A NEW METHOD FOR COMPOSING CLASSIFICATION RULES: AR+OPTBP

Humar KAHRAMANLI[a, *] and Novruz ALLAHVERDİ[b]
[a, *] Selçuk Üniversitesi, Teknik Eğitim Fakültesi Elektronik ve Bilgisayar Eğitimi Bölümü , Konya, TÜRKİYE,
E-mail: hkahramanli@selcuk.edu.tr
[b] Selçuk Üniversitesi, Teknik Eğitim Fakültesi Elektronik ve Bilgisayar Eğitimi Bölümü , Konya, TÜRKİYE,
E-mail: noval@selcuk.edu.tr

Abstract

Although Artificial Neural Network (ANN) usually reaches high classification accuracy, the obtained results sometimes may be incomprehensible. This fact is causing a serious problem in data mining applications. The rules that are derived from ANN are needed to be formed to solve this problem and various methods have been improved to extract these rules. In our previous works Artificial Immune Systems (AIS) algorithm has been used to extract rules from trained ANN and has been applied to various databases [11, 41-43]. In this study, association rules have been composed using Apriori algorithm and transactions, which provide these rules, were eliminated. This provides shrinking database. Then ANN has been trained and used Opt-aiNET for composing rule set. It's been observed that this method increased classification accuracy despite decreasing number of rules.

Keywords: Association Rules, Artificial Neural Networks, Artificial Immune Systems, Optimization, Rule Extraction.

1. Introduction

Data mining is a process of inductively analyzing data to find interesting patterns and previously unknown relationships in the data. Typically, these relationships can be translated into rules that are used to predict future events or to provide knowledge about interrelationships among data [1]. Association rule discovery is an important task in data mining, which aims to find the correlations among items in a transactional database [2]. Association is a set of items found in a database, which provides useful and actionable insights into the structure of the data [3]. Classification is another important task in data mining, which aims to predict the classes of future data objects [2]. Classification is the task of building a model (classifier) from a training data set of given class labels in order to classify data of unknown class labels by the determined model. While in a typical data the class labels must be categorical attributes, the remaining, so-called predictor attributes can contain continuous or numerical features [4]

In the previous years, researchers developed varieties of classification algorithms. The most commonly used classification algorithms are NewID, AC2, CAL5, CN2, C4.5, IndCART, Bayes Tree, ITrule, k-nearest neighbor, radial basis function, Naive Bayes, Polytrees, Kohonen self-organizing net, LVQ, Hard K means and Fuzzy C-Means. One of the most commonly used classifier techniques is artificial neural networks. Neural networks usually provide high classification accuracy. However, the knowledge acquired by such models is generally incomprehensible for humans. This fact is a major obstacle in data mining applications, in which ultimately understandable patterns (like classification rules) are very important [5]. Trained ANNs are often viewed as 'black boxes' which map input data onto an output representation through a number of mathematically weighted connections between layers of neurons [6]. ANN generally suffers from a drawback—they operate as a black box and they do not have an explanatory facility to justify their predictions [7]. ANN can generalize knowledge formed during training and they are relatively insensible to noise included in data, but they do not produce the explanation of their final decision [8]. Thus, knowledge captured by neural networks is not transparent to users and cannot be verified by domain experts [9]. This may cause some problems in practice. To solve this problem, researchers are interested in developing a humanly understandable representation for neural networks. This can be achieved by extracting production rules from trained neural Networks [9]. Rule extraction techniques seek to clarify to the user how the network arrived at its decision [10].
In our previous work, a method for mining classification rules named as OPTBP had been presented [11]. In this study, an AR+OPTBP method was proposed for mining classification rules. This method consists of three-stages. In the first stage, the association rules have been discovered for the classes and data, which provide these rules, were eliminated. This provides the training time to become a little short. The Apriori algorithm is used for mining association rules. In the second stage, neural network has been trained. In the third stage, Opt-aiNET has been executed for extraction rules from this ANN.

