International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

An Optimized Associative Classifier for Incremental Data Based On Non-Trivial Data Insertion

Ramesh R, Saravanan V, Manikandan R

Abstract: ***Associative classification (AC) is an interesting approach in the domain of data mining which makes use of the association rules for building a classification system, which are easy for interpretation by the end user. The previous work [1] showed excellent performance in a static large data base but there existed a question of same performance when applied in an incremental data. Many of the Associative Classification methods have left the problem of data insertion and optimization unattended that results in serious performance degradation. To resolve this issue, we used new technique C-NTDI for building a classifier when there is an insertion of data that take place in a non-trivial fashion in the initial data that are used for updating the classification rules and thereafter to apply the PPCE technique for the generating of rules and further Proportion of Frequency occurrence count with BAT Algorithm (PFOCBA) is applied for optimizing the rules that are generated. The experiments were conducted on 6 different incremental data sets and we found that the proposed technique outperforms other methods such as ACIM, E-ACIM, Fast Update (FUP), Galois Lattice theory (GLT) and New Fast Update (NFUP) in terms of accuracy and time complexity.***

Key Terms: Associative Classification, Optimization, Time Complexity, Incremental data, Galois Lattice Theory, Fast Update and New Fast Update

I. INTRODUCTION

Later improvements of data innovation and PC systems caused the generation of huge amounts of databases. These databases regularly contain covered up valuable data that can be made used for the decision making and global arranging. Subsequently, effectively finding and overseeing the valuable data from thesis expansive databases gotten to be a need. One of the most found instruments which finds and extricates knowledge from distinctive sorts of information is data-mining. The significance of data-mining is developing quickly in later a long time since it can be utilized for a few distinctive errands counting classification, clustering, regression, ACs and analyzing the outliers.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author

Ramesh R *, Assistant Professor, Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore and Research Scholar-Bharathiar University

Saravanan V, dean- Computer Studies, Dr. SNS College of Arts and Science, Coimbatore.

Manikandan R, Research Scholar , Anna University , Chennai, India.

Incremental learning is a great challenges that are related with data mining errands particularly associative rules and classifications. In classifications setting, the issue includes overhauling the model of classification (classifier) at whatever point the initial data gets into updation. In most applications that include stock trade, online real-time transactions, retail promotions, and financial products, data gets updated on a day by day premise, and thus dealing with the incremental learning issue gets to be significant in these applications. In most certifiable applications like securities exchange trade, online exchange, retail promoting, and banking, information for the most part are refreshed regularly, just as, new information are produced and old information might be out of date with the advancement of time. Thus, productive gradual refreshing calculations are required for support of the found affiliation principles to abstain from re-trying mining all in all refreshed database and in this manner dealing with the steady learning issue ends up critical in these applications [2]

Many incremental methodologies were proposed for handling the problem of Associative Classifications such as Fast Update (FUP) [3], FUP2 [4], Insertion, Deletion and Updating [5], Galois Lattice theory [6], and New Fast Update (NFUP) [7]. Moreover, only a little attention was paid to problems in classification particularly in associative classification [8] and in the rule induction methods [9], researchers have paid small consideration to the incremental database issue. Advance, since the process of classification may be considered a much common errand in data-mining and features an incredible volume of vital applications where data are frequently collected from these applications on day by day, week after week or month to month there exists great need to create or at slightest improve the current classification strategies to handle the incremental learning issue. Within the final few a long time an approach that coordinating affiliation run the show and classification called Associative Classification (AC) was proposed [10]. Many existing studies as in [11] have provided enough evidence that the ACRs are more accurate in building the classifiers when compared to other methods such as induction of rules and decision trees. Further, few of the applications such as image classification [12] and text classification [13] have also used the AC as it generates only simple IF_THEN rules which are easier for interpretation. In the previous work [1] Enhanced Equivalence Fuzzy Class Rule tree (EEFCR-tree) is proposed in the past work. The major shortcoming of FCARs Miner is that when the count of the rules constrained in a given class overwhelms, there seems to exist a degrade in the performance.

