Evaluation Average Total Data Set Learning Machine on the Meta Heuristic Algorithm

Maysam Toghraee[1], Mrziyeh Bahrami[2], Hamid parvin[3,] Sepideh Ghahramanzadeh Namin[4], Elham Safari jokandan[5,] Mohammad Esmaeili[6], Farhad Rad[7]

[1]Faculty of engineering, department of computer science, Islamic Azad University, Yasouj, Iran

[2]Faculty of engineering, department of Computer software, Islamic Azad University, Ardabil, Iran

[3]Faculty of engineering, department of computer science, Islamic Azad University, Yasouj, Iran

[4]Faculty of engineering, department of Computer software, Islamic Azad University, Ardabil, Iran

[5]Faculty of engineering, department of computer science, Islamic Azad University, Ardabil, Iran

[6]Faculty of engineering, department of computer science, Islamic Azad University, Ardabil, Iran

[7]Faculty of engineering, department of computer science, Islamic Azad University, Yasouj, Iran

Abstract
Now a days, developing the science and technology and technology tools, the ability of reviewing and saving the important data has been provided. It is needed to have knowledge for searching the data to reach the necessary useful results. The scope is to study the predictive role and usage domain of data mining in medical science and suggesting a frame for creating, assessing and exploiting the data mining patterns in this field. As it has been found out from previous researches that assessing methods cannot be used to specify the data discrepancies, our suggestion is a new approach for assessing the data similarities to find out the relations between the variation in data and stability in selection. Therefore we have chosen Meta heuristic methods to be able to choose the best and the cluster algorithms among a set of algorithms.

Keywords: feature selection, data mining, algorithm cluster, heuristic methods

1. INTROUDUCTION
Manual processing of data is impracticable because of Fast growing technologies. Even worse, machine learning techniques and knowledge extraction seem to be unable against this big data base. Vast dimensions management is one of the common challenges for extracting knowledge and machine learning.

It is clear that however these methods are suitable for efficiency, they are slow and are complicated in Calculation. Since these methods teach a classifying model for each assessment of efficiency function, these methods are appropriate for the environments with subset generator

Similar to the optimization algorithm of nature imitation. Utilized methods of classified model during learning the educational data set, feature selection is done implicitly.

An example of this kind of feature selection methods is decision tree classifying model. As these methods are explicitly part of feature selection methods, we do not use these methods.

Filter methods, Set of different subsets of features are placed according to a factor.
Then we can choose their best. The problem is that these methods try to find the best subset of features out of candidate subset 2^n . Producing all of them is hard and inapplicable but since calculating the efficiency function in these methods is mostly fast (the accuracy of classified model is not on selected features), they are appropriate for environments with subset generators similar to optimization algorithms of nature imitation.

2.1 Efficiency function
To calculate the efficiency function, we should first calculate the relationship of each feature with other features and label.

After calculation of each feature with other features and label, selecting features is done regarding to relationship amount of each feature with other features and labels. To do this, methods depended on evolutional algorithms have been represented for selecting subsets of features, in this chapter we discuss about the efficiency function of this algorithms. We are seeking for 2 targets in feature selecting. First, an evolutional algorithm is looking for subset of features which has the most relationship with label and second the selecting features have the least relationship with other selecting features. Both are demonstrated in evolutional algorithms of efficiency function explicitly.

Now, we use the following relation which shows the variation of selecting features and label similarity of selecting features to calculate the efficiency of this chromosome.

(a) Using subspaces determined for clusters or branches by user, (b) using feature selection methods or decreasing dimensions like analyzing main factors and finally (c) using subspace clustering or subspace classifying methods.

We discuss about the feature selection methods (b) in this report. A lot of solutions and algorithm have been represented for feature selection issue. A lot of solutions and algorithms have been represented for feature selection issue, some of which are 30 or 40 years old. The problem about algorithms when represented was their calculating feature. However, fast computers and big saving sources have made this problem unimportant, beside, big data sets for new issues has made it important to find a fast algorithm for this issue. Feature selection has 2 types:

a)supervised feature: Labels are used during feature selection algorithm(Zhao &Liu 2007)

b)unsupervised feature :Labels are not used during feature selection algorithm(G.D 2008)

Research domain is just limited to the supervised feature selection while labels are used during feature selection algorithm.

