Comparative Analysis of Association Rule–Based Collaborative Filtering for Medical Decision Support and Clinical Recommendation Systems

Dr. Lukas Weber[1]*
Dr. Sofia Marinova[1]
Prof. Amina El-Sayed[1]
[1] University of Cape Town, Department of Biomedical Informatics and Clinical Data Analytics, Cape Town, South Africa

ABSTRACT

The World Wide Web have brought us an overabundant knowledge in varied fields and as a result of the data or information overloading, it is very arduous to find out related data. So, Recommendation System comes into existence. The main goal of this system is to recommend the best suitable items to the user or customer. The suggestions pertinent to decision making processes, like what things to obtain, which new music to listen to, which on-line latest news to search, or which image is best one from all. The advantages of recommendation system depend on efficiency of the system. The efficiency can be measured in terms of reliability, accuracy, flexibility. The main aim of the proposed system is to generate the rules based on the association mining and try to improve the accuracy of the system.

Keywords: *Recommendation system, collaborative filtering, content-based filtering, Hybrid filtering, Hybrid recommendation system.*

I. INTRODUCTION

Recommender systems main aim to help user's access and find user's needed info from giant collections, by automatically finding and implying products of possible user interest based on observed the user's preferences. Nowadays with the fast growing of internet, people try to search more and more things on internet due to this, there is an abundant amount of information available that results into the information overloading. So to overcome the problem, the propose system tries to recommend the best suitable items based on the association rules [Jiawei Han and MichelineKamber (1)].

II. RELATED WORK

Step 1:
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 sizes of data are used.

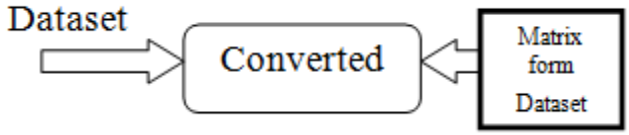

Fig 1: Matrix Conversion

In below table pre-process dataset is shown.

Step 2:

Table 1: Pre- Processed Dataset

User Id	Pearson	Famingo	Instru -ction book	Tech max
276727	1	0	0	0
276729	1	1	1	0
276738	1	0	0	1
276734	0	0	0	1

Frequent item-set generation

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 1.1 shows that the result of given input.

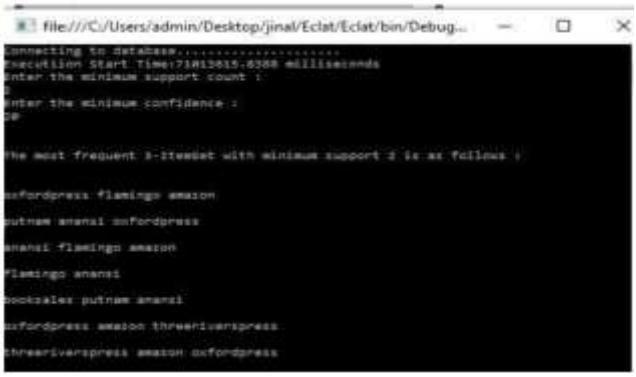

Fig 1.1: Frequent item-set with minimum support

Step 3:

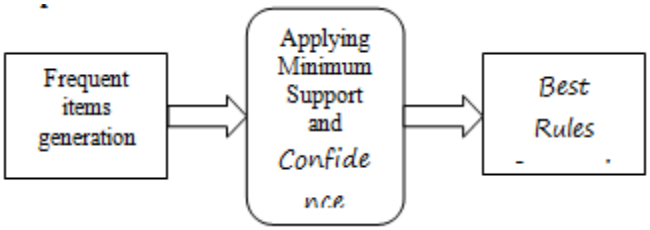

Fig 1.2: Matrix Conversion

Fig 1.3: Best Rule generation

Fig 1.4: Rules Generation based on frequent item set
After the applying ECLAT (Equivalent Class Transformation) algorithm association rules are generated. In figure 1.3 shows that the execution end time and total time. To find out the execution time this formula is used;
Execution time = ending time – starting time

Comparison of Eclat algorithm and Apriori algorithm;
Confidence = 20 and Support=2