To take care of this issue Proportion of Constraint Class Estimation (PPCE) calculation for mining Enhanced Proportion Equivalence Fuzzy Constraint Class Association Rules (EPEFCARs) is developed so as to spare memory utilization, time complexity and accuracy. At that point, Proportion Frequency Occurrence check with Bat Algorithm (PFOCBA) is then presented for pruning the rules. At long last, an efficient technique was proposed for mining PEFCARs rules. Anyway the procedure was not a best fit when applied to an incremental data gets updated along the time. This issue is addressed by making use of non-trivial data insertion method (C-NTDI) to build the classifier when an insertion operation is done on the actual dataset that are used to construct the old and finally the FOCBA is applied to get the optimized rule generated. The remaining of the manuscript flows as follows. Chapter II gives a gist of the literature available in the specific area of research considered. Chapter III gives the architecture and working of the proposed methodology and Chapter IV explains the data set considered and the experimental setup. Chapter V discusses the various results obtained over the experiments through tabulations and graph. Chapter VI concludes the paper with scope for future research.

II. RELATED WORKS

Mohammed H et al [14] [2014] managed the Data addition issue inside the incremental data in AC mining. Especially, adjustment of a existing AC calculation named CBA was made to treat a single part of the database that is incremental. which is the insertion of data along the time. The new calculation called ACIM was proposed. Exploratory outcomes against six data sets from the UCI repository showed that the method projected decreases the time complexity whenever contrasted with CBA, and nearly determines a similar accuracy.

Raghuram Bhukya et al[15] [2014] Considered the significance of mining in incremental data and ACRs in a distributive environment and introduced a classification model for incremental data in a database that is partitioned horizontally. The experiments that were performed using the synthesized data produced better results.

Mustafa An et al[16] [2015] presented a changed ACIM calculation called Enhanced ACIM (E-ACIM). This method manages the problem of data insertion in AC context. The E-ACIM is aggressive and increasingly effective in computational time contrasted and ACIM and CBA and nearly gives a similar accuracy to both.

Vincent To-Yee Ng et al[17] [2015] intended to enhance the I/O exertion for discovering ACs. The method partitions the databases into two sections and all exchanges will be compacted with the assistance of a transaction that has a reference and found in the smallest partition. The method also compared the proposed technique and a typical compression methods which is the binary-compression. Observational results demonstrates that the proposed method performs well both in decreasing the space used for storage and the I/O procedure required to locate the huge item sets for affiliation rules.

Neda Abdelhamid et al[18] [2016] identified that AC experiences recognizable difficulties some of which have been acquired from ARs and others have been come about because of constructing the classifier stage. These difficulties are not constrained to the large amount of candidate rules found, the extremely huge classifiers determined the powerlessness to deal with multi-label data, and the plan of pruning the rules, ranking and prediction. The work features the regular difficulties looked by AC calculations that are as yet continued. The work additionally opens the entryway for new researchers to further examine these difficulties planning to upgrade the general execution of this methodology and increment it relevance in research areas.

B. Subbulakshmi et al [19] [2016] proposed a system that spotlights on the incremental mining which is utilized in the greater part of the ongoing applications. It utilizes ICCAR method by creating the ICCR rule Tree to produce the Constraint based CARs . If there should be an occurrence of any data insertion, the ICCR Tree is refreshed for the new arrangement of records without re-scanning. Safety treshold is utilized as a metric to decide if the dataset to be re-filtered or not. Additionally, since the hubs of the ICCR Tree is made to store the distinctions of the item identifiers, it expends less memory when contrasted with existing affiliated classifiers.

Darshana H. Patel et al [20] [2017] concentrates existing methods for Associative Classification which is a functioning territory of research in DM. Studies on pruning strategies were carried out for removing the rules that are not important and different estimates used to assess the ACs. It has been presumed that CBA is a basic and effective procedure that produces precise principles. Further, attempts to discover parameters like support and confidence was made for ACs as for number of rules as a few applications request interpretable outcomes which is reliant on the quantity of generated rules. The parametric values obtained from the experiment can be used for the improvement in associative classification.

Armando Segatori et al [21] [2018] propose a productive fuzzy AC approach in the MapReduce framework. The methodology abuses a novel dispersed discretizer dependent on Fuzzy entropy. At that point, a lot of candidate Acs that are fuzzy in nature are created by utilizing a distributed of the notable FP-Growth calculation. At long last, this set is pruned by utilizing three deliberately adjusted kinds of pruning. The method was experimented in a Hadoop frame work. Featured that, in spite of the fact that the accuracies result to be similar, the time complexity, assessed as far as the total count of the rules, no of the classifiers that are created by the fuzzy approach is much lower than non-fuzzy classifiers.

