5-2017

Detection and Classification of EEG Epileptiform Transients with RBF Networks using Hilbert Huang Transform-derived Features

Raghu Jagadeesha
Clemson University, rjagade@clemson.edu

Detection and Classification of EEG Epileptiform Transients with RBF Networks using Hilbert Huang Transform-derived Features

---

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Computer Engineering

---

by
Raghu Jagadeesha
May 2017

Accepted by:
Dr. Robert Schalkoff, Committee Chair
Dr. Brian Dean
Dr. Carl Baum

Abstract

Diagnosis of epilepsy or epileptic transients AEP (Abnormal Epileptiform Paroxysmal) is tedious, but important, and an expensive process. The process involves trained neurologists going over the patients EEG records looking for epileptiform discharge like events and classifying it as AEP (Abnormal Epileptiform Paroxysmal) or non-AEP. The objective of this research is to automate the process of detecting such events and classifying them into AEP(definitely an Epileptiform Transient) and non-AEPs (unlikely an epileptiform transient). The problem is approached in two separate steps and cascaded to validate and analyze the performance of the overall system.

The first step is a detection problem to find the Epileptiform like transients (ETs) from the Electroencephalograph (EEG) of a patient. A Radial basis function-based neural network has been trained using a training set consisting of examples from both classes (ETs and non-ETs). The ETs are the yellow boxes which are marked by expert neurologists. There are no particular examples of non-ETs and any data not annotated by experts can be considered to be examples of non-ETs.

The second step is classification of the detected ETs also known as yellow boxes, into AEPs or non-AEPs. A similar Radial basis function-based neural network has been trained using the ETs marked and classified into AEPs and non-AEPs manually by seven expert neurologists. The annotations or yellow boxes along with the contextual signal was used to extract features using the Hilbert Huang Transform.

The system is validated by considering an entire epoch of the patient EEG and potential ETs are identified using the detector. The potential ETs marked by the detector are classified into AEPs and non-AEPs and compared against the annotations marked by the experts.

Acknowledgments

I would like to convey my sincere gratitude towards Dr. Robert J. Schalkoff for his continuous support and guidance throughout my research. I am thankful for Dr. Jonathan Halford for providing us with all the necessary data for this project. I am also thankful to Dr. Gabriel Rilling for open sourcing the implementation of Empirical mode decomposition and providing an insightful explanation of the same. I would like to thank Dr. Brian Dean and Dr. Carl Baum for being part of the Committee. Finally, I would like to thank my research team members for their encouragement and continuous support.

Table of Contents

Title Page . . . . . **i**

Abstract . . . . . **ii**

Acknowledgments . . . . . **iii**

List of Tables . . . . . **vi**

List of Figures . . . . . **vii**

1 Introduction **1**
1.1 Overview of the problem 1
1.2 Previous Work 2
1.3 Approach 3

2 Data Acquisition and Preprocessing **5**
2.1 Electroencephalogram 5
2.2 Data Acquisition 7

3 Outline of HHT and RBFNN **16**
3.1 Hilbert Huang Transform (HHT) 16
3.2 Radial Basis Function Neural Network 25

4 Classification of Yellow Boxes **32**
4.1 Data 32
4.2 Classifier Neural Network 33
4.3 Results 35
4.4 Conclusion 43

5 Detection of Epileptiform Transients **44**
5.1 Training Data 44
5.2 Neural Network for Detection 46
5.3 Results 46
5.4 Conclusion 52

6 Overall System Validation **53**
6.1 Method 53
6.2 Performance Measures 55
6.3 Results 65
6.4 Conclusion 66

7 Future Work **67**

Bibliography **69**

List of Tables

2.1 Nomenclature of Electrodes. 6
2.2 Confidence factors used for annotations. 8
2.3 Yellow Box duration statistics 8

4.1 Distribution of samples during k-fold validation 36
4.2 Performance Metrics for each Trial of k-fold Validation 42
4.3 Average Performance Metrics of k-fold validation 43

5.1 Performance Metrics for each Trial of k-fold Validation in Detection 50
5.2 Average Performance Metrics of k-fold validation 51

6.1 List of files for Validation 54
6.2 Results of Validation 65
6.3 Performance metrics of Validation 66

List of Figures

1.1 Block diagram of the classification stage Neural Network 4
1.2 Block diagram of the ET Detection Neural Network 4

2.1 10-20 electrode placement system 6
2.2 Examples of class 205 11
2.3 Examples of class 204 12
2.4 Examples of class 203 13
2.5 Examples of class 202 14
2.6 Examples of class 201 15

3.1 Examples of IMFs from class 201 and 202 19
3.2 Examples of IMFs from class 204 and 205 20
3.3 A Basic RBF Network 27
3.4 Representation of Multilayer RBF Network 28

4.1 Histogram of Yellow Boxes 33
4.2 Histogram of length of annotations 34
4.3 Classification phase error plot for k = 4 37
4.4 Error plot for k = 5 37
4.5 Error plot for k = 6 38
4.6 Error plot for k = 7 38
4.7 Error plot for k = 8 39
4.8 Error plot for k = 9 39
4.9 Error plot for k = 10 40
4.10 Confusion matrix 41

5.1 Detection phase error plot for k = 4 47
5.2 Error plot for k = 5 47
5.3 Error plot for k = 6 48
5.4 Error plot for k = 7 48
5.5 Error plot for k = 8 49
5.6 Error plot for k = 9 49
5.7 Error plot for k = 10 51

6.1 Block diagram of ET detection system 54
6.2 Validation on file no.12 on channel 'O1 - T5' 57
6.3 A closer look on file no.12 validation 57
6.4 validation on file no.36 on channel 'FP1 - Avg' 58
6.5 A closer look on file no.36 validation 58
6.6 validation on file no.65 on channel 'Fp1 - F3' 59
6.7 A closer look on file no.65 59
6.8 Validation on file no.89 on channel 'Fp1 - F3' 60

6.9 An example of a false positive of non-AEP in file no.89 60
6.10 Validation on file no.95 on channel 'F3 - C3' 61
6.11 A closer look on file no.95 validation 61
6.12 Validation on file no.124 on channel 'Fp2 - F4' 62
6.13 A closer look on file no.124 validation 62
6.14 Validation on file no.160 on channel 'O2 - O1' 63
6.15 A closer look on file no.160 validation 63
6.16 Validation on file no.159 on channel 'T5 - Avg' 64
6.17 Examples of validation on file no.24 with no yellow boxes 64

Chapter 1

Introduction

1.1 Overview of the problem

Epileptic seizure is a transient occurrence of signs or symptoms due to abnormal excessive or synchronous neuronal activity in the brain. Epilepsy is a disease characterized by an enduring predisposition to generate epileptic seizures and by the neurobiological, cognitive, psychological and social consequences of this condition [8]. This chronic disease is characterized by recurrent, unprovoked seizures. Although the symptoms of a seizure may affect any part of the body, electrical events that produce the symptoms occur in the brain. The location of that event, how it spreads, how much of the brain is affected, and the duration it lasts all have profound effects. These factors determine the characteristics of a seizure and its impact on the individual [9]. According to the epilepsy foundation, epilepsy is the fourth most common neurological disorder and affects people of all ages.

The routine scalp electroencephalogram (rsEEG) is an important test for diagnosing epilepsy which records the electrical activity in the brain. EEG machine records the brains activity as a series of signal traces. Each trace corresponds to different region of the brain. The presence of epileptiform transients (ETs) also known as the spike or sharp wave discharges in EEG is evidence of epilepsy [1] [14]. Due to the wide variety of morphologies of ETs and their similarities to normal activities of the brain and other artifacts (i.e. extra cerebral potentials from muscles, eyes, heart etc.) make it difficult to detect with high accuracy [11]. These spikes are typically 20-70 ms in duration and are followed by sharp waves of 70-200 ms duration. Some of the ETs are more complex consisting

of spikes followed by a slow wave lasting 150-350 ms called as spike-and-slow-wave-complex [46].

ETs are usually detected by expert physicians who visually inspect the patients EEG recordings. This process is highly time consuming, especially in the case of long term EEG recordings, e.g. 24-hour continuous ambulatory monitoring studies. Also, the considerable disagreement in the detection of ETs by experts leading to increased chances of misdiagnosis [46] [27] [17]. It is therefore necessary to develop an efficient and reliable technique to automatically detect the Epileptiform Transients.

1.2 Previous Work

A lot of work has been done in the area of ET detection and classification in EEG using both raw data and feature engineering. Most of these strategies can be roughly divided into the following categories. (1) Orthogonal transforms which captures the rhythmic change of EEG and is used to detect spike-and-wave complex [27]. (2)Template matching where templates of epileptiform waves are used and cross correlation between EEG segment and ET templates are computed and decision is made based on threshold [35]. (3) Expert System methods mimic human visual inspection by considering parameters such as amplitude, width, slope of EEG waves and compare them against a threshold, additional information such as spatial and temporal context of such signals are also considered to reject artifacts sharing similarities to ETs [11] [12]. (4) Raw EEG signals along with few parameters are considered and combined to form a weighted input to discriminative classifiers [10] [43] [24] [44] [28]. (5) Time-frequency analysis in detection of Epileptiform EEG has also been widely used by great number of researchers [38] [36] [22]. Each method provides some unique advantages but none of them alone can fulfill the requirement of Epileptiform detection.

It is very important to consider temporal information of signals in performing analysis of EEG signals for any applications [11] [12]. Feature extraction from EEG signals using empirical mode decomposition (EMD) which provides effective time-frequency analysis of non-stationary signals with support vector machines (SVMs) classifier has shown good results [34]. In 1998, Norden E. Huang et al, proposed a new method for analyzing non-stationary data [19] [4]. This technique is known as the Huang-Hilbert transform(HHT). This technique has been recently applied to different applications to analyze EEG signals including Epileptic seizure detection and classification [32] [33] which utilize EMD as a preprocessing step in computing Hilbert transforms. In parallel, the parameters such as

energy, variance, peaks, sharp-spike wave complexes and duration are used to extract features and radial basis functions (RBF) neural networks are used as post classifiers to predict the epilepsy risk levels in EEG signals due to its structural simplicity and faster learning abilities of locally tuned neurons [15].

1.3 Approach

The diagnosis of epilepsy is usually a very time consuming procedure usually performed manually by expert neurologists. The EEG recordings are visually examined and suspected ETs are marked. Then the marked regions also known as the annotations or yellow boxes are classified into AEPs or non-AEPs. These ETs can be seen on different channels depending on the location of occurrence of such events in the brain. The main objective of this research is to automate the entire process of detection of Epileptiform like Transients (ETs) in the given EEG data and classify them in to AEPs (Abnormal Epileptiform Paroxysmal) and non-AEPs. This problem is approached as two separate steps.

In the first step, an entire 30 second recording of EEG on one montage and channel is considered. The suspected ETs are marked using a pre-trained Radial Basis Function (RBF) based neural network. The data for training the network is taken from the annotations (yellow boxes) marked by experts which represents the examples of the class ETs. The examples for the other class, non-ETs, are taken randomly by considering all the non-annotated data (which represents a majority class) and different balancing techniques are applied to have equal number of examples from both classes.

In the second step, the marked or suspected ETs are classified into AEPs or non-AEPs using a different radial basis function based neural network. This network is trained using the annotated data provided by experts which is assigned into one of the two classes using the five confidence levels (201 to 205) with 201 indicating that the annotation is definitely not an epileptiform transient and 205 indicating that the annotation is definitely an epileptic transient as explained in Table. 2.2. The overview of the model used for classification and detection is shown in Figure. 1.1 and 1.2

The data for detection and classification networks is balanced and divided into training and test set. The test set is used to analyze the performance of the classifiers. These trained networks are then used in validating the system by considering a test EEG data and the performance is analyzed.

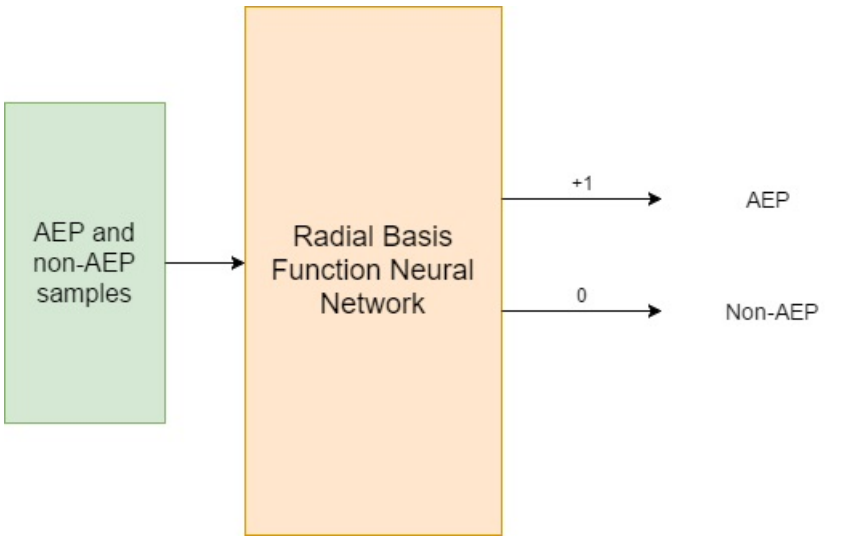


Figure 1.1: Block diagram of the classification stage Neural Network

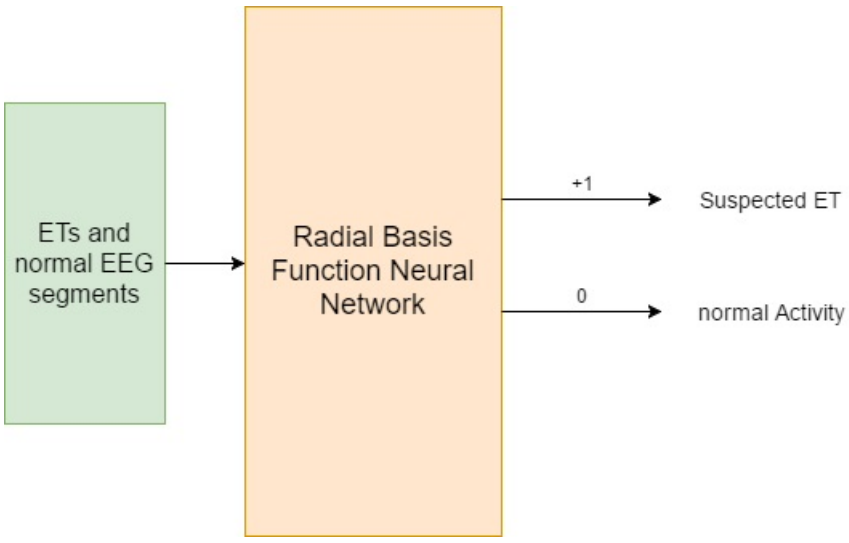


Figure 1.2: Block diagram of the ET Detection Neural Network

Chapter 2

Data Acquisition and Preprocessing

2.1 Electroencephalogram

Electroencephalography(EEG) is a noninvasive electrophysiological monitoring method of recording the electrical activity in the brain. These electrical activities in the brain are due to the voltage fluctuations resulting from the ionic current within the neurons in the brain [31]. Brain waves or neural oscillations are observed in EEG signals and has several diagnostic applications. EEGs are most commonly used in diagnosis of epilepsy. Prior to the current sophisticated techniques such as Magnetic Resonance Imaging (MRI) and computed Tomography (CT), EEGs were also used for diagnosis of tumors and focal brain disorders. It is also used to diagnose sleep disorders, coma and brain death.

A routine clinical EEG recording typically lasts 20-30 minutes and is recorded from the scalp electrodes. An internationally accepted standard of placement of electrodes is the 10-20 system. The distance between the electrodes is either 10 to 20 percent of the front-back or left-right distance of the skull [40]. The nomenclature is based on the brain lobe and numbers indicate whether the electrode is attached to the left or right hemisphere. Diagrammatic representation of 10-20 placement system on the scalp is shown in Figure. 2.1.

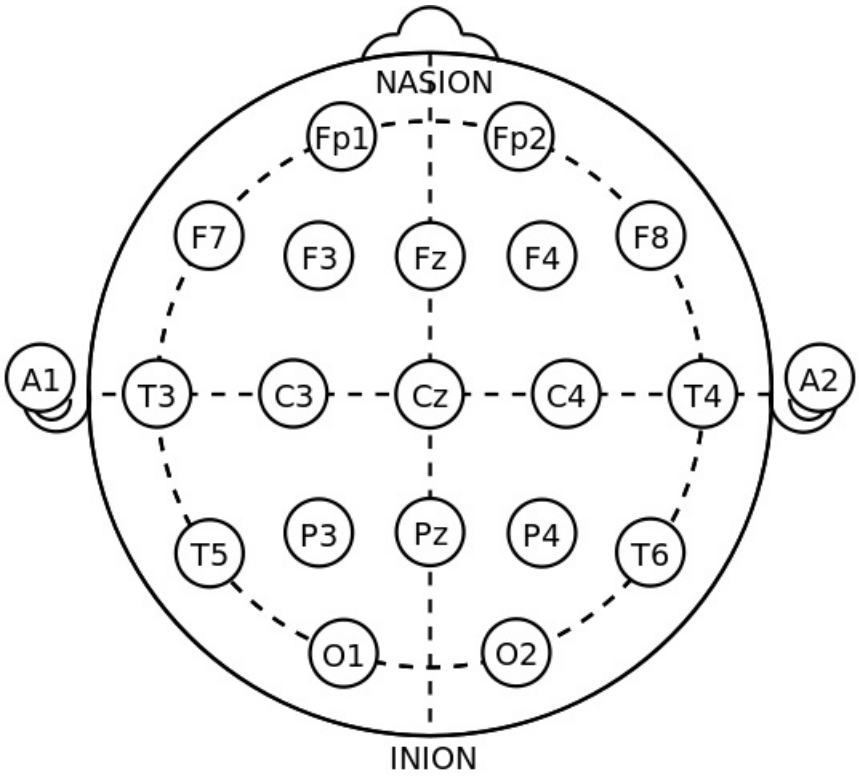


Figure 2.1: 10-20 electrode placement system

Electrode	**Lobe**
F	*Frontal*
T	*Temporal*
C	*Central**
P	*Parietal*
O	*Occipital*

Table 2.1: Nomenclature of Electrodes.

- The *Central lobe is not an actual lobe and is used only for identification purposes.
- Even numbers refer to electrode positions on the right hemisphere.
- Odd numbers refer to electrode positions on the left hemisphere.
- "z" refers to electrode positions on the mid line on the scalp.

A total of 21 electrodes are placed on different positions of the scalp as shown in Figure. 2.1. EEG recordings can be displayed in several formats called montages. A montage consists of a set of channels, where each channel is a difference between the voltage at two electrodes. There are three different types of montages used by neurologists for visual inspection of the EEG signal. [5]

- Bipolar: It is the difference between any two electrodes and montages contain a series of differential data of a series of electrode pairs.
- Referential: Each channel is the difference between a certain electrode and a common reference electrode connected at A1 or A2 or a mathematical average of electrodes attached to both earlobes.
- Average Reference: The output of all the electrodes are averaged and this signal is used as common reference for each channel.

There is no specific rule in selection of montages and the option of choosing a specific montage is based on the preference of each neurologist. The annotations in our data were marked on 47 unique channels on 7 different montages.

2.2 Data Acquisition

2.2.1 Raw Data

EEG of 200 patients were recorded for a duration of 30 seconds. This data was collected using the international standard of 10-20 system described in the previous section. For convenience, the data from all 200 patients were concatenated to form a single file of around 100 minutes duration and was provided to 18 board certified (American Board of Clinical Neurophysiology) neurologists. The data was visually inspected by experts and the epileptiform discharges were marked. This entire process was done in three phases.

Scores	Definition	Class
201	*Definitely not an Epileptiform Transient*	non-AEP
202	*Mostly not an Epileptiform Transient*	
203	*Can go either way*	not sure
204	*Mostly an Epileptiform Transient*	AEP
205	*Definitely an Epileptiform Transient*	

Table 2.2: Confidence factors used for annotations.

- Phase 1: The experts were asked to mark the location of the suspected epileptiform discharges (ETs) present in the entire 100 minute recording.
- Phase 2: All the markings or annotations were collected and any event marked by at least 2 experts were considered [6]. Approximately one month later, all the previously marked annotations were asked to be categorized into epileptiform or non-epileptiform transients (AEPs and non-AEPs) on a 5 point scale.
- Phase 3: To ensure reliability, another opinion or clarification was rendered for any inconsistencies in opinion from the previous two phases.

After the 3 phases, based on consistency, a list of most consistent or 'best 7' experts was formed. The annotations from the list of 'best 7' was then used as training data or ground-truth for the research.

Out of the 200 patient files, all the annotations came from 91 files and the rest 109 files that had no annotations was later used to extract the training data to represent the non-ET signals. Each annotation or epileptic transients vary in their duration depending on whether the ET is a spike or sharp wave or spike-slow wave complex etc [21]. Some statistical information about the duration of epileptiform transients are shown in the table 2.3 below.

Annotations	***Length***
Minimum	11
Mean	49
Median	39
Maximum	316

Table 2.3: Yellow Box duration statistics

2.2.2 Data Formats

The EEG recordings of 200 patients were recorded in three different sampling rates: 200 Hz, 256 Hz and 512 Hz based on the EEG machine setting. This data was filtered with a notch filter of 60 Hz to remove any interference of surrounding electrical signals or noise. This data is then stored in the European Data Format (.edf) [46]. Each (.edf) file has two objects (header and record). The header consists of all the information about the particular file such as sampling rate, signal information, electrodes etc. and the record consists of the actual signal recordings from the electrodes. This record data from all 200 files was combined to form a single 100 minute long file for annotations from the experts. The annotations from the experts were then obtained in the comma separated value (.csv) file containing the following details:

- Annotation ID: A unique ID corresponding to each annotation.
- Dataset ID: Number corresponding to the patient to which the annotation belongs. (1-200)
- Start Second: The start time of each event in the 100 minute long record
- End Second: The end time of each event in the 100 minute long record
- Montage ID: The number corresponding to the montage used by experts
- Channel Num: The channel on which the event was marked.
- User ID: Filename matching the dataset ID
- Classification score: Score assigned (from the 5 point scale) to the particular event by each experts

Additionally the information regarding the montage and channel numbers and the corresponding electrode combinations in the 10-20 system is also obtained for research purposes.

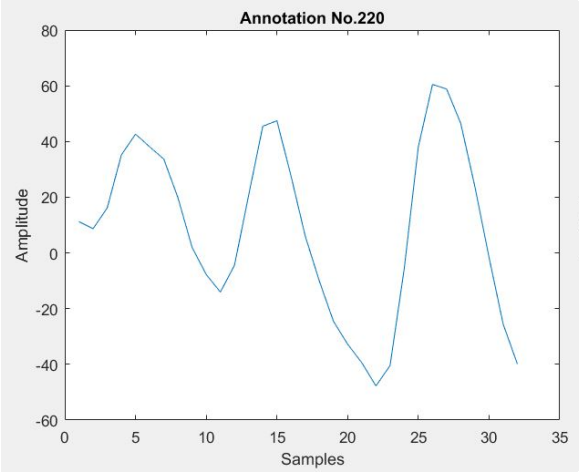
2.2.3 Data Processing

All the above information from various files was converted into a single unified database and stored in the 'mat' file for further processing. The matlab struct "overall_db" with following fields was created.

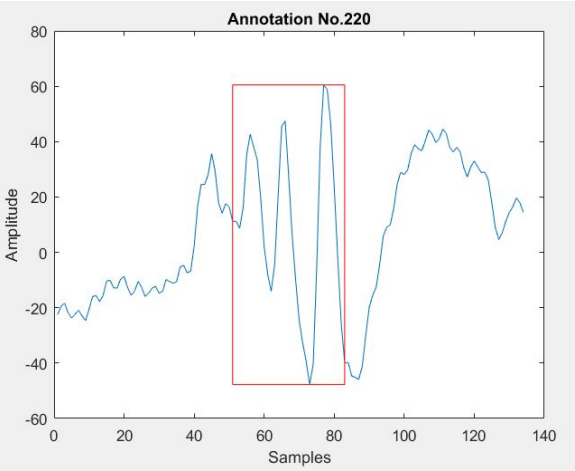
- Fields 1-6: These fields contain the information about the annotations, the annotation ID, start time, end time, total time, montage ID, Channel number.
- Fields 7-13: These fields contain the scores assigned to particular annotation by best 7 experts.
- Field 14: "montage_signal" consists of the entire 30 second recording of the particular channel on which the epileptiform transient was marked.
- Field 15: This field spike contains each annotation individually
- Field 16,17: These 2 fields "prespike" and "postspike" of around 50 samples before and after the yellow box, is the precontextual and postcontextual signal. This was used in feature extraction of ETs.
- Field: 18: contains information about the sampling frequency of the particular file to which the annotation belongs.
- Field 19: The field "inferred_groundtruth" states a single true value for the annotation. This was assigned by taking votes of the scores (5 point scale) and the true value was decided to be the score that had the maximum votes.

"Overlap": This struct contains information about all the start and end times belonging to a single patient and the number of annotations in that file.

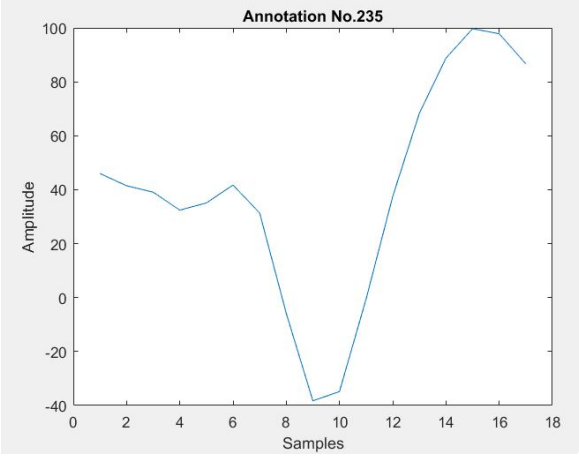
These variables contains all the basic information about each annotation which can be easily queried and can be directly used for further processing and feature engineering without need to understand the details of data acquisition. A few examples of the annotations belonging to both classes AEP and non-AEP with its confidence score levels are shown in left side of the Figure. 2.2 to 2.6. The "pre and post contextual" information along with the annotation are shown to the right. It can be observed that, even the transients belonging to the same class differs in its length and morphology.



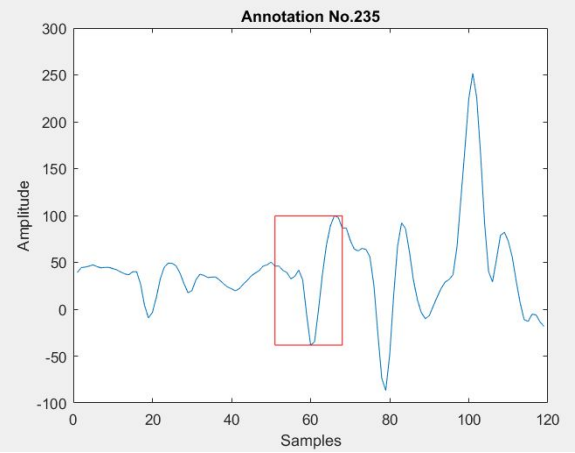
(a) Annotated signal



(b) Annotation with contextual data

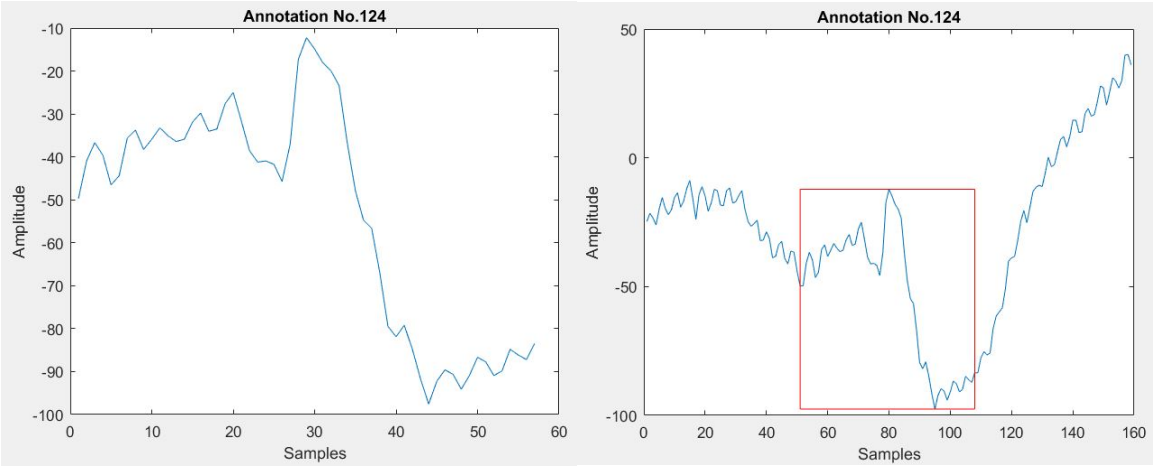


(c) Annotated signal



(d) Annotation with contextual data

Figure 2.2: 2 examples of Annotations taken from file 36 and file 65. Belongs to class AEP with a confidence level of 205 indicating that its definitely an epileptiform discharge.



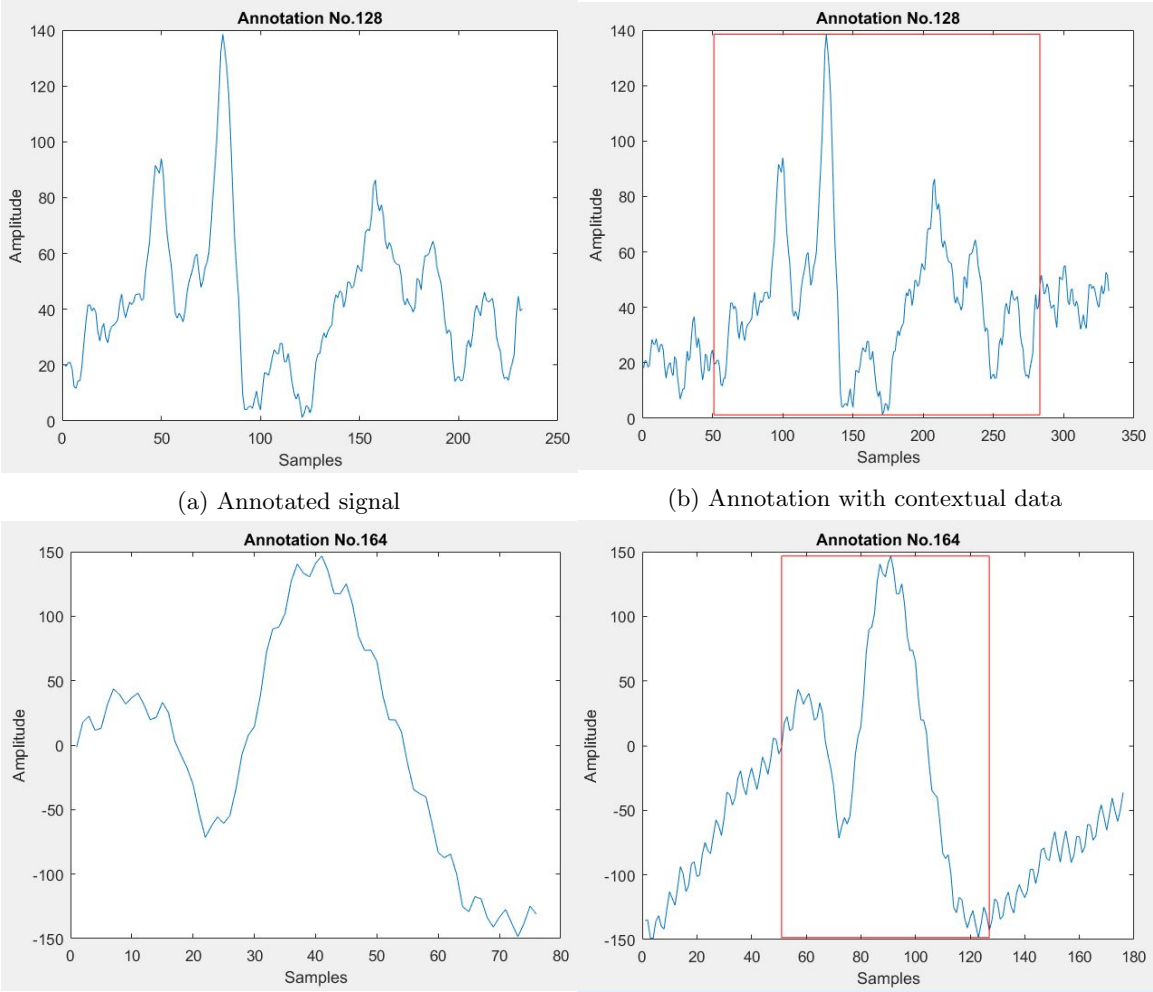
(a) Annotated signal

(b) Annotation with contextual data

(c) Annotated signal

(d) Annotation with contextual data

Figure 2.3: 2 examples of Annotations taken from file 138 and file 71. Belongs to class AEP with a confidence level of 204 which indicates the transient is most likely an epileptiform discharge.



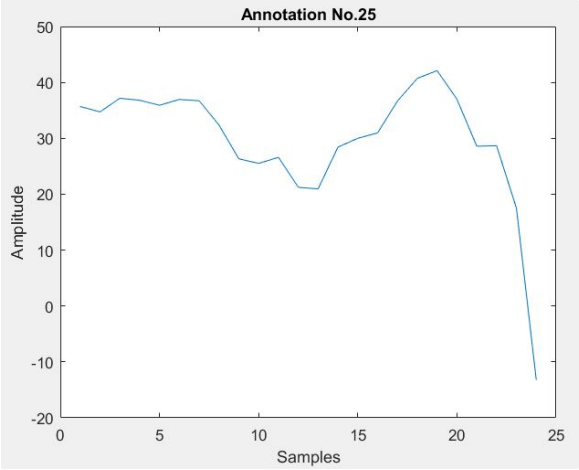
(a) Annotated signal

(b) Annotation with contextual data

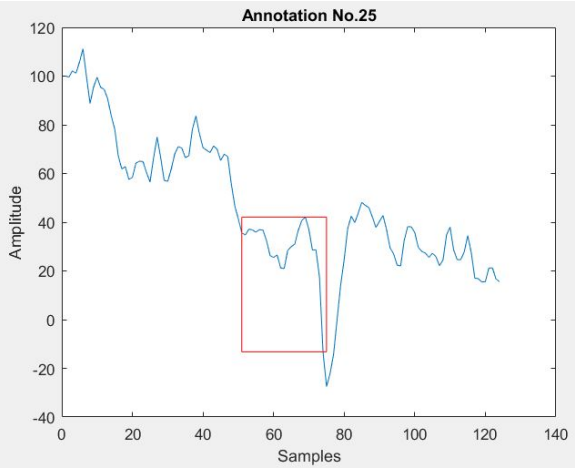
(c) Annotated signal

(d) Annotation with contextual data

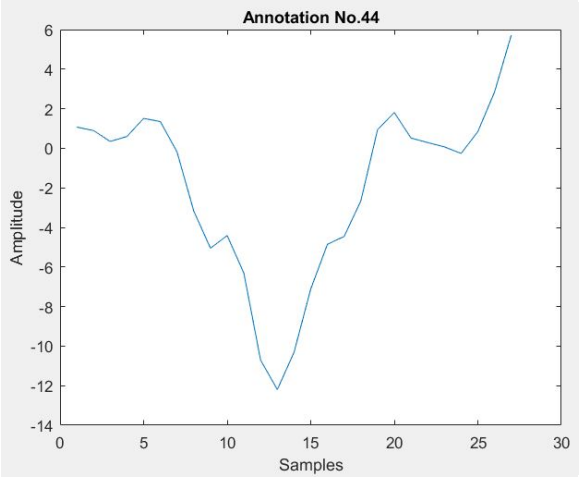
Figure 2.4: 2 examples of Annotations taken from file 108 and file 164, is scored as 203, this confidence level indicates that the experts are unsure whether it is an AEP or non-AEP.



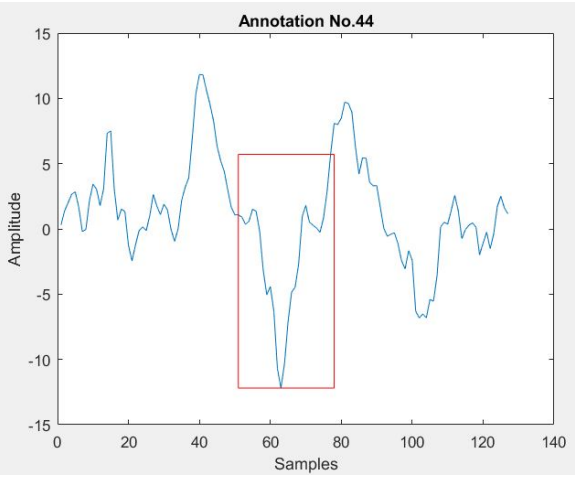
(a) Annotated signal



(b) Annotation with contextual data

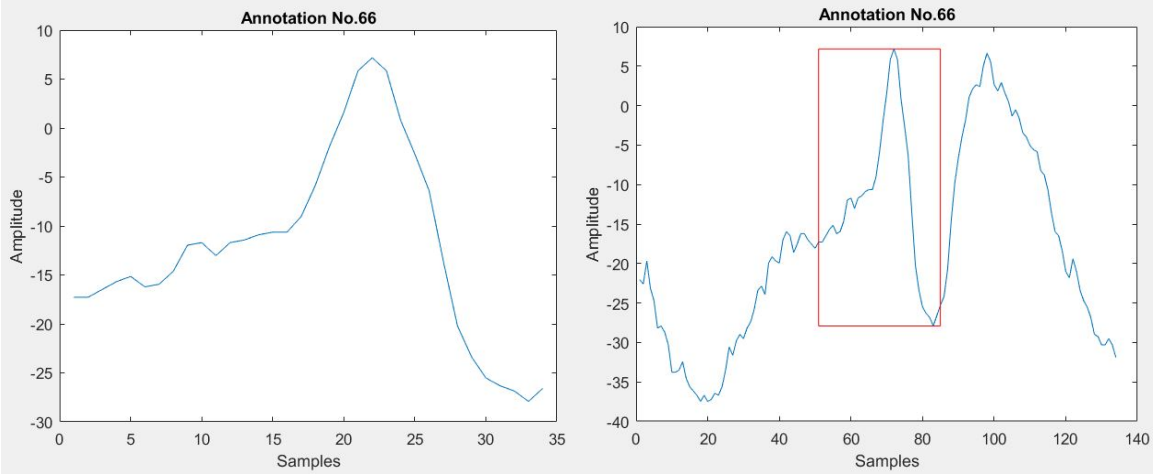


(c) Annotated signal



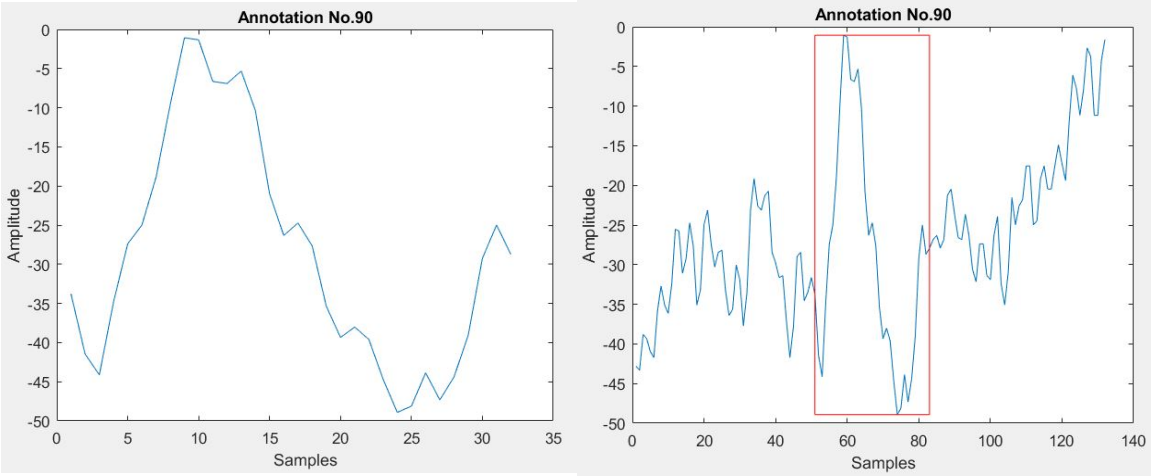
(d) Annotation with contextual data

Figure 2.5: 2 examples of Annotations taken from file 160 and file 106. Belongs to class non-AEP with a confidence level of 202 indicating that the signal is most likely not an epileptiform discharge



(a) Annotated signal

(b) Annotation with contextual data



(c) Annotated signal

(d) Annotation with contextual data

Figure 2.6: 2 examples of Annotations taken from file 30 and file 27. Belongs to class non-AEP with a confidence level of 201 indicating that the signal is definitely not an epileptiform discharge.

Chapter 3

Outline of HHT and RBFNN

3.1 Hilbert Huang Transform (HHT)

Traditional data-analysis methods are all based on stationary assumptions of a signal. Fourier transforms, for example, considers a fixed *'a priori'* basis of orthogonal functions and the transform does not depend on the nature of the analyzed signal. This kind of analysis is not very well suited for non-stationary signals. Wavelet transforms use a preset of basis functions, i.e. a mother wavelet used in the transform can be selected. Every component resulting from a wavelet transform has parameters that determine the scale and level or time for which it solves the problem, this addresses some of the constraints of non-stationarity of the signal.

Most natural physical processes are non-stationary and necessitates an adaptive basis function based on the nature of the analyzed signal. The Hilbert-Huang transform (HHT) by Huang et al. [19] provides a method to effectively analyze the complex signals which decomposes the signal into intrinsic mode functions (IMF) and obtain instantaneous frequency data. Most recently, HHT has also been used in the analysis of EEG signals and epileptic seizure detection [32] [30] [45].

HHT decomposes the signal in time domain and preserves the characteristics of the varying frequency. A brief description is provided by N.E.Huang in his book [18] which is also shown below. Hilbert-Huang transform consists of two parts as described below:

1. Empirical mode decomposition (EMD)

2. Hilbert Spectral Analysis (HSA)

3.1.1 Empirical mode Decomposition (EMD)

EMD is a method of decomposing the signal into finite number of components which form a complete and nearly orthogonal basis for the original signal, these are called intrinsic mode functions (IMFs) [19] [29], and are the functions into which the signal is decomposed. They are in time domain and are of same length as the original signal and hence the varying frequency information is preserved. Each intrinsic mode represents simple oscillations which will have the same number of extrema and zero crossings. Also, the oscillations will also be symmetric with respect to the local mean [19]. An IMF represents a simple oscillatory mode as a counterpart of the simple harmonic function but with variable amplitude and frequency as functions of time.

An IMF is constrained by the following conditions:

- In the entire signal, the number of extrema and the number of zero-crossings must be either equal or differ at most by one,
- At any point the mean value of the envelope defined by local maxima and envelope defined by local minima is zero [18].

The process of obtaining IMFs is called sifting, as explained below.

1. Identify all the local extrema from the signal $x(t)$.
2. Generate its upper and lower envelopes, $X_{u_{en}}(t)$ and $X_{l_{en}}(t)$ using cubic spline interpolation.
3. Calculate point by point mean from the upper and lower envelopes, i.e. $m_1(t)$

$$m_1(t) = mean(X_{u_{en}}(t) + X_{l_{en}}(t)) \tag{3.1}$$

4. Extract the difference between the data and $m_1(t)$ which is the first component $h_1(t)$

$$h_1(t) = x(t) - m_1(t) \tag{3.2}$$

   Test if the two conditions of IMFs are satisfied by $h_1(t)$. If it meets the two conditions, the first IMF is derived.

5. If $h_1(t)$ does not satisfy the conditions of IMF, it is considered to be a proto-IMF and in the next step it is treated as data, can be written as,

$$h_{11}(t) = h_1(t) - m_{11}(t) \tag{3.3}$$

where $m_{11}(t)$ is the mean from the upper and the lower envelopes generated from treating $h_1(t)$ as data and $h_{11}(t)$ is the new proto-IMF.

6. The above steps are repeated k times until an IMF is derived, given by,

$$h_{1k}(t) = h_{1k-1}(t) - m_{1k}(t) \tag{3.4}$$

$$c_1(t) = h_{1k}(t) \tag{3.5}$$

where $c_1(t)$ from Eqn.(3.5) is the first derived intrinsic mode function (IMF), $h_{1k-1}(t)$ and $m_{1k}(t)$ are the proto-IMF and mean of the envelopes from the previous step.

7. Now $c_1(t)$ can be separated from the rest of the data by

$$r_1(t) = x(t) - c_1(t) \tag{3.6}$$

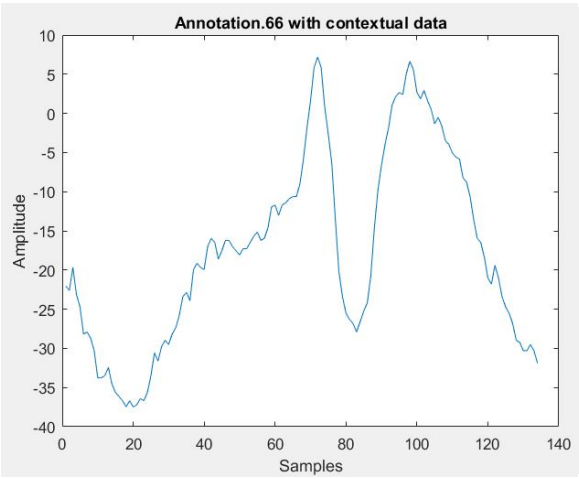
$r_1(t)$ is residue and can still contain longer period variations in the data, which can be assumed as new data and new IMFs are derived.

8. This sifting process is repeated until a predetermined criteria is met, i.e. when the residue is a monotonic function, from which no more IMFs can be extracted, which can be written as,

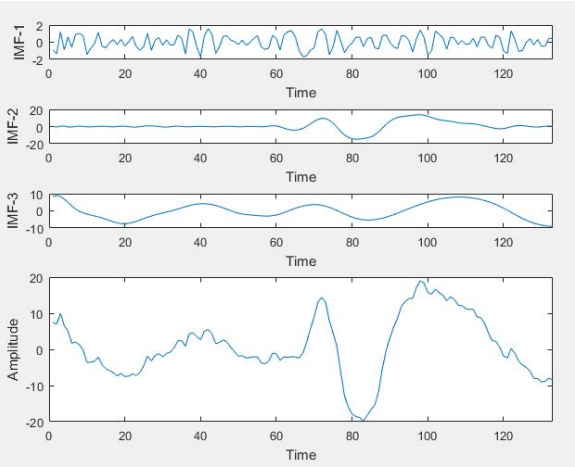
$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \tag{3.7}$$

where n is the number of IMFs, $r_n(t)$ denotes the final residue, which represents the mean trend or constant.

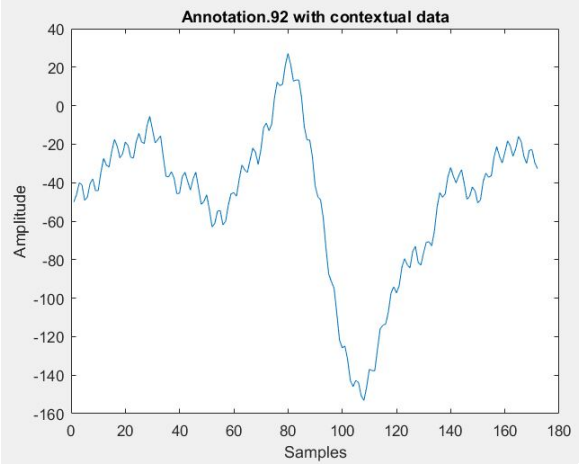
A few examples of the IMFs extracted on the EEG signals are shown in Figure.3.1 and 3.2. It can be seen that the IMFs are of the same length as the original signal and the sum of first three IMFs closely resemble the original signal, and hence, the time information is preserved. The phase and frequency of each IMFs obtained helps in the time-frequency analysis, which is an important information in distinguishing ETs from non-ETs.



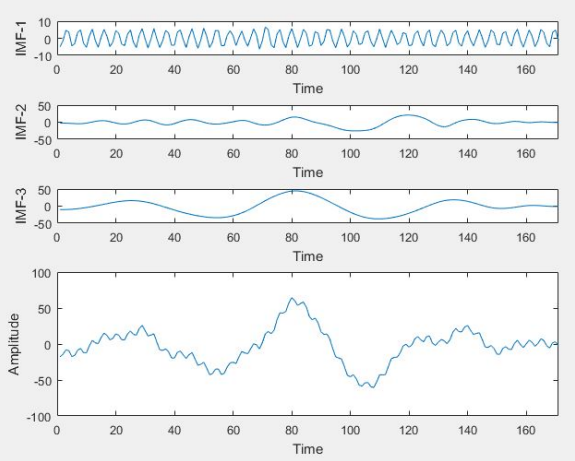
(a) Annotation example of class 201



(b) IMFs from the EEG yellow box

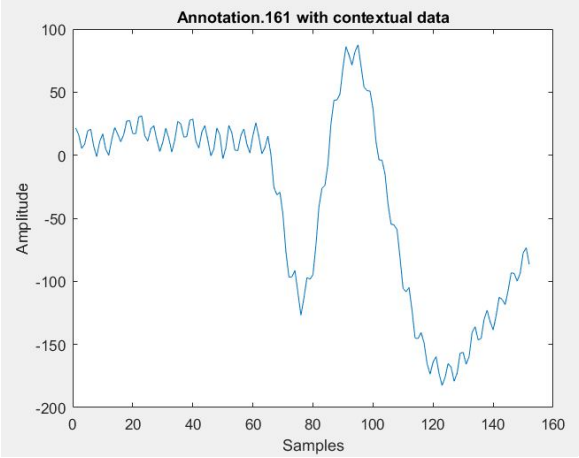


(c) Annotation example of class 202

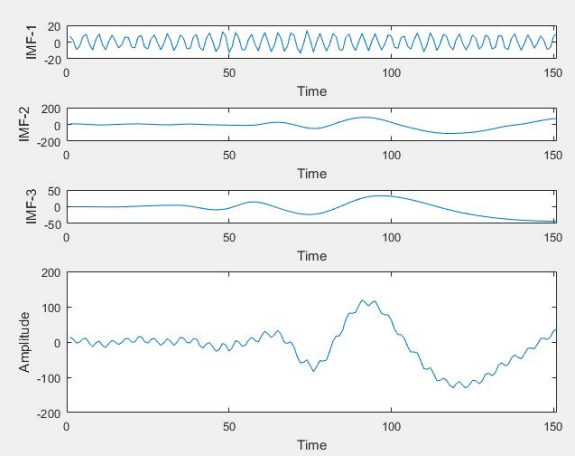


(d) IMFs from the EEG yellow box

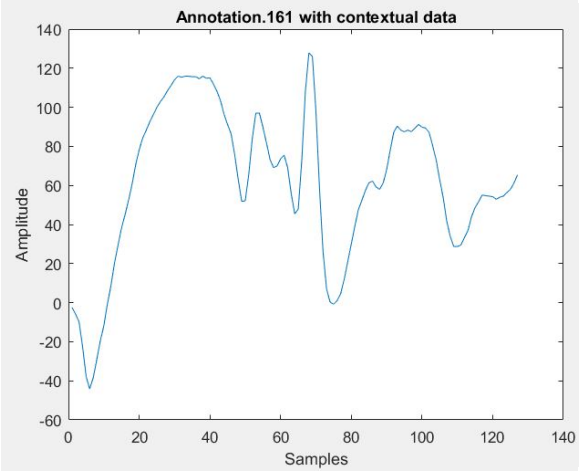
Figure 3.1: Examples of IMFs extracted from annotations belonging to class 201 and 202 respectively. It can be seen that the sum of the first three IMFs preserves most of the time series information.



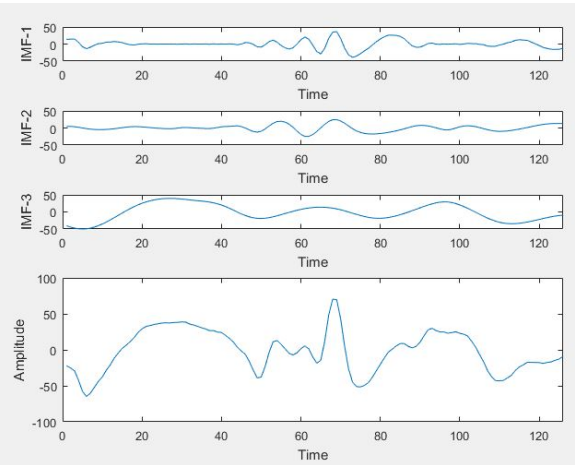
(a) Annotation example of class 204



(b) IMFs from the EEG yellow box



(c) Annotation example of class 205



(d) IMFs from the EEG yellow box

Figure 3.2: Examples of IMFs belonging to annotations from class 204 and 205 respectively. The three IMFs preserve most of the time series information and the frequency analysis of each IMFs provide both the time and frequency information of the entire signal.

3.1.2 Hilbert Spectral Analysis

Hilbert spectral analysis is a method applying the Hilbert transform to compute the instantaneous frequency [19]. The Hilbert transform creates an analytic signal by defining an imaginary component and provides instantaneous amplitude and phase with respect to time. The Hilbert transform is applied to each IMF and instantaneous frequency is computed. Given a real signal $x(t)$, its analytic signal $z(t)$ is given by,

$$z(t) = x(t) + ix_H(t) \tag{3.8}$$

Where $x_H(t)$ is the Hilbert transform of $x(t)$ [26], i.e.

$$x_H(t) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \left(\int_{t-1/\epsilon}^{t-\epsilon} \frac{x(\tau)}{t-\tau} d\tau + \int_{t+\epsilon}^{t+1/\epsilon} \frac{x(\tau)}{t-\tau} d\tau \right) \tag{3.9}$$

The complex representation of the signal in its polar form is written as,

$$Z(t) = A(t)e^{i\varphi(t)} \tag{3.10}$$

Where,

$$A(t) = \sqrt{x(t)^2 + x_H(t)^2} \tag{3.11}$$

$$\varphi(t) = \arctan\left(\frac{x_H(t)}{x(t)}\right) \tag{3.12}$$

The time derivative of the polar form is

$$z'(t) = A(t)e^{i\varphi(t)}(iw(t)) + e^{i\varphi(t)}A'(t) \tag{3.13}$$

Where w(t) is the instantaneous angular frequency, which by definition is the time derivative of the instantaneous angle.

$$w(t) = \varphi'(t) = \frac{d}{dt}\left(\arctan\left(\frac{x_H(t)}{x(t)}\right)\right) \tag{3.14}$$

$$f(t) = \frac{w(t)}{2\pi} \tag{3.15}$$

After performing the Hilbert transform on each IMF, the original signal can be expressed as the real part in the following form.

$$x(t) = Re\sum_{j=1}^{n} A_j(t)e^{i\int (w_j(t))dt} \tag{3.16}$$

where $w_j(t)$ is the derived instantaneous frequency of each IMF except the residue $r_n(t)$ which is usually constant.

This gives both the amplitude and frequency of each component as functions of time. This frequency-time distribution of the amplitude is called Hilbert amplitude spectrum or Hilbert spectrum. [18]

With the above method, intrinsic functions (IMFs) of the EEG signal are extracted for the annotated part along with the "pre and post" contextual signal surrounding the annotated signal.

3.1.2.1 Feature Extraction

The following features were extracted from each IMF based upon previous work. [20]

Weighted Mean Frequency: It is one of the most useful feature of HHT which is also called the Hilbert weighted frequency [32]. Weighted frequency is computed by taking the weighted average of instantaneous frequency with instantaneous amplitude acting as weights. Weighted frequency is the indicative of the energy distribution in each IMF. This frequency w_f gives an idea about the mean frequency using instantaneous frequency f and amplitude A over an interval from the point index 1 to point index l.

$$w_f = \frac{\sum_{i=1}^{l}(A(i)f^2(i))}{\sum_{j=1}^{l}(A(j)f(j))} \tag{3.17}$$

Temporal Statistics of Analytic IMFs: The statistical features of IMF are useful features due to the asymmetry, dispersion and concentration around the mean. The IMFs from healthy EEGs and non-healthy EEGs during interictal periods of epilepsy tend be different [34]. These differences can be analyzed using the statistics of the IMFs. Hence, statistical features (mean μ, variance σ, skewness β and kurtosis κ) are computed for each IMF.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{3.18}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \mu)^2} \tag{3.19}$$

$$\beta = \frac{1}{N}\sum_{i=1}^{N}(\frac{y_i - \mu}{\sigma})^3 \tag{3.20}$$

$$\kappa = \frac{1}{N}\sum_{i=1}^{N}(\frac{y_i - \mu}{\sigma})^4 \tag{3.21}$$

Mean absolute deviation in each IMF: Average or mean absolute deviation is the mean of absolute deviations of a set of data about the data's mean. For a sample size N and the mean distribution x, the mean absolute deviation is defined by (MAD) [20].

$$MD = \frac{1}{N}\sum_{i=1}^{N}|x_i - \mu| \tag{3.22}$$

where N is the number of samples in the IMF.

Interquartile range in each IMF: Interquartile deviation is a measure of the statistical dispersion being equal to the difference between 75^{th} and 25^{th} percentiles. It is a measure of variability, based on dividing the data set into quartiles [20] [42].

$$IQR = Q3 - Q1 \tag{3.23}$$

Where $Q1$ and $Q3$ are the first and third quartile respectively.

Spectral Statistics of IMFs: Previous research suggests that epileptic seizures give rise to changes in certain frequency bands [2] [41]. The spectral features extracted from IMFs thus provide information about the physiology of the EEG signals. Power spectral density [34] has been calculated for each IMFs and statistics of PSD from each IMFs are used as features.

$$P(f) = \sum_{-\infty}^{\infty} r_y[n]e^{-jwn} \tag{3.24}$$

Where f is the frequency bin and $r_y[n]$ represents the autocorrelation of $y[n]$ given by ,

$$r_y[n] = E(y[n]y*[n]) \tag{3.25}$$

Spectral Entropy: Entropy is defined as the measure of degree of disorder [23]. In the context of EEG signals, entropy can be considered to be a measure of the complexity of the time series. Spectral entropy uses the power spectrum of the signal to estimate the regularity of the time series. Spectral entropy is evaluated using the normalized Shannon entropy applied to the power spectral density of each IMF [39]. If the power spectral estimate for each frequency is denoted by P_f, then the spectral entropy is calculated using the formula:

$$ShEn = \sum_f (p_f log(p_f)) \tag{3.26}$$

Where p_f is the normalized power spectral estimate calculated using the total power $\sum P_f$ and dividing the power level corresponding to each frequency by the total power.

Weiner Entropy: Weiner entropy or the flatness is mostly used in the characterization of the audio signals. A higher spectral flatness indicates that the power is distributed uniformly across the spectral bands and a low spectral flatness indicates that the power is concentrated on smaller number of bands. It is given as the ratio between the geometric mean of the power spectral density estimate of the IMFs to the arithmetic mean of the power spectral density estimate of the IMFs.

$$WnEn = \frac{(\prod_f p_f)^{\frac{1}{N}}}{\sum_f p_f} \tag{3.27}$$

Spectral Centroid: Spectral centroid is used to characterize a spectrum which indicates where the center of mass of the spectrum is. When applied to IMFs extracted from the EEG signals it represents the centroid of the frequencies in each IMF. The centroid frequencies of the IMFs extracted from EEG signals form distinctive groups for seizure and nonseizure EEG signals [32].

The centroid frequency is calculated using the formula:

$$C_s = \frac{\sum_f f P_f}{\sum_f P_f} \tag{3.28}$$

where P_f is the power corresponding to the f^{th} frequency bin.

Variation Coefficient: Like the spectral centroid, spectral variation coefficient can also be used for characterization of EEG signals. Spectral variation in the IMFs is said to be different for normal and pathological EEG signals [34]. The variation coefficient can be calculated using the formula:

$$\sigma_s^2 = \frac{\sum_f (f - C_s)^2 P_f}{\sum_f P_f} \tag{3.29}$$

Where C_s is the spectral centroid of the IMF.

Spectral Skew: Skewness is the third order moment and it is a measure of the asymmetry of a distribution. Skewness of each IMF is computed as

$$\beta_s = \frac{\sum_f (\frac{f - C_s}{\sigma_s})^3 P_f}{\sum_f P_f} \tag{3.30}$$

The number of IMFs generated by the EEG signals varies based on the length of the signal and the intrinsic mode oscillations present in the signal. A visual investigation indicated that most of the properties of the original signal is preserved in the first few IMFs and hence the above mentioned features are extracted from the first three IMFs resulting in a 33 dimensional feature vector which is used for this research.

The list of features extracted from each IMF is shown below.

$$Features_{IMF_j} = [w_f, \mu, \sigma, \beta, \kappa, MD, IQR, SpEn, C_s, \sigma_s, \beta_s]^T, 1 \leq j \leq 3 \tag{3.31}$$

3.2 Radial Basis Function Neural Network

3.2.1 Basic RBF Network

Radial basis function network is a feed-forward network with a modified hidden layer and training algorithm which may be used for mapping. A basic RBF network consists of single hidden layer of locally sensitive units, i.e. the unit's response decreases monotonically with the distance of input vector i, from the units receptive field center w_{ij} [37].

3.2.1.1 Network Architecture

An RBF network consists of hierarchy of units, that are organized to form consecutive layers. The input is fed to these interconnected layers through the input layer, and output layer (in most cases) consists of linear units. There exists intermediate or hidden layers between the input and output layers. The RBF hidden layer employs a Gaussian activation function with euclidean distance measure which may be described as,

$$o_{ij} = f(net_{ij}) = e^{(\frac{-||i-w_{ij}||^2}{2\sigma^2})} \tag{3.32}$$

where, net activation of the j^{th} unit in the i^{th} layer, net_{ij} is given by,

$$net_{ij} = \frac{-||i-w_{ij}||^2}{2\sigma^2} \tag{3.33}$$

w_{ij} the weight vector or unit center corresponding to the j^{th} unit in the i^{th} layer.

The unit has maximum net activation net_{ij} and correspondingly maximum output o_{ij} when $i = w_{ij}$, (i.e. $net_{ij} = 0$). In practice, for a given input, only a fraction of the hidden units will be active (net activation, $net_{ij} \geq 0$) that leads to a non-zero output [37].

Output Units are in most cases simple linear units which is just the weighted linear combination of the output from the previous layer, given by

$$o_{ij} = net_{ij} = \sum_{j=1}^{n} w_{ij}^{h-o} o_{i-1j} \tag{3.34}$$

where, o_{i-1j} are the outputs of j^{th} unit from the $i-1^{th}$ layer (hidden layer) and w_{ij}^{h-o} are weights connecting the output of the hidden layer to output layer.

A block diagram of simple RBF network is shown in Figure.3.3

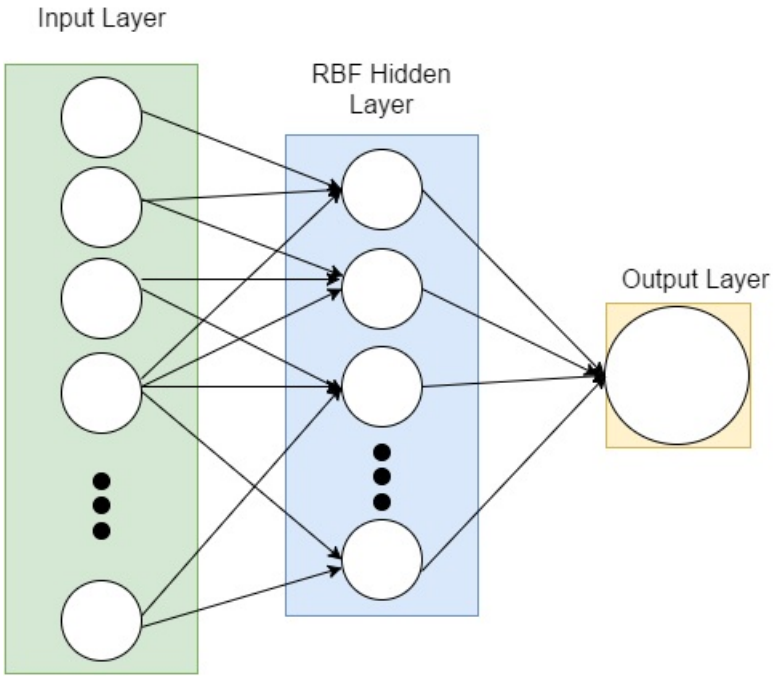


Figure 3.3: A Basic RBF Network

3.2.1.2 Network training

The process of training involves two steps:

1. Determination of RBF unit centers which can be accomplished in different ways.
   - Using the K-means clustering, initialized from randomly chosen points from the training set.
   - Using the Kohonens SOFM which forms individual neural clusters self-organized to reflect input pattern similarity [37]. Then select unit weight vectors from the map as the RBF unit centers.
2. The training of RBF network consists of determining weights connecting RBF to output layer which can be trained using the psuedoinverse or gradient descent methods [37] [25].

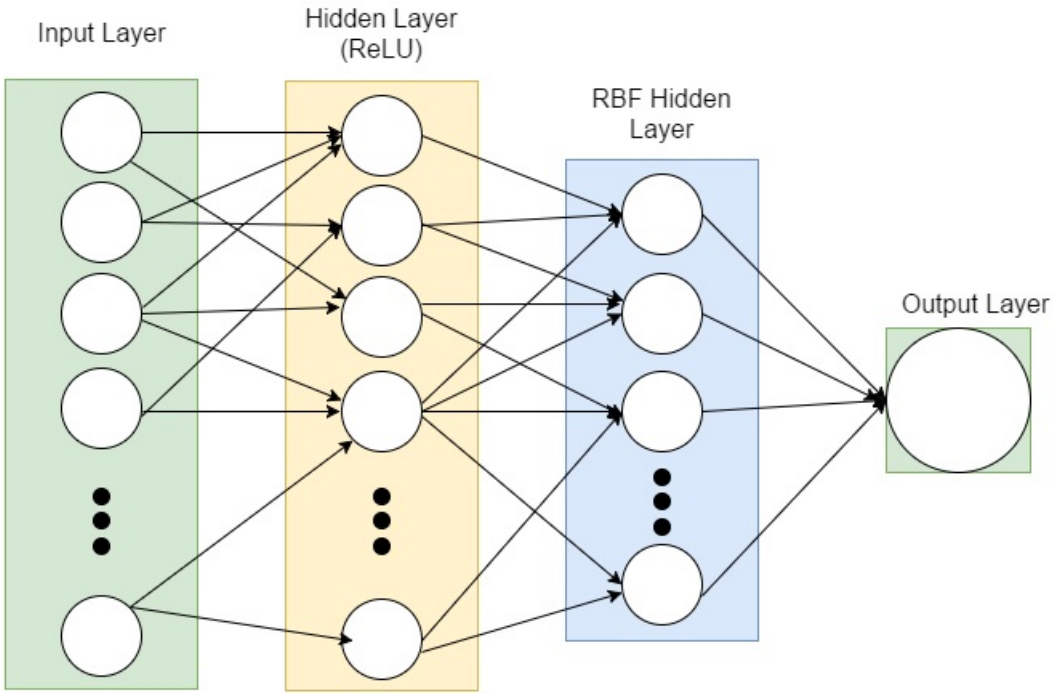


Figure 3.4: Representation of Multilayer RBF Network

3.2.2 Multilayer RBF Network

A multilayer RBF network is designed by adding an extra hidden layer before the RBF layer. Although there is no clear proof of significant improvement in the results by adding extra layers in the RBF network, the advantage of using deep networks is that it can compactly represent a significantly larger group of functions than simpler networks. In this method an extra hidden layer with RELU non-linearity [13] has been added in between the input and the RBF hidden layer. The network is trained using back-propagation algorithm [37].

The block diagram representation of the network is shown in Figure.3.4

3.2.2.1 L2-Regularization

The number of iterations to train the network is a hyper parameter which is to be chosen carefully to achieve small error on the training data but also make sure the network generalizes well on the test data. Training of the network involves adjusting the network weights to minimize the cost function, so that, for the training data, the predicted output closely match the known correct

outputs. Unfortunately, if the network is over-trained the network adjusts the weights so as to perfectly predict the correct values on every training example. These weights when used on test data that's previously unseen by the network would perform poorly. This is called over-fitting. The over-fitted weights tend to be large and dominate the output prediction [7].

In order to avoid the overfitting it is possible to regularize the cost function. L2-Regularization technique restricts the values of the weights by penalizing the sum of squared values of the weights to avoid any weight dominating the output prediction [13]. This regularization term in the update of weights causes weight decay in proportion to its size. Weight decay is an additional term in the weight update rule that causes the weight to exponentially decay to zero, if no other update is scheduled [7].

3.2.2.2 Hidden Layer with ReLU nonlinearity

The multilayer RBF network uses ReLU nonlinearity in the hidden layer between input and RBF layer. Rectified linear units (ReLU) use the activation function given by $g(z) = \max(0, z)$. ReLU units are easy to optimize because they are similar to linear units. The derivatives through a rectified linear unit remain large whenever the unit is active. The gradients are consistent and the second order effects are almost zero everywhere and the derivative of the rectifying operation is one when the unit is active. Thus the gradient direction is far more useful for learning than when compared to the other activation functions that introduce second-order effects [13].

Formulae for forward and back-propagation used in the algorithm are given below:

Forward Pass hidden layer:

$$net_{IH} = (x.w_H + b_H) \tag{3.35}$$

$$o_{IH} = \max(0, net_{IH}) \tag{3.36}$$

Forward pass RBF layer :

$$o_{HR} = e^{net_{HR}} \tag{3.37}$$

where net_{HR} is,

$$net_{HR} = \frac{-||o_{IH} - \mu||^2}{2\sigma^2} \tag{3.38}$$

Forward pass output layer:

$$o_{RO} = o_{HR}.w_O + b_O \tag{3.39}$$

Sum of squared error loss function is considered in the training of the network, given by:

$$loss = \frac{1}{2}(o_{RO} - t)^2 + \frac{1}{2}(reg * ||w_O^2|| + reg * ||w_H^2||) \tag{3.40}$$

where reg is the L2-regularization coefficient.

$$do_{RO} = (o_{RO} - t) \tag{3.41}$$

Backward pass output layer:

$$dw_O = o_{HR}^T.do_{RO} \tag{3.42}$$

$$do_{HR} = do_{RO}.w_O^T \tag{3.43}$$

$$db_O = \sum_{i=1}^{m} do_{RO} \tag{3.44}$$

where M is the number of training samples in each epoch.

Backward pass RBF layer:

$$d\mu_{_j} = \sum_{i=1}^{M} do_{HR}.o_{HR}.\frac{(o_{IH_i} - \mu_{_j})}{\sigma_{_j}^2} \tag{3.45}$$

$$d\sigma_{_j} = \sum_{i=1}^{M} -do_{HR}.o_{HR}.\frac{||o_{IH_i} - \mu_{_j}||^2}{\sigma_{_j}^3} \tag{3.46}$$

$$do_I H_{_i} = -\sum_{j=1}^{k} do_{HR}.o_{HR}.\frac{(o_{IH_i} - \mu_{_j})}{\sigma_{_j}^2} \tag{3.47}$$

Backward pass ReLU Hidden layer:

$$dnet_{IH} = (do_{IH}(o_{IH} <= 0) = 0) \tag{3.48}$$

$$dx = dnet_{IH}.w_H^T \tag{3.49}$$

$$dw_H = x^T.dnet_{IH} \tag{3.50}$$

$$db = \sum_{i=1}^{M} dnet_{IH} \tag{3.51}$$

Where M is the number of training examples, k is the number of RBF centers in the network. The choice of hyper parameters like the number of receptive units k in the RBF layer, the learning rate η were decided based on the cross-validation tests covered in the subsequent chapters. The above network was used for classification of epileptic transients in both stages (detection and classification).

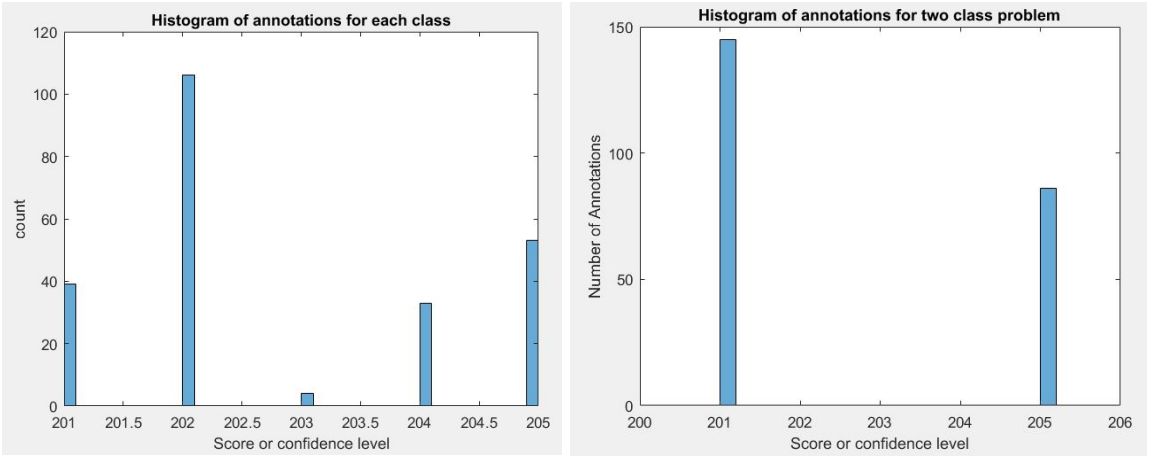
Chapter 4

Classification of Yellow Boxes

Classification of the yellow boxed signals involves mapping the annotations or yellow boxes into AEP or non-AEP classes. As mentioned in the previous chapter, a multilayer RBF neural network (Figure.3.4) is used for this purpose. Neural network is trained using the annotated data provided by the most consistent seven experts from the EEG recordings of 200 patients.

4.1 Data

As explained in chapter 2, EEG recordings from 200 patients were collected and provided to experts for visual inspection and identification of the annotations. The scores given by seven most consistent experts were considered. The scores were on a 5 point scale from 201 to 205. This problem is modified into a two class problem of AEP and non-AEP. The scores of 204/205, which indicated the annotation to be mostly AEP, is considered to be class AEP and the scores of 201/202, which indicated the annotations to be mostly non-AEP, are considered to be class non-AEP. The data with score 203 had no consensus between the experts and hence is left out from this study. There are a total of 231 annotations out of which 86 belong to class AEP and 145 to class non-AEP. This imbalanced dataset was balanced using max distance criteria [16] and then divided into training and test set.



(a) Count of annotations from each class 201 to 205 (b) Count of annotations for a two class problem

Figure 4.1: Histogram of Yellow Boxes

The visual identification of epileptic transients are performed by experts by comparing the signal with the surrounding data. Thus the contextual information around the transients are also included to compute the features from the annotations. A typical length of an annotation lasts around 200ms which corresponds to around 50 samples at sampling frequency of 256Hz. Therefore, a window size of 50 samples of pre-contextual and post-contextual signals has been considered. A histogram plot of the length of annotations are shown in Figure.4.2

4.2 Classifier Neural Network

4.2.1 Design

A multilayer Radial basis function neural network is trained to classify the epileptiform transients into two classes (AEP and non-AEP). The hidden layer between the input units and the RBF layer consists of a Rectified Linear Unit activation(ReLU) [13] non-linearity, explained in the previous chapter. The RBF layer unit receptive centers are chosen using the k-means algorithm, the activations of the RBF units are computed based on the distance between the test input and the RBF unit's center. The output layer consists of a single unit which is a linear weighted combination of the output from the RBF layer. The desired output of +1 corresponds to an AEP and an output of 0 is classified as non-AEP. The number of centers for the RBF is chosen based on the performance of the RBF network. The number of ReLU units in the hidden layer is the equal to the dimension

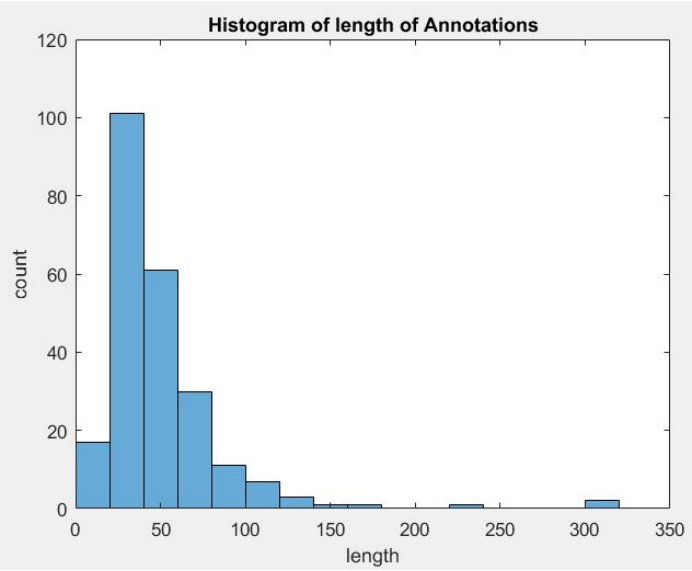


Figure 4.2: Histogram of length of annotations

of the input feature vector.

4.2.2 Training

The training of the network involves two steps as described in Chapter 3. First step is to choose the centers or receptive field of the RBF units which is computed using the standard k-means algorithm with correlation distance measure. The centroids from the k-means algorithm was considered as the RBF receptive field centers. The entire network is then trained by back propagating the error.

4.2.3 Cross Validation

Cross validation is a method of validating the model and to gain an insight of how well the model performs on an independent dataset. In a k-fold cross validation, the dataset is divided into k groups or subsets and k-1 subsets are used for training the network. The other remaining subset is used as test set to analyze the performance of the network. The process is repeated k times with each of the k subsets acting as test set exactly once. This gives a statistical prediction as to how the model would perform on an independent data or an instance from the real problem. Cross validation combines the results of model and averages them to provide a more accurate estimate of

the performance of the model.

The dataset is balanced using the max distance criteria [16] and the balanced dataset is used in cross validation to contain equal representations of AEPs and non-AEPs. The value of k was varied from 4 to 10. The table 4.1 shows the number of samples from each class (AEP and non-AEP) in the subset.

4.2.4 Parameters

Initially the weights of the units with ReLU activation and the output unit is initialized with values of magnitude between 0.005 to 0.05. The centers for the RBF is chosen by k-means algorithm with a k value of 8. The network is trained for 15000 iterations with a learning rate of 0.001. The value for L2-Regularization to prevent the network from over fitting was chosen to be 0.0001.

4.3 Results

The training of neural network can be seen on the graph plot of Error with respect to training iterations. The graph has been plotted for each value of k-fold. The error plot shows the error decreasing with respect to the iterations on the k^{th} training set. The Figure. 4.3 to 4.9 shows the error plot for value of k from 4 to 10. This ensures the network trains well to discriminate the AEPs from the non-AEPs.

No. of Folds	*Trial*	*AEP samples*	*non-AEP samples*	*Ratio Training:Test*	*Total*
4	**1**	**21**	**21**	**3.09**	**42**
	2	**22**	**22**	**2.90**	**44**
	3	**22**	**22**	**2.90**	**44**
	4	**21**	**21**	**3.09**	**42**
5	**1**	**17**	**17**	**4.05**	**34**
	2	**18**	**18**	**3.77**	**36**
	3	**17**	**17**	**4.05**	**34**
	4	**17**	**17**	**4.05**	**34**
	5	**17**	**17**	**4.05**	**34**
6	**1**	**14**	**14**	**5.14**	**28**
	2	**15**	**15**	**4.73**	**30**
	3	**15**	**15**	**4.73**	**30**
	4	**14**	**14**	**5.14**	**28**
	5	**14**	**14**	**5.14**	**28**
	6	**14**	**14**	**5.14**	**28**
7	**1**	**12**	**12**	**6.16**	**24**
	2	**13**	**13**	**5.61**	**26**
	3	**13**	**13**	**5.61**	**26**
	4	**12**	**12**	**6.16**	**24**
	5	**12**	**12**	**6.16**	**24**
	6	**12**	**12**	**6.16**	**24**
	7	**12**	**12**	**6.16**	**24**
8	**1**	**10**	**10**	**7.60**	**20**
	2	**11**	**11**	**6.81**	**22**
	3	**11**	**11**	**6.81**	**22**
	4	**11**	**11**	**6.81**	**22**
	5	**11**	**11**	**6.81**	**22**
	6	**11**	**11**	**6.81**	**22**
	7	**11**	**11**	**6.81**	**22**
	8	**10**	**10**	**7.60**	**20**
9	**1**	**9**	**9**	**8.55**	**18**
	2	**10**	**10**	**7.60**	**20**
	3	**10**	**10**	**7.60**	**20**
	4	**10**	**10**	**7.60**	**20**
	5	**10**	**10**	**7.60**	**20**
	6	**10**	**10**	**7.60**	**20**
	7	**9**	**9**	**8.55**	**18**
	8	**9**	**9**	**8.55**	**18**
	9	**9**	**9**	**8.55**	**18**
10	**1**	**8**	**8**	**9.75**	**16**
	2	**9**	**9**	**8.55**	**18**
	3	**9**	**9**	**8.55**	**18**
	4	**9**	**9**	**8.55**	**18**
	5	**9**	**9**	**8.55**	**18**
	6	**9**	**9**	**8.55**	**18**
	7	**9**	**9**	**8.55**	**18**
	8	**8**	**8**	**9.75**	**16**
	9	**8**	**8**	**9.75**	**16**
	10	**8**	**8**	**9.75**	**16**

Table 4.1: Distribution of samples during k-fold validation

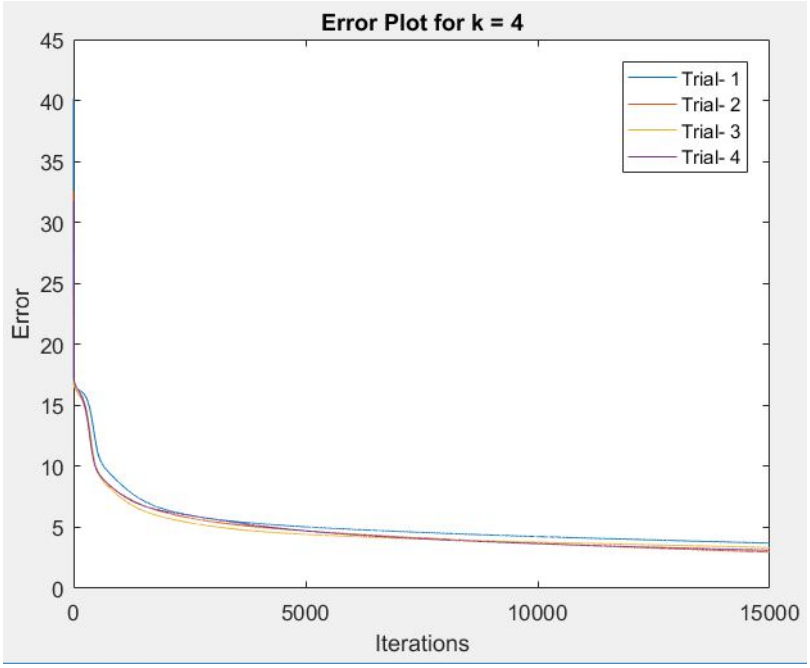


Figure 4.3: Error plot for k = 4

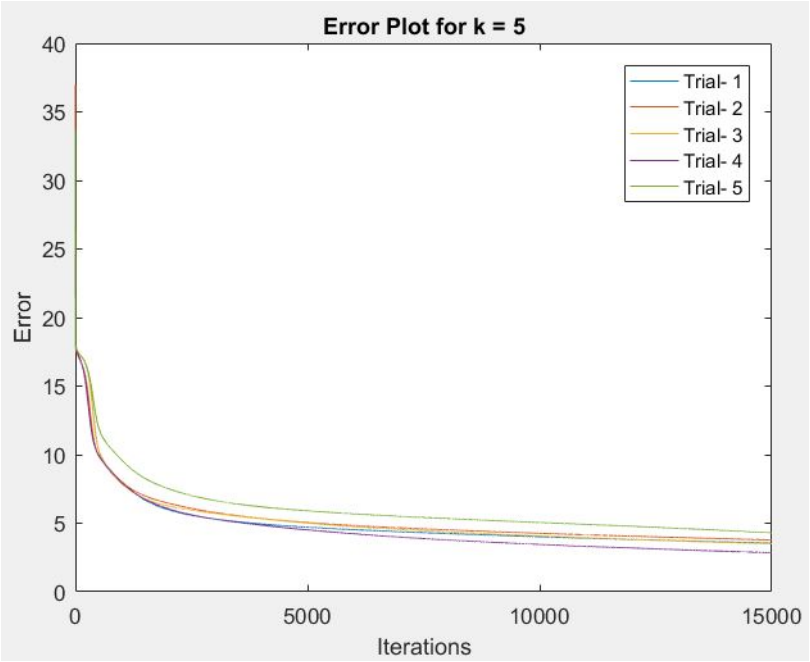


Figure 4.4: Error plot for k = 5

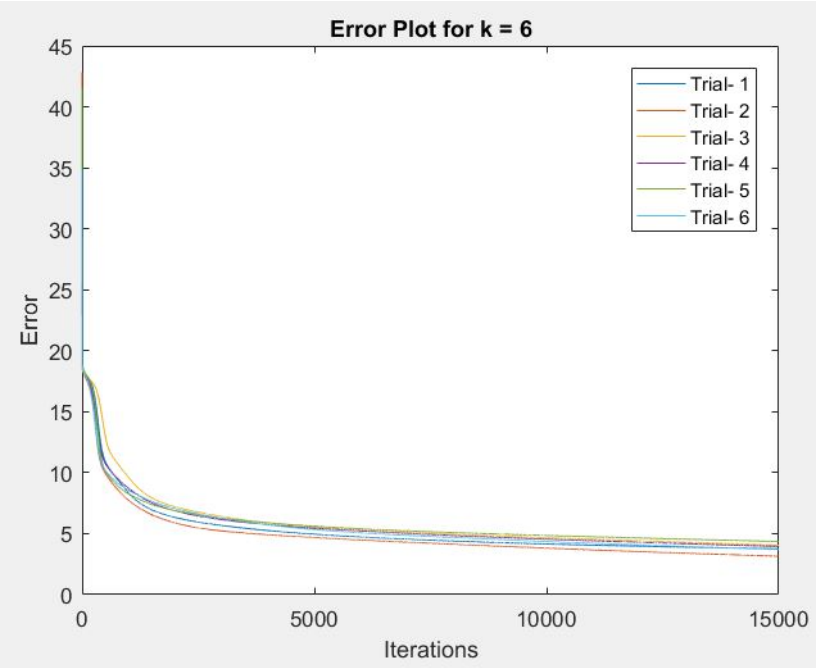


Figure 4.5: Error plot for k = 6

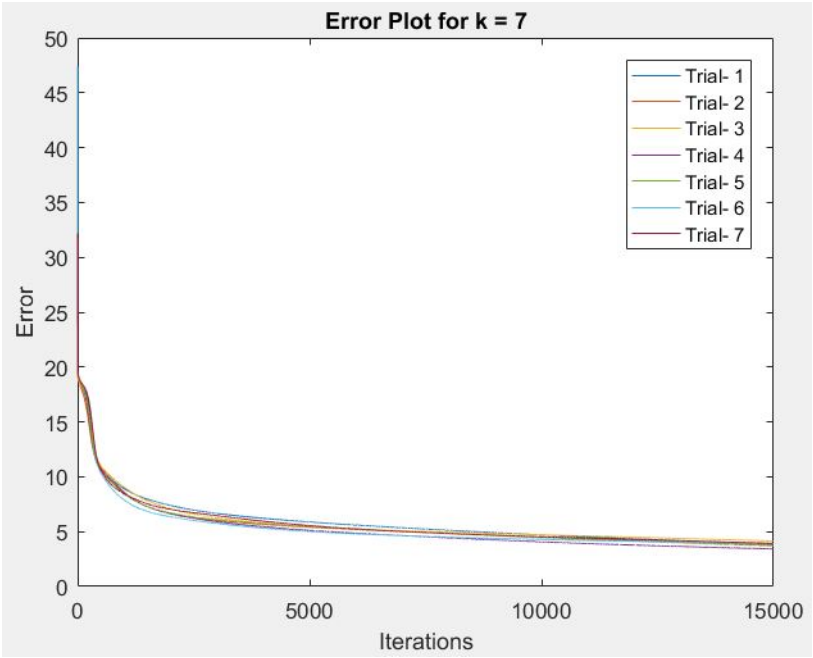


Figure 4.6: Error plot for k = 7

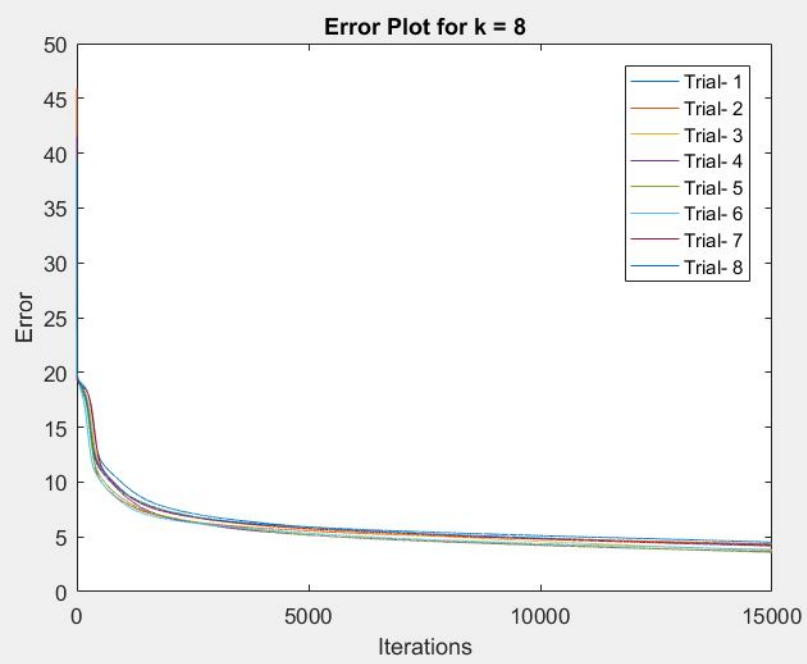


Figure 4.7: Error plot for k = 8

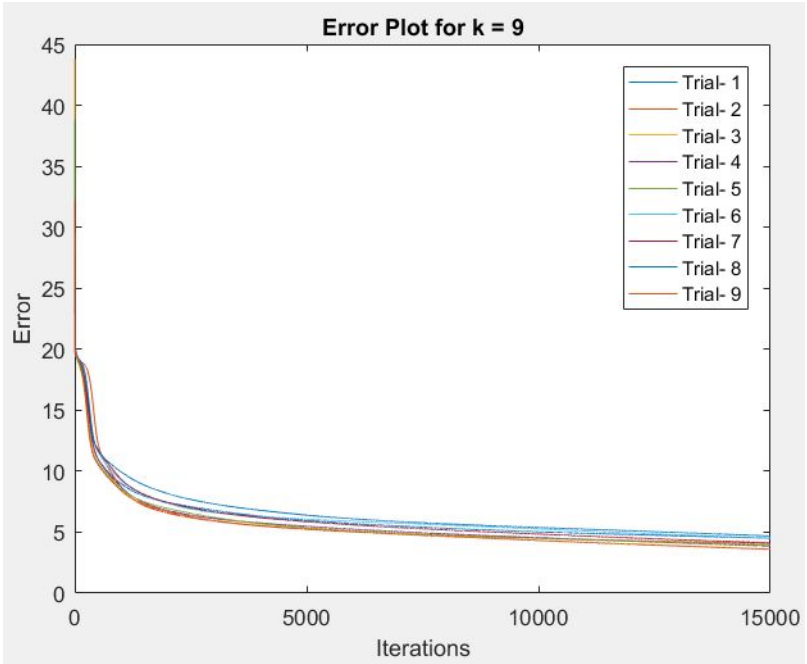


Figure 4.8: Error plot for k = 9

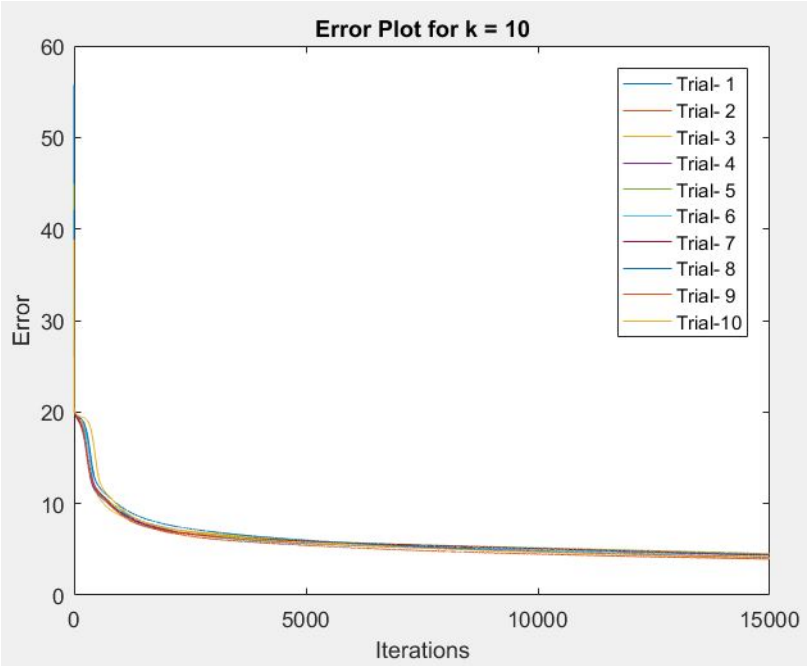


Figure 4.9: Error plot for k = 10

To analyze the performance of the model a confusion matrix is drawn which yields a better estimate of the performance than by just computing the accuracy which might be misleading. Confusion matrix as shown in the Figure. 4.10 consists of 4 values described below.

- An output is **True Positive** when the model predicts the input to be an AEP and the groundtruth for the input is also an AEP.
- An output is **True Negative** when the model predicts the input to be non-AEP and the groundtruth for the input is also non-AEP.
- An output is **False Positive** when the model predicts the input to be an AEP but the groundtruth of the input is actually non-AEP. Its also called the type I error.
- An output is **False negative** when the model predicts the input to be non-AEP but the groundtruth of the input is actually an AEP. Its also called the type II error.

From the confusion matrix the following performance metrics were calculated.

- **Sensitivity** is the rate at which the positives are correctly identified. Its also called the true

		Ground Truth	
		Positive	Negative
Prediction	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Figure 4.10: Confusion matrix

positive rate.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.1}$$

- **Specificity** is the rate at which the negatives are correctly identified. Its also called the true negative rate.

$$Specificity = \frac{TN}{TN + FP} \tag{4.2}$$

- **Precision** is the fraction of the positives that are correctly identified over all the positives predicted by the model.

$$Precision = \frac{TP}{TP + FP} \tag{4.3}$$

- **Accuracy** is the number of correct predictions from all the predictions by the model. It can be written as the ratio of correctly identified samples (True positives and True negatives) to the total number of samples.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4.4}$$

Table. 4.2 shows the performance measures of the network for each trial of the k-fold validation. Table. 4.3 shows the average performance metrics of the model for each fold.

No. of Folds	*Trial*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*Accuracy (%)*
4	**1**	**90.48**	**85.71**	**86.36**	**88.10**
	2	**86.36**	**77.27**	**79.17**	**81.82**
	3	**77.27**	**86.36**	**85.00**	**81.82**
	4	**85.71**	**80.95**	**81.82**	**83.33**
5	**1**	**64.71**	**82.35**	**78.57**	**73.53**
	2	**83.33**	**94.44**	**93.75**	**88.89**
	3	**88.24**	**76.47**	**78.95**	**82.35**
	4	**76.47**	**76.47**	**76.47**	**76.47**
	5	**100.00**	**94.12**	**94.44**	**97.06**
6	**1**	**78.57**	**71.43**	**73.33**	**75.00**
	2	**86.67**	**60.00**	**68.42**	**73.33**
	3	**100.00**	**86.67**	**88.24**	**93.33**
	4	**78.57**	**100.00**	**100.00**	**89.29**
	5	**100.00**	**71.43**	**77.78**	**85.71**
	6	**64.29**	**85.71**	**81.82**	**75.00**
7	**1**	**100.00**	**66.67**	**75.00**	**83.33**
	2	**76.92**	**92.31**	**90.91**	**84.62**
	3	**92.31**	**84.62**	**85.71**	**88.46**
	4	**83.33**	**83.33**	**83.33**	**83.33**
	5	**91.67**	**75.00**	**78.57**	**83.33**
	6	**75.00**	**75.00**	**75.00**	**75.00**
	7	**83.33**	**100.00**	**100.00**	**91.67**
8	**1**	**80.00**	**90.00**	**88.89**	**85.00**
	2	**72.73**	**72.73**	**72.73**	**72.73**
	3	**100.00**	**100.00**	**100.00**	**100.00**
	4	**72.73**	**63.64**	**66.67**	**68.18**
	5	**81.82**	**90.91**	**90.00**	**86.36**
	6	**81.82**	**81.82**	**81.82**	**81.82**
	7	**100.00**	**81.82**	**84.62**	**90.91**
	8	**90.00**	**80.00**	**81.82**	**85.00**
9	**1**	**100.00**	**88.89**	**90.00**	**94.44**
	2	**80.00**	**80.00**	**80.00**	**80.00**
	3	**80.00**	**70.00**	**72.73**	**75.00**
	4	**100.00**	**80.00**	**83.33**	**90.00**
	5	**70.00**	**80.00**	**77.78**	**75.00**
	6	**90.00**	**100.00**	**100.00**	**95.00**
	7	**88.89**	**88.89**	**88.89**	**88.89**
	8	**100.00**	**100.00**	**100.00**	**100.00**
	9	**77.78**	**66.67**	**70.00**	**72.22**
10	**1**	**87.50**	**87.50**	**87.50**	**87.50**
	2	**88.89**	**88.89**	**88.89**	**88.89**
	3	**77.78**	**88.89**	**87.50**	**83.33**
	4	**88.89**	**77.78**	**80.00**	**83.33**
	5	**100.00**	**77.78**	**81.82**	**88.89**
	6	**77.78**	**77.78**	**77.78**	**77.78**
	7	**66.67**	**100.00**	**100.00**	**83.33**
	8	**87.50**	**75.00**	**77.78**	**81.25**
	9	**75.00**	**87.50**	**85.71**	**81.25**
	10	**87.50**	**87.50**	**87.50**	**87.50**

Table 4.2: Performance Metrics for each Trial of k-fold Validation

No. of Folds	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*Accuracy (%)*
4	**84.95**	**82.57**	**83.08**	**83.76**
5	**82.54**	**84.77**	**84.43**	**83.66**
6	**84.68**	**79.20**	**81.59**	**81.94**
7	**86.08**	**82.41**	**84.07**	**84.24**
8	**84.88**	**82.61**	**83.31**	**83.75**
9	**87.40**	**83.82**	**84.74**	**85.61**
10	**83.75**	**84.86**	**85.44**	**84.30**

Table 4.3: Average Performance Metrics of k-fold validation

4.4 Conclusion

From Table. 4.3 it can be seen that k = 9 provides the best results in terms of sensitivity and specificity. This indicates that as the sample size of the training set increase, the model showed better results. The results showed that the model with the considered features had an average of 84% sensitivity to 83% specificity and is able to distinguish majority of the AEP and non-AEP transients correctly. Investigation of other features along with considering different techniques could result in higher ideal performance.

Chapter 5

Detection of Epileptiform Transients

The aim of this project is to automate the system of detecting and classifying the annotations into AEP and non-AEP. Usually this process, as explained in Chapter-2, is performed manually by experts who mark or yellow box the epileptiform like transients and classify them into AEP and non-AEPs. The process of detection is performed by training a multilayer RBF neural network with segments of normal EEG data extracted from the non-annotated part of the signal to classify an epileptiform like transient from that of the normal EEG segment.

5.1 Training Data

5.1.1 Data Synthesis

The data for training the neural network had to be indicative of epileptiform transients and non-epileptic signals. For examples of epileptiform transients, the annotations or yellow boxes given by experts are considered. However, there is no specific definition nor morphological properties that can specifically describe a non-epileptic signal. So, any part of the signal that is not annotated can be considered to be a non-epileptic signal and that provides us with enormous amount of examples indicative of non-epileptic signals.

As mentioned in Chapter-2, EEG recordings of 200 patients, each of 30 seconds duration

were recorded. The annotations that represented class ET came from only the 91 files out of 200, and the rest 109 files had no annotations and hence any segment from any channel is an indicator of non-epileptic signal and can be used as training data. To synthesize a good dataset for training involves taking few precautions. First of all, the data belonging to different patients were recorded at different sampling rates, so the EEG recordings are frequency normalized to 256 Hz. Also, there are 47 unique combination of electrodes or channels which are used by experts in identification of the epileptic transients. The data from 109 files on those 47 channels are primarily considered as the source of non-epileptic signals. These recordings represent only the normal brain activity and might have the possibility of not containing any artifacts. In order to account for this, the EEG data from other 91 EEG files is also considered. All the data in each channel other than the annotated data segments fall under the class non-ET which might be normal brain activity or artifacts. Around 270000 non-et signals has been extracted from 47 different channels in 200 EEG files. Like the length of the epileptiform transients, the length of the non-ET signal is also considered to be variable with a similar length distribution as discussed in Table. 2.3 in Chapter-2. The 'pre' and 'post' contextual information of around 200ms is also considered along with the non-ET signal for computing the features.

5.1.2 Balancing the data

There are around 270000 extracted samples of non-et class while there are only 235 samples of class ET. This represents a hugely imbalanced dataset. In literature, [16] several techniques have been proposed to balance the dataset such as oversampling, under-sampling, random sampling etc. Some techniques based on the distance measures between the samples of two classes have also been suggested but does not work very well for this problem may be due to the fact that non-ET signals look very similar to epileptiform transients. So, in order to balance the data informed under-sampling has been used. Around 120 examples are taken from 109 files that do not have any annotations and the other 115 examples from the remaining files which contains epileptiform transients.

5.1.3 Features

As mentioned in Chapter-3, Empirical mode decomposition and Hilbert Huang transform has been used in the extraction of features at both stages. In addition to the 33 dimensional feature

from the EEG data segment, 'Weiner Entropy' and the ratio of the weighted mean frequency of the 'pre' and 'post' contextual signal to the weighted mean frequency of the annotation are also used as additional features only for the detection stage resulting in 42 dimensional feature vector.

5.2 Neural Network for Detection

A multilayer RBF network as described in classification phase is also used in the detection phase. The number of RBF unit centers, number of training iterations, learning rate and L2-regularization value are chosen based on the performance of the model in the cross validation tests. The desired output of +1 here corresponds to a suspected epileptic transient which is to be yellow boxed and an output of 0 is classified as non-epileptiform transient.

The weights of the units with ReLU activation and the output unit is initialized with values of magnitude between 0.005 to 0.05 and the RBF unit centers is chosen by using standard k-means algorithm with a k value of 8. The network is trained for 25000 iterations with a learning rate of 0.0005. The value for L2 regularization to prevent the network from over fitting is set to 0.0001.

5.3 Results

The training of neural network can be seen on the graph of Error with respect to training iterations. To analyze the performance of the detection model a k-fold cross validation analysis is performed with k values from 4 to 10. The graph has been plotted for each value of k-fold. The error plot shows the error decreasing with the iterations on the k^{th} training set. The Figure. 5.1 to 5.7 shows the error plot for value of k from 4 to 10. Table. 5.1 shows the performance results for each trial of the k-fold validation and Table. 5.2 shows the average performance results of the model for each fold.

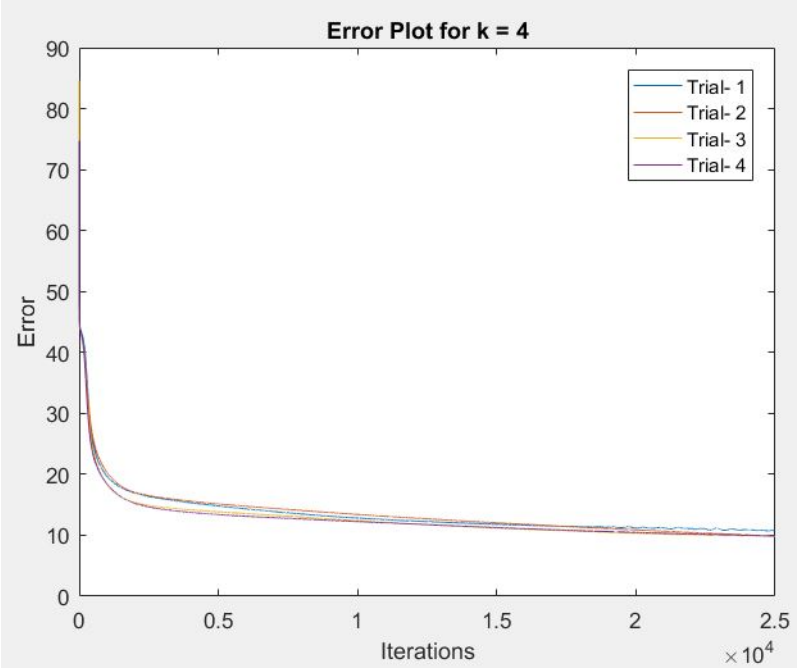


Figure 5.1: Error plot for k = 4

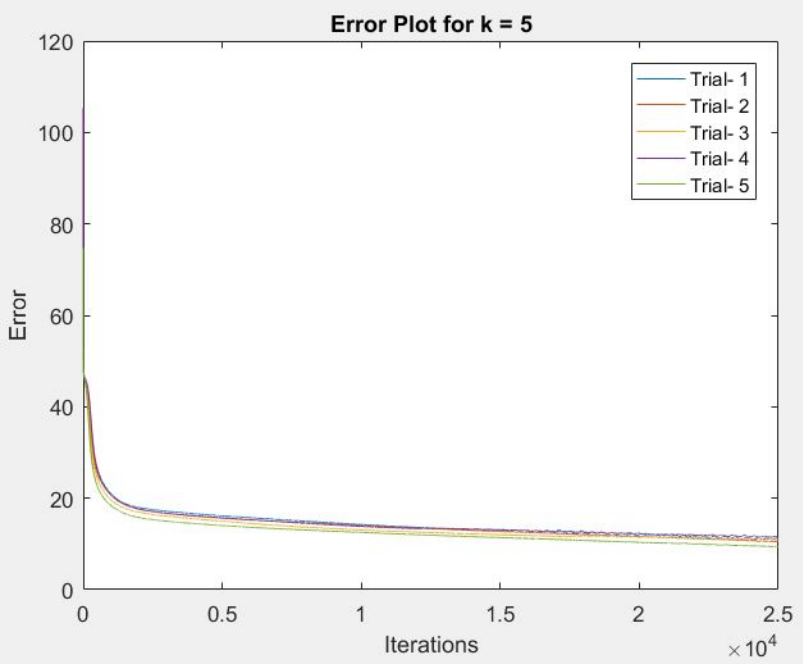


Figure 5.2: Error plot for k = 5

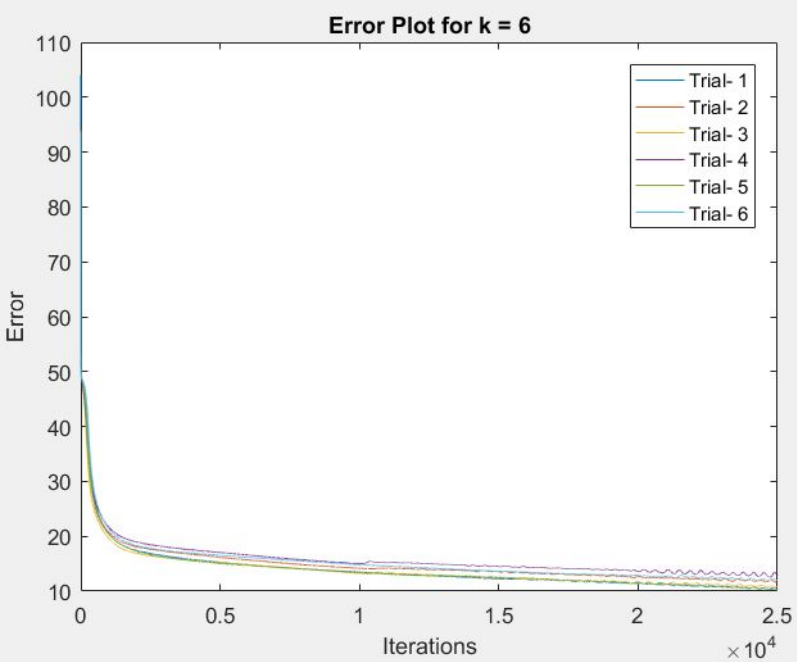


Figure 5.3: Error plot for k = 6

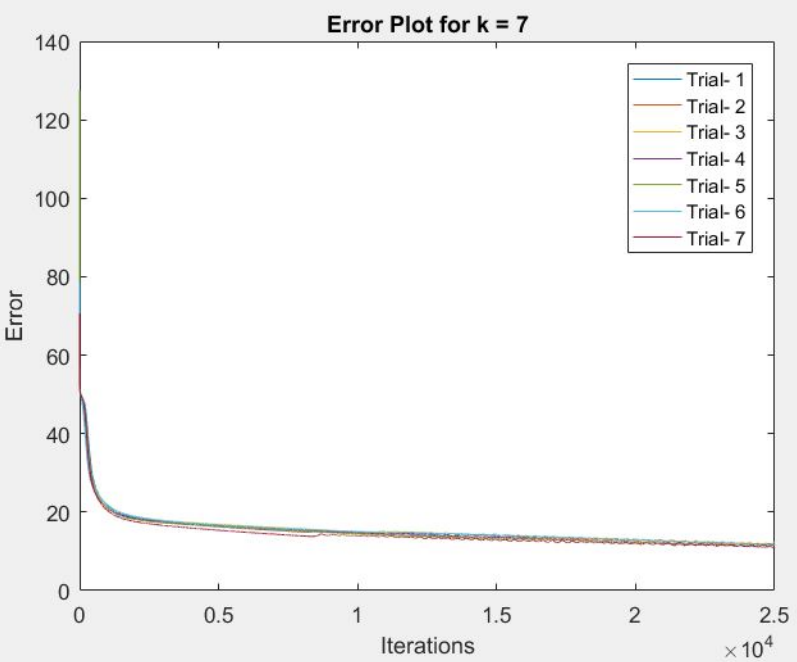


Figure 5.4: Error plot for k = 7

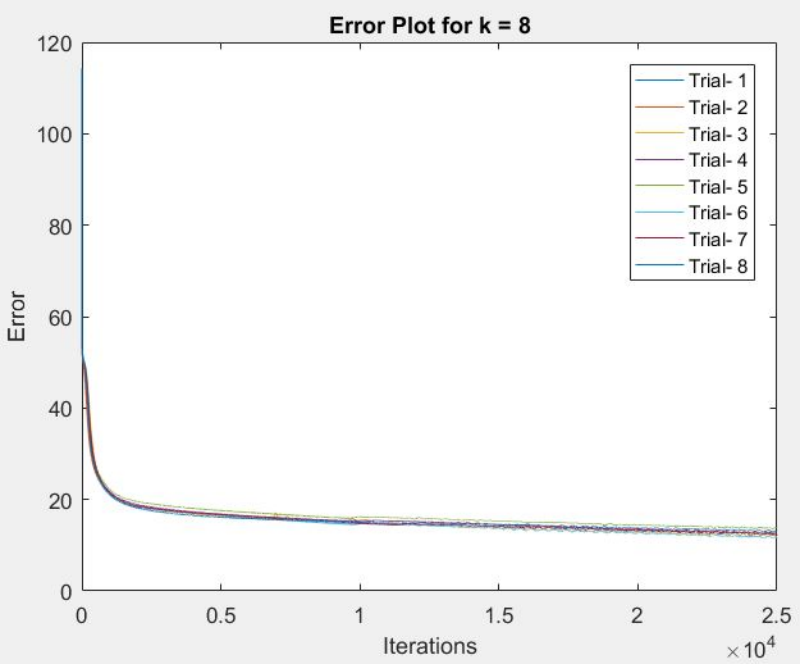


Figure 5.5: Error plot for k = 8

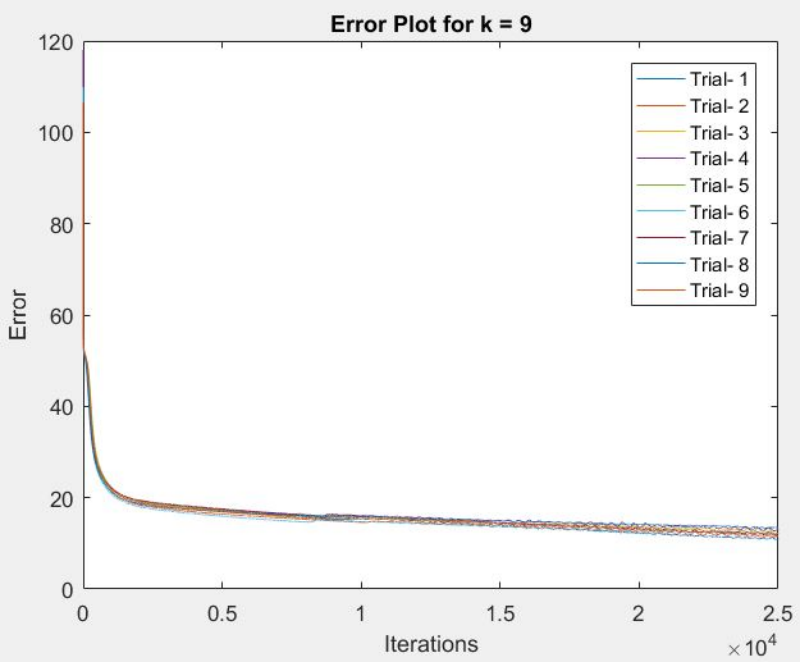


Figure 5.6: Error plot for k = 9

No. of Folds	*Trial*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*Accuracy (%)*
4	**1**	**92.98**	**82.46**	**84.13**	**87.72**
	2	**91.38**	**89.66**	**89.83**	**90.52**
	3	**82.76**	**87.93**	**87.27**	**85.34**
	4	**84.48**	**79.31**	**80.33**	**81.90**
5	**1**	**95.65**	**82.61**	**84.62**	**89.13**
	2	**82.98**	**95.74**	**95.12**	**89.36**
	3	**93.48**	**89.13**	**89.58**	**91.30**
	4	**86.96**	**91.30**	**90.91**	**89.13**
	5	**95.65**	**71.74**	**77.19**	**83.70**
6	**1**	**89.47**	**84.21**	**85.00**	**86.84**
	2	**94.87**	**87.18**	**88.10**	**91.03**
	3	**82.05**	**89.74**	**88.89**	**85.90**
	4	**92.31**	**89.74**	**90.00**	**91.03**
	5	**84.21**	**86.84**	**86.49**	**85.53**
	6	**86.84**	**97.37**	**97.06**	**92.11**
7	**1**	**90.91**	**96.97**	**96.77**	**93.94**
	2	**87.88**	**81.82**	**82.86**	**84.85**
	3	**90.91**	**93.94**	**93.75**	**92.42**
	4	**90.91**	**93.94**	**93.75**	**92.42**
	5	**96.97**	**81.82**	**84.21**	**89.39**
	6	**90.91**	**93.94**	**93.75**	**92.42**
	7	**87.88**	**78.79**	**80.56**	**83.33**
8	**1**	**89.29**	**85.71**	**86.21**	**87.50**
	2	**96.55**	**86.21**	**87.50**	**91.38**
	3	**89.66**	**75.86**	**78.79**	**82.76**
	4	**89.66**	**89.66**	**89.66**	**89.66**
	5	**86.21**	**96.55**	**96.15**	**91.38**
	6	**89.66**	**93.10**	**92.86**	**91.38**
	7	**89.66**	**79.31**	**81.25**	**84.48**
	8	**96.55**	**89.66**	**90.32**	**93.10**
9	**1**	**92.00**	**84.00**	**85.19**	**88.00**
	2	**88.46**	**84.62**	**85.19**	**86.54**
	3	**96.15**	**76.92**	**80.65**	**86.54**
	4	**84.62**	**96.15**	**90.38**	**90.38**
	5	**84.62**	**92.31**	**91.67**	**88.46**
	6	**80.77**	**73.08**	**75.00**	**76.92**
	7	**96.15**	**96.15**	**96.15**	**96.15**
	8	**92.00**	**72.00**	**76.67**	**82.00**
	9	**96.00**	**96.00**	**96.00**	**96.00**
10	**1**	**86.96**	**82.61**	**83.33**	**84.78**
	2	**83.33**	**87.50**	**86.96**	**85.42**
	3	**86.96**	**82.61**	**83.33**	**84.78**
	4	**86.96**	**91.30**	**90.91**	**89.13**
	5	**95.65**	**91.30**	**91.67**	**93.48**
	6	**86.96**	**73.91**	**76.92**	**80.43**
	7	**95.65**	**91.30**	**91.67**	**93.48**
	8	**91.30**	**73.91**	**77.78**	**82.61**
	9	**82.61**	**82.61**	**82.61**	**82.61**
	10	**86.96**	**95.65**	**95.24**	**91.30**

Table 5.1: Performance Metrics for each Trial of k-fold Validation in Detection

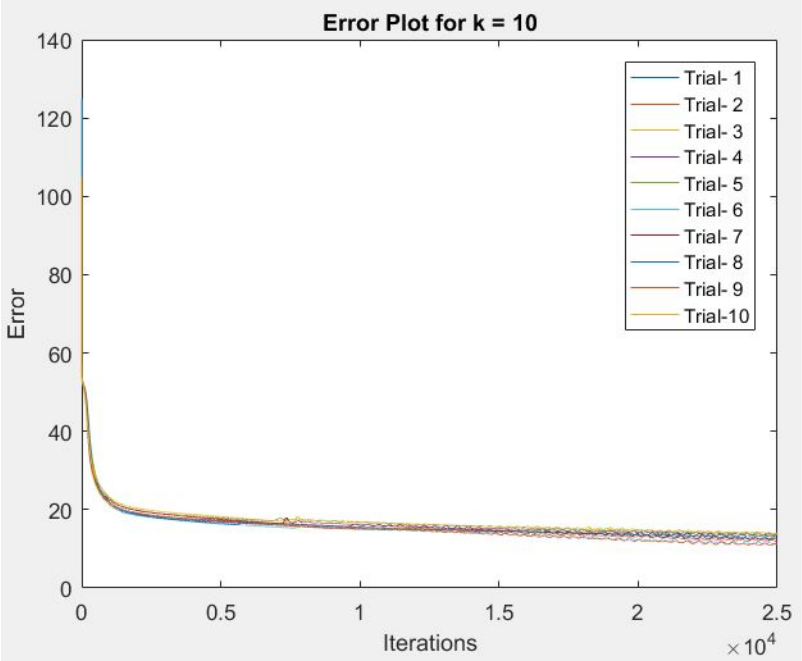


Figure 5.7: Error plot for k = 10

No. of Folds	*Sensitivity*	*Specificity*	*Precision*	*Accuracy*
4	**87.90**	**84.83**	**85.38**	**86.36**
5	**90.94**	**86.10**	**87.48**	**88.52**
6	**89.29**	**89.18**	**89.25**	**88.73**
7	**90.90**	**88.74**	**89.37**	**89.82**
8	**90.90**	**87.00**	**87.84**	**88.95**
9	**90.08**	**85.69**	**86.90**	**87.88**
10	**88.33**	**85.27**	**86.04**	**86.80**

Table 5.2: Average Performance Metrics of k-fold validation

5.4 Conclusion

In order to accomplish detection of yellow boxing of epileptic transients using a neural network, the data that best represents the both classes were extracted. The values of sensitivity and specificity are around 89% and 86% which indicates that the model is performing well in classifying the epileptiform transients from the non-epileptic data segments extracted from the EEG signal. The results are slightly better compared to the classification phase. This may be due to the larger training set. These detection and classification stage neural networks trained using the extracted data are then cascaded to automatically detect the epileptiform transients on an EEG channel waveform.

Chapter 6

Overall System Validation

6.1 Method

The main objective of this project is to build a system to automatically identify epileptiform transients and classify them into AEPs and non-AEPs. The neural networks trained separately for classification of ETs and the non-epileptiform segments and then classifying the ETs into AEPs and non-AEPs, as described in the previous chapters, are used in cascade. The block diagram representation of the entire system is shown in the Fig. 6.1.

To analyze the performance of the entire system, a subset of 10 patient files (chosen randomly based on the number of annotations in each file), contributing about 25% (57 of the 235 annotations) of the total annotations marked by the experts, are considered and run through the system. In contrast to the extracted/synthesized training data used previously in Chapter-5, here, the actual EEG channel waveform is considered. Each file may contain multiple annotations and each annotations could be on different channels. From the prior information, there are about 47 different channels on which the annotations are identified by experts. So, all the channels have to be checked for identifying all the annotations. However, to validate the performance of the system, the channel having the maximum number of annotations is only considered. This also helps to compare the performance of the system with expert's opinion. Table. 6.1 shows the files considered for validation along with the number of ETs in that file.

A sliding window of 1/3rd overlap is run through the entire epoch of 30 seconds on the selected channel. Features are extracted from each windowed EEG segment and checked for epilep-

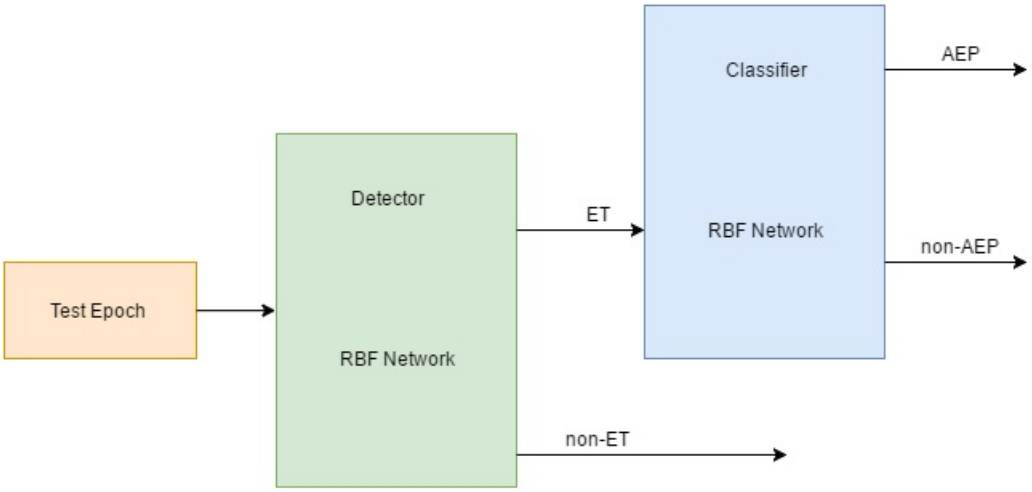


Figure 6.1: Block diagram of ET detection system

	Annotations by experts		
File	***AEPs***	***non-AEPs***	***Total***
12	**4**	**3**	**7**
36	**8**	**0**	**8**
65	**12**	**0**	**12**
89	**10**	**0**	**10**
95	**4**	**3**	**7**
124	**0**	**6**	**6**
160	**3**	**4**	**7**
24	**No Epileptiform Transients**		
84	**No Epileptiform Transients**		
159	**No Epileptiform Transients**		

Table 6.1: List of files for Validation

tiform transients. If the segment is a suspected epileptiform transient, the same data segment is passed through the second neural network to classify the suspected epileptiform transient into an AEP or non-AEP.

A window length similar to annotations has been used rather than using a fixed length window to account for variable length of the annotations. This also reduced the number of false positives greatly. The pre and post contextual signal lengths are fixed to 50 samples each, similar to the individual detector and classifier stages. A black box is drawn across all EEG segments suspected to be an epileptiform transient at the detector stage. These segments are then run through the classifier and a yellow box is drawn if the epileptiform is identified as an AEP.

To compare with the ground-truth of the annotations, a green box is drawn across AEPs and red box is drawn across non-AEPs. The classifier and detector stage neural networks are trained using all the data excluding the file that is considered for validation, and the neural network with best performance is used for validating the signal.

6.2 Performance Measures

Performance measures like sensitivity, specificity and precision helps in analyzing classifier when there are approximately equal number of samples from both classes. In this case, each channel of EEG is recorded for a duration of 30 seconds and sampled at 256Hz. The number of epileptiform transients in the channel for entire epoch is much smaller than the non-epileptiform segments. So, performance measures like sensitivity and specificity could be misleading. The following performance indicators were defined to quantify the performance of the overall system.

- **B-Selectivity(Boxing selectivity)** - This measure indicates the model's selectivity in marking the data segments. It is defined by the ratio of the number of EEG events that are actually present on the channel to the number of EEG events marked by the system on the channel. Ideally this value should be 1 for a system which has no false positives and a system with low selectivity value indicates the presence of lot of false positives.

$$\text{B-selectivity} = \frac{\text{Total number of events actually present on a channel}}{\text{Total number of events detected by the system}} \tag{6.1}$$

- **Inverse B-score(Inverse Boxing score)** - The system's ability to capture all the EEG

events that are actually present on the channel. It is given by one minus the ratio of number of EEG events missed by the system to the number of EEG events actually present on the channel. Ideally it should be 1 for a system which detects all the events.

$$\text{Inverse B-score} = 1 - \frac{\text{Number of EEG events missed}}{\text{Total Number of EEG events actually present on a channel}} \tag{6.2}$$

- **M-Score(Marking score)** - The system's ability to correctly classify the detected events into AEP or non-AEP. It can be given by the ratio of the number of correctly classified AEP/non-AEP to the number of AEP/non-AEP present and detected by the system. Ideally it should be 1 for a system which is able to correctly classify the events into AEP/non-AEP correctly.

$$\text{M-score} = \frac{\text{Number of events(AEP and non-AEPs) correctly classified}}{\text{Total number of events (AEPs and non-AEPs) actually present and detected by the system}} \tag{6.3}$$

- **M-Sensitivity(Marking sensitivity)** The system's ability to correctly identify the detected events as an AEP. It can be defined by the ratio of number of AEPs present on the channel and correctly predicted by the system to the total number of events detected as an AEP. Ideally should be 1 for a system which has correctly classified all the AEPs with no false positive AEPs.

$$\text{M-sensitivity} = \frac{\text{Number of AEPs actually present on the channel and correctly classified by the system}}{\text{Number of AEPs detected by the system}} \tag{6.4}$$

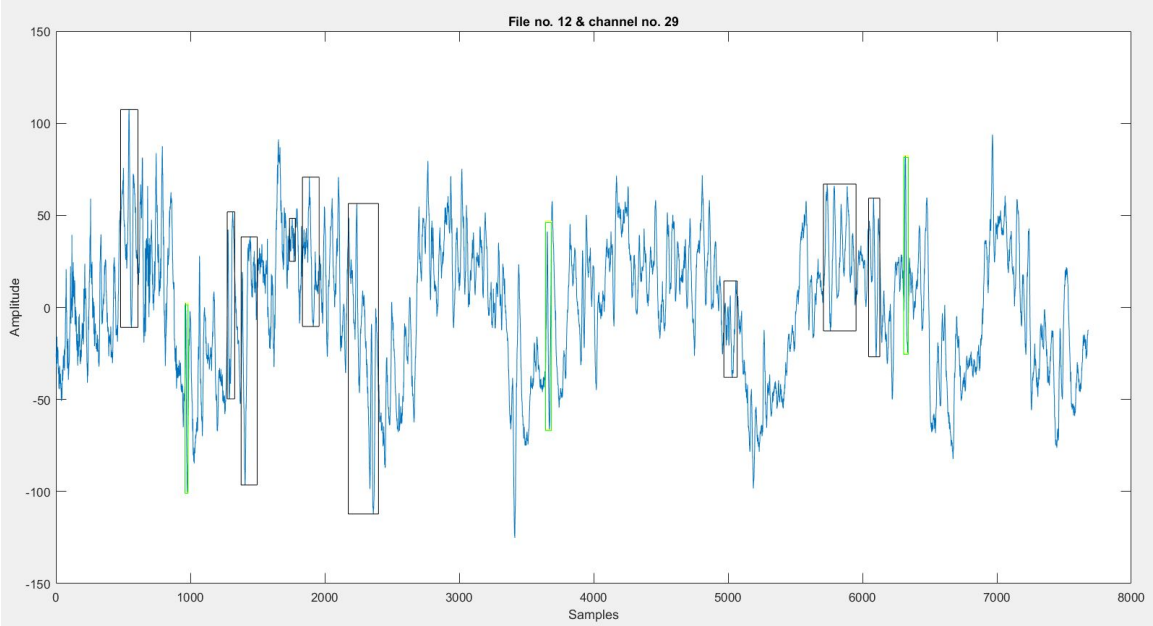


Figure 6.2: Validation on file no.12 on channel 'O1 - T5'

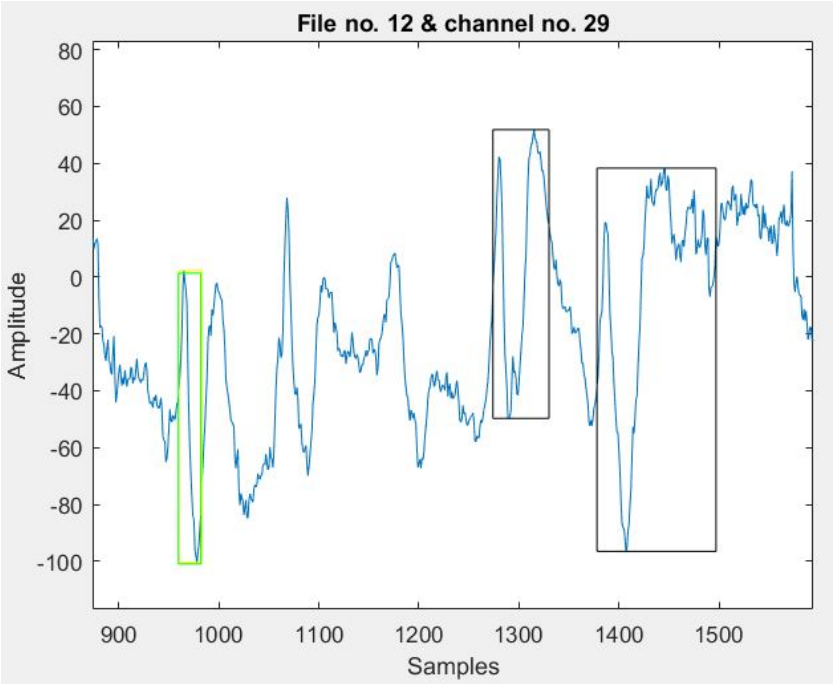


Figure 6.3: A closer look on annotation in file no.12 on channel 'O1-T5' showing the identification of AEP transient.

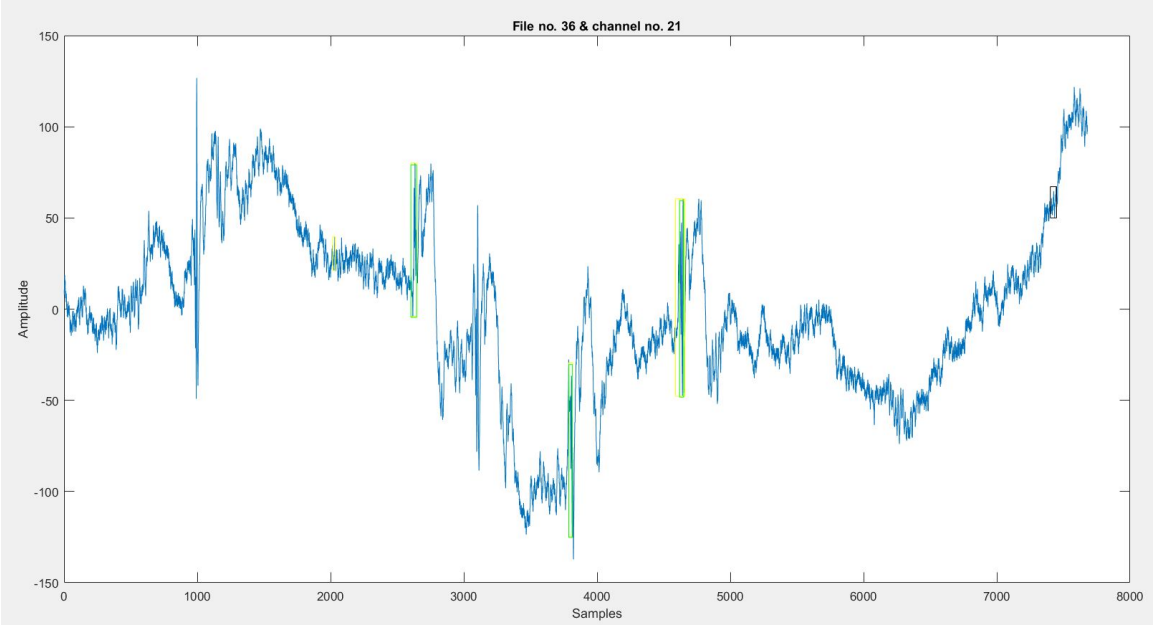
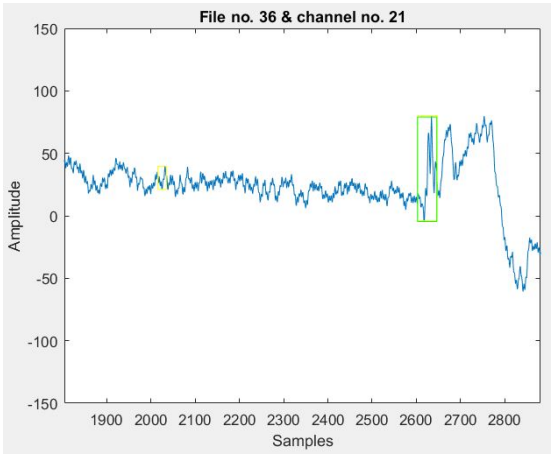
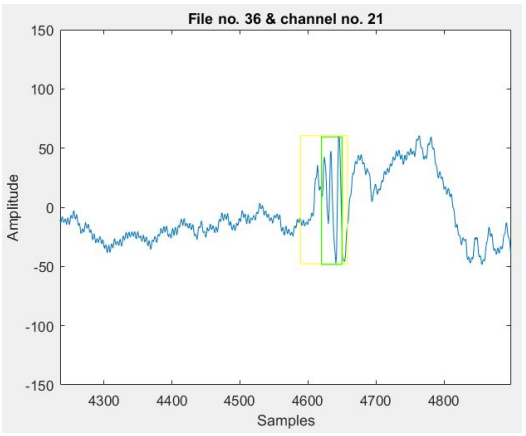


Figure 6.4: validation on file no.36 on channel 'FP1 - Avg'



(a) AEP transient detection



(b) A false positive overlapped window on the actual annotation

Figure 6.5: A closer look on file no.36 which has been successful in identifying the all EEG segment into AEP.

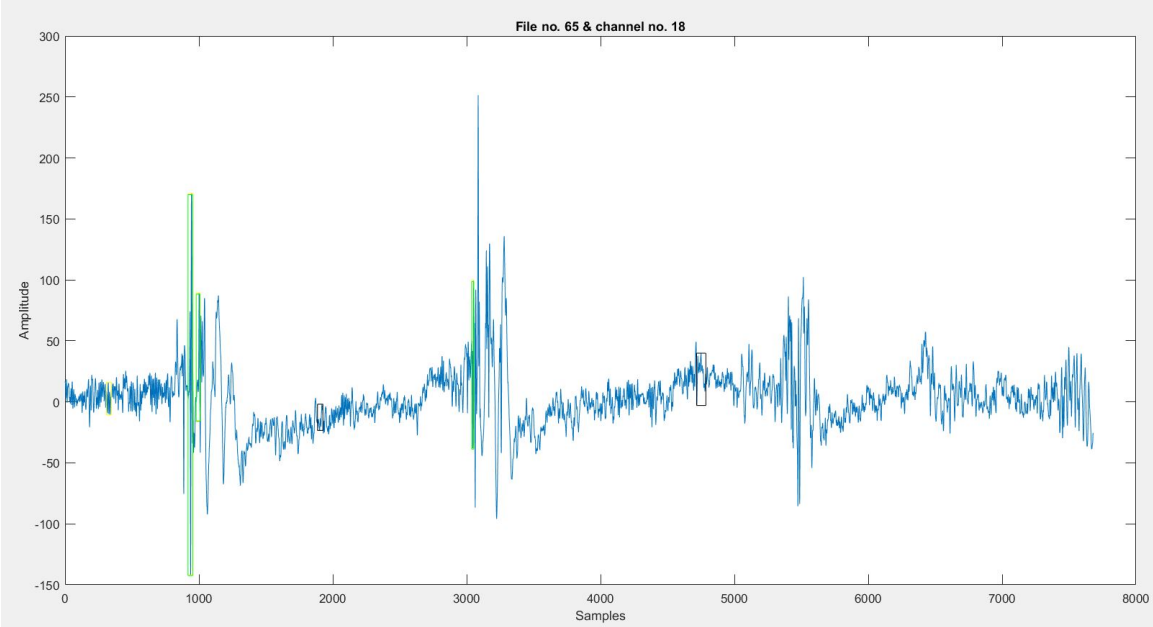
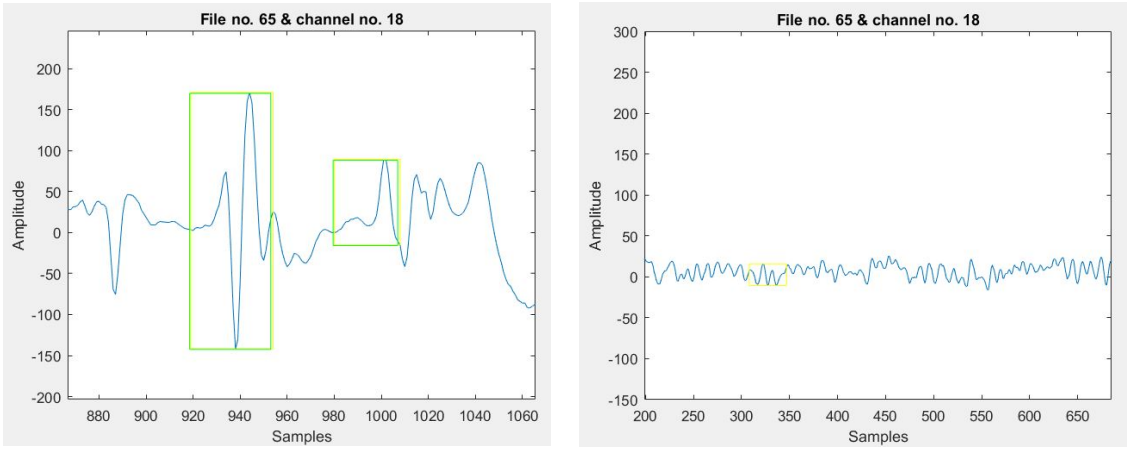


Figure 6.6: validation on file no.65 on channel 'Fp1 - F3'



(a) AEP transient detection

(b) False positive AEP transient predicted by system

Figure 6.7: A closer look on file no.65 which has correctly identified AEP transients and a false positive AEP.

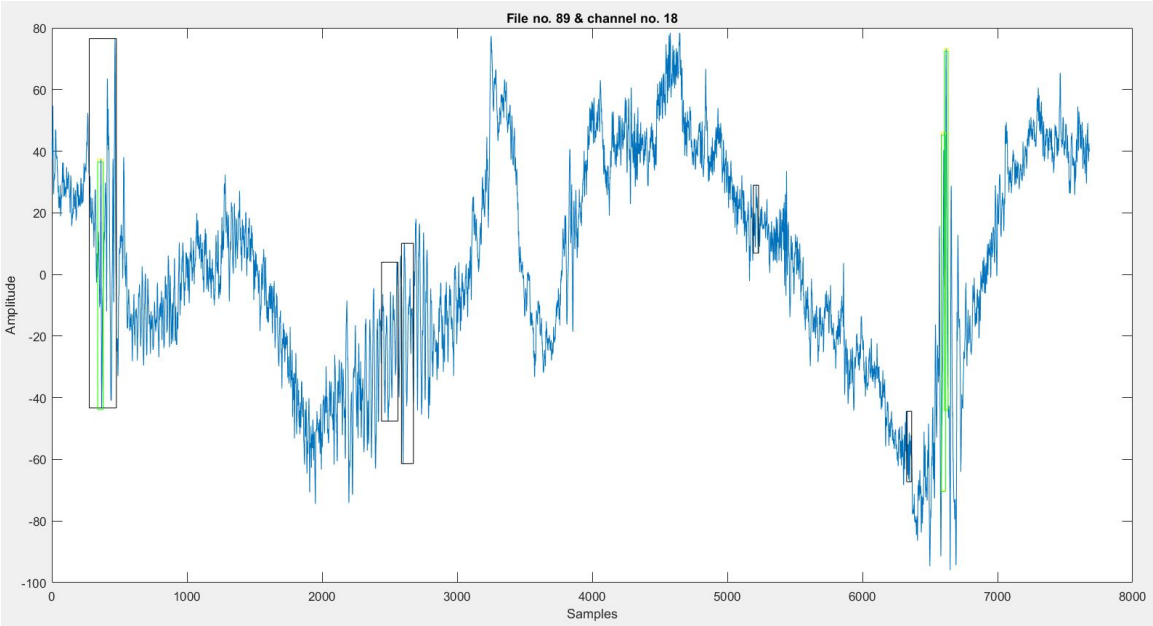


Figure 6.8: Validation on file no.89 on channel 'Fp1 - F3'

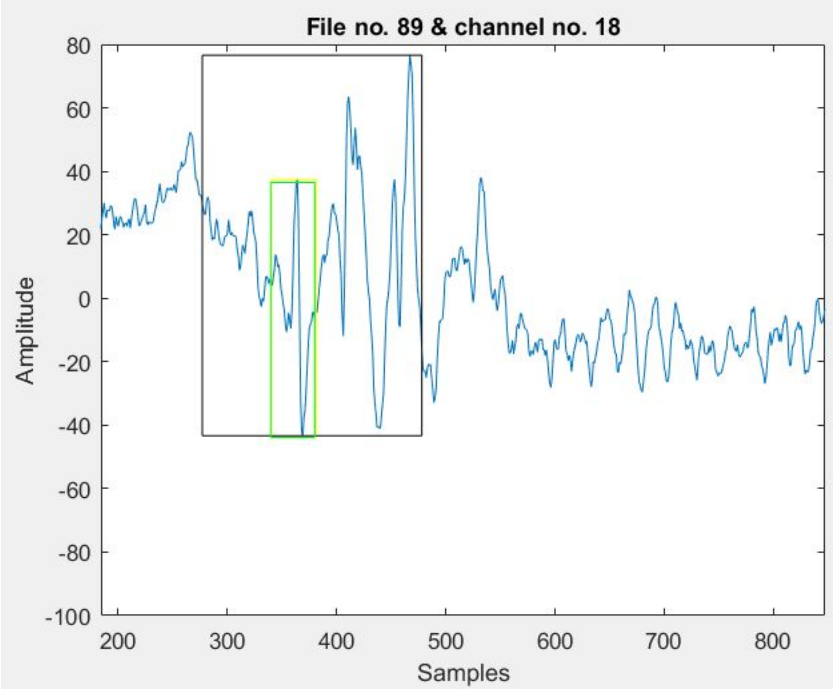


Figure 6.9: An example of a false positive of non-AEP in file no.89

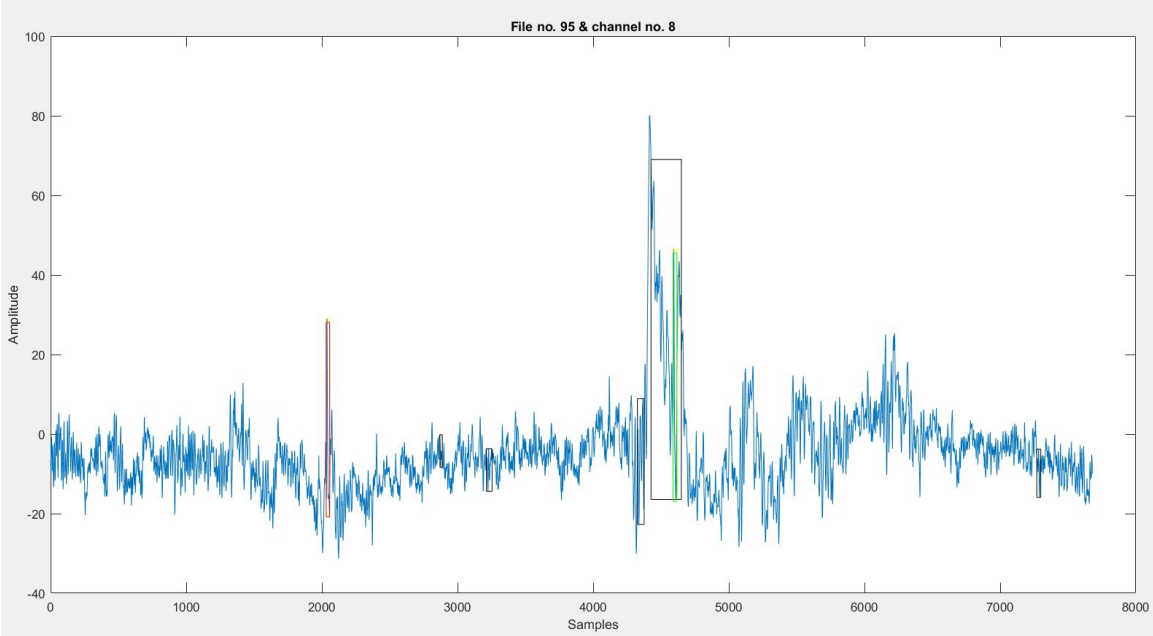
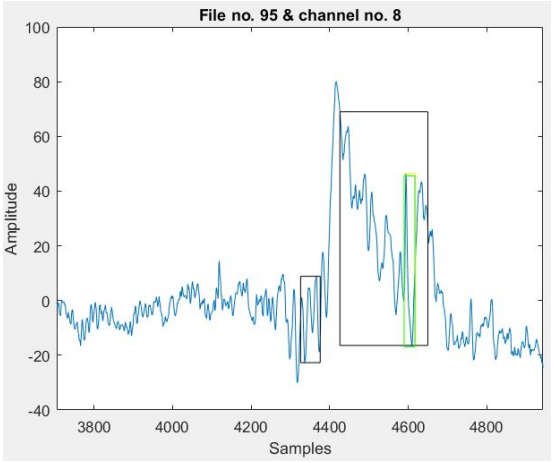
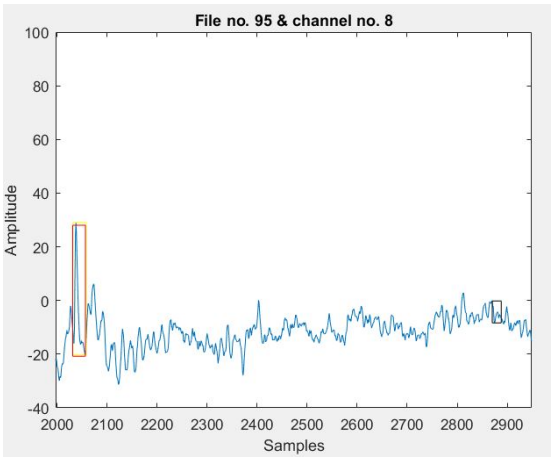


Figure 6.10: Validation on file no.95 on channel 'F3 - C3'



(a) AEP detected correctly



(b) non-AEP misclassified

Figure 6.11: Validation closer look on channel 'F3 - C3' having two annotations one AEP and one non-AEP. Misclassification of non-AEP segment is shown in fig. 6.11b.

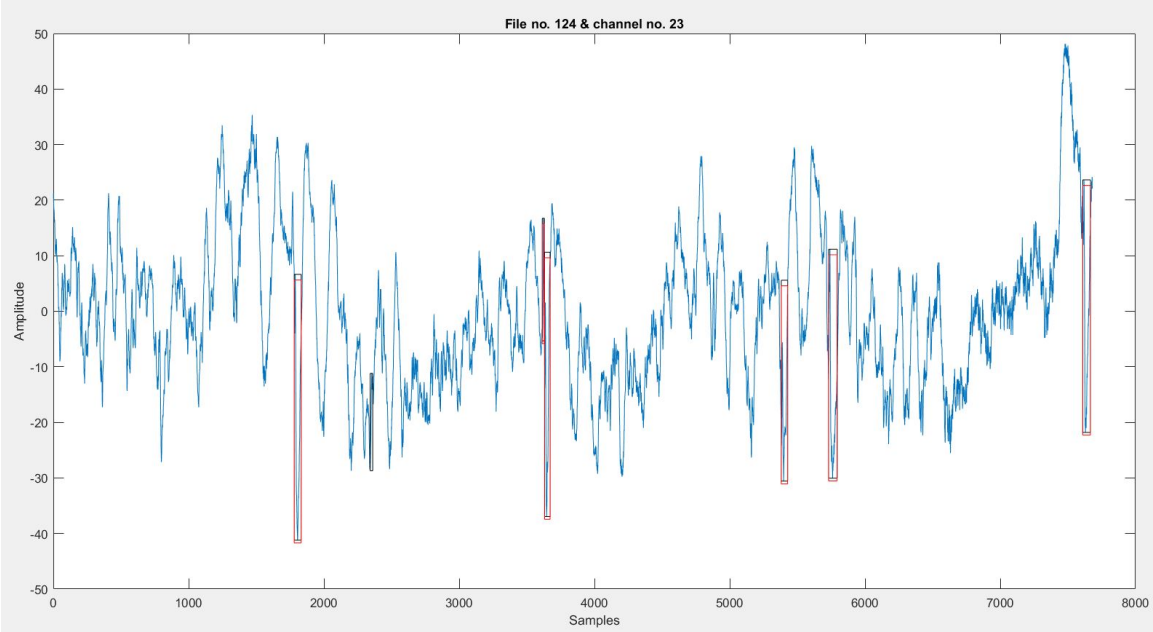


Figure 6.12: Validation on file no.124 on channel 'Fp2 - F4'

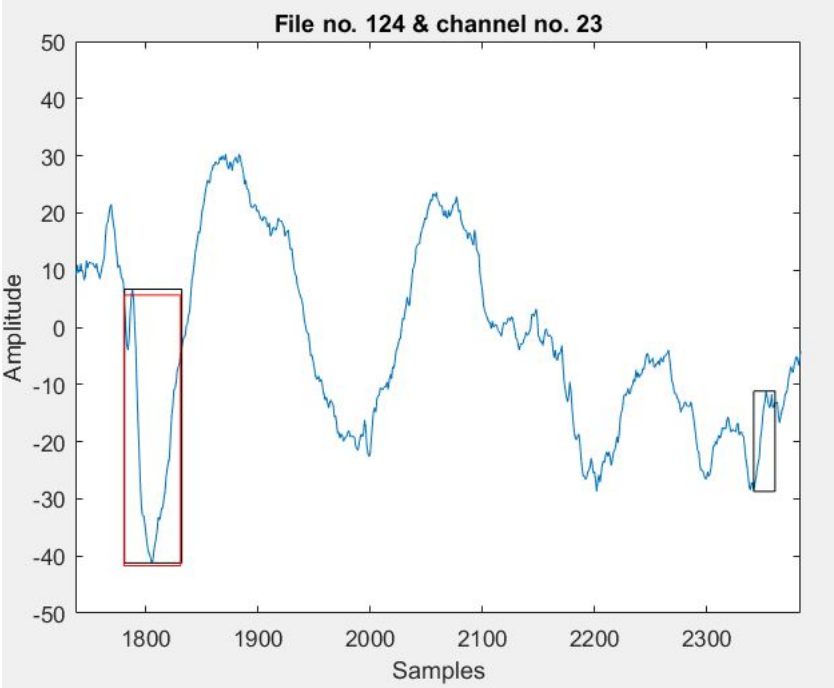


Figure 6.13: A closer look of validation on channel 'Fp2 - F4' having six non-AEP transients all of them correctly identified

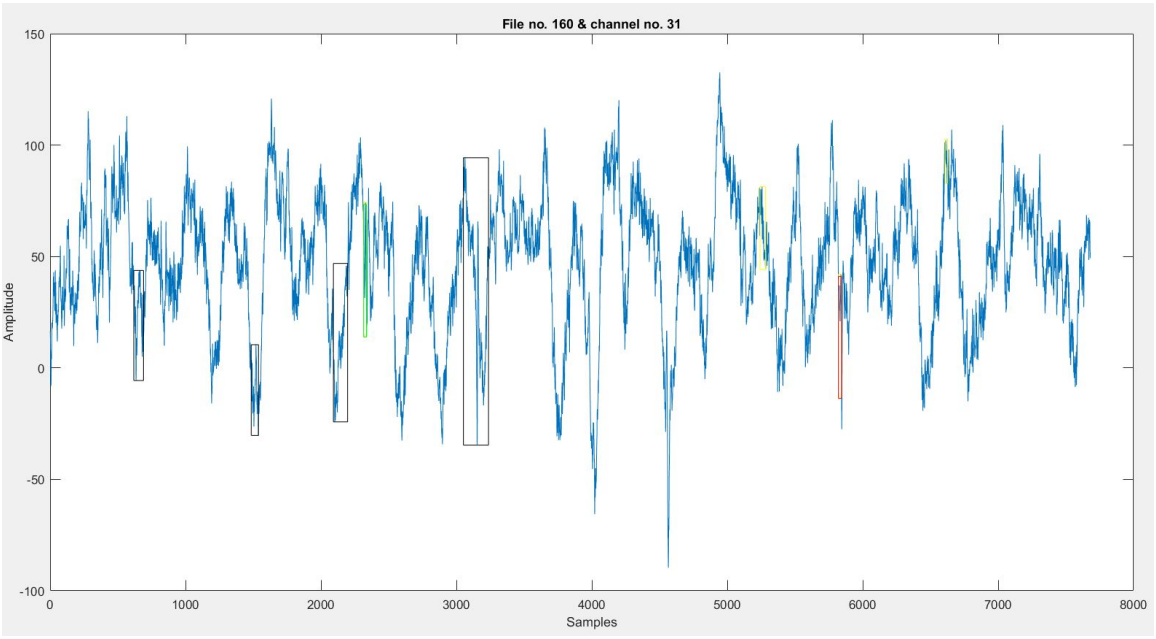
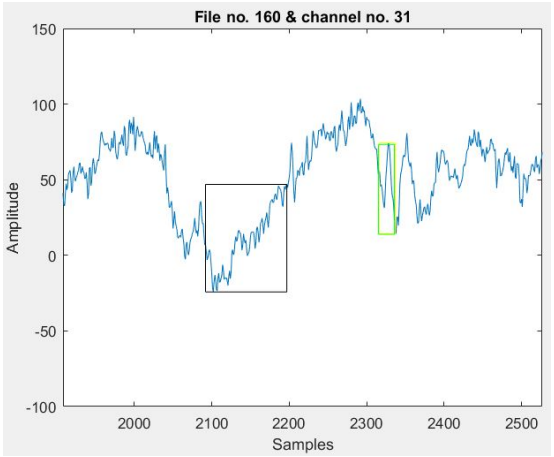
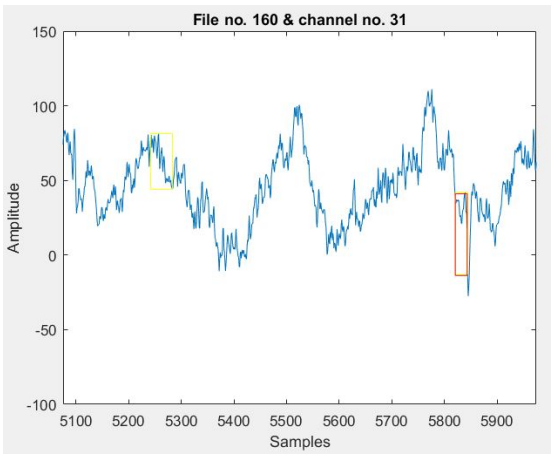


Figure 6.14: Validation on file no.160 on channel 'O2 - O1'



(a) AEP detected correctly



(b) non-AEP misclassified

Figure 6.15: Validation closer look on channel 'O2 - O1' having two annotations one AEP and one non-AEP. Misclassification of non-AEP segment and a false positive AEP transient predicted by the system is shown in fig. 6.15b.

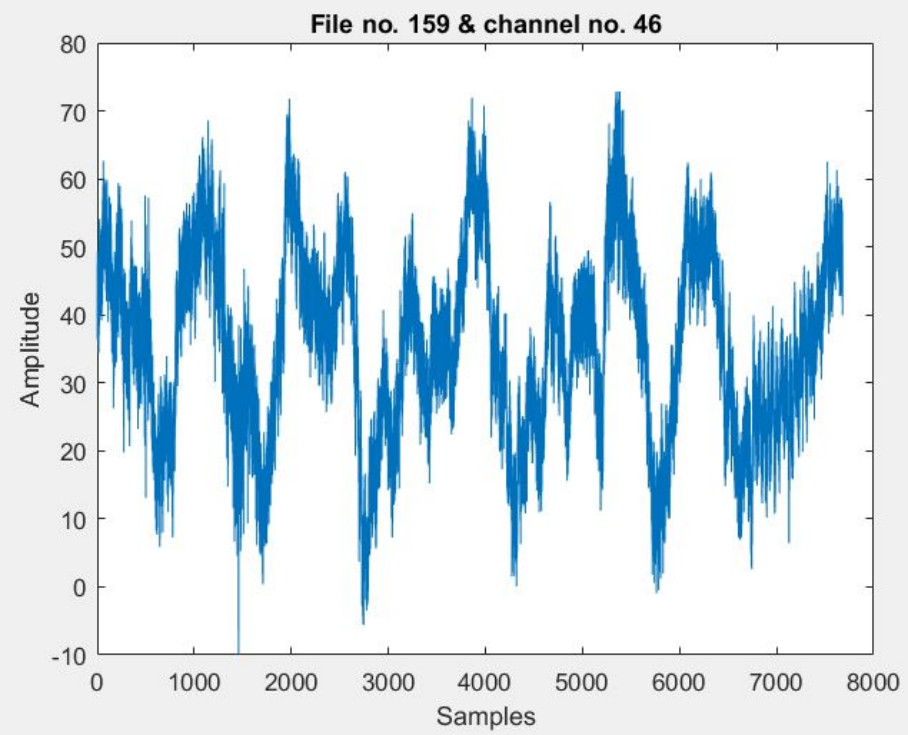
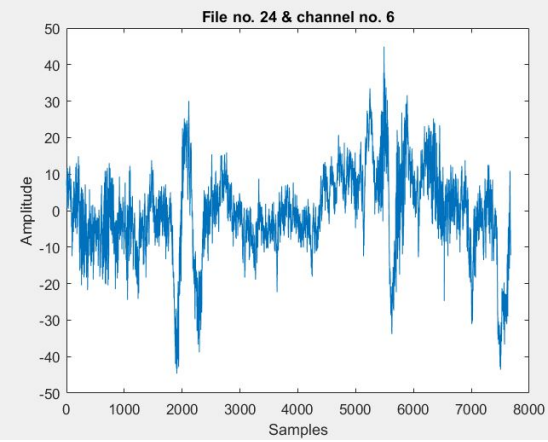
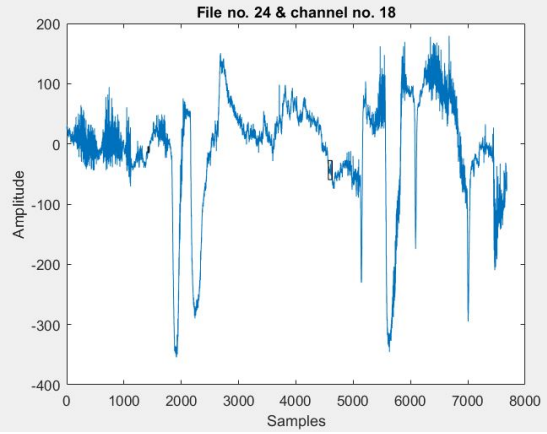


Figure 6.16: Validation on file no.159 on channel 'T5 - Avg'



(a) Validation on channel 'Cz - Pz'



(b) Validation on channel 'Fp1 - F3'

Figure 6.17: Examples of validation performed on file no.24 on different channels. No suspected epileptiform transients is predicted by the system in channel 'Cz - Pz'. In the fig. 6.17b the system detects a false positive non-AEP transient.

6.3 Results

Table. 6.2 and Table. 6.3 shows the system's response on EEG channel data considered for validation and the performance measures of the system. System did a good job of detecting all the actual annotations present in the signal. The classifier also performed well in correctly classifying the marked epileptiform transients as AEP and non-AEP in most cases.

System's performance on file 124 is good with only one false positive non-AEP transient and the system's performance on file no.95 was bad with 57% B-score indicating that the system could only identify 57% of the actual annotations present in the EEG data. Also, the system did not mark any annotations on a few files of healthy patients and also the false positives on such files are mainly non-AEPs.

The overall performance of the system in the validation of the random EEG signal is lesser than the individual performances of the network in the classifier and detection stages due to marking a large number of false positives. The marking score provided overall performance of the system by considering the annotations on all the channels.

File	***Black Boxes***	***Yellow Boxes***	***Green Boxes***	***Red Boxes***
12	**9**	**3**	**3**	**0**
36	**1**	**5**	**3**	**0**
65	**2**	**3**	**3**	**0**
89	**5**	**3**	**3**	**0**
95	**5**	**2**	**1**	**1**
124	**7**	**0**	**0**	**6**
160	**4**	**4**	**1**	**1**

Table 6.2: Results of Validation

File	*B-Selectivity (%)*	*B-Score (%)*	*M-Score (%)*	*M-Sensitivity (%)*
12	**18.75**	**100**	**57.14**	**100**
36	**50**	**100**	**87.5**	**60**
65	**60**	**91.66**	**91.66**	**100**
89	**37.50**	**100**	**90**	**100**
95	**28.57**	**57.14**	**28.57**	**50**
124	**85.71**	**100**	**100**	**-**
160	**25**	**100**	**71.42**	**25**

Table 6.3: Performance metrics of Validation

6.4 Conclusion

The main aim of this project to automatically detect and classify the epileptiform transients was accomplished by a cascaded design of two radial basis function neural networks. The individual neural networks in the classification and detection stages were trained on the features extracted from the data provided by best seven experts. The results of validation on a random EEG signal was not as good as expected and this may be due to the technique of windowing method adopted on the validation data. The system was able to correctly identify most of the AEPs on the validation data, although with a lot of false positives. The system's performance on the healthy EEG signal by not marking any segment to be a suspected epileptiform transient is a good indicator of the significance of the features extracted using EMD and HHT technique.

There were a lot of false positives marked as suspected Epileptiform like transients, similar to the results in the previous efforts [3] [46]. However, most of the errors or false positives on the non-annotated data were marked to be as non-AEP transients which indicates that these transients are mostly non-epileptic discharges. Another type of error or false positive in which the annotation was detected but misclassified was marking non-AEPs to be AEPs which even though not desirable, is better than the error of marking AEP transients as non-AEPs.

Chapter 7

Future Work

The main inference from this research was the importance of time-frequency information in the EEG signals. Hilbert-Huang transform and empirical mode decomposition served as good feature extraction techniques in the analysis of EEG signals. HHT has gained some recognition in the recent times and substantial research is being done in the area of EEG seizure detection and BCI applications using HHT and wavelet transforms. Unfortunately, a proper theoretical base is yet to be established. All the results are case-by-case comparisons conducted empirically [18]. Some of the outstanding mathematical problems in HHT that needs further investigation are, selection of the best IMFs extracted from the EEG signals, spline implementation problems in the process of sifting and extracting the IMFs.

The problem under consideration is challenging due to the fact that AEP transients are very similar to other normal electrocortical signals and artifacts. The understanding of the structural nuances of epileptiform transients in the way the expert neurologists examine is vital. Investigation of morphological properties of the epileptiform transients and the relation of pre and post contextual signals to the transients is an area worth considering which may lead to extraction of more localized and dependable features.

Although there are different standard databases for EEG signals for seizure detection [41], there is not much information on the interictal spikes complexes and no standard databases with such pre seizure spikes. An example dataset, with signals that define a proper non-epileptiform signals like artifacts and normal brain activities, need to be developed under the supervision of experts.

Other techniques such as use of wavelets or power spectral density estimates are being explored to compute interesting features. Also, use of raw signal in deep convolutional networks are being explored which may lead to better results.

Bibliography

[1] N. Acir, I. Oztura, M. Kuntalp, B. Baklan, and C. Guzelis. Automatic detection of epileptiform events in eeg by a three-stage procedure based on articial neural networks. *IEEE Transactions on Biomedical Engineering*, pages 52(1):30–40, Jan 2005.

[2] H. Adeli, S Ghosh Dastidar, and N. Dadmehr. A wavelet-chaos methodology for analysis of eegs and eeg sub-bands to detect seizure and epilepsy. *IEEE Trans. Biomed. Eng.*, pages 54(2):205–211, Feb. 2007.

[3] Kartikeya Shrikant Bagalore. Studying the use of hidden markov models in the detection and classification of eeg epileptiform transients using lpc features. Master's thesis, Clemson University, December 2016.

[4] V. Bajaj and R.B. Pachori. Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of eeg signals. *Biomed. Eng. Lett.*, pages 3:17–21, 2013.

[5] R. Cooper, J.W. Osselton, and J.C. Shaw. *EEG Technology.* Butterworths,, 1969.

[6] Brian C. Dean, Chad G. Waters, and Jonathan J. Halford MD. Eegnet: A web platform for collaborative eeg research. in graduate research and discovery symposium (grads). http://tigerprints.clemson.edu/grads_symposium/54/, 2013.

[7] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification.* Wiley., 2001.

[8] Epilepsy Foundation. Url. http://www.epilepsy.com/article/2014/4/revised-definition-epilepsy, April 2014.

[9] Epilepsy Foundation. What is Epilepsy. http://www.epilepsy.com/learn/epilepsy-101/what-epilepsy, 2014.

[10] A.J. Gabor, R.R. Leach, and F.U. Dowla. Automated seizure detection using a self-organizing neural network. *Electroenceph. Clinic. Neurophysiol.*, pages 99:257–266, 1996.

[11] John R. Glover, N. Raghavan, Periklis Y. Ktonas, and J.D. Frost. Context-based automated detection of epileptogenic sharp transients in the eeg: elimination of false positives. *IEEE Transactions on Biomedical Engineering*, pages 36(5):519–527, 1989.

[12] J.R. Glover, P.Y. Ktonas, N. Raghavan, J.M. Urunuela, S.S. Velamuri, and E.L. Reilly. A multichannel signal processor for the detection of epileptogenic sharp transients in the eeg. *IEEE Trans. Biomed. Eng.*, pages BME–33:1121–1128, 1992.

[13] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. Book in preparation for MIT Press, 2016.

[14] Jonathan J. Halford, Robert J. Schalkoff, Jing Zhou, Selim R.Benbadis, William O. Tatum, Robert P. Turner, Saurabh R Sinha, Nathan B. Fountain, Amir Arain, and Paul B. Pritchard et al. Standardized database development for eeg epileptiform transient detection: Eegnet scoring system and machine learning analysis. *Journal of neuroscience methods*, pages 212(2):308–316, 2013.

[15] Dr.R. Harikumar and T. Vijayakumar. A comparison of elman and radial basis function (rbf) neural networks in optimization of fuzzy outputs for epilepsy risk levels classification from eeg signals. *International Journal of Soft Computing and Engineering.*, pages 2(6):295–303, 2013.

[16] H. He and E. Garcia. Learning from imbalanced data. *IEEE Trans. on Knowledge and Data Eng.*, pages 21(9):1263–1284, Sept. 2008.

[17] W.E. Hostetler, H.J. Doller, and R.W. Homan. Assessment of a computer program to detect epileptiform spikes. *Electroencephalogr Clin Neurophysiol*, pages 83:1–11, July 1992.

[18] N.E. Huang and S.S. Shen (Eds.). *Introduction to the HilbertHuang transform and its related mathematical problems.* World Scientific Publishing Company, 2005.

[19] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and hilbert spectrum for non-linear and non-stationary time series analysis. *Proc. Roy. Soc.*, pages 454:903–995, 1998.

[20] S.J. Husain and K.S. Rao. An artificial neural network model for classification of epileptic seizures using huang-hilbert transform. *International Journal on Soft Computing (IJSC)*, pages 23–33, Nov 2014.

[21] K.P. Indiradevi, Elizabeth Elias, P.S. Sathidevi, S. Dinesh Nayak, and K. Radhakrishnan. A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram. *Computers in Biology and Medicine.*, pages 38(7):805–816, April. 2008.

[22] T. Kalayci and O. Ozdamar. Wavelet preprocessing for automated neural network detection of eeg spikes. *IEEE Eng. Med. Biol. Mag.*, pages 14:160–166, 1995.

[23] N. Kannathal, M.L Choo, U.R. Acharya, and P. Sadasivan. Entropies for detection of epilepsy in eeg. *Comput. Methods Programs Biomed.*, pages 80:187–194, 2005.

[24] C.W. Ko and H.W. Chung. Automatic spike detection via an artificial neural network using raw eeg data: Effects of data preparation and implications in the limitations of online recognition. *Clin. Neurophysiol.*, pages 111:477–481, 2000.

[25] N. Krayiannis. Reformulated radial basis neural networks trained by gradient descent. *IEEE Trans. on Neural Networks.*, pages 10(3):657–671, 1999.

[26] F.R. Kschischang. The hilbert transform. University of Toronto, 2006.

[27] P.Y. Ktonas. *Automated spike and sharp wave (ssw) detection.* EEG handbook, 1987.

[28] Fabien Lotte, Marco congedo, Anatole L. 'ecuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classication algorithms for eeg-based braincomputer interfaces. *Journal of neural engineering*, page 4(2):R1, 2007.

[29] Lambert Max, Engroff Andrew, Dyer Matt, and Byer Ben. Empirical mode decomposition. http://www.clear.rice.edu/elec301/Projects02/empiricalMode/process.html.

[30] D.B. Melges, A.F.C. Infantosil, F.R. Ferreira, and D.B. Rosas. Using the discrete hilbert transform for the comparison between trac alternant and high voltage slow patterns extracted from full-term neonatal eeg. *IFMBE Proceedings*, pages 14(2):1111–1114.

[31] E. Niedermeyer and F.L. da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincot Williams and Wilkins., 2004.

[32] R.J. Oweis and E.W. Abdulhay. Seizure classification in eeg signals utilizing hilberthuang transform. *BioMedical Engineering OnLine*, 2011.

[33] R.B. Pachori and V. Bajaj. Analysis of normal and epileptic seizure eeg signals using empirical mode decomposition. *Computer Methods and Programs in Biomedicine*, pages 104:373–381, 2011.

[34] F. Riaz, A. Hassan, S. Rehman, I.K. Niazi, and K. Dremstrup. Emd based temporal and spectral features for the classification of eeg signals using supervised learning. *IEEE Trans Neural Syst Rehabil Eng*, pages 107–124, 2013.

[35] R. Sankar and J. Nator. Automatic computer analysis of transients in eeg. *Comput. Biol. Med.*, pages 22(6):407–442, 1992.

[36] F. Sartorettoa and M. Ermani. Automatic detection of epileptiform activity by single-level wavelet analysis. *Clin. Neurophysiol.*, pages 110:239–249, 1999.

[37] Robert J. Schalkoff. *Artificial Neural Networks.* McGraw-Hill Higher Education, 1997.

[38] S.J. Schiff, A. Aldroubi, M. Unser, and S. Sato. Fast wavelet transformation of eeg. *Electroenceph. Clin. Neurophysiol.*, pages 92:442–455, 1994.

[39] R. Sharma, R.B Pachori, and U.R. Acharya. Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. *MDPI Entropy*, pages 17:669–691, 2015.

[40] Trans Cranial Technologies. 10-20 system positioning manual. https://www.transcranial.com/local/manuals/10_20_pos_man_v1_0_pdf.pdf, 2012.

[41] A. Tzallas, M. Tsipouras, and D. Fotiadis. Automatic seizure detection based on time-frequency analysis and articial neural networks. *Comput. Intell. Neurosci.*, page 18, 2007.

[42] Graham Upton and Ian Cook. *Understanding Statistics.* Oxford University Press., 1996.

[43] W.R.S. Webber, R.P. Lesser, R.T. Richardson, and K. Wilson. An approach to seizure detection using an artificial neural network (ann). *Electroenceph. Clinic. Neurophysiol.*, pages 98:250–272, 1996.

[44] S.B. Wilson, C.A. Turner, R.G. Emerson, and M.L. Scheuer. Spikedetection ii: Automatic, perception-based detection and clustering,. *Clin. Neurophysiol.*, pages 110:404–411, 1999.

[45] J.V. Zaena. Adaptive tracking of eeg oscillations. *Neuroscience Methods.*, pages 186:97–106, 2010.

[46] Jing Zhou. *A study of automatic detection and classication of EEG Epileptiform Transients.* PhD thesis, Clemson University, August 2014.