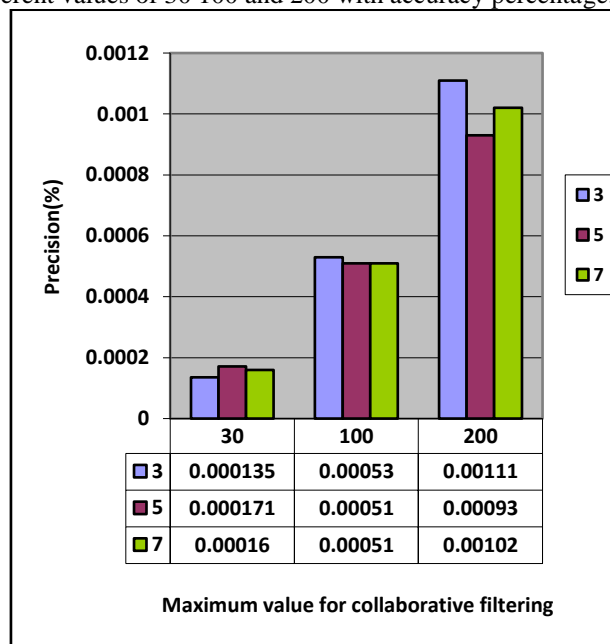


Figure 7. Graph representing Precision value of Hybrid Algorithm

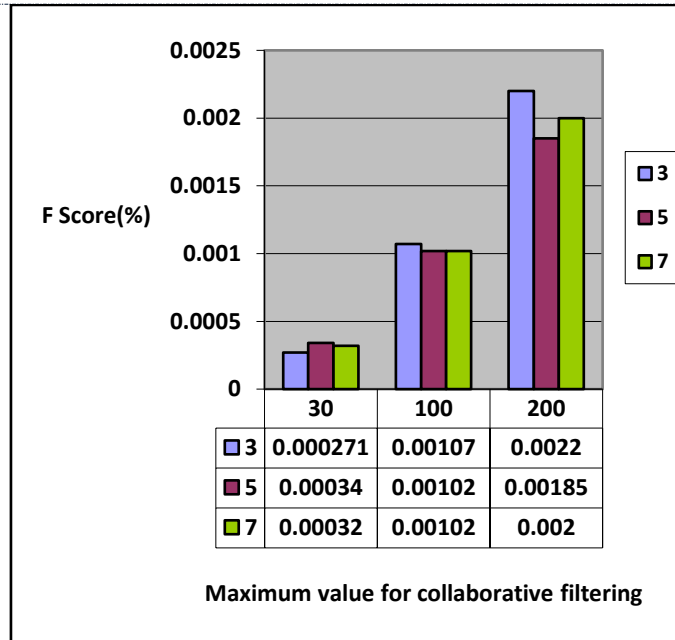


Figure 8. Graph representing F Score of Hybrid Algorithm

Below the table is showing the comparison results of proposed work with the base paper work which is showing the accuracy on the given value of maximum value of collaborative filtering:

Table 2. COMPARISON TABLE

Maximum value of Collaborative Filtering	Accuracy (Base paper results)	Accuracy (Proposed work results)
10	76	90
20	75	84
30	73	80
40	71	79
50	70	76

Accuracy on the given value of maximum result of collaborative filtering through Collaborative filtering and association rule mining algorithms which used in base work can be easily seen given in below graph.

Below graph is showing the accuracy on the given value of maximum result of collaborative filtering of base paper:

