

# Sequential Covering Strategy Based Classification Approach Using Ant Colony Optimization

Revathi.B , Chidambaram.S

Dept of Information Technology, National Engineering College, Kovilpatti, Tamilnadu, India.

Dept of Information Technology, National Engineering College, Kovilpatti, Tamilnadu, India.

**Abstract**— Ant Colony Optimization (ACO) algorithm has been implemented to discover a list of classification rules. In the proposed system, a sequential covering strategy for ACO Classification algorithm used to remove the problem of interaction. In this finding the “best” rule that accounts for a part of the training data. Adding the best rule to the induced rule set and removing the data it covers. This iterative process continues until no training instance remains. Heuristic function is used for selecting a best rule and calculating the predictive accuracy. In the proposed method, weather-nominal data set has been used in order to calculate the predictive accuracy.

**Keywords**— Data Mining, Ant Colony Optimization, rule based classification, Classification and Sequential covering strategy.

## I.INTRODUCTION

Data Mining, the extraction of hidden predictive information from large amount of databases, is a powerful technology with great potential to help companies focus on the most important information in their data warehouses techniques. Data mining tools predict future trends and behaviours, allowing businesses to make proactive. In

essence, the main goal of data mining is to extract the knowledge from data. Data mining is the interdisciplinary field, whose core is at the intersection of machine learning, statistics and databases [1].

Also the Data mining tools can answer the business questions that traditionally were too consuming to resolve .They scour databases for hidden data’s finding predictive information that experts may miss because it lies outside their expectations. Data mining techniques are the result of a very big process of research and product development.

Classification [2], is one of the most frequently occurring tasks of decision making. Many decision making problems in a variety of domains, such as medical sciences, engineering, management sciences, and human sciences. It can be considered as the classification problems. Most popular examples of classification, credit scoring, speech recognition, and character recognition. In data mining, various classification techniques are used to involve and to solve the stochastic as well as probabilistic problems.

Optimization plays a fundamental role in Data mining through Support vector machine (SVM), rule based classification and many of the techniques. Optimization is the combination of formulation and solution.

In this paper an Ant Colony Optimization is used to evaluate the various datasets and to find which value is the good one for making the rule. From this we have to discover the new algorithm called Sequential covering strategy based Ant colony Optimization (Seq-ACO). It is used to describe if –then rule based classification rules. If-then rules [5], are the basis for some of the most popular concept description languages used in machine learning. They allow knowledge extracted from the various dataset to be represented in a form that is easy for people to understand. Finally it will compare which value is the best one and provide the best accuracy. It will done under in the process of the Optimization

This paper is structured as follows: In section II presents the algorithm of Ant colony Optimization. Section III

presents the Methodology of the proposed work. Section IV discussed about the proposed work Sequential Covering Strategy based ACO. The experimental results are presented in Section V. Section VI concludes the paper.

**II.LITERATURE SURVEY**

In this survey various rule based classification algorithms are discussed and compared with the proposed work. U. Fayyad and K. Irani, [15], they presented the classification learning algorithms that generate decision trees or rules using information entropy minimization heuristic for discretizing continuous valued attributes. The result serves to give a better understanding of the heuristic measure, to point out the behavior of the information entropy process desirable properties that justify its usage in a formal sense, ant to improve the efficiency of calculating continuous valued attributes for cut value selection.

G.Venturini [2] decribed a genetic learning system called Supervised inductive learning algorithm, which learns attributes based rules from a set of pre-classified examples. The rule learning reduces greatly the possible rule search space when compared to the genetic Michigan approach. R. Parpinelli, H. Lopes, and A. Freitas [4] they prosed ant colony based system for data mining. The main objective of this algorithm is the extraction of classification rules to be applied to hidden data. It is used to discover such rules is inspired in the behavior of a real ant colony, as well as some concepts of information theory and data mining.

K. Salama, A. Abdelbar, and A. Freitas [9] proposed five extensions to Ant Miner .At first they utilize multiple types of pheromone, one for each permitted rule class and they use a quality contrast intensifier to magnify the high quality of rules. They allow the use of logical negation operator in the antecedents of constructed rules. They incorporate the stubborn ants; an ACO variation in which an ant is allowed to take into consideration .Finally the ant used its own values.