Purpose of feature selection is decreasing the dimensions of a set to a smaller subset of the set's features. This subset of main features is known as the best subset of features; a target function is used as being better factor which can be different targets according to the primary hypothesis. This being better can be measured, for example accuracy in classifying. We are seeking for the below targets using local searching methods and imitating the nature searching algorithm; a) discuss about the feature selection efficiency methods.
b) Explain the assessing algorithm of this research
c) looking for tests and results from these methods using real data sets.

2. EFFICIENCY FUNCTION METHODS

Different types of efficiency function for various subsets includes (1) wrapper methods, (2) embedded methods, (3) filter methods.

Wrapper methods: To asses each of the candidate feature subsets, a classifying model is done on those features (instruct and produce). We consider the accuracy of the trained classifier on a distinct experimental set as the efficiency amount of that candidate subset.

Chromosome of the population is specified according to fitness function we specified. Now we select two chromosomes for reproduction using an appropriate method (A method in which the probability of chromosome selection with better fitness amount is more than a chromosome with less fitness amount). Then using these two chromosome, we create a new chromosome (mating). We change some genes of some chromosomes having a specified probability. Selecting, recombination and mutation steps make a new population of chromosomes (new generation). If the chromosomes tend to the demanded answer, reproduction process stops. Otherwise creating a generation out of the previous

generation continues until we reach to a desired answer or ending the algorithm, condition.

3.2 particles swarm optimization algorithm

In 1995, Kennedy and Eberhard offered particle swarm optimization algorithm for the first time as a non-deterministic search methods for functional optimization (Kennedy &et al, 1995). This algorithm has got inspired from the collective movement of the birds that were seeking for food. A group of birds are looking for food in a space randomly. There is only one piece of food in the discussed space. None of the birds know the location of food. One of the best strategies can be following a bird who is closer to the food. This strategy is the basis of the algorithm (M. carvalho& et al, 2006). Each solution, which is called a particle, is the same as a bird in the algorithm of a mass movement of birds. Each particle has a fitness value which is determined by a fitness function. Particle swarm works based on this principle. In each moment each particle sets its location according to the best location in which it has been located in the searching atmosphere and in the best location which is in its neighborhood. The time passing. $X_i^{(t)}$ Determines the itch position of the particle in the tth time. Also, each particle needs a speed for moving in the space. $V_i^{(t)}$ is the speed of ith particle in t^{th} time. Adding speed to the position of each particle, a new position can be considered for the particle. The function of updating the position of particle is as below:

$$X_i^{(t+1)} = X_i^{(t)} + V_i^{(t+1)}, where \ X_i^{(t)} \sim u(X_{min}, X_{max}) (3\text{-}2\text{-}1)$$

The suitability of a particle in the search space is evaluated by the fitness function.

$$fit_{ch} = \sum_{i=1}^{f} lesser\left(max_{k=1}^{f}\left(|cor(X_i, X_k)| \times and(ch_i, ch_k)\right), th_2\right) \times \alpha + greater(|cor(X_i, T)|, th_1) \qquad (1-2)$$

Where f is the amount of features and α is the big positive number, th_1 and th_2 are two thresholds which should be adjusted by the user, ch_j shows Ith of chromosome $_{,//}$ shows the absolute value, $cor(X_j, X_k)$ shows the relationship of I th and k th features , and (ch_j, ch_k) show the logical operator (output is 1 when both inputs are 1,otherwise function output is 0) , (a,b) is greater than1 if $a \geq b$,otherwise function output is 0 and function(a,b) is less than 1, if $a \leq b$. Otherwise function output is 0 .

