
Review of Salam and Irshad (SI) Algorithm in Different Fields

Aamna Irshad¹ and Irshad Ullah²

¹National University of Sciences and Technology Islamabad, Pakistan, Mobile :+92 300 936 3476
E-mail: aamna.saleem@s3h.nust.edu.pk

²Education Department, Government of Khyber Pakhtunkhwa, Pakistan, Mobile: +92 300 936 3476
E-mail: irshadullah79@gmail.com

Abstract: Association rule mining is a subfield of data mining. Some of the famous algorithms used for association rule mining are Apriority, frequent pattern mining. SI algorithm or SI algorithm is one among them. In this research a review was performed to analyze the use of SI algorithm in different field of life. The objective of this research is to highlight different uses of the algorithm in the said field. In methodology the papers and work containing the work by SI algorithm was analyze. Conclusion was made on the basis of analysis and discussion. Discussion was made about the efficient use of SI algorithm. Literature from 2008 to 2019 containing the work in the form of presentations and published was analyzed.

Keywords: Association rule mining, Apriority, SI, Data Mining, Frequent item set.

Introduction

AIS algorithm was the first used and introduced for association rule mining problem in the era of 90's. This technique was used to find out association rules in market basket data [1]. This was the first published algorithm for the purpose to find out association between set of items and give the frequent item sets as a result. The process of this algorithm for the production of new item sets, it's make many passes on the entire databases. The process is continued till the production of all large data sets [2]. Apriori algorithm is also used for the purpose to produce the large data set but not necessary to scan the entire database in all passes [2]. Same like that SI algorithm scan the database to produce the frequent item sets and the large item sets but to scan the whole database in all passes is not necessary. This algorithm gives a new way to generate the largest item sets [3]. Instead of generating the largest item sets on the basis of similarity this technique used dissimilarity to produce the frequent item sets. The algorithm gives a new dimension for the generation of large frequent item sets. This work was presented and submitted for the first time in 2008 for the award of MS degree in Computer Sciences at City University Peshawar [3]. This work was also presented and published at 4th international Conference of Statistical Sciences held at University of Gujrat. The work was published and appreciated by the organizers of the conference [4].

SI algorithm was used latterly for different purpose like it is used in the form of presentation for the work that how computer can help in Engineering Data analysis. The algorithm was used to generate largest item sets in the field of Engineering data. This work was presented in a conference in 2011 arranged by Islamic countries society of statistical sciences held and arranged at National Conference of Business Administration and Economics Lahore, Pakistan [5]. The work was presented and it is showed that the algorithm can be used for the purpose of large item sets by using a data sets used in previous research of Irshad Ullah submitted at City University Peshawar. On the basis of generated data and frequent itemset suggestions were made [3]. This was the 8th international conference on recent advances in statistics. The same algorithm was used for the analysis of student's data. This was also presented in the form of presentation [5]. The student's data was analyzed by this algorithm [6]. Again frequent item sets produced were not only on the basis of dissimilarity but was compared with Apriori algorithm [3] for the purpose and accuracy. On the basis of results produced suggestions and conclusion were made and drawn. Future work was also proposed in the last section.

In this conference [7] in other presentation was made for the purpose of population study. The data was provided to the algorithm for the production of frequent item sets. To make an accurate analysis and reliable results production size of the data base was

changed. By making change in size of data base the results and the largest frequent item sets were the same. For more accuracy purpose input value was also changed, but it was seen that the output produced by the algorithm remain same. It produces the same largest frequent item sets like with the previous input value which is the requirement of the algorithm. But it does not make any effect on the results. Such type of experiments makes at sure that SI algorithm [3] is a reliable algorithm for the production of frequent and large item sets. Here the researcher also shows that data mining algorithm can be used for many purpose not only for the market basket data but even for the data from different fields. Researcher use at in different areas of life. This was done for the purpose to generate some fruitful and meaningful patterns with the help of data mining algorithm.

Further researcher uses at for the purpose of logical reasoning. Data was provided to the algorithm in different stages. Initially the size of database was small. In the next step the size of database was made large. The size of database was changed gradually. The produced results and frequent item sets remain unchanged. The research was accepted and presented in 8th International Conference held at NCBA&E in 2011[8].

In 2012 the researcher arranges and organizes the data. The data was executed by giving to SI algorithm. This research was accepted and presented at 9th International Conference on Statistical Sciences. The conference was arranged at NCBA&E. Research was presented in the form of three presentations “Data Quality enhancement using data mining approach”, “Data Quality Management using Data Mining Technique” and” Computer Science algorithm in space science”. The research suggested and draws the conclusions on the basis of experiments in the said fields. Different suggestions were made that how these fields can get benefit from the results of the data. It is highlighted that how frequent item sets can be generated in these fields. How experts can arrange and organize their data. Further how the people from these fields can get benefit of this data. It is suggested that by arranging data in such large item sets it is possible to use it for different purpose [9-11]. This research shows that data mining algorithms can be used for different purpose by arranging the data in proper format. It also infers from the research that using of these algorithms make it possible for the experts to use at for their required purpose.

