Music Recommendation system using Hybrid Data mining

Tarang Raval[1], Yask patel[2]
[1]M.Tech Pursuing, department of Computer Engineering, parul university, vadodara, india
[2]Asst.Prof, HoD of Diploma department of Computer Engineering, parul university, vadodara, india
Email – [1]TarangRaval1992@gmail.com, [2]Yask.patel@paruluniversity.ac.in

Abstract: *The web and the internet have presented to us a rich data in different fields due to the information overloading. It's very hard to determine relevant content so, Recommended System enters into existence. The principle of recommended system is to provide relevant information to end user. The recommendations highly relevant to decision-making techniques, like what to buy, which music to hear, which online news and article to read, or which movie is most likely one. Recommendation system use method is collaborative filtering, Content-based filtering, and hybrid approach. The main Problem is that is "data sparsity" and "cold start problem". To overcome this Problem we proposed a new method combining the clustering and association rule mining to achieve better performance. The Clustering is applied to the user's and item matrix set into k clusters by applying the k-means algorithm and Association mining is used for generating the rules. Over experimental results obtained the proposed recommendation algorithms perform better and solve the challenges such as data sparsity and cold start problem.*

Keywords: *recommendation system, Clustering, Association mining, K-means*

1. INTRODUCTION:

Recommended systems have become indeed popular in crisp years and are used in various World Wide Web applications. Recommended Systems (RSs) are software tools that are used to give suggestions to user according to their need[8]. The suggestions associate mutually various decision-making processes, one as which items to bought for a song, what music to listen, Item is the general term used to denote what the system recommends to users.

A recommender system is the information filtering that applies data analysis techniques to the problem of helping customers find the products they would like to purchase by producing a predicted likeness score or a list of recommended products for a given customer[10]. Recommender systems work from a specific type of information filtering system technique that attempts to recommend information items (TV program/show/episode, Movie, books, news, images, music, web pages, scientific literature etc.) and social elements (e.g. events or groups, people) that are likely to be of interest to the user[8]. Broadly speaking, a RS suggests to a user those items that might be of user's interest**.**

Recommendation systems method are classified into 3 approaches i.e. collaborative, content based or knowledge-based, and hybrid approach [8].

Collaborative-filtering plays A significant role in the recommendation process also because of that, The collaborative filtering centered on items obtains a likeness of different items from their scores provided by same consumers and make advice in line with the similarity of items. The collaborative selection based on consumers acquires similarity of different consumers using their analysis of the same items and make recommended in line with the similarity of consumers [8].

Content-Based filtering determined by user's previous choices. The item information and a profile of the user's orientation play a natural part in Content-based filtering. in the recommendation process, the engine compares items that were already rated by an individual with items that they would not rate and looks for similarities. The items which are highly rated will be suggested to you [8].

The hybrid approach is a combination of content based and collaborative fileting approach [8].The hybrid filtering approach is introduced to overcome some very that is associated with above filtering methods such as cold start problem, overspecialization problem, and sparsity problem.

Recommending good music can be a difficult task. The song a user might prefer to listen can be based on numerous factors like mood, location, user preferences etc. In our paper, we tried to build a system for personalized music recommendation based on their user preferences which can be determined from their music listening patterns. we try to predict which song might be liked next based on currently rated songs. we consider the songs which are currently being played by the user as factors. Then analyze previous playlists those were played by users. We try to figure out which songs the user might prefer to listen with what types of songs. The proposed method mainly applies an association mining over clustering to recommend the best suitable items to the user by generating better rules. The

Clustering is applied to the group of user's set into k clusters by applying the k-means algorithm and Association mining is used for generating the rules. We also propose a music discovery service as users might also like to listen to new music, Association Rule mining we find which song a user might prefer to listen that are among the currently available songs, and we take the support/confidence value. Now we pick each song from this list and based on the confidence of new songs, we recommend the next song.

2. ASSOCIATION RULE MINING:

Association rule mining discovers interesting association and correlation relationship among rich data set of items. The the exposure of interesting correlation among a large amount of business transaction data helps in various business decision making processes, such as customer shopping behavior analysis [16]. One of the typical examples of association rule mining is Market basket analysis. In market basket analysis customer buying habit is analyzed for finding an association between different items, the client keeps accordingly in their shopping cart.

Let I = {i1, i2…... ,im} be a set of items. An association rule can be represented by this form A→B, where A □ I, B⁻ I and A ⁻ B=⁻ [16]. Association rule extract the pattern from the database based on the two measures minimum support and the minimum confidence. The support and confidence measures are described as stated in [16].

Support: The minimum percentage of instances in the a database that contains all items listed in a given association rule. The rule A→B holds in the transaction set D with support s, where s is the percentage of the transactions in D containing A⁻ B[16]

Support (A□B) =P (A⁻B) (16)

Confidence: Given a rule of the form "if A then B", confidence is defined as the conditional probability that B is true when A is known to be true. Thus, the rule (A→B) has confidence c in the transaction set D, where c is the percentage of transactions in D containing A that also contain B[16].