3. Evaluation Algorithms

3.1 Genetics Algorithm

The main idea of evolutional algorithms were presented in 1960 by Rothenberg. Genetic algorithms are derived from this type of algorithm. In fact the computer search methods based on optimization algorithms based on the genetic

International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 6, Issue 4, July- August 2017   **ISSN 2278-6856**

structure of chromosome's by john Holland(1970) was introduced at the university of Michigan(A.mehdi;1386). The most extensive definitions of genetic algorithm are from Goldberg's book:" Genetic algorithms is machine learning model, its behavior is an example of the evolutionary processes in nature". Genetic algorithm is one the strongest methods derived from the nature which seeks for the problem randomly. This search is seeking for creating better answers in each generation compared to the answers of previous generations which represents one of the best forms of number optimization on the issues of science and engineering (masoudian & et al;2007). We explain the genetics algorithm steps based on (C. Meyers &et al; 1996): First we answer the question by defining a chromosome structure (coding). Introducing the fitness function, we explain the quality of the given answers in each chromosome numerically. Then we generate some chromosomes randomly (or semi-random. And the chromosomes are known as the initial population. We have some answers for the problem in this step which have lower quality. The quality of each

In order to produce a selected food situation already, bee algorithm uses the above phrase:

$$V_{i,j} = x_{i,j} + \emptyset_{i,j}(x_{i,j} + x_{k,j}) \qquad (3\text{-}3\text{-}2)$$

Where $K \in [B_n]$ and $j \in [D]$ are selected randomly. Although K is determined randomly, but must be different from $i.\emptyset_{i,j}$ is a random number between [-1,1].This variable controls generating the neighbor food source around the Xi,j ad shows the comparing changes of the food source positions by the honey bee visually. More the difference between Xi,j and Xk,j decreases, deviation from Xi,j situation decreases. If the parameters generated by these operations are more than their preset amounts, this parameter Parameters can be selected as an acceptable value (D.Karaboga; 2007).

$$X_i^j = X_{min}^j + rand[0,1](X_{max}^j - X_{min}^j) \qquad (3\text{-}3\text{-}3)$$

After generating and fitting the new population new selection is done out of the new population. These are continued as long as the number of iterations of the algorithm ends.

3.4. Big Bang Algorithm

First algorithm was introduced by Eksin and Erol (I.Eksin&K.Erol, 2006). This algorithm uses the phenomenon of the big bang and then the contraction of the universe in the center of gravity. This algorithm has higher speed of convergence compared to other algorithms. Particles are scattered in searching space randomly like other evolutional algorithms which is called big bang phase (C.V. camp, 2007).

Each particle has a position which determines what the particle coordinates in the search space is. Position of the particle changes by its movement in phase, all the particles accumulate around the center of gravity. This phase acts

like a converging operator which is calculated through the following equation:

$$X_i^{(k)} = \frac{\sum_{j=1}^{N} \frac{x_i^{(k,j)}}{f_j}}{\sum_{j=1}^{N} \frac{1}{f_j}} \qquad i = 1,2,3,4,\dots,c \qquad (3\text{-}4\text{-}1)$$

Where $X_i^{(K)}$, is the ith particle of the center of gravity in k^{th} repetition? $X_i^{(k,j)}$ is the Ith component of jth particle produced in k^{th} repetition.

3.3. Colony Bees Artificial Algorithm

Ants algorithms based on intellectual foundation can be simply stated in one sentence: Ants select the best way between different ways of reaching food among the barriers in nature. The short way is always chosen. Ants secrete a substance called pheromone after finding food which is seen white after the rain. They find their way finding the pheromone way. Bees are in three groups in colony bee's algorithm bees:

Worker bees, the audience and the pioneers (scout). A honey bee stays in dance region to make a decision for choosing a food source is called searching honey bee and and a honey bee, which is looking for the specified is called worker bee. A bee which searches randomly is called pioneer honey bee (scout). In the bee's algorithm, worker bees are half of the population and the other half are searching bees. For every food source there is only one worker bees, in other words, the number of worker bees around the hive equals the number of food sources. The worker bees who are tired of working in the food supply are leading wasp's browser. The main steps of the algorithm is given below:

Initialization
Repeat
    a) The location of worker bees in food supplies in memory
    b) The location of searching bees in food sources in memory
    c) Sending the pioneer wasps bees to search for new food sources.
(Until the desired situation gets achieved).

