UNIVERSITY OF SOUTHAMPTON

Predicting Epileptic Seizures from Electroencephalography

by

Ahmed A. Mubaraki

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the

Faculty of Engineering and Physical Sciences

School of Electronics and Computer Science

August 2020

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCE
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Ahmed A. Mubaraki

Epilepsy is a neurological disorder that is characterised by repeated seizures. The sudden onset of a seizure affects a patient's quality of life. Therefore, predicting an epileptic seizure in advance can improve their life by giving them warning and thus avoiding serious accidents. In this work, two general prediction models are formulated using the electroencephalography (EEG) signals of patients with Temporal Lobe Epilepsy (TLE) and Absence seizures.

EEG is the most common technique to map brain functions. Studying brain functions and how the brain regions interact is essential to understand the basis of several neurodegenerative diseases. Functional brain connectivity, as derived from multichannel EEG, is currently used as a tool to understand how the various brain regions interact with each other during a cognitive task. Researchers started to study the functional brain network by analysing the EEG data captured. Because of the high level of synchronization observed during a seizure, synchronization measures are logically the best way to assess the dynamic change in functional brain connectivity.

In the current work, Phase Locking Value (PLV), Phase Lag Index (PLI) and Synchronization likelihood (SL) were used to create functional brain connectivity networks. The networks were characterized by nine graph-theoretic parameters (assortativity coefficient, transitivity, clustering coefficient, strength of node, modularity, betweenness centrality, characteristic path length, global efficiency and radius). A machine-learning framework was used to extract the patterns that the patients' data had in common to build the prediction models.

Both general prediction models were formulated using PLI and SL connectivity networks. They achieved sensitivity (both 100%) and a false prediction rate of 0.00001/h and 0.01/h, with a maximum prediction time of 19 and 40 minutes, respectively.

Contents

List of Figures

List of Tables

Abbreviations

$ADHD$	Attention Deficit Hyperactivity Disorder
ANN	Artificial Neural Network
AR	Average Reference
BC	Betweenness Centrality
$BOLD$	Blood-Oxygen-Level-Dependent
BSS	Blind Source Separation
CFS	Cerebrospinal Fluid
CPL	Characteristic Path Length
CSP	Common Spatial Pattern
DWT	Discrete Wavelet Transformation
EEG	Electroencephalography
EMG	Electromyogram
EOG	Electrooculogram
FIR	Finite Impulse Response
$fMRI$	Functional Magnetic Resonance Imag
FPR	False Prediction Rate
ICA	Independent Component Analysis
$iEEG$	Intracranial Electroencephalography
KNN	K-Nearest Neighbour
LDA	Linear Discriminant Analysis
LM	Linked Mastoid
LR	Left-mastoid Reference
MRI	Magnetic Resonance Image
$NA-MEMD$	Noise-Asseisd Multivariate Empirical Mode Decomposition
PCA	Principal Component Analysis
PLI	Phase Lag Index
PLV	Phase Locking Value
SL	Synchronization Likelihood
$STFT$	Short-Time Fourier Transform
SVM	Support Vector Machine
TLE	Temporal Lobe Epilepsy

UNIVERSITY OF Southampton

Academic Thesis: Declaration Of Authorship

I, Ahmed Ali Mubaraki declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Practical Adaptive Security for Resource-Constrained IoT Devices

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: .....

Date: .....

Acknowledgements

I would like to thank my family and my friends for their support. Also I would like to thank my supervisor Prof. Koushik Maharatna for his support and his patience. Finally, I would like to thank Dr. Dniel Konn, Dr. David Allen, Dr. Simona Pellacani and Dr. Federico Sicca for providing the data of my work

Chapter 1

Introduction

Neurological disorders originate from dysfunction of neuronal system and significantly affect cognitive and motor functions. They can be broadly categorised into Neurodevelopmental disorders – the disorders that are set from childhood and continue throughout the life, and neurodegenerative disorders that may result from a number of factors, like brain injury, genetic signature, aging etc. While Autism, Attention Deficit Hyperactivity Disorder (ADHD) fall within the first category, the diseases in the second category include epilepsy, Parkinson's, Alzheimer's, etc. Out of these disorders epilepsy is particularly disruptive to a patient's life as acute epileptic seizures may occur even when a person is in apparently healthy state. Such seizures may associate with complete loss of consciousness and muscle stiffening which could be life threatening depending upon the activity the person is involved at the moment of seizure – imagine a seizure occurs when a person is driving a car, which will certainly result into a life-threatening incidence. In the UK, people diagnosed with epilepsy are required to hand in driving licenses (Service 2015). Currently about 50 million of people around the world suffer from epilepsy (Organization 2019).The increasing rates of people with epilepsy make it imperative to find new methods to diagnose and treat it.

1.1 Motivation

Epilepsy is one of the common neurological diseases affecting the brain and therefore investigating the brain functions will contribute in developing new techniques for its diagnosis and treatment. It is defined as a clinical condition whereby the brain is affected by repeated seizures (epileptic seizures). These epileptic seizures originate from abnormal electrical activity of neurons. Anti-epileptic drugs are one of the most important means used to tackle epilepsy. However, approximately one-third of the patients are not benefited with the use of anti-epileptics (Carney et al. 2011). To diagnose epilepsy and then find an appropriate treatment, researchers began to study the brain and how

it works. The early detection of the epileptic seizures is necessary in order to prevent the occurrence of series of these seizures. This early detection helps to deliver the right intervention at the right time and contribute to develop new methods for treatment. Although extensive research has been carried out on predicitng the epileptic seizures, no single model exists that can predict more than one type of epilepsy. In addition, these models have limitations represented in the low percentages of the sensitivity (the rate of correct predicting seizures) as in (Li et al. 2013, Aarabi & He 2012)(75.8% and 79.9% respectively) and high false prediction rate (Iqbal et al. 2015, Cho et al. 2016)(0.38/h and 0.5/h respectively). Table 1.1 illustrates some of these models with their associated performance. Our motivation is to develop a novel model that can predict seizure with high sensitivity and low false prediction rate while maximizing the prediction time for an impending seizure episode.

Reference of the Model	**Sensitivity (%)**	**FPR (/h)**	**Predicting Time Rate (minutes)**	**Amount of Data Analysed (hours)**
Li et al. (2013)	75.8	0.09	49.7	582
Aarabi & He (2012)	79.9	0.17	50	316
Chu et al. (2017)	86.67	0.367	45.3	583
Cho et al. (2016)	82.44	0.5	5	844
Iqbal et al. (2015)	100	0.38	60	-

TABLE 1.1: Prediction Models with their Associated Performance and Predicitng Time. FRP refers to False Prediction Rate

1.2 Aims and Objectives

The aim of study is to develop a model for prediction of epileptic seizures using functional brain connectivity networks in a machine learning framework with low false prediction rate. The thesis has the following objectives:

- The first objective of the research is trying to predict epileptic seizures for patients with Temporal Lobe Epilepsy (TLE) by applying each one of the following functional connectivity approaches (Phase Locking Value, Phase Lag Index and Synchronization Likelihood) separately.
- Predicting epileptic seizure for patients with other type of epilepsy seizures using the same three functional connectivity approaches.
- Finally, based on the results of the previous two objectives, a generalized prediction algorithm will be developed to predict more than one type of epilepsy. It is a new challenge because the current prediction methods focus only one type of epilepsy.

1.3 Contribution

The following are the new contributions of this work:

- Two predictive models have been formulated for TLE subjects using Phase Locking Value (PLV). They achieved high sensitivity (94%, 92%) and low false prediction rate (0.04/h, 0.03/h) with maximum prediction time 12 minutes. As well as two predictive models have been formulated to predict the epileptic seizures of TLE subjects using Phase Lag Index (PLI) and Synchronized Likelihood (SL) with sensitivity (88%, 87%) and false prediction rate (0.4/h, 0.17/h) respectively. Both models achieved maximum prediction time 24 minutes.
- For patients with absence seizures, three predictive models have been formulated using the three functional connectivity measures (PLV, PLI and SL). They achieved high sensitivity (100%, 99%, 99%) and maximum prediction time (41.8 mins, 17.43 mins, 41.9 mins) respectively.
- Finally two general predictive models have been formulated to predict both types of epileptic seizures using PLI and SL. The first models achieved sensitivity (100%) and false prediction rate (0.00001/h) with maximum prediction times (5 mins, 19 mins) for TLE and Absence subjects respectively. The second model achieved sensitivity (100%) and false prediction rate (0.1/h) with maximum prediction time (5 mins, 40 mins) for TLE and Absence subjects respectively.

1.4 Thesis Outline

Chapter 2 provides the background and a literature review of brain connectivity and the techniques used to map the brain. It also describes EEG signals and how they are measured. The various techniques to predict epileptic seizures mentioned in this chapter have prompted the development of this study's proposed method, and a description is provided of its synchronisation approaches (PLV, PLI and SL). The chapter reviews graph theory and its role in predicting seizures, in addition to the machine-learning algorithms used in this study. Chapter 3 describes the techniques proposed to predict epileptic seizures along with the data source used in this work and the pre-processing procedure steps applied to obtain clean EEG data. Chapter 4 presents and describes the results of applying the proposed method with the PLV approach to TLE subjects. Chapter 5 presents the results of formulating predictive models for TLE subjects using both PLI-based and SL-based brain connectivity networks. In Chapter 6, the three functional connectivity formulation approaches; PLV, PLI and SL; were again used in formulating predictive models for patients with absence seizures. Chapter 7 illustrates the two general predicative models that have been formulated based on the results

obtained in Chapters 4, 5, and 6. Finally, Chapter 8 provides a conclusion, with reference to possible future work.

Chapter 2

Literature Review and Background

To obtain fundamental knowledge about epilepsy that may lead to alternative, more effective treatment, many studies have begun by analysing brain activity, then the research has moved in two directions: the early detection of the onset of an epileptic seizure and its prediction. Methods have been developed to detect the onset of an epileptic seizure, to give an automatic diagnosis without the need for a doctor's involvement. Moreover, knowing in advance that there will be an epileptic seizure improves the quality of a patient's life by giving them some warning, enabling them to avoid many of the problems that arise from its sudden onset. For decades, scientists have worked on developing methods to predict epileptic seizures. They rely on analysing brain activity using known techniques of brain mapping (e.g. MRI and EEG) giving patients the opportunity to minimise the dangers and doctors a chance to schedule the appropriate treatment. Some of these methods have yielded promising results, yet they have not exhibited the same level of precision for all patients due to physiological differences.

Studying the prediction of epileptic seizures is more important than the detection, because they can already be detected by a doctor using proven diagnostic methods, while predicting them involves observations of sudden or abnormal changes in brain function, which can contribute in avoiding problems during seizure occurring. Accordingly, the current study aims to predict epileptic seizures. This chapter provides the essential background and describes the state-of-the-art techniques used to explore brain connectivity. Existing methods of predicting epileptic seizures using EEG are also discussed. Phase synchronisation phenomenon between EEG electrodes, which is considered to be one of the most common approaches to capture brain synchronisation, especially in epilepsy, is described. Finally, the fundamentals of graph theory are reviewed, with its potential for its application in analysing the brain connectivity network for prediction of epilepsy.

2.1 Epilepsy

Epilepsy originates from uncontrolled synchronisation of neuron populations. It is classified by type of seizure: partial epilepsy; or generalised epilepsy. Each is identified by characteristics such as the seizure's origin and the associated symptoms (Fisher et al. 1997). In partial (focal) epilepsy, the seizure begins in one hemisphere of the brain and is not usually accompanied by loss of consciousness. In contrast, generalised epilepsy spreads to all parts of the brain, and this type is usually hereditary. It is accompanied by loss of consciousness (Fisher et al. 1997). Partial epilepsy can be simple, without loss of consciousness, or complex, accompanied by a loss of awareness, while generalised epilepsy can be one of several types. An absence seizure, characterised by loss of consciousness without convulsion, lasts for only a short time (2 to 10s). Atonic-clonic seizure is another type, and it lasts from one to 10 minutes. It is characterised by sudden muscle limpness and can be accompanied by body stiffening, as in the tonic seizure type. This last type, tonic seizure, lasts less than 20 seconds. In the myoclonic seizure type, muscle jerking can occur.

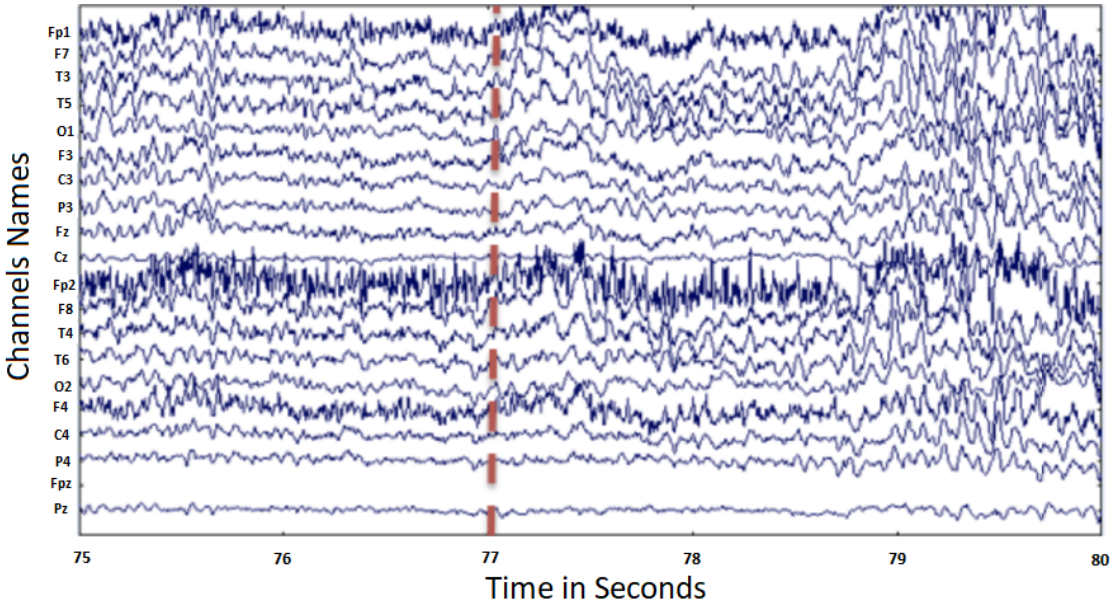


FIGURE 2.1: It is illustration of EEG signals captured by electrodes; the red vertical dotted line refers to seizure onset

The types of epilepsy vary in terms of their duration, symptoms and place of origin. The epilepsy may be linked to a specific region in the brain, as in the case of temporal lobe epilepsy. One of the most important symptoms associated with epilepsy is recurring seizures. However, not all seizures signify epilepsy, as other causes may be responsible, such as brain injury or lack of oxygen at birth. The seizures are certainly the chief problem associated with epilepsy, and they occur suddenly. They endanger the

patient, especially when exercising, driving a car or performing similar activities. Figure 2.1 presents the brain activity of a patient with temporal lobe epilepsy, as captured by EEG signal. The vertical axis shows the channel name, and each of them refers to the region of the brain where the EEG signal was captured. The horizontal axes represents the capture time of that signal. The dashed red line represents the onset of the epileptic seizure. During the epileptic seizure, the brain regions synchronise with each other uncontrollably, starting to work together at the same frequency. Although various techniques could be used to predict an epileptic seizure the fact that its origins are in the synchronisation of regions of the brain could be exploited to provide early warning by capturing this abnormal synchronisation. The direct way to measure it is through functional brain connectivity network analysis.

2.2 Brain Connectivity

The human brain is divided into regions, composed of neurons, and each region is responsible for conducting a particular task. The brain connectivity network, in essence, illustrates how the brain regions are connected and interact when a specific task is performed. Brain connectivity can be classified as: functional connectivity; structural (physical) connectivity; or effective connectivity (Figure 2.2) (Niso et al. 2013). The usual ways to capture brain data for formulating various types of brain connectivity networks are MRI and its variants, and EEG. These are described in the following subsections.

2.2.1 Structural connectivity

Structural (anatomical) connectivity refers to the physical connections that link the various brain areas. These can be synaptic contact between the neurons, or what can be described as white matter tracts (sets of fibre tracts). MRI is employed to map the brain's structure, and it uses the interaction between radio waves and magnetic fields to create a detailed image of the brain's structure. A variant of MRI, diffusion MRI (dMRI), can be used in conjunction with the technique of tractography to find the white matter tracts (Sporns 2010) between different brain regions.

2.2.2 Functional connectivity

Functional connectivity refers to the temporal correlation of the interaction between brain regions, which are spatially connected/distributed (Friston 2011). It can be calculated by determining the extracted phase synchronisation between them. The various methods to measure it in the brain depend on specific techniques such as EEG and functional MRI (fMRI). Functional connectivity can be calculated at the cortical level, as

well as at scalp level. At the cortical level, it can be extracted using fMRI to detect the change in oxygen level in each region, based on Blood-Oxygen-Level-Dependent (BOLD) signals. The technique produces a map showing which parts of the brain are involved in a particular task. At the scalp level, functional connectivity is calculated using EEG by measuring the electrical activity in the brain regions as they interact. Using electrodes (metal discs) placed on the scalp, it picks up the signals resulting from brain cell firing. Functional brain connectivity is usually represented as a graph with brain areas as nodes and the strength of their correlation as the arcs.

2.2.3 Effective connectivity

Effective connectivity refers to the measures of the causal (directional) influences of the brain regions, building up a directed graph whose links represent the direct effects of each region on another. Effective connectivity can map the brain using the same techniques as in functional connectivity formulation, such as EEG and fMRI. The existing methods to calculate effective connectivity are mainly Granger causality or model-based dynamic causal modelling (DCM) (Niso et al. 2013).

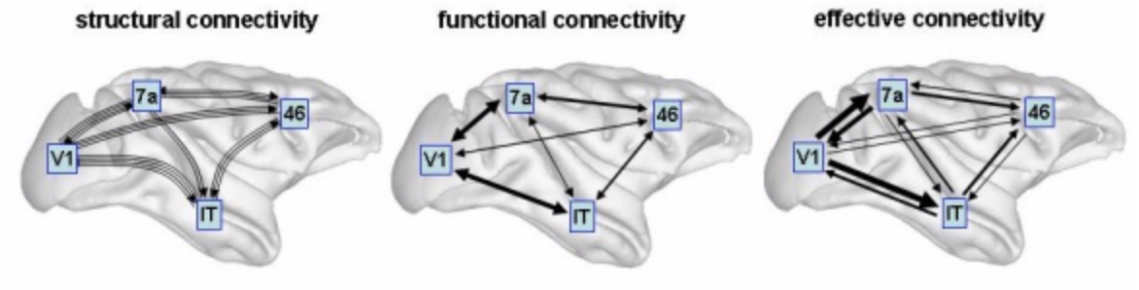


FIGURE 2.2: Three types of brain connectivity (redrawn from (Sporns 2007)).

Generally speaking, structural connectivity, the measurement of which involves the MRI technique, is not time dependent, as its temporal changes are slower than those of functional connectivity. The change of structural connectivity can occur due to brain injury or acquired experiences during ageing (Scholz et al. 2009). MRI to map the brain's structure has excellent spatial resolution, which means that they relate to physical location (the whole brain). Consequently, the fMRI technique used to measure functional connectivity has extremely high spatial resolution and can thus record signals from all parts of the brain. It is considered to be a safe technique to estimate brain function with no risk involved (non-invasive). Its drawbacks are its substantial cost and the fact that patients must lie down and not move at all if a clear image is to be obtained. Moreover, fMRI depends on the BOLD method for measuring activated areas in the brain, and therefore is slow to respond to neural activity (seconds after the activity). Therefore, its time resolution is worse than EEG.

EEG can estimate the dynamic change in brain function in milliseconds, therefore it has a high temporal resolution. However, its spatial resolution is low, as it records signals at the scalp and cannot distinguish precisely which activated regions are being captured by an electrode. On the other hand, effective connectivity provides the correlation between two regions and at the same time identifies which affects the other. These properties are extremely effective in the diagnosis of neurological diseases such as epilepsy, where determining the origin of the seizure and how it spreads is paramount. However effective brain connectivity is formed either with fMRI or EEG data, and therefore suffers either the spatial or temporal resolution problem as functional brain connectivity.

The study of brain connectivity is the first step for specific methods aimed at diagnosing epilepsy and finding the appropriate methods with which to predict the occurrence of epileptic seizures. During such seizures, a high level of synchronisation among the regions of the brain is observed due to the abnormal firing of neurons. Calculating functional connectivity is the most appropriate way to capture this type of synchronisation, and the best and most usual technique to measure it, by virtue of its high temporal resolution, detecting changes in the order of milliseconds, is EEG (Noachtar & Rémi 2009). It also involves lower cost than fMRI and is also non-invasive. Apart from being cheaper, EEG is portable, so it can capture brain activity even outside the clinic, making it the logical choice for developing a prediction method. In the next section, EEG's measurement of the electrical activity of the brain is discussed in detail.

2.3 Electroencephalography (EEG) and brain waves

EEG is a non-invasive system to measure the electrical response of the brain. It uses electrodes to pick up the small electrical signals resulting from brain cells firing. These electrodes are metal discs (sensors) placed on the a subject's scalp, arranged in specific positions (in a placement system) to record the brain's signals and display them on a monitor as channels. The international 10-20 system is the standard approach to placement, whereby the electrodes are placed at a specific distance of 10% or 20% relative to four points: naison for the nose position; inion; and the two preauricular points at the ears (Figure 2.3) (Herwig et al. 2003). Each electrode bears a label consisting of a letter and number, the letter referring to the part of the brain (e.g. F denotes frontal), while an even number represents the right hemisphere of the brain and an odd number the left.

In 1875, Caton was the first to attempt to record brain activity (electrical signals) using a galvanometer. He did so by placing two electrodes on the scalp of each of a group of monkeys and rabbits (Caton 1875, Niedermeyer & da Silva 2005). Berger (1924) was the first to use an EEG machine on a human La Vaque (1999). In 1930, he started to use the approach with epilepsy patients to determine the patterns that distinguished

their seizure type, and then this technique of capturing brain activity became widely adopted, and was soon being used in the diagnosis of many neurological disorders. EEG is commonly used in diagnosing epileptic seizures, sleep disorders and behaviour change Noachtar & Rémi (2009). In addition to the conditions mentioned above, EEG is used to recognise any abnormal patterns in brain waves after injury or before the onset of a seizure.

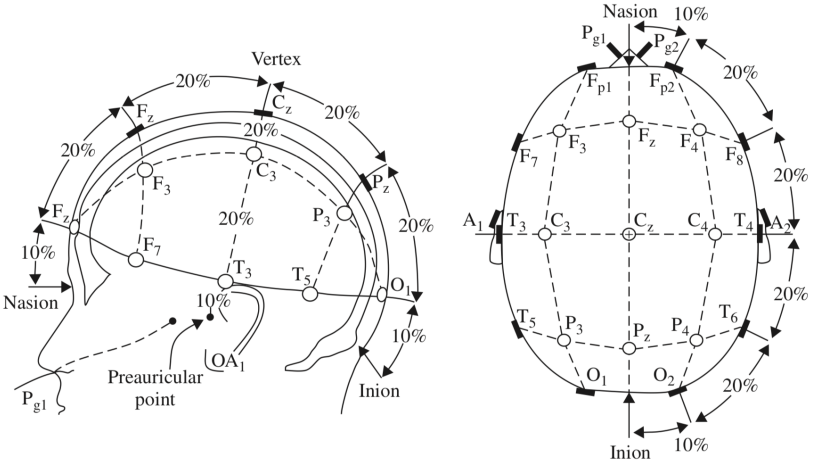


FIGURE 2.3: The International 10-20 System of Electrodes Replacement. It is illustration of the head from the left and above. The shortcuts of letters refer to the following; A, ear lobe; C, central; F, frontal; FP, frontal polar; O, occipital; P, parietal; Pg, nasopharyngeal. (redrawn from (Sanei & Chambers 2013)).

Since an EEG signal measured at the scalp is very small, of the order of μV, a differential amplifier is used to increase its amplitude before processing. The amplifier records the input of two electrical signals and identifies the difference between them. Next, a standardised display pattern called montage is used to choose which two electrodes should be connected to the same amplifier. There are three types of montage: common reference; average reference; and bipolar (Thatcher et al. 2004). In the first, a common electrode is used as the reference for all the electrodes, and will usually be one of the earlobe electrodes (A1, A2) or the Cz electrode at the centre, but this cannot be used if the participant is asleep because some artifacts waves can be seen in channel Cz such as saw tooth and sleep spindles which are distributed across all channels(Grigg-Damberger & Foldvary-Schaefer 2012). The second, the average reference montage, produces the average of all electrodes values (note: the researcher can choose which are to be included). In the third, the bipolar montage, the electrodes are linked together in a sequential pattern from the front to the back of the brain, or even transversely. Its disadvantage is that cancellation occurs when the activity potentials of the two electrodes are identical and, as a result, the channel will show as flat.

EEG has an amplitude of around 100 μV when measured at the scalp. Typically, EEG waves are divided into several frequency bands, as summarised in the following table.

EEG Band	Frequency Range (Hz)	Significance
Alpha	8 - 13	Appears in awake adults while meditating or in a state of relaxation
Beta	13 - 30	Appears during daily activities
Gamma	30 - 42	Related to emotions and feelings such as fear and anxiety
Delta	0.5 - 4	Observed during deep sleep
Theta	4 - 8	Observed in the state of drowsiness or deep meditation

TABLE 2.1: Brain waves used with their related frequencies range and significance

2.4 State-of-the-Art Methods for Epileptic Seizure Detection/Prediction

Several methods have been proposed to detect or predict seizures, using the various approaches in time, frequency or network domain.

2.4.1 Time Domain

Some research predict or detect seizures by measuring the change in certain features of EEG signals over time. They use either univariate measures, where a single EEG signal such as amplitude value is involved, or bivariate measures to check the relation between two signals, such as correlation. Webber et al. (1996) calculated the amplitude, slope and curvature of EEG for detecting seizures using an Artificial Neural Network (ANN) classifier. The method achieved 76% sensitivity with a false prediction rate of 1/h. Gabor (1998) used the spectral features of EEG, also with a self-organising neural network, and detected seizures with a sensitivity of 92.8% and a false prediction rate 1.35/h. Similarly, Alkan et al. (2005) later used these features for absence seizure detection, this time with two classifiers (ANN and logistic regression). Petersen et al. (2011) later used them with Support Vector Machine (SVM) to do the same by checking the spike of the wave, achieving 99.1% sensitivity and a false prediction rate of 0.5/h.Zandi et al. (2010) proposed a method exploiting the features of EEG itself. This involved calculating the positive zero-crossing for a data epoch. It was applied to 3 patients, and predicted 86% of seizures (sensitivity) in three patients with temporal lobe epilepsy, with an average prediction of onset of 20.8 minutes and a false prediction rate of 0.12/hr. The method had varying sensitivity when applied to three channels or five channels (85.7% and 71.43% respectively) of EEG recording. A seizure prediction method using a combination of five nonlinear measures (correlation dimension, correlation entropy, noise level, Lempel-Ziv complexity and largest Lyapunov exponent) to measure the independence

of two signals was proposed in (Aarabi & He 2012). It attained an average sensitivity of 79.9% and 90.2%, with an average false prediction rate of 0.17 and 0.11/h, respectively for two seizure prediction periods of 30 and 50 minutes each, tested with 11 patients. This method involves Intracranial Electroencephalography (iEEG), which requires an extensive surgical procedure. Iqbal et al. (2015) proposed a prediction method using two features (entropy and approximate entropy). They used a linear classifier to detect the pre-ictal (before seizure) and ictal (during seizure), and the predicted seizure onset ranged from 5 to 60 minutes, and the false detection rate from 0.38/h to 1.00/h. A method of prediction through detecting the spike rate was proposed in (Li et al. 2013). It calculates this rate using a morphology filter. It was applied to 21 patients and achieved a sensitivity of 75.8%, with an average false prediction rate of 0.09/h. It uses a univariate feature, and this is computed separately for each EEG channel. Although several univariate measures have been proposed for the purpose, bivariate measures have shown a more reliable seizure prediction result (Mormann et al. 2005). In addition, in Li et al. (2013) approach, the accuracy of prediction achieved was low. Le Van Quyen et al. (2001) used scalp-EEG/video recordings of 23 patients with temporal lobe epilepsy, checking 18 minutes before the onset of a seizure. Their method is based on measuring nonlinear similarity. It transfers the signal to the phase information, then calculates the cross-correlation between them. They achieved an average predicted time to onset of 7 minutes. Relatedly, Ouyang et al. (2007) proposed a method calculating the similarity of index based on the wavelet transform. A decrease in the similarity of index value was noted before a seizure. It achieved accuracy of 91%, with a prediction of onset of about 9 minutes. It performed better with data in the beta band (13–30 Hz).

2.4.2 Frequency Domain

The other approach to detecting or predicting a seizure uses the frequency domain. Such methods filter the signal from the time domain to the frequency domain using techniques such as Fourier transform (Rubinov & Sporns 2010). A great deal of signal information is hidden in the time domain, and extracting it is possible only by studying its frequency. Short-Time Fourier Transform (STFT) and wavelet transform are used to change the signal to a frequency domain. Wavelet transform appears to solve the problem of STFT, whereby the use of a fixed size of window affects the resolution of frequency (Sanei & Chambers 2013). The frequency domain was used by Pradhan et al. (1996) to detect seizures. They used wavelet transformation features with an ANN classifier with five patients and achieved 97% sensitivity, with a specificity of 89.5%. The frequency domain was also used in (Wang et al. 2013) in predicting seizures, this time using the Lyapunov exponent, where the correlation dimension was extracted from the wavelet transformation with the K-nearest Neighbour (KNN) classifier. It was applied to 10 patients and achieved 73% sensitivity, with 67% specificity. Hung et al. (2010) proposed a method to predict seizures for 11 patients with temporal lobe epilepsy, using

Discrete Wavelet Transformation (DWT) with the correlation dimension and coefficient. It achieved 87% sensitivity, with a false prediction rate of 0.24/h and an average of 27 minutes predicted onset time.

2.4.3 Network Domain

The third prediction approach is through the network domain. Its methods try to create a brain connectivity network using various connectivity measures. Functional brain connectivity have been used with EEG to create correlations, and coherence and synchronisation indices. The goal of the work in the network domain is exploring changes in the network structure in brain. He & Yang (2016) proposed a method to detect seizures in patients with temporal lobe epilepsy. They used a phase synchronisation index (phase locking value) to construct the functional connectivity network. The method explored the features of the network in order to detect seizures, and it was tested with patients with temporal lobe epilepsy. It showed the difference in the features' values at seizure, especially in data in the gamma band. However, the method is limited to detecting the onset of a seizure. Another study in (Cho et al. 2016) checked the value of PLV of the data in the gamma band, using Noise-Assisted Multivariate Empirical Mode Decomposition (NA-MEMD). NA-MEMD filtering algorithm was used to decompose spectral components from the data, and 21 epileptic children were tested by this method. Two periods (preitcal and interictal which are the period before and between seizures respectively) were then classified using the SVM classifier, achieving a rate of 83.17%. The highest PLV was achieved with NA-MEMD, which is considered to be an indicator that predicts seizures. Graph theory forms the basis of the method for prediction proposed by Haddad et al. (2014), who used the high correlation between EEG electrodes in voltage peaks in the delta band. It was applied to 12 patients with temporal lobe epilepsy and achieved a detection accuracy of 72%, with a false prediction rate of zero and 30 minutes prediction of onset. Mormann et al. (2000) used phase synchronisation with patients with temporal lobe epilepsy. It was noted that synchronisation appeared to decrease prior to a seizure. This study only explored the changes in the phase synchronisation immediately before seizure and during those periods free of seizures, without making any predictions. However, it serves as an indicator of the significance of using phase synchronisation in predicting seizures.

The approach of network was selected in the thesis to develop a predictive model using functional brain conectivity networks. Table 2.2 illustratres the current methods for prediction within the three differnt domains. It provides a comparison through providing the performance of these methods with the patients number and epilepsy type used. Additionally, it illustrates the maximum predicting time achieved by each method.

Reference	Method	Epilepsy Type	No.of Patients Used	Description
Zandi et al. (2010)	Positive zero-crossing	Tmporal Lobe	3	20.8 mins as prediction time, sensitivity of 86% and FPR 0.12/h
Aarabi & He (2012)	Combination of linear measures	Focal neocortical epilepsy	11	Prediction time 50 mins, sensitivity 79.9%, and FPR 0.17/h
Iqbal et al. (2015)	Entropy and approximate Entropy	Intractable seizures	4	Prediction time 60 mins, sensitivity 100% and FPR 0.38/h
Li et al. (2013)	Spike rate	Intractable focal epilepsy	21	Prediction time 49.7 mins, Sensitivity 75.8% and FPR 0.09/h
Le Van Quyen et al. (2001)	Nonlinear similarity	Temporal Lobe	23	Prediction time 7 mins
Ouyang et al. (2007)	Similarity index	Temporal Lobe	16	Prediction time 9 mins and sensitivity 91%
Cho et al. (2016)	NA-MEMD with Phase Locking Value	Undefined	21	Prediction time 5 mins, sensitivity 82.44% and FPR 0.5/h
Chu et al. (2017)	Attractor state analysis	Undefined	16	Prediction time 45.3 mins, sensitivity 86.67% and FPR 0.367/h
Myers et al. (2016)	Phase and Amplitude Lock Values	Undefined	10	Sensitivity 76.8% and FPR 0.167/h
Usman et al. (2017)	Machine-learning methods	Undefined	22	Prediction time 23.6 mins and sensitivity 92.23%
Alotaiby et al. (2017)	CSP and LDA	Intractable seizures	24	Prediction time 68.71 mins, sensitivity 89% and FPR 0.39/h

TABLE 2.2: Comparison of the current methods of predicting epileptic seizures.

2.5 Functional Brain Connectivity and Prediction

Existing methods to predict seizures take various approaches (time domain, frequency domain and network domain), but they all have limitations (see Table 2.2). They are

limited to the prediction of a single type of epilepsy and have low sensitivity. Therefore, a proposed method has been developed to predict seizures using functional brain connectivity networks. Three approaches - PLV, PLI, and SL - used to measure the functional brain connectivity. A description of these three approaches is provided in the following subsections.

2.5.1 Theory of PLV

PLV is one of the techniques employed to extract the features of functional brain connectivity by estimating the long-range phase synchronisation between two signals – in the case of EEG, signals captured at multiple electrodes (Lachaux et al. 1999),(Bola & Sabel 2015). First, both pre-processed EEG signals undergo the Hilbert transform to calculate their instantaneous phase values, which are subsequently used to establish the phase difference between them, typically expressed within the range of $-\pi$ to π. Mathematically, PLV is expressed as in (Equation 2.1) (Bola & Sabel 2015).

$$PLV_t = \frac{1}{N}\sum_{n=1}^{N}|exp(j\theta(t,n))| \tag{2.1}$$

where, $\theta(t,n)$ represents the phase difference between the two signals at time t, and n denotes the n^{th} trial(the sample number where within each window). PLV value ranges from 0 to 1, signifying the range from no synchronisation to perfect synchronisation. As PLV outputs measures related to time, using it to formulate functional brain connectivity enables the researcher to analyse its temporal behaviour, which may result in important features for predicting the onset of an impending seizure. However, a fundamental problem with PLV is the volume conduction effect, which needs to be considered during functional connectivity network formulation.

To explain the volume conduction problem, it is has to know how EEG signals are extracted by the electrodes. EEG can record the dynamic change in brain functions in milliseconds, and therefore has a high temporal resolution. However, because it records the signals at the scalp and they pass through multiple layers with different insulator current, it has low spatial resolution and cannot distinguish where the recorded signals are being transmitted from. Figure 2.4 shows the nature of propagating electrical signals from their source through biological tissue. The tissue acts as conductor or insulator, so a signal can either be propagated through or be reflected out in another direction. This leads to the problem whereby several electrodes may record a different signal from the same original impulse. The other challenge is that the amplitude of the electrical signal decreases with distance from the source of the signal. For more details about PLV, see (Lachaux et al. 1999),(Niso et al. 2013).

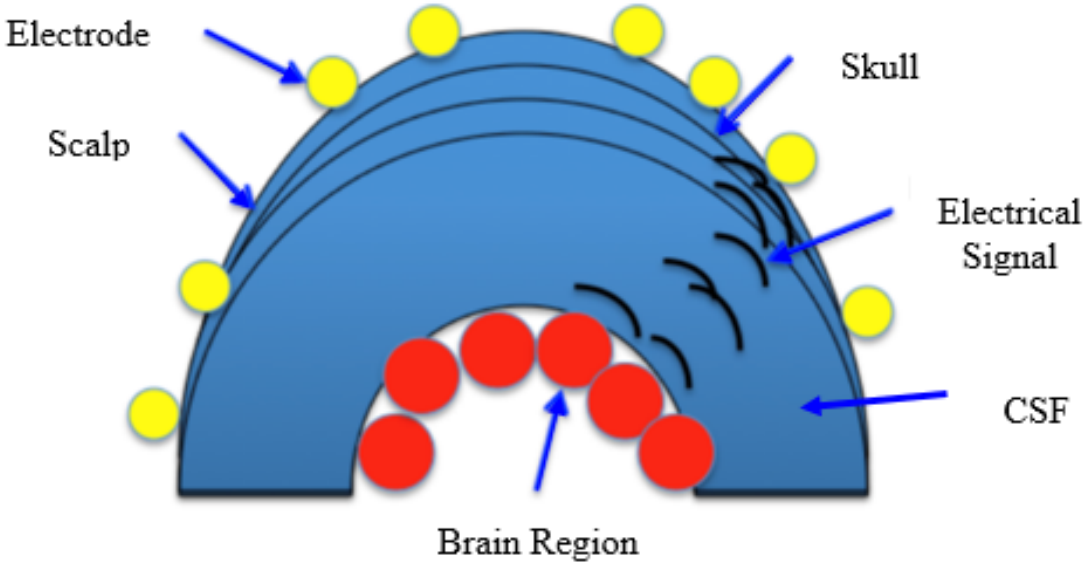


FIGURE 2.4: Propagation of signal from source to electrode. CSF refers to the Cerebrospinal fluid found in the brain.

2.5.2 Theory of PLI

PLI was created by Stam & Reijneveld (2007) to handle these problems of volume conduction and active reference electrode. It is a measure of the asymmetry in the distribution of the phase difference between two signals captured from the same source (common source), which leads to fake correlation. Several active reference electrodes may share the problem and thus provide further fake correlations (Stam & Reijneveld 2007). The results after PLI are less affected by either problem.

PLI assumes that there is a systematic, nonzero phase lag that reflects the real interaction between two brain sources and cannot be interpreted by volume conduction. This phase lag can be found by using the asymmetry of the distribution of the phase difference of the two signals. If the likelihood of phase difference distribution in the interval $[0, \pi]$ is equal to the likelihood of the interval $[0, -\pi]$, which means symmetry distribution, then the phase difference will be eliminated. This is what is meant by phase difference centred around 0 mod π. Following is the equation used to calculate PLI value (Stam et al. 2007)

$$PLI = |\langle sign\left[\Delta\theta\left(t_k\right)\right]\rangle| \tag{2.2}$$

$\Delta\theta\left(t_k\right)$ is phase difference of the time series, where $k = 1, ..., N$ is the sample number within each window and N is total number of these samples inside each window (the sliding window used to calculate the connectivity network). PLI values ranging from 0 to 1. One indicates a phase locking value with a phase difference other than 0 mod π.

Zero refers to either no coupling or coupling with phase difference of around 0 mod π. Using different montages could be slightly affect the value of PLI.

