# EEG SIGNAL CLASSIFICATION USING PRINCIPAL FEATURE ANALYSIS AND ARTIFICIAL NEURAL NETWORK FOR BRAIN DISEASE DIAGNOSIS

**Abinaya A[1], Bhavya T[2],Evangeline T[3],Monica M[4]**

*[1,2,3,4] Final year, Department of Information Technology,*

*Panimalar Engineering College, Chennai, (India)*

**ABSTRACT**

*The paper proposes an automated system for classification of brain signals using EEG signals by adopting concepts such as artificial neural networks and principle feature analysis or principle component analysis(PCA). The artificial neural network, an artificial illustration of the human brain that tries to imitate its learning process that will be used to classify the brain EEG signal that is tumor case or epilepsy case or normal. The manual analysis of the signal is time consuming, inaccurate and requires highly qualified professionals to avoid diagnostic errors. The detection of the brain tumor is a challenging problem, due to the nature of the tumor cells. PCA is a way of identifying patterns in input data, and expressing the data in such a way as to highlight their similarities and differences. The back propagation network is used finally for classifying the pattern of tumor and normal EEG. The probability of correct classification has been increased by using soft computing techniques like Principal Component Analysis with neural network.*

***Keywords- Artificial Neural Network, Back Propagation Network, EEG Signals, Epilepsy, Principle Component Analysis / Principle Feature Analysis.***

## I. INTRODUCTION

This system of automated EEG classification of various signals from EEG sensor is always needed in medical environment to get accurate results on the brain activity conditions. Since EEG brain disease diagnosis system has A major impact on valuable human life, it is very much necessary to maintain the system accurately and there must be no possibility for risking human life. Existing method of brain disease diagnosis uses a EEG monitor to receive and record the signal in which the doctor or a lab technician detects and collects the features of brain signal and finally decides on the result of whether the person has a normal or epileptic brain signal .This technique is practically impossible to do because of the large number of patients ,lack of time and in emergency situations where the professionals are pressurized to provide results as soon as possible .Brain–computer interfaces (BCI) is a direct communication pathway between human and brain.bci is often used for assisting, augmenting, directing humans about the sensory motor functions of brain. BCI is a part of the neuroprosthetics which connects the neural system to a device. The electroencephalogram (EEG) is defined as the electrical activity, recorded from the scalp surface after being picked up by metal electrodes and conductive media. It measures the voltage fluctuations resulting from ionic current flows within neurons of the brain .The EEG is a non-invasive technique and practically does not harm the patient under test. When brain cells (neurons) are activated, local current flows are produced in brain. EEG measures mostly the currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex of the brain. Differences of electrical potentials are caused by summed postsynaptic graded potentials from pyramidal cells that create electrical dipoles between body of neuron(soma) and neural branches (apical dendrites).The electrical current of brain consists mostly of Na+, Ca++, K+ and Cl- ions which are pumped through channels in neuron membranes in the direction shown by membrane potential. The microscopic picture is more sophisticated, including various types of synapses involving variety of neurotransmitters. The large populations of active neurons can generate electrical activity recordable on the head surface. There lies a layer between electrode and neuronal layers in which current penetrates through skin, skull and several other layers. Only weak electrical signals detected by the scalp electrodes are massively amplified, and then displayed on paper or stored in PC. Due to capability to reflect both the normal and abnormal electrical activity of the brain, it has been found to be a very powerful tool in the field of diagnosis in neurology and clinical neurophysiology.

