EEG Signal Sorting & ANN Principal Features Analysis for Brain Disease Diagnosis

Sangita V. Darandale, M.B. Tadwalkar

Abstract*— Automatic support system for EEG signal classification for brain diseases diagnosis is proposed in this study. The artificial neural network is used to diagnosis disease like epilepsy by classifying the EEG Signal. The manual analysis of the signal require more time. Also there is a requirement of intensive trained person, to minimize the diagnostics errors. Back Propagation Network with data processing techniques was employed. Decision is based on two stages: feature extraction using Principal Component Analysis and the classification using Back Propagation Network (BPN).The training performance as well as classification accuracy is evaluated for Back Propagation classifier performance. Back Propagation Network classifier is used for high speed and accuracy.

Index Terms*— EEG signals, Classification algorithms, Back propagation network, epilepsy disease.

I. INTRODUCTION

Automated classification and detection of different medical signals is motivated by the necessity of high accuracy when dealing with a human life. Also, the computer assistance is required in medical institutions to improve the disease diagnosis results of human. There is a need of good software technique to minimize the double reading cost. Conventional methods of monitoring and diagnosing the diseases rely on detecting the presence of particular features by a human observer. Due to huge number of patients in intensive care units, continuous monitoring is required by using latest techniques.

Brain–Computer Interfaces (BCI) is the best feasible way of providing the communication between the human and the system by means of brain signals. The signal classification module is composed of the obtained EEG signal features extraction and the transformation of these signals into device instructions. The predicted EEG drives the classification to some precise feature extraction methods.

Electroencephalography is a medical imaging technique that reads scalp electrical activity generated by brain structures. The electroencephalogram (EEG) is defined as electrical activity of an alternating type recorded from the scalp surface after being picked up by metal electrodes and conductive media.

The EEG measured directly from the cortical surface is called electrocardiogram while when using depth probes it is called electro gram. In this article, we will refer only to EEG measured from the head surface. Thus electroencephalographic reading is a completely non-invasive procedure that can be applied repeatedly to patients, normal adults, and children with virtually no risk or limitation.

When brain cells (neurons) are activated, local current flows are produced. EEG measures mostly the currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. Differences of electrical potentials are caused by summed postsynaptic graded potentials from pyramidal cells that create electrical dipoles between soma (body of neuron) and apical dendrites (neural branches). Brain electrical current consists mostly of Na+, K+, Ca++, and Cl ions that are pumped through channels in neuron membranes in the direction governed by membrane potential. The detailed microscopic picture is more sophisticated, including different types of synapses involving variety of neurotransmitters. Only large populations of active neurons can generate electrical activity recordable on the head surface. Between electrode and neuronal layers current penetrates through skin, skull and several other layers. Weak electrical signals detected by the scalp electrodes are massively amplified, and then displayed on paper or stored to computer memory. Due to capability to reflect both the normal and abnormal electrical activity of the brain, EEG has been found to be a very powerful tool in the field of neurology and clinical neurophysiology.

II. PRINCIPLE COMPONENT ANALYSIS

PCA is known a Principle Component Analysis – this is a statistical analytical tool that is used to explore, sort and group data. What PCA does is take a large number of correlated (interrelated) variables and transform this data into a smaller number of uncorrelated variables (principal components) while retaining maximal amount of variation, thus making it easier to operate the data and make predictions. Or as Smith (2002) puts it “PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data.”

According to Jolliffe (2002) it is generally accepted that PCA was first described by Karl Pearson in 1901. In his article ” On lines and planes of closest fit to systems of points in space,” Pearson (1901) discusses the graphical representation of data and lines that best represent the with the maximum axis of the correlation ellipsoid”. He also states that the analysis used in his

Manuscript Received on November 2014.

Sangita V. Darandale, Department of Electronics and Telecomm Engineering, JSPMs, Jayawantrao Sawant College of Engg, Hadapsar, Pune 28, India.

