Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

Speech Emotion Recognition Using Fuzzy Logic Classifier

Daniar aghsanavard
Department of Electrical Computer and Biomedical Engineering,
Qazvin Branch Islamic Azad University
Qazvin,Iran
Email: daniaraghsanavard@gmail.com
Dr . Vahid rostami
Department of Electrical Computer and Biomedical Engineering,
Qazvin Branch Islamic Azad University
Qazvin,Iran
Email:vh_rostami@yahoo.com

ABSTRACT

Over the last two decades, emotions, speech recognition and signal processing have been one of the most significant issues in the adoption of techniques to detect them. Each method has advantages and disadvantages. This paper tries to suggest fuzzy speech emotion recognition based on the classification of speech's signals in order to better recognition along with a higher speed. In this system, the use of fuzzy logic system with 5 layers, which is the combination of neural progressive network and algorithm optimization of firefly, first, speech samples have been given to input of fuzzy orbit and then, signals will be investigated and primary classified in a fuzzy framework. In this model, a pattern of signals will be created for each class of signals, which results in reduction of signal data dimension as well as easier speech recognition. The obtained experimental results show that our proposed method (categorized by firefly), improves recognition of utterances.

Keywords - **speech emotion recognition, fuzzy logic, Fly-FNN, firefly, noise- taking, progressive neural network.**

Date of Submission: July 05, 2015 Date of Acceptance: Aug 12, 2015

1. INTRODUCTION

Over the past, speech emotion recognition has been a significant issue in signal processing debates and various methods have been adopted to distinguish it that each one has advantages and disadvantages. For instances,

particle swarm optimization (PSO) merged with fuzzy neural networks (FNNs), here after called the PSO-FNN method, back propagation algorithm–back-propagation algorithm for feed forward neural network training BP-FNN methods did not have the recognition ability of similar signals and by rapid extracting of signal features cause less compared with similar signals, in addition, surrounding noises terminate the specific characteristics of signal. One of the shortcomings of fuzzy network (FNN) is that it faced with massive amount of speech signals that it causes the processing time become longer and also, the optimization order of the neural network in speech recognition is important in order to have an optimization order of neural network and its corresponding weights (1-3).

In this article, we present a new method by FNN classification which removes the pervious flaws such as low-speed signal processing and the problems related to the recognition of closed signal features and it acts with greater speed and higher accuracy in speech recognition. In this method, by considering test patterns and sensitive edges of signal word, a pattern will be presented to terminate environment noises.

High speed of speech recognition, 1) using becoming small sent signals base by fuzzy orbit, 2) high accuracy in speech emotion recognition, 3) separation of signals will be done using clustering of firefly and by making noisy classes for each speech emotion, we will be able to recognize emotions of speech in noisy situation and removing noise with primal signals will be done.

In the previous methods for speech emotion recognition, the original signal was used to remove noise that may leads to damage to signal. This damage is due to noise incorrect detection of signal. In the taking-noise part of our proposed method for removing noise from speech emotion recognition, after possible recognition of speech emotion, we compare input signal with trained signals and if there is any similarity between signals, by using noisy classes and trained and noiseless signals, we remove the noise of input signal. By our presented method in this paper, we do the speech emotion recognition by checking fuzzy circuit output as well as comparison closed classes with each other, which in addition to increasing of accuracy it have higher speed.

We encounter with uncertainty of words in the recognition of speech emotion that will be expressed in the various positions. For this problem, we use a 5-layer fuzzy circuit in order to help us in uncertainty. Classification in speech emotion helps us to choose the closer class. This leads to a

Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

higher rate of detection of emotion in speech as well as the initial guess in order to obtain accurate recognition of speech emotion by accurate comparison of fuzzy circuit output.

2. DESIGN OF FUZZY SEGMENT

As mentioned above, our proposed method for speech emotion recognition is composed of 5 layers which according to figure 1, X_i are the input of speech's signals and y_i are the output of circuit. As you can see, the outputs size is smaller than our inputs. The long of these outputs is depended on the number of rules that have been established in fuzzy circuit. These rules depend on various signals which having different sizes. The outputs for each input speech signal in continuous mode, allocate some as key or signal indicator.

