A Novel Hybrid Framework using Evolutionary Computing and Swarm Intelligence for Rule Mining in the medical domain

Veenu Mangat
Assisstant Professor
U.I.E.T, Panjab University
Chandigarh

ABSTRACT
Modern medicine generates a huge quantity of information daily which is stored in the medical databases. Extracting useful knowledge and providing scientific decision-making for the diagnosis and treatment of disease from the database has become a necessity. The preferred data mining functionality is association rule mining as rules are simple to understand and infer. For a rule based system to be usable in the medical domain, it must exhibit high predictive accuracy and be comprehensible. This paper surveys the various techniques for rule mining in the medical domain, identifies gaps and proposes a novel hybrid framework for efficient rule mining. A pilot study conducted over medical data paved the way for the framework. The output of the system can be used to discover new associations, validate previous findings or for the task of classification. Section I discusses the association rule mining problem. Section II discusses traditional approaches to rule mining. Section III lists the gaps in research. Section IV describes our proposed framework which includes a novel interestingness measure embedded in mining process to make it tailored to medical domain. Section V concludes the paper.

General Terms

Data mining, soft computing

Keywords

Association rule mining, evolutionary algorithms, genetic algorithms, Particle Swarm Optimization, Ant Colony Optimization

1. INTRODUCTION

Data Mining consists of the (semi-)automatic extraction of knowledge from data. The discovered knowledge should be accurate and comprehensible. Knowledge comprehensibility can be achieved by using high-level knowledge representations. A natural choice is association rules. Association rule mining aims to extract interesting correlations, frequent patterns, associations or causal structures among sets of items in the transaction databases or other data repositories. An association rule is an implication in the form of X->Y, where X, Y ϵ I are sets of items called itemsets, and $X \cap Y = \phi$; I=I1, I2, … ,Im is a set of m distinct attributes, T is a transaction/record that contains a set of items such that T ϵ I, D is a database with different transaction records Ts. X is called antecedent while Y is called consequent, the rule means X implies Y.There are two important basic measures for association rules, support and confidence. Support(s) of an association rule is defined as the percentage/fraction of records that contain $X \cup Y$ to the total number of records in the database. Confidence of an association rule is defined as the percentage/fraction of the number of transactions that contain $X \cup Y$ to the total number of records that contain X. Association rule mining problem is to find out association rules that satisfy the predefined minimum support and confidence from a given database. The problem is usually decomposed into two sub problems. One is to find those itemsets whose occurrences exceed a predefined threshold in the database; those itemsets are called frequent or large itemsets. The second problem is to generate association rules from those large itemsets with the constraints of minimal confidence. Since the second sub problem is quite simple, most of the research focuses on the first sub problem. The first sub-problem can be further divided into two sub-problems: candidate large itemsets generation process and frequent itemsets generation process. Those itemsets that are expected or have the hope to be large or frequent are called candidate itemsets. In many cases, the algorithms generate an extremely large number of association rules, often in thousands or even millions. Further, the association rules are sometimes very large. It is nearly impossible for the end users to comprehend or validate such large number of complex association rules, thereby limiting the usefulness of the data mining results.

Soft computing is a paradigm that deals with imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost. An evolutionary algorithm (EA) is a subset of evolutionary computation; a generic population-based metaheuristic optimization algorithm. An EA uses some mechanisms inspired by biological evolution: reproduction, mutation, recombination, and selection. Swarm Intelligence is an innovative distributed intelligent paradigm for solving optimization problems that originally took its inspiration from biological examples. Data mining and swarm optimization may seem that they do not have many properties in common. However, they can be used together to form a method which often leads to a target solution to the problem, even when other methods would be too expensive or difficult to implement [1-8].

2. RULE MINING METHODS

The AIS algorithm was the first algorithm proposed for mining association rules [9]. In this algorithm, only one item consequent association rules and too many candidate itemsets that finally turn out to be small are generated. This requires more space and wastes much effort. At the same time this algorithm requires too many passes over the whole database. Apriori is more efficient during the candidate generation process [10]. Apriori uses pruning techniques to avoid measuring certain itemsets, while guaranteeing completeness. But it has two drawbacks. One is the complex candidate generation process that uses most of the time, space and memory.Another bottleneck is the multiple scanning of the database.