Prof. M.B.Tadwalkar, Department of Electronics and Telecomm Engineering, JSPMs,Jayawantrao Sawant College of Engg, Hadapsar, Pune 28, India.

Paper can be applied to multiple variables. However, PCA was not widely used until the development of computers. It is not really feasible to do PCA by hand when number of variables is greater than four, but it is exactly for larger amount of variables that PCA is really useful, so the full potential of PCA could not be used until after the spreading of computers (Jolliffe, 2002).

According to Jolliffe (2002) significant contributions to the development of PCA were made by Hotelling (1933) and Girshick (1936; 1939) before the expansion in the interest towards PCA. In1960s as the interest in PCA rose, important contributors were Anderson (1963) with a theoretical discussion, Rao (1964) with numerous new ideas concerning uses, interpretations and extensions of PCA, Gower (1966) with discussion about links between PCA and other statistical techniques and Jeffers (1967) with a practical application in two case studies. The following is a standard derivation of principal components presented by Jolliffe (2002).To derive the form of the PCs, consider first $\alpha'1x$; the vector a1 maximizes. It is clear that, as it stands, the maximum will not be achieved for finite $\alpha1$ so a normalization constraint must be imposed. The constraint used in the derivation is$\alpha1\alpha1 = 1$, that is, the sum of squares of elements of $\alpha1$ equals 1. Other constraints may be more useful in other circumstances, and can easily be substituted later on. However, the use of constraints other than $\alpha'1\alpha1$ = constant in the derivation leads to a more difficult optimization problem, and it will produce a set of derived variables different from the principal components.

To maximize subject to $\alpha1\alpha1 = 1$, the standard approach is to use the technique of Lagrange multipliers. Maximize where λ is a Lagrange multiplier. Differentiation with respect to $\alpha1$ gives or Where Ip is the (p x p) identity matrix. Thus, ë is an Eigen value of and á1 is the corresponding eigenvector. To decide which of the p eigenvectors gives á'1x with maximum variance, note that the quantity to be maximized is so ë must be as large as possible. Thus, á1 is the eigenvector corresponding to the largest eigen value of, and the largest eigen value. In general, the kth PC of x is a' kx and, where ëk is the kth largest eigen value of, and ák is the corresponding eigenvector.

Shlens (2009) derives an algebraic solution to PCA based on an important property of number of measurement types and n is the number of samples. The goal is summarized as follows:

Find some orthonormal matrix P in Y = PX such that is a diagonal matrix. The rows of P are the principal components of X. He begins by rewriting CY in terms of the unknown variable. Note that they have identified the covariance matrix of X in the last line. The plan is to recognize that any symmetric matrix A is diagonalzed by an orthogonal matrix of its eigenvectors. For a symmetric matrix A => A=EDET, where D is a diagonal matrix and E is a matrix of eigenvectors of A arranged as columns. Now comes the trick. They select the matrix P to be a matrix where each row pi is an eigenvector of. By this selection, with this relation and A (P^{-1} = PT) we can finish evaluating CY. It is evident that the choice of P diagonalizes CY. This was the goal for PCA. We can summarize the results of PCA in the matrices P and CY.

The principal components of X are the eigenvectors of the ith diagonal value of CY is the variance of X along pi. In practice computing PCA of a data set X entails (1) subtracting off the mean of each measurement type and computing the eigenvectors of CX.

A) Artificial Neural Network (ANN):

A neural net is an artificial illustration of the human brain that tries to imitate its learning process. ANN is an interrelated group of artificial neurons that uses a mathematical model or computational model for information processing. ANN is a network of simple processing elements which can demonstrate complex overall performance, determined by the connections between the processing elements and element parameters. ANN is an adaptive system that changes its structure based on external or internal information that flows through the network. ANN computing approach to information processing primarily involves a learning process with an ANN architecture that adaptively responds to inputs according to a learning rule. After the NN has learned, the trained network can be used to execute certain tasks depending on the exact purpose. The talent to learn by example and simplify are the principal characteristics of ANN. Classification of signals is done by using this ANN to obtain the correct classification percentage. ANN is learned using the back propagation algorithm in which the errors for the units of the hidden layer are determined by back propagating the errors of the units of the output layer. It is a systematic method of training multi-layer ANNs. It contains an input layer, at least one intermediate/hidden layer and an output layer in its network. Some of the ANN learning parameters are Threshold, Goal, Epoch, Sigmoid function, Training type and Number of Hidden layers.

