Machine Learning Approach for Epileptic Seizure Detection

Mirwais Farahi

Department of Post-graduate Studies and Research in Computer Science

Mangalore University

Supervisor

Dr. Doreswamy

In partial fulfillment of the requirements for the degree of

Master of Science in Computer Science

July, 2019

Acknowledgements

I take this opportunity to express my sincere thanks to my supervisor Dr. Doreswamy, Professor, Department of PostGraduate Studies and Research in Computer Science, Mangalore University, for providing me this project and other relevant infrastructures work in the Department, for his all-round guidance, timely help at every stage of this project and the valuable suggestions and unlimited support.

I would like to express of gratitude to Dr. Manjaiah D H, Chairman and Professor, Department of PostGraduate Studies and Research in Computer Science, Mangalore University, for according persimmon and for proving all kinds of infrastructure falsities in the department to carry out this project successfully.

It is my gratitude to express my thanks to Dr. B.H. Shekar, Professor, Dr. H L.Shashirekha Professor, and Mr. Prakash M, Assistant Professor, Department of Postgraduate Studies and Research in Computer Science, an Mr.Ibrahim, Mr. Mohammad Kazim and all other research scholars, who helped me in all ways of walk through programming and coding the project in Python and other guests and visiting faculties for their kind guidance and supports directly and indirectly during project tenure in the department.

I would like to express my heart full gratitude to my parents, friends, well-wishers who have always inspired and blessed me, including those whom I may have inadvertently failed to mention. Last but not the least; it gives me immense pleasure and a great sense of fulfillment to thank all those who have directly and indirectly helped in the completion of my project work.

Mirwais Farahi

Abstract

Epilepsy is the name of a neurological disorder of the human brain, which is characterized by chronic disorders and occurs at random to interrupt the normal function of the brain. The diagnosis and analysis of epileptic seizure is made with the help of Electroencephalography (EEG). In order to detect seizure, it involves the interpretation of long EEG records by the expert physicians, which is time-consuming and need high human efforts. Thus, this study aims to construct an automatic seizure detection system to analyze epileptic EEG signals. The CHB-MIT Scalp EEG recording of patients is used in this work for the experiment purpose. The Welch Fast Fourier Transform is used to convert time domain features to the frequency domain. The statistical features are extracted respectively in the time domain and frequency domain. The ANOVA based feature selection is used to deduct variables. The Random Undersampling (RUS) and Synthetic Minority Oversampling Technique (SMOTE) methods are used to solve the data imbalance problem. Four machine learning algorithms, including decision tree classifier (DTC), extra-decision tree classifier (EDTC), Linear Discriminant Analysis Classifier(LDAC), Quadratic Discriminant Classifier(QDC), Random Forest Classifier (RFC), Gradient Boosting Classifier (GBC), Multilayer Perceptron Classifier (MLPC), and Stochastic Gradient Descent Classifier (SGDC) are used to classify the data. As a result, the per-

formance of the proposed classifier is 99.48% of accuracy, 99.79% of sensitivity, and 99.17% of specificity. The system might be a helpful tool for doctors to make a more reliable and objective analysis of patient EEG records.

Contents

List of Figures

List of Tables

Abbreviations

EEG	Electroencephalogram
EMG	Electromyogram
Linear Discriminant Analysis	LDA
Stochastic Gradient Descent	SGD
classification and Regression Tree	CART
Multilayer Perceptron	MLP
Random Forests	RF
Extremely Randomized Decision Trees	EDT
Quadratic Discriminant Classifier	QDC
Gradient Boosting Classifier	GBC
Root Mean Square	RMS
Coefficient of Variation	COV
Power Spectral Density	PSD
Analysis of Variance	ANOVA
Random Under-sampling	RUS
Synthetic Minority Over-sampling Technique	SMOTE

Chapter 1

Introduction

1.1 Problem Statement

Epilepsy is the name of a neurological disorder of the human brain, which is characterized by chronic disorders and occurs at random to interrupt the normal function of the brain. [11; 12]. As per World Health Organization (WHO, 2017) fact sheet around 50 million people are suffering from epilepsy. Three fourth of the people living with the epilepsy are not getting treatment and they commonly suffer from stigma and discrimination [13; 14; 15].

Seizures disorder can be partial (focal) limited to one part of the brain, or they can be general that exits in both halves of the brain. During a partial seizure, the signs are partial tic or numbness, losing the ability of expression and involuntary behaviors of talking to himself, scratching, walking around, blinking and chewing. In a generalized seizure, the most common type is the tonic-clonic. It causes upward gaze, cyanotic lips, spasticity, stiff limbs, and uncontrollable drooling. [16; 17].

The diagnosis and analysis of epileptic seizure is made with the help of Electroencephalography (EEG). The physicians and scientists use EEG to study brain functions and diagnose the neurological disorder since it contains physiological information of the brain.The standard EEG signals will appear as spiking waves during seizure activities [3]. In order to detect seizure, it involves the interpretation of long EEG records by the expert physicians, which is time-consuming and need high human efforts. Thus, an automatic seizure detection system is required to reduce the volume of data for the physicians. It will help experts only to study those parts of the EEG data that is seizure effected. Some studies have conducted on EEG signals classification into normal and seizure states. The majority of previous studies on seizure detection and prediction have concentrated on patient-specific predictors, where a classifier is trained on one person and tested on the same person [18; 19; 20; 21; 22]. The aim of this study is to categorize EEG records in significant states of normal and seizure.

1.2 Objectives and Contributions

The objective of the proposed study is to advance existing studies and provide an automatic epilepsy detection and prediction system with a less computational cost for clinical uses and applications. The objectives are categorized as follows:

1. To find the best combination of EEG signal processing methods which could generate an intelligent time/frequency feature extractor capable of:
    - Robustness to artifacts
    - Handling non-linearity
    - Generating a feature set which brings out maximum representation of the signal
    - Low memory consumption
    - Real-time signal processing capability

2. To classify epileptic and non-epileptic EEG signals using supervised machine learning algorithms.

3. To compare the performance of various supervised machine learning classifiers and find out the most accurate classification model

1.3 Overview of the Thesis

Chapter 1 describes an in-depth overview of the problem followed by the thesis objectives.

Chapter 2 describes the theoretical background of Electroencephalogram (EEG) signals and its technical background. Then epilepsy, along with its types, is explained in details. In this chapter, the significance of EEG signals and various applications of EEG signals, such as EEG in epilepsy diagnosis, EEG in brain-computer interfaces (BCI) and computer-aided EEG diagnosis are studied.

In chapter 3, various supervised machine learning algorithms such as linear machine learning algorithms, non-linear machine learning algorithms, Neural Networks, and different ensemble methods are studied.

Chapter 4 contains the literature review in which the previous works are explained in details. This chapter has divided the literature into three sections of the international review, national review, and the studies on CHB-MIT EEG dataset.

Chapter 5 describes the dataset used in this work, followed by the proposed work methodology. Additionally, the experiment details and results of different classifiers are explained. The result is also compared with some existing studies.

Chapter 6 concluded this research. Furthermore, the future direction of the proposed work is presented in this chapter.

Chapter 2

Theoretical Background

2.1 Electroencephalogram (EEG) and its Background

2.1.1 Introduction

Electroencephalography (EEG) is the measurement of potentials that reflect the electrical activity of the living brain. It is medical imaging available technique that provides evidence of how the brain functions over time.

The first electrical currents in the brain was introduced to the world in 1875 by an English physician named Richard Caton. Caton observed the existence of EEG in the brains of rabbits and monkeys. In 1924 a neurologist Hans Berger from the University of Jena in Germany discovered the EEG recording of the human brain using his ordinary radio equipment called galvanometer. Based on Berger approach, the weak electric currents generated in the human brain can be recorded without opening the skull and represented graphically on a piece of paper. The recorded signals observed by Berger describes the functional status of the brain, such as sleep, lack of oxygen, and in brain diseases such as epilepsy. He used the

German word elektrenkephalogramm to describe the graphical representations of the electric currents generated in the brain. He suggested that brain activity changes in a consistent and recognizable way based on the functional status of the subject, such as sleep, anaesthesia, and epilepsy [23].

2.1.2 EEG Signal Generation in the Brain

An EEG signal is a measurement of currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. When brain cells (neurons) are activated, the synaptic currents are produced within the dendrites. This current generates a magnetic field measurable by electromyogram (EMG) machines and a secondary electrical field over the scalp measurable by EEG systems.

Differences of electrical potentials are caused by summed postsynaptic graded potentials from pyramidal cells that create electrical dipoles between the soma (body of a neuron) and apical dendrites, which branch from neurons (Figure 2.1). The current in the brain is generated mostly by pumping the positive ions of sodium, Na + , potassium, K + ,calcium, Ca ++ , and the negative ion of chlorine, Cl , through the neuron membranes in the direction governed by the membrane potential [1; 2].

The human head consists of different layers including the scalp, skull, brain (Figure 2.2), and many other thin layers in between. The skull attenuates the signals approximately one hundred times more than the soft tissue. On the other hand, most of the noise is generated either within the brain (internal noise) or over the scalp (system noise or external noise). Therefore, only large populations of active neurons can generate enough potential to be recordable using the scalp electrodes. These signals are later amplified greatly for display purpose.

Approximately 10 Billions neurons are developed at birth when the central

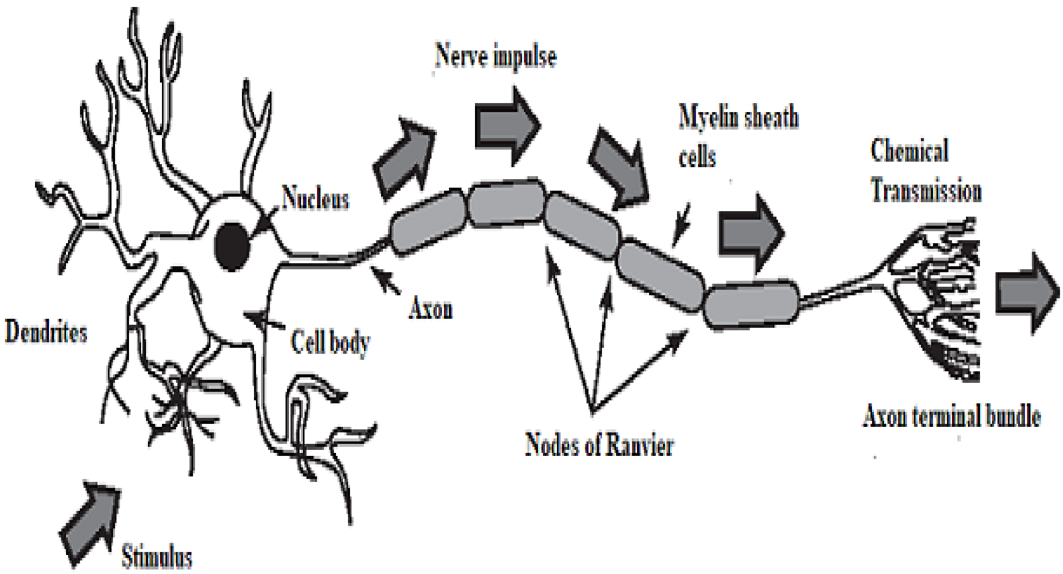


Figure 2.1: Structure of a neuron (Ref. [1]

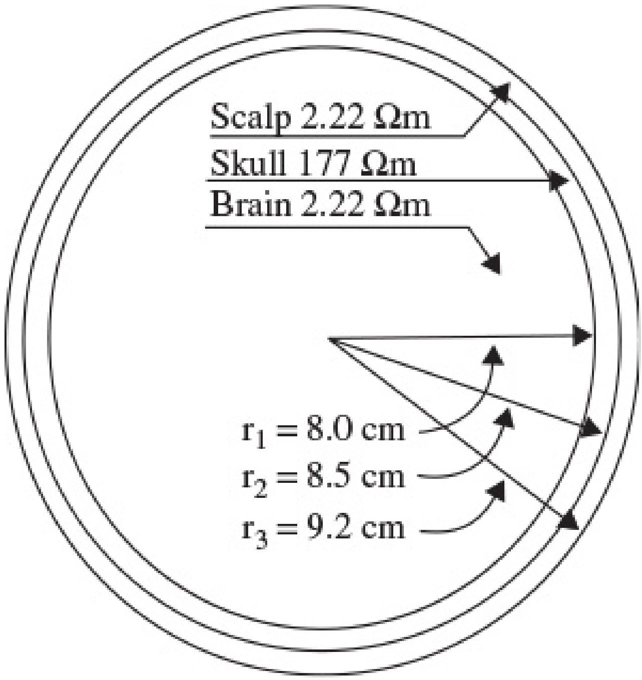


Figure 2.2: The three main layers of the brain including their approximate resistivities and thicknesses (= ohm) (Ref. [2]

nervous system (CNS) becomes complete and functional [2; 24]. This makes an average of 10,0000 neurons per cubic mm. Neurons are interconnected into neural nets through synapses. Adults have approximately 5×10^4 synapses. The number of synapses per neuron increases with age, whereas the number of neurons decreases with age. From an anatomical point of view the brain may be divided into three parts: the cerebrum, cerebellum, and brain stem. The cerebrum includes the regions for movement initiation, conscious awareness of sensation, complex analysis, and expression of emotions and behaviour. The cerebellum coordinates voluntary movements of muscles and maintains balance. The brain stem controls involuntary functions such as respiration, heart regulation, biorhythms, and neurohormone and hormone sections [2; 25]. Based on the above section it is clear that the study of EEGs paves the way for diagnosis of many neurological disorders and other abnormalities in the human body.

2.1.3 Applications of EEG

According to R.D. Bickford [26] the recorded EEG signals from humans and animals are used for investigation of the following clinical problems:

- monitoring alertness, coma, and brain death
- locating areas of damage following head injury, stroke, and tumour
- testing afferent pathways (by evoked potentials)
- monitoring cognitive engagement (alpha rhythm)
- producing biofeedback situations
- controlling anaesthesia depth (servo anaesthesia)
- investigating epilepsy and locating seizure origin

- testing epilepsy drug effects
- assisting in experimental cortical excision of epileptic focus
- monitoring the brain development
- testing drugs for convulsive effects
- investigating sleep disorders and physiology
- investigating mental disorders
- providing a hybrid data recording system together with other imaging modalities.

2.1.4 Classifications of EEG

The EEGs are measured in terms of frequency. Frequency is the number of cycles in second (in Hz), and it is one of the most important criteria for estimating abnormalities in clinical EEGs. There are five basic groups of brain waves categorized by their different frequency ranges [3]. The frequency bands from low to high frequencies are distinguished as follows:

- delta (0.5-4 Hz)
- theta (4-8 Hz)
- alpha (8-13 Hz)
- beta (13-30 Hz)
- gamma(>30 Hz)

