INTERNATIONAL JOURNAL OF ADVANCED INNOVATIVE TECHNOLOGY IN ENGINEERING

Published by Global Advanced Research Publication House Journal

Home page: www.ijaite.co.in

Association Rule Mining Using Genetic Algorithm

[1]Devyani Parate, [2]Bharti Shende, [3]Tinal Ganeshkar, [4]Pratiksha Meshram, [5]Megha Gaikwad, [6]Tejasvi Moon, [7]Prof. Vaishali Gedam

[1,2,3,4,5,6,7]Department of Computer science and Engineering, Nagpur Institute of Technology, Nagpur, Maharashtra, India

[1]dparate.2001@gmail.com, [2]shendebharti08@gmail.com, [3]tinalganeshkar52010@gmail.com, 4meshpratiksha008@gmail.com, [5]meghagaikwad173@gmail.com, [6]moonteju11@gmail.com

Article History

Received on: 10 April 2022

Revised on: 28 April 2022

Accepted on: 18 May 2022

Keywords: Association Rule Mining, Data Mining, Transactional Dataset

e-ISSN: 2455-6491

Production and hosted by
www.garph.org

ABSTRACT

Association Rule Mining is the main curiosity area for plenty of researchers for many years. It is the backbone of data mining. Relationships are discovered among different items in the Database. In this paper proposed IMLMS-GA association rule mining based on min-max algorithm and MLMS formula. In this method we used a multi-level multiple support of data table as 0 and 1. The divided process reduces the scanning time of database. The proposed algorithm is a combination of MLMS and min-max algorithm. Support length key is a vector value given by the transaction data set.

1. INTRODUCTION

Data mining is the step-by-step study and scrutiny of the KDD process (Knowledge Discovery and Data Mining). It is the process to extract exciting and useful (understood, formerly unidentified and constructive) information or patterns from mega information repositories such as data of warehouses, relational databases, etc. The motto of the data mining process is to take out information from a data set and alter it into a comprehensible and clear structured manner for further use. Due to its broader applicability and acceptability, Data mining has attracted much interest in database communities. The issues of mining association rules from the transactional database were introduced in. The theory aims to find regular patterns, exciting correlations, and links among sets of items in the data repositories

or transaction databases. Association rules are broadly used in controlling inventory, diagnosis in the medical field, market and risk management industry, drug testing industries, etc.

2. PROPOSED METHODOLOGY

In this section discuss proposed algorithm for optimization of association rule mining, the proposed algorithm resolves the problem of negative rule generation and also optimized the process of rule generation. Negative association rule mining is a great challenge for large dataset. In the generation of valid rules association existing algorithm or method generate a series of negative rules, which generated rule affected a performance of association rule mining. In the process of rule generation various multi objective associations rule mining algorithm is proposed but all these are not solve. In this Paper we proposed MLMS-GA of association rule mining with min-max algorithm.

Steps of algorithm (MLMS-GA)

1. Scanning of database used flowing steps

Some standard notation of pseudo code of algorithm such as D dataset, K level MLMS, Ls generation candidate

K = MLMS dataset (D)

n = Number of multiple level block

For i = 1 to n loop

Scan_k (Ki ∈k)

Li = gen__itemsets (ki)

For (i = 2; L^{j} i≠ φ, j = 1,2....,n; i++)

$Ci^{G} = \cup j = 1,2,...nL_{i}^{j}$

End;

For i = 1 to n

scan_kmap (ki∈K)

For all items C ∈CG generate block (C, ki)

End;

LG = {c ∈CG|}

2. Generate multiple support vector value for selection process

for all transaction LG do

generate count table TC

L_1 =(frequent 1-itemsets);

$C_2 = L_1 \infty L_{1;}$

L_2 ={cEC_2 | sup(c)≥MinSupNum};

For(k=3;L_{k-1} ≠Ø ;k++)do begin

For (j=k;j≤m;j++)do

Generate $CIV_{ij}^{k-1;}$

C_k=candidate_gen(L_{k-1})

L_k ={cEC_k| sup(c)≥MinSupNum};\

End

3. Set of rule is generated

Return L = U' $L_{k;}$

Candidate_gen(frequent itemset L_{k-1})

a. for all(K-1)-itemsetlE L_{k-1} do

b. for all i_jE $L_{k-1\ do}$

c. //S is the result of the formula(2)

