iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** IV **Month of publication:** April 2016

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Performance of Novel Based Fuzzy Association Rules to Reduced Computational Data Sets by Row Counts

Rohit Miri[1], Priyanka Tripathi[2], Vinay Soni[3], Asha Miri[4]
[1,3,4]Department of Computer Science & Engineering, [2]Department of Computer Technology
[1,2,3]Dr. C.V. Raman University, Kota, Bilaspur, India (C.G.)
NITTTR, Bhopal (M.P.)
[4]Govt. Polytechnic College, Bilaspur, India (C.G.)

Abstract*—It is a big challenge to predict the result of large data sets quickly. To overcome this problem we have proposed a novel algorithm to reduce the large data sets to smaller one without affecting the result of fuzzy association rule by Row count algorithm. Row counts and column counts play a vital role to output the result of association rule. It behaves like a two way filter of the data. Further the comparison is done among Different fuzzy association rule ,A fuzzy AprioriTid mining algorithm with reduced computational time [9] and CPDA Based Fuzzy association rules for learning achievement mining, 2009 international conference on machine learning and computing IPCSIT [15] (mention specific names) and proposed algorithm. The performance of each is shown using graphs and data sets.

Keywords*—Fuzzy Logic,Data Mining, fuzzy association rule.

I. INTRODUCTION

Fuzzy logic and data mining are the two most important techniques to decide logistics in this era. Fuzzy logic uses linguistic variables to provide most appropriate solution to the problems having uncertainty in it. Data mining is used to identify the pattern from large data sets, by applying association mining algorithm to find the relationships among the item sets. This paper finds out the range of linguistic variable by using suitable statistical measures viz standard deviation and mean in place of assuming its values. Secondly fuzzy logic contains the distinct values by applying membership function.

II. LITERATURE REVIEW

Different approaches have been used by the researchers to reduce the time complexity of fuzzy association rule.

The item sets that are frequently present in particular transaction id are chooses by Mohammed Al-Maoleg1[1] . Its algorithm works on low support .3 and reduces the time complexity of the program very easily. Zhiyong ma [4] et al converts all the item sets into Boolean matrix by using CP tree method and reduces the time for the task. ArpnaShrivastava [5] et al have used the codes for all the items and remove the duplication by using data cleansing technique. This is also most efficient as compared to simple Apriori algorithm.

Another approach reduces the operational time carried out by Apriori algorithm by using artificial Bee colony optimization method (FABCO) is given by K. Sathesh Kumar and M. Hemalatha [3] .Mehmet Kaya et al [24] has worked out an efficient algorithm by carrind out mining fuzzy clustering algorithm (CURE). They found out the centroid by CURE for triangular membership function , so that they can range the fuzzy membership method correctly and also reduces the computational time.

Agrawal and his co-worker carried out some mining algorithm based on the large data sets, which also find association mining rule [9]. These algorithms break the mining steps into two phases. In the first phase candidate of item sets are obtained and counted by scanning the transactions. The number of item set must support the minimum pre-defined threshold value called minimum support. Then later we make the pair of item sets and apply the association rule for getting the required output. Srikant and Agrawal also proposed partitioned based mining association algorithm. Cai at al proposed weighted mining rule of data sets [19]. Yue et al, extended the fuzzy concept based on vectors [22].Most of them are find out the range of triangular fuzzy membership function directly , means they assumed the range of linguistic variable. But on my paper We have find out the range of linguistic variable by using mean and standard deviation.