III. PROPOSED METHEDOLOGY

A. Problem Definition:

Existing AC calculations mine the data in training as entire so as to deliver the result. At the point when data based operations that includes Insertion , Updation or deletion happen on the data that are being trained, current calculations need to examine the total data under training once again so as to indicate the changes. Further the information are gathered in most of the application areas and are growing exponentially with time, the data under training can also quickly develop. Because of that, the expense of the scan done whenever the training data is updated and so as to refresh the ACs is exorbitant with respect to time complexity.

Incremental AC calculations, which will retract the mining outcomes and just consider information records which are refreshed, are a progressively proficient methodology, which can prompt a tremendous reduction in the time complexity.

In-order to understand the concept of data insertion problem, in an Associative Classification, Consider a sample training set B , in which the said operations are subjected to take place.

1. Adding can be done which increases the size from B to B+
2. Deleting will make the number of records reduced to B- from B
3. Updation occurs by Addition of B+ records and removal of B- records

The aftereffect of any of the activities portrayed above on B is data set that is updated B'. The inquiry is the manner by which the result (rules) of the initial data B can be updated to ponder changes done in B without performing broad calculations. This issue can be separated further into sub-issues as per the conceivable ruleitems contained in B in the wake of playing out a data manipulation. For instance, rule things in T can be isolated into the groups subsequent to the embeddings of new records (B+):

1. The ruleitems which occurs often in B and never in B+
2.The ruleitems which occurs often in B+ and never in B
3.The ruleitems which occurs neither often in B+ nor in B

The rule items in gatherings 1 and 2 can be distinguished in a direct way. For example, on the off chance that standard thing Y is visit in B, at that point it's help include in the refreshed preparing information (B'), Y'count = Ycount+Y+ tally, where Y count is known and Y+ check can be gotten in the wake of filtering B+ . The test is to discover visit ruleitems that are not visit in T however visit in T+ since these standard things are not decided subsequent to filtering B or B+

B. The Proposed Non-Trivial Data Insertion Method

Figure 1 illustrates the developed Non-Trivial Data Insertion Method algorithm which are sequentially given below

1. Constructing the classifier from TD , at that point testing, and inferring the accuracy of the classification
2. Selecting a stratified random example (dsi) based on new information that are inserted to D (di) where this example speaks to just 10%. For reasonable choice, the picked 10% are stratified which means the class conveyances inside the new included lines are considered to be the chosen tests
3. Making use of the dsi sample in a random manner as a testing set for the classifier in the initial step
4. Comparing the precision which acquired in 3 with that of 1. On the off chance that the precision of stage 3 is more noteworthy than or equal to 1 from which we keep the classifier delivered at 1
5. Forms another classifier utilizing ACs dependent on a incremental mining technique.
6. To apply the FOCBA on the created guidelines for streamlining so as to limit the running time.

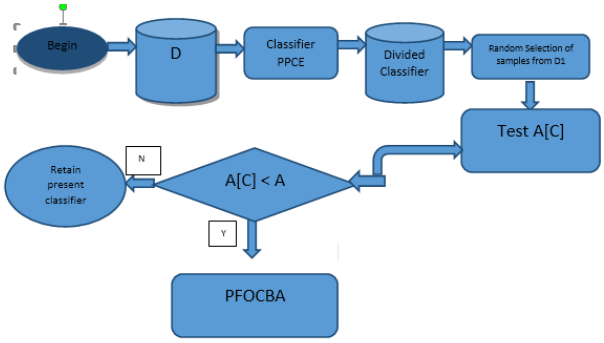

Figure 1 – Proposed NTDI method

a. General Description :

Table 1 shows the methods and general description of the proposed algorithm

Table 1 – Description of the symbols and methods

Symbol	Explanation
I	No of Incremental Data
D	The actual Database
d_i	The dataset incremented
ds_i	The random sample
A	Accuracy of the present classifier
a	Accuracy of the classifier for ds_i
FOCBA()	Optimizing the CARs using [1]

1 Inputs: Initial DB, min-sup, min-conf
2 Output: Set of Classifiers (C)
3 build_classifier_CBA(D)
4 A='Accuracy_of_classifier'(D);
5 do until (data Updation is nil) {
6 dsi= SRS(di)
7 for each r Є dsi {14 classifier- testing(r)
8 a = Accuracy-of-classifier (dsi) }
9 if (a < A) then {PFOCBA (di U D) }
10 Else {D=D U di } } }

