

AN ALGORITHM FOR MINING USABLE RULES USING A HOLISTIC SWARM BASED APPROACH

¹Veenu Mangat and ²Renu Vig

¹Information Technology, University Institute of Engineering and Technology,
Sector 25, Panjab University, Chandigarh, 160014, India

²University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Received 2013-09-02; Revised 2013-10-10; Accepted 2013-12-09

ABSTRACT

Evolutionary algorithms are capable of finding near optimal solutions to problems which are intractable to solve using conventional methods. One such problem is to accurately classify patients using rule mining methodology while controlling the size of output rules. A massive amount of data pertaining to medicine is generated and recorded daily. Uncovering useful knowledge and assisting decision makers in the diagnosis and treatment of diseases from this vast data has become imperative. Association rule mining is an obvious choice for representing this previously hidden information as rules are simple to understand and infer. These rules can be used to understand the etiology of diseases and classify patients based on recorded characteristics. The interestingness of such an algorithm for rule mining will be determined by its accuracy and ability to produce easily understandable rules. This study applies latest improvements in swarm intelligence to devise a novel strategy for rule mining that exhibits high predictive accuracy and comprehensibility. It has been applied over four medical datasets to classify patients as fit or unfit. The paper begins with an explanation of rule mining functionality and concept of swarm intelligence. The current techniques for rule mining in the medical domain are surveyed and their shortcomings are identified. This is followed by a description of the proposed algorithm which includes a novel rule discovery procedure and a novel rule list selection criterion. The results of the proposed algorithm thus obtained, are compared with the other best known approaches. Finally, the future scope of work in this area is briefly discussed.

Keywords: Association Rules, Particle Swarm Optimisation, Ant Colony Optimisation, Quality Function

1. INTRODUCTION

Data Mining is an automated process that analyses and explores huge amounts of data in order to unravel novel and consistent trends and relationships between entities or their attributes. These previously unknown patterns are then validated before being output by applying them over new data. The discovered knowledge should be accurate and comprehensible. High level knowledge representations like association rules are the most intuitive choice to achieve knowledge comprehensibility. Association rule mining is a data

mining functionality that discovers correlations, frequent patterns, associations or causal structures within characteristics or dimensions of data present in the databases. An association rule is an implication in the form of $X \rightarrow Y$, where X, Y are attribute-value pairs and $X \cap Y = \phi$. X is called antecedent while Y is called consequent. The meaning of rule is that presence of X implies presence of Y with a measure of certainty associated with it. If Y relates to a single attribute value pair, such an association can be used by mining engine to classify records. There are various subjective and objective criteria to evaluate a rule's worth for a data

Corresponding Author: Veenu Mangat, Information Technology, University Institute of Engineering and Technology,
Sector 25, Panjab University, Chandigarh, 160014, India Tel: 91-9814475450

mining system user. Two basic objective criteria are support and confidence. Support of an association rule is computed by taking a count of those records that contain both antecedent and consequent and dividing by total number of records present. Confidence of an association rule is computed by taking a count of those records that contain both antecedent and consequent and dividing by total number of records that contain the antecedent. Support measures coverage and confidence measures certainty. Acceptable thresholds for minimum values of support and confidence are predefined as percentages. Different applications set various thresholds on these measures and only those rules are output which satisfy these minimum thresholds. The first step in association rule mining problem is to find those terms which occur more than predefined support number of times in the database. These are called frequent terms. The second step combines these terms into rules and outputs those rules that occur more than confidence threshold number of times in the database. Finding frequent terms is done incrementally. Those terms which may be frequent based on count of their components, are known as candidate terms. As the size of modern medical datasets is very large and constantly growing, most of the known algorithms output an overwhelmingly large number of association rules when used over these databases. The problem is aggravated by the fact that each association rule contains multiple terms within itself. This makes it virtually impossible for the end users, medical practitioners in this case, to understand or verify such a huge number of big and complex association rules. Hence these mined rules and consequently knowledge discovery systems are seldom adopted.

Evolutionary methods of soft computing paradigm can perform well in problems with vast search spaces and produce near optimal solutions. Soft computing is a paradigm that trades an exact optimal solution which might be expensive to find, for an approximate solution that is low cost and guaranteed to be found. It uses concepts of imprecision, uncertainty, partial truth and approximation. An Evolutionary Algorithm (EA) is a generic population-based metaheuristic optimization algorithm. It uses some concepts derived from evolution of biological beings like reproduction, mutation, recombination and selection. Swarm Intelligence is a part of these innovative techniques and it uses concept of distributed adaptive self learning for finding a solution to optimization problems. This is based on behaviour of various biological species of animals. These techniques can be used in data mining to find a target solution to the problem when other methods are difficult to implement.

