Audio Classification based on Association and Hybrid Optimization Technique

Gurdeep Singh
Research Scholar, Department of CSE
Shri Guru Granth Sahib World University
Fatehgarh Sahib (Punjab), India

Shruti Aggarwal
Assistant Professor, Department of CSE
Shri Guru Granth Sahib World University
Fatehgarh Sahib (Punjab), India

ABSTRACT

In pattern recognition areas and data mining, audio data classification is a most important topic. This paper describes a new classification method, where **O**ptimal **C**lassification **R**ule **E**xtraction for multi-class **A**udio **D**ata (O-cREAD). This classification method uses a new hybrid optimization approach for extracting optimal classification rules, and then these optimal rules are further used for classifying multi-class testing audio data to their respective classes with better accuracy. The optimal classification rule extraction is a two-step process. In the first step, frequent itemsets are generated by the hybrid apriori algorithm and generates classification rules using the association concept. Next, a new hybrid optimization approach is used for optimizing classification rules of classification method. The new hybrid optimization approach is based on Ant Colony Optimization (ACO) and Multi-Objective Genetic Algorithm (MOGA). The best feature of classification method (O-cREAD) is that size of classification rules can be dramatically reduced and produce more sophisticated or complicated rules to improve classification accuracy for classifying a real audio dataset into their respective classes.

Keywords
Audio Data, Classification Rules, Hybrid Apriori Algorithm (H-AA), Genetic Algorithm (GA), Ant colony optimization (ACO), Hybrid optimization approach.

1. INTRODUCTION
Audio mining is a fast expending field due to wide growth of the internet (i.e. web). Audio classification is recognition based difficult process, so it needs more efficient and accurate procedures for this task. Classification using association or associative classification [14] is a rule-based approach generates the classification rules with measuring the association between the data. In associative classification approach there are various classic methods like apriori, FP-tree for discovering the frequent itemsets for the dataset, then generate classification rules using the association concept form these frequent itemsets. The frequent itemsets are also efficiently discovered by a hybrid apriori algorithm [9].

In recent years, much attention has been paid to optimization approaches like genetic algorithm (GA) and ant colony optimization (ACO) for optimization of the rules generated from frequent itemsets. GA [4-5] was proposed by John Holland in 1975 and one of the oldest optimization approach, based on natural evolution. GA is population based heuristic search method with some operator (selection, crossover, & mutation) and control parameters (Chromosome Encoding, Population Size, Maximum Generation, Selection Method, Crossover Probability, Mutation Probability, and Fitness

Function). On the other hand, ACO [6-8] was proposed by Marco Dorigo in early 1990. The ACO is a population based Meta-heuristic search method and inspired by the behaviour of real ants for food search with optimum paths. In ACO there are also some operators (Pheromone Update & Measure, Trail Evaporation) and control parameters (Number of ants, Iterations, and Pheromone Evaporation Rate) are used for generating optimal solution of a problem. The main objective of any optimization process for classification rules to minimize the number of classification rules or maximize the accuracy of classification rules.

This paper proposed a new classification method, where optimal classification rule extraction for multi-class audio data (O-cREAD).The optimal classification rules are extracted by following these steps:

a) Find frequent itemsets from the transaction database of each class using the hybrid apriori algorithm and generate classification rules using the association concept form these frequent itemsets of each class.

b) The hybrid optimization approach is used for each class's original classification rules to extract optimal classification rules for classification method.

By using the hybrid optimization approach in classification method, it dramatically reduced the size of classification rules and produces optimal classification rules with the better classification accuracy rate for classifying a real audio dataset into their respective classes.

This paper is organized as: Section 2 provide the overview about work related to this proposed work. Section 3 describes the basic idea and concept for extract optimal classification rules from transaction databases. Section 4 describes the procedure of O-cREAD method for classification. Section 5 explains the audio dataset used in this study and the observation, experimental concept of evaluating performance of classification method using O-cREAD. Section 6 discusses the various experimental results and performances. At last, section 7 summaries concussions of this study and future work.