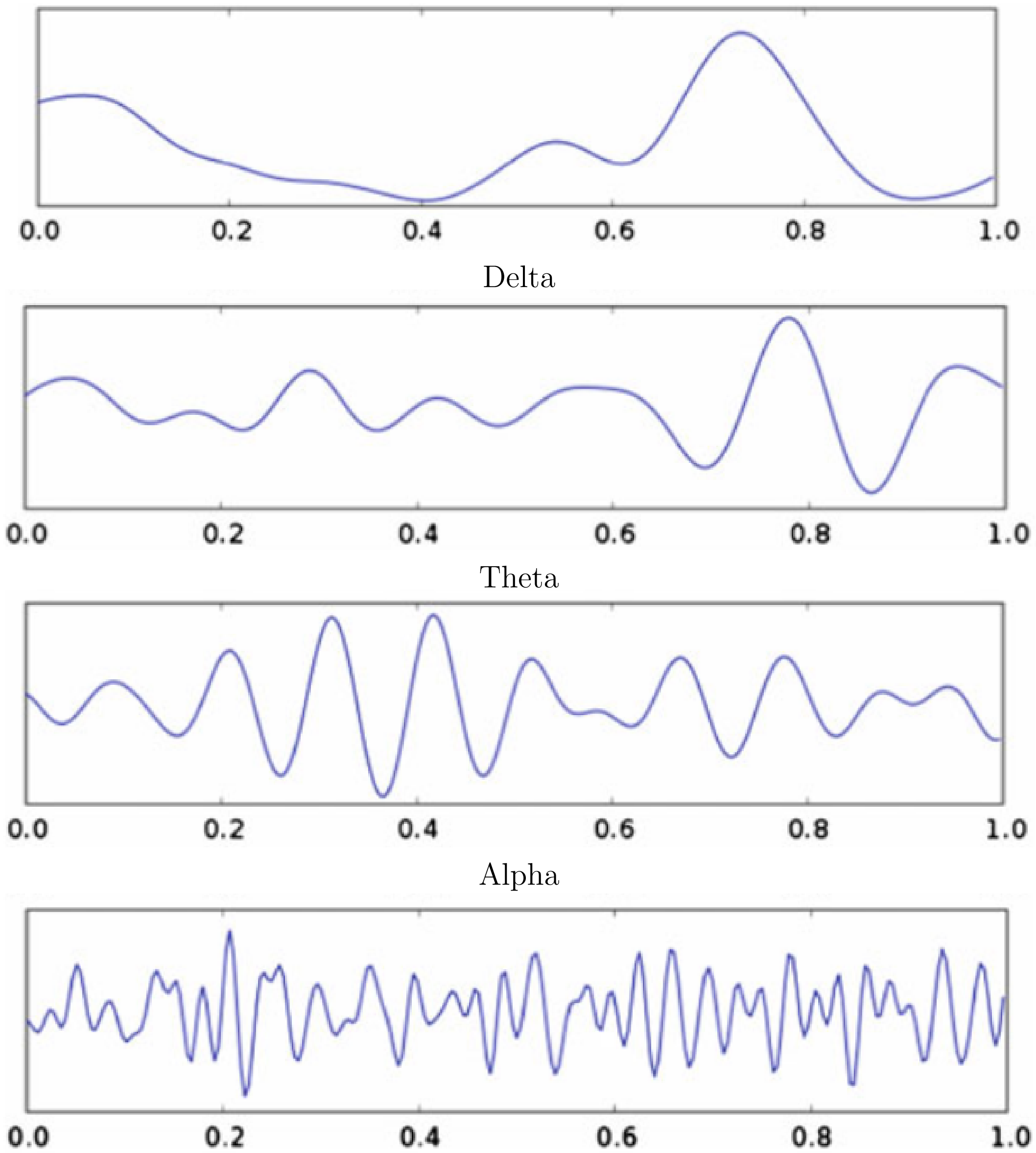
Most abnormal brain states are in higher frequencies such as epilepsy. Figure 2.3 describes examples of these EEG wave bands. The frequency range 0.5-4 Hz belongs to the delta wave in which the shape is the lowest in wave and the highest in amplitude. It is associated with waking state, deep sleep, and brain disorders. The theta wave is between 4 and 8 Hz frequency range with an amplitude usually greater than 20 V. It is associated with unconscious material, creative inspiration and deep meditation, and emotional stress, especially frustration or disappointment. The alpha wave lies in 8 to 13 Hz range, usually with 3050 m V amplitude. The alpha arises from intense mental activity, stress, and tension. The beta is in the 1330 Hz frequency range. It is the lowest amplitude and varying frequencies symmetrically on both sides in the frontal area. The beta wave is generated when the brain is aroused and actively engaged in mental activities, such as active attentions, and focusing on the outside world or solving concrete problems. The gamma wave is having the frequency from 30 Hz and up. It is sometimes defined as having a maximal frequency around 80 or 100 Hz. It is associated with various cognitive and motor functions.

2.1.5 Artefacts in EEG

Electrical signal sequences in the EEG that are having higher amplitude and different shape relative to signal sequences that are suffering by any large contamination is called artefact. The artefact in the EEG may be either patient-related or technical [4].

Patient-related artefacts: Patient-related artefacts are unwanted patient generated and physiological signals that disturb the EEG pattern. The most common sources of patient-related artefacts are:

- any minor body movements



Beta

Figure 2.3: examples of EEG wave bands (Ref. [3]

- EMG
- ECG (pulse, pace-maker)
- movement of eyes
- gamma (¿30 Hz)
- sweating

Technical artefacts: Technical artefacts are due to environmental interference or equipment problems during recording of the EEG signal. The most common sources of technical artefacts are:

- 50/60 Hz
- impedance fluctuation
- cable movements
- broken wire contacts
- too much electrode paste/jelly or dried pieces
- low battery

The artefacts from the EEG signals can be removed automatically or by the trained experts. Additional electrodes may be useful for better recording of EEG signals and free of different physiological artefacts. Technical artefacts, such as AC power line noise, can be decreased by shorter electrode wires and decreasing electrode impedance.

2.1.6 Equipment for EEG Recording

EEG recording systems basically consists of:

- electrodes with conductive media: To read the EEG signal from the head surface.
- amplifiers with filters: It brings the microvolt EEG signals in a range where they can be digitized correctly.
- A/D converter: To convert analog signals to digital signals.
- recording device: Computers or other relevant devices are used to store and display the EEG data.

EEG is the measurement of potential changes over time in a basic electric circuit carrying between the active electrode and reference electrode [4; 27]. A third electrode called ground electrode is used to get differential voltage by subtracting the voltages showing at active and reference electrodes. For single-channel EEG measurement minimal configuration is one active electrode, one (or two connected together) reference electrode, and one ground electrode. In the case of multi-channel configuration 128 or 256 active electrodes are used. Figure 2.4 shows a set of the equipment used for EEG measurement.

2.1.7 10-20 EEG recording technique

To record the EEG signals from the scalp of brain, electrodes are placed based on some standard methods. One of the most well-known approaches is 10-20 electrode placement system. The 10-20 system is an internationally recognized electrode placement method used to describe the location of scalp electrodes [28]. The "10" and "20" describes the distance between neighboring electrodes as 10 or 20 percent of the total right-left or front-back distance of the skull. Figure 2.5

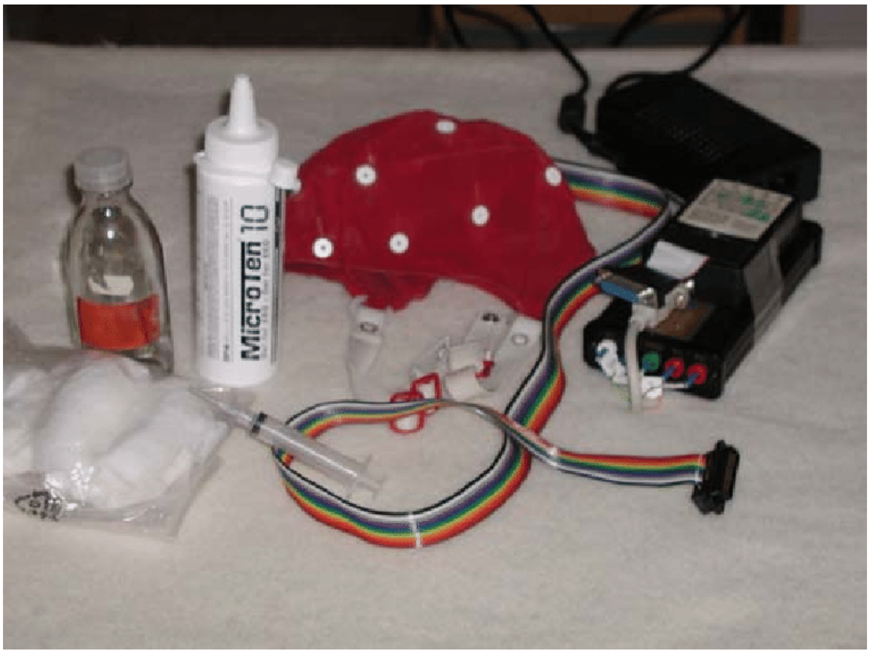


Figure 2.4: Equipment for EEG measurement: injection, amplifier unit, conductive jelly, electrode cap, and aid for disinfection. (Ref. [4]

shows the electrodes location, according to the 10-20 international system which covers all regions of the brain. A letter is used to identify each position of the lobe and a number to identify the hemisphere position. The letters F, P, T, C and O stand for Frontal, Parietal, Temporal, Central and Occipital, respectively. Odd numbers point to electrode locations on the left hemisphere, whereas even numbers point to those on the right hemisphere. A "z" refers to an electrode disposed on the midline. Since an EEG voltage signal represents the difference between voltages at two electrodes, the machine may be set up in various ways to display the EEG signal for reading. For this purpose, different placement methods of electrodes refer to montages, which are used to monitor the EEG signals. The following montages are used to monitor the EEG signals.

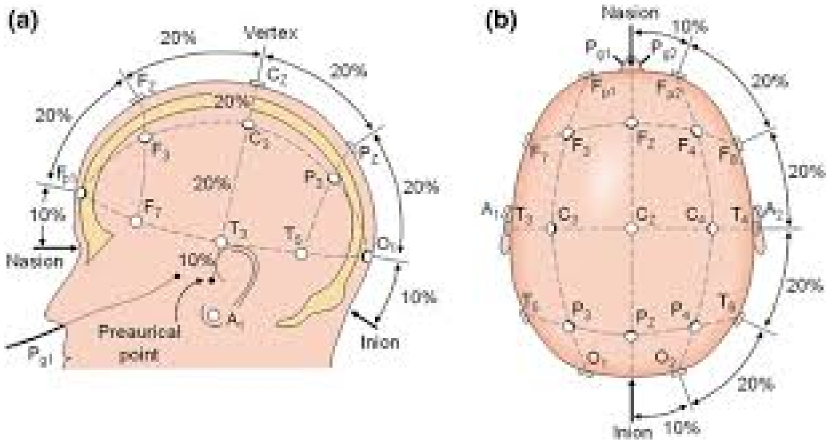


Figure 2.5: The 1020 international electrode placement system (Ref. [3]

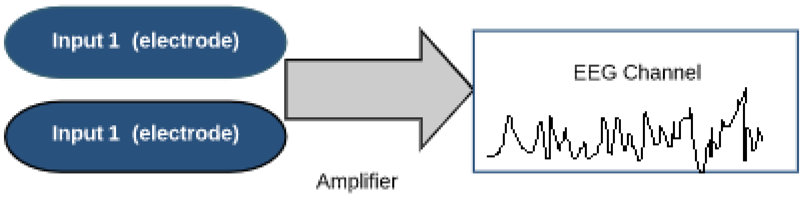


Figure 2.6: A single channel of EEG (Ref. [3]

Bipolar montage

In the bipolar montage, two electrodes usually create a channel, as shown in Figure 2.6. The bipolar montage consists of a series of channels, where each channel is identified by the difference of adjacent electrodes [3; 29; 30], as shown in Figure 2.7. For example, Figure 2.8 represents the diagram of a bipolar montage, where the channel "Fp1-F3" shows the difference in the voltage between the Fp1 and F3 electrodes. Furthermore, the next channel is "F3-C3", which shows the difference between voltages of F3 and C3, and so on, within the entire array of electrodes.

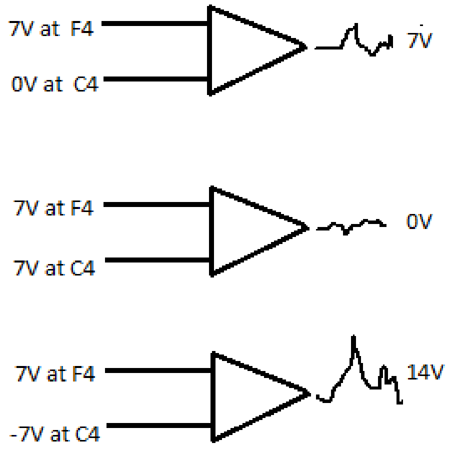


Figure 2.7: Example of a channel as the difference between two adjacent electrodes (Ref. [3]

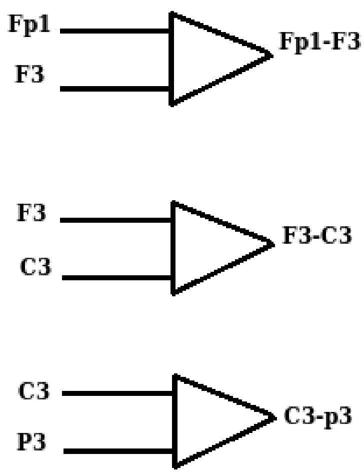


Figure 2.8: Example of bipolar montage (Ref. [3]

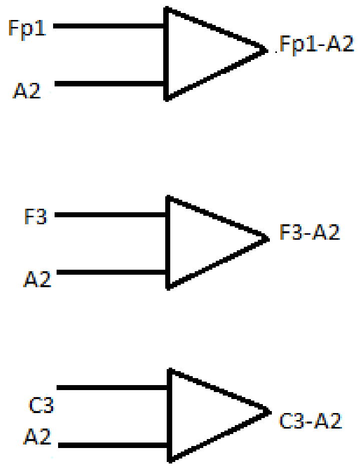


Figure 2.9: Example of referential montage (Ref. [3]

Referential montage

In referential montage, each channel is described as the difference between an active (recorded) electrode and a designated reference electrode [3; 29; 30]. There is a common reference electrode for all the channel. For example electrode A2 in Figure 2.9 is a reference electrode. There is no standard location for this reference electrode in referential montage and it is in a different location than the active electrodes.

Average reference montage

In average reference montage, the average of outputs of all amplifiers is taken, and this average signal is used as the common reference for each channel [3; 30]. Figure 2.10 shows an example of average reference montage.

Laplacian montage

Each channel is defined as the difference between an electrode and a weighted average of the surrounding electrodes [3; 30].

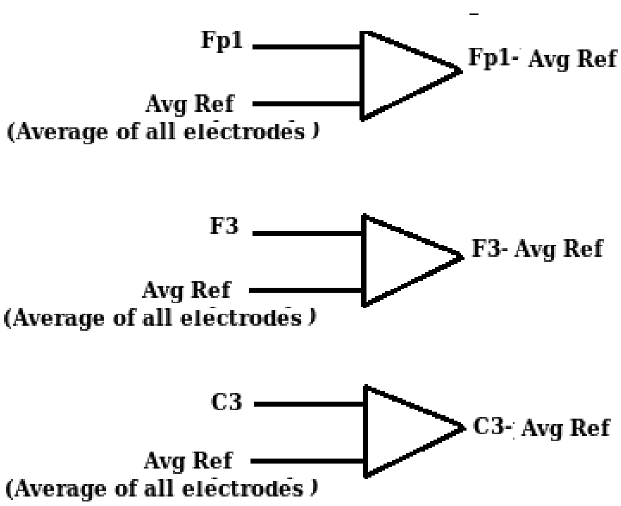


Figure 2.10: Example of average referential montage (Ref. [3]

2.2 Epilepsy and Its Background

2.2.1 Introduction to Epilepsy

Epilepsy is the name of a neurological disorder of the human brain, which is characterized by chronic disorders and occurs at random to interrupt the normal function of the brain, termed as epileptic seizures. [11; 12]. As per World Health Organization (WHO, 2017) fact sheet around 50 million people are suffering from epilepsy. Around 80% people with epilepsy brain disorder are from low- and middle-income countries. Nearly 75% of people with epilepsy in low- and middle-income countries are not receiving the necessary treatment. [13; 14; 15]. Seizure is defined as unreasonable electrical discharges interrupts the normal function of the brain, which includes low blood sugar, alcohol or drug withdrawal, high fever, or a brain concussion. Under these situations, anyone can have one or more seizures. However, a person having two or more unprovoked seizures is considered epileptic patient. There are many possible causes of epilepsy; it may develop an abnormality in brain wiring, strokes, tumors, an imbalance of nerve-signaling

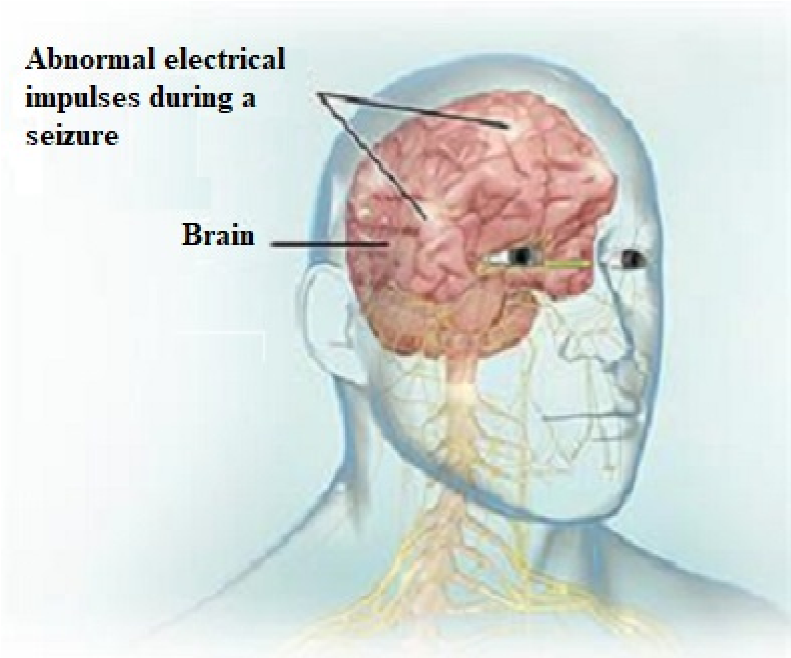


Figure 2.11: Abnormal electrical impulses of the brain during a seizure (Ref. [3]

chemicals called neurotransmitters, or some combination of these factors. In most of the cases, there may be no detectable cause for epileptic patient. Figure 2.11 shows an image of abnormal electrical impulses during a seizure. Neurons usually generate electrochemical impulses that act on other neurons, muscles, and glands to create human feelings, thoughts, and actions. In epilepsy, the regular pattern of neuronal activity becomes disrupted, causing strange emotions, sensations, and behaviors, or sometimes loss of consciousness, convulsions, and muscle spasms. During a seizure period, neurons may fire as many as 500 times a second, much faster than normal (1100V). In some people, this occurs only occasionally; for others, it may occur up to hundreds of times a day.