2.5.3 Theory of SL

From the various generalised measures available, temporal SL is most appropriate to build a prediction method. It is considered to be the most commonly used with neurological data (Pijnenburg et al. 2004, Ahmadlou et al. 2012, Rosales et al. 2015). Several indexes of generalised synchronisation have been proposed, as in (Arnhold et al. 1999), which relies on the change in the average distance between points from their mutual and nearest neighbours. This has succeeded in detecting nonlinear dependency between signals. However, detecting asymmetry interdependency is less easy for this index, as it is sensitive to signal amplitude, thus any measurement of asymmetry may reflect the dissimilar degrees of freedom in the two systems used (Arnhold et al. 1999). SL has been proposed to overcome this problem (Stam & Van Dijk 2002). It relies on detecting a common pattern occurring simultaneously in two signals. In other words, if one signal has a pattern of activity, another signal may tend to repeat that pattern at the same time, so SL measures how strongly one signal (channel) is synchronised at a specific time to all other signals (channels) (Montez et al. 2006). The equations of SL for two points i and j for Channel k is defined as follows:

$$S_{k,i} = \frac{H_{i,j} - 1}{M - 1} \tag{2.3}$$

Where $H_{i,j}$ is the number of channels for closer embedded vectors and M is total number of channels. If the Euclidean distance $|X_{k,i} - X_{k,j}|$ ($X_{k,i}$ is the embedded vector created for channel k at time i) is equal or greater than the critical distance $\varepsilon_{k,i}$ then the synchronisation likelihood ($S_{k,i,j}$) equal zero; otherwise the synchronization likelihood is calculated from the following equation (Niso et al. 2013).

$$S_{k,i} = \frac{1}{2\,(w_2 - w_1)} \sum_{j=1}^{N} S_{k,i,j} \tag{2.4}$$

where $w_1 < |i - j| < w_2$. This equation measures the SL for Channel k at time point i.

2.6 Graph-theoretic Features of Functional Connectivity Networks

Once a functional connectivity network is formulated, using PLV, PLI or SL , it can be represented in the form of a graphical network in which each electrode represents a node and the edges between two nodes represent the strength of their connection. The

essential characteristics of such graph networks may be analysed using several graph-theoretic measures (features) to understand their structural and temporal properties. These features mainly fall into two categories: network aggregate, such as clustering coefficient, transitivity, modularity, characteristic path length, and so on, capturing the properties of the entire network; and node-specific features, such as betweenness centrality, strength of node, and so on, capturing the connectivity properties of individual nodes (Rubinov & Sporns 2010). In our study to develop a predictive model for patients with TLE and Absence seizures, we concentrated on the set of features described below (Rubinov & Sporns 2010). Table 2.3 illustrates describes the network features used in the thesis with their mathematical equations.

Graph theory measure	**Type**	**Equation**	**Description**
Characteristic path length	Network Aggregation	$L = \frac{1}{n}\sum_{i\in N} L_i = \frac{1}{n}\sum_{i\in N} \frac{\sum_{j\in N, j\neq i} d_{ij}}{n-1}$	Measures the average shortest paths between all node pairs in the graph
Global efficiency	Network Aggregation	$E = \frac{1}{n}\sum_{i\in N} E_i = \frac{1}{n}\sum_{i\in N} \frac{\sum_{j\in N, j\neq i} d_{ij}^{-1}}{n-1}$	Presents the average inverse of all shortest paths between graph nodes
Transitivity	Network Aggregation	$T = \frac{\sum_{i\in N} 2t_i}{\sum_{i\in N} k_i(k_i-1)}$	Measures the fraction (probability) of the number of connections composed of three nodes (triangular connection) compared to the total number of triplet connections of nodes in the network
Modularity	Netowrk Aggregation	$Q = \sum_{u\in M}\left[e_{uu}\left(\sum_{u\in M} e_{uv}\right)^2\right]$	Shows the strength of a network and is divided into groups (modules)
Assortativity Coefficient	Node-specific	$r = \frac{l^{-1}\sum_{(i,j)\in L)} k_i k_j - \left[\sum_{(i,j)\in L} \frac{1}{n}(k_i+k_j)\right]^2}{l^{-1}\sum_{(i,j)\in L}(k_i^2+k_j^2) - \left[\sum_{(i,j)\in L} \frac{1}{n}(k_i+k_j)\right]^2}$	Measures the trend of each node in terms of how it connects with nodes of the same degree
Betweenness centrality	Node-specific	$b_i = \frac{1}{(n-1)(n-2)}\sum_{(h,j)\in N, h\neq i, h\neq j, i\neq j} \frac{p_{hj}(i)}{p_{hj}}$	Measures the importance or influence of the node in the network
Strength of node	Node-specific		Measured by calculating the total weights of the connected edges linked to the node
Clustering coefficient	Netowrk Aggregation	$C = \frac{1}{n}\sum_{i\in N} \frac{2t_i}{k_i(k_i-1)}$	It measures the tendency of nodes to be clustered together
Raduis	Node-specific		the minimal eccentricity (where eccentricity is maximum distance of one node from other node).

TABLE 2.3: Where n represents the number of node, d_{ij}represents the shortest path from$i$ to j, t_i represent the number of triangle around node i and k_i represents the degree of node i ,e_{uv} represents the propotion of all links that connect nodes in module u with nodes in module v. p_{hj} is number of shortest paths between h and i while $p_{hj}(i)$ is number of shortest paths between h and j passes through i.

2.7 Machine Learning Framework

Machine learning is the part of artificial intelligence responsible for the significant development in most new technologies such as robots and self-driving cars (Cortes & Vapnik 1995). It has also been applied to the prediction of epilepsy. Amongst the services it provides is pattern recognition and decision-making. It uses algorithms to learn from the incoming data and make decisions. Machine learning can be divided into supervised learning, such as a classification algorithm, and unsupervised learning, such as clustering, in which the algorithm tries to find hidden patterns. K-means and k-medoids are the two clustering algorithms used in this current work. By contrast, classification algorithms learn from the input data provided with class labels. support-vector networks (SVM) is widely used in two-group classification problems. Each data point in the input, which is considered a training set for the algorithm, is assigned to one of the classes. Table 2.4 provides a description of all the classifiers used in this study thesis.

Classifier Name	**Description**
k-nearest Neighbour (KNN)	Used for classification and regression. It depends on finding the nearest k-neighbour to a selected data point, so if k=2 then the selected point will be within the class of the two nearest neighbours. It is easy to implement and can solve problems of multiclass.
Decision Tree	It is fast and can handle large sets of data. It works by splitting the set of input features using the tree model. Complex Tree was used in the current work.
Random Forest	It works as an ensemble of Decision Trees. It selects random subsets of input features for the decision tree. A prediction is made by averaging the predictions of each decision tree.
Bagged Tree	It is an ensemble of Fine Decision Tree. It is often very accurate.
Support-Vector Machine (SVM)	It uses a hyperplane to separate input data points by maximising the margin between the classes using the closet points (support vectors) from each class. One of its advantages is its resistance to overfitting, and it can work with a small quantity of data. Linear SVM is used in the current work, which is linear separation.
Logistic Regression	It uses a statistical technique. It is a binary classification that works by calculating the probability in relation of the input features to the target class. It is easy to implement.
Linear Discriminant Analysis (LDA)	It works by creating linear boundaries between classes by calculating the probability of the input with each class. It works with multiple classes.

Table 2.4: Description of all the classifiers used in this study thesis

2.8 Conclusion

In this chapter, different brain connectivity approaches used in studying the brain activity have been illustrated. Because of the high level of synchronisation noticed during seizure, Functional brain connectivity was the most appropriate way to measure it. Among the functional brain connectivity modalities used - fMRI and EEG - using EEG was the best technique to measure the synchronisation owing to its high time resolution. Although, the plenty of predicting methods have been developed using EEG, they have limitations represented in low senstivity and high false prediction rate (see Table 2.2). In this thesis, functional connectivity with the graph theory have been used to develop a model that can predict seizures with high sensitivity and low false prediction rate with maximizing the prediction time.

Chapter 3

Research Methodology

As has been discussed in the previous chapter, there were several prediction methods developed. However, none of them predicts all types of epileptic seizures with acceptable accuracy. This is due to the difference between the types of epilepsy in terms of their origin and the duration of the epileptic seizures. In each type, brain activity varies in terms of how the brain regions interact for manifested symptoms. Nonetheless, focusing on one type of epilepsy may give a significant chance of building a robust prediction method and then it could be generalized for other types of epilepsy. In this work, we propose models to predict seizures for subjects with Temporal Lobe Epilepsy (TLE) and with absence seizures. This chapter presents the research methodology adopted in this work. In essence our methodology is based on finding temporal variation in the functional brain connectivity networks that serve as features for building predictive models in a machine-learning framework. Connectivity networks are formed using Phase Locking Value (PLV), Phase Lag Index (PLI) and Synchronization Likelihood (SL) for data in each frequency band of the recorded EEG signals. Each of the three approaches listed above to be used in building functional networks was applied to develop models for predicting seizures separately. Finally, a general model using the network features for predicting the epileptic seizures in the three approaches has been formulated.

3.1 Overview of Methodology

The proposed methods are based on creating functional brain connectivity networks by calculating the synchronization between the EEG signals captured at different EEG electrodes. A pre-processing step has been applied to the acquired EEG signals in order to obtain clear data before calculating the functional connectivity networks. Artefacts and noise have been separated from the data before conducting any analysis. The values of connectivity represent the weighted edges of the created undirected networks, while each EEG electrode, which captures electrical activity in a brain region, represents a

node of the networks. The dynamic change in the value of the functional connectivity is tracked over time by analysing the change in the values of the network features. A machine-learning framework is used to divide data from extracted network features into groups that share a specific pattern of network feature values in an attempt to find a common pattern (group) close to an epileptic seizure shared among the epileptic subjects. Figure 3.1 illustrates the procedural steps of the proposed method which is described in detail in the following subsections.

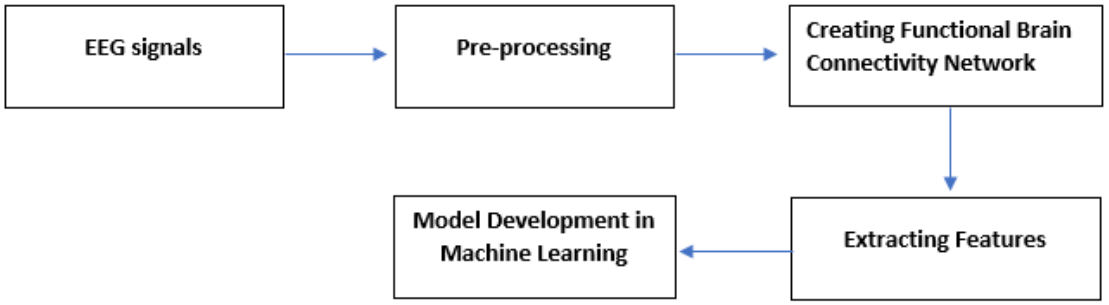


FIGURE 3.1: The procedural steps of the methodology

3.1.1 Data Collection

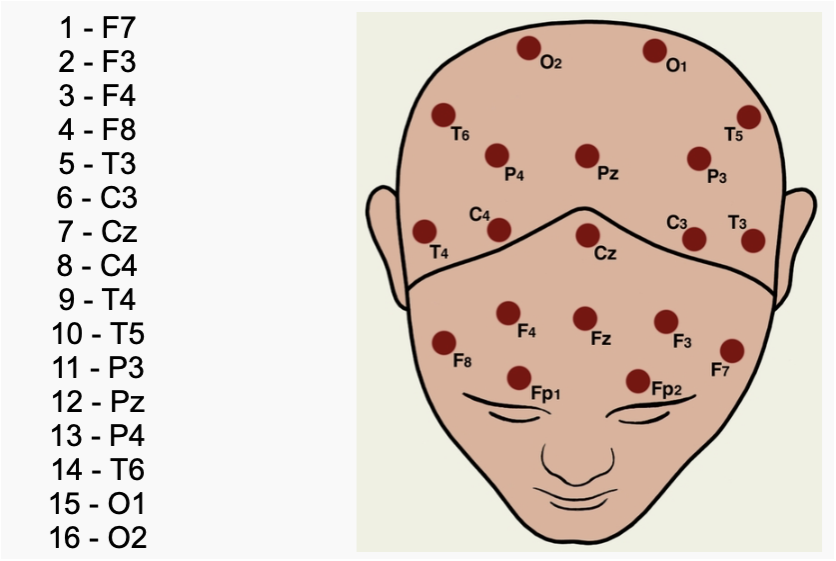


FIGURE 3.2: Electrode replacement and the order of the 16 channels of the 10 healthy subjects (redrawn from (EEG.Healthy.Schizophrenia 2019))

24 EEG files from 10 adult TLE surgical subjects (right hemisphere origin) were used in conjunction with 10 healthy subjects (each resulting in 1 EEG data file). The data were sourced by Southampton University Hospital NHS Trust (SUHT), UK and Fondazione Stella Maris, Pisa, Italy. The data were collected by the clinicians from the respective

hospitals as part of regular clinical follow-up, properly anonymized and shared for secondary analysis only. The data was recorded at a sampling rate of 200 Hz using 20 channels international 10-20 EEG system. The data for the 10 healthy subjects were obtained from Neurophysiology and Neuro database (EEG.Healthy.Schizophrenia 2019) that was recorded with 16 channels (F7, F3, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) at a sampling rate of 128 Hz (see Figure 3.2). The seizure recording times of the 10 TLE subjects are shown in the Table 3.1. The table illustrates the recording time of 24 seizure onsets of 10 TLE subjects from the start of data recording. A training set of 12 EEG data – six EEG data belonging to four TLE and six EEG data from healthy subjects - was selected as a training set to formulate the predictive models. The six EEG data of TLE were selected because they have the most common longest recording time before seizure, which is 5 minutes.

Patient no.	**Seizure no.**	**Recording time before seizure (seconds)**
1	1	52.9
	2	24.1
2	1	75.1
	2	70.6
3	1	164.1
	2	44.1
4	1	1053.8
	2	1444.4
	3	739.5
5	1	75.4
	2	1801
6	1	44
	2	81
	3	130
7	1	77
	2	20
8	1	58.3
	2	29.1
	3	26
	4	30
	5	21
9	1	424
	2	549
10	1	344

TABLE 3.1: Illustration of 24 Seizures Time of 10 TLE Subjects from Start of EEG Data

On the other side, 32 EEG data belonging to nine patients with absence seizures have been used. Two of the nine patients had no recorded seizures. The EEG data were recorded using 22 channels for five patients and 21 for four patients at sampling rate of 200 Hz. Only 19 channels that are in common to all patients were used, as shown in Figure 3.3.

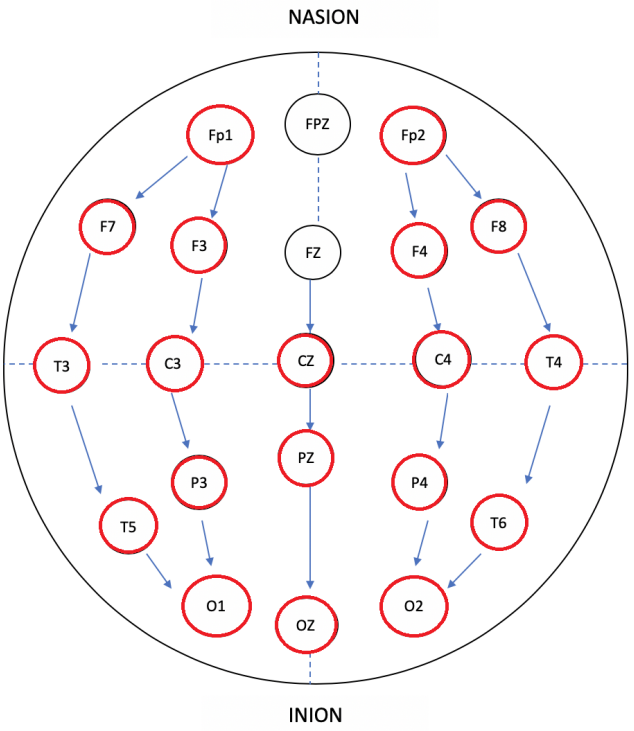


FIGURE 3.3: The 19 channels used with all patients; They are created as bipolar using the electrodes shown with red circles

The number of seizures detected for each of the seven patients and their minimum and maximum recording time before seizure onset are shown in the Table 3.2. The training set was divided into two groups to develop a predictive model. Group 1 contains nine recording of EEG data of 6 patients while Group 2 has two data with no seizures recorded. Four minutes of recording time before the seizures was selected to search for a common pattern. Group 2 of the training set was used along with Group 1 to provide a comparison regarding the possible abnormal change that might have happened at any time before the seizure onset.

Patient no.	**No. of Seizures Used**	**Minimum and Maximum Recording Time Before Seizure Onset(seconds)**
1	5	100 - 929
2	6	201 - 980
3	2	572 - 1225
4	1	67
5	6	247 - 2519
6	7	133 - 800
7	3	192 - 429

TABLE 3.2: Minimum and maximum recording time before seizure onset for 7 patients with Absence Seizure. The column three refers to the minimum and maximum time of all seizures of each patient

Table 3.3 summaries the amount of data used for each epilepsy type with thier number of seizures. It illustrates the amount of data analysed before seizures and how much data have been used as training set.

Epilepsy Type	**Total Data Analysed (hours)**	**Data Analysed before Seizure (hours)**	**Data Used as Training (hours)**	**Number of Seizures Used**
TLE	3.9	2	1	24
Absence seizures	16.3	5	3.1	30

TABLE 3.3: Summary of the amount of data used in the thesis

3.1.2 Pre-processing Data

Pre-processing of the EEG data has been done in MATLAB to remove unwanted noise and artefacts. A 6^{th} order Finite Impulse Response (FIR) notch filter has been used to remove 50Hz power line noise and then the signal was low-pass filtered with cut-off frequency 100Hz to remove high frequency noise. Artefacts were removed using Independent Component Analysis (ICA). The resulting EEG signals were then re-referenced using average referencing approach. The re-referenced EEGs were then broken down into canonical frequency bands - alpha, beta, gamma, delta and theta ($\alpha, \beta, \gamma, \delta$, and θ respectively) using band pass filtering in the range of [8 13] Hz, [13 30] Hz, [30 42] Hz, and [0.5 4] Hz, [4 8] Hz respectively. Figure 3-4 illustrates the order of pre-processing steps applied to EEG data. Next, steps applied to pre-processing EEG data will be detailed.

3.1.2.1 Artefacts and Noise Separation

First, two types of noises were removed. These noises are power line noise at 50Hz and high frequency noise. The power line noise was removed by using a FIR notch filter of 6^{th} order. The other noise to be removed is high frequency. The upper limit of the frequencies of the human brain waves is 100Hz. Therefore, any frequency higher than 100Hz represents a noise caused by sources of signal other than brain activity. Low-pass FIR filter with cut-off frequency of 99Hz was applied to remove this noise.

The second noise that needs to be removed is that from artefacts. The process of artefact removing involves trying to separate the artefacts without distorting the actual EEG signal. EEG signals captured for brain activity are mixed with other electrical activities that are derived from different non-brain sources. These non-cerebral-origin electrical activities are considered artefacts. Such artefacts could stem from inside the body of the patient (cardiac, muscular and ocular artefacts) or outside the body (equipment).

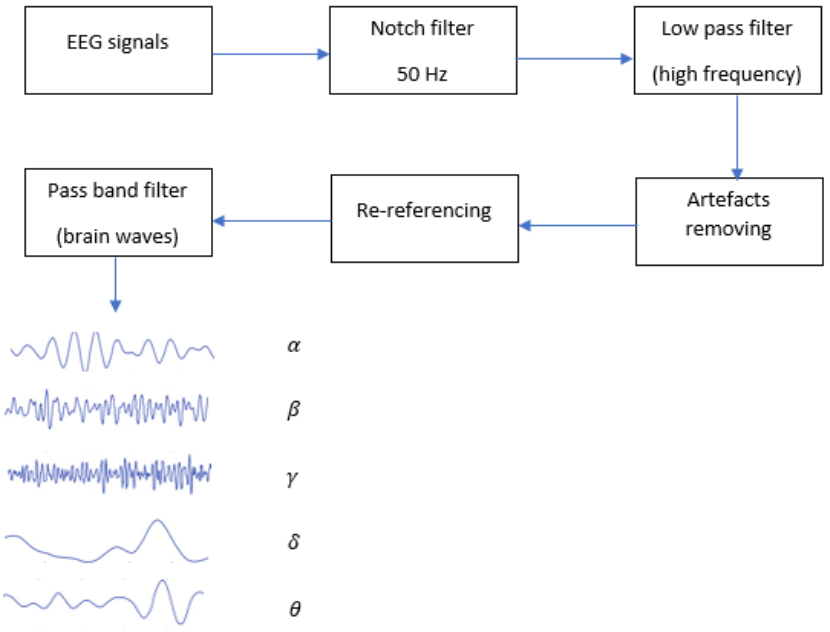
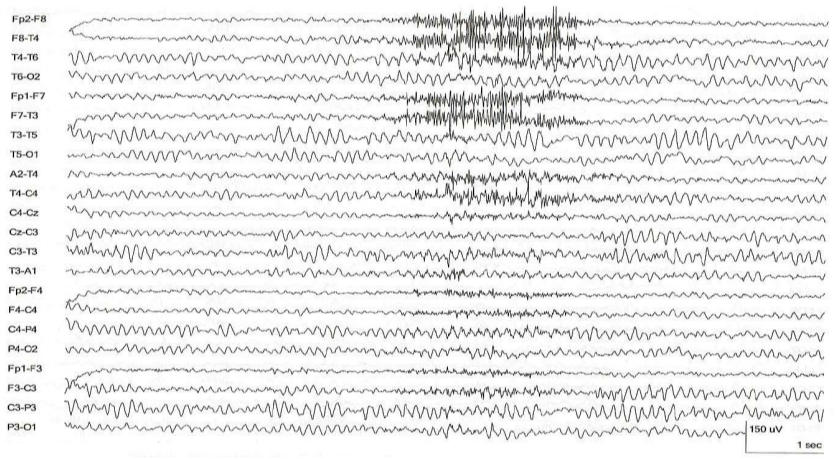


FIGURE 3.4: Pre-processing steps applied to EEG signals

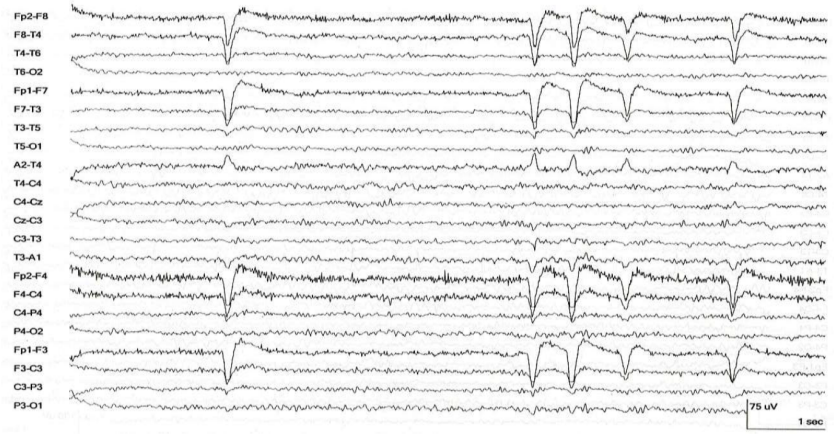
Figure 3.5 presents three types of artefacts: (a) muscular artefacts (EMG); (b) and (c) eye-blink and lateral eye-movement artefacts (EOG), respectively. Eye blinking produces ocular artefacts because of the rapid movement. It appears as a slow wave. Meanwhile, lateral eye-movement artefacts originate because of the horizontal movement of the eye, which can be noticed in the frontal electrodes such as F7 (odd numbers represent the left side of the brain).

Over the past years, many algorithms and methods have been proposed to remove all types of artefacts, such as ICA. In general, all algorithms apply either filtering and regression or decomposing (separating) data. The drawback of the algorithms using regression fail if there is no reference channel, such as a channel recording muscular movement (Islam et al. 2016). On the other hand, decomposing separates EEG data into components and then removes artefacts from each of the components. Blind source separation (BSS) is an example of a separation method, which is considered one of the most common methods used (Urigüen & Garcia-Zapirain 2015). It is used to divide data into spatial components. Figure 3.6 illustrates the main principles of BSS work.

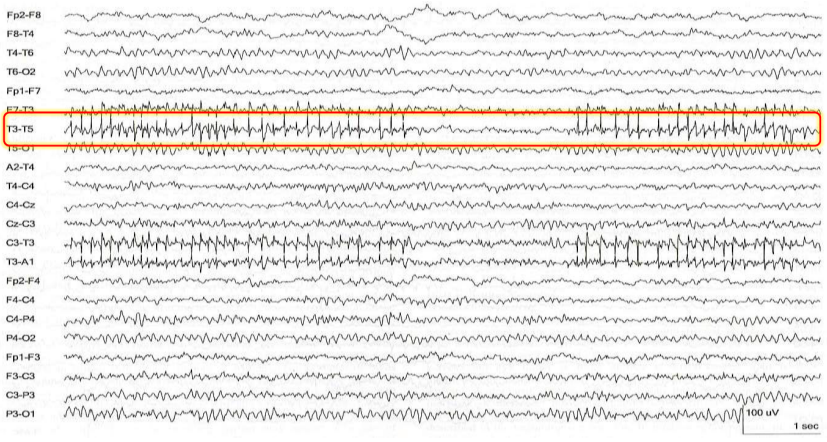
ICA is the BSS method mostly used by researchers to separate artefacts (Naik & Kumar 2011). ICA separates multichannel signals into independent components, assuming that these components are non-Gaussian signals and statistically independent (Sweeney et al. 2012). It works by seeking maximally independent sources (Crespo-Garcia et al. 2008). Equation 3.1 illustrates the general model of ICA as obseved data or matrix decomposition (Coben & Evans 2010, Lu et al. 2019).



(a)



(b)



(c)

FIGURE 3.5: Examples of three types of artefacts. (a) muscular artefacts. (b) eye blink artefacts (c) lateral eye movement artefacts(redrawn from (Marella 2012))

$$X = AS \quad where \quad x_i(t) = \sum_{i=1}^{n} a_i s_i(t) \tag{3.1}$$

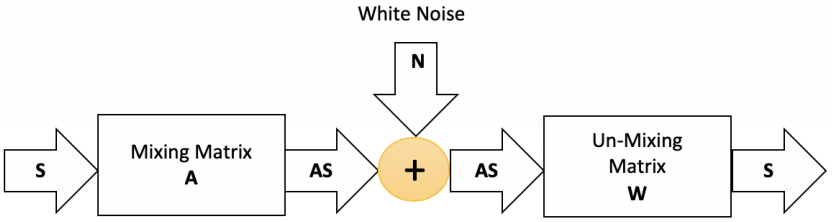


FIGURE 3.6: Illustration of the technique of blind source separation, where S represents a linear mixture of sources, N the white noise, and W an uncorrelated mixture of components estimated from a mixture of A sources (redrawn from (Islam et al. 2016))

where $i = 1, ..., n$ (n is the total number of signals) and $s_i(t)$ represents the statistically independent i^{th} signal at time t. $x(t)$ is the observed signals which are mixture of sources S (different signals representing the brain activity with other artefact sources) captured by electrodes. Each electrode has a different mixture of sources. A in the equation represents the unknown mixing matrix (a BSS principle that estimates the sources without information of source signals). ICA decomposes signals by finding the maximally independent sources of activity in the brain (eye movement and muscular artefacts are examples of these activities, two sources said to be statistically independent if each source does not provide information on the other) (Naik & Kumar 2011).

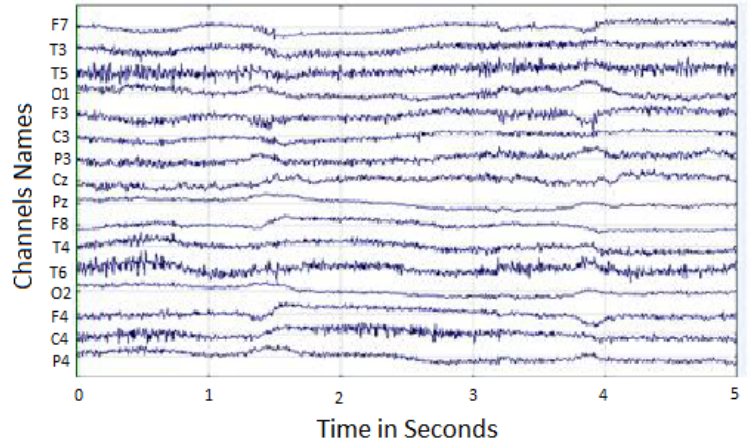
In the present work, Principal Component Analysis (PCA), a decomposing method, has been used to make a comparison with ICA to assess the efficiency of the ICA method. PCA uses orthogonal transformation to extract uncorrelated components (principal components). Its separation method is based on maximising the variance of the orthogonal transformation, in contrast to ICA (which maximises non-Gaussian (normality)) (Tibaduiza et al. 2012). Figure 3.7, (a) represents EEG signals with artefacts, while (b) is a representation of the clear signals, showings how clear a signal is after applying the ICA method. The Y-axis shows the channel names while the x-axis shows the time in seconds. FastICA has been used to apply ICA. It is a popular ICA algorithm characterised by finding the maximally non-Gaussian components (Hyvarinen 1999).

For comparison Figure 3.8 shows the application of the PCA method to the same EEG signals. It can be seen in (b) that EEG signals still contain artefacts compared to ICA. Therefore, ICA has been selected to remove artefacts in the proposed method.

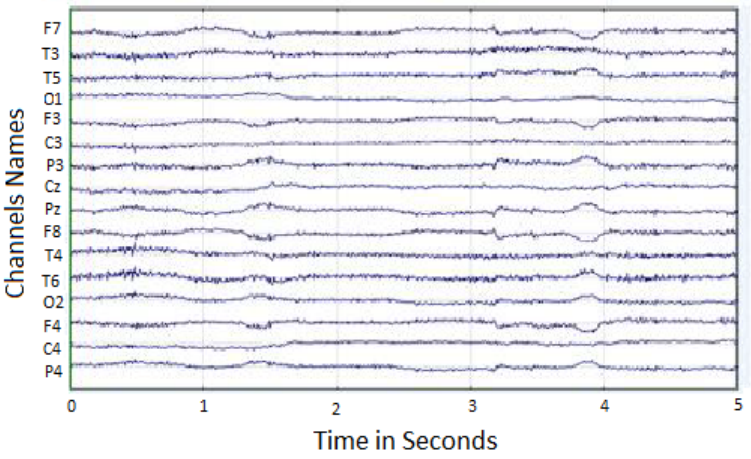
At this point, the EEG signals have been cleaned of noise and artefacts. As a next step, the clean EEG signals will be re-referenced and then filtered into different brain waves.

3.1.2.2 Re-referencing EEG Data

EEG scalp recording uses the technique of a differential amplifier in producing channels. It uses the difference between two values of two electrodes to produce channels. One of



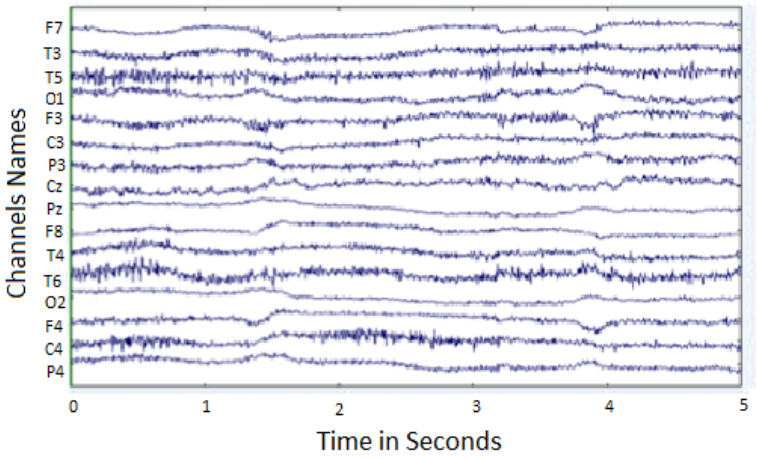
(a) EEG Signals with Artefacts



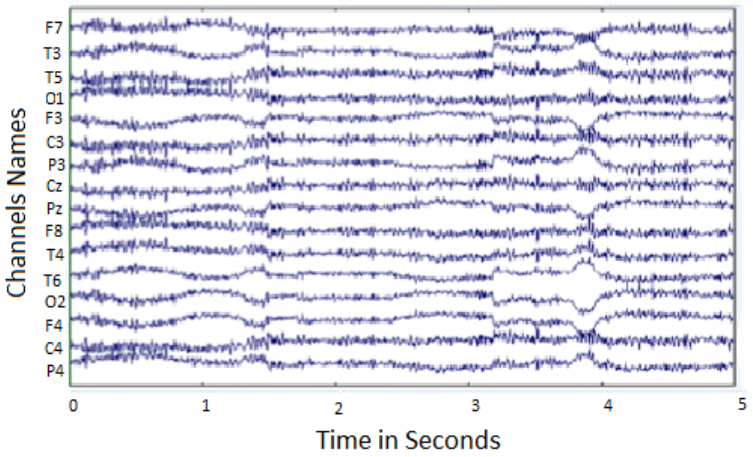
(b) EEG Signals without Artefacts

FIGURE 3.7: Illustration of artefact removal using ICA. (a) EEG signal with artefacts (b) clear EEG after ICA was applied

the two electrodes is an active electrode and the other is the reference. Choosing the best reference depends on the purpose of analysis. The reference has a direct impact on the functional network studies (Huang et al. 2017). Different examples of references include the left mastoid reference (LR), which uses the right earlobe as a reference, and linked mastoids(LM), which is an average of both earlobes. The Average Reference (AR) is another choice of reference, which concerns the average value of all electrodes. Using LR as a reference reduces the amplitude of the signals captured at electrodes, especially those which are close to the reference, while using the linked earlobes will avoid any asymmetry produced by the LR reference. The AR avoids the asymmetry from LR and LM references. AR is the best choice with respect to analysis based on correlation or synchronisation, as it will avoid any unwanted correlation made by the reference activity (Dien 1998). As a result, the common average was applied to EEG data. All pre-processing steps have been carried out using MATLAB as a framework, which provides an integrated environment for all functions needed to analyse EEG data.



(a) EEG Signals with Artefacts



(b) EEG Signals without Artefacts

FIGURE 3.8: Illustration of artefact removal using PCA. (a) EEG signal with artefacts (b) clear EEG after PCA was applied

This, in turn, facilitates the process of analysis.

After re-referencing step, the signal was filtered into five different frequency bands. FIR band pass filter was applied to EEG signals using pass band edges of [8 13] Hz, [13 30] Hz, [30 42] Hz and [0.5 4] Hz, and [4 8]Hz with cut-offs frequency [7 14] Hz, [12 31] Hz, [29 43] Hz, [0.4 5] Hz, and [3 9] Hz for alpha, beta, gamma, delta and theta respectively. Filter order was automatic assigned by EEGLAB.

3.1.3 Network creation and parameters extraction

Most of the current methods proposed to predict seizures use univariate measure while some use bivariate measures. Regardless of which one can predict better (Yadollahpour & Jalilifar 2015), it is known that the most significant phenomenon noticed during seizure is the high level of uncontrolled synchronization between brain regions. As a

result, detecting this synchronization may provide a better predictive result. Here, undirected weighted functional brain connectivity networks have been constructed using PLV, PLI and SL at each frequency band using 1s sliding window from the start of the data to the point of clinical onset of seizures. In these created networks, each EEG channel represents a region of the brain wherein electrical signals are captured. These brain regions represent the nodes of the network. Meanwhile, the value of calculated functional connectivity represents the weighted edges of the network. Figure 3.9 shows an example of creating undirected weighted network of PLV using 21 channels.

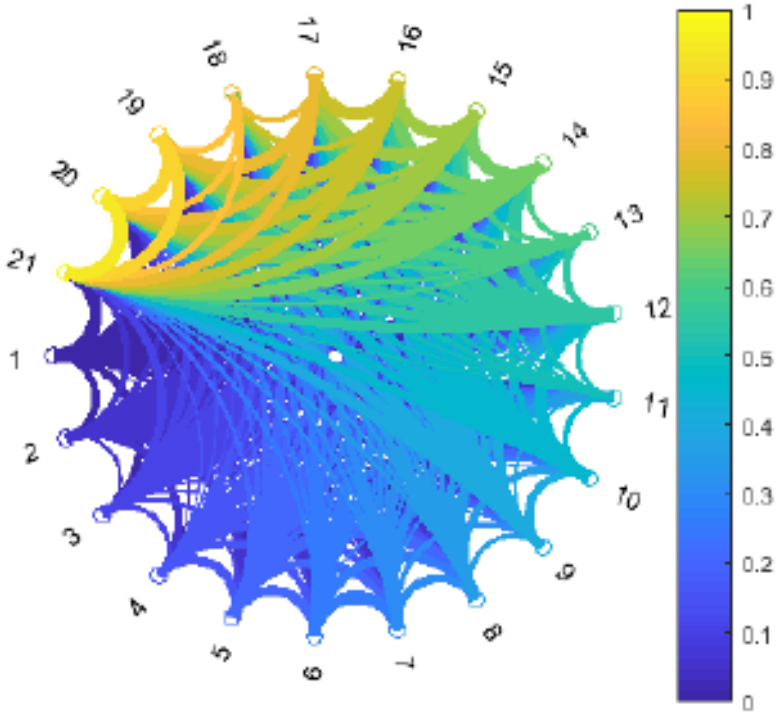


FIGURE 3.9: Functional connectivity created using phase locking value (PLV) for 21 channels. The colour bar illustrates how strong PLV value is

Subsequently, the created networks are characterised using a set of graph-theoretic features, such as described in Chapter 2. They are used to characterise the dynamic change in the brain network structure, wherein the prediction of epileptic seizures is made based on the variation in the network structure. Some of the graph-theoretic parameters describe the property of the entire network, such as characteristic path length, global efficiency, and transitivity. Meanwhile, other parameters describe a specific node in the network, such as betweenness centrality. For more details of the graph-theoretic parameters and how they are calculated see Table 2.3.

The big question is how the temporal variation of connectivity networks is captured. A sliding-window technique is used to find the temporal variation. This approach is used with most detection and prediction problems to find out features of continuous data (Esteller et al. 2001). In the proposed method, a sliding window of 1-second

duration with overlapping is used to create subsequent sliding windows. Two overlapping values were used in the current work, one sample of samples contained within 1 seconds (for example, sampling rate of 200 Hz has 200 samples within each second) and 50% overlapping. For each window, the functional connectivity value is calculated to create a network. Figure 3.10 illustrates how the approach of sliding window is used with one sample overlapping to capture the temporal variation. The red rectangles represent windows of fixed length (1s). While t refers to the starting time of window, that equal the sample number for $t = t_i, t_{(i+1)}, \ldots\ldots\ldots.., t_{(i+n)}$, while n is the total number of samples. For every time window for each canonical band, a functional connectivity network is created. αN_i refers to the created functional connectivity network for alpha band of window i. As a result, the temporal variation will be presented as consecutive number series of windows. Each created network is characterized by extracting its features using several graph-theoretic parameters. All values of features of the networks will be used to describe the structure of each network, which changes over time. The set of features of each network will be used in clustering. The clustering algorithm will divides the set of networks based on their features into clusters. Each network is assigned to a specific cluster based on its features values.