Fig 1.4: Result of Eclat algorithm

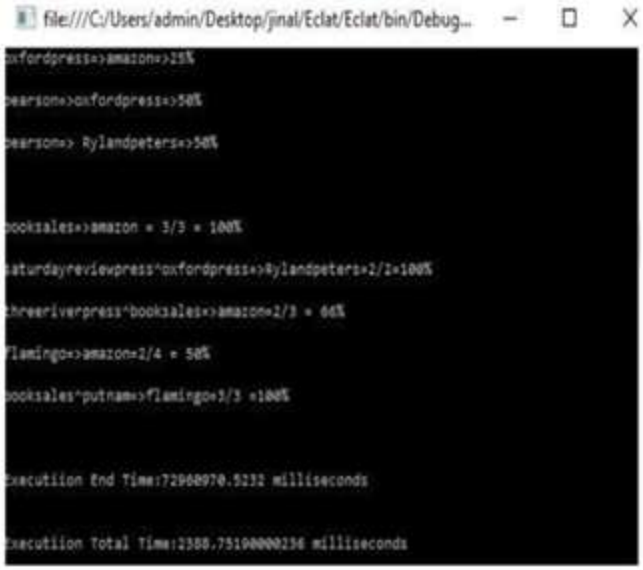

Fig 1.5: Count minimum support- confidence

Apriori
Apriori algorithm The Apriori algorithm developed is a great achievement in the history of mining association rules. It is by far the most well-known association rule algorithm. This technique uses the property that any subset of a large item set must be a large item set. Also, it is assumed that items within an item set are kept in lexicographic order. The fundamental differences of this algorithm from the AIS and SETM algorithms are the way of generating candidate item sets and the selection of candidate item sets for counting. As mentioned earlier, in both the AIS and SETM algorithms, the common item sets between large item sets of the previous pass and items of a transaction are obtained. These common item sets are extended with other individual items in the transaction to generate candidate item sets. However, those individual items may not be large

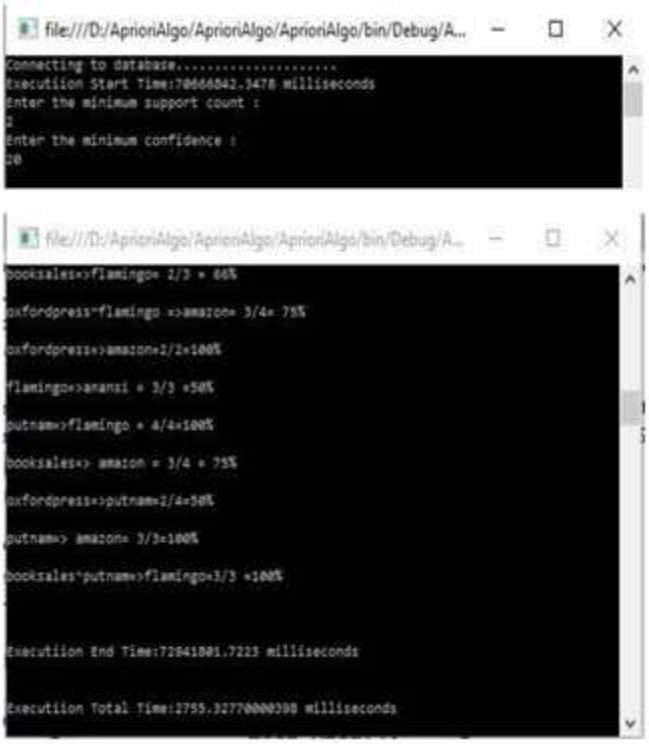

Fig 1.6: Result of Apriori

Comparison with ApriorivsEclat with minimum confidence =40 and support=2;
Eclat

An Eclat is a depth-first search algorithm which refers set intersection. Vertical database layout is referred for illustration. All transactions are not listed explicitly but each item is stored together with its cover and uses the intersection based approach to compute the support of an item set. The support of an item set A can be easily calculated by cover's intersection of any two subsets Y, Z ⊆ A, such that Y U Z = A. Candidate generation of Eclat uses only the join step of Apriori ,

Following figure present the result of Eclat algorithm where as minimum confidence is 40 and support is 2