2.2.2 Seizure Types

The International League Against Epilepsy (ILAE) introduced an approach to group seizures, which helps doctors to identify a person's seizures and prescribe the most appropriate treatments.
Seizures are classified based on the following factors:

- Which part of the brain is affected (onset)
- whether or not a person's awareness is affected
- Whether or not seizures involve other symptoms, such as movement

The two broad categories of epileptic seizures are generalized and partial seizures.

1. **Focal or Partial Seizures**
   Focal onset seizures start when abnormal electrical brain function occurs in a large part of one hemisphere or just a small area of the brain. It is also called partial seizure. With focal onset seizure, a person may experience an aura (a warning of another seizure). The most common aura involves impending doom, euphoria, and fear. Changes in the sense of smell, visual changes, and hearing abnormalities can also be auras. Two types of focal seizures include:

   (a) Simple focal seizures
   Simple focal seizures occur in one part of the brain, and the symptoms depend on which side of the brain is affected. If the seizure starts in the occipital lobe, the back part of the brain that is involved with vision. The person sight and muscles are altered. Simple focal seizure is limited to a group of isolated muscles group, such as fingers, or larger muscles in the legs and arms. The patient may also experience nausea, sweating, or become pale.

(b) Complex focal seizures

Complex focal seizures usually occur in the temporal lobe of the brain (The area of the brain that is involved with memory and emotion functions). The person may look awake but loses consciousness and may have different unusual behaviors, such as lip smacking, gagging, screaming, crying, laughing, and running. The person may feel tired or sleepy after regaining consciousness. This is called the postictal period.

2. **Generalized Seizures**

Generalized seizures affect both sides of the brain. There is loss of consciousness and a postictal state after the seizure occurs. Types of generalized seizures include the following:

(a) Absence seizures (also called petit mal seizures)

Absence seizures or petit mal seizures are characterized by a small, altered state of consciousness and staring events. Typically, the person mouth or face may twitch, or the eyes may blink rapidly. The absence seizure may not continue more than 30 sec. After the seizure, the person may not remember what just happened and may go on with his or her activities, acting as though nothing occurred. These seizures may happen several times a day. These types of seizures almost always start between ages 4 to 12 years.

(b) Atonic Seizures

Atonic seizures (also called drop attacks) involve a sudden loss of muscle tone, and the person is limp and unresponsive. For example, in atonic seizure, a person may fall from a standing position or suddenly drop his or her head.

(c) Generalized tonic-clonic Seizures (GTC or also called grand mal seizures)
Five distinct phases characterize generalized tonic-clonic seizures. The arms, legs, and body will flex (contract), extend (straighten out), and tremor (shake), followed by a clonic period (contraction and relaxation of the muscles) and the postictal period. Not all of these phases may occur in every case with this type of seizure. During the postictal period, the person may have problems with vision or speech, feeling sleepy, and may have a bad headache, body aches or fatigue.

(d) Infantile spasms Seizure
Infantile spasms seizure disorder occurs in infants within six months of age. The child in awake state or before going to sleep has a high chance of getting this type of seizure. In this seizure, the child usually has short periods of movements of the neck, or trunk, or legs that last for a few seconds. A child may have hundreds of Infantile spasms seizures a day, which can have a serious impact on infants growth and development.

(e) Febrile Seizures
Febrile seizures are not epilepsy and associated with fever. These seizures are usually seen in children between six months and five years of age. These types of seizures that last less than 15 minutes are called simple, and typically do not have long-term neurological consequences. Seizures lasting more than 15 minutes are called complex, and there may be long-term neurological effects in the child.

2.2.3 Causes of a Seizure

A person may have one or many different types of seizures. The exact cause of seizure may not be known but the most common seizures are caused by the following:

1. In babies and infants:
    - Birth trauma
    - Fever or infection
    - Congenital (present at birth) problems
    - Metabolic or chemical imbalances in the body
2. In children, adolescents, and adults:
    - Alcohol or drugs
    - Infection
    - Head trauma
    - Genetic factors
    - Congenital conditions
    - Progressive brain disease
    - Alzheimer's disease
    - Unknown reasons
    - Stroke
    - Unknown reasons

Other possible causes of seizures may include the following:

- Neurological problems
- Brain tumor
- Use of illicit drugs
- Drug withdrawal

2.2.4 symptoms of a Seizure

A person may have different symptoms depending on the type of seizure. General symptoms or warning signs may include:

- Jerking movements of the arms and legs
- Staring
- Loss of consciousness
- Stiffening of the body
- Loss of bowel or bladder control
- Breathing problems or breathing stops
- Not responding to noise or words for brief periods
- Falling suddenly for no apparent reason, especially when associated with loss of consciousness
- Nodding the head rhythmically, when associated with loss of awareness or even loss of consciousness
- Appearing confused or in a haz
- Periods of rapid eye blinking and staring

During the seizure, the person's breathing may not be normal, and lips may become bluish.

2.2.5 Treatment of a seizure

Treatments for a seizure by a doctor may depend on the following factors:

- A person age, overall health, and medical history
- Type of the seizure
- Frequency of the seizures
- A person tolerance for specific medications, procedures, or therapies
- Expectations for the course of the condition

The treatments include

1. **Medications:** There are different types of medications applied to a patient to treat seizures and epilepsy. While taking medications, various tests may be applied to an epileptic patient to monitor the effectiveness of the medication. These tests include:
   - Blood test
   - Urine test
   - EEG recording from the patient brain
2. **Vagus nerve stimulation (VNS):** VNS is usually used for people over age 12 who have partial seizures that are not well-controlled by other approaches, such as medication. To control seizure, VNS sends small pulses of energy from vagus nerve (a large nerve in the neck) to the brain. The procedure is to place a small battery into the chest wall surgically, and then small wires are attached to the battery placed under the skin around the vagus nerve. Then the battery is programmed to send energy impulses every minute to the brain. When the person having epilepsy feels a seizure

coming on, may activate the impulses by holding a small magnet over the battery. In most cases, this may help to halt the seizure.
Some of the most common side effects that may occur with the use of VNS include:

- Hoarseness
- Pain or discomfort in the throat
- Change in voice

3. **Surgery:** Surgery of epileptic patient is done by removing that affected part of the brain by seizure. The person will be considered for surgery based on the following conditions.

   - The person seizure is not able to be controlled with medications.
   - The person is having a partial seizure
   - Seizure affected area of the person's brain is not having an important task, such as vision, memory, or speech.

   Surgery for epilepsy and seizures is quite tricky and performed by a specialized surgical unit. The operation may remove the part of the brain that is seizure affected.

2.3 Applications and Significance of EEG Signals in Health Research

2.3.1 EEG in Epilepsy Diagnosis

The diagnosis and analysis of epileptic seizure is made with the help of EEG [3]. EEG continues to play a vital role in the diagnosis and management of patients with seizure disorders. It is a convenient and relatively inexpensive way to explain the physiological manifestations of abnormal cortical excitability that carries epilepsy [3; 10]. The epileptic seizure can partial, encompass a discrete part of the brain or generalized seizure, the complete cerebral mass of the brain.The physicians and scientists use EEG to study brain functions and diagnose the neurological disorder since it contains physiological information of the brain.The standard EEG signals will appear as spiking waves during seizure activities [3].Two categories ictal (during an epileptic seizure) and interictal (between seizures) of epileptic seizure can be detected from an EEG signal. Usually, the onset of a clinical seizure is characterized by a sudden change of frequency in the EEG computation. It is generally within the alpha signal frequency band with a gradual reduction in frequency but increases in amplitude during the seizure period. It may or may not be spiky in shape. EEG helps to define seizure types and epilepsy symptoms in patients with epilepsy, thereby assisting in determining anti-epileptic medication choice and the prediction of diagnosis. Another vital contribution of EEG findings is to discover the multi-axial diagnosis of epilepsy, in terms of whether the seizure disorder is idiopathic or symptomatic, focal or generalized, or part of a specific epilepsy syndrome [10].Figure 2.12 shows usual patterns of EEG waves for normal, partial seizure and generalized seizure. Epilepsy syndrome associated with specific EEG features exists in early life or childhood [10]. Table 2.1 shows epilepsy syndromes in early stages of life. Some syndrome are well accepted; others are uncertain or may not be included in the current International League Against Epilepsy (ILAE) classification systems because of lack of sufficient data. These classifications are work in progress by developments in genetics, imaging, and molecular biology. In some cases, the epilepsy syndrome may only become visible over time, which necessitate regular electro-clinical test.

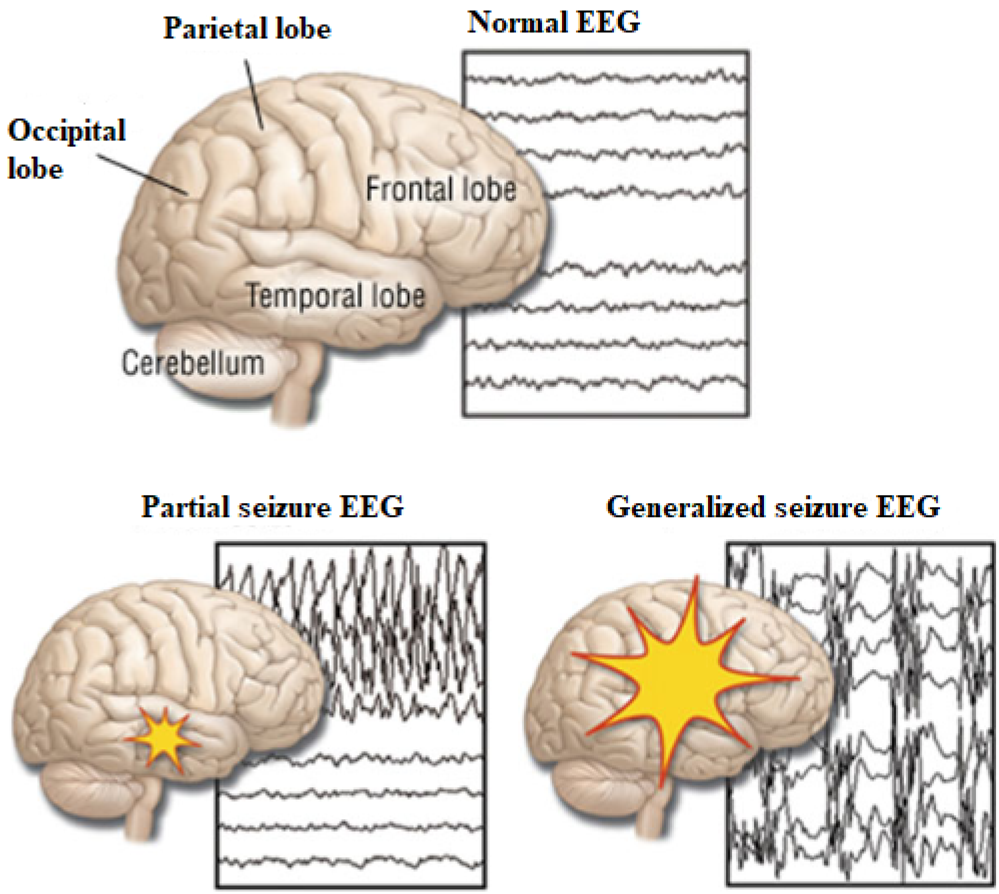


Figure 2.12: usual patterns of EEG waves for normal, partial seizure and generalized seizure (Ref. [3]

Table 2.1: Epilepsy syndromes in early life (Ref. [10])

Syndrome	Age of onset	Clinical	EEG
Benign idiopathic neonatal seizures	27 days (5th day fits)	Partial/clonic/ apnoeic	Trace pointu alternans (75%)
Early myoclonic epilepsy	First month of life	Myoclonus, partial motor, tonic spasms	Burst suppression
Early infantile epileptic encephalopathy (Otohara syndrome)	13 months	onic spasms; may be focal motor, generalised tonic-clonic seizure,myoclonic jerk	Burst suppression
Severe myoclonic epilepsy of infancy (Dravet syndrome)	Within first year	Febrile: afebrile seizures, myoclonus , atypical absences	Generalised pike wave photosensitivity (40%)
West syndrome	47 months	Infantile spasms	Hypsarrhythmia

2.3.2 EEG in Brain-Computer Interface (BCI)

Brain-Computer Interface (BCI) is an application and a new emergent field for EEG signals analysis. BCI is a technology that works as an interface between a brain and a computer. Brain-Computer Interface (BCI) is an application and a new emergent field for EEG signals analysis. BCI is a technology that works as an interface between a brain and a computer. It is a computer-based system that collects brain signals, analyze them, and interprets them into control signals (commands) that are sent to an output device to carry out the desired operation (see Figure 2.13). To estimate EEG signals, an electrode cap is placed on the head of a person. To command the device a person imagines a particular task, such as the composing of words or movement of limbs. These tasks alter the patterns of EEG signals. Computers detect and classify these patterns into various tasks in order to control a device (such as a wheelchair) or a computer application (e.g.,

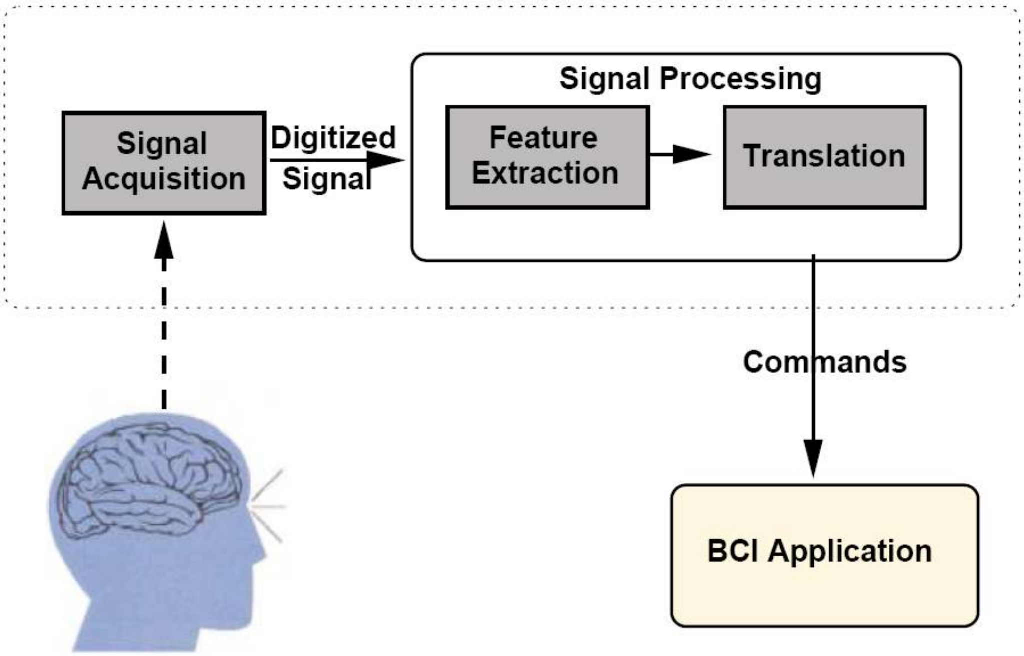


Figure 2.13: A general diagram of a BCI system. (Ref. [5]

a cursor movement).In general a BCI system consists of following six steps:

1. **Brain activity measurement:** The brain activities are recorded as EEG signals by placing electrodes on the brain. The brain activities can be recorded in two ways of invasive and non-invasive. In the invasive method, EEG signals are recorded from electrodes implanted over the brain cortex, surgery is required to implants electrodes on the brain. In the non-invasive approach, electrodes are placed on the scalp (outside the head) to record EEG signals [31].

2. **Preprocessing:** Preprocessing is a process applied on data to improve signal quality without losing information. In this step, the EEG signals are filtered to clean, denoise, and remove unwanted data from the signals to enhance the important information embedded in the signals [3; 32].

3. **Feature extraction:** The EEG recorded signals from the brain are characterized by specific features. In the feature extraction step, EEG signals are described by a few important values called features [3; 32].

4. **Translation into a command/application:** In this step, mental state of the brain is identified, and EEG signals are interpreted into control signals (commands) in order to monitor a given application, such as a robot or a computer.

5. **Feedback:** Finally, the user is provided with feedback about the identified mental state, which helps the user control his/her brain activities. The overall objective of feedback is to enhance the users performance.

A BCI can only recognize and classify particular patterns of an activity in continuous brain EEG signals that are related to a specific task or event. The mental strategy determines what a user has to do in order to produce brain patterns that the BCI system can translate. It is the foundation of any brain-computer communication system. There are two most common mental strategies: Motor imagery (MI), which is the imagination of a movement without actually performing the movement, and selective (focused) attention, which require external stimuli (can be auditory or somatosensory) provided by a BCI system.
The BCI systems can also be developed for health sector, such as epilepsy detection systems.

