Analysis of EEG Signal for the Detection of Brain Abnormalities

M.Kalaivani
PG Scholar
Department of Computer Science and Engineering – PG
National Engineering College
Kovilpatti, Tamilnadu

V.Kalaivani, Ph.D
Associate Professor(SG)
Department of Computer Science and Engineering – PG
National Engineering College
Kovilpatti, Tamilnadu

V.Anusuya Devi
Assistant Professor
Department of Computer Science and Engineering – PG
National Engineering College
Kovilpatti, Tamilnadu

ABSTRACT
In the field of medical science, one of the major ongoing researches is the diagnosis of the abnormalities in brain. The Electroencephalogram (EEG) is a tool for measuring the brain activity which reflects the condition of the brain. EEG is very effective tool for understanding the complex behaviour of the brain. The aim of this study is to classify the EEG signal as normal or abnormal. It is proposed to develop an automated system for the classification of brain abnormalities. The proposed system includes pre-processing, feature extraction, feature selection and classification. In pre-processing the noises are removed. The discrete wavelet transform is used to decompose the EEG signal into sub-band signals. The feature extraction methods are used to extract the time domain and frequency domain features of the EEG signal.

General Terms
Methodology for Information in brain abnormality using EEG Signal.

Keywords
Electroencephalogram, brain diseases, wavelet transform, EEG waves, feature extraction

1. INTRODUCTION
A disease is an abnormal condition that affects the body of an organism**.** Any deviation from the normal structure of a body part or organ is displayed by a characteristic set of symptoms and sign. Electroencephalogram is used for detecting the brain diseases. Electroencephalogram is the recording of electrical activity of the brain from scalp. It measures the voltage fluctuations resulting from ionic current flows within the neurons of the brain. Diagnostic applications generally focus on spectral content of EEG that is the type of neural oscillations that can be observed in EEG signals. EEG is painless and harmless. And it does not pass any electricity into your brain or body. The EEG signals are commonly decomposed into five EEG sub-bands: delta, theta, alpha, beta and gamma. Alpha waves are rhythmic and its frequency range is from 8 to 13 Hz. The amplitude of the alpha wave is low. Each region of the brain has the characteristic of alpha rhythm but mostly it is recorded from the occipital and parietal regions. It oscillates from adult in awake and relaxed state with eyes closed.

Beta waves are irregular and its frequency range is greater than 13 Hz. The amplitude of the beta wave is very low. It is mostly recorded from temporal and frontal lobe. It oscillates from during the deep sleep, mental activity and is associated with remembering. Delta waves are rhythmic and its frequency range is 4 to 7 Hz. The amplitude of the delta wave is high. It oscillates from the children in sleep state, drowsy adult and emotional distress occipital lobe. Theta waves are slow and its frequency range is less than 3.5 Hz. The amplitude of the theta wave is low-medium. It oscillates from adult and normal sleep rhythm. Gamma waves are the fastest brainwave frequency and its frequency range is from 31 to 100 with the smallest amplitude.

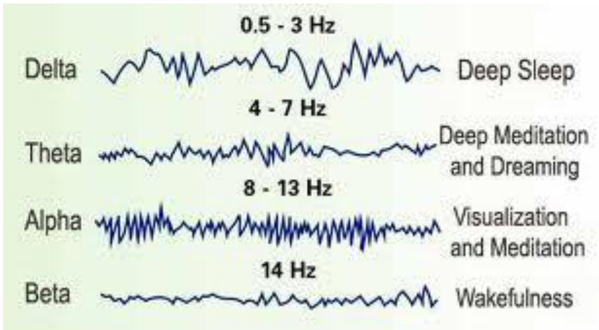

Figure 1 Normal EEG waves

In the proposed work the EEG signals are given as input to the pre processing. From the pre processing the discrete wavelet transform are used to remove noises and the EEG signal are decomposed into five sub-band signals. The non linear parameters (time and frequency) were extracted from each of the six EEG signals (original EEG, delta, theta, alpha, beta and gamma). A genetic algorithm was used to extract the best features from the extracted time and frequency domain features. Then the classifier is used to classify the given EEG signal as normal or abnormal.

