IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"A NOVEL RECOMMENDATION TECHNIQUE TO OVERCOME THE SPARSITY PROBLEM BASED ON THE COLLABORATIVE FILTERING "

Prashant Pathak [1], Dr. Manmohan Singh [2]

[1,] *M.Tech Scholar, Dept. of Computer Science & Engineering, RKDF School of Engineering, Indore*
[2]*Associate Professor, Dept. of Computer Science & Engineering, RKDF School of Engineering, Indore*

ABSTRACT

Presently a day's Recommendation systems are generally utilized as a part of online business. Now a day's large portion of the general population utilize internet business for obtaining the items. Web business includes assortment of items so consumer can not select which items suites the best so hence, Recommendation system was introduced. With the help of Recommender system we can determine about consumer interests and consequently propose the best proposed items to the consumer. Recommender systems assume a critical part in sifting and modifying the coveted data. The principle point of the proposed work is to determine the best appropriate items to the client. So to suggest the most effective acceptable things higher run extraction and selection is needed. The proposed work basically applies an association rules for higher extraction and selection to suggest most appropriate item to user and get higher performance accuracy.

Keyword: Recommendation System, Recommendation Methods, Type of Recommendation System.

I. INTRODUCTION

In regular daily existence, individuals depend on recommendations from other individuals by talked words, reference letters, news reports from news media, general overviews, travel guides and so on so recommendations assumes a critical part in finding the best things. A recommender framework is the data sifting that applies information examination methods to the issue of helping clients discover the items they might want to buy by delivering an anticipated resemblance score or a rundown of prescribed items for a given client. Recommender frameworks work from a particular kind of data sifting framework method that endeavour's to suggest data things (films, television program/appear/episode, music, books, news, pictures, website pages, logical writing and so forth.) or social components (e.g. individuals, occasions or gatherings) that are probably going to hold any importance with the user[1]. There are for the most part 3 sorts of strategy i.e content based filtering, collaborative based filtering , Hybrid filtering. The Content based filtering techniques are situated in light of a depiction of the thing and a profile of the client's inclination. In Collaborative based filtering strategies depend on gathering and breaking down a lot of data on clients' practices, exercises or inclinations and anticipating what clients will like in light of their likeness to different clients.

The main purpose of proposed solution is to recommend the best suitable items to the user to recommend the best suitable items. The proposed approach is design for solving the data sparsity problem in recommender systems. The approach is to make improvement in current apriori algorithm to apply association mining algorithm on different clustering groups. And to generate strong rules which results best suitable item to the user.
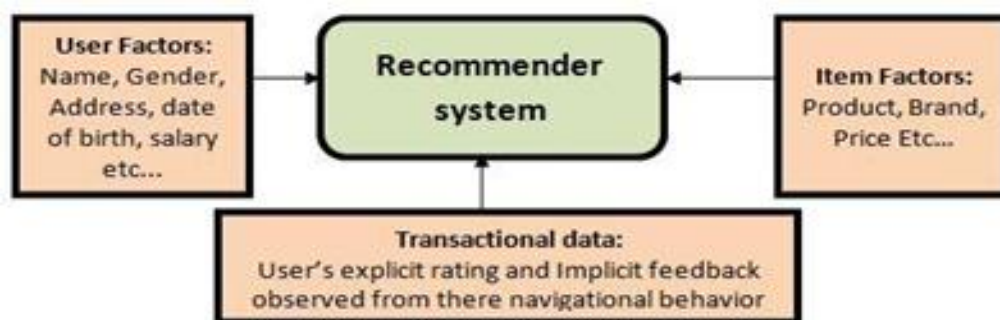
Figure 1: Overview of Recommender System

II. PROBLEM DOMAIN AND OBJECTIVE

Data mining and Recommendation system are quickly developing exploration region in the field of research today. In recommendation system When client look through any content, the system need to prescribe the best appropriate things to the client, so to find the best reasonable things better rule extraction and rule choice is required. So applying the association mining algorithm which leads to better rules extraction.