## II. RELATED WORKS

Principal component analysis (PCA), first described by Karl Pearson in 1901, is a statistical and analytical procedure that uses an orthogonal transformation to convert a large set of observations includes possibly correlated(inter-related) variables into a set of values of linearly uncorrelated variables (principal components).The number of principal components is less than or equal to the number of original variables. In this transformation, the first principal component has the largest possible variance ie., maximal amount of variability in the input data, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. The principal components are usually orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric. Retaining maximal amount of variation makes it easier to operate the data and make predictions. Depending on the area of application, it is also named the discrete Karhunen–Loève transform (KLT) in signal processing, the Hotelling transform in multivariate quality control, proper orthogonal decomposition (POD) in mechanical engineering, singular value decomposition(SVD) of X (Golub and Van Loan, 1983), eigen value decomposition (EVD) of $X^TX$ in linear algebra, factor analysis , and empirical modal analysis in structural dynamics. A BCI based on MI translates the subject's motor intention into a control signal through classifying EEG. Multichannel EEG are usually needed for spatial pattern identification and therefore MI based BCI is still in form of lab demonstration. Or as Yijun Wang and Shangkai Gao (2005)[1] put it, "A classification algorithm can be developed by combining LDA and RP . They are independent in frequency domain and can lead to significant performance gain."According to Aleˇs Proch´azka, Jarom´ır Kukal and Oldˇrich Vyˇsata (2008)[2] in their article, they denote that " DWT can be used for signal processing and feature extraction as an alternative to the DFT. This signal classification assumed the knowledge of the range of the number of classes to apply a self-creating classification method to find their optimal value and to exclude the possibility of dead neurons."L.M.Patnaik and Ohil K. Manyam (2008)[3] specify that "Automated detection process can be developed by using WT for feature extraction and ANN for classification. This provides many advantages among them are low cost, non-stop monitoring and faster diagnosis." A new EEG classification was developed by Min Han and Leilei Sun(2010)[4] . They put it as "A new EEG classification can be based on RVM and AR model. Good performance of RVM employing AR coefficients increase interest in the prediction of epilepsy."EEG signal classification as stated by Abdul hamit Subasi and M.Ismail Gursoy(2010)[5], is by using PCA , ICA, LDA and SVM. It acts as a promising tool for intelligent diagnosis. According to Indu Sekhar Samant, Guru Kalyan Kanungo, Santosh Kumar Mishra(2012)[7] in their paper propose "To use Least Mean Square algorithm to remove the artifact in the EEG signal. This proposed method can facilitate the doctor to detect the breast cancer in the early stage of diagnosis as well as classify the total cancer affected area." Epileptic EEG and its classification as given by Nilima Mohite , Rajveer Shastri , Arnab Das(2014)[8], is "To extract features by combining DWT , EMD & bispectrum analysis. Results of all four methods shows that the EMD has better variance and thus the potential of classifying normal and Seizure EEGs is more using EMD than other methods."

### 2.1 Artificial Neural Network (ANN)

Artificial neural network (ANN) is a machine learning approach that models human brain and consists of a number of artificial neurons It is a computational model for information processing. Neurons in ANNs tend to have fewer connections than biological neurons, each receiving a number of inputs.

An activation function is applied to these inputs which results in activation level of neurons, which is nothing but the output value of the neuron. Knowledge about the brain's learning task is given in the form of examples called training examples.

A Neural Network is clearly specified by:

(i) Neuron model: the information processing unit of the NN, Fig 1.2 gives the artificial neuron model.

(ii) A architecture: a set of neurons and links connecting all neurons. Each link has a specified weight.

A learning algorithm is mainly used for training the NN by modifying the weights in order to model a particular learning task correctly on the training examples. The aim is to obtain a NN that is trained and generalizes well. It should behave correctly on new instances of the learning task .The neuron of brain is the basic information processing unit of a NN.

It consists of:

(i) A set of links, with weight, describes the neuron inputs, with weights of $W_1, W_2, \ldots, W_m$

(ii) An adder function (linear combiner) for computing the weighted sum of the inputs:

in real numbers $u = \sum_{j=1}^{m} w_j x_j$

(iii) Activation function $\varphi$ for limiting the amplitude of the neuron output. Here 'b' denotes bias.

$$y = \varphi(u + b)$$

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. Most systems have three layers. The first layer of input neurons used to send data through synapses to the second layer of neurons, and then send through more synapses to the third layer of output neurons.

More complex systems have more layers of neurons with some having increased layers of input neurons and output neurons. The synapses store parameters called "weights" for manipulating data in the calculations.