2. RELATED WORK
Some literature survey has been focused for the pre-processing of EEG signals, Feature extraction, Feature selection and Classification methods. Siuly [1] has proposed a cross correlation based LS-SVM [4] [6] for improving the classification accuracy of EEG signals. Sabeti M [2] uses the discrete wavelet transform for preprocessing [4] [9] and genetic algorithm, which is used to select the best features from the extracted features. The two classifiers SVM [4] and LDA are used to classify the EEG signal abnormalities. Stevenson N J [3] has developed the automated grading system for EEG abnormality in neonates. Multiple linear discriminant classifier are used to classify the EEG abnormality in neonates with HIE. Marcus [5] has presented the time-frequency distributions of EEG signals. Here the SVM are used to classify the epilepsy from EEG signals.

Nandish M [7] has proposed the classification of EEG signals based on the neural networks. Salih Gunes [8] has discussed that, the Fast Fourier Transform for pre-processing. The combination of KNN and Decision Tree classifiers to classify the EEG signals. Umut Orhan [9] has focused the Multilayar perceptron neural network for EEG signal classification. Parvinnia E [10] has presented the adaptive method named weighted distance nearest neighbour algorithm is applied for EEG signal classification.

ANALYSIS OF EEG SIGNALS

The main goal of proposed work is to analyze the EEG signal for the detection of brain abnormalities. This system involves the process such as EEG signal pre-processing, feature extraction and classification. The modules of the proposed system are:

1. Pre-processing
2. Feature extraction
3. Feature selection
4. Classification

The first module deals with the EEG signal pre-processing method. It is be used to remove the noises from the signal. The next module extracts the EEG signal features from decomposed signal. Then the relevant features are selected from the extracted features. The selected features are given as inputs to the classification process. The classification method is mainly used to analyse the EEG signal and it classifies the signal into normal or abnormal. The Figure 6 shows frame work for analysis of EEG signal. This work is implemented by using MATLAB.

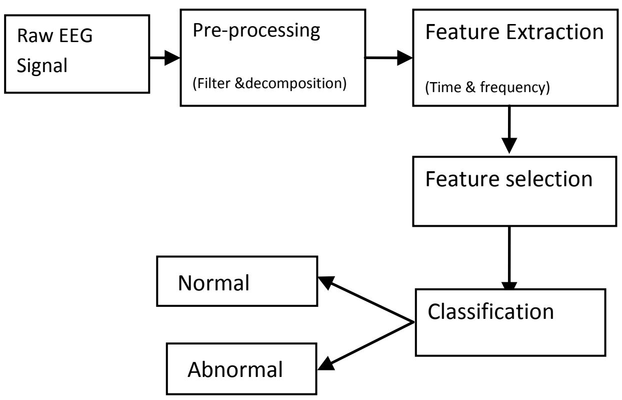

Figure 2 Frame work for analysis of EEG signal

EEG DATABASE

The raw EEG signal is collected from the physionet database. (http://www.physionet.org/cgi-bin/atm/ATM)

3. PROPOSED SYSTEM

3.1 EEG Signal Pre-processing

The raw EEG signal contains some noises that occur due to eye blinking, muscle artifacts and breathing deeply at the testing time. These noises affect the edge function of the EEG signals and the structure of the wave form. The noises are removed by the discrete wavelet transform which decomposes the full-band signal into sub-band signals. The process of the discrete wavelet transform is as follows:

a) The EEG signal is processed with the deubechies wavelet which is used to remove the noises and decompose the signal into sub-bands signals.

b) Based on the frequency range the sub-band signals are separated as delta, theta, alpha, beta and gamma.

c) After the decomposition, the noises are reduced then the Error rate is calculated.

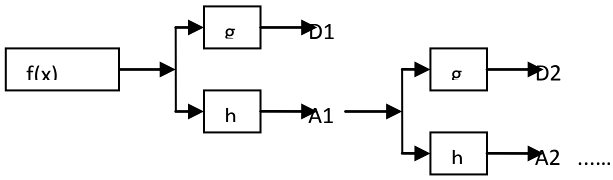

Figure 3 sub-band decomposition of a EEG signal by using Discrete Wavelet Transform

The figure 3 represents the Discrete Wavelet Transform was used to decompose the EEG signal into sub-band signals. The discrete wavelet function splits the signal into detail coefficient (higher level frequency) and approximation coefficient (low level frequency). The approximation coefficient values are chosen because it mainly reduces the noises. After eight level of decomposition, the EEG was decomposed into five EEG sub-bands that approximation to delta (0-4Hz), theta (4-8Hz), alpha (8-15Hz), beta (15-30Hz) and gamma (30-100Hz).