➢ Objective of this work is to provide an approach to recommend the best suitable items to the customer.
➢ To overcome the data sparsity problem in recommendation system.
➢ To Improve the Rule extraction and selection by applying an association rules on Pre-processed Data.

III. METHODOLOGY

3.1 PROBLEM DEFINITION

This proposed solution is based mostly on recommendation system that recommends completely different things to users. This technique can advocate books to users. This technique can offer additional precise results as compared to the present systems. The present system works on individual users' rating. This might be someday useless for the users who have completely different style from the recommendations shown by the system as each user might have different tastes. This technique calculates the similarities between different users and so advocate book to them as per the ratings given by the different users of similar tastes. This can offer a precise recommendation to the user.

3.2 EXISTING SYSTEM

Recommender systems usually manufacture an inventory of recommendations in one amongst 2 ways that – through collaborative filtering or through content-based filtering (also referred to as the personality-based approach).Collaborative filtering approaches build a model from a user's past behaviour (items antecedently purchased or designated and/or numerical ratings given to those items) also as similar choices created by different users. This model is then accustomed predict things (or ratings for items) that the user might have associate interest in. Content-based filtering approaches utilize a series of separate characteristics of associate item so as to suggest further things with similar properties. These approaches are usually combined named as Hybrid Recommender Systems.

Challenges & Limitation in Recommender systems:

1. Cold Start Problem
2. Data Sparsity Problem
3. Scalability of the approach
4. Gray Sheep

5. Accuracy of the prediction
6. New user: when there's not enough information to build a solid profile for a user, the recommendation could not be provided correctly
7. Limited content analysis: if the content does not contain enough information to discriminate the items precisely, the recommendation will be not precisely at the end

3.3 PROPOSED APPROACH

3.3.1 PROPOSED SYSTEM

- In this research, we are generating association rules which recommends suitable item to user for selection.
- Define the pre-processed book data set on which contains user based rating against the books
- Apply the association mining algorithm i.e. proposed algorithm on this pre-processed book data set
- By applying the proposed association mining algorithm generate strong rules
- Based on the generated the strong rules recommends the suitable item to the user

3.3.2 ADVANTAGES OF PROPOSED SYSTEM

- The propose scheme allows to recommend suitable item to the user based on the behaviour of the user.
- The propose scheme allows user independence i.e. the content suitable for user profile are suggested
- The propose scheme allows transparency i.e. recommendation based on the unknown users have taste like you
- The propose scheme resolved the data sparsity problem occur in recommendation system

3.4 PROPOSED ALGORITHAM

The main purpose of proposed algorithm is to recommend the best suitable items to the user. The proposed algorithm generates the strong rules of pre-processed data to recommend suitable item.
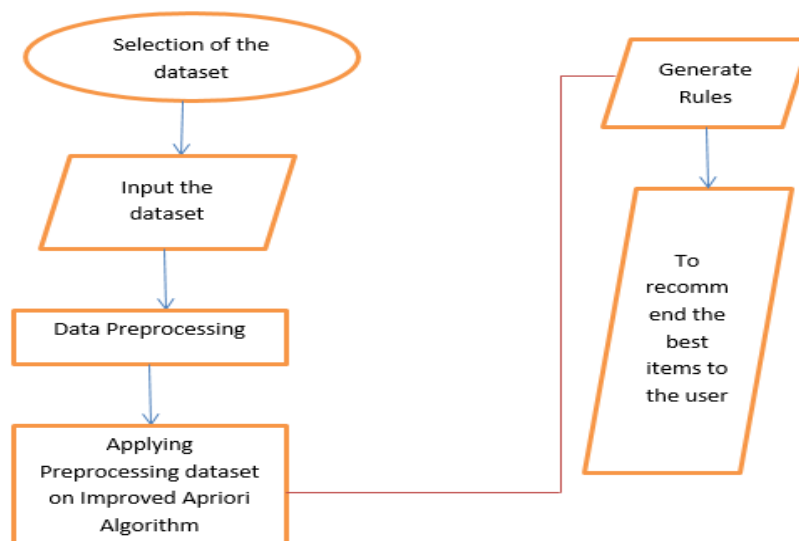