A searching bee chooses the food source regarding the possibility related to that source, P_i, which is calculated by the following phrase:

$$P_i = \frac{fit_i}{\sum_{n=1}^{sn} fit_n} \qquad (3\text{-}3\text{-}1)$$

Where fit_i is the fitness value of the solution i by worker bees. That is in proportion to the amount of nectar in food source in position i. SN is the number of food sources that equals the number of worker bees (EB) or searching bees

International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 6, Issue 4, July- August 2017   **ISSN 2278-6856**

(OB).). In this method, the worker bees interchanges their information with the searching bees.

Displaying the patterns refers to the number of classes, available samples and the number, type and scale of features in a clustering algorithm. Some of these data are not controlled by the user. Selecting feature, Identifying process of a subset is one of the most effective features for using in clustering and feature extracting is, the process of changing some available features and generating new features. Both of these techniques are used in order to achieve a suitable set of features and enhancing the performance of clustering. Adjacent of samples usually is measured by a function of distance between the pair of input patterns. Various criterion are used to measure distances in various fields. A simple measuring criteria such as Euclidean distance is usually used to represent the dissimilarity between the two models. Identifying conceptual similarities between the patterns, evaluation criteria could no longer be used. Grouping step has various methods. The results of the clustering can be in from of hard segmentation (dividing data within separate groups) or phase segmentation (each sample is located in various groups with different membership grades). Hierarchical clustering algorithm represent nested chains of different divisions on the basis of dividing and combining clusters by similarity function. Non-hierarchical clustering algorithm gives a division of the samples. There are other techniques, such as possibility clustering and graph theory clustering. Summarization of data is the extraction process of a precise and simple of the original data set. In clustering, a sample of data summarization could be a short description for each cluster is usually referred to as per-pattern model. For example, the center of gravity for each cluster could be a brief description of the sample within that cluster.

4. Evaluation Methods

In this section, the results of applying the proposed method on different data sets and used parameters has been reported. In this section, analysis and interpretation of the results and efficiency of the proposed algorithm also briefly discussed. In this research validation includes experimental validation. In experimental validation, the efficiency of proposed algorithm is discussed in comparison with other algorithm in several real data sets. UCI standard data sets are the used datasets(it is on the machine learning website) is that almost all the result of recent studies in the

f_j is the target function for point j and N is the number of points or particles and C the number of controlling variables respectively, after determining the particles center of gravity, the new position of particles can be calculated using the following equation :

$$X_i^{(k+1,j)} = X_i^{(k)} + (X_{imax} - X_{imin}) \times \alpha_1 \times r \times \left(\frac{1}{k+1}\right) \quad (3\text{-}4\text{-}2)$$

r is the random number, X_{imax} and X_{imin} respectively are lower and upper constraints to limit in the above equation . α_1 is also a parameter for limiting the search space.

3.5. Hill climbing Algorithm

In Hill climbing algorithm, at first an answer to the question is generated randomly. Then in a loop until the stop condition is not established for the algorithms, a number of neighbors are generated for the current mode of production and the best one is chosen among them and replaces the current mode. (Of course, another definition for the hill climbing has been stated). In general, the optimality of the answer to the algorithm is local. To run Hill climbing algorithm we need two functions: performance function and neighbor function. The performance function determines optimality of the answer. 8 minister of guards on board pair's queens returns. Neighbor function also produces current mode neighbors. In problem 8 queens are for generating neighbor modes, each of them are chosen and move once upward or downward. This means that in the worst mode of each case, there will be 16 neighboring modes that in each repetition of the loop as the best answer will be replaced neighbors. Ends when there is not a better mode compared to the current mode.

3.6. Clustering algorithm

A general trend for clustering process includes the following steps:

1. Displaying the patterns that usually involves selecting or extracting the feature.
2. Defining an assessment criteria of the similarity according to the data domain.
3. Clustering or Grouping process.
4. Summarization of data if needed.
5. Validation of system.
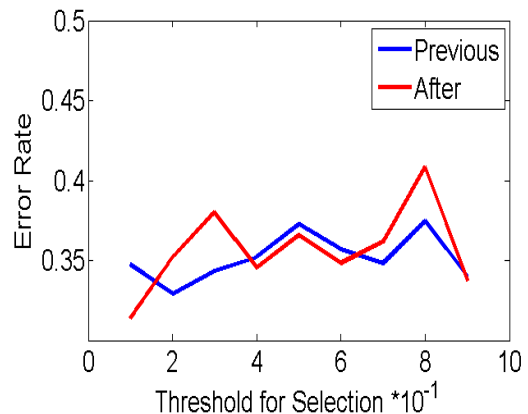
5. Results from Algorithms