3.2 Feature Extraction

The extraction methods are used to reduce the dimensionality of the features. Extracted features represent the characteristics of original signal without redundancy. The features can be extracted from the EEG signal in two different domains such as Time domain features (TDF) and Frequency domain features (FDF).

3.2.1 Time domain features

Time domain analysis process consists of statistical calculations. The time domain features are: Mean, Median, Mode, Standard deviation, Maximum and Minimum. These time domain features are calculated for the reconstructed EEG signal amplitude and time duration.

i. Mean

Mean corresponds to the centre of a set of value. The Mean is calculated for each and every sub-band signals.

$$Mean = \frac{1}{N}\sum_{i=1}^{n} x_i \qquad (3)$$

ii. Standard deviation

Standard deviation is a simple measure of the variability of a data set. The Standard deviation is the root-mean-square (RMS) deviation of its values from the mean.

$$Std = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x)^2}{N-1}} \qquad (4)$$

iii. Maximum and Minimum

The maximum and minimum values are used to describe the range of observation in the reconstructed signal.

3.2.2 Frequency domain features

The frequency domain features are the power values of each channel from the frequency band. Some of the frequency domain features are Band power, Fractal Dimension and Energy.

Band power describes how the power of a signal or time series is distributed with frequency. Fractal dimension is used to approximate dimension of a signal.

3.3 Feature Selection

Feature selection method is the process to select the relevant features by eliminating features with little or no predictive information. To find a feature subset that produces higher classification accuracy and used to reduce the training time. GA is the process to select the relevant features. GA starts with the initial population of individuals, which represents a possible solution to optimization problems. The evolution process governed by selection, crossover, and mutation rules. The mutation and crossover operators keep the diversity of the population. GA deals with the large search space efficiently.

3.4 Classification

Classification is a data miningtechnique that assigns data in a collection of target categories and classes. The goal of classification is to accurately predict the labels for each class in the data. In the proposed system, the selected features are used as inputs to the k-means classifier, based on the selected features vector the signal are classified as normal and abnormal. It is used to reduce the training time and increases the classifier performance.

<u>Algorithmic steps for k-means clustering</u>

Let X = {$x_1, x_2, x_3, \ldots\ldots, x_n$} be the set of data points and V = {$v_1, v_2, \ldots\ldots, v_c$} be the set of centers.

1) Select *'c'* data points as cluster centers for initialization.
2) Calculate the distance between each data point and cluster centers.
3) Assign the data point to the cluster center whose distance from the cluster center is minimum of all the cluster centers.
4) Update the new cluster center using,
$V_i = \left(\frac{1}{C_i}\right)\sum_{j=1}^{C_i} x_i$(3.10)

where, *'c_i'* represents the number of data points in i^{th} cluster.

5) Recalculate the distance between each data point and the newly obtained cluster centers.
6) Goto step 3 until the cluster centers no longer move.

4 SIMULATION ENVIRONMENTS

The implementation results contain raw EEG signal, EEG signal de-noising process, Feature extraction process and Classification process. The results of each module are given below:

4.1.1 Results of Pre-processing

The raw EEG signal contains some noises that occur due to eye blinking, muscular artifacts and deep breathing at testing time. The low pass filter is used to reduce the noises. It provides a smoother form of a signal removing the short term fluctuations and leaving the longer term trend.

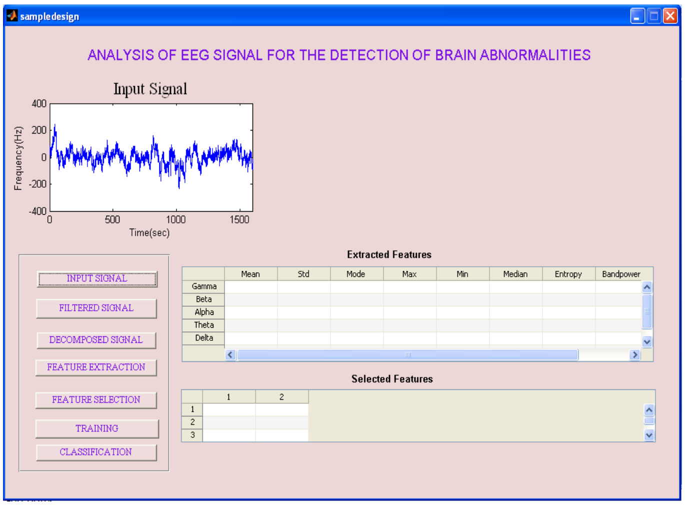

Figure4.1Input EEG signal

Figure 4.1 shows the actual EEG signal. This EEG signal is taken from the Physionet EEG database. The x axis contains the time duration and y axis contains frequency.