FP-Tree [11], frequent pattern mining, generates frequent itemsets with only two passes over the database and without any candidate generation process. FP-tree is an extended prefix-tree structure storing crucial, quantitative information about frequent patterns. Only frequent length-1 items will have nodes in the tree, and the tree nodes are arranged in such a way that more frequently occurring nodes will have better chances of sharing nodes than less frequently occurring ones. FP-Tree scales much better than Apriori.

The problem with FP-Tree is that it is difficult to use in an interactive mining system because when users change the threshold of support, it may lead to repetition of the whole mining process. Another limitation is that FP-Tree is not suitable for incremental mining.

TreeProjection is another efficient algorithm proposed in [12]. The general idea of TreeProjection is that it constructs a lexicographical tree and projects a large database into a set of reduced, item-based sub-databases based on the frequent patterns mined so far. An efficient algorithm for mining association rules, PRICES [13] was then presented. The approach reduces large itemset generation time, known to be the most time-consuming step, by scanning the database only once and using logical operations in the process.

Another algorithm for generating large frequent candidate sets is proposed in [14], which is called Matrix Algorithm. The algorithm generates a matrix with entries 1 or 0 by passing over the database only once, and then the frequent candidate sets are obtained from the resulting matrix. Finally association rules are mined from the frequent candidate sets. Experimental results confirm that the proposed algorithm is more effective than Apriori Algorithm.

Sampling can be used in association rule mining algorithm effectively [15-18]. The approach can be divided into two phases. During phase 1, a sample of the database is obtained and all associations in the sample are found. These results are then validated against the entire database.

A parallelization of Apriori to shared nothing machines, each with its own partition of the database called FDM algorithm was discussed in [19]. A distributed version of this called DDM was presented in [20]. Another efficient parallel algorithm FPM (Fast Parallel Mining) for mining association rules on a shared-nothing parallel system has been proposed by [21]. It adopts the count distribution approach and has incorporated two powerful candidate pruning techniques, i.e., distributed pruning and global pruning. It has a simple communication scheme which performs only one round of message exchange in each iteration. A new algorithm called Data Allocation Algorithm (DAA) was proposed in [22] that uses Principal Component Analysis to improve the data distribution prior to FPM.

The focus of [23] was improving the efficiency of constraint-based frequent pattern mining by using dataset filtering techniques. A specific type of constraints called category-based as well as the associated algorithm for constrained rule mining based on Apriori was discussed in [24]. A novel algorithm, called IDTE, which can incrementally update the discovered generalized association rules when the taxonomy of items is evolved with new transactions insertion to the database was introduced [25].

Temporal association rules can be more useful and informative than basic association rules. For example the stage of the disease for a patient varies with respect to time. Monitoring patient's infections by knowing the temporal order in which some symptoms occurred and how long they lasted, is required. In [26] an algorithm for mining periodical patterns and sequential patterns was introduced.

The application of evolutionary approaches to rule mining started in [27] wherein the authors proposed an algorithm called AntMiner for extracting rules from a medical database as a decision aid. They showed that the proposed algorithm achieved good predictive accuracy and a reduced number of rules at the same time. The main drawback that was cited was the computational cost.

Mapping of medical data to transaction format was discussed in [28]. The combinatorial nature of association rules was discussed and methods to preprocess medical data were explored. The suitability of association rules for decision making support was further reinforced in [29] wherein rule mining was used to identify diseases that occurred together commonly, or less commonly than their individual frequencies in the population would predict.

The use of association rule mining with constraints and neural network technique based on back propagation to detect tumor was explored in [30]. Authors reported that neural network technique was less sensitive to database imbalance but had high training costs. ARM obtained better results on balanced database.