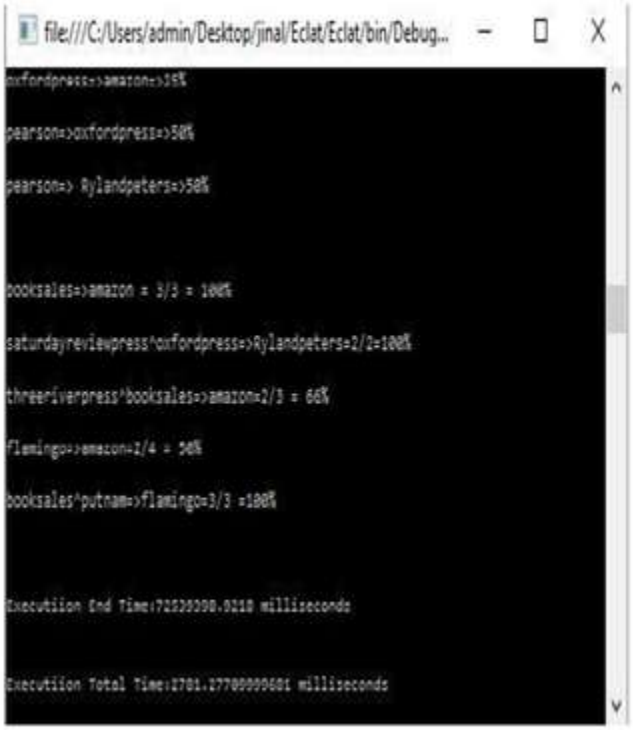

Fig 1.7: Result of Eclat for min support 2 and confidence 40

Apriori

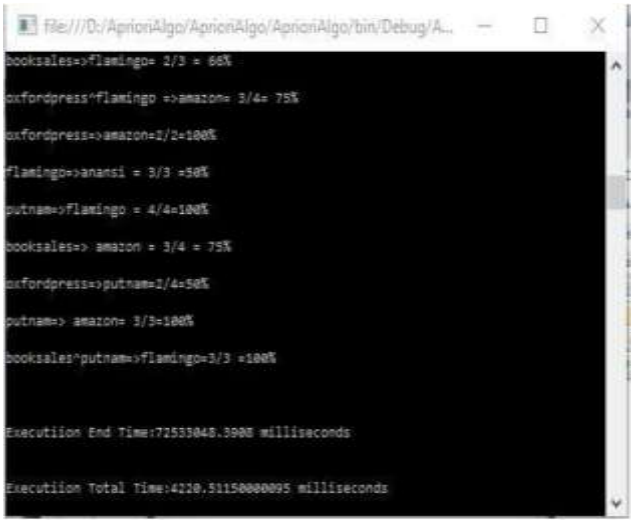

Fig 1.8:Result of Apriori for min support 2 and confidence 40

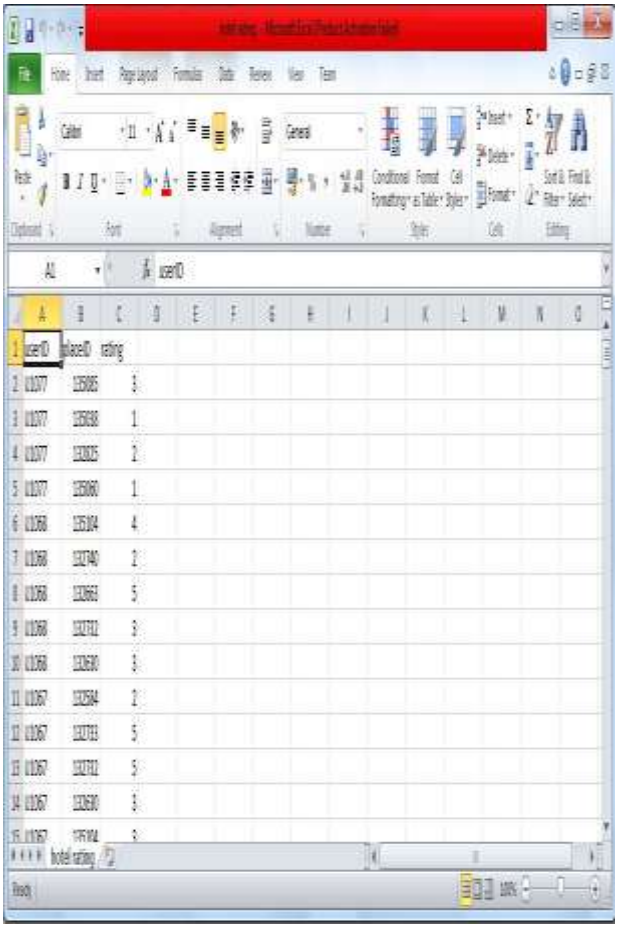

Fig 1.9 Data Simple Set

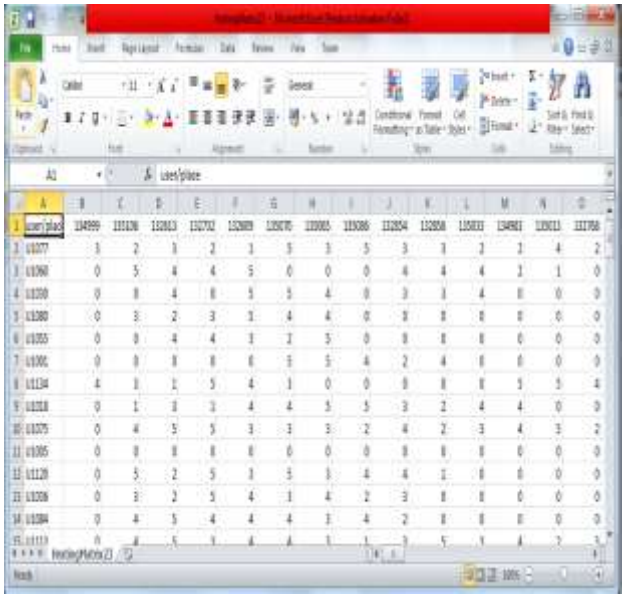

Fig 1.10 Data Set

Table 2: Comparison between Apriori and Eclat

confidence	Apriori	Eclat
20	2755	2388
30	3277	2514
40	4220	2781
50	5023	2615
75	6621	3066

Result in graph form;

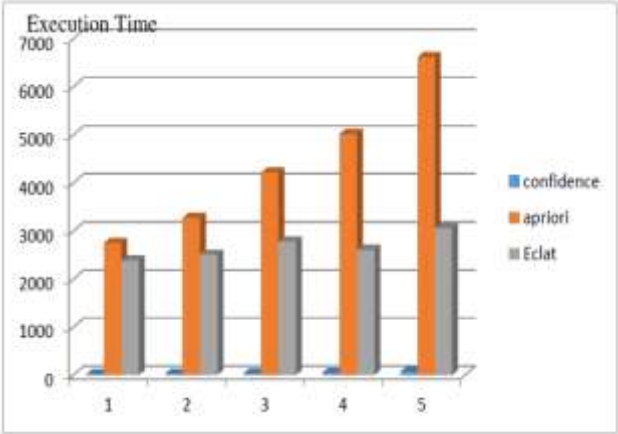

Fig 1.11: Comparison graph between Apriori & Eclat

III. CONCLUSION

Recommendation systems help to recommend the best suitable items to the users. The proposed system deals with the improving the accuracy and sparsity problem. The proposed system applies association mining for better rules generation. Initially the standard dataset is taken for evaluation by the system. The dataset is pre-processed and the pre-processed data is given to the Eclat algorithm which is used for finding the best rules for better recommendation. The proposed algorithm will also try to improve the sparsity.

REFERENCES

1. *Mohammad HamidiEsfahani, Farid Khosh Alhan - "New Hybrid Recommendation System Based On C-Means Clustering Method " -IEEE-2013 5th conference of information Knowledge.*
2. *Qi Wang, Wei Cao and Yun Liu– "A Novel Clustering Based Collaborative Filtering Recommendation System Algorithm " –Springer 2014*
3. *Ahmed Mohammed K. Alsalama – "A Hybrid Recommendation System Based On Association Rules" – ISSR-2014*
4. *Hazem Hajj, Wassim El-Hajj, Lama Nachman – "A Hybrid Approach with Collaborative Filtering for Recommender Systems" –IEEE 2013.*