Confidence (A→B) = P (B|A) (16)

3. OUR PROPOSED METHOD:

Our proposed approach is mainly divided into three sections. First, it is to convert the raw data into user-item rating matrix. On that Applying Clustering algorithm on user-item rating matrix and will result in cluster groups. Then applying association mining algorithm for rules generation and based on the rules generated will try to generate strong rules and with the help of strong rules best items will be recommended to the user.

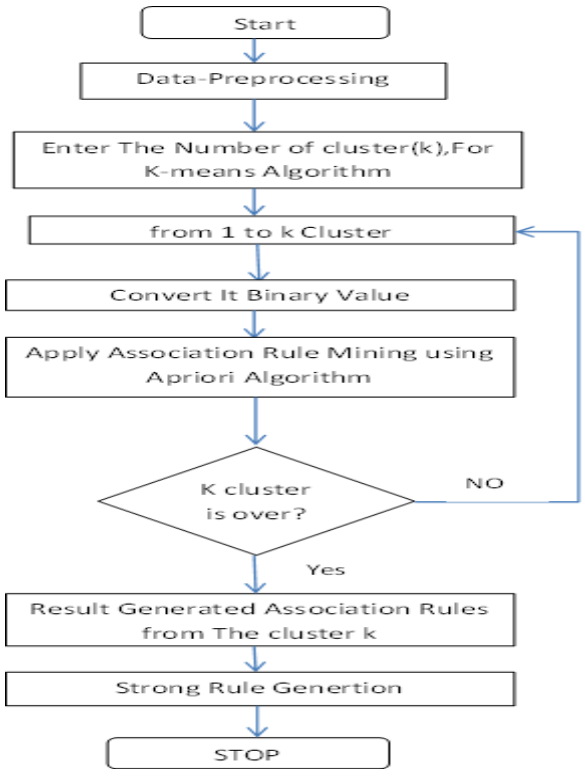

Figure 1. Proposed Flow Diagram

Step 1: Here dataset is, It contain user id and item name, its rated value field. On this data set preprocessing is applied. So, it will generate the user-item matrix.

Table I: Dataset type and Preprocessing on it generate user item matrix

userID	songID	Rating
U1	S1	3
U2	s2	1
U3	S2	4
U3	S1	5

⇩

userID	S1	S2
U1	3	0
U2	0	1
U3	5	4

Step 2: After Preprocessing on this apply clustering, K-means clustering applied on preprocesses data and generates it into a similar user group.

Table 2:Preprocess Data into cluster

userID	S1	S2
U1	3	0
U2	0	1
U3	5	4

⇩

userID	S1	S2
U1	3	0
U2	0	1
U3	5	4

Step 3: Each Cluster data is converted into Boolean value, convert the numerical data to Boolean of the cluster data, Because of finding out the efficient rules generation.

Table 2: cluster data Convert into Binary

userID	S1	S2
U1	3	0
U2	0	1

⇩

userID	S1	S2
U1	TRUE	FALSE
U2	FALSE	TRUE

Step 4: Apply Association rule Mining on cluster data for frequent rule generation, Apriori Algorithm use for Frequent Rule generation. enter the minimum number of support and confidence. So, it will generate the frequent item-set with minimum support which is given as an input.

Item1->Item2
Item1,Item2->Item3

Step 5: Extract the Strong Rule from the rule in this best rule is recommended to user. Strong rule find the rule is high confidence and its rule have include of rated item.in this rule find the strong ruleand recommended to user.

4. EXPERIMENTAL AND ANLYSIS:

This section presents an experimental study of our proposed framework. It describes the experimental setup, presents the experiment results, and finally it summarizes our observation.

A. Dataset Preferences

Experiments is to prove the performance of the proposed recommendation method. In this paper, a popular dataset, namely Movielens [7] dataset, is used for conducting the experiments. The Movielens is the closely utilized dataset to indicate the performance of recommendation algorithms. This dataset contains 100,000 ratings on a scale from 1 to 5 of 1682 movies by 943 users. In this dataset, users have rated at least 20 items.

B. Effect of Cluster No

Our Approach have Clustering and association Mining on that no of clustering Group experiment of number of different cluster apply and measures time on that number of cluster. Based on that say that more cluster on number based on data it required less time.

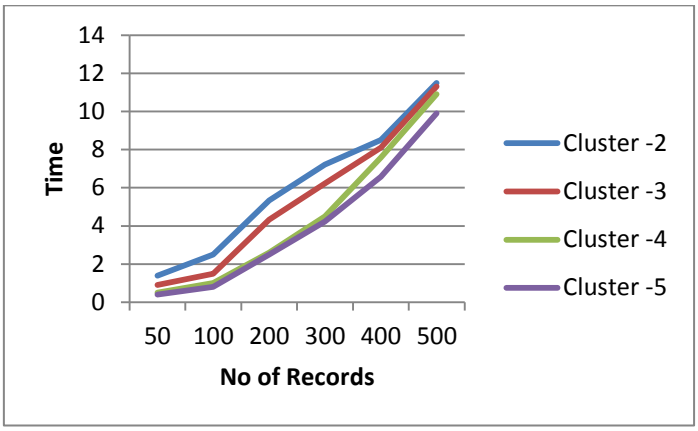

Figure 2 Time Required with apply no of Cluster based on No of Record

Figure 2 shows the no of Records Is m×n matrix like 50 rows and 50 column, based on this experiment the cluster no is higher than time is decreasing of executing proposed method.