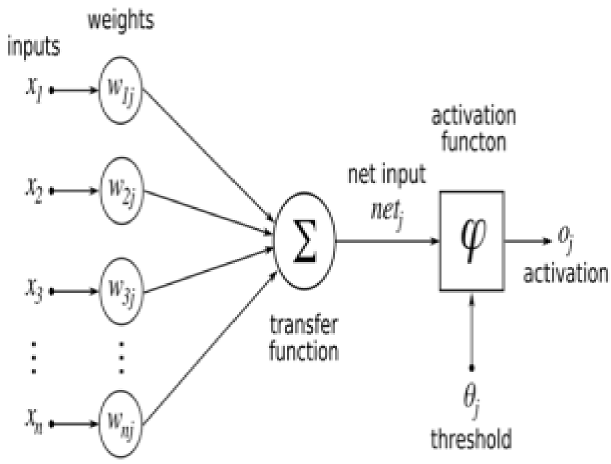
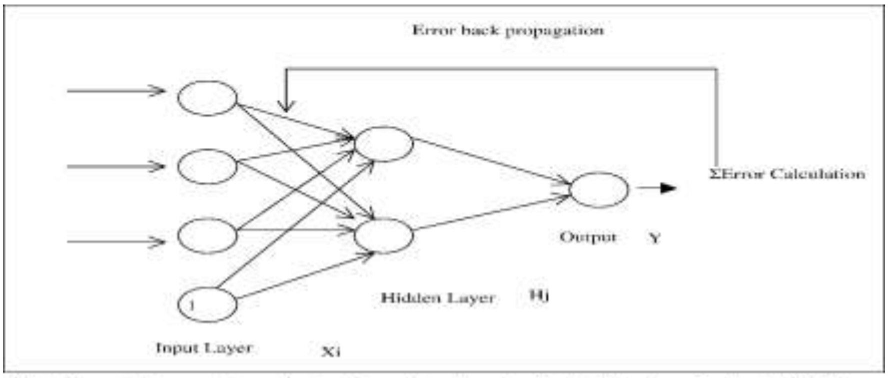



**Fig.1.1 shows the basic diagram of Artificial Neural Network**

ANN is a neural network of simple processing elements to demonstrate complex overall performance of brain, based on the connections between the processing elements and their parameters.ANN is an adaptive system that changes its structure based on external or internal information when flowing through the network. Neural computing approach to information processing mainly involves a learning process with an ANN architecture that adaptively responds to inputs according to a learning rule of the network. After the NN has learned, the trained network can be used to execute certain tasks depending on the exact purpose for the variation . The ability to learn with the help of example and simplify are the main characteristics of ANN. Classification of signals is performed using ANN to obtain the correct classification percentage. ANN is learned using the back propagation algorithm where 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 way of training multi-layer ANNs. It contains an input layer, at least one intermediate or hidden layer and an output layer in its network. Some ANN learning parameters are Threshold, Goal, Epoch, Sigmoid function, Training type and Number of Hidden layers.

## 2.2 Back Propagation Network

The Back propagation algorithm mainly searches for weight values of the data that minimize the total error of the network over the large set of training examples (training set).




**Notes:** The weight connecting node *i* in the input layer to node *j* in the hidden layer is denoted by *Wji*, and the weight connecting node *j* to the output node is represented by *Vj*

**Fig: 1.2 Back Propagation Training Algorithm**

Back propagation network consists of the repeated application of the following two passes:

(i) Forward pass: In this pass, the network is activated on one example and the error of (each neuron of) the output layer is computed.

(ii) Backward pass: In this pass, the network error is used for changing by updating the weights. The error is propagated backwards from the output layer through the network layer after layer.

This is done by repeatedly computing the local gradient of all neuron individually

Back propagation adjusts the weights of the Neural Network in order to minimize the network total mean squared error.

Consider a network of three layers. We use i to represent nodes in input layer, j to represent nodes in hidden layer and k represent nodes in output layer.$w_{ij}$ refers to weight of individual connection between a node in input layer and node in hidden layer.