F. Otero, A. Freitas [7] they presented a paper an extension of Ant Miner (coping with continuous attributes ,which incorporates an entropy based discretization method. n future it would be interested to extend the entropy based discretization method used in the rule construction process to allow the creation of intervals between the lower and upper bound of the intervals.

**III.ANT COLONY OPTIMIZATION (ACO)**

Ant Colony Optimization [3] successfully applied to discover the list of classification if-then rules. It is one of the techniques for solving computational problems which can be used to find the good paths through graphs based on the state of real ant’s behavior.

It supports both nominal and numeric attributes. Rule induction is an area of machine learning in which if-then production rules are extracted from the set of observations. It is used to simulate the behaviour of real ants using a colony of real ants which cooperate in finding good solutions to optimization problems. It is used to find the shortest path between the food and source. This is the general behaviour of an ant. This way of approach is used

the Ant colony optimization.

In this algorithm, first it starts with an empty rule list. [11]Each ant starts with a rule with no term in its antecedent (empty rule), and adds one term at a time to its current partial rule. The current partial rule constructed by an ant corresponds to the current partial path followed by that ant.. An individual ant constructs candidate solutions by starting with an empty solution and then iteratively adding solution components until a complete candidate solution is generated. The ACO algorithm supports both nominal and numeric attributes. [13]The algorithm consists of two loops. Outer (while loop)-where a single rule in each iteration is added to the discovered rule list. Inner (repeat-until loop)-where an ant in each iteration constructs a rule as follows. Each ant first probabilistically select a one term at a time in a current partial rule.

If there is any duplicating terms are involved in the iteration, using the pruning technique to remove the irrelevant terms and conditions. It is used to improve the accuracy and also the quality of the rule will be updated in the iterative process. The inner loop checks the rule quality and updates the quality in the correspondent rule list. This process continues until there is no item remaining in the training data set.

**IV.METHODOLOGY**

For Optimization using Ant Colony Algorithm the following sample datasets are taken to produce the list of rules. They include Weather, Breast cancer, contact lenses and so on. This algorithm applied in the rule induction with sequential covering algorithms, which will provide the best list of rules instead of using list of rules. The proposed methodology follows the ACO technique then it calls it as Seq-ACO.

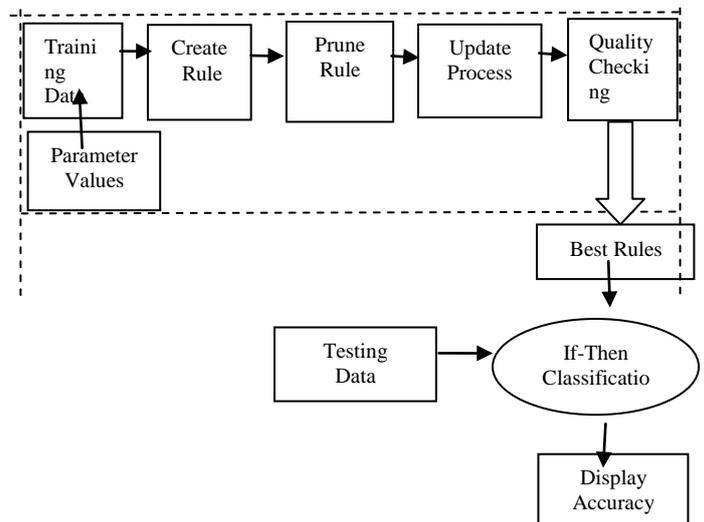


Fig1: Block diagram of proposed work

The overall Block diagram of our proposed work is to create the possible if-then rules from the overall example training data evaluated based on the quality. At first, we are getting the input parameter from the user. From this, the ACO automatically generate the if-then rules. Ants will create the rules. It checks whether if there is any irrelevant terms involved in the training datasets or not. If appear means, use the pruning technique and also updates the process. The quality between the two Ants are measured which will produce the maximum accuracy. The ACO consider the three parts: A.ACO rule Creation/Cross validation, B.ACO rule updating, and C. System Parameters used.

