Original Research Article **Neurobiology**

International Journal of Pharma and Bio Sciences

ISSN
0975-6299

A COMPARATIVE ANALYSIS OF FEATURE EXTRACTION AND MACHINE LEARNING BASED CLASSIFIER FOR EEG SIGNAL CLASSIFICATION

SANDEEP KUMAR SATAPATHY*[1], SATCHIDANANDA DEHURI[2] AND ALOK KUMAR JAGADEV[3]

[1]*Department of Computer Science & Engineering, ITER, S'O'A University, Bhubaneswar, Odisha*
[2]*Department of Information & Communication Technology, Fakir Mohan University, Balasore, Odisha*
[3]*School of Computer Engineering, KIIT University, Bhubaneswar, Odisha*

ABSTRACT

Electroencephalogram (EEG) is a trial did on the brain to record the electrical activity inside it. The neural structure of the brain can consist of various neurons in terms of lacs or crores. These neurons communicate by colliding among themselves and communicating data to each other. This collision leads to the generation of the very small amount of electricity. The electrical signal generated can then be recorded and carefully studied to solve many neurological disorder diseases for example epilepsy. About 1% of the total population in the world are affected by this disease. In this study, the behavior of the EEG signals was analyzed by extracting the required important features, as well as classifying the extracted features to detect epileptic seizures. This analysis was done using different machine learning techniques such as Multilayer Perceptron Neural Network (MLPNN), Radial Basis Function Neural Network (RBFNN) and Support Vector Machines (SVM).

KEYWORDS: Electroencephalogram (EEG), Discrete Wavelet Transform, Continuous Wavelet Transform, Thresholding Technique, Classification

SANDEEP KUMAR SATAPATHY
Department of Computer Science & Engineering, ITER, S'O'A University, Bhubaneswar, Odisha

INTRODUCTION

The Electroencephalogram (EEG) signal is usually used for the purpose of recording the electrical activities of the brain signal that typically arises in the human brain. An EEG signal can be defined as the amount of currents that flow during the synaptic excitations of the dendrites in the cerebral cortex.[1] The synaptic currents are formed within the dendrites during the creation of the brain cells. As a result of which a magnetic field measurement is created from the currents by electromyography (EMG) machines and a secondary electrical field over the scalp measurable by EEG systems. The recording of the electrical activity is essentially caused by embedding electrodes on the scalp for 20-40 minutes, which measures the voltage fluctuations in the brain.[2] The source of electric charge is the neurons in the brain, so they exchange ions with the extracellular milieu. Ions of the same charge ward off each other and in this manner they are forced out of the neurons when a number of ions are driven out at the same time they promote each other and form a path that is known as volume conduction.[3] When this wave reaches the electrode they push or force the ions along the airfoil of the electrode which create potential difference and this voltage difference recorded over time gives EEG signals.Human scalp EEG was intuitive in 1920 by the German physician that measured the brain activities on the scalp and since then the interpretations of the EEG signal patterns have been a challenging topic. There are certain unwanted signals that are generated during the EEG signal and are recorded over time which are otherwise called as artifacts. These artifacts are later pre-processed as they create noise. The artifacts are basically separated into two types: physiological and non- physiological; the former occur due to the some activities of the patient's body like eye movement, respiration, blinking of the optics and the latter take place due to some trouble in the EEG machine like wrong placement of electrodes, environment sources, etc.[4-5] Different types of feature reduction techniques are used to remove these artifacts.The wave patterns generated in EEG signals are of various types: delta waves, theta waves, alpha waves, beta waves, gamma waves and lastly mu waves.[2-3, 6- 8] Delta waves ranges up to 4 Hz and is mainly seen in normal adults while they are in sleeping form; theta waves ranges from 4Hz to 7 Hz and is mostly seen in young children when they are drowsy or in arousing condition like during meditation; alpha waves ranges from 8 Hz to 12 Hz and is seen in adults while they are closing their eyes or during relaxation period; beta waves ranges from 12 Hz to 30 Hz and is observed in adults and children who are very active and keep their eyes open for a longer time; gamma waves ranges from 30 Hz to 100 Hz and is largely found when neurons of different populations are brought together for a function and; mu waves ranges from 8 Hz to 13 Hz and it partially overlaps with other frequencies.EEG has a good compact of clinical uses like to differentiate epileptic seizures from others like fainting, or migraine variants to serve as a trial for brain death, differentiate organic encephalographic from primary psychiatric syndromes like catatonia, etc. Recently, classifying epileptic seizure and non epileptic-seizure signals have attracted quite a considerable attention in the area of machine learning community. Some usage towards this is as follows: Multilayer Perceptron Neural Network (MLPNN), Adaptive Neuro-fuzzy Inference System (ANFIS), Radial Basis Function Neural Network (RBFNN) and Recurrent Neural Network (RNN) for epileptic seizure identification. Apart from these, some other machine learning algorithms have also been built up like learning vector quantization (LVQ) and support vector machine (SVM) for identification of epileptic seizure successfully. Here, few important machine learning techniques have been implemented such as MLPNN, RBFNN, SVM, etc. Apart from this, MLPNN is compared with different learning methods such as Back-propagation, Resilient-propagation, Quick-propagation and Manhattan update rule, etc. Similarly, RBFNN is constructed with different basis functions such as Gaussian function, Multi-quadric and Inverse multi-quadric function. Furthermore, SVM is designed with several kernel functions such as Linear, Polynomial and RBF kernel functions etc. With this preliminary introduction to the EEG signal and its significant contribution towards the growth of automatic medical diagnosis system, the focus of this paper is set out as follows. In following section , the methodology of EEG recording is discussed. Next sections deals about the wavelet transform in general and its usage in signal decomposition and extraction. Thresholding techniques for signal de-noising and machine learning techniques for epileptic seizure identification are discussed in the remaining sections respectively. In the last section an overall observed conclusion and summary is drawn.

EEG RECORDING METHODOLOGY

It has been often found that the EEG signals are typically non-Gaussian, non-stationary and having a non- linear nature. The EEG machine consists of four components electrodes, amplifiers, filters and recording unit.[3] An electrode is basically metallic which is utilized and intended for carrying the electrical movement of the brain to the input circuit of the amplifier in the EEG signal.[4] An electrolyte is a conducting solution normally a gel or paste or fluid of living tissue when the electrode is under the skin.[5] In EEG recording, the electrodes are positioned on the scalp after preparing the scalp area by light abrasion to lessen impedance due to dead skin cells. In a normal recording, 19 electrodes are placed on the scalp. In case of high density recording up to 256 electrodes can be placed. [3,6] The electrodes are located in precise locations which are recognized by 10-20 international system or 10-10 international system. From the electrodes positioned on the scalp, the EEG signals are recorded. Brain is divided into two segments or hemispheres, right and left where each hemisphere has four lobes: frontal, parietal, occipital and temporal. The association between the position of the electrode and the region of the cerebral cortex is where the system is based. Each site is represented by a letter and a number where, the letter denotes the lobes and the number denotes the hemisphere location. The right hemisphere and the left hemisphere can be depicted by even numbers and odd numbers correspondingly. Each electrode is linked to the input differential amplifier; a common system reference electrode that is related to the other input of the differential amplifier. The voltages between the active electrodes and the reference can be strengthened by these amplifiers. The EEG signal is then conceded through a series of low pass and high pass filter for filtering purposes. There are usually two types of EEG recording: monopolar and bipolar, where the former picks up the difference in voltage between an active electrode on the scalp and a reference electrode on the ear lobe and the latter gives the dissimilarity occurring in the voltage between two scalp electrodes.

WAVELET TRANSFORM FOR SIGNAL DECOMPOSITION

There are various methods for analyzing the EEG signals like thresholding technique, wavelet transform, Fast Fourier Transform (FFT) etc. In our study, we will be particularly discussing the thresholding technique along with the wavelet approach as well as will provide some suitable results that were obtained using the wavelet approach.

(i) Wavelet Transform Approach

The EEG signal consists of oscillations of different frequency range, spatial distribution that are related to different states of brain operation. This oscillation represents the activities of neural network. EEG signal being a non-stationary signal is not so simple to get transitory and dissimilar features from the signal without any proper methodology being applied to it. Non-stationary signals can be defined as those signals whose statistical distinctiveness fluctuate with time. Various methods like Fourier transform and parametric power method can be used for EEG signal analysis. Fourier transforms of time domain signal gives frequency domain depiction. Hence, it was not possible to provide the information about the time evolution of the signal frequencies using the Fast Fourier Transform (FFT). The real cause is that when one is in time domain then we cannot get information about the frequency and hence the same happens when one is in the frequency domain. Besides, the output of FFT suffered from large noise sensitivity and parametric power spectrum estimation methods which were not suited for frequency distribution.[9,10,11] The classic Fourier transform was also not suited for studying these kinds of signals as they cannot furnish any data on how the frequency changes over time and at what times these frequency components occur.[9] In non-stationary signals, short time Fourier transform can be used, but it do have a time-frequency resolution problem. The wavelet technique is an influential tool for investigating small-scale oscillations of the brain signals. The wavelet transform was projected in the late 1980s, as it was the most effective technique for analysis of non-stationary signal over spectral analysis that include both time and frequency domain and is well suited in locating the transient events occurring during signal recording.[12] It offers an efficient mode of interpreting signal on time frequency domain using the variable sized windows. The long-time windows are utilized for low-frequency resolution and short-time windows for high-frequency. This makes the wavelet transform more appropriate for analyzing the irregular data patterns. The wavelet transform decomposes the signal into different scales with different levels of resolution by dilating the mother wavelet.[13] It is defined as shown in equation (1):

$$\phi_{a,b}(t) = \frac{1}{\sqrt{a}}\phi\left(\frac{t-a}{b}\right) \qquad (1)$$

The parameters $a, b \in R, a \neq 0$ measure the degree of compression and the time position of the wavelet respectively. When |a| < 1, then $\phi_{a,b}(t)$ is the compressed version of the mother wavelet and relates mostly to upper frequencies. When |a| > 1, then $\phi_{a,b}(t)$ has a larger time-width than mother wavelet $\phi(t)$ and relates to lower frequencies. This is the main cause for using wavelets in signal processing and time-frequency signal analysis. The Wavelet Transform uses multi-solution technique for analyzing unusual frequencies with diverse results. This technique is intended to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Akin *et al.* proposed a methodology that was applied to wavelet transform from which the artificial neural network was derived and the accuracy found was 97% as in epileptic cases, 98% for healthy cases, and 93% for pathological cases that were considered for testing.[14] X. Li *et al.* applied the wavelet technique and suggested that wavelet transform is a constructive tool to scrutinize the EEG signals with the epileptic seizures.[15] Ganesan *et al.* have applied the technique in which the signals were decomposed into sub- bands and after evaluating the data from 81 patients, they correctly detected epileptic events that were 90.0% and that of non-epileptic events were 98.0%.[16] Wavelet transform is classified into two forms: Discrete Wavelet Transform (DWT) and Continuous Wavelet Transform (CWT).

(ii) Discrete Wavelet Transform

Crochiere *et al* suggested a signal that relates to sub-band coding and pyramidal coding or multi-resolution analysis.[17] This is known as Discrete Wavelet Transform (DWT). Multi-resolution filter banks and special wavelet filters were used by DWT to study and reconstruct the signals. It includes iteration of filters with rescaling. Successive low pass and high pass filtering of the discrete time-domain signal is needed for the transformation process. The resolution of the signal (that is a measure of the quantity of detailed information in the signal) is determined by the filtering process, and the scale is determined by up sampling and down sampling.[11,13] This was called the *Mallat algorithm* or *Mallat-tree decomposition*.[13] This algorithm for the signal can be broken down into the approximations and details. The input signal is converted into one high-pass wavelet coefficient series and one low-pass wavelet coefficient series (of length reduced to half). DWT can also be used for compressing the information.

(iii) Continuous Wavelet Transform

A continuous wavelet transform is used to divide a continuous-time function into wavelets.[18,19,20] It maps or convert the original signal that is a function of one independent variable *t* (time) into a function of two independent variables *a* and *b* (time and frequency). Unlike Fourier transform, the continuous wavelet transform possess the capability to create a time-frequency demonstration of a signal that offers quite a good time and frequency localization.[13,18-19] The drawbackof CWT is that it provides redundant information from which the original signal is not possible to recover, since *a* and *b* are continuous.

FEATURE EXTRACTION METHODOLOGY

(i) DWT for Feature Extraction

From the data existing at a rectangular window of length 256 discrete data were chosen to form a single EEG segment.[21] It is really significant to select a suitable wavelet and the number of decompositions for efficient wavelet transforms. The wavelet coefficients can be computed using Daubechies wavelet of order 2, 3, 4, etc. This technique is more efficient because, its smoothing features are more appropriate to detect changes in EEG signal. In the present study, the EEG signals were decomposed into details d1, d2, d3, d4 and one approximation a4. For four detailed coefficients we get 247 coefficients (129+66+34+18) and 18 for the approximation coefficient. From these coefficients the following important features can be extracted:

- Maximum
- Mode
- Minimum
- Standard Deviation
- Mean
- Mean Absolute Deviation
- Median
- Median Absolute Deviation
- Range
- L2 Norm
- L1 Norm
- Max Norm

Figure 1 show the statistical feature extraction from wavelength co-efficient and Figure 2 shows the discrete wavelet transform of a sample EEG signal using DB2 and decomposed up to level 4. Each set A to E has been decomposed into 5 coefficients (4 detailed and 1 approximations). From each coefficient 4 statistical features have been extracted. So all total a dataset of size 500x20 is being constructed.

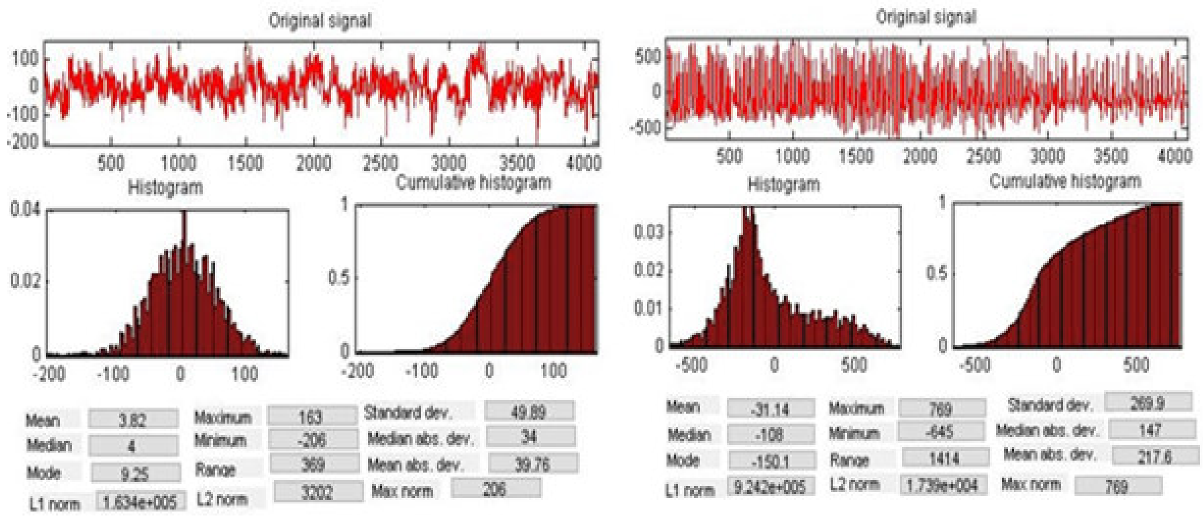


Figure 1
Statistical feature extraction using DWT

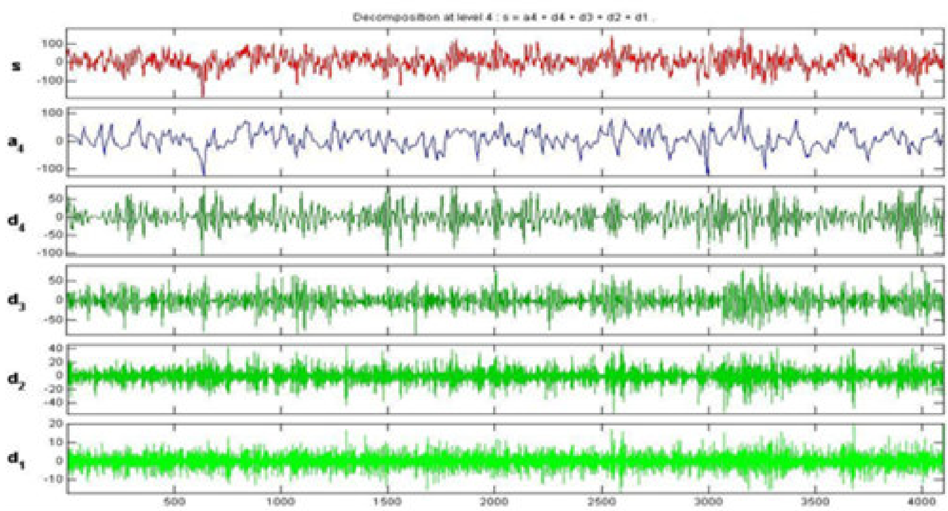


Figure 2
Discrete wavelet transforms of a sample EEG signal using DB2 and decomposed up to level 4

(ii) ***Thresholding for De-noising***

After getting the coefficients, certain methodologies have to be applied for de-noising (shown in Figure 3) that means removing the noise from the signal so as to get a noise-free signal.[17] In this technique, a threshold value is selected and is compared with the detail and approximate coefficients obtained using DWT. The two types of thresholding techniques are hard thresholding and soft thresholding de-noising (wavelet shrinkage de-noising), that provide a narrative way to lessen noise in the signal.[22] When compared, it was found that soft thresholding was the best in reducing noise.[22-24] Donoho *et al.* introduced soft thresholding as a powerful tool for de-noising signals. Hard thresholding is the standard process of setting zero to the coefficients whose values are lower than the threshold.[22] Soft thresholding involves first setting the coefficients to zero whose values are lower than the threshold, then shrinking the non zero coefficients towards 0. [23-25] The threshold value can be manually selected or can be obtained by the method proposed by Donoho *et al.*[25] The thresholding extracts the major coefficients by setting to zero and also for the coefficients where the absolute value is below a certain threshold level.

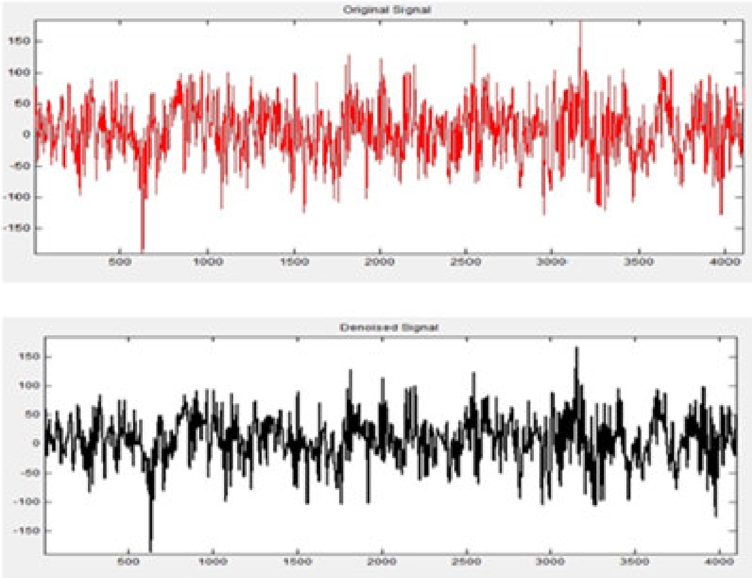


Figure 4
Denoising of EEG Signal Using Soft Thresholding Technique

(iii) ***Other Feature Extraction Techniques***

Like DWT, there are many other techniques for extracting features from an EEG data set.[26] the features that can be extracted from the techniques include:

a. Fractal Dimension (Higuchi and Petrosian)
b. Hurst Exponent
c. Spectral, Approximation and SVD Entropy
d. Detrended Fluctuation Analysis
e. Hjorth Mobility and Complexity

These features extracted provides us a sector to explore the EEG dataset in a more detailed way for the purpose of classification. Hence, along with the features that have been extracted from DWT, we have also used these nine features for our experimental analysis process of classification. The techniques have been elaborated and stated as given below. For mathematical discussions, let us consider the signal for which features are being extracted is

a. Fractal Dimension [2,28]

It is one of the important feature of a signal that may hold some data around the geometrical shape at different scales. This information's can be extracted using different methods as proposed by Petrosian and Higuchi and named as Petrosian Fractal Dimension (PFD) and Higuchi Fractal Dimension (HFD). Equation 2 shows formula for calculating PFD.

$$PFD = \frac{\log_{10} S}{\log_{10} S + \log_{10}\left(\dfrac{S}{S + 0.4 S_{\varphi}}\right)} \qquad (2)$$

Where, S is the series length and S_{φ} ***is the number of sign changes in the signal.***

Similarly, HFD is the slope of line that best fits the curve of *ln*(*Z*(*k*)) and *ln*(1/*k*) where, *Z*(*k*) is defined in equation 3.

$$Z(k) = \frac{\sum_{i=1}^{k} Z(i,k)}{k} \qquad \text{where,}$$

$$Z(m,k)=\frac{\sum_{i=1}^{\lfloor (N-m)/k \rfloor}\left|X_{m+ik}-X_{m+(i-1)k}\right|(N-1)}{\lfloor (N-m)/k \rfloor k} \quad (3)$$

This algorithm constructs k new series from original series as shown below:

$X_m, X_{m+k}, X_{m+2k}, ...X_{m+\lfloor (N-m)/k \rfloor k}$ here m=1,2,...k

b. Hurst Exponent (HE) [29]

It is generally used as a measure of long term memory of time series data. It can be calculated by first calculating deviation from mean of the time series and then by calculating the rescaled range statistics (*R*/*S*). First, we have to calculate the accumulated deviation from mean of time series within range *T* as shown in equation 4.

$$X(t,P)=\sum_{i=1}^{t} x_i-\bar{x} \quad \text{where } \bar{x}=\frac{1}{p}\sum_{i=1}^{p} x_i \text{ , t € [1,...N],} \quad (4)$$

Then *R*(*P*)/*S*(*P*) is calculated as per the formula shown in equation 5.

$$\frac{R(P)}{S(P)}=\frac{\max(X(t,p))-\min(X(t,p))}{\sqrt{(1/P)\sum_{t=1}^{P}\left[X_t-\bar{X}\right]^2}} \quad (5)$$

The Hurst Exponent is calculated as the slope of line produced by *ln*(*R*(*P*)/*S*(*P*)) versus *ln*(*P*).

c. Spectral Entropy (SE), Approximation Entropy (AE) and SVD Entropy (SVDE) [31-33]

Entropy can be measured as the spread of data. Data with broad or flat probability distribution has a high entropy and vice-versa. This is one of the statistical description of variations in EEG signal. Spectral entropy can be defined in terms of Power Spectral Intensity (PSI) and Relative Intensity Ratio (RIR) as shown in equation 6.

$$SI=\frac{-1}{\log(K)}\sum_{i=1}^{K} RIR_i \log RIR_i \quad (6)$$

Where $RIR_j=\frac{PSI_j}{\sum_{K=1}^{K-1} PSI_k}$ and $PSI_k=\sum_{i=\lfloor N(f_k/f_s)\rfloor}^{\left\lfloor N\left(f_k+\frac{1}{f_s}\right)\right\rfloor}\left|X_i\right|$, *k*=1, 2, ...*K*-1.

Where, f_s is the sampling rate X_i denotes FFT of time series x_i. f_1 to f_K represents *K* slices of frequency band of equal or unequal widths. Similarly, *approximation entropy* is a statistical parameter computed for a time series data. SVD entropy defines an entropy measure by the help of Singular Value Decomposition.

d. Detrended Fluctuation Analysis (DFA) [34]

It is another important feature extracted for analysis of signals with scale invariant structure. It is a way for determining statistical self-affinity of a signal. The exponents obtained are almost similar to Hurst exponent.

e. Hjorth Mobility and Complexity (HJM, HJC) [35]

Hjorth parameters generally describe the statistical properties of a signal. This is a very popular signal analysis method proposed by Hjorth in 1970, used for analyzing electroencephalogram signals. It has mainly three kinds of parameters such as activity, mobility and complexity. Here, mobility and complexity are used for the analysis of the EEG signal that uses the activity parameter. Mathematically, it can be defined as shown in equation 7.

$$Mobility=\sqrt{B2/AVG}\text{ ; } Complexity=\sqrt{(B4*AVG)/(B2*B2)} \quad (7)$$

Where, $AVG=\sum x_i/N$, $B2=\sum d_i/N$, , $B4=\sum\frac{(d_i-d_i-1)^2}{N}$ $d_i=x_i-x_{i-1}$,

After the features have been extracted, the next important task is to design a classifier model for classifying the seizure and non-seizure signals. For this task different classifier models such as MLPNN, RBFNN, SVM, etc were considered. These techniques have been briefly described in the next section.

MACHINE LEARNING BASED CLASSIFIERS

It has been found that about 1% of the people in world are usually affected by epilepsy and the main reason behind epilepsy is the occurrence of recurrent seizures. So, if the EEG signal will be carefully observed then epilepsy can be detected in a person. After extracting the important features from EEG signal using above mentioned techniques, several methods can be used to design an efficient classifier which can classify epileptic and non-epileptic seizures. The following figure (figure 4) describes the model for different feature extraction and classification methods used for EEG signal in epilepsy identification. After this, we have compared the performance of different classifier based upon their accuracy.

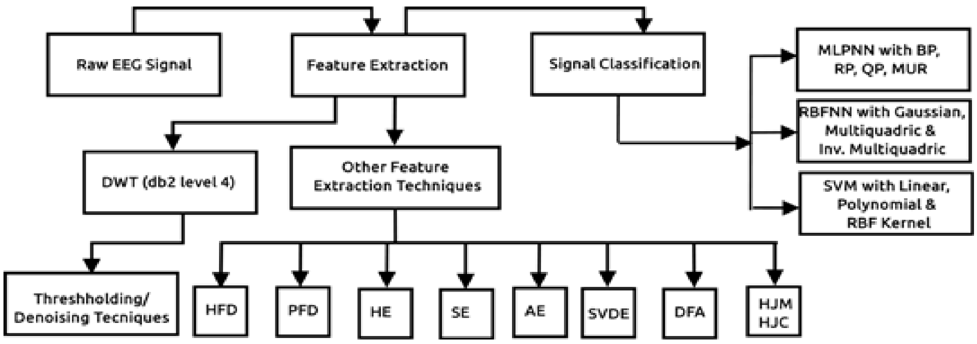


Figure 4
Model for feature extraction and classification of EEG signal

(i) ***Machine Learning Algorithms***

Machine learning is a type of artificial intelligence (AI) that provides computers with the aptitude to study without being overtly programmed. Machine learning focuses on the growth of computer programs that can teach themselves to grow and change when exposed to new data. It examines how to automatically learn to arrive at accurate predictions based on past observations. It classifies examples into giving set of categories as described in Figure 5. In the following sections, some of the machine learning techniques have been discussed for EEG signal classification along with their performance measurement.

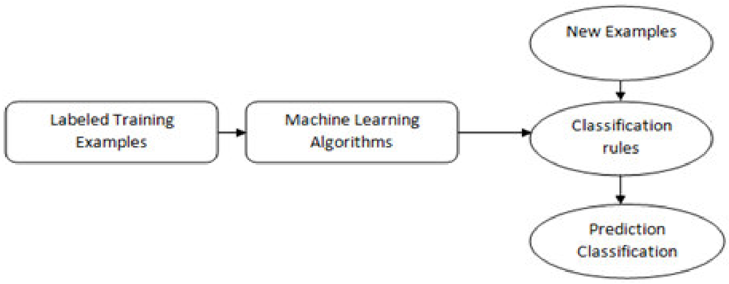


Figure 5
Procedure for Machine Learning Algorithms

a. Multilayer Perceptron Neural Network (MLPNN)

MLPNN is the most commonly and popularly used neural network structure.[36] It is an extension of the single layer perceptron that mainly consists of a set of source nodes called input layer, one or more hidden layer nodes and an output layer. These types of networks have been used to resolve many pattern recognition problems by training them in supervised manner by using a highly popular algorithm based on error correction rule called an error back propagation algorithm. Guler *et al.* have found an accuracy of 93.63% by applying the MLPNN along with using feature extraction techniques like DWT and Lyapunov Exponents.[37] Guo *et al.* have found an accuracy of 99.6% by applying MLPNN along with line length features based on wavelet transform. They have found this accuracy taking the dataset of normal and epileptic persons.[38] They found an accuracy of 97.7% by considering all the situations in a normal person that is by taking set A, B, C and D as one class and considering set E as another class which is from epilepsy affected person. Kashtiban *et al.* have applied wavelet transform as well as derived the exact class by using WARD and dendogram graph. And then applied MLPNN for classification for which they got an accuracy of 98.67%.[39] Orhan *et al.* analyzed the signal using DWT and then the wavelet coefficients are clustered using *k*-means for each frequency sub band.[40] The probability distribution was computed according to the distribution of wavelet coefficients to the cluster and then used as input to MLPNN model. They found an accuracy of 100% for two class problem used for only dataset A & E.

b. Support Vector Machines **(***SVM***)**

SVM was first proposed by Vapnik which has been widely used for classification, regression and density estimation.[41] It usually maps the input patterns into a higher dimensional feature space through some nonlinear mapping. These are supervised learning models with related learning algorithms that analyze data and distinguish patterns. It is very important to select a proper kernel function for efficient classification using SVM. Guler *et al.*, have proposed an efficient multiclass SVM for EEG signal classification.[37] They had used PNN and MLPNN as benchmark for their performance. They did feature extraction by using wavelet coefficients and Lyapunov exponents and got the accuracy of the classifier up to 99.28% which is better compared to PNN and MLPNN. Subasi *et al.*, projected an proficient classification technique for epileptic seizure detection using SVM.[42] They had applied SVM along with DWT and three dimension reduction techniques such as PCA, ICA and LDA for which they found accuracies of 98.75%, 99.5% and 100% respectively.

c. Radial Basis Function Neural Network (RBFNN)

RBF networks are also a type of feed-forward network trained using a supervised training algorithm.[43] The primary advantage of RBF network is; it has only one hidden layer. The RBF network, usually trains much faster than back-propagation networks. This form of network is less susceptible to problems with non-stationary inputs because of the behavior of radial basis function hidden units.

RESULTS AND ANALYSIS

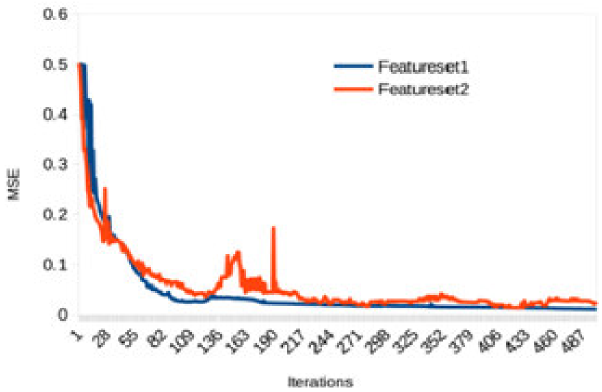
This part reports the comparative analysis of different machine learning based classification techniques such as MLPNN, RBFNN and SVM for epileptic seizure identification. Table 1 shows the value of different parameters used for different classification techniques. Table 2 depicts the comparison of different parametric measures for MLPNN with Resilient-propagation, Back-propagation, Manhattan Update Rule and Quick-propagation training algorithms. In these tables, two feature sets, Featureset1 (Includes features like HFD, PFD, HE, SE, AE, SVDE, DFA, HJM, HJC) and Featureset2 (Includes features extracted from DWT technique) are used. The different parametric measures used for comparison are like Testing Accuracy, Specificity, Sensitivity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), False Positive Rate (FPR), False Discovery Rate (FDR), F1-score and Time in seconds. From this table, we can conclude that Resilient-propagation always provides better results as compared to the other propagation techniques. Except that, in some situations Featureset1 can be considered as more effective features compared to Featureset2. Figure 6 shows the Mean square error graph comparison of two different feature sets with MLPNN using Resilient and Back-propagation. Figure 7 shows the Mean square error graph comparison of two different feature sets with MLPNN using Manhattan update rule and Quick-propagation. From this analysis, we can aver that Quick - propagation cannot be used efficiently for any of the feature sets. Here, all the experiments have been validated by using *k*-fold cross validation. Here, *k* value is taken as 10. Hence, the whole dataset has been divided into 10 subsets. In each iteration one subset is used for testing purpose and the rest 9 subsets are used for training purpose. At the end, the average accuracy has been recorded as the final result.

Table 1
Parameter description used for MLPNN, SVM and RBFNN

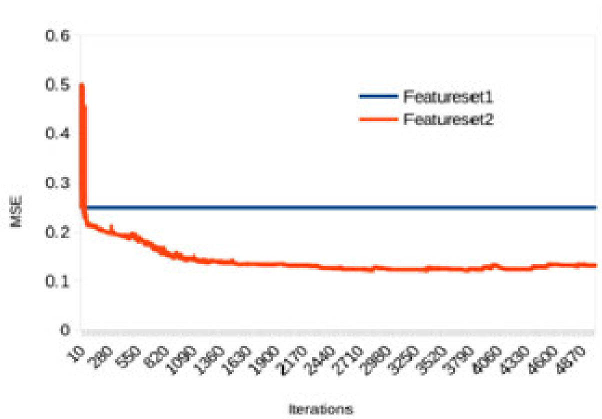
Classification Techniques	Required Parameters and values
MLPNN/BP	Activation Function - Sigmoid Learning Rate = 0.7 Momentum Coefficient = 0.8 Input Bias - Yes
MLPNN/RPROP	Activation Function - Sigmoid Learning Rate = NA Momentum Coefficient = NA Input Bias – Yes
MLPNN/MUR	Activation Function - Sigmoid Learning Rate = 0.001 Momentum Coefficient = NA Input Bias - Yes
SVM	Kernel Type – Linear, Polynomial, Radial Basis Function Penalty Factor = 1.0
RBFNN	Basis Function – Inverse Multiquadric, Multiquadric, Gaussian Center & Spread Selection – Random Training Type– SVD (Singular Value Decomposition)

Table 2
Performance assessment of MLPNN with different training algorithms

Performance Measures	MLPNN with Resilient-Propagation		MLPNN with Back-Propagation		MLPNN with Manhattan Update		MLPNN with Quick-Propagation	
	Featureset1	Featureset2	Featureset1	Featureset2	Featureset1	Featureset2	Featureset1	Featureset2
Accuracy	99.5%	98.5%	50.0%	84.0%	93.5%	50.5%	51.0%	50.0%
Specificity	99.009%	97.087%	NaN	86.170%	88.495%	50.251%	50.735%	50.0%
Sensitivity	100.0%	100.0%	50.0%	82.075%	100.0%	100.0%	51.562%	NaN
PPV	99.0%	97.0%	100.0%	87.0%	87.0%	1.0%	33.0%	0.0%
NPV	100.0%	100.0%	0.0%	81.0%	100.0%	100.0%	69.0%	100.0%
FPR	0.99	2.912	NaN	13.829	11.504	49.748	49.264	50.0
FDR	1.0	3.0	0.0	13.0	13.0	99.0	67.0	100.0
F1-Score	99.497%	98.477%	66.66%	84.466%	93.048%	1.98%	1.98%	0.0%
Time	8.083sec	2.209sec	16.594sec	13.506sec	5.099sec	10.534sec	10.534sec	11.415sec

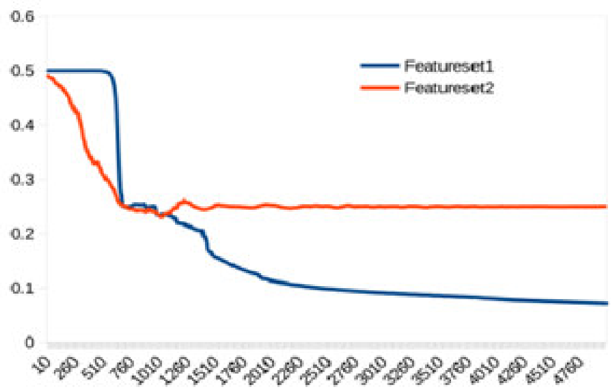


MLPNN with Resilient-propagation for Featureset1 and Featureset2

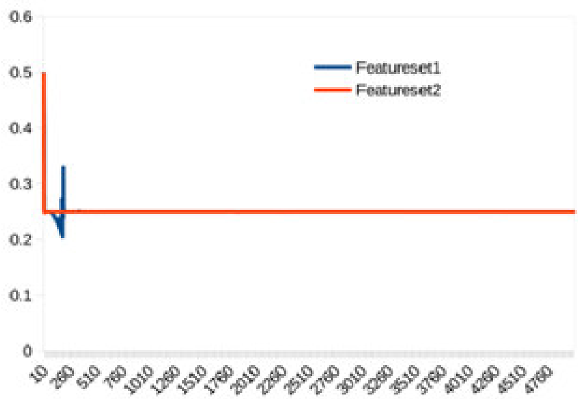


MLPNN with Back-propagation for Featureset1 and Featureset2

Figure 6
Mean Square Error graph for MLPNN with Resilient and Back-propagation algorithm



MLPNN with MHUP for Featureset1 Featureset2



MLPNN with Quiick-propagation for and Featureset1 and Featureset2

Figure 7
Mean Square Error graph for MLPNN with Manhattan Update Rule and Quick-propagation algorithm

Table 3 shows the comparison of different parametric measures for RBFNN with different basis functions such as Inverse Multiquadric, Multiquadric and Gaussian function. From this table, it can be clearly inferred that for EEG signal analysis MLPNN is more efficient than RBFNN. Also, somehow the performances of both the feature sets with RBFNN are with less difference. Again, all the same parameters are used for this comparison. In this technique, featureset 1 has bit high performance as compared to featureset 2. RBFNNs are basically designed with simple architecture and fewer parameters. Hence, it is more effective for any application for which there can be some hybridization techniques that is used to enhance the performance of RBFNN.

Table 3
Performance comparison of RBFNN with different Basis Functions

Performance Measures	**RBFNN with Inverse Multiquadric Function**		**RBFNN with Multiquadric Function**		**MLPNN with Gaussian Function**	
	Featureset1	**Featureset2**	**Featureset1**	**Featureset2**	**Featureset1**	**Featureset2**
Accuracy	80.0%	73.0%	77.0%	71.5%	79.5%	66.5%
Specificity	77.272%	88.33%	71.428%	87.719%	79.797%	71.428%
Sensitivity	83.33%	66.428%	86.486%	65.034%	79.207%	63.414%
PPV	75.0%	93.0%	64.0%	93.0%	80.0%	78.0%
NPV	85.0%	53.0%	90.0%	50.0%	79.0%	55.0%
FPR	0.227	0.116	0.285	0.122	0.202	0.285
FDR	0.25	0.07	0.36	0.07	0.2	0.22
F1-Score	78.947%	77.5%	73.563%	0.765	79.601%	69.955%
Time	2.957sec	4.349sec	1.632sec	3.446sec	2.759sec	10.063sec

Table 4 shows the comparison of different parametric measures for SVM with different kernel functions such as Linear, Polynomial and RBF kernel functions. From this table, it can be clearly understood that featureset 1 cannot give more efficient result as featureset 2 has a performance of 100%.

Table 4
Performance comparison of SVM with different Kernel Functions

Performance Measures	SVM with Linear Kernel Function		SVM with Polynomial Kernel Function		SVM with RBF Kernel Function	
	Featureset1	Featureset2	Featureset1	Featureset2	Featureset1	Featureset2
Accuracy	92.0%	100.0%	72.0%	100.0%	80.0%	100.0%
Specificity	86.206%	100.0%	92.307%	100.0%	77.272%	100.0%
Sensitivity	100.0%	100.0%	64.864%	100.0%	83.333%	100.0%
PPV	84.0%	100.0%	96.0%	100.0%	75.0%	100.0%
NPV	100.0%	100.0%	48.0%	100.0%	85.0%	100.0%
FPR	13.793	0.0	7.692	0.0	22.727	0.0
FDR	16.0	0.0	4.0	0.0	25.0	0.0
F1-Score	91.304%	100.0%	77.419%	100.0%	78.947%	100.0%
Time	10.54sec	1.85sec	38.261sec	1.597sec	2.004sec	1.309sec

CONCLUSION

We briefly covered the feature extraction techniques for EEG signal using discrete wavelet transform and others. EEG signal has many applications in clinical studies, like automatic epileptic seizure detection, sleep disorder detection, etc. This is a case of non-stationary signal which cannot be examined by using FFT or any other technique. It is most suitable and efficient to apply DWT for the analysis of EEG signal and soft thresholding technique for de-noising the signal. After this analysis, the statistical features can be extracted which can be used for classification purposes. In that respect, there are several classification models; nevertheless, some machine learning models that is used for EEG signal analysis.

CONFLICT OF INTEREST

I hereby declare that no case of animal study conduction has been done in accordance of the relevant ethical committee.