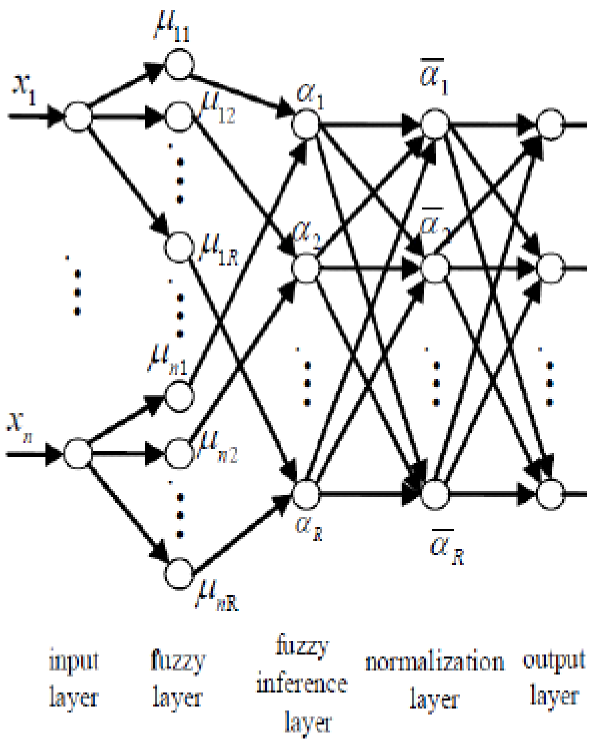


Figure 1: Fuzzy diagram

2-1. FUZZY MODELLING INPUT

Equation 1, is a triangular membership function where a, b, c are the placement locations on signals.
In this equation, the fuzzy inference system receives the inputs and it determines the membership degree of inputs for each fuzzy sets.

$$u(x;a,b,c)=\max\left(\min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),o\right) \qquad (1)$$

2-2. APPLYING FUZZY OPERATORS

After fuzzy making, the degree of correctness of each of component parts of premises (inputs) is determined.
Equation 2 is the calculate function of fuzzy interface.

$$\prod_{i=1}^{R}\mu_{ij} * \frac{1}{N_{adj}} \qquad (2)$$

In which R is indicator of the number of parts of fuzzy operator which has been created to incorporate, by which parts accuracy degree and number production will be provided as accuracy degree of the assumption part: $N_{ad\,j}$ is equivalent coefficient that we consider it equals with n/4 in this article: n is the number of network input signals.

2-3. APPLYING IMPLICATION METHOD

Membership function determines the result of fuzzy sets. Process input is indicator of one number and its output is a fuzzy set. According to equation 3, the node input jth equals with:

$$\bar{a}_j = a_j / \sum_{i=1}^{R} a_i \qquad (3)$$

2-4. AGGREGATION OF OUTPUTS

Since in a fuzzy interface, decisions will be made based on the assessment of all the rules, we combine them in these layers and according to equation 4, kth node output equals with:

$$y_k = \sum_{j=1}^{R} W_{jk}\bar{a}_j , k = 1,2,A,R \qquad (4)$$

In which Wjk is the weight of each rule and it is applied on the value obtained from the assumption. After creating output signals, for better recognition, we classified them using firefly algorithm. This method is significantly faster than other clustering method which we will show the accuracy degree of this method.

3. EXTRACT FEATURES FROM INPUT SIGNALS

One of the important parts in identification systems of speech emotion is selecting important features of word. Feature extraction will be done based on partitioning into short space namely frame. PITCH signal has useful information in conjunction with the word because it is the result of vibrations for sound production. PITCH signal will be created due to the vibrations occurred in larynx.
Vibration rate in vocal fold is known as the fundamental frequency. The next important feature for recognition of speech emotion is energy. It is important because change in energy of speech emotion signal is important based on emphasize on the importance of word. MFCC is considered as the spectral features in automatic recognition of emotion in speech and in speech emotion recognition.

Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

4. CLUSTERING FUZZY OUTPUT SIGNALS USING FIREFLY ALGORITHM

Clustering is an important unsupervised classification technique in which a series of patterns (usually vectors in a multidimensional space) are classified in clusters based on similarity measurements such as: Euclidean distance, Mahalanvbys distance, Chertoff distance, etc. Clusters are often used for various applications such as data analysis, image analysis, data mining and other engineering and scientific disciplines. Clustering algorithm is divided into two groups: 1) hierarchical clustering 2) partitioned clustering. Hierarchical is a hierarchical structure of clusters which is made by splitting a large cluster into smaller clusters and then merge smaller clusters with respect to the obtained nearest gravity centre. Here, there are two main methods for hierarchical clustering: a) separation method that divides a large cluster into two or smaller cluster, b) compressed method in which a large built by combining two or more smaller clusters. Partitioned cluster tries to divide data set into a set of discrete clusters without hierarchical structure. In most cases, the partitioned cluster algorithms uses clustering algorithm based on prototype, where each cluster is represented by its canter. The objective function (error squared function) is the sum of patterns distance from centre. In this paper, we further use partitioned clustering to generate cluster canters. So, we classify data sets using this cluster.