*A. Aco Rule Creation/Cross Validation*

The ACO Rule creation is done by the Ant Colony Classification algorithm .It is applied for ACO Rule Creation/Cross Validation/Rule Updating. The Classification algorithm classifies the rules and then creates rule creation process. It aims at extracting the if-then classification rules [3]. The if-then rules is of the form IF (term1) AND (term2) .....AND (term N) THEN (class C). Each term in the rule is a triple (attribute, operator, values).The “If” part is the rule antecedent and the “Then” part is the rule consequent, which represents the class to be predicted by the rule. The rule is created by an ant. At each iteration, a rule is created by an ACO procedure.

The goal of Ant-Miner [12], is to extract classification rules from data. The algorithm is presented in Figure 1.

```

ALGORITHM I HIGH LEVEL PSEUDO CODE OF ANT-MINER
Input: training examples
Output: discovered list of rules
examples ← training examples
list ← ∅ ;
while |examples| > maximum uncovered do
  InitializePheromones();
  ComputeHeuristicInformation(examples);
  rulegb ← ∅ ;
  m ← 0;
  while m < maximum iterations and not stagnation do
    ruleib ← ∅ ;
    for n ← 1 to colony_size do
      rulen ← CreateRule(examples);
      Prune(rulen);
      if Quality(rulen) > Quality(ruleib) then
        ruleib ← rulen;
      end if
    end for
    UpdatePheromones(ruleib);
    if Quality(ruleib) > Quality( rulegb ) then
      rulegb ← ruleib ;
    end if
    m ← m+1;
  end while
  examples ← examples – Covered( rulegb ,examples);
  list ← list+ rulegb ;
end while
return list;
    
```

Classification rule starts with an empty rule list and iteratively adds one rule at a time to that list until the number of uncovered training examples is greater than the user specified maximum value. In order to construct a rule, a single ant starts with an empty rule (no terms in its antecedent) and adds one term at a time to the rule antecedent. It probabilistically chooses a term to be added to their current partial rule based on the values of the amount of pheromone ( $\tau$ ) and a problem dependent Heuristic function ( $\eta$ ) associated with this term. [4], A pheromone value is associated with each possible term – i.e. each possible triple (attribute, operator, value). This process is called “rule construction”. Once this rule construction process is being finished, the rule is created by an ant. The Heuristic function is expressed as,

$$entropy(a_i = v_{ij}) \equiv \sum_{c=1}^k -P(c/a_i = v_{ij}) \log_2 p(c/a_i = v_{ij}) \quad (1)$$

From this equation, the  $term_{ij}$  involves a measure of entropy where every  $term_{ij}$  has the form of  $p(c/a_i = v_{ij})$  which is the probability of observing class ‘c’, which satisfies the condition  $a_i = v_{ij}$  and k is the number of classes. Hence the rule is being created. The ACO algorithm is added to the classifier section and hence classifies the if- then rules. An Ant Probabilistically selects an item defined by the following formula as shown below:

$$P_v(t) = \frac{\tau_v(g) \eta_v(g)}{\sum_{i=1}^{a_i} X_i \sum_{j=1}^{b_j} [\tau_{ij}(g) \eta_{ij}(g)]} \quad (2)$$

Where ‘a’ and ‘b’ are the number of attributes, and ‘t’ is the time duration of the ant.  $\tau$  - The amount of pheromone updating and  $\eta$  is the problem dependent heuristic value. X is the selection of an item from the current partial rule. Hence the rules will be added under the rules category of the classifier panel.

After completion of the Rule creation process the ant automatically detects whether the irrelevant terms (or) items are present in the rule or not. If present, the Pruning technique is used to delete those terms. The ACO algorithm is inserted in the Data mining tool and it is evaluated by the process of Cross validation and hence the updating process is done. In the classifier section, the added ACO algorithm is selected. Then the ACO cross validation action is performed. It iteratively chooses any number of cross validation for both training and testing data’s. The Parameter values are taken in the cross validation section. It evaluates the best analysis process.