Figure1 average error rate from total data set in Genetic algorithm

In genetic algorithm on average all dataset, pervious error and after error rate situation in this thresholds better. Because they are nearly. In threshold 0.2 and 0.9 pervious and after error rate is 0.35 and this showed that situation in this threshold is better. Because situation in thresholds is low.
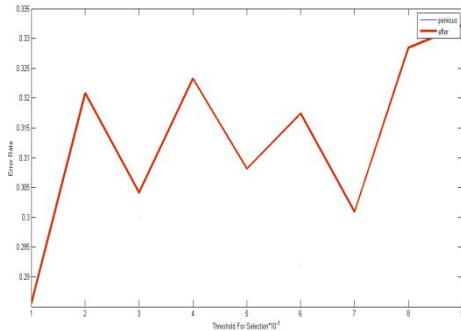

Figure2 average error rate from total data set in hill climbing algorithm

In hill climbing algorithm on average all dataset, situation pervious error and after error isn't favorable. And pervious error rate in this situation is INF. And pervious error rate and after error rate aren't nearly. Because after error rate in threshold 0.1 is low and this situation for mode average isn't favorable.

Field of data mining in the world are reported using this data set. These actual data are often standard. Experimental results of the proposed method and other methods with the world's valid criterion such as error size the feature set size and chosen will be reported.

Table1 datasets used in the first experiment in this thesis. Starred data set"*" are real data sets.

Dataset Name	# of data items	# of features	# of classes	Data distribution per clusters
Breast Cancer*	404	9	2	444-239
Bupa*	345	6	2	145-200
Glass*	214	9	6	70-76-17-13-9-29
Galaxy*	323	4	7	51-28-46-38-80-45-35
SAHeart*	462	9	2	160-302
Ionosphere*	351	34	2	126-225
Iris*	150	4	3	50-50-50
Wine*	178	13	3	59-71-48
Yeast*	1484	8	10	463-5-35-44-51-163-244-429-20-30

It should be mentioned that all results represented in this thesis are achieved, because the test results are so strong and extendable, with an average of 10 separate performance. In the first experiments have been performance on several real data sets. The actual data sets have been derived from machine learning website (Newman, 1998). These data sets is presented in the table above. Test results of the standard features of these data sets have been reported. In other words, each of the properties of this dataset with zero mean and variance one, N (0, 1) are standard. Thresholds th_1 and th_2 changing from 0, 1 to 0, 9 and training set size from 10 percent to 90 percent, different levels of accuracy are achieved. Notice that each performance is done10 times repeatedly. In all cases the maximum population size should be adjusted on 50 and the maximum generation size on 50.
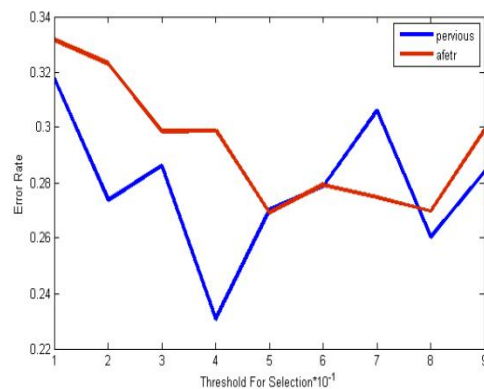

Figure5 average error rate from total data set in particle swarm optimization algorithm

In particle swarm optimization algorithm on average all dataset, situation is favorable. And in threshold 0.4 error rate has low. Pervious error and after error are near. But in threshold 0.1 error rate than to other thresholds is height. In threshold 0.9 situation error arte goes on height.
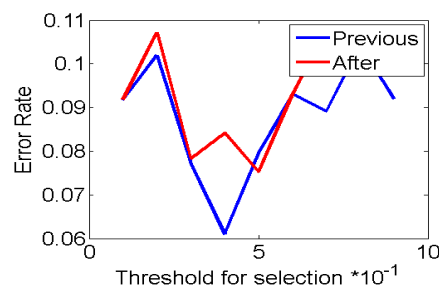