Mining association rule was also performed by the [21] T. P. Hong, C. S. Kuo, and S. C. Chi, “Mining association rules from quantitative data,” Intelligent Data Analysis. The aim of his research is digging out the essential or useful item from very large data

set. Authors want to improve the data mining algorithm. We used [9] Tzung-Pei Honga,Chan-Sheng Kuob,Shyue-Liang Wangc "A fuzzy AprioriTid mining algorithm with reduced computational time and [15] JrShianchen hung liehchou, ChingHsue Cheng, jen-ya Wang CPDA Based Fuzzy association rules for learning achievement mining, 2009 international conference on machine learning and computing IPCSIT vol3 (2011) IACSIT PRESS SINGAPUR for comparison with proposed algorithm. They have done on fuzzy mining association rule to reduce the computational time. They all used the mining association rule for doing the task, the TRApriori mining association technique is used from the paper [16] E Ramaraj, K Rameshkumar, N Venkatesan"A better performed transaction Reduction algorithm for mining frequent item set from large voluminous database .

III. PROPOSED ALGORITHM

STEP 1: Find out the ranges of linguistic variable for triangular fuzzy membership function of data sets by using standard deviation and mean.

STEP 2 Create or build triangular membership function and define or assume the range of linguistic variable with the help of standard deviation and mean of given data sets i.e low middle and high.

STEP 3: Find out the fuzzy values of given data sets

According to step2, the degree of membership for each data set is calculated.

In this, we have used triangular membership function because of its easiness and computational efficiency. We can also use Gaussian membership function.

We can categories it's as Low, Middle and high. Thus we have used three fuzzy membership values produced for each attributes.

STEP4: Count its row values and column values of fuzzy values data sets. i.e

Row Values: it is a valid value present in the row accept zero.

Column Values: it is a valid value present in the column accept zero.

STEP5: Select the row which has maximum number of row values or contains at least 60% of count value.

Count Value: Row values/ Total no of items

STEP6: Support values of the output datasets should be greater than the remaining data sets.

Select the attributes which has a sum of the fuzzy values is > then the Greatest value of any one of the selected attributes .

STEP7: Select only those values that satisfy the step 5 and 6. This data set is important or essential data sets used for building fuzzy mining association rules. Discard the remaining data sets because they are not important for mining association rule.

STEP8: Select the column which has more than 60% of count value. This column values plays an important role in making data mining association rules.

STEP9: Apply data mining fuzzy association rule on the above important or useful data sets and find the result of important data sets.

IV. DATA ANALYSIS

In data analysis part we compared the various algorithm or methodology to our latest proposed algorithm Titled "Novel Based Fuzzy Association Rules for Reduced Computational Data Sets by Row Counts ". We compared on the basis of space or data complexity, output and performance of each published paper or various proposed technology. We applied same algorithm to all the published paper and found that our proposed algorithm worked well in all the aspects regarding data complexity, providing same output and performance. It also reduces the complexity between candidate item sets. Firstly we mentioned or explained the latest proposed methodology and then later we applied our methodology to different published paper. Lastly we concluded the output or result of all the analysis done. The past or published paper data analysis is done by no of lines takes to perform the task. The below mentioned data analysis is based on the single data or a variable required the space or memory to perform the particular task. The below mentioned data analysis is based on the single data or a variable required the space or memory to perform the particular task. We have calculated the no of single data used or required to complete our task by both the algorithm. First one is Published paper used algorithm and second one is by our proposed algorithm "Novel Algorithm To Reduced Computational Data Sets for Fuzzy Association Rule (By Row Counts) ". Later we have mentioned the performance of the entire or all algorithms by comparing with our latest proposed algorithm by using row count technology. The result or analysis are as follows:-