4-1. FIREFLY ALGORITHM

Fireflies are the insects which shine on. We use these three following desirable rules for easier description of firefly algorithm.

- All fireflies are from one gender so that one firefly attracts other fireflies irrespective of its gender.
- An important and interesting behaviour of fireflies is that the firefly which shines brighter attracts prey and divides food with others.
- Brightness is a measure of attractiveness of a firefly. So, firefly moves towards his neighbour who shines brighter.

Firefly algorithm (FA) is population-based algorithm. This algorithm seeks to find an optimal global objective function based on the exploratory behaviour of firefly. In FA, the physical (agents or firefly) are randomly distributed in problem space. Agents are known as firefly and the quality of light is called light intensity. Each firefly is absorbed by other brighter neighbours. Attractiveness decreases with increasing distance between them. If any fireflies are not brighter than others, then they will move randomly. In applying clustering FA are decision variable of clusters canter. The objective is the sum of Euclidean distance of all training data samples in N-dimensional space. Agents based on this objective function, are distributed and randomly and will be initially quantified.

Firefly two-phase algorithm is as follows:

Changes in light intensity: Light intensity change: light intensity is related to the target values. So that the problem of minimizing and maximizing of light intensity, if fireflies emit more light so they can absorb most of their fellow and if they have less radiation so they attract less fellow. We assume that we have a set of n factorial (firefly) and X_i is a solution for Ith silkworm that f (X_i) shows its fitness value. Here, light intensity of a firefly chosen to reflect the current position x and also fitness value f(X).

$$\mathrm{I}_i = f\left(x_i\right) \qquad 1 \le i \le n. \tag{5}$$

The move towards more attractive firefly: amount of light which is radiated by a fascinating firefly is visible by his/her adjacent firefly. Each firefly has specified absorption which it determines the ability's degree of firefly in absorption of other cluster members. However, the absorption degree of firefly is related to r_{ji} which is indicator of distance between two fireflies i.e. ith and jth and are in the location of X_i and X_j, respectively. r_{ij} is defined as follows:

$$R_{ij} = \left\| \mathrm{X}_i - \mathrm{X}_j \right\| \tag{6}$$

The attractiveness function of a firefly is determined by the following equation:

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{7}$$

In which, absorption degree is indicated in r=0 and is called light absorption coefficient.

The movement of ith firefly which is located at position X_i and is moving towards brighter jth firefly which is located at X_j, is shown by equation 8:

$$x_i(t+1) = x_i(t) + \beta_0 e^{\gamma r^2}\left(x_j - x_i\right) \tag{8}$$

4-2 CLUSTERING ALGORITHM OF FIREFLY

Clustering methods have been developed regardless of the signals in groups or classes based on unsupervised learning. Training data sets in unsupervised technique are grouped based on the numerical information in data (such as clusters canters) and then they will be matched by analyst of informational classes. The data sets that we track contain class information for each data. Therefore, the main objective is finding the cluster centres by minimizing the objective function (the total distance of patterns of cluster centres). For N object in giving problem, the purpose is minimizing the square sum of Euclidean distance between all patterns and assigning each pattern to one of cluster centre K. Clustering objective of error square sum is calculated through equation (9):

$$J(K) = \sum_{k=1}^{k} \sum_{i \in c_k} \left(x_i - c_k\right) \tag{9}$$

In equation 9, K i.e. the number of clusters for n (pattern= I X_i, 1, 2,…….n) is the location of ith pattern and c_k (c= 1,2,……K) is kth of cluster canter which is calculated by this equation (10):

$$c_k = \sum_{i \in c_k} \frac{x_i}{n_k} \tag{10}$$

Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

In equation 10, n_k is the number of patterns in Kth cluster. Data sets are allocated to clusters in clustering analysis so that patterns are classified based on some similarity criteria in a cluster. Similarity measuring criteria is used to assess distance between patterns. Cluster canters are decision variables that are achieved by minimizing the sum of Euclidean distance on all training samples in n-dimensional space.
The objective function for i-th pattern is calculated by equation 11:

$$f_i = \frac{1}{D_{Train}} \sum_{j=1}^{D_{Train}} d\left(x_j, p_i^{CL_{known}(x_j)}\right) \qquad (11)$$

