بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

نَرْفَعُ دَرَجَاتٍ مَّن نَّشَاءُ وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ

صدق الله العلي العظيم

(سورة يوسف / الاية 76)

AUTOMATIC EPILEPTIC SEIZURE DETECTION IN EEG BASED ON AN EFFICIENT MACHINE LEARNING TECHNIQUE

THESIS

SUBMITTED TO THE COMMUNICATION TECHNIQUES ENGINEERING DEPARTMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER

BY

AMAL SALMAN ABDULHUSSEIN

Supervised by

Prof. Dr. Ahmad T. Abdulsadda

2021/11

Supervisor Certification

I certify that this thesis titled " **Automatic Epileptic Seizure Detection in EEGs Based on an Efficient Machine Learning Technique** " which is being submitted by Name of student was prepared under our supervision at insert name of Department, name of college, AL-Furat Al-Awsat Technical University, as a partial fulfilment of the requirements for the degree of master.

Signature:

Name: Prof. Dr. Ahmad T. Abdulsadda

(Supervisor)

DATE: / /2021

Signature:

In view of the available recommendation, I forward this thesis for debate by the examining committee.

Signature:

Name: Prof. Dr. Ahmad T. Abdulsadda

(Head of the Communication Techniques Engineering Dept.)

Date: / /2021

Committee Report

We certify that we have read this thesis titled **"Automatic epileptic seizure detection in EEGs based on an efficient machine learning technique"** which is being submitted by **AMAL SALMAN ABDULHUSSEIN** and as Examining Committee, examined the student in its contents. In our opinion, the thesis is adequate for the award of the degree of Master of Technical in Communication Engineering.

Signature:

Name: **Prof. Dr. Ahmad T. Abdulsadda**

(Supervisor)
Date: / / 2021

Signature:

Name: **Asst. Prof. Dr. Bashar Jabbar Hamza**

(Member)
Date: / / 2021

Signature:

Name: **Dr. Ahmed Al Hilli**

(Member)
Date: / / 2021

Signature:

Name: **Prof. Dr. Hawraa Hassan Abbas**

(Chairman)
Date: / / 2021

Approval of the Engineering Technical College-Najaf

Signature:

Name: **Asst. Prof. Dr. Hassanain Ghani Hameed**

Dean of Engineering Technical College-Najaf

Date: / / 2021

Linguistic Certification

This is to certify that this thesis entitled " **Automatic epileptic seizure detection in EEGs based on an efficient machine learning technique** " was reviewed linguistically. Its language was amended to meet the style of the English language.

Signature:

Name:

Date: / /2021

Abstract

Epilepsy is a dangerous disease that may affect humans of all ages about 1% of the people in the world. A human brain generates messages at all times. A seizure occurs when this communication breaks down or when abnormal events occur. Scalp Electroencephalogram (EEG) and internal Electroencephalogram (iEEG) are two monitoring technologies that can be used to track this uncommon behavior. Due to the EEG and iEEG are low amplitude, complexity and non-stationary signal, the odds of missing a seizure increase. Even though these monitoring devices identify seizures, there are still some challenges, such as accurate seizure detection and fast seizure detection. In this thesis, the advantage of techniques are used to solve the two challenges and study the effect of fast detection on accuracy. Firstly, two datasets (Bonn and CHB-MIT) are used and two groups of features (feature extracted without delay and with delay) are applied that are extracted from the EEG signal directly or the features extracted after converting the EEG signal to the frequency domain by using a new approach which is Sliding Discrete Fourier Transform (SDFT). This is due to the fact that signal data contains a variety of features that aid seizure detection. Secondly, after assessing the performance of multiple classifiers, the most appropriate have identified and also the effective classifier for detecting seizures from EEG and iEEG datasets. Thirdly, we analyses the importance of each feature and select the most features have effect on seizure detection. This study appears that the SDFT feature has high importance for seizure detection and the random forest (RF) and Artificial neural network (ANN) are better than decision tree (DT) and adaptive-network-based fuzzy inference system (ANFIS). Finally this study appear the maximum acuuracy was 100% with one sec delay for iEEG datasets and 96% with one sec delay for scalp EEG dataset.

Dedicated

To the first teacher in life, who learned and encouraged me to be learned and educated, in blessing of his breath and his existence in my life… **My beloved father**…Sadly, you passed away before sharing with me the moment of achieving your and my dream, I wish you could share it with me my dear father…. May Allah have mercy on your soul.

To the light of my eyes, the spark of my path, and the biggest role model in my life…**My beloved mother**…I went through the paths of knowledge in blessing of your prayer hymns…May Allah protect you and prolong your life.

To my supporter and the lifelong companion… **My precious husband…Ali**… .

To the most precious gifts in my life…My precious and lovely children…**My tittle princess… Asma**… **and my heroes…Hassan and Hussain**…I love you all.

To whom their existence, is the biggest happiness in my life…My …. **Brother…Sisters**.

To whom always support and encourage me during my study …**My family in law**.

To all **Epileptic patients** across the world.

Acknowledgement

In the beginning I thank my god Allah, for facilitating and helping me in completing my master's research. Praise and thanks to Allah.