2.3.3 Computer-Aided EEG Diagnosis

EEG signals are complex, and analyzing them in size collected from a large number of patients makes estimation time-consuming. Therefore recently, the computer-aided diagnosis systems are proposed to make it possible to conduct an automatic neurophysiological estimation for detection of abnormalities from

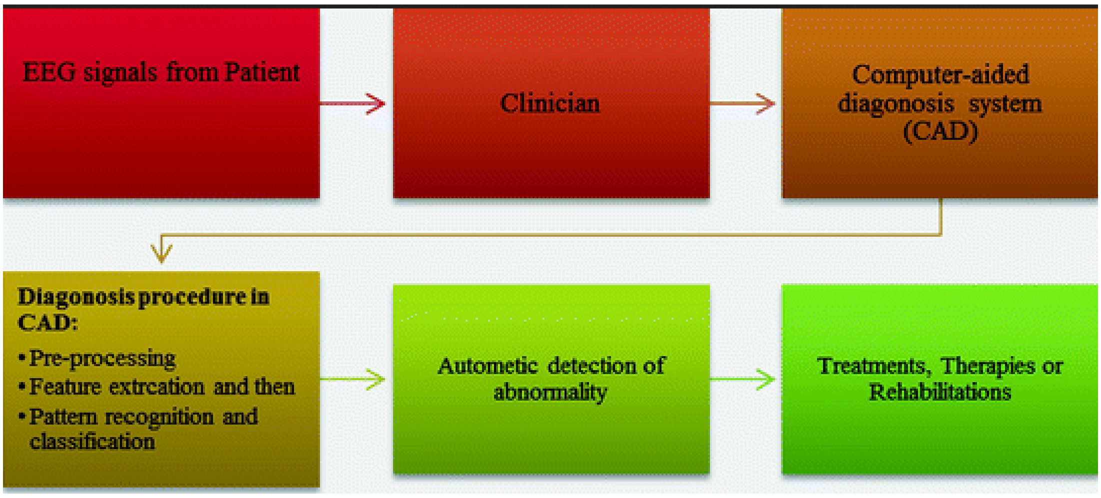


Figure 2.14: A general diagram of a BCI system. (Ref. [5]

EEG signals [3]. The CAD system consists of three main steps(Figure 2.14): pre-processing, where acquired EEG data are processed to denoise and remove unwanted data to reduce the complexity and computation time of the CAD algorithms. Feature extraction, where biomarkers of disease identification are extracted from the original data. Finally, classification step, where the extracted features are used in the classifier model as input for assigning the candidate to one of the possible classes according to the output of a classifier, such as healthy or normal. Generally, a CAD system can be classified as one of the two types. When a CAD system involves classifying all candidates into two classes, such as abnormal and normal candidates, it is called a two-class categorization system. On the other hand, if a CAD system can categorize unknown cases more than two into several types of abnormalities, it is called a multi-class categorization system.

Chapter 3

Supervised Machine Learning algorithms for EEG-based Applications

3.1 Machine Learning Algorithms

Machine learning algorithms are learning from a target function f that best maps input variables X to an output variable Y as defined in the following equation.

$$Y = f(X) + e$$

There is also error (e) that is not dependent of the input data X. This error might be error such as not having enough attributes to sufficiently characterize the best mapping from X to Y. This error is called irreducible error because no matter how good we get at estimating the target function (f), we cannot reduce this error. This error could be not having enough features to sufficiently describe the best mapping from X to Y. The error rate is called irreducible error because

it is not reducible even the target function f is correctly well estimated.
The machine learning algorithms can be of types, such as Supervised, Unsupervised, and Semi-Supervised Learning algorithms.

3.1.1 Supervised Machine Learning

Supervised learning is defined as an algorithm to learn the mapping function from the input variables (X), and an output variable (Y).

$$Y = f(X) + e$$

The intention is to approximate the mapping function extremely well that when there is new input data (X), the model predicts the output variables (Y) for that data.Supervised learning problems can be further grouped into regression and classification problems.

1. **Classification:** A classification problem is when the output variable is a category, such as a disease or no disease.

2. **Regression:** In a regression problem, the output variable is a real value, such as weight, age.

3.1.2 Unsupervised Machine Learning

Unsupervised machine learning is used when there is only input data (X) and no related output variables. The purpose of unsupervised learning is to model the distributed underlying structure in the data in order to learn more about the data. Unsupervised learning problems can be further classified into clustering and association problems.

1. **Clustering:** A clustering technique is used to discover the inherent groupings in the data, such as grouping employees by on working projects.

2. **Association:** An Association technique is used to discover rules that describe large portions of data, such as people that buy A item also tend to buy B item.

3.1.3 Semi-Supervised Machine Learning

Semi-supervised machine learning algorithms are used when there is a large amount of input data (X), and only some of the data is labeled (Y). These algorithms lie in between both supervised and unsupervised learning algorithms. For example, in a photo archive where only some of the images are labeled (e.g., cat, dog, person) and the majority of images are unlabeled. Various real-world machine learning problems fall into this area. The reason is that the data can be expensive or time-consuming to label data as it may need access to domain experts. Whereas unlabeled data is inexpensive and simple to collect and store.

3.2 The Bias-Variance Trade-Off

An algorithm in supervised machine learning is trained with data to develop a model. The purpose of any supervised machine learning algorithm is to best measure the mapping target function (f) for the output variable (Y) given the input data (X). While learning from the data, there will always be an error rate. This prediction error for any machine learning algorithm can be of one of the following three types.

1. Bais error

2. Variance error

3. Ireducable error

3.2.1 Bias Error

Bias is the simplifying assumptions made by a model to make the target function easier to learn. A model with High bias is having on average predicted values far from the real values. This model will be too simple and could not capture the complexity of the data. Usually, parametric machine learning algorithms have a high bias making them fast to learn and easier to understand but usually less flexible. In contrast, they have a lower predictive performance on complex problems that fail to meet the simplifying assumptions of the algorithms bias.

- **Low Bias:** Implies less assumptions about the form of the target function.
- **High-Bias:** Implies more assumptions about the form of the target function

Examples of low-bias machine learning algorithms include k-nearest neighbors, decision trees, and support vector machines. Examples of high-bias machine learning algorithms include linear discriminant analysis, linear regression, and logistic regression.

3.2.2 Variance Error

Variance is the measure that shows how the target function estimation change if the model is trained on different unseen data. It happens when the model performs well on the trained dataset but does not act well on a dataset that it is not seen, like a validation dataset or test dataset. Variance tells us how scattered are the predicted values from the real value.

- **Low Variance:** Proposes small changes to the measure of the target function with changes to the training dataset (e.g., test dataset).

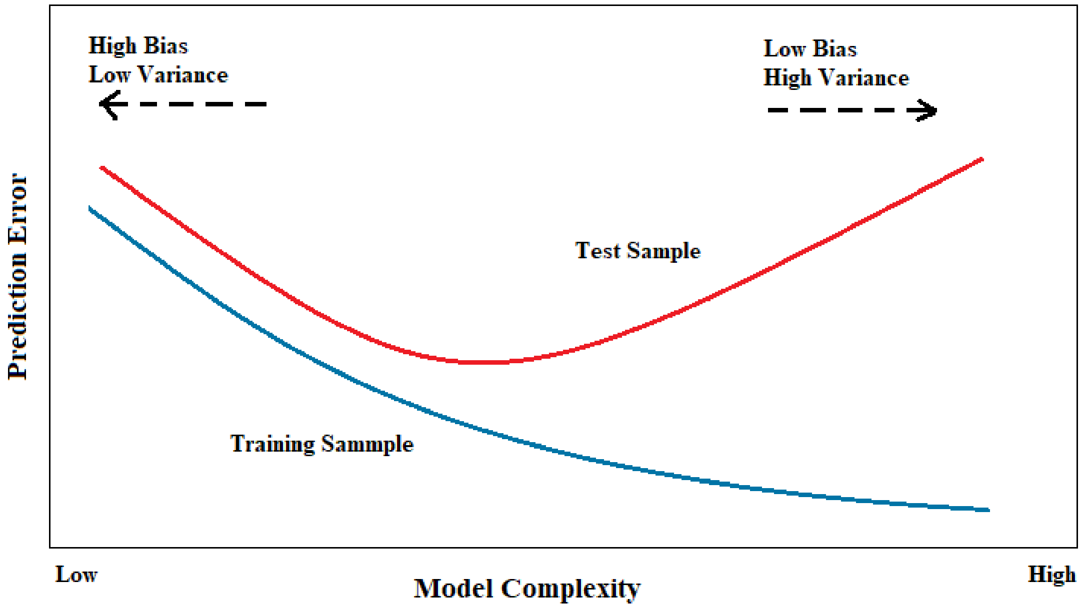


Figure 3.1: Bias and Variance (Ref. [6])

- **High Variance:** Proposes large changes to the measure of the target function with changes to the training dataset (e.g., test dataset).

Commonly nonparametric machine learning algorithms that have much flexibility have a high variance. For instance, decision trees have a high variance, that is even higher if the trees are not pruned well before use. Examples of low-variance machine learning algorithms include logistic regression, linear regression, linear discriminant analysis. Examples of high-variant machine learning algorithms are k-nearest neighbors, decision trees, and support vector machines.
Figure 3.1 shows a model with high bias is straightforward. In contrast, a model with high variance tries to fit most of the data points, which makes the model complex.

3.2.3 Bias-Variance Trade-Off

The purpose of supervised machine learning algorithms are to obtain low bias and low variance, with good performance.

- Parametric (or linear) machine learning algorithms usually have a high bias but a low variance.
- Parametric (or linear) machine learning algorithms usually have a high bias but a low bias but a high variance.

The bias and variance have an inverse relationship. By increasing the bias, variance value will be decreased. machine learning algorithms should be well parameterized to balance bias and variance. Below two examples describe the configuring the bias-variance trade-off for the k-nearest neighbors and super vector machines algorithms.

3.3 Classification Algorithms for EEG-based Applications

3.3.1 Introduction

The EEG signals are used to record brain activities. The EEG signals of a patient's brain are interpreted to detect the brain abnormalities. To achieve this goal, classification machine learning algorithms can be used. In this study, we focus on machine learning algorithms which well plays for seizure detection and classification problems. The following categories of algorithms are explained in this research.

- Linear classification algorithms,

- Non-linear classification algorithms
- Neural Networks

3.3.2 Linear Machine Learning Algorithms

The model for a linear algorithm (also called parametric algorithm) is built with a fixed number of parameters, which are independent of the number of training data. The number of parameters in linear learning model will not change even there is a change in the training data. The parametric algorithms are high in speed, simpler, and less data is required to train the model.

3.3.2.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a parametric algorithm. LDA algorithm uses hyperplanes to split the data describing the different classes [7]. For a binary classification problem, The class of a feature vector is based on the vector placed on which side of the hyperplane, as shown in Figure 3.2. LDA considers a normal distribution of the data, with equal covariance matrix for both classes. The hyperplane is achieved by seeking the projection that maximizes the distance between the two classes 'means' and minimize the interclass 'variance'.

3.3.2.2 Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is a simple yet very effective method to discriminative learning of linear machine learning classifiers under convex loss functions such as (linear) logistic regression and support vector machines [33]. The advantages of SGD include:

- Efficiency
- Ease of implementation (lots of opportunities for code tuning).

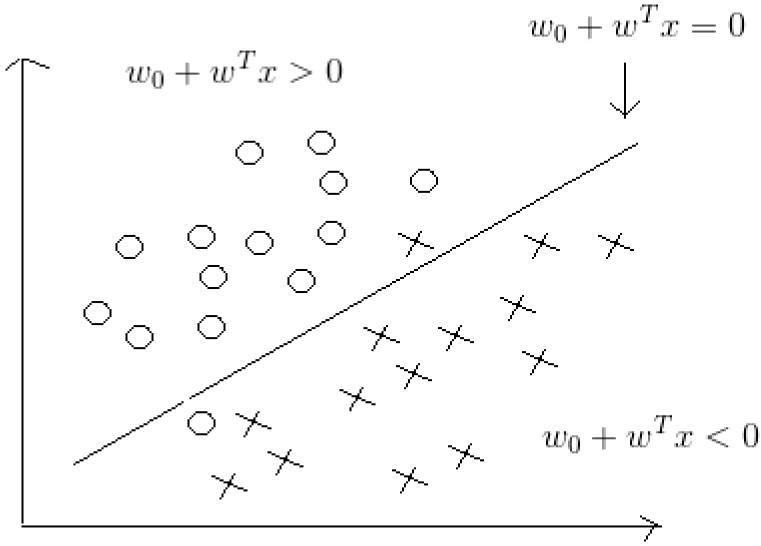


Figure 3.2: A hyperplane which split two classes (Ref. [7])

The disadvantages of SGD are:

- SGD needs a number of hyperparameters such as the the number of iterations and regularization parameter.
- Sensitive to feature scaling.

3.3.3 Non-linear Machine Leaning Algorithms

3.3.3.1 Decision Trees

The DT is a top-down induction machine learning algorithm. The DT algorithm aims to build a tree that has leaves that are homogeneous as possible. Leo Breiman introduces the classification and regression tree (CART) [34]. The significant advantages of the CART algorithm are high classifying speed, strong learning ability, and simple construction [35]. The design structure for the CART model is a binary tree. Each node represents a single input variable and a separation point on that variable. The leaf nodes of the decision tree hold an output

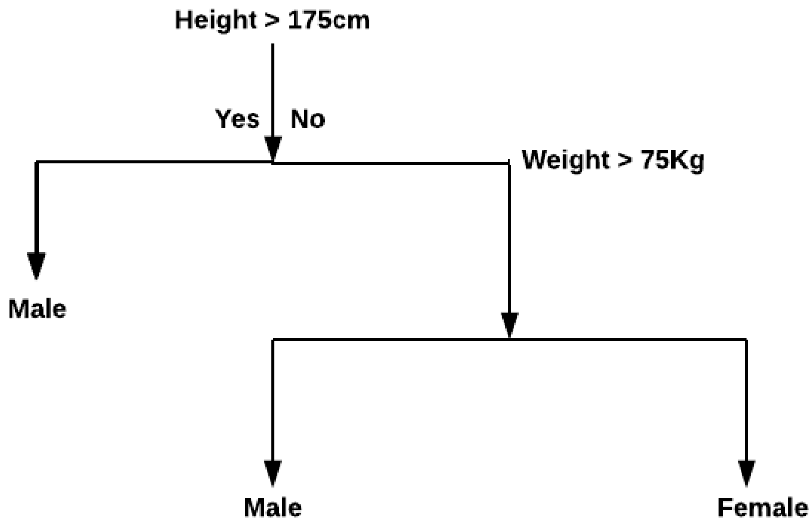


Figure 3.3: Example of a Graph Representation of a Decision Tree

variable used to make a prediction. Figure 3.3 shows a binary decision tree given two inputs variables of height and weight and the output as sex (male or female) in leaf nodes. The decision tree can also be represented as a set of rules; for example, below shown the Figure 3.3 graph as a set of rules.

If Height > 175 cm Then Male
If Height <= 175 cm AND Weight > 75 kg Then Male
If Height <= 175 cm AND Weight <= 75 kg Then Female

Cost Functions for Decision Tree

A decision tree is a process of splitting up the input space in a greedy approach called recursive binary splitting. The recursive binary approach is a numerical procedure where different split points are tried and tested using a cost function for the data. The aim is to minimize the cost function, and the lowest cost value is considered as the best for selection. The process of evaluating all input variables along with all possible split points is performed in a greedy manner, and the best split point is chosen. There are two different cost functions for regression and classification decision trees. For regression decision tree, the minimized cost function is selected to choose split points based on the sum of squared error across all training data that fall within the rectangle.

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.1}$$

Where y is the actual output, and $\hat{y}$ is the predicted output for the sample data. For the classification problem, the Gini index cost function is used.

$$G = \sum_{i=1}^{n} p_k \times (1 - p_k) \tag{3.2}$$

In equation 3.2 G is the Gini index over all classes of data, p_k is the proportion of training samples with class k in the interested rectangle.

3.3.4 Neural Networks

3.3.4.1 Multi Layer Perceptron

A multilayer perceptron is (MLP) is composed of several layers of neurons. The layers are an input layer, one or several hidden layers, and an output layer. The input of each neuron is connected to the output of the previous layer's neurons. The neurons of the output layers describe the classes of the feature vectors [7]. A MLP is shown in Figure 3.4.