The role of evolutionary algorithms for rule mining and knowledge discovery was elaborated in [31]. The use of genetic algorithms for rule mining was done and it was found that GAs cope well with attribute interaction. GA has been used as a wrapper to select attributes for a constructive neural network. Genetic Programming also has the potential to create derived attributes with greater predictive power.

A framework for pushing support constraints into the Apriori algorithm was presented in [32]. This Adaptive Apriori method does not require a minimum support for all itemsets. [33] was a seminal paper that proposed the use of Particle Swarm Optimization as a new tool for data mining. Three different Particle Swarm Data Mining Algorithms were implemented and tested against a Genetic Algorithm and a Tree Induction Algorithm (J48).From the obtained results, Particle Swarm Optimizers proved to be a suitable candidate for rule discovery tasks. The results obtained in these domains seem to indicate that Particle Swarm Data Mining Algorithms are competitive, not only with other evolutionary techniques, but also with industry standard algorithms such as the J48 and c4.5 algorithm, and can be successfully applied to more demanding problem domains.

An evolutionary method to mine time series database was discussed in [34]. The algorithm uses sub tree swapping crossover, tree generating mutation and reproduction as genetic operators. Individuals are chosen for participation in new generations using tournament selection. It was shown empirically that this method is capable of producing rules with good predictive power and with a high level of interestingness. It can discover real-world relationships between time series.

Another method for rule mining based on an evolutionary (GA) approach –EGAR was described in [35] 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 [36]. The method addressed issues such as missing value management and interactive rule extraction. 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. It was found that C4.5 gave higher accuracy but at the cost of a more complex rule set and lower generalization. PSO shows a good convergence speed and is faster than ACO while having same performance. Convergence of C4.5 is very fast, but not recommended for small databases. 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.

In [37], search constraints are introduced to find only medically significant association rules and make search more efficient. Association rules are compared to predictive rules mined with decision trees. Results provide evidence that decision trees are less effective than constrained association rules to predict disease with several related target attributes, due to low confidence factors (i.e. low reliability), slight over fitting, rule complexity for unrestricted trees (i.e. long rules) and data set fragmentation (i.e. small data subsets).

The shuffled frog leaping(SFL) algorithm for classification rule mining, as discussed in [38], is a new robust evolutionary algorithm based on the local search and the shuffling processes. Authors contend that other approaches based on statistical, rough set, neural networks, etc. though algorithmically strong, do not provide intelligible rules. In classification rule mining, the SFL algorithm has the ability to perform a flexible robust search for a good combination of terms(logical conditions) involving values of the predictor attributes. SFL has been shown to give better performance than Antminer and GA.

A hybrid particle swarm optimization/ant colony optimization (PSO/ACO) algorithm for the discovery of classification rules was proposed in [39]. Unlike a conventional PSO algorithm, this hybrid algorithm can directly cope with nominal attributes, without converting nominal values into binary numbers in a preprocessing phase. Authors compared the algorithm to an industry standard algorithm PART and its reduced version, coping only with continuous data, to the new classification algorithm for continuous data based on differential evolution. The results show that proposed algorithm is very competitive in terms of accuracy to PART and that it produces significantly simpler (smaller) rule sets, a desirable result in data mining—where the goal is to discover knowledge that is not only accurate but also comprehensible to the user. The results also show that the reduced PSO version for continuous attributes provides a slight increase in accuracy when compared to the differential evolution variant.

Association rule mining has been applied for discovering hyperlipidemia form biochemistry blood parameters in [40].PSO/ACO approach to knowledge discovery was successfully applied in a pharmacovigilance context in [41]. The approach showed high accuracy to detect presence of previously undetected causal relationships between therapeutics, patient characteristics and adverse events.

A real-life time series data of muscular activities of human participants obtained from multiple Electromyogram(EMG) sensors was considered and patterns were discovered in these EMG datastreams in [42]. Evaluation with large sets of time series data from multiple EMG sensors demonstrated that a two-stage approach speeds up the process of finding association rules in such multidimensional environment as compared to other methods and scales up linearly in terms of number of time series involved.