2. RELATED WORK
In 2010, *Okada et al.* [10] proposed a new rule-based classification method for multi-class audio data. This method is also known as cREAD method (Classification Rule Extraction for Audio Data). This classification method takes the advantage of a closed itemset mining algorithm to achieve efficient and fast extraction of classification rules. In the cREAD method, the search space of classification rules are dramatically reduced by adding the "class label item" with the each transaction in the transaction database.
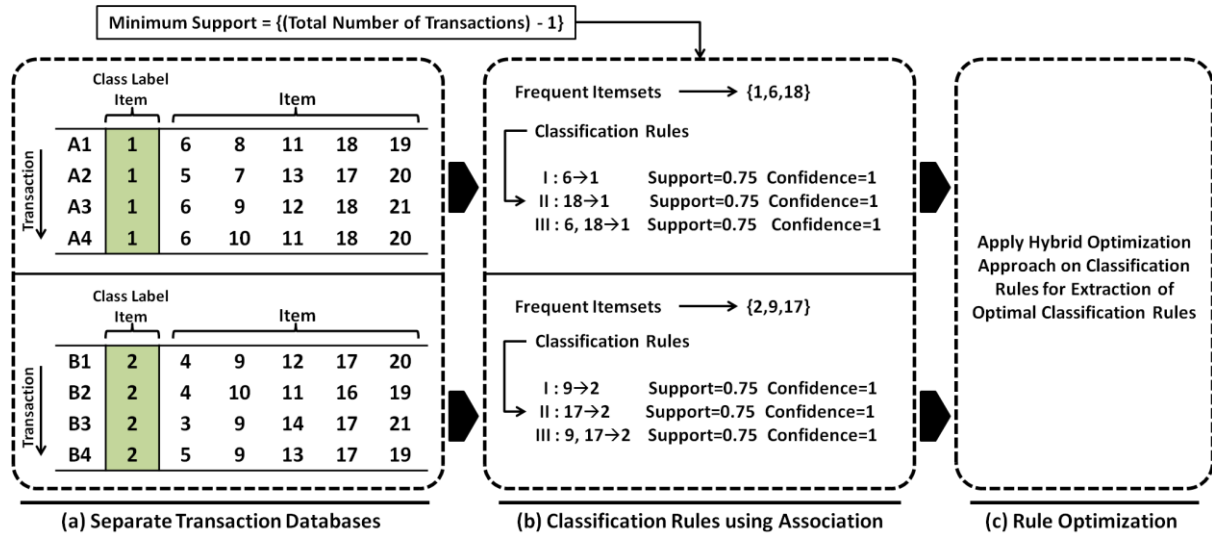
Fig 1: Basic idea to extract optimal classification rules

3. BASIC IDEA TO EXTRACT OPTIMAL CLASSIFICATION RULES

3.1 Frequent Itemsets

A frequent itemset [1] [14] is set of items that are frequently appearing in a transaction database with satisfying a predefined minimum support value. A transaction database is represented as shown in figure 1(a), where each row represents a transaction and each column represents an item includes in that transaction. In this method, here two different transaction databases are used and each transaction database includes a unique 'class label item', this item is belong to a particular class (i.e. class 1 and class 2) as shown in figure 1(a). In other word, each class having its own transaction database.

3.2 Hybrid Apriori Algorithm

A hybrid apriori algorithm (H-AA) [9] is used to find frequent itemsets from the transaction database. The frequent itemsets are finding by the H-AA which is the combination of transaction reduction technique and double pruning method for reducing the candidate itemsets. H-AA applies to the transaction table of each class for finding frequent itemsets. As shown in figure 1(a), there are two separate transaction databases and each transaction database having a 'class label item' to represent a particular class. When applying H-AA with minimum support 3 on the transaction database of each class, then it generated frequent itemsets of the transaction database of each class are as shown in figure 1(b).