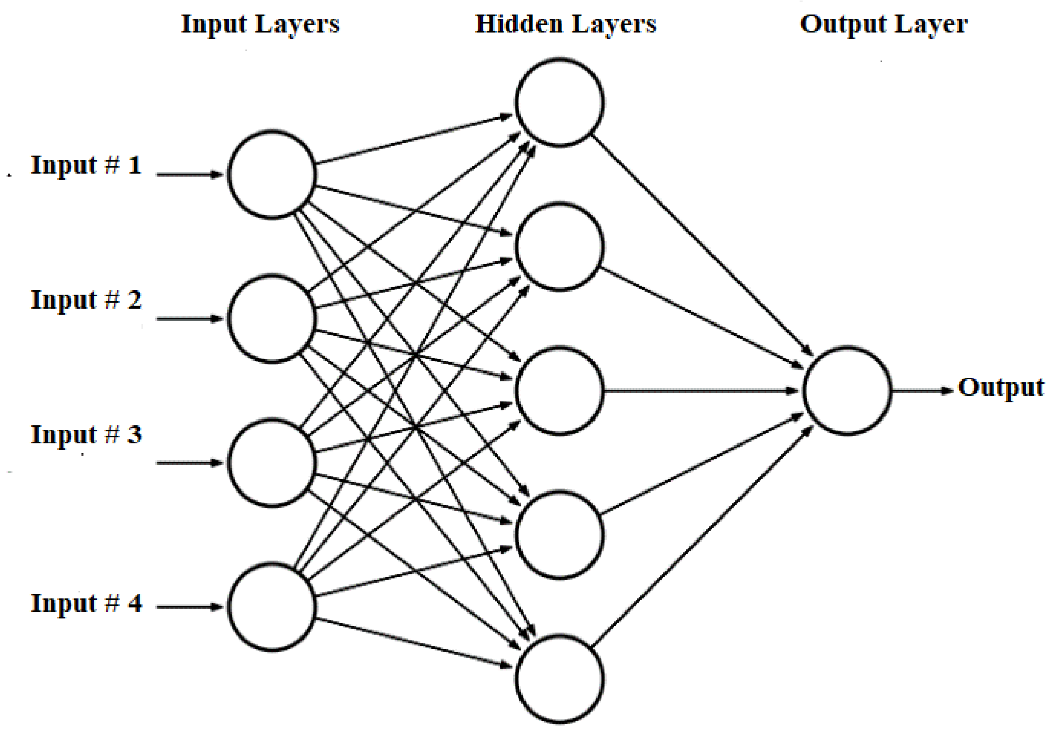


Figure 3.4: Layers of MLP neural network (Ref. [8])

3.3.5 Ensemble Methods

An ensemble is a set of machine learning classifiers defined as base classifiers, where each classifier learn a target function. The individual predictions of classifiers are combined to classify new samples. Ensemble techniques are used to improve the generalization performance of a set of classifiers.

3.3.5.1 General Ensemble Procedure

Let S indicate the original training data, N indicate the number of base classifiers, and T be the test data.

Algorithm 1 General Ensemble Procedure

for *i = 1 to N* **do**
  Create training set S_i from S
  Build a base classifier C_i from S
end
for *each test record x in T* **do**
  $C^*(x) = Vote(C_1, C_2, ..., C_N(x))$
end

3.3.5.2 Advantages of Ensemble Methods

1. **Statistical Aspect:** A set of machine learning classifiers with the same training performances may have different generalization performances. Therefore, the combined outputs of several classifiers lessen the risk of selecting a weak performing classifier.

2. **Volume of Data:** A single classifier may not able to handle large amounts of data. The ensemble methods could help to train several classifiers on different partitions of the data. Additionally, They can be used as resampling techniques on small amounts of data.

3. **Divide-and-conquer:** Divide-and-conquer-based methods break a complex task into many simple tasks and solve them collectively [36]. The ensemble methods use a hierarchy of divide-and-conquer-based approaches to develop an optimized framework for a set of classifiers.

3.3.5.3 Error Rate in Ensemble Methods

The base classifiers are independent, and their errors are uncorrelated. The error rate of the ensemble classifier is measured as follows.

$$E_{ensemble} = \sum_{i=1}^{N} (i^N) \times e^i (1-e)^{N-i} \tag{3.3}$$

Where e is the error rate of individual classifiers.
Consider an ensemble of 25 binary classifiers, each having an error rate of 0.35. Then, the error rate of the ensemble classifier is 0.06.

$$E_{ensemble} = \sum_{i=1}^{25} (i^{25}) \times 0.35^i (1-e)^{25-i} = 0.06$$

Which is considerably lower than the base classifiers.

3.3.5.4 Types of Ensemble Methods

Two families of ensemble methods are normally recognized:

1. **Average Ensemble Methods:** In averaging techniques, several classifiers are built independently, and then the average of their predictions is estimated. The combined classifier (or ensemble classifier) is normally better than any of the single base classifier because its variance is reduced. The most common are Bagging methods, and Forests of randomized trees.

2. **boosting Ensemble Methods:** Base classifiers are built sequentially, and

each classifier tries to reduce the bias of the combined classifier. The motivation is to combine multiple, poorly performed classifiers to produce a powerful ensemble. The types can be AdaBoost, and Gradient Tree Boosting.

In this study only Random forest trees, Extremely Randomized Decision Trees, and Gradient Boosting are investigated.

3.3.5.5 Random Forests

Random decision forests (RF) are an ensemble learning technique for classification, regression by constructing a number of decision trees at training time and taking the mode of the classes for the classification and mean prediction of the individual trees for the regression[37]. RF solved the overfitting drawback of the decision tree on the training data[38]. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them [39].

3.3.5.6 Extremely Randomized Decision Trees

The building of Extremely randomized decision trees (Extra-trees) differs from the standard decision trees. When searching for the best split to separate the samples of a node into two groups, splits are drawn randomly for the selected features, and the best split among them is chosen. IT has two main differences with other tree-based ensemble techniques that it divides nodes by choosing cut-points fully at random without using bagging and that it uses the whole learning sample (rather than a bootstrap replica) to grow the trees [40].

3.3.5.7 Gradient Boosting

Gradient boosting is a machine learning method for regression and classification problems. It builds a prediction model in the form of an ensemble of weak prediction models, usually decision trees. Gradient boosting ensemble of trees builds highly robust, competitive, interpretable procedures for both regression and classification problems [41].

Chapter 4

Literature Review

4.1 A review of previous works

In this section, a technical review of epileptic seizure detection and prediction systems are studied. The review is categorized into three sections as International review and Other datasets, national review (in India) and related dataset (CHB-MIT Scalp EEG), the dataset used in this study.

4.1.1 International Review

In [42] described a semi-automatic method for temporal lobe seizures lateralization using raw scalp EEG signals. They have used two Hjorth parameters to estimate the quadratic mean and dominant frequency of signals on each side of the brain for lateralization of the seizure onset and mean values are segmented. The means of segmented variables were used to characterize the seizure by a point in a (frequency, amplitude) plane, and six criteria were proposed for the partitioning of this plane for lateralization. The procedure was applied to 45 patients (85 seizures) in which the two most powerful criteria (Cartesian C3r and polar C5r) achieved 96% and 97% correct lateralization. The method produced satisfactory results, easy to interpret, the setting of procedure parameters was simple, and the approach was robust to artifacts.

In [43] a Computer Aided Diagnostic (CAD) based automatic characterization of EEG signals into normal, interictal, and ictal is proposed, in which fewer number of features are extracted with an acceptable computational cost. In this work hundred samples, each in normal, interictal, and ictal categories are decomposed into wavelet coefficients using Wavelet Packet Decomposition (WPD). Then eigenValues are extracted from the resultant wavelet coefficients using Principle Component Analysis (PCA). Significant eigenValues, selected using ANOVA test, are used to train and test several supervised classifiers using the 10-fold stratified cross-validation technique, in which the Decision Tree (DT) 93.3%, Fuzzy classifier (Fuzzy) 96.7%, , K-Nearest Neighbor (KNN) 93.3%, Naive Bayes Classifier (NBC) 82%, Radial Basis Probabilistic Neural Network (RBPNN) 92.7%, and 99% classification accuracy is obtained using Gaussian Mixture Model (GMM) classifier which is reported the highest accuracy. The EEG dataset used in this work was taken from the artifact free EEG time series data available at the University of Bonn.

In [44] a Cloud-based Deep Learning of Big EEG Data for Epileptic Seizure Prediction system is proposed. Feature extraction from the EEG signals because of the non-stationary nature of signals, normal and seizure pattern very across different patients is a challenging task. Thus, a group of manual feature extraction is not practical. Moreover, when using implanted electrodes for brain recording, massive amounts of data are produced. To address these challenges, a cloud-based BCI system for the analysis of EEG is presented. First, a dimension reduction using Principal and Independent Component Analysis to increase the classification accuracy as well as to decrease the energy, computation time, and the communication bandwidth techniques are used. Second, a stacked auto-encoder as a deep-learning structure to analyze EEG signals for the epileptic seizure prediction is trained in two steps. The developed deep-learning methods have the capability for unsupervised feature extraction and, therefore, represent a suitable substitute to manual feature-extraction techniques for classification purposes. These methods extract high-level, complex abstractions for data representations through a hierarchical learning process prediction system. Third, the BCI system is implemented in the cloud as safe storage with high computational resources for big EEG data generated by implanted electrodes. To study the accuracy and performance, the system is evaluated and compared to other methods on a benchmark clinical epilepsy dataset in which the proposed method resultant accuracy is 94% compared to Random Forest 75%, Linear SVM 71%, Non-linear SVM 78%, MLP-Neural Network 68%. The key benefit of the proposed method centers upon the analysis and learning allowed from massive amounts of unsupervised data, making it a practical method for developing a patient-based seizure. A cloud-based deep-learning method that is able to perform seizure prediction under such circumstances has immediate applicability in the present day.

In [45] optimized deep learning, big data and seizure prediction and localization from EEG and Electrocorticography (ECoG), or intracranial electroencephalography (iEEG) data via Internet of Things are proposed which is an extended method of previous work [44]. By leveraging the potential of cloud computing and deep learning, they developed and deployed BCI seizure prediction and localization from scalp EEG and ECoG big data. First, a new method for epileptic seizure prediction and localization of the seizure focus is presented. Second, an extended optimization approach on existing deep-learning structures, Stacked Autoencoder and Convolutional Neural Network (CNN), is proposed based on principal component analysis (PCA), independent component analysis (ICA), and Differential Search Algorithm (DSA). Third, a cloud-computing solution (i.e., Internet of Things (IoT)), is developed to define the proposed structures for real-time processing, automatic computing, and storage of big data. To study the accuracy and performance, the system is evaluated and compared to other methods on a benchmark clinical epilepsy dataset in which the proposed (Optimized CNN) method resultant accuracy is 96% compared to Opt. Stacked Autoencoder 94%, Random Forest 75%, Linear SVM 71%, Non-linear SVM 78%, and MLP-Neural Network 68%.

In [46] classification of epilepsy EEG signals using wavelet-based envelope analysis (EA), and neural network ensemble (NNE) for detecting normal, interictal, and epileptic signals is proposed. The discrete wavelet transform (DWT) in combination with the EA method, is developed to extract significant features from the EEG signals. Moreover, a powerful network model called NNE is explicitly designed to the task of epilepsy detection. To evaluate the performance of the presented algorithm effectively, different classifiers and feature extracting techniques are considered in this work. The experimental results have shown that the introduced scheme achieved satisfying recognition accuracy of 98.78% compare to DWT-based EA features Linear Discriminant Analysis (LDA) 92.67% and entropies features Fuzzy algorithm 98.10% respectively.

In [47], EEG signal processing for epilepsy and autism spectrum disorder diagnosis is studied. In this work, different EEG feature extraction and classification techniques for assisting epilepsy and autism spectrum disorder (ASD) diagnosis are presented. The University of Bonn dataset, the MIT dataset, the King Abdulaziz University dataset, and their own EEG recordings (46 subjects) are used to evaluate the proposed methods, with different numbers of EEG channels and electrode positions, which increases the problem complexity. First, the EEG signal is preprocessed to remove major artifacts before being decomposed into several EEG sub-bands using a discrete-wavelet-transform (DWT). Two nonlinear methods were studied, namely, Shannon entropy and largest Lyapunov exponent, which measure complexity and chaoticity in the EEG recording, in addition to the two conventional methods (namely, standard deviation and band power). Then cross-validation approach is used to measure synchronization between EEG channels, which may reveal an abnormality in communication between brain regions. The extracted features are then classified using several classification (namely, ANN, SVM, KNN, and LDA) methods in which the combination of DWT, Shannon entropy, and k-nearest neighbor techniques achieved the best results, with an overall accuracy of 94.6% on the three-class (multi-channel) classification problem (normal vs. epilepsy. vs. autism).

In [48], an automated system for epilepsy detection using EEG brain signals based on deep learning approach is described. They proposed a system that is an ensemble of pyramidal one-dimensional convolutional neural network (P-1D-CNN) model in which the model learns the internal structure of data and outperforms hand-engineering techniques. The P-1D-CNN works on the concept of refinement approach, and it involves 61% fewer parameters compared to standard CNN models with better generalization. Further to overcome the limitations of the small amount of data, they proposed two augmentation schemes in which the data is augmented by splitting the given full-length EEG signals into small signals using a fixed size window; each small signal is used as an independent instance for learning CNN models. Then the system is tested on the University of Bonn dataset, a benchmark dataset; in almost all the cases concerning epilepsy detection, it gives an accuracy of 99.1 0.9% and outperforms the state-of-the-art systems. Besides, to strength, the P-1D-CNN model requires 61% less memory space, and its detection time is concise (¡ 0.000481 s), as such, it is suitable for the real-time clinical setting.

In [49], a deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals is proposed. Traditionally, neurologists employ direct visual inspection to identify epileptic form abnormalities, which can be time-consuming. Therefore, in this work, a computer-aided diagnosis (CAD) system in which 13-layer deep convolutional neural network (CNN) is implemented to detect normal, pre-ictal, and seizure classes in EEG signals. In this approach, each EEG signal is normalized with Z-score normalization, zero mean and standard deviation of 1 before feeding into the 1-D deep convolutional network (CNN) for training and testing using 10-fold stratified cross-validation technique. The sampling rate of the EEG signal is set at 173.61 Hz. The proposed technique achieved an accuracy, specificity, and sensitivity of 88.67%, 90.00%, and 95.00%, respectively at Bonn University, Germany database. The performance (accuracy, sensitivity, and specificity) of the proposed model is slightly lower than some of the works (99% on Average other latest works). The advantage of the model presented in this paper; however, separate steps of features extraction and features selection are not required. Nevertheless, the main drawback of this work is the lack of vast EEG database. The proposed algorithm requires a diversity of data to obtain optimum performance. The performance of this technique can be improved by applying a bagging algorithm and increasing the number of samples.

In [50] research work, a new modified particle swarm optimization (PSO) algorithm has been proposed to train the radial basis function neural network

(RBFNN) more efficiently to classify the epileptic seizures. Also, this technique is examined on a different dataset, e.g., Eye state prediction. This technique is also compared with a few other available techniques such as Gradient descent, convention PSO) by rigorous and thorough practical implementations, and experimental results. Thus, it is proved that the proposed technique outperformed the other existing techniques. In this research work, DWT is only utilized for analysis and statistical feature extraction from EEG datasets for epilepsy. For eye state prediction, the dataset was already in the format for classification. First, the EEG signals are preprocessed with the discrete wavelet transform (DWT). Then to classify the EEG signal, they used a radial basis function neural network (RBFNN). The network is trained to optimize the mean square error (MSE) by using a modified particle swarm optimization (PSO) algorithm. The key idea behind the modification of PSO is to introduce a method to overcome the problem of slow searching in and around the global optimum solution. The effectiveness of this procedure is verified by an experimental analysis on a benchmark dataset which is publicly available. The result of this experimental analysis revealed that the improvement in the algorithm is significant with respect to RBF trained by gradient descent and canonical PSO. Here, two classes of EEG signals were considered: the first being an epileptic and the other being non-epileptic. The proposed method produced a maximum accuracy of 99% as compared to the other techniques. In this research study, two different types of datasets are used. One of them is an EEG dataset for epileptic seizure identification, and the other one is an EEG dataset for Eye state prediction.

In [51], an efficient procedure that provides an accurate classification of EEG signals for early detection of epileptic seizures is presented. Many tools such as artificial neural networks (ANNs), gradient-based algorithms, genetic algorithms (Gas), feature extraction with discrete wavelet transform (DWT), and fuzzy relations for dimension reduction are used. In this research work, ANNs are trained by the gradient-based algorithms and GAs considering early stopping, cross-validation and information criteria. In order to ensure an accurate classification performance, the automated multi-resolution signal processing technique splits EEG signals into the detailed partitions with different bandwidths and then decomposes them into detail and approximation coefficients through DWT at the different decomposition levels. Thus, some specific latent features that characterize the nonlinear and dynamical structures of EEG signals are acquired from these coefficients. Then fuzzy relations are used to reduce the dimensions of the feature matrix and to detect the epileptic behaviors in EEG signals, these selected features are processed by ANNs based cross-entropy and information criteria. The method produced an accuracy of 99.5% on average.