The word network in the term 'artificial neural network' refers to the inter–connections between the neurons in the different layers of each system. An example system has three layers. The first layer has input neurons which send data via synapses to the second layer of neurons, and then via more synapses to the third layer of output neurons.

More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons. The synapses store parameters called "weights" that manipulate the data in the calculations. Fig.1. shows the basic diagram of Artificial Neural Network

B) Back Propagation Network:

The back propagation algorithm is used to compute the necessary corrections, after choosing the weights of the network randomly. The algorithm can be decomposed in the following four steps.

a) ***Feed-forward computation***

b) ***Back propagation to the output layer***

c) ***Back propagation to the hidden layer***

d) ***Weight updates***

The algorithm is stopped when the value of the error function has become sufficiently small.

The fig.2 shows the notation for three layered network. The connection we're interested in is between neuron A (a hidden layer neuron) and neuron B (an output neuron) and has the weight WAB. The diagram also shows another connection, between neuron A and C, but we'll return to that later. The algorithm works like this:

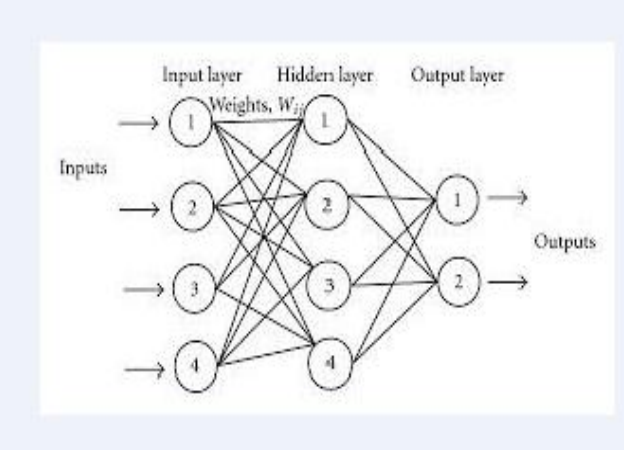


Figure 1: Basic Diagram of Artificial Neural Network

1) *First apply the inputs to the network and work out the output:–* remember this Initial output could be anything, as the initial weights were random numbers.

2) *Next work out the error for neuron B. The error is what you want:* What you actually get, in other words*:*

Error B = Output B (1-OutputB) (Target B – Output B)

The "Output (1-Output)" term is necessary in the equation because of the Sigmoid Function – if we were only using a threshold neuron it would just be (Target –Output).

3) *Change the weight. Let W+AB be the new (trained) weight and WAB be the initial weight.*

W+AB = WAB + (Error B x Output A)

Notice that it is the output of the connecting neuron (neuron A) we use (not B). We update all the weights in the output layer in this way.

4) *Calculate the Errors for the hidden layer neurons. Unlike the output layer we can't*

Calculate these directly (because we don't have a Target), so we Back Propagate them from the output layer (hence the name of the algorithm). This is done by taking the Errors from the output neurons and running them back through the weights to get the hidden layer errors. For example if neuron A is connected as shown to B and C then we take the errors from B and C to generate an error for A.

Error A = Output A (1 - Output A) (Error B WAB + Error C WAC)

Again, the factor "Output (1 - Output)" is present because of the sigmoid squashing function.

5) *Having obtained the Error for the hidden layer neurons now proceed as in stage 3*

To change the hidden layer weights. By repeating this method we can train a network of any number of layers.