*B.ACO Rule Updating*

ACO Rule Updating is another part of ACO procedure. After completing the Cross validation process, the ACO updates the pheromone and heuristic information. Pheromone trails are first initialized for updating the pheromone values. (i.e., in which path the rule will be

correctly classified and that value will be calculated and then updated). Based on the pheromone updating process, the ant chooses the path. The updating process based on the probability of an ant is expressed by,

$$P_{v_j} = \frac{\tau(t, v_i, v_j) \eta_{v_j}}{\sum_{k=1}^{F_{v_k}} \tau(t, v_i, v_j) \eta_{v_k}} \quad (3)$$

Where,  $\tau(t, v_i, v_j)$  is the amount of pheromone associated with the entry  $\tau(t, v_i, v_j)$  in the pheromone matrix,  $\eta_{v_j}$  is the heuristic information associated with one term path  $v_j$ , and  $F_{v_k}$  is the set of adding overall items of  $v_i$ . The ant will find the best rule based on the amount of pheromone value. If one item is selected in the pheromone matrix, all the remaining items are added in the form of summation, all the items are divided by the amount of probability for selecting one item. Finally pheromone trails are updated using the best rule quality. The current iteration and the best rule added so far across all iterations are stored and updated. The pheromone updating process is to cope with  $term_{ij}$  (a term that represents a continuous attribute  $a_i$ ). After this, the ACO rule creation Quality will be checked. The Quality function  $Q$  is defined as,

$$Q = \frac{Truepositive}{Covered} \quad (4)$$

True positive is the number of correctly classified records and covered is the total number of correctly classified records for both training and testing data's.  $Q$  provides the overall testing set accuracy, total number of terms and the total number of rules. The Overall results of the ACO Cross validation, taking the dataset records for each and every cross validation randomly generates the testing data records.

### C. System parameters

These parameters are used to analyze the best one in the overall cross validation. The parameters are as follows:

#### 1, Number of Ants:

This is the first parameter in the ACO algorithm. It creates the rule iteratively. The ant takes the input from the user; and user defined number will be generated. At first the rule is empty and after adding the term in the rule one by one in the iterative process, the rule selection is based on quality basis.

#### 2, Minimum number of cases per rule:

The above parameter takes the smaller number of terms (i.e., one or two terms are taken) for a single rule. Minimum numbers of terms (or) conditions are covered by a single rule, by which the rule quality will be improved.

#### 3, Number of iterations:

Numbers of iterations are needed to find the best rule in the list of rules. The Rule Creation, Rule pruning, and rules addition are performed.

Copyright to IJIRSET

#### 4, Maximum uncovered cases:

When the number of iterations reaches the user defined maximum uncovered value, the iteration process is stopped.

#### 5, Convergence test:

The convergence test finds the best rule in the overall list of rule. This test produces the best rule.

## V. SEQUENTIAL COVERING STRATEGY BASED ACO (Seq - ACO)

In Ant-Miner Sequential Covering Strategy is work as follows, applying the Seq-ACO algorithm in ACO classification provides the best list of rules. This algorithm is followed by the Pittsburgh approach which is one of the rule based classification algorithm. It is followed by the ACO procedure which discovers the complete list of rules (i.e. the best list of rules produced the overall iterations). It also contains the if-then part. It is used to find the best list of rules given in the training data records. For each iteration, the rule is constructed by the ACO and the examples that are covered by the rule are included and the remaining example data are removed from the dataset. This iterative process is continued until the training data becomes empty.

The following algorithm is used to describe the Sequential Covering Strategy based ACO. The algorithm is presented in Figure 2.

#### ALGORITHM 2 HIGH LEVEL PSEUDO CODE OF SEQ-ACO

Input: training examples

Output: best discovered list of rules

InitializePheromones();

$list_{gb} \leftarrow \varnothing$  ;

$m \leftarrow 0$ ;

while  $m < \text{maximum iterations and not stagnation do}$

$list_{ib} \leftarrow \varnothing$  ;

for  $n \leftarrow 1$  to colony\_size do

examples  $\leftarrow$  all training examples;

$list_n \leftarrow \varnothing$  ;

while  $|examples| > \text{maximum uncovered do}$

ComputeHeuristicInformation(examples);

$rule_n \leftarrow \text{CreateRule(examples)}$ ;

Prune( $rule_n$ );

Examples  $\leftarrow$  examples - Covered(  $rule_n$ , examples);

$list_n \leftarrow list_n + rule_n$  ;

end while

if  $Quality(rule_n) > Quality(rule_{ib})$  then

$list_{ib} \leftarrow list_n$ ;

end if

end for

UpdatePheromones( $list_{ib}$ );

if  $Quality(list_{ib}) > Quality(list_{gb})$  then

$list_{gb} \leftarrow list_{ib}$  ;

end if

$m \leftarrow m+1$ ;

end while

return  $list_{gb}$  ;