Volume 4 Issue IV, April 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Result Of Data Analysis														
Existing Techniques with their no. of variables or Data Required for calculating the Candidate Item Sets.								Proposed Algorithm with their no. of variables or Data Required for calculating the candidate Item sets.					Performance
S. N	Journal Name	Paper Title	Technology Uses	C1	C2	C3	Total	C1	C2	C3	Total	Output	
1	2009 International Conference on Machine Learning and Computing IPCSIT vol.3 (2011) © (2011) IACSIT Press, Singapore	CPDA Based Fuzzy Association Rules for Learning Achievement Mining	CPDA	20	42	48	110	10	40	50	100	Same	Both the approaches give the same output but our approaches takes less space or reduce the data complexity.
2	ScienceDirect IEEE International Conference on Fuzzy Systems Volume 5, Issue 1,	A fuzzy AprioriTid mining algorithm with reduced computational time	AprioriTid	65	102	-	167	8	46	-	54	Same	Both the approaches give the same output but our approaches takes less space or reduce the data complexity more efficiently.
3	International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3679-3683	TRApriori Classification Based Algorithm by Fuzzy Techniques to Reduced Time Complexity	TRApriori	22	48	51	121	14	51	12	77	Different	Both the approaches reduce the data complexity but novel algorithm by row counts seems to be more efficient. Both the approaches provide the different output.
4	International Journal for Research in Applied Science & Engineering Technology (IJRASET)-2015	Exploration of Novel Algorithm for Reduced Computational Time by Using Fuzzy Classification Technique in Data Mining	Novel Algorithm	22	42	27	91	8	38	-	46	Different	Both the approaches reduce the data complexity but novel algorithm by row counts seems to be more efficient. Both the approaches provide the different output.

Table no 1.1 Data analysis of algorithms

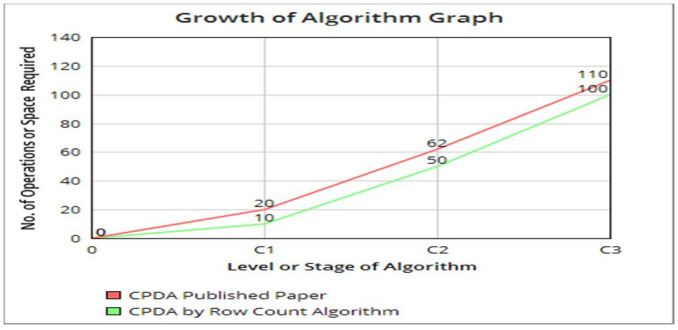


Figure No 1.1 CPDA [15] Vs Novel algorithm by row count

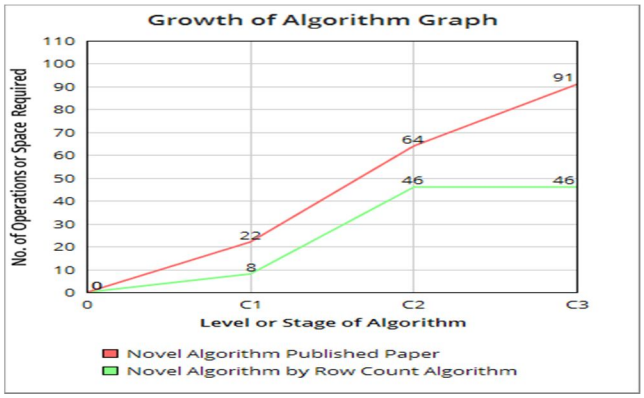


Figure No 1.2 Novel algorithm [26] Vs Novel algorithm by row count(Proposed)