The back propagation algorithm is used to compute the necessary corrections, after choosing the weights of the network randomly. The algorithm can be decomposed in the following four steps: Feed-forward computation, Back propagation to the output layer, Back propagation to the hidden layer and Weight updates. Fuzzy logic is a form of many valued logic, it deals with reasoning that is approximate rather than fixed and exact.

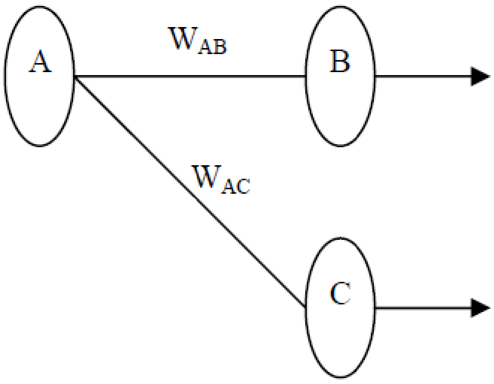


Figure 2: Single Connection learning in Back Propagation Neural Network

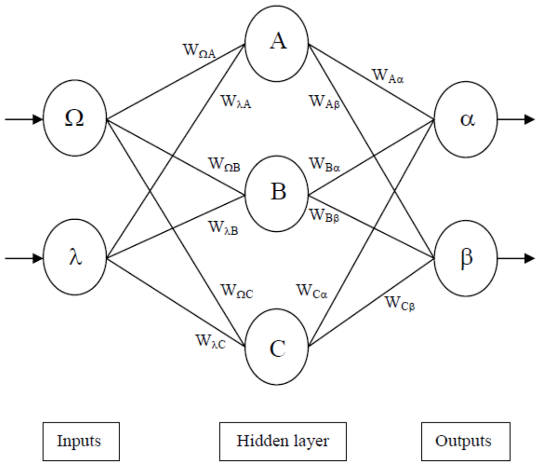


Figure 3: Figure for Back Propagation Network

This may well have left some doubt in your mind about the operation, so let's clear that up by explicitly showing all the calculations for a full sized network with 2 inputs, 3 hidden layer neurons and 2 output neurons as shown in figure 3. W + represents the new, recalculated, weight, whereas W (without the superscript) represents the old weight.

III. PROPOSED WORK

A. Block Diagram:

The proposed block diagram is shown in fig 4.which include the 2steps.

1) *Feature Extraction(PCA)*

The signals obtained from the electrodes are given to Principal Component Analysis for dimensionality reduction to remove the redundant variables in the data and the classified using Neural Network classifier with back propagation. In PCA the better classification of signals is obtained for the learning parameters like epochs as 1000, number of hidden layers as 3, goal as 0.01, and sigmoidal function as tensing, threshold das 0.5 and training type.

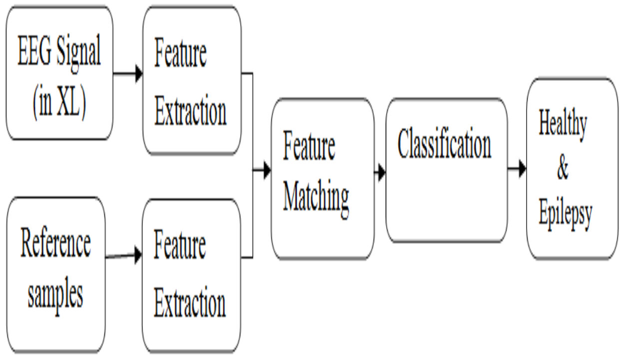


Figure 4: Proposed system Block diagram

The brain signals are trained using Neural Network and the training is shown in Fig.4. During the classification of the mental tasks using Neural Network classifier, the data is misclassified at the output i.e., the percentage of correct classification is low Similarly during the classification of the mental tasks using Principal Component Analysis with Neural Network classifier, the data is perfectly classified at the output i.e., the percentage of correct classification is good because of the reduction of the redundant variables in the dataset. The comparison of the results of Neural Network classifier and Principal Component Analysis with Neural Network classifier is tabulated in Table 1 to show the variation of mean square error during training, mean square error during testing, computation time and the percentage of correctly classified data for both type of classification.