Figure 2 - Pseudo code of the proposed C-NDTI

b. Working :

The pseudo code of the proposed method is shown below

As indicated by Figure no 2, in lines 3-4, the clustering framework is developed utilizing CBA calculation from the first database (D), and after that the precision is determined and put away in factor A. In the line 5, the calculation chooses a stratified arbitrary example (dsi) from the steady informational index (di) where this example speaks to just 10% of the gradual information, that has nontrivial nature in size. For reasonable determination, the picked 10% columns are stratified significance the class dispersions inside the new included lines are to be considered and hence the chose test contains all class marks. The precision got from this test is compared to CBA.

In the event that the new exactness is better, at that point the delivered classifier at line 4 or 8 is kept unaltered. Generally another classifier will work from the gradual information and unique information by applying ACIM calculation [9]. This procedure is rehashed at whatever point the preparation informational index D gets refreshed (Updated data)

Therefore, in light of the above mentioned, the main classifier worked through first database D; when addition data di are inserted to the first dataset. The sample taken in random was selected, tried utilizing the current classifier only if exactness and the present classifier function admirably satisfactory on the incremental data; at that point proposed calculation hold the current classifier. Something else, (for example the exactness of the present classifier on arbitrary example is not exactly the past precision) ACIM will be utilized to assemble another classifier utilizing the incremental data and the exact data D by using the classifier which inherent the past stage. Once the ACRs are obtained as a result of the proposed method, The, PFOCBA [1] is applied for optimizing the pruned rules which satisfies the class constraints the most for better performance in terms of time complexity.

C. Performance Measures

The performance are measured as far as the accuracy is concerned in the following way

i. Classification Accuracy

Generally accuracy is what often meant as Classification Accuracy which is defined as the ratio between the number of predictions that are correct with the overall count of the predictions.

$$Accuracy = \frac{No\ of\ Correct\ Predictions}{Total\ no\ of\ Predictions} \quad (1)$$

ii. Logarithmic Loss

Logarithmic Loss or Log Loss works with the concept of penalizing the classifications that are not true. It often works along with the multi level classification. When these are considered, the classifier must be assigned a probability for all the class in all the samples . in case if there are N number of samples that belong to M no of classes, then the Log loss is computed as

$$Logarithmic\ Loss = \frac{-1}{N} \sum_{i=1}^{N} M \sum_{j=1}^{M} N\, y_{ij} \ast \log P_{ij} \quad (2)$$

F1 Score

F1 Score is calculated as the Harmonic Mean in-between the precision and the recall ranging from 0 to 1.It indicates the correctness of the classifier and the strength of the classifier. It is calculated as follows

$$F1 = 2 \ast \frac{1}{\frac{1}{precision} + \frac{1}{recall}} \quad (3)$$

Where Precision is the count of the positive results that are correct divided with that of the reults that are positive and results that are negative.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (4)$$

and Recall is the count of the results that are positive which are divided by the count of all appropriate samples and are denoted below as

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (5)$$

In terms of Time complexity, the performance metrics considered the average case time which is calculated as follows

Let $T_1\ (N), T_2\ (N), T_3\ (N), \ldots, T_N\ (N)$ be the execution time for all possible inputs of the size ‘ n’and $P_1\ (n), P_2\ (n), P_3\ (n), \ldots P_n\ (n)$ be the probabilities of these inputs , then the average-case time complexity is measured as

$$P_1\ (n) T_1\ (n) + P_2\ (n) T_2\ (n) + \ldots + P_n\ (n) T_n\ (n-1) \quad (6)$$

IV DATA AND EXPERIMENTAL SETUP

The information were taken from the UCI store [13] . These information are gradual in nature and are Car eval. ", "led7", "pageblocks ", "pen digits ", "waveform ", and "wine qu." . The choice of these informational indexes depends on the size and the information quality since we have searched for medium to huge size preparing informational collection since we isolated the dataset into unique and steady information, additionally the above informational indexes don't contain missing qualities. The total attributes of the information are found in [13] . Table 1 gives a review about the information considered for the trial

Table 3 – Data Set Characteristics

Data Set	No of Attributes	No of Records	Class number
Car Evaluation	6	1598	4
Page Block	10	2658	5
Pen Digit	16	1102	10
Wav form	21	10025	3
Wine Quality	12	4685	7
Led 7	7	7805	10

All the UCI data set are divided into couple of partitions the first for training and the second for the incremental learning. Then, further bifurcation of data was made on the incremental data into a set of five blocks for ensuring the testing purpose and specifically to test on multiple insertions in the data used for the training over the process of evaluation.The minsup is set to 1% and the minconf is set to 50% before the process of classification. The experiments were simulated in the MATLAB environment with other minimum system configuration.