In [52], a novel approach for automatic epilepsy seizure detection based on EEG analysis that exploits the underlying non-linear nature of EGG data. In this research, two main contributions are presented and validated. As the first contribution, one of the proposed systems consists of a non-linear implementation of a Support Vector Machine (SVM) classifier that makes use of well-known techniques and non-linear kernels to optimize the feature extraction and learned classification model. The second contribution is inspired by a successful model for Natural Language Understanding and Computer Vision, known as Bag-of-Words (BoW), which is adapted to the field and added into the framework for extracting a non-linear feature representation of the input data in an unsupervised manner. The use of public datasets caters for comparison purposes whereas the private one shows the performance of the system under realistic circumstances of noise, artifacts, and signals of different amplitudes. Moreover, the proposed solution has been compared to state-of-the-art works not only for pre-processed and public datasets but also with the private datasets. The mean F1-measure (85.59%) shows a 10% improvement over the second-best (DWT+ANN Tzallas et al. (2007) 76.13%) ranked method, including cross-dataset experiments. The obtained results prove the robustness of the proposed solution to more realistic and variable conditions.

In [53] article, they applied convolutional neural networks to different intracranial and scalp electroencephalogram (EEG) datasets and proposed a generalized retrospective and patient-specific seizure prediction method. In this work, they used the short-time Fourier transform on 30-s EEG windows to extract information in both the frequency domain and the time domain. The algorithm automatically generates optimized features for each patient to best classify preictal and interictal segments. The method can be applied to any other patient from any dataset without the need for manual feature extraction. The proposed approach achieves a sensitivity of 81.4%, 81.2%, and 75% and a false prediction rate of 0.06/h, 0.16/h, and 0.21/h on the Freiburg Hospital intracranial EEG dataset, the Boston Childrens HospitalMIT scalp EEG dataset, and the American Epilepsy Society Seizure Prediction Challenge dataset, respectively. The prediction method is also statistically better than an unspecific random predictor for most of the patients in all three datasets.

In [54] parallel processing of massive EEG data with MapReduce algorithm is studied. To increase efficiency and performance, this research developed parallel ensemble empirical mode decomposition (EEMD) neural signal processing with MapReduce. They implemented the parallel EEMD with Hadoop in an advanced cyberinfrastructure. Test results and performance evaluation show that parallel

EEMD can significantly improve the performance of neural signal processing.

In [55], a random ensemble learning for EEG seizure detection and classification is proposed. It is developed in a multi-tier distributed computing infrastructure and a semantic linked data superstructure. The proposed work enables two operation scenarios. First, big data analysis using a cloud computing paradigm. Second, interactive and adaptive prediction using real-time brain state and relevant data sets for training and refining brain state prediction. In this work, EEG signals are pre-processed from noise and artifacts using wavelet and independent component analysis in which time/frequency-based features are extracted, and an automatic method infinite component analysis (I-ICA) is used to select independent features. The feature space is divided into sub-spaces via random selection, and combinations of classifiers (SVM, multilayer perceptron (MLP) neural network and extended k-nearest neighbors (k-NN), called extended nearest neighbor (ENN)) are used on each subspace to classify the input data as non-epileptic or epileptic. Using majority voting (MV), the output with the highest number of votes is chosen. To evaluate the solution, a benchmark ECoG of eight patients with temporal and extra-temporal epilepsy is implemented in a distributed computing framework as a multi-tier cloud-computing architecture. Using leave-one-out cross-validation, the accuracy, sensitivity, specificity, and both false positive and false negative ratios of the proposed method are found to be 0.97, 0.98, 0.96, 0.04, and 0.02, respectively. In this work, the clinical ECoG dataset of eight epileptic patients, including 104interictal (normal) segments and 104 ictal (seizure) developed by the University of Pennsylvania and the Mayo Clinic, and also sponsored by the American Epilepsy Society is used.

Table 4.1: National level review of previous studies

Author	Place	Methodology	Performance Parameters
Varsha Harpale et al. [56]	Pune University	Statistical, Wavelet Transform and Fuzzy classifier	Classification Accuracy: 96.48% TPR: 96.52%, FPR: 0.352 Pre-seizure Accuracy: 96.02%
Sandeep Kumar et al. [50]	S'O'A University, Bhubaneswar, Fakir Mohan University, KIIT University, Odisha	Particle Swarm Optimization (PSO), Discrete Wavelet Transform (DWT), Radial Basis Function Neural Network (RBF)	Classification Accuracy: 99%
Debdeep Sikdara et al. [57]	Indian Institute of Technology (IIT) Kharagpur	wavelet and Support Vector Machine (SVM)	Classification Accuracy: 99.6%
Kavya Devarajan et al. [58]	Anna University	wavelet-based functional mixed-effect model	Detection Accuracy: 87.5%.
P.Ramina et al. [59]	Bharathidasan University, Tamilnadu	Artificial Neural Network	Prediction Accuracy: 99.9%,
Raj Sadaye et al. [60]	Dwarkadas J. Sanghvi College of Engineering Mumbai	Power spectrum and Support Vector machine (SVM)	Prediction Accuracy: 70%
Sivakumaran N et al. [61]	National Institute of Technology, Amrita Institute of Medical Sciences, Chennai Medical College Hospital and Research Centre.	statistical features, discrete Fourier Transforms, Wavelet Transforms and Hilbert transforms with ICA and machine learning algorithms (ANN, Clustering, pattern recognition, Knowledge-based rules, Data mining and other classification)	Detection Accuracy: 100%
A. Sivasangari et al. [62]	Sathyabama University, Chennai	continuous wavelet transform (CWT) for feature extraction, Fuzzy entropy, and SVM	Prediction Accuracy: 99.19%

4.1.2 National Level Review

The table 4.1 provides a review of national level studies in India on seizure epileptic detection and prediction systems.

4.1.3 Related Dataset Reviews

In this section, papers related to CHB-MIT Scalp EEG Database are reviewed. In [56] an adaptive method for feature selection and extraction for the classification of pre-ictal and seizure state of EEG signals using time and frequency domain features is described. In the proposed work, from raw EEG data time-domain statistical and pattern adapted wavelet transform based frequency-domain features are extracted. The artifacts are removed from the EEG signal using independent component analysis (ICA). The final feature vector is generated using ANOVA feature selection method and then fuzzy classifier is applied. The results presented are with 96.48% of seizure classification accuracy, 96.52% True Positive

Rate (TPR) and 0.352 of False Positive Rate (FPR) of the system for seizure detection and 96.02% accuracy for identifying the pre-seizure state. The Pre-seizure state is identified at 13110 s before the actual onset seizure occurs. The classification accuracy can be further improved with advanced machine learning techniques. Due to variation in seizure characteristic training, a system is a challenging task. Further, the system performances can be improved by selecting precise features, advanced and most suitable classifier and selection of high-performance non-specific algorithm. Prediction of seizures well before the time is still a task can take up for further research.

In [63] paper a patient-specific epileptic seizure detection method based on Common Spatial Pattern (CSP) is described. The features are extracted from EEG signals through projection on a CSP projection matrix. Then the extracted features are trained on SVM classifier. This approach has obtained 100% an average sensitivity, 1.17 an average false alarm, and 7.02 seconds an average detection latency time.

In [64], which is the next paper of [63] published by the same author. In this paper, the author applied the same feature extraction method on 23 patients of MIT-Scalp EEG dataset. Then the linear discriminant analysis (LDA) classifier is used. This approach has achieved an average sensitivity of 0.89, an average false prediction rate of 0.39, and an average prediction time of 68.71 minutes using a 120-minute prediction horizon.

In [65] patient-specific method is presented based on k-means unsupervised algorithm to cluster the recording into two distinct clusters of seizure and non-seizure data. The algorithm achieved 91.43% accuracy.

In [66] a patient non-specific strategy for seizure detection based on Stationary Wavelet Transform (SWT) is reported. The linear discriminant analysis (LDA) and neural network (NN) classifiers are trained on the extracted features from SWT. For patient-specific, the LDA and NN classifiers obtained on average sensitivity of 92.6% and 79.9% respectively. For patient non-specific on average sensitivity of 87.5%, and specificity of 99.9%.

In [67], an automatic seizure detection system is described. Features are extracted using local binary pattern (LBP) operator. The K-nearest neighbor classifier is applied on extracted features from LBP algorithm. The classifier obtained a mean accuracy of 99.7% and a mean specificity of 99.8% with an average false detection rate of 0.47/h and sensitivity of 99.2% respectively.

In [68] work a machine learning system for automated whole-brain seizure detection method is presented. In this study, the raw EEG data is filtered using second-order Butterworth filters. Then statistical features are extracted, and several classifiers are used to detect seizure from data. The K-NN obtained sensitivity of 88%, specificity of 88% and 93% for the area under the curve, which is higher compared to other used classifiers.

Chapter 5

Materials and Methodology

5.1 Introduction

In order to detect seizure, it involves the interpretation of long EEG records by the expert physicians, which is time-consuming and need high human efforts. Thus, an automatic seizure detection system is required to reduce the volume of data for the physicians. It will help experts only to study those parts of the EEG data that is seizure effected. Some studies have conducted on EEG signals classification into normal and seizure states. The majority of previous studies on seizure detection and prediction have concentrated on patient-specific predictors, where a classifier is trained on one person and tested on the same person [18; 19; 20; 21; 22]. The aim of this study is to categorize EEG records in significant states of normal and seizure as shown in Figure 5.1.

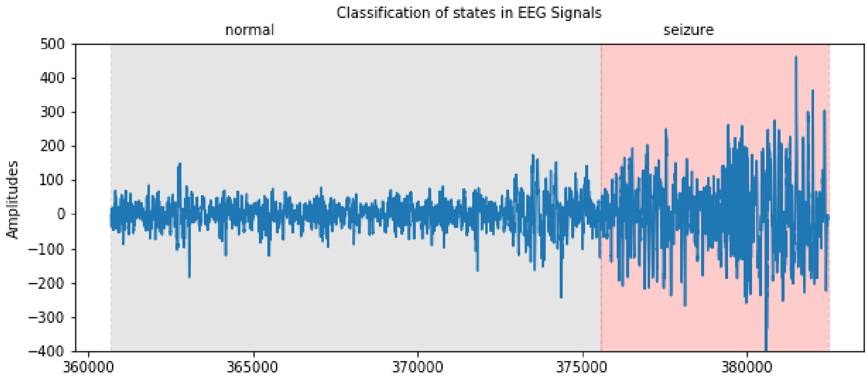


Figure 5.1: EEG signals with sampling Frequency of 256

5.2 Materials and Methodology

5.2.1 Materials

The CHB-MIT EEG scalp dataset from the Childrens Hospital Boston [69] consists of EEG recordings is used in this work. The dataset, grouped into 23 cases, from 22 pediatric subjects (5 males, ages 3-22; and 17 females, ages 1.5-19) recorded at various time in 654 .edf files with intractable seizure. The files are having from one to four hours recordings of digitized EEG signals with sampling rate of 256 and 16-bit resolution. The international 10-20 system of EEG electrode positions and nomenclature with 23 channels is used for these recordings. For the experiment, 15 subjects data and only the files seizure records are used as shown in table 5.1.

Table 5.1: CHB-MIT EEG used patients

Patient	Gender	Age (years)	Total Seizures Files	Total used files
chb01	Female	11	7	6 seizure files
chb02	Male	11	3	3 seizure files
chb03	Female	14	7	7 seizure files
chb04	Male	22	4	3 seizure files
chb05	Female	7	5	5 seizure files
chb07	Female	14.5	3	3 seizure files
chb08	Male	3.5	5	5 seizure files
chb09	Female	10	3	1 seizure files
chb10	Male	3	7	7 seizure files
chb11	Female	12	3	3 seizure files
chb15	Male	16	14	9 seizure files
chb16	Female	7	6	6 seizure files
chb17	Female	12	3	3 seizure files
chb19	Female	19	3	2 seizure files
chb22	Female	9	3	3 seizure files

5.2.2 Feature Extractions

1. Time Domain Features

Regularity (periodicity), repeatability, amplitude variation, and synchronicity are significant time domain features, which differentiate epileptic seizure from the normal EEG signal [56]. The following statistical features are extracted from the time domain EEG signals.

Maximum Amplitude

Maximum amplitude value is defined as

$$max = \max_{t=0}^{N-1} x(t) \tag{5.1}$$

Mean

Amplitude variation is analyzed by the mean power of the EEG signal. It is calculated as the average value of the signal, mathematically represented as

$$a = \frac{\sum x}{n} \tag{5.2}$$

where 'a' is mean of the signal, x is the value of data at a particular instant and n number of data samples.

Standard Deviation

It is the measure of dispersion of signal at a particular instant from the mean of the signal. Standard Deviation is defined as

$$s = \sqrt{\frac{\sum(x-a)^2}{n-1}} \tag{5.3}$$

where 's' is the standard deviation of the signal at a particular instant of time.

Coefficient of Variation (COV)

COV is the measure of variability and regularity of the EEG signal. It is the ratio of the standard deviation to the mean of the signal. The equation is defined as

$$v = \frac{s}{a} \tag{5.4}$$

where 'v' is the Coefficient of Variation of the signal at a particular instant of time.

Root Mean Square (RMS)

RMS of a signal x(t) is computed as the square root of the average of the squared value of the signal, mathematically described as

$$rms = \sqrt{\frac{1}{N}\sum_{t=0}^{N-1} x(t)^2} \tag{5.5}$$

Skewness

It is a measure of the asymmetrical spread of a signal about its mean value. mathematically represented as

$$sk = \sqrt{\frac{\sum(a-x)^3}{(n-1)s^3}} \tag{5.6}$$

Kurtosis

Kurtosis is a measure of the distribution of signal with its peak value such as higher values indicates a higher, sharper peak; lower values indicate a lower, less distinct peak. It is the fourth-moment calculation of a signal and described as

$$k = \frac{E(x-a)^4}{s^4} \tag{5.7}$$

2. Frequency Domain Features

Spectral analysis of EEG signals using Welch method

Fast Fourier Transform (FFT) is a fast algorithm for calculating discrete Fourier transforms. In this study, power spectral density (PSD) is computed using Welch's FFT method. Welch's based FFT is an effective non-parametric signal processing method in frequency domain. The advantages of this method are the enhanced speed with a reduction in computations time and storage over all other available methods in real-time applications [9; 70]. PSD estimation based on the Welch method algorithm is described as follows,

- The input signal $X(t)$ is divided into N overlapping segments.
- The specified window is applied to each segment.
- Discrete Fourier transforms based on FFT is used to the windowed data.
- Each periodogram of new windowed data segment is estimated which is modified periodogram.
- Taking the average of periodograms to obtain spectral density

The mathematical procedure for computing Welch's method is as follows. Let

$$X_K(t) = X(t + (K-1)D) \tag{5.8}$$

Where, X(t) denotes the data segments with the starting point of these segments D of length L and K number of such segments; $X_1(t), ..., X_K(t)$, and that they cover the entire record, i.e. $X_K(t) = X(t + (K-1)D)$. This segmenting is illustrated in Figure 5.2. The method of estimation is as follows. For each segment of length L we calculate a modified periodogram. That is, we select a data window $W(t), t = 0, ..., L-1$, and form the sequence.

We then take the finite Fourier Transforms of these sequence. Here

$$A_K(n) = \frac{1}{L} \sum_{L-1}^{t=0} X_K(t) W(t) e^{-2kijn/L} \tag{5.9}$$

Where, $i = (-1)^{1/2}$. Finally, we obtain the K modified periodograms

$$I_k(f_n) = \frac{L}{U} \left| A_k(n) \right|^2, \quad k = 1, 2, ..., K \tag{5.10}$$

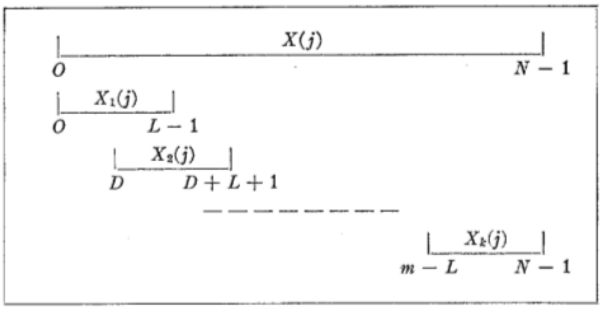


Figure 5.2: An Illustration of EEG records segmentation (Ref. [9])

Where

$$f_n = \frac{n}{L}, \quad n = 0, 1, ..., L/2 \tag{5.11}$$

and

$$U = \frac{1}{L}\sum_{t=0}^{L-1} W^2(t) \tag{5.12}$$

The spectral estimate is the average of these periodograms

$$P(f_n) = \frac{1}{K}\sum_{k=1}^{K} I_k(f_n) \tag{5.13}$$

Hence, we have a spectral estimator $P(f_n)$ with a resultant spectral window whose area is unity and whose width is of the order of $1/L$

5.2.3 Data Preprocessing

Normalization

Normalization is a pre-processing scaling technique to convert continues data into discrete and find new range from an existing range. [71]. It will help for better classification of the problem[72]. In this work, Min-Mix a linear normalization technique is used and the data is normalized or ranged between 0 and 1. The method is mathematically described as follows.

$$Normalized(X) = \frac{X - min\ value\ of\ X}{max\ value\ of\ X - min\ value\ of\ X} \tag{5.14}$$

Feature Selection

Feature selection is also known as variable selection used to reduce the dimensionality of the data and to keep the number of features as low as possible. It is an effective method to decrease computational cost, response time, and enhance the classification accuracy of the algorithms [73; 74]. In this work, Analysis of variance (ANOVA) hypothesis test is used to select relevant features. ANOVA testing is a suitable method for non-stationary data such as EEG, fMRI, and MEG [56; 75]. ANOVA hypothesis test signifies about the difference in multiple groups of data based on mean and variance. It is a statistical test on variance used to compare the differences between two or more means of the samples in the experiment [76].