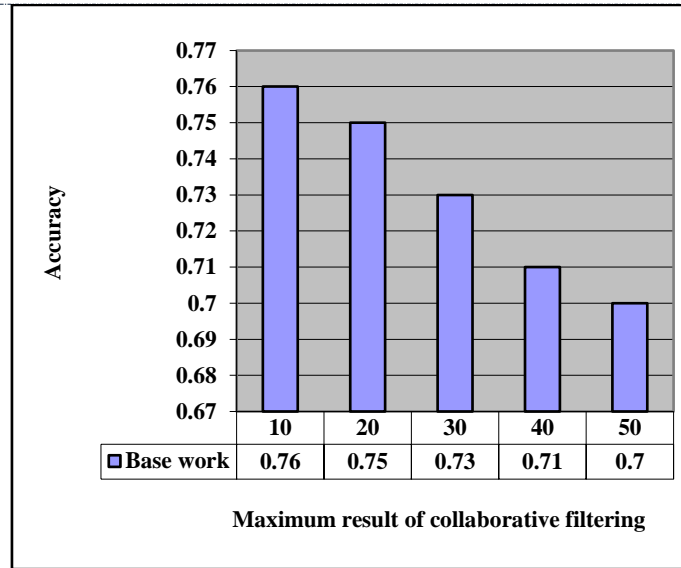


Figure 10. Accuracy of base paper

Accuracy on the given value of maximum result of collaborative filtering through k-mean clustering and collaborative filtering algorithms as my proposed work can be easily seen in given below graph. Below graph is showing the accuracy of proposed work on the given value of maximum result of collaborative filtering:

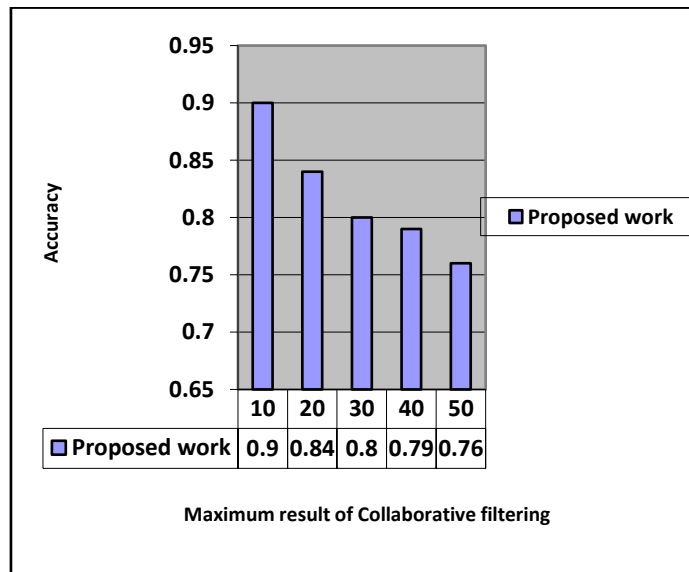


Figure 11. Accuracy of proposed work

VII. CONCLUSION

This system of recommendation is simple and convenient for user to use and reduce searching efforts, while using method by avoiding answers of complicated questions and other method depending on rating of data. The complete work signifies that the similar users are based on the relevant distance. The similarity is based on the maximum size of cluster and relevant value.

VIII. FUTURE WORK

Several limitations are overcome in proposed work using method of collaborative filtering and K-mean, which provides better performance. In terms of accuracy K-mean performs better for most probably accurate



recommendation. The proposed method can be used in real world. Computation can be done to speed up the system. According to the datasets and parallel computing clustering can be altered and more accuracy can be achieved, which is the future extension that can be used in future work.

IX. REFERENCE

1. T. Zuva, S.O. Ojo, S. M. Ngwira, K. Zuva, "A Survey of Recommender System Techniques, Challenges," International Journal of Emerging Technology and Advanced Engineering, vol. 2, no.11, pp. 382-386, November 2012.
2. A.S. Tewari, A. Kumar, and A.G. Barman, "Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining." Advance Computing Conference (IACC), 2014 IEEE International. IEEE, pp. 500 - 503, 2014.
3. Zhao kai ,Lu Peng-Yu , " Improved Collaborative filtering approach based on User similarity Combination",IEEE International Conference on Management Science & Engineering ,Helsinki, Finland ,2014,pp 238-241.
4. Pijitra Jomsri, "Book recommendation System for Digital Library based on User profile by using Association rule", Thailand, IEEE, 2014, pp 130- 134.
5. Anand Shanker Tewari, Kumari Priyanka , "Book Re-commendation system based on Collaborative filtering and association rule mining for college students",IEEE,2014,pp 135-138.
6. Nirav M. Khetra, ShrutiB.Yagnik, "A Collaborative Approach for Web Personalized Recommendation System", IJEDR, Volume 2, No. 4, 2014, pp 3761-3766.