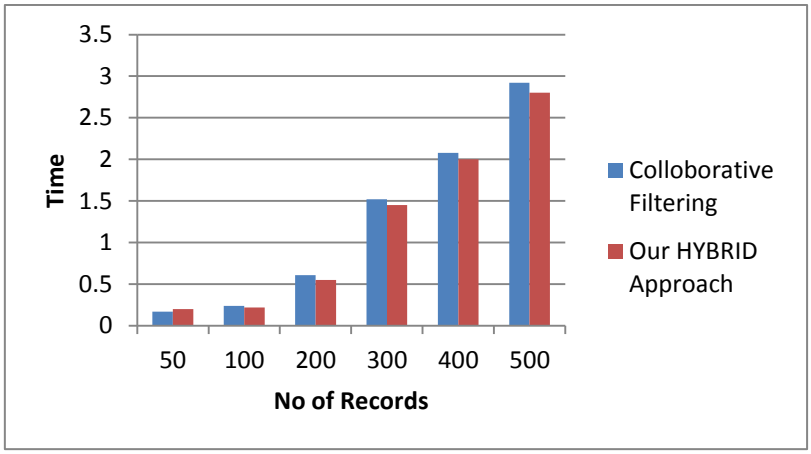

Figure 3 shows that time required of executing proposed Method compare with collaborative filtering. Our method is Executing with More Number of cluster is take less or equal time required for executing method.

5. CONCLUSION:

In this paper we have propose a new hybrid based recommendation method that combines Clustering and Association mining hybrid Approach. Here we have used a k-means clustering to cluster the user profiles and on the cluster data applying algorithm to generate the best rules for the recommending the items to the user. The experiments demonstrate that our method achieves the improvement in time and also reduce the sparsity. The future work is to improve the accuracy to the system by enhancing the clustering algorithm

REFERENCES:

1. JinHyun Jooa,SangWon Bang,GeunDuk Park,"Implementation of a Recommendation System using Association Rules and Collaborative Filtering.", Procedia Computer Science, Aug 2016, pp 944–952.
2. Gurpreet Singh, Rajdavinder Singh Boparai, "A Novel Hybrid Music Recommendation System using K-Means Clustering and PLSA.", Indian Journal of Science and Technology, July 2016.
3. Chuen-He Liou, "Personalized Article Recommendation Based on Student's Rating Mechanism in an Online Discussion Forum .", System Sciences (HICSS), 2016 49th Hawaii International Conference, 5-8 Jan. 2016.
4. Peng Yu, " Collaborative filtering recommendation algorithm based on both user and item ", Computer Science and Network Technology (ICCSNT), 4th International Conference on, 19-20 Dec. 2015.
5. Anand Shanker Tewari, Kumari Priyanka, "Book Recommendation System Based on Collaborative Filtering and Association Rule Mining for College Students", Contemporary Computing and Informatics (IC3I), 2015 International Conference on, 27-29 Nov. 2015.
6. Lin Ziqi, Ni Wancheng, Zhang Haidong, "A K-medoids Algorithm Based Method to Alleviate the Data Sparsity in Collaborative Filtering .", Control Conference (CCC), 34th Chinese, 28-30 July 2015.
7. Mohammad Soleymani, Anna Aljanaki. Frans Wiering, "Content-Based Music Recommendation Using Underlying Music Preference Structure.", Multimedia and Expo (ICME), 2015 IEEE International Conference, 29 June-3 July 2015.
8. Paritosh Nagarnaik, Prof. A.Thomas, "Survey on Recommendation System Methods", International Conference On Eletronics And Communication System (ICECS), 2015 IEEE International, 26-27 Feb. 2015
9. Xuejiao Han, Wenqian Shang, Shuchao Feng, "The Design and Implementation of Personalized News Recommendation System", Computer and Information Science (ICIS), 2015 IEEE/ACIS 14th International Conference on, 28 June-1 July 2015.
10. Poonam B. Thorat, R. M. Goudar, Sunita Barve, "Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System ", International Journal of Computer Applications,Volume 110 – No. 4, January 2015
11. Anand Shanker Tewari, Abhay Kumar, Asim Gopal 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, 21-22 Feb. 2014.
12. X. Luo, Y. Xia, Q. Zhu, "Applying the learning rate adaptation to the matrix factorization based collaborative filtering", Knowledge Based Systems 37 (2013) 154–164.
13. J. B. Schafer, D. Frankowski, J. Herlocker, S. Sen,"Collaborative filtering recommender systems", in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, 2007, pp. 291–324 .
14. C.C. Aggarwal, Recommender Systems the textbook
15. M. Pazzani, "A framework for collaborative, content-based, and demographic filtering", Artificial Intelligence Review-Special Issue on Data Mining on the Internet 13 (5-6) (1999) 393–408.
16. J. Han, M.Kamber, Data Mining : Concepts and Techniques, The Morgan Kaufmann Series, 2001.