V RESULTS AND DISCUSSIONS

Table 4 shows the tabulation of the comparison of the classification accuracy of different state of the art methods. The proposed C-NDTI method shows 98.23 % accuracy in the car evaluation data set and 98.15% accuracy in page Block data set. The Pen digit data produced 99.19% accuracy and an accuracy of 98.65% is obtained for Wav form data. The wine wine quality and Led 7 also showed a significant accuracy of 99.12% and 98.60% which is substantially more than other methods considered.

Table 4 - Comparison of Classification Accuracy

Data Set	Classification Accuracy %					
	C-NTDI	FUP	N-FUP	ACIM	E-ACIM	GLT
Car Evaluation	98.23	94.12	96.23	97.12	96.12	92.57
Page Block	98.15	94.35	96.15	97.85	96.32	92.75
Pen Digit	99.19	95.36	94.62	95.15	97.25	93.18
Wav form	98.65	94.27	95.54	97.26	96.18	94.96
Wine Quality	99.12	95.85	92.18	96.84	97.19	95.41
Led 7	98.60	94.28	91.32	96.25	94.28	92.85

Figure 3 gives the graphical representation of the results obtained in terms of classification accuracy depicted in Table 4. It is evident from the graph that the proposed method C-NTDI outperformed the other methods in all the datasets that were taken for the experiment.

Table 5 shows the tabulation of the comparison of the Logarithmic Loss of different state of the art methods. The proposed C-NDTI method shows minimum loss value of 0.051 in the car evaluation data set and 0.041 loss value in Page Block data set. The Pen digit data produced 0.125 loss value and the same was 0.068 for Wav form data. The wine quality and Led 7 also showed a very minimum loss of 0.124 and 0.543 which is substantially less than all the other methods considered.

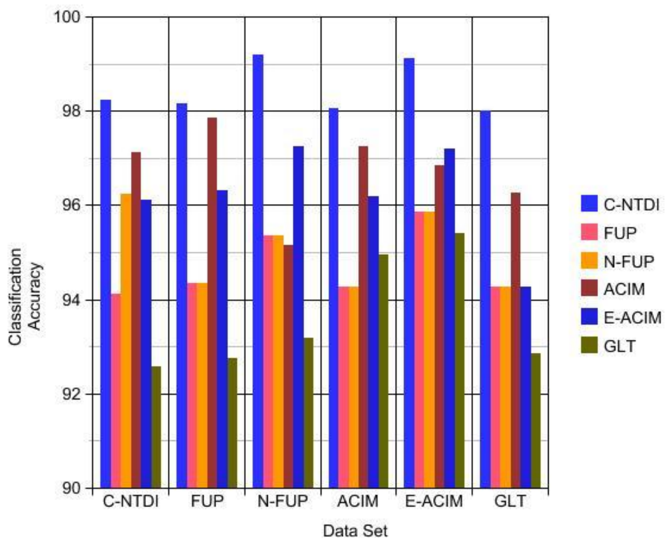

Figure 3 – Comparison of Classification Accuracy

Table 5 – Comparison of Logarithmic Loss

Data Set	Logarithmic Loss					
	C-NTDI	FUP	N-FUP	ACIM	E-ACIM	GLT
Car Evaluation	0.051	0.123	0.184	0.162	0.212	0.298
Page Block	0.041	0.086	0.097	0.165	0.054	0.214
Pen Digit	0.125	0.289	0.314	0.241	0.185	0.368
Wav form	0.068	0.095	0.126	0.184	0.195	0.214
Wine Quality	0.124	0.258	0.354	0.310	0.268	0.190
Led 7	0.543	0.654	0.987	0.121	0.825	0.781

Figure 4 gives the graphical representation of the results obtained in terms of Logarithmic Loss depicted in Table 5. It is evident from the graph that the proposed method C-NTDI produced a minimum loss than other methods in all the datasets that were taken for the experiment.