3.3 Classification Rule Generation using Association

Classification Rules [14] are completely different from association rules, but represented in the same structure. The basic concept of classification rules is finding some relationship between some attributes or in attribute values, and also representing decision support between them. As shown in figure 2, classification rules [11-12] are represented in the IF – THEN structure, and IF and THEN part is also known as Antecedent & Consequent respectively. In the classification rules, consequent part contains values of only one attribute (Predefined Attribute) but on the antecedent part there is no such restriction (N- Predicting Attribute). The major difference between classification rules and association

rules is that, in the association rules any attribute and any number of attributes may appear in either side.
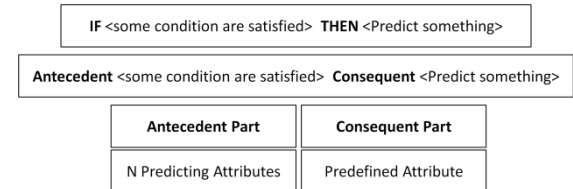

Fig 2: Basic concept of classification rules

The only restriction is that the two parts should not have a common attribute. Find support and confidence of each classification rule for choosing best rules. As shown in figure 1(b), the classification rules based on association are generated from frequent itemsets of each class's transaction database. These classification rules are also known as original classification rules.

3.4 Hybrid Optimization Approach

Optimal classification rules are extracted from original classification by a hybrid optimization approach as shown in figure 1(c). The original classification rules generated from frequent itemsets using association. This new approach is based on Ant Colony Optimization (ACO) [7] and Multi-Objective Genetic Algorithm (MOGA) [3]. The steps of new optimization approach are as follows:

1. Initialization of number of generations, population size, crossover probability, mutation probability, number of ants & stop criteria (Food).
2. Calculate the initial population fitness value.
3. Select a strong population lies between first top-N fitness values (population size) & discard weak population.
4. Then select pairs of chromosomes for mating, with the crossover probability for exchange parts of the selected chromosomes and create two offspring. If the mutation probability true, then randomly change the gene values in the two offspring chromosomes.
5. After step 4, calculate confidence and interestingness of new population (new route).
6. If number of generation count is equal to maximum generations then going to step 7, else go to step 2.

7. Select path by calculating mean of confidence and interestingness of new population (new route), and select shortest path by comparing all the paths and update pheromone trial.
8. If stopping criteria (Food) met then stop (Find a solution), else go to step 9.
9. If the number of ant count is equal to maximum ants then going to step10, else go to step 2.
10. There is no solution find for the problem (Food).

The representation of each chromosome is in real numbers. As shown in figure 3, fitness function [3] of new hybrid optimization approach is calculated using support, confidence, comprehensibility, & interestingness.

Fig 3: Fitness function of hybrid optimization approach

The comprehensibility means if a number of conditions in the antecedent are less, then the rule is assumed as a simple rule.

Interestingness means how much rule is surprising for the users. In fitness function *W1, W2, W3 & W4* are user defined weights and fitness value should be in the range of 0 to 1.

4. METHOD

The methodology of classification method divided into three phases. The 'phase I' represents the extraction of optimal classification rules, and the 'phase II' represents the preprocessing of unknown or testing audio samples and the 'phase III' represents the classification of unknown or testing audio samples to their respective classes. Figure 4 illustrates the procedure of O-cREAD method for extraction of optimal classification rules for classification method to classifying multi-class audio data.

4.1 Pre-Processing

Consider N_1 and N_2 samples of audio data of class 1 and class 2 respectively for rule extraction. Each audio sample of class 1 & class 2 is preprocessed by following steps:

a) Convert audio digital sample (any format file) into a waveform.
b) The waveform is then converted into the frequency domain by Fourier transform and having *p* frequency points of audio sample.
c) Power of each frequency point of audio sample is normalized by discretization according to a maximum power of audio sample.
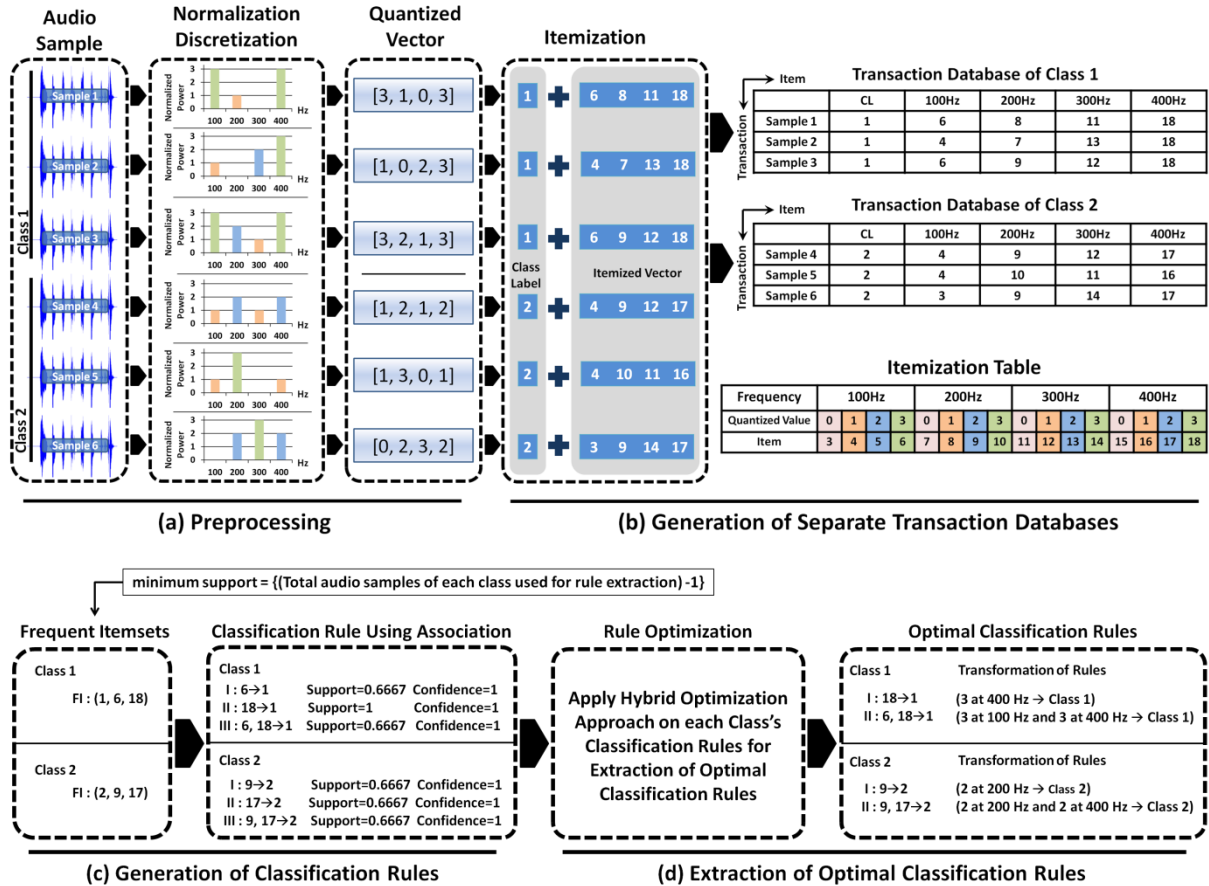d) Last, normalized power is quantized into *q* levels.

Fig 4: Procedure of O-cREAD Method

4.2 Generation of Separate Transaction Databases

Transaction database is generated by an itemization table as shown in figure 4(b); there are some colour codes are used to represent a pair of each quantized value & item in the each frequency point. Then, according to the itemization table, transform each quantized vector into an itemized vector with a class label item. At last, these itemized vectors are summarized into the separate transaction databases according to their class labels. In the end, the two separate transaction databases are generated for class 1 and class 2 as shown in figure 4(b).

4.3 Generation of Classification Rules

Individually apply H-AA to the transaction database of each class for finding frequent itemsets with minimum support equals to the total number of audio samples of that class are used for rule extraction minus one. Then, generate classification rules using the association concept from these frequent itemsets of each class. Measure the support and confidence value of each classification rule as shown in figure 4(c).

