Pokhara University

Nepal Engineering College

A Final Year Project Report on

CLASSIFICATION OF EEG SIGNALS EXTRACTED PRIOR TO THE EPILEPTIC SEIZURE TO DETECT ITS ONSET FOR PATIENT SPECIFIC CASE

[BEX 014-11]

Submitted by:

Madhab Prasad Pandey (014-421)

Bhoj Raj Thapa (014-444)

Submitted to:

Department of Electronics and Communication Engineering

Nepal Engineering College

Bhaktapur, Nepal

December, 2018

CLASSIFICATION OF EEG SIGNALS EXTRACTED PRIOR TO THE EPILEPTIC SEIZURE TO DETECT ITS ONSET FOR PATIENT SPECIFIC CASE

by

Madhab Prasad Pandey (014-421)

Bhoj Raj Thapa (014-444)

Under the supervision of

Asst. Prof. Sachin Shrestha

A project report submitted to the

Department of Electronics and Communication Engineering

in partial fulfillment of the requirements for the degree of

Bachelor of Engineering in Electronics and Communication

Department of Electronics and Communication Engineering

Nepal Engineering College

Bhaktapur, Nepal

December, 2018

Nepal Engineering College

Department of Electronics and Communication Engineering

CERTIFICATE

The undersigned certify that they have read and recommended to the Department of Electronics and Communication Engineering for acceptance, a progress project report entitled "Classification of EEG signals extracted prior to the epileptic seizure to detect its onset for patient specific case", submitted by Madhab Prasad Pandey and Bhoj Raj Thapa in partial fulfillment of the requirement for the Bachelor's degree in Electronics and Communication Engineering.

Supervisor:

…………………………

Er. Sachin Shrestha

Assistant Professor

Nepal Engineering College

External Examiner:

…………………………

Er. Rajan Sharma

Senior Engineer

Nepal Telecom

External Examiner:

……………………………

Er. Devendra Kathayat

Deputy Head, Department of Electronics and Communication Engineering

Himalaya College of Engineering

ACKNOWLEDGEMENT

We would like to express our greatest gratitude to the people who have helped and supported us. We are so much grateful to our project supervisor and also the project coordinator, Asst. Prof. Sachin Shrestha and all our respected teachers for supporting and motivating us with much valuable ideas and providing the documents and references to work on this project. We would like to thank our head of department, Assistant Prof. Rupesh Dahi Shrestha for providing this valuable opportunity to do this project on the topic, "Classification of EEG signals extracted prior to the epileptic seizure to detect its onset for patient specific case". The critical explanation about the feasibility of the project suggested by our teacher and friends was booster for the initiation of the project. We appreciate everyone for all the support and direction they have given for our project.

ABSTRACT

The project "Classification of EEG signals extracted prior to the epileptic seizure to detect its onset for patient specific case" is basically focused to detect the epileptic seizures for epileptic people. The mechanism of detection of epileptic seizure are clinically shown to act over seconds to minutes before seizure onset. The unpredictable nature of seizures poses risks for the individual with epilepsy. The automatic detection of oncoming seizures, before their actual onset, can facilitate timely intervention and hence minimize the risks. The early detection of the epileptic seizure is possible based on the various technical parameters. This project presents the concept of wavelet transform (Discrete Wavelet Transform) for the energy calculation of EEG signals coming from the epileptic people and Artificial Neural Network concept for the classification. Frequency transformation of EEG signals has helped for the energy calculation of the EEG signal on both frequency and time domain. For this, we implemented Wavelet Transform on EEG sample signals (obtained from the CHB-MIT database) to transform signals into frequency domain and analyze the change in energy level which helps to detect the pattern of EEG signals before the epilepsy. Then, we used Artificial Neural Network for the classification technique. We performed the process of energy calculation and implementation of DWT, and classification using ANN of EEG sample signals on the platform called MATLAB (Matrix Laboratory).

The proper calculation of the change in energy while and before the onset of the seizure is described in this report. Using FFT, we did the transformation of time domain signal to the frequency signal. Due to the advantages of DWT over FFT, we used DWT for further processing. We calculated frequency domain energy of the first patient's EEG signal samples using DWT and then applied ANN technique for the classification. First patient had 41 files which consists of 1 hour of EEG data in each. In which, 7 files have samples with seizure and remaining files don't have samples with seizure. We implemented DWT for 41 files which consists of 1 hours of EEG data each. After the DWT implementation and classification using Neural Network Tool of MATLAB, we got the 97.5 % accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION

ANN	Artificial Neural Network
CHB-MIT	Children’s Hospital Boston
CT	Computed Tomography
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
EEG	Electroencephalography
EMG	Electromyography
FFT	Fast Fourier Transform
Hz	Hertz
LFP	Local Field Potential
MRA	Multi Resolution Analysis
MRI	Magnetic Resonance Imaging
ROC	Receiver Operating Characteristics
STFT	Short Time Fourier Transform
SVD	Singular Value Decomposition
SVM	Support Vector Machine
WT	Wavelet Transform

CHAPTER 1 INTRODUCTION

1.1 Background and Statement of Problem

Human brain is one of the most complex systems in the universe. Signals are generated as the form of ionic movements in the human brain, those signals are generated differently with different activity of human or external excitation. An Epilepsy is a disease caused as a result of disorder of the nervous system and is unpredictable. Epileptic seizures can be detected from electroencephalogram (EEG) data using computational methods. A set of features can be computed from the EEG data, the features set includes time domain, frequency domain and nonlinear features among them. Significant diagnostic information can be obtained from the frequency distribution of epileptic EEG. Basically the seizure detection is a classification between normal and seizure EEG. There are various signal processing computational methods among them a Wavelet Transform is used, because there was information loss about time domain and gave only spectral information in the frequency domain in Fast Fourier Transform (FFT). And Short Time Fourier Transform (STFT) does not give multi-resolution information of the signals. Wavelet Transform is a type of time-frequency analysis, which provides information about both frequency and time within signals. By dividing the signal into small time interval features can be analyzed, classified.

1.1.1 Epilepsy

Epilepsy is a common brain disorder that, according to estimate of the world health organization, affects almost 60 million people around the world. Nearly 80% of the people with epilepsy are found in developing regions. Epilepsy responds to treatment about 70% of the time, yet three fourth of effected people in the developing countries do not get the treatment they need. Approximately one in every 100 persons will experience a seizure at some time in their life. Epilepsy is characterized by the recurrent and sudden incidence of epileptic seizures which can lead to dangerous and possibly life threating situations. The seizures are the result of a transient and unexpected electrical disturbance of the brain and excessive neuronal discharge that is evident in the electroencephalogram (EEG) representative of the electrical activity of the brain. The causes of epilepsy are unknown and are undetectable to long periods. In epilepsy, seizures tend to recur, and have no immediate underlying cause while seizures that occur due to a specific cause are not deemed to represent epilepsy. Since disturbance of consciousness and sudden loss of motor control often occur without any warning, the ability to predict epileptic seizures would reduce patient's anxiety, thus improving quality of life and safety considerably. In this light in the absence of completely controlling a patient's epilepsy, seizure prediction is an important aim of clinical management and treatment. The epilepsy seizures cause medical treatments complex and making patient life not secure due to its unexpected occurrence. The onset detection can be implement on monitoring system, thus can be continuously monitor.

1.1.2 EEG (Electroencephalography)

Our brain is made up of hundreds of billions of brain cells called neurons. Neurons have axons that are neuro transmitters and dendrites that receives the signals. Dendrites of the neuron receives the neuro transmitters from the axons of the other neurons that causes the electrical polarity change inside of the neuron. This polarity change is what the EEG records. It's the post-synaptic dendritic currents from the cortical pyramidal cells. The activity of one single neuron is way too small to detect using EEG equipment but when thousands of neurons work in group for high potential, an area of the neurons work in groups is called LFP (Local Field Potential). EEG signals are typically measure in microvolts. More details about EEG testing is on. There are five types of EEG wavelengths which are determined on the basis of frequency oscillations. These are:

- Delta: 1-3 Hz
- Theta: 4-7 Hz
- Alpha: 8-12 Hz
- Beta: 13-30 Hz
- Gamma: 31-50 Hz

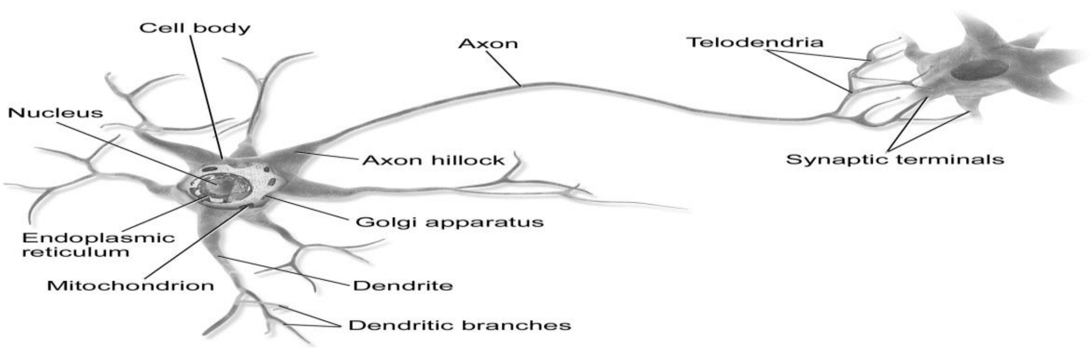

Figure 1.1 Anatomy of multipolar neuron (Adapted from https://en.wikipedia.org/wiki/Neuron#/media/File:Blausen_0657_MultipolarNeuron.png)

Electroencephalography signal is the recording of electrical activity of brain that provides valuable information of the brain function and neurological disorder. In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time as recorded from multiple electrodes placed on the scalp. In neurology, the main diagnostic application of EEG is in the case of epilepsy, as epileptic activity can create clear abnormalities on a standard EEG study. A secondary clinical use of EEG is in the diagnosis of coma encephalopathy, and brain death. EEG used to be a first-line method for the diagnosis of tumors, stroke and other focal brain disorder, but this use has decreased with the advent of anatomical imaging techniques such as MRI and CT.

EEG can be used in intensive care units for brain function monitoring:

- To monitor for non-convulsive seizures/non-convulsive status epileptics.
- To monitor the effect of sedative/anesthesia in patients in medically induced coma (for treatment of refractory seizures or increased intracranial pressure).
- To monitor secondary brain damage in conditions such as subarachnoid hemorrhage (currently a research method).
- When someone is in a coma, an EEG may be performed to determine the level of brain activity.

If a patient with epilepsy is being considered for respective surgery, it is often necessary to localize the focus (source) of the epileptic brain activity with a resolution greater than what is provided by scalp EEG.

1.2 Objective and Scopes

With the help of different EEG sample signals from CHB-MIT database, we will try to meet the following objectives:

1. To implement DWT to EEG signals.
2. To classify EEG signal using Artificial Neural Network (ANN) prior to epileptic seizures based on the extracted features.

The recorded EEG signal prior to epilepsy seizure can classify and extract the features based on the several mathematical computations such as Fast Fourier Transform (FFT), Wavelet

Transform (WT) etc. We will use Wavelet Transform (WT), and automatic epileptic seizure detection can be implemented by developing the algorithms.

1.3 Applications

The cause of epilepsy is unknown, occurs unpredictably and irregularly. Patient can suffering from life threating conditions. Thus early detection of epileptic seizures will be much more helpful for various purpose. Some of the applications of prior detection of epileptic seizures are mentioned as follows:

- Onset detection of epileptic seizures can help to protect the unfavorable conditions such as life threating conditions.
- Onset detection of epileptic seizures can help for effectiveness of medical treatments, thus it is helpful for the doctor.
- Accurate detection and logging of seizures can be used to improve the diagnostic yield from patient monitoring during epilepsy surgery evaluation and to improve understanding of epilepsy as a dynamical disease.
- The calculation of energy can help the analysis and classification of the signal.
- Frequency transformation of time signal gives us the information about frequency components present in the signal from which processes like filtering, etc. can be done.

1.4 Overview of the Report

This report mainly highlights about the problem of epileptic seizure in the present world with the statistical data from the calculation of energy. The basic idea about the implementation of energy calculation method of time domain using energy calculation formula for time domain and frequency domain signal can give the direction to solve the problem of epileptic seizure after the analysis of the EEG signal. The introduction part gives the explanation about motivation of doing this project and objectives of our project.

Second chapter corresponds to the study about subject we have done for our project. And third chapter explains methods of energy calculation and classification of EEG signals. It also includes block diagram of our project and explanation. The works that are done to meet the objective of our project is presented in the fourth chapter.

CHAPTER 2 LITERATURE REVIEW

This project includes signal processing for the extraction of the features for the detection of epileptic seizure prior to the onset. There are many methods for the signal processing that have been used until now. Transformation of the time domain signal to the frequency signal and analyzing those signals is also a very important part of this project. Classification, on the other hand, is a critical part of the detection.

2.1 Fourier Transform

Signal processing refers to the use and implementation of the signals and Fourier transform is the first thing we would want to use for that. Fourier transform decomposes the function of time into the frequencies that make it up, in a way similar to how a musical chord can be expressed as the frequency of its constituent notes. Basically any time domain signal can be broken in a collection of sinusoidal signals. In this way, lengthy and noisy EEG recording can be conveniently plotted in a frequency power spectrum. By doing so, hidden features can become apparent. The subject of statistical signal processing does not usually apply the Fourier transformation to the signal itself. Even if a real signal is indeed transient, it has been found in practice advisable to modal a signal by a function which is stationary in the sense that its characteristic properties are constant over all time. But there is no time information in Fourier transformed signal and no frequency information in time domain single. In simple words, FT does not describe in which time certain frequency exists. This might lead us to the fact that use of FT in signal processing of FFT is less as FT of two entirely different signals can be similar.

2.2 Short Time Fourier Transform

The short-time Fourier transform (STFT) is a Fourier-related transform used to determine the sinusoidal frequency and phase content of local sections of a signals as it changes over time. The main procedure of computing STFTs is to divide a longer time signal into shorter segments of equal length and then compute the Fourier transform separately on each shorter segments.

STFT had a fixed segment where non stationary waves are considered stationary and hence different frequencies were resolved equally in STFT. This resolving gives rise to what we call the resolution problem. The Heisenberg's Uncertainty principle can be applied to time frequency information of the signal. One cannot know what spectral components exist at what instance of times but what one can know are the time intervals in which certain band of frequencies exist. This is called the resolution problem. The size of window where non-stationary signals are considered stationary now defines the better resolution for either time or frequency information of the signal. For a narrow window the time resolution is better and the frequency resolution is poor. Similarly, the wide window has a good frequency resolution and poor time resolution. That's why STFT is less used for this project.

2.3 Wavelet Transform

The wavelet transform is similar to the Fourier transform (or much more to the windowed Fourier transform) with a completely different merit function. The main difference is: Fourier transform decomposes the signal into sines and cosines, in contrary the wavelet transform uses function that are localized in both the real and Fourier space. Generally wavelet transform can be expressed by the following equation:

$$F(a,b) = \int_{-\infty}^{\infty} f(x)\varphi^{*}_{(a,b)}\ x\, dx \qquad (1)$$

W.T gives a complete three dimensional information about any signals i.e. what different frequency components are present in any signal, what are their respective amplitudes and at time axis where these different frequency components exist. More details about wavelet is on [1].

MRA (Multi-resolution Analysis) is the design method of most of the practically relevant discrete wavelet transforms and the justification for the algorithm of the fast wavelet transform. MRA which is employed in WT is designed to give good time resolution and poor frequency resolution at high frequency and good frequency resolution and poor time resolution at low frequencies. This approach makes sense especially when the signal has high frequencies for short durations and low frequencies for short duration and low frequency components for long duration. EEG signals have this same feature and hence WT is well suited for the processing of EEG samples.

In continuous wavelet transform (CWT), a given signal of finite energy is projected on a continuous family of frequency bands. [2] The admissibility condition is used to first analyze and then reconstruct a signal without loss of information. Discrete wavelet transform is just another form of representing the signal and does not change the information content in the signal. The wavelet series is simply a sampled version of the continuous wavelet transform and the information provided by it is highly redundant as far as the reconstruction of the signal is concerned.

2.4 Artificial Neural Network

One of the classifiers for the classification of EEG signal as normal or epileptic is an artificial neural network. [3] It is an interconnected group of nodes. In machine learning, ANNs are family of statistical learning algorithms. It is an electronic model based on the neural structure of the brain. Basically, AANs are computing systems made up of large number of filmy interconnected adaptive processing elements (neurons). These elements perform massively parallel computations for data processing and knowledge representation. Learning in ANNs is accomplished through special training algorithm. ANNs can be trained to recognize the non-linear models and patterns develop during training. ANNs have evolved as a powerful tool for classification, pattern recognition, prediction as well as pattern completion. It is an inspiration from the biological neurons. [10] Neural network is not an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs.

Neural network [9] is a compact interconnection of billions of neuron with trillion of interconnections between them. ANNs are the biologically inspired simulations performed on the computer to perform specific tasks like clustering, classification, pattern recognition etc. ANNs is biologically inspired network of artificial neurons configured to perform specific tasks. It acquired knowledge through learning.

For a basic idea of how a deep learning neural network [11] learns, imagine a factory line. After the raw materials (the data set) are input, they are then passed down the conveyer belt, with each subsequent stop or layer extracting a different set of high-level features. If the network is intended to recognize an object, the first layer might analyze the brightness of its pixels. The next layer could then identify any edges in the image, based on lines of similar pixels. After this another layer may recognize textures and shapes, and so on. By the time fourth or fifth layer is reached, the deep learning net will have created complex feature detectors. It can figure out that certain image elements (such as a pair of eyes, a nose, and a mouth) are commonly found together. Once this is done, the researchers who have trained the network can give labels to the output, and then use back-propagation to correct any mistakes which have been made.

After a while, the network can carry out its own classification tasks without needing humans to help every time.

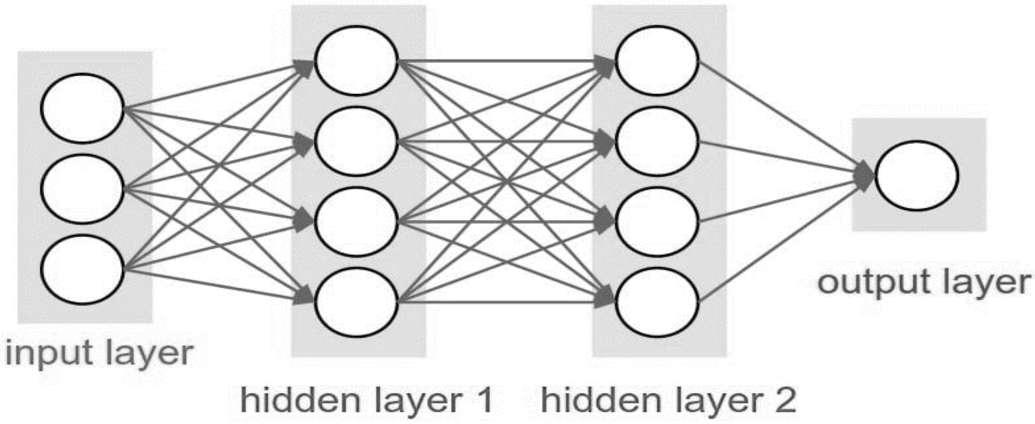

Figure 2. 2 Artificial Neural Network (Source: https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/)

There are multiple types of neural network, each of which come with their own specific cases and levels of complexity. The most basic type of neural network is called feed-forward neural network, in which information travels in only one direction from input to output. A more widely used type of network is the recurrent neural network, in which data can flow in multiple directions. These neural networks possess greater learning abilities and are widely employed for more complex tasks such as learning handwriting or language recognition. There are also convolutional, Boltzmann machine, Hopfield networks, and a variety of others.

2.5 Other Works

The other methods used for the classification and detection of EEG signals are as follows:

1. B. Al-Bokhity,Dalia Nashat and T.M. Nazmy [3] interpreted the method of epileptic seizure onset detection using the soft computing techniques such as short time Fourier transform for de-noising the signal throughout computing the STFT of the noisy signal. The EEG signal were classified by using the Neural Network classifier.
2. Milica Milosevic [4] had done one automated detection of epileptic seizures in pediatric patients based on accelerometry and surface electromyography. On the basis of her report feature extracted from surface EMG signals can be divided in three big groups: time domain features, frequency features and entropy features. Many features are the same features as features calculated from ACM signals.

3. The project done by Aditya Sundar and Chinmay Das [5] was based on the wavelet based signal processing technique used for the feature extraction. Features like sample entropy, approximate entropy, recurrence quantifiers, wavelet energy, etc. of the main signals as well the decomposed signals were extracted. Many different classifying techniques were used. The random forest algorithm was found to give the most accurate results. Scalograms, time frequency plots, box-plots, ROC curves were among the other tools used to extract information from the EEG signals.
4. Ujjwal Kc, Ghanshyam Kshetri, and Saroj Koirala [6] had done the project on epileptic seizure onset detection which explores the ability of energy calculation using DWT to detect the onset. From their project, it was concluded that the epileptic seizures have higher energy. Our project has the extensive objective of this project.

CHAPTER 3 METHODOLOGY

The pre-processing of EEG signals using FFT and calculation of associated energy for the classification of the EEG data on the basis of supervised learning will be used. There were multiple methods used for the detection of onset of the epileptic seizure in the past. But we used the method represented by the block diagram of figure 3.1. We will mainly focus on the EEG signals in which onset of the epilepsy is present and analyze and process for the time domain EEG signals as well as frequency domain. Our main task would be the processing of those signals under the circumstances of some transform methods like FT, STFT, Discrete Wavelet Transform, etc. When we extract features from the high frequency component using DWT, we take total length of the high frequency components. We investigate the strength of each transformation and decomposition for the capability of classification and accuracy as well as the combined strength through different feature extractions. The classification will leave to the result of epileptic or non-epileptic EEG signals.

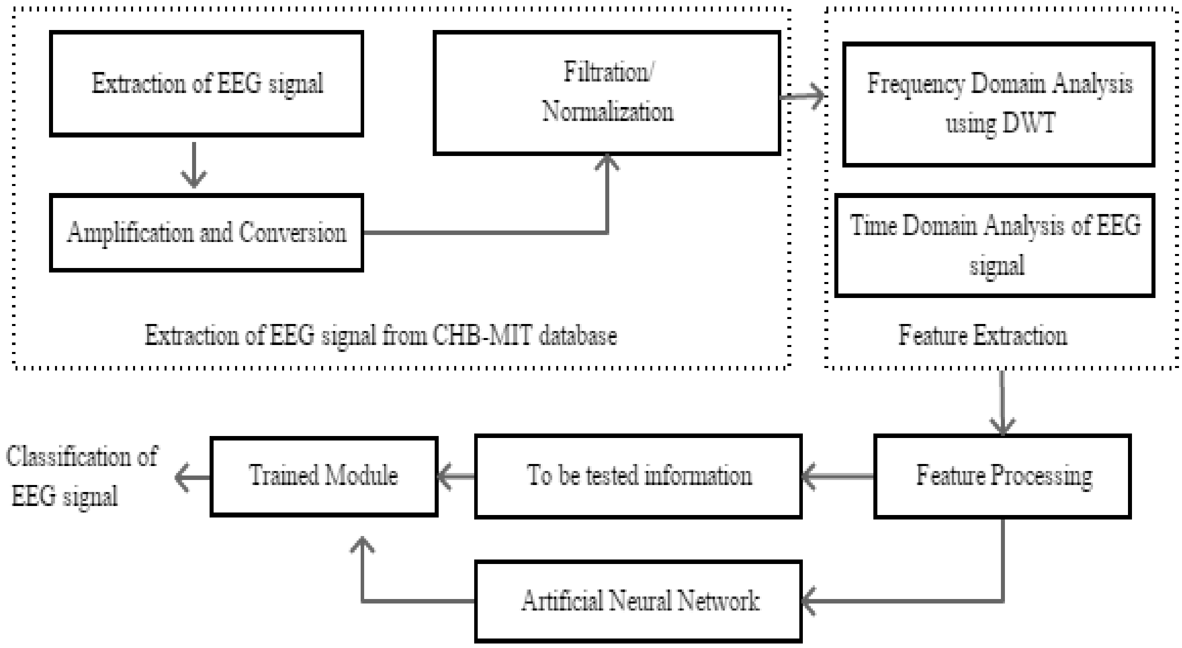

Figure 3.1 Block Diagram of proposed scheme for classification of EEG signal

3.1 Pre-processing of the EEG Signal

This step is exploited in order to eliminate the influence of disturbance/noise of EEG signal. The noises can be divided into two parts; one is physiological which is raised from the body and another is non-physiological artifact that comes environment and instruments. There is several type of physiological noises as muscle noise, pulse noise and eye blinking. The non-physiological noises are power line and sweat noises. Filter can be used to remove noise interference which is also called normalization but in our case the EEG signals from CHB-MIT are already normalized and preprocessing is done. Normalization process is necessary to standardize all the features to the same level. This makes the calculation simpler. Amplification is done using amplifier for the better analysis of the small amplitude signals.

3.2 Wavelet Transform

Abnormalities in EEG data during serious neurological diseases such as epilepsy are too subtle to detect using conventional techniques. The techniques that have been applied to address this problem include the analysis of EEG signals for the detection of epileptic seizures using the autocorrelation function, time domain features, frequency domain features, time frequency analysis, nonlinear time series analysis, and wavelet transform. However, the result of various studies have demonstrated that the wavelet transform is the most promising method for the extracting features from the EEG signals. The adaptive time-scale representation and decomposition of signal into different frequency sub-bands presents an efficient signal analysis method without introducing a calculation burden.

Instead of working on a single scale (time or frequency), it can work on multi-scale basis. Therefore, multi-scale feature of WT allows the decomposition of signal into number of scales. In analysis of signals using DWT, is very important to choose a suitable wavelet and a number of levels of decomposition. Based on dominant frequency components of the signal the number of levels of decomposition is chosen. The procedure of multi-resolution decomposition of the signal x[n] is shown in Figure 3.2.

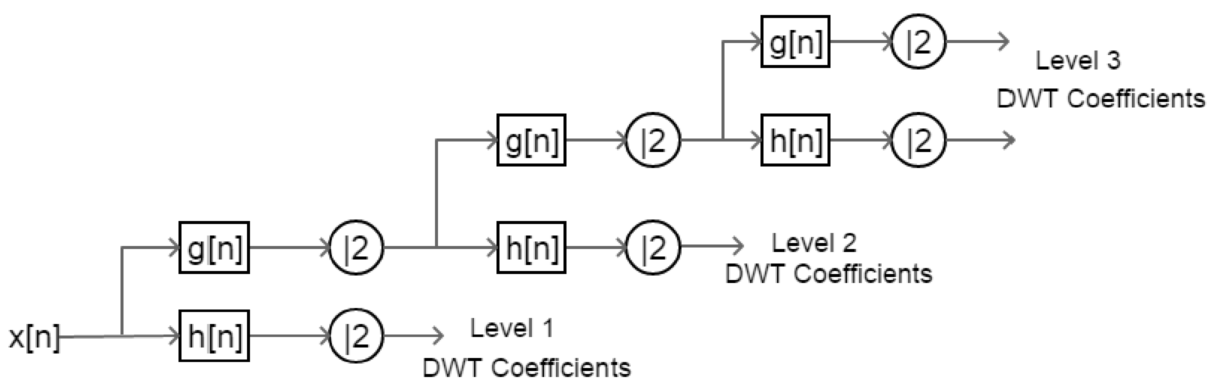

Figure 3.2 Sub Band Coding Algorithm for DWT implementation

Use of DWT over other transforms are as follows:

1. DWT gives time and frequency information.
2. Flexible: DFT is just based on cosine and sine of the function whereas DWT has different types of bases.
3. DWT can reconstruct complex and non-stationary (signal where is change in in the properties of signals) signals as well.
4. Wavelets are far better to de-noise the signal than conventional transforms like FFT, FT etc.
5. The loss is less while reconstructing the signal.
6. DWT is more of decomposition rather than filtering.

3.2.1 Daubechies 8 Discrete Wavelet Transform

The most commonly used set of discrete wavelet transforms was formulated by the Belgian mathematician Ingrid Daubechies in 1998. This formulation is based on the use of recurrence relations to generate progressively finer discrete sampling of an implicit mother wavelet function; each resolution is twice that of the previous scale. We are using Daubechies 8 discrete wavelet transform for our project. Properties of db8 are asymmetric, orthogonal and biorthogonal. The wavelet and scaling function of db8 are as follows:

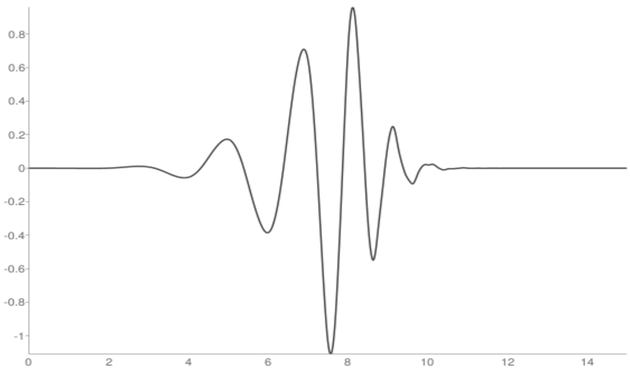

Figure 3.3 Wavelet function (Source: http://wavelets.pybytes.com/wavelet/db8/)

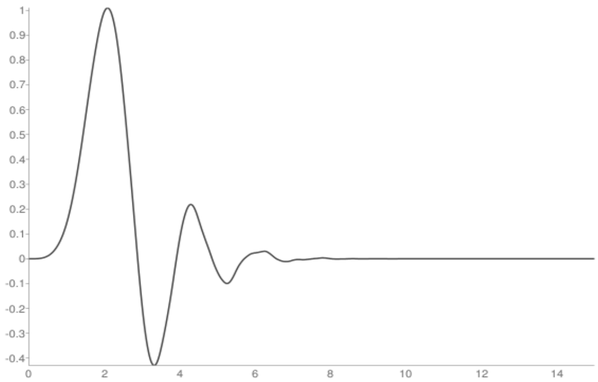

Figure 3.4 Scaling function (Source: http://wavelets.pybytes.com/wavelet/db8/)

3.3 Feature Extraction

For the classification purpose, relevant features needed to be extracted from EEG signals. Computational burden should be reduced and false positive rate should be lowered for sensitivity while preserving high prediction ability. Discrete wavelet transform may be a good technique for the feature extraction.

Then the frequency domain analysis is done using listed parameters:

Mean [6]: It is the arithmetic mean of the calculated energies of detailed coefficients for different frequency ranges.

$$\text{Mean} = \frac{\sum x}{n} \tag{4}$$

Where, x = energy values

N=total number of samples

Variance [6]: Variance measures the variability of the data values and is calculated as:

$$\text{VAR} = \frac{1}{N-1}\sum_{n=1}^{N} xn^2 \tag{5}$$

Where, N = total number of samples

xn =energy values for nth valued energy

Power [6]: It is the measure of average power any desired frequency range.

Log Detector [6]: It is a statistical parameter which is calculated as:

$$\text{L} = e^{\frac{1}{N}\sum_{n=1}^{N}\log(|xn|)} \tag{6}$$

Where, N= total number of samples

xn = energy values for nth valued energy

Energy calculation is also done using DWT method. In similar way, sample entropy and approximate are also some of the features of EEG signal that can be calculated.

3.3.1 Parseval's Theorem

In mathematics, Parseval's theorem usually refers to the result that Fourier transform is unitary that the sum of the square of a function is equal to the sum of the square of its transform. [13]

It originates from a 1799 theorem about series by Marc-Antoine Parseval, which was later applied to the Fourier series. Parseval's theorem is used to evaluate the energy carried by a signal both in time and frequency domain, which is basically energy conservation.

Theorem: For any $x \in C^N$

$$|x|^2 = \frac{1}{N}|X|^2 \qquad (2)$$

I.e.,

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^2 \qquad (3)$$

3.4 Classification and Detection

After the feature extraction process of the EEG signal extracted from CHB-MIT database, the most prior mechanism is to interpret those features and selection of the features. The extracted features are then analyzed and classified on the basis of supervised learning classification technique. To extract the features using ANN, we reshape energy values of the EEG signals into a two dimensional matrix. We make an input matrix for ANN using frequency transformed energy values. Then we will train the machine as per our features to detect the unknown EEG signal if that is epileptic or non-epileptic. So, training and testing will be our summarized process of the project which will be done using Neural Network Toolbox of MATLAB.

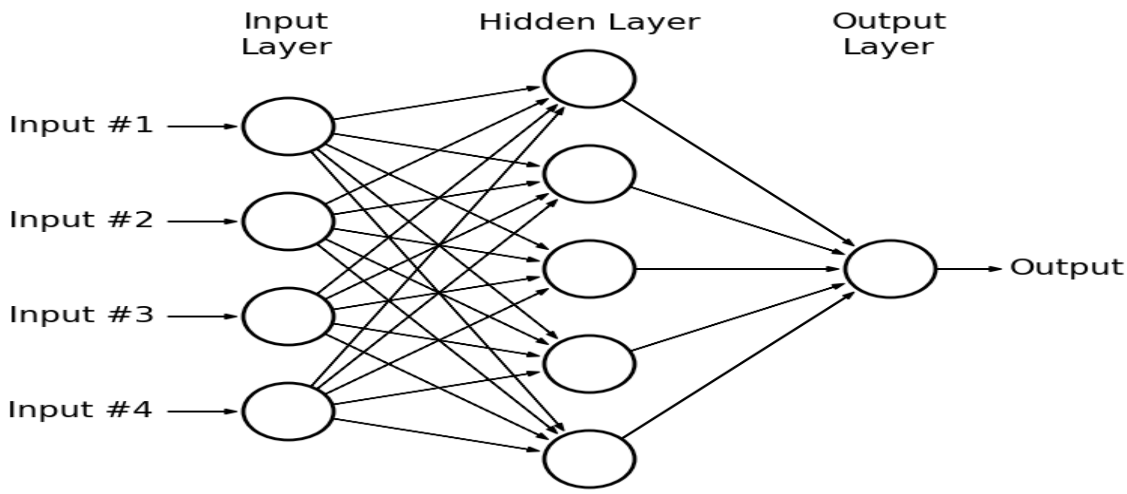

Figure 3. 5 Artificial Neural Network (Source: https://github.com/Aerolyzer/Aerolyzer/wiki/Neural-Network-Logan)

As in the figure 3.5, ANN takes large number of inputs as a matrix and then that goes into a lot of hidden layers which will approximate or train the network for proper output. And gives single output i.e. epileptic or non-epileptic.

We used built in Neural Network Toolbox using command line of MATLAB. We used pattern recognition and classification app wizard of NTT which uses two-layer feed-forward algorithm to train, test and validate the network. Our main objective is to train, validate and test the input data to the targeted data so that a pattern can be created. The Neural Pattern Recognition app will help you select data, create and train a network, and evaluate its performance using cross-entropy and confusion matrices.

Algorithms used for ANN are:

Data Division: Random i.e. it randomly divides the input data for testing, validation and training.

Training: Scaled Conjugate Gradient [15] is the supervised learning algorithm for feed-forward neural networks, and is a member of the class of conjugate gradient method.

Performance: Cross-Entropy i.e. it calculates a network performance given targets and outputs, with optional performance weights and other parameters. Minimizing cross entropy leads to good classifiers.

Calculations: MEX (Minimum Excluded) is the smallest value from the whole set that does not belong to the subset.

3.5 MATLAB

MATLAB stands for matrix laboratory which is a multi-paradigm numerical computing environment which allows matrix manipulation, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages including C, C++, C#, Java, Fortran, Python. We will be using MATLAB's built in function for transformations like FFT, DWT etc. We will use MATLAB (R2016a) for our project. We used MATLAB because of the availability of the functions that can be used for the calculation and manipulation of the non-linear EEG signals.

CHAPTER 4 RESULTS

This project was done by diving into the different sub-tasks which will eventually be summed up and give the required target of the project. This project starts from the collection of EEG samples from the CHB-MIT database for the development of the system which has the EEG information to detect the onset of epileptic seizures.

4.1 Collection of EEG Samples from the CHB-MIT Database

In this project, we are classifying the EEG signal extracted prior to the epileptic seizure to detect its onset for patient specific case. The CHB-MIT database, collected at the Children's Hospital Boston, consists of EEG recordings from pediatric subjects with intractable seizures. We are not extracting the EEG signal from epileptic people. Instead, we are using the EEG signal samples from the CHB-MIT database. This will save our time to do the preprocessing part of the EEG signal.

We can extract the signal from the website as per our requirement. In the website [12] there are basically recordings, grouped into 23 cases, were collected from 22 subjects (5 males, ages 3-22; and 17 females, ages 1.5-19). (Case chb21 was obtained 1.5 years after case chb01, from the same female subject.) The file subject-info contains the gender and age of each subject. (Case chb24 was added to this collection in December 2010, and is not currently included in subject-info.) Each case (chb01, chb02, etc.) contains between 9 and 42 continuous .edf files from a single subject. Hardware limitations in gaps between consecutively-numbered .edf files, during which the signals were not recorded; in most cases, the gaps are 10 seconds or less, but occasionally there are much longer gaps. Dates in the original files have been replaced by surrogate dates, but the time relationship between the individual files belonging to each case have been preserved. In most cases, .edf files contain exactly one hour of digitized EEG signals, although those belonging to case chb10 are two hours long, and those belonging to cases chb04, chb06, chb07, chb08 and chb23 are four hours long; occasionally, files in which seizures are recorded shorter.

We can extract the signal in .mat file format because we are using the MATLAB for processing the signal. All the related information was carried out while formatting the database of EEG signals. The signal has the information about the starting and ending of the epileptic seizure. We will be using only a sample signal for now to analyze the features. We are using the information of a sample signal chb01_03 for the normal signal processing and the information are in tabular form below in table 4.1:

Table 4.1 Basic Information about EEG sample from CHB-MIT

Source: record chbmit/chb01/chb01_03.edf	
Sampling Frequency	256 Hz
Sampling Interval	0.00390626 second
Duration	1 hour
File Start Time	13:43:04
Seizure Start Time	2996 seconds
Seizure End Time	3036 seconds
Signal	FP1-F7

Above signal information like sampling frequency, sampling interval, gain is almost similar to the other information. Duration, file start time, seizure start time, seizure end time and electrode can be different for different cases and different processing of signal.

Table 4.2 Subject-Info of CHB-MIT database

S.N.	Case	Gender	Age (Years)
1	chb01	Female	11
2	chb02	Male	11
3	chb03	Female	14
4	Chb04	Male	22
5	Chb05	Female	7
6	Chb06	Female	1.5
7	Chb07	Female	14.5
8	Chb08	Male	3.5
9	Chb09	Female	10
10	Chb10	Male	3
11	Chb11	Female	12
12	Chb12	Female	2
13	Chb13	Female	3
14	Chb14	Female	9
15	Chb15	Male	16

16	Chb16	Female	7
17	Chb17	Female	12
18	Chb18	Female	18
19	Chb19	Female	19
20	Chb20	Female	6
21	Chb21	Female	13
22	Chb22	Female	9
23	Chb23	Female	6

The table 4.2 contains the information of every patient in terms of age and gender. We have used almost all the signal for the processing and analysis of the EEG signals during our project.

4.2 Energy Calculation of Time and Frequency Domain of EEG signal using Fast Fourier Transform (FFT)

We calculated time domain energy of the EEG signal and for the frequency domain energy, we used Fast Fourier Transform (FFT) method. We first transformed the signal into frequency domain and then we classified the whole signal into different frequency band as per the requirement. Then we applied the formula for both calculations (time domain as well as frequency domain):

$$\text{Energy} = \frac{\sum_{n=1}^{N} X(n)^2}{N} \qquad (7)$$

Where N= number of samples

We did the following processes during implementation of FFT:

1. Testing of Sinusoidal wave of frequency 20 Hz, sampling frequency of 256 Hz, and Amplitude of 1 volt using MATLAB function getEnergyOfBand which gives 7*1 matrix containing energy values of 6 bands and time domain signal respectively.
2. We used 1024 point FFT and shifting technique with the window size of 256 samples and shifting of 128 samples.
3. Energy calculation of epileptic and non-epileptic EEG signals which gave us the result that there is clear difference between energy plots of both the signals.

4.3 Energy Calculation of EEG Signal using Discrete Wavelet Transform (DWT)

For the detection of onset of epileptic seizures, we will be using the concept of associated energy along with frequency analysis through DWT. We are using DWT because of its advantages over Fast Fourier Transform. We have been writing codes to implement DWT using the wavelet called Daubechies 8 so that EEG samples can be decomposed into wavelets for frequency spectrum analysis which could be further coupled with associated energy concept.

DWT actually decomposes original signal into two different frequency groups for 1st level decomposition. For our EEG signals, it has 256 sampling frequency which means it has maximum of 128 Hz frequency component. Applying DWT in the EEG signal gives two different component called approximation coefficients and details coefficients. Approximation coefficient is found from passing the original signal into low pass filter and detail coefficients is found from passing the original signal into high pass filter. Once, the signal x(t), extracted EEG signal is passed through DWT method, it will give us two coefficients which are decomposed into two frequency groups which are 0-64 Hz and 64 to 128 Hz. Further, 0-64 Hz decomposed signal is decomposed into two frequency groups which are 0-32 Hz and 32-64 Hz. In the same way, the decomposition of signal is done as per the requirement of frequency band. In our case, we do 5 level decomposition to get 6 frequency bands.

We have made a function called energyDWT in MATLAB which decomposes the given signal into 5 levels. We used Daubechies 8 wavelet for the DWT decomposition. And the output of the signal is calculated in the form of matrix of 6 columns and 1 row which gives the following information:

Table 4.3 Output matrix of the function energyDWT

Energy0_4	Energy4_8	Energy8_16	Energy16_32	Energy32_64	Energy64_128

The decomposition is done in following way:

1. 1st level decomposition: 0-128 Hz signal is decomposed into 0-64 Hz and 64-128 Hz
2. 2nd level decomposition: 0-64 Hz signal is decomposed into 0-32 Hz and 32-64 Hz
3. 3rd level decomposition: 0-32 Hz single is decomposed into 0-16 Hz and 16-32 Hz

4. 4^{th} level decomposition: 0-16 Hz signal is decomposed into 0-8 Hz and 8 to 16 Hz
5. 5^{th} level decomposition: 0-8 Hz is decomposed into 0-4 Hz and 4-8 Hz

With every level of decomposition, number of samples in frequency band decreases by half. After decomposition, we applied the following formula to calculate the energy of every required decomposed frequency bands:

$$\text{Energy} = \frac{\sum_{n=1}^{N} X(n)^2}{N} \qquad (7)$$

Where N= number of samples in the signal

4.3.1 Testing of DWT Method for Energy Calculation

Implementation of DWT using Daubechies 8 wavelet is done in sinusoidal signal, to make sure that the method will give the right result. We used a sinusoidal signal of following properties:

1. Sampling frequency (fs) = 256
2. Frequency (f) = 20
3. Expression = sin(2*pi*f*n) where f is frequency, n is number of samples which runs from 1/fs to 1 with the gap of 1/fs

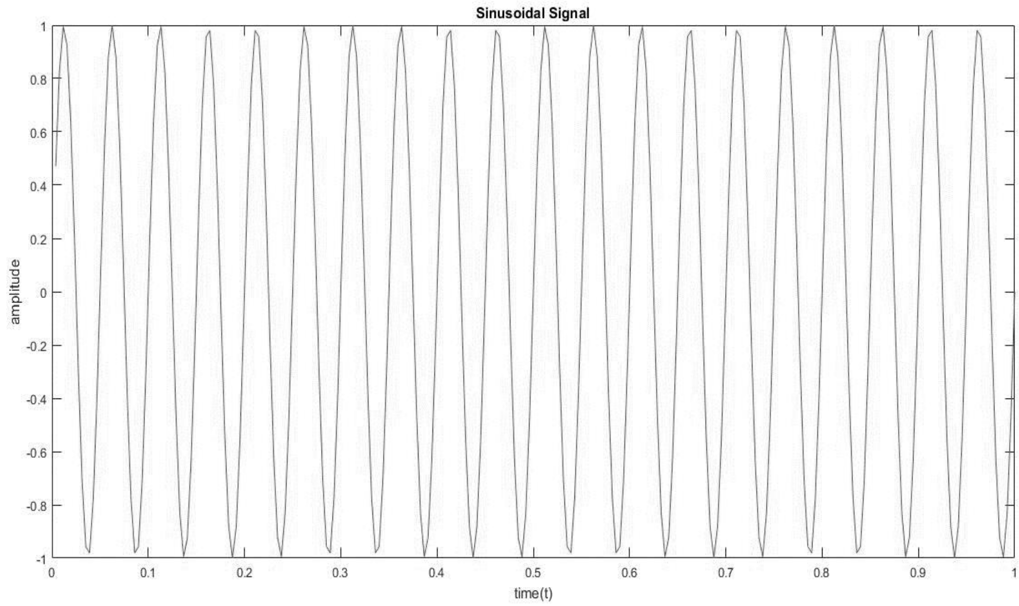

Figure 4. 1 Sinusoidal of Frequency 20 Hz

Then, we gave the above sinusoidal signal as an input to the function energyDWT. Now the output of that function is given as in matrix form of 6*1 size. Which is given below:

Table 4.4 Energy Values of different frequency ranges after DWT decomposition

S.N.	Frequency Band (Hz)	Energy of Coefficients (Joule)
1	0 – 4	0.72828
2	4 – 8	0.29418
3	8 – 16	1.7839
4	16 - 32	2.979
5	32 – 64	0.04837
6	64 – 128	0.00056228

From table 4.4, we can tell that the sinusoidal signal has the highest energy i.e. 2.979 in the frequency band 16 to 32 Hz. Which can validate the fact that the above sinusoidal has the frequency of 20 Hz.

4.3.2 DWT Implementation on EEG Signal Extracted from First Patient

The first patient is our main patient of which we are studying the characteristics of EEG signal during or before the onset of epileptic seizure. According to the CHB-MIT database, patient one is 11 years old girl. The database contains 41 different EEG sample file in which 1 hour of EEG samples are taken in almost every file. The combination of EEG signal of 41 files becomes almost a continuous record of 41 hours. The data sampling rate of all the EEG signal is 256 Hz. Each 1 hour file consists of 23 leads' EEG samples. So, one file of first patient consists of 23*921600 data as a matrix. The EEG signal files that consists of epileptic seizure and non-epileptic seizure according to the database are given in the table below:

Table 4.5 Epileptic Seizure presence information for patient one

File Name	Seizure Presence	File Name	Seizure Presence
Chb01_01	Absent	Chb01_22	Absent
Chb01_02	Absent	Chb01_23	Absent
Chb01_03	Present	Chb01_24	Absent
Chb01_04	Present	Chb01_25	Absent
Chb01_05	Absent	Chb01_26	Present
Chb01_06	Absent	Chb01_27	Absent
Chb01_07	Absent	Chb01_29	Absent
Chb01_08	Absent	Chb01_30	Absent

Chb01_09	Absent	Chb01_31	Absent
Chb01_10	Absent	Chb01_32	Absent
Chb01_11	Absent	Chb01_33	Absent
Chb01_12	Absent	Chb01_34	Absent
Chb01_13	Absent	Chb01_36	Absent
Chb01_14	Absent	Chb01_38	Absent
Chb01_15	Present	Chb01_39	Absent
Chb01_16	Present	Chb01_40	Absent
Chb01_17	Absent	Chb01_41	Absent
Chb01_18	Present	Chb01_42	Absent
Chb01_19	Absent	Chb01_43	Absent
Chb01_20	Absent	Chb01_46	Absent
Chb01_21	Absent		

With each file containing one hour of data, we then applied the same DWT function we used for the testing of sinusoidal signal. We divided the whole one hour of signal into different windows with window size of 256 samples and step size of 128 samples. We used shifted window technique because we wanted to reduce the edge effect.

For the DWT implementation on patient one, we started from the file 3rd because that file has the seizure present in it. The reason behind starting from the 3rd file is because we wanted to see how the epileptic seizure behave. One hour of sample is divided into different windows using window shifting windows technique. We got 7199 windows for one hour of samples. DWT implementation on each window gives us the 6 different frequency bands' energy in a matrix of 1*6 size and for 7199 windows, we get 7199*6 matrix as the output for every lead. Each lead has 6 different energy of different frequency band and they are plotted.

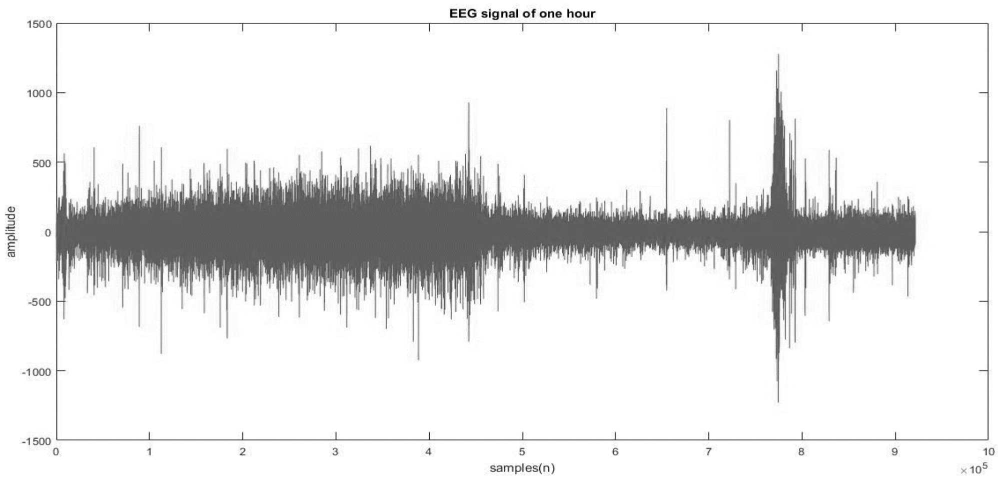

Figure 4.2 EEG signal of chb01_03_edfm.mat of lead FP1-F7

The above figure is the EEG signal contains 921600 EEG samples of lead 1st of file 3rd. In file 3rd, there is one epileptic seizure present which starts from 2996th second to 3036th second for about 40 seconds. Now, the above signal is divided into 7199 windows using window shifting technique and then calculated the energy for all 6 bands. The figures seen below has the x-axis representing the window number and y-axis representing the energy value of corresponding windows.

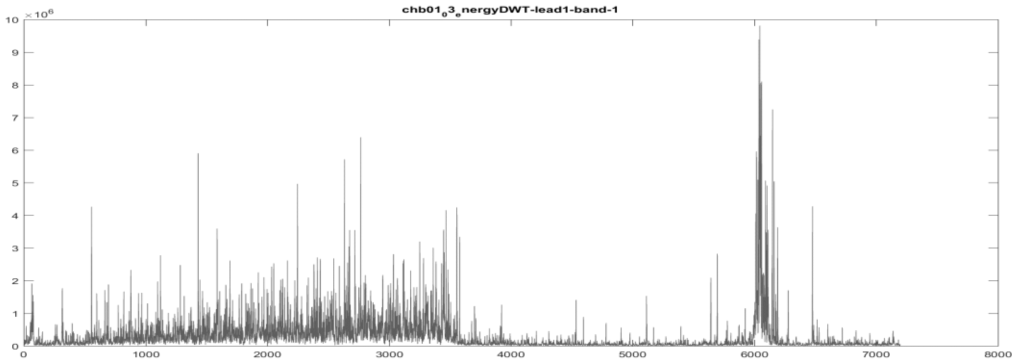

Figure 4.3 Energy spectrum of first frequency band i.e. 0 to 4 Hz

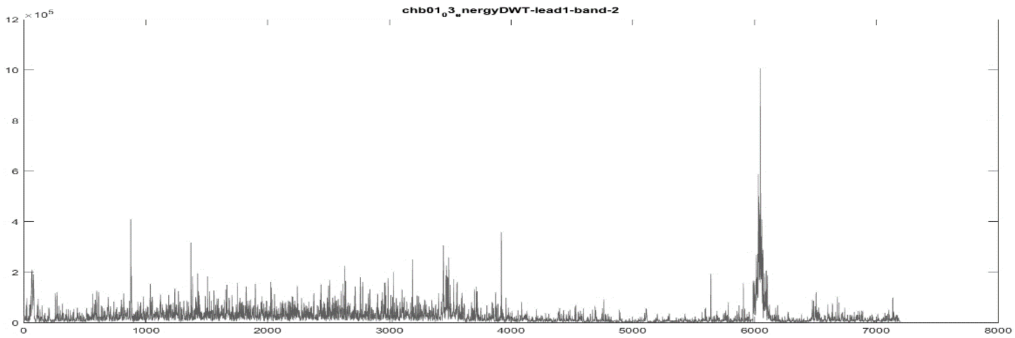

Figure 4.4 Energy spectrum of second frequency band i.e. 4 to 8 Hz

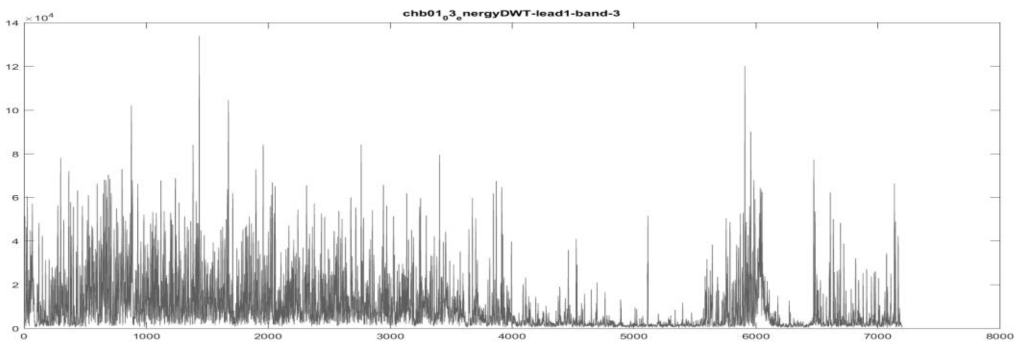

Figure 4.5 Energy spectrum of third frequency band i.e. 8 to 16 Hz

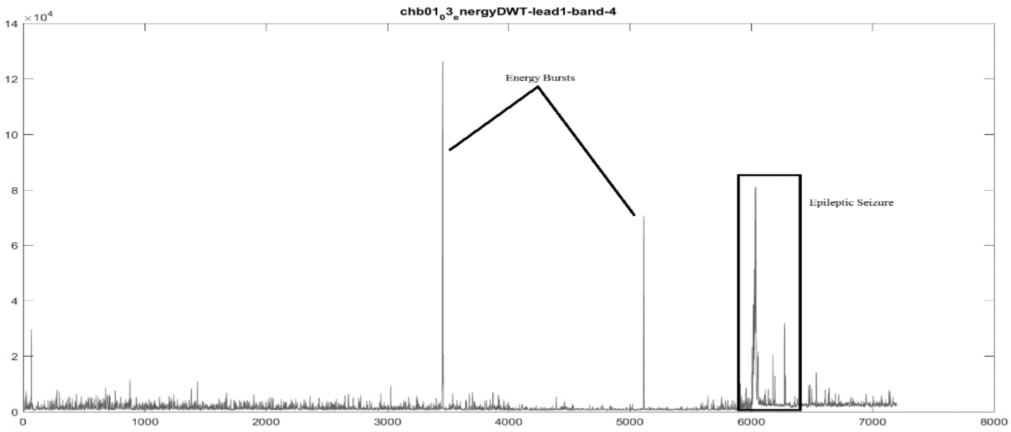

Figure 4.6 Energy spectrum of third frequency band i.e. 16 to 32 Hz

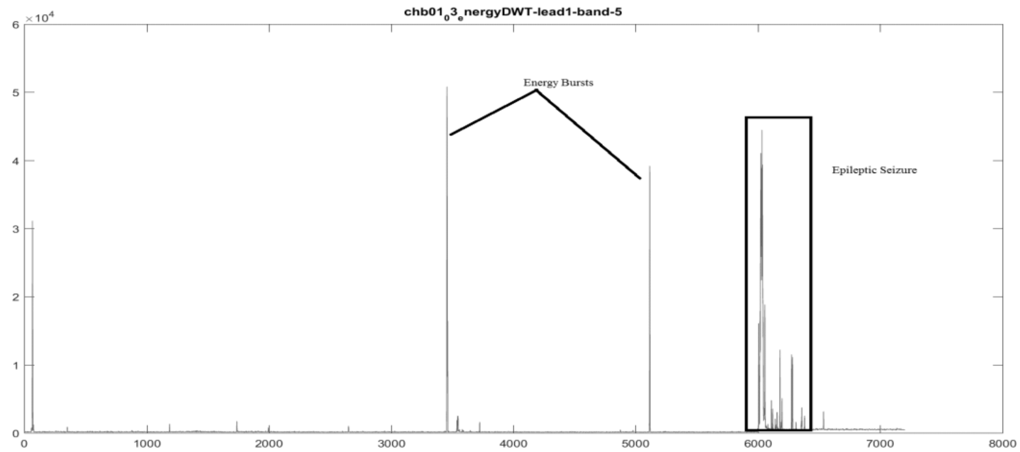

Figure 4.7 Energy spectrum of fourth frequency band i.e. 32 to 64 Hz

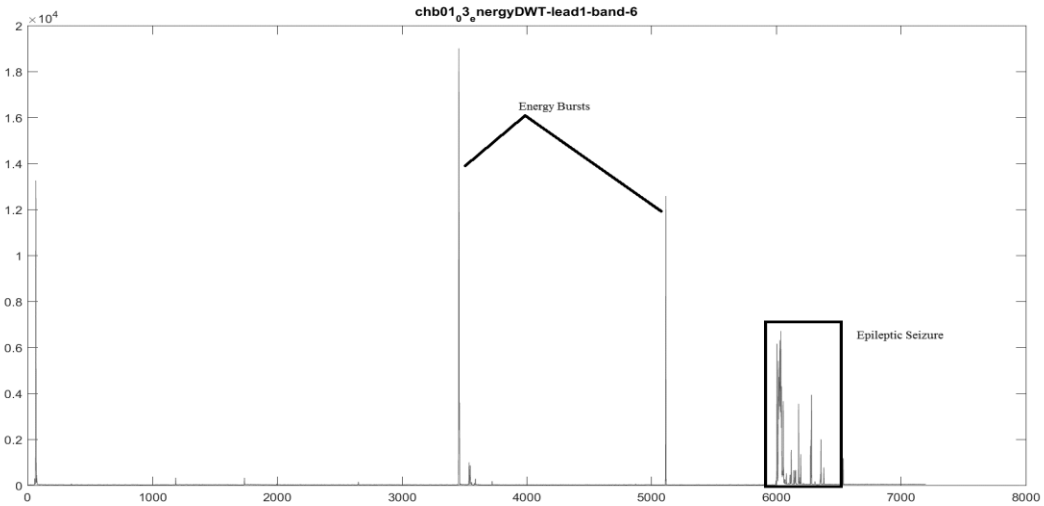

Figure 4.8 Energy spectrum of fifth frequency band i.e. 64 to 128 Hz

From the above 6 figures, we can see that energy spectrum of first, second and third frequency band does not show the clear distinction between seizure and other energy bursts. In fourth, fifth and sixth frequency bands we can clearly see the distinct energy bursts before the onset of epileptic seizure. For the third file, there are two energy bursts before the onset of epileptic seizure. These energy bursts can be the pattern that we can look for in other lead and files of first patient. We have done the same procedure for all the files of first patient and for all the 23 leads. And after the implementation, each EEG signal gives the 6 energy spectrum of individual frequency bands. And, we plotted each energy spectrum for all 943 signals and got almost 5500+ figures. After getting all the energy spectrum as 5500+ figures, we analyzed each figure. For the proper analysis, we made a binary condition where 1 means the figure contains the same pattern i.e. energy bursts before the epileptic seizure and 0 means there is not distinct pattern of epileptic seizure. Then, we went through every figure and put them in a table which gives all the information for one file. The logical analysis of 3rd file is shown in table 4.6.

Table 4.6 Logical representation of presence of seizure pattern for 3rd file of patient one

Leads	Bands					
	0-4 Hz	4-8 Hz	8-16 Hz	16-32 Hz	32-64 Hz	6-128Hz
1	0	0	0	1	1	1
2	0	0	0	1	1	1
3	0	0	0	1	1	1
4	0	0	0	0	1	1
5	0	0	0	1	1	1
6	0	0	0	1	1	1
7	0	0	0	0	1	1
8	0	0	0	0	1	1
9	0	0	0	0	1	1
10	0	0	0	0	1	1
11	0	0	0	0	1	1
12	0	0	0	0	1	1
13	0	0	0	1	1	1
14	0	0	0	1	1	1
15	0	0	0	1	1	1
16	0	0	0	0	1	1
17	0	0	0	0	0	1
18	0	0	0	0	0	1
19	0	0	0	1	1	1
20	0	0	0	1	1	1
21	0	0	0	1	1	1
22	0	0	0	1	1	1
23	0	0	0	1	1	1

From the above figure it can be clearly said that the seizure pattern can be seen in 4^{th}, 5^{th} and 6^{th} frequency band. Similarly, we did the analysis for all the 41 files of patient one and the results are presented in table 4.7.

Table 4.7 Analysis of 41 files of first patients

SN	File Name	Presence or Absence of Epilepsy	Lead	Band	Remarks
1	Chb01_01	Absent	Except lead 5	4^{th} ,5^{th} , and 6^{th}	
2	Chb01_02	Absent	Except leads 3,4,7 and 18		
3	Chb01_03	Present		4^{th} ,5^{th}, and 6^{th} bands	
4	Chb01_04	Present	Except lead 1	4^{th} ,5^{th} , and 6^{th} bands	High energy around 6000^{th} and 3500^{th} window
5	Chb01_05	Absent	Except lead 13	5^{th} and 6^{th}	For some file 1500^{th} ,2000^{th},3000^{th} window
6	Chb01_06	Absent	Except lead 7	3^{rd} , 4^{th}, 5^{th} , and 6^{th}	High energy bursts just below 4000^{th} and 7000^{th} window
7	Chb01_07	Absent		Bursts in random band	Bursts at 200^{th} and 1000^{th}
8	Chb01_08	Absent			Bursts around 5500^{th}
9	Chb01_09	Absent			
10	Chb01_10	Absent		Bursts in 1^{st} band	100-400^{th} of the window
11	Chb01_11	Absent			There is some random bursts only
12	Chb01_12	Absent			Some bursts around 4500^{th} and 5000^{th}
13	Chb01_13	Absent		Bursts in 3^{rd} and 4^{th} band	Energy bursts around 3000^{th}
14	Chb01_14	Absent	For very few leads	Only thin bursts for 5^{th} and 6^{th} band	Around 2000^{th}
15	Chb01_15	Present	Except 16^{th} and 18^{th} lead	4^{th} ,5^{th} , and 6^{th} bands	3600^{th} ,4500^{th} and 5000^{th} of the window
16	Chb01_16	Present	Except 16^{th} and 19^{th} leads	4^{th} ,5^{th} , and 6^{th} bands	Around 2100^{th} and 2500^{th}

17	Chb01_17	Absent	Lead 16		Less magnitude around 500 to 2500^{th} of the window
18	Chb01_18	Present	Except 11, 12 and 16	4^{th} ,5^{th} , and 6^{th} bands	Bursts are in two region first 4500^{th} ,5000^{th},7000^{th} and second 1000^{th},2000^{th} and 2500^{th} of the window
19	Chb01_19	Absent	Absent in 12 and 16 lead only	Bursts present in 5^{th} and 6^{th} band	Present around 3200^{th},4500^{th},and 5800^{th} of the window
20	Chb01_20	Absent	No clear bursts in 13^{th},14^{th} ,21th and 22th leads	1^{st} ,5^{th} , and 6^{th} bands	Around 2000^{th}, 3500^{th} and 4000^{th} of the window
21	Chb01_21	Absent		High Energy in 5^{th} and 6^{th} band	High energy at 700^{th} and 800^{th} of the window
22	Chb01_22	Absent			
23	Chb01_23	Absent		Bursts present in 3^{rd} band	There is energy bursts in 7000^{th} of the window
24	Chb01_24	Absent			No energy bursts and pattern as well
25	Chb01_25	Absent			But bursts are present around 4800^{th} and 7000^{th} in some leads these does not show any pattern
26	Chb01_26	Present		Nearly similar 5^{th} and 6^{th}	Bursts are mainly in 3500-3800^{th} and 700-1200^{th} window
27	Chb01_27	Absent		High energy in 3^{rd} band	Around 1000^{th} of the window
29	Chb01_29	Absent	Seizure pattern in few leads	Present in 2^{nd} and 3^{rd} band	Bursts are around 500^{th},1200^{th},5000^{th} of the window
30	Chb01_30	Absent			Energy bursts similar to the seizure pattern in around 3000^{th} window but can’t be taken as seizure pattern
31	Chb01_31	Absent			Randomly seen pattern of energy bursts in 5500^{th} window in few leads and bands.

32	Chb01_32	Absent			There are some bursts of energy but not in deterministic manner
33	Chb01_33	Absent			There is randomly seen energy bursts so can be neglected
34	Chb01_34	Absent	10^{th} and 18^{th} leads	5^{th} and 6^{th} bands	There is seizure pattern and bursts are in around 2800^{th}, 4900^{th} and 6900^{th} window
36	Chb01_36	Absent	Seizure pattern present in 10th and 11st	Mainly seen in 5^{th} and 6^{th} band	There is seizure pattern seen in around 3500^{th} and 4700^{th} window
38	Chb01_38	Absent	Pattern can be seen except in 11^{th} , 12^{th}, 16^{th} and 17^{th} lead	Pattern is mainly seen in 5^{th} and 6^{th} band	Only one bursts of energy (in around 6600^{th} window)
39	Chb01_39	Absent	Pattern can be seen except in 6^{th} , 7^{th}, 8^{th} and $11^{th,}$ 12^{th}, and 16^{th} lead	The pattern can be seen in 5^{th} and 6^{th} band	Seizure pattern seen in four bursts of energy (in around 2000^{th}, 2500^{th}, 4800^{th} and 6900^{th} window)
40	Chb01_40	Absent		5^{th} and 6^{th} band	Two bursts are in 2300^{th} and 4000^{th} window
41	Chb01_41	Absent	Except 7^{th}, 8^{th}, 9^{th}, 10^{th}, 11^{th}, 12^{th}, 16^{th}, 17^{th}, and 18^{th} leads	5^{th} and 6^{th} band	Windows with bursts are 2500^{th}, 4100^{th} and 5200^{th}
42	Chb01_42	Absent			
43	Chb01_43	Absent	Except in 9^{th}, 10^{th}, 17^{th} and 18^{th} leads	5^{th} and 6^{th} band	Four bursts around 9000^{th}, 3500^{th}, 4600^{th} and 5500^{th} window
46	Chb01_46	Absent			There is energy burst in very few cases (10-15 cases of 1180) and they are randomly seen

We have some files in which there are seizure pattern like 3^{rd}, 4^{th}, 15^{th}, 16^{th}, 18^{th}, 21^{st} and 26^{th} but there is seizure pattern in 17^{th}, 19^{th} and 20th file as well. The reason behind the presence of seizure pattern in other non-epileptic EEG signal can be the effect of epileptic seizure remains for the time after the seizure. We also found the same pattern in file no. 36, 37, 38, 39, 40 and

41 in which there is no seizure according to the CHB-MIT database. There are no file with epileptic seizure after 26th file.

In 17th, 19th and 20th file, we checked for the magnitude of energy by plotting 5th and 6th frequency band of 15th to 21st files together. Then, we saw no particular difference in magnitude in non-epileptic signals as well. We mainly did for the 5th and 6th band because there is high presence of seizure pattern in these files as well as in epileptic seizure containing files. In other frequency bands there were no significant seizure pattern that could be used for the analysis. For example:

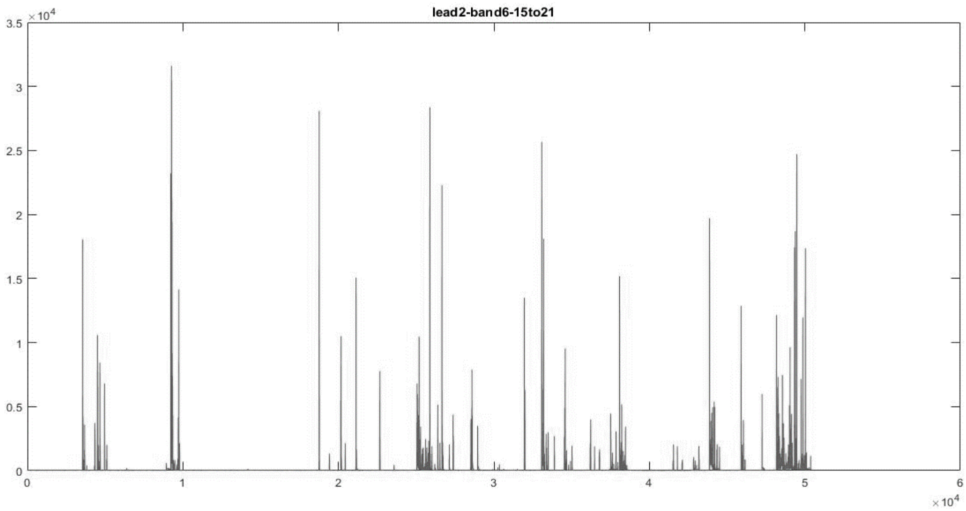

Figure 4.9 Energy spectrum of band 6 of first patient from 15th to 21st file

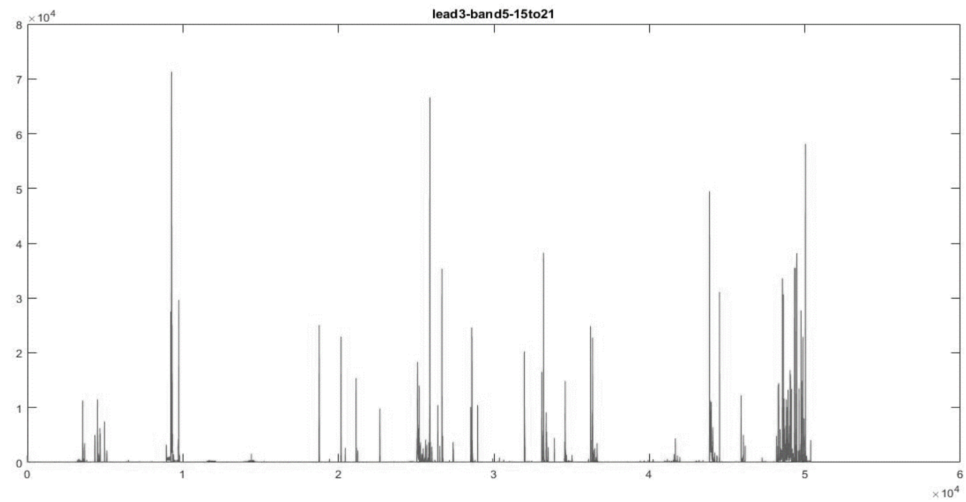

Figure 4.10 Energy spectrum of band 5 of first patient from 15th to 21st file

From the above figure, we can conclude that there is no particular pattern difference between non-epileptic and epileptic signal. We have to analyze more in detail to find the reason of these energy bursts and seizure pattern alike pattern coming in between files/signals.

In 36^{th}, 37^{th}, 38^{th}, 39^{th}, 40^{th}, 41th and 43^{rd} file, there is similar pattern seen as in epileptic seizure which should be analyzed carefully to find the reason behind that. In the figures below, band 1, 2, 3, 4, 5 and 6 represents 0-4 Hz, 4-8 Hz, 8-16 Hz, 16-32 Hz, 32-64 Hz and 64-128 Hz respectively.

Similar logical represented tables of all the files tell if the epileptic pattern is present in the figure plotted for the each band of each lead or not. 1 represents the presence of epileptic seizure pattern and 0 and empty space represents the absence of epileptic seizure pattern. We have to analyze why this seizure patterns are seen in non-epileptic signals too. We have to find if that epileptic seizure pattern is really the pattern for epileptic seizure or just any random energy bursts seen in epileptic seizure present signal. Because, if the pattern we are assuming as epileptic pattern is really the pattern of seizure then that should not be seen in the non-epileptic signals.

For the verification, if the seizure pattern we had seen in the energy spectrum of first patient is true or not, we checked all other patients' signals and energy spectrum. We did the same DWT method using Daubechies 8 wavelet to find the energy spectrum of other patients. There were some energy bursts in some patients' EEG signals and in some cases there was the same pattern as epileptic signal. Which shows irregularities in results that we have been searching for. As our main objective is to classify the EEG signals of first patient, we used that analysis only for the reference.

4.4 Classification using Artificial Neural Network

We used built in Neural Network Toolbox using command line. First of all, we tested the program using an example for pattern recognition of Iris Flower.

For Iris flower, we have a dataset called iris_dataset in MATLAB which contains two variables called targets and inputs. Target matrix is of size 3*150 which contains first 50 elements of 1^{st} row as 1, 51 to 100 elements of second row as 1 and last 50 elements of third row as 1 and remaining as 0 and input matrix of size 4*150. As we know, target matrix should always have value either 0 or 1. Then, we made a neural network of 1 hidden layer using 'patternet' command of MATLAB. We divided the dataset for training, validation and testing in 70%, 15% and 15% respectively. Then, we trained the network with given inputs and targets for Iris Flower which gave us 98.7% of accuracy. 18 number of epochs were used.

The same program was applied to our EEG data from 41 different files. We used 400 energy data i.e. 3.3 minutes of both 5^{th} and 6^{th} band of 22 leads because there was unique pattern in those bands which can be used as a classification parameter. We made a matrix of size 18400*40 called inputDataForANN. This matrix contains the energy data as shown in the following table:

Table No 4.8: Matrix inputDataForANN of size 18400*40

S.N.	Chb01_01(1)		…………	Chb01_46(40)	
	Leads	Bands	…………	Leads	Bands
1	1	5^{th} band (400 energy data)	…………	1	5^{th} band (400 energy data)
800		6^{th} band (400 energy data)	…………		6^{th} band (400 energy data)
801	2	5^{th} band(400 energy data)	…………	2	5^{th} band(400 energy data)
1600		6^{th} band (400 energy data)	…………		6^{th} band (400 energy data)
….	…………	…………	…………	…………	………..
….	…………	…………	…………	…………	…………
18001	23	5^{th} band (400 energy data)	…………	23	5^{th} band (400 energy data)
18400		6^{th} band (400 energy data)	…………		6^{th} band (400 energy data)

Then, for target matrix, we used 1*40 matrix which contains ones in 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 5^{th}, 6^{th} and 21^{st} matrix element and remaining elements are zeros. In target matrix, 1 represent epileptic EEG signal and 0 represents the non-epileptic EEG signal. We used 1 hidden layer and 50 nodes for better computation and accuracy. The division of the data is:

Training: 70%

Validation: 15%

Testing: 15%

But the selection of data was random. From figure 4.11, during training, network used 14 iterations. Neural Network shows that there are 18400 inputs and should be one output. These are the results after execution of the program. As we can see in the progress bar, current gradient value is 0.0626 and initial value is 0.750 and current value is $1*10^{-6}$. 6 validation checks were done since we used default setup.

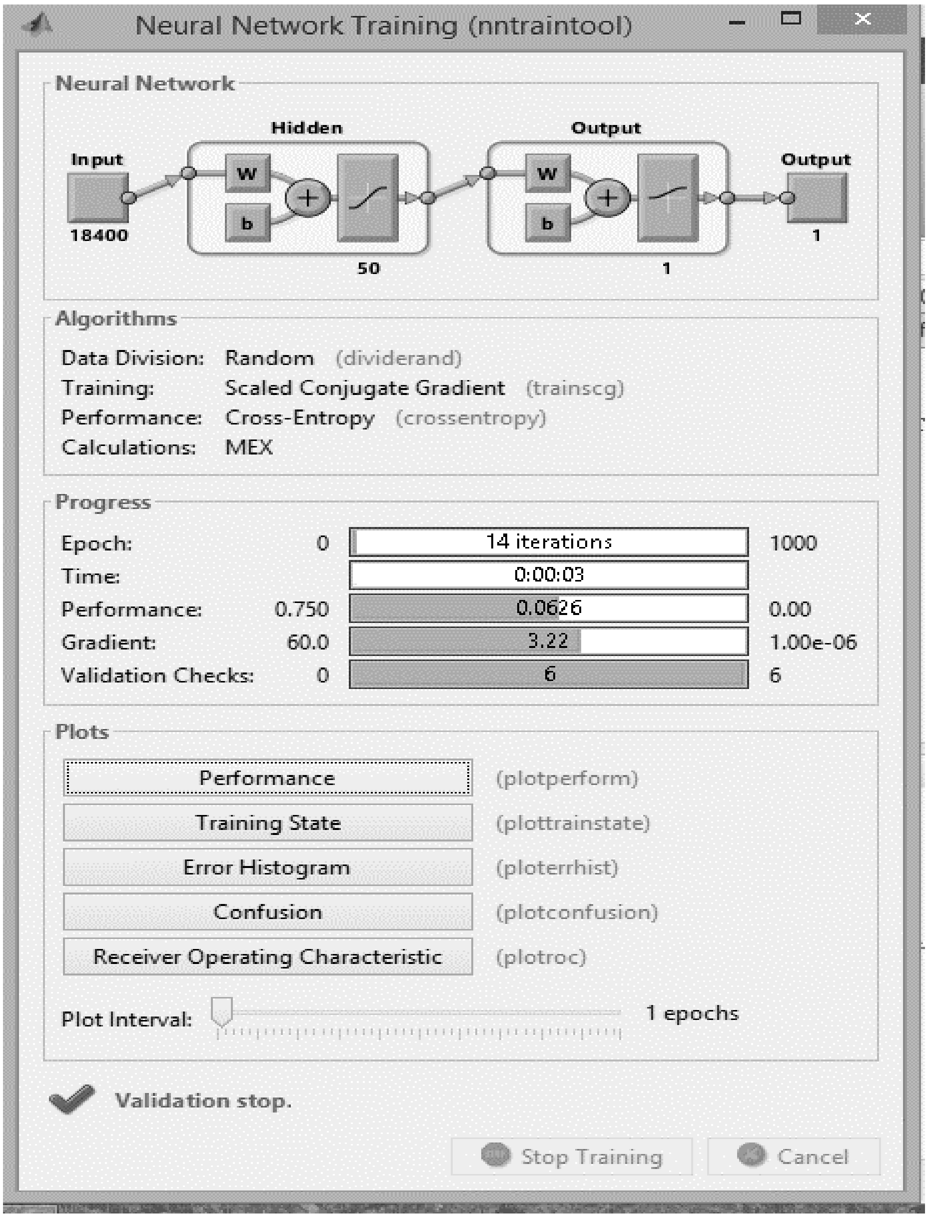

Figure 4. 12 Neural Network Training Tool for classification

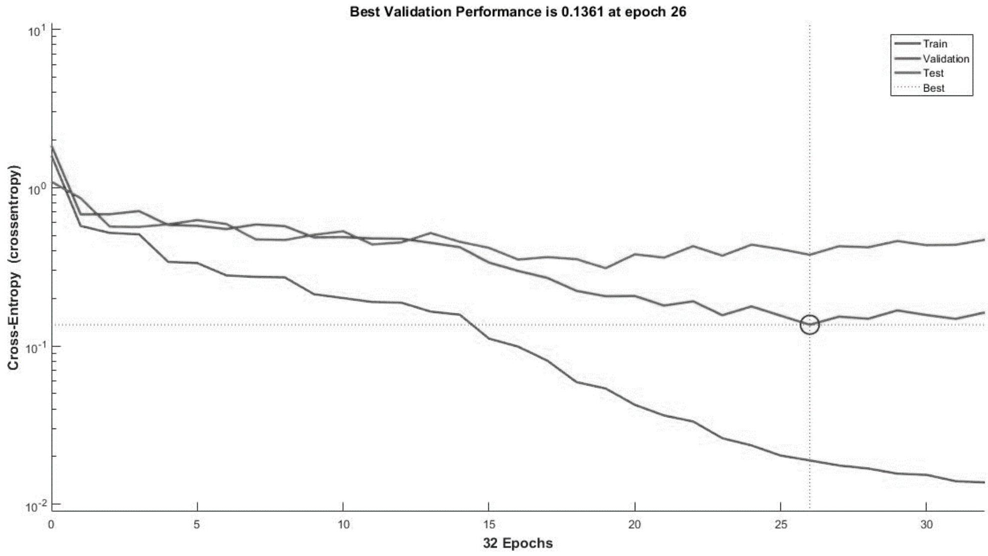

Figure 4. 12 Performance Plot for classification

From the figure 4.12, performance is calculated using cross-entropy which is improving with increasing epochs. And the best validation performance is 0.136 at epoch 26. Performance plot gives us the information to resolve the over-fitting problem. Gradient is needed in the calculation of the weights to be used in neural network. So, there is different weights in each epochs or iterations which can be seen in figure 4.13.

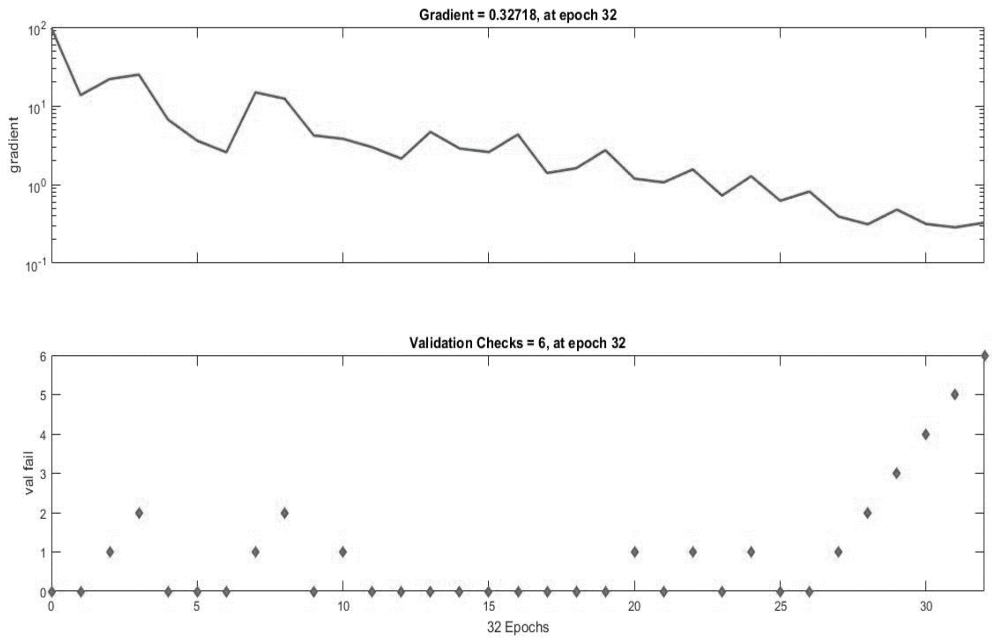

Figure 4.13 Training State

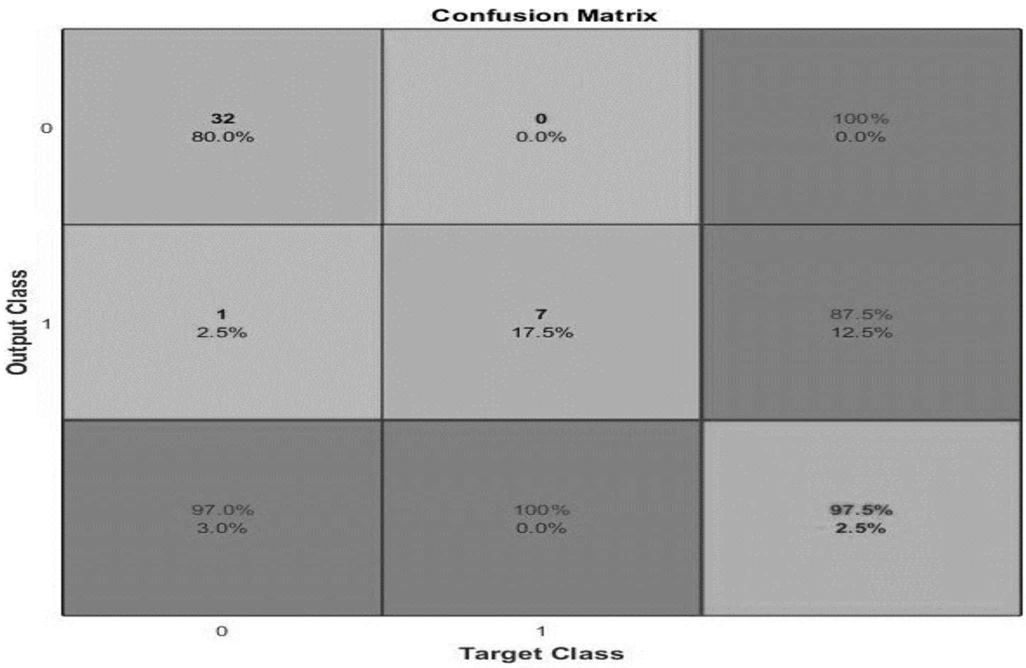

Figure 4.14 Confusion Matrix

From the figure 4.14, there were 33 EEG signals which were not epileptic, out of which 32 EEG signals have been classified as non-epileptic and 1 EEG signal appeared as epileptic. In the similar way, out of 7 epileptic EEG signals, all signals have been verified as epileptic EEG signals. So, the overall accuracy has been found 97.5 %.

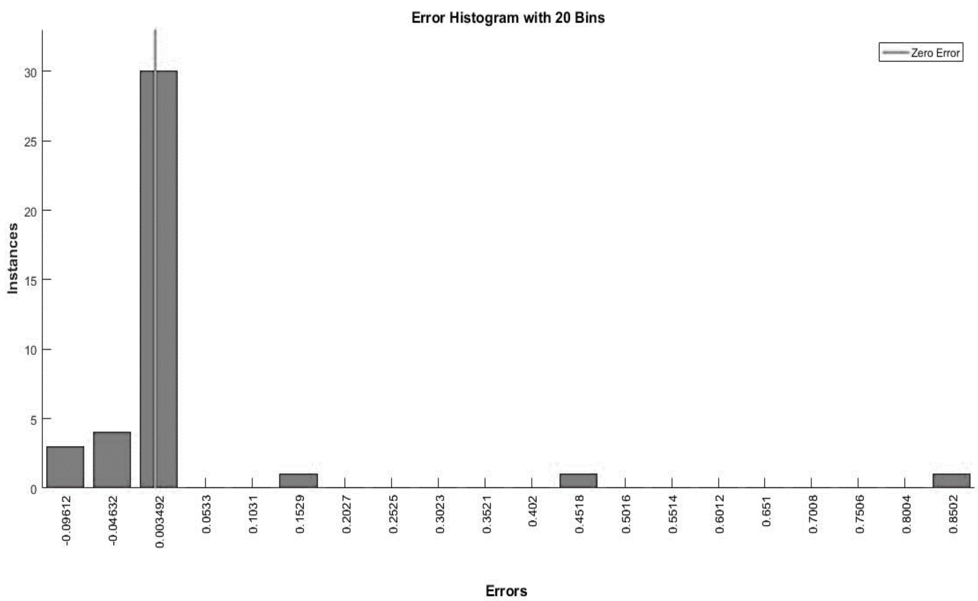

Figure 4. 15 Error Histogram

We have 20 bins according to the figure 4.15 which means there are 20 vertical bars we are observing on the graph. The total error ranges from -0.09612 (leftmost bin) to 0.08502 (rightmost bin). The width of each bins is {0.08502-(-0.09612)}/20 which equals to 0.009057. 30 datasets have an error that lies in the range from (0.003492-0.009057) to (0.003492+0.009057).

4.5 Discussion on Result

Overall summary of all the graphs that were taken using FFT can be described as, all the graphs shows the almost same characteristics. The characteristics that can be observed in the above graphs is that when the shifting of windows goes on increasing, the energy of the window goes on increasing because we have used only 20 seconds before the start of the epilepsy. After the shifting window reaches to the start of epilepsy the energy starts to increase. The magnitude of energy is highest for the 0 to 4 Hz frequency band which tells that the signal might have greater components in that range. Regardless of magnitude of the energy, all the frequency band and time domain signal has shown the same characteristics which is when increasing the window shifting to the starting of seizure results in increasing of the energy of the time and frequency domain signal's energy.

After the implementation of DWT, from the first patient's data and energy spectrum of each bands for every lead we could tell that there is energy pattern seen before the epilepsy happens. There is some energy bursts seen before epileptic seizure and large amount of energy is seen during epileptic seizure. For first patient, there are energy in non-epileptic signals as well. First patient has epileptic seizure at 3rd, 4th, 15th, 16th, 18th, 21st and 26th file but we could see the same pattern that we assumed as seizure pattern in 17th, 19th, 20th, 36th, 37th, 38th, 39th, 40th, 41st and 43rd file as well. 17th, 19th and 20th pattern could be said that it is the effect of seizure that happened just before this period. There may be the energy burst seen way before the epileptic seizure in 20th file. But for other non-epileptic files in which there was epileptic seizure pattern, the reason behind that epileptic seizure patterns are very susceptible although there is random occurrence of energy bursts. And for pattern verification we used other patients EEG signal. We can see that there is pattern seen in epileptic files. The same pattern we saw in the first patient's energy spectrum is seen in the 2nd, 3rd and 4th patient's energy spectrum as well. But for other patients there is very much random occurrence of energy bursts and seizure pattern which will be studied and analyzed further.

After ANN implementation for classification has been done, we used 40 files of first patient to train the network in which 7 files were epileptic and 33 files were non-epileptic. Since, we used random division algorithm, the accuracy was varying at different instances in the range of 85% to 97.5%.

CHAPTER 5 CONCLUSION

We have done the energy calculation of time domain signal and frequency domain signal. We transformed the time domain signal to frequency domain signal by the Fast Fourier Transform mechanism and Discrete Wavelet Transform. Signal which have the presence of epileptic seizure were extracted from the CHB-MIT database and then we selected the whole signal and files for the analysis and energy calculation. We used mainly the first patients EEG signal and other patients' EEG signal which only had epileptic seizure. We divided the selected signal for testing to different windows of size 256 samples with shifting windows technique of 128 samples. We calculated the energy for each windows and analyzed for the different frequency band using DWT implementation. From the results, we can conclude that the energy of the signal starts to rise with the start of epileptic seizure and in frequency domain there is energy bursts seen just before the onset of epileptic seizure. From the analysis of first patient's all files we found a pattern and for further pattern verification we used other patient's EEG signal as well as first patient's other signals.

Then, we implemented Artificial Neural Network tool of MATLAB to classify the EEG signals as epileptic or non-epileptic. We trained the network using input data which we got from DWT implementation on all 40 files of first patient. We divided 40 files' data into 75% for training, 15% for validation and 15% for testing. But the selection of data was done randomly. Finally, we got the accuracy of 85% to 97.5 %.

5.1 Recommendations

The above project can be enhanced as follows:

1. The same DWT implementation and classification using ANN can be done for other patients and other EEG signals.
2. Different features like mean, log detector, variance, entropy etc. can be extracted and processed.
3. Different classification algorithms like Linear Discriminant Analysis (LDA) [14], Support Vector Machine [7] etc. can be used.

REFERENCES

[1] Data Processing & Analyzing. (2018, Feb 18th). Retrieved from http://gwyddion.net/documentation/user-guide-en/wavelet-transform.html#wavelet-transform-ref-1

[2] Nitendra Kumar,Khursheed Alam and Abdul Hasan Siddiqi. (2017, August 5th). Wavelet Transform for Classification of EEG Signal using SVN and ANN

[3] B. Al-Bokhity, Dalia Nashat and T. M. Nazmy. (2017 Nov 1st). Accuracy Enhancement of the Epileptic Seizure Detection in EEG signals.

[4] Milica Milosevic. (2015 May). Automated Detection of Epileptic Seizures in pediatric patients based on accelerometry and surface electromyography.

[5] Aditya Sundar, Chinmay Das. (2015 July). Matlab Analysis of EEG signals for diagnosis of epileptic seizures.

[6] Ujjwal KC, Ghansyam Kshetri, Saroj Koirala. (2015 February). Epileptic Seizures Onset Detection.

[7] Mohammad Zavid Parvez,Manoranjan Paul. (2015 October). Prediction and Detection of Epileptic seizure by analyzing EEG signals.

[8] Support Vector Machine. (2018, Feb 25th). Retrieved from http://scikit-learn.org/stable/modules/svm.html

[9] Overview and application of Artificial Neural Networks. (2018, Feb 25th). Retrieved from https://medium.com/@xenonstack/overview-of-artificial-neural-networks-and-its-applications-2525c1addff7

[10] Artificial Neural Network. (2018, Nov 19th). Retrieved from https://en.wikipedia.org/wiki/Artificial_neural_network

[11] What is an artificial neural network? (2018, Sep 13th). Retrieved from https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

[12] CHB-MIT Scalp EEG Database. (2018, Feb 25th). Retrieved from https://physionet.org/physiobank/database/chbmit/

[13] Paseval’s theorem. (2018, Nov 20th). Retrieved from https://en.wikipedia.org/wiki/Parseval%27s_theorem

[14] Lorena Orosco, Agustina Garces Correa and Eric Laciar Leber. (2011). Epileptic Seizure Detection Based on Empirical Mode Decompostion of EEG Signals.

[15] Neural Network Toolbox. (2018, Nov 20th). Retrieved from http://matlab.izmiran.ru/help/toolbox/nnet/backpr59.html