If for every r(1≤r≤k) such that S[r]≥k-1 then

L_1 = (frequent 1-itemsets);

$C_2 = L_1 \infty L_{1;}$

L_2 = {cEC_2 | sup(c)≥MinSupNum};

For (k=3;L_{k-1} ≠Ø ;k++)do begin

For (j=k;j≤m;j++)do

Generate $CIV_{ij}^{k-1;}$

C_k=candidate_gen(L_{k-1})

4. Check MLMS value of table

5If rule is not MLMS go to selection process

6. Else optimized rule is generated.

7. Exit

a.) Data Encoding

The process of data in min-max algorithm needs some data encoding technique for representation of data. In this technique used binary encoding technique.

b.) Fitness function

The population selection of Min-max Algorithm is a design of Fitness Function:

$$m(S) = \frac{Ai}{wi} + \frac{Bi}{L \times (1 - wi)}$$

Ai = {frequent item support}

Wi_1= {level of Wight value of MLMS}

Bi = {those value or Data infrequent}

The Min-max operators determine the search capability and convergence of the algorithm. Min-max operators hold the selection crossover and mutation on the population and generate the new population. In this algorithm it restore each chromosome in the population to the corresponding rule, and then calculate selection probability pi for each rule based on above formula.

3. DATAFLOW DIAGRAM

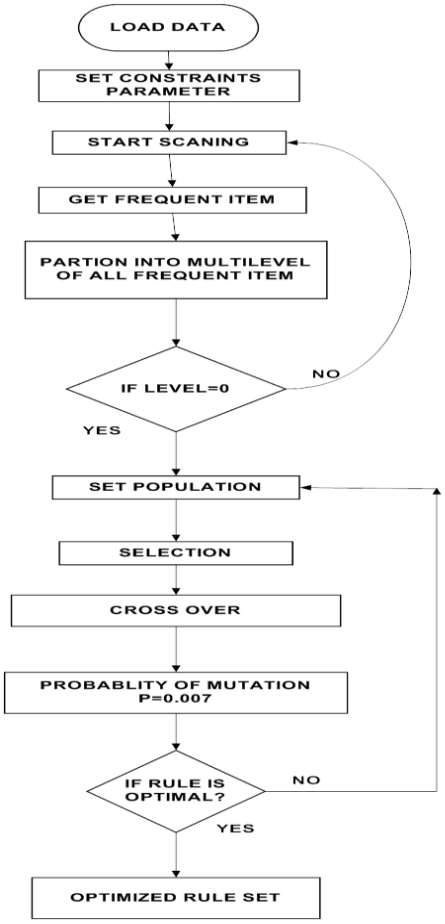

Figure 1: workflow of proposed System

3. EXPERIMENTAL RESULT

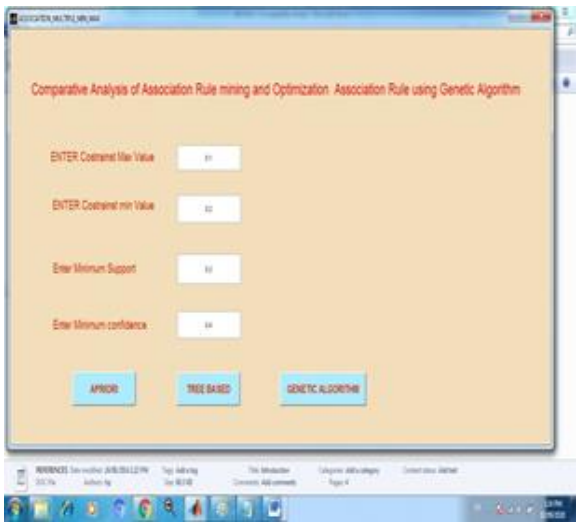

Figure 2: shows that the main window initially empty

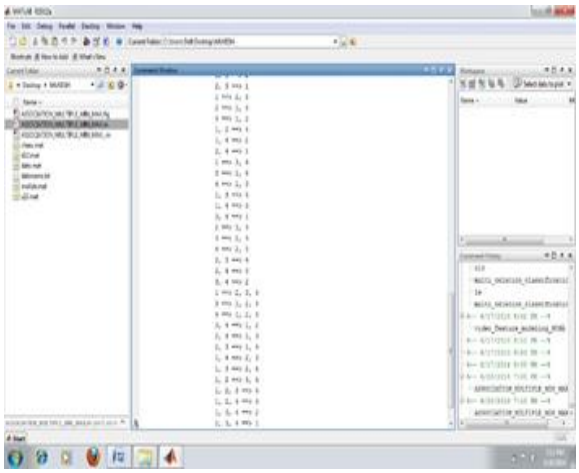

Figure 3:Rule generation by Apriori method with the no. of generated rule is 48

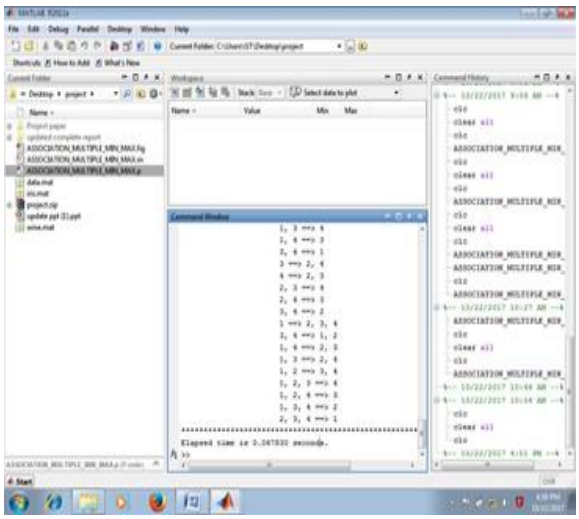

Figure 4: Rule generation by Tree method with the no. of generated rule is 36.

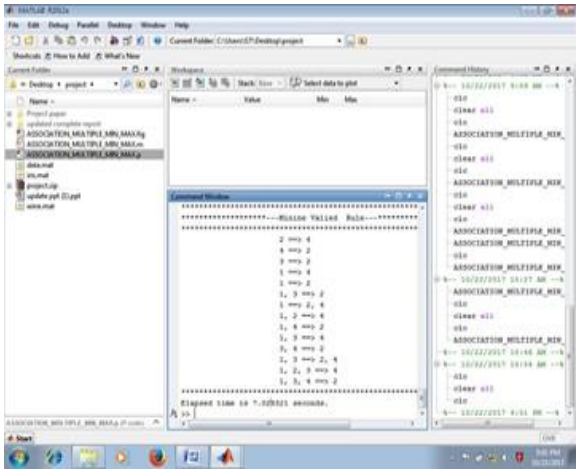

Figure 5: Rule generation by Genetic Algorithm with the no. of generated 14

CONCLUSION AND FUTURE SCOPE