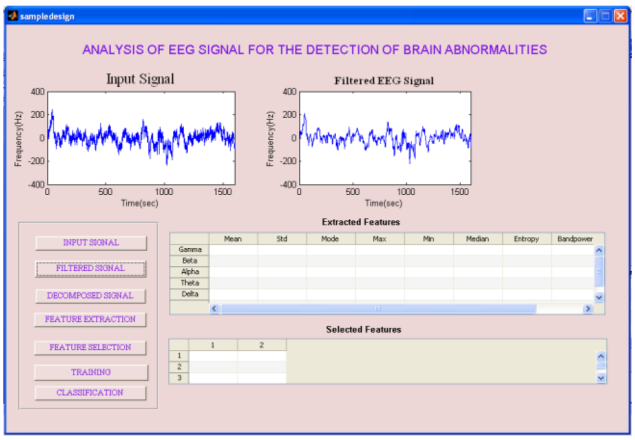

Figure 4.2 De-noised signal

Figure 4.2 shows the de-noised signal. Low-pass filters are used to remove the noises and provide a smoother form of a signal, removing the short-term fluctuations, and leaving the longer-term trend.

The results of de-noised signal are decomposed by discrete wavelet transform. DB8 mainly based on the mother wavelet DWT function. The eight levels of DB8 decomposed approximation coefficient is based on de-noised signal.

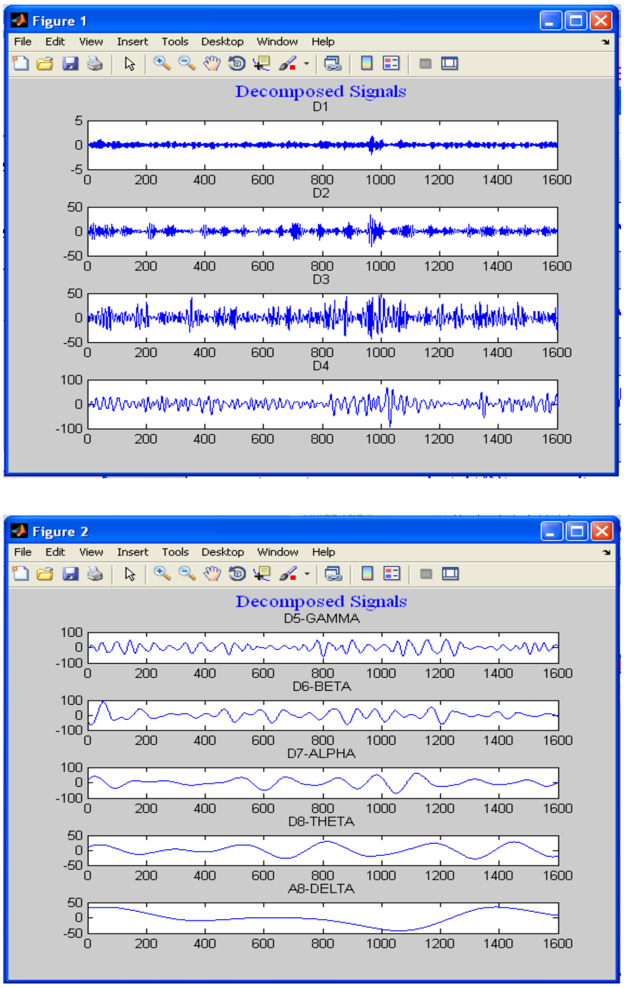

Figure 4.3 Decomposed signals

Figure 4.3 shows the decomposed signal, based on the frequency range, the approximation coefficient isseparated as sub-band signals such as gamma, beta, alpha, theta and delta.