2. CONVENTIONAL RULE MINING METHODS

2.1. Review of Literature

Silva and Neves (2004) was a seminal paper that suggested using Particle Swarm Optimization for performing data mining tasks. Variants based on Particle Swarm technique were developed and used for classification. Results were compared against a Genetic Algorithm and a Tree Induction Algorithm (J48). Particle Swarm Optimizers gave promising results in terms of accuracy. This was an indicator that PSO based techniques have the potential to compete with other evolutionary techniques as well as industry adopted algorithms like J48 and C4.5 algorithm. They should be tested in more problem domains to determine their suitability for mining.

Another method for rule mining based on an evolutionary (GA) approach-EGAR was described in (Kwasnicka and Switalski, 2005) and compared with FP tree method. It was found that FPtree works well for discrete attributes whereas EGAR performs better with mix of discrete and continuous attributes.

An advanced swarm intelligence data mining algorithm was proposed in (Ghannad-Rezaie *et al.*, 2006). The method has additional functions to handle missing tuples and/or unavailable attribute values. It also facilitated discovery of rules based on dynamic user input. The method was applied to select candidates for surgery for temporal lobe epilepsy. Four algorithms: decision tree, ant colony miner, PSO miner and the proposed hybrid PSO, were compared. C4.5 gave better average predictive accuracy but the rule set was larger and individual rules were bulky and did not cover much data. PSO converges faster than ACO but has similar accuracy values. C4.5 converges very fast, but may lead to overfitting. It should be used only on large balanced datasets. The hybrid PSO is faster than other evolutionary variants but has slightly more memory usage. It uses a combination of support vector machines and radial basis functions in conjunction with PSO.

Ordonez (2006) introduced an efficient provision to add domain specific constraints to guide search for rules. Decision Tree Induction (DTI) was performed over same dataset. It was proven that constrained association rules were more effective than predictive rules of DTI in predicting disease with various characteristics that are related to each other. DTI rules have lower confidence, hence lower certainty and result in immensely long rules since trees are unrestricted. Also they suffer from the problem of some overfitting and data set fragmentation.

A hybrid method that combines Particle Swarm Optimization and Ant Colony Optimization (PSO/ACO) was introduced in (Holden and Freitas, 2008) for mining rules that can be used for classification. The disadvantage of using PSO algorithm was that nominal values had to be converted into binary numbers before mining. The hybrid algorithm suggested eliminates the need for this preprocessing phase. The algorithm was compared to PART, which is an industry standard algorithm. Authors also compared performance of module handling only continuous data to another new classification algorithm based on differential evolution. Results indicate that this hybrid algorithm gives very good accuracy measure and outputs simpler (smaller) rule sets, thus achieving twin goals of correctness and understandability. The results also show that the PSO component that handles continuous data gives slightly higher accuracy than the differential evolution algorithm.

Association rule mining has been applied for discovering hyperlipidemia from biochemistry blood parameters in (Dogan and Turkoglu, 2008). The philosophy of combining PSO and ACO approaches to mine data was also used in a pharmacovigilance context in (Sordo *et al.*, 2009). The approach was able to extract previously hidden cause-effect relationships between therapeutic care, patient attributes and external events. These patterns were detected with a high degree of accuracy.

The standard GA has been enhanced with features like duplicate check to perform rule discovery in data mining in (Cattral *et al.*, 2009). But there were still issues to be addressed. All quantitative attributes are not easy to partition or discretize even with user input. Also the minimum support level cannot be easily defined. If the number of attributes and their values is very large, it leads to a large search space and very big output rule set. The proposed EARMGA algorithm can mine high quality rules without the user specifying minimum support or confidence threshold levels (Yan *et al.*, 2009).

Databases with vast number of attributes each having thousands of values are difficult to mine due to the sheer volume of the data. This issue can be dealt with by combining a traditional Genetic Programming based mining method with a specially designed Genetic Algorithm (GA). The process starts by dividing the input database into many small databases. Each of these small databases make up an individual and together they all form a population. The traditional GNP based mining method is applied to extract association rules for each of these

individuals. Lastly, GA with some special operators is run for several generations and the population is evolved iteratively. The results show that this combined method allows discovering association rules from huge data-dense databases directly and more efficiently than the traditional GNP method alone (Gonzales *et al.*, 2009).