2) *Fuzzy Logic(FL)*

Fuzzy logic is a form of many valued logic, it deals with reasoning that is approximate rather than fixed and exact. Compared to traditional binary sets (where variables may take on true or false values) fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. Furthermore, when linguistic variables are used, these degrees may be managed by specific functions. Irrationality can be described in terms of what is known as the fuzzjective. The term "fuzzy logic" was introduced with the 1965 proposal of fuzzy set theory by Lotfi A. Zadeh. Fuzzy logic has been applied to many fields, from control theory to artificial intelligence. Fuzzy logics however had been studied since the 1920s as infinite-valued logics notably by Łukasiewicz and Tarski.

3) *Classification(BPN)*

Here the classification is done by back propagation method.

The back propagation algorithm is used to compute the necessary corrections, after choosing the weights of the network randomly. The algorithm can be decomposed in the following four steps: Feed-forward computation, Back propagation to the output layer, Back propagation to the hidden layer and Weight updates.

The algorithm is stopped when the value of the error function has become sufficiently small.

B. Experimental Results:

1) *Mean of entire db:*

We can calculate about a data set and can calculate the mean of the sample. The mean of a sample is given by the formula:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (1)$$

Notice the symbol $\bar{X}$ (said "X bar") to indicate the mean of the set. All this formula says is "Add up all the numbers and then divide by how many there are".

The performance of classifier can be evaluated through following parameters, Sensitivity: It measures the proportion of actual positives which are correctly identified

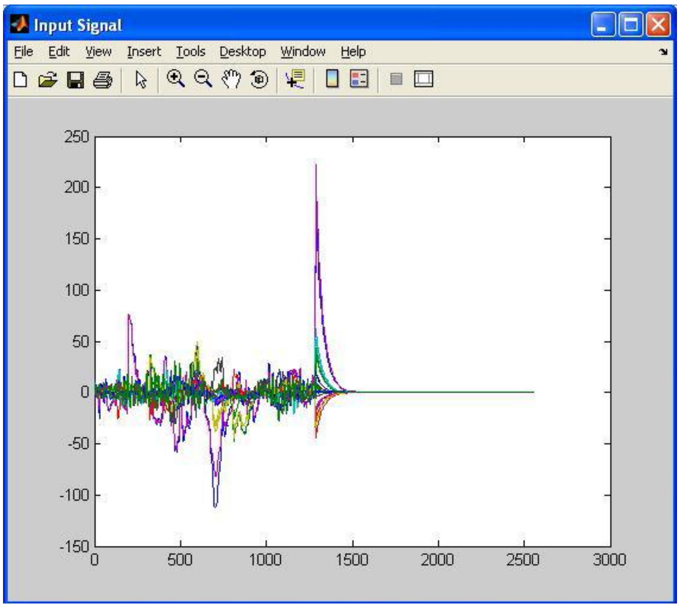
Sensitivity = Tp /(Tp + Fn)

Where, Tp = True Positive: Abnormality correctly classified as abnormal Fn = False negative: Abnormality incorrectly classified as normal Specificity: It measures the proportion of negatives which are correctly identified.

Specificity = Tn / (Fp + Tn)

Where, Fp = False Positive: Normal incorrectly classified as Abnormal Tn = True negative: Normal correctly classified as normal Total accuracy: (Tp+ Tn)/ (Tp +Tn +Fp +Fn).

Fig.5 shows the EEG waveform of normal patient, Fig.6 shows the EEG waveform of Epileptic patient, Fig.7 shows EEG of Patient during seizures and Fig.8 shows the performance plot.