The following equation is used for deriving the output value Yj of node j

$$Yj = \frac{1}{1+e^{-X_j}}$$

where, $X_j = \sum x_i \, . \, w_{ij} - \theta_j$ , $1 \leq i \leq n$; n is the number of inputs to node j, and $\theta_j$ is threshold for node j.

The back propagation algorithm is used to compute the necessary corrections, after choosing the weights of the network randomly.

The algorithm can be decomposed into the following four steps.

- Feed-forward computation
- Back propagation - output layer
- Back propagation - hidden layer
- Weight updates

The algorithm continues until the value of the error function has become sufficiently small. The fig.2 shows the notation for three layered network. Consider the connection between neuron A (a hidden layer neuron) and neuron B (an output neuron) and has the weight of WAB. The diagram specifies another connection, between neuron A and C.

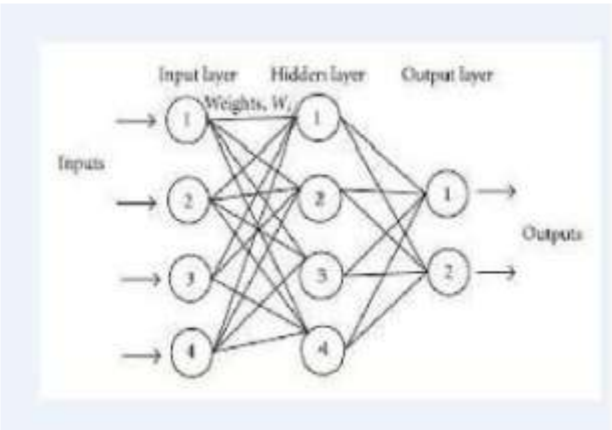



**Fig 2: Basic Diagram of Artificial NN**

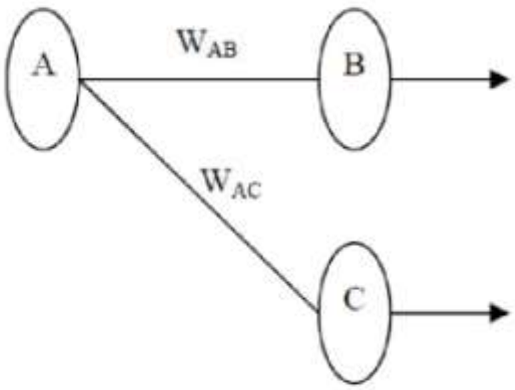



**Fig 3: Single Connection learning in BPN**

The algorithm works like this:

1) *First apply the inputs to the network and find the output:*

Note that this Initial output could be anything, since the initial weights were random numbers.

2) *Next find the error for individual neuron B. The error is what you want and What you actually get, in formula:*

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 each of the hidden layer neurons. But in output layer we can't calculate these directly (because we don't have a specified Target), so we Back Propagate them from the output layer (hence it is named so).*

This is done by taking the Errors from the output neurons and running them back through the weights to get error in the hidden layer. 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) *After obtaining the Error for the hidden layer neurons now proceed as in stage 3*

To change the hidden layer weights of nodes. By repeating this method we can train a network with 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.

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

### 3.1 Block Diagram

The block diagram of the entire project is shown in fig 4.1. It is composed of four major techniques