2. Literature review

Mining association rules is a popular and well researched method for discovering interesting relations between variables in large databases [12]. Agrawal et al. [13] presented the first algorithm using the support-confidence framework to mine frequent itemsets and association rules. Agrawal and Srikant [14] developed Apriori algorithm, which is a level-wise, breadth-first algorithm, which counts transactions. Liao et al. [15] proposed the Apriori algorithm as a methodology of association rules for data mining, which is implemented for mining marketing map knowledge from customers. Lazcorreta et al. [16] had been introduced a new method towards automatic personalized recommendation based on the behavior of a single user in accordance with all other users in web-based information systems. Karabatak and Ince [17] are used association rules for reducing for reducing the dimension of breast cancer database and ANN for intelligent classification.

In the literature, there are many different approaches for the rule extraction. One of the first rule extraction techniques from neural networks was proposed by Gallant [18]. He was working on connectionist expert systems. In this work, each ANN node represents a conceptual entity. Towell and Shavlik [19] showed how to use ANNs for rule refinement. The algorithm was called SUBSET, which is based on the analysis of the weights that make a specific neuron active. Alexander and Mozer [20] developed a rule extraction method, based on connection weights, that supposes activation functions showing approximately Boolean behavior. Sethi and Yoo [21] developed a rule extraction method based on the connection weights. Keedwell et al. [6, 22] developed a system in which a genetic algorithm is used to search for rules in the ANN input space. Garcetz et al. [23] presented a method to extract non-monotonic rules from ANNs formed by discrete input units. Snyders and Omlin [24] compared the performance of symbolic rules extracted from ANNs trained with and without adaptive bias, giving empirical results for a molecular biology problem. Jiang et al. [25] proposed that combines ANNs and rule learning. The proposed algorithm utilizes a ANN ensemble as the front-end process, which generates abundant training instances for the back-end rule learning process. Setiono et al. [26] presented an approach for extracting rules from ANNs trained in regression problems. Elalfi et al. [27] presented an algorithm for extracting rules from databases via trained ANN using genetic algorithm.

In summary, most of the approaches described in the literature have basically two motivations. On the one hand, some authors noticed the need for simplification of neural networks to facilitate the rule extraction process, and are in favor of using specialized training schemes and architectures to perform such task. The assumption underlying these approaches is that neural networks can help the extraction of interesting rules. On the other hand, some papers have proposed algorithms mainly intended to clarify the knowledge encoded in previously trained ANNs [5].

In our previous work, we presented a method for rule extraction from trained neural networks using artificial immune systems named as OPTBP [11]. This study has been focused on the problem for generating classification rules. The study on rule extraction from trained ANN is based on the work of Elalfi et al. [27] and presents algorithm for extracting rules from neural network using artificial immune systems with discovering preprocessing of association rules.

In this study, an AR + OPTBP method was proposed for mining classification rules. This method consists of three-stages. In the first stage, the association rules have been discovered for classes and datas, which provide this rules, has been eliminated. This provides the training time to become a little shorter. The Apriori algorithm has been used for mining association rules. In the second stage, neural network has been trained. In the third stage, Opt-aiNET has been executed for extraction rules from this ANN.

3. Study Environment and Background Theories

3.1. Association rules and Apriori algorithm

The problem of mining association rules from database of transactions was introduced by Agrawal et al. [13]. Let *A* be a set of items, $X \subset A$, *T* is a database of transactions and *n* is the number of transactions. The support of an itemset *X,* is defined as follows:

$$\sup(X) = \frac{freq(X)}{n} \qquad (1)$$

where *freq(X)* is the number of transactions in which *X* occurs as a subset. A rule is an implication of the form $X \rightarrow Y$, where $X, Y \subset A$ and $X \bigcap Y = \varnothing$. *X* and *Y* are called as antecedent and consequent of the rule respectively. The support of the rule is expressed as follows:

$$\sup(X \rightarrow Y) = \frac{freq(X \bigcup Y)}{n} \qquad (2)$$

The confidence of the rule is defined as follows:

$$conf(X \rightarrow Y) = \frac{\sup\left(X \rightarrow Y\right)}{\sup(X)} \qquad (3)$$

The Apriori algorithm has been used to generate association rules. The Apriori algorithm works iteratively. It first finds the set of large 1-item sets, and then set of 2-itemsets, and so on. The number of scan over the transaction database is as many as the length of the maximal item set [17]. The algorithm works is as follows [28]:

The algorithm finds the frequent sets *L* in database *D*.