Data Sampling

Data sampling methods are used to balance imbalanced data. In this study, two sampling approaches are applied.

1. Random Under-sampling (RUS): The classes of data which has higher proportion would be deleted randomly until the proportion of each class leads to balance. This approach decreases the majority class by randomly removing data points from the majority class [77].

2. Synthetic Minority Over-sampling Technique (SMOTE): SMOTE is an over-sampling method in which the minority class data points are increased [78].

This method randomly selects samples from the minority class and creates synthetic new examples, which are between selected and adjacent samples.

5.2.4 Classification

In this study, powerful algorithms explained in previous chapter are adopted for classification purpose. These include the decision tree classifier (DTC), extra-decision tree classifier (EDTC), linear discriminant classifier(LDC), quadratic discriminant classifier(QDC), random forest classifier (RFC), Gradient Boosting classifier (GBC), and Multi-layer Perceptron Classifier (MLPC), and Stochastic Gradient Descent classifier (SGDC).

5.2.5 Validation methods

In order to determine the overall accuracy of each of the classifiers Stratified K-Fold cross-validation technique is used on 70% of randomly selected observations to train the algorithms and 30% of randomly selected test cases to test the algorithms. For model selection, the Grid-Search CV method is used to tune over specified parameters values for the classification algorithms.

5.3 Experiments and Results

5.3.1 Research Process

The procedure of this study is described in Figure 5.3. It includes the stages of signal processing, data pre-processing and classifier construction.

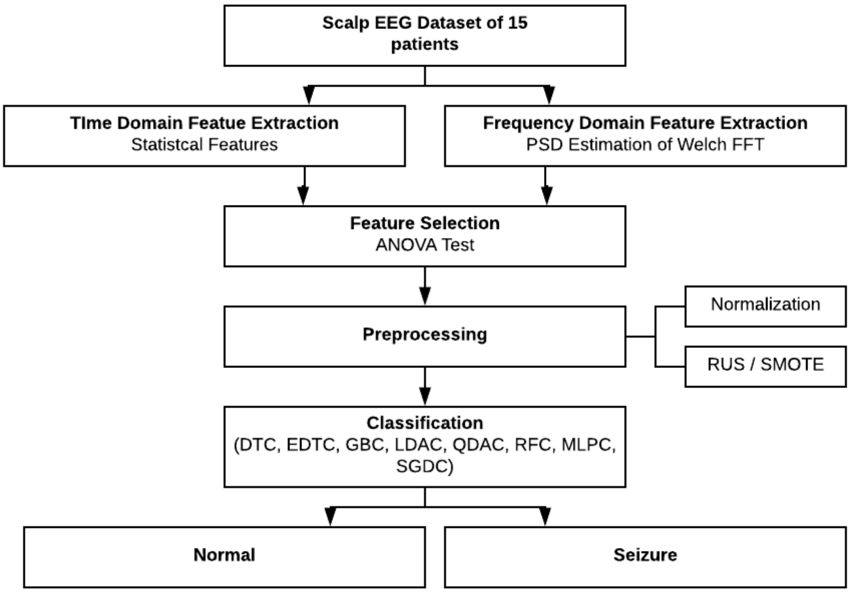
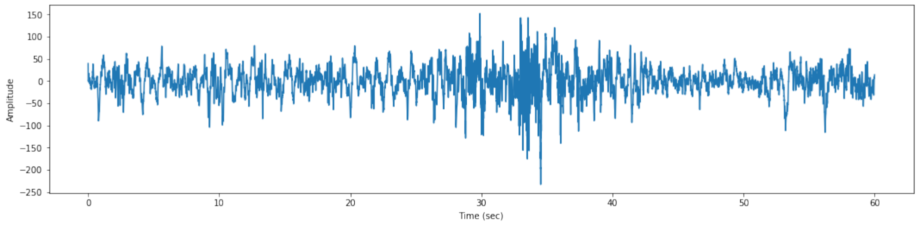


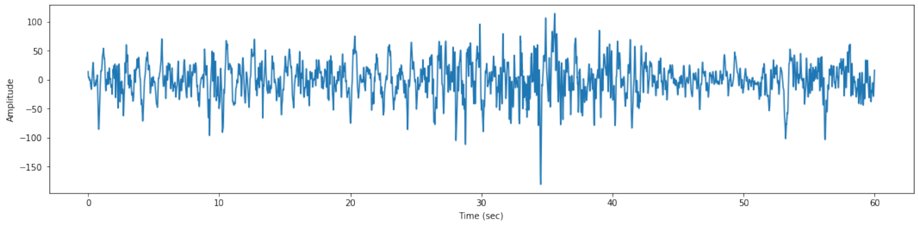
Figure 5.3: Research Process

5.3.2 Signal Processing and Feature Extraction

In this study, 6 channels (T8-P8, F3-C3, FP2-F8, F7-T7, P8-O2, T7-P7) are extracted for the experiment from multi-channel EEG data (23 channels in this case) from 15 subjects as it holds the most regular epileptic seizure activities, and has less noise compared to other channels [69; 79]. After extraction of the T8-P8, F3-C3, FP2-F8, F7-T7, P8-O2, and T7-P7 channels, the Savitzky-Golay filter is applied on each second of the EEG data to remove noise from the EEG signals due to the sudden eye movement or muscle tightening. Figure 5.4 presents an original 60 seconds EEG signal versus the signal form processed through the Savitzky-Golay filter. Additionally, only the first 60 seconds of ictal data is used from each seizure record that lasts longer than 60 seconds due to the outliers containing in the records.



Before Filtering



After Filtering

Figure 5.4: The signal filtering charts

5.3.3 Feature Extraction

Time-Domain Features: In time-domain, maximum amplitude, mean, standard deviation (SD), coefficient of variation (COV), root mean square (RMS), skewness, and kurtosis features are extracted from the EEG signals to classify epilepsy.

Frequency-Domain Features: FFT method is applied to convert time-domain EEG signal to frequency-domain. In this study, the PSD using Welch's method FFT is calculated by dividing the data into overlapping segments of a 'hanning' window with one-sided spectrum for each EEG window of length 1 second (each of 256 segments). The obtained spectrum is divided into four sub-bands: Delta signal (0 Hz4 Hz), Theta signal (4Hz8 Hz), Alpha signal (813 Hz) and Beta signal (1330 Hz), then the following statistical features are extracted in each sub-band as features for epilepsy classification in frequency-domain.

- Mean value estimated for frequency ranges $(\delta, \theta, \alpha, \beta)$ and and complete frequency bandwidth using Equation 5.2

- Standard Deviation estimated for frequency ranges $(\delta, \theta, \alpha, \beta)$ and complete frequency bandwidth using Equation 5.3
- Root Mean Square (RMS) estimated for frequency ranges $(\delta, \theta, \alpha, \beta)$ and on all data using Equation 5.5
- Skewness estimated for frequency ranges $(\delta, \theta, \alpha, \beta)$ and and complete frequency bandwidth using Equation 5.6
- Kurtosis estimated for frequency ranges $(\delta, \theta, \alpha, \beta)$ and on all data using Equation 5.7

5.3.4 Data Preprocessing

Feature Selection

In this work, ANOVA F-distribution is an appropriate feature extraction method used to select the most dominant features from two categories of features (Normal and seizure). F-distribution is measured as the ratio between group variance and within group variance. If a value of F is large, then the value of P is observed to be less. The significance level, P is traditionally 1% or 5% [80]. In this study, the P-value is selected as 5% level of significance. All the time and frequency domain features are tested using ANOVA to find the value of P and according to its values, its significance is decided for feature selection. If the P value is less than 0.05, then the feature is relevant and can be used for classification.

Normalization

The obtained features from EEG data through the feature extraction steps are required to be classified by different classifiers. In order to improve classifiers accuracy, the values of extracted features need to be transferred to a standard scale, without distorting differences in the ranges of values. Hence, the features of each data should be normalized. In this study, the features are normalized to a scale of 0 and 1.

Data Sampling

After signal segmentation, this study has 300,175 normal activities and 3582 seizure activities. As the ratio of positive and negative cases differs widely, it is an imbalanced dataset. In order to avoid misclassification problems, this study applies RUS and SMOTE techniques to the data.

- **RUS:** RUS is a random under-sampling technique, in which the majority class is under-sampled by randomly selecting samples from the majority class. In this study, from 300,175 normal activities, 3582 samples are randomly selected to balance the ratio between normal and seizure activities.

- **SMOTE:** SMOTE is an oversampling technique to synthesize the minority class samples depending on k minority class nearest neighbors. In this study, SMOTE is used on seizure activities, and the nearest neighbors parameter is set to 5. That is, in the seizure data, select sample X and randomly selected one of the samples from five seizure samples that is more adjacent to X to produce a new synthesized sample. In this study, the seizure class is over-sampled to 300,175 samples to balance the ratio between normal and seizure classes.

The machine learning classifiers are separately trained with both RUS and SMOTE data sampling techniques.

Table 5.2: Classification accuracy after using RUS

Classifier	Accuracy (%)
DTC	83.07
EDTC	90.42
GBC	90.84
LDAC	76.74
MLPC	80.23
QDAC	71.58
RFC	89.25
SGDC	72.70

5.3.5 Construction of Classification

Data sampling techniques are used to improve imbalanced data. Data sampling techniques could be divided into undersampling and oversampling. In this study, in order to avoid misclassification, both under-sampling and oversampling are used to balance the data. At first, the number of non-epileptic samples is undersampled using RUS. The 3582 samples are randomly selected from 300,175 samples to balance the dataset. Then the classifiers are constructed with DTC, EDTC, GBC, LDAC, QDAC, MLPC, RFC, and SGDC. The accuracy of classification constructed by RUS technique is shown in table 5.2. In this study, SMOTE is also used. It is an oversampling technique to bring the ratio to 50:50. The 3582 samples are synthesized to balance the data. Then the classifiers are trained with the balanced data. Table 5.3 shows the accuracy of constructed classifiers after using SMOTE. Our results show that the SMOTE performed better compared to the RUS technique. RUS reduce the size of the dataset, and it is less computationally expensive in terms of implementation than SMOTE, but it may result in loss of relevant information. The accuracy of SMOTE technique is almost 10% higher in most classifiers compared to RUS. Therefore, our recommendation is to use SMOTE. It helps to balance the data, and also there is no loss of relevant information in data.

Table 5.3: Classification accuracy after using SMOTE

Classifier	Accuracy
DTC	96.54
EDTC	99.48
GBC	95.04
LDAC	78.95
MLPC	99.24
QDAC	64.39
RFC	99.23
SGDC	78.70

Table 5.4: Evaluation of classification models after using RUS

Classifier	Accuracy (%)	Specificity (%)	Sensitivity (%)
DTC	83.07	82.49	83.66
EDTC	90.42	90.59	90.25
GBC	90.84	91.84	89.89
LDAC	76.74	82.66	72.71
MLPC	80.23	89.27	74.66
QDAC	71.58	87.30	65.30
RFC	89.25	91.24	87.48
SGDC	72.70	70.42	75.46

5.3.6 Evaluation of Classification Models

The performance and ability of classifiers are measured using several common classification indicators, such as accuracy, specificity, and sensitivity. In this study, accuracy, specificity, and sensitivity respectively evaluate the performance of nine classification models. The results are shown in Table 5.4 and 5.5.

5.4 Comparison

The performance of our proposed approach is compared with previous methods proposed for seizure detection by different researchers with the CHB-MIT EEG scalp dataset, although other methods are tested with different conditions, such as

Table 5.5: Evaluation of classification models after using SMOTE

Classifier	Accuracy (%)	Specificity (%)	Sensitivity (%)
DTC	96.54	95.47	97.66
EDTC	99.48	99.17	99.79
GBC	95.04	94.70	95.39
LDAC	78.95	81.52	76.75
MLPC	99.24	98.69	99.79
QDAC	64.39	91.67	58.68
RFC	99.23	98.93	99.53
SGDC	78.70	81.23	76.54

a different selection of EEG records from CHB-MIT EEG scalp dataset, different prediction horizons, etc. Table 5.6 shows the performance of our proposed method compared to various other studies.

Table 5.6: Comparison of performance with previous studies

Study	Method	Accuracy	Specificity	Sensitivity
[63]	Common Spatial Pattern (CSP) and SVM	Not shown	Not shown	100 %
[64]	Common Spatial Pattern (CSP) and LDA	Not shown	Not shown	89 %
[81]	Variational Bayesian Gaussian Mixture Model of Zero-Crossing Intervals	Not shown	Not shown	88.34%
[65]	k-means unsupervised algorithm	91.43%	Not shown	Not shown
[66]	Stationary Wavelet Transform (SWT)	Not shown	Not shown	92.6%
[67]	local binary pattern (LBP) operator and K-NN classifier	99.7%	99.8%	99.2%
[68]	K-NN classifier	Not shown	88%	88%
[56]	pattern adapted wavelet transform and fuzzy classifier	96.48%	Not shown	Not shown
[82]	Spectral based analysis	Not shown	Not shown	86.67%
[83]	Recurrent convolutional network	99%	99%	84%
[84]	Convolutional neural networks (CNNs)	Not shown	Not shown	86.29%
This work	Machine Learning Algorithms and Welch FFT	**99.48%**	**99.17%**	**99.79%**

Chapter 6

Conclusion and Future Work

6.1 Summary of the Thesis

In this research, we have presented an epileptic seizure detection system based on Welch FFT and supervised machine learning algorithms. Welch FFT is used to process EEG signals in this work. The statistical features are extracted respectively in time domain and frequency domain. Then ANOVA based feature selection is used to deduct variables. The undersampling and oversampling methods are used after feature selection to balance the EEG data. Several powerful supervised machine learning algorithms are trained with the data. The results present 99.48% of accuracy, 99.79% of sensitivity, and 99.17% of specificity using EDTC classifier for seizure detection, which shows an improvement on existing studies. The system performances can be developed by selecting very precise features, advanced and more suitable machine learning algorithms.

6.2 Future Directions

Future work will investigate the use of more advanced algorithms, despite the good performance of the classifiers studied in this research. Furthermore, none of the existing studies proposed a generalized method for epileptic seizure prediction

[53]. Moreover, there is a need for modern features extraction methods to provide dominant features for prediction algorithms is a challenging task where many of the existing researches rely on handcraft feature extraction and/or tailored feature extraction methods, performed for each patient independently, which is not practical. Besides, when using implanted electrodes for brain recording, massive amounts of data are produced. Thus, big data analysis using cloud computing is required for safe storage and high computational resources for real-time processing [85]. To put in brief, the future work is categorized as follows:

1. Hybrid and multi-model feature extraction techniques capable of handling artefacts, improving low signal-to-noise and distortion-ratio (SNDR), handling non-linearity, handling non-stationarity, low memory consumption , and real-time signal processing capability.