Several additions to the standard GA intended to enhance the ability for an evolutionary system to perform rule discovery in data mining have been elaborated in [43]. While some of the methods, such as Automatic Consequent Balancing, are designed specifically for the task of rule discovery, other additions such as the Lifetime Duplicate Check can be used to extend most implementations of GA. Several of these techniques are applied between generations in intergenerational processing stage. Experiments have shown that the system performs better, on average, than many algorithms like C4.5, logistic regression, Naïve Bayesian, etc. described in other published work.

Another evolutionary method for quantitative association rule mining is mainly motivated by (1) partition of quantitative attribute is not easy for every attribute and every user, (2) users, and even experts, usually feel difficult to specify the minimum-support, (3) the search space might be very large when we face quantitative attributes, and (4) the rules returned might be too many to deal with. The developed EARMGA algorithm can mine high quality rules without the user specifying minimum support or confidence threshold levels [44].

Large and dense databases with a huge amount of attributes can be mined through the combination of conventional GNP based mining method and a specially designed genetic algorithm (GA). The strategy consists of the division of a large and dense database into many small databases. These small databases are considered as individuals and form a population. Then the conventional GNP based mining method is applied to extract association rules for each of these individuals. Finally, the population is evolved through several generations using GA with special genetic operators. The results show that this method allows extracting association rules from large and dense databases directly and more efficiently than the conventional GNP method [45].

One of the major problems in pattern mining is the explosion of the number of results. The problem can be solved by using the Minimum Description Length principle (MDL): the best set of patterns is that set that compresses the database best. Using the Krimp algorithm for frequent itemset mining and classification, a dramatic reduction in the number of returned frequent item sets is obtained. These selections, called code tables, are of high quality. This is shown with compression ratios, swap-randomization, and the accuracies of the code table-based Krimp classifier, all obtained on a wide range of datasets. The algorithm shows high stability w.r.t. different candidate sets and it is parameter-free. There are many data mining tasks for which it can be used eg: frequent itemset mining while preserving privacy [46].

A data mining tool (DIFACONN-miner) was used for generating accurate classification rules for classifying causes of defects

[47]. DIFACONNminer uses differential evolution (DE) algorithm for training ANNs and touring ant colony optimization (TACO) algorithm for generating classification rules. Fitness of ANN structure is evaluated according to a multiple objective function which consists of three performance measures namely error of ANN, number of rules and training accuracy. It was proven that DIFACONN-miner is able to produce accurate and effective classification rules.

A new evolutionary image classification algorithm namely Simplified Swarm Optimization (SSO) has been proposed recently. Image classification has faced a problem where the number of possible different combination of variables is very high. The performance of SSO, Particle Swarm Optimization (PSO) and Support Vector Machine (SVM) has been compared and analyzed. Unlike PSO, SSO does not need to use the velocity parameter and inertia weight; it uses only one random number and three predefined parameters to update each of the particle's position. SSO provides better classification accuracy than PSO and SVM. It has a potential to achieve higher classification result with smaller number of particles. Furthermore, PSO needs to allocate more memory than SSO for each particle to achieve better performance [48].

3. GAPS IN RESEARCH

a) Dirty Data:
Data in the real world is inherently dirty containing duplicate tuples, missing values, etc. The system needs to be designed to automatically deal with dirty data and avoid returning rules involving missing information.

b) Support for different data types:
No single efficient scheme that works on all textual data types has been reported in literature. The algorithm should be such that it directly supports binary, nominal, categorical and continuous attributes.

c) High dimensionality small datasets:
The 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.

d) Interestingness framework:
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.

e) Trivial patterns:
Certain combinations of attributes may be trivial and should be grouped or eliminated based on previous knowledge of the domain.

f) Explanatory model:
The discovered knowledge should not be a black box which makes predictions without explaining them. The user may not trust such a system.

g) Predictive accuracy:
Last but not the least, no single approach reported in literature 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.

4. PROPOSED FRAMEWORK

To overcome the problems cited in previous section, it is proposed that a hybrid framework be evolved for performing rule mining.