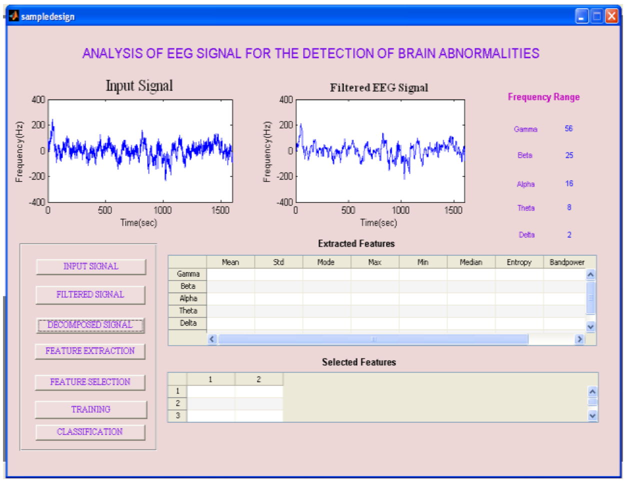

Figure 4.4Frequency range

Figure 4.4 shows the frequency range of the decomposed signals.

4.1.2 Results of Feature Extraction Process

The EEG signals are non-linear, it represents time vs. frequencies. So the Time domain and Frequency domain features are extracted. The EEG signal contains five types of waves such as delta, Theta, Alpha, Beta and Gamma waves. The features can be extracted from the EEG signal in two different domains such as Time Domain Features (TDF) and Frequency Domain Features (FDF).

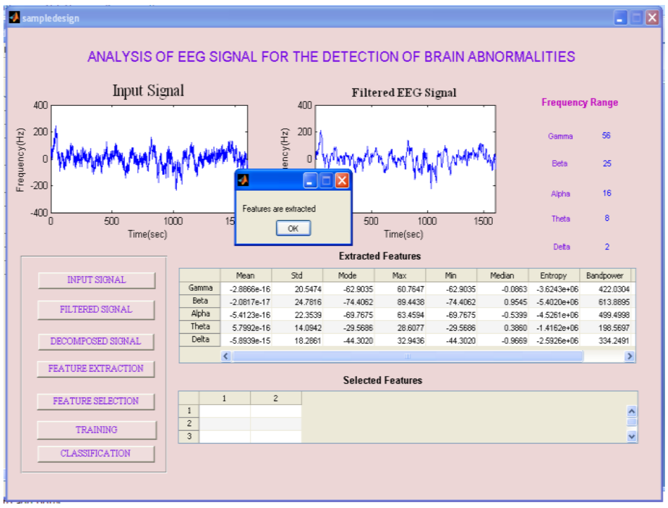

Figure 4.5Extracted Features

Figure 4.5 shows the extracted features from the each decomposed signals. The time domain features such as mean, standard deviation, median, mode, max, min and entropy are extracted. The frequency domain features such as power and energy are extracted.

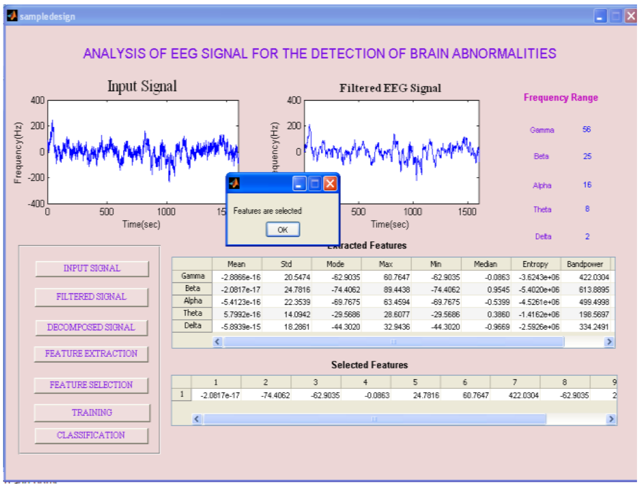

Figure 4.6Selected Features

Figure 4.6 shows the selected features from the extracted features. The alpha band features are selected from the extracted features of all other sub band features.

4.1.3 Results of Classification process

K-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. K-means clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean. First choose the k centers, one for each cluster. The next step is to take each data belonging to a given data set and associate it to the nearest center. Recalculate the distance between each data point and new obtained cluster centers.If no data point was reassigned then stop the process.