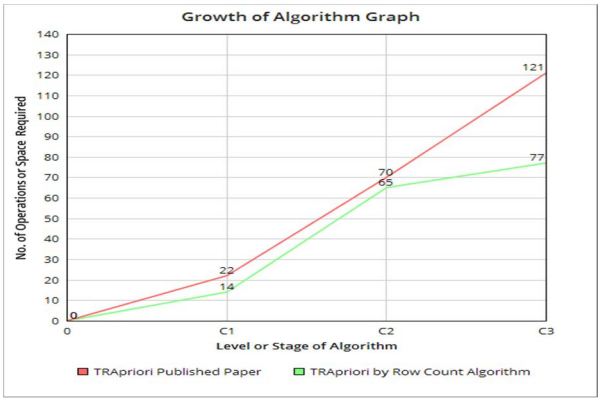


Figure No 1.3 TRApriorialgorithm[25] Vs Novel algorithm by Row Count

Volume 4 Issue IV, April 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

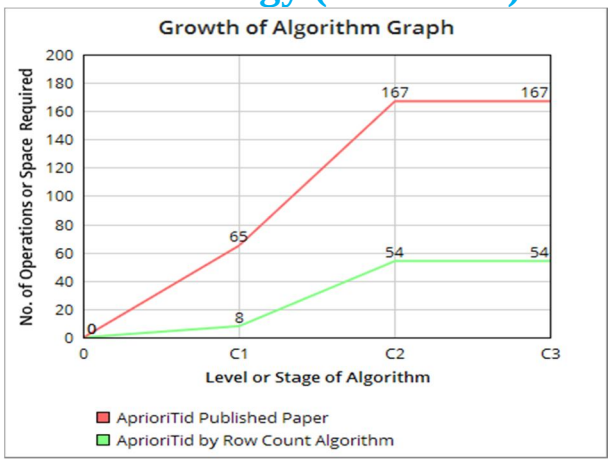


Figure No 1.4AprioriTid algorithm [9] Vs Novel algorithm by row count

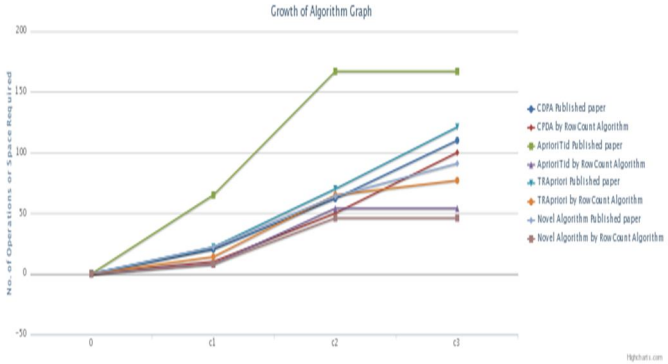


Figure No 1.5 Performance Graph of all the Technology/Algorithm

These graphs show that proposed algorithm required less space to store the data. The performance is better as compared to all the above mentioned algorithms in terms of reduced space complexity .The Proposed algorithm also get the important data sets from large data sets. This will perform the fuzzy mining association in fast way and also same for prediction of past data sets.

V. CONCLUSION

Performance of the proposed algorithm is better than the Existing Fuzzy Association Rule regarding Space Complexity , Time Complexity and Prediction of Past Data Sets.

It a simple Novel algorithm to find the fuzzy association rule for large data sets.

Prediction of results is more accurate for any data sets.

It makes mining algorithm simple due to smaller data sets. i.e. Novel algorithm reduces the large data sets into smaller data sets. We apply mining association rule after this task. Further for future work, we predict the result of association rule without applying the Data Mining association algorithm based on the fuzzy classification or clustering technique.