Figuer. 5: EEG of Normal Patient

The EEG measured directly from the cortical surface is called electrocardiogram while when using depth probes it is called electro gram. In this article, we will refer only to EEG measured from the head surface. Thus electroencephalographic reading is a completely non-invasive procedure that can be applied repeatedly to patients, normal adults, and children with virtually no risk or limitation. To maximize subject to α1α1 = 1, the standard approach is to use the technique of Lagrange multipliers. Find some orthonormal matrix P in Y = PX such that is a diagonal matrix

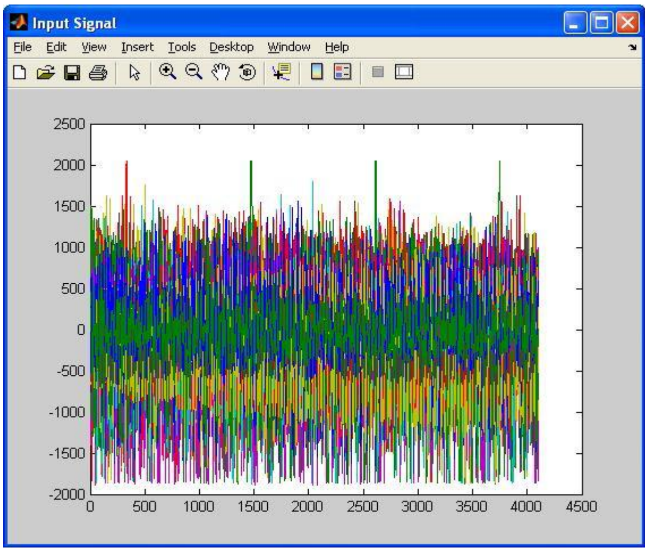


Figure 6: EEG of Epileptic Patient

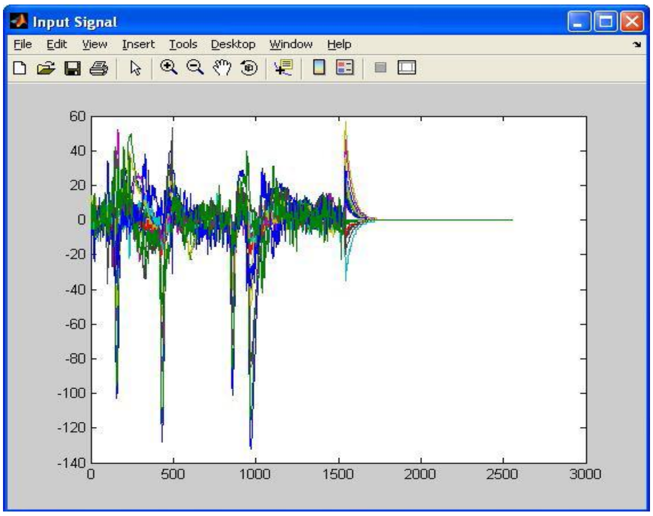


Figure 7: EEG of Patient during seizures

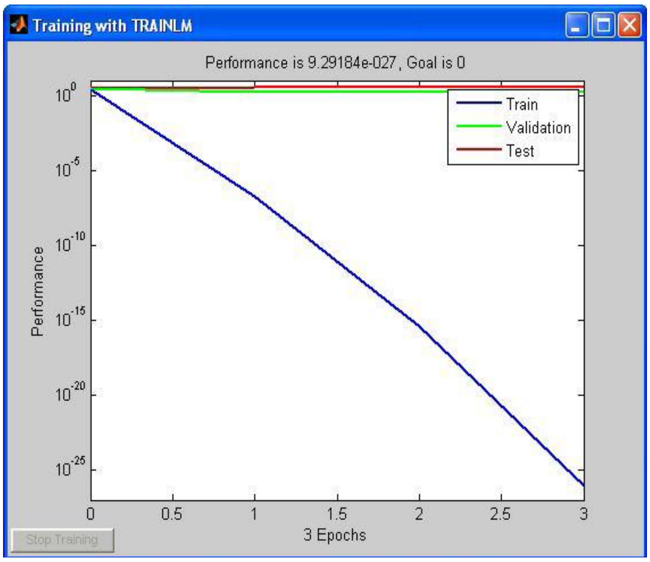


Figure 8: Performance Plot

IV. CONCLUSION

The BCI problem is solved using neural network. The mental task classification is enhanced by several kinds of pre-processing, to generate the input data of the neural network. The soft computing techniques are employed for the classification of the EEG signals as the techniques are intended to model and make possible solutions to real world tribulations. The probability of correct classification has been increased by using soft computing techniques like Principal Component Analysis with neural network and Fuzzy Logic.

REFERENCES

1. K. Sivasankari† and K. Thanushkodi "An Improved EEG Signal Classification Using Neural Network with the Consequence of ICA and STFT"VOL-9,1060-1071,JEET 2014
2. Neelam Rout "Analysis and Classification Technique Based On ANN for EEG Signals" International Journal of Computer Science and Information Technologies, Vol.5, , 2014
3. Kottaimalai R, EEG "Signal Classification using Principal Component Analysis with Neural Network in Brain Computer Interface Applications " 2013 IEEE International Conference on EmergingTrends in Computing, Communication and Nanotechnology (ICECCN 2013)
4. Sharan reddy, P.K. Kulkarni, "EEG signal classification for Epilepsy Seizure Detection using Improved approximate Entropy" International Journal of Public Health Science (IJPHS) Vol. 2, No. 1, March 2013.