As stated earlier, a primary issue in adoption of rule mining systems is the enormous number of rules output as result. The problem can be tackled effectively by using the Minimum Description Length principle (MDL): Choose the result set that is the most compressed. Using the Krimp algorithm for frequent itemset mining and classification, a dramatic reduction in the number of returned frequent item sets is obtained. This algorithm outputs code-tables which are small and accurate. Krimp classifier when executed over a large range of datasets, gives good accuracies and compression ratios. It has been tested by swapping and randomizing data values. The algorithm does not involve elaborate parameter setting and is very stable over skewed and imbalanced datasets. There are many data mining tasks for which it can be used eg: frequent itemset mining while preserving privacy (Vreeken *et al.*, 2011).

Many variants of the Ant-Miner in terms of heuristic information, pheromone update, rule construction and pruning procedures have been reviewed in (Martens *et al.*, 2011).

In order to find causes of defects, a data mining tool (DIFACONN-miner) was used for generating classification rules (Baykasoglu *et al.*, 2011). Differential Evolution (DE) algorithm is used for training phase of an Artificial Neural Network. The multi-objective function used by neural network is derived from three parameters: error of ANN, number of rules and accuracy of training phase. Then classification rules are generated using Touring Ant Colony Optimization (TACO) algorithm. Authors proved through experimentation that DIFACONN-miner generates useful and correct classification rules.

The evolutionary technique has also been applied to classification of images. It is a perfect candidate as the number of possible combinations of features of images is extremely high. Most of these features contribute to the classification step. An algorithm namely Simplified Swarm Optimization (SSO) has been developed recently. SSO, Particle Swarm Optimization (PSO) and Support Vector Machine (SVM) have been applied for classification of images. SSO is simple to use since it uses only one random number and three predefined parameters to update each of the particle's position. PSO requires several parameter settings like inertia weight

and velocity update. It has been shown that SSO gives better precision and recall metrics than PSO and SVM. It uses lesser memory and time to provide comparable accuracy as it employs lesser number of particles. Specifically, PSO uses substantially more memory resources than SSO per particle to achieve higher performance (Wahid, 2011).

Four techniques based on swarm intelligence were studied and implemented in (Mangat, 2011) and they provided good accuracy values when compared to other traditional non SI based mining techniques. A new quality criterion combined with Shuffled frog leaping technique, showed very good results.

ABCMiner, an algorithm to mine rules using concepts from Artificial Bees Colony, was suggested in (Shukran *et al.*, 2011). The algorithm modified the search strategy to model the real behaviour of worker and scout bees. The rule format and fitness function were also redefined. Accuracy of this method was quite high for classification. The original Ant-Miner algorithm was modified to handle continuous attributes during the rule construction process in a technique called cAnt-Miner. It discretizes the continuous data values for each such attribute dynamically and then applies ant colony based pheromone updation and rule construction processes. The advantage is that it can handle nominal and continuous data without needing a separate data discretization technique in a preprocessing phase (Fernando *et al.*, 2008).

This algorithm was further extended to evaluate the quality of a rule set overall in addition to quality of individual rules in (Fernando *et al.*, 2012). The pheromone matrix used by the ACO algorithm is modified to include additional information. This information stores the identification value of each tour which basically gives the order in which rules should be created and generated so that overall rule list quality can also be maximized. Application of particle swarm optimization to suggest suitable threshold values for rule mining over real world stock database has been done in (Kuo *et al.*, 2011; Nouaouria, 2013) discusses a particle swarm classification technique that successfully deals with data containing a large number of attributes which in turn are of various types like discrete, continuous. It has been shown to be a promising technique when compared with other popular classification techniques.

2.2. Observations

A review of the current literature suggests that there is no single efficient scheme for rule mining that works

on all textual data types. The method should directly support binary, nominal, categorical and continuous attributes. Additionally, medical datasets are usually of high dimensionality and typically contain a few thousand records. As the search space becomes larger, the computational feasibility should not get lost. This suggests an evolutionary kind of approach. Support and confidence framework by itself is not enough to prune out uninteresting rules as medical data needs to find rules with low values of support. The requirement is to reduce the number of false positives. This will require some modification of the fitness function defined in terms of support and confidence.