After the initial data collection, cleaning of data to handle missing values can be done by substituting the null with most frequently encountered value if number of missing values is less than user defined threshold, or by ignoring the tuple if it is more than threshold. Attribute discretization for attributes with large ranges and attribute construction for related attributes can be done by consulting with medical practitioners.

The evolutionary approach using Genetic algorithms can then be used for feature selection. This can be provided as an input to a hybrid algorithm which is a combination of ACO, PSO and/or SFLA, and can deal with all types of attributes. In a pilot study conducted in [49], 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. These results have been obtained without optimizing any algorithm parameters and may be further improved. To tailor the system to medical domain, a new fitness function can be embedded into the process. The GA component can use the accuracy parameter as defined using specificity and sensitivity. Post filtering of rules using measures of lift and conviction may be done.

The major advantage of using GAs in the discovery of prediction rules is that they perform a global search based on the greedy approach. The function of genetic algorithm is as follows:

1) Selection: The selection of the member from the population can be done with the help of Roulette Wheel sampling method. Roulette Wheel selection is a process of choosing members from the population of chromosomes in a way that is proportional to their fitness.

2) Crossover is performed by selecting a random gene along the length of the chromosomes and swapping all the genes after that point.

3) Mutation alters the new solutions so as to add stochasticity in the search for better solutions. This is the chance that a bit within a chromosome will be flipped (0 becomes 1, 1 becomes 0). Whenever chromosomes are chosen from the population the algorithm first checks to see if crossover should be applied and then the algorithm iterates down the length of each chromosome mutating the bits if applicable.

The population is ranked with the help of fitness function. We apply genetic algorithm on the selected population from the database and compute the fitness function after each step until termination. It is known that higher the values of true positives(TP), and lower the values of false positives(FP) and false negatives (FN), the better is the rule.

Confidence Factor, CF = TP / (TP + FN)

Completeness, Comp=TP/ (TP+FP)

Fitness=CF *Comp (1)

The SI based component can use a novel fitness function defined in Eq. (3) that penalizes false positives severely. The ACO/PSO and/or SFLA algorithm with Precision Fitness uses a sequential covering approach to discover one classification rule at a time.

AlgoCover is :

RS = {}

FOR EACH class C

TS = {All training examples belonging to any class}

WHILE (Number of uncovered training examples belonging to class C >MaxUncovExampPerClass)

Run the AlgoCore to discover best nominal rule predicting class C called Rule

Run the standard PSO algorithm to add continuous terms to Rule, and return the best discovered rule BestRule
Prune BestRule
RS = RS ∪ BestRule
TS = TS −{training examples covered by discovered rule}
ENDWHILE
END FOR
Order rules in RS by descending Quality
Prune RS removing unnecessary terms and/or rules

A single iteration of this loop only discovers rules based on nominal attributes, returning the best discovered rule. For the continuous part of the rule, a conventional PSO algorithm with constriction is used. The vector to be optimized consists of two dimensions per continuous attribute, one for an upper bound and one for a lower bound. At every particle evaluation, the vector is converted to a set of terms and added to Rule produced by the algorithm for fitness evaluation. If two bounds cross over, both terms are omitted from decoded rule, but Personal Best position is still updated in those dimensions using :

$$vid = \chi\,(vid + c1\varphi1(Pid-xid) + c2\varphi2(Pgd - xid)))$$
$$xid=xid+vid \qquad (2)$$

where vid is the dimension velocity, xid is the particle position, Pid is the particle's previous best position, Pgd is the best position in the neighborhood, χ is constriction coefficient, $\varphi1$ and $\varphi2$ are random weights, c1 and c2 are constants.