This is the method to extract the rules from the decision tree. Each branch in the tree corresponds to a rule that has its own consequent and conjunction of the attribute value pairs. The Rule Sets induced by sequential Covering Algorithms may be ordered and unorderd. That is the best rule is chosen by entropy criterion methodology such as accuracy. Highly Specific rules may have high accuracy on given training example instances. Normally it performs a

Num of Ants	Min number of cases per rule	Max uncovered cases	Convergence test	Num of iterations	Percentage(%)
6	2	5	4	3	45
4	2	2	5	3	80
2	4	5	6	3	50
2	5	6	7	3	30
5	4	4	2	3	45
4	2	5	4	3	80
5	4	2	4	3	60
6	2	4	6	3	35
2	5	4	7	3	45
6	4	2	3	3	40

greedy search this searching category keeps track of the most promising candidate rules according to the current top rated hypothesis. Also this strategy that keeps looking for the locally best candidates.

In this global iteration the best rule will be produced by the Seq-ACO. The Sequential ACO is used to provide smaller number of greater rules. Rules are discovered once at a time, the resulting rule will not affect the previous set of rules because the search space is changed due to the removal of training examples covered by the previous rules. This approach is used in order to solve the problem of rule interaction. Once a rule is created and pruned, the old training examples covered by the rule are removed and the new rule which is created is added to the current list of rules. It also uses the same user defined parameter.

When an ant finishes the rule list creation process, the iteration best list is updated. Newly created list has a better quality than the iteration best list. The quality of the iteration best list is greater than the quality of the global best list. Finally the Seq-ACO provides better predictive accuracy from the best list of rules.

Why we are using the rule based classification method means? It is very easy to understand the Knowledge ‘content’, explanation for the reasoning is easily shown, modification is easy, provided the rule base structured well, and finally uncertainty can incorporated into the knowledge base.

**VI. EXPERIMENTAL RESULTS**

The major difference between the sequential ACO and the ACO is to provide the best list of rules instead of producing the list of rules. By comparing with ACO algorithm Seq-ACO provides good performance results. For example, ACO, Seq-ACO analysis is weather-nominal dataset. In this there are 5 Attributes are used. They are Play, temperature, humidity, outlook, windy. The first

attribute is predictable one. Rule can be generated either from a decision tree or directly from the training data using sequential covering algorithms. The rule induction using sequential covering is one of the important algorithms, which will produce the best list of rules. In this algorithm using the rule based classification. The graph represented as the analysis of both ACO and Seq-ACO. For the analysis process 5 parameters have been used. Changes are done only for the convergence test values and the remaining all parameter values are assigned 3.

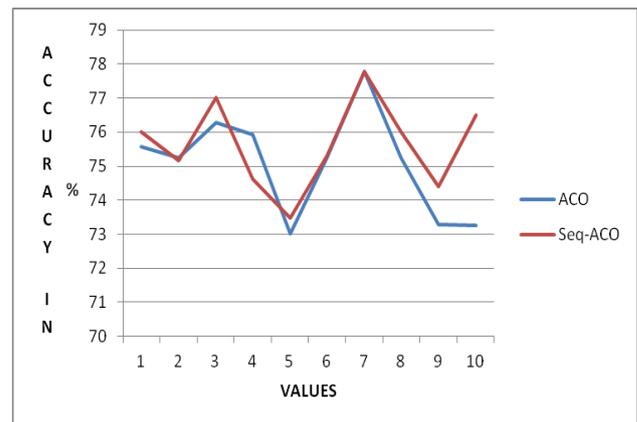


Fig. 1 Comparisons results for ACO and SeqACO

TABLE I.

**AVERAGE COMPUTATIONAL RESULTS FOR ACO**

Num of Ants	Min number of cases per rule	Max uncovered cases	Convergence test	Num of iterations	Percentage(%)
6	2	5	4	3	45
4	2	2	5	3	50
2	4	5	6	3	35
2	5	6	7	3	45
5	4	4	2	3	35
4	2	5	4	3	60
5	4	2	4	3	45
6	2	4	6	3	55
2	5	4	7	3	40
6	4	2	3	3	35

TABLE II.