Also an explanatory model is required. The system should not be an opaque box which takes input and gives output. The user may not trust such a system. It should be able to explain predictions and help user in interactive exploration that is flexible. Last but not the least, no single technique consistently outperformed the other in terms of accuracy, though swarm based approaches gave good results. Results were highly dependent on the datasets. A hybrid approach might help in this regard.

3. PROPOSED METHOD: HOSE

Hybrid algorithms which combine concepts from ACO and PSO can deal with all types of attributes. In a pilot study conducted in (Mangat, 2011), these methods have shown reasonably good accuracies in the range of 91-94% while maintaining the comprehensibility of the rules as measured using size of rule and rule sets. To tailor the system to medical domain, a new fitness function can be embedded into the process. The choice to combine evolutionary methods to mine association rules is justified since they are able to conduct a global search of the vast domain space using greedy technique.

The proposed algorithm: HOSE (Hybrid Optimization based on Swarms) is based on ACO/PSO and uses a sequential covering approach to discover classification rules one by one.

Pseudocode of HOSE:

```

AlgoOuter
RSglobal = { }
RSlocal = { }
FOR EACH class C
  TDB= {Training dataset including all records of all
  classes}
  IF

```

```

(Remaining training records of class C
>MaxUncovExampPerClass)
then
Execute AlgoInner(TDB)/*returns Rule*/
Run modified PSO/*expands Rule to include
continuous terms, returns highest quality rule
called FitRule*/
Clip FitRule
RSlocal = RSlocal  $\cup$  FitRule
TDB = TDB - {training records covered by
FitRule}
END IF
END FOR
Clip RSlocal
Sort rules in RSlocal by descending Quality
If Quality(RSlocal)>Quality(RSglobal)
then
RSglobal=RSlocal
Return RSglobal
Nominal attributes are handled by the AlgoInner:
Initialise individuals in population
k=1
DO WHILE (k< MaxIterations)
FOR every particle x
Set Rule Rx = "IF {null} THEN C"
FOR every attribute a of x
Apply roulette selection and set state to 0 or 1
If (state=1) then
add respective attribute-value pair to Rx;
LOOP
Calculate Quality Qx of Rx
P = x's past best state
Qp = P's quality
IF Qx>Qp
Qp = Qx
P = x
END IF
LOOP
FOR every particle x
P = x's past best state
N = the best state of any neighbor of x
FOR every attribute a of x
IF Pa = Na THEN increase pheromone value of
of Na in the current xd by Qp
ELSE IF Pa = off AND seeding term for xa  $\neq$  Na
THEN
Increase pheromone entry for state=0 in xa by Qp
ELSE
Increase pheromone entry in current xa of Na by
Qp

```

```

END IF
Normalize pheromone entries
LOOP
LOOP
LOOP
k=k+1
RETURN fittest rule discovered

```

The rule discovery, fitness evaluation and clipping of individual terms is the same as for ACO/PSO with PF (Mangat, 2011).

We suggest a modified PSO to handle the continuous attributes. Diversity is introduced by selecting the exemplar particles from a prespecified region or neighbourhood rather than randomly. All other particles' past best information is used to update a particle's velocity. The velocity updating is done according to a vector from the particle's region only. But the regions are reconstructed at fixed prespecified points of time in the execution of the algorithm. This mechanism ensures a good balance between the exploration and exploitation properties of the algorithm and avoids premature or delayed convergence.

The particle learns simultaneously from its own best known position as well as the global best. We update only one term or dimension and not all the dimensions of all the velocities of the particles simultaneously. We construct a vector for each particle which indicates which other particle's personal best should this particle learn from. $F_i = |f_{i1}, f_{i2}, \dots, f_{id}|$.

If the fitness of a particle does not increase for a fixed number of iterations (parameter iter) then a random number between 0 and 1 is chosen. If this number is greater than P_c , then $fid = i$. If it is less than P_c , then the particle i learns from some other particle's personal best in the same region as given by vector F_i . P_c is parameter that controls how frequently learning occurs.