Table 6 shows the tabulation of the comparison of the F1 score obtained in a 10 point scale from different methods. The proposed C-NDTI method shows a maximum F1 score of 9.88 in the car evaluation data set and 8.98 in Page Block data set. The Pen digit data produced 9.35 as F1 score and it was 8.90 for Wav form data. The wine quality and Led 7 also produced a great F1 values of 9.5 and 9.68 which is substantially great than all the other methods considered.

Table 6 – Comparison of F1 Scores

Data Set	F1 Score on a 10 point scale					
	C-NTDI	FUP	N-FUP	ACIM	E-ACIM	GLT
Car Evaluation	9.88	7.15	8.02	8.12	8.01	8.51
Page Block	8.98	7.25	7.09	8.85	8.08	7.68
Pen Digit	9.35	8.84	8.04	7.45	8.09	8.32
Wav form	8.90	8.01	7.15	8.25	8.10	7.21
Wine Quality	9.50	7.23	8.21	8.23	8.17	8.84
Led 7	9.68	8.15	8.09	9.10	8.29	7.25

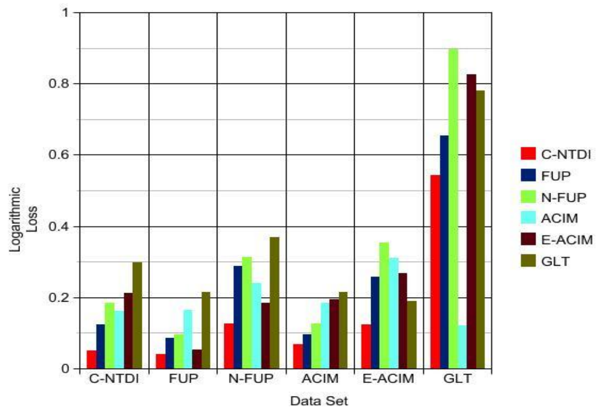

Figure 4 – Comparison of Logarithmic Loss

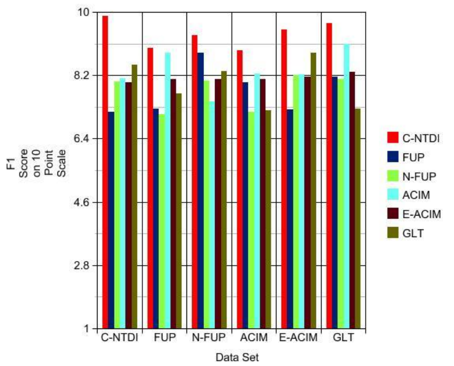

Figure 5 – Comparison of F1 values

Figure 5 gives the graphical representation of the results obtained in terms of F1 score depicted in Table 6. It is evident from the graph that the proposed method C-NTDI produced a maximum F1 score than other methods in all the datasets that were taken for the experiment.

VI. COCLUSION AND FUTURESCOPE

This research is carried out as an extension of the previous work [1] which lagged performance when applied to incremental database**.** The problem of insertion is addressed here by proposing the C-NTDI which solves the problem using Non trivial data insertion. The PFOCBA proposed in the former work is made used for producing optimized ACRs. The results obtained were substantially better in terms of Classification accuracy, Logarithmic loss and F1 score when compared with the state of the art algorithms. The future work aims to apply these identifications in a distributed big data environment and to study the performance deviations if any.

REFERENCES

1 Ramesh, R. & Saravanan, V.. (2018). Proportion frequency occurrence count with bat algorithm (FOCBA) for rule optimization and mining of proportion equivalence Fuzzy Constraint Class Association Rules (PEFCARs). Periodicals of Engineering and Natural Sciences. 6. 305-325. 10.21533/pen.v6i1.278

2 Nath B., Bhattacharyya D. K., and GhoshA., “Incremental association rule mining: a survey,” WIREs Data Mining KnowlDiscov, vol. 3, No. 3, 2013, pp. 157–169.

3 Liu B., Hsu W. and Ma Y., “Integrating classification and association rule mining”, in proceedings of the KDD, 1998, pp. 80-86, New York, NY.

4 Toshi C. and Neelabh S., “Incremental Mining on Association Rules”, Research Inventy: International Journal of Engineering and Science, Vol. 1, No. 11, 2012, pp. 31-33.

5 Nath B., Bhattacharyya D. K., and GhoshA., “Incremental association rule mining: a survey,” WIREs Data Mining KnowlDiscov, vol. 3, No. 3, 2013, pp. 157–169.