Researcher use at in 2013 in the form of presentation in 10th International Conference on

Statistical Sciences held at Lahore Lead University. Presentations were made on “Data Mining algorithms in the field of Cancer”. Data was analyzed and results were produced. The produced results were presented in the front of experts. Conclusion was drawn. The results show that these algorithms can be used in the field of cancer or for the data in cancer field an effective way.

Data Extraction by Computer algorithms another presentation accepted was made in 2014 in International Conference held at National University of Sciences and Technology. SI algorithm was used for the generation of results. Data was given and provided to the algorithm gradually. Results produced were discussed and presented in front of experts. Suggestions were made and conclusion was drawn on the basis of the results.

In other presentation was made that how computer can help the special people an International Conference at Superior University Lahore in 2014. Outcome was discussed and suggestions were made for the welfare of special people [14].

Conclusion

Review was made on the research work carried out through SI algorithm. Review shows that research carried out through this algorithm can be used for the betterment of the society. This is the work done from 2008 to 2018. Discussion made shows that SI algorithm can be used in a variety of ways. It shows that SI algorithm is very efficient in different areas. By using suggestions of the research expert can use it for their purpose. In future this work can be extended in different ways. SI algorithm can have used in a new field and in new ways. Further algorithm if use in a better way will be useful for the stakeholders of the study.

References

1. Agrawal, R. Imienski, T. and Swami, N. A. (1993). Association Rules Between Sets of Items in Large Databases, Washington, D.C.: Proceedings of the ACM SIGMOD International Conference on Management of Data.
2. Agrawal, R. and Srikant, R. (1994) Fast Algorithms for Mining Association Rules in Large Databases, Santiago.: Proceedings of the Twentieth International Conference on Very Large Databases.
3. Ullah, I. (2008). Dissimilarity based mining for finding frequent item set. Thesis submitted for the partial fulfillment of Master degree in

-
- Computer Science. Peshawar. City University of Science and Technology.
4. Salam, A. Rahman, S. and Ullah, I. (2008). Dissimilarity based Mining for finding frequent itemset. Gujrat.: Proceedings of the 4th international Conference, University of Gujrat.
 5. Ullah, I. (2011). Data analysis in the field of Engineering. Lahore.: Presentation 8th International Conference of Statistical Sciences. NCBA&E.
 6. Ullah, I. (2011). Data Mining and Students analysis. Lahore.: Presentation 8th International Conference of Statistical Sciences. NCBA&E.
 7. Ullah, I. (2011). Computer Science and Population Study, Presentation 8th International Conference of Statistical Sciences, NCBA&E, Lahore.
 8. Ullah, I. (2011). Logical Reasoning and Data Mining algorithm, Presentation 8th International Conference of Statistical Sciences, NCBA&E, Lahore.
 9. Ullah, I. (2012). Data Quality Enhancement using Data Mining Approach, Presentation 9th International Conference on Statistical Sciences, NCBA&E, ISOSS.
 10. Ullah, I. (2012). Data Quality Management using Data Mining Techniques, 9th International Conference on Statistical Sciences, NCBA&E, ISOSS.
 11. Ullah, I. (2012). Computer Science algorithms in Space Science, 9th International Conference on Statistical Sciences, NCBA & E, ISOSS.
 12. Ullah, I. (2013). Data Mining algorithms in the field of cancer. Presentation at 10th International Conference on Statistical Sciences, Lahore Leads University.
 13. Ullah, I. (2014). Data Extraction by Computer Algorithm. Presentation, Research Workshop. National University of Sciences and Technology, Islamabad,
 14. Ullah, I. (2014). How Computer can help special People, Presentation at International Conference on Management Research. Superior University, Lahore.
 15. Ullah, I. (2015). Computer Systems in terms of fulfilling local and Global Challenges. Doctoral Thesis submitted to Concordia College and University Delaware.
 16. Ullah, I. and Khan, S. A.(2016). Computer Science algorithms in Education. Presentation in 14th International Conference on Statistical Sciences at Jinnah Sindh Medical University Karachi.
 17. Ullah, I. and Khan, S. A. (2017). Classroom Arrangement through Computer: A New Approach Global Journal of Business and Social Science Review 5 (1) 18 – 22.
 18. Ullah, I. and Khan, S. A.(2018).Educational Data Interpretation by Algorithms Scientific Journal “Herald National Academy of Managerial Staff of Culture and Arts”.