In equation 11, D_{Train} is the number of training data set is used for normalizing the sum and is placed between (0.0, 1.0) and is defined as a class in which samples belong to it in according with database.
Note that decision variables are clusters centres in FA algorithm. The objective function in firefly algorithm is determined by equation 11. For a data set, n indicate the number of data points, d indicate the dimension of problem and c indicate the number of classes. A signal data point belongs to one of c classes.

5- RECOGNITION OF NOISE IN CLUSTERING

In this clustering method, because each signal is in a special class, and noise is an undeleted factor in environment, so, we have a noise in classes in each pattern which decrease the recognition accuracy. For a specific pattern to noise is recognized in a class and to have fewer errors in the speech emotion recognition, in our training pattern for each individual, we put the average of variation in expressing a word as circuit input pattern and we save it in a special class which is called noisy class. Whatever the number of noisy classes is more, it helps more to noise recognition in the environment.
For example, we consider expression sample of number 3. Note that we obtain these signals in 12 different states and with different noise. According to figure 2, noise at the beginning and the end of and also in signal of uttered word is perfectly clear.

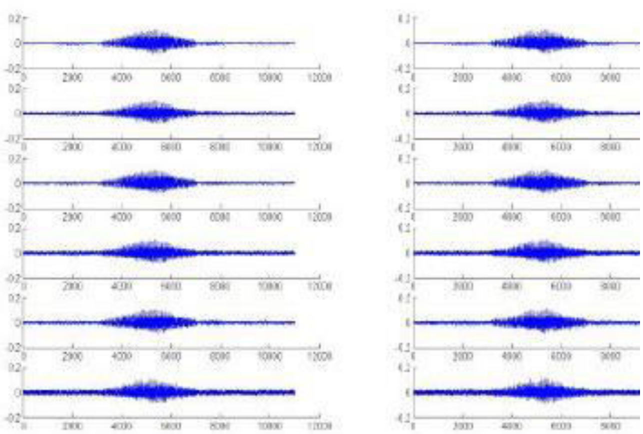
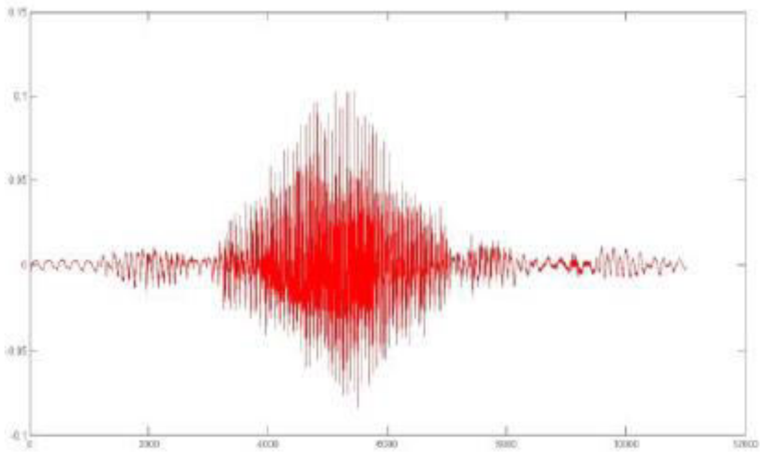


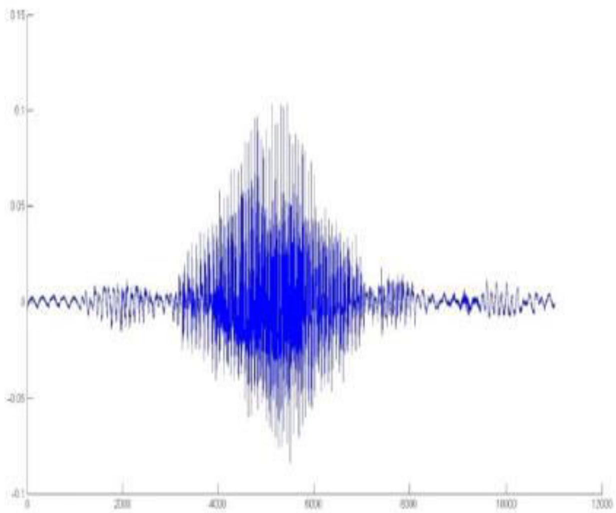
Figure 2: expressing sample of number 3

In taking -noise part of signals, first, we find the mean difference of noiseless signals with noisy signals. This difference determines the amount of signal noise and then we give it to circuit as a noisy class sample in order to be saved in a noise class related to that word