5. Baha Sen,Musa Peker, "Novel approaches for automated epileptic diagnosis using FCBF selection and classification algorithms", Turkish journal of Electrical Engineering and Computer Science,2013
6. Zarita Zainuddin1, Lai Kee Huong1, Ong Pauline1,2 "Reliable epileptic seizure detection using an improved wavelet neural network " Australasian Medical Journal [AMJ 2013]
7. Satyanarayana Vollala & Karnakar Gulla]"Automatic detection of epilepsy EEG using Neural Networks" International Journal of Internet Computing ISSN No: 2231 – 6965, Vol.1, ISS- 3 2012
8. Kavita Mahajan, M. R. Vargantwar, Sangita M. Rajput "Classification of EEG using PCA, ICA and Neural Network " International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-1, October 2011
9. S. Aydn," Determination of auto agressive model orders for seizure detection."Turkish Journal of Electrical Engineering & Computer Science,Vol.18,pp.23-30,2010
10. T. Fathima, M. Bedeeuzzaman, O. Farooq, U.K. Yusuf, "Wavelet based features for epileptic seizure detection", MES Journal of Technology and Management, pp. 108-112, 2010
11. Forrest Sheng Bao , Donald Yu-Chun Lie, and Yuanlin Zhang " A New Approach to Automated Epileptic Diagnosis Using EEG and Probabilistic Neural Network"CSAI,2008
12. N. Kannathal, M. Choo, U. Acharya, P. Sadasivan, \Entropies for detection of epilepsy in EEG", Computer Methods and Programs in Biomedicine, Vol. 80, pp. 187,94, 2005.
13. V. Srinivasan, C. Eswaran, N. Sriraam, Artical neural network based epileptic detection using time domain and frequency domain features", Journal of Medical systems, Vol. 29, pp. 647/660, 2005.
14. A. Subaşı, A. Alkan, E. Köklükaya, "Wavelet neural network classification of EEG signals", Teknoloji, Vol. 7, pp. 71-80, 2004 (in Turkish).
15. M. Akın, M.A. Arserim, M.K. Kıymık, İ. Türkoğlu, "A new approach for diagnosing epilepsy by using wavelet transform and neural networks", Proceedings of the 23rd Annual EMBS International Conference, İstanbul, pp. 1596-1599, 2001