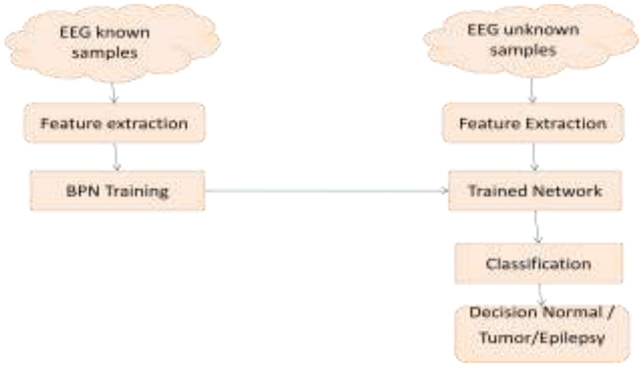



**Fig 3.1: Proposed system Block diagram**

### 3.1.1 Input acquisition

The analysis of brain signals plays an important role in classification and diagnosis of different brain diseases. MATLAB provides a user friendly and interactive graphic user interface (GUI) allowing users to flexibly and interactively process their high-density EEG dataset. EEGLAB is a MATLAB toolbox distributed under the free GNU GPL license for processing data from electroencephalography (EEG) and other electrophysiological signals. EEGLAB also implements independent component analysis (ICA) and several techniques. EEGLAB allows users to access their electrophysiological data by importing it into about 20 binary file formats, processing the data, performing single trials, and then ICA. Repeated ICA components may be subtracted from the data. Sometimes, ICA components representing brain activity may be further processed and analyzed. EEGLAB allows users to group data from several areas, and cluster their unique components. The input retrieval is through EEGLAB software installed in the personal computer which is used to convert the .eeg file generated by the EEG monitor to .txt file which in turn can be easily exported to the excel sheet that can then be used as input. Block diagram is in Figure 4.2)

### 3.1.2 Feature Extraction (PCA)

The electrodes generate the signals from patient's brain and give them to the Principal Component Analysis for dimensionality reduction for removing the redundant variables in the data by converting the multiband signal to a single-band signal and these are classified using Neural Network classifier with back propagation.

In PCA the better classification of signals is obtained for the learning variables like epochs as 1000, number of hidden layers as 3, goal of it as 0.01, and sigmoid function as a threshold of 0.5,tensing and training type.

A data reduction method will be applied to each signal for converting multi band to single band data using PCA. PCA is used to reduce the large dimensionality of the data and multi spectral band reduction through extracting features like covariance, Eigen values and vectors. It is useful for discriminating the pattern of different signals with limited features. The signal reduction is used to explain the majority of its variability compared to multiband features. It is also named as Karhunen-Loève transform or proper orthogonal decomposition.

The algorithm of PCA:

Step-1: Input the Samples

Step-2: Compute Mean value as

$$M=sum(I_{ij})./N$$

Step-3: Find the Difference Matrix,

$$D= I-M$$

Step-4: Calculate Covariance, C=D*D’

Step-5: Compute the Eigen Vectors

$$[V\ D]= eig(C)$$

Step-6: Obtain Features

The brain signals are trained using Neural Network. During the classification of the mental tasks using Neural Network classifier, the data is misclassified at the output that is, 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 correctly and accurately classified at the output . The percentage of accurate classification is high because of 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 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 types of classification.

### 3.1.3 Fuzzy Logic (FL)

Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which most of the modern computers are running. The idea of fuzzy logic was first introduced by Dr. Lotfi Zadeh from the University of California at Berkeley in the 1960s. Dr. Zadeh was dealing with the problem of computer understanding of natural language. Natural language (like most other activities in life and indeed the universe) is not easily translated into the absolute terms i.e., binary form 0 and 1. (Whether is it possible to describe everything in binary terms is a shocking question, but in reality much data we might want to feed a computer are in states between 0 and 1 ).

### 3.1.4 Experimental Results

Mean of entire db:

We can calculate the mean of the sample. The mean of a sample is given by the formula:

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

Notice the symbol (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".