Condition based association rule mining is great advantage over conventional rule generation technique. The conventional rule generation technique used some standard algorithm. The proposed algorithm is very promising in the field of association rule mining. The proposed algorithm has multiple constraints such as genetic algorithm and sine and cosine function. The value of sine and cosine increases the process of algorithm work as normal apriori and tree based techniques.

.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

FUNDING SUPPORT

The author declare that they have no funding support for this study.

REFERENCES

[1] [1] Xiaobing Liu , Kun Zhai, Witold Pedrycz "An improved association rules mining method" Expert Systems with Applications 2012, Pp 1362–1374.

[2] [2] Ying-Ho Liu "Mining frequent patterns from univariate uncertain data" Data & Knowledge Engineering, 2012. Pp 47-68.

[3] [3] Li Guang-yuan, Cao Dan yang, Guo Jian-wei "Association Rules Mining with Multiple Constraints" Elsevier ltd. 2011, Pp 1678-1683.

[4] [4] Idheb Mohamad Ali O. Swesi, Azuraliza Abu Bakar, Anis Suhailis Abdul Kadir "Mining Positive and Negative Association Rules from Interesting Frequent and Infrequent Item sets" IEEE 9th International Conference on Fuzzy Systems and Knowledge Discovery, 2012. Pp 650-655.

[5] [5] Huang QingLan, Duan LongZhen "Multi-level association rule mining based on clustering partition" IEEE Third International Conference on Intelligent System Design and Engineering Applications, 2013. Pp 982-986.

[6] [6] Vidhu Singhal, Gopal Pandey "A Web Based Recommendation Using Association Rule and Clustering" International Journal of Computer & Communication Engineering Research, 2013. Pp 1-5.

[7] [7] Gavin Shaw, Yue Xu, Shlomo Geva "Interestingness Measures for Multi-Level Association Rules' Proceedings of the 14th Australasian Document Computing Symposium, 2009. Pp 2-9.