In application test, we compare uttered word (number 3) with collected data set explained in part 4. Then, we fetch the noise class of this signal and we decrease the difference with input signal in frequency domain, we compare obtained signal with primal signal, if they are equal or having few difference so the signal of noise will be removed, if not, it will be considered as a noise signal and mean of that signal will be added to the mean of the pervious signals and then another noisy class will be created. In 3- a figure, we observe the main signal and in 3- b figure we observe the noise which its background has been completely deleted also the main quality of speech emotion along with its quality and details is observed.



3- a figure: the primal signal



3-b figure: corrected signal by using of noisy classification

5-1 noise deletion method of signal

In this part, noise deletion is done based on the explanation in part 4. Our introduced firefly clustering algorithm has classification of data or clustering, to do this, we consider noisy and noiseless class as two

Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

separated part. In mentioned method, we do clustering with two classes of noise and noiseless signal. After determining of clustering centres, we compare input signals with canters and we relate them to the closest class. If input signal is coincided so we conclude that there is not any noise in the system, but if there will be coincided on noisy class so we continue the clustering process in 12 noisy classes. That is, we seek a particular noisy pattern for that word. After determining the noisy class by considering the relationship that is in input signal, we can delete the related noise.

6- FA CLUSTERING AND PARAMETER SETTINGS

Fireflies are randomly quantified in search space. Values of the parameters used in the algorithm are as follows:
The number of firefly N equals with 20.
The initial absorption rate equals with 1.
The light absorption coefficient equals with 1.
The number of nasal T equals with 100.
Test and performance
In this paper, we use 45 subjects, 25 females and 20 males, to establish a database. 30 Persian words and the Persian pronunciation of numbers from 1 to 30 per person were done and 6 kinds of pronunciations have been done for each word. The speed of sampling frequency is considered 11.025 kHz. This algorithm has been done in 4 noise model in the presence of Gaussian white noise, with signal degree to input noise db 15, 20 db,25 db, 30 db and a noiseless model. We implement the proposed algorithm, we do clustering after entering speech emotion signals to fuzzy circuit and after obtaining output signals (the output of fuzzy circuit). This clustering for one speech places in a set of the pronunciation of that number (for expressing word 3, 45 classes of letter 3) and for every class, 6 various samples of signal of one person) and we consider a noisy signal class for each speech from each person. In investigating speech emotion of a speaker, the class of that speaker is created and is compared with other classes in order to select the closest class. Then, to more accurate recognise, we compare fuzzy output circuit with founded classes fuzzy circuit output, if any similarity observed between two signals, we investigate the noise between this signal and we remove the signal noise by mentioned method in part 3-4 and then we update the noise of noisy classes.
In table 1, the output of algorithm with respect to above noisy models is compared with other algorithm and the effect of FAFC-FNN method in increasing the accuracy and efficiency of speech emotion recognition has been specified.

CONCLUSION

In this paper, we used a combining method of progressive neural networks and clustering fuzzy output signals based on firefly algorithm in order to recognizing speech emotion.

In this method, the input signals are given to a fuzzy system and these signals are clustered to better recognise the emotional speech signals. The clustering process is performed based on firefly algorithm. Then, we do taking-noise the clustering signal with generated noisy class in order to recognise the final signal of speech emotion regardless of noise. This combining method has advantages such as:
High speed in processing because of reduction in dimension of studied signals in 5 layers fuzzy circuit, high accuracy of speech emotion recognition due to the use of fuzzy circuit output signals clustering method that the stability of this system in front of natural noise presented in the environment is due to noise exposure in a separate class which is easily recognisable.
Table 1: voice recognition and comparing FAFC-FNN, PSO-FNN, and BP-FNN algorithm with each other

EXAMPLES FOLLOW:

.