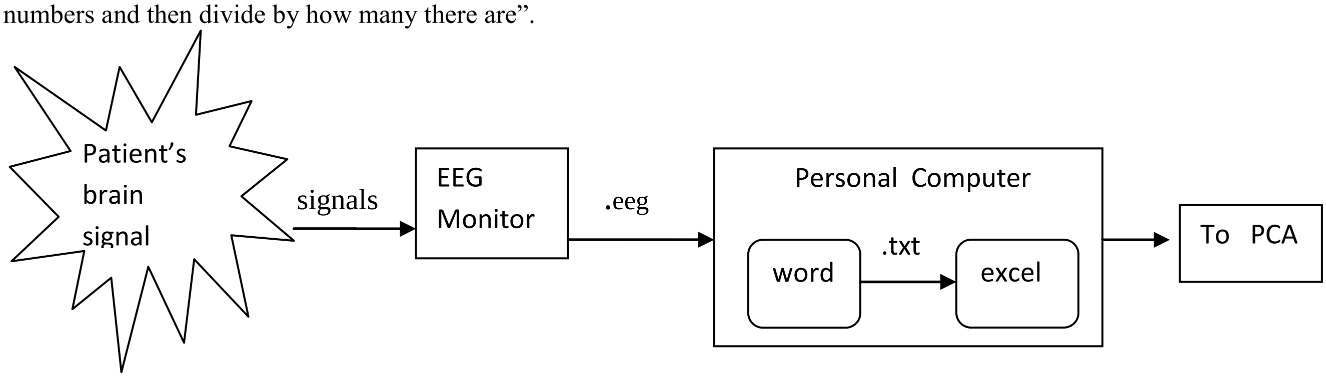



**Fig 4.2: Input acquisition**

**3.2. Classification (BPN)**

The classification of EEG signals is done here by using back propagation method. The back propagation algorithm plays a key role in computing the necessary corrections, after choosing the weights of the network in random. The algorithm can be categorized into the following four steps:

i) Feed-forward computation

ii) Back propagation - output layer

iii) Back propagation - hidden layer

iv) Weight updates

The algorithm continues until the value of the error function has become sufficiently small.

**3.2.1Forward path**

i. Initialize weights.
ii. Choose activation function
iii. Apply i/p. Calculate o/p of hidden layer which is the input to output layer.
iv. Then calculate o/p of output layer.
v. Calculate error.

**3.2.2Reverse path**

i. Adjust the weight of the o/p layer based on the error using delta learning rule.
ii. Adjust the weight of the hidden layer based on the weight of the output layer.
iii. This is continued until error is minimized.

**3.3Performance metrics**

The performance metrics of BPN classifier can be calculated using the following variables,

Sensitivity: It is the one that is used to measures the proportion of 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 is the one that measures the proportion of actual 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)