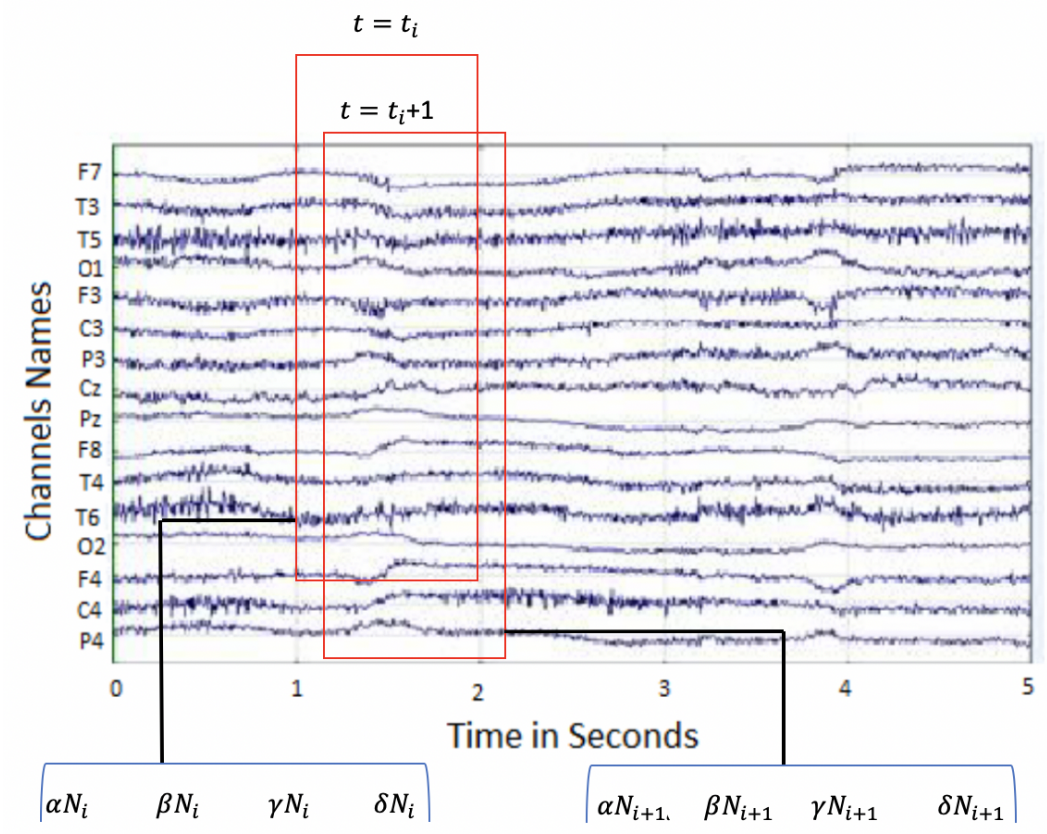


FIGURE 3.10: Sliding-window approach to calculate functional connectivity for four different frequency bands; t is the starting time of window, t_i is the sample number within the network of i^{th} window , $i = 1, 2, \ldots\ldots n$ where n is the total number of windows. αN_i is network calculated for alpha at i^{th} trial

3.1.4 Approach for Formulating Predictive Algorithm

The temporal dynamics of the graph-theoretic parameters for the created networks were used in a machine-learning framework to develop the target model. Clustering techniques were used to find the hidden structure (pattern) inside the data by segmenting data into clusters in which elements in the same cluster are more similar to each other than elements in other clusters. Figure 3.11 is an example of clustering a set of networks with specific feature values extracted from EEG signals. Since all created networks have been clustered (P1, P2 and P3) based on their features, each network is assigned to one of these three clusters.

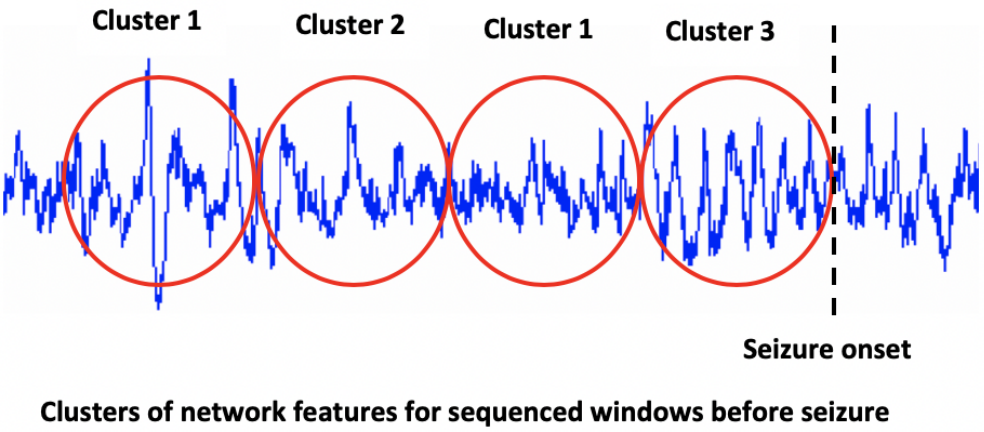


FIGURE 3.11: Clustering the extracted features from EEG data before seizure onset time into three clusters; each circle represents a window of created connectivity network with 1 seconds length

Figure 3.11 shows three groups of graph-theoretic parameters belonging to three created networks. The variation inside each red circle refers to the EEG signals within 1 second were used to create the functional connectivity network and then characterised using graph-theoretic parameters. These groups are three different clusters where each cluster is composed of a different pattern of network feature values. The cluster which appears close to seizure onset could be a sign of abnormal change happening before seizure. This features space clustering is applied to different patients to see which cluster they share and its time of occurrence for each patient. Finding a shared cluster close to seizure for all patients, indicates a common change occurring before seizure. This change may start early and at different time points for different patients. As a result, using clustering will help in formulating a prediction model. In the current work, k-means and k-mediods are used to cluster data because they can work with data of a significant size. Before clustering data, it is necessary to normalise the features. Normalising is used when features have different units or when one or more features has a large variance value compared to the other features. Normalising is used to prevent certain features which have large variance values compared to other features, dominating distance measures,

(e.g. the Euclidean distance measure) (Suarez-Alvarez et al. 2012). Therefore, the features were normalised so that their values were between zero and one. Equation 3.2 has been used for normalising features:

$$output_{norm} = \frac{(A - max(A)}{(max(A) - min(A))} \tag{3.2}$$

where A represents the matrix or vector of feature, and $min(A)$ or $max(A)$ represent the minimum or maximum value of all matrices respectively. Next, the idea of the two clustering algorithms (k-means and k-mediods) is described.

3.1.4.1 K-means and K-mediods Clustering Algorithms

The k-means is one of the unsupervised learning pattern recognition approaches. It depends on dividing data into groups after providing the number of clusters as prior knowledge. It starts by randomly defining the centroid points which equal the number of clusters predefined. Then it seeks to calculate the distance between each data point (observations) and each centroid using one of the distance measures. Euclidean is one of the common measures used (Shahid et al. 2009). The k-means clustering technique works iteratively by assigning each data point to the cluster of which the distance value to its centroid is the smallest. In other words, it measures the dissimilarity between data vectors (sets of extracted features). Therefore, it tries to minimise the cost function iteratively. Equation 3.3 (Kassambara 2017) illustrates the cost function of cluster u_j which represents the sum of squared distance of the data points to their centre points in the selected clusters. μ_j is the mean of the elements of j^{th} cluster(centroid point) while x_i is i^{th} data point wanted to calaculate the distance between it and μ_j .

$$Cost\_function(u_j) = \sum_{x_i \in j} \|x_i - \mu_j\|^2 \tag{3.3}$$

On the other side, K-mediods work in a similar way to k-means but the centre points in the clusters are used as mediods. They minimise the sum of dissimilarity compared to the k-means, which minimise the sum of the Euclidean distance. This makes it robust against outliers and noise (Arora et al. 2016).

Choosing the optimal number of clusters is a big issue; some algorithms must be provided with number of clusters as prior knowledge, such as k-means. Unfortunately, the optimal number of clusters is subjective, and depends on the distance measure and the parameters used in each clustering algorithm. Three common methods are used to calculate the optimal number of clusters: the elbow method, the silhouette method, and the gap statistic method.

The elbow method is one of the methods used to select the optimal number of clusters. This method compares the average of all points of variance within the same cluster to another cluster. The process chooses the optimal number based on the average value of the total variance within that cluster. Thus, choosing the optimal number of clusters will be at the point when the average value of the following cluster does not show any reduction. If the average value at each cluster is plotted, the number of clusters is chosen at the point where a knee or sharp angle is shown. The elbow method does not always identify the optimal number and can fail to show any elbow. Therefore, checking different methods is necessary to assess data and know if they can be clustered. On occasions, the elbow method does not give a clear result (ambiguous). Therefore, the average silhouette could be another choice when it comes to verifying the optimal number of clusters (Rousseeuw 1987).

The silhouette of data observation instance is the measure of how similar an observation is to other data observations in the same cluster, and the extent to which the observation instance (the set of features values for one connectivity network) is dissimilar to the neighbouring cluster. The method works by calculating the average silhouette at different numbers of clusters (Equation 3.4) (Han et al. 2012):

$$S_i = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{3.4}$$

where $a(i)$ is the average dissimilarity of point i to all points in the same cluster, which is calculated using a distance measure. Moreover, $b(i)$ is the minimum average dissimilarity of point i to all other points in the neighbouring cluster. After trying different clusters, the method chooses the neighbouring cluster that provides the smallest similarity. The silhouette value is between -1 and 1. If the silhouette value close to 1 means that the data observation instance is in the correct cluster, while a value close to -1 suggests that the data observation instance is in the wrong cluster. Rescaling features before calculating the silhouette could help to maximise the value of the silhouette at the correct number of clusters (Kaufman & Rousseeuw 2009).

The third method used to obtain the optimal number of clusters is the statistical gap. This method differs from previous methods, which depend on a null hypothesis. It starts by creating different numbers of random samples from original data. The value of the within sum square (similarity) is calculated for both sets of data – namely original and random samples – at different numbers of clusters. The statistic gap represents the deviation of the within-cluster sum-square value of the original data from value of the random samples. The optimal number of clusters will be chosen based on finding the smallest cluster k where the gap value at cluster $k + 1$ is larger or equal. This can be calculated as follows (Kassambara 2017):

$$Gap(k) = \frac{1}{B}\sum_{b=1}^{B} log(W_{kb}^{*}) - log(W_k) \tag{3.5}$$

where W_k is the total variation within cluster k and B the number of random samples. $k = [1, 2, 3, \ldots .N]$, where N is the number of clusters. W_{kb}^{*} is the total variation within cluster for the corresponding random sample.

Finally, 30 indices, which represent criteria to evaluate the number of clusters have been used (Caliński & Harabasz 1974). The description of these indices and how they are used to evaluate the number of clusters can be found in the research conducted in (Liu et al. 2010). These criteria are used to confirm the optimal number of clusters selected by the three methods or whether these methods failed to identify an optimal number of clusters. The optimal number of clusters was chosen based on the common decision made by these 30 indices. All previous methods have been carried out using MATLAB, which was provided as an integrated environment for all methods.

3.1.5 Performance Metrics Calculation for Classifier Models

To assess the performance of the created models in the thesis, three performance measures have been used - sensitivity, specificity, and false prediction rate. They were calcluated using the metrics of confusion matrix created for each model. Because of the small number of seizures for each patient, the number of common patterns appearing was considered in the thesis. On other word,the window based method rather than event based method was used for measuring detections and the number of 'events' (positive and negative) which are present for the performance metrics calculation. Table 3.4 illustrates the matrix used measure the performance of prediction models. Tge following are defintions of these metris:

- **True Positive**: the number of windows correctly identified as common pattern.
- **False Positive**: the number of windows incorrectly identified as common pattern.
- **True Negative**: the number of windows incorrectly identified as common pattern.
- **False Negative**: the number of windows incorrectly identified as non-pattern.

The sensitivity which measures the rate of windows that model has predicted correctly as common pattern while specificity measures the rate of correct predicted windows as non-pattern. False prediction rate measures the average number of windows predicted incorrectly as a common pattern of abnormal change per hour. The following equations illustrate how performance measures were calculated using the metrics of confusion matrix.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	TP	FP
	Non-pattern	FN	TN

TABLE 3.4: Representation of a Confusion Matrix; the shortcut letters refers to True Positive, False Positive, True Negative, False Negative for TP, FP, TN, and FN respectively

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.6}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.7}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.8}$$

$$FalsePredictionRate = \frac{FP}{FP + TN} \quad within\ each\ hour \tag{3.9}$$

Confidence intervals of sensitivity was applied to test the effect of sample size on the reported sensitivity for each model. The percentage of 95% was used for the confidence interval (Casson & Rodriguez-Villegas 2011). It means there are 95% possibility that sensitivity within the interval.

3.2 Summary

In this chapter, the procedural steps of the proposed method have been described. After the EEG signal has been cleaned of noise and artefacts using the ICA method, the synchronisation between brain regions was detected by calculating one of the three functional connectivity measures mentioned in chapter 2. From these values, the connectivity networks were created and characterised using graph-theoretic parameters. Finally, the clustering approach was taken to find the optimal pattern (cluster) for the temporal variation of the connectivity networks. In the next chapter, the results of using PLV in creating the connectivity networks to formulate predictive models are described.

Chapter 4

Predicting TLE Seizures Using PLV

In conjunction with six healthy subjects, 6 EEG records from 4 adult TLE surgical subjects (right hemisphere origin) were used as a training set (for more details on the data, see Chapter 3). The pre-processing steps described in Chapter 3 were applied to the EEG data, which were filtered into four bands – alpha, beta, gamma and delta. Undirected weighted functional brain connectivity networks were constructed using PLV at each frequency band, using a one-second sliding window with a one-sample overlap from the start of the data to the point of clinical onset of the seizure.

4.1 Analysis of the created networks

In this section, the results of applying the proposed method are presented, starting by exploring the dynamic changes in the created functional connectivity networks and investigating their temporal topographical changes using unsupervised clustering, leading towards the construction of a predictive model.

4.1.1 Selection of threshold in the connectivity networks

After creating the PLV-based functional connectivity network, a proportional thresholding (van den Heuvel et al. 2017) was applied to retain only the dominant connections. To select the appropriate value of proportional thresholding, the small-worldness measure signifying the structural randomness of a network was used as the determining factor using the training data (Bassett & Bullmore 2017, Humphries & Gurney 2008, Watts & Strogatz 1998). This compares the value of the clustering coefficient and the characteristic path length of the network to the value of the same network measures in a

random network with the same number of nodes and edges, using the following equation to calculate the small-world property (Watts & Strogatz 1998, Rubinov & Sporns 2010):

$$S = \frac{\frac{C}{C_{rand}}}{\frac{L}{L_{rand}}} \tag{4.1}$$

where, C and C_{rand} are the clustering coefficients of the original network and that of a random one, respectively, and, L and L_{rand} are the characteristic path lengths of the original and the random network, respectively. To calculate C_{rand} and L_{rand}, the following equations were used (Rubinov & Sporns 2010):

$$L_{rand} \approx \frac{lnN}{ln\,\langle k\rangle} \tag{4.2}$$

$$C_{rand} \approx \frac{\langle k\rangle}{N} \tag{4.3}$$

where $\langle k\rangle$ is the average degree of nodes, and N is the total number of nodes in the network. If the calculated value of S exceeded exceeds 1, then the structure of network tends to be a small-worldness one – a high clustering coefficient and low characteristic path length – and vice versa. Figure 4.1, (a) and (b) show the measure of the small-worldness values with respect to the percentages of proportional thresholding for seven TLE and six healthy subjects, respectively. It is apparent that the random structure of network ($S \leq 1$) for both TLE and healthy subjects starts after a 90% proportional threshold for both TLE and healthy subjects. As a result, 90% has been selected to be the proportional threshold applied to all created the connectivity networks that were created either in training or testing set.

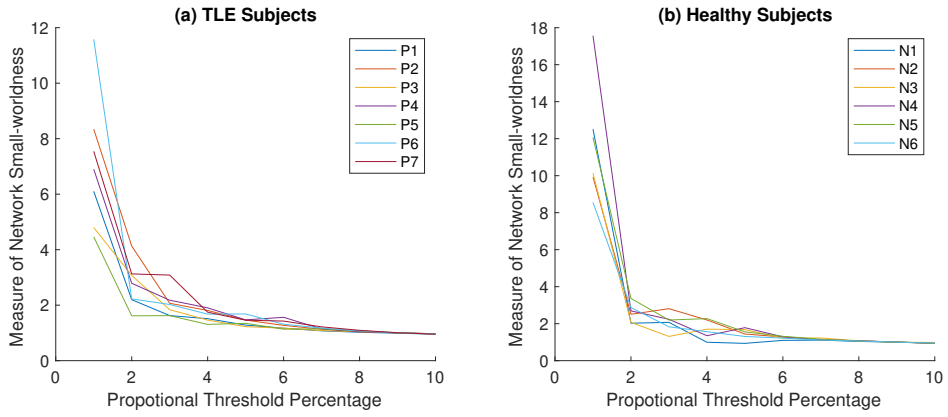


FIGURE 4.1: Small world-ness values at various proportional threshold percentages for: (a) 7 patients; and (b) 6 non-patients.

4.1.2 Exploring the change in PLV-based networks

The temporal change in PLV value before seizure onset had been explored for the patients in the training set, as well as for the six healthy subjects, in the four bands. Figure 4.2 shows the value of PLV at four points of time – 60 seconds before, 30 seconds before, 10 seconds before and immediately before seizure onset – for one patient, in the gamma band and using 19 channels. The colour maps show the variation in PLV, and light-yellow indicates values above 0.9 in some channels. Although this increase (to above 0.9) in PLV value was evident in some channels before seizure onset, it was not common among all the patients in the four bands.

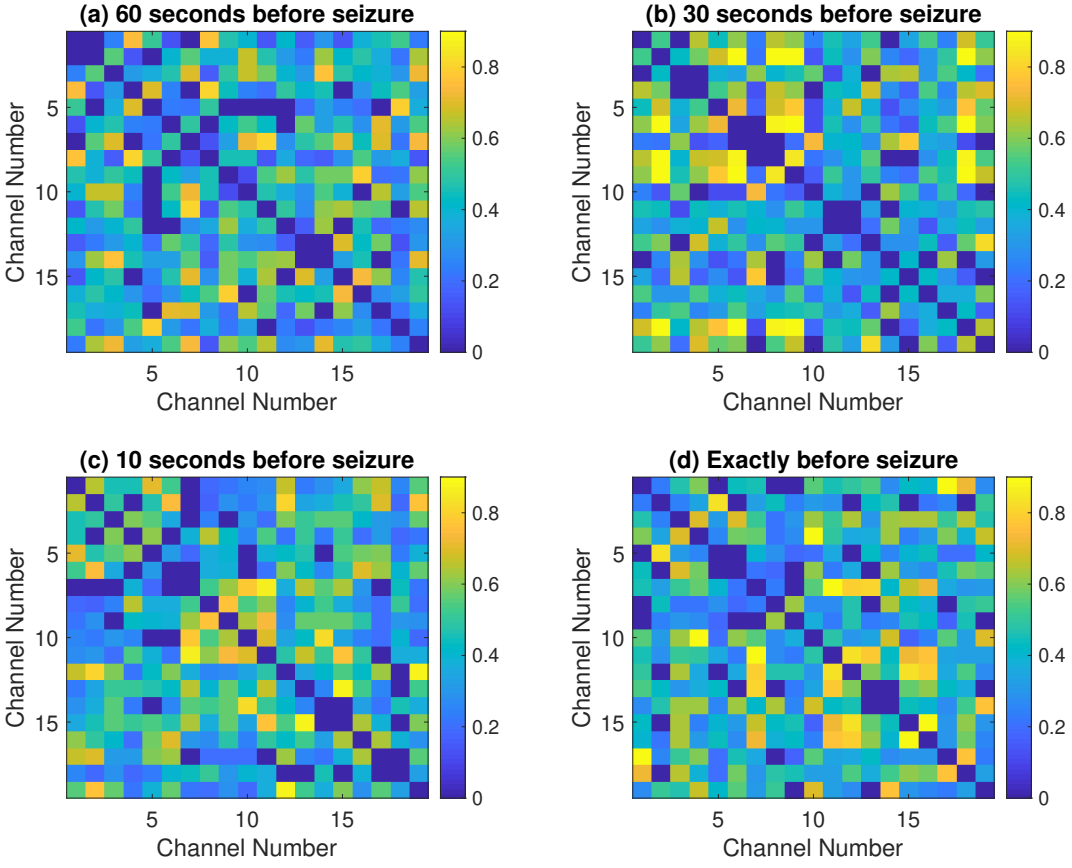


FIGURE 4.2: Temporal change in PLV value of data in gamma band in one patient at four points in time; (a) 60 seconds before; (b) 30 seconds before; (c) 10 seconds before; and (d) immediately before seizure onset.

To evaluate the temporal change in PLV values further and to extract a possible common pattern, two statistical measures of PLV - the average PLV values of all nodes (channels) within each window and the total number of the synchronised nodes where their PLV values are above 0.9 - were calculated for patients in the training set and the six healthy

subjects. Table 4.1 shows the results of the two measures' application within the four bands.

Subject	**Average PLV value**				**No. of synchronised nodes with PLV above 0.9**			
	Alpha	Beta	Gamma	Delta	Alpha	Beta	Gamma	Delta
P1	0.69	0.58	0.67	0.51	152	84	148	34
P2	0.66	0.58	0.57	0.52	114	48	64	22
P3	0.75	0.62	0.57	0.65	220	66	52	86
P4	0.56	0.59	0.60	0.55	52	50	58	26
P5	0.58	0.64	0.62	0.56	46	62	136	58
P6	0.68	0.63	0.68	0.53	122	90	120	36
N1	0.81	0.73	0.64	0.60	198	110	710	42
N2	0.75	0.56	0.49	0.58	170	42	20	18
N3	0.83	0.71	0.62	0.74	216	108	76	104
N4	0.82	0.70	0.58	0.59	216	104	48	30
N5	0.83	0.65	0.51	0.71	216	58	18	104
N6	0.80	0.63	0.50	0.54	196	60	22	28

TABLE 4.1: Trends in values of the two measures for the training subjects, including 6 healthy subjects, within four bands

All values in the table represent the maximum value reached for each subject. It is apparent that neither of the calculated measures in the four bands show a common trend value among patients, so checking the temporal change of PLV provided no pattern.

Subsequently, network features analysis was applied to find out if there is any common pattern in the feature domain.

4.1.3 Exploring features of the networks

Eight features (as mentioned in Chapter 2) were used to characterize the networks over four frequency bands (alpha, beta, gamma and delta), as discussed in chapter 3, as these are the most common features in studying and diagnosing various neurological disorders (Kodinariya & Makwana 2013, Rousseeuw 1987, Mohajer et al. 2011, Liu et al. 2010) such as epilepsy.

The dynamic change in the eight features was shown visually in order to find any common pattern. Figure 4.3 shows an example of the characteristic path length value (CPL) for six seizures in four patients, calculated for data in four frequency bands at exactly 10 seconds before the seizure onset. The x-axis denotes the window number - each window represents a connectivity network created over one second. It is apparent that there is no abnormal pattern of change (either an increase or a decrease) exhibited in the four frequency bands by all patients. Similarly, in Figure 4.4, exploration of the transitivity values revealed no common pattern. In visualizing the temporal change in the remaining

features before seizure onset, only random variations are shown. As a result, exploring changes in the features values did not show any common abnormal change. Accordingly, a clustering approach was taken to detect any possible common yet hidden pattern among patients.

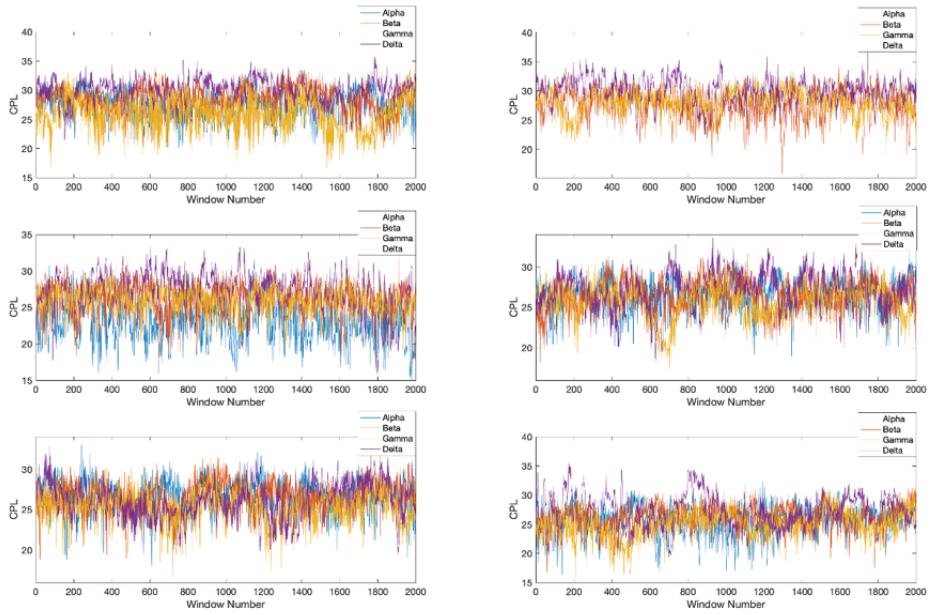


FIGURE 4.3: Characteristic path length (CPL) of 6 seizures in 4 patients at 10 seconds before seizure onset: calculated by sliding window of 1 second's duration, with 1 sample overlap (1 second = 200 samples)

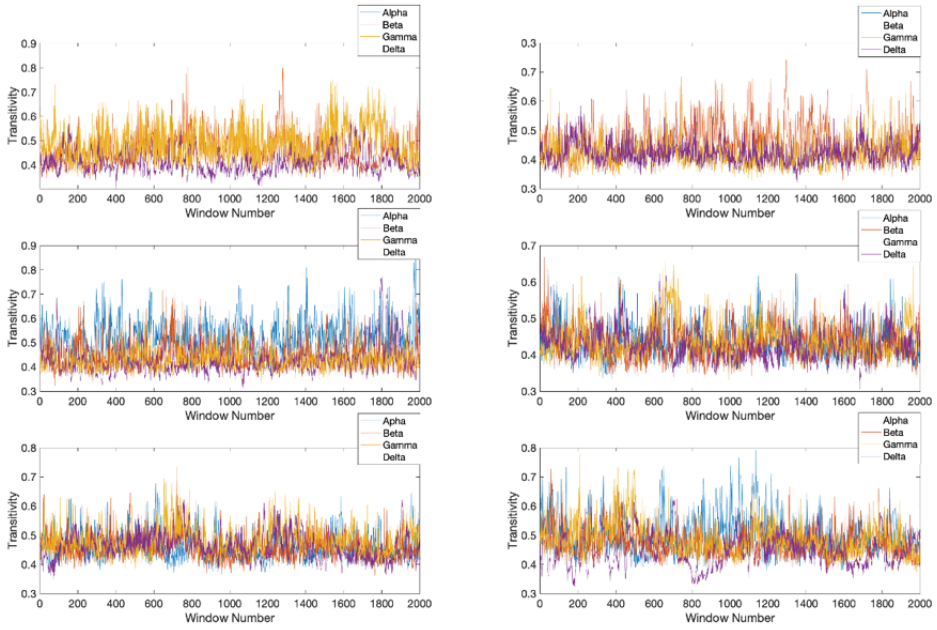


FIGURE 4.4: Transitivity values of 4 frequency bands in six patients, at 10 seconds before a seizure

4.1.4 Deriving a common pattern using clustering

To detect the presence of a common pattern, unsupervised pattern recognition using k-means and k-medoids (Arora et al. 2016) techniques were employed. Both algorithms work by minimizing the distance between the point designated as a centre and all points within the same cluster. Euclidean distance was used as the distance measure. To nullify any possible dominating effects of the features with large variance, feature normalization was the first step. Both algorithms needed to be provided with the number of clusters as prior knowledge. Therefore, three popular methods were used to find the optimal number of clusters: the elbow method (Kodinariya & Makwana 2013); the silhouette method (Rousseeuw 1987); and the gap statistic method (Mohajer et al. 2011). The clustering results from the above-mentioned methods for k-means and k-medoids are summarized in Table 4.2.

		Alpha	**Beta**	**Gamma**	**Delta**
k-means	Elbow	-	-	-	-
	Silhouette	2	2	3	2
	Gap	9	3	8	8
k-medoids	Elbow	-	-	-	-
	Silhouette	2	2	3	2
	Gap Statistic	2	2	3	2

TABLE 4.2: Optimal number of clusters using various evaluation methodss

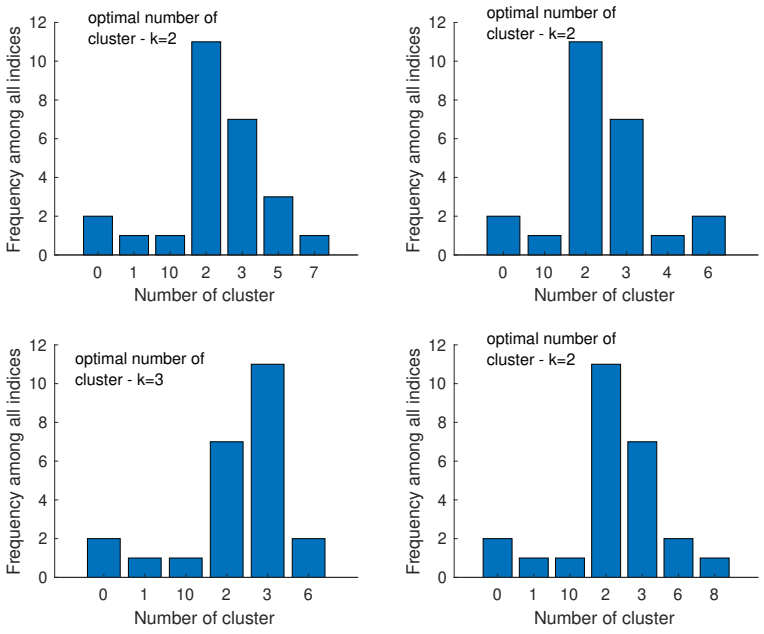


FIGURE 4.5: Frequency of the optimal number of clusters within each cluster, as produced by the 30 indices.

As shown in Table 4.2, no clear optimal number of clusters was obtained by k-means. Therefore, 30 indices, as described in an earlier study (Liu et al. 2010), served as the criteria by which to evaluate this number of clusters. The choice of the optimal number of clusters is based on the most frequent number provided by using these indices. Figure 4.5 shows a histogram of these 30 indices used for k-means clustering. It can be seen that the data in all bands show that two clusters is the optimal number, apart from that in the gamma band, which show three clusters. From Table 4.2 and Figure 4.5, it is apparent that both k-means and k-medoids arrive at the same number of clusters, which is two in the alpha, beta, and delta bands and three in the gamma band. As a result, both algorithms were used in formulating the predictive algorithm for epileptic seizures.

4.1.5 Formulating a predictive algorithm

After clustering the data, a feature selection step reduced the feature space dimensionality by retaining only the most significant features. Random forest classifier was used for this purpose, with cluster number serving as classification label. Figure 4.6 illustrates the five most important features of the data in the beta band, clustered by k-means, as shown in red. It can be seen that the characteristic path lengths and assortativity coefficients are the most significant features in classifying the data. These five features, from the TLE data only, were used to build a classifier model. The data corresponding to five-minute point before a seizure for these six TLE were used for this purpose.

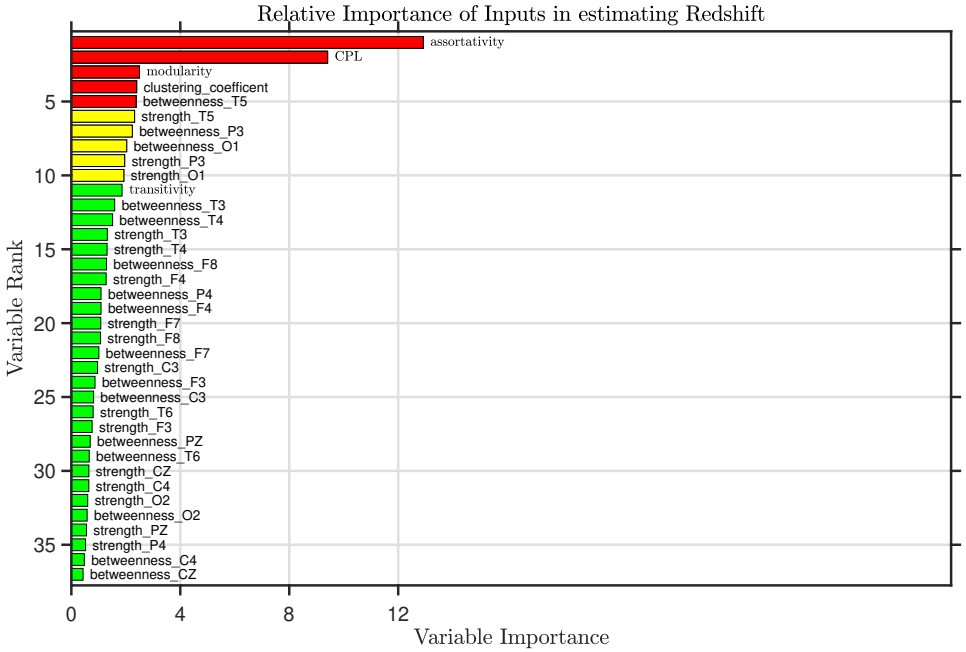


FIGURE 4.6: Feature ranking of data in the beta band, clustered by k-means.

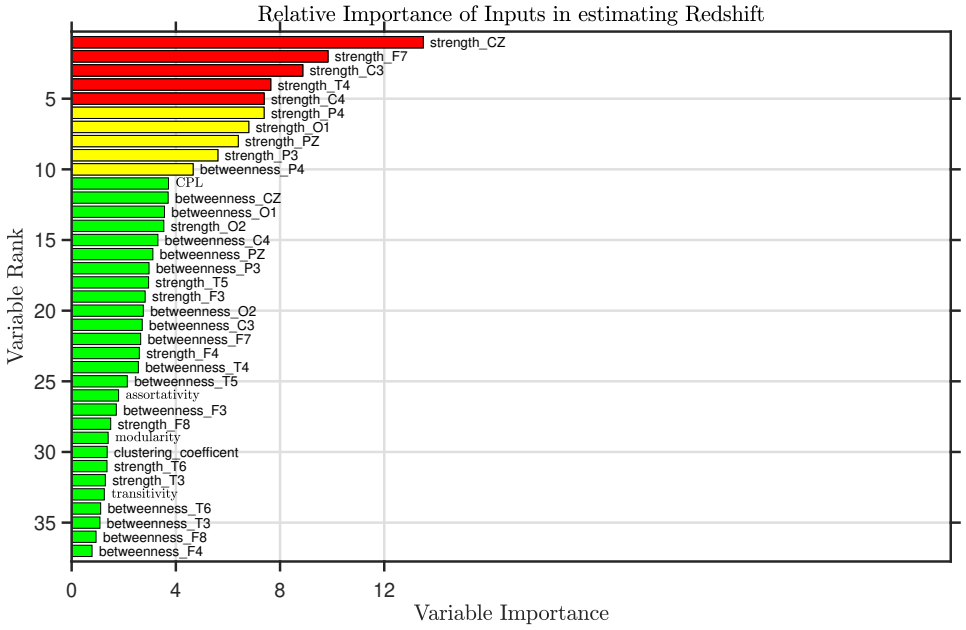


FIGURE 4.7: Feature ranking of data in the beta band, clustered by k-mediods.

Five classifier algorithms were employed with the features for k-means and k-medoids clustering, as their top-ranked features differed, as shown in Figure 4.6 and 4.7. The classification results for both are shown in Table 4.3. Five-fold cross-validation then calculated the accuracy and sensitivity of the classifiers.

Clssifuer	**k-means**			**k-medoids**		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Complex Tree	97.2%	96%	97.7%%	97%	96%	97.5%
LDA	96.3%	90%	99.2%	96.3%	90%	99.2%
SVM	100%	100%	99.9%	100%	99.9%	100%
KNN	98.2%	98%	98.4%	98%	97.9%	98%
Logistic Regression	100%	99.9%	100%	100%	99%	100%

TABLE 4.3: Performance results of 5 classifiers applied to clustered features using k-means and k-medoids

It can be seen that both the Logistic Regression and SVM classifiers in both k-means and k-medoids exhibited 100% accuracy and had a slightly higher sensitivity than the others. Therefore, these two were chosen to build the predictive model and were applied to all the TLE and the six healthy subjects within each band of clustered features. The output clusters of each subject produced by the classifier models were averaged by determining the most frequent in the output classes (created by the classifier model) during each second. These averaged results were represented graphically to seek any pattern that was shared by all patients. For the features clustered by k-means, one was found only within the gamma band – the presence of Cluster 3 for a duration of longer than five seconds at different time-points before onset of a seizure in all TLE, whereas in healthy subjects it lasted for less than five seconds. By contrast, there was no common pattern among TLE subjects using the features clustered by k-medoids. Therefore, the common pattern in the gamma band, which was found using k-means, was selected to build the predictive algorithm.

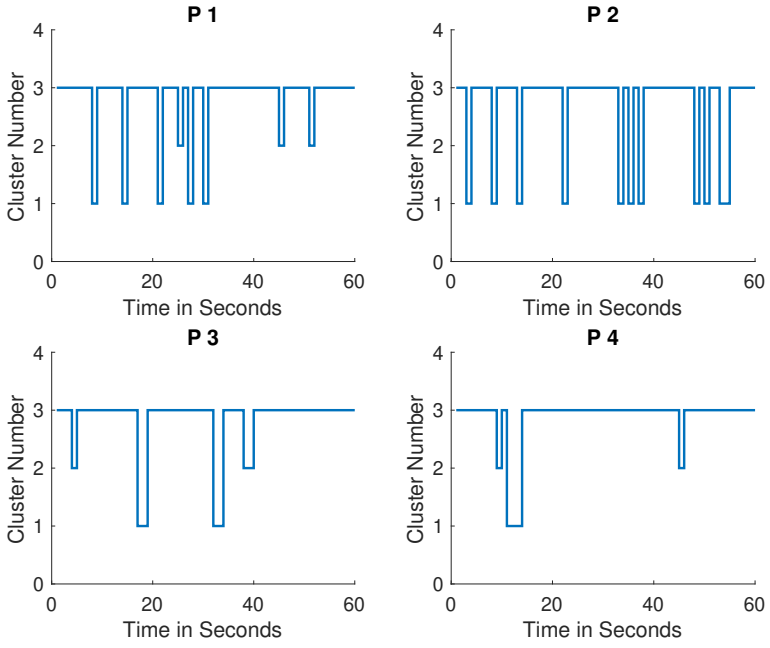


FIGURE 4.8: Transition of the three clusters in four TLE subjects before seizure onset.

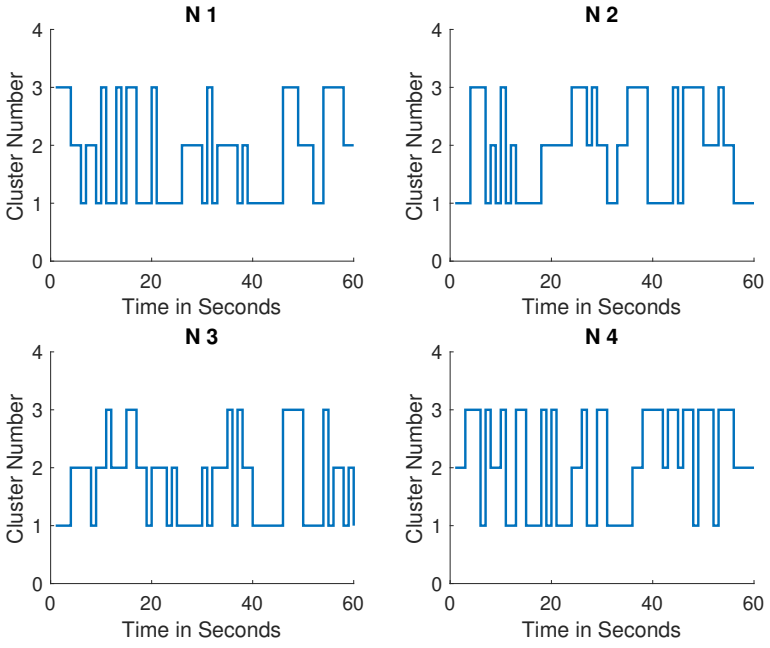


FIGURE 4.9: Transition of the three clusters in four healthy subjects.

Figure 4.8 and 4.9 show the temporal transition of the three clusters of four TLE and four healthy subjects, respectively. The presence of Cluster 3 for longer than than 5 seconds in all the patients is evident in Figure 4.8, for the TLE subject, yet not in Figure 4.9, for the healthy subjects. Table 4.4 presents the mean value of the five most important features of data in the gamma band, showing how features' values vary from one cluster to another. It is evident that Cluster 3 is characterized by a higher value of CPL, with lower values for transitivity, clustering coefficient and strength of the two nodes T5 and P3.