Figure6 average error rate from total data set in cluster algorithm

In cluster algorithm on average all dataset pervious error and after error situation in threshold 0.4 error rate have low, but in this situation, on average all dataset in this cluster algorithm then to other algorithms are better. Because in this mode, errors are nearly. After error rate and pervious error rate in threshold 0.4 situation were very better. Situation pervious errors in all thresholds level have

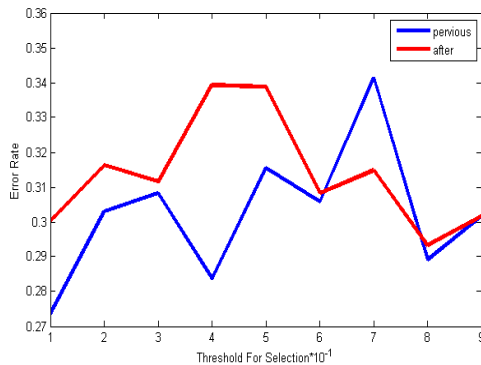lowest. From these experiments obtained, we conclude that the situation clustering algorithm is better.

Figure3 average error rate from total data set in artificial bee's algorithm

In bees algorithm on average all dataset, situation pervious error rate and after error rate is favorable and they error rate have low, and in threshold 0.1 minimum has pervious error rate. But in other thresholds situation isn't favorable. Just in threshold 0.9 they have collision but error rate is height.
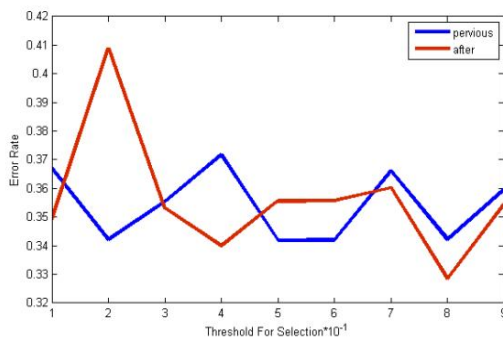

Figure4 average error rate from total data set in big bang algorithm

In the big bang algorithm on average all dataset previous error and after error just in threshold 0.8 error rate than in other threshold is better. And situation in all threshold is favorable. Because pervious and after error situation have nearly. But error rate haven't lowest and this for average all data isn't Appropriate. But speed in this algorithm than to other algorithms is height.

6. Conclusion

In this study, the efficiency of the proposed algorithm, advantages, challenges and innovations have been discussed briefly. A set of feature selecting method based on the collective intelligence methods has been presented. It was shown experimentally that these methods can have no decrease in classifying quality, selecting almost 80% of features .Beside ,these experiments showed that clustering method is the best way in finding appropriate features for classifying .

7. Future Works

Feature selection methods is one of the most challenging and the most important activities in developing machine learning and patterns recognition. Feature selection is one of the issues which has been discussed in machine learning and also pattern statistical recognition. This issue is important in many usages (e.g. classification), since there are a lot of features in these usages. Many of them are useless or are not informative**.** Eliminating these features does not change the informative content but it effects on calculating feature of the mentioned usage Also it helps to save much useless information with useful data. Not elimination of the waste features makes some dimensional problems. Dimensional problem says that when the dimensions increase, it is possible that the 2 data (or sample) get far from each other. It means that the higher the dimensions (features), probably the most distance between them randomly. As the result, those samples are affected by the dimension and the distance between those samples are estimated much more. It makes the distance between both samples less representing the real distances. So, the quality of classifying or clustering are unpleasantly unreal and drop. It can be stated in another way. It can be said that some clusters or branches in feature's atmosphere are more coherent with some special features; three general ways have been submit to overcome the above dimension problem: As future works, we should accomplish the data sets on hierarchical and partition clustering algorithm as it selects the final cluster qualitatively for us.