- Find frequent set L_{k-1}.
- Join Step.
  - o C_k is generated by joining L_{k-1} with itself
- Prune Step.
  - o Any (k−1) -itemset that is not frequent cannot be a subset of a frequent k -itemset, hence should be removed.

where

- (C_k: Candidate itemset of size k)
- (L_k: frequent itemset of size k)

1)

2)

3) Apriori Pseudocode

Apriori $\left(T, \varepsilon\right)$

$L_1 \leftarrow \{$ *large 1-itemsets that appear in more than* ε *transactions* $\}$

$k \leftarrow 2$

while $L_{k-1} \neq \varnothing$

$C_k \leftarrow$ *Generate* $\left(L_{k-1}\right)$

for transactions $t \in T$

$C_t \leftarrow \text{Subset } (C_k, t)$

for candidates $c \in C_t$

$count[c] \leftarrow count[c] + 1$

$L_k \leftarrow \{c \in C_k \mid count[c] \geq \varepsilon\}$

$k \leftarrow k + 1$

return $\bigcup_k L_k$

3.2. Backpropagation

ANN's have always been regarded as the most powerful and universal predictor of all of the various kinds [29]. Considering a typical backpropagation network, there are three layers namely input, output and at least one hidden layer. Each neuron in a layer is connected with all the neurons of consecutive layer. There is no connection between the neurons in the same layer or like a type of feedback. Backpropagation is a technique based on supervised learning and is used for training artificial neural networks. First description of it was by P. Werbos in 1974, and later development of it by D. E. Rumelhart, G. E. Hinton and R. J. Williams in 1986 [30]. The way of backpropagation to learn is to process a set of training samples iteratively, to compare the network's prediction for each sample with the actual known class label. In order to minimize the mean squared error between the network's prediction and the actual class, the weights are modified for each training sample [26]. The output of *j* unit is calculated by using sigmoid activation function as follow:

$$O_j = \frac{1}{1 + e^{-\sum_i w_{ij} O_i + \theta_j}} \qquad (4)$$

where w_{ij} is the weight of the connection from unit *i* in the previous layer to unit *j*; O_i is the output of unit *i* from the previous layer; and θ_j is the threshold of the unit.
Weights are updated by the following equation:

$$w_{ij} = w_{ij} + l * E_j * O_i \qquad (5)$$

where, l is the learning rate. For a unit *j* in the hidden layer, the error is computed as follows:

$$E_j = O_j (1 - O_j) \sum_k E_k w_{jk} \qquad (6)$$

For a unit *j* in the output layer, the error is computed as follows:

$$E_j = O_j (1 - O_j)(T_j - O_j) \qquad (7)$$

3.3. Artificial Immune System

Immune systems are naturally existing mechanisms which are responsible for detecting and coping with intruders in living organisms [31]. The main purpose of the immune system is to recognize all cells (or molecules) within the body and categorize those cells as self or non self [32] and protect the organism against disease-causing cells called pathogens and to eliminate malfunctioning cells [33]. All elements recognizable by the immune system are called antigens [33]. There are two types of antigens: self and non-self. Non-self antigens are disease-causing elements, whereas self-antigens are harmless to the body [34]. There are two major groups of immune cells: B-cells and T-cells which helps in recognizing an almost limitless range of antigenic patterns. It was discovered that people who had been inoculated against diseases contained certain agents that could in some way bind to other infectious agents. These agents were named antibodies [35].

AIS is a computational technique inspired by ideas coming from immunology and used to develop adaptive systems capable to solve different domain problems [31]. The AIS [35] have become popular over the last year. Applications of AIS include pattern recognition, fault and anomaly detection, data mining and classification, scheduling, machine learning, autonomous navigation, search and optimization areas [36].

The acronym opt-aiNET stands for "Optimization version of an Artificial Immune Network" [35]. It is a particular type of artificial immune system developed to solve optimization problems [37]. Opt-aiNET is capable of either unimodal or multimodal [38].

The opt-aiNET is a valuable tool for solving a wide range of optimization problems for two main reasons:

1. It presents a good balance between exploration and exploitation of the seach-space;
2. Differently from other evolutionary proposals, it contains a mechanism devised to regulate population size and to maintain the diversity [37].