4.4 Extraction of Optimal Classification Rule

This section extract optimal classification rules of class 1 and class 2 using the hybrid optimization approach as shown in figure 4(d). In this rule optimization process, all optimal classification rules are extracted from original classification rules. Next, these extracted optimal classification rules are transformed into again quantized values with the help of itemization table.

4.5 Pre-Processing of Testing Audio Samples

Suppose M_k unknown or testing audio samples are needed to classify into their respective classes, then firstly each unknown or testing sample is pre-processed according to these following steps:

a) Convert audio digital sample (any format file) into a waveform.

b) The waveform is then converted into the frequency domain by the Fourier transform and having *p* frequency points of audio sample.

c) Power of each frequency point of audio sample is normalized by discretization according to a maximum power of audio sample.

d) Last, normalized power is quantized into *q* levels.

Measure the each quantized value of each frequency point of each audio sample & create quantized vector database of all these unknown or testing audio samples.

4.6 Classification

After extracting optimal classification rules & preprocessing of unknown or testing audio samples, then the final step is the classification of unknown or testing audio samples by using k-nearest neighbor (k-NN) approach [14] and optimal classification rules here act as training set and unknown or testing audio samples quantized vector database act as testing set. In this approach Euclidian distance is calculated between training set and each testing set value. The classification of each unknown or testing audio sample is based upon majority vote.

5. EXPERIMENT

5.1 Dataset

In this proposed work, here telephone based Speaker identification dataset produced by *Azarias et al.* [15] is used to evaluate the performance of classification method. This dataset was collected for the purpose of speaker identification and consists of two classes: female and male participants (Indian National) from different backgrounds. Each participant reads some lines of digits and one digit at a time. All digits were read in English & all audio files in WAV format (e.g. file name→123345.wav). The pre-processing of all these audio samples; first generate waveforms, then these waveforms data is transformed into the frequency domain by Fourier analysis in the frequencies 0-8000 Hz and in a resolution of 15.625 Hz (512 frequency points).

5.2 Evaluation of Classification Method

In the execution of classification method, 83 audio samples (i.e. 35 female & 48 male) are used for classification and 12 audio samples are used from each class for extraction of optimal classification rules. The minimum support of hybrid apriori algorithm is equals to 11for finding frequent itemsets form transaction table of each class. A variation analysis of this classification method done by performing three different types of experiments:

a) Concentrate only on condition (food)

b) Concentrate only on the number of generations

c) Concentrate only on the number of ants

In the first experiment, analysis is concentrating only on condition (food). In the second experiment, analysis is concentrating only on the number of generations. In the third experiment, analysis is concentrating only on the number of ants. All these experiments are executed with some observation of control parameters.

This classification method using optimal classification rules (Optimal-CR) is compared with; classification method using frequent itemsets as classification rules (FI-CR), and classification method using original classification rules (Original-CR).

6. RESULTS AND DISCUSSION

The first experimental (concentrating only on condition (food)) as shown in figure 5(a), this graph represents the classification accuracy of classification method (O-cREAD) and the average classification accuracy of classification method is 68.67±6.12% and average error rate is 31.33±6.12%. In figure 5(b), this graph provides the description about the average sensitivity & specificity of classification method are 55.42±15.71% & 78.33±19.79% respectively and average gmean factor for analysis sensitivity & specificity is 63.57±7.13%. The average precision & recall of classification method are 68.28±9.09% & 55.42±15.71% and average f-measure factor for analysis of precision & recall is 59.46±5.56% as shown in figure 5(c). In the first experiment, the control parameter (condition (food)) of classification method (O-cREAD) continuously increased by 2.