The particles are prevented from fully converging using the Min-Max system. After the BestRule has been generated it is then added to the rule set after being pruned using ACO's pruning method. Nominal attributes are handled by the AlgoCore:

Initialise population
REPEAT for MaxInterations
FOR every particle x
Set Rule Rx = "IF {null} THEN C"
FOR every dimension d in x
Use roulette selection to choose whether state should be set to off or on. If it is on then the corresponding attribute-value pair set in the initialization will be added to Rx; otherwise nothing will be added.
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 ever held by a neighbor of x according to N's quality Q_N
FOR every dimension d in x
IF Pd = Nd THEN pheromone entry corresponding to the value of Nd in the current xd is increased by Qp
ELSE IF Pd = off AND seeding term for xd ≠ Nd THEN
pheromone entry for the off state in xd is increased by Qp
ELSE
pheromone entry corresponding to the value of Nd in the current xd is increased by Qp
END IF
Normalize pheromone entries
LOOP
LOOP
LOOP
RETURN best rule discovered

Each particle has four neighbours. Initially, pheromone state in each dimension is set to 0.9 for on and 0.1 for off. Quality, Q is defined using Precision as given by:

Laplace-corrected Precision = (1+TP)/(1+TP+FP)
If TP<MinTP, Q=Laplace-Corrected Precision*0.1,
ELSE Q=Laplace-Corrected Precision (3)

whereMinTP is the least number of correctly covered examples that a rule has to cover [14].

The above algorithm can be replaced with another evolutionary algorithm like Shuffled frog-leaping algorithm (SFL) which is a new memetic meta-heuristic algorithm with efficient mathematical function and global search capability. The pseudocode of SFL algorithm [21][24] is as follows:

Begin
Generate random population of P solutions (frogs)
For each individual i that belongs to P : calculate fitness (i)
Sort the population P in descending order of their fitness
Divide P into m memeplexes
For each memeplex
Determine the best and worst frogs
Improve the worst frog position using (8)
NewpositionXi= CurrentpositionXi + NewVelocityVi (4)
Repeat for a predefined specific number of iterations
End for
Combine the evolved memeplexes
Sort the population P in descending order of their fitness
Check if termination=true
End;

An initial population of P frogs is created randomly. The structure of an individual frog for rule mining problem is composed of a set of attribute values. The velocity of individual i corresponds to the attribute update quantity covering all attribute values, the velocity of each individual is also created at random. In the next step, the frogs are sorted in a descending order according to their fitness. Fitness is defined as given by (3). The entire population is divided into m memeplexes, each containing n frogs. In this process, the first frog goes to the first memeplex, the second frog goes to the second memeplex, frog m goes to the mth memeplex, and frog m+1 goes back to the first memeplex, etc.

Within each memeplex, PSO is applied to improve only the frog with the worst fitness in each cycle. If no improvement becomes possible in this case, then a new solution is randomly generated to replace that frog.

5 . CONCLUSION AND FUTURE WORK

The proposed framework should be implemented to evaluate its effectiveness. A dataset like the Heart Disease data set available at the UCI machine learning repository may be used as input to check effectiveness in the medical domain. Cross validation to check the predictive accuracy and validation by the medical practitioner to check whether the number of terms in the rules and number of rules in rule set is acceptable and understandable, should be done.

6. REFERENCES

[1] Abraham A, Ramos V, 2003. Web Usage Mining Using Artificial Ant Colony Clustering and Genetic Programming, IEEE Congress on Evolutionary Computation (CEC 2003), Australia, IEEE Press, ISBN 0780378040, pp 1384-1391.

[2] Admane L, Benatchba K, Koudil M, Siad L, Maziz S, 2006. AntPart: an algorithm unsupervised classification problem using ants, Journal of Applied Mathematics and Computation, Volume 180, Issue 1, pp 16-28.

[3] Oliveira LS, Britto AS Jr., Sabourin R, 2005. Improving Cascading Classifiers with Particle Swarm Optimization. In Proceedings of International Conference on Document Analysis and Recognition ICDAR 2005, Seoul, South Korea, pp 570-574.

[4] Omran, M., 2005. Particle Swarm optimization methods for pattern Recognition and Image Processing, Ph.D. Thesis, University of Pretoria.

[5] Shelokar PS, Jayaraman VK, 2004. An ant colony classifier system: application to some process engineering problems, Computers & Chemical Engineering Journal, pp 1577-1584.

[6] Shi Y, et al, 2005. Co-evolutionary particle swarm optimization to solving minmax problems. In Proceedings of the IEEE Conference on Evolutionary Computation, Hawai, pp 1682-1687.