The opt-aiNET algorithm borrows ideas from two main theories about how the immune system operates, namely, clonal selection and the immune network theory [37]. Depending upon the extent of the infection, a large number of B-cells and T-cells may be required to handle the infection successfully and effectively. The size of subpopulations of these cells is controlled by a process termed clonal selection [33]. The clonal selection is the theory used to explain how an immune response is mounted when a non-self antigenic pattern is recognized by a B-cell [39]. It establish the idea that only those cells capable of recognizing an antigenic stimulus will proliferate and differentiate into effector cells, thus being selected against those that do not [35]. In brief, when a B-cell receptor (*antibody*) recognizes a nonself antigen with a certain affinity, it is selected to proliferate and it produces antibodies in high volumes. Proliferation in the case of immune cells is asexual, a mitotic process; the cells divide themselves (there is no crossover). During reproduction, the B-cell progenies (*clones*) undergo a mutation process with high rates (*hypermutation*) that, together with a strong selective pressure result in B-cells with antigenic receptors presenting higher affinities with selective antigen. This whole process of mutation and selection is known as *Affinity maturation or Immune response*. In addition to differentiating into antibody producing cells, the activated B-cells with high antigenic affinities are selected to become memory cells with long life spans. These memory cells are pre-eminent in future response to this same antigenic pattern, or similar one. The mainly features of the clonal selection principle are affinity proportional reproduction and mutation. In other words, the proliferation rate of each immune cell is proportional to its affinity with the selective antigen. The higher affinity, the higher number of offspring generated. The mutation suffered by each immune cell during reproduction is inversely proportional to the affinity of the cell receptor with the antigen. The higher affinity, the smaller mutation, and vice versa [35].

The immune network theory was proposed by N. K. Jerne in 1974 [35]. This theory describes the immune system as being composed of cells and molecules that interact with each other in a network-like from. These self-interaction patterns suggest a dynamic immune system with eigen-behaviors even in the absence of foreign stimulation (antigens). An antigen would thus be responsible for disturbing a self-organizing and self-sustainable system [37].

The opt-aiNET algorthm can be described as follows [35]:

1. *Initialization:* create an initial random population of network antibodies;
2. *Local search:* while stopping criterion is not met, do:
   - *Clonal expansion:* for each network antibody, determine its fitness and normalize the vector of fitnesses. Generate a *clone* for each antibody, i.e., a set of antibodies which are the exact copies of their antibody;
   - *Affinity maturation:* mutate each clone inversely proportionally to the fitness of its parent antibody that is kept unmutated. For each mutated clone, select the antibody with highest fitness, and calculate the average fitness of the selected antibodies;
   - *Local convergence:* if the average fitness of the population does not vary significantly from one iteration to the other, go to the next step; else, return to Step 2;
3. *Network interactions:* determine the affinity (similarity) between each pair of network antibodies;
4. *Network suppression:* eliminate all network antibodies whose affinity is less than a pre-specified threshold, and determine the number of remaining antibodies in the network; these are named memory antibodies;
5. *Diversity:* introduce a number of new randomly generated antibodies into the network and return to Step 2.

5. Evaluation

In our previous work, a method named as OPTBP for mining classification rules had been presented [11]. In this study, an AR + OPTBP method has been proposed for mining classification rules. This method consists of three-stages:

1- Minig assosiation rules and eliminating;
2-Classification of data;
3- Rule extraction.

In the first stage, the association rules which were discovered for classes and data, that provide these rules have been eliminated. This provides the training time to become a little shorter. The Apriori algorithm has been used for mining association rules. Data elemination, which provide association rules, has been inspired by the work of Karabatak and Ince [17]. In the second stage, neural network has been trained. In the third stage, Opt-aiNET has been executed for extraction rules from this ANN. Data coding, classification and rule extraction are same as the previous study [11]. After training, a nonlinear function that depends on input vector X , has been obtained [11]

$$C(X)=\left(1+e^{-\sum_{j=1}^{k}\left(1+e^{-\sum_{i=1}^{m}x_i*w_{ij}+\theta_j}\right)^{-1}*v_j+\xi}\right)^{-1} \quad (8)$$

The used dataset is the Cleveland heart disease from UCI Machine Learning Repository [40]. This dataset contains 303 samples that are taken from patients with heart problem. The dataset has 13 attributes. The dataset has two classes and the classes are coded as 0 and 1 for absence and presence, respectively. The attributes have different range values in the database and these ranges of the data can be seen in Table 1. The sub-intervals used to coding of attribute values are summarized in the third column of the Table 1.