REFERENCES

1. Sanei S, Chambers JA. EEG Signal Processing. Wiley Publications.
2. Niedermeyer E, Lopes da Silva F. Electroencephalography: Basic Principles. Clinical Applications and Related Fields. Lippincot Williams & Wilkins. 2004.
3. Abou-Khalil B, Musilus KE. Atlas of EEG & Seizure Semiology. Elsevier Publication. 2005.
4. Towle VL, Bolaños J, Suarez D, Tan K, Grzeszczuk R, Levin DN, Cakmur R, Frank SA, Spire JP. The spatial location of EEG electrodes: locating the best fitting sphere relative to cortical anatomy. Electroencephalography and Clinical Neurophysiology. 2003; 86 (1): 1–6.
5. Aurlien H, Gjerde IO, Aarseth JH, Karlsen B, Skeidsvoll H, Gilhus NE. EEG background activity described by a large computerized database. Clinical Neurophysiology. 2004; 115 (3): 665–673.
6. Creutzfeldt OD, Watanabe S, Lux HD. Relations between EEG phenomena and potentials of single cortical cells. I. Evoked responses after thalamic and epicortical stimulation. Electroencephalography and Clinical Neurophysiology. 2006; 20(1): 1–18.
7. Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. Electroencephalography and Clinical Neurophysiology. 1999; 110 (11): 1842–1857.
8. Oberman LM, Hubbard EM, McCleery JP, Altschuler EL, Ramachandran VS, Pineda JA. EEG Evidence for mirror neuron dysfunction in autism spectrum disorders. Cognitive Brain Research. 2005; 24 (2): 190–198.
9. Bronzino JD. Biomedical Engineering Handbook. 2000; New York: CRC Press LLC, (1), 2nd edition.
10. Bradbury D. Volunteer's Guide to an EEG. Institute of Neurology Wellcome, Department of Imaging Neuroscience, *http://www.fil.ion.ucl.ac.uk/EEGvolunteerguide.pdf* .
11. Ocak H. Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm. Signal Processing. 2008; 88: 1858–1867.
12. Sharanreddy M and Kulkarni PK. Review of Significant Research on EEG based Automated Detection of Epilepsy Seizures & Brain Tumor. International Journal of Scientific & Engineering Research. 2011; 2(8): 1-9.
13. Sifuzzaman M, Islam MR and Ali MZ. Application of Wavelet Transform and its Advantages Compared to Fourier Transform. Journal of Physical Sciences. 2009; 13: 121-134.
14. Akin M, Arserim MA, Kiymik MK, Turkoglu I. A New Approach for Diagnosing Epilepsy by using Wavelet Transform and Neural Networks. Proc. of 23rd Annual Int. Conference on Engineering in Medicine and Biology Society. 2001; 2: 1596-1599.
15. Li X. EEG Analysis with Epileptic Seizures using Wavelet Transform. Dept. of Automation and Computer-Aided Engineering, Chinese University of Hong Kong. 2002.
16. Ganesan M, Sumesh EP, Vidhyalavanya R. Multi-Stage, Multi-Resolution Method for Automatic. International Journal of Signal Processing, Image Processing and Pattern Recognition. 2010; 3(2): 33-40.
17. Durand S, Froment J. Artifacts Free Signal Denoising with Wavelets.P roc. International Conference on Acoustics, Speech and Signal Processing. 2001; 6: 3685-3688.
18. Arab MR, Suratgar AA, Martínez Hernández VM, Ashtiani AR. Electroencephalogram Signals Processing for the Diagnosis of Petit mal and Grand mal Epilepsies Using an Artificial Neural Network. Journal of Applied Research and Technology. 2010; 8(1): 120-129.
19. NazariKousarrizi MR, Asadi Ghanbari A, Teshnehlab A, Aliyari M, Gharaviri A. Feature Extraction and Classification of EEG Signals using Wavelet Transform SVM and Artificial Neural Networks for Brain Computer Interfaces. Proc. of International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing. 2009: 352-355.
20. Ganesh A. Classification and Clustering of Brain Seizure activity using Wavelet transform and Radial Basis Neural Network. Journal of Mathematics. 2012; 1(1): 1001–1008.
21. *www.meb.unibonn.de/epileptologie/science/physik/eegdata*
22. Qian J. Denoising by Wavelet Transform. Article. 2001.
23. Donoho DL. De-noising by soft-thresholding. IEEE Trans. on Inf. Theory. 1995; 41(3): 613-627.
24. Polikar R. The Wavelet Tutorial Part IV. Multiresolution Analysis: The Discrete Wavelet Transform. Part.3. 1996.
25. Sumithra MG, Thanushkodi BK. Performance Evaluation of Different Thresholding Methods in Time Adaptive Wavelet Based Speech Enhancement. International Journal of Engineering and Technology. 2009; 1(5): 439-447.
26. Bao FS, Liu X and Zhang C. PyEEG: An Open Source Python Module for EEG/MEG Feature Extraction. Computational Intelligence and Neuroscience. 2011: 1-7.
27. Petrosian A. Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns. Proc. of the 8th IEEE Symposium on Computer-Based Medical Systems. 1995: 212–217.
28. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D. 1988; 31(2): 277–283.
29. Balli T and Palaniappan R. A combined linear & nonlinear approach for classification of epileptic EEG signals. Proc. of the 4th International IEEE/EMBS Conference on Neural Engineering (NER '09). 2009: 714–717.

30. James CJ and Lowe D. Extracting multisource brain activity from a single electromagnetic channel. Artificial Intelligence in Medicine. 2003; 28(1): 89–104, 2003.
31. Inouye T, Shinosaki K, Sakamoto H. Quantification of EEG irregularity by use of the entropy of the power spectrum. Electroencephalography and Clinical Neurophysiology. 1991; 79(3): 204–210.
32. Pincus SM, Gladstone IM and Ehrenkranz RA. A regularity statistic for medical data analysis. Journal of Clinical Monitoring and Computing. 1991; 7(4): 335–345.
33. Roberts SJ, Penny W and Rezek I. Temporal and spatial complexity measures for electroencephalogram based brain-computer interfacing. Medical and Biological Engineering and Computing. 1999; 37(1): 93–98.
34. Peng CK, Havlin S, Stanley HE and Goldberger AL. Quantification of scaling exponents and crossover phenomena in non-stationary heartbeat time series. Chaos. 1995; 5(1): 82–87.
35. Hjorth B. EEG analysis based on time domain properties. Electroencephalography and Clinical Neurophysiology. 1970; 29(3): 306–310.
36. Ishwinder K. S, Nair M, Shukla H and Sandhu S. S. Artificial Neural Network: As emerging diagnostic tool for breast cancer. Int. Jour. Pharma & Bio-Sciences. 2015; 5(3): 29-41.
37. Guler I and Beyli EDU. Multi-class Support Vector Machines for EEG-Signals Classification. IEEE Transactions on Information Technology in Biomedicine. 2007; 11(2): 117-126.
38. Guo L, Rivero D, Dorado J, Rabunal JR, Pazos . Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. Journal of Neuroscience Methods. 2010; 191:101–109.
39. Kashtiban AM, Razmi H, Kozehkonan MK. Combined LVQ neural network and multivariate statistical method employing wavelet coefficient for EEG signal classification. IEEE International Conference on Mechatronics (ICM). 2011:809-814.
40. Orhan U, Hekim M, Ozer M. EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Systems with Applications. 2011; 38: 13475–13481.
41. Vapnik V. The Nature of Statistical Learning Theory. 1995: Springer-Verlag.
42. Subasi A, Ismail Gursoy M. EEG signal classification using PCA, ICA, LDA and Support Vector Machines. Expert Systems with Applications. 2010; 37: 8659–8666.
43. Venkatesan P, Thomas T. A comparative study of Neural Network architectures for predicting gene expression in M. Tuberculosis. Int J Pharm Bio Sci 2013; 4(1): (B) 965 – 972.