The second experimental (concentrating only on the number of generations), here control parameter (number of generations) of classification method (O-cREAD) continuously increased by 2 and the number of ants is constant at 100. Then, as shown in figure 6(a), average classification accuracy of classification method is

71.68±3.40% and average error rate is 28.32±3.40% of classification method. The average sensitivity & specificity of classification method are 54.99±5.66% & 83.84±5.53% respectively and average gmean factor for analysis sensitivity & specificity is 67.75±3.74% as shown in figure 6(b). The

average precision & recall of classification method are 71.90±7.00% & 54.99±5.66% and average f-measure factor for analysis of precision & recall is 62.03±4.64% as shown in figure 6(c).
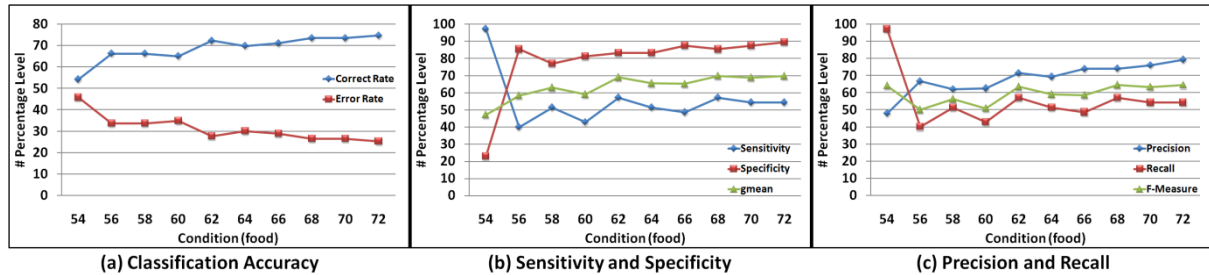

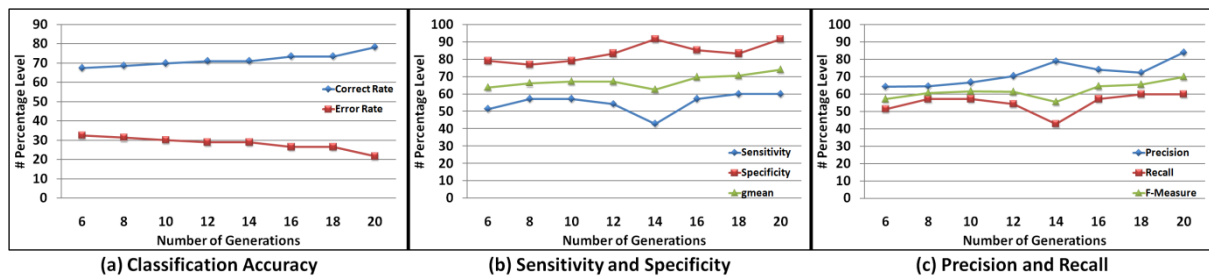**Fig 5: Performance of classification method (condition (food))**

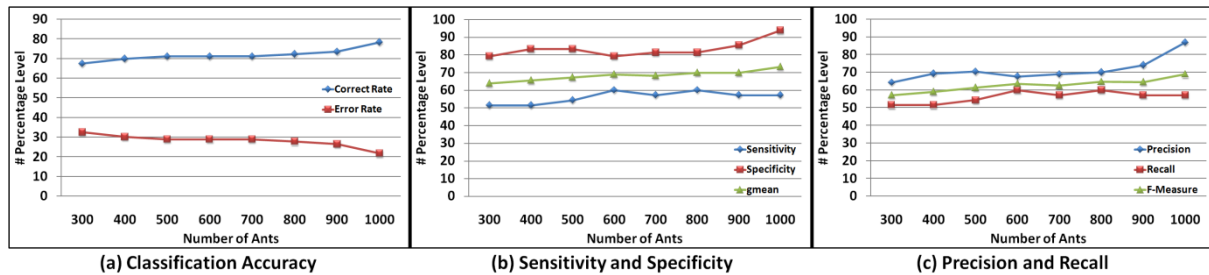**Fig 6: Performance of classification method (number of generations)**

Fig 7: Performance of classification method (number of ants)

Type of Classification Rules Use	Number of Classification Rules	Correct Rate (%) (TP+TN)/(TP+FN+FP+TN)	Error Rate (%) (FP+FN)/(TP+FN+FP+TN)	Sensitivity (%) TP / (TP+FN)	Specificity (%) TN / (TN+FP)	G-mean (%) √(Sensitivity)*(Specificity)	Precision (%) TP/(TP+FP)	Recall (%) TP / (TP+FN)	F-Measure (%) 2*Precision*Recall/(Precision+Recall)
FI-CR	2	66.26	33.74	82.58	54.16	66.99	56.86	82.58	67.44
Original-CR	1086	45.78	54.22	100	6.25	25	43.75	100	60.87
Optimal-CR	20	78.31	21.69	57.14	93.75	73.19	86.95	57.14	68.96