6 D. Cheung, J. Han, V. Ng, and C. Wong, "Maintenance of Discovered Association Rules in Large Databases: An Incremental Updating Technique," in Proc. 12th the International Conference on Data Engineering, New Orleans, 1996, pp.106-114.

7 D. Cheung, S. Lee, and B. Kao, "A general Incremental Technique for Mining Discovered Association Rules," in Proc. 5th International Conference on Database System for Advanced Applications, Melbourne, 1997, pp. 185-194

8 P. Tsai, C. Lee and A. Chen, "An efficient approach for incremental association rule mining," in Proc. 3rd Pacific-Asia Conference on Methodologies for Knowledge Discovery and Data Mining, London, 1999, pp. 74-83.

9 V. Petko, M. Rokia, R. Mohamed, and G. Hacene Robert (2003),"Incremental Maintenance of Association Rule Bases," in Proc. 2nd Intl. Workshop on Data Mining and Discrete Mathematics, San Francisco, 2003, 12 p.

10 C. Chang, Y. Li, and J. Lee, "An Efficient Algorithm for Incremental Mining of Association Rules," in Proc. 15th International Workshop on Research Issues in Data Engineering: Stream Data Mining and Applications, Washington, 2005, pp.3-10.

11 M. Antonie, O. Zaïane, and A. Coman, "Associative classifiers for medical images," Lecture Notes in Artificial Intelligence 2797, Mining Multimedia and Complex Data, Springer-Verlag, 2003, pp. 68-83

12 Y. Yoon and G. Lee, "Practical application of associative classifier for document classification," in Proc. 2nd Asia Information Retrieval Symposium, Jeju-island, Korea, 2005, pp. 467-478.

13 Y. Yoon and G. Lee, "Practical application of associative classifier for document classification," in Proc. 2nd Asia Information Retrieval Symposium, Jeju-island, Korea, 2005, pp. 467-478.

14 Mohammed H. Alnababteh, M. Alfyoumi, A. Aljumah, and J. Ababneh , Associative Classification Based on Incremental Mining (ACIM), International Journal of Computer Theory and Engineering, Vol. 6, No. 2, April 2014

15 R. Bhukya and J. Gyani, "Incremental associative classification on distributed databases," International Conference for Convergence for Technology-2014, Pune, 2014, pp. 1-6.

16 Mustafa A. Al-Fayoumi , Enhanced Associative classification based on incremental mining Algorithm (E-ACIM), IJCSI International Journal of Computer Science Issues, Volume 12, Issue 1, No 1, January 2015 ISSN : 1694-0814

17 Vincent To-Yee Ng, Jacky Man-Lee Wong and P. Bao, "Incremental mining of association patterns on compressed data," Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569), Vancouver, BC, Canada, 2001, pp. 441-446 vol.1.

18 Neda Abdelhamid, Ahmad Abdul Jabbar , Fadi Thabtah , Associative Classification Co mmon Research Challenges , 2016 45th International Conference on Parallel Processing Workshops

19 B. Subbulakshmi and M. Monisha, "Incremental constraint class association rule mining of student performance dataset," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2016, pp. 1-5

20 H. Patel, N. K. Radhika and A. R. Vasant, "Associative classification: A comprehensive analysis and empirical evaluation," 2017 Nirma University International Conference on Engineering (NUiCONE), Ahmedabad, 2017, pp. 1-5.

21 Armando Segatori, Alessio Bechini, Pietro Ducange, and Francesco Marcelloni , A Distributed Fuzzy Associative Classifier for Big Data, IEEE TRANSACTIONS ON CYBERNETICS, VOL. 48, NO. 9, SEPTEMBER 2018

AUTHORS PROFILE

Ramesh R is a research scholar of Bharathiar university Coimbatore and presently working as an Assistant professor in the department of Computer Science in Sri Krishna Arts and Science college , Coimbatore. His area of interest includes data mining, Associative rule mining , Fuzzy set theories and Bigdata analytics.

Saravanan V is presently working as the Dean-Computer Studies in Dr.SNS Colelge of Arts and Science. He has nearly 60 publications to his credit. His area of interest and specialization includes Data Mining, Data Warehousing, Automated mining and Software Engineering. He has guided many research scholars leading to PhD degree

Manikandan R is a research scholar in Anna University Chennai and his area of interest are data mining, opinion and sentiment analysis and classification theory