Figure 2: Proposed Algorithm

3.4.1 ALGORITHAM FLOW

As shown in figure 2 the algorithm works as follows:

Step 1: Input of the dataset

Step 2: Data pre-processing

a. In this step the pre-processing on dataset is done by applying proposed association mining rule

Step 3: Applying pre-processing dataset on improved Apriori Algorithm

a. In this step the proposed improve algorithm is apply on the pre-processed dataset in step 2
b. The mathematical operation are performed in this step to calculate the confidence of each item

Step 4: Generate Rules

a. In this step the proposed algorithm which is run in step 4 generates the strong rules based on confidence.
b. Based on the generated rules best item recommend to the user

3.5 PROPOSED CONTRIBUTION

1. This research suggest the algorithm to recommend the best item to the user
2. Examine improvements in performance accuracy
3. Determine the data sparsity problem, and proposed mathematical model to overcome this problem

3.6 PROPOSED ARCHITECTURE
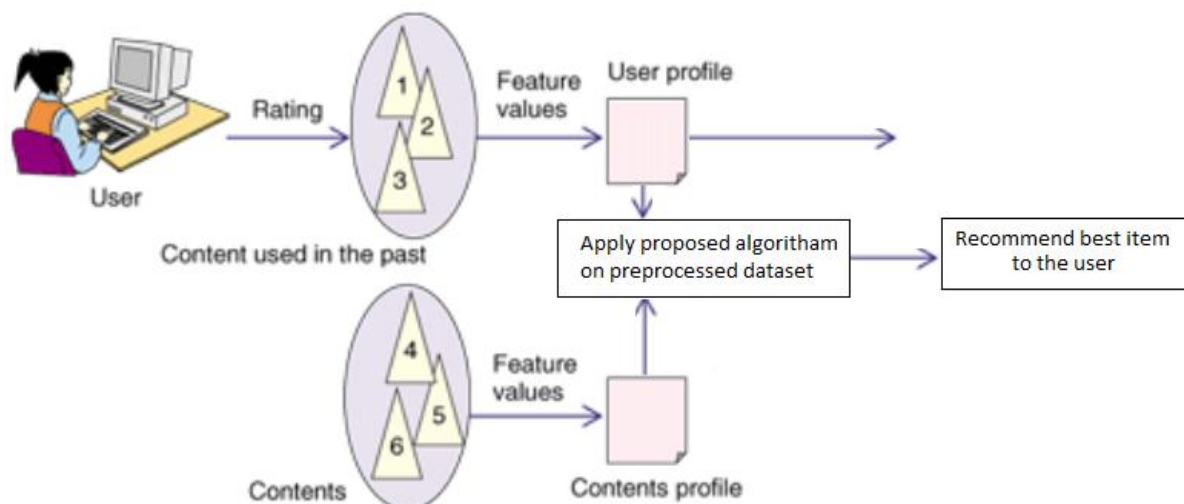
Figure 3: Architecture of Recommendation System

3.7 MATHEMATICAL MODEL

The proposed mathematical model takes input as preprocessed data. Once preprocessed data receive for each item we are converting dataset value to Boolean value for each item rating. Then based on support item set are generated. The items are selected from generated item set based on confidence formula calculation.

Step 1: Consider user matrix dataset for item rating

User/Item	R1	R2	R3	R4	R5
U2	5	0	3	1	0
U4	0	4	4	4	4
U7	4	4	4	3	0
U10	5	0	0	3	3

Step 2: Converting the exiting dataset into Boolean values. 0 rating means false and 1-5 rating means true.