REFERENCES

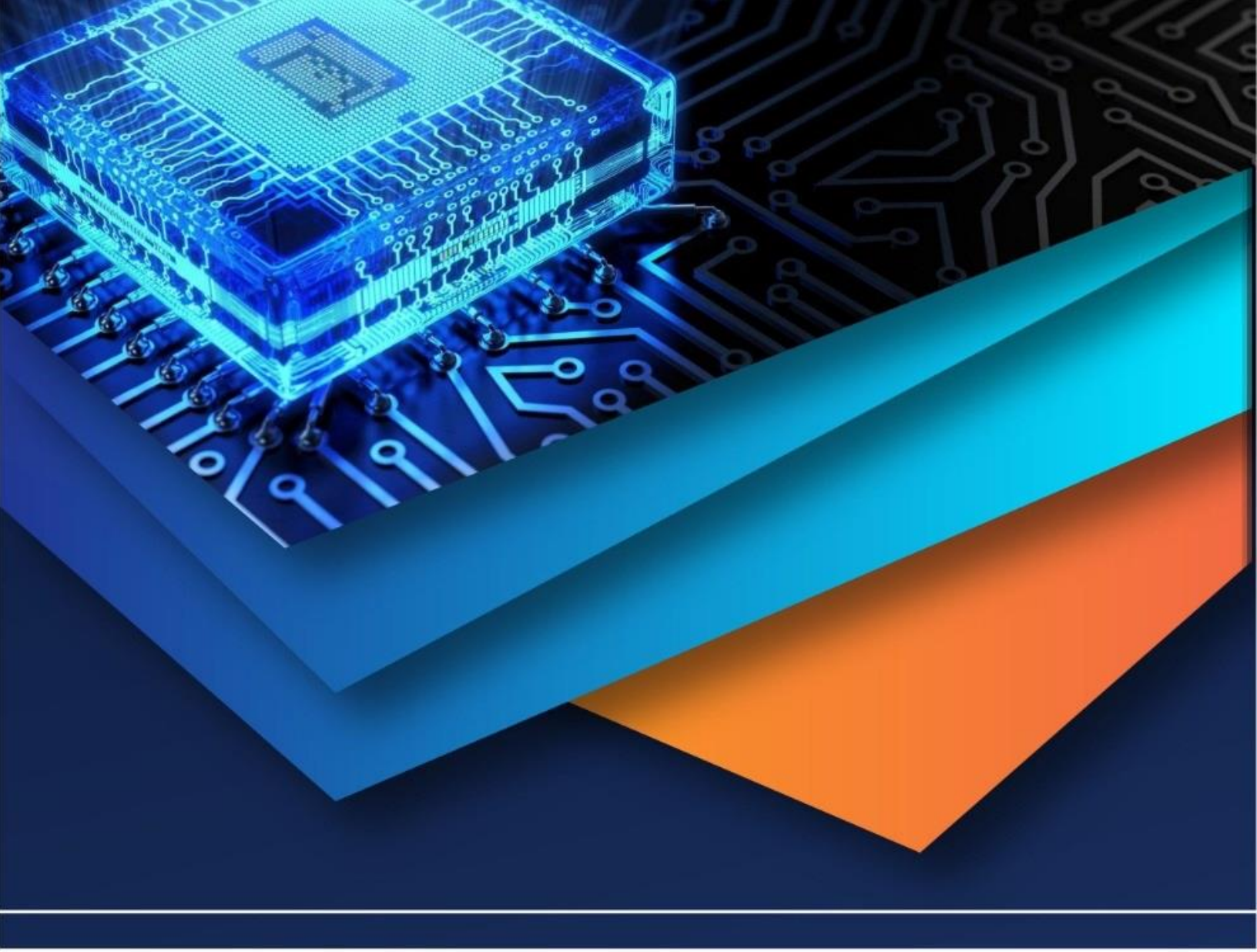
[1] Mohammed Al-Maolegi ,Bassamarkok “An Improved Apriorialgorithm For Association Rules” International Journal On Natural Language Computing Volume 3 , No-1 , Febuary 2014.

[2] N S Nithya and K Duraiswami “Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface” vol 39 part 1 ,febuary 2014 , pp 39-52 .

[3] K Sathesh Kumar and M Hemlatha “An innovative potential on rule optimization using fuzzy artificial bee colony” Research journal of applied science , engineering and Technology.. volume 2627-2633 , ISSN:2040-7459, E-issn:2040-7467, April 05 2014.

[4] zhiyong MA , QingyingQiu and PeieenFeng “CPM algorithm for mining associaotion rules from databases of engineering design+n instances.” Journal of information science and engineering vol30 , 463-481 (2014)