Table 1. Attribute names, range values and coding of the attributes for Cleveland Heart Disease database

Attribute	Range	Subintervals	No. Of inputs
Age	29 - 77	[29, 50] , (50,60) , [60, 77]	3
Sex	Male , Female	{male} , {female}	2
Chest Pain Type	angina , asympt , notang , abnang	{angina} , {asympt} , {notang} , {abnang}	4
resting blood pressure	94 - 200	[94, 115] , (115, 140) , [140, 200]	3
serum cholesterol in mg/dl	126 - 564	[126, 220] , (220, 300) , [300, 564]	3
fasting blood sugar >120 mg/dl	0 , 1	{0} , {1}	2
resting electro cardiographic results	norm , abn , hyper	{norm} , {abn} , {hyper}	3

maximum heart rate achieved	71 – 202	[71, 120] , (120, 150) , [150, 202]	3
exercise induced angina	0 , 1	{0} , {1}	2
oldpeak=ST depression induced by exercise relative to rest	0 – 6.2	[0, 0.6] , (0.6, 1.6), [1.6, 6.2]	3
the slope of the peak exercise ST segment	up , flat , down	{up} , {flat} , {down}	3
number of major vessels (0-3) colored by fluoroscopy	0 – 3	{0} , {1} , {2} , {3}	4
thal	norm, fixed, rever	{norm} , {fixed } , {rever}	3

To solve the problem, firstly an Apriori algorithm has been executed. As a result of Apriori, four rules have been discovered. All of rules have %100 confidence. The discovered rules have been presented in Table 2. As a result of this stage 47 transactions have been eliminated.

Table 2. The discovered Association rules for the

No	Rule
1.	IF age $\in[29,\ 50]$ & sex=female & rest_ecr=true Then healthy
2.	IF age $\in[29,\ 50]$ & cpt= abnang & oldpeak $\in[0,\ 0.6]$ Then healthy
3.	IF age $\in[29,\ 50]$ & max_hra $\in[71,\ 120]$ Then sick
4.	IF age $\in[29,\ 50]$ & number_maj_ves=3 Then sick

Secondly, a neural network has been constructed. Each attribute value has been coded as a binary string for being used as input to the network. Table 1 has been used for coding of attribute values. With the coding scheme that is shown in Table 1, we had a total of 38 binary inputs. As the patients were classified into two classes, a single unit of the output layer was sufficient. The targeted output was 1 if the patient belonged to Class 1, and 0, otherwise. The number of neurons in the hidden layer has been taken as five.

Thirdly, the Opt-aiNET algorithm has been applied to solve the equation (8) and in order to get the vectors, which maximizes or minimizes that function. Multiplying factor is 0.5 and mutation rate is 10. The Opt-aiNET has been then run with a population of 20 for 20000 generation for each classification. All parameters have been chosen empirically for the best convergence rate between the actual and desired output. Both the maximum and minimum of output antibodies have been determined and will be translated into rules. Classification accuracy, of the proposed system is 97.7%.

6. Conclusion

In our previous work, a method named as OPTBP for rule extraction from trained neural networks using artificial immune systems, have been presented [11]. In this study, an AR + NN + AIS method named as AR+OPTBP has been proposed for mining classification rules. This method consists of three-stages:

1- Minig assosiation rules and eliminating;
2-Classification of data;
3- Rule extraction.

The dataset which is same as the application part of the previous study, are used in this study [11]. Eventually, increment of accuracy and decrement of number of the extracted rules have been seen when AR+OPTBP has been used. This also means decrement of the decision time. In spite of the number of rules decreased relatively with respect to the previous study, it's not reasonable. Therefore, in future works, classification with fewer rules by decreasing the number of rules has been aimed.