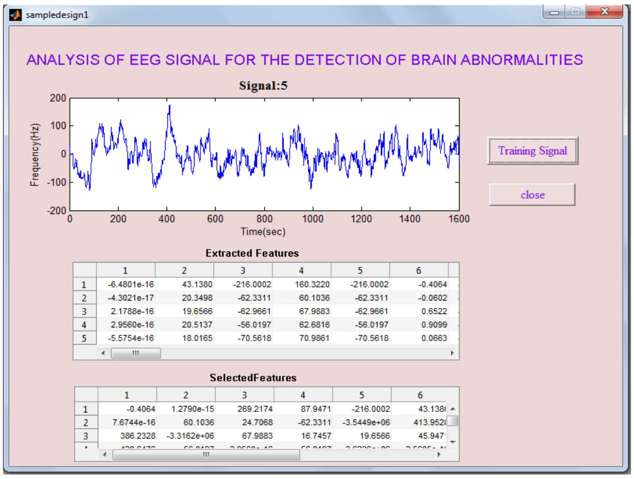

Figure 4.7 Trained Signals

Figure 4.7 shows the features of trained signals. For the classification, 12 dataset is used to the training process. From the trained signals the time and frequency domain features are extracted. Based on the test feature selection, it can select the relevant features from the extracted features.

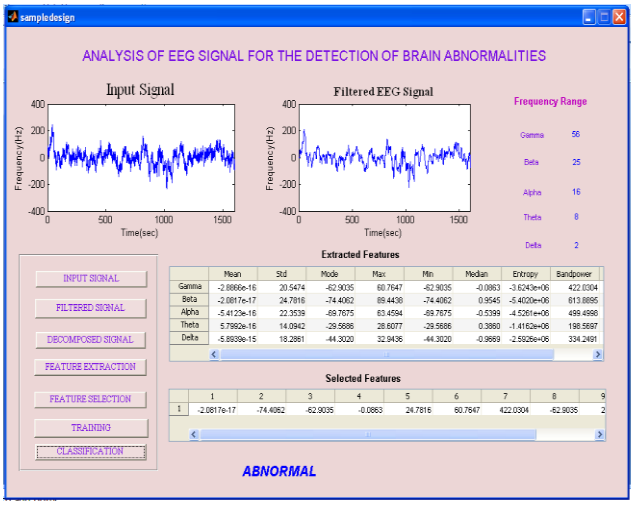

Figure 4.8Classification of Abnormal Signal

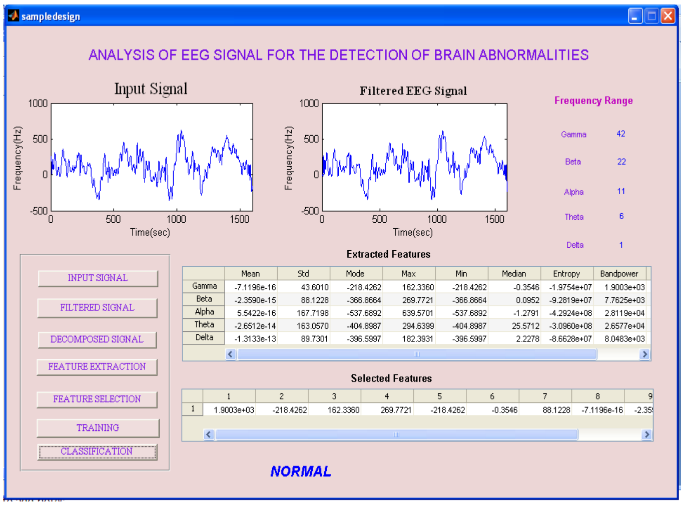

Figure 4.9Classification of Normal Signal

Figure 4.8 and 4.9 show the classification of the EEG signal. The K-means classifier classifies the given signalas normal and abnormal.It can classify the data into two clusters based on the similarity of the data. The two clusters are represented as cluster1 (normal) and cluster2 (abnormal). The advantage of the K-means is Fast and easier to understand. It gives best result when data set are distinct or well separated from each other.

4.2 PERFORMANCE ANALYSIS

The performance of the proposed models was evaluated using the following statistical measure. The accuracy of a classifier on a given test dataset is the percentage of test dataset that are correctly classified by the classifier.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100\%$$

Where,

TP – True Positive (An Abnormal signal is correctly identified)

TN – True Negative (A Normal signal is correctly identified)

FP – False Positive (A Normal signal is incorrectly identified)

FN – False Negative (An Abnormal signal is incorrectly identified)

Table 4.1 summarized the results when each sub band features are applied to classifier as input. In this table each row shows the result of three classifiers. The number of selected features is about 8 in the classifiers.

Table 4.1 shows the highest accuracy, 74% has been achieved when alpha band feature is applied to the classifiers as the input. For better comparison the bar chat of results has been provided in Figure 4.10.

Table 4.1 – Results of classification accuracy for each sub band features

Classifier/Feature	Delta	Theta	Alpha	Beta
KNN	66.6	70	70	66.6
LDA	66.6	70	73.3	70
K-Means	58.3	75	83.3	66.6

According to Fig 4.10 alpha band has highest accuracy in normal and abnormal groups in all classifiers, because the dataset contains mostly depression in alpha bands.

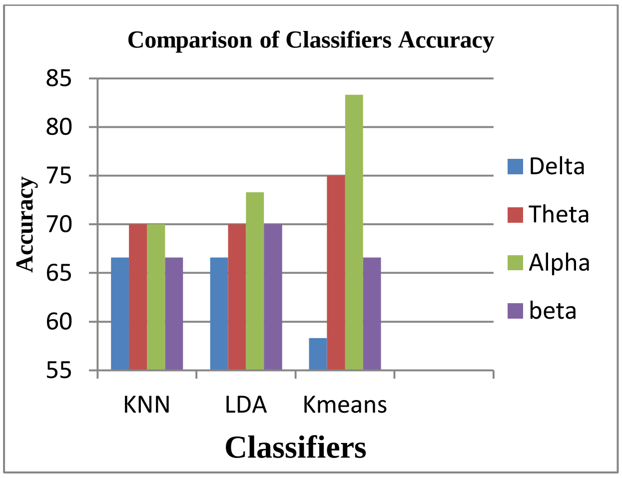

Figure 4.10 Comparison of classifiers Accuracy

4. CONCLUSION

The analysis of EEG signal for the detection of brain abnormalities is a difficult process. So the PC based automatic system is needed for the detection of brain abnormalities.Our proposed work can be a useful tool in studying normal and abnormal patients. The time and frequency domain features were extracted. The K-Means classifier wasused for classification. The alpha band has achieved the highest accuracy to classify the normal and abnormal in the EEG signals.

5. REFERENCES

[1] AbdulhamitSubasi and Ismail M Gursoy (2010)'EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines' Elsevier Transactions on Expert Systems with applications Vol.37pp.8659-8666.

[2] Abdulhamit and Subasi (2005) 'Epileptic Seizure Detection Using Dynamic Wavelet Network' ElsevierTransactions on Expert Systems with Applications Vol.29 pp.343-355.

[3] AbdulhamitSubasi and Ergun Ercelebi (2005) 'Classification of EEG Signals Using Neural Network and Logistic Regression' Elsevier Transactions onComputer Methods and Programs in Biomedicine Vol.78 pp.87-99.

[4] BehshadHosseinifarda, Mohammad Hassan Moradia and Reza Rostamib (2013) 'Classifying Depression Patients and Normal Subjects Using Machine Learning Techniques and Nonlinear Features from EEG Signal', Elsevier Transactions on Computer methods and Programs Vol.109 pp.339-345.

[5] ClodoaldoA.M. Limaa and Andre L.V. Coelho (2011) 'Kernel Machines for Epilepsy Diagnosis via EEG Signal Classification', Elsevier Transactions on Artificial Intelligence in Medicine Vol.53 pp.83-95.

[6] Clodoaldo A.M. Lima, Andre L.V. Coelho andMarcioEisencraft(2010) 'Tackling EEG Signal Classification with Least Squares Support Vector Machines: A Sensitivity Analysis Study', Elsevier Transactions on Computers in Biology and Medicine Vol.40 pp.705-714.

[7] Clodoaldo A.M. Lima, André L.V. Coelho and Sandro Chagas(2009) 'Automatic EEG Signal Classification for Epilepsy Diagnosis With Relevance Vector Machines', Elsevier Transactions on Expert Systems with Applications Vol.36 pp.10054-10059.

[8] Deng Wang, Duoqian Miao and Chen Xie (2011) 'Best Basis-based Wavelet Packet Entropy Feature Extraction and Hierarchical EEG Classification for Epileptic Detection', Elsevier Transactions on Expert systems with Applications Vol.38 pp.14314-14320.

[9] Kai-Cheng Hsu andSung-Nien Yu (2010) 'Detection of Seizures in EEG Using Subband Nonlinear Parameters and Genetic Algorithm', ELSEVIER Transactions on Biology and Medicine Vol.40 pp.823-230.

[10] KhadijehSadatnezhad, Reza Boostani and Ahmad Ghanizadeh (2011)'Classification of BMD and ADHD Patients Using Their EEG Signals', Elsevier Transactions on Expert Systems with ApplicationsVol.38 pp.1956-1963.