	Cluster 1	**Cluster 2**	**Cluster 3**
CPL	0.1380	0.5358	0.6714
Transitivity	0.3333	0.3086	0.2008
Clustering Coefficient	0.5242	0.4870	0.4266
Strength of Node T5	0.5621	0.5531	0.4575
Strength of Node P3	0.5733	0.6051	0.4603

TABLE 4.4: Mean values of five features of data in the gamma band in each cluster, conducted by k-means.

4.1.6 Statistical analysis

To characterize the dynamic property of the above-mentioned features, five descriptive statistics for each were calculated: mean; standard deviation; skewness; kurtosis; and

coefficient of variance. The aim was to find statistical trends shared by all patients, which could be used as new features in formulating the predictive algorithm.

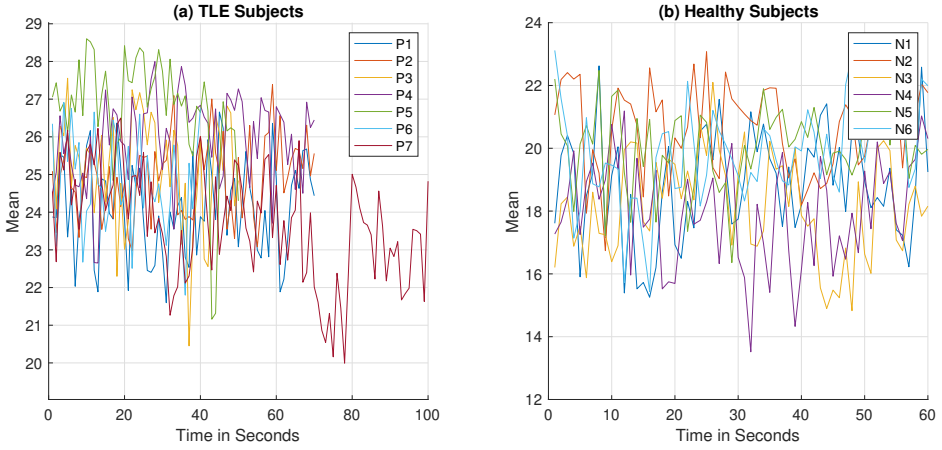


FIGURE 4.10: Mean of CPL value of 10 TLE and 6 healthy subjects.

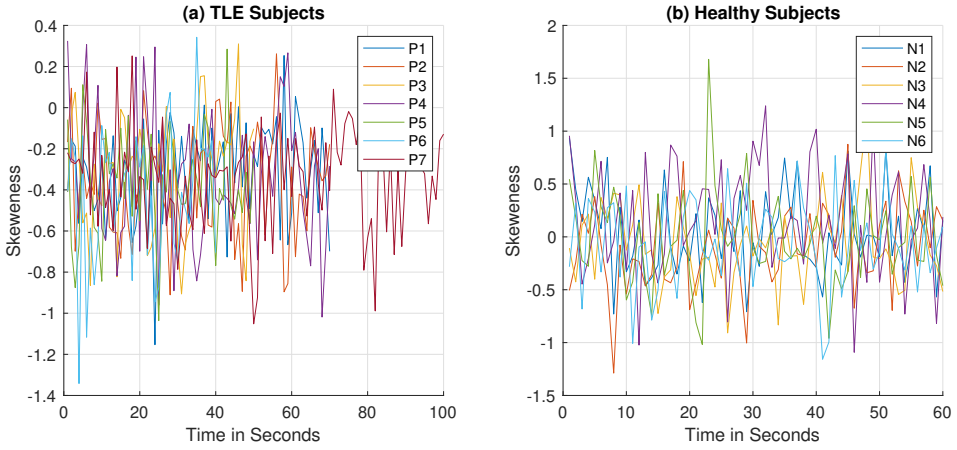


FIGURE 4.11: Skewness of CPL value of 10 TLE and 6 healthy subjects.

After applying these five statistical measures to all the five extracted features in the gamma band for the TLE and healthy subjects, only the feature of characteristic path length exhibited a statistical trend. This analysis shows that, for all the TLE subjects, for CPL only mean, skewness and coefficient of variance (mean/standard deviation) exhibit statistical trends: mean >25, skewness <0.5 and coefficient of variance <0.16. Figure 4.10, 4.11 and 4.12 show the three that show a trend in their values. In all figures, (a) and (b) refer to the seven TLE and the six healthy subjects, respectively. As shown in the figures, TLE subjects experience a varying length of time before seizure onset, while the healthy subjects all show one minute. A binary recoding of the clusters was applied by using these trends to formulate the predictive model according to the following rule: 1 – when all the three trends are simultaneously satisfied in the incoming

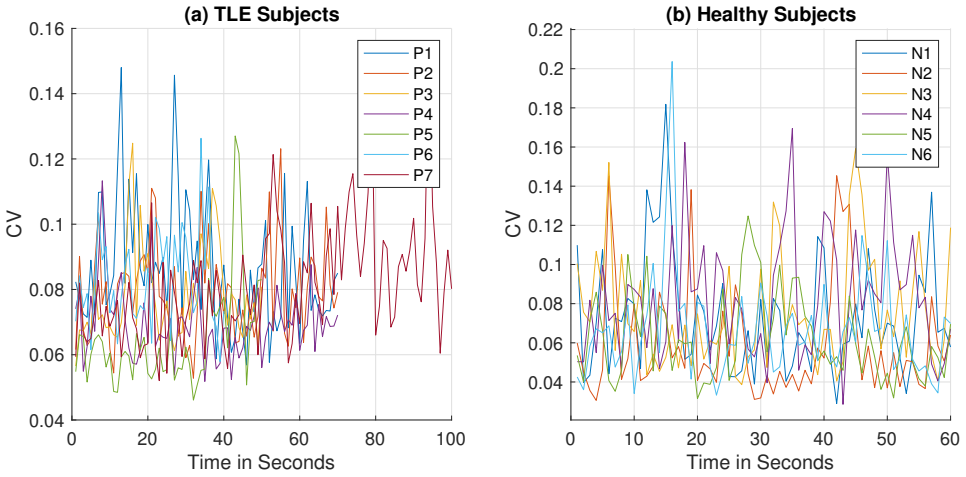


FIGURE 4.12: Coefficient of Variance of CPL value of 10 TLE and 6 healthy subjects.

data; and 0 – when at least one of the above mentioned trends is not satisfied. In addition, another measure was used, J, signifying the maximum number of occurrences of Cluster 3, which was calculated as follows:

$$J = W \times S \tag{4.4}$$

where S is the average value of characteristic path length at each second and W is calculated as follows:

$$W = \frac{(FD_{C1} + FD_{C2})}{FD_{C3}} \tag{4.5}$$

where, FD_* denotes the frequency of occurrence of the $cluster^*$. FD_{C3} was set to one, if it did not appear during the selected time window, to avoid dividing by zero in Equation 4.5. In Figure 4.13, (a) and (b) illustrate the value of J for the TLE and healthy subjects respectively. J is less than one when $FD_{C3} > FD_{C1} + FD_{C2}$, and vice versa.

The value of J and the recoded trend value (0 or 1) were then used as features to build the predictive classification model based on the following: at every second, the values of these two features extracted from the incoming data are fed into a classifier to classify two classes: seizures(when J value less than one and the value of satisfied trends is one); and non-seizures. The earliest time point at which the classifier output is labelled as 'seizure' is considered the prediction time point. The overall prediction time is calculated by subtracting this time point from the actual seizure time observed in the raw data.

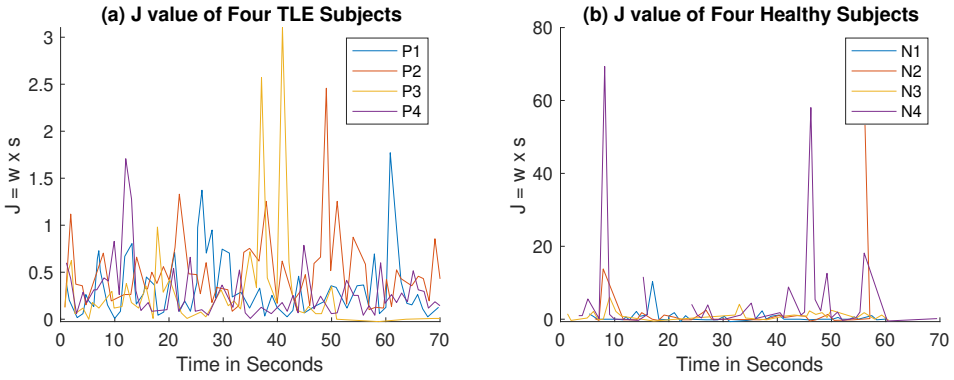


FIGURE 4.13: J value, applied to TLE and healthy subjects.

4.1.7 Prediction Performance

The length of EEG data of each patients used as training set was 5 minutes (30 minutes in total). Because of 1 sample used as overlapping, the total number of windows have been analysed is 360000. They are equal to 100 hours as a time of analysis. Here, the features values used in classification were calculated for each seconds (1 second =200 windows). As a result, 1800 windows (0.67 hours) were used in building the predictive model - 1000 windows assigned as pattern. Table 4.5 and Table 4.6 present the confusion matrices used to calculate the performance of Linear SVM and Logistic Regression classifier models respectively as described in Chapter 3.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	940	24
	Non-pattern	60	776

TABLE 4.5: Confusion Matrix of Linear SVM Model

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	920	16
	Non-pattern	80	784

TABLE 4.6: Confusion Matrix of Logistic Regression Model

Both the Linear SVM and Logistic Regression classifier models were applied to 24 seizure episodes in 10 patients, and the five-fold cross-validation approach was adopted to build the model and verify it. Both classifier models achieved 96% accuracy rate, with 94% and 92% sensitivity and 97% and 98% specificity, for Linear SVM and Logistic Regression respectively. Furthermore, the data for the six healthy subjects (containing no seizure episodes) were used with the classifiers and, in all cases, they worked correctly and made no seizure classification. The rate number of incorrect predicted windows as

pattern (0.03 and 0.02 for Linear SVM and Logistic Regression classifier models respectivley) was divded by the amount data used (0.67 hours) to extract false prediction rate in hour (0.04/h and 0.03/h for Linear SVM and Logistic Regression classifier models respectively). False prediction rate was calculated here to evaluate the performance of predictive model before apply it to the remaining patients. In Figure 4.14 (a) and (b) provide the 95% confidence interval of sensitivity for the first and second models respectively. They show 95% probability that the true sensitivity value ranges within the interval. It was checked with different size of test data (number of windows used).

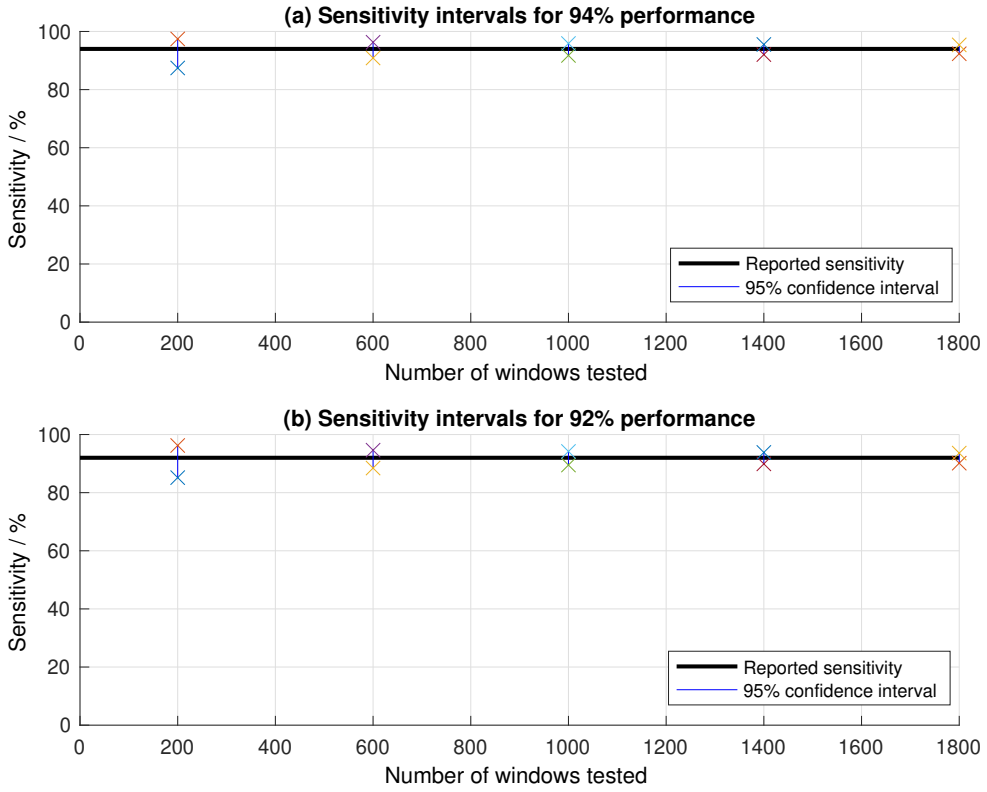


FIGURE 4.14: Estimated 95% confidence intervals of the sensitivity of the two models within different number of windows.

The patient-by-patient prediction time result is summarized in Table 4.7, and it shows that the maximum prediction time obtained was 12 minutes (Patient 3), while the minimum prediction time was 40 seconds. Although the minimum prediction time seems very low, it is because, in this particular case, the data were recorded for only 44 seconds before the seizure, which is considered to be insufficient for good prediction. This trend was also observed for the other records tested with the model (see Table 4.7), showing the trend whereby the longer the recorded data before the actual seizure, the better is the prediction time. Therefore, these models are expected to demonstrate better prediction in a continuous-monitoring scenario.

Patient no.	**Length of data before actual seizure (minutes)**	**Prediction time using Logistic Regression (minutes)**	**Prediction time using Linear SVM (minutes)**
1	5	4.99	4.98
2	5	5	5
3	24.07.4	12	12
4	1.25	1.17	1.17
5	1.26	1.16	1.16
6	1.28	1.17	1.17
7	0.88	0.83	0.83
8	0.73	0.67	0.67
9	0.97	0.83	0.83
10	2.74	1.67	1.67

TABLE 4.7: Prediction time of two classifiers for 10 TLE subjects.

4.2 Comparison

Table 4.8 compares the two models proposed in this study to nine previously published methods. It can be seen that the proposed models achieved a maximum prediction time of 12 minutes, which was the case for Patient 3 and, compared to the other methods, this prediction time seems to be low. However, from a clinical point of view, this is considered to be more practical. This is because, from this perspective, too early a prediction of a seizure (such as 60 (Iqbal et al. 2015) or 50 (Aarabi & He 2012) minutes before the actual seizure) is prone to misrepresentation, as there are several changes in information processing in the brain that could arise in the intervening period. Therefore, it is considered that a practical seizure prediction time should be neither be too long, nor too short (of the order of seconds) before an actual seizure. From this point of view, the prediction time obtained in these models is more practical than the others. This is demonstrated also by the high sensitivity obtained by the two proposed models, compared earlier models (Aarabi & He 2012, Li et al. 2013, Myers et al. 2016) and close to that of the study (Iqbal et al. 2015) where the prediction time was calculated for only four subjects, compared to the 10 TLE subjects in the proposed models. Also, the models proposed here have a higher specificity than those in earlier studies (Myers et al. 2016) and (Cho et al. 2016) and a lower false prediction rate (all seizures used have been predicted while no seizure was detected for healthy subjects) than those in (Alotaiby et al. 2017) and (Chu et al. 2017). However, the results are still based on only a small number of patients (10) and, although promising, need to undergo a larger patient trial.

Model	Sensitivity (%)	Specificity (%)	FPR (/h)	Prediction time (minutes)	No. of patients
Zandi et al. (2010)	86	-	0.12	20.8	3
Aarabi & He (2012)	79.9	-	0.17	50	11
Iqbal et al. (2015)	100	-	0.38	60	4
Li et al. (2013)	75.8	-	0.09	49.7	21
Cho et al. (2016)	82.44	82.76	0.5	5	21
Chu et al. (2017)	86.67	-	0.367	45.3	16
Myers et al. (2016)	76.8	88.3	0.167	-	10
Usman et al. (2017)	92.23	-	-	23.6	22
Alotaiby et al. (2017)	89	-	0.39	68.71	24
Model 1 (Linear SVM)	94	97	0.04	12	10
Model 2 (Logistic Regression)	92	98	0.03	12	10

TABLE 4.8: Comparison of the two models proposed to nine existing models. FPR refers to False Prediction Rate

4.3 Conclusion

The models proposed predict the epileptic seizures of 10 TLE patients by analysing the functional brain connectivity networks created using PLV. A common pattern among patients' data was found only in the gamma band – the presence of Cluster 3 for longer than five seconds at different points in time before onset of a seizure in all TLE patients, whereas in the healthy subjects it lasted for less than five seconds. It was characterized by an increased value of the characteristic path length. When using this observation within the framework of Linear SVM and Logistic Regression classifiers, it achieved seizure prediction with 94% and 92% sensitivity and 97% and 98% specificity, respectively. For healthy subjects, neither model detected any seizure episodes, which was the expected result. Both predictive models showed 12 minutes as the maximum prediction time prior to the clinical seizure onset, based on the available data in this work. The minimum prediction time in the models, in essence, depends on the length of the data recorded before the actual seizure episode, and therefore the models are expected to work well when applied in a continuous monitoring scenario. The combination of high sensitivity and specificity and a clinically practical prediction time makes these models good candidates for further exploration with a larger population, with the aim of translating them into real-life clinical practice.

Chapter 5

Predicting TLE Seizures Using PLI and SL

One of the objectives of this research is to develop models to predict seizure in TLE subjects using three synchronisation measures – PLV, PLI, and SL. The previous chapter showed the possibility of using PLV-based connectivity networks in developing a predictive model. The two models formulated were able to predict all the seizures, giving a maximum prediction time of 12 minutes and achieving a false prediction rate of 0.04/h and 0.03/h respectively. However, the biggest drawback of PLV-based method is that they are susceptible to volume conduction which may result into false synchronisation. On the other hand, PLI and SL are more robust in this respect. Therefore, this chapter is aimed at developing models using PLI and SL and compare their outcomes with the result of the previous chapter.

The present chapter has been divided into three sections. It starts by dealing with the result of using PLI in formulating a predictive model, and then illustrates the result for SL, ending with a discussion on the results of all three measures.

5.1 Result of prediction using PLI

EEG data of the 10 TLE subjects used with PLV were explored to develop a PLI-based predictive model. The number of seizures checked for each patient was increased to at least two. Furthermore, EEG data for six healthy subjects were used alongside the patients' data to assess the difference. A pre-processing step was used (see Figure 3.4) to clean the data of any noise or artefacts. Here, EEG signals were filtered into five different bands – alpha, beta, gamma, delta, and theta – instead of the four used with PLV. PLI-based connectivity networks were created for each 1-second sliding window. Moving the

one window sample, as applied previously with PLV, did not provide additional information but instead raised the computation time for creating the functional connectivity networks. Accordingly, the overlap of the sliding window was modified to 0.5 seconds (50% overlap) instead of one sample. That means there are 599 windows for each EEG data in the training set with total number equal 3594 windows. They are equal 1 hour as a time of analysis.Similar to PLV, the created networks were characterized using nine features – strength of node, modularity, transitivity, clustering coefficient, characteristic path length, density, betweenness centrality and assortativity – in addition to the global efficiency and raduis features. Four types of classifiers (SVM, logistic regression, linear discriminant analysis and Bagged Tree) were selected to build a predictive model. All four had their own way of learning and, consequently, may give prediction models of varying accuracy. The result of using PLI in predicting seizures will be presented later in this section.

5.1.1 Exploring the change in PLI-based networks

Functional brain connectivity networks were created for 24 seizure files of 10 TLE subjects. The six EEG data of four TLE subjects who had 5 minutes recording time before seizure onset were selected for the training set. They were used in analysing and formulating the predictive model. Furthermore, the functional brain connectivity networks of six EEG data of healthy subjects were created for comparison to the structure change in the patients' networks. To explore the dynamic change of PLI, functional connectivity networks were created before and during seizure for all five bands. The temporal variation in PLI values was checked for the six EEG data 10 seconds before and during seizures at the beginning, then it was increased until it reached 60 seconds before seizure onset. Only the data in the beta and gamma bands showed a slight increase in PLI value before and during seizures for all six seizures at several points of time. By contrast, the data within these two bands for the six healthy subjects showed no increase in PLI value.

Figure 5.1 depicts the colour maps to illustrate the dynamic change in PLI value of data in the beta band. They show the change in PLI value at three points in time within 10 seconds prior to and during seizure. A slight increase in the value of PLI – which exceeded 0.9, as illustrated by the colour bar – is exceeded at one point, shown in bright yellow. Figure 5.2 exhibits the same noticeable increase but for the data in the gamma band for the same patient. In both figures, it is apparent that the PLI value of most channels exceeded 0.9 at some point either before or during a seizure.

To evaluate this increase in PLI value and determine whether it is a sign of abnormal change common to all patients, two statistical values of PLI value were calculated: the average value of the PLI of all synchronised nodes (channels) within each window; and the total number of these synchronised nodes whose PLI value exceeded 0.9. Both measures were calculated for the six EEG data of the four TLE subjects, as well as

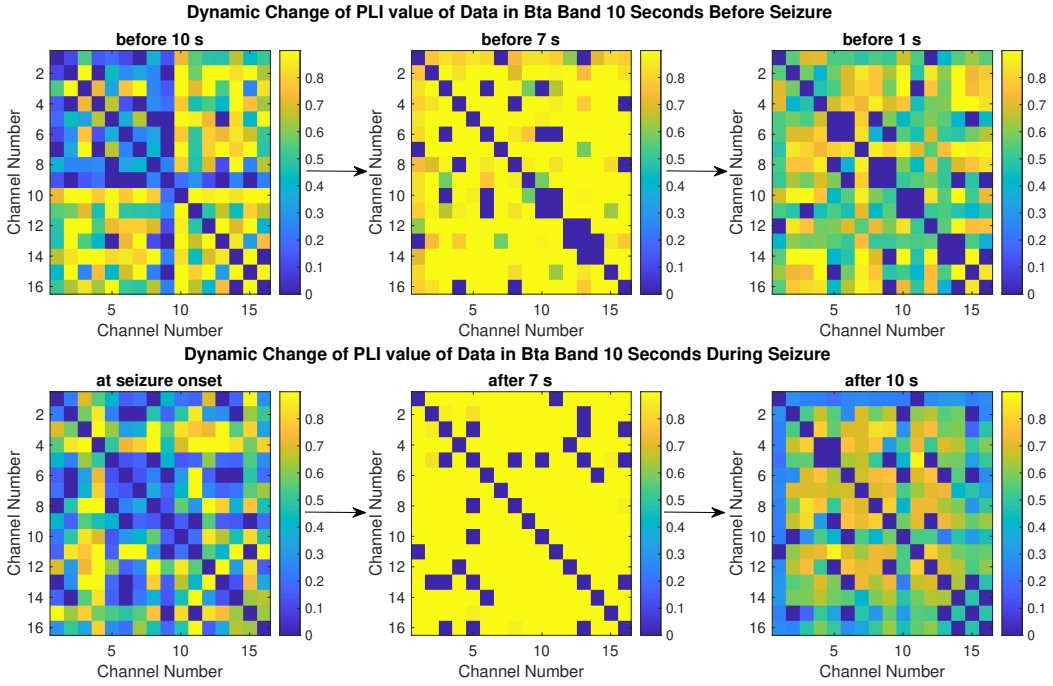


FIGURE 5.1: Temporal change of PLI value before and during seizures in beta within 10 seconds.

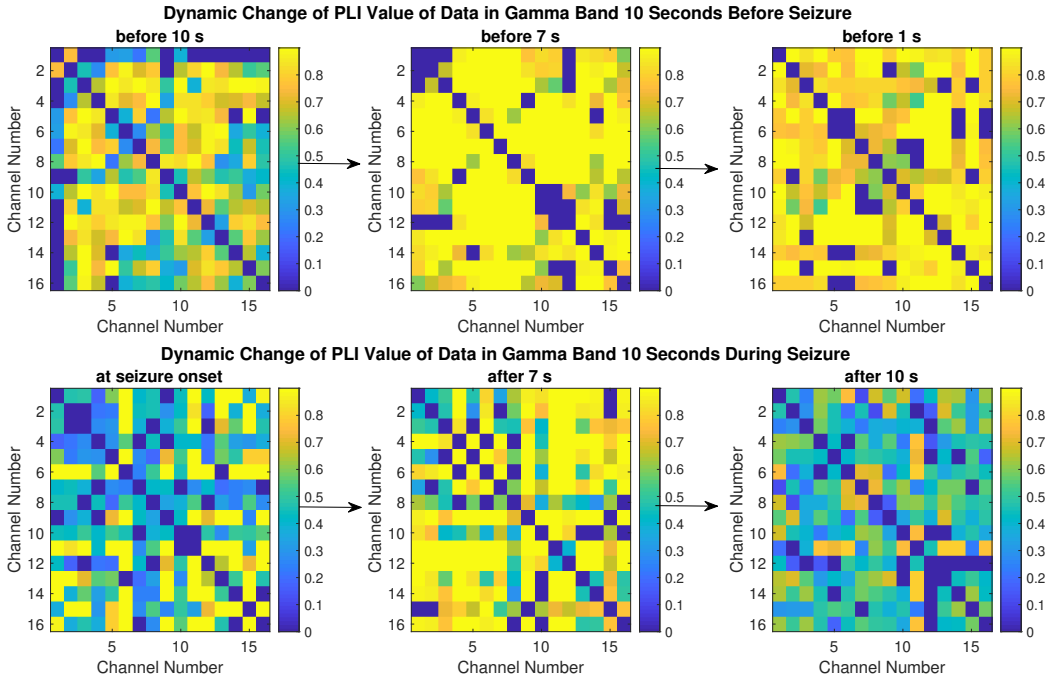


FIGURE 5.2: Temporal change in PLI value before and during a seizures in gamma band, within 10 seconds of onset.

for the six healthy subjects. The aim of the calculation was to seek a trend in value shared by the patients to distinguish them from healthy subjects. After checking the temporal variation of both measures, it was found that only the measure of total number of synchronised nodes with a PLI value above 0.9 showed a trend in value shared by the patients (TLE subjects). Figure 5.3 and 5.4 illustrate the temporal change for three TLE and six healthy subjects for data in the gamma band. The windows shown in Figure 5.4 represent one minute overall, which is the total length of recording data of all healthy subjects. It is apparent from the two figures that the trend value – which exceeded 200 – is the same in both the TLE and healthy subjects. Therefore, it cannot be used to formulate a predictive model. Next, the extracted features of the networks were explored to identify any abnormal changes.

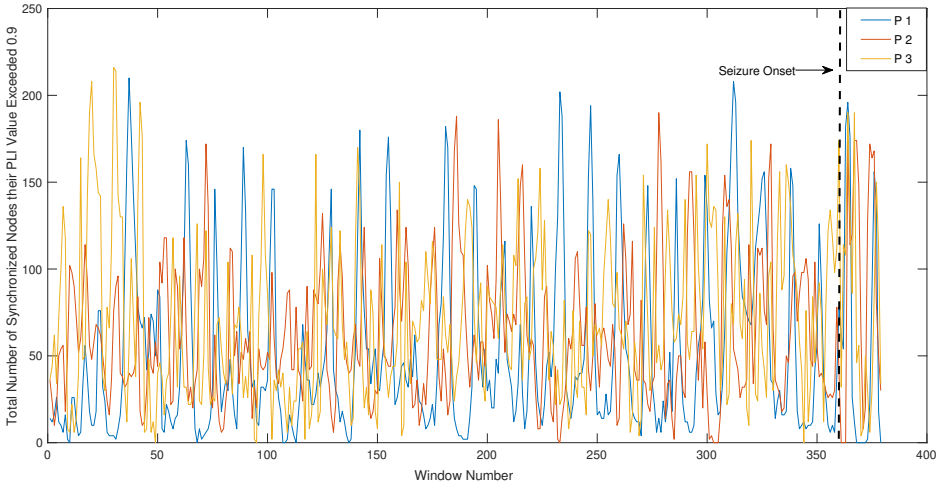


FIGURE 5.3: Temporal change in the total number of nodes with PLI values exceeding 0.9 for three TLE subjects in the gamma band.

5.1.2 Exploring the network features

Since checking the temporal change in PLI value either before or during a seizure did not reveal any clear abnormal change common to all patients, the features of the created functional connectivity networks within each of the five bands were subsequently extracted. In the current work, nine features were extracted for all bands: Strength of Node, Assortativity, Betweenness Centrality, Clustering Coefficient, Transitivity, Modularity, Characteristic path length (CPL), Global Efficiency and Radius. An examination of the temporal change of the features' values was conducted for the six EEG data before and during seizures to search for any pattern of sudden falls or rises and to compare it to the features' values in the six healthy subjects. Figure 5.5 shows the temporal change in CPL value before and during epileptic seizure for the beta band. The x-axis represents the window number at which the feature value was calculated. The red dashed vertical

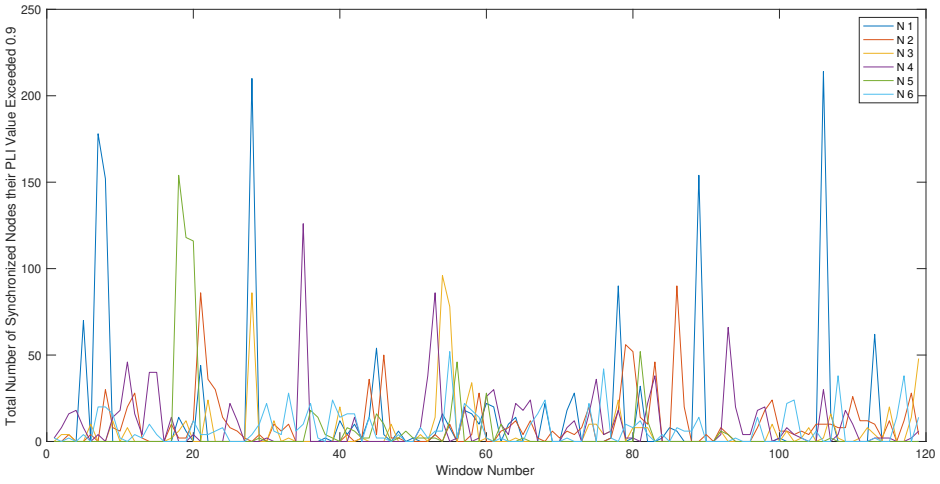


FIGURE 5.4: Temporal change of the total number of nodes with PLI values exceeding 0.9 for six healthy subjects in the gamma band.

line marks the beginning of an epileptic seizure. The figure compares the variation in the CPL value of data in the beta band before and during seizure within 10 seconds. It can be seen how different they are in all patients during seizure. In addition, the same level of increase in CPL value was observed in the data of healthy subjects. Similarly, assessment of CPL variation in the alpha, gamma and delta bands revealed the same result. Therefore, tracing the temporal change in CPL showed no common abnormal change within the five bands.

Checking the temporal change in betweenness centrality of the nodes started with four channels in and around the right temporal lobe – T4, T6, C4 and F4. Both Figure 5.6 and 5.7 illustrate the betweenness centrality values for two nodes represented in the right temporal lobe area – T4 and T6 – for data in the beta band for six TLE and six healthy subjects, respectively. The x-axis represents the window number, where a window number of zero indicates a window 5 minutes before onset of the seizure. The green vertical lines in Figure 5-6 represent the start of the seizure. Both TLE and healthy subjects exhibited peaks exceeding 100, but these figures were below 80 during seizure. After checking these four channels around the right temporal lobe, no clear abnormal change was found in patients' values before seizure, while the same was true for subjects without epilepsy. In conclusion, exploring the betweenness centrality value did not reveal any sign that could be used to predict a seizure.

Similarly, the remaining features of the networks were checked but, unfortunately, exploring the temporal changes of these features showed no abnormal change which can be used as a common pattern. Following this, clustering was applied to figure out any hidden pattern. The procedure of clustering and formulating the predictive model is explained in the following subsection.

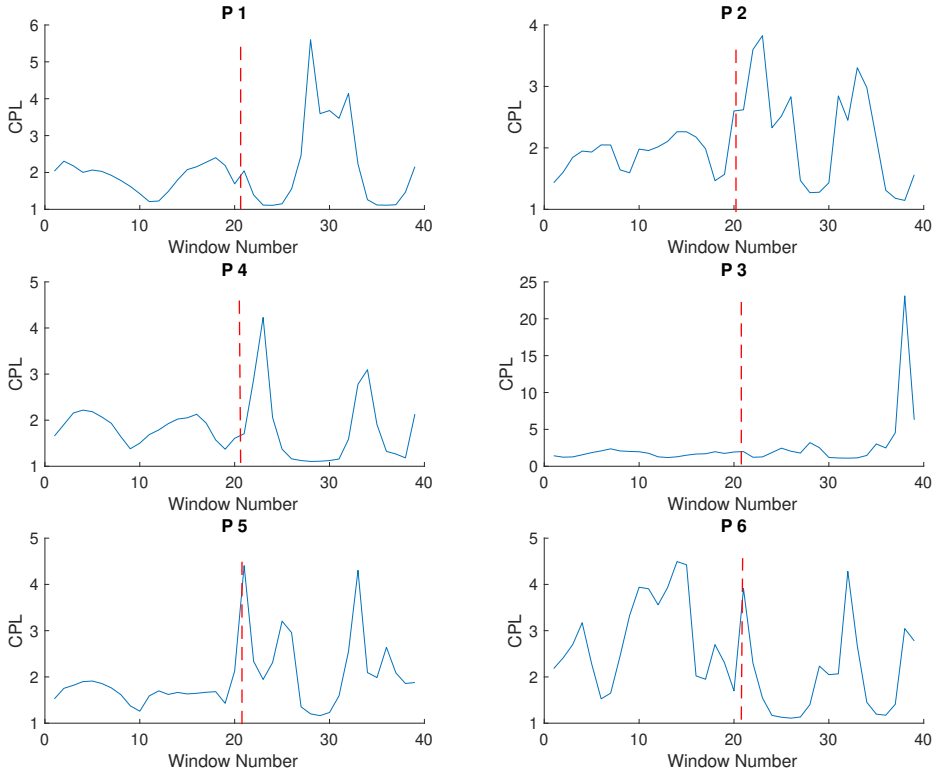


FIGURE 5.5: Characteristic path length value for six TLE subjects 10 seconds before and during seizure.

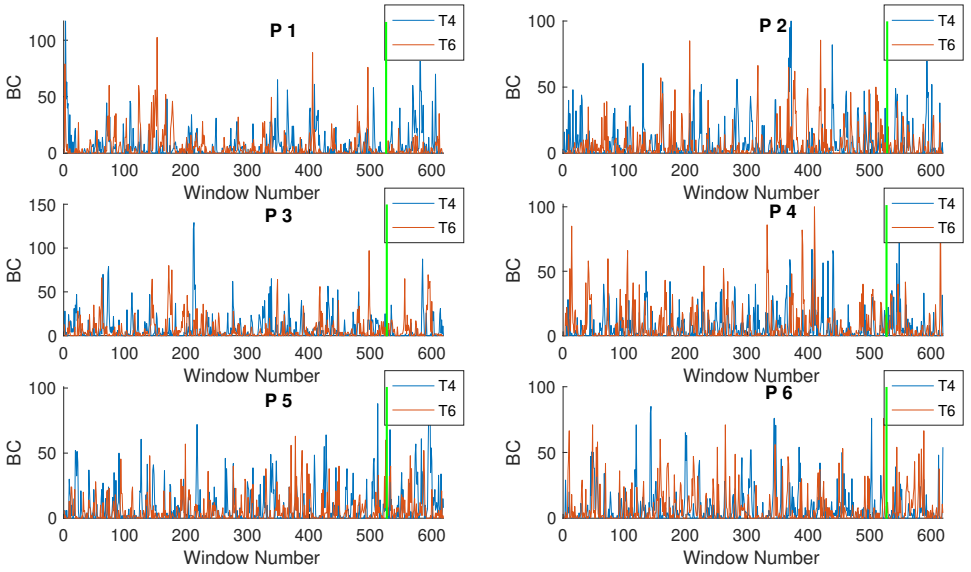


FIGURE 5.6: Betweenness centrality of data in beta band for six TLE subjects, 5 minutes before seizure onset.

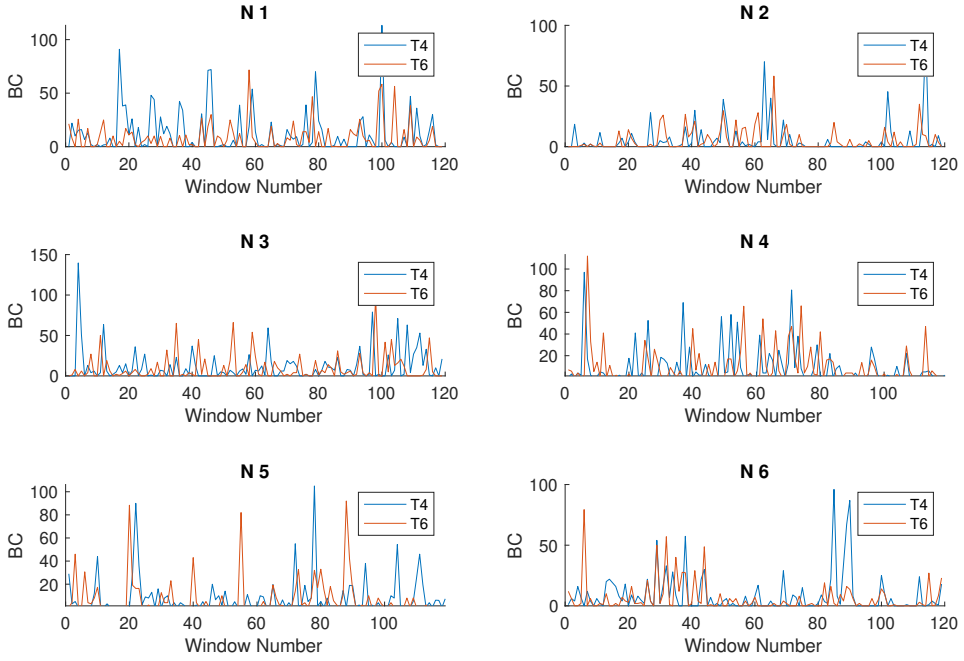


FIGURE 5.7: Betweenness centrality of data in the beta band for six healthy subjects.

5.1.3 Clustering data and predicting seizures

As the next step, a clustering process was applied using K-means. Before clustering, the features' dimensions were reduced by deleting one of the two strongest correlated features. Doing this removed any bias arising from clustering. The strength of the node features of all 16 channels used showed a high correlation to the characteristic path length and modularity features. Therefore, they were deleted. The nodes used to calculate betweenness centrality were reduced to four: two represent the right temporal lobe of the brain (T4 and T6) and two around this area (F4 and C4). They were selected for checking because the patients in the study had temporal lobe epilepsy originating in the right lobe. As a result of the reduction in the features' dimensions, eight features were selected – CPL, Modularity, Transitivity, four nodes of Betweenness Centrality (T4, T6, F4, C4) and Assortativity – and clustered for each band. The optimal number of clusters was calculated using the elbow, silhouette and gap statistical methods as before. Extracting the common pattern for the clustered features in each band separately did not reveal any shared pattern among the patients. Therefore, a combination of the features of the bands was applied, starting with the beta and gamma bands, because of the slight change in PLI values that was observed earlier. The optimal number of clusters of using both beta and gamma bands' features was calculated using three methods –

namely elbow, silhouette and gap statistics – and they arrived at three, two and one, respectively.