Any connected topology can be used to define regions. The length of region controls number of particles in it. The vector to be optimized consists of two terms or dimensions for every continuous attribute that specifies the range for this attribute. Everytime the fitness evaluation of particle is done, the vector is transformed to a set of terms that are added to Rule produced by the algorithm. If upper and lower limits are crossing, then both terms are removed from rule. Updation of Personal Best position is done for those dimensions using Equation (1):

$$vid = \chi \left(\begin{matrix} vid + c1\phi1(P_{id}^{fid} - xid) \\ +c2\phi2(Pgd - xid) \end{matrix} \right) \quad xid = xid + vid \quad (1)$$

where, vid is the dimension velocity, xid is the particle position, P_{id}^{fid} denotes the corresponding dimension d of the ith particle's own pbest or the exemplar's pbest, Pgd is the best position in the neighborhood, χ is constriction coefficient, $\phi1$ and $\phi2$ are random weights, c1 and c2 are constants.

A particle operates within its own region. A random particle is picked as seed initially. It has to be from same region. Other particles set their initial values to a uniformly distributed position between the value of this former seed's continuous attribute and add it to the range for that attribute (for upper bound) and at a uniformly distributed position between seed's value and deduct it from range for that attribute (for lower bound).

Quality, Q of a rule is computed using Laplace-corrected Precision (LP) Equation (2):

$$LP = (1 + TP) / (1 + TP + FP) \text{ If } TP < \text{MinTP}, \quad (2)$$

$$Q = LP * 0.1, \text{ ELSE } Q = LP$$

where, MinTP is the lower threshold for number of correctly covered examples by a rule (Mangat, 2011).

The second modification is to compute quality of not just the individual rules in rule set, but also compute and evaluate the complete rule set quality. This is done using function Equation (3):

$$\text{Quality}(RS) = 1 - (\sum_{r=1}^N (TPn + FPn)) \quad (3)$$

$$U_{CF}(FPn, TPn + FPn) / S$$

where, TPn, FPn are the number of true positives and false positives respectively of nth rule, S is number of records in training set and U_{CF} is the error rate (Fernando *et al.*, 2012).

4. EXPERIMENTAL SETUP

4.1. Database

We have used four publicly available datasets to check the performance of our algorithm: Dermatology, Parkinsons, Pima and Indian Liver Patient Dataset. These datasets contain a good mix of binary, nominal and continuous attributes. Dermatology dataset

contains 33 nominal and 1 continuous attribute and 6 classes. Parkinsons contains 22 continuous attributes and 2 classes. Pima diabetes dataset has 8 continuous attributes and 2 classes and Indian Liver Patient dataset contains 10 real attributes and 2 classes. The task is to determine whether a patient has the respective ailment or disease or not and to use rule based methodology for this classification.

4.2. Parameter Settings

Similar parameter settings have been used as in hybrid ACO/PSO with PF (Mangat, 2011). For the ACO component, parameter values were set as given: Number of Ants = 500, Minimum number of records a rule covers = 10, maximum uncovered examples of each class = 15 and number of rules to test if ant has converged = 20. For PSO component, number of particles = 20 and number of iterations = 50. The minimum number of true positives to apply quality modification = 15. Typical values were set for constriction factor $\chi = 0.729$, social and personal learning coefficients, $c1 = c2 = 2.05$. Maximum number of uncovered examples per class was set to 15. We assume a ring topology where the length of region is taken as 5. Regrouping is done after every 7 iterations. The learning probability P_c is varied from 0 to 0.5. *Iter* parameter is set to 15.

5. RESULTS AND ANALYSIS

A comparison of our proposed HOSE algorithm has been done with well known algorithms: c4.5, PART, ACO/PSO with PF and cAnt-Miner. We used three criteria to compare and analyze performance of rule mining algorithms under consideration. The first criterion is predictive accuracy which is defined in terms of cross validation accuracy rate. It is computed as quotient between number of test examples correctly classified and the total number of test records. A k-fold cross validation was used with value of $k = 10$. Since rules need to be comprehensible and not just correct and smaller rules are intuitively simpler to understand, the other two criteria relate to size of output. The number of rules in a rule set is second criterion and the number of attribute value combinations or conditions per rule is the final criterion.

Table 1-4 summarize the results obtained by the c4.5, PART, ACO/PSO with PF, cAnt-Miner and HOSE over dermatology, parkinsons, pima and Indian liver patient datasets, respectively.