Vocabulary	Training Algorithm	SNR(dB)				clean	Maximum Difference	Average Recognition Rate
		15	20	25	30			
10	**FAFC-FNN**	**97**	**96**	**97**	**97**	**97.9**	**26**	**96.98**
	PSO-FNN	94.8	95.7	96.7	96.2	96.7	1.9	96.0
	BP-FNN	95.2	94.8	94.8	94.8	95.2	0.4	95.0
20	**FAFC-FNN**	**96.2**	**96.7**	**97.1**	**96.4**	**97.6**	**1.4**	**96.8**
	PSO-FNN	96.4	95.5	96.4	95.7	97.6	2.1	96.3
	BP-FNN	94.5	95.2	95.5	96.4	96.7	2.2	95.7
30	**FAFC-FNN**	**96.1**	**96.4**	**97.3**	**97.6**	**97.9**	**1.8**	**97.1**
	PSO-FNN	95.6	96.3	97.3	97.5	97.6	1.9	96.9
	BP-FNN	93.2	94.3	96.1	94.8	96.8	3.6	95.0
40	**FAFC-FNN**	**96.1**	**97.1**	**97.7**	**97.9**	**98.3**	**2.2**	**97.3**
	PSO-FNN	95.5	96.8	97.6	97.5	97.9	2.4	97.0
	BP-FNN	92.6	93.3	95.5	95.9	95.6	3.0	94.6
50	**FAFC-FNN**	**95.8**	**97.1**	**96.8**	**97.3**	**97.8**	**2.0**	**97.0**
	PSO-FNN	95	96.7	96.6	97.3	97.8	2.8	96.7
	BP-FNN	92.2	93.3	94.9	94.2	95.2	3.0	94.0

Volume: 07 Issue: 04 Pages: 2817-2822 (2016) ISSN: 0975-0290

REFERENCES

[1] Gaoyun Li, “The Study of the Speech Recognition Based on theOptimized RBF Neural Network”, Master Degree Thesis of TaiyuanUniversity of Technology, Taiyuan, 2009.

[2] Xue Ying Zhang, Peng Wang, “Improved T-S Fuzzy NeuralNetwork in Application of Speech Recognition System”, ComputerEngineering and Applications, vol. 45, April 2009, pp 246-248.

[3] Yang Wu, Wu Yuan, Wu Zhongru, “Forecast Model Study Based on Fuzzy Neural Network” ,Water Resources and Power, vol. 22, pp 63-65, Mar, 2004.

[4] M. SenthilArumugam, M. V. C. Rao, Alan W.C.Tan, “A novel andeffective particle swarm optimization like algorithm withextrapolation 2009.

[5] Otoacoustic emissions (OAEs) Portal [Online]2006 [2006 Nov]; Availble fromURL:http://www.oae.it.

[6] Xiucheng Dong, Yunyuan Zhao, Yunyun Xu, Zhang Zhang and Peng Shi. Design of pso fuzzy neural network control for ball and plate system. International Journal of InnovativeComputing, Information and Control, 2011, pp. 7091-7103 .

[7] Jing-Ru Zhang , Jun Zhang , Tat-Ming Lok , Michael R. Lyu . A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training. Elsevier , Applied Mathematics and Computation 185 (2007) 1026–1037.

[8] Y. Marinakis, M. Marinaki, M. Doumpos, N. Matsatsinis, C. Zopounidis, Ahybrid stochastic genetic-GRASP algorithm for clustering analysis, OperationalResearch An International Journal 8 (1) (2008) 33–46.

[9] D. J. Livingstone Artificial Neural Networks: Methods and Applications, 2011 :Humana Press

[10] C.-F. Juang , T.-C. Chen and W.-Y. Cheng "Speedup of implementing fuzzy neural networks with high-dimensional inputs through parallel processing on graphic processing units", IEEE Trans. Fuzzy Syst., vol. 19, no. 4, pp.717 -728 2011

[11] C.-J. Lin , C.-F. Wu and C.-Y. Lee "Design of a recurrent functional neural fuzzy network using modified differential evolution", Int. J. Innovative Comput., Inform. Control, vol. 7, no. 2, pp.669 -684 2011

[12] G. Karypis. (2010) Karypis Lab, http://glaros.dtc.umn.edu/gkhome/cluto/ cluto/download

[13] Calkins K, Devaskar SU. Fetal origins of adult disease. Curr Probl Pediatr Adolesc Health Care 2011; 41:158-176.

[14] Plata, J.L., Rubio, F.M., " Basic theory of the Magnetic Resonance Sounding method"Boletin Geologicoy minero, July –September 2007 , vol