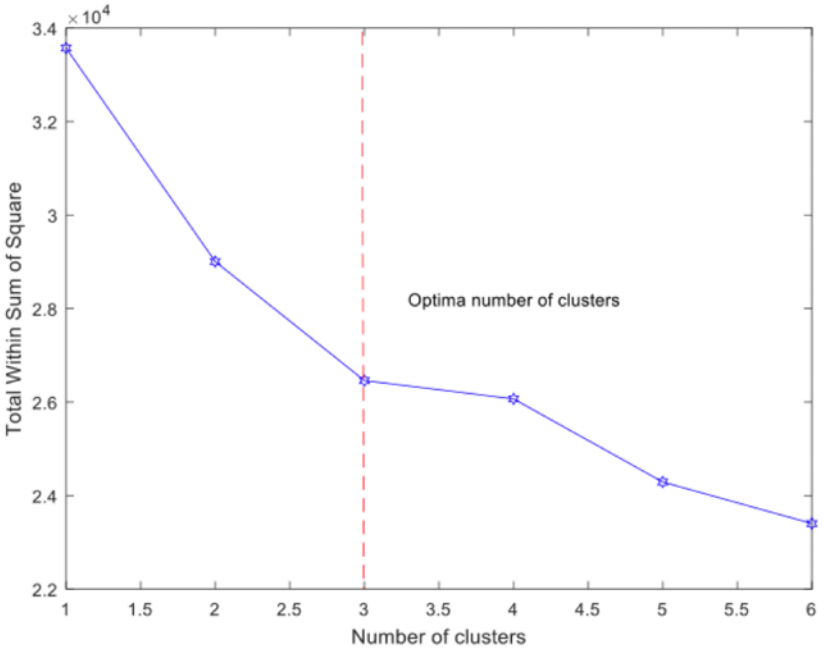


FIGURE 5.8: Elbow method for extracted features in both beta and gamma bands.

Figure 5.8 shows the optimal number of clusters as three, using the elbow method. Such an elbow can be seen at 3 on the x-axis. For further confirmation of the optimal number of clusters, the 30 indices mentioned in (Liu et al. 2010) were applied to assess the number of clusters. Each of these 30 indices has its own way of finding the optimal number of clusters, leading to various answers. A histogram of the frequency of index for each cluster (Figure 5.9) shows that the most frequently selected clusters by the 30 indices was three. This confirms the results achieved previously by the elbow method. Therefore, the extracted features were grouped into three clusters.

Each cluster of the three was tested to extract the common pattern (cluster) shared among patients. Every cluster was tested for all TLE and healthy subjects. In each, the process of testing was constructed by building a classifier model using samples of data from the patients in that cluster, labelled number 1. The data of the extracted features of the six healthy subjects were added to the classifier with the label 0. Four classifier models were chosen to classify the prepared data mentioned earlier in this chapter, and composed of TLE subjects' data of the cluster with the label 1, and label 0 for the healthy subjects.

Measure	**LDA**	**Logistic Regression**	**SVM**	**Bagged Tree**
Sensitivity	84%	75 %	88%	88%
Specificity	78%	79%	80.5%	84%

TABLE 5.1: Sensitivity and specificity measures of four classifier models.

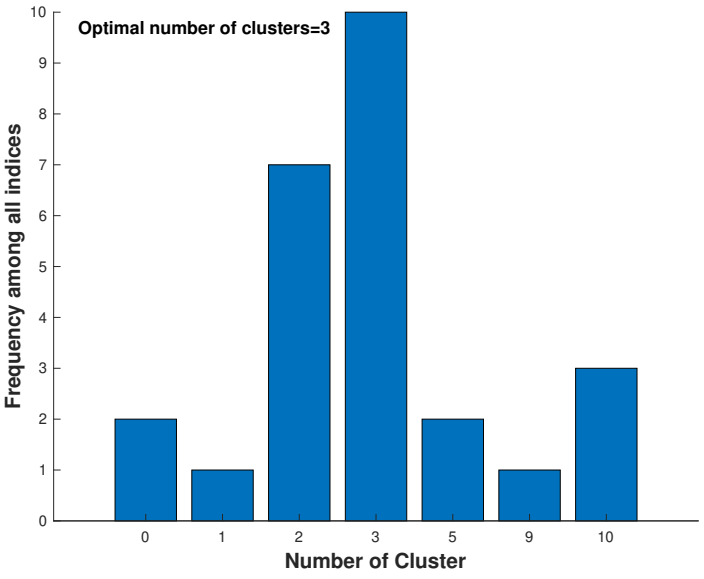


FIGURE 5.9: Histogram of results of applying 30 indices in the clustering evaluation.

Table 5.1 shows both the specificity and sensitivity of all four classifiers for Cluster 1. It illustrates that the Bagged Tree showed slightly better prediction statistics (sensitivity and specificity). These were calculated from the confusion matrices extracted after evaluating the performance of classifiers by applying the five-fold cross-validation method.

Simple statistical analysis was used to describe the networks' structure during each cluster by calculating the mean values. These values were used to compare the common pattern (Cluster 1) to the other two clusters. For Cluster 1, the statistical values of the extracted features of data in the gamma band reported significantly more distinguishable values than those in the beta band. Table 5.2 presents the results obtained from calculating the mean value of network features of data in the gamma band, which was used with data in the beta band to create the predictive model. It shows the mean value for three clusters predicted using Bagged Tree model, which achieved high accuracy in addition to good sensitivity and specificity measures. It provides a comparison of the three clusters and shows what happened to the network structure in the common pattern.

It can be seen from the data in Table 5.2 that the mean of CPL value of Cluster 1 was higher than that of the other two clusters. This indicates that the structure of the network changed, making it the longest path to transfer data through the brain regions. According to this increase in the path lengths, the mean of the betweenness centrality value of the nodes of the right temporal lobe (the origin of seizure) exhibited higher values in T4, T6 and the other two nodes around them (F4 and C4). This refers to increases in the activity in these regions of the brain, which interprets the increase

Feature	Cluster 1	Cluster 2	Cluster 3
Assortativity	-0.06723	-0.09461	-0.09597
Betweenness centrality F4	11.66197	8.697298	7.14617
Betweenness centrality C4	10.44352	8.534289	7.564803
Betweenness centrality T4	11.54533	8.043288	8.075776
Betweenness centrality T6	10.56221	7.554554	7.722671
Transitivity	0.417287	0.629029	0.61843
Modularity	0.075775	0.024779	0.027528
Characteristic path length	2.233076	1.588615	1.625887

TABLE 5.2: Mean value of the three clusters of data in the gamma band.

in the network path lengths. Similarly, the mean of modularity of Cluster 1 (common pattern) had the highest value, while the transitivity value had the lowest. A possible explanation for the transitivity value might be the lack of transitive triples in the network because of the greater betweenness centrality value of nodes in the right temporal lobes.

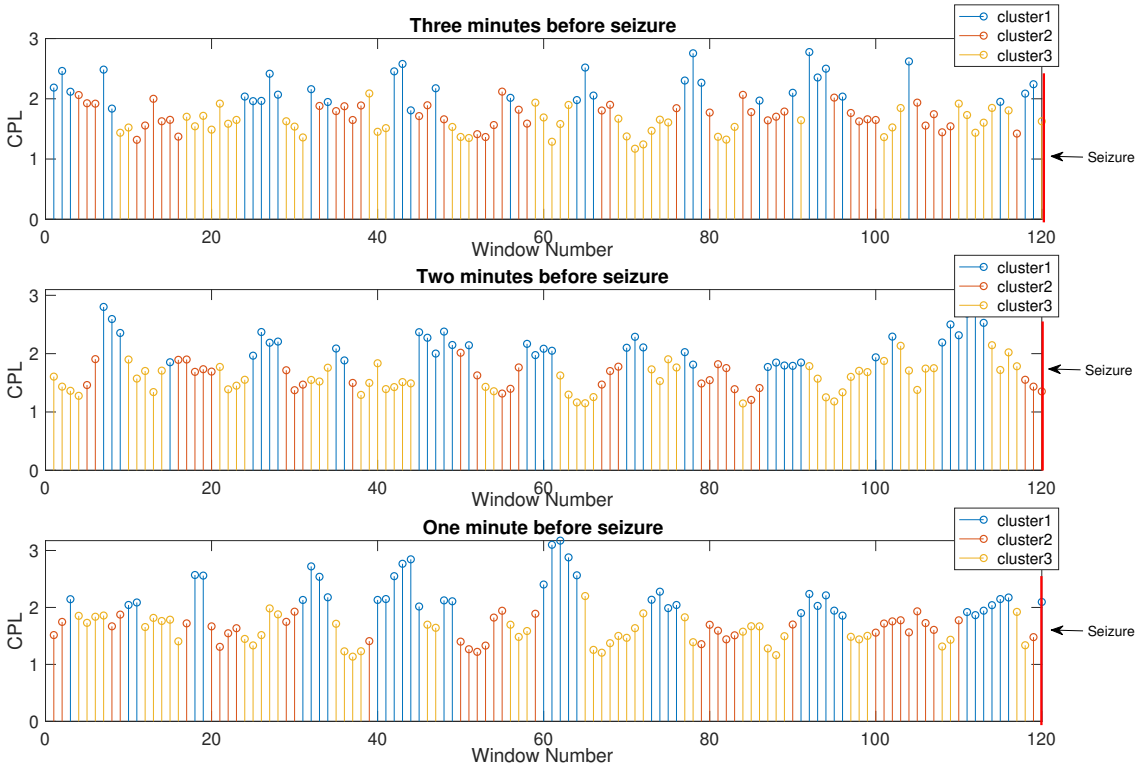


FIGURE 5.10: Variation in characteristic path length value of data in gamma band of Patient 1 for three clusters at 3 minutes before seizure.

Figure 5.10 and 5.11 are examples to show the variation in the characteristic path length value of two patients for three clusters over 1 minute at three points of time – 3 minutes, 2 minutes and 1 minute before seizure onset. The x-axes represent the number of windows, which is 120 windows per minute. Each window is 1 second in duration,

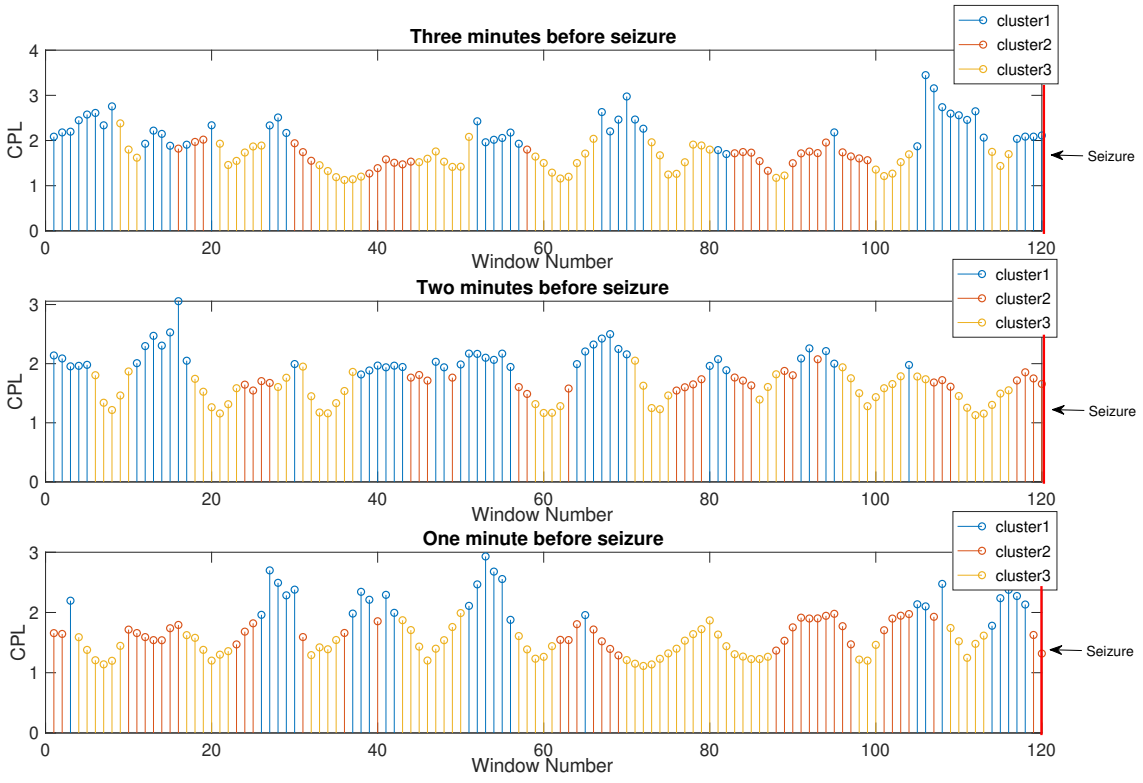


FIGURE 5.11: Variation in characteristic path length value of data in gamma band of Patient 2 for three clusters at 3 minutes before seizure.

with 50% overlap, which means 120 windows for 1 minute. It is apparent from both how the CPL value of Cluster 1 (the common pattern) was higher than that of the others. Both illustrate the time of Cluster 1's occurrence and how it was close to the onset of the seizure. The increase in the CPL value of Cluster 1 can be seen in both patients, confirming that the abnormal increase in CPL value is common to the patients. These procedures were undertaken to describe the network structure of the common pattern (Cluster 1) by exploring its features and comparing them to other clusters.

The four classifier models built to predict Cluster 1 were applied to 24 EEG data of 10 TLE and 10 healthy subjects. As mentioned previously, the classifier models were built using the 16 extracted features of Cluster 1 of the six seizure files of four TLE subjects as a training set, labelled 1. The extracted features of the six healthy subjects used in building a training set were labelled 0.

Table 5.3 illustrates the prediction time in minutes for 10 TLE subjects. All four models predicted Cluster 1 (common pattern) in all patients at around the same time. Bagged Tree model had the best sensitivity and specificity values (see Table 5.1), so it is the best choice to predict epileptic seizures for new patients. The amount of data used in building the classifier model was 1437 windows (0.4 hour) containing 717 windows for

Patient no.	Time before seizure onset (seconds)	LDA (minutes)	Logistic Regression (minutes)	SVM (minutes)	Bagged Tree (minutes)
1	300	4.96	4.88	4.96	4.96
2	300	4.96	4.96	4.96	4.96
3	300	5	5	5	5
4	70	1.08	1.08	1.17	1.12
5	70	0.83	0.83	0.83	0.83
6	50	0.3	0.3	0.3	0.33
7	40	0.25	0.27	0.27	0.27
8	100	0.57	0.58	0.67	0.57
9	180	1.67	1.67	1.67	1.67
10	20	0.8	0.8	0.8	0.8

TABLE 5.3: Prediction time of patients for common pattern using four classifiers.

the common pattern (cluster 1). The sensitivity was calculated using the metrics of the confusion matrix as seen in Table 5.4. The rate of incorrect predicted windows was equal 0.16. Because the amount of data used in building the model is 1437 windows (0.4 hour), the false prediction rate is 0.4/h (It calculated through dividing 0.16 by 0.4).

Figure 5.12 provides the 95% confidence interval of sensitivity for Bagged Tree model. It shows 95% probability that the true sensitivity value ranges within the interval. It was checked with different size of test data (number of windows used). It is apparent that the interval of sensitivity decrease with increasing the total number of tested windows.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	631	115
	Non-pattern	86	605

TABLE 5.4: Confusion Matrix of Bagged Tree Model

The maximum prediction time was revealed for Patient 3, with a recording time before seizure of 300 seconds. However, this is not the maximum time as, to create the training set, a limit of 5 minutes was imposed for this patient: their whole recording time before seizure onset was assessed to identify the maximum prediction time of 24 minutes. This time was explored at three different points of time: 24 minutes, 15 minutes and 5 minutes before seizure. Each part was 1 minute in length. The maximum prediction time was 23.99 minutes before seizure, and the pattern was detected at both 15 and 5 minutes beforehand. One of the issues noticed while exploring these three points in time was that the total number of records (number of windows) detected in the pattern increased as the seizure was approached. The miss-prediction rate was calculated for 10 healthy subjects by the four classifier models. Table 5.5 presents the rate of miss-prediction windows of 10 healthy subjects as non-pattern using four classifier models.

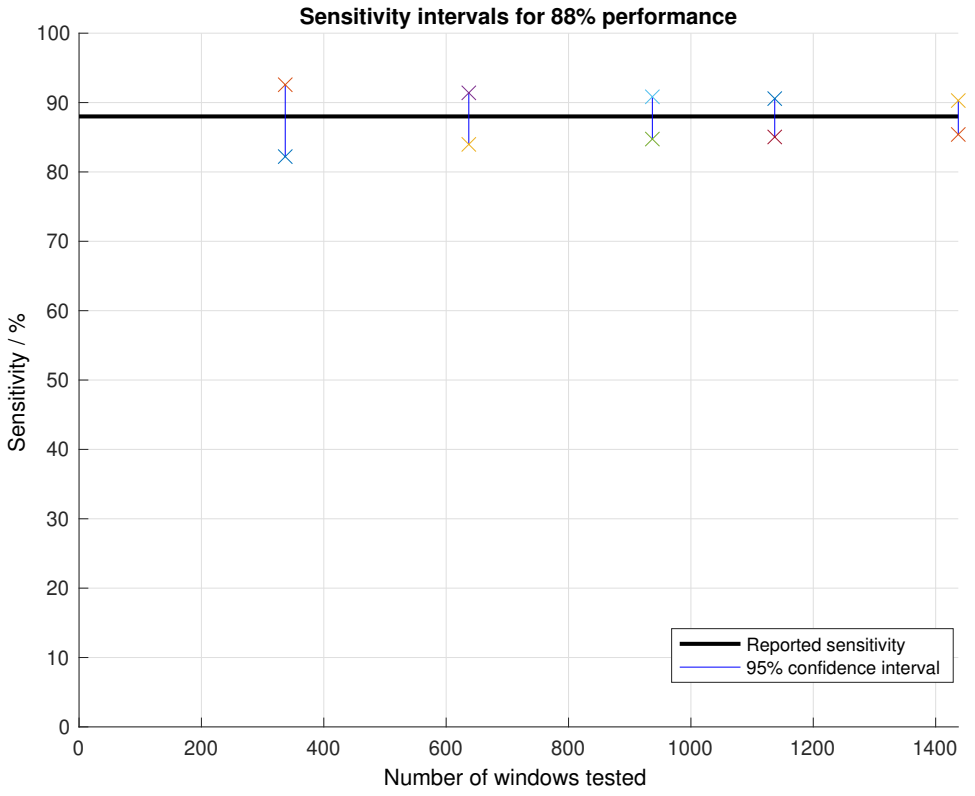


FIGURE 5.12: Estimated 95% confidence intervals of the sensitivity of the Bagged Tree model within different number of windows.

Subject	LDA	Logistic Regression	SVM	Bagged Tree
1	0.3	0.18	0.13	0
2	0.31	0.22	0.18	0
3	0.3	0.32	0.24	0
4	0.29	0.23	0.16	0
5	0.28	0.26	0.13	0
6	0.28	0.22	0.18	0
7	0.27	0.27	0.19	0.15
8	0.21	0.15	0.01	0.11
9	0.31	0.28	0.23	0.23
10	0.26	0.12	0.17	0.07

TABLE 5.5: The miss-prediction rate windows of 10 healthy subjects as non-pattern, using four classifier algorithms.

The miss-prediction rate for all 10 people without epilepsy was 0.056 for Bagged Tree classifier. Bagged Tree model exhibited a low miss-prediction rate, compared to other classifier models, and the best performance, sensitivity and specificity measures as seen in Table 5.1. It achieved 0.4/h as false prediction rate. The following section gives the results of predicting epileptic seizures using the synchronisation likelihood approach for

the same TLE patients.

5.2 The result of predicting using synchronisation likelihood (SL)

The SL approach was used to create a model to predict epileptic seizures. The same steps as those proposed in Figure 3.1 were taken. The model using SL was applied to the data in four frequency bands (alpha, beta, gamma and delta). In addition, the theta band was used with SL because the results produced by Douw et al. (2010) showed that it can be a more significant predictor in the diagnosis of epilepsy. The proportional threshold was applied to the network constructed to choose the percentage of the strongest connection preserved. Nine graph features were used for prediction: strength of node, betweenness centrality, global efficiency, characteristic path length, transitivity, modularity, assortativity and raduis). The four classifier models used with PLI were applied here to predict epileptic seizures. The following subsections describe the exploration of temporal change in SL within each window, before and during seizure.

5.2.1 Exploring the change in synchronisation likelihood (SL)-based networks

The temporal dynamic of the SL was explored 60 seconds before seizure, 10 seconds before and during the seizure in five bands of data at each point. For the data in the alpha band, there was no clear variation in the value of SL, or even an abnormal change.

In contrast, data in beta and gamma band exhibited a variation in the value of SL, shown as an increase in the connectivity value at a specific time in all six seizure files of the four TLE subjects used as a training set at 10 seconds before seizure. Figure 5.13 shows an example of this increase for one of the patients in the gamma band. Unfortunately, this increase was not detected in the remaining TLE subjects. The remaining data in the delta and theta bands did not show any noticeable increase either before or during a seizure. For further verification of the dynamic change in SL value, the two statistical measures used earlier with PLI in this chapter were subsequently applied.

Figure 5.14 shows the average value of SL before and during seizure for the data in the beta, gamma, delta and theta bands. The figures show the data of three TLE subjects as an example. The x-axes present the window number where SL average was calculated. In the data of the four bands, the SL value of Patient 1 exhibited the highest average value, apart from in its gamma band. In Patient 2, the value of the data in the gamma band increased at the start of the seizure then fell after 5 seconds relative to other patients. From these observations, which were not common to the patients, checking

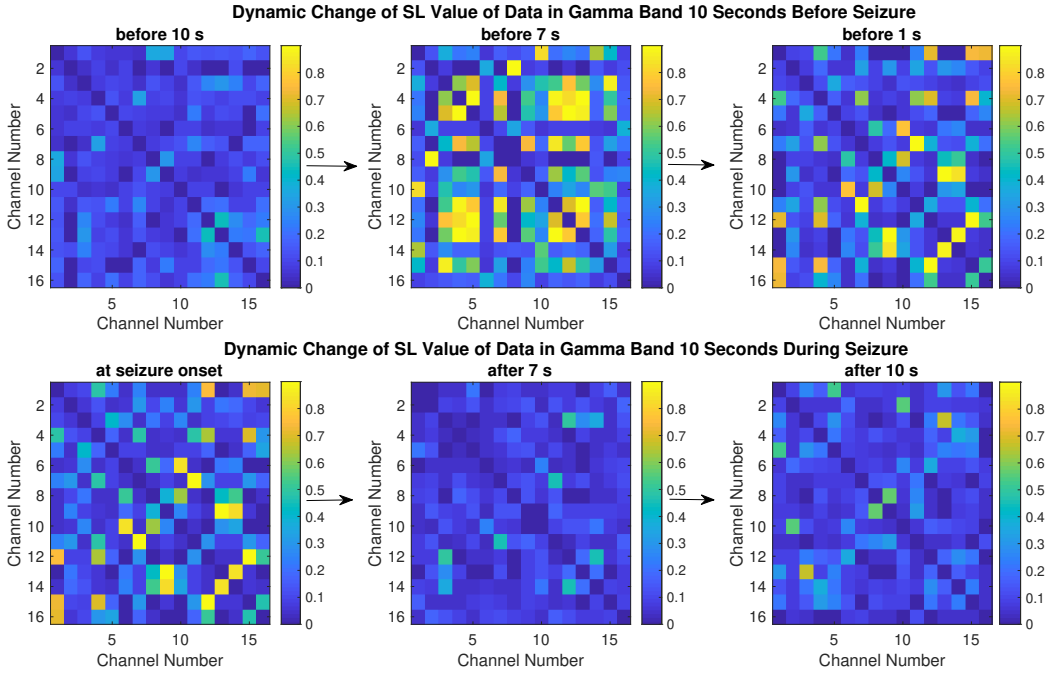


FIGURE 5.13: Variation in SL value of data in gamma band 10 seconds before and during seizure.

the temporal change in the average value of SL does not detect any abnormal change that is shared by all patients.

The second way to assess the increase in SL value is to calculate the total number of synchronised nodes in each network (window) whose SL value exceeds 0.8. Because the increased level of SL value was less than PLI, 0.8 was selected as the threshold here. Only connectivity values above 0.8 were considered for the synchronised nodes in each window. Figure 5.15 shows the dynamic change in the total number of synchronised nodes with a value above 0.8 in each window at exactly 5 minutes before seizure.

It can be seen in the figure on the left, for data in the beta band, that only one patient of the three had peaks representing high values of SL in around 25 nodes out of 256. The figure on the right represents the data in the gamma band, and it shows only two patients of the three had a high increase in SL value for around 25 nodes. It is apparent that the total number of synchronised nodes with an SL value exceeding 0.8 decreased a couple of seconds before seizure for the data in the beta band. The small total number of synchronised nodes with SL value exceeding 0.8 may be linked to the existing high SL value for a few nodes (channels) in the brain network related to the seizure's origin. Moreover, the increase might be shown earlier, in other patients. Therefore, if this small number is to be considered as an abnormal change for the purpose of prediction, it needs to be evident in all patients. Additionally, the total number of high synchronised nodes,

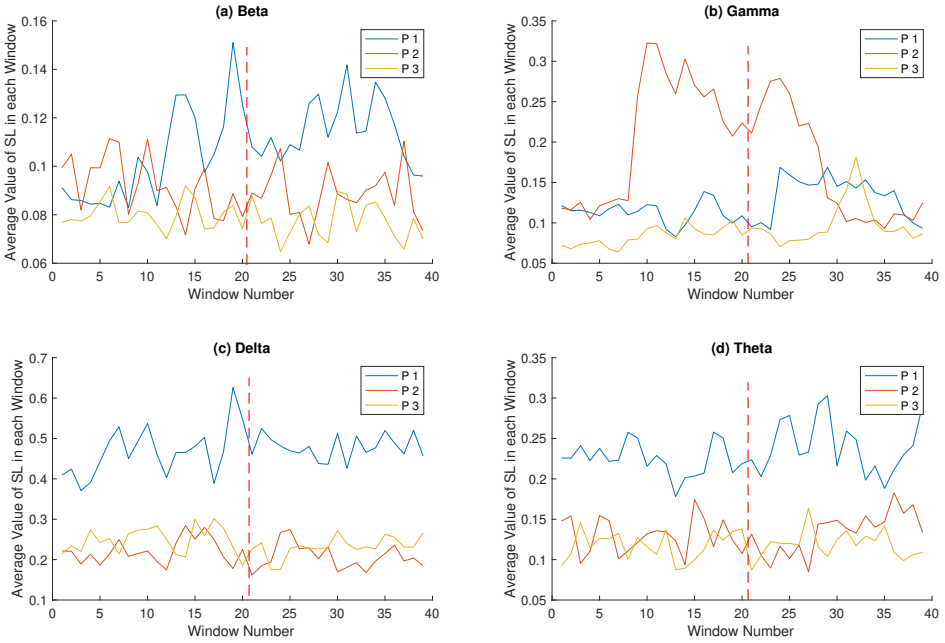


FIGURE 5.14: Average SL value in each window of data in four bands at 10 seconds before and during seizure. The dashed red line represents the start of seizure. The total number of windows at 10 seconds before and during seizures was 38 windows (Window 19 refers to the calculated SL average window immediately before seizure).

which was around 25, can be used as abnormal change, if it is not presented in healthy subjects.

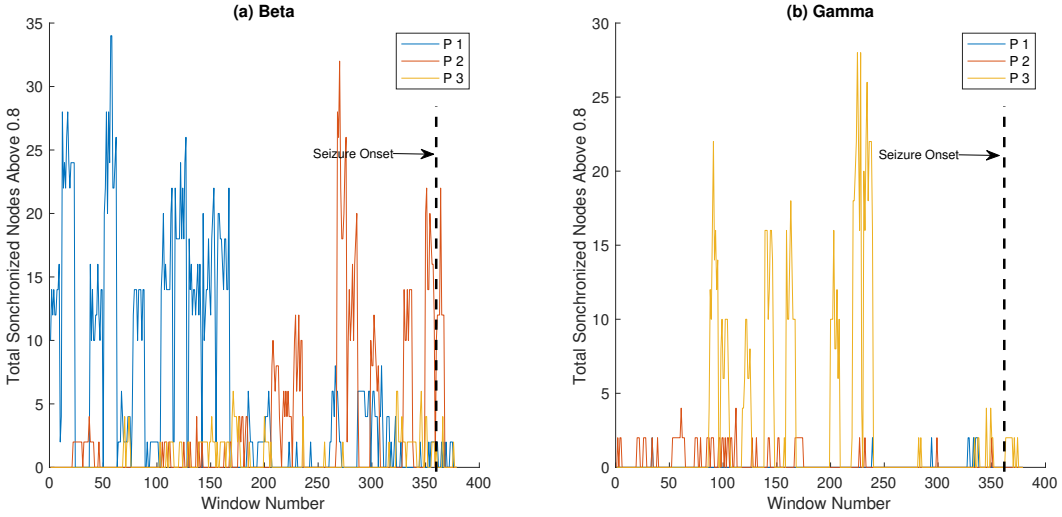


FIGURE 5.15: Total number of nodes with SL values exceeding 0.8 in three TLE subjects for data in the beta and gamma bands: the dashed black line refers to the beginning of the seizure.

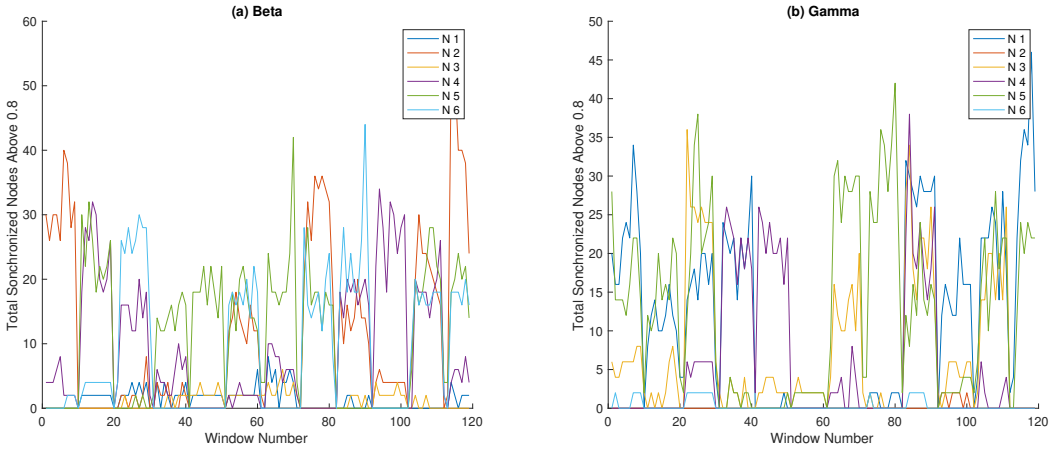


FIGURE 5.16: Total number of nodes with SL values exceeding 0.8 in six healthy subjects for data in the beta and gamma bands

The six healthy subjects in Figure 5.16 showed peaks in the total number of synchronised nodes with SL above 0.8, exceeding 30. This value is considered higher than those observed in patients. Therefore, both statistical measures failed to show a common abnormal change among patients. The next step was to extract features for the networks created using SL.

5.2.2 Exploring network features

Nine features were extracted to describe the temporal change of the network structure. Figure 5.17 shows the CPL value variation within 5 minutes before onset of seizure for six TLE subjects. The x-axes refer to the window number where CPL value was calculated (e.g. Window no. 1 is 5 minutes before seizure onset).

By comparing the CPL values of the TLE subjects with the values of the healthy subjects in Figure 5.18, it isshown that CPL values of healthy subjects reached a similar trend value. In addition, there was no common pattern of values (increasing or decreasing) among patients. Similarly, the data of both the TLE and healthy subjects in the remaining bands showed no visible pattern.

The results obtained by exploring the variation in the transitivity and modularity values of data in both the beta and gamma bands showed random variation in values without presenting any abnormal behaviour that is shared among patients. Figure 5.19 illustrates these randomised values of the transitivity of data of six TLE subjects in the beta band within 5 minutes of the onset of the epileptic seizure. It was noticed that there was no common pattern of change among patients. For the betweenness centrality feature, four nodes of the network that represent the right temporal lobe (the origin of seizure) and around it were checked.

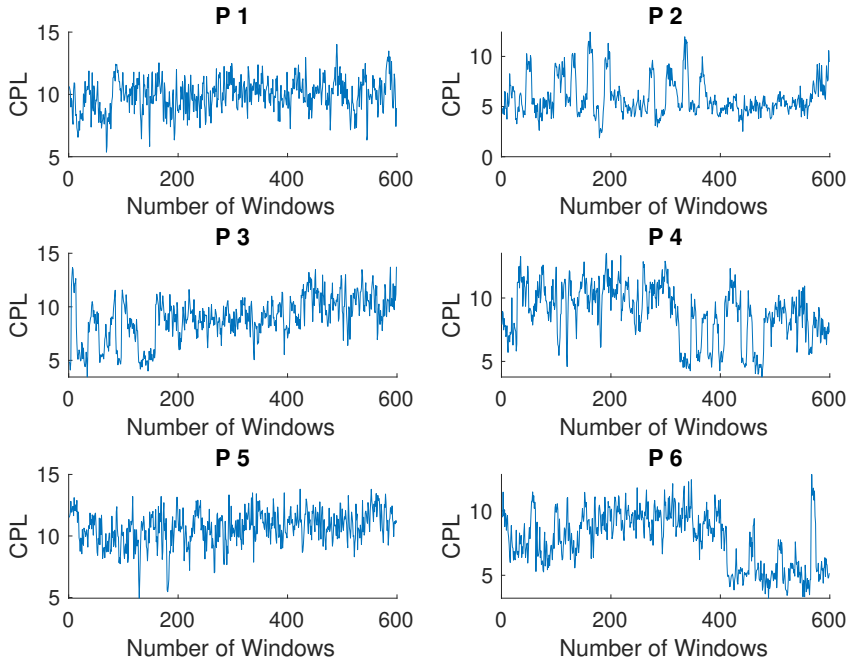


FIGURE 5.17: Variation in CPL value of data in beta band, 5 minutes before seizure for six TLE: Window no. 600 represents the last window calculated before seizure onset.

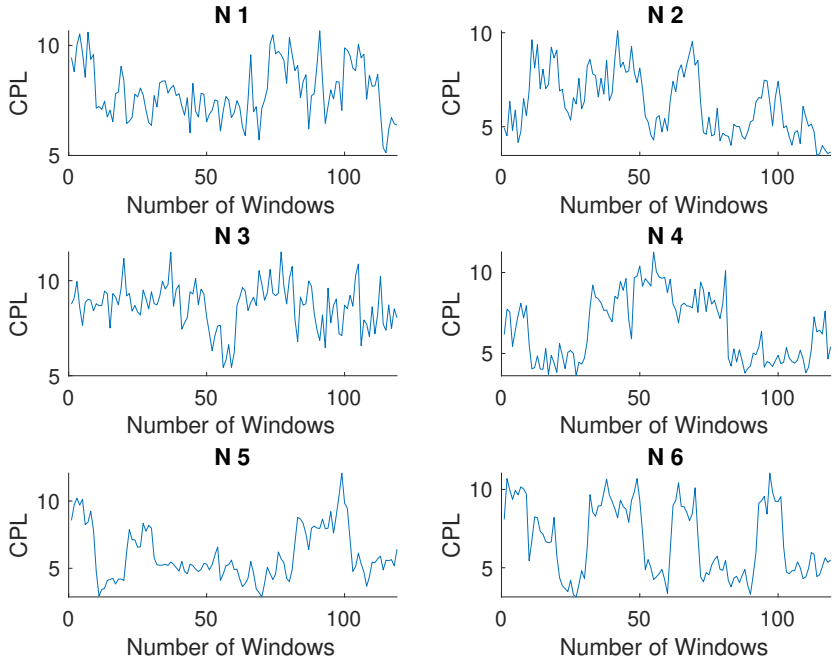


FIGURE 5.18: Variation in CPL value of data in beta band, 5 minutes before seizure for six healthy subjects.

Figure 5.20 presents the variation in betweenness centrality value of nodes T4 and T6 for six TLE subjects in the beta band at 5 minutes before the onset of the epileptic seizure.

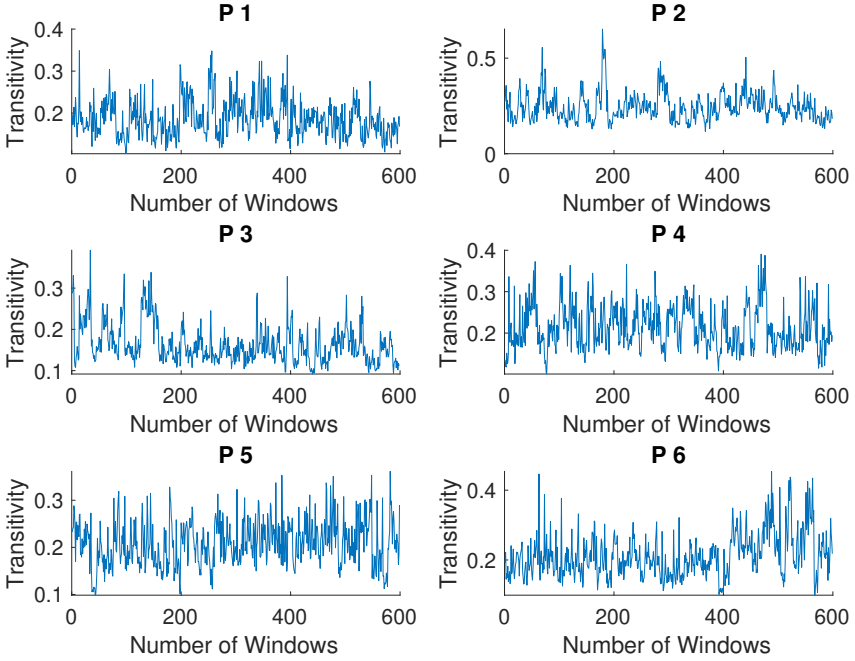


FIGURE 5.19: Transitivity value of data in beta band for 6 patients, at 5 minutes before seizure.

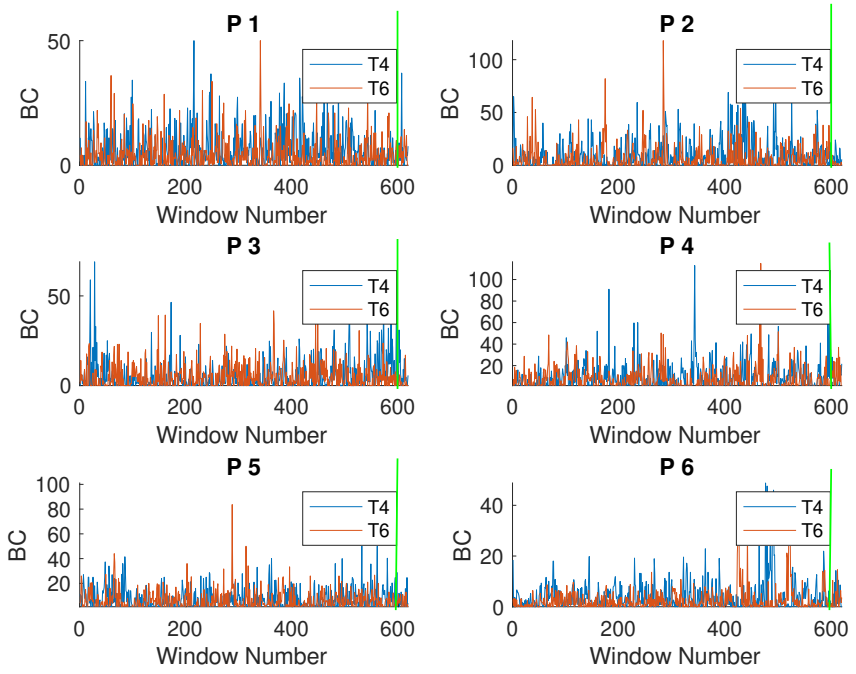


FIGURE 5.20: Betweenness centrality of data in beta band for six TLE subjects, 5 minutes before seizure onset.

The green line indicates the onset, where it can be seen that the value of betweenness was under 50 in all patients. The presence of peaks prior to epileptic seizures, which

is the increase in the value of the betweenness centrality to more than 100, cannot be considered as abnormal behaviour used in creating a prediction method for seizures, because there are peaks that exceed 100 in the data of healthy subjects.

None of the remaining features (clustering coefficient, radius, and global efficiency) showed a common pattern among patients. As a result, checking the network features did not contribute to finding any abnormal change for use in predicting seizures. In a subsequent step, the data of the extracted features were clustered to find any hidden pattern.

5.2.3 Clustering data and predicting seizures

Before applying clustering to the extracted features, they were normalised to eliminate any bias arising from the range of values of features, which could affect the clustering. Following this, one of the two most correlated features was eliminated. As a result, the features were used to characterise networks and included in the clustering process were assortativity, betweenness centrality of each of the 16 channels (network nodes), clustering coefficient, transitivity, modularity, characteristic path length, and global efficiency. Therefore, the total number of features from data in all five bands was 110 features. The same six TLE subjects were used as a training set with PLI selected to extract a hidden pattern common to the patients. Because of the variation noticed in the value of SL, represented by a slight change in the value (increasing) of the data in both the beta and gamma bands only, the features of these bands were checked together first in clustering.

K-means was selected to cluster the data, because it can deal with extremely large amount of data. It needs to be provided with the number of clusters as prior knowledge, and so the elbow method was used to select the optimal number of clusters. Unfortunately, there was no clear 'elbow', referring to the optimal number of clusters, by using the extracted features of data in the beta and gamma bands together. Hence, the features of the data in all five bands were selected for clustering. The optimal numbers of clusters of data in the five bands, using the elbow method, were 10, 10, 8, 10 and 7 for the data in the beta, gamma, delta and theta bands, respectively.

For each of the data sets in the five bands, a classifier model was created to check the possibility of its presence in TLE and healthy subjects. Four classifier models were used (LDA, Logistic Regression, SVM and Bagged Tree). A label of 1 was attached to the data of the TLE subjects belonging to the selected cluster, while 0 was attached to the healthy subjects' data. By applying all classifiers to predict all clusters of data in all bands, Cluster 10 of the data in the delta band was dominant, becoming the common pattern. It was shown in all TLE subjects and in only four of the 10 healthy subjects in the study.

Cluster	LDA	Logistic Regression	SVM	Bagged Tree
1	87.1%	86.6%	88.6%	87.8%
2	66.2%	66.9%	90.8%	90.5%
3	53.1%	82%	91.9%	92.5%
4	82.3%	82.4%	92.8%	91.7%
5	82.3%	83.7%	94.7%	94.7%
6	82.5%	85%	95.9%	95.6%
7	84.6%	85.7%	94.3%	93.4%
8	84.8%	86.1%	96%	96.2%
9	87.1%	88.8%	96.9%	96.3%
10	88.5%	92.6%	92.7%	92.7%

TABLE 5.6: Performance (accuracy) of the four classifier models on data in delta band.

Table 5.6 presents the performance (accuracy) of the four classifier models used in predicting 10 clusters of data in the delta band. They were assessed using the five-fold cross-validation method. The table shows the closest percentage values of all classifiers, with a slightly high value for Bagged Tree and SVM. The sensitivity and specificity measures were calculated for the four classifiers to predict Cluster 10 of the data in the delta band (see Table 5.7).

no.	LDA		Logistic Regression		SVM		Bagged Tree	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
1	82%	89%	72%	91%	80%	92%	71%	93%
2	84%	93%	77%	95%	81%	95%	76%	95%
3	84%	86%	71%	93%	78%	94%	72%	95%
4	81%	90%	66%	92%	75%	92%	92%	86%
5	90%	90%	80%	93%	86%	91%	79%	94%
6	85%	91%	76%	93%	84%	92%	76%	95%
7	74%	90%	71%	92%	74%	92%	52%	95%
8	89%	90%	80%	93%	82%	93%	77%	95%
9	81%	91%	72%	92%	80%	94%	78%	93%
10	87%	91%	77%	93%	78%	94%	76%	95%

TABLE 5.7: Sensitivity and specificity of the four classifiers for 10 Clusters. first column refers to the number of cluster

As seen in Table 5.7, the specificity in all classifiers was higher than the sensitivity, which means less prediction of the selected cluster. The difference between specificity and sensitivity was large, in some cases. Therefore, a trade-off of these two measures was considered when the classifier had been selected for the remaining TLE and healthy subjects. For example, SVM in Table 5.7 was selected to predict Cluster 1, and LDA was chosen to predict Cluster 10.

Table 5.8 compares Cluster 10 with three clusters for data in the delta band, using the mean value of seven network features of the nine used to describe connectivity networks.

	Cluster 1	Cluster 2	Cluster 5	Cluster 10
CPL	0.1460	0.2216	0.2828	0.2802
Modularity	0.2092	0.0996	-0.0789	-0.2547
Transitivity	-0.3096	-0.3161	-0.1449	-0.0030
Clustering coefficient	-0.1276	-0.1894	-0.2593	-0.3519
Betweenness of node T4	-0.2631	-0.1575	-0.1345	2.1257
Betweenness of node T6	-0.1185	-0.1065	-0.2224	0.0681
Assortativity	-0.1030	0.1914	-0.0942	-0.0944

TABLE 5.8: Mean values of five features in four clusters.

As can be seen, Cluster 10 exhibited high values of betweenness centrality for nodes T4 and T6, compared to other clusters. Similarly, the transitivity value of Cluster 10 was slightly higher than that of others.

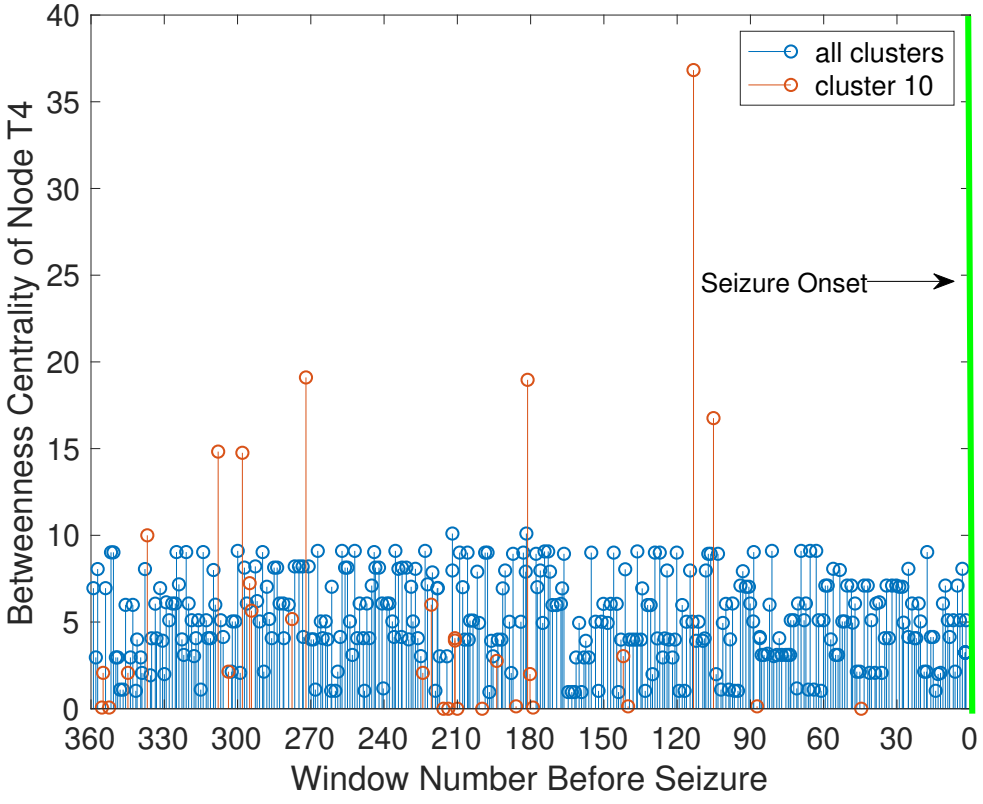


FIGURE 5.21: Variation in betweenness value of node T4 within Cluster 10 (common pattern) compared to other clusters.

Figure 5.21 shows the variation in betweenness value of node T4 within Cluster 10 (common pattern) compared to other clusters. It shows the point in time where Cluster 10 occurred. The highest peaks of values within Cluster 10 are represented by a red line. This increase in the value at these points before seizure, which represents an abnormal change, refers to the existing uncontrolled synchronisation around this node. Because

the patients used in the current study had temporal lobe epilepsy originating in the right lobe, this sudden increase in the betweenness value of node T4 can be considered as a sign of an oncoming seizure. The proposed model created using the LDA classifier was selected to predict epileptic seizures of the remaining patients, based on the range of values of sensitivity and specificity shown in Table 5.7 for Cluster 10.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	998	65
	Non-pattern	149	655

TABLE 5.9: Confusion Matrix of LDA Model

Table 5.9 illustrates the metrics values of confusion matrix used in calculating sensitivity and specificity. The amount of data used in building the classifier model was 1867 windows (0.52 hour) containing 1147 windows for common pattern (cluster 10). The sensitivity was calculated using the metrics of the confusion matrix as seen in Table 5.9. The total rate of incorrect predicted windows was equal 0.09. Because the amount of data used in building the model is 0.52 hour, the false prediction rate is 0.17/h. It was calculated throuh dividing 0.09 by 0.52.

Again, the 95% confidence interval of sensitivity has been estimated. The same notice was showed with PLI, here the amount of tested windows affect the sensitivity value (Figure 5.22).

Patient no.	**Time before seizure onset (seconds)**	**Maximum prediction time (minutes)**	**Minimum prediction time (seconds)**	**seizures used**	**seizures predicted**
1	300	4.43	8	2	2
2	300	4.89	1	1	1
3	300	4.95	1	3	3
4	70	2.89	4.5	2	2
5	70	1.11	6	2	2
6	50	0.97	5	2	2
7	40	0.73	8	2	2
8	100	0.39	6	3	3
9	180	2.16	3	2	2
10	20	0.63	2	4	4

TABLE 5.10: Prediction time of Cluster 10 for 23 seizures of 10 TLE subjects.

Table 5.10 shows the prediction time of 10 TLE subjects used to predict the common pattern (Cluster 10). The third and fourth columns illustrate the maximum and minimum prediction time for each patient, respectively. As a result, the average of maximum predicting time for all patients was 1.7 minutes, while the minimum was 9.6 seconds. A total of 5 minutes' recording time before seizure for Patient 3 was chosen in building the

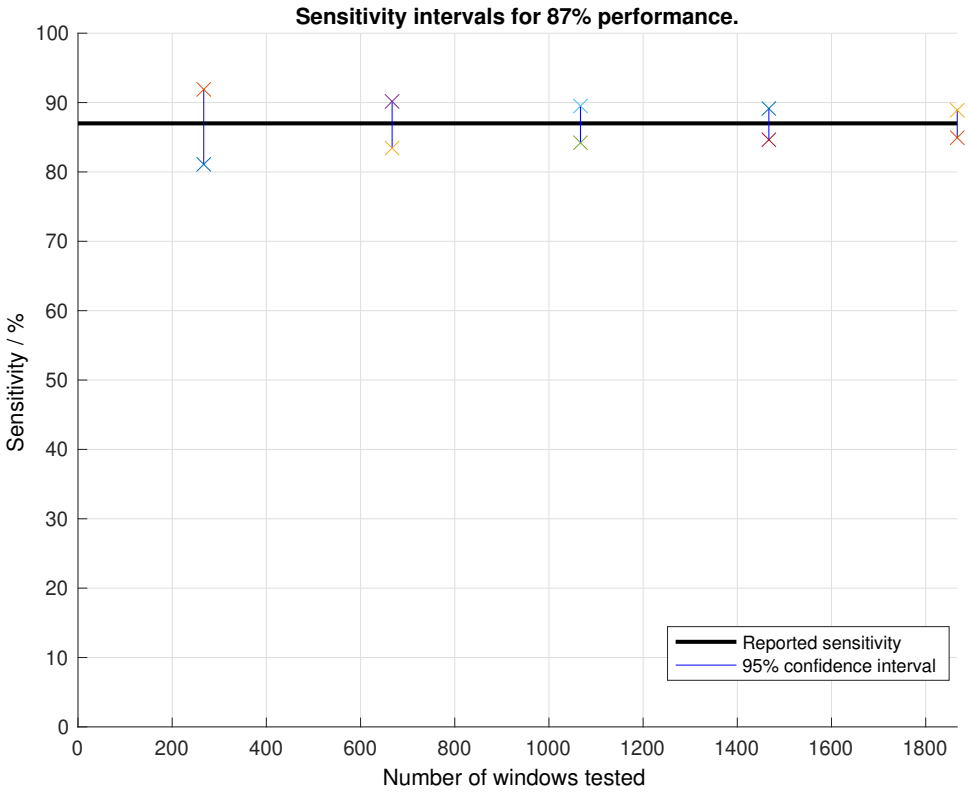


FIGURE 5.22: Estimated 95% confidence intervals of the sensitivity of LDA model within different number of windows.

training set of the classifier model for prediction, as seen in Table 5.10. However, Patient 3 had a recording time before seizure of more than 24 minutes, but only 5 minutes before seizure was chosen in building the training set of the classifier model for predicting. To see the maximum prediction time for Cluster 10, the first 3 minutes of Patient 3's recording time were checked. The result showed that Cluster 10 could be predicted 24 minutes before seizure.

The miss-prediction rate was calculated for each healthy subject which was 0.17 as average of all. The false prediction rate of LDA classifier model was 0.17/h, which is lower than that of the predicted model in one study (Alotaiby et al. 2017) (0.39/h) and equal to that found in another (Aarabi & He 2012) (0.17/h). In general, the proposed predictive model using LDA did not demonstrate a substantially better performance – high sensitivity and low specificity – than that proposed using PLI. However, the above model is considered substantial on the basis of its results produced (24 minutes prediction time, and a low false prediction rate of 0.17/h) compared to those in other studies (Chu et al. 2017, Cho et al. 2016), which achieved 0.367/h and 0.5/h respectively.

5.3 Comparison of performance between PLV, PLI and SL approaches

As mentioned in Chapter 1, three approaches (PLV, PLI and SL) were chosen to assess the uncontrolled synchronisation and to create a connectivity network to predict seizures. After applying PLV in the previous chapter to predict seizure, the PLI approach was used here to predict seizures for the same patients. Four classifier models were used separately to formulate a predictive model. They employed the network features of data in both the beta and gamma bands. They were selected for their slightly elevated PLI value, observed only in the data of these two bands before seizure. They were applied to 10 TLE and 10 healthy subjects to predict each cluster. Bagged Tree model was the best model, with high specificity, sensitivity and accuracy (see Chapter 2 for more describing about model). Cluster 1 was the common pattern, with a false prediction rate of 0.4/h. The most interesting finding was the high value of CPL in Cluster 1 compared to other clusters. Figure 5.10 and 5.11 illustrate this increase in the CPL value. This increase in the CPL value during Cluster 1 refers to changes in the network structure. A possible explanation may be the increase in brain activity represented by a high level of uncontrolled synchronisation. It is interesting to note that the betweenness centrality value of the right temporal lobe nodes (T4, T6) was higher within Cluster 1 (see Table 5.2). The predictive model achieved a maximum prediction time of 23.99 minutes. In comparison with earlier methods mentioned in Table 2.2, the proposed model achieved a substantial prediction time of 23.99 minutes before seizure. Indeed, it achieved a prediction time longer than those found in other studies (Cho et al. 2016, Ouyang et al. 2007, Le Van Quyen et al. 2001). It was a logically accepted time, compared to the 60 minutes achieved in (Iqbal et al. 2015). Additionally, the sensitivity of the prediction method was 88%, higher than that in (Cho et al. 2016, Li et al. 2013, Aarabi & He 2012). The false prediction rate was 0.4/hr, lower than that provided by one existing model (Cho et al. 2016) (0.5/hr). The model achieved sensitivity (88%) for 24 seizures used which are all successfully predicted. As a result, the model created using PLI was substantially better than the current proposed models in Table 2.2. One of its limitations was the number of seizures studied for each patient, which was only one.

A predictive model was created using the SL approach and applied to the same TLE subjects used with PLI. The extracted features of data in the five bands were explored (alpha, beta, gamma, delta and theta). The four classifier models used with PLI were employed to formulate a predictive model using the clustered features of the created networks. They used to predict each cluster in the data of each band. Cluster 10 of the delta band was predicted in all patients and in only four of the 10 healthy subjects. Conversely, other clusters were predicted in both the TLE and healthy subjects. Cluster

10 of the data in the delta band (the common pattern) was predicted substantially better by the LDA classifier, in terms of its sensitivity and specificity measures (see Table 5.7).

The most interesting result was the value of the betweenness centrality of the right temporal lobe nodes T4 and T6 (seizure origin) within Cluster 10, as shown for node T4 in Figure 5.20. Similarly, this increase in the value of betweenness centrality was found in the common pattern predicted by the PLI approach. A possible explanation may be the presence of abnormal brain activity in the right temporal lobe represented by uncontrolled synchronisation. The predictive method achieved an average prediction time of 1.7 minutes and a maximum of 24 minutes for Patient 3. Again, the model gave a substantial prediction time of 24 minutes compared with the time in other studies Aarabi & He (2012), Iqbal et al. (2015). The false prediction rate was low (0.17/hr) compared with the methods mentioned in Table 2.2. In addition, the number of patients used in the current method was higher than used in Zandi et al. (2010), Iqbal et al. (2015). As a result, the predictive model created using SL is substantially better than existing models. Table 5.11 summarise the results of the models developed using the three approaches.

Approach	**Sensitivity (%)**	**Specificity (%)**	**Network Features used in formulating model**	**Classifier models used**
PLV	94,92	97,98	CPL in gamma band	Linear SVM and Logistic Regression
PLI	88	84	CPL, Modularity, Betweenness Centrality of nodes (T4, T6, P4, and C4), and Assortativity for both beta and gamma bands	Bagged Tree
SL	87	91	CPL, Modularity, Transitivity, Clustering Coefficient, Assortativity, Betweenness Centrality of nodes (T4, T6, P4, C4) in delta	Linear Discriminant Analysis

TABLE 5.11: Summary of all prediction models created in this study.

5.4 Conclusion

All the prediction models created by the three approaches (PLV, PLI and SL) were tested on data from TLE patients. All achieved a substantial performance - a high sensitivity and low FPR. However, the sample of patients used – 10 patients – compared

to that used in existing studies (Chu et al. 2017, Li et al. 2013) (16 and 21 respectively) is considered small. Another is the short duration of the recording time before seizure onset in most patients – less than 5 minutes.

By the end of this chapter, the first objective of the thesis has been achieved. The other objective of the study was to use these approaches to predict epileptic seizures for patients with generalised epilepsy- absence seizures in our case. The next chapter gives the results of predicting absence seizures using PLV, PLI and SL approches.

Chapter 6

Predicting Absence Seizures Using PLV, PLI and SL

As mentioned in Chapter 1, the main goal of this study is to establish the feasibility of formulating a generalised model to predict epileptic seizures of both types (partial and generalized epilepsy). To achieve this, a work has been done to create models for predicting epileptic seizures for each type. These prediction models are now complete for partial epilepsy, using EEG signals of TLE subjects from each of the three approaches of estimating synchronization (PLV, PLI and SL) (as described in Chapters 4 and 5). In this chapter, models are formulated to predict epileptic seizures for generalized epilepsy by using the EEG signals of patients with Absence seizures - a form of generalized epilepsy. It explores the procedures adopted in building the models through each of the three approaches mentioned above.

6.1 Predicting absence seizures using PLV

All the pre-processing steps employed with the TLE subjects described in chapters 4 and 5 have been applied to the patients with generalized epilepsy. After this, the first step was to check the temporal variation in the PLV values of patients in Groups 1 and 2 (see Chapter 3 for more details). Two measures were calculated to assess the change before and during a seizure: the average of the PLV values of all nodes in the network within each window of length 1 second; and the total number of synchronized nodes whose PLV value exceeds 0.9 within each window. This has been explored for each of the five bands – alpha, beta, gamma, delta and theta.

The temporal changes in the average values that were observed were assessed by examining the values of the nine EEG data of patients in Group 1 with those in Group 2. The comparison sought to find a trend, shared among the nine EEG data used as training

set, that was absent from the data of patients in Group 2. Table 6.1 below summarizes the results.

Band	**PLV average**	**Total no. of synchronized nodes exceeding 0.9**
Alpha	No pattern	No pattern
Beta	No pattern	No pattern
Gamma	No pattern	No pattern
Delta	No pattern	All patients >40
Theta	No pattern	All patients >45

TABLE 6.1: PLV average and total number of nodes with a PLV above 0.9, in each of the five bands.

No common trend in the values appeared among patients in Group 1 regarding the temporal change in the average PLV value in each window. On the other hand, checking the total number of nodes whose PLV value exceeded 0.9 in the delta and theta bands only showed trend in values. These values were detected in all nine EEG data in Group 1. They were checked during seizure in both bands, and the results showed that the same pattern appears both before and during seizure, in only the delta band. Nonetheless, the extracted features of both bands were used to find a common pattern and build a prediction model. Figure 6.1 shows the results of measuring just the delta band data for the two patients from each group of the training set. The window number shown in Figure 6.1 (b) represents the total length of recorded data (82 seconds). It is apparent from Figure 6.1 (a) that the value before and during seizure exceeded the trend of 40, while in (b) it remained lower. This observation is considered later in formulating the prediction model.

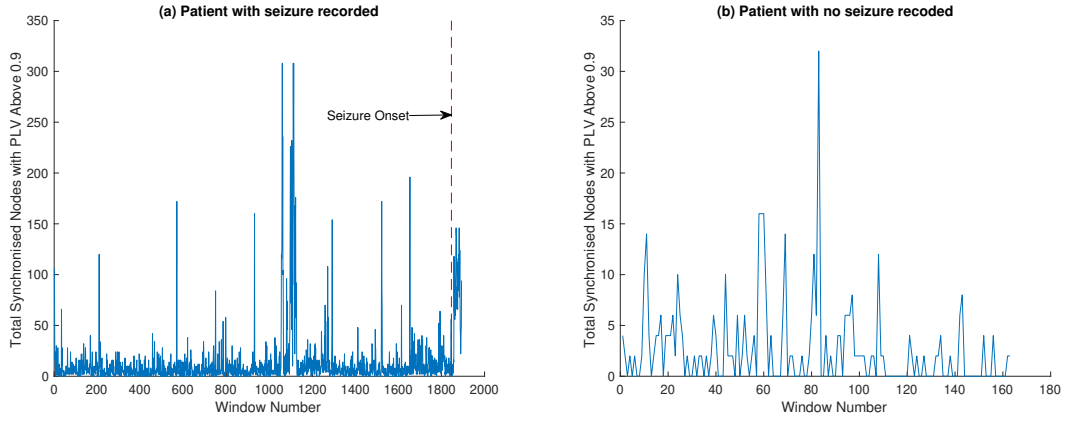


FIGURE 6.1: Result of calculating the total synchronised nodes with a PLV value above 0.9 for data in delta band, for both groups of patients. (a) refers to group 1 of patients with seizures recorded and (b) refers to group 2 of patients with no seizures recorded

6.1.1 Extracting graph theoretic parameters of PLV-based networks

Because of the trend noticed in the values in these patients in both the delta and theta bands, their PLV-based networks were characterized by extracting eight features, as shown in Table 6.2, demonstrating which displayed a trend that was common to only the patients in Group 1, not in Group 2.

	Delta	**Theta**
Radius	-	-
Global Efficiency	> 0.7	>0.65
CPL	<1.6	-
Modularity	-	-
Transitivity	>0.7	-
Betweenness Centrality	Nodes CZ-PZ, FP2-F8, P4-O2, T4-T6, and T5-O1 exceeded 60	Nodes FP2-F4, T4-T6, and T5-O1 exceeded 65
Assortativity	$>=0.1$	-
Strength of Node	-	-

TABLE 6.2: Trend values of the extracted features in delta and theta bands.

It is apparent that the common trends lie in global efficiency, CPL, transitivity and assortativity, in the delta band. In addition, the betweenness centrality of five nodes exhibit a shared trend. On the other hand, for the data in the theta band, global efficiency and betweenness centrality show a common trend in only three nodes. The features exhibiting trends in values were selected to build the prediction model. The total number of synchronised nodes whose PLV values exceeded 0.9 – as checked earlier – were added only to the features of delta band, because the same trend was detected both before and during a seizure.

6.1.2 Formulating a prediction model using PLV

Only the network features showing a shared trend were used in the clustering process for both bands to extract the hidden pattern, as shown in Table 6.2. Global efficiency was removed from the data in the delta band because of its strong correlation to other features. Therefore, the features used with the delta band were: Assortativity, Betweenness centrality of five nodes: FP2-F8, P4-O2, T4-T6 and T5-O1, Transitivity, CPL and the total number of synchronized nodes whose PLV value is above 0.9 in each window. Only global efficiency and the betweenness centrality of nodes were used for the data in the theta band, as shown in Table 6.2. K-means was used to cluster the data in both bands. The optimal number of clusters was four, as indicated by the elbow method, for data in the delta band, while the method failed to establish an optimal number of clusters in the theta band. Therefore, the silhouette and gap statistic methods were applied to both. The result was 2 and 20 clusters for data in the delta band, and 2 and

11 clusters for the data in the theta band. Because there was no consensus achieved by these methods, the 30 indices used in Chapter 4 were applied to determine the optimal number of clusters. The results of applying all 30 indices to data in the delta and theta bands respectively leads to four and two clusters.

Three classifier models were applied, using the cluster number as a label for classification. They were applied to the nine EEG data records for the six patients used in Group 1 and to the two patients in Group 2 group used as training set. For data in theta band, no common pattern was found. For the delta band, Clusters 1 and 4 are shown only for the two patients without seizures. The remaining clusters – Clusters 2 and 3 – were not apparent in these two patients, and are described in Table 6.3, showing the mean values of the features in each. It is apparent that Clusters 2 and 3 have lower values of CPL and betweenness centrality than Clusters 1 and 4. The highest total numbers of synchronized nodes whose PLV value exceeds 0.9 are in Clusters 2 and 3.

Network feature	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**
Assortativity	-0.04	-0.03	-0.05	-0.03
Betweenness Centrality of node (P4-O2)	7.45	5.85	2.63	6.82
Betweenness Centrality of node (FP2-F8)	8.49	4.47	5.02	6.33
Betweenness Centrality of node (T4-T6)	9.42	5.85	4.55	7.77
Betweenness Centrality of nodef (T5-O1)	8.27	6.08	1.45	7.59
Betweenness Centrality of node (CZ-PZ)	8.65	6.09	1.33	7.77
Transitivity	0.43	0.65	0.85	0.54
CPL	2.62	1.82	1.33	2.1
No. of nodes with PLV above 0.9	4.38	64.47	202.12	24.18

TABLE 6.3: Mean values of the clustering features within the fourth clusters.

Applying the three classifier models to patients in Group 1 showed that only Cluster 2 was predicted in all patients. Therefore, Cluster 2 was selected as the common pattern. Table 6.4 shows the performance of the three classifiers in terms of accuracy, sensitivity and specificity for data in the delta band.

Classifier	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
Complex Tree	100	100	100
Linear SVM	99.75	97	97
Bagged Tree	99.9	100	98

TABLE 6.4: Performance of three classifier models for clustered features in delta band.

Figure 6.2 presents the results of applying the classifier models to the patients in Group 1. It shows a transition in the four clusters during seizures. It is apparent that Cluster 2 is evident in all seizures, and for all patients is present both before and during the onset of a seizure.

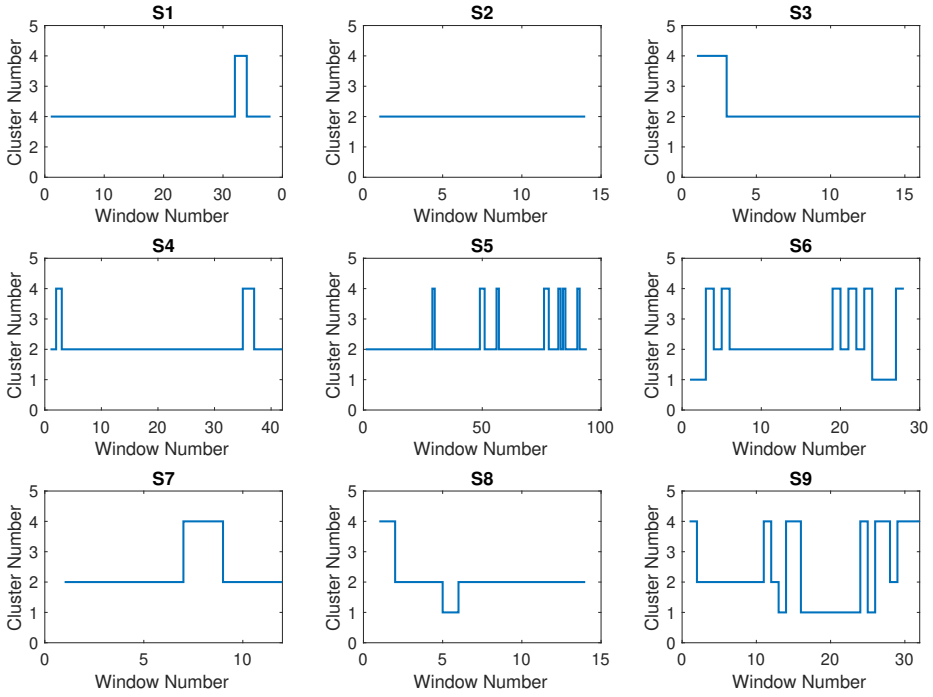


FIGURE 6.2: Illustration of transition of the four clusters during 9 seizures for data in delta band. The shortcut s1 refers to the seizure file one

Based on the results in Table 6.4, the Complex Tree model was used to predict Cluster 2 for patients' remaining seizures. Among 11188 clustered windows (3.1 hours), only 446 windows were assigned as cluster 2. All windows of cluster 2 were completely predicted by the classifier model leading to 100% as sensitivity and specificity. 6.5 presents the metrics of the confusion matrix used in calculating the performance of the model. Figure 6.3 illustrates the 95% confidence interval of sensitivity. The classifier model was applied to 30 seizures in seven patients. Table 6.6 illustrates the maximum prediction time achieved for each patient, alongside the time recorded before the seizure. This time relates to the seizure with the longest prediction time, for each patient.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	446	1
	Non-pattern	0	10741

TABLE 6.5: Confusion Matrix of Complex Tree Model

It is apparent that the longest prediction time was 41.8 minutes, observed for Patient 5. All seizures in the data were predicted by the developed model.

The next section gives the result of predicting Absence seizures using the PLI approach.

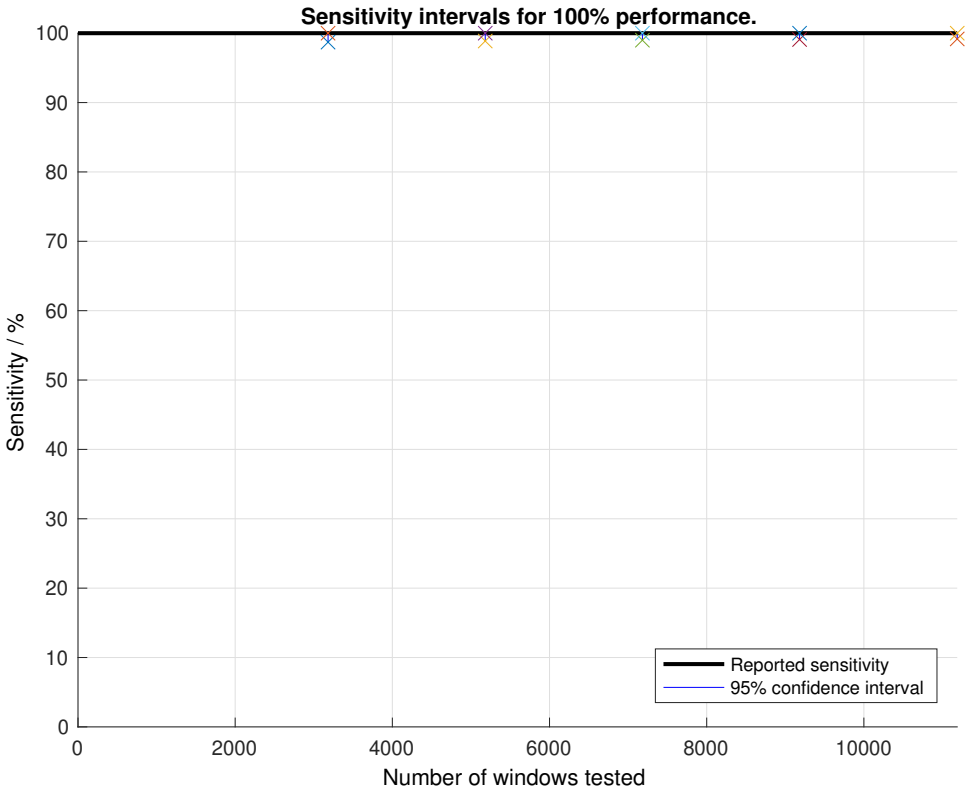


FIGURE 6.3: Estimated 95% confidence intervals of the sensitivity of the Complex Tree model within different number of windows.

Patient no.	Seizures used	Seizures predicted	Prediction time (minutes)	Time before Seizure onset (seconds)
1	5	5	15.47	929
2	6	6	16.32	980
3	2	2	18.43	1225
4	1	1	0.96	67
5	6	6	41.8	2519
6	7	7	13.3	800
7	3	3	6.99	429

TABLE 6.6: Maximum prediction time for all seven patients, with their recorded time before seizure .

6.2 Predicting Absence Seizures Using PLI

The temporal variation in PLI values was examined for the data of the two groups of patients used as a training set – data on nine seizures of six patients and two records for patients with no recorded seizures. It was observed that the value of PLI for patients in the two groups has not exceeded 0.9 like PLV. Therefore, 0.8 was selected to be

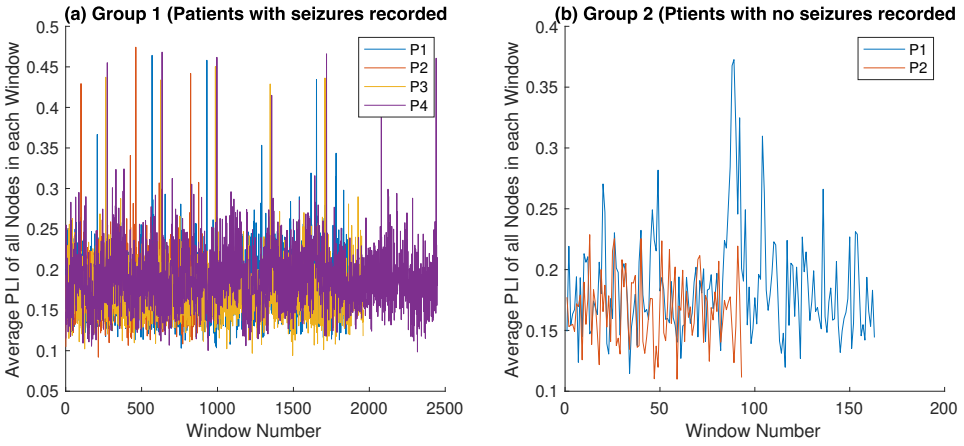


FIGURE 6.4: Application of the average PLI measure of all nodes for patients in the two groups, using data in the theta band.

the threshold value used to compare the dynamic change of PLI values. Exploring the dynamic change in the total number of synchronized nodes whose PLI value exceeded 0.8 revealed no pattern common to the nine seizures of the six patients. On the other hand, calculating the average value of PLI of all nodes within each window showed a trend in values of 0.4 in data in both the alpha and theta bands. Only before seizure onset were these values shared by all patients in Group 1 – patients with seizures – but not in Group 2 – patients with no seizures recorded. Figure 6.4 (a) and (b) show the average PLI of all nodes in the theta band for both groups. It is apparent that in Group 1 the values exceeded 0.4, while in Group 2 they did not attain that value. Based on previous observations, the network features in both the alpha and theta bands of the patients were extracted to explore any abnormal change.

6.2.1 Extracting theoretic graph parameters of the PLI-based networks

	Alpha	**Theta**
Radius	9	10
Global efficiency	0.5	0.5
CPL	6.5	8
Modularity	-	0.14
Transitivity	0.55	0.45
Clustering coefficient	0.55	-
Betweenness centrality	Nodes (F4-C4, C4-P4, FP2-F8, T6-O2, FP1-F3, F3-C3, C3-P3, FP1-F7) exceeded 104	Nodes (C4-P4, F8-T4, FP1-F3, F3-C3, C3-P3, P3-O1, FP1-F7, T3-T5, PZ-OZ) exceeded 108
Assortativity	0.04	0.02

TABLE 6.7: shows the trend in the features' values common to all patients.

All the features in Table 6.7, in alpha and theta, that had a common trend were used to build a prediction model. The features in the remaining bands were extracted for use only if the features in alpha and theta bands had failed in identify a common pattern. Figure 6.5 (a) and (b) shows how the trend value of CPL of data in the alpha band is exceeded by patients with recorded seizures.

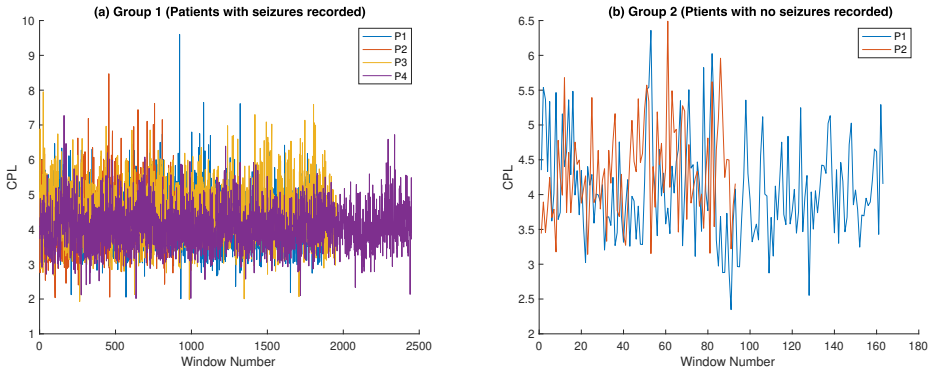


FIGURE 6.5: CPL value of data in the alpha band for both groups of patients.

6.2.2 Formulating a prediction model using PLI

Next, the extracted features were clustered, using k-means, to identify any hidden patterns. Only features with shared trends were used in this process. The features' dimensions were reduced by removing one of the two strongest correlated features. After removing the radius and clustering coefficient, 12 features remained for the data in the alpha band – CPL, transitivity, assortativity, global efficiency of betweenness centrality for eight nodes - while there were 14 features for theta – CPL, global efficiency, transitivity, assortativity, modularity and betweenness centrality for nine nodes. The features in the alpha band formed 10 clusters- the optimal number determined by the 30 indices mentioned in Chapter 4- and two in the theta band. Unfortunately, no pattern was found that was common to both bands. As the next step, the features of both the alpha and theta bands were combined, and formed two clusters. However, this combination also failed to reveal any common pattern. As a consequence, all the extracted features of all bands were used in clustering. Table 6.9 shows the optimal number of clusters in each band using the three clustering methods.

	Alpha	**Beta**	**Gamma**	**Delta**	**Theta**
Elbow	-	-	-	-	-
Silhouette	2	2	2	2	2
Gap statistic	1	1	1	1	1

TABLE 6.8: Optimal number of clusters in the five bands, as established by the three methods.

It is apparent that the elbow method failed to determine an optimal number of clusters, while the others arrived at different numbers, for all bands. Therefore, 30 indices were applied, as described in Chapter 4, and it is found that two was optimal for each band. Consequently, the extracted features of each band were formed into two clusters. The logistic regression model classified the data, using the cluster number as a label. Clustering revealed no common pattern in the bands, apart from in the delta band. There was a pattern shared by all nine data records of the six patients used as training set, compared to the two patients without any seizures recorded. This appeared close to the seizure's onset, as seen in Figure 6.6 between the two dashed red lines, starting and ending with Cluster 2. It appeared at 180 seconds before seizure for eight patients' data, apart from one, which persisted longer.

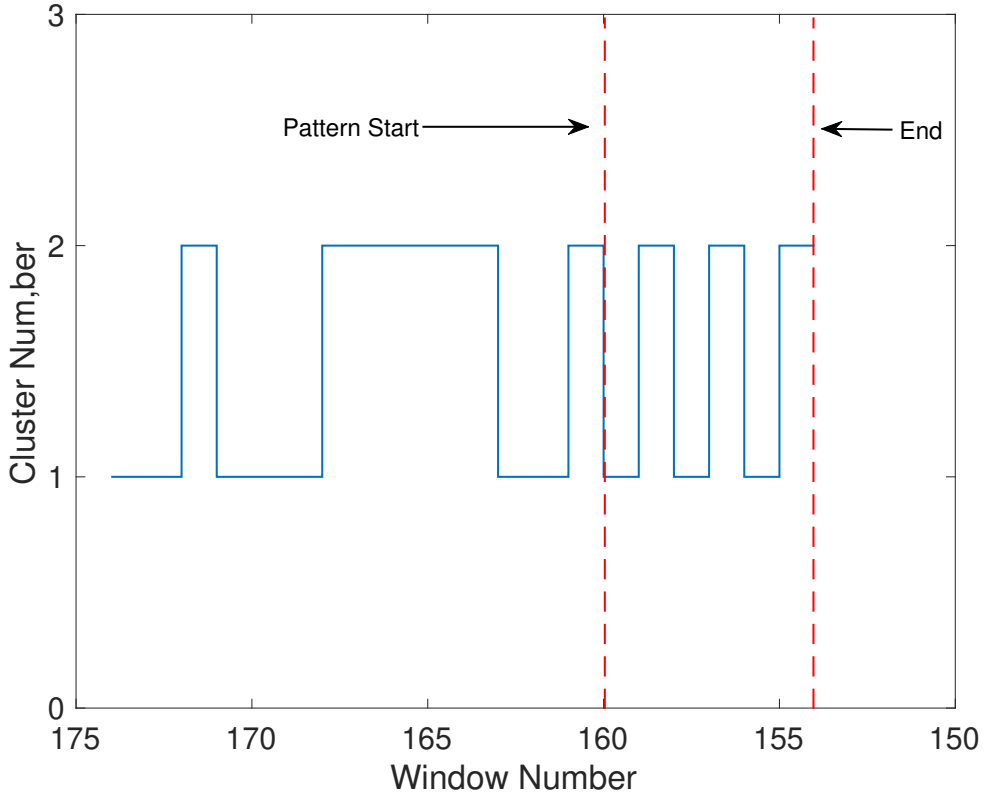


FIGURE 6.6: Common pattern in the transition of the two clusters.

It is evident from the figure that the pattern takes place between Windows 1154 and 1160 (the window number represents distance from seizure onset). It was checked in the remaining seizures data of seven patients (21 seizures) by applying the Logistic Regression classifier model. The model achieved 98% specificity and 99% sensitivity. Figure 6.8 shows the procedures adopted to predict the seizures and their time of onset. It starts with a series of sliding windows seven windows long and compares their values to the common pattern [2 1 2 1 2 1 2], referring to the cluster numbers. Applying this

procedure to all patients resulted in just one prediction failure in 30 seizures in all seven patients. The classifier model was built by asigning value 1 to the windows showed the common pattern in the training set and 0 for the other windows. The total amount of windows used in building the model were 11182 - 1394 windows for the common pattern. Table 6.9 illustrates the metrics of the confusion matrix used to calcluate sensitivity and specificity of the model. Figure 6.7 illustrates the 95% confidence interval of sensitivity.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	1380	188
	Non-pattern	14	9600

TABLE 6.9: Confusion Matrix of Logistic Regression Model

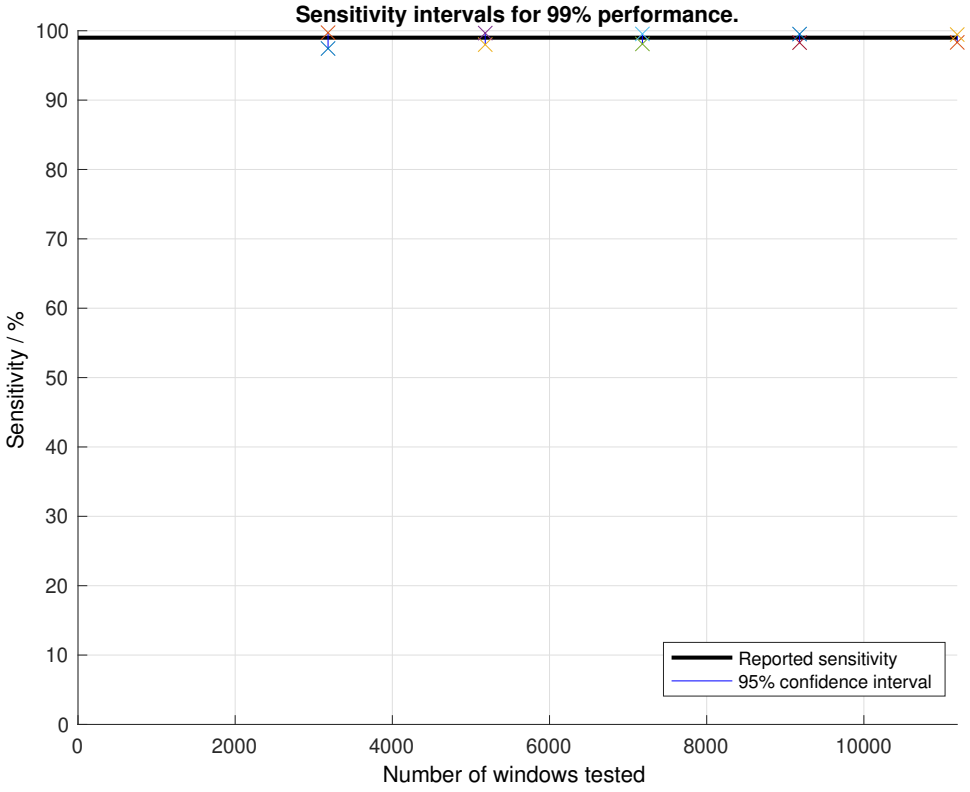


FIGURE 6.7: Estimated 95% confidence intervals of the sensitivity of the Logistic Regression model within different number of windows.

The values of the features in the common pattern for the two clusters are described in Table 6.10 in terms of the maximum and minimum values of five features. Cluster 2, which is shown at the start and end of the common pattern, exhibited a higher CPL value than Cluster 1. After incorporating the CPL values within the pattern, a high peak was observed at the start in all patients (Figure 6.9). This confirms the results in Table 6.10.

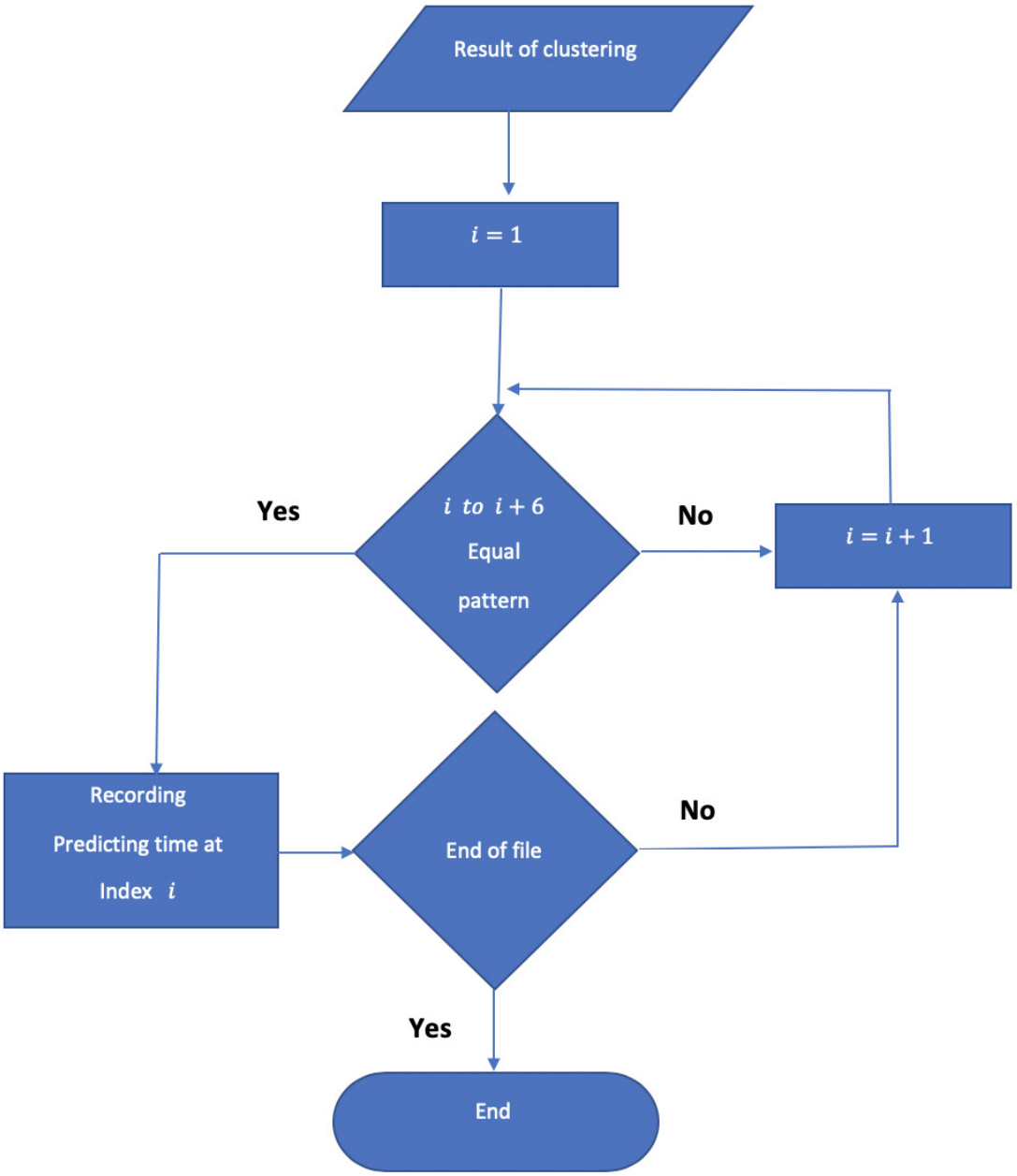


FIGURE 6.8: Flowchart of the steps to predict the pattern in all patients.

	Cluster 1			**Cluster 2**		
	Max	Min	Avg	Max	Min	Avg
Global Efficiency	0.668	0.260	0.358	0.529	0.163	0.269
CPL	4.53	1.65	3.39	7.31	2.16	4.51
Modularity	0.172	-9.8	0.063	0.180	0.024	0.091
Transitivity	0.640	0.234	0.370	0.422	0.159	0.289
Assortativity	0.039	-0.181	-0.062	0.059	-0.142	-0.060

TABLE 6.10: Maximum and minimum values of network features of the two clusters.

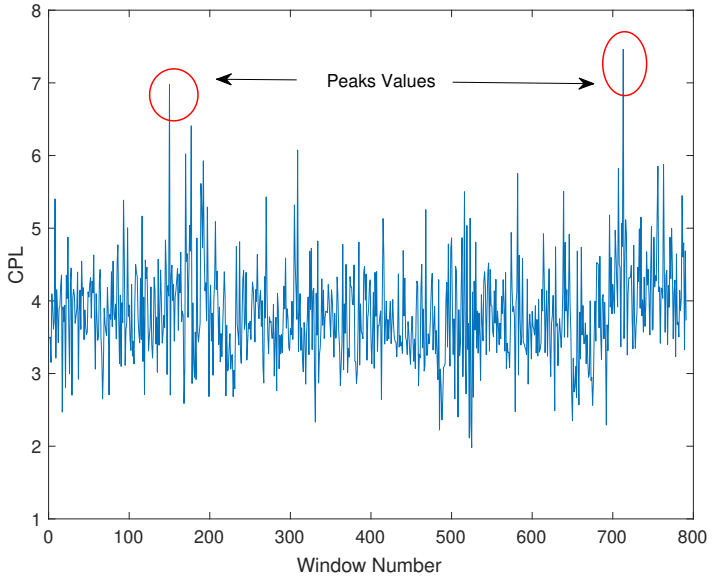


FIGURE 6.9: Values of peaks of CPL in pattern, at two points in time.

Figure 6.9 shows with a red circle the peak in CPL value at the beginning of the common pattern, at two points of time.

The prediction time was calculated for all predicted seizures. Table 6.11 shows the prediction time for seven patients, with the number of seizures predicted and the time recorded before the seizure onset.

Patient no.	**Seizures used**	**Seizures predicted**	**Prediction time (minutes)**	**Time before Seizure onset (seconds)**
1	5	5	6.53	480
2	5	5	11.78	862
3	1	0	-	67
4	2	2	3.01	572
5	6	6	4.66	294
6	7	7	17.43	800
7	2	2	2.09	192

TABLE 6.11: Predicted results of all seven patients.

The maximum prediction time was 17.43 minutes, in Patient 6; that is, 800 seconds before seizure onset. There was only one prediction failure, for Patient 3, of the total 28 seizures examined.

The next section provides the results of using SL-based connectivity networks to build a prediction model.

6.3 Predicting Absence Seizures Using SL

The two statistical measures used with PLI were again applied here to check the temporal variation in SL value for the two groups of patients in the training set. They used to find a common trend in either of the two measures' values among patients with seizures that is not present in subjects without any recorded seizures. A common trend was seen in the total number of synchronized nodes of SL value above 0.8 in data in the beta band, which exceeded two, while the temporal change in SL average in all nodes within each window showed a common trend of 0.48 in data in the delta band. Figure 6.10 illustrates these two trends: (a) the trend value of total number of synchronised nodes above 0.8 for four patients in Group 1 - it is clear that values for the beta band are over two; and (b) the trend value (0.48) is exceeded by these same patients for data in the delta band. These trends are used later with network features to help in building the prediction model.

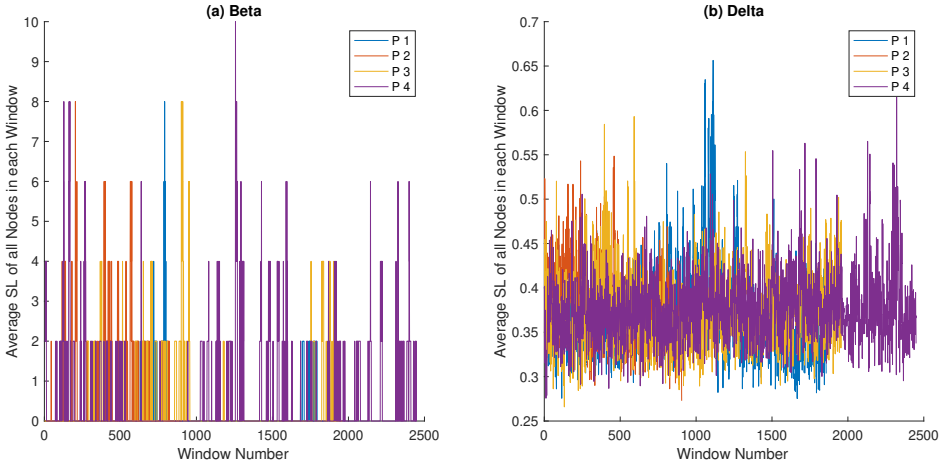


FIGURE 6.10: Two trends in the two measures for data patients, in beta and delta respectively.

6.3.1 Extracting the graph theoretic parameters of the SL-based networks

All the features of the network used with PLI were extracted here, for all five bands. Table 6.12 shows the trend values of the network features in both beta and delta. The table gives the trend values exceeded by the patients' data in Group 1 of the training set, in both the beta and delta bands. Besides the statistical values of the two measures provided in the previous subsection, only features with the trends in the beta and delta bands were used to build the prediction model.

	Beta	Delta
Radius	> 18	-
Global efficiency	-	> 0.55
CPL	> 13.5	-
Modularity	> 0.3	-
Transitivity	-	> 0.75
Betweenness centrality	5 nodes exceeded 40	10 nodes exceeded 17
Assortativity	-	-
Strength of node	-	-

TABLE 6.12: Trend values of network features in the beta and delta bands.

6.3.2 Formulating a prediction model using SL

To build a prediction model based on the previous results, the features extracted from the beta and delta bands, which showed the trends, were clustered using k-means. Six and eight clusters were formed for the data in the beta and delta bands, respectively. Unfortunately, no common pattern appeared for the data in either. Consequently, all the features extracted from the five bands were used in clustering. The optimal number of clusters was two, for each band. There was a common pattern only in data in the beta and gamma bands. In beta, all features were used in clustering apart from radius, because of its strong correlation to the other features. Cluster 2 showed a pattern of a duration of more than 10 seconds that was not found in the two patients in Group 2. During this pattern, the CPL value in all patients in Group 1 at some point exceeded 10, while in patients of Group 2 it reached less than 10. Figure 6.11 and Figure 6.12 show examples of patients with and without seizures, respectively.

Because a 50% overlap was used with the sliding windows, in the above figures each 20 windows represents 10 seconds. Figure 6.11 and Figure 6.12 show that in the common pattern – starting at Window 5 – the CPL value can either exceed or remain under 10. As a result of this split, a binary recording was applied using the cluster number with the CPL value to build a model with the following rule: 1- When the cluster number is two and the CPL value is higher than 10; and 0- When at least one of the two values is not satisfied. These binary values were used in the classification. Two models were used: Linear SVM and Logistic Regression. They achieved (99%, 100%) sensitivity and (100%, 75%) specificity, respectively. Based on these results, Linear SVM was chosen to be applied to the remaining patients. The total amount of data used in building the two models was 11188 windows (3.1 hours). The total number of windows used for the common pattern was 3502. Table 6.13 and Table 6.14 illustrate the metrics of the confusion matrices used to calculate the performance of Linear SVM and Logisitic Regression models respectively.

The result of applying Linear SVM model showed eight prediction failures of the 30 seizures in the seven patients, as shown in Table 6.15. It shows the prediction time

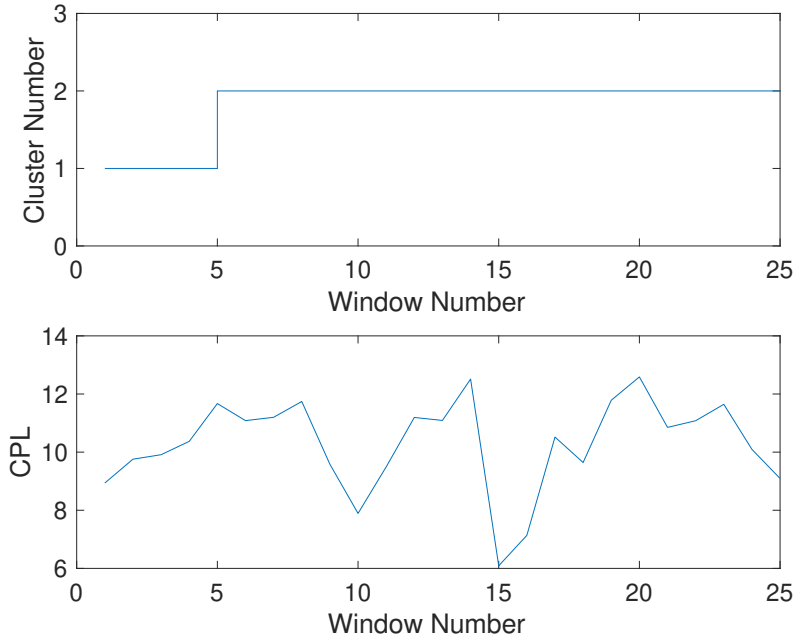


FIGURE 6.11: Common pattern in beta band, with corresponding CPL value, for patients with recorded seizures.

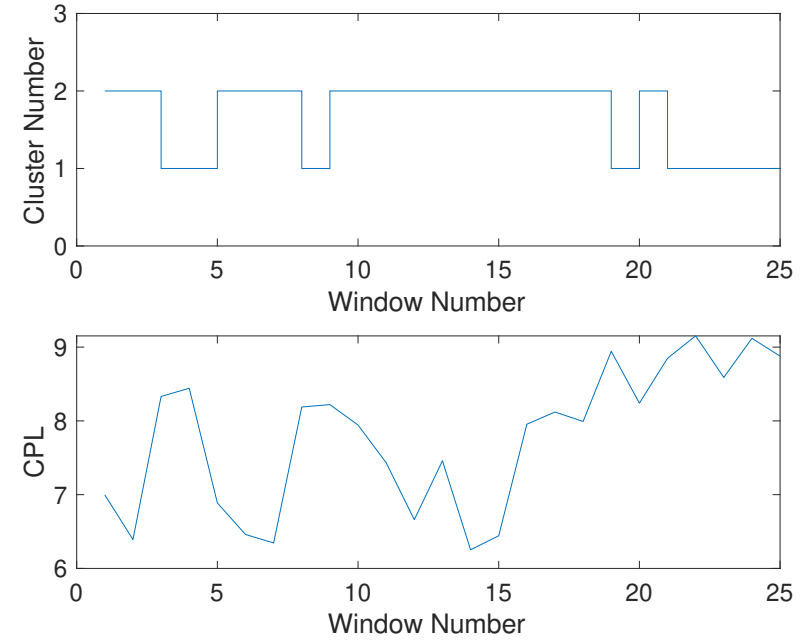


FIGURE 6.12: Common pattern in beta band, with corresponding CPL value, for patients without recorded seizures.

with the recorded time before seizure onset and the number of predicted seizures. The maximum prediction time was 15.68 minutes, in Patient 7.

The data in the gamma band were clustered into two clusters, after removing the radius feature due to its strong correlation. CPL values of higher than 13 were observed in all

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	3466	4
	Non-pattern	36	7682

TABLE 6.13: Confusion Matrix of Linear SVM Model for data in beta band

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	3497	1920
	Non-pattern	5	5766

TABLE 6.14: Confusion Matrix of Logistic Regression Model for data in beta band

Patient no.	**Seizures used**	**Seizures predicted**	**Prediction time (minutes)**	**Time before Seizure onset (seconds)**
1	5	4	6.91	929
2	6	5	12.98	862
3	2	0	-	-
4	1	1	1.05	67
5	6	4	6.03	366
6	7	6	10.52	708
7	3	2	15.68	1082

TABLE 6.15: Information of predicting results of all patients for data in beta band.

patients during Cluster 2, and this was not demonstrated by the two patients without any recorded seizures.

It is apparent from Figure 6.13 that the CPL value can exceed 13 during Cluster 2 in patients with recorded seizures, while it remains under 13 in those without recorded seizures as seen in Figure 6.14. Based on this observation, a binary recording similar to that used before was applied, using values of CPL greater than 13 to build a prediction model. Linear SVM was used to classify, achieving 99% sensitivity and 99% specificity. Table 6.16 illustrates the confusion matrix used to calculate the performce of the model using 2847 out of 11188 windows for the common pattern. Figure 6.15 illustrates the 95% confidence interval of sensitivity.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	2819	87
	Non-pattern	28	8254

TABLE 6.16: Confusion Matrix of Linear SVM Model for data in gamma band

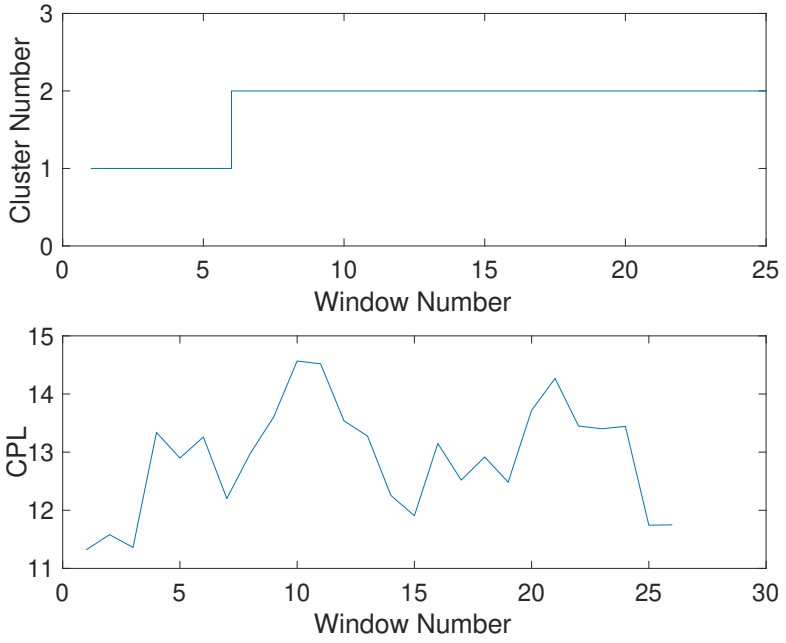


FIGURE 6.13: Common pattern in the gamma band, with corresponding CPL values, for patients with recorded seizures.

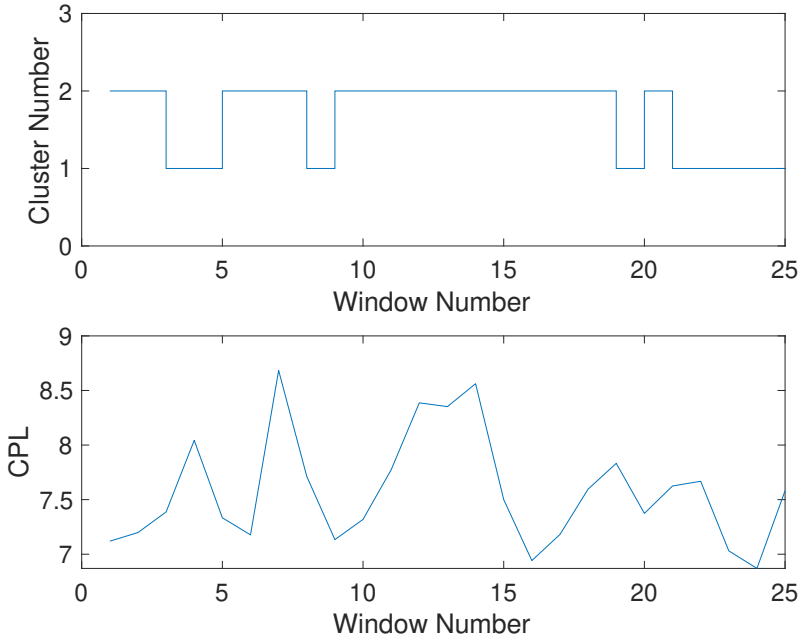


FIGURE 6.14: Common pattern in the gamma band, with corresponding CPL values, for patients without recorded seizures.

When applied to the remaining patients, there were four prediction failures out of 30 in the data for seven patients. Table 6.17 gives the result of prediction seizures for all seven, showing the prediction time and the number of predicted seizures. The maximum prediction time was 41.9 minutes, shown by Patient 5.

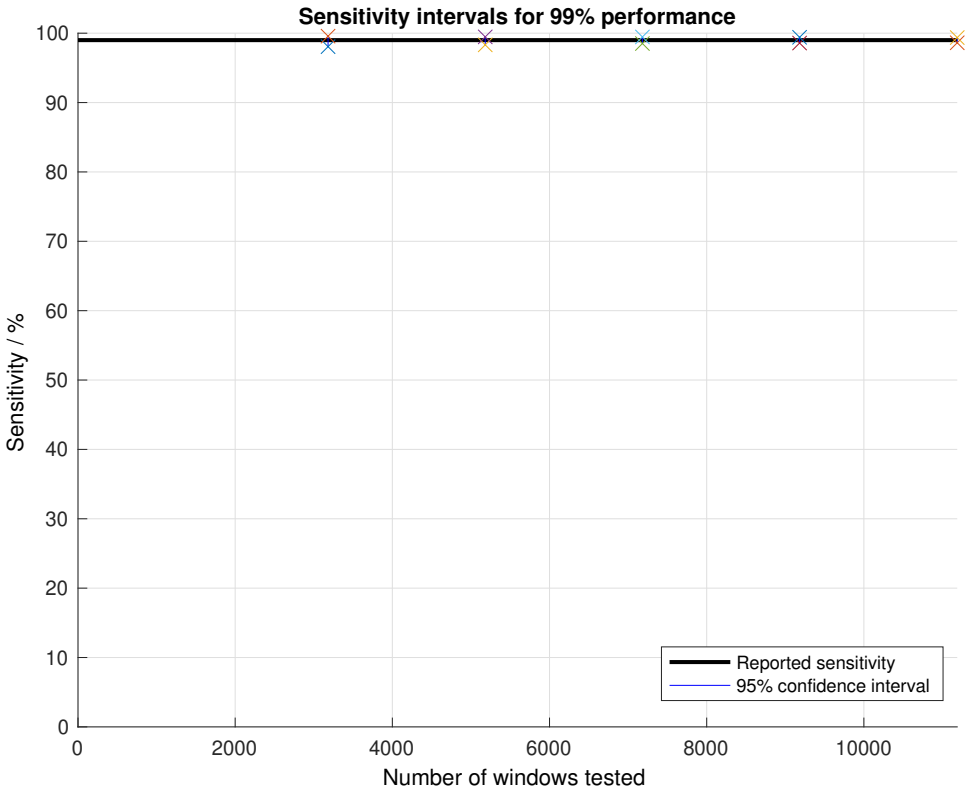


FIGURE 6.15: Estimated 95% confidence intervals of the sensitivity of Linear SVM model within different number of windows.

Patient no.	Seizures used	Seizures predicted	Prediction time (minutes)	Time before Seizure onset (seconds)
1	5	5	14.93	929
2	6	6	15.73	980
3	2	1	15.71	1225
4	1	1	0.96	67
5	6	6	41.9	2519
6	7	4	5.04	708
7	3	3	13.77	1082

TABLE 6.17: Information of predicting results of all patients for data in gamma band.

Compared to the models created for the data in the beta band, this prediction model for the gamma band had fewer prediction failures, representing an improvement.

6.4 Comparison of the models' performance

This chapter present the results of predicting Absence seizures using three synchronization approaches for constructing functional brain connectivity networks – PLV, PLI and SL. For PLV-based connectivity networks, the prediction model was based on a pattern in the features of delta band found to be common to all patients, where CPL has a low value during seizure, as shown in Table 6.3. Complex Tree classifier was used to build a prediction model that achieved 100% as sensitivity and specificity. All seizures used were predicted, with a maximum prediction time of 41.8 minutes (see Table 6.6), and it successfully showed no seizures for the two subjects without recorded seizures.

For PLI-based connectivity networks, a common pattern was found for data in the delta band, characterized by a high CPL value at the start of the pattern, as shown in Figure 6.9. Logistic Regression classifier was used to build the prediction model, which achieved 99% and 98% for sensitivity and specificity, respectively, giving a maximum prediction time of 17.43 minutes (for patient with 800 seconds before seizure). This is less than that achieved by the model using the PLV-based connectivity network (for patient with 2519 seconds before seizure), as shown in Table 6.11 and Table 6.6, respectively. The prediction model succeeded in predicting all the seizures apart from one.

Finally, a prediction model using SL-based connectivity networks was built for data in the gamma band. The linear SVM classifier was used to formulate a model that achieved 99% for both sensitivity and specificity. It used a common pattern characterized by the high values of CPL, exceeding 13. The model predicted 26 seizures out of 30, with a maximum prediction time of 41.9 minutes for the patient with the longest recorded time before seizure onset (2519 seconds) (Table 6.17).

6.5 Conclusion

The atypical variation in CPL was seen in prediction models of both types of epilepsy that used the three synchronization measures of PLV, PLI and SL. Consequently, this observation helped later in building a general prediction model for both types of epilepsy examined in this studying.

All prediction models using the three synchronization measures – PLV, PLI and SL - for Absence seizures achieved a high sensitivity (100%, 99% and 99%, respectively), higher than that of models in previous studies (Cho et al. 2016, Myers et al. 2016, Li et al. 2013) (82.44%, 76.8% and 75.8%, respectively). It was observed that, in the TLE models, the maximum prediction time depends on the length of the data recorded before seizure onset, as seen in Table 4.7 and Table 5.10, whereas in the models of Absence

Approach	Sensitivity (%)	Specificity (%)	Network Features used in formulating model	Classifier models used	Prediction horizon (minutes)
PLV	100	100	CPL, Global Efficiency, Transitivity, Assortativity and Betweenness Centrality of Nodes CZ-PZ, FP2-F8, P4-O2, T4-T6, and T5-O1 in delta band	Complex Tree	41.8
PLI	99	98	Global Efficiency, CPL, Transitivity, Assortativity, in delta band	Logistic Regression	17.43
SL	99	99	Transitivity, Global Efficiency, Betweenness Centrality Modularity and Assortativity in gamma band	Linear SVM	41.9

TABLE 6.18: Summary of all prediction models created for absence seizures.

seizures there is no relationship (see Table 6.11 and Table 6.17). Table 6.18 provides a comparison of performace for the developed models using the three approaches -PLV, PLI and SL for patients with absence seizures. However the significant performance provided by these models, they are limited to the sample of patients used -7 patients and the length of recording time from the start of data to the seizure onset - there is only one patient with maximum recording time of 20 minutes.

After creating prediction models for both types of epilepsy - both TLE and Absence – a general prediction model was formulated on the basis of their results.

Chapter 7

Formulating a General Model for Epileptic Seizure Prediction

Earlier chapters built models to predict both TLE and Absence types of epilepsy, based on analysing the theoretic graph parameters of the functional brain connectivity networks constructed. These networks were created using three measures of functional connectivity - PLV, PLI and SL - to assess the uncontrolled synchronization between neural populations. Each model achieves a higher sensitivity and specificity than models in previous works. In this chapter, a general prediction model is formulated to predict both types of epilepsy that are examined in this study. To do this, the features were selected that contributed most significantly to these models' predictions of seizures. Table 5.11 and Table 6.18 summarize all the models created for the prediction of both types of epilepsy through the three measures - PLV, PLI and SL. These tables show their sensitivity, specificity and classifier model used, in addition to the network features that contributes to their construction.

Three criteria in selecting the network features were adopted to build the general prediction model:

- Only features common in both types of epilepsy.
- Only features showing an abnormal change.
- Only features in bands shared by both types of epilepsy. For example, although in different bands - gamma and delta - CPL was common to both PLV-based prediction models in the two types, as seen in Table 5.11 and Table 6.18. Thus, CPL was selected to build the prediction model.

Based on these criteria, the feasibility of building a prediction model was explored for each of the three measures to establish a common pattern among patients with both types of epilepsy. In preparation, various scenarios were examined:

- All selected bands, with all common features.
- All selected bands, with one feature.
- One band, with all common features.
- One band, with one common feature.

For each scenario, the features of training set – six seizures files of 4 TLE and nine seizures file of Absence – were used as a training set. The total windows used in the training set was 14782 - 3594 windows for TLE and 11188 windows for absence seizure- which are equal 4.1 hours in total. The features were clustered using the k-means algorithm after ascertaining the optimal number of clusters. Classifier models were formulated using the clustered features and applied to the remaining patients and six healthy subjects. If a pattern was found in common to all types of patients, or even within each type of epilepsy, the prediction time was calculated, along with the sensitivity and specificity of the prediction model. The two possible scenarios could each have either two patterns – seizure or no seizure - or three patterns - TLE, Absence or no seizure. Figure 7.1 shows one of these two possible scenarios, containing three patterns.

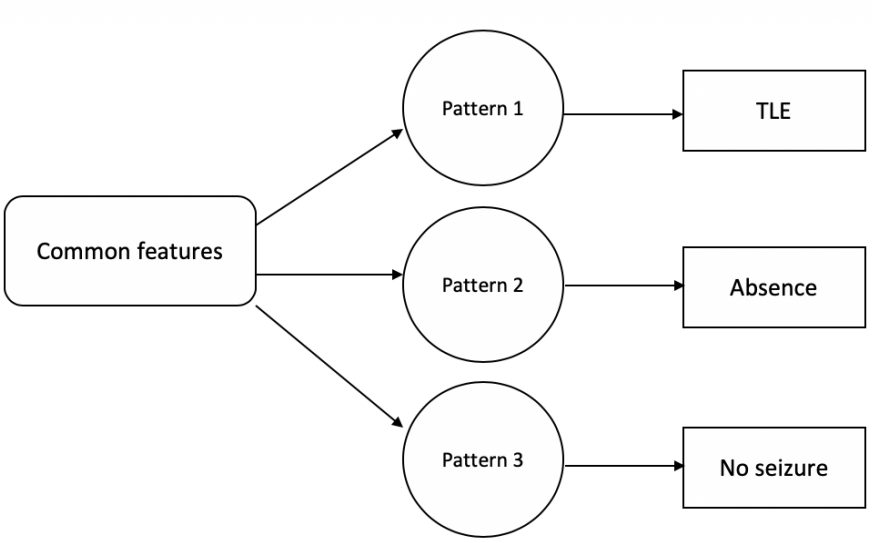


FIGURE 7.1: First possible senario in formulating the general prediction model.

After applying the criteria, the CPL values in the gamma and delta bands for the PLV measure for TLE and Absence patients, respectively, were selected to build the prediction model. The features were formed into three clusters, based on the results of applying the 30 indices. Applying all the possible scenarios to formulate a prediction model using PLV features of both epilepsy types revealed no common pattern. However, using just the features of PLI and SL led to common patterns in both. The following sections describe the process of building prediction models using PLI and SL, along with their prediction times.

7.1 Formulating General Predictive Model Using PLI

The beta, gamma and delta bands were selected to create the prediction model, using the common features of PLI, CPL and assortativity. Only one scenario – using CPL in the gamma band, for both epilepsy types – showed a common pattern for use in the model. The features were formed into clusters, after calculating the optimal number of clusters using four methods. The elbow method failed to find the number of clusters, while the gap statistic and silhouette methods arrived at one and two clusters, respectively. Therefore 30 indices were used to evaluate the number of clusters, which was two. The linear SVM classifier classified the two clusters of data, using the cluster number as the label, and achieved 100% sensitivity and specificity for both. The classifier was applied to data on six seizures files for four TLE patients, nine seizures files for six Absence patients and for six healthy subjects. As a result, three patterns emerged: TLE pattern; Absence pattern; and no seizure pattern. The first pattern comprised Cluster 1 lasting for two seconds or more. It was demonstrated by only the TLE subjects in the training set, as shown in Figure 7.2.

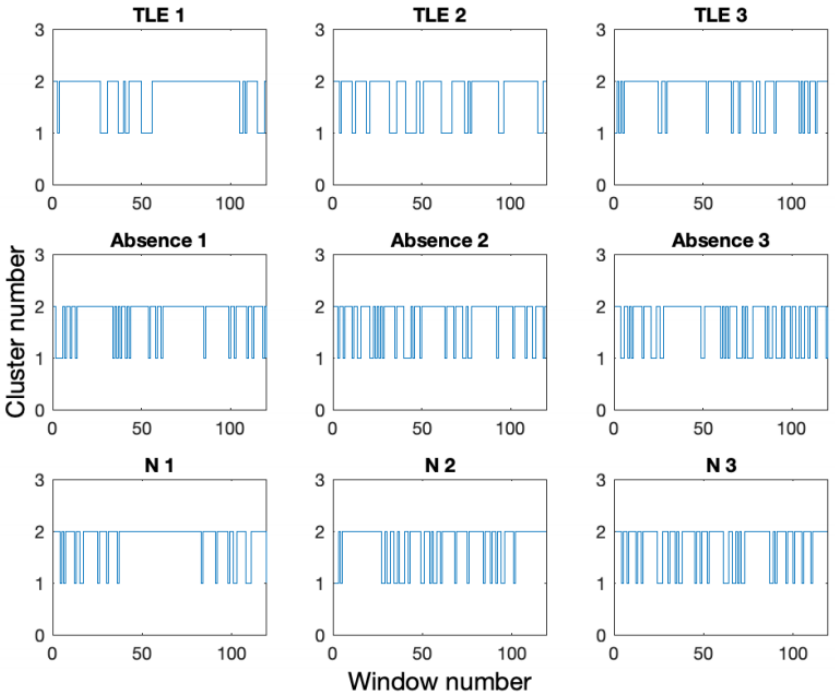


FIGURE 7.2: Transition of the two clusters for three patients of both epilepsy types, and three healthy subjects.

Figure 7.2 shows the transition of the two clusters for TLE, Absence and healthy subjects in Rows 1, 2 and 3, respectively. It is apparent that Pattern 1 – Cluster 1 lasting for two seconds or more – is present only in TLE patients. Every two windows represent one second, as 50% overlap was used when the synchronization values of functional brain

connectivity networks were calculated. The values of CPL of TLE patients during the pattern have been drawn to characterize Pattern 1.

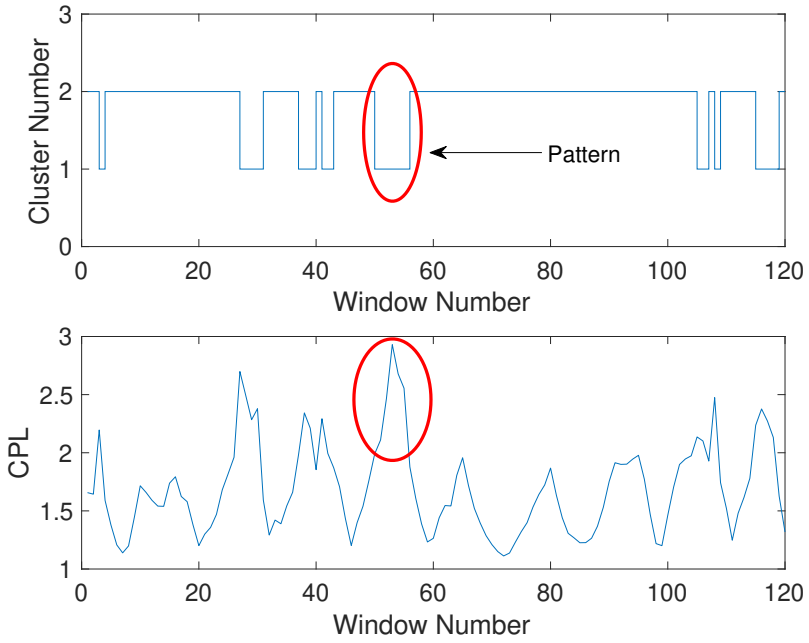


FIGURE 7.3: CPL value of TLE subjects throughout two clusters.

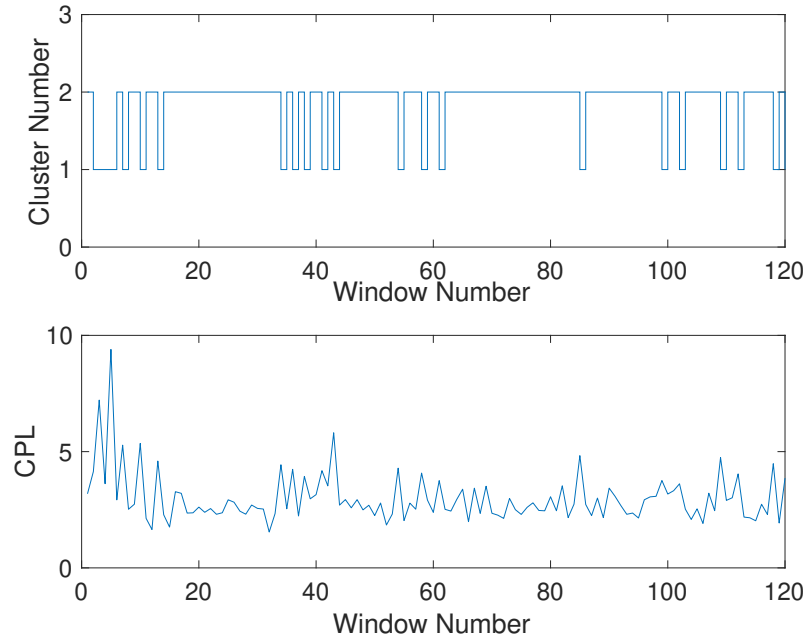


FIGURE 7.4: CPL value of Absence subjects throughout two clusters.

The Pattern 1 values for CPL during Cluster 1 were compared to the actual values in Absence patients and healthy subjects. This showed that CPL values within Cluster 1 for TLE patients lie between 2 and 3. Figure 7.3, 7.4, and 7.5 show the CPL values

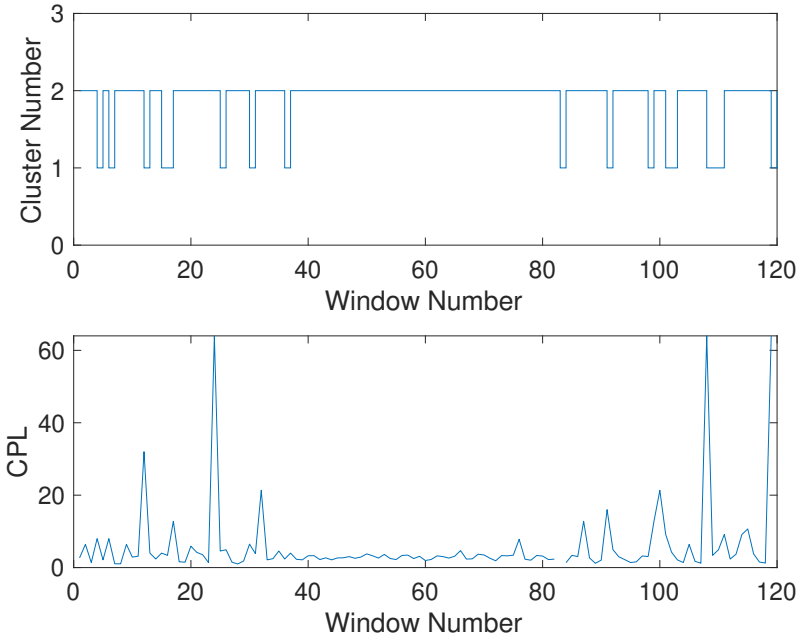


FIGURE 7.5: CPL value of healthy subjects throughout two clusters.

for the two clusters for TLE, Absence and healthy subjects, respectively. In Figure 7.3, Pattern 1 and its corresponding CPL values are indicated by red ovals. It is evident that during Cluster 1 CPL ranges from 2 and 3 for TLE patients, yet exceeds 3 in both Absence patients and healthy subjects, as shown in Figure 7.4 and Figure 7.5.

This observation was used later to build the model to predict Pattern 1. The range of CPL value was checked in the data for patients of both types of epilepsy and also the healthy subjects. It was found only in TLE subjects, confirming that this range of values indeed characterizes Pattern 1. For further confirmation, the CPL values of two Absence patients without recorded seizures were checked, and Pattern 1, as indicated by CPL values of between 2 and 3 in Cluster 1, was found absent. Subsequently, another pattern for Absence patients was found, represented by the transition between the two clusters, as shown in Figure 7.6.

As with Pattern 1, a comparison of the CPL value in Pattern 2 was undertaken among all subjects – TLE, Absence and healthy. It was observed that during Pattern 2, only in Absence patients does the value of CPL exceed 10 in Cluster 1. This was used to build the general model to predict Pattern 2. Based on the previous results, the third pattern - no seizures - was considered for the remaining data, unrelated to Patterns 1 and 2. The general prediction model was formulated using three classes - 1, 2 and 3 - for TLE, Absence and no seizures patterns, respectively.

Using the Bagged Tree classifier, the proposed model achieved a sensitivity and specificity of 100% for all three patterns. The total amount of data used in building the model were 14782 windows (4.1 hours). Table 7.1 illustrates the confusion matrix of the three

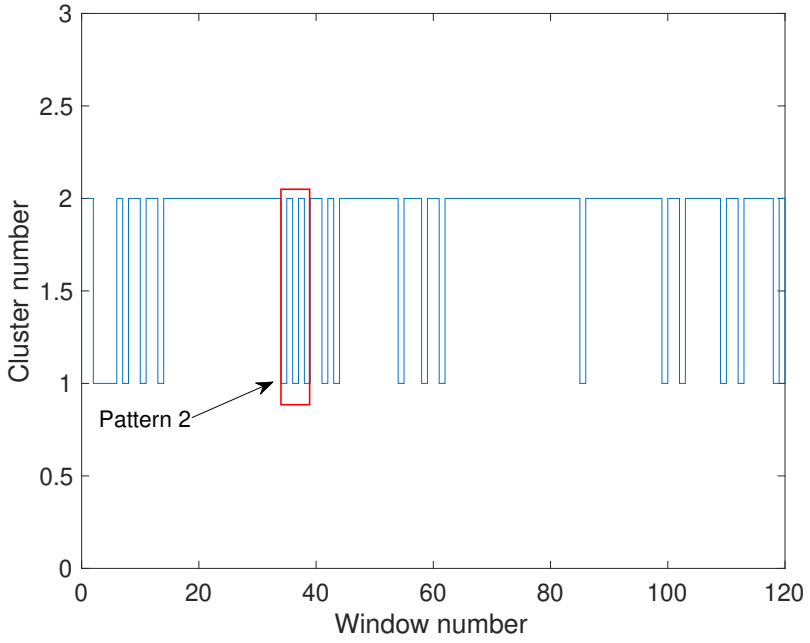


FIGURE 7.6: Pattern 2, related to Absence patients.

patterns. The performance of the model was calculated by combining the performance of three patterns as following:

- Sensitivity of the model=(774+13+13993)(775+13+13994)=100%
- Specificity of the model=(14006+14769+787)/(14007+14769+788)=100%
- False Prediction Rate= Fale Positive Rate /amount of data used =(2/29564)/4.1 =0.00001/h.

		Actual		
Predicted		Pattern1	pattern2	Pattern3
	Pattern1	774	0	1
	Pattern2	0	13	0
	Pattern3	1	0	13993

TABLE 7.1: Confusion Matrix of Bagged Tree Model for data in gamma band

It was applied to 24 seizures files of TLE, 30 seizures files of Absence epilepsy, 10 files of healthy subjects and the data on two Absence patients without recorded seizures. The model succeeded in predicting Pattern 1 for 24 seizures of TLE subjects without missing any, while it predicted Pattern 2 for 21 seizures of Absence patients and missed nine, and successfully predicted no seizures for the two files of patients with Absence epilepsy yet with no recorded seizures. For Pattern 3, it resulted in prediction failure for three healthy subjects of the 10 used. Table 7.2 gives the maximum prediction time in both

types of epilepsy, and the third column shows the time recorded before seizure onset only for the subject with the longest prediction time.

Patient no.	**Maximum predicting time(seconds)**	**Time recording before seizure onset (seconds)**	**Seizures used**	**Seizures predicted**
TLE-patient 1	229	300	2	2
TLE-patient 2	291	300	1	1
TLE-patient 3	299	300	3	3
TLE-patient 4	67	70	2	2
TLE-patient 5	68	70	1	1
TLE-patient 6	63	70	2	2
TLE-patient 7	45	50	2	2
TLE-patient 8	128	130	3	3
TLE-patient 9	97	100	2	2
TLE-patient 10	48	50	5	5
Absence-patient 1	525	929	5	4
Absence-patient 2	791	980	6	4
Absence-patient 3	1120	1225	2	2
Absence-patient 4	-	67	1	0
Absence-patient 5	624	730	6	4
Absence-patient 6	724	800	7	4
Absence-patient 7	1062	1082	3	3

TABLE 7.2: Prediction time for all patient types, using the PLI model.

From the table, it is clear that of the TLE patients Patients 1 and 3 show the maximum prediction time with the PLI model (299 seconds = 5 minutes), while Patient 3, with Absence seizures, shows the longest time overall (1120 seconds = 19 minutes). The model has a false prediction rate of 0.00001/h.

7.2 Formulating a General Predictive Model Using SL

In the previous section, a general prediction model was created using PLI features. In this section, similar scenarios were again applied, this time to create a general prediction model using SL features. In SL, the features common to both types of epilepsy were transitivity in the gamma and delta bands, as in Table 5.11 and Table 6.18. After checking all the possible scenarios, it was found that only transitivity in the delta band could contribute to a prediction model. Thus, it was extracted from six seizure files for four TLE patients and nine seizure files of six Absence patients. The elbow method failed to determine the optimal number of clusters to be used, while the Silhouette and Gap Statistic methods arrived at two and one, respectively, thus showed no consensus. The 30 indices which contain the three earlier methods were employed and, as a result, the data were grouped by k-means into two clusters.

Two classifier models - Linear SVM and Logistic Regression – classified the data, using the cluster numbers as labels. Both achieved 100% as sensitivity and specificity. The result of applying these two classifier models was plotted to identify any common pattern. One of the most significant contributions in formulating the prediction model is the value of transitivity that is observed during Cluster 1, which exceeded 0.73 in both TLE and Absence patients data, compared to that of the four healthy subjects. For further verification, the transitivity value of two files of Absence patients – who had had no recorded seizures - were also checked (Figure 7.7).

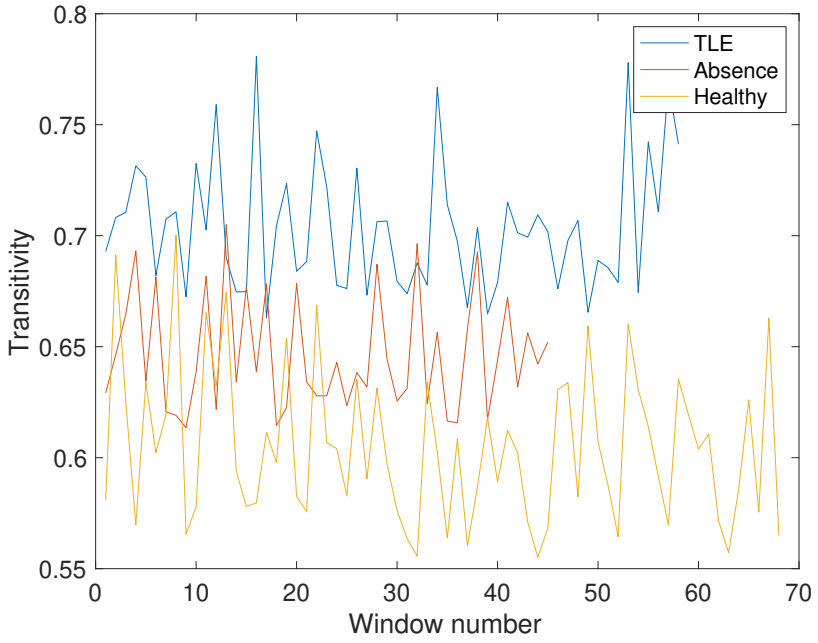


FIGURE 7.7: Transitivity value during Cluster 1 for TLE patients, Absence patients and healthy subjects.

The pattern of transitivity value was not apparent in the data for these two patients, therefore, this observation was used to create a training dataset for the classification of two classes – seizure and no seizure. All samples of data with transitivity values of more than 0.73 were set to one, while the others were set to zero. A general prediction model was created using Linear SVM, achieving 100% sensitivity and 94.4% specificity. The total amount of data used to build the model was 14782 windows (4.1 hours) represent the training set (6 EEG data of TLE and 9 EEG data of Absence seizures subjects). There were 14442 windows out of 14782 were assigned to the common pattern. The total rate of incorrect predicted windows as pattern equal 0.06 which was calculated using the equation 3.9. The False Prediction Rate was 0.01/h which is calculated through dividing 0.06 by the total amount of hours used in the in building the model (4.1 hours). Table 7.3 illustrates the confusion matrix created to calculate the performance of the model. Figure 7.8 illustrates the 95% confidence interval of sensitivity.

		Actual	
Predicted		Pattern	Non-pattern
	Pattern	14442	19
	Non-pattern	0	321

TABLE 7.3: Confusion Matrix of Linear SVM Model for data in gamma band

The model created was applied to the remaining patients. Table 7.4 shows the maximum prediction time in both types of patients, along with the recorded time before seizure onset. The model predicted TLE seizures with a maximum time of 298 seconds (5 minutes) for Patient 1, with 300 seconds of recorded time before seizure onset. on the other side, the model predicted Absence seizures with a maximum time of 2,392 seconds (40 minutes) for Patient 15, with 2,519 seconds of recorded time before seizure onset.

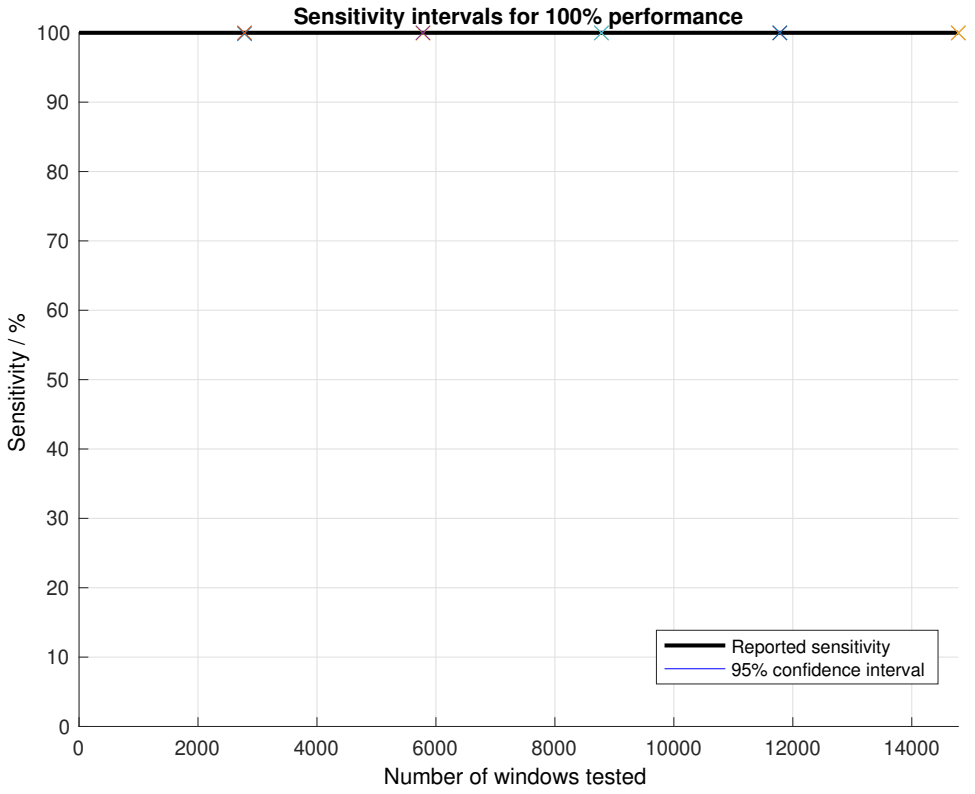


FIGURE 7.8: Estimated 95% confidence intervals of the sensitivity of Linear SVM model within different number of windows.

7.3 Performance Comparison and Conclusion

In this chapter, two general prediction models were formulated from the findings of Chapters 4, 5 and 6. These were based on the features common to both types of epilepsy

Patient name	Maximum predicting time(seconds)	Time recording before seizure onset (seconds)	Seizures used	Seizures predicted
TLE-patient 1	298	300	2	2
TLE-patient 2	253	300	1	1
TLE-patient 3	294	300	3	3
TLE-patient 4	65	70	2	2
TLE-patient 5	60	70	1	1
TLE-patient 6	62	70	2	2
TLE-patient 7	44	50	2	2
TLE-patient 8	123	130	3	3
TLE-patient 9	99	100	2	2
TLE-patient 10	28	50	5	5
Absence-patient 1	857	929	5	3
Absence-patient 2	782	980	6	6
Absence-patient 3	1112	1225	2	2
Absence-patient 4	-	67	1	0
Absence-patient 5	2392	2519	6	6
Absence-patient 6	723	744	7	6
Absence-patien t7	623	1082	3	2

TABLE 7.4: Prediction time using SL model on all subject types.

and used the PLI and SL prediction models (see Table 5.11 and Table 6.18).

In the PLI prediction model, CPL and assortativity in the beta, gamma and delta bands were the features extracted to detect the three patterns, characterized by variations in the CPL value in the gamma band (see Table 5.11 and Table 6.18). The model achieved 100% sensitivity and specificity with a false prediction rate of 0.00001/h. As seen in Table 7.2, the maximum prediction time was 299 seconds (5 minutes), as demonstrated by two TLE subjects, while the prediction time was 1,120 seconds (19 minutes), as demonstrated by an Absence subject. The maximum prediction time here is dependent on the length of recorded data before the seizure's onset.

In the SL prediction model, transitivity in the beta, gamma and delta bands was the feature used to identify the common pattern. Two patterns - seizure and no seizure – in the trend of transitivity value were shown only by patients with both types of epilepsy. The prediction model was formulated using Linear SVM, with 100% and 94.4% as sensitivity and specificity respectively. Its false prediction rate was 0.01/h, lower than that in the first model. As seen in Table 7.4, the maximum prediction time was 2,392 seconds (40 minutes) for the fifth patient with Absence seizures, while it was 298 seconds (5 minutes) for the first TLE subject. Table 7.5 provides a comparison between the two general predictive models.

In general, both prediction models achieved a higher sensitivity (100%) than those in

Model	**Sensitvity (%)**	**Specificity (%)**	**Network features used**	**Classifier**	**FPR (/h)**	**Predicting time** (mins)
PLI-based model	100	100	CPL in gamma band	Bagged Tree	0.00001	19
SL-based model	100	94.4	Transitivity in delta band	Linear SVM	0.01	40

TABLE 7.5: Summary of all prediction models created in this study.

previous studies (Chu et al. 2017, Li et al. 2013)(86.67%, 75.8%). The maximum prediction time of the second general prediction model was greater (40 minutes) than that of the first. The false prediction rates (0.00001/h and 0.01/h for the first and second model respectively) were lower than the models in previous studies (Chu et al. 2017, Alotaiby et al. 2017, Iqbal et al. 2015) (0.367/h, 0.39/h and 0.38/h, respectively). The limitation of both general models is that they use only one seizure type for each type of partial and generalized epilepsy: TLE and Absence epilepsy, respectively.

Chapter 8

Conclusions and future work

8.1 Conclusion

The main goal of this thesis was to formulate a general prediction model for epileptic seizures by applying functional brain connectivity networks with machine learning to achieve a low false prediction rate. The aim was to develop this model to contribute in predicting both TLE and Absence-type seizures.

This goal was achieved in Chapter 7 through formulating two general prediction models with a lower false prediction rate (0.00001/h and 0.01/h) than those in earlier studies (Iqbal et al. 2015, Chu et al. 2017, Alotaiby et al. 2017, Zandi et al. 2010) (0.38/h, 0.367/h, 0.39/h and 0.12/h, respectively). Additionally, the earlier models mentioned in Chapter 2 have low sensitivity. This is a limitation that has been improved by this research: both the general models developed achieved the high sensitivity of 100%.

Both general prediction models were developed using three synchronisation measures - PLV, PLI, and SL - for the two types of epilepsy. All the models improved the limitations of those mentioned in the literature by achieving high sensitivity combined with a low false prediction rate and long prediction time. In doing this, all objectives of the research were achieved, as described in Chapters 4, 5, 6 and 7. All the prediction models developed contribute to the field of predicting epileptic seizures in terms of improved performance. In addition, the general prediction models are the first to predict more than one type of epileptic seizure.

The models are, however, limited by the sample of patients available, which is relatively small. Moreover, the duration of recording before the onset of seizures was short, in many of the patients' data, and the study's TLE subjects all experienced seizures originating from only the right temporal lobe. Further, the general models developed were applied only to one sub-type in each of the two types of epilepsy. This is because the database available during the study featured only patients with right TLE and Absence seizures.

8.2 Future works proposed

The two general models proposed in the thesis will contribute to the study area of developing a closed-loop system for preempting epilepsy. In these systems, the doctor provides appropriate treatment according to the data collected by the prediction model. The research can be extended to implement further work, as follows:

- The prediction models developed in this study achieved a strong performance – low FPR and high sensitivity - but are limited by the data available on the patients used - 10 TLE and seven Absence subjects. Therefore, in future work the models should be applied to a larger sample to assess their validity, checking if their performance is altered in terms of either the false prediction rate or decreased sensitivity.

- The sample of patients available to the study had only one TLE subject with a maximum of 40 minutes recorded time before seizure onset, while the remaining subjects had less than 5 minutes. Similarly, the available data on subjects with Absence seizures had a maximum of 20 minutes recorded time, and for only one subject. Further work needs to be done to assess the relationship between the extent of the recorded time before seizure onset and the duration of the prediction time that is achieved. This can be done by exploring the data of patients with a more prolonged recording before seizure than those in this study.

- In the research, only data on patients with right TLE was available for prediction. Therefore, further work might explore the possibility using the developed models to predict seizures in patients with left TLE, making sure that the models can predict TLE in general.

- The epileptic seizures types examined in this study were used to represent just two sub-types of the two general types of epilepsy - generalized and partial. As future work, the prediction of seizures in patients with the remaining types of generalized epilepsy such as, tonic-clonic, atonic, clonic, myoclinic seizures; and partial epilepsy such as, simple and complex partial seizures could be attempted using the general models developed in this research. By doing this, we can explore the possibility of generalizing these prediction models to all types of epileptic seizure.

- Finally, in a future research, it would be valuable to explore the possibility of formulating high-performance prediction models using effective brain connectivity measures in contrast to those used in this study, for instance Granger's Causality and Dynamic Causal Modeling (DCM) (to provide the causality between two times series).

Bibliography

Aarabi, A. & He, B. (2012), 'A rule-based seizure prediction method for focal neocortical epilepsy', *Clinical Neurophysiology* **123**(6), 1111–1122.

Ahmadlou, M., Adeli, H. & Adeli, A. (2012), 'Fractality analysis of frontal brain in major depressive disorder', *International Journal of Psychophysiology* **85**(2), 206–211.

Alkan, A., Koklukaya, E. & Subasi, A. (2005), 'Automatic seizure detection in eeg using logistic regression and artificial neural network', *Journal of Neuroscience Methods* **148**(2), 167–176.

Alotaiby, T. N., Alshebeili, S. A., Alotaibi, F. M. & Alrshoud, S. R. (2017), 'Epileptic seizure prediction using CSP and LDA for scalp EEG signals', *Computational intelligence and neuroscience* **2017**.

Arnhold, J., Grassberger, P., Lehnertz, K. & Elger, C. E. (1999), 'A robust method for detecting interdependences: application to intracranially recorded EEG', *Physica D: Nonlinear Phenomena* **134**(4), 419–430.

Arora, P., Varshney, S. et al. (2016), 'Analysis of k-means and k-medoids algorithm for big data', *Procedia Computer Science* **78**, 507–512.

Bassett, D. S. & Bullmore, E. T. (2017), 'Small-world brain networks revisited', *The Neuroscientist* **23**(5), 499–516.

Bola, M. & Sabel, B. A. (2015), 'Dynamic reorganization of brain functional networks during cognition', *Neuroimage* **114**, 398–413.

Caliński, T. & Harabasz, J. (1974), 'A dendrite method for cluster analysis', *Communications in Statistics-theory and Methods* **3**(1), 1–27.

Carney, P. R., Myers, S. & Geyer, J. D. (2011), 'Seizure prediction: methods', *Epilepsy & behavior* **22**, S94–S101.

Casson, A. J. & Rodriguez-Villegas, E. (2011), 'Interfacing biology and circuits: Quantification and performance metrics', *CMOS biomicrosystems: Where electronics meet biology* pp. 1–32.

Caton, R. (1875), 'Electrical currents of the brain', *The Journal of Nervous and Mental Disease* **2**(4), 610.

Cho, D., Min, B., Kim, J. & Lee, B. (2016), 'Eeg-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **25**(8), 1309–1318.

Chu, H., Chung, C. K., Jeong, W. & Cho, K.-H. (2017), 'Predicting epileptic seizures from scalp eeg based on attractor state analysis', *Computer methods and programs in biomedicine* **143**, 75–87.

Coben, R. & Evans, J. R. (2010), *Neurofeedback and neuromodulation techniques and applications*, Academic Press.

Cortes, C. & Vapnik, V. (1995), 'Support-vector networks', *Machine learning* **20**(3), 273–297.

Crespo-Garcia, M., Atienza, M. & Cantero, J. L. (2008), 'Muscle artifact removal from human sleep EEG by using independent component analysis', *Annals of biomedical engineering* **36**(3), 467–475.

Dien, J. (1998), 'Issues in the application of the average reference: Review, critiques, and recommendations', *Behavior Research Methods, Instruments, & Computers* **30**(1), 34–43.

Douw, L., De Groot, M., Van Dellen, E., Heimans, J. J., Ronner, H. E., Stam, C. J. & Reijneveld, J. C. (2010), "functional connectivity'is a sensitive predictor of epilepsy diagnosis after the first seizure', *PLoS One* **5**(5), e10839.

EEG.Healthy.Schizophrenia (2019), 'Eeg of healthy adolescents and adolescents with symptoms of schizophrenia', http://brain.bio.msu.rueeg_schizophrenia.htm. Accessed on 2019-07-14.

Esteller, R., Echauz, J., D'alessandro, A., Vachtsevanos, G. & Litt, B. (2001), Feature parameter optimization for seizure detection/prediction, *in* '2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society', Vol. 2, IEEE, pp. 1711–1714.

Fisher, R., Stein, A. & Karis, J. (1997), 'Epilepsy for the neuroradiologist.', *American journal of neuroradiology* **18**(5), 851–863.

Friston, K. J. (2011), 'Functional and effective connectivity: a review', *Brain connectivity* **1**(1), 13–36.

Gabor, A. (1998), 'Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies'.

Grigg-Damberger, M. & Foldvary-Schaefer, N. R. (2012), *Sleep-related Epilepsy and Electroencephalography, An Issue of Sleep Medicine Clinics-E-Book*, Vol. 7, Elsevier Health Sciences.

Haddad, T., Talbi, L., Lakhssassi, A., Naim, B.-H. & Aouini, S. (2014), Epilepsy seizure prediction using graph theory, *in* '2014 IEEE 12th International New Circuits and Systems Conference (NEWCAS)', IEEE, pp. 293–296.

Han, J., Kamber, M. & Pei, J. (2012), 'Data mining: concepts and techniques, waltham, MA', *Morgan Kaufman Publishers* **10**, 978–1.

He, Y. & Yang, F. (2016), Temporal evolution analysis of functional connectivity in epilepsy based on weighted complex networks, *in* '2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)', IEEE, pp. 1–3.

Herwig, U., Satrapi, P. & Schönfeldt-Lecuona, C. (2003), 'Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation', *Brain topography* **16**(2), 95–99.

Huang, Y., Zhang, J., Cui, Y., Yang, G., He, L., Liu, Q. & Yin, G. (2017), 'How different EEG references influence sensor level functional connectivity graphs', *Frontiers in neuroscience* **11**, 368.

Humphries, M. D. & Gurney, K. (2008), 'Network 'small-world-ness': a quantitative method for determining canonical network equivalence', *PloS one* **3**(4), e0002051.

Hung, S.-H., Chao, C.-F., Wang, S.-K., Lin, B.-S. & Lin, C.-T. (2010), Vlsi implementation for epileptic seizure prediction system based on wavelet and chaos theory, *in* 'TENCON 2010-2010 IEEE Region 10 Conference', IEEE, pp. 364–368.

Hyvarinen, A. (1999), 'Fast and robust fixed-point algorithms for independent component analysis', *IEEE transactions on Neural Networks* **10**(3), 626–634.

Iqbal, S., Khan, Y. U. & Farooq, O. (2015), Nonlinear analysis of eeg for seizure prediction, *in* '2015 Annual IEEE India Conference (INDICON)', IEEE, pp. 1–5.

Islam, M. K., Rastegarnia, A. & Yang, Z. (2016), 'Les méthodes de détection et de rejet d'artefact de l'EEG de scalp: revue de littérature', *Neurophysiologie Clinique* **46**(4-5), 287–305.

Kassambara, A. (2017), *Practical guide to cluster analysis in R: Unsupervised machine learning*, Vol. 1, Sthda.

Kaufman, L. & Rousseeuw, P. J. (2009), *Finding groups in data: an introduction to cluster analysis*, Vol. 344, John Wiley & Sons.

Kodinariya, T. M. & Makwana, P. R. (2013), 'Review on determining number of cluster in k-means clustering', *International Journal* **1**(6), 90–95.

La Vaque, T. (1999), 'The history of eeg hans berger: psychophysiologist. a historical vignette', *Journal of Neurotherapy* **3**(2), 1–9.

Lachaux, J.-P., Rodriguez, E., Martinerie, J. & Varela, F. J. (1999), 'Measuring phase synchrony in brain signals', *Human brain mapping* **8**(4), 194–208.

Le Van Quyen, M., Martinerie, J., Navarro, V., Boon, P., D'Havé, M., Adam, C., Renault, B., Varela, F. & Baulac, M. (2001), 'Anticipation of epileptic seizures from standard eeg recordings', *The Lancet* **357**(9251), 183–188.

Li, S., Zhou, W., Yuan, Q. & Liu, Y. (2013), 'Seizure prediction using spike rate of intracranial eeg', *IEEE transactions on neural systems and rehabilitation engineering* **21**(6), 880–886.

Liu, Y., Li, Z., Xiong, H., Gao, X. & Wu, J. (2010), Understanding of internal clustering validation measures, *in* '2010 IEEE International Conference on Data Mining', IEEE, pp. 911–916.

Lu, J., Cheng, W., He, D. & Zi, Y. (2019), 'A novel underdetermined blind source separation method with noise and unknown source number', *Journal of Sound and Vibration* **457**, 67–91.

Marella, S. (2012), 'Eeg artifacts', https://www.slideshare.net/anjupaed/eeg-artifacts. Accessed on 2012-12-10.

Mohajer, M., Englmeier, K.-H. & Schmid, V. J. (2011), 'A comparison of gap statistic definitions with and without logarithm function', *arXiv preprint arXiv:1103.4767* .

Montez, T., Linkenkaer-Hansen, K., van Dijk, B. W. & Stam, C. J. (2006), 'Synchronization likelihood with explicit time-frequency priors', *Neuroimage* **33**(4), 1117–1125.

Mormann, F., Kreuz, T., Rieke, C., Andrzejak, R. G., Kraskov, A., David, P., Elger, C. E. & Lehnertz, K. (2005), 'On the predictability of epileptic seizures', *Clinical neurophysiology* **116**(3), 569–587.

Mormann, F., Lehnertz, K., David, P. & Elger, C. E. (2000), 'Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients', *Physica D: Nonlinear Phenomena* **144**(3-4), 358–369.

Myers, M. H., Padmanabha, A., Hossain, G., de Jongh Curry, A. L. & Blaha, C. D. (2016), 'Seizure prediction and detection via phase and amplitude lock values', *Frontiers in human neuroscience* **10**, 80.

Naik, G. R. & Kumar, D. K. (2011), 'An overview of independent component analysis and its applications', *Informatica* **35**(1).

Niedermeyer, E. & da Silva, F. L. (2005), *Electroencephalography: basic principles, clinical applications, and related fields*, Lippincott Williams & Wilkins.

Niso, G., Bruña, R., Pereda, E., Gutiérrez, R., Bajo, R., Maestú, F. & Del-Pozo, F. (2013), 'HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity', *Neuroinformatics* **11**(4), 405–434.

Noachtar, S. & Rémi, J. (2009), 'The role of eeg in epilepsy: a critical review', *Epilepsy & Behavior* **15**(1), 22–33.

Organization, W. H. (2019), 'Epilepsy', https://www.who.int/news-room/fact-sheets/detail/epileps. Accessed on 2019-15-07.

Ouyang, G., Li, X., Li, Y. & Guan, X. (2007), 'Application of wavelet-based similarity analysis to epileptic seizures prediction', *Computers in Biology and medicine* **37**(4), 430–437.

Petersen, E. B., Duun-Henriksen, J., Mazzaretto, A., Kjær, T. W., Thomsen, C. E. & Sorensen, H. B. (2011), Generic single-channel detection of absence seizures, *in* '2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society', IEEE, pp. 4820–4823.

Pijnenburg, Y., Vd Made, Y., Van Walsum, A. V. C., Knol, D., Scheltens, P. & Stam, C. (2004), 'EEG synchronization likelihood in mild cognitive impairment and alzheimer's disease during a working memory task', *Clinical neurophysiology* **115**(6), 1332–1339.

Pradhan, N., Sadasivan, P. & Arunodaya, G. (1996), 'Detection of seizure activity in eeg by an artificial neural network: A preliminary study', *Computers and Biomedical Research* **29**(4), 303–313.

Rosales, F., Garcia-Dopico, A., Bajo, R. & Nevado, A. (2015), 'An efficient implementation of the synchronization likelihood algorithm for functional connectivity', *Neuroinformatics* **13**(2), 245–258.

Rousseeuw, P. J. (1987), 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis', *Journal of computational and applied mathematics* **20**, 53–65.

Rubinov, M. & Sporns, O. (2010), 'Complex network measures of brain connectivity: uses and interpretations', *Neuroimage* **52**(3), 1059–1069.

Sanei, S. & Chambers, J. A. (2013), *EEG signal processing*, John Wiley & Sons.

Scholz, J., Klein, M. C., Behrens, T. E. & Johansen-Berg, H. (2009), 'Training induces changes in white-matter architecture', *Nature neuroscience* **12**(11), 1370–1371.

Service, G. D. (2015), 'Epilepsy and driving', https://www.gov.uk/epilepsy-and-driving. Accessed on 15-07-2020.

Shahid, R., Bertazzon, S., Knudtson, M. L. & Ghali, W. A. (2009), 'Comparison of distance measures in spatial analytical modeling for health service planning', *BMC health services research* **9**(1), 1–14.

Sporns, O. (2007), 'Brain connectivity', *Scholarpedia* **2**(10), 4695.

Sporns, O. (2010), *Networks of the Brain*, MIT press.

Stam, C. J., Nolte, G. & Daffertshofer, A. (2007), 'Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources', *Human brain mapping* **28**(11), 1178–1193.

Stam, C. J. & Reijneveld, J. C. (2007), 'Graph theoretical analysis of complex networks in the brain', *Nonlinear biomedical physics* **1**(1), 3.

Stam, C. J. & Van Dijk, B. (2002), 'Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets', *Physica D: Nonlinear Phenomena* **163**(3-4), 236–251.

Suarez-Alvarez, M. M., Pham, D.-T., Prostov, M. Y. & Prostov, Y. I. (2012), 'Statistical approach to normalization of feature vectors and clustering of mixed datasets', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **468**(2145), 2630–2651.

Sweeney, K. T., Ward, T. E. & McLoone, S. F. (2012), 'Artifact removal in physiological signals—practices and possibilities', *IEEE transactions on information technology in biomedicine* **16**(3), 488–500.

Thatcher, R. W., Biver, C. J., North, D. M. & Thatcher, C. R. W. (2004), 'EEG and brain connectivity: A tutorial', *Unpublished manuscript* .

Tibaduiza, D., Mujica, L., Anaya, M., Rodellar, J. & Güemes, A. (2012), Principal component analysis vs independent component analysis for damage detection, *in* 'Proceedings of the sixth European workshop on Structural Health Monitoring', Vol. 2, pp. 3–6.

Urigüen, J. A. & Garcia-Zapirain, B. (2015), 'EEG artifact removal—state-of-the-art and guidelines', *Journal of neural engineering* **12**(3), 031001.

Usman, S. M., Usman, M. & Fong, S. (2017), 'Epileptic seizures prediction using machine learning methods', *Computational and mathematical methods in medicine* **2017**.

van den Heuvel, M. P., de Lange, S. C., Zalesky, A., Seguin, C., Yeo, B. T. & Schmidt, R. (2017), 'Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations', *Neuroimage* **152**, 437–449.

Wang, S., Chaovalitwongse, W. A. & Wong, S. (2013), 'Online seizure prediction using an adaptive learning approach', *IEEE transactions on knowledge and data engineering* **25**(12), 2854–2866.

Watts, D. J. & Strogatz, S. H. (1998), 'Collective dynamics of 'small-world'networks', *nature* **393**(6684), 440–442.

Webber, W., Lesser, R. P., Richardson, R. T. & Wilson, K. (1996), 'An approach to seizure detection using an artificial neural network (ann)', *Electroencephalography and clinical Neurophysiology* **98**(4), 250–272.

Yadollahpour, A. & Jalilifar, M. (2015), 'Seizure prediction methods: a review of the current predicting techniques', *Biomedical and Pharmacology Journal* **7**(1), 153–162.

Zandi, A. S., Tafreshi, R., Javidan, M. & Dumont, G. A. (2010), Predicting temporal lobe epileptic seizures based on zero-crossing interval analysis in scalp eeg, *in* '2010 Annual International Conference of the IEEE Engineering in Medicine and Biology', IEEE, pp. 5537–5540.