AVERAGE COMPUTATIONAL RESULTS FOR SEQ-ACO

Num of Ants	Min number of cases per rule	Max uncovered cases	Convergence test	Num of iterations	Percentage(%)
6	2	5	4	3	45
4	2	2	5	3	80
2	4	5	6	3	50
2	5	6	7	3	30
5	4	4	2	3	45
4	2	5	4	3	80
5	4	2	4	3	60
6	2	4	6	3	35
2	5	4	7	3	45
6	4	2	3	3	40

VII.CONCLUSION

The proposed method consists of two modules i.e. ACO rule Creation/Cross validation/Rule Updating and Sequential covering strategy based ACO. At first the ACO algorithm is applied for discovering the list of rules. By using this approach the best rule is found by the Seq-ACO. The sequential covering strategy based Ant Colony Optimization has been tested for various dataset records, which supports only nominal and continuous attributes. Sequential ACO is used to provide smaller number of greater rules. Rules are discovered once at a time, the resulting rule will not affect the previous set of rules because the search space is changed due to the removal of training examples covered by the previous rules. It is mainly used to mitigate the problem of rule interaction. That is why the accuracy of Seq-ACO is

greater than that of ACO. Both the modules are completed and it is concluded that the Seq-ACO is the most accurate algorithm, which achieves statistically and significantly higher predictive accuracy than ACO.

REFERENCES

- [1] Fernando E.B. Otero, Alex A. Freitas, and Colin G. Johnson “A New Sequential Covering Strategy for Inducing classification rules with Ant Colony Algorithms” IEEE transactions on Evolutionary Computation, vol,17, no.1, February 2013.
- [2] G Venturini, “SIA: A supervised inductive algorithm with genetic search for learning attributes based concepts,” in Proc. Mach. Learn.: ECML, 1993, pp. 280–296
- [3] Frank.E and Witten.I, “Generating accurate rule sets without global optimization,” in Proc. 15th Int. Conf. Mach. Learn., 1998, pp. 144–151.
- [4] R. Parpinelli, H. Lopes, and A. Freitas, “An ant colony based system for data Mining applications to medical data”, in Proc.GECCO, 2001, PP.791-798.
- [5] R. Parpinelli, H. Lopes, and A. Freitas, “Data mining with an ant colony optimization algorithms”, IEEE Trans. Evol.Computat,Vol.6, no. 4,pp.321-322, Aug 2002.
- [6] B.Liu, H.Abbass, and B.McKay, “Density based heuristic for rule discovery With ant-miner”, in Proc, 6th AJWIS, 2002,pp .180-184. 2004.
- [7] F.Otero,A.Freitas, and C.Johnson, “cAnt-Miner: An ant colony classification Algorithm to cope with continuous attributes”, in Proc , 6th Int. Conf. Swarm Intell.(ANTS), vol.5217 ,2008 .pp. 48-59.
- [8] F. Otero, A.Freitas, and C. Johnson,” Handling continuous attributes in ant Colony classification algorithms”, in Proc .IEEE Symp. CIDM,Mar-Apr 2009, pp.225-231.10.
- [9] K.Salama,A.Abdelbar, and A.Freitas, “Multiple Pheromone types and other Extensions to the ant-miner classification rule discovery algorithm”, Swarm Intell., vol.5 nos.3-, pp.149-182 ,2011.
- [10] G.Piatetsky- Sharpio and W. Frawley, Knowledge Discovery in Databases. Cambridge, MA: AAAI Press,1991, pp 1-525.
- [11] V.Dhar, D.Chou, and F.Provost, “Discovering interesting patterns for investment decision making with GLOWER: A genetic learner overlaid with entropy reduction,”Data mining Knowledge Discovery, vol.4, no.4, pp.251-280,2000.
- [12] J.Quinlan, C4.5: Progrmas for Machine learning, San Mateo,CA: Morgan Kaufmann, 1993, pp,1-302.
- [13] J.Quinlan, “Improved use of continuous attributes in C4.5,”Artiff.Intell.Res.,vol.4, pp.77-90 Jan.1996.
- [14] M.Dorigo,G.D .Caro and L.Gambardella,”Ant algorithms for discrete Optimization,”Artiff.Life.vol.5, no.2, pp.137-172,1999.
- [15] U.Fayyad.U and Irani.K, “On the handling of continuous Valued attributes in decision tree generation,” Mach. Learn., vol.8,no.1,pp.87–102,Jan.1992.