User/Item	R1	R2	R3	R4	R5
U2	True	False	True	True	False
U4	False	True	True	True	True
U7	True	True	True	True	False
U10	True	False	False	True	True

Step 3: Generate frequent Item Set 1

R1 = True	{ U2, U7, U10 }
R2 = True	{ U4, U7 }
R3 = True	{U2,U4,U7}
R4 =True	{U2,U4,U7,U10}
R5= True	{U4,U10}
R1=False	{ U4}
R2 =False	{U2,U10}
R3 =False	{U2,U10}
R4 =False	{ }
R5= False	{U2,U7}

Step 4: Generate frequent Item Set 2

R1=True,R2 =True	{ U7}
R1 = True,R2 =False	{U2,U10}
R1 =True,R3 =True	{U2,U7}
R1=True,R5 =False	{U2,U7}
R1=False,R5 = True	{U4,U7}
R2=True,R3=True	{U4,U7}
R2=True,R4=True	{U4,U10}
R3=True,R4=True	{U2,U4,U7}
R3=True,R5=False	{U2,U7}
R4=True,R5=False	{U2,U7}
R4=True,R5=True	{U4,U10}

Step 5: Generate frequent Item Set 3

R1=True, R3 = True, R4 =True	{ u2,u7}
R1=True,R3=True,R5=False	{u2,u10}

Step 6: Generate the rules by applying formula of confidence as below.

Formula of confidence = Support (A U B) / Support (A)
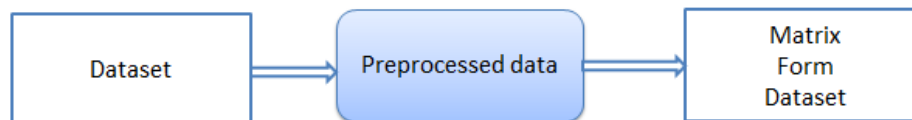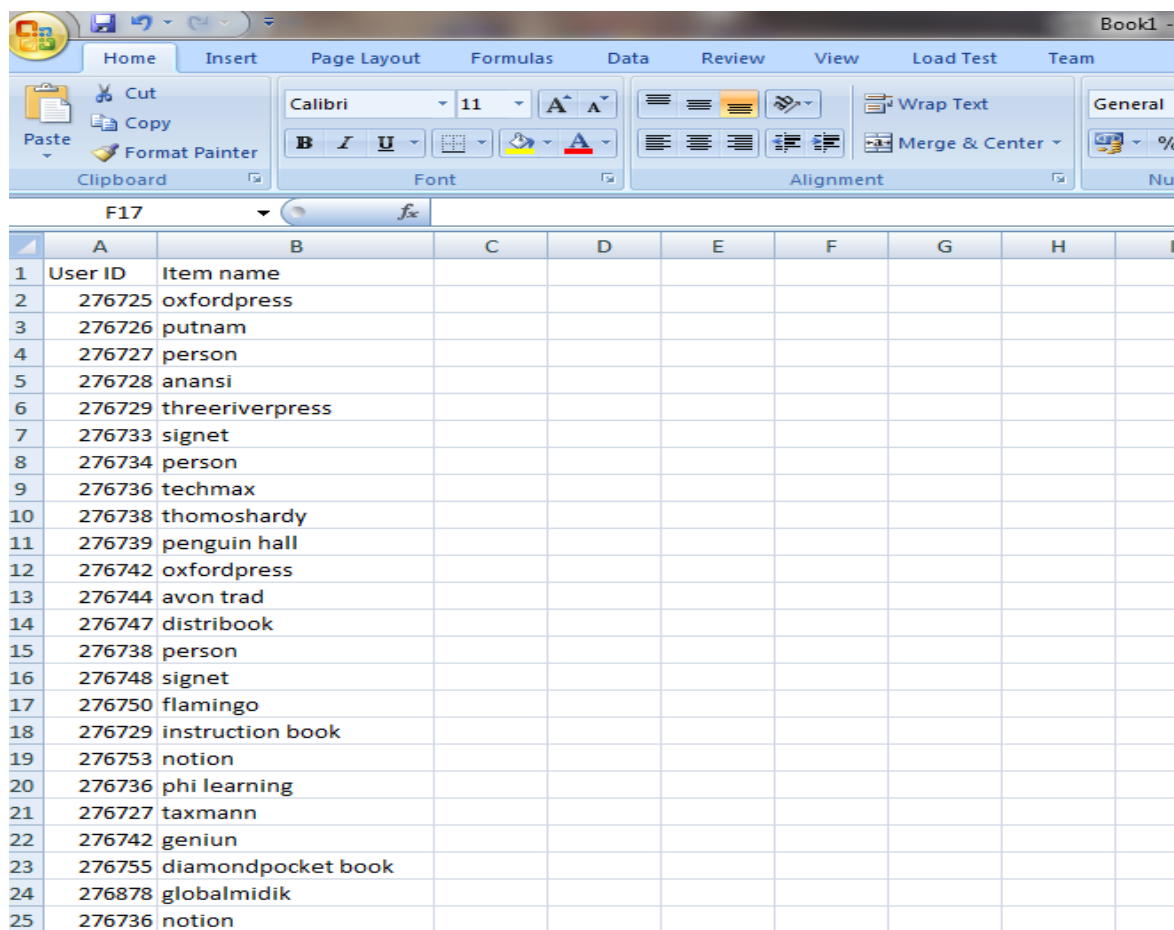
Here it satisfy the minimum support and minimum confidence so rules generated as follows