FI-CR → Frequent Itemsets as Classification Rules, **Original-CR** → Original Classification Rules, **Optimal-CR** → Optimal Classification Rules

Table 1: Comparison analysis of different types of classification rules
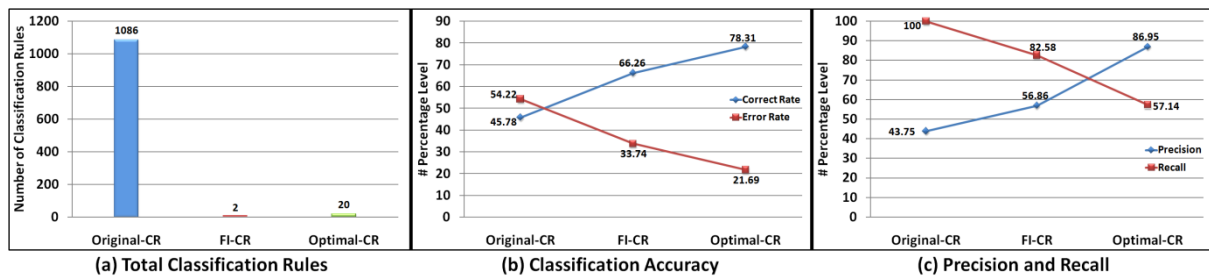

Fig 8: Effect of classification method using optimal-CR

The third experimental (concentrating only on the number of ants), here control parameter (number of ants) of classification method (O-cREAD) continuously increased by 2 and the number of generation is constant at 10. Then, as shown in figure 7(a) average classification accuracy of classification method is $71.83\pm3.15\%$ and average error rate is $28.17\pm3.15\%$ of classification method. The average sensitivity & specificity of classification method are $56.06\pm3.39\%$ & $83.33\pm4.72\%$ respectively and average gmean factor for analysis sensitivity & specificity is $68.30\pm2.89\%$ as shown in figure 7(b). The average precision & recall of classification method are $71.45\pm6.83\%$ & $56.06\pm3.39\%$ and average f-measure factor for analysis of precision & recall is $62.70\pm3.65\%$ as shown in figure 7(c).

Above experimental results show that performance of classification method (O-cREAD) is depending on different control parameters of the hybrid optimization approach for extracting optimal classification rules for classification of multi-class audio data.

The comparison analysis is based upon different types of classification rules used for a classification method. The experiments are implemented in sequence; Firstly, a classification method using FI-CR, a second classification method using original-CR (i.e. generated from frequent itemsets using the association concept), and last classification method using optimal-CR. The experimental results of these classification methods are as shown in table 1, where classification using original-CR and FI-CR are viewing as worst case & average case respectively. The classification method using optimal-CR is the best case as compared to the classification method using original-CR and FI-CR.

As shown in figure 8(a), the classification method using original-CR then rule size is 1086 and these rules provide 45.78% accuracy for classifying multi-class audio dataset. The classification method using FI-CR then rules are 2 and classification accuracy is 66.26%, on the other hand when the classification method using optimal-CR then rule size is 20 with 78.31% accuracy for classifying classifying multi-class audio dataset. As shown in figure 8(b), it provides an overview about the classification accuracy at different cases. The classification accuracy of classification method using optimal-CR is 78.31% as compared to; classification method using original-CR its improvement is 32.53% and classification method using FI-CR its improvement is 12.05%. On the other hand, the error rate of classification method using optimal-CR is 21.69% as compared to; classification method using original-CR its decreases with 32.53% and classification method using FI-CR and its decreases with 12.05%. As shown in figure 8(c), it provides an overview about the precision at different cases. The precision of classification method using optimal-CR is 86.95% as compared to; classification method using original-CR its improvement is 43.20% and classification method using FI-CR its improvement is 30.09%. On the other hand, the recall of classification method using optimal-CR is 57.14% as compared to; classification method using original-CR its decreases with 42.86% and classification method using FI-CR its decreases with 25.44%.