2. Generalized and patient non-specific seizure detection and prediction system for the following problems:

   - Epileptic and non-epileptic EEG signals detection and classification
   - Classification between normal, pre-ictal and in-ictal stages of seizure.
   - Finding best channels pair for classification of seizures
   - Epilepsy seizure lateralization (function localization)
   - Classification of epilepsy seizures between children and adults.
   - Computation of seizure activities duration
   - Seizure activities visualization

3. Cloud-based BCI system for EGG big data and seizure prediction

4. Parallel processing of EEG big data with MapReduce

References

[1] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller, "Information transfer rate in a five-classes brain-computer interface," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 9, no. 3, pp. 283–288, 2001. ix, 5, 6

[2] S. Sanei and J. A. Chambers, "Eeg signal processing," 2007. ix, 5, 6, 7

[3] S. Siuly, Y. Li, and Y. Zhang, "Eeg signal analysis and classification," *IEEE Transactions on Neural Systems and Rehabilitaiton Engineering*, vol. 11, pp. 141–144, 2016. ix, 2, 8, 10, 14, 15, 16, 17, 18, 25, 26, 27, 29, 30, 31

[4] M. Teplan *et al.*, "Fundamentals of eeg measurement," *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002. ix, 9, 12, 13

[5] M. Kołodziej, A. Majkowski, and R. Rak, "Matlab fe_toolbox-an universal utility for feature extraction of eeg signals for bci realization," *Przeglad Elektrotechniczny*, vol. 1, 2010. ix, 29, 31

[6] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics New York, 2001. ix, 36

[7] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007. x, 38, 39, 41

[8] A. Shaker, A. H. Ali, M. Hamed, T. Saleh, and A. Emad, "Assessment of artificial neural network for bathymetry estimation using high resolution satellite imagery in shallow lakes: Case study el burullus lake.," 2015. x, 42

[9] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967. x, 65, 66

[10] S. Smith, "Eeg in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2–ii7, 2005. xi, 26, 28

[11] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel Jr, "Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005. 1, 17

[12] S. Supriya, S. Siuly, and Y. Zhang, "Automatic epilepsy detection from eeg introducing a new edge weight method in the complex network," *Electronics Letters*, vol. 52, no. 17, pp. 1430–1432, 2016. 1, 17

[13] S. Fazel, A. Wolf, N. Långström, C. R. Newton, and P. Lichtenstein, "Premature mortality in epilepsy and the role of psychiatric comorbidity: a total population study," *The Lancet*, vol. 382, no. 9905, pp. 1646–1654, 2013. 1, 17

[14] G. T. Keusch, J. Wilentz, and A. Kleinman, "Stigma and global health: developing a research agenda," *The Lancet*, vol. 367, no. 9509, pp. 525–527, 2006. 1, 17

[15] D. Hirtz, D. Thurman, K. Gwinn-Hardy, M. Mohamed, A. Chaudhuri, and R. Zalutsky, "How common are the common neurologic disorders?," *Neurology*, vol. 68, no. 5, pp. 326–337, 2007. 1, 17

[16] J. I. Sirven, "Epilepsy: a spectrum disorder," *Cold Spring Harbor perspectives in medicine*, vol. 5, no. 9, p. a022848, 2015. 1

[17] J. M.-T. Wu, M.-H. Tsai, C.-T. Hsu, H.-C. Huang, and H.-C. Chen, "Intelligent signal classifier for brain epileptic eeg based on decision tree, multilayer perceptron and over-sampling approach," in *Future of Information and Communication Conference*, pp. 11–24, Springer, 2019. 1

[18] J. Yoo and M. A. B. Altaf, "Machine-based patient-specific seizure classification system," Aug. 21 2014. US Patent App. 14/182,570. 2, 60

[19] L. Ayoubian, H. Lacoma, and J. Gotman, "Automatic seizure detection in seeg using high frequency activities in wavelet domain," *Medical engineering & physics*, vol. 35, no. 3, pp. 319–328, 2013. 2, 60

[20] R. Hopfengärtner, B. S. Kasper, W. Graf, S. Gollwitzer, G. Kreiselmeyer, H. Stefan, and H. Hamer, "Automatic seizure detection in long-term scalp eeg using an adaptive thresholding technique: a validation study for clinical routine," *Clinical Neurophysiology*, vol. 125, no. 7, pp. 1346–1352, 2014. 2, 60

[21] B. Hunyadi, M. Signoretto, W. Van Paesschen, J. A. Suykens, S. Van Huffel, and M. De Vos, "Incorporating structural information from the multichannel eeg improves patient-specific seizure detection," *Clinical Neurophysiology*, vol. 123, no. 12, pp. 2352–2361, 2012. 2, 60

[22] J. Yoo, L. Yan, D. El-Damak, M. B. Altaf, A. Shoeb, H.-J. Yoo, and A. Chandrakasan, "An 8-channel scalable eeg acquisition soc with fully integrated

patient-specific seizure classification and recording processor," in *2012 IEEE International Solid-State Circuits Conference*, pp. 292–294, IEEE, 2012. 2, 60

[23] J. D. Bronzino, "Principles of electroencephalography," *The biomedical engineering handbook*, vol. 1, 1995. 5

[24] J. R. Wolpaw, "Brain-computer interfaces (bcis) for communication and control: a mini-review," in *Supplements to Clinical neurophysiology*, vol. 57, pp. 607–613, Elsevier, 2004. 7

[25] N. Birbaumer, "Brain–computer-interface research: coming of age," *Clinical neurophysiology*, vol. 117, no. 3, pp. 479–483, 2006. 7

[26] R. Bickford, "Electroencephalography: encyclopedia of neuroscience," 1987. 7

[27] G. Kondraske, "Neurophysiological measurements," *Biomedical Engineering and Instrumentation*, pp. 138–179, 1986. 12

[28] G. H. Klem, H. O. Lüders, H. Jasper, C. Elger, *et al.*, "The ten-twenty electrode system of the international federation," *Electroencephalogr Clin Neurophysiol*, vol. 52, no. 3, pp. 3–6, 1999. 12

[29] 14, 16

[30] B. J. Fisch and R. Spehlmann, "Eeg primer: basic principles of digital and analog eeg," *Elsevier Health Sciences*, pp. 23–27, 1999. 14, 16

[31] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002. 29

[32] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals," *Journal of Neural engineering*, vol. 4, no. 2, p. R32, 2007. 29, 30

[33] T. Hauck, *scikit-learn Cookbook*. Packt Publishing Ltd, 2014. 38

[34] L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classification and regression trees (wadsworth, belmont, ca)," *ISBN-13*, pp. 978–0412048418, 1984. 39

[35] M. Brijain, R. Patel, M. Kushik, and K. Rana, "A survey on decision tree algorithm for classification," 2014. 39

[36] M. Asafuddoula, B. Verma, and M. Zhang, "A divide-and-conquer-based ensemble classifier learning by means of many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 5, pp. 762–777, 2017. 44

[37] T. K. Ho, "Random decision forests," in *Document analysis and recognition, 1995., proceedings of the third international conference on*, vol. 1, pp. 278–282, IEEE, 1995. 45

[38] M. M. A. Abdallh, K. H. Bilal, and A. Babiker, "Machine learning algorithms," *International Journal of Engineering, Applied and Management Sciences Paradigms*, vol. 36, no. 01, pp. 17–27, 2016. 45

[39] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001. 45

[40] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3–42, 2006. 45

[41] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001. 46

[42] T. Cecchin, R. Ranta, L. Koessler, O. Caspary, H. Vespignani, and L. Maillard, "Seizure lateralization in scalp eeg using hjorth parameters," *Clinical Neurophysiology*, vol. 121, no. 3, pp. 290–300, 2010. 47

[43] U. R. Acharya, S. V. Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic eeg activities in wavelet framework," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9072–9078, 2012. 48

[44] M.-P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, and D. Pompili, "Cloud-based deep learning of big eeg data for epileptic seizure prediction," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1151–1155, IEEE, 2016. 48, 49

[45] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for eeg big data and seizure prediction bci via internet of things," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 392–404, 2017. 49

[46] M. Li, W. Chen, and T. Zhang, "b," *Biomedical Signal Processing and Control*, vol. 31, pp. 357–365, 2017. 50

[47] S. Ibrahim, R. Djemal, and A. Alsuwailem, "Electroencephalography (eeg) signal processing for epilepsy and autism spectrum disorder diagnosis," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 16–26, 2018. 50

[48] I. Ullah, M. Hussain, H. Aboalsamh, *et al.*, "An automated system for epilepsy detection using eeg brain signals based on deep learning approach," *Expert Systems with Applications*, vol. 107, pp. 61–71, 2018. 51

[49] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals," *Computers in biology and medicine*, vol. 100, pp. 270–278, 2018. 52

[50] S. K. Satapathy, S. Dehuri, and A. K. Jagadev, "Eeg signal classification using pso trained rbf neural network for epilepsy identification," *Informatics in Medicine Unlocked*, vol. 6, pp. 1–11, 2017. 52, 57

[51] O. Kocadagli and R. Langari, "Classification of eeg signals for epileptic seizures using hybrid artificial neural networks based wavelet transforms and fuzzy relations," *Expert Systems with Applications*, vol. 88, pp. 419–434, 2017. 53

[52] J. Martinez-del Rincon, M. J. Santofimia, X. del Toro, J. Barba, F. Romero, P. Navas, and J. C. Lopez, "Non-linear classifiers applied to eeg analysis for epilepsy seizure detection," *Expert Systems with Applications*, vol. 86, pp. 99–112, 2017. 54

[53] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Networks*, vol. 105, pp. 104–111, 2018. 55, 78

[54] Y. Yao, J. Chang, and K. Xia, "A case of parallel eeg data processing upon a beowulf cluster," in *2009 15th International Conference on Parallel and Distributed Systems*, pp. 799–803, IEEE, 2009. 55

[55] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Random ensemble learning for eeg classification," *Artificial intelligence in medicine*, vol. 84, pp. 146–158, 2018. 56

[56] V. Harpale and V. Bairagi, "An adaptive method for feature selection and extraction for classification of epileptic eeg signal in significant states," *Journal of King Saud University-Computer and Information Sciences*, 2018. 57, 62, 67, 76

[57] D. Sikdar, R. Roy, and M. Mahadevappa, "Epilepsy and seizure characterisation by multifractal analysis of eeg subbands," *Biomedical Signal Processing and Control*, vol. 41, pp. 264–270, 2018. 57

[58] K. Devarajan, E. Jyostna, K. Jayasri, and V. Balasampath, "Eeg-based epilepsy detection and prediction," *International Journal of Engineering and Technology*, vol. 6, no. 3, p. 212, 2014. 57

[59] P. Ramina and M. Vanitha, "Fast and effective real time seizure prediction on streaming eeg signals," 57

[60] R. Sadaye, S. Parekh, J. Samuel, and K. Deulkar, "Epileptic seizure prediction using power spectrum and amplitude analysis of beta band of eeg signals," *International Journal of Computer Applications*, vol. 155, no. 9, 2016. 57

[61] N. Sivakumaran, E. Sathish, K. Radhakrishnan, S. Raghavan, M. Rajajeyakumar, *et al.*, "Importance of classification algorithms in epileptic seizure detection," *J. Bioeng. Biomed. Sci*, vol. 6, no. 3, 2016. 57

[62] F. Jesudas, S. Nagarajan, S. SureshBabu, D. Maheswar, G. Ranganath, N. Srimath, N. Murugan, R. Ezhilarasi, R. Minu, B. Sahu, *et al.*, "Available online at www. sciencedirect. com," 57

[63] T. N. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, F. M. Alotaibi, K. Aljibrin, S. R. Alrshod, I. M. Alkhanin, and N. Alrajhi, "A common spatial pat-

tern approach for scalp eeg seizure detection," in *2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, pp. 1–5, IEEE, 2016. 58, 76

[64] T. N. Alotaiby, S. A. Alshebeili, F. M. Alotaibi, and S. R. Alrshoud, "Epileptic seizure prediction using csp and lda for scalp eeg signals," *Computational intelligence and neuroscience*, vol. 2017, 2017. 58, 76

[65] S. Chakrabarti, A. Swetapadma, P. K. Pattnaik, and T. Samajdar, "Pediatric seizure prediction from eeg signals based on unsupervised learning techniques using various distance measures," in *2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech)*, pp. 1–5, IEEE, 2017. 58, 76

[66] L. Orosco, A. G. Correa, P. Diez, and E. Laciar, "Patient non-specific algorithm for seizures detection in scalp eeg," *Computers in biology and medicine*, vol. 71, pp. 128–134, 2016. 59, 76

[67] P. M. Shanir, K. A. Khan, Y. U. Khan, O. Farooq, and H. Adeli, "Automatic seizure detection based on morphological features using one-dimensional local binary pattern on long-term eeg," *Clinical EEG and neuroscience*, vol. 49, no. 5, pp. 351–362, 2018. 59, 76

[68] P. Fergus, A. Hussain, D. Hignett, D. Al-Jumeily, K. Abdel-Aziz, and H. Hamdan, "A machine learning system for automated whole-brain seizure detection," *Applied Computing and Informatics*, vol. 12, no. 1, pp. 70–89, 2016. 59, 76

[69] A. H. Shoeb, *Application of machine learning to epileptic seizure onset detection and treatment.* PhD thesis, Massachusetts Institute of Technology, 2009. 61, 69

[70] P. Akhilesh Krishna and J. Andrews, "Psd computation using modified welch algorithm," *Int J Sci Res Eng Technol (IJSRET)*, vol. 4, no. 9, pp. 951–954, 2015. 65

[71] L. Al Shalabi, Z. Shaaban, and B. Kasasbeh, "Data mining: A preprocessing engine," *Journal of Computer Science*, vol. 2, no. 9, pp. 735–739, 2006. 66

[72] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv preprint arXiv:1503.06462*, 2015. 66

[73] M. Bennasar, Y. Hicks, and R. Setchi, "Feature selection using joint mutual information maximisation," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8520–8532, 2015. 67

[74] S.-H. Lee and J. S. Lim, "Minimum feature selection for epileptic seizure classification using wavelet-based feature extraction and a fuzzy neural network," *Applied Mathematics & Information Sciences*, vol. 8, no. 3, p. 1295, 2014. 67

[75] W. Penny and R. Henson, "Analysis of variance," *Statistical parametric mapping: The analysis of functional brain images*, pp. 166–177, 2006. 67

[76] T. T. Chowdhury, S. A. Fattah, and C. Shahnaz, "Classification of seizure and non-seizure activity in seizure patients using time-frequency domain processing of gamma band eeg signals," in *2017 4th International Conference on Advances in Electrical Engineering (ICAEE)*, pp. 537–540, IEEE, 2017. 67

[77] A. Liu, J. Ghosh, and C. E. Martin, "Generative oversampling for mining imbalanced datasets.," in *DMIN*, pp. 66–72, 2007. 67

[78] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002. 67

[79] L.-L. Chen, J. Zhang, J.-Z. Zou, C.-J. Zhao, and G.-S. Wang, "A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection," *Biomedical Signal Processing and Control*, vol. 10, pp. 1–10, 2014. 69

[80] R. Nuzzo, "Scientific method: statistical errors," *Nature News*, vol. 506, no. 7487, p. 150, 2014. 71

[81] A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting epileptic seizures in scalp eeg based on a variational bayesian gaussian mixture model of zero-crossing intervals," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 5, pp. 1401–1413, 2013. 76

[82] H. Chu, C. K. Chung, W. Jeong, and K.-H. Cho, "Predicting epileptic seizures from scalp eeg based on attractor state analysis," *Computer methods and programs in biomedicine*, vol. 143, pp. 75–87, 2017. 76

[83] W. Liang, H. Pei, Q. Cai, and Y. Wang, "Scalp eeg epileptogenic zone recognition and localization based on long-term recurrent convolutional network," *Neurocomputing*, 2019. 76

[84] R. Yuvaraj, J. Thomas, T. Kluge, and J. Dauwels, "A deep learning scheme for automatic seizure detection from long-term scalp eeg," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 368–372, IEEE, 2018. 76

[85] S. Krishnan and Y. Athavale, "Trends in biomedical signal feature extraction," *Biomedical Signal Processing and Control*, vol. 43, pp. 41–63, 2018. 78

[86] N. Hazarika, J. Z. Chen, A. C. Tsoi, and A. Sergejew, "Classification of eeg signals using the wavelet transform," in *Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on*, vol. 1, pp. 89–92, IEEE, 1997.

[87] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," *Journal of neuroscience methods*, vol. 123, no. 1, pp. 69–87, 2003.

[88] M. Schröder, T. N. Lal, T. Hinterberger, M. Bogdan, N. J. Hill, N. Birbaumer, W. Rosenstiel, and B. Schölkopf, "Robust eeg channel selection across subjects for brain-computer interfaces," *EURASIP Journal on Applied Signal Processing*, vol. 2005, pp. 3103–3112, 2005.

[89] E. Alickovic, J. Kevric, and A. Subasi, "Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction," *Biomedical Signal Processing and Control*, vol. 39, pp. 94–102, 2018.

[90] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, "A novel eeg feature extraction method using hjorth parameter," *International Journal of Electronics and Electrical Engineering*, vol. 2, no. 2, pp. 106–110, 2014.

[91] S. M. Usman, M. Usman, and S. Fong, "Epileptic seizures prediction using machine learning methods," *Computational and mathematical methods in medicine*, vol. 2017, 2017.

[92] A. D. Bhople and P. Tijare, "Fast fourier transform based classification of epileptic seizure using artificial neural network," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 4, 2012.

[93] D. R. Meenakshi and A. Singh, "Frequency analysis of healthy & epileptic seizure in eeg using fast fourier transform," *International Journal of Engineering Research and General Science*, vol. 2, no. 4, pp. 683–691, 2014.

[94] Q. I. Rahman and G. Schmeisser, "Characterization of the speed of convergence of the trapezoidal rule," *Numerische Mathematik*, vol. 57, no. 1, pp. 123–138, 1990.

[95] R. Acharya, O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, "Non-linear analysis of eeg signals at various sleep stages," *Computer methods and programs in biomedicine*, vol. 80, no. 1, pp. 37–45, 2005.