References

[1] Rajasethupathy, K., Scime, A., Rajasethupathy, K.S., & Murray, G.R. Finding ‘‘persistent rules’’: Combining association and classification results. Expert Systems with Applications, doi:10.1016/j.eswa.2008.06.090, 2008.

[2] Thabtah, F.A., & Cowling, P.I. A greedy classification algorithm based on association rule. Applied Soft Computing, 7, 1102–1111, 2007.

[3] Hahsler, M. A Model-Based Frequency Constraint for Mining Associations from Transaction Data. Data Mining and Knowledge Discovery, 13, 137–166, 2006.

[4] Pach, F.P., Gyenesei, A., & Abonyi, J. Compact fuzzy association rule-based classifier. Expert Systems with Applications, 34, 2406–2416, 2008.

[5] Hruschka, E. R., & Ebecken, N. F. F. Extracting rules from multilayer perceptrons in classification problems: A clustering-based approach. Neurocomputing, 70, 384–397, 2006.

[6] Keedwell, E., Narayanan, A., & Savic, D. Evolving rules from neural networks trained on continuous data. Evolutionary Computation. In Proceedings of the 2000 congress on evolutionary computation, 2000.

[7] Tan, S.C., Lim, C.P., & Rao, M.V.C. A hybrid neural network model for rule generation and its application to process fault detection and diagnosis. Engineering Applications of Artificial Intelligence, 20, 2, 203-213, 2007.

[8] Markowska-Kaczmara, U., & Trelak, W. Fuzzy logic and evolutionary algorithm—two techniques in rule extraction from neural Networks. Neurocomputing, 63, 359–379, 2005.

[9] Huang, S. H., & Xing, H. Extract intelligeble and concise fuzzy rules from neural networks. Fuzzy Sets and Systems, 132, 233–243, 2002.

[10] Plikynas, D. Decision Rules Extraction from Neural Network: A Modified Pedagogical Approach. Informacines Technologijos ir Valdymas, 2, 31, 2004.

[11] Kahramanlı, H. & Allahverdi, N. Kalp Hastalıkları Veri Tabanı için Sınıflandırma Kurallarının Bulunması,