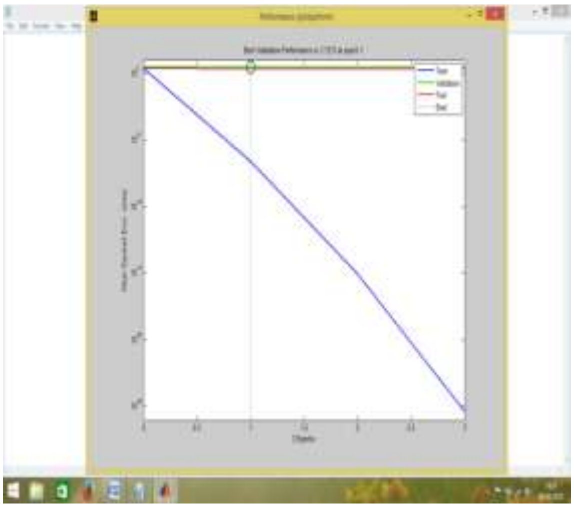


**Fig 5: Performance Plot**

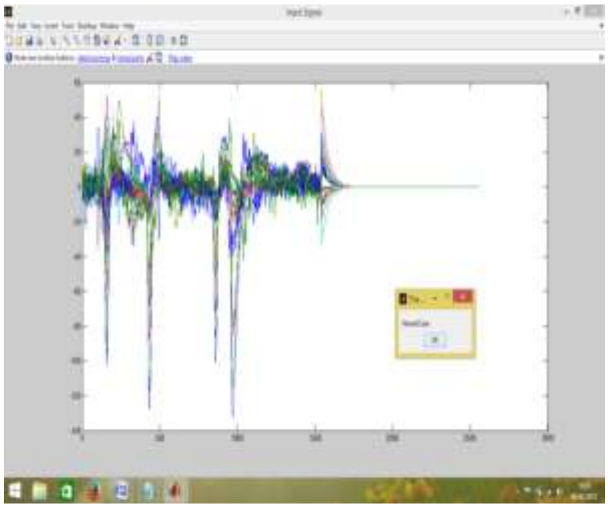


**Fig 6: EEG of Normal patient**

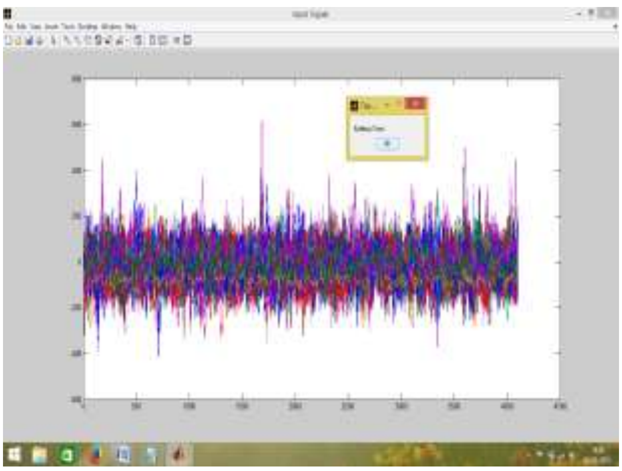


**Fig 7: EEG of Epileptic Patient**

## IV CONCLUSION

Human brain is judgmental ,neural networks has become the solution for the BCI problem .The task are based on brain signals classification is improved by several methods of pre-processing techniques, to generate the input data for the classification of signal with the help of neural networking concept. The soft computing techniques are used for the classification of the EEG signals as they are used to model and make accurate and possible solutions to real world circumstances. The probability of accurate classification has been increased by using soft computing techniques like Principal Component Analysis with artificial neural network and Fuzzy Logic techniques.

## V ACKNOWLEDGEMENT

We thank the Department of Information Technology of Panimalar Engineering College for permitting to use the computational facilities by providing us most suitable environment for research and development. we are especially thankful to Dr. D.Karunkuzhali, Associate Professor; Mrs.S. Mary Rexcy Asha, Associate Professor and Dr.M. Helda Mercy, HOD-Department of Information Technology, Panimalar Engineering College for their contribution and support for the successful completion of this paper.

## REFERENCES

[1] Yijun Wang and Shangkai Gao "Common spatial pattern method for channel selection in motor imagery based BCI", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005.

[2] Aleˇs Proch´azka, Jarom´ır Kukal and Oldˇrich Vyˇsata "Wavelet Transform Use for Feature Extraction and EEG Signal Segments Classification" Institute of Chemical Technology in Prague Department of Computing and Control Engineering,2008.

[3] L.M.Patnaik and Ohil K. Manyam "Epileptic EEG detection using NN and post classification", computer methods and programs in biomedicine 91 (2008)100–109.

[4] Min Han and Leilei Sun "EEG signal classification for epilepsy diagnosis based on AR model and RVM", International Conference on Intelligent Control and Information Processing August 13-15, 2010 - Dalian, China.

[5] Abdulhamit Subasi and M.Ismail Gursoy "EEG signal classification using PCA , ICA, LDA and SVM", Expert Systems with Applications 37 (2010) 8659–8666.

[6] Gerwin Schalk and Eric C. Leuthardt "BCI using Electrocaticograph signals", IEEE REVIEWS IN BIOMEDICAL ENGINEERING,VOL.4,2011

[7] Indu Sekhar Samant, Guru Kalyan Kanungo, Santosh Kumar Mishra "Desired EEG Signals For Detecting Brain Tumour Using LMS And FFN", International Journal of Engineering Trends and Technology-Volume3Issue6- 2012.

[8] Nilima Mohite , Rajveer Shastri , Arnab Das "Epileptic Electroencephalogram Classification", International Conference on Communication and Signal Processing, April 3-5, 2014, India.

[9] R. J.Deshmukh and R.S Khule "Brain Tumor Detection Using Artificial Neural Network Fuzzy Inference System (ANFIS)" , International Journal of Computer Applications Technology and Research Volume 3– Issue 3, 150 - 154, 2014, ISSN: 2319–865.

[10] S. Saraswathi and Dr. S. Shenbaga Devi "Detection of Epileptic Seizure", International Journal of Engineering Research & Technology (IJERT) Vol. 3 Issue 9, September- 2014.