[5] Arpnashrivastava , R . C Jain and A K shrivatava "Generating 3rd level association rule using fast aprioriimplementation."British journal of mathematics and computer science. Volume 4(2):241-251, 2014
[6] Neelukhare ,neeruadlakha, K R pardasani "An algorithm for mining multidentional fuzzy association rule" international journal of computer science and information security vol 5, no-1 ,2009
[7] Thomas sudkomp "Refinement of temporal constraints in fuzzy association " volume 48(2008) 14 June 2007
[8] Tiberiu Ban , Fuzzy computing for complexity level of evaluation tests. StudiaUniv, Babes-Bolai, Informatics, Volume LVIII number 1, 2013
[9] Tzung-Pei Honga,Chan-Sheng Kuob,Shyue-Liang Wangc "A fuzzy AprioriTid mining algorithm with reduced computational time" Volume 5, Issue 1, December 2004, Pages 1–1
[10] Tzung-pei-Hang, Guo-Cheng lan, Yi-Hsinlin and Shing-tai pan. An effective gradual data reduction strategy for fuzzy item set mining international journal of fuzzy systems ,vol 15 no2 June 2013.
[11] Usha Rani, R VijayaPrakkash, Dr A. Gavardhan,Mining Multi-level association Rules using fuzzy, International journal of emerging technology and advanced engineering. Volume 3, issue 8, August 2013.
[12] RuchiBhargava ,Shrikant lade , Effective positive negative association rule mining using improved frequent pattern tree. International journal of advanced research in computer science and software engineering.Volume3 , issue 4 April 2013 .
[13] DharmendraKumar , Deepak Bhardwaj rise of data mining: Current and Future application areas. IJCSI international journal of computer science issue vol 8 issue 5 No1 September 2011.
[14] KanuPatel ,Vatsalshal , jitendrapatel , jayna donga comparision of various association rule mining algorithm on frequent item sets. International journal of advanced research in computer Science and software engineering Vol-3 issue II Nov 2013 .
[15] JrShianchen hung liehchou, ChingHsue Cheng, jen-ya Wang CPDA Based Fuzzy association rules for learning achievement mining, 2009 international conference on machine learning and computing IPCSIT vol3 (2011) IACSIT PRESS SINGAPUR.
[16] E Ramaraj, K Rameshkumar, N Venkatesan"A better performed transaction Reduction algorithm for mining frequent item set from large voluminous database"Volume 5, Issue 1, December 2004, Pages 1–10
[17] R. Agrawal, T. Imielinksi and A. Swami, "Mining association rules between sets Submitted manuscript 24 of items in large database," The 1993 ACM SIGMOD Conference, Washington DC,USA, 1993.
[18] R. Agrawal, T. Imielinksi and A. Swami, "Database mining: a performance perspective," IEEE Transactions on Knowledge and Data Engineering, Vol. 5, No.6, 1993, pp. 914-925.
[19] C. H. Cai, W. C. Fu, C. H. Cheng and W. W. Kwong, "Mining association rules with weighted items," The International Database Engineering and Applications Symposium, 1998, pp. 68-77.
[20] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," Fuzzy Sets and Systems, Vol. 84, 1996, pp. 33-47.
[21] T. P. Hong, C. S. Kuo, and S. C. Chi, "Mining association rules from quantitative data," Intelligent Data Analysis, Vol. 3, No. 5, 1999, pp. 363-376. A. Kandel, Fuzzy Expert Systems, CRC Press, Boca Raton, 1992, pp.8-19. Submitted manuscript 26
[22] S. Yue, E. Tsang, D. Yeung and D. Shi, "Mining fuzzy association rules with weighted items," The IEEE International Conference on Systems, Man and Cybernetics, 2000, pp. 1906-1911.
[23] L. A. Zadeh, "Fuzzy sets," Information and Control, Vol. 8, No. 3, 1965, pp. 338-353.
[24] Mehmet Kaya1,*, Reda Alhajj1,**, Faruk Polat2, and Ahmet Arslan3 "Efficient Automated MiningofFuzzyAssociation Rules "DEXA 2002, LNCS 2453, pp. 133-142, 2002.
[25] Rohit Miri , Priyanka Tripathi , S R Tandan "TRApriori Classification Based Algorithm by Fuzzy Techniques to Reduced Time Complexity" International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3679-3683.
[26] Rohit Miri , Priyanka Tripathi , S R Tandan "Novel Algorithm for Reduced Computational Data by Using Fuzzy Classification and Data Mining Techniques"ACM Digital Library ISBN: 978-1-4503-3216-3 ,Nov 2014



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY