

Technological Educational Institute of Crete  
Department of Informatics Engineering

# **Machine learning data preparation for epileptic seizures prediction**

Andriana Theodorakopoulou

Submitted in part fulfilment of the requirements for the degree of  
Master of Science

Approved by: Dr. Giorgos Papadourakis

June 2017



ANDRIANA F. THEODORAKOPOULOU

2017

Creative Commons License

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

## **Abstract**

Epilepsy is a chronic neurological disorder that affects approximately 50 million people worldwide. These intractable seizures postures are a serious risk of injury, restrict the self-sufficiency and mobility of a person. They also result in social isolation and have a severe financial impact on the individuals. Prognosis of the above postures has been a research topic for the last 30 years.

A key task when you want to build an sophisticated analytic model using machine learning or deep learning technique is the integration and preparation (data cleansing and feature engineering) of the data sets. This step consists approximately 80 percent of the whole analytics project. In this thesis, we focus on data preparation of Electroencephalography (EEG) signals in order to extract accurate features and therefore predict epileptic seizures.

The data preparation is then applied on the CHB-MIT Scalp EEG database which was collected at the Childrens Hospital Boston and is consisted of scalp Electroencephalogram recordings from pediatric subjects with intractable seizures. In this database, all subjects were monitored for up to several days following withdrawal of anti-seizure medication.

In conclusion, this thesis implements different combination of data preprocessing algorithms to generate features for seizures prediction, which, in turn, can lead to substantial advances of neuroscience.

## **Acknowledgements**

I would like to express my gratitude to my thesis advisor Dr. Giorgos Papadourakis of the Department of Informatics Engineering at Technological Educational Institute of Crete. He consistently steered me in the right direction whenever he thought I needed it.

Finally, I must express my very profound gratitude to my family as well as to Ioannis Deligiannis for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Thank you.

# Contents

<b>Copyrights</b>	<b>i</b>
<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Why prediction is important? . . . . .	2
1.2 Search methodology . . . . .	2
1.3 Data preparation in machine learning projects . . . . .	3
1.4 Objective of the thesis . . . . .	4
1.5 Structure of the thesis . . . . .	4
<b>2 Theoretical Background</b>	<b>7</b>
2.1 A review of techniques in seizure detection and prediction . . . . .	7
2.1.1 Seizure prediction versus early seizure detection . . . . .	11
2.1.2 Seizure Prediction Methodology . . . . .	11
2.2 Human Brain Anatomy . . . . .	12
2.3 What is epilepsy? . . . . .	13
2.3.1 Categories of Epileptic Disorders . . . . .	15
2.3.2 Treatment for Epileptic Seizures . . . . .	16
2.3.3 Can Epileptic Seizures Be Predicted? . . . . .	17
2.4 Technical Background of Electroencephalography . . . . .	17
2.4.1 Impedance and Electrodes . . . . .	18
2.4.2 10-20 Electrode Placement System . . . . .	19
2.4.3 EEG Montages . . . . .	20
<b>3 Data Preparation</b>	<b>21</b>
3.1 Data Selection . . . . .	21
3.1.1 CHB-MIT Scalp EEG Database . . . . .	21
3.1.2 Data labels . . . . .	24
3.2 Preprocess and data cleansing . . . . .	28
3.2.1 Data sampling and filtering . . . . .	29
3.2.2 Missing values and outliers . . . . .	29
3.3 Ictal file extraction . . . . .	33

3.4	Remove Artifacts . . . . .	35
3.4.1	Discrete Wavelet Transform . . . . .	35
<b>4</b>	<b>Feature Engineering</b>	<b>39</b>
4.1	Moving Window Analysis . . . . .	39
4.2	Feature selection . . . . .	40
4.3	Brain Wave Patterns . . . . .	41
4.4	Linear Methods . . . . .	47
4.4.1	Hjorth Parameters . . . . .	47
4.5	Nonlinear Methods . . . . .	49
4.5.1	Coherence . . . . .	49
4.5.2	Entropies . . . . .	52
<b>5</b>	<b>Conclusion</b>	<b>55</b>
5.1	Summary of Thesis Achievements . . . . .	55
5.2	Applications and Future Directions . . . . .	56
<b>Bibliography</b>		<b>57</b>
.1	Appendix A : Brain waves . . . . .	70
.1.1	Normal activity at Subject Chb01 . . . . .	70
.1.2	Differences between pre-ictal and ictal state at Subject Chb01 . . . . .	71
.1.3	Differences between pre-ictal and ictal state at Subject Chb24 . . . . .	72
.2	Appendix B: Entropies . . . . .	73
.2.1	ShanEntropy & LogEntropy at Subject Chb10 . . . . .	73
.2.2	ShanEntropy & LogEntropy at Subject Chb01 . . . . .	74
.3	Appendix C: Coherence . . . . .	75
.3.1	Normal coherence at Subject Chb10 . . . . .	75
.3.2	Differences between pre-ictal and ictal state at Subject Chb10 . . . . .	76
.3.3	Deferencies between pre-ictal and ictal state at Subject Chb24 . . . . .	77
.4	Appendix D: Hjorth parameters . . . . .	78
.4.1	Normal Hjorth parameters at Subject Chb10 . . . . .	78
.4.2	Differences between pre-ictal and ictal state at Subject Chb02 . . . . .	79
.4.3	Differences between pre-ictal and ictal state at Subject Chb24 . . . . .	80

# List of Tables

## 2 Theoretical Background

2.1 Achievements of seizure prediction methods developed to date. Listed by the authors, the year of implementation since 1998 to 2015, the number of the false prediction rate according to by the total duration of interictal EEG data, the mean prediction time (MPT), sensitivity, and the rate of false predictions per hour. . . . .	9
2.2 Some of the epilepsy symptoms according to state (Prodromal Symptoms, Auras, Ictal, and Postictal) . . . . .	15
2.3 Types Of Epilepsy . . . . .	16

## 3 Data Preparation

3.1 Patient details of the CHB-MIT Scalp EEG Database. Subject: Patient, Gender: female (F) or male (M), Duration of Records: total duration of eeg signals (seizure and non-seizure), Ictal Duration: the sum of seizure duration as annotated by clinical experts. .	23
3.2 Channels details according to Bipolar Montage "Double Banana" . . . . .	27

## 4 Feature Engineering

4.1 Brain Waves (Gamma, Beta, Alpha, Theta, and Delta) . . . . .	42
--	----

# List of Figures

## 1 Introduction

1.1 Data Preparation as a step in the Knowledge Discovery Process . . . . .	4
---	---

## 2 Theoretical Background

2.1 Schematic representation of seizure detection BCI system . . . . .	10
2.2 Illustration of a correct prediction . . . . .	12
2.3 Left side: Illustrates the major parts of brain. Right side: Brain lobes . . . . .	13
2.4 In EEG analysis, the net electrical field from dipoles generated by cortical neurons are the most relevant features to capture by the electrodes. . . . .	18
2.5 10-20 SystEach site has a letter to identify the lobe and a number to identify the hemisphere location. Even numbers (2,4,6,8) refer to electrode positions on the right hemisphere, whereas odd numbers (1,3,5,7) refer to those on the left hemisphere. A "z" (zero) refers to an electrode placed on the midline. . . . .	19

## 3 Data Preparation

3.1 The International 10-20 system of surface electrode placement. The abbreviation stands for C= Central, F= Frontal, FP = Frontal-pole, O= Occipital, P= Parietal, and T= Temporal. Odd numbers represent the left-hand side electrodes, even numbers represent the right-hand side electrodes and letter Z the central line. . . . .	24
3.2 Electrical field mapping with bipolar montage. Negative potential at T4 is seen as a phase reversal between the F8-14 and T4-15 derivations. . . . .	25
3.3 Left temporal, left parasagittal, right parasagittal, right temporal. . . . .	26
3.4 Channels according to Bipolar Montage "Double Banana". Example of a seizure within the scalp EEG of Patient A. The seizure, which begins at 2996 seconds and ending at 3036 seconds. According to the graph, it is observed that the seizure comes from the Right Temporal Lobe. . . . .	28
3.5 An example of filtering in subject chb03 using a FIR low pass filter butterworth . . .	29
3.6 The signal appears mainly on frontal area (Fp1 and Fp2). The signal is symmetric between two hemispheres. . . . .	30
3.7 Eye movement make as well a slow signal (<4Hz) corresponding to mechanical movement.	31

3.8	Muscular artifacts. Muscular activity give high frequency signals ( $>13\text{Hz}$ ), often much higher than cerebral signals. . . . .	32
3.9	Electrode artifacts. Sometimes, wire or electrode movements create a low frequency artifact ( $<2\text{Hz}$ ) on one electrode. . . . .	33
3.10	EEG recording for subject Chb01 with epileptic seizure, Seizure Start Time: 2996 seconds Seizure End Time: 3036 seconds. Pre-ictal (yellow), ictal (red) and post-ictal (green) . . . . .	34
3.11	An example of eye blink artifacts in channel Fp1-F7 in 10 mins (Brainstorm toolbox)	35
3.12	Discrete wavelet transform (DWT) performed in 1 hour recording of channel FP1-F7 .	38

## 4 Feature Engineering

4.1	Seizure Prediction from EEG time series. Continuous multi-channel EEG recordings are analyzed by means of a moving-window analysis. The data covered by the orange window is transformed into a single value in the time profile of a multivariate characterizing measure. When this time profile crosses a certain pre-defined threshold, an alarm is issued [16]. . . . .	40
4.2	Gamma waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	43
4.3	Beta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd image) . . . . .	43
4.4	Alpha waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	44
4.5	Theta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	44
4.6	Delta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	45
4.7	The inner product between Delta and gamma waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	45
4.8	The inner product between Gamma and Beta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column) . . . . .	46
4.9	Results of Hjorth Parameters between interictal, preictal and ictal (seizure) . . . . .	49
4.10	An illustration of coherence between interictal, preictal and seizure . . . . .	50

4.11 Differences of coherence between preictal and seizure . . . . .	52
4.12 Shannon and Log energy Entropy in overall signal . . . . .	54

# Chapter 1

## Introduction

Epilepsy is a chronic neurological disorder in which nerve cell activity becomes disrupted, causing seizures or periods of abnormal behavior, influences sensations and sometimes results even in loss of consciousness. These intractable seizure postures are a serious risk of injury, restrict the self-sufficiency and mobility of a person.

Epileptic people state that : "Its the variability that really makes it so stressful. You never know when it is going to be chaos again and youll have one. Just because this morning is terrific doesnt mean tonight is going to be terrific, either behavior-wise, medication-wise, or any otherwise. So, it is the unpredictability of it that is really nerve-racking to live with." [1]

While medication and surgery can, to some degree, relieve the symptoms, these treatments fail to help all patients. Population wide, approximately 1% is suffering from epilepsy. Current treatment is inefficient on about 30% among those people [2]. Additionally, unforeseen seizures can put people with epilepsy at risk during everyday activities such as biking and driving.

A far as the diagnosis is concerned, this can be made using an electroencephalograph (EEG). EEG is a painless test which records the electrical activity of the brain. The results of the EEG recordings help doctors diagnose and make decisions about the appropriate treatment. Most epileptic seizures are controlled through drug therapy. In some cases where drugs and diet do not work, surgery may be used. The type of treatment followed depends on many factors, including the frequency and severity of crises, the age of individual, the general health, and the medical history.

The sudden and apparently unpredictable nature of epileptic seizures is one of the most disabling

aspects of epilepsy. Therefore, there has been a growing research interest, especially within the last 30 years, in seizure detection and prediction from EEG recordings. It is believed that developing a method capable of predicting the occurrence of seizures from the electroencephalogram (EEG) of epilepsy patients will open new therapeutic possibilities. Since the 1970s, studies on the predictability of seizures have advanced from preliminary reports of seizure precursors to controlled studies implementing prediction algorithms to continuous daily EEG recordings. Most of the seizure detection algorithms are patient specific; that is, they are applied to the patient for which training data are extracted. This is intuitively meaningful as each patient has a different nature for its EEG signals. On the other hand efforts have also been made to develop seizure prediction algorithms. Although most of the studies published in the 1990s and around the turn of millennium yielded rather encouraging results, more recent evaluations could not reproduce their findings, thus raising a debate about the validity and reliability of those studies.

## 1.1 Why prediction is important?

For many people with epilepsy, seizures appear at random times and greatly disrupt their cognitive and emotional state. The ability to predict epileptic seizures will significantly improve life quality of the epileptics by giving them a warning of an imminent crisis so they can take safety measures. Given this ability of some epilepsy patients to predict their own seizures through cognitive or behavioral signs (termed as "prodromes"), it is believed that although seizures typically seem to warn through an accidental process there is an underlying deterministic or dynamical component to the process of seizure generation that could be reliably detected for the purposes of seizure prediction.

## 1.2 Search methodology

The objective of this work is to examine the state of epileptic seizure prediction research by reviewing the literature, identifying current trends, describing the challenges that approach ESP diffusion, presenting open research questions and future directions.

The search process covered journal articles and conference papers (excluding surveys and literature reviews) which were available in four major electronic databases; PMC Digital Library, IEEE Explorer, Springer- Link, and ScienceDirect. In the initial stage, we identified relevant papers by analyzing

publications title and abstract. In the second stage, a full-text analysis was undertaken to discover and record the particular technologies reported in each of the relevant papers. After the second filtering stage, the primary studies were subsequently divided based on their keywords.

Each paper was carefully analyzed and classified into single categories according to its content. The broad categories were: challenges, technology, applications, overview/survey and future directions. The search was conducted in August 2016; therefore, results that were indexed after this date have not been included in this study.

### 1.3 Data preparation in machine learning projects

Data Preparation is the core of data science and it is crucial for any data analysis. It involves data cleansing and feature engineering. Usually, this takes 60 to 80 percent of the whole analytical pipeline. However, it is a mandatory task in order to get the best accuracy from machine learning algorithms on your datasets [3].

- **Data Cleansing:** is the process of altering data in a given storage resource to make sure that it is accurate and correct. Data cleansing is also known as data cleaning or data scrubbing. It includes many different functions such as:

Basics (select, filter, removal of duplicates,etc), Sampling (balanced, stratified), Data Partitioning (create training, validation, test data set), Transformations (normalization, standardization, scaling, pivoting), Binning (count-based, handling of missing values as its own group), Data Replacement (cutting, splitting, merging), Weighting and Selection (attribute weighting, automatic optimization, etc), Attribute Generation (ID generation), Imputation (replacement of missing observations by using statistical algorithms).

- **Feature Engineering:** Selection of the right attributes to be analyzed. One uses domain knowledge of the data to select or create attributes that make machine learning algorithms work. Feature engineering is fundamental to the application of machine learning, and is both difficult and expensive. It includes different functions like:

Brainstorming or testing of features, Feature selection, Validation of how the features work with your model, Improvement of features if needed and the return to brainstorming / creation of more features until the work is done.

Both data cleansing and feature engineering are part of data preparation and fundamental to the application of machine learning and deep learning algorithms. They are also complicated and time-consuming.

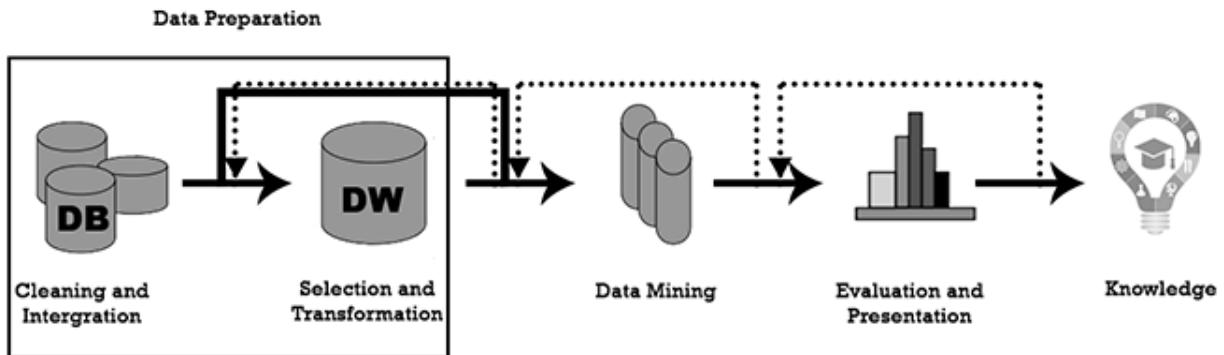


Figure 1.1: Data Preparation as a step in the Knowledge Discovery Process

## 1.4 Objective of the thesis

This study aims to investigate and extract features for epilepsy prediction seizures with emphasize on data preparation of EEG signals for machine learning algorithms. It is hoped that it will also contribute to better understanding the nature of epilepsy seizure and improve prediction methods.

The data preparation is applied on the CHB-MIT Scalp EEG database which was collected at the Childrens Hospital Boston (CHB). This database is consisted of scalp EEG recordings from pediatric subjects with intractable seizures. In order to reach these objectives, a survey was carried out and the results of the experiments were recorded.

As a future step towards clinical applications, the suggested prediction method should be tested on large datasets of EEG recordings from patients with intractable epilepsy. Its application to other forms of this type of epilepsy is conceivable but will require revisiting the training procedure and be re-evaluating its performance and statistical validity.

## 1.5 Structure of the thesis

The thesis is divided into four chapters covering the theoretical background, the preprocessing of data, features extraction techniques, classification and conclusion followed by references. A brief description

of each chapter is provided in the following paragraph.

**The Chapter of Theoretical Background** in chapter 2, we discuss the nature of epilepsy and the categories of epileptic disorders. It also contains a brief analysis of electroencephalogram (EEG) which is analyzed in details: the genesis of the electrical pulse from the brain, the recording of this electrical activity, and the analysis of the EEG waveform produced.

**The Chapter of Data Preparation** in chapter 2, we describe the process of selecting and analyzing the data to be used for the implementation of the project. One of the fundamental purposes of data preparation is to ensure that data being prepared for analysis are accurate and consistent, so the results for the features extraction will be valid. Additionally, we include the analysis of EEG records from the database. Analyzing the placement of the electrodes in the patient's cranium will enable us to move to the next stage of the data preprocessing. Removal of the artifacts also describes the procedure of resampling as well as the techniques for removal artifacts. The process of correcting inaccuracies and joining data sets constitutes a big part of the data preparation process.

**The Chapter of Features Extraction Techniques** In this chapter become an investigation of the optimal features, selects the right attributes to analyze. The domain knowledge of the data to select or create attributes that make machine learning algorithms work. The input EEG signal scanned with sliding window technique and calculated features where extracted from each second, the output of the signal stage is a label for each features value.

**Finally the Chapter of summary and conclusion** contains brief discussion about the topic which may require further study and investigation as well as future work and implementation of the robotic platform. The thesis is concluded in chapter 6 with references.



# Chapter 2

## Theoretical Background

This chapter provides an overview of the human brain and describes the disease of epilepsy, and its impact on it.

### 2.1 A review of techniques in seizure detection and prediction

The attempt to predict epilepsy seizures has been a research topic for the last 30 years. The multidisciplinary fields of neuroscience, computer science, microelectronics, bioengineering, and neurophysiology have contributed to build and improve an efficient Brain Computer Interface system steadily. Being able to predict seizures and couple this information with state of the art technology will allow patients to take action before the occurrence of the seizure and thus, minimizing potential risk [4]. Developing such prediction devices will be extremely beneficial for the epileptic patients since they will be able to constantly monitor their state autonomously and perhaps prevent their hospitalization.

The first attempts for seizure detection and prediction were carried out by Viglione and Walsh [5]. They tried to find seizure precursors using linear approaches for absence seizure EEGs. They defined absence seizure as a specific category of seizure which is associated with brief loss of consciousness. Rogowski et al. [6] and later Salant et al. [7] were able to detect changes 6 seconds before seizure onset, using an autoregressive model of the neuronal activity. Siegel et al. [8] found changes among 1-minute epochs before the seizure and conducted further analysis on the spike occurrence rates in the EEG, indicating decreased focal spike-rate along with an increased rate of bilateral spikes before the seizure. Following studies by Le Van Quyen et al. [9] compared pre-seizure dynamic variations to

those of interictal (between seizures) EEG and developed a dynamical similarity index which seemed to decrease before seizures. Another groundbreaking discovery was made by Iaesemidis et al. [10]. By using the Lyapunov exponent and an open window analysis they revealed a chaotic behavior in invasive EEG and a decrease in this behavior before the seizure. [11]

Considering the nature of epilepsy, there are many types of seizures. This can sometimes be a challenging task to address, especially when one considers that some of the epileptic syndromes are difficult to be characterized and therefore are being classified as particular category [12]. In addition, understanding of the underlying mechanisms leading to seizures and the origin of a seizure in each case is still under investigation.

Up to now, several investigations based on nonlinear time series analysis have been carried out on intracranial and surface EEG that provided data with promising results. It has been claimed that seizures can be predicted at least 20 minutes beforehand, maybe up to 1 hour and 30 minutes before the onset of temporal lobe epilepsy. Table 1.1 summarizes the accomplishments of seizure prediction methods developed by a chronological overview with relevant characteristics.

Table 2.1: Achievements of seizure prediction methods developed to date. Listed by the authors, the year of implementation since 1998 to 2015, the number of the false prediction rate according to by the total duration of interictal EEG data, the mean prediction time (MPT), sensitivity, and the rate of false predictions per hour.

Year	Authors	Method	MPT (min)	(FP/h)	sensitivity (%)
1998	Lehnertz and Elger [13]	Effective correlation dimension	11.5	0	94
1999	Le Van Quyen et al. [9]	Similarity index	5.75	-	83
2000	Le Van Quyen et al. [14]	Similarity index	4.45	-	94
2001	Iasemidis et al. [15]	Lyapunov exponent	49.1	-	91
2002	Schindler et al. [16]	LIFU (surface EEG)	4-330	0.3	100
2003	Mormann et al. [17]	Phase coherence, lin. cross corr	86/102	-	86
2004	VP Nigam, D Graupe. [18]	Nonlinear filtering	-	-	97.2
2005	Srinivasan et al. [19]	Time and frequency features, ANN	-	0	99.6
2006	Guler and Ubeyli [20]	PSD features, modified mixture of experts	-	-	98.6
2007	Tzallas et al. [21]	Timefrequency analysis, ANN	-	-	96.3
2008	Polat and Gunes [22]	PCA-FFT, AIRS classifier	2.7	0	100
2009	Ocak [23]	Wavelet transform, ApEn	-	-	94.85
2010	Kumar et al.[24]	Entropy measures, Recurrent Elman network (REN)	-	-	99.75
2011	Fathima et al. [25]	Discrete wavelet transform	-	-	99.5
2012	Martis et al.[26]	EMD features	-	-	95.33
2013	Alam et al.[27]	EMD statistics and ANN	-	-	100
2014	Nilufer et al.[28]	Hilbert Huang transform, Bayesian Classifiers	-	-	96.55
2015	Martis et al.[29]	wavelet leader based scaling and cumulant estimation	-	-	80.5

There are numerous techniques and algorithms for analysis and classification of bio signals, 1 or 2-dimensional, in time or frequency distribution. The majority of algorithms follow the structure of selecting and applying feature extraction methods on the EEG signal and using classification methods to conclude the diagnosis. The seizure detection method of epilepsy can be made on a single or multi-channel way. Single-channel seizure detection requires selecting the channel receiving the strongest EEG signal collected from the closest point to the seizure spot. Some algorithms create models for normal and abnormal EEG signals of the patients and use these models in the training process.

EEG feature extraction has been studied from early time and lots of advanced techniques and transformations have been proposed for accurate and fast EEG feature extraction. For example, discrete Fourier transform (DFT) and discrete wavelet transform (DWT) have found popularity in seizure detection and prediction applications also the power spectral features, higher order spectral methods, nonlinear transformations such as Lyapunov exponents have been used as appropriate sources for feature extraction. The classifying methods which have been proposed during the last decade include Fuzzy Logic methods, Artificial Neural Network, Hidden Markov Model, Support Vector Machines, Cluster analysis, with each approach exhibiting its own advantages and disadvantages. Tzallas et al. presented a classification of EEG seizure detection methods into pattern recognition, morphological analysis, parametric, decomposition, clustering and data mining methods [30].

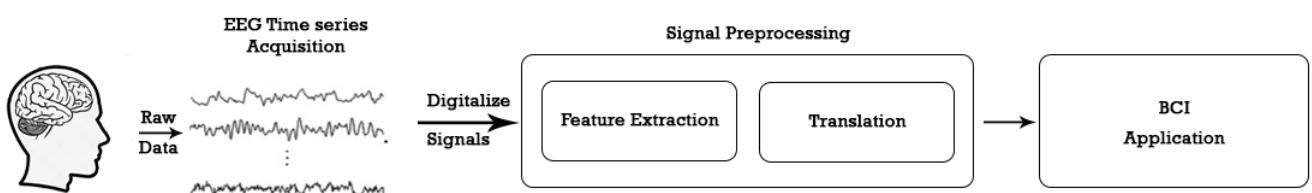


Figure 2.1: Schematic representation of seizure detection BCI system

The combination of signal processing with the electronic devices serves as a primary root for the development of various biomedical applications. Ideally, a seizure prediction method/algorithm has to guess an impending epileptic seizure by raising an alarm in advance of the seizure onset [13-21]. A perfect prediction method will be able to show the exact point in time when a seizure will occur.

### **2.1.1 Seizure prediction versus early seizure detection**

It is important to clarify the differences between seizure detection studies from seizure prediction studies. *Seizure detection* relates to the automatic recognition of seizures presented before or after the actual onset, usually in a short prediction window of a few seconds long. *Seizure prediction* represents the automatic recognition of seizures well in advance of the actual onset where the prediction window can be several minutes long [31]. The difference between the above two is important because of their applications are likely to be different.

### **2.1.2 Seizure Prediction Methodology**

Before the occurrence of a seizure episode, there are some clinical symptoms which have been shown to present. Among these are an increase in oxygen availability, cerebral blood flow, blood oxygen dependent signals, changes in heart rate etc [32][33][34]. In addition to these changes, it is believed that an increased number of critical interactions between neurons in the focal area unfolds over time [11]. This concept allows researchers to study electroencephalography in an alternative way, find correlations of these processes and identify the situation before the (pre-ictal state) seizure.

A seizure prediction method has to forecast an impending epileptic seizure by raising an alarm in advance of the seizure onset. A perfect prediction method indicates the exact point in time when a seizure occurs. This ideal behavior is not expected for current prediction methods that analyze EEG data.

The uncertainty can be considered by use of the seizure occurrence period (SOP) which is determined as a period during which the seizure is to be expected. Also, to permit a therapeutic intervention, a minimum window of time among the alarm raised by the prediction algorithm this window of time is called seizure prediction horizon (SPH) [35]. Considering the two-time periods SPH and SOP, a correct prediction is defined as follows: after the alarm, during SPH, no seizure has occurred yet. During SOP, a seizure occurs. The accurate time of seizure onset may vary within SOP, through reflecting the possibility of the prediction. Seizures outside of any seizure occurrence period (SOP) are not predicted by the system and therefore are classified as false negatives. Alarms without a seizures during SOP are considered as false predictions. The measures which describe a prediction method performance of SPH and SOP are sensitivity and false prediction rate. The first is defined as the fraction of correctly

predicted seizures within the total seizures. The second corresponds to the number of false predictions per time interval. It should be mentioned that the above measures are not independent.



Figure 2.2: Illustration of a correct prediction

## 2.2 Human Brain Anatomy

The human brain is the most important and major part of the central nervous system is one of the largest organs in the body. It consists of more than 100 billion nerve cells that communicate with trillions of connections called synapses. It is also consisted of a layer of tissue called the meninges and the skull (cranium) which helps to protect the brain from injury.

The brain is made up of many specialized areas that work together:

- The *cortex* is the outermost layer of brain cells. Thinking and voluntary movements begin in the cortex.
- The *brain stem* is between the spinal cord and the rest of the brain. All areas of the nervous system connect with the brain stem. It regulates heart rate, breathing, blood pressure, digestion, sweating, level of alertness and ability to sleep. It also contains reflex centres that control swallowing, coughing, sneezing and vomiting.
- The *basal ganglia* are a cluster of structures in the center of the brain. The basal ganglia coordinate messages between multiple other brain areas. The basal ganglia are associated with a variety of functions including: control of voluntary motor movements, procedural learning, routine behaviors or "habits" such as teeth grinding, eye movements, cognition, and emotion.
- The *cerebellum* is at the base at the back of the brain. The cerebellum is responsible for coordination and maintaining of posture, muscle tone and balance.

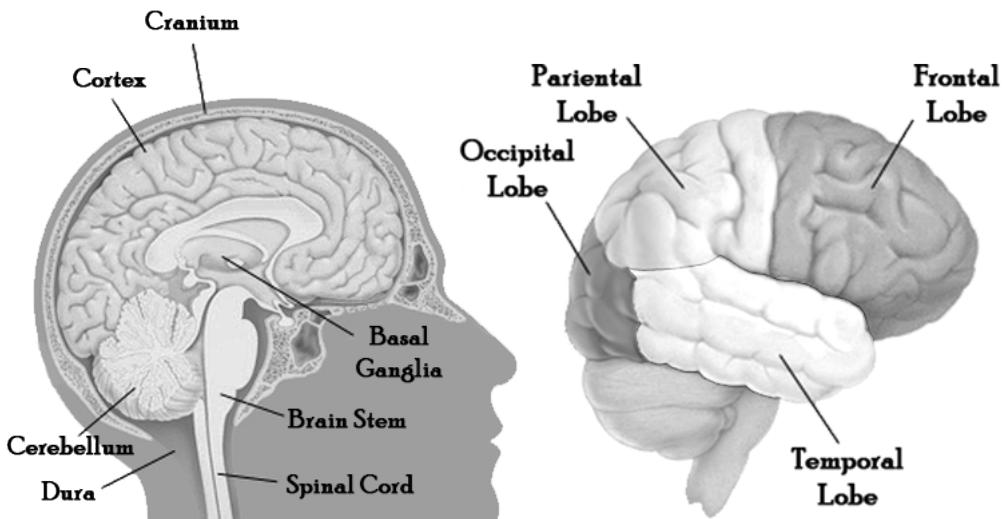


Figure 2.3: Left side: Illustrates the major parts of brain. Right side: Brain lobes

The human brain is also divided into four different lobes:

- The frontal lobes are responsible for problem solving and judgment and motor function.
- The parietal lobes manage sensation, handwriting, and body position.
- The temporal lobes are involved with memory and hearing.
- The occipital lobes contain the brain's visual processing system.

## 2.3 What is epilepsy?

Epilepsy is a paroxysmal disorder which produces small and intense outbursts of brain electrical activity during which clusters of nerve cells, or neurons, in the brain sometimes signal abnormally and cause seizures. Neurons normally generate electrical and chemical signals that act on other neurons, glands, and muscles to produce human thoughts, feelings, and actions. During a seizure, many neurons fire (signal) at the same time as many as 500 times a second, much faster than normal [36]. The chemical or structural disorder of neurons causes epilepsy in the brain an abnormal electrical explosion activity that results in the appearance of crisis. The crisis is temporary and affects muscle control, movement, speech or vision. People with epilepsy have recurrent crises, which usually occur without any warning and for no apparent reason.

When a *seizure* takes place, a disruption in the normal pattern of these electrical impulses in the brain is observed, which in turn is caused by the brain cells firing simultaneously at a much faster

rate. Sometimes, the part of the body affected can be traced back to a specific area within the brain where excess electrical activity is taking place. A seizure may alter behaviour, consciousness, movement, perception and/or sensation.

Epilepsy can be considered a spectrum disorder because of its different causes, different seizure types, its ability to vary in severity and impact from person to person, and its range of co-existing conditions. Some people may have convulsions (sudden onset of repetitive general contraction of muscles) and lose consciousness. Others may simply stop what they are doing, have a brief lapse of awareness, and stare into space for a short period [36]. Some people have seizures very infrequently, while others may experience hundreds of them each day. Different types of epilepsy can result from a variety of causes. Recent adoption of the term the epilepsies underscores the diversity of types and causes.

Anyone can develop epilepsy. About 2.3 million adults and more than 450,000 children and adolescents in the United States currently live with epilepsy. Each year, an estimated 150,000 people are diagnosed with epilepsy [36]. Epilepsy affects both males and females of all races, ethnic Predicting Epileptic Seizure from EEG data using HHT and Neural Network backgrounds, and ages. In the United States alone, the annual costs associated with the epilepsies are estimated to be \$15.5 billion in direct medical expenses and lost or reduced earnings and productivity. The majority of those diagnosed with epilepsy have seizures that can be controlled with drug therapies and surgery. However, as much as 30 to 40 percent of people with epilepsy continue to have seizures because available treatments do not completely control their seizures (called intractable or medication resistant epilepsy).

**States of Epilepsy Seizures and Symptoms** A seizure often has six separate phases: *Prodromal Symptoms, Auras, Pre-ictal, Ictal, Inter-ictal, and Postictal Stages*.

- *The prodromal stage*, includes mostly emotional signals. This prodromal group of symptoms befalls days or hours before a seizure follows. This prodromal group of symptoms befalls days or hours before a seizure follows.
- In an *aura*, changes in activity, emotions, hearing, smell, taste, visual perception are involved. Auras are a small partial seizure that is often followed by a larger event. They usually come a few seconds to a few minutes before the actual seizure.
- *Pre-ictal period*, refers to the state immediately before the actual seizure, stroke, or headache, though it has recently come to light that some characteristics of this stage (such as visual auras)

are actually the beginnings of the ictal state.

- *Ictal period*, refers to a physiologic state or event such as a seizure, stroke, or headache. The word originates from the Latin ictus, meaning a blow or a stroke. In electroencephalography (EEG), the recording during a seizure is said to be "ictal".
- *Inter-ictal period*, refers to the period between seizures, or convulsions, that are characteristic of an epilepsy disorder. For most people with epilepsy, the interictal state corresponds to more than 99% of their life. The interictal period is often used by neurologists when diagnosing epilepsy since an EEG trace will often show small interictal spiking and other abnormalities known by neurologists as subclinical seizures. Interictal EEG discharges are those abnormal waveforms not associated with seizure symptoms.
- The *Post-ictal period*, occurs after the ictus or active stage of the seizure. As the seizure ends, the postictal phase occurs. This is the healing period after the seizure. Some people recover immediately, while others may take minutes to hours to feel like their usual self.

Table 2.2: Some of the epilepsy symptoms according to state (Prodromal Symptoms, Auras, Ictal, and Postictal)

Prodromal Symptoms	Auras	Ictal Phase	Postictal Phase
Affection	Biting of tongue	Blurry vision	Anxiety
Depression	Blinking of eyes	Confused	Depression
Difficulty concentrating	Change in skin color	loses consciousness	Exhaustion
Ecstatic feelings	Difficulty breathing	Deja vu	Fear
Headache	Dreamlike experiences	Feeling detached	Headache
Insomnia	Lack of movement	Flashing lights	Memory loss
Mood changes	Unpleasant smells	Heart racing	Sleepiness

### 2.3.1 Categories of Epileptic Disorders

During seizure, the patients are possible to exhibit different characteristics which depending on the type of epilepsy and the proportion of the brain which affected. The duration of epileptic crisis may hold for a few seconds to a few minutes. The categorization based off the type each crisis, the percentage of the brain involved in the particular disorder, the clinical symptoms and the form of electroencephalogram (EEG). The classification of epileptic crisis represented in above table 2.2.

The seizures can be divided into two categories the focal and primary generalized seizures, partial seizures originate from the focal region of the brain, nevertheless they can be extended rapidly to other areas of the brain creating secondarily generalized seizures. Such crises are divided into those that affect and impair consciousness (complex partial), where there is a high correlation between the symptoms presented by the patient and the part of the brain that begins abnormal discharge of the brain and that do not affect or harm consciousness (simple partial). The last one as known as temporal sclerosis is the most common type of seizures in adults and more severe about to treatment. It is possible to have a warning sign, called aura precedes the loss or reduction of consciousness these crises during less than three minutes and patients can be seen to keep their conscience. After the end of the crisis following his sleep patient, they achieve full recovery of consciousness.

Table 2.3: Types Of Epilepsy

<b>Types of Epilepsy</b>	<b>Generalized Epilepsy</b>	<b>Partial Epilepsy</b>
Idiopathic (genetic causes)	<ul style="list-style-type: none"> <li>- Childhood absence epilepsy</li> <li>- Juvenile myoclonic epilepsy</li> <li>- Grand-mal seizures on awakening Others</li> </ul>	<ul style="list-style-type: none"> <li>- Benign focal epilepsy of childhood</li> </ul>
Symptomatic (cause unknown) or cryptogenic (cause unknown)	<ul style="list-style-type: none"> <li>- West syndrome</li> <li>- Lennox-Gastaut syndrome</li> <li>- Others</li> </ul>	<ul style="list-style-type: none"> <li>- Temporal lobe epilepsy</li> <li>- Frontal lobe epilepsy</li> </ul>

### 2.3.2 Treatment for Epileptic Seizures

When epilepsy is diagnosed, it is important to begin treatment as soon as possible. For about 70% of those diagnosed with epilepsy, seizures can be controlled with medicines and surgical techniques. Some drugs are more efficient for specific types of seizures [36]. A patient with seizures, particularly those that are not easily controlled, may require to see a neurologist and take a specifically trained to treat epilepsy. In some children, special diets may help to control seizures when medications are either not effective or cause serious side effects. While epilepsy cannot be remedied, for some people the seizures can be controlled with medication, diet, devices, or surgery. Most seizures do not provoke

brain damage, but ongoing uncontrolled seizures may cause harm. It is not uncommon for people with epilepsy, especially children, to develop behavioral and emotional problems in association with seizures. For many people with epilepsy, the risk of seizures limits their autonomy (some states resist drivers licenses to people with epilepsy) and recreational activities [36].

### **2.3.3 Can Epileptic Seizures Be Predicted?**

Recent studies showed evidence of premonitory signs in 6.2% of 500 epileptic patients. Another study found that 50% of 562 patients observed auras before seizures [37]. These clinical investigations give the inspiration to search for premonitory changes on electroencephalographic (EEG) recordings from the brain. The epileptic brain has transitions between different states of activity: from normal interictal to preictal then ictal (seizure) and postictal, before recovering to the interictal state [38].

Despite the current lack of a complete neurological understanding of the preictal brain state, which is patient and conditions specific, researchers increasingly hypothesize that brainwave synchronization patterns might differentiate interictal, preictal and ictal states. The specific seizure prediction task becomes a classification problem where one aims at separating between interictal and preictal patterns of brain activity. Ictal and postictal states are discarded from the classification because the task is not to detect experiencing seizures to warn the patient or clinician about future ones.

## **2.4 Technical Background of Electroencephalography**

The brain waves have always been an interesting area for many kinds of research. The electroencephalogram (EEG) is such a time series of wave-like voltage fluctuations that record the activities of large ensembles of neurons in the brain cortex. Approximately 80 years ago, Hans Berger, a German scientist discovered electroencephalography (EEG). An electroencephalogram is a noninvasive test used to measure the electrical activity in the cerebrum, EEG usually used to diagnose epilepsy, which causes abnormalities in EEG readings. Changes from the normal pattern of electrical activity can show certain conditions, such as seizures.

The EEG graph represents the voltage differences between two different locations over time, the signal is recorded at the scalp but actually arises from pyramidal neurons located in the cortical layers, arranged in columns perpendicular to the brains surface. The sum of thousands of synchronized

postsynaptic potentials will generate electric currents recorded at the surface [39]. The scalp EEG can only detect electrical potentials generated near the brain surface. Moreover, at least  $6 \text{ cm}^2$  (or, roughly, one square inch) of the cortex with synchronous neuronal activity is needed to create a scalp potential. This surface activity is synchronized and modulated by complex neuronal networks involving interactions between the cortex and deep structures of the brain, mainly the thalamus, as well as other cortical areas. This will produce the rhythmicity and the waves of the brain activity seen on the EEG [39].

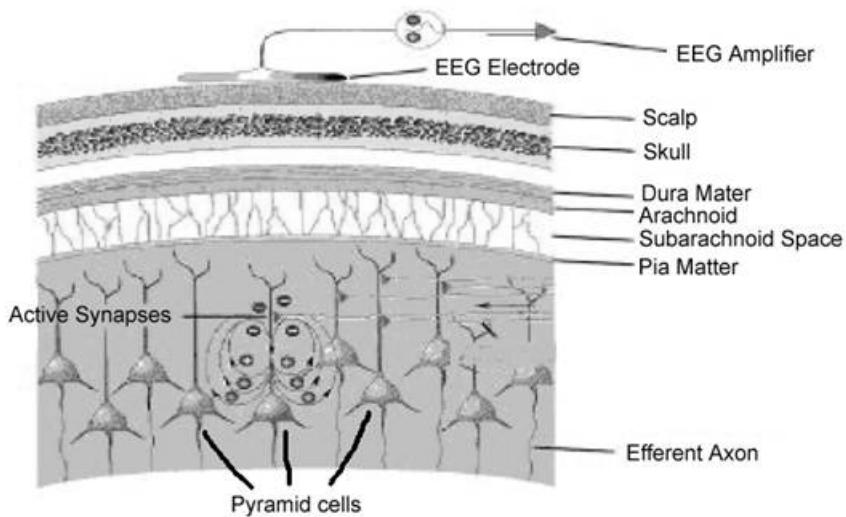


Figure 2.4: In EEG analysis, the net electrical field from dipoles generated by cortical neurons are the most relevant features to capture by the electrodes.

#### 2.4.1 Impedance and Electrodes

The potentials recorded on scalp EEG are measured in microvolt (V), typically 10–100, while, as a comparison, the electrocardiogram measures are in millivolt (mV). These voltage signals have to cross several layers and potential electrical barriers to reach the surface, such as the cerebrospinal fluid, the dura, the skull, and finally the scalp as presented in figure 2.2. Then, the current will go through electrodes and wires, which may also be considered as electrical obstacles. All these structures produce an opposition to the electric signal. In an alternating current (AC) circuit, this is called impedance and it is measured in  $\Omega$ . According to the International Federation of Clinical Neurophysiology (IFCN) guidelines and the American Clinical Neurophysiology Society (ACNS) recommendations, electrodes impedances should be checked before every recording and should not exceed  $5 \text{ K}\Omega$  [40]. In practice, one can accept values up to  $10 \text{ K}\Omega$ . Obtaining low and uniform impedances throughout all the electrodes

is crucial to obtain a reliable EEG signal and avoid artifacts. On the other hand, impedances less than  $1\text{ K}\Omega$  may indicate a possible shortcut between electrodes and should be specifically addressed [39].

#### 2.4.2 10-20 Electrode Placement System

The international 10-20 system is based upon a separation of the head from nasion (the point between the eyes on top of the nose) over cranium (top of the scalp) to inion (the bump on the lower back part of the skull). The 10-20 refers to the percentages that demarcate the electrode positions on this line that traverses the scalp. The electrodes of an EEG are put in specific locations on the patient's head, which means that several electrodes can record the activity from several and known areas of the cranium. When the technician or doctor is studying the results of the EEG, can recognize what brain activity is happening, and in which particular part of the brain it is happening in [36].

To work out where to put the electrodes, a special arrangement called the 10-20 system is used. Each electrode is put either 10 or 20 per cent of the total distance between specific points on the head, done by measuring the persons head and marking the position with a soft pencil. Each electrode has a number; all the odd numbers are on the left side of the head, and the even numbers on the right [36]. The electrodes also have a letter, depending on the area of brain that it is recording from: F for frontal lobe, T for temporal lobe, P for parietal and O for occipital lobes. The letter Z is used for the line of electrodes sited on the midline of the head.

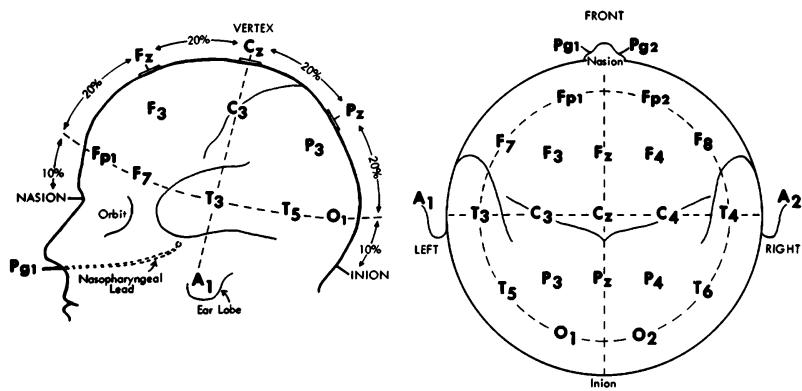


Figure 2.5: 10-20 SystEach site has a letter to identify the lobe and a number to identify the hemisphere location. Even numbers (2,4,6,8) refer to electrode positions on the right hemisphere, whereas odd numbers (1,3,5,7) refer to those on the left hemisphere. A "z" (zero) refers to an electrode placed on the midline.

### 2.4.3 EEG Montages

As we say before an EEG voltage signal represents a difference between the voltages at two electrodes, the representation of the EEG for the reading Electroencephalograph machine may be set up in one with different ways. The representation of the channels is referred to as a montage all signals are typically digitized and stored in a particular montage; since any montage can be constructed mathematically from any other, the EEG can be viewed by the Electroencephalograph in any display montage that is desired.

The EEG can be monitored with either a bipolar montage or a referential:

- **Bipolar montage:** In our case, we use this type of montage, each channel (waveform) represents the difference between two adjacent electrodes. The entire montage consists of a series of these channels. For example, the channel "Fp1-F3" represents the difference in voltage between the Fp1 electrode and the F3 electrode. The next channel in the montage, "F3-C3," represents the voltage difference between F3 and C3, and so on through the entire array of electrodes [41].
- **Referential montage:** Each channel represents the difference between a certain electrode and a designated reference electrode. There is no standard position for this reference; it is, however, at a different position than the "recording" electrodes [41]. Midline positions are often used because they do not amplify the signal in one hemisphere vs. the other. Another popular reference is "linked ears," which is a physical or mathematical average of electrodes attached to both earlobes or mastoids.

In this project, used an 23-channel eeg cap for the collection and analysis of EEG signals with Bipolar montage Which is discussed in detail in Chapter 3 "Data Preprocessing".

# Chapter 3

## Data Preparation

This chapter describes the process of selecting and analyzing the data in advance of algorithmic and machine learning implementation of experiments. One of the fundamental purposes of data preparation is to ensure that data being prepared for analysis is accurate and consistent, so the results for the features extraction will be valid. Additionally, describes the procedure of resampling as well as the techniques for removal artifacts. The process of correcting inaccuracies and joining data sets constitutes a big part of the data preparation process.

### 3.1 Data Selection

The CHB-MIT Scalp EEG Database is one of the most cited resources used in prediction detection experiments. It is also one of the few publicly available invasive EEG datasets. The database contains 24 hour-long continuous pre-surgical invasive EEG recordings of 22 patients suffering from epilepsy. The patients are from a wide range of varying age, sex, seizure type and seizure locality, but they all suffer from focal medically intractable epilepsy and were admitted for pre-surgical evaluation at the Childrens Hospital Boston.

#### 3.1.1 CHB-MIT Scalp EEG Database

This database was collected at the Childrens Hospital Boston (CHB) in collaboration with the Massachusetts Institute of Technology (MIT), consists of scalp EEG recordings from pediatric subjects

with intractable seizures [42] [43]. The database includes 22 subjects (5 males, ages 322; and 17 females, ages 1.519) which were monitored for up to several days following withdrawal of anti-seizure medication. The recordings were grouped into 24 cases containing continuous scalp EEG recordings from a single patient, except cases 1 and 21 which were collected from the same patient within a two-year interval. The resulting database includes a total duration of 940 hours of mainly continuous EEG activity, in which the patients experienced 198 events that were classified as epileptic seizures by clinical experts. For each seizure, the experts indicated the earliest time interval where seizure-related EEG changes could be visually detected, as well as the equivalent seizure end time interval. More details about the CHB-MIT database are presented below in Table 3.1.

The recordings were organized in files that include 1 hour long EEG signals (2 or 4 hrs. per file in some cases) following the European Data Format (.EDF), which is widely used for the storage of biological or physical multichannel signals. EDF files start with a header which contains some general information, such as patient identification, start/ end time of record, etc., and technical specification of each signal (e.g. sampling rate, filtering, the name of the electrode) coded as ASCII characters. The header is followed by the data records that contain digitized EEG signal samples formatted as little-endian 16-bit integers. All signals were sampled at 256 Hz and digitized at 16-bit resolution.

Table 3.1: Patient details of the CHB-MIT Scalp EEG Database. Subject: Patient, Gender: female (F) or male (M), Duration of Records: total duration of eeg signals (seizure and non-seizure), Ictal Duration: the sum of seizure duration as annotated by clinical experts.

Subject	Gender	Age	Number of Seizures	Duration of Records (hour: min: sec)	Ictal Duration(min)
chb01	F	11	7	40:33:08	7.37
chb02	M	11	3	35:15:59	2.87
chb03	F	14	7	38:00:06	6.7
chb04	F	22	4	156:03:54	6.3
chb05	F	7	5	39:00:10	9.3
chb06	F	1.5	10	66:44:06	2.55
chb07	F	14.5	3	67:52:18	5.42
chb08	M	3.5	5	20:00:23	15.32
chb09	F	10	4	67:52:18	4.6
chb10	M	3	7	50:01:24	4.45
chb11	F	12	3	34:47:37	13.43
chb12	F	2	40	23:41:40	24.58
chb13	F	3	12	33:00:00	8.92
chb14	F	9	18	26:00:00	2.82
chb15	M	16	20	40:00:36	33.2
chb16	F	7	10	19:00:00	1.4
chb17	F	12	3	21:00:24	4.88
chb18	F	18	6	35:38:05	5.28
chb19	F	19	3	29:55:46	3.93
chb20	F	6	8	27:36:06	4.9
chb21	F	13	3	32:49:49	3.32
chb22	F	9	7	31:00:11	3.4
chb23	F	6	16	26:33:30	7.04

In a few records and cases, other signals such as the Electrocardiogram (ECG) and Vagal Nerve Stimulus (VNS) were also included. Finally, in order to protect patients' privacy, all protected health information (PHI) in the original files has been replaced with surrogate entries. The recording dates have also been replaced. However, the time relationships between the individual files belonging to each case have been preserved. The data were segmented into one, two and four hour long records.

### 3.1.2 Data labels

In order to perform measurements for feature extraction, it is needed to localize the channels from the dataset recordings, which contain the raw brain waves with some technical specification of each signal. Most files contain 23 EEG signals (24 or 26 in a few cases) from surface electrodes that were placed on the patient's head following the International 10-20 system of electrodes position which give good coverage of cortical potentials generated by superficial cortex.

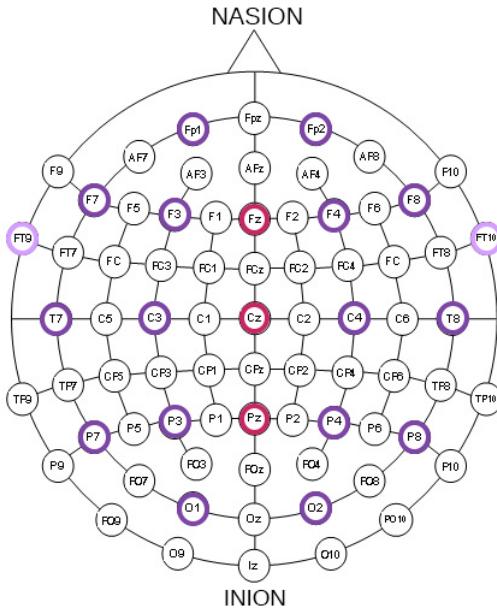


Figure 3.1: The International 10-20 system of surface electrode placement. The abbreviation stands for C= Central, F= Frontal, FP = Frontal-pole, O= Occipital, P= Parietal, and T= Temporal. Odd numbers represent the left-hand side electrodes, even numbers represent the right-hand side electrodes and letter Z the central line.

We must also say that EEG recordings use differential amplifiers that subtract one voltage signal from another and create patterns of connections between electrodes and amplifiers. The way these signals are displayed is called "montages," from the French word which means "putting together", in the same way that a video or film montage is an assembly of video clips. Each montage is an organized collection of channels derived from two electrodes (and hence called a "derivation") [44].

What montage to use in each situation depends on the nature of EEG activity and its location. The standard montages, including referential montages (where each "exploring electrode" is compared to a single or paired "reference electrode") and bipolar montages in which closely related electrode pairs are linked together in chains. Each channel (waveform) represents the difference between two adjacent electrodes. The entire montage consists of a series of these channels. For example, the channel "Fp1-F7" represents the difference in voltage between the Fp1 electrode and the F7 electrode. The next channel in the montage, "F7-T7," represents the voltage difference between F7 and T7 and so on through the entire array of electrodes.

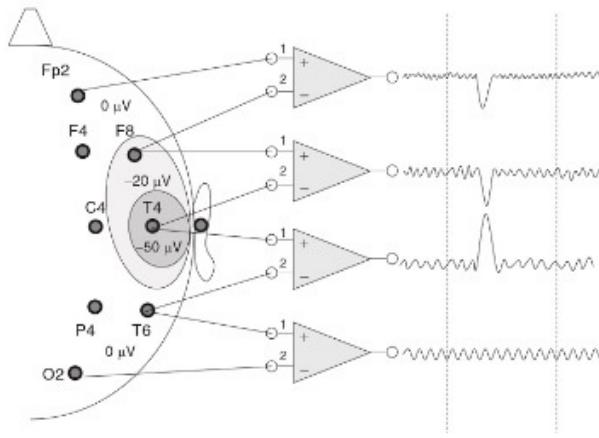


Figure 3.2: Electrical field mapping with bipolar montage. Negative potential at T4 is seen as a phase reversal between the F8-14 and T4-15 derivations.

In our case the data have been recorded with the bipolar montage method. The temporal chain montage we have used is part of the standard longitudinal bipolar montage, which is commonly referred to as the "double banana" due to its two crescent-moon (banana)-shaped outlines formed by the pairs of temporal and parasagittal chains (Fig. 3.3). This is one of the more popular ways to view routine EEG data, by providing views that allow quick front-to-back and side-to-side comparisons.

This montage can be arranged for viewing in several different ways. The chains can be arranged left to right (left temporal, left parasagittal, right parasagittal, right temporal) or be grouped by paired localization (left temporal, right temporal, left parasagittal, right parasagittal). The mid-sagittal chain is sometimes placed between the left and right side chains, at the beginning or at the end of the montage. In British publications, right is often on top, while left on top is more common in the United States. There is no "correct" arrangement, and each lab may choose a preferred display arrangement [44]. Knowing how the channels are positioned in the patient's cranium, we could divide the corresponding tables in the Matlab environment so that we have specific areas of the brain that

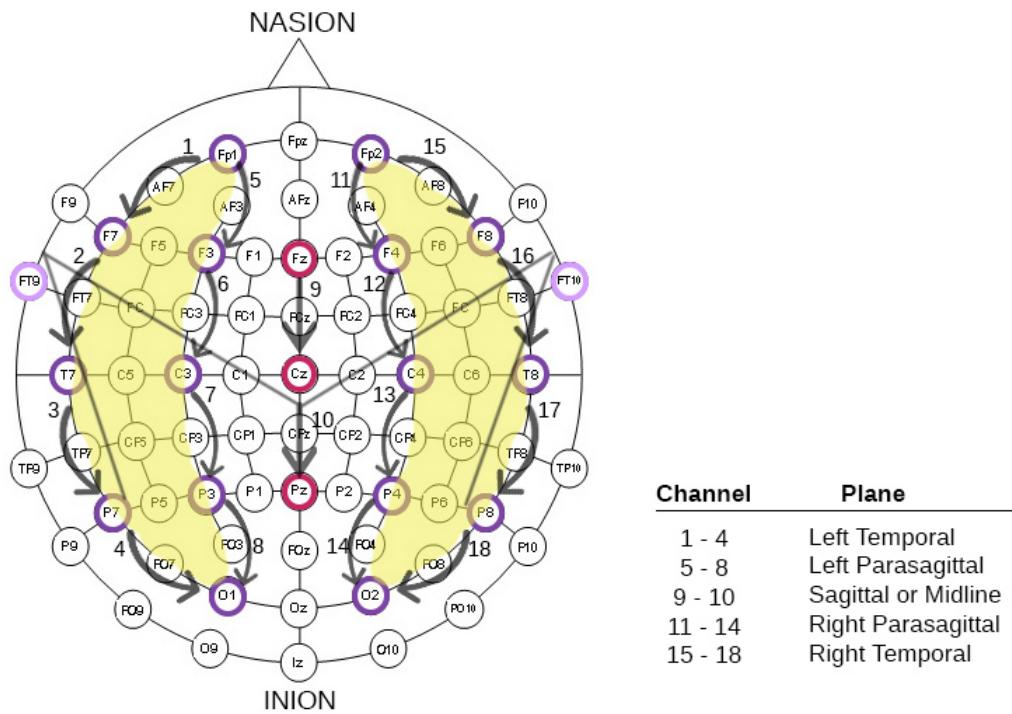


Figure 3.3: Left temporal, left parasagittal, right parasagittal, right temporal.

allow quick side-to-side comparisons to draw appropriate results. In the above Table 3.2 are described in detail the channels locations.

Table 3.2: Channels details according to Bipolar Montage "Double Banana"

<b>Channel Number</b>	<b>Differential Electrodes</b>	<b>Longitudinal bipolar montage double banana</b>
1	FP1-F7	
2	F7-T7	Left Temporal
3	T7-P7	
4	P7-O1	
5	FP1-F3	
6	F3-C3	Left Parasagittal
7	C3-P3	
8	P3-O1	
9	FP2-F4	
10	F4-C4	Right Parasagittal
11	C4-P4	
12	P4-O2	
13	FP2-F8	
14	F8-T8	Right Temporal
15	T8-P8	
16	P8-O2	
17	FZ-CZ	Sagittal or Midline
18	CZ-PZ	
19	P7-T7	
20	T7-FT9	Mid-Sagittal
21	FT9-FT10	
22	FT10-T8	
23	T8-P8	

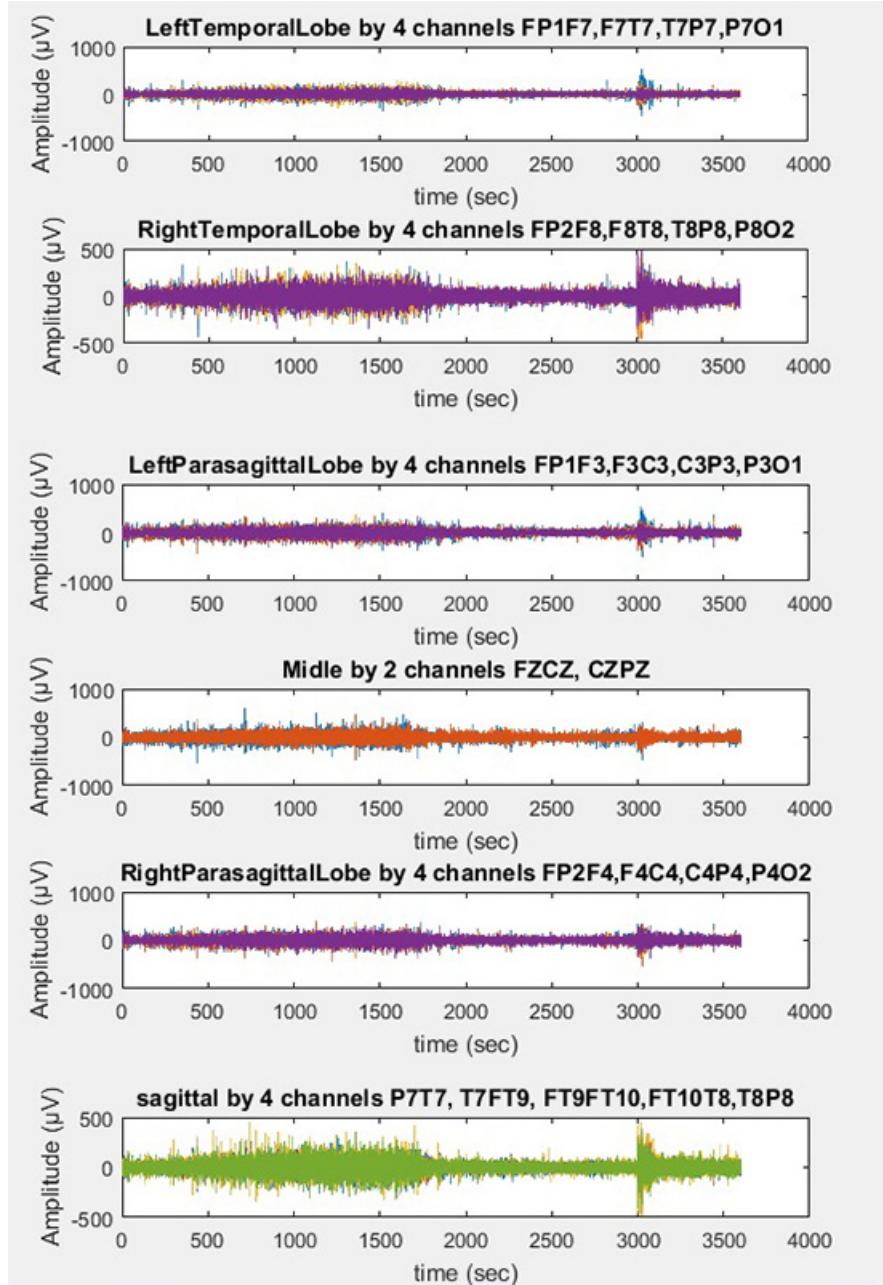


Figure 3.4: Channels according to Bipolar Montage "Double Banana". Example of a seizure within the scalp EEG of Patient A. The seizure, which begins at 2996 seconds and ends at 3036 seconds. According to the graph, it is observed that the seizure comes from the Right Temporal Lobe.

### 3.2 Preprocess and data cleansing

Correct preparation of epileptic data, before they enter the imaging sequence, is not only critical, it is the cornerstone on which the rest of processing sequence will be based. In biomedical field use the signal processing techniques for better interpretation of the measurements resulting from the biological/physiological signals.

### 3.2.1 Data sampling and filtering

Firstly, we resample the EEG recordings **to 128 Hz** instead to 512 Hz for convenience and lower volume of data processed. Additional to resample the data are to save memory, to improve a method's robustness.

Furthermore, from the data we removed the frequencies 0 - 4 Hz (delta band) owing to noise in most channels, as well as all frequencies above the 45 Hz using a low-pass filter. Lowpass filters remove the airborne and omnipresent 45 Hz hum from electrical channels and other background noise that interferes with the reading of micro voltages on the scalp.

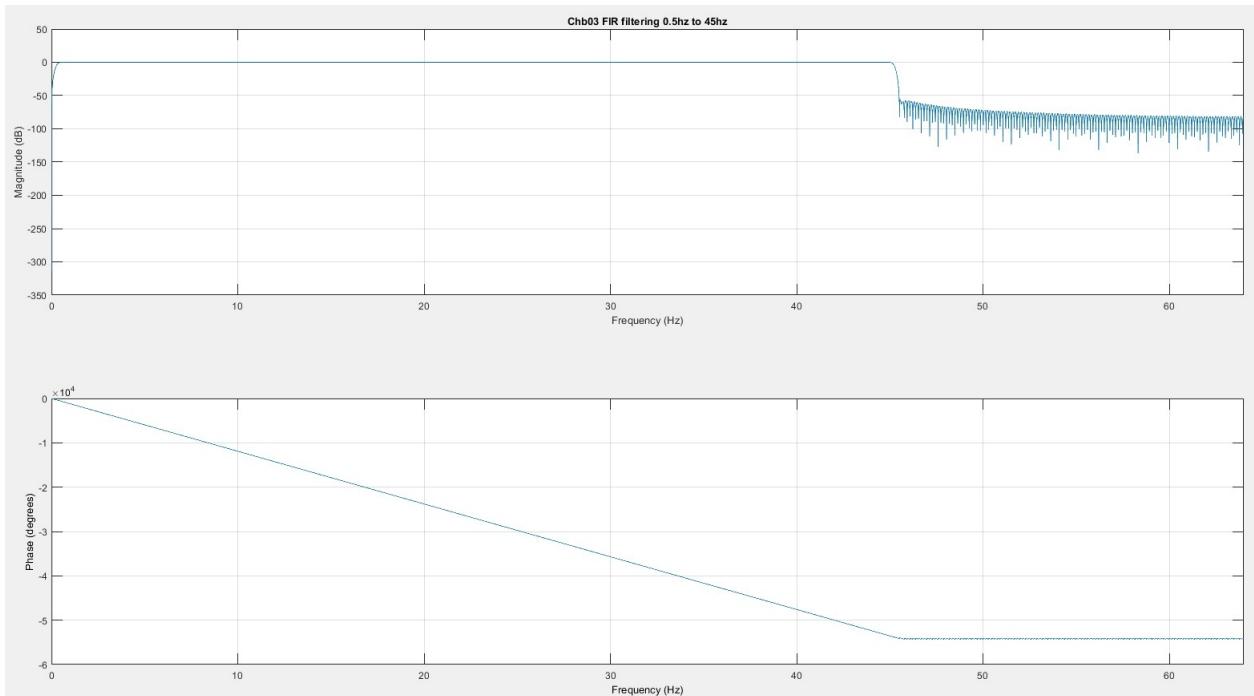


Figure 3.5: An example of filtering in subject chb03 using a FIR low pass filter butterworth

### 3.2.2 Missing values and outliers

The common practice for handling outliers in a dataset is to represent them in a way, which makes them usable for the learning model, causing the least structural damage to the dataset. At the same time its important to take extreme care when replacing outliers with numerical values in order to avoid introducing bias.

Artifacts in EEG recordings are forms of outliers and are considered as disturbances in a measured brain-signal that do not originate from the brain. The different sources of artifacts are classified into

external and internal categories. External artifacts often result from unsatisfactory technology such as exceeding measurement range of signals and disconnection of the electrode box. Internal artifacts arise from body activities due to movements or bioelectrical potentials. As a result, the potential between electrodes changes from effects such as eye movement or muscular activity, can cause an artifact.

The artifacts are signals with no cerebral origin. Instead they stem from ocular, muscular and mechanical activities. The figures below show the nature of the artifacts in the EEG signal.

#### Eye Blinks (ocular)

Vertical eye movements typically are observed with blinks. A blink causes the positive pole (ie, cornea) to move closer to frontopolar (Fp1-Fp2) electrodes, producing symmetric downward deflections. During downward eye movement the positive pole (ie, cornea) of the globe moves away from frontopolar electrodes, producing an upward deflection best recorded in channels 1 and 5 in the bipolar longitudinal montage. Eye blinks make a slow signal (<4Hz) corresponding to mechanical movement of the eyes.

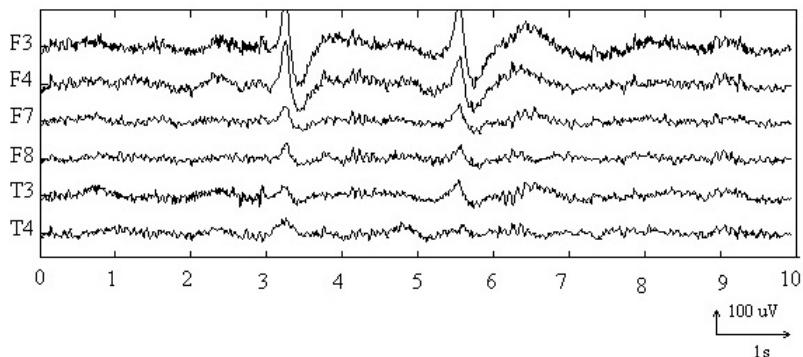


Figure 3.6: The signal appears mainly on frontal area (Fp1 and Fp2). The signal is symmetric between two hemispheres.

#### Eye movement

Lateral eye movements affect mostly lateral frontal electrodes F7 and F8 (see images below). During a left lateral eye movement the positive pole of the globe moves toward F7 and away from F8. As we see in the figure below there are differences between the signals. Channels near the frontal lobe have high fluctuations which means that there is noise from the eye movement.

The eyes make a dipole, which go closer or away from some electrodes and create the signal. The signal appears mainly at the frontal and temporal area. It is more propagated than blinks. This signal is anti-symmetric between two hemispheres.

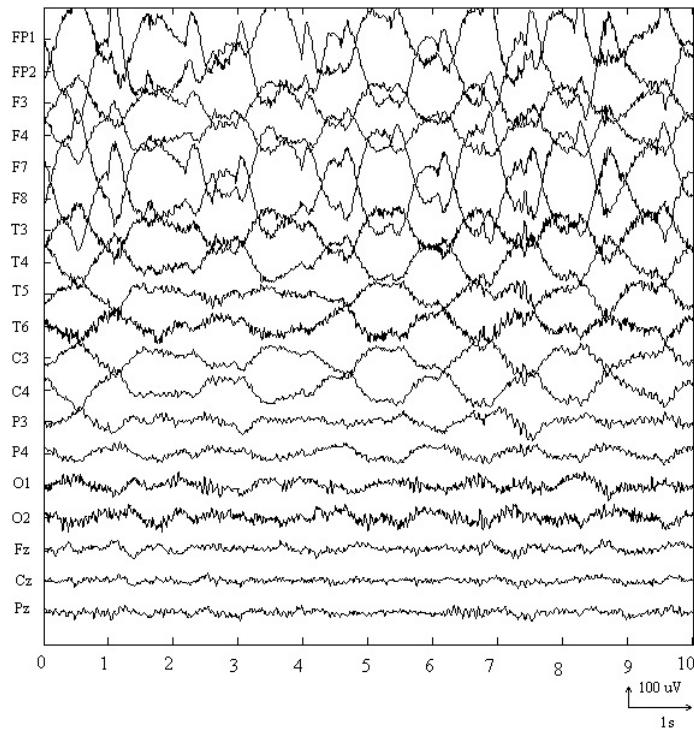


Figure 3.7: Eye movement make as well a slow signal ( $<4\text{Hz}$ ) corresponding to mechanical movement.

### Muscular artifacts

Muscular activity give high frequency signals ( $>13\text{Hz}$ ), often much higher than cerebral signals. The main head muscle is the jaw which can create an important signal in temporal area (seconds 0 to 5). Frontal muscles can appear as well since they are located just under the electrodes (seconds 5 to 10).

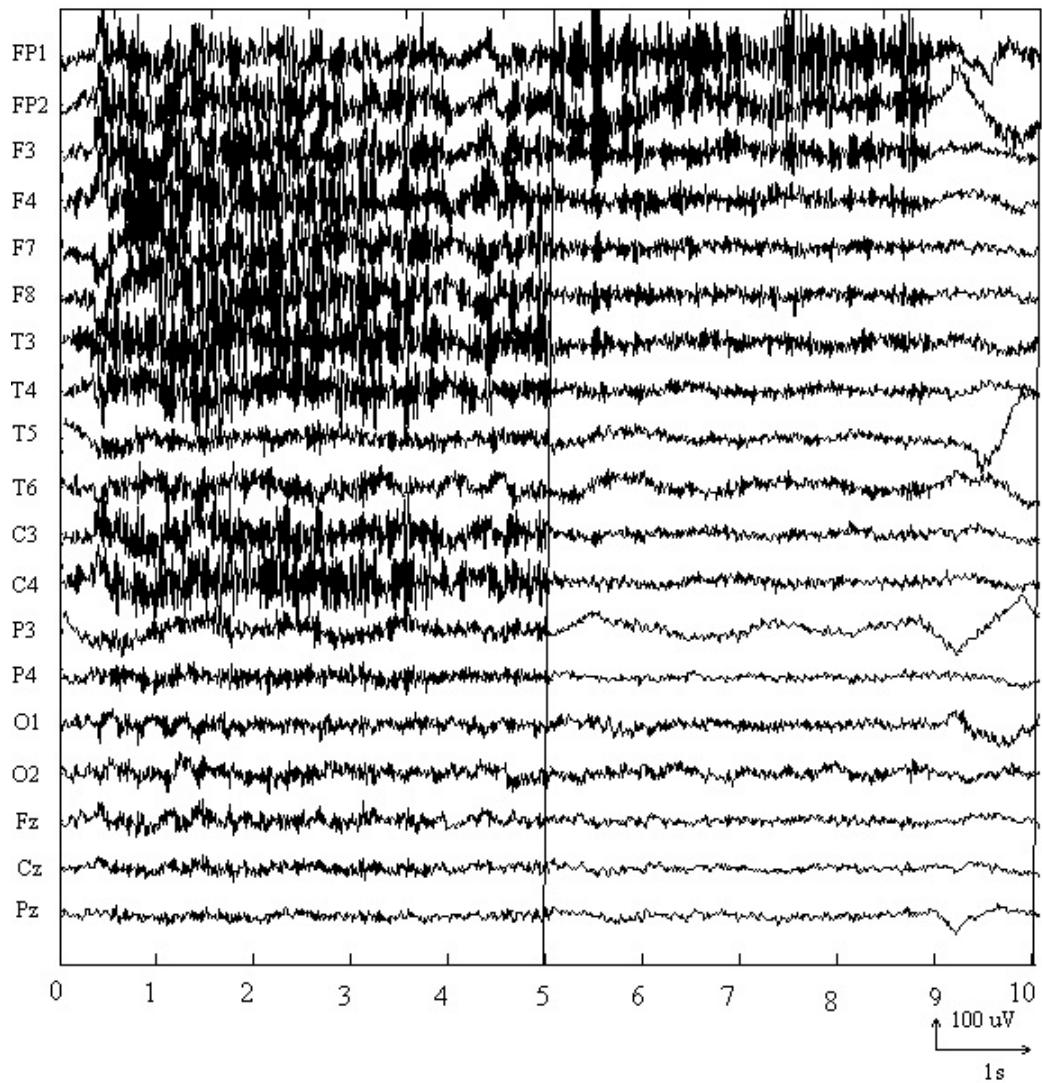


Figure 3.8: Muscular artifacts. Muscular activity give high frequency signals ( $>13\text{Hz}$ ), often much higher than cerebral signals.

### Electrode artifacts

Sometimes, wire or electrode movements create a low frequency artifact (<2Hz) on one electrode. The signal has often high amplitude. It is also possible that a mechanical artifact appears following to a head movement and on this case a signal appear on several electrodes.

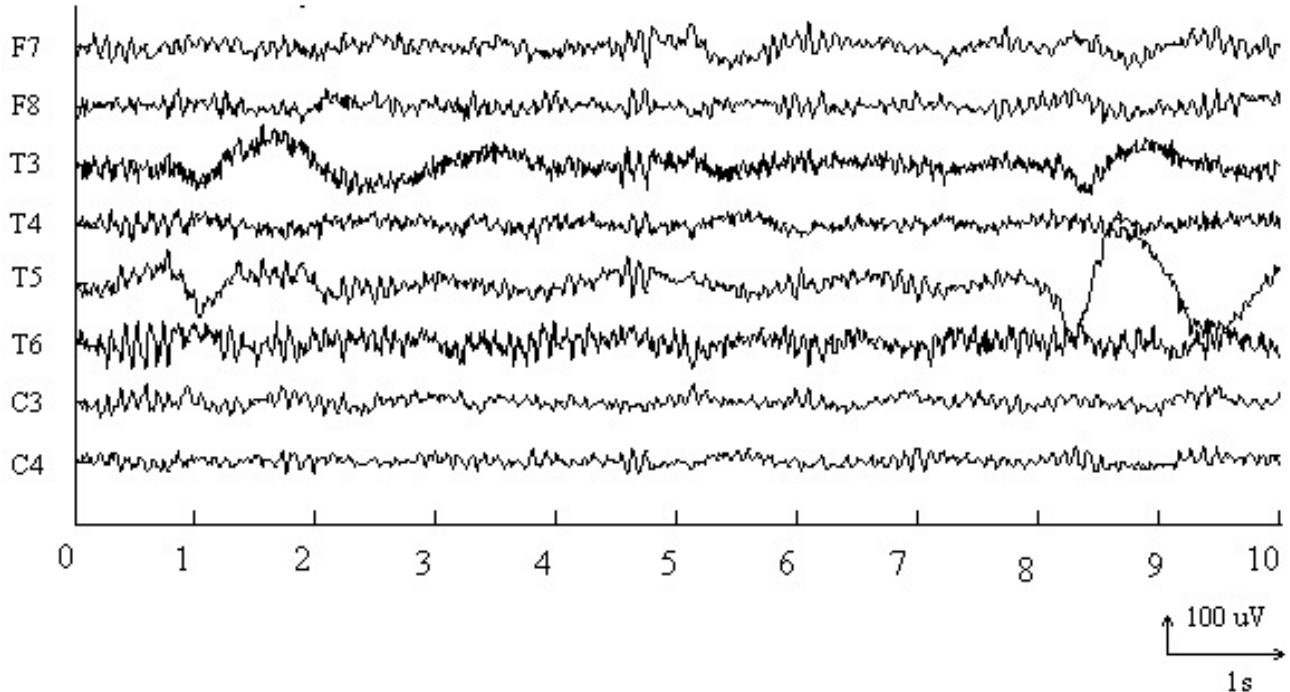


Figure 3.9: Electrode artifacts. Sometimes, wire or electrode movements create a low frequency artifact (<2Hz) on one electrode.

### **3.3 Ictal file extraction**

The EEG data for each patient in CHB-MIT scalp database is organised into separate .EDF files, each file containing 1 hour long recordings of a single channel as mentioned in 3.1.1. Each file contains a combination of ictal and non-ictal data. The bellow EEG graph shows the states between pre-ictal (yellow), ictal (red) and post-ictal (green) during one hour of monitoring.

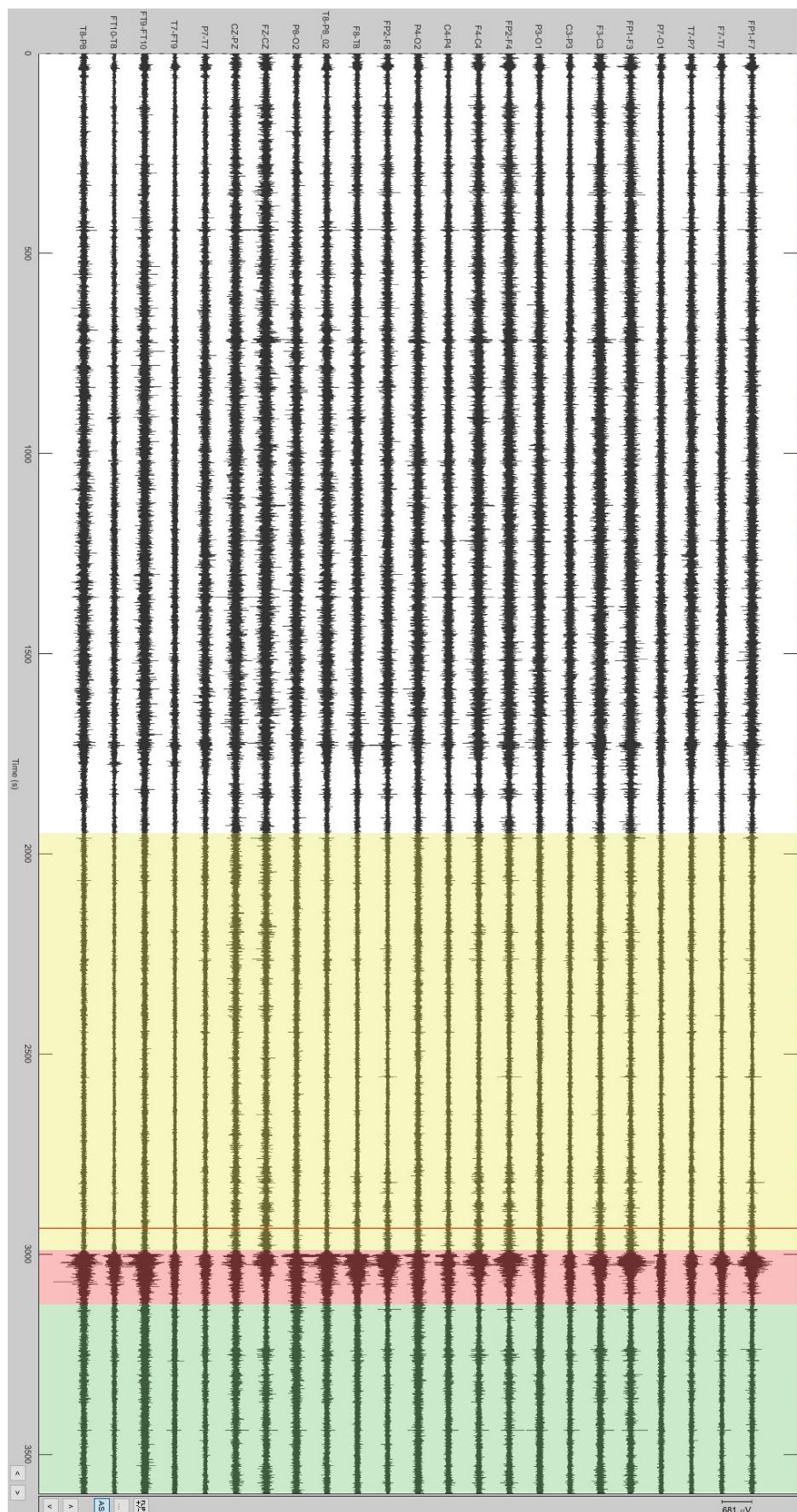


Figure 3.10: EEG recording for subject Chb01 with epileptic seizure, Seizure Start Time: 2996 seconds Seizure End Time: 3036 seconds. Pre-ictal (yellow), ictal (red) and post-ictal (green)

### 3.4 Remove Artifacts

As we can see in the above figures, the noises of artifacts affect the EEG signals and the structure of the waveform [45]. EEG is a complicated and non-stationary signal and its characteristics are spatio-temporally dependent. Based on these properties, Discrete Wavelet Transform (DWT) is chosen in this work for pre-analyzing the epileptic EEG signal. The noises are removed by the DWT, which decomposes the full-band signal into sub-band signals. The process of the transformation is as follows:

- The EEG signal is processed with the deubechies wavelet (Wavelet Families) which is used to remove the noises and after that decompose the signal into sub-bands.
- Based on the frequency range the sub-band signals are classified as delta, theta, alpha, beta, and gamma.
- After the decomposition, the noises are reduced then the Error rate is estimated.

Also, the eye blinks and movements detected and rejected with Brainstorm toolbox for Matlab.

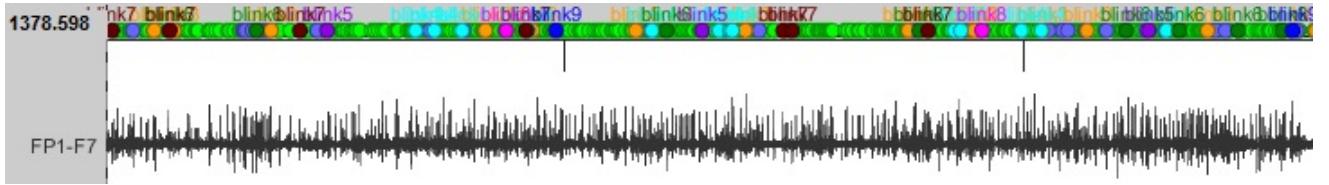


Figure 3.11: An example of eye blink artifacts in channel Fp1-F7 in 10 mins (Brainstorm toolbox)

#### 3.4.1 Discrete Wavelet Transform

Discrete wavelet transform can be used for easy and fast denoising of a noisy signal. If we take only a limited number of highest coefficients of the discrete wavelet transform spectrum, and we perform an inverse transform on the same wavelet basis, we can obtain more or less denoised signal. There are several ways how to choose the coefficients that will be kept [46]. The wavelet transform is similar to the Fourier transform (or much more to the windowed Fourier transform) with a quite different merit function [47]. The central difference is this: Fourier transform decomposes the signal into sines and cosines, i.e. the functions localized in Fourier space; in antithesis, the wavelet transform uses functions that are surrounded in both the real and Fourier space. The following equation can represent the wavelet transform:

$$F(a, b) = \int_{-\infty}^{\infty} f(x)\psi_{a,b}^*(x)dx$$

where the \* is the complex conjugate symbol and function  $\psi$  is some function. This function can be chosen arbitrarily provided that it obeys certain rules.

The Wavelet transform is an infinite set of various transforms, depending on the merit function used for its computation. Here we show only the division based on the wavelet orthogonality. We can use orthogonal wavelets for discrete wavelet transform development and non-orthogonal wavelets for continuous wavelet transform development. These two transforms have the following properties:

- The discrete wavelet transform returns a data vector of the same length as the input is. Usually, even in this vector, many data are almost zero. This corresponds to the fact that it decomposes into a set of wavelets (functions) that are orthogonal to its translations and scaling. Therefore, we decompose such a signal to a same or lower number of the wavelet coefficient spectrum as is the number of signal data points. Such a wavelet spectrum is very good for signal processing and compression, for example, as we get no redundant information here.
- The continuous wavelet transform in contrary returns an array one dimension larger than the input data. For a 1D data we obtain an image of the time-frequency plane. We can easily see the signal frequencies evolution during the duration of the signal and compare the spectrum with other signals spectra. As here is used the non-orthogonal set of wavelets, data are highly correlated, so big redundancy is seen here. This helps to see the results in a more humane form.

The discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translations obeying some defined rules. In other words, this transform decomposes the signal into mutually orthogonal set of wavelets, which is the main difference from the continuous wavelet transform (CWT), or its implementation for the discrete time series sometimes called discrete-time continuous wavelet transform (DT-CWT).

The wavelet can be constructed from a scaling function which describes its scaling properties. The restriction that the scaling functions must be orthogonal to its discrete translations implies some mathematical conditions on them which are mentioned everywhere, e.g. the dilation equation can be applied to the EEG signal as follow.

$$\phi(x) = \sum_{\kappa=-\infty}^{\infty} a_{\kappa}\phi(S_x - \kappa)$$

where S is a scaling factor (usually chosen as 2). Moreover, the area between the function must be normalized and scaling function must be orthogonal to its integer translations, i.e.

$$\int_{-\infty}^{\infty} \phi(x)\phi(x+l)dx = \delta_{0,l}$$

The decomposition of the signal into different frequency bands is simply obtained by successive high pass and low pass filtering of the time domain signal. The original signal  $x[n]$  is first passed through a half band high pass filter  $g[n]$  and a low pass filter  $h[n]$ . After filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of  $p/2$  radians instead of  $p$ . The signal can therefore be down sampled by 2, simply by discarding every other sample. This constitutes one level of decomposition and can mathematically be expressed as follows:

$$h_{high} |\kappa| = \sum_n x[n] g[2\kappa - n]$$

$$h_{low} |\kappa| = \sum_n h[2\kappa - n]$$

Where  $y_{high}[k]$  and  $y_{low}[k]$  are the outputs of the high pass and low pass filters, respectively, after down sampling by 2.

This decomposition halves the time resolution since only half the number of samples now characterizes the entire signal. However, this operation doubles the frequency resolution, since the frequency band of the signal now spans only half the previous frequency band, effectively reducing the uncertainty in the frequency by half. The above procedure can be repeated for further decomposition.

The method provides an adaptive approach for optimal signal representation for the purpose of compression and can thus be applied to any one-dimensional biomedical signal. A classical approach for optimizing the basis function set is to search for the best wavelet packet decomposition.

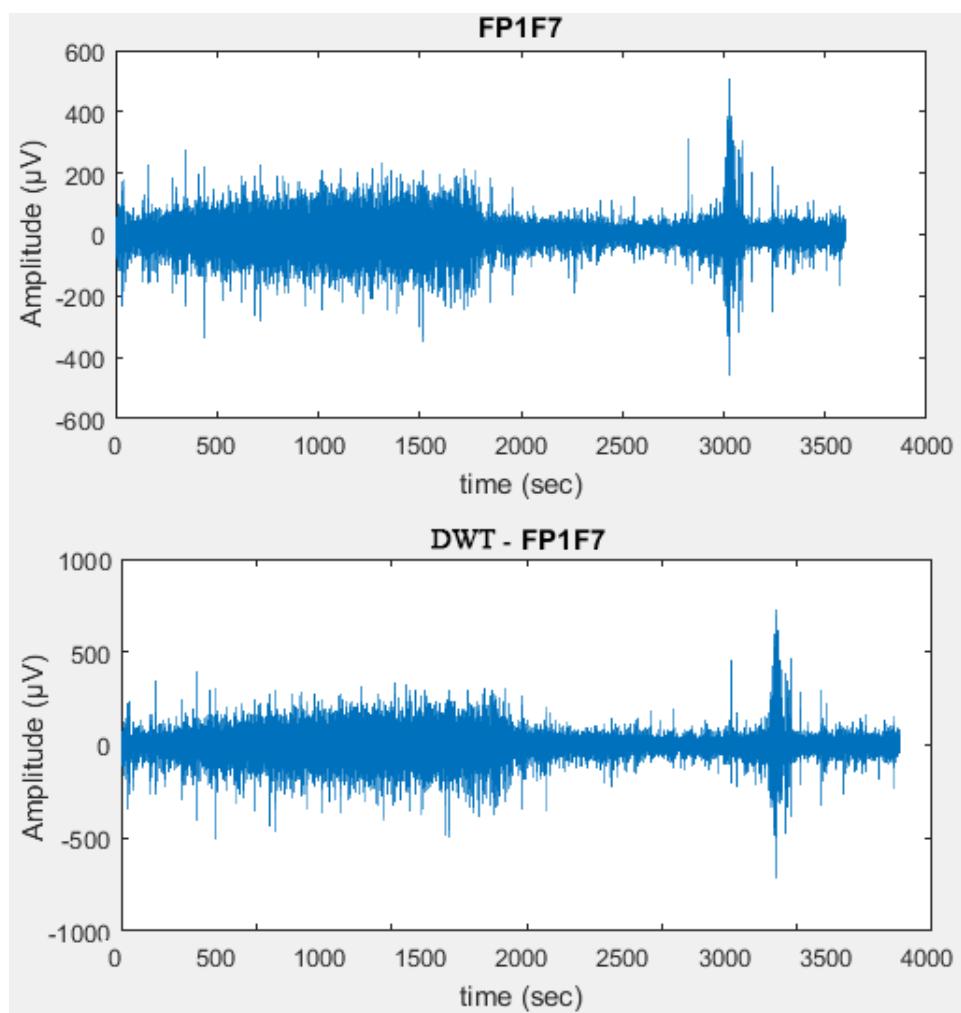


Figure 3.12: Discrete wavelet transform (DWT) performed in 1 hour recording of channel FP1-F7

# Chapter 4

## Feature Engineering

In this chapter become an investigation of the optimal features, selects the right attributes to analyze. The domain knowledge of the data to select or create attributes that make machine learning algorithms work. Following artefact removal, features were then extracted between pre-ictal and ictal epochs of each patient, in order to build datasets for training the predictive models. The input EEG signal scanned with sliding window technique and calculated features where extracted from each second, the output of the signal stage is a label for each features value. For the implementation of these stages, we used Matlab built-in functions.

### 4.1 Moving Window Analysis

Most EEG-based prediction methods use a moving-window analysis in which some linear and nonlinear characterizing measure is calculated from a window with a predefined length this technique was used in our dataset.

The duration of these analysis windows typically varies between 10 and 40 s. Depending on whether the measure is used to characterize a single EEG channel or associations between two or more channels, it is introduced to as a univariate, bivariate or multivariate measure, respectively. The moving-window analysis hence renders time profiles of a characterizing measure for different channels or channel combinations. Also, when the time profile of some characterizing measure rises above a predefined threshold, the prediction algorithm issues an alarm that is classified as true or false depending on whether it is followed by a seizure within a specified warning time termed the prediction horizon.

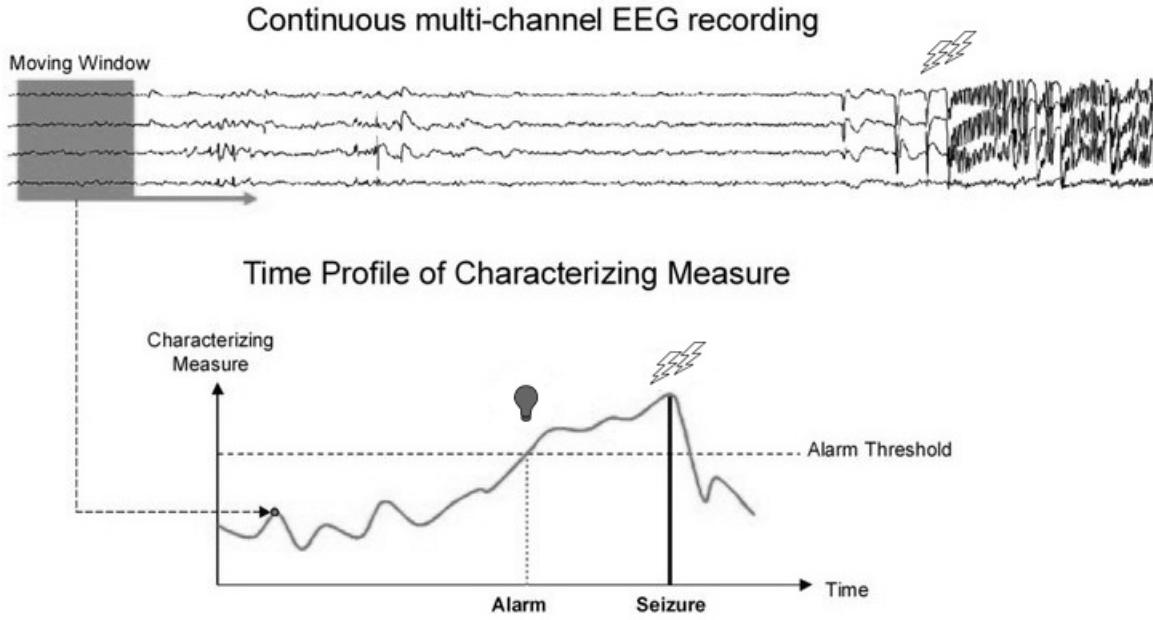


Figure 4.1: Seizure Prediction from EEG time series. Continuous multi-channel EEG recordings are analyzed by means of a moving-window analysis. The data covered by the orange window is transformed into a single value in the time profile of a multivariate characterizing measure. When this time profile crosses a certain pre-defined threshold, an alarm is issued [16].

## 4.2 Feature selection

"Features" are the values which define some relevant properties of the acquired signals. Feature extraction is the most important aspect of any classification process. The extracted feature contains the discriminatory information hence the accuracy of feature extraction affects the accuracy of the classification. Feature extraction can be done in both time and frequency domain.

Feature extraction for classification is to seek a transformation or mapping from original features to a new feature space which can maximize the separability of different classes. Feature extraction for classification is to investigate a transformation or mapping from original features to a new feature space which can maximize the separability of different classes. A classification problem cannot be properly solved if important interactions and relationships between the original features are not taken into consideration. Thus, many types of research agreed that the feature extraction is the most important key to any pattern recognition and classification problem. In most cases, feature extraction is done by a human based on the researcher's knowledge, experience, and/or intuition.

Some of feature for epilepsy prediction are: Statistical Measures, Hjorth Parameters, Accumulated Energy, Autoregressive Models, SVM, Kalman Filtering, Lyapunov Exponent, Dynamical Similarity Index, Correlation Dimension, Entropies, Phase Synchronization. In the present thesis we used some

of the above characteristics (Hjorth Parameters, Entropies, Phase Synchronization-Coherence).

### 4.3 Brain Wave Patterns

Epilepsy leaves its signature in the EEG signals. The detection of seizures occurring in the EEGs is an important component in the diagnosis and treatment of epilepsy. The separation of two important categories of abnormal activity can be observed in an EEG signal: ictal (during an epileptic seizure) and inter-ictal (between seizures). Often, the onset of a clinical seizure is characterized by a sudden change of frequency in the EEG measurement. It is normally within the alpha wave frequency band with slow reduction in frequency but increase in amplitude during the seizure period. It may or may not be spiky in shape [48]. For assisting the diagnosis and treatment of epilepsy or neurological disease, this dissertation aims to develop methods that can identify the epileptic EEG signals during seizure activity and also during seizure-free time.

It is important to know that all humans display five different types of electrical patterns or brain waves over the cortex. Each brain wave has a purpose and helps serve us in optimal mental functioning. If one of the five types of brain waves is either overproduced or underproduced in our brain, it can cause problems [49]. For this reason, it is important to understand that no single brain wave is better or more optimal than the others. Below is described in details the five brain waves in order of highest frequency to lowest are as follows: *Gamma, Beta, Alpha, Theta, and Delta*.

Table 4.1: Brain Waves (Gamma, Beta, Alpha, Theta, and Delta)

	<b>Gamma</b>	<b>Beta</b>	<b>Alpha</b>	<b>Theta</b>	<b>Delta</b>
<b>Frequency range (Hz)</b>	40 - 100	12 - 40	8 - 12	4 - 8	0 - 4
<b>Too much</b>	Anxiety, high arousal, stress	Adrenaline, anxiety, high arousal, inability to relax, stress	Daydreaming, inability to focus, too relaxed	Depression, hyperactivity, impulsivity, inattentiveness	Brain injuries, learning problems, inability to think
<b>Too little</b>	Depression, learning disabilities	Daydreaming, depression, poor cognition	Anxiety, high stress, insomnia, OCD	Anxiety, poor emotional awareness, stress	Inability to rejuvenate body, inability to revitalize the brain, poor sleep
<b>Optimal</b>	Binding senses, cognition, information processing, learning, perception, REM sleep	Conscious focus, memory, problem solving	Relaxation	Creativity, emotional connection, intuition, relaxation	Immune system, natural healing, restorative / deep sleep
<b>Increase Waves</b>	Meditation	Coffee, energy drinks, various stimulants	Alcohol, marijuana, relaxants, some antidepressants	Depressants	Depressants, sleep

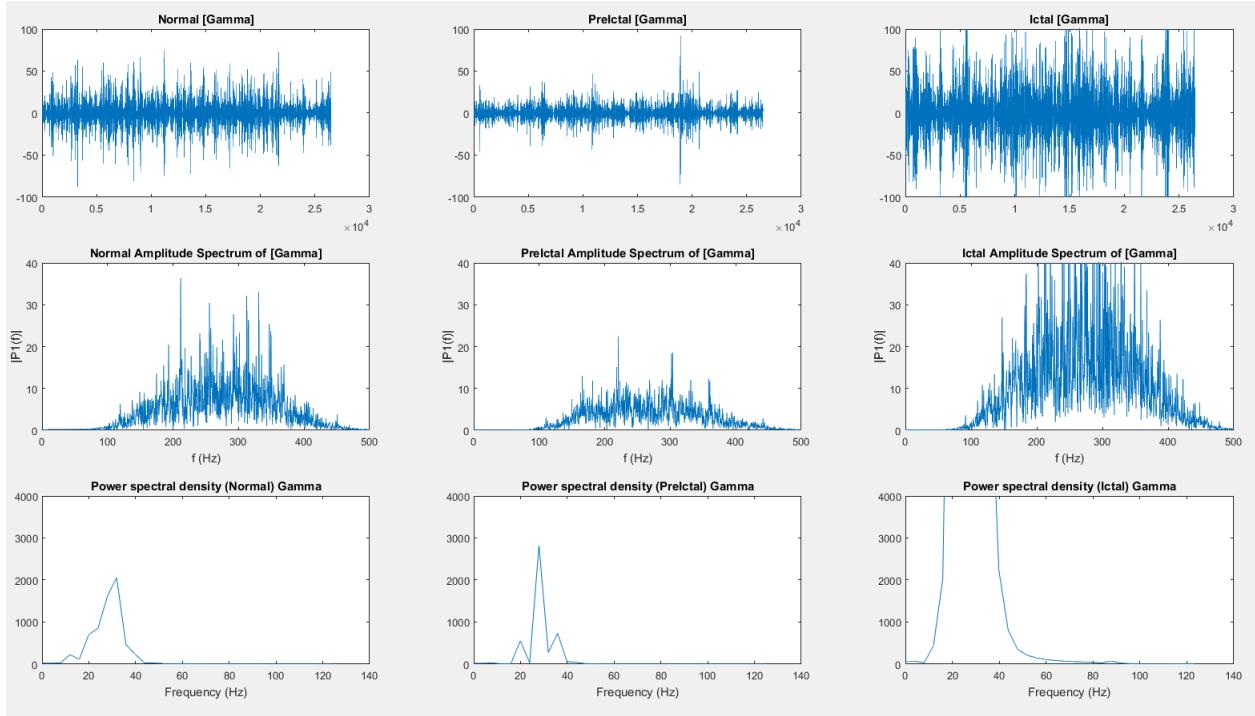


Figure 4.2: Gamma waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

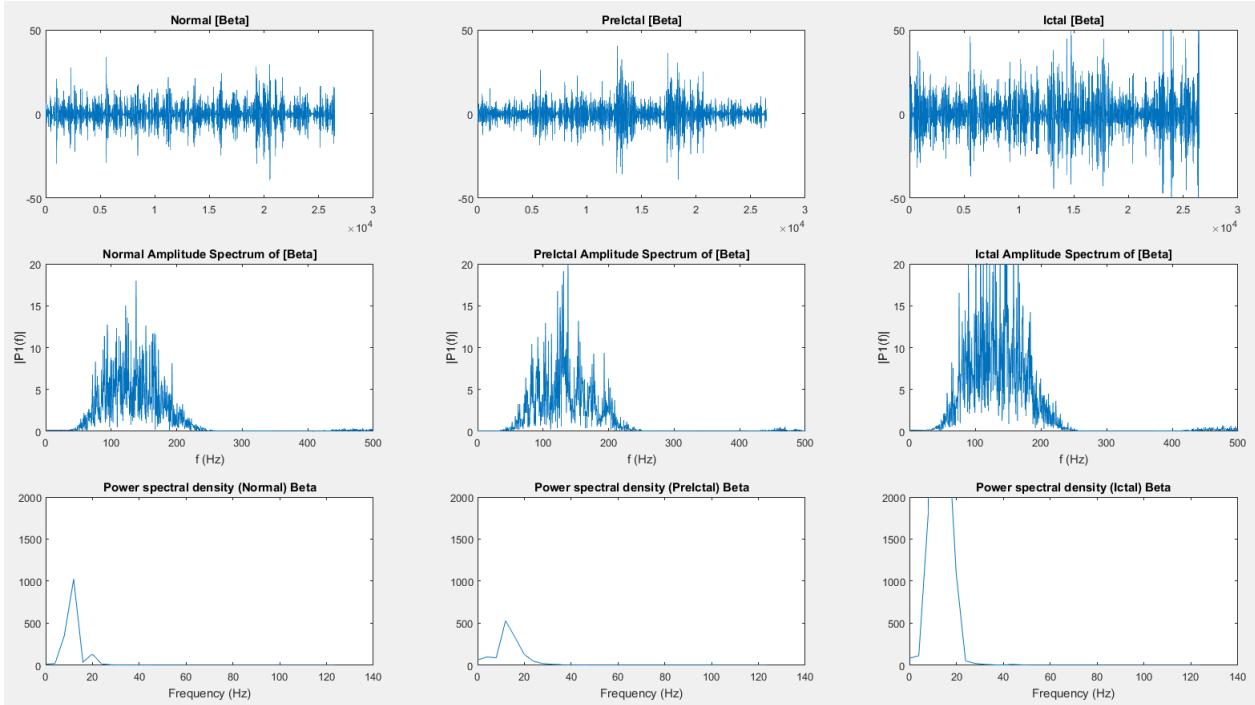


Figure 4.3: Beta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd image)

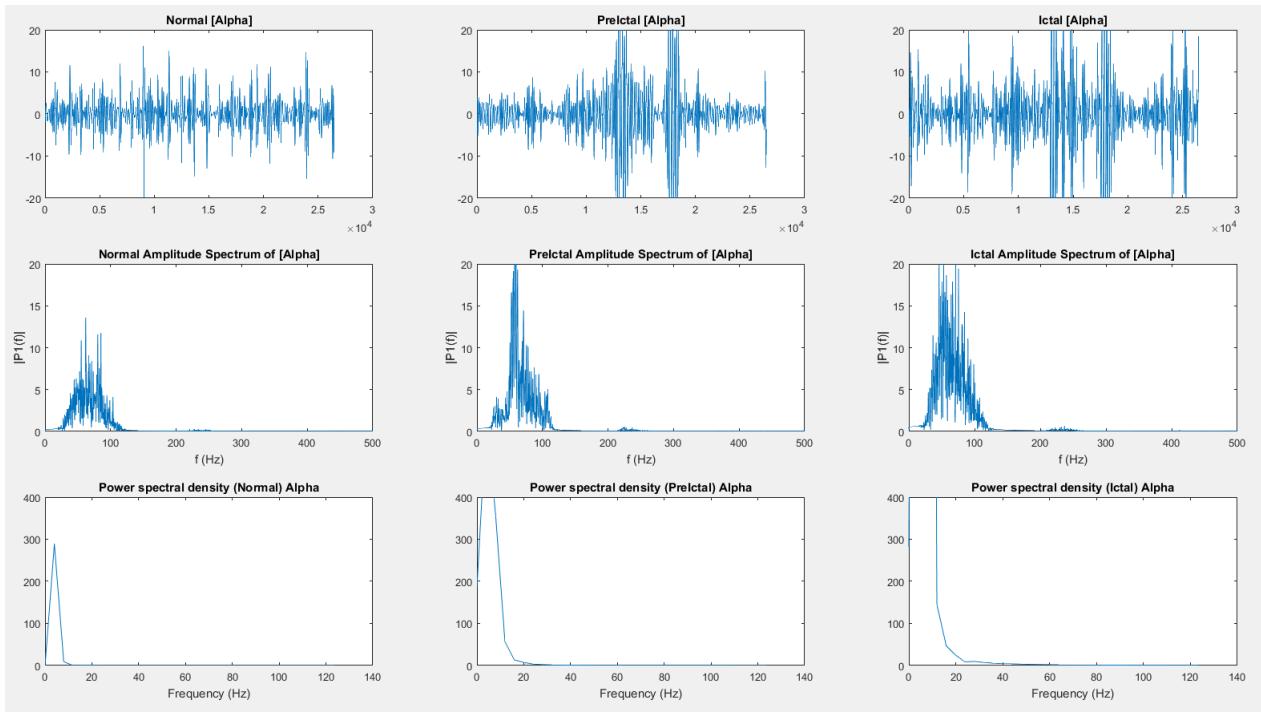


Figure 4.4: Alpha waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

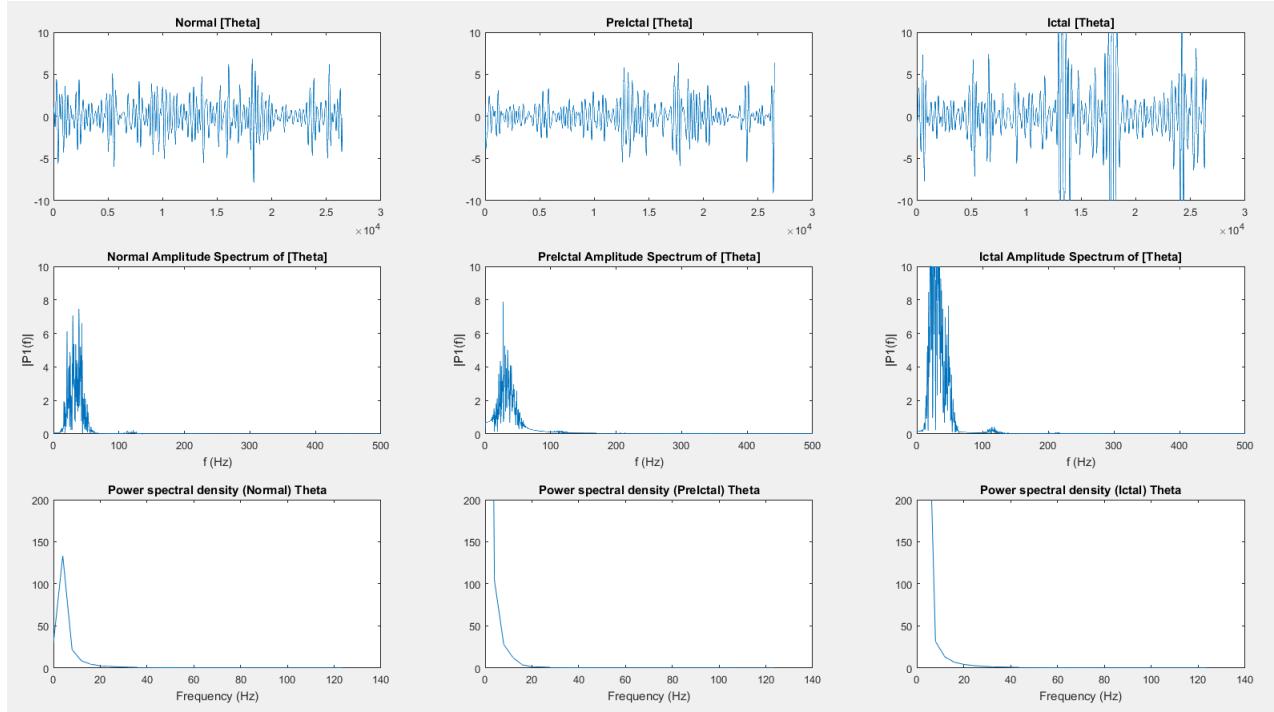


Figure 4.5: Theta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

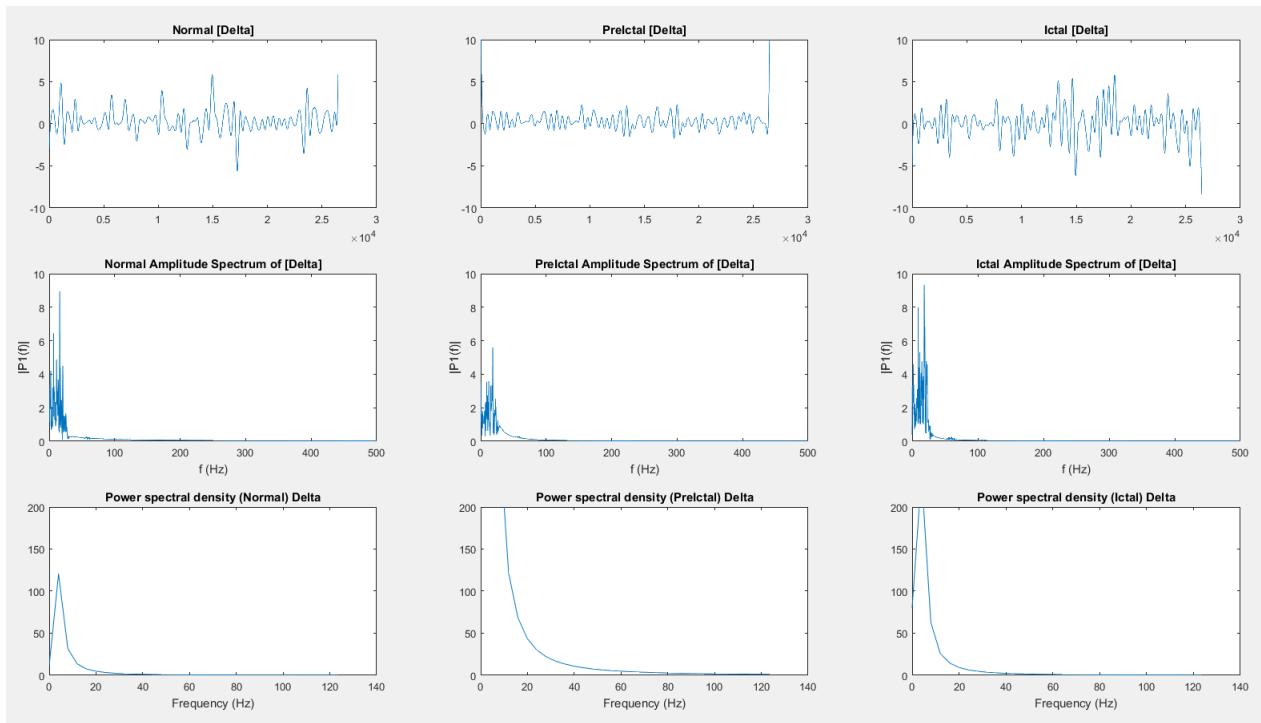


Figure 4.6: Delta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

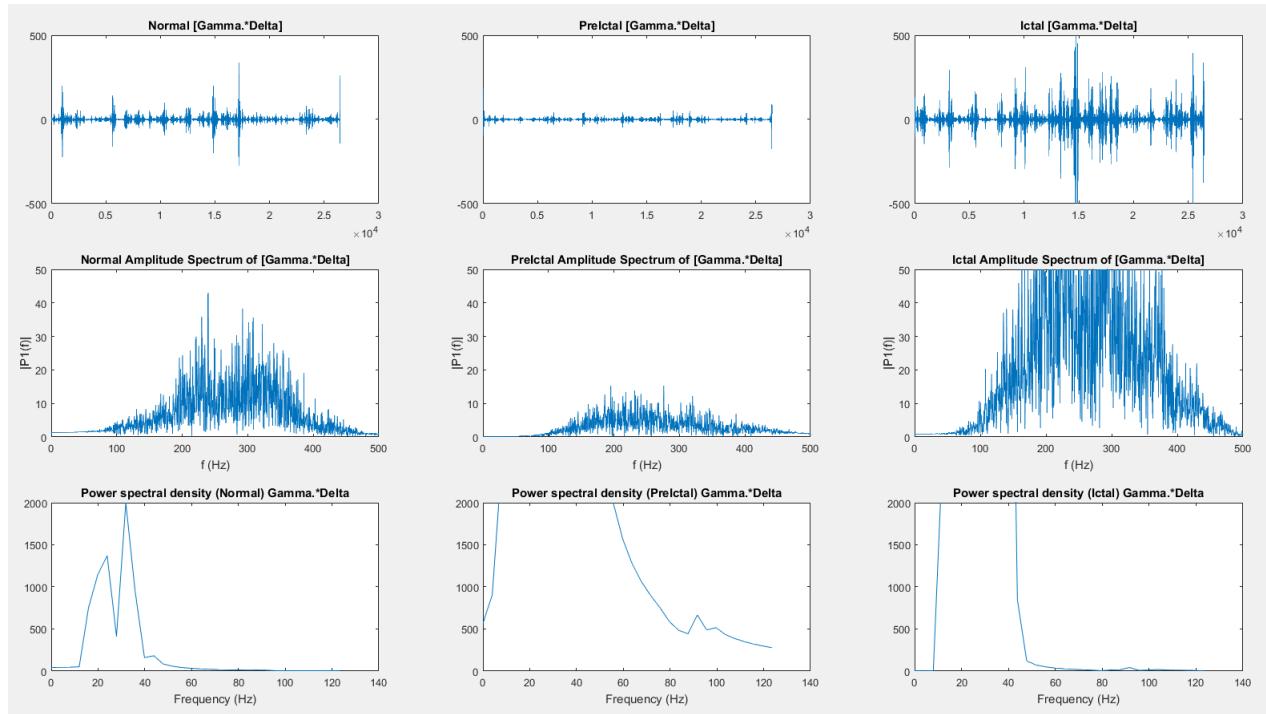


Figure 4.7: The inner product between Delta and gamma waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

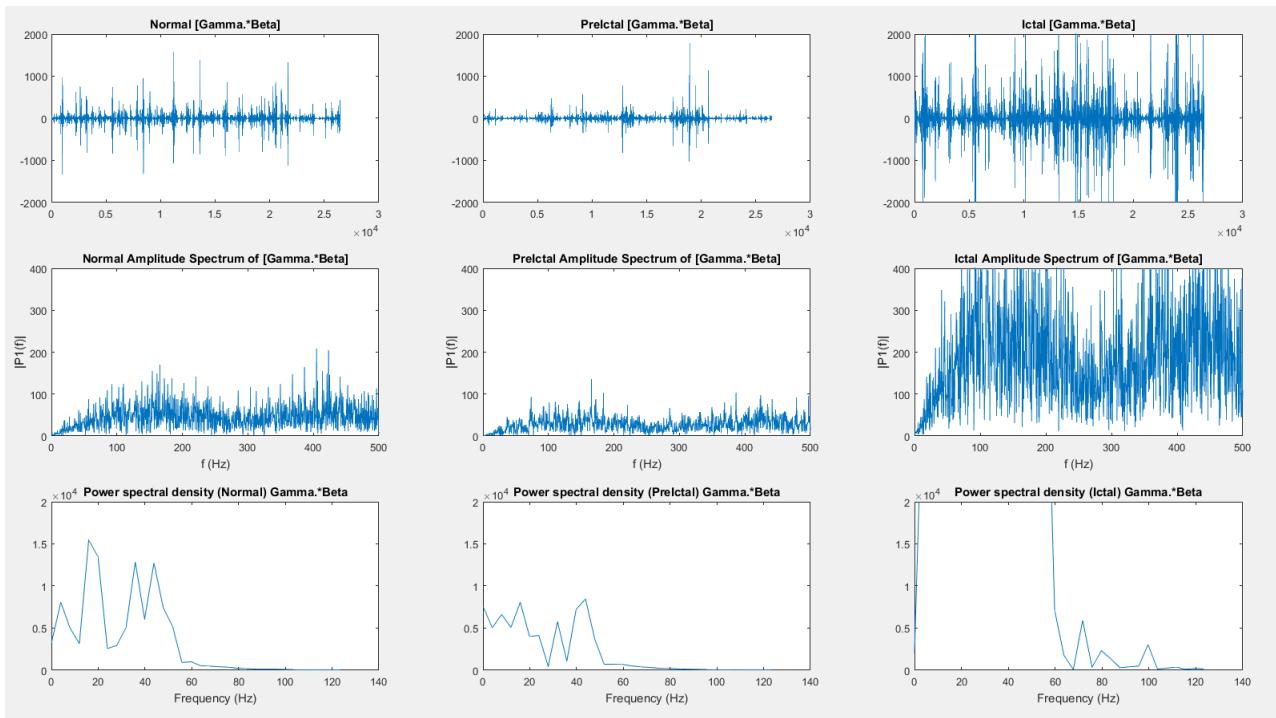


Figure 4.8: The inner product between Gamma and Beta waveband during normal state (1st column), pre ictal state (2nd column) and seizure state (3rd column)

- **Gamma waves ( $\gamma$ )** These are involved in higher processing tasks as well as cognitive functioning. Gamma waves are important for learning, memory and information processing. It is thought that the 40 Hz gamma wave is important for the binding of our senses in regards to perception and are involved in learning new material. It has been found that individuals who are mentally challenged and have learning disabilities tend to have lower gamma activity than average.
- **Beta waves( $\beta$ )**, ranging from 13 to 30 Hz, are associated to deep thinking, high concentration level and anxious state. They have large frequency band compared with others. Beta waves originate from central area of the brain and front side of head. Amp <20 mV and lower in elderly.
- **Alpha waves ( $\alpha$ )**, are in the frequency range from 7.5 Hz to 12 Hz. These types of waves originate from occipital lobe and backside of the head. Alpha waves dominate in relaxed and calm mental states while being awake. They have higher amplitude compared with other waves. Amp 30-60 mV, Right hemisphere >Left hemisphere by 20-50%, drowsiness, Age 60-80, a=9.5 Hz; a <8 in elderly suggests. Slow (sub-harmonic, 4-5 Hz) and fast alpha (16-20 Hz).
- **Theta waves ( $\theta$ )**, are in the frequency range from 3.5 to 7.5 Hz. They originate from central, temporal and parietal parts of head. High level of theta waves generally occurs in abnormal adults, usually one with AD/HD. They arise during creative thinking, stressed and deep meditating state.
- **Delta waves ( $\delta$ )**, are in the frequency range from 0.5 to 3.5 Hz. They are the slowest waves compared to others. Generally, occur in deep sleep and sometimes when awake. They also occur in coma mental state in awake state is considered to be a serious phenomenon.

## 4.4 Linear Methods

### 4.4.1 Hjorth Parameters

The Hjorth parameter can be used as a good feature in real-time EEG applications. Hjorth parameters define the temporal dynamics of a signal  $X(t)$ , by using three measures that are the mobility, the activity, and the complexity, indicating the statistical feature of a continuous EEG signal in time domain.

The first Hjorth parameter is a measure represents the variance of signals amplitude, (variance of the signal) the value of Activity returns a large/small value if the high-frequency components of the signal exist many/few [50]. Mathematically defined as:

$$Activity(y) = \sum_{i=1}^{N_s} \frac{(y(i)-a)^2}{N_s}$$

The mobility parameter estimates the mean frequency, it represents by square root of the ratio between the variances of the first derivative  $y'(t)$  and the amplitude (variance of the signal)  $y(t)$ .

$$Mobility(y) = \sqrt{\frac{var(y')}{var(y)}}$$

The third Hjorth parameter, called complexity, is an estimate of the bandwidth of the signal and is defined as:

$$Complexity(y) = \frac{Mobility(y')}{Mobility(y)}$$

For the above functions,  $y$  is the signal,  $y'$  is the derivative of the signal,  $N_s$  is the number of samples in the window, and  $a$  is the mean of the signal within the window. The Hjorth parameters have been utilized as EEG features by many researchers working in the field of EEG-based BCI systems, Mormann et al. used Hjorth parameters among others as features for seizure prediction. Mobility has also been used followed by SVM classification was achieving better false positive rates (fpr) in comparison with plain spectral analysis.

To examine the Hjorth parameters, we need to separate the EEG signal into epochs. The epochs which selected are divided into three primary categories, the "Inter-Ictal" of 20 mins period (several hours prior to the seizure), the "Pre-Ictal" of 10 mins period (shortly before the seizure) and the "Ictal" of 40 secs period (epileptic seizure).

The results show that the activity has lower values shortly before the crisis, in contrast to the prices in the inter-ictal state. Also, the complexity and mobility values are low in the pre-ictal state, as opposed to the inter-ictal state. Also, the complexity presents significant high prices during the seizure. In the Appendix D: displays more examples of different patients.

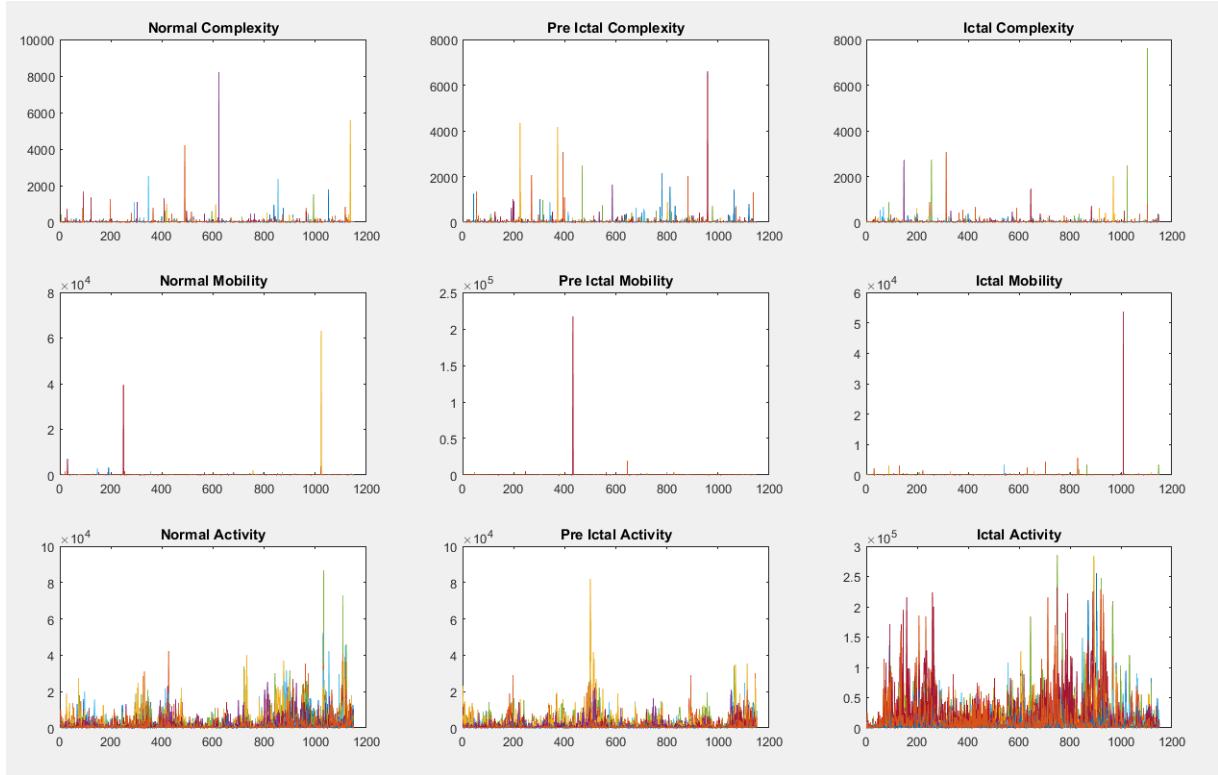


Figure 4.9: Results of Hjorth Parameters between interictal, preictal and ictal (seizure)

## 4.5 Nonlinear Methods

Most of the nonlinear methods exploit the reconstruction of a time series  $x(i)$ ,  $i = 1, 2, \dots, N$  in phase space domain forming the  $m$  dimensional time delayed vectors

$$x_m(i) = \{x(i), x(i + 1t), \dots, x(i + (m - 1)t)\}$$

where  $m$  is the embedding dimension and  $t$  is the time delay. This reconstruction conveys important information about the nonlinear dynamics of a signal and it is used to many methods some of them described below.

### 4.5.1 Coherence

Coherence analysis can detect the coordination of EEG rhythms between brain areas. Is a method developed on the base of classic coherence analysis and signals joint time-frequency representations in recent years [51]. It was used to extract transient characteristics of interactions among brain areas. Describes the temporal, spatial and frequency relationships of brain activities. In this analysis

discuss the frequency-varying coherence of EEG (Electroencephalogram) to examine the coordination mechanism of the brain. The power spectral density (PSD) (Welch method) is the frequency-varying method to examine the coordination mechanism of brain areas.

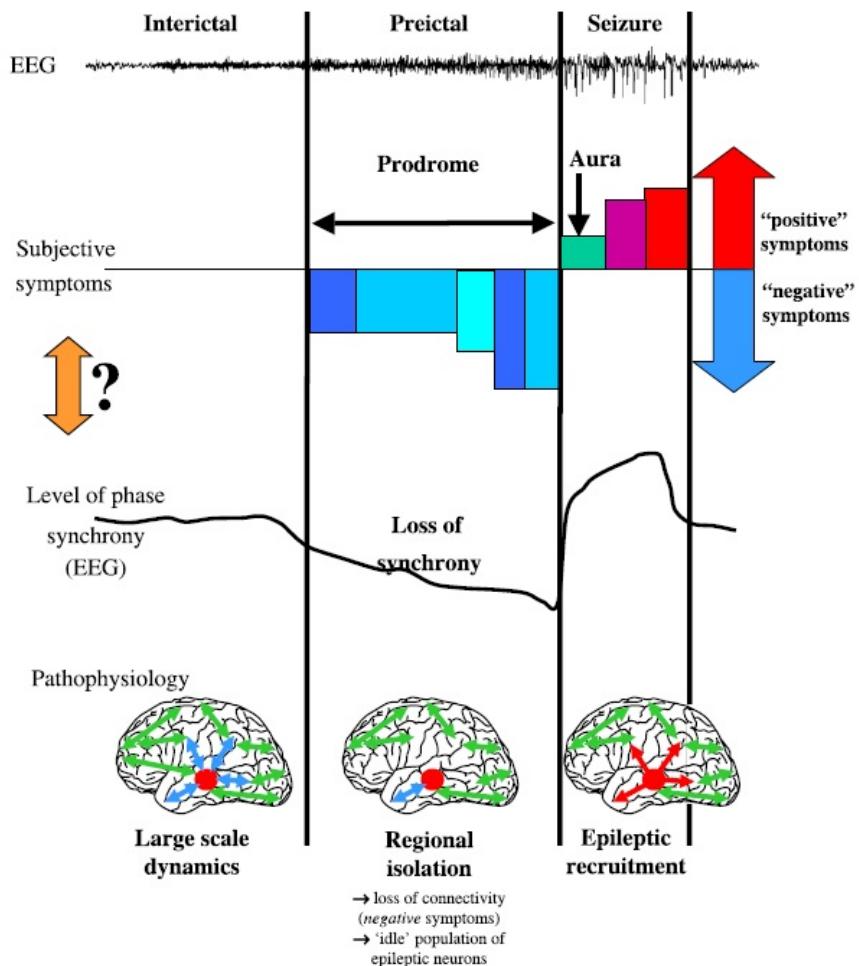


Figure 4.10: An illustration of coherence between interictal, preictal and seizure

### Power spectral density

Power spectral density (PSD) is the frequency response of a random or periodic signal. It describes to us where the average power is distributed as a function of frequency. Indicates which frequencies variations are strong and at which frequencies variations are weak. The unit of PSD is energy per frequency, can obtain energy within a particular frequency range by integrating PSD with that frequency range. The power of a signal in a given frequency band can calculate by integrating over positive and negative frequencies. The description of power spectral density generalizes in a straight manner to finite time series with  $1 \leq n \leq N$ , such as signal sampled at discrete times  $x_n = x(n\Delta t)$  for a total measurement period  $T = N\Delta t$ .

$$S(e^{jw}) = \frac{1}{2\pi N} \left| \sum_{n=1}^N x_n e^{-jwn} \right|^2$$

Equation (1), gives the equation for both methods. In a real-world application, one would typically average the single-measurement PSD on several repetitions of the measurement to obtain a more accurate estimate of theoretical PSD of the physical process underlying the proper measures.

### **Welch Method**

The power spectral density (PSD) of the input signal vector using Welch's averaged modified method of spectral estimation. The power spectral density is calculated in units of power per radians per sample.

### **Coherence Estimation**

The coherence analysis is a kind of study based on spectrum analysis, which can describe the synchronization of electric brain activities of different frequencies between brain areas [52]. But the nonstationary of EEG data made the classic coherence analysis not able to explore the dynamical properties of the brain rhythms. Hence, from classic coherence analysis and signals time-frequency representations, time-varying coherence methods had been developed and applied to certain cognitive tasks or pathology and healing researches.

The coherence analysis is based on the power spectrum estimation. The coherence between two signals  $x(t)$  and  $y(t)$  is a real-valued function that is defined as, magnitude squared value of the cross power spectrum divided by the product of power of the spectra of both signals are completely correlated. Equation (2) is an equitation for analysis of coherence of EEG signals.

$$\text{Coherence}(f) = \frac{|\text{Cross-Spectrum}(f)XY|^2}{(\text{Auto-Spectrum}(f)(X)\text{Auto-Spectrum}(f)(Y))}$$

### **Coherent / Non-coherent**

When the phase difference and amplitude ratio between two signals remain unchanged; the coefficient equals one, it means that two signals are more coherent, otherwise to zero, the two signals are non-coherent. When two areas of the brain are more synchronous to each other the result is nearer to 1 that is, they are more coherent to each other, otherwise non-coherent means closer to 0 [52].

As we said in chapter 2 before a seizure (preictal period) occur appears some prodromal symptoms which reflected in the coherence. During the seizure, observed hyper- synchronization in contrast with

the preictal period which appears a progressive decrease in synchrony between the EEG channels from the epileptogenic focus and the channels from surrounding areas. The epileptic neurons located in the epileptogenic zone, as they lose their large scale connections to brain dynamics during the preictal period, and as they may lose the inhibitory control from surrounding areas, become idle and can be recruited to build the seizure.

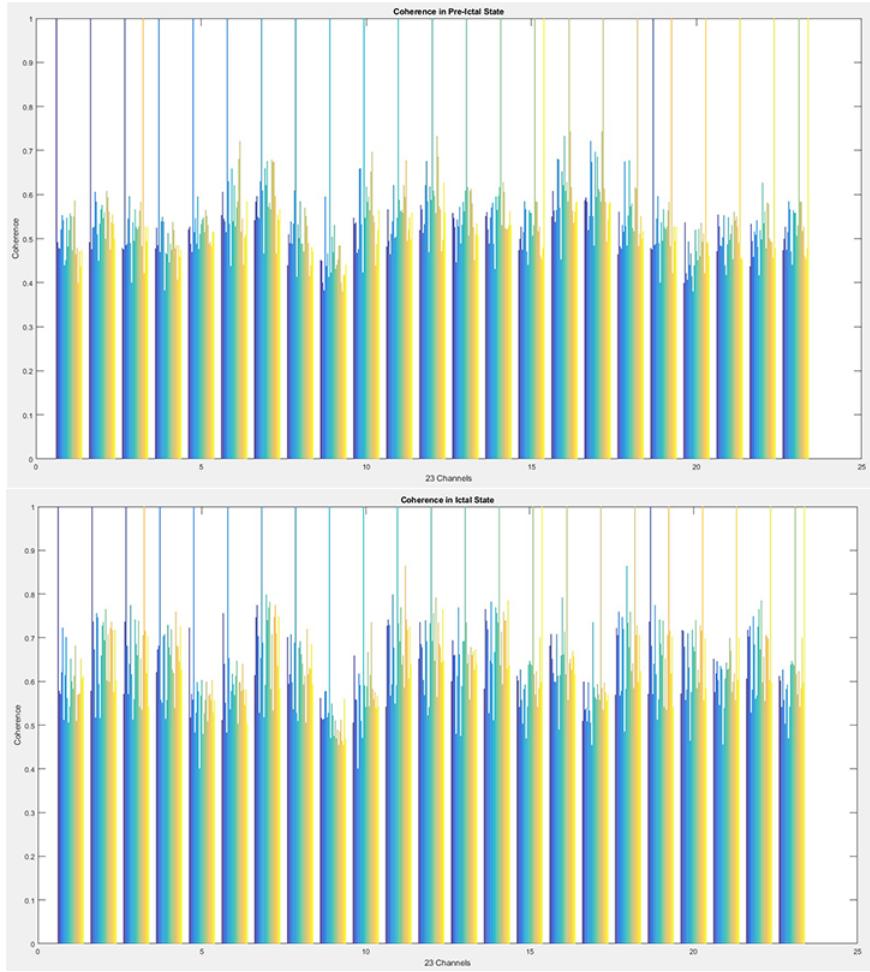


Figure 4.11: Differences of coherence between preictal and seizure

In the Appendix C: Coherence, shows more results between the normal, pre-ictal and Ictal situations among two different subjects where it is observed that there is no high coherence before the epileptic seizure in contrast during the epileptic seizure.

#### 4.5.2 Entropies

The choice of entropies for feature extraction arises from the fact that epileptic EEG signals present a high rate of periodicity, decreasing randomness hence, the measure of information during epilepsy [53]. The generalized concept of entropy is related when a system exhibits high entropy values equal

to high levels of disorder of a system, whereas low values describe a more ordered system, capable of producing more work. Conceptually, entropy has to do with how much information is carried by a signal. We believe that entropy contains useful information and features that can be used for seizure prediction. The computation of each entropy index is described as follows.

### **Shannon and Log Entropy**

Entropy-based wavelet packet decomposition presented by Coifman and Wickerhauser [54] are used to compute the Shannon Entropy and Log Entropy. Conceptually, entropy has to do with how much information is carried by a signal. In other words, entropy can provide analysis with how much randomness is in the signal[55]. In general, the entropy of a finite length discrete random variable,  $x=[x(0) \ x(1) \dots \ x(N-1)]$  with probability distribution function denoted by  $p(x)$  is defined by:

$$H(x) = - \sum_{i=0}^{N-1} p_i(x) \log_2(p_i(x))$$

by where i indicates one of the discrete states. This entropy is larger if each discrete state has about the same probability of occurrence.

Shannon Entropy or information theory introduced by Shannon is a non-linear technique, is considered the value of uncertainty within a random variable. If the entropy is 0, that means that the system is predictable. Due to the chaotic nature of seizure in relation non-seizure activity, the information theory of time series was recommended as a metric of seizure detection and a recognize measure for stress or fatigue (prodromal symptoms). Is a functional of the  $p(x)$  in the sense of expectation in form:

$$H_{ShanEn}(x) = -E \{ \log_2 (p(x)) \}$$

This expression means that the ShanEn is the average information content of x. The ShanEn comes from the field of Information Theory such that it measures the degree of uncertainty that exists in the system as follows,

$$H_{ShanEn}(x) = - \sum_{i=0}^{N-1} (p_i(x))^2 (\log_2(p_i(x)))^2$$

and the Log Entropy of x is given by,

$$H_{LogEn}(x) = - \sum_{i=0}^{N-1} (\log_2(p_i(x)))^2$$

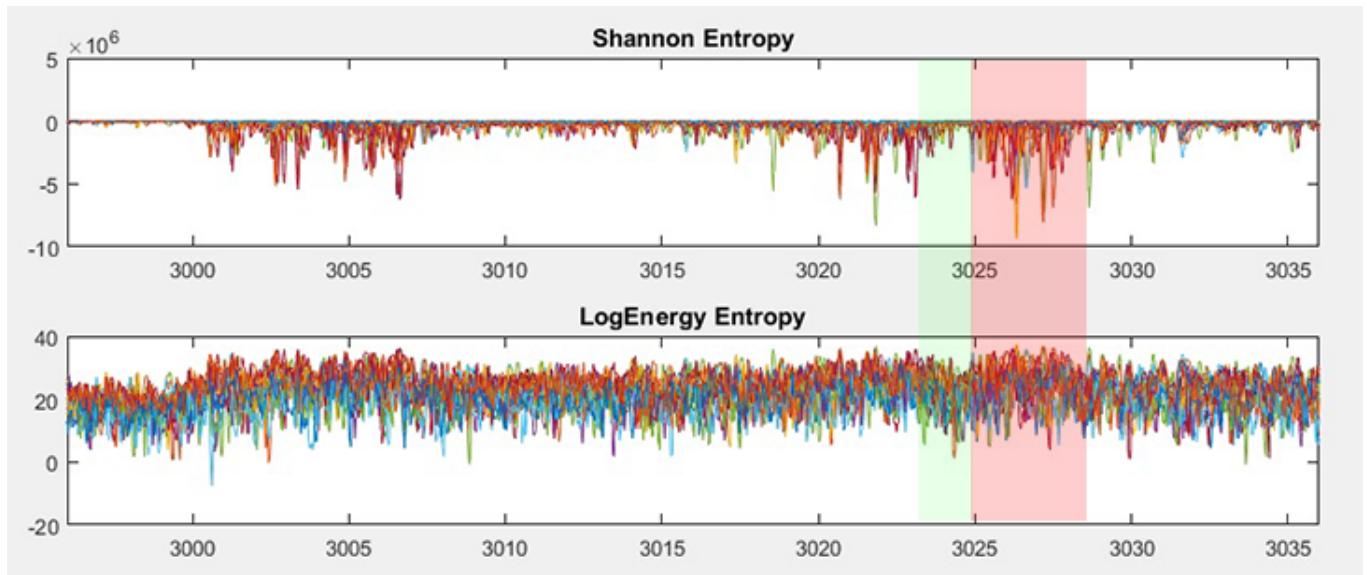


Figure 4.12: Shannon and Log energy Entropy in overall signal

The analysis was also carried out on measurements indicating the entropy of Shannon and Log Energy where a large entropy occurred a few minutes before the onset of the seizure and during of the seizure.

# **Chapter 5**

## **Conclusion**

This section presents the results obtained from the overall process of EEG signal analysis; we conclude the thesis with a summary of its goals and scientific contributions followed by suggested improvements and directions for future work.

### **5.1 Summary of Thesis Achievements**

The electroencephalogram (EEG) has been widely used in senior cognitive research for brain functions, mental states, and emotions. The EEG itself emits a very complex and irregular signal, the analysis of which is a complicated process that requires special attention to avoid distorted results. Analyzing larger volumes of information could result in more scientifically reliable results. Therefore, there is a increasing need for collection of large datasets.

So, the purpose of this dissertation was to present the overall process of data preparation. We worked on the proper selection of specific features for machine learning algorithms, which would help us provide maximum accuracy in the data collection. That way we sought to offer ground for further analysis and research into seizure prediction with EEG. The results obtained using the above features within algorithms lead to reliable results, which were also confirmed by all the current methods can detect the onset of the seizure.

In Chapter "Data Preparation," we describe the steps conducted to collect and analyze the data. This chapter takes a large part (almost 70%)of the thesis, because proper data preparation is crucial before any classification algorithm takes place. This is a process that involves many different tasks and cannot

be fully automated. Furthermore, the steps followed include the accuracy of data, their import to Matlab and data transformation by downsampling, filtering, and cleaning them from outliers(artifacts, missing values, etc).

In Chapter "Feature Selection" , we tested some features for seizure prediction, whose results were quite encouragingly as compared to other studies. The above features were applied to epochs of the signal between inter-ictal, pre-ictal and ictal. More specifically, the coherence indicated that the majority of the patients had lower values before the epileptic seizure than during the crisis. Also, the values of the complexity and mobility of Hjorth parameters showed that the above values are very low before the seizure than under normal conditions. In addition, the values of the activity were lower just before the seizure. An analysis was also carried out on measurements indicating the entropy of Shannon, where a large entropy occurred a few seconds after the onset of the crisis. Finally, we studied the brain waves patterns. Those showed important changes in frequency fluctuation just before the crisis, where most of the subjects appeared to be stressed according to the table 3.1 that is presented in Chapter 3 "Data preparation."

The results that came out of this study were also verified by already published studies. Therefore we can be assured that the procedure followed was performed correctly and accurately.

At this point, it should be noted that there are some limitations in the field of prognosis of epilepsy. Some fundamental factors which make the classification a difficult task are the different types of epilepsy disorders, that may be included in a dataset, and the age difference between youths and adults patients. Another highly important factor is that there is no ideal algorithm to work and to adapt in more than one database.

## 5.2 Applications and Future Directions

In the research for seizure prediction there is a general effort to develop a reliable prediction method that can meet the standards of clinical application. Some approaches that could be used to the analysis of electroencephalographic (EEG) signal and reduce the dimension of data are Principal Component Analysis (PCA) and Independent Component Analysis (ICA). Then these features can be utilized as an input to machine learning algorithms, such as support vector machine (SVM). Another technique is Sparse Coding which can find a set of basis vectors  $\phi_i$  in a way that can represent an input vector

$\mathbf{x}$  as a linear combination of these basis vectors:

$$x = \sum_{i=1}^k a_i \phi_i$$

While techniques such as Principal Component Analysis (PCA) allow us to learn a complete set of basis vectors efficiently, we wish to learn an over-complete set of basis vectors to represent input vectors  $x \in R^n$  (i.e. such that  $k > n$ ). The advantage of having an over-complete basis is that our basis vectors are more able to capture structures and patterns inherent in the input data. However, with an over-complete basis, the coefficients  $a_i$  are no longer uniquely determined by the input vector  $\mathbf{x}$ . Therefore, in sparse coding, it is introduced the additional criterion of **sparsity** to resolve the degeneracy from over-completeness. The Dictionary Learning can also be used, because it is appropriate for the representation of given classes of signals and multisensor data. Moreover, the dimensionality reduction based on dictionary representation can be extended to address specific tasks such as data analysis or classification, when the learning includes a class separability criteria in the objective function. The benefits of dictionary learning clearly show that a proper understanding of causes underlying the sensed world is key to task-specific representation of relevant information in high-dimensional data sets. All above techniques could be a novel approach in future work where deep learning algorithms will be applied in epilepsy onset seizure detection.



# Bibliography

- [1] J. Murray, “Coping with the uncertainty of uncontrolled epilepsy,” *Seizure*, vol. 2, no. 3, pp. 167–178, 1993.
- [2] D. J. Thurman, E. Beghi, C. E. Begley, A. T. Berg, J. R. Buchhalter, D. Ding, D. C. Hesdorffer, W. A. Hauser, L. Kazis, R. Kobau, *et al.*, “Standards for epidemiologic studies and surveillance of epilepsy,” *Epilepsia*, vol. 52, no. s7, pp. 2–26, 2011.
- [3] S. P. Kai Whner, “Data preprocessing vs. data wrangling in machine learning projects.”
- [4] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [5] B. Litt and J. Echaz, “Prediction of epileptic seizures,” *The Lancet Neurology*, vol. 1, no. 1, pp. 22–30, 2002.
- [6] Z. Rogowski, I. Gath, and E. Bental, “On the prediction of epileptic seizures,” *Biological cybernetics*, vol. 42, no. 1, pp. 9–15, 1981.
- [7] Y. Salant, I. Gath, and O. Henriksen, “Prediction of epileptic seizures from two-channel eeg,” *Medical and Biological Engineering and Computing*, vol. 36, no. 5, pp. 549–556, 1998.
- [8] A. Siegel, C. L. Grady, and A. F. Mirsky, “Prediction of spike-wave bursts in absence epilepsy by eeg power-spectrum signals,” *Epilepsia*, vol. 23, no. 1, pp. 47–60, 1982.
- [9] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, “Anticipating epileptic seizures in real time by a non-linear analysis of similarity between eeg recordings,” *Neuroreport*, vol. 10, no. 10, pp. 2149–2155, 1999.

- [10] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams, "Phase space topography and the lyapunov exponent of electrocorticograms in partial seizures," *Brain topography*, vol. 2, no. 3, pp. 187–201, 1990.
- [11] N. Moghim and D. W. Corne, "Predicting epileptic seizures in advance," *PLoS one*, vol. 9, no. 6, p. e99334, 2014.
- [12] P. Aswathappa and M. Ambar, "Separation of background activity of epileptic eeg using artifical wavelet transforms techniques," 2014.
- [13] K. Lehnertz and C. E. Elger, "Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity," *Physical review letters*, vol. 80, no. 22, p. 5019, 1998.
- [14] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clemenceau, and F. Varela, "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures," *European Journal of Neuroscience*, vol. 12, no. 6, pp. 2124–2134, 2000.
- [15] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D.-S. Shiu, "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *Journal of combinatorial optimization*, vol. 5, no. 1, pp. 9–26, 2001.
- [16] K. Schindler, R. Wiest, M. Kollar, and F. Donati, "Eeg analysis with simulated neuronal cell models helps to detect pre-seizure changes," *Clinical neurophysiology*, vol. 113, no. 4, pp. 604–614, 2002.
- [17] F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Physical Review E*, vol. 67, no. 2, p. 021912, 2003.
- [18] V. P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Research*, vol. 26, no. 1, pp. 55–60, 2004.
- [19] V. Srinivasan, C. Eswaran, Sriraam, and N, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, no. 6, pp. 647–660, 2005.
- [20] İ. Güler and E. D. Übeyli, "Expert systems for time-varying biomedical signals using eigenvector methods," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1045–1058, 2007.

- [21] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “The use of time-frequency distributions for epileptic seizure detection in eeg recordings,” in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 3–6, IEEE, 2007.
- [22] K. Polat and S. Güneş, “Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and fft method based new hybrid automated identification system for classification of eeg signals,” *Expert Systems with Applications*, vol. 34, no. 3, pp. 2039–2048, 2008.
- [23] H. Ocak, “Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.
- [24] S. P. Kumar, N. Sriraam, P. Benakop, and B. Jinaga, “Entropies based detection of epileptic seizures with artificial neural network classifiers,” *Expert Systems with Applications*, vol. 37, no. 4, pp. 3284–3291, 2010.
- [25] T. Fathima, M. Bedeeuzzaman, O. Farooq, and Y. U. Khan, “Wavelet based features for epileptic seizure detection,” *MES Journal of Technology and Management*, vol. 2, no. 1, pp. 108–112, 2011.
- [26] R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. K. Ng, and L. Tong, “Application of empirical mode decomposition (emd) for automated detection of epilepsy using eeg signals,” *International journal of neural systems*, vol. 22, no. 06, p. 1250027, 2012.
- [27] S. S. Alam and M. I. H. Bhuiyan, “Detection of seizure and epilepsy using higher order statistics in the emd domain,” *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 2, pp. 312–318, 2013.
- [28] A.-O. Boudraa and J.-C. Cexus, “Emd-based signal filtering,” *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, 2007.
- [29] K. Gadhoumi, J. Gotman, and J. M. Lina, “Scale invariance properties of intracerebral eeg improve seizure prediction in mesial temporal lobe epilepsy,” *PloS one*, vol. 10, no. 4, p. e0121182, 2015.
- [30] T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. A. El-Samie, “Eeg seizure detection and prediction algorithms: a survey,” *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 183, 2014.

- [31] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, “Seizure prediction: the long and winding road,” *Brain*, vol. 130, no. 2, pp. 314–333, 2007.
- [32] C. Baumgartner, W. Serles, F. Leutmezer, E. Pataraia, *et al.*, “Preictal spect in temporal lobe epilepsy: regional cerebral blood flow is increased prior to electroencephalography-seizure onset,” *The Journal of Nuclear Medicine*, vol. 39, no. 6, p. 978, 1998.
- [33] P. Federico, D. F. Abbott, R. S. Briellmann, A. S. Harvey, and G. D. Jackson, “Functional mri of the pre-ictal state,” *Brain*, vol. 128, no. 8, pp. 1811–1817, 2005.
- [34] D. Kerem and A. Geva, “Forecasting epilepsy from the heart rate signal,” *Medical and biological engineering and computing*, vol. 43, no. 2, pp. 230–239, 2005.
- [35] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, “Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic,” *Physica D: nonlinear phenomena*, vol. 194, no. 3, pp. 357–368, 2004.
- [36] N. I. of Neurological Disorders and Stroke, “The epilepsies and seizures: Hope through research.”
- [37] A. Schulze-Bonhage, C. Kurth, A. Carius, B. J. Steinhoff, and T. Mayer, “Seizure anticipation by patients with focal and generalized epilepsy: a multicentre assessment of premonitory symptoms,” *Epilepsy research*, vol. 70, no. 1, pp. 83–88, 2006.
- [38] P. Rajna, B. Clemens, E. Csibri, E. Dobos, A. Geregely, M. Gottschal, I. György, A. Horvath, F. Horvath, L. Mezőfi, *et al.*, “Hungarian multicentre epidemiologic study of the warning and initial symptoms (prodrome, aura) of epileptic seizures,” *Seizure*, vol. 6, no. 5, pp. 361–368, 1997.
- [39] V. Alvarez and A. O. Rossetti, “Electroencephalography and evoked potentials: Technical background,” in *Clinical Neurophysiology in Disorders of Consciousness*, pp. 7–23, Springer, 2015.
- [40] G. Deuschl, A. Eisen, *et al.*, “Recommendations for the practice of clinical neurophysiology: guidelines of the international federation of clinical neurophysiology,” 1999.
- [41] B. Researches, “Epilepsy awareness program - eeg resources.”
- [42] Physionet, “Chb-mit scalp eeg database.”
- [43] D. I. Fotiadis, “Handbook of research on trends in the diagnosis and treatment of chronic conditions,” *IGI Global book series advances in Medical Diagnosis, Treatment, and Care*, 2016.

- [44] L. J. Greenfield, J. D. Geyer, and P. R. Carney, *Reading EEGs: a practical approach*. Lippincott Williams & Wilkins, 2012.
- [45] M. Kalaivani, V. Kalaivani, and V. A. Devi, “Analysis of eeg signal for the detection of brain abnormalities,” *at International Journal of Computer Applications® year*, 2014.
- [46] Gwyddion, “Wavelet transform.”
- [47] R. University, “Multiresolution analysis: The discrete wavelet transform.”
- [48] A. Goshvarpour, H. Ebrahimnezhad, and A. Goshvarpour, “Classification of epileptic eeg signals using time-delay neural networks and probabilistic neural networks,” *International Journal of Information Engineering and Electronic Business*, vol. 5, no. 1, p. 59, 2013.
- [49] M. H. Daily, “5 types of brain waves frequencies: Gamma, beta, alpha, theta, delta.”
- [50] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, “A novel eeg feature extraction method using hjorth parameter,” *International Journal of Electronics and Electrical Engineering*, vol. 2, no. 2, pp. 106–110, 2014.
- [51] S. A. Unde and R. Shriram, “Coherence analysis of eeg signal using power spectral density,” in *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pp. 871–874, IEEE, 2014.
- [52] S. A. Unde and R. Shriram, “Psd based coherence analysis of eeg signals for stroop task,” *International Journal of Computer Applications*, vol. 95, no. 16, 2014.
- [53] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadashivan, “Entropies for detection of epilepsy in eeg,” *Computer methods and programs in biomedicine*, vol. 80, no. 3, pp. 187–194, 2005.
- [54] R. R. Coifman and M. V. Wickerhauser, “Entropy-based algorithms for best basis selection,” *IEEE Transactions on information theory*, vol. 38, no. 2, pp. 713–718, 1992.
- [55] S. Aydin, H. M. Saraoğlu, and S. Kara, “Log energy entropy-based eeg classification with multilayer neural networks in seizure,” *Annals of biomedical engineering*, vol. 37, no. 12, p. 2626, 2009.



# Bibliography

- [1] J. Murray, “Coping with the uncertainty of uncontrolled epilepsy,” *Seizure*, vol. 2, no. 3, pp. 167–178, 1993.
- [2] D. J. Thurman, E. Beghi, C. E. Begley, A. T. Berg, J. R. Buchhalter, D. Ding, D. C. Hesdorffer, W. A. Hauser, L. Kazis, R. Kobau, *et al.*, “Standards for epidemiologic studies and surveillance of epilepsy,” *Epilepsia*, vol. 52, no. s7, pp. 2–26, 2011.
- [3] S. P. Kai Whner, “Data preprocessing vs. data wrangling in machine learning projects.”
- [4] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [5] B. Litt and J. Echaz, “Prediction of epileptic seizures,” *The Lancet Neurology*, vol. 1, no. 1, pp. 22–30, 2002.
- [6] Z. Rogowski, I. Gath, and E. Bental, “On the prediction of epileptic seizures,” *Biological cybernetics*, vol. 42, no. 1, pp. 9–15, 1981.
- [7] Y. Salant, I. Gath, and O. Henriksen, “Prediction of epileptic seizures from two-channel eeg,” *Medical and Biological Engineering and Computing*, vol. 36, no. 5, pp. 549–556, 1998.
- [8] A. Siegel, C. L. Grady, and A. F. Mirsky, “Prediction of spike-wave bursts in absence epilepsy by eeg power-spectrum signals,” *Epilepsia*, vol. 23, no. 1, pp. 47–60, 1982.
- [9] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, “Anticipating epileptic seizures in real time by a non-linear analysis of similarity between eeg recordings,” *Neuroreport*, vol. 10, no. 10, pp. 2149–2155, 1999.

- [10] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams, "Phase space topography and the lyapunov exponent of electrocorticograms in partial seizures," *Brain topography*, vol. 2, no. 3, pp. 187–201, 1990.
- [11] N. Moghim and D. W. Corne, "Predicting epileptic seizures in advance," *PLoS one*, vol. 9, no. 6, p. e99334, 2014.
- [12] P. Aswathappa and M. Ambar, "Separation of background activity of epileptic eeg using artifical wavelet transforms techniques," 2014.
- [13] K. Lehnertz and C. E. Elger, "Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity," *Physical review letters*, vol. 80, no. 22, p. 5019, 1998.
- [14] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clemenceau, and F. Varela, "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures," *European Journal of Neuroscience*, vol. 12, no. 6, pp. 2124–2134, 2000.
- [15] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D.-S. Shiu, "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *Journal of combinatorial optimization*, vol. 5, no. 1, pp. 9–26, 2001.
- [16] K. Schindler, R. Wiest, M. Kollar, and F. Donati, "Eeg analysis with simulated neuronal cell models helps to detect pre-seizure changes," *Clinical neurophysiology*, vol. 113, no. 4, pp. 604–614, 2002.
- [17] F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Physical Review E*, vol. 67, no. 2, p. 021912, 2003.
- [18] V. P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Research*, vol. 26, no. 1, pp. 55–60, 2004.
- [19] V. Srinivasan, C. Eswaran, Sriraam, and N, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, no. 6, pp. 647–660, 2005.
- [20] İ. Güler and E. D. Übeyli, "Expert systems for time-varying biomedical signals using eigenvector methods," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1045–1058, 2007.

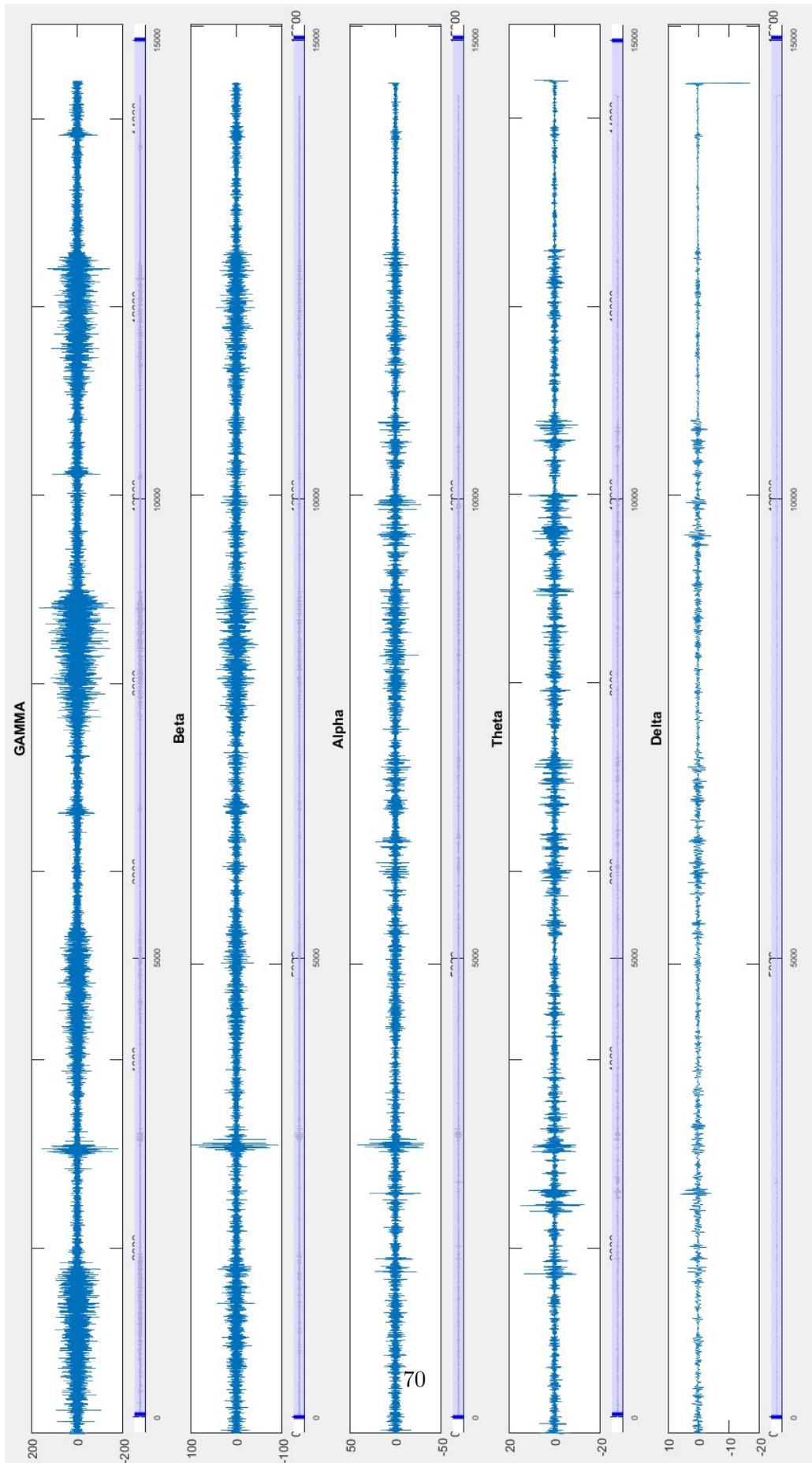
- [21] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “The use of time-frequency distributions for epileptic seizure detection in eeg recordings,” in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 3–6, IEEE, 2007.
- [22] K. Polat and S. Güneş, “Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and fft method based new hybrid automated identification system for classification of eeg signals,” *Expert Systems with Applications*, vol. 34, no. 3, pp. 2039–2048, 2008.
- [23] H. Ocak, “Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.
- [24] S. P. Kumar, N. Sriraam, P. Benakop, and B. Jinaga, “Entropies based detection of epileptic seizures with artificial neural network classifiers,” *Expert Systems with Applications*, vol. 37, no. 4, pp. 3284–3291, 2010.
- [25] T. Fathima, M. Bedeeuzzaman, O. Farooq, and Y. U. Khan, “Wavelet based features for epileptic seizure detection,” *MES Journal of Technology and Management*, vol. 2, no. 1, pp. 108–112, 2011.
- [26] R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. K. Ng, and L. Tong, “Application of empirical mode decomposition (emd) for automated detection of epilepsy using eeg signals,” *International journal of neural systems*, vol. 22, no. 06, p. 1250027, 2012.
- [27] S. S. Alam and M. I. H. Bhuiyan, “Detection of seizure and epilepsy using higher order statistics in the emd domain,” *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 2, pp. 312–318, 2013.
- [28] A.-O. Boudraa and J.-C. Cexus, “Emd-based signal filtering,” *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, 2007.
- [29] K. Gadhoumi, J. Gotman, and J. M. Lina, “Scale invariance properties of intracerebral eeg improve seizure prediction in mesial temporal lobe epilepsy,” *PloS one*, vol. 10, no. 4, p. e0121182, 2015.
- [30] T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. A. El-Samie, “Eeg seizure detection and prediction algorithms: a survey,” *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 183, 2014.

- [31] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, “Seizure prediction: the long and winding road,” *Brain*, vol. 130, no. 2, pp. 314–333, 2007.
- [32] C. Baumgartner, W. Serles, F. Leutmezer, E. Pataraia, *et al.*, “Preictal spect in temporal lobe epilepsy: regional cerebral blood flow is increased prior to electroencephalography-seizure onset,” *The Journal of Nuclear Medicine*, vol. 39, no. 6, p. 978, 1998.
- [33] P. Federico, D. F. Abbott, R. S. Briellmann, A. S. Harvey, and G. D. Jackson, “Functional mri of the pre-ictal state,” *Brain*, vol. 128, no. 8, pp. 1811–1817, 2005.
- [34] D. Kerem and A. Geva, “Forecasting epilepsy from the heart rate signal,” *Medical and biological engineering and computing*, vol. 43, no. 2, pp. 230–239, 2005.
- [35] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, “Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic,” *Physica D: nonlinear phenomena*, vol. 194, no. 3, pp. 357–368, 2004.
- [36] N. I. of Neurological Disorders and Stroke, “The epilepsies and seizures: Hope through research.”
- [37] A. Schulze-Bonhage, C. Kurth, A. Carius, B. J. Steinhoff, and T. Mayer, “Seizure anticipation by patients with focal and generalized epilepsy: a multicentre assessment of premonitory symptoms,” *Epilepsy research*, vol. 70, no. 1, pp. 83–88, 2006.
- [38] P. Rajna, B. Clemens, E. Csibri, E. Dobos, A. Geregely, M. Gottschal, I. György, A. Horvath, F. Horvath, L. Mezőfi, *et al.*, “Hungarian multicentre epidemiologic study of the warning and initial symptoms (prodrome, aura) of epileptic seizures,” *Seizure*, vol. 6, no. 5, pp. 361–368, 1997.
- [39] V. Alvarez and A. O. Rossetti, “Electroencephalography and evoked potentials: Technical background,” in *Clinical Neurophysiology in Disorders of Consciousness*, pp. 7–23, Springer, 2015.
- [40] G. Deuschl, A. Eisen, *et al.*, “Recommendations for the practice of clinical neurophysiology: guidelines of the international federation of clinical neurophysiology,” 1999.
- [41] B. Researches, “Epilepsy awareness program - eeg resources.”
- [42] Physionet, “Chb-mit scalp eeg database.”
- [43] D. I. Fotiadis, “Handbook of research on trends in the diagnosis and treatment of chronic conditions,” *IGI Global book series advances in Medical Diagnosis, Treatment, and Care*, 2016.

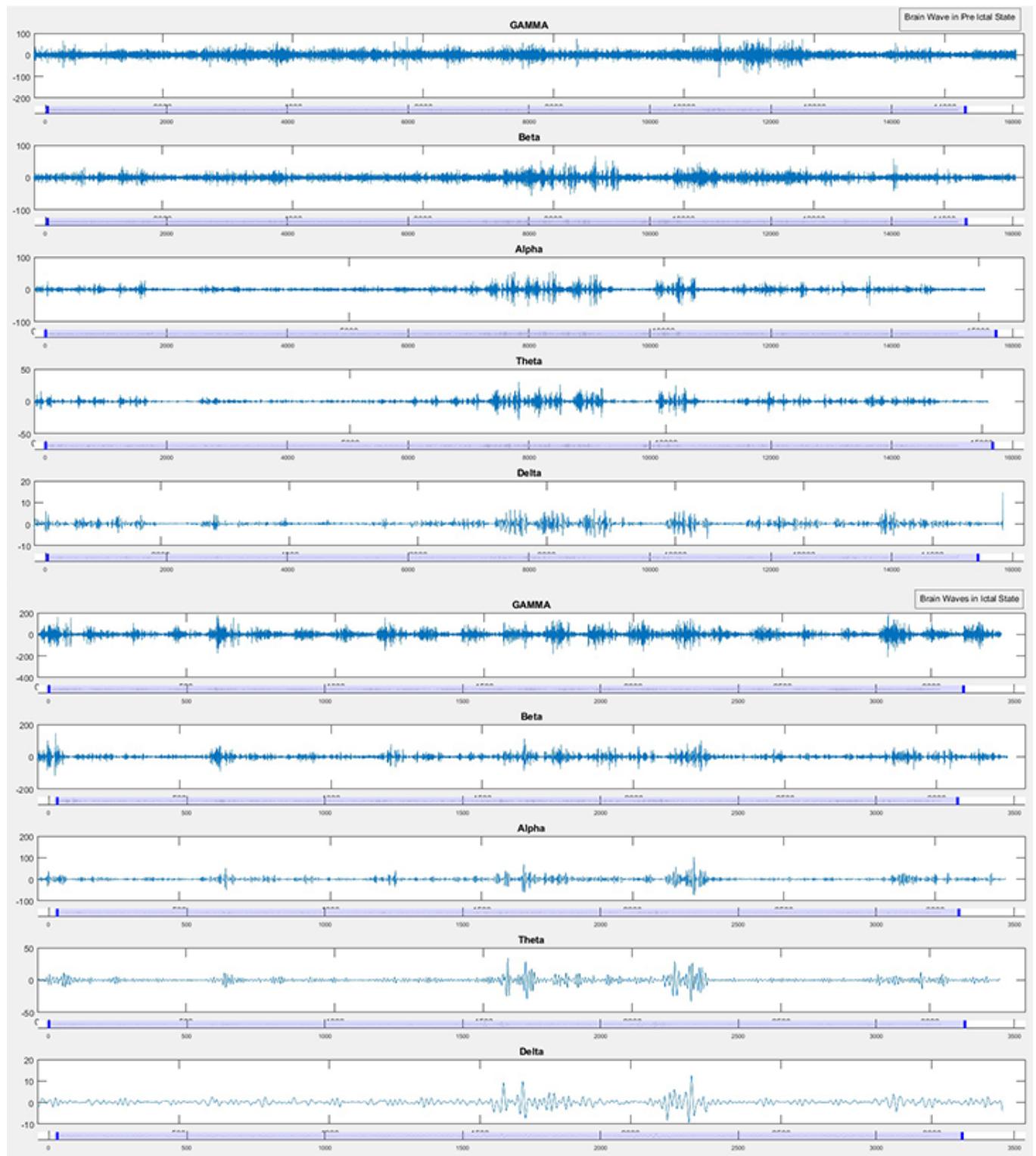
- [44] L. J. Greenfield, J. D. Geyer, and P. R. Carney, *Reading EEGs: a practical approach*. Lippincott Williams & Wilkins, 2012.
- [45] M. Kalaivani, V. Kalaivani, and V. A. Devi, “Analysis of eeg signal for the detection of brain abnormalities,” *at International Journal of Computer Applications® year*, 2014.
- [46] Gwyddion, “Wavelet transform.”
- [47] R. University, “Multiresolution analysis: The discrete wavelet transform.”
- [48] A. Goshvarpour, H. Ebrahimnezhad, and A. Goshvarpour, “Classification of epileptic eeg signals using time-delay neural networks and probabilistic neural networks,” *International Journal of Information Engineering and Electronic Business*, vol. 5, no. 1, p. 59, 2013.
- [49] M. H. Daily, “5 types of brain waves frequencies: Gamma, beta, alpha, theta, delta.”
- [50] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, “A novel eeg feature extraction method using hjorth parameter,” *International Journal of Electronics and Electrical Engineering*, vol. 2, no. 2, pp. 106–110, 2014.
- [51] S. A. Unde and R. Shriram, “Coherence analysis of eeg signal using power spectral density,” in *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pp. 871–874, IEEE, 2014.
- [52] S. A. Unde and R. Shriram, “Psd based coherence analysis of eeg signals for stroop task,” *International Journal of Computer Applications*, vol. 95, no. 16, 2014.
- [53] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan, “Entropies for detection of epilepsy in eeg,” *Computer methods and programs in biomedicine*, vol. 80, no. 3, pp. 187–194, 2005.
- [54] R. R. Coifman and M. V. Wickerhauser, “Entropy-based algorithms for best basis selection,” *IEEE Transactions on information theory*, vol. 38, no. 2, pp. 713–718, 1992.
- [55] S. Aydin, H. M. Saraoğlu, and S. Kara, “Log energy entropy-based eeg classification with multilayer neural networks in seizure,” *Annals of biomedical engineering*, vol. 37, no. 12, p. 2626, 2009.

## .1 Appendix A : Brain waves

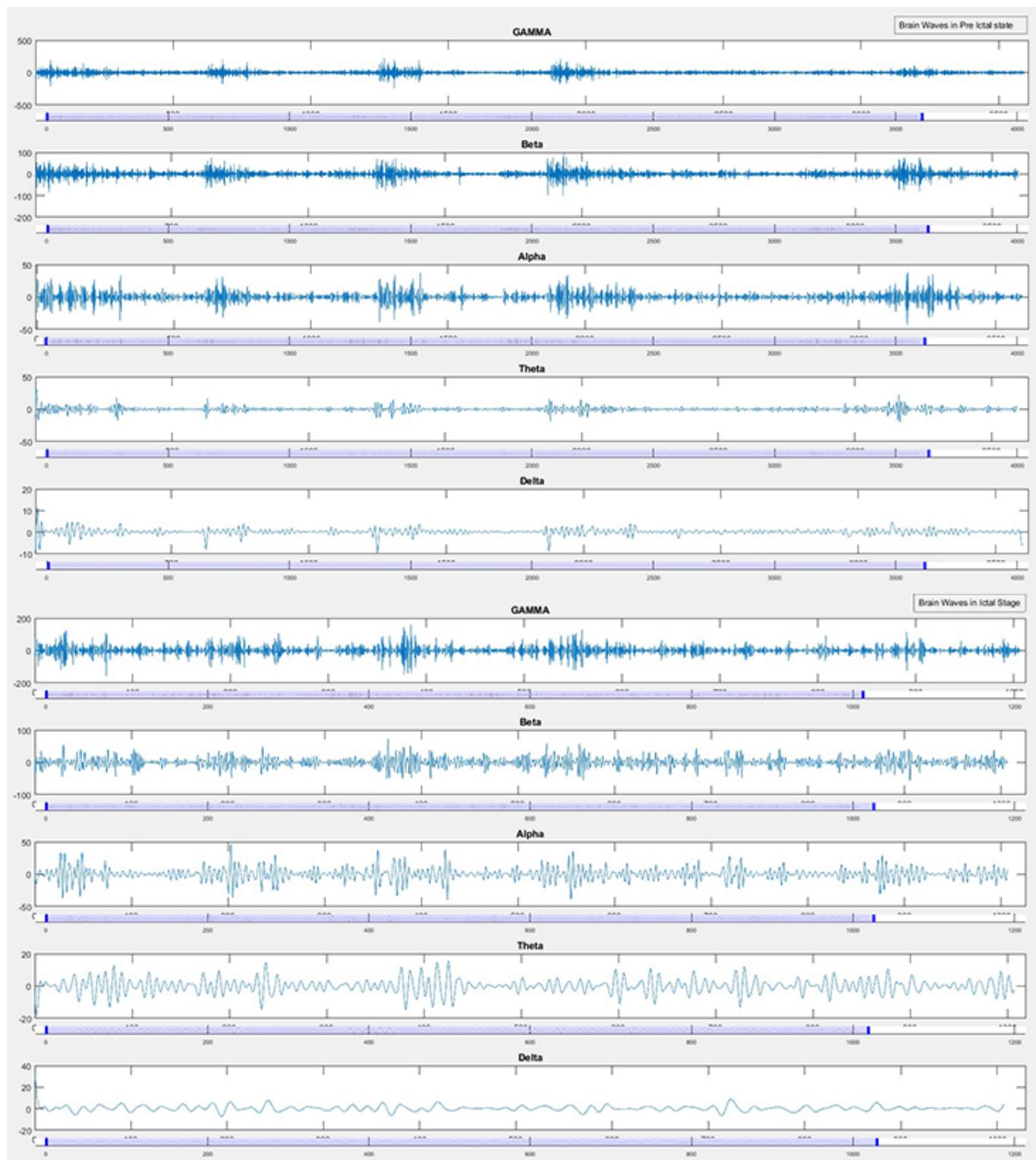
### .1.1 Normal activity at Subject Chb01



### .1.2 Differences between pre-ictal and ictal state at Subject Chb01

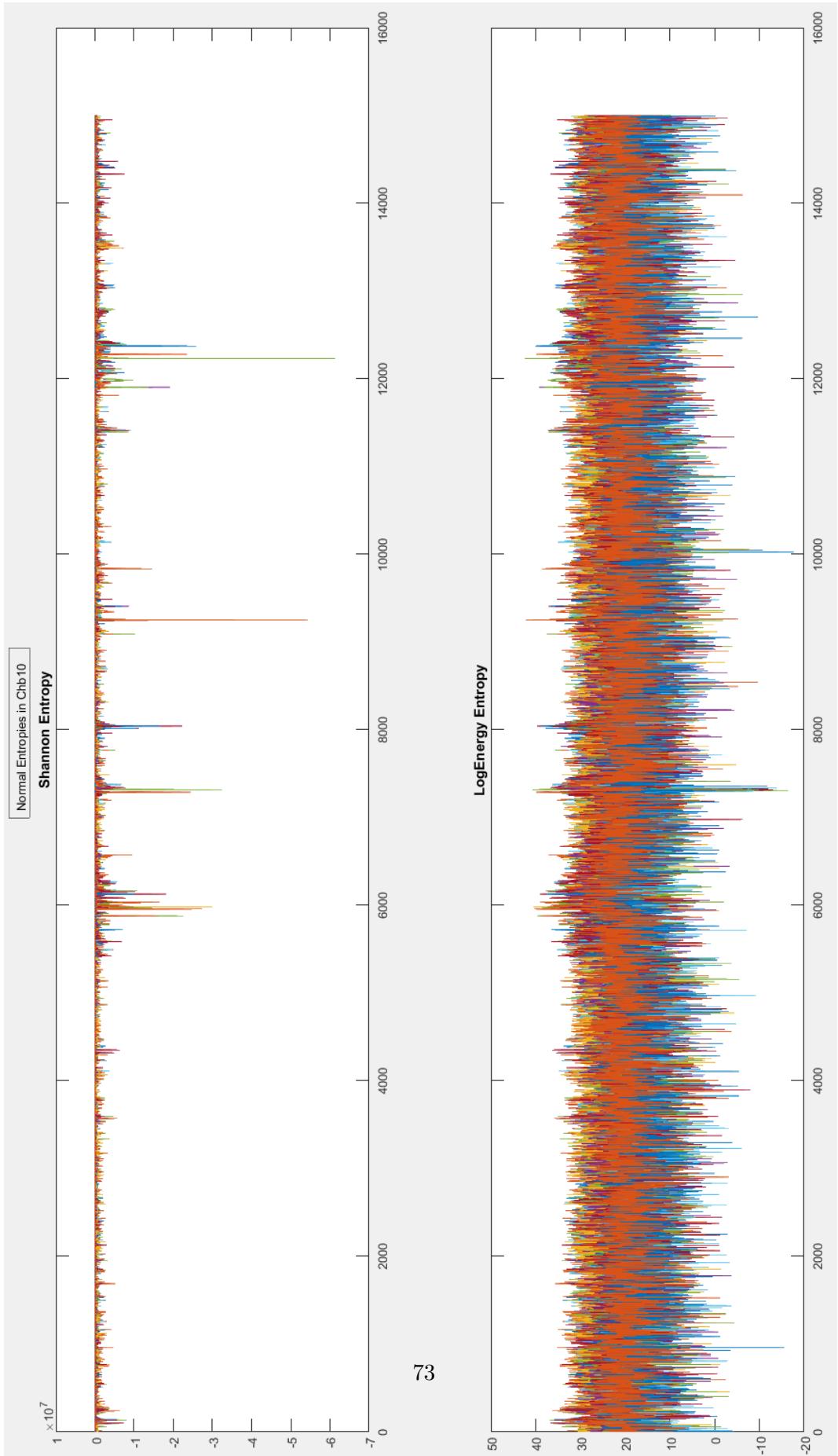


### .1.3 Differences between pre-ictal and ictal state at Subject Chb24



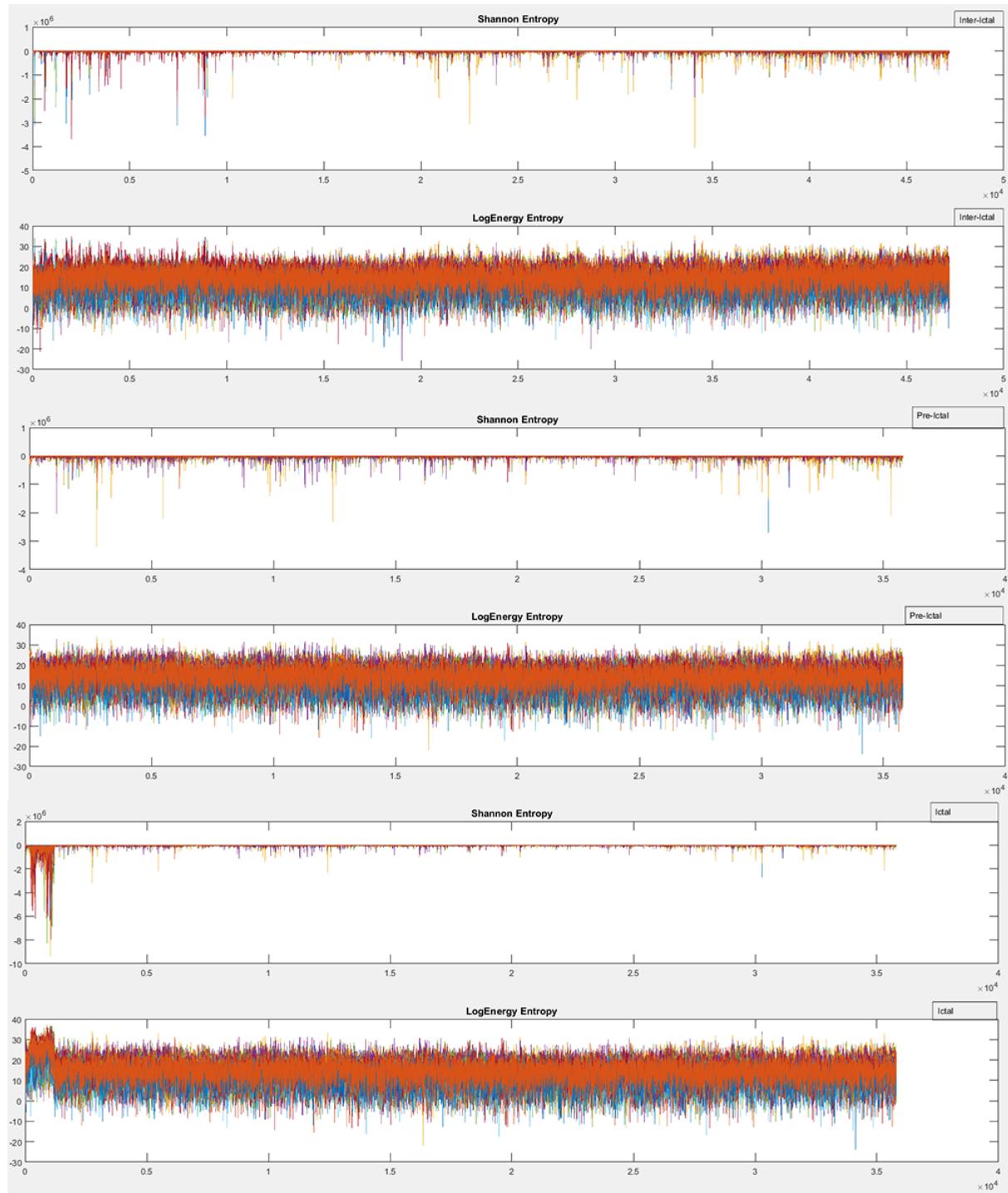
## .2 Appendix B: Entropies

### .2.1 ShanEntropy & LogEntropy at Subject Chb10



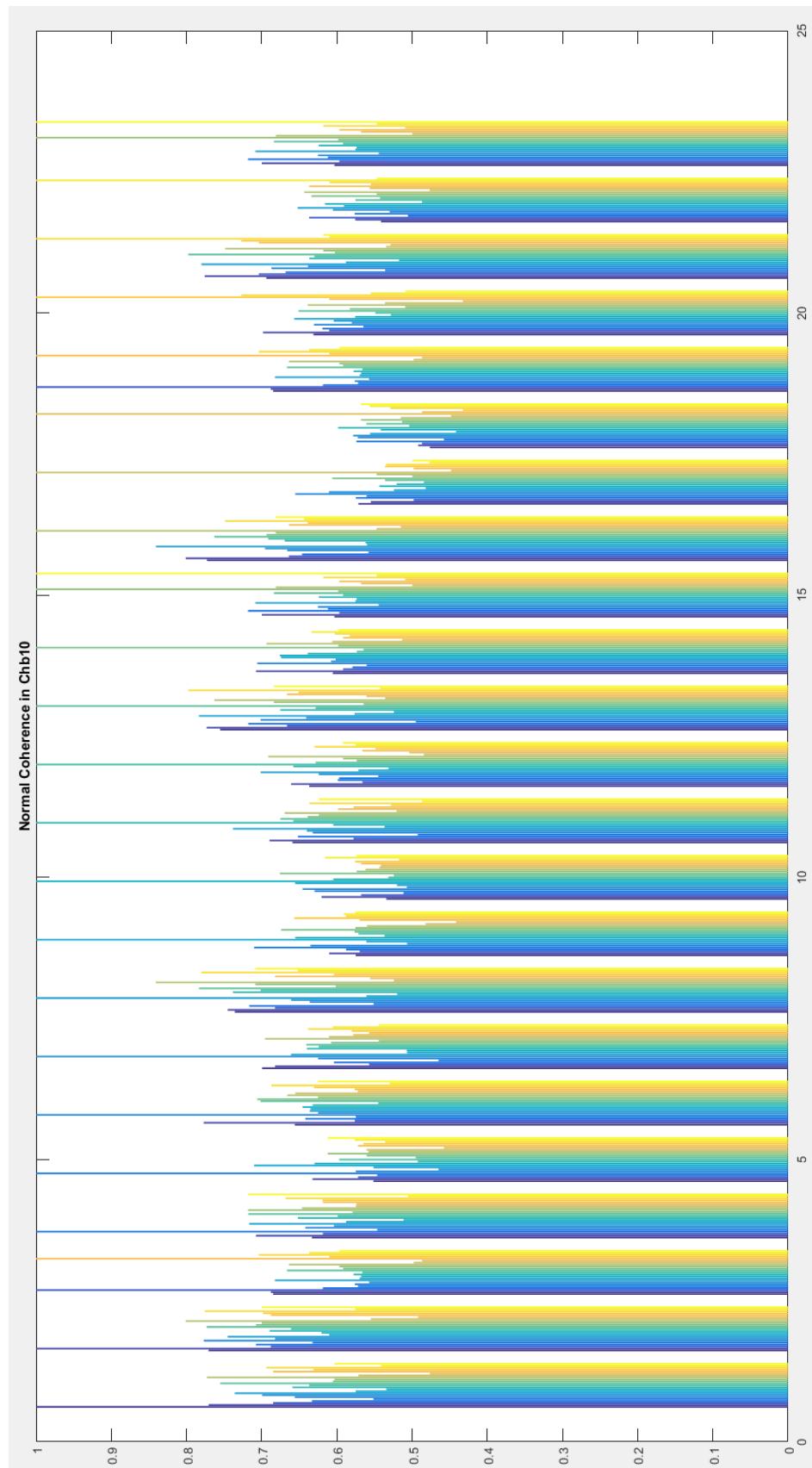
The red marker displays the seizure, the green one display the pre-seizure activity between two entropies.

### .2.2 ShanEntropy & LogEntropy at Subject Chb01

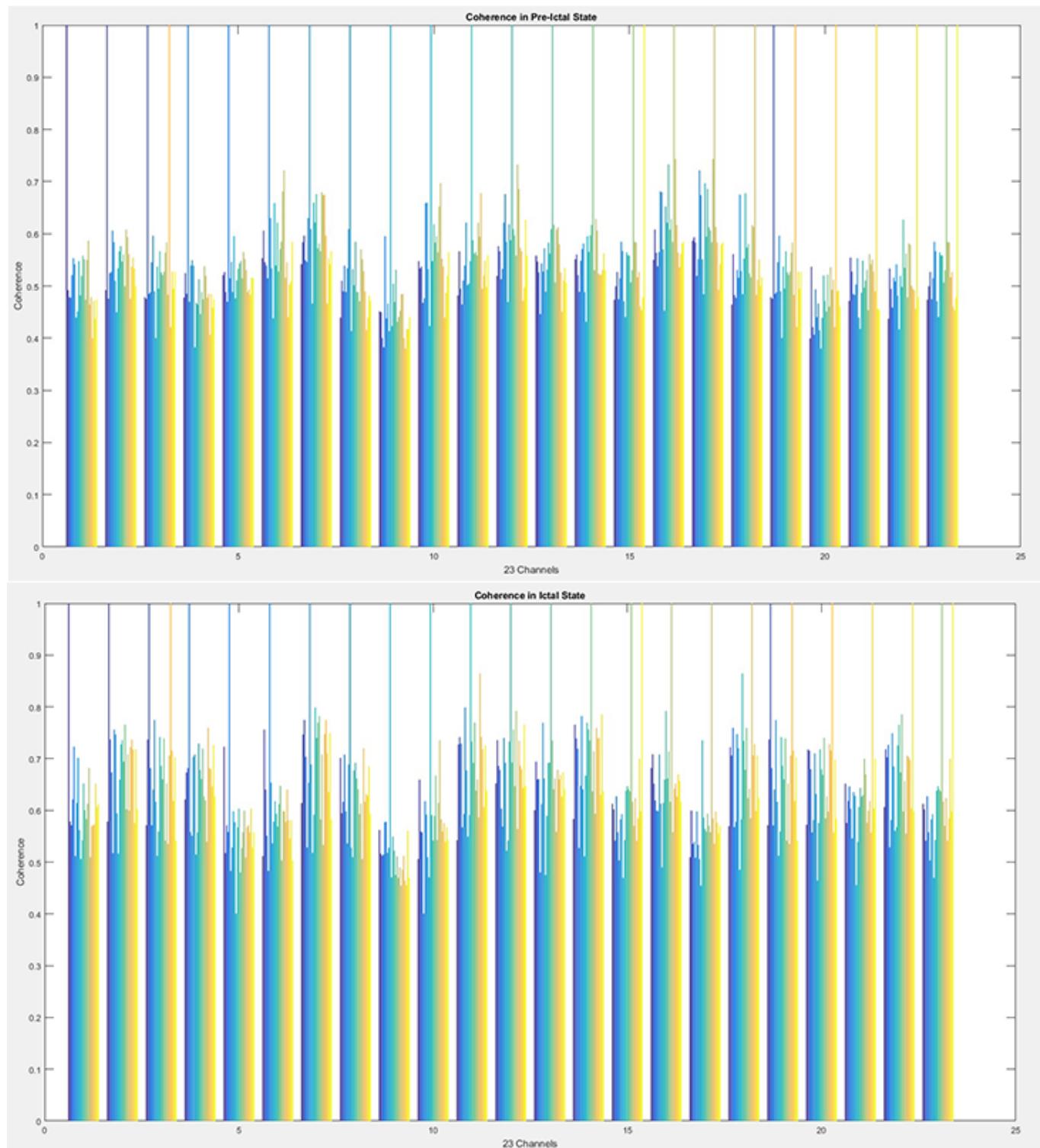


### .3 Appendix C: Coherence

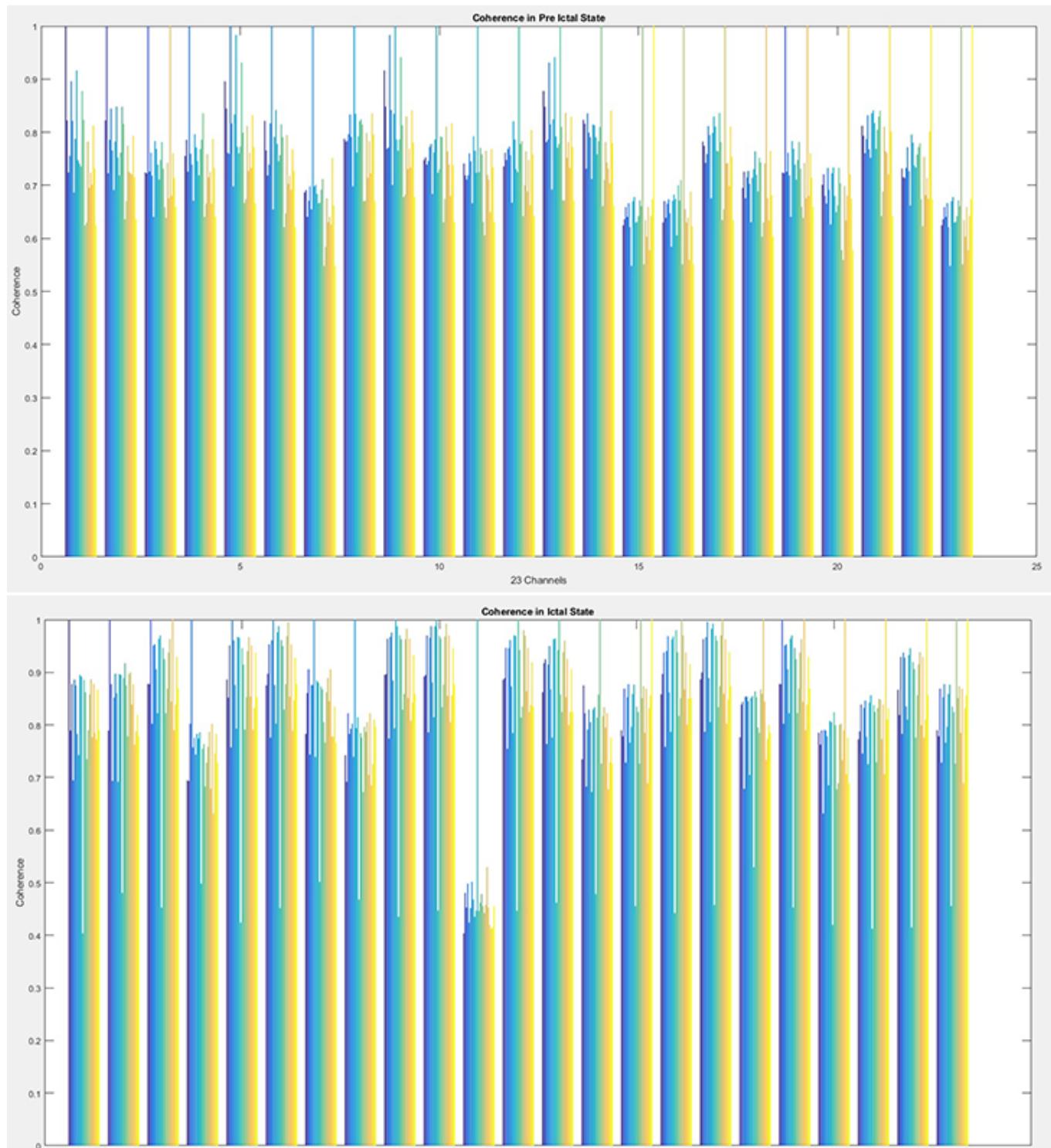
#### .3.1 Normal coherence at Subject Chb10



### .3.2 Differences between pre-ictal and ictal state at Subject Chb10

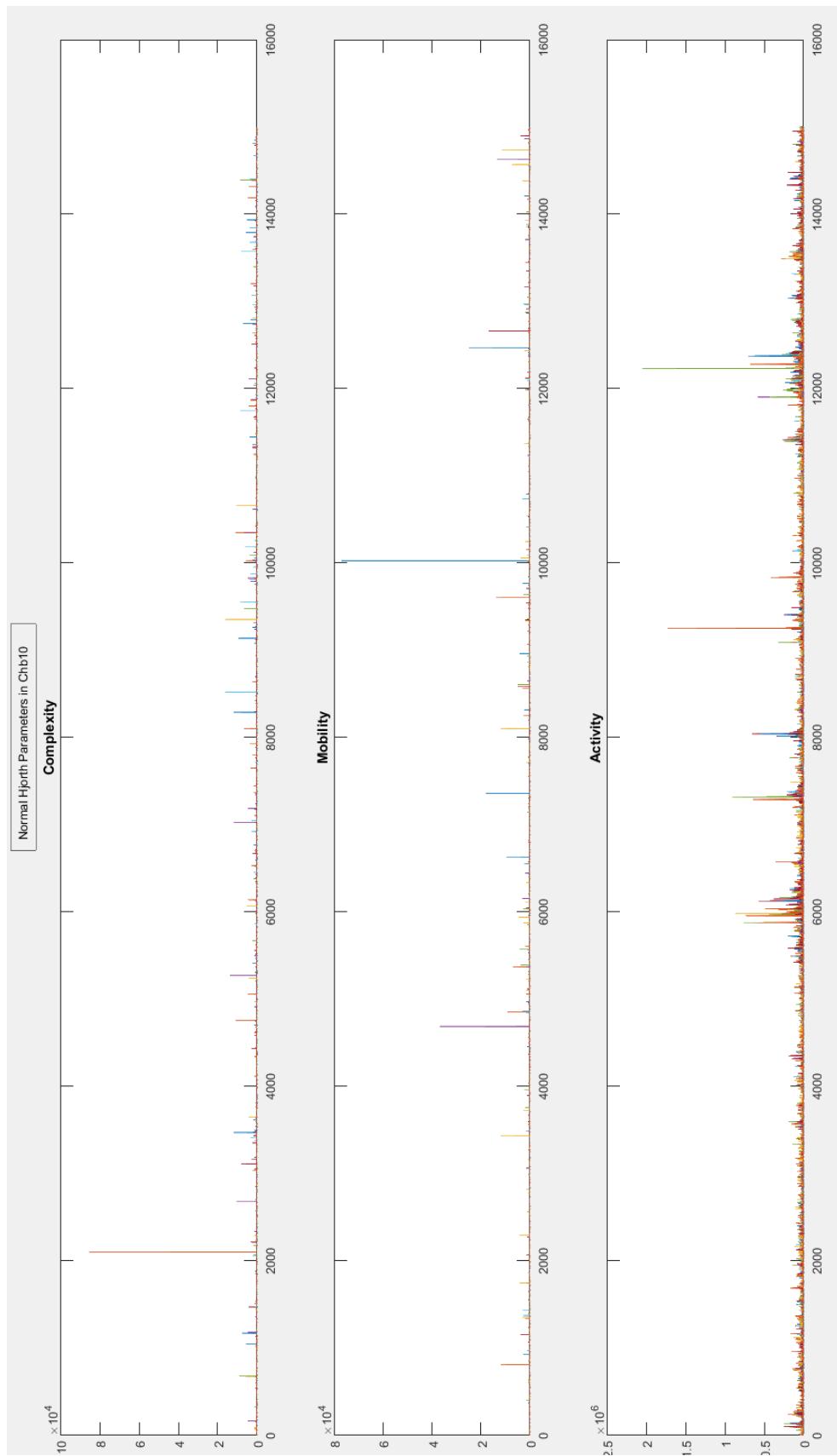


### .3.3 Deferencies between pre-ictal and ictal state at Subject Chb24

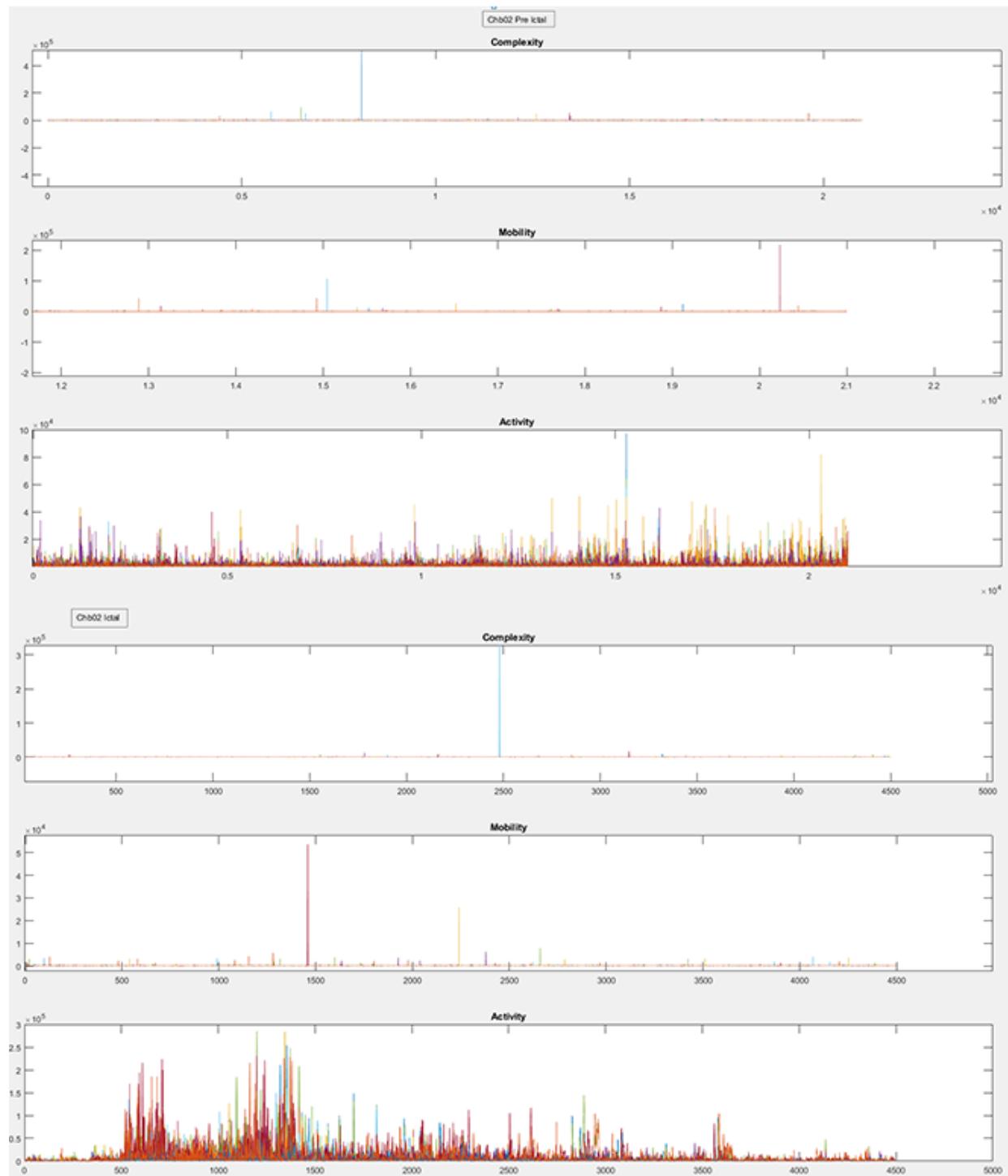


## .4 Appendix D: Hjorth parameters

### .4.1 Normal Hjorth parameters at Subject Chb10



#### .4.2 Differences between pre-ictal and ictal state at Subject Chb02



#### .4.3 Differences between pre-ictal and ictal state at Subject Chb24