IEEE 16. Sinyal İşleme, İletişim ve Uygulamaları Kurultayı, 2008.
[12] Hahsler, M., Buchta, C., & Hornik, K. Selective Association Rule Generation. Computational Statistics, 23, 303–315, 2008.
[13] Agrawal, R., Imielinski, T., & Swami, A. Mining association rules between sets of items in large databases. In Proceedings of the ACM SIGMOD International Conference on Management of Data, 207–216, Washington D.C, 1993.
[14] Agrawal, R. & Srikant, R., (1994). Fast Algorithms for Mining Association Rules. In JB Bocca, M Jarke, C Zaniolo (eds.), Proc. 20th Int. Conf. Very Large Data Bases, VLDB, 487–499. Morgan Kaufmann.
[15] Liao, S-H., Chang, W-J., & Lee, C-C. Mining marketing maps for business alliances. Expert Systems with Applications, 35, 1338–1350, 2008.
[16] Lazcorreta, E., Botella, F., & Ferna´ndez-Caballero, A. Towards personalized recommendation by two-step modified Apriori data mining algorithm. Expert Systems with Applications, 35, 1422–1429, 2008.
[17] Karabatak, M., & Ince, M.C. An expert system for detection of breast cancer based on association rules and neural network. Expert Systems with Applications, doi:10.1016/j.eswa.2008.02.064, 2008.
[18] Gallant, S. I. Connection expert systems. Communications of the ACM, 31(2), 152–169, 1988.
[19] Towell, G. G., & Shavlik, J. Extracting refined rules from knowledge-based neural networks. Machine Learning, 13, 71–101, 1993.
[20] Alexander, J. A., & Mozer, M. C. Template-based algorithm for connectionist rule extraction. In G. Tesauro, D. Touetzky, & T. Leen (Eds.). Advances in neural information processing systems, 7. Cambridge, MA: MIT Press, 1995.
[21] Sethi, I., & Yoo, J. Multi-valued logic mapping of neurons in feedforward networks. Engineering Intelligent Systems, 4(4), 153–243, 1996.
[22] Keedwell, E., Narayanan, A., & Savic, D. Creating rules from trained neural networks using genetic algorithms. International Journal of Computers Systeming Signals (IJCSS), 1(1), 30–42, 2000.
[23] Garcez, A. S. D., Broda, K., & Gabbay, D. M.. Symbolic knowledge extraction from trained neural networks: A sound approach. Applied Intelligence, 125, 155–207, 2001.
[24] Snyders, S., & Omlin, C. Rule extraction from knowledge-based neural networks with adaptive inductive bias. In Proceedings of the eighth international conference on neural information processing (ICONIP), 1, 143–148, 2001.
[25] Jiang, Y., Zhou, Z., & Chen, Z. Rule learning based on neural network ensemble. In Proceedings of the international joint conference on neural Networks, 1416–1420, Honolulu, 2002.
[26] Setiono, R., Leow, W. K., & Zuarada, J. M. Extraction of rules from artificial neural networks for nonlinear regression. IEEE Transactions on Neural Networks, 13(3), 564–577, 2002.
[27] Elalfi, A. E., Haque, R., & Elalami, M. E. Extracting rules from trained neural network using GA for managing E-business. Applied Soft Computing, 4, 2004.
[28] http://en.wikipedia.org/wiki/Apriori_algorithm, last accessed : 10 January 2009.
[29] C. Shin, U.T. Yun, H.K. Kim, & S.C. Park, “A hybrid Approach of Neural Network and Memory-Based Learning to Data Mining”, IEEE Transactions on Neural Networks, 11 (3), 637-646, 2000.
[30] http://en.wikipedia.org/wiki/Backpropagation, last accessed: 10 January 2009.
[31] Seredynski, F., & Bouvry, P. Anomaly detection in TCP/IP networks using immune systems paradigm. Computer Communications, 30, 740–749, 2007.
[32] A. Kalinli & N. Karaboga, “Artificial immune algorithm for IIR filter design”, Engineering Applications of Artificial Intelligence 18, 919-929, 2005.
[33] P. Musilek, A. Lau, M. Reformat & L. Wyard-Scott, “Immune Programming”, Information Sciences 176, 972-1002, 2006.
[34] A. Kumar, A. Prakash, R. Shankar, & M.K. Tiwari, “Psycho-Clonal algorithm based approach to solve continuous flow shop scheduling problem”, Expert System with Applications 31, 504-514, 2006.
[35] L.N. de Castro & J. Timmis, “Artificial Immune Systems: A New Computational Intelligence Approach”, Springer. UK, 2002.
[36] T. Hou, C. Su & H. Chang, “Using neural Networks and immune algorithms to find the optimal parameters for an IC wire bonding process”, Expert System with Applications 34, 427-436, 2008.
[37] R.R. de Attux, L.T. Duarte, R. Ferrari, C.M. Panazio, L.N. de Castro, F.J. Von Zuben & J.M.T. Romano, “MLP-Based Equalization and Pre-Distortion Using An Artificial Immune Network”, IEEE Workshop on Machine Learning for Signal Processing, 177-182, 2005.
[38] J. Timmis & C. Edmonds, “A Comment on opt-AiNET: An Immune Network Algorithm for Optimisation”, D. Kalyanmoy et.al. (eds), Genetic and Evolutionary Computation, Vol. 3102, Lecture Notes in Computer Science, Springer, 308-317, 2004.
[39] G.L. Ada & G.J.V. Nossal, “The clonal selection theory”, Sci. Am. 257 (2), 50-57, 1987.
[40] [http://www.ics.uci.edu/~mlearn/MLRepository.html], Irvine,CA: University of California, Department of Information and Computer Science, 1988, last accessed: 12 January 2007
[41] Kahramanli H. & Allahverdi, N., Rule extraction from trained adaptive neural networks using artificial immune systems, Expert Systems with Applications, Vol. 36, 2009
[42] Kahramanli H. & Allahverdi, N., Extracting rules for classification problems: AIS based approach, Expert Systems with Applications, Vol. 36, 2009, In press.
[43] Kahramanli H. & Allahverdi, N., Mining Classification Rules for Liver Disorders, International Journal of Mathematics and Computers in Simulation, Vol. 3, 2009.