[7] Skopos C, et al, 2005. Particle swarm optimization: an efficient method for tracing periodic orbits in three-dimensional galactic potential, Monthly Notices of RAS, vol. 359, pp 251-260.

[8] Weng SS, Liu YH, 2006. Mining time series data for segmentation by using Ant Colony Optimization, European Journal of Operational Research, pp 921-937.

[9] Agrawal, R., Imielinski, T., and Swami, A. N., 1993. Mining association rules between sets of items in large databases, Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp 207-216.

[10] Agrawal, R. and Srikant, R., 1994. Fast algorithms for mining association rules. In Proceedings of 20th International Conference on Very Large Data Bases, pp 487-499.

[11] Han, J. and Pei, J. , 2000. Mining frequent patterns by pattern-growth: methodology and implications, ACM SIGKDD Explorations Newsletter 2, pp 14-20.

[12] Agarwal, R. and Prasad V., 2000. A tree projection algorithm for generation of frequent itemsets, Journal of Parallel and Distributed Computing.

[13] Wang, C., Tjortjis, C., 2004. PRICES: An Efficient Algorithm for Mining Association Rules, Lecture Notes in Computer Science, Volume 3177, pp 352 – 358.

[14] Yuan, Y., 2005. Matrix Algorithm for Mining Association Rules, LNCS, Vol. 3644, pp 370 – 379.

[15] Toivonen, H., 2000. Sampling large databases for association rules, The VLDB Journal, pp 134-145.

[16] Parthasarathy, S., 2002. Efficient Progressive Sampling for Association Rules. In Proceedings of IEEE Conference on Data Mining 2002, pp 354-361.

[17] Chuang, K., Chen, M., Yang, W., 2005. Progressive Sampling for Association Rules Based on Sampling Error Estimation, Lecture Notes in Computer Science, Volume 3518, pp 505 – 515.

[18] Li, Y., Gopalan, R., 2004. Effective Sampling for Mining Association Rules, Lecture Notes in Computer Science, Volume 3339, Jan 2004, Pages 391 – 401.

[19] Cheung, D., et al, 2000. A fast distributed algorithm for mining association rules. In Proceedings of International Conerence on Parallel and Distributed Information Systems, pp. 31 - 44.

[20] Schuster, A., 2001. Communication-efficient distributed mining of association rules. In Proceedings of ACM SIGMOD International Conference on Management of Data, pp. 473-484.

[21] Cheung, D., Xiao, Y, 1998. Effect of data skewness in parallel mining of association rules, Lecture Notes in Computer Science, Volume 1394, Aug 1998, Pages 48 – 60.

[22] Manning, A., Keane, J., Data Allocation Algorithm for Parallel Association Rule Discovery, Lecture Notes in Computer Science, Volume 2035, pp 413-420.

[24] Tien Dung Do, Siu Cheung Hui, Alvis Fong, 2003. Mining Frequent Itemsets with Category-Based Constraints, Lecture Notes in Computer Science, Volume 2843, pp. 76 – 86.

[25] Tseng, M., Lin, W., Jeng, R., 2005. Maintenance of Generalized Association Rules Under Transaction Update and Taxonomy Evolution, Lecture Notes in Computer Science, Volume 3589, pp 336 – 345.

[26] Verma, K., Vyas, R., 2005. Temporal Approach to Association Rule Mining Using T-Tree and P-Tree, Lecture Notes in Computer Science, Volume 3587, pp. 651 – 659.

[27] Parpinelli, R.S., Lopes, H.S., and Freitas, A.A., 2002. Data Mining with an Ant Colony Optimization Algorithm, IEEE Transactions on Evolutionary Computation, special issue on Ant Colony algorithms, pp 321-332.

[28] Carlos Ordonez, Cesar Santana, Levien de Braal, 2000. Discovering interesting association rules in medical data. In Proceedings of the ACM SIGMODD workshop on Research Issues in Data Mining and Knowledge Discovery.