R1 = True -> R5 = False: Confidence 66%, R3 = True -> R5 = False: Confidence 66%

IV. IMPLEMENTATION

4.1 SELECTION OF DATASET

Here book dataset is used. It contain user id and item name field. On this dataset pre-processing is applied. So, it will generate user-item matrix. Here 50 x 50 size of data are used.

Figure 4: Book Dataset

Figure 5: Preprocessed Dataset

4.2 GENERATION OF FREQUENT ITEM DATASET

Now, 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. In following fig 6 shows that the result of given input.
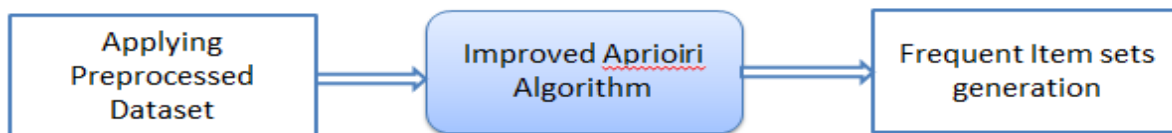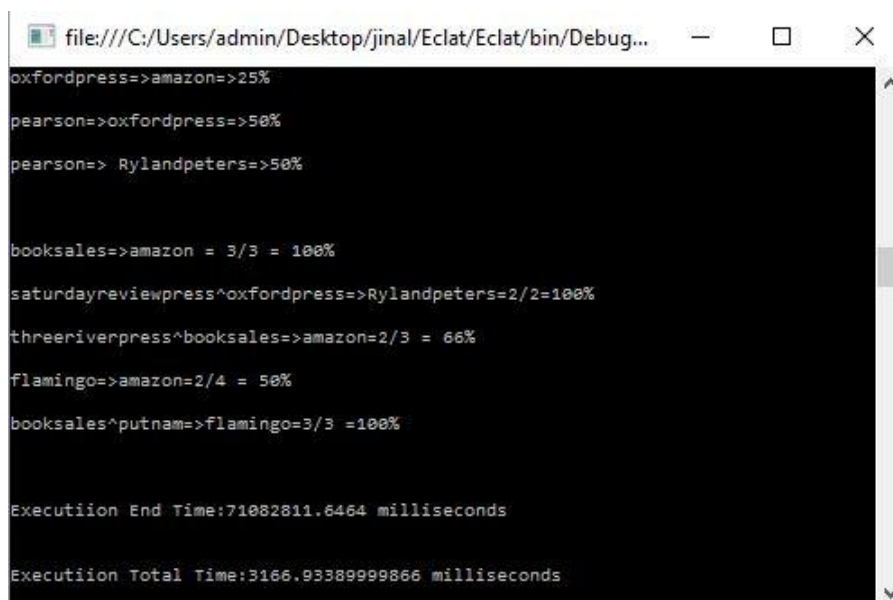
Figure 6: Frequent item set with minimum support