Table 1. Comparison between C4.5, PART, ACO/PSO with PF, cANTMINER, hose over dermatology dataset (Average values)

Algorithm	Accuracy	No. of terms in a rule	No. of rules in ruleset
C4.5	93.45+-1.22	45.00+-1.89	19.60+-0.34
PART	94.26 +-1.17	25.00+-1.12	9.60+-0.34
ACO/PSO with PF	92.84+-0.12	30.70+-0.30	10.29+-0.08
cAntMiner	92.32+-0.31	20.40+-0.37	9.24+-0.10
HOSE	93.62+-0.22	19.00+-0.20	10.12+-0.26

Table 2. Comparison between c4.5, PART, ACO/PSO with PF, cANTMINER, hose over parkinsons dataset (average values)

Algorithm	Accuracy	No. of terms in a rule	No. of rules in ruleset
C4.5	83.47+-2.22	18.90+-1.29	8.80+-0.42
PART	86.05+-2.47	11.10+-0.50	6.40+-0.34
ACO/PSO with PF	86.87+-0.60	10.33+-0.45	4.79+-0.03
cAntMiner	86.98+-0.65	9.19+-0.07	7.11+-0.07
HOSE	87.24+-0.23	8.89+-0.30	5.34+-0.12

Table 3. Comparison between c4.5, part, ACO/PSO with PF, cANTMINER, hose over Pima dataset (average values)

Algorithm	Accuracy	No. of terms in a rule	No. of rules in ruleset
C4.5	74.32+-1.73	36.60+-4.09	12.50+-0.92
PART	71.73+-1.71	14.60+-0.60	7.50+-0.34
ACO/PSO with PF	73.10+-0.31	26.93+-2.46	33.55+-0.13
cAntMiner	74.81+-0.18	13.17+-0.15	14.89+-0.09
HOSE	74.68+-0.88	10.52+-0.08	8.43+-0.06

Table 4. Comparison between c4.5, PART, ACO/PSO with PF, cAntMiner, HOSE over Indian liver dataset (average values)

Algorithm	Accuracy	No. of terms in a rule	No. of rules in ruleset
C4.5	94.13+-2.2	4.86+-0.22	16.20+-0.12
PART	93.84+-1.65	4.72+-0.18	15.83+-0.08
ACO/PSO with PF	96.21+-1.25	4.69+-0.21	18.52+-0.46
cAntMiner	95.13+-0.50	3.24+-0.06	8.32+-0.10
HOSE	96.42+-1.20	2.12+-0.04	6.86+-0.06

6. CONCLUSION

This study discusses a proposed algorithm that can mine quality association rules to be used for classification of patients. The algorithm uses a hybridization of ant colony optimization and particle swarm optimization techniques to mine categorical, binary and continuous data. These traditional techniques have been modified to achieve better results. Firstly, proposed algorithm implements improvements in the traditional PSO, which allow learning of particle in a region of the search space i.e., exploitation and then reformation of regions to allow for exploration. The initial particle is not chosen randomly. The search can be better guided this way. Secondly, the quality of the entire set of rules as well as each individual rule in rule set is evaluated. This should contribute to the

accuracy of the entire rule list and be able to handle effect of rule interactions within the same list. It will eliminate rules that are not increasing the quality of the entire rule set. Proposed algorithm: HOSE has been compared with other well known algorithms over four sample datasets. It performs best in terms of understandability of output rule set and shows comparable results for accuracy parameter as well. One possible further research direction is to explore the effect of region formation and learning probability parameters on the performance of the algorithm. Also different topologies like star, mesh for the regions may be required for good performance over other datasets. In order to check the stability of the proposed method, it may be used over a dataset containing noisy values. Also, the effect of using different functions to evaluate rule list quality, can be explored.