[29] Goldacre, Kurina et al, 2000. Use of large medical databases to study associations between diseases, International Journal Q J Medicine, pp 669-675.

[30] Maria-Luiza Antonie, Osmar R. Za¨ıane, Alexandru Coman, , 2001. Application of Data Mining Techniques for Medical Image Classification. In Proceedings of ACM/ SIGKDD 2001.

[31] Alex A. Frietas, 2003. A survey of evolutionary algorithms for data mining and knowledge discovery, Advances in Evolutionary Computing, Springer-Verlag , pp 819-845.

[32] Ke Wang, Yu He, and Jiawei Han, 2003. Pushing Support Constraints Into Association Rules Mining, IEEE Transactions on Knowledge and Data Engineering, Vol. 15 .

[33] Sousa, T., Silva, A., Neves, 2004. A Particle Swarm based Data Mining Algorithms for classification tasks, Journal of Parallel Computing, pp 767–783.

[34] Hetland, Saetrom, 2005. Evolutionary rule mining in time series database, Machine learning Journal, Springer-Verlag,Vol 58, pp 107-125.

[35] Halina Kwasnicka and Kajetan Switalski, 2005. Discovery of association rules from medical data - classical and evolutionary approaches. In Proceedings of Meeting of Polish Information Processing Society Conference, pp. 163–177.

[36] M. Ghannad-Rezaie,et al, 2006. Medical Data Mining using Particle Swarm Optimization for Temporal Lobe Epilepsy. In Proceedings of IEEE Congress on Evolutionary Computation.

[37] Carlos Ordonez, 2006. Comparing Association Rules and Decision Trees for Disease Prediction. In Proceedings of ACM HIKM'06, USA.

[38] Yin, Cheng, Zhou, 2008. An efficient SFL based classification rule mining algorithm. In Proceedings of the IEEE International symposium on IT in medicine.

[39] Holden, N., Freitas, A., 2008. A Hybrid PSO/ACO Algorithm for Discovering Classification Rules in Data Mining, Journal of Artificial Evolution and Applications (JAEA).

[40] Sengul Dogan and Ibrahim Turkoglu, 2008. Diagnosing hyperlipidemia using association rules, Mathematical and Computational Applications, Vol. 13, No. 3, pp. 193-202.

[41] Margarita Sordo, Shawn N. Murphy, A PSO/ACO Approach to Knowledge Discovery in a Pharmacovigilance Context, Harvard Medical School, Charlestown, MA. USA,

[42] Gaurav N., Pradhan B. Prabhakaran, 2009. Association Rule Mining in multiple multidimensional time series data. In Proceedings of ICME 2009.

[43] Robert Cattral,et al, 2009. Techniques for Evolutionary Rule Discovery in Data Mining. In Proceedings of IEEE Congress on Evolutionary Computation.

[44] Xiaowei Yan, et al, 2009. Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support, Journal on Expert Systems with Applications, Expert Systems with Applications 36, pp 3066–3076.

[45] Eloy Gonzales, et al, 2010. Combination of Two Evolutionary Methods for Mining Association Rules in Large and Dense Databases, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.13 No.5, pp 561-72.

[46] Jilles Vreeken, Matthijs van Leeuwen, Arno Siebes, 2011. KRIMP: mining itemsets that compress, Data Mining and Knowledge Discovery Journal, No 23, pp 169–214.

[47] Adil Baykasoglu, Lale Ozbakir, Sinem Kulluk, 2011. Classifying defect factors in fabric production via DIFACONN-miner: A case study, Elsevier Journal on Expert Systems with Applications 38, 2011, pp. 11321–11328.

[48] Noorhaniza Wahid, 2011. A Novel Image Classification Algorithm Using Swarm-Based Technique for Image Database, Ubiquitous Computing and Multimedia Applications, Springer-Verlag Journal, Part II, Vol 151, pp. 460–470.

[49] Veenu Mangat, 2011. Natural Intelligence based knowledge discovery for medical practitioners, Intelligent Automation and Systems Engineering, Lecture Notes in Electrical Engineering, Vol 103, pp. 53-64.