5. *Haojun Sun –"A New Item Clustering based Collaborative Filtering Approach." –IEEE 2012.*
6. *R.Srinivasa Raju I.Kali Pradeep I.Bhagyasri "Recommender Systems for E-commerce: Novel Parameters and Issues".-IJARCSSE 2013*
7. *Robin Burke "Hybrid Recommender Systems: Survey and Experiments"–ACM 2002*
8. *UjwalaH.Wanaskar, SheetalR.Vij, Debajyoti Mukhopadhyay "A Hybrid Web Recommendation System based on the Improved Association Rule Mining Algorithm".-Arxiv journal 2013*
9. *S.Vijayarani,P.Sathya – "Mining Frequent Item Sets over Data Streams using Éclat Algorithm".*
10. *Soumi Ghosh, Sanjay Kumar - "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms".-(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, 2013*
11. *Sunita B. Aher, L.M.R.J. Lobo - Combination of machine learning algorithms for recommendation of*

courses in E-Learning System based on historical data-Springer.

12. *Sunita B. Aher - Data preparation strategy in e-learning systemusing association rule algorithm, International Journal of Computer Applications 41 (3) (2012) 35-40.*
13. *Ismail Sengor Altingovde, Özlem Nurcan Subakan- Cluster searching strategies for collaborative recommendation systems-Elsevier 2012*
14. *Weiyang Lin Sergio A. Alvarez- Collaborative Recommendation via Adaptive Association Rule Mining*
15. *Mustansar Ali Ghazanfar - Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems- Elsevier-2013*
16. *Roman Neruda- Clustering Based Classification in Data Mining Method Recommendation –IEEE 2013*
17. *Addressing cold start problem in recommender systems using association rules and clustering technique-IEEE 2013.*