7. REFERENCES

- Baykasoglu, A., L. Ozbakir and S. Kulluk, 2011. Classifying defect factors in fabric production via DIFACONN-miner: A case study. *J. Expert Syst. Applic.*, 38: 11321-11328. DOI: 10.1016/j.eswa.2011.02.182
- Catral, R., F. Oppacher and K.J.L. Graham, 2009. Techniques for evolutionary rule discovery in data mining. *Proceedings of the IEEE Congress on Evolutionary Computation*, May 18-21, IEEE Xplore Press, Trondheim, pp: 1737-1744. DOI: 10.1109/CEC.2009.4983151
- Dogan, S. and I. Turkoglu, 2008. Diagnosing hyperlipidemia using association rules. *Mathem. Comput. Applic.*, 13: 193-202.
- Fernando, E.B., A.A. Freitas and C.G. Johnson, 2008. cAnt-miner: An ant colony classification algorithm to cope with continuous attributes. *Ant Colony Optim. Swarm Intell.*, 5217: 48-59. DOI: 10.1007/978-3-540-87527-7_5
- Fernando, E.B., O. Alex A. Freitas and C.G. Johnson, 2012. A new sequential covering strategy for inducing classification rules with ant colony algorithms. *IEEE Trans. Evol. Comput.*, 17: 64-76. DOI: 10.1109/TEVC.2012.2185846
- Ghannad-Rezaie, M., H. Soltanain-Zadeh, M.R. Siadat and K.V. Elisevich, 2006. Medical data mining using particle swarm optimization for temporal lobe epilepsy. *Proceedings of the IEEE Congress on Evolutionary Computation*, Jul. 16-21, IEEE Xplore Press, Vancouver, BC., pp: 761-768. DOI: 10.1109/CEC.2006.1688388
- Gonzales, E., K. Taboada, S. Mabu, K. Shimada and K. Hirasawa, 2009. Combination of two evolutionary methods for mining association rules in large and dense databases. *J. Adv. Comput. Intell. Intelligent Inform.*, 13: 561-572.
- Holden, N. and A. Freitas, 2008. A Hybrid PSO/ACO algorithm for discovering classification rules in data mining. *J. Artif. Evol. Applic.*, 2008: 316145-316155. DOI: 10.1155/2008/316145
- Kuo, R.J., C.M. Chao and Y.T. Chiu, 2011. Application of particle swarm optimization to association rule mining. *Applied Soft Comput.*, 11: 326-336. DOI: 10.1016/j.asoc.2009.11.023
- Kwasnicka, H. and K. Switalski, 2005. Discovery of association rules from medical data-classical and evolutionary approaches. *Proceedings of the Meeting of Polish Information Processing Society Conference, (PSC' 05)*, pp: 163-177.
- Mangat, V., 2011. Natural intelligence based automatic knowledge discovery for medical practitioners. *Intell. Automat. Syst. Eng.*, 103: 53-64. DOI: 10.1007/978-1-4614-0373-9_5
- Martens, L., M. Chambers, M. Sturm, D. Kessner and F. Levander *et al.*, 2011. mzML--a community standard for mass spectrometry data. *Mol. Cell Proteomics*. DOI: 10.1074/mcp.R110.000133, PMID: 20716697
- Nouaouria, N., 2013. Particle swarm classification: A survey and positioning. *Patt. Recogn.*, 46: 2028-2044. DOI: 10.1016/j.patcog.2012.12.011
- Ordóñez, C., 2006. Comparing association rules and decision trees for disease prediction. *Proceedings of the International Workshop on Healthcare Information and Knowledge Management*, Nov. 05-11, ACM Press, New York, USA., pp: 17-24. DOI: 10.1145/1183568.1183573
- Shukran, M.A.M., Y.Y. Chung, W.C. Yeh, N. Wahid and A.M.A. Zaidi, 2011. Artificial bee colony based data mining algorithms for classification tasks. *Modern Applied Sci.*, 5: 217-231. DOI: 10.5539/mas.v5n4p217
- Silva, S.T. and A. Neves, 2004. Particle swarm based data mining algorithms for classification tasks. *J. Parallel Comput.*, 30: 767-783. DOI: 10.1016/j.parco.2003.12.015
- Sordo, M., N. Shawn and A. Murphy, 2009. PSO/ACO approach to knowledge discovery in a pharmacovigilance context. *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, Jul. 08-12, ACM Press, New York, USA., pp: 2679-2684. DOI: 10.1145/1570256.1570382
- Vreeken, J., M. Van Leeuwen and A. Siebes, 2011. KRIMP: Mining itemsets that compress. *Data Min. Knowl. Discov. J.*, 23: 169-214. DOI: 10.1007/s10618-010-0202-x
- Wahid, N., 2011. A novel Image Classification Algorithm Using Swarm-Based Technique for Image Database. In: *Ubiquitous Computing and Multimedia Applications*, Kim, T.H., H. Adeli, R.J. Robles and M. Balitanas (Eds.), Springer, Berlin Heidelberg, ISBN-10: 3642209971, pp: 460-470.
- Yan, X., C. Zhang and S. Zhang, 2009. Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support. *J. Expert Syst. Applic.*, 36: 3066-3076. DOI: 10.1016/j.eswa.2008.01.028