So, on the basis of above comparison analysis it shows that the classification method using optimal classification rules (O-cREAD) for classification of audio dataset is more efficient as compared to the other two methods.

7. CONCLUSION

This paper proposed a new classification method (O-cREAD) for multi-class audio data and this method is based on the new hybrid optimization approach to extract optimal classification rules. The main objective of the hybrid optimization approach in this O-cREAD method is to reduce classification rules in size and to maximize the classification accuracy rate of these rules. The classification accuracy of a classification method using optimal-CR presented the best accuracy rate among the other methods using original-CR and FI-CR. Although classification accuracy rate of O-cREAD method highly depends upon control parameter settings. By some further procedural improvements like more efficient pre-processing, complexity reduction and data compression method can be make this proposed O-cREAD method to a classical classification approach for multi-class audio data.

8. REFERENCES

[1] R. Agarwal, R. Srikant, "Fast algorithm for Mining Association Rules", In Proceeding of 20th VLDB Conference, Santiago, Chile, 1994.

[2] N. Leavitt, "Let's Hear It for Audio Mining", IEEE Computer Magazine on Technology News, ISSN: 0018-9162, Volume 35, Issue 10, October 2002, pages 23-25.

[3] M. K. Gupta, G. Sikka, "Association Rules Extraction Using Muti-Objective Feature of Genetic Algorithm", Proceeding of World Congress on Engineering and Computer Science, Vol. 2, 23-25 October, 2013, San Francisco, USA.

[4] J. Holland, "Adaption in Natural and Artificial System", University of Michigan Press, Ann Arbor, Michigan; Reissued by MIT press 1992.

[5] J. F. Frenzel, "Genetic Algorithms a New Breed of Optimization",IEEE, October 1993, page 21-24.

[6] M. Dorigo, V. Maniezzo, A. Colorni, "Ant System: Optimization by a colony of cooperating agents", IEEE Transactions on Systems, Man, and Cybernetics-part B, Vol. 26, No. 1, 1996, page 29-41.

[7] M. Dorigo, M. Birattari, T. Stiitzle, "Ant Colony Optimization Artificial Ants as a Computational Intelligence Technique", IEEE Computational Intelligence Magazine, November 2006, page 28-39.

[8] M. Dorigo, G. D.Caro, L. M. Gambardella, "Ant Algorithms for Discrete Optimization", Artificial Life, MIT Press, 1999.

[9] R. Kaur, and S. Aggarwal, "Association Rule Mining for Punjabi Text" International journal of Computer Science and Technology (IJCST). Volume 4, Issue 1, Jan - March 2013, pages 404-406.

[10] Y. Okada, T. Tada, K. Fukuta and T. Nagashima, "Audio classification based on a closed itemset mining algorithm" International Conference on Computer Information Systems and Industrial Management Applications (CISIM), IEEE, 2010, pages 60-65.

[11] A. Ghosh, B. Nath, "Multi-Objective Rule Mining Using Genetic Algorithm", Information Science, Elsevier, Vol. 163, 2004, page 123-133.

[12] S. Ghosh, S, Biswas, D. Sarkar, P. P. Sarkar, "Association Rule Mining Algorithm and Genetic Algorithm: A Comparative study", Third International Conference on Emerging Applications of Information Technology (EAIT), IEEE, 2012, page 202-205.

[13] R. L. Haupt, S. E. Haupt, "Practical Genetic Algorithms" A Wiley-Interscience Publications, USA, 2004.

[14] J. Han, M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, USA, 2001.

[15] A. Reda, S. Panjwani, E. Cutrell, "Hyke: A Low-cost Remote Attendance Tracking System for Developing Regions", The 5th ACM Workshop on Networked Systems for Developing Regions (NSDR), June 2011.