4.3 RULES GENERATION OF FREQUENT ITEMSET
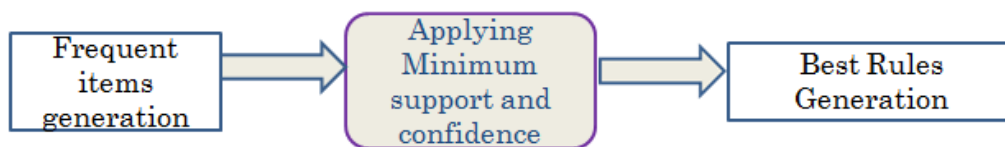
Figure 7: Rules Generation based on frequent item set

After the applying proposed algorithm association rules are generated. In figure 7 shows that the execution end time and total time. To find out the execution time following formula is used:

Execution Time = Ending Time – Starting Time

V. CONCLUSION

The three recommender systems have their advantages and disadvantages in performing their job. Most of the limitations in each one of the approaches can be complimented by the other. A good recommender system should be able to provide positive and relevant recommendations from time to time and also provide alternative recommendations to break the fatigue of the users seeing the same items in the recommendation list. In proposed method, we have select book dataset, after that preprocessing the dataset so, it generate user-item matrix. The association mining rules apply on different clustering groups after that apply proposed algorithm to recommend the best items to the user. This will result in a very efficient recommendation system with has its own intelligence to predict the best interest of the user and hence provide recommendations with high accuracy. Therefore, proposed algorithms and techniques reduce memory computational complexity eventually eliminate synchronization problems.

REFERENCES

1. J.A. Konstan, J. Riedl, "Recommender systems: from algorithms to user experience User Model User-Adapt Interact", 22 (2012), pp. 101-123 CrossRefView Record in Scopus.
2. C. Pan, W. Li, "Research paper recommendation with topic analysis", In Computer Design and Applications IEEE, 4 (2010) pp. V4-264.

3. Pu P, Chen L, Hu R., "A user-centric evaluation framework for recommender systems", In: Proceedings of the fifth ACM conference on Recommender Systems (RecSys'11), ACM, New York, NY, USA; 2011. p. 57–164.

4. Hu R, Pu P., "Potential acceptance issues of personality-ASED recommender systems", In: Proceedings of ACM conference on recommender systems (RecSys'09), New York City, NY, USA; October 2009. p. 22–5.

5. B. Pathak, R. Garfinkel, R. Gopal, R. Venkatesan, F. Yin, "Empirical analysis of the impact of recommender systems on sales", J Manage Inform Syst, 27 (2) (2010), pp. 159-188

6. D. Fox Harrell, "Phantasmal Media: An Approach to Imagination, Computation, and Expression"

7. Suresh Kumar Gorakala, "Building Recommendation Engines"

8. G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in Recommender Systems Handbook: A Complete Guide for Research Scientists and Practitioners. Springer, 2010.

9. S. Aciar, "Mining context information from consumers reviews," 2010.

10. Elgohary, H. Nomir, I. Sabek, M. Samir, M. Badawy, and N. A. Yousri, "Wiki-rec: A semantic-based recommendation system using wikipedia as an ontology," in Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on, 29 2010.

11. E. Peis, J. M. Morales-del Castillo, and J. A. Delgado-Lpez, "Semantic recommender systems. analysis of the state of the topic." Seoul, South Korea: Department of Computer Science, YonseiUniverisity, 2008.

12. S. Berkovsky, T. Kuflik, and F. Ricci, "Cross-domain mediation in collaborative filtering," in Proceedings of the 11th international conference on User Modeling, ser. UM '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 355–359

13. R. Kanawati and H. Karoui, "A p2p collaborative bibliography recommender system," in Proceedings of the 2009 Fourth International Conference on Internet and Web Applications and Services. Washington, DC, USA: IEEE Computer Society, 2009, pp. 90–96. [Online]. Available: http://portal.acm.org/citation.cfm?id=1585689.1586218

14. R. Baraglia, M. Mordacchini, P. Dazzi, and L. Ricci, "A p2p recommender system based on gossip overlays (prego)," in Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on, 2010, pp. 83 –90.

15. C.-Z. Yang, I.-X. Chen, and P.-J. Wu, "Cross-lingual news group recommendation using cluster-based cross-training." Taiwan: Computational Linguistics and Chinese Language Processing, 2008, pp. 41–60.