REFRENCE

[1]. Algorithm", PATAT, pp. 528–531. ISBN 80-210-3726-1.(2006),
[2]. Marki F and Vogel M and Fischer M. "Process Plan optimization using a Genetic Algorithm", PATAT, pp. 528–531. ISBN 80-210-3726-1.(2006).
[3]. Masoudian S and ESTEKI A,." Design schedule automatically using genetic algorithm ", thesis, university of Isfahan. .(1386).
[4]. Matlab version 7.4.0.287(R2012a), 29 january 2012, U.S.Patents Carol Meyers and James B. Orlin, "Very Large-Scale Neighborhood Search Techniques in Timetabling Problems", PATAT 2011, pp. 36–52. ISBN 80-210-3726-1. (2011).
[5]. Mehdi A." Introduction to genetic algorithm and application", Tehran: Bell Press naghos. (1386).
[6]. Murata S and Kurova H. Self-Organization of Biological Systems. (2012).
[7]. Neumann J, C and Schnar G. S. Combined SVM-based feature selection and classification, Machine Learning, Vol. 61, No. 3, pp. 129 – 150.(2005).
[8]. Perzina R. "Solving the University Timetabling Problem with Optimized Enrolment of Students by aParallel Self-adaptive Genetic Algorithm",PATAT 2006, pp. 264–280. ISBN 80-210-3726-1. (2006).
[9]. Pham D. T and Ghanbarzadeh A and Koc E nad Otri S and Rahim S., Zaidi M. The Bees Algorithm –A

Novel Tool for Complex Optimization Problems. (2006).

[10]. Susana M and Vieira J and Sousa M.C. Fuzzy criteria for feature selection, Fuzzy Sets and Systems, Vol. 189, No. 1, pp. 1–18. (2012).

[11]..Zhao Z and Liu H. Semi-supervised feature selection via spectral analysis. In Proceedings of SIAM International Conference on Data Mining (SDM). (2007).

[12]..Camp C.V. "Design of space trusses using Big Bang – Big Crunch optimization", Journal of Structural Engineering, vol.133, Issue 7, pp.999-1008,(2007).

[13]..Dy J. Unsupervised feature selection. Computational Methods of Feature Selectionpages ,19-39. , (2008).

[14]..Erol Osman k and Eksin I. "New optimization method : Big Bang-Big Crunch", Elsevier, Advances in Engineering Software 37 pp: 106–111 .(2006).

[15]..Guyon I and Elise A. An introduction to variable and feature selection. Journal of Machine Learning Research,.3:1157-1182, (2003).

[16]..Jain k ,. Dubes R. C. Algorithms for Clustering Data, Prentice Hall, Englewood Cliffs. (1988).

[17]..Karaboga D, Ozturk C. A novel clustering approach: Artificial Bee Colony (ABC) algorithm. (2011).

[18]..Kennedy J and Eberhart R, C. "Particle Swarm Optimization", Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ,. pp: 1942- 948. (1995).

[19]..Lampinen J   and Laksone J and Oja E. Pattern recognition. In editor, Image Processing and Pattern Recognition, volume 5 of Neural Network Systems Techniques and Applications, pages 1- 59. Academic Press. (1998).

[20]..Ladha L and Scholar R, Depa T, LFeature Selection Methods and Algorithms, International Journal on Computer Science and Engineering (IJCSE), Vol. 3, No. 5, pp. 1787–1797. (2011).

[21]..Marki F and Vogel M and Fischer M. "Process Plan optimization using a Genetic

Marziyeh. Bahrami received her M.Sc. degrees in Computer Software Engineering with highest distinction from the Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Ardabil, Iran, in 2016. She is the co-author of a book on Advanced Data Mining (in Persian) in 2017. Her researches is related to Distributed Systems with particular emphasis on Body Sensor Networks, Big Data, Internet of Things and Cloud Computing.

AUTHOR

MAYSAM TOGHRAEE first author is this article. Maysam toghraee in a family toghraeee semiromi was born in 1987 in semirom and primary and secondy education the in city end ,and after goes in Esfahan in university shahid ahrafi esfahani succeeded to a bachelor's Degree in since computer in 2010. and went in university yasouj in field of engineering software computer continuing education, and my own thesis in the field of data mining with matlab software was defanes.