First, even though, I know there is no word can express my thankful feeling toward my supervisor, I would like to express deeply my gratitude to Prof.Ahmad Taha for accepting me in his group, it is such an honour to be one of your students. I’m grateful for your advice, support, and guidance, and for your wide and deep scientific knowledge that transformed my narrow limits. I’m so proud to be under your supervision and have your name on my thesis, it is such an honour!

Also my grateful for my faculty (Department of Communication Engineering, Engineering Technical College/Najaf, Al-Furat Al- Awsat Technical University).

I would like to thank all of my friends and colleagues.

DECLARATION

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

1August 2021 AMAL SALMAN ABDULHUSSEIN

Table of Content

List of Tables

Table **Page**

List of figures

List of Abbreviations

Symbol	Description
RNS	Responsive Nerve Stimulation
EEG	ElectroEncephaloGram
iEEG	Internal EEG
FFNN	FeedForward Neural Network
ANFIS	Adaptive-Network-based Fuzzy Inference System
DT	Decision Tree
RF	Random forest
DSTFT	Discrete Short Time Fourier Transform
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
MLP	Multilayer Perceptron
DWT	Discrete Wavelet Transform
EMD	Empirical Mode Decomposition
PSR	Phase Space Representation
PCA	Principle Component Analysis
CHB-MIT	Children's Hospital Boston- Massachusetts Institute of Technology
LDA	Linear Discriminant Analysis
NB	NAIVE BAYES
LR	Logistic Regression
WPD	Wavelet Packet Decomposition
ANN	Artificial Neural Network
LDAG-SVM	Layered Directed Acyclic Graph Support Vector Machine
AE	Approximate Entropy
LLE	Largest Lyapunov Exponent
CD	Correlation Dimension
SE	Spectral Entropy
PE	Permutation Entropy
TQWT	Tunable-Q Wavelet Transform

MEMD	Multivariate Empirical Mode Decomposition
MGA	Modified Genetic Algorithm
PFEN	Permutation Fuzzy Entropy
MAV	Mean Absolute Value
AVP	Average Power
SD	Standard Deviation
GE	Genetic Algorithm
Edf	European Data Format
PHI	Protected Health Information
VNS	Vagal Nerve Stimulus
STFT	Short-Time Fourier Transform
CNN	Convolutional Neural Networks
SDFT	Sliding Discrete Fourier Transform
DSP	Digital Signal Processing
DFT	Discrete Fourier Transform
IIR	Infinite Impulse Response
MDI	Mean Decrease in Impurity
ROC	Receiver Operating Characteristics
ZCI1	Zero Crossing for First Derivative
ZC2	Zero Crossing for Second Derivative
NMM	No. of Peaks
LL	Line Length
MXI	Maximum Value
MNI	Minimum Value
d1	First Derivative
d2	Second Derivative
Absd1	Absolute of First Derivative
Absd2	Absolute of Second Derivative
SFS	Sequential Feature Selection
IOT	Internet of things
PSO	Practical Swarm Optimization

List of Symbols

Symbol	Definition	Unit
X_n	SDFT amplitude	
x	Sample value in the time domain	volt
k	frequency-domain index	------
m	The index of the first sample in the window	------
M	Window's size	------
$\hat{n}$	The index of the first sample for specific window from time-series signal.	------
n	time-domain index	------
x'	The first derivative for the signal	volt
x''	The second derivative for the signal	volt
L	Line length value	volt
var	Variance	volt
En	Energy	joule
x_j	The node input	------
w_{ij}	The weight for each node	------
y_i	The result from the summation of the multiplication of each node input with its weight	------
s	Sigmoid function	------
y	The output from the ANFIS classifier	------
w_i	The adaptive parameters for the ANFIS classifier	------
G	Gini impurity	------
p_i	Probability	------

CHAPTER ONE

Introduction and Literature Review

1.1.Introduction

Epilepsy is a potentially fatal disease that is caused by a neurological disorder that could affect the people at different ages, jeopardizing the individual's daily life due to repeating seizures[1]. Around 0.6- 0.8 percent of the world's population have suffered from epilepsy, which can impose numerous restrictions on their daily lives and even put their lives in danger. Generally, there are three main types that are used to seizure treatment

- Some of them are treatable with antiepileptic drugs, but these medications have side effects, and approximately 30% of epilepsy patients did not respond to drug treatment[2].
- The other method is to surgically remove the region of the brain that generates epileptic activity; in this case, the location of the epileptic region should be determined prior to the operation [3] [4].
- Another epilepsy terminal technique is responsive nerve stimulation (RNS), which is designed to continuously monitor intracranial electroencephalogram (iEEG) activity and then rapidly respond to high-frequency neurostimulation to "disrupt" the network synchronization required for seizure progression and spread [5].

Electroencephalogram (EEG) signals are widely used for assessing brain functions and for epilepsy investigation because they contain a wealth of information about the electrical activity of the brain[6]. The EEG signals were recorded from the scalp of three distinct subject groups: (a) healthy subjects, (b) epileptic subjects, during a seizure-free interval, and (c) epileptic subjects during a seizure. Visual analysis of long-term EEG signals by expert people is the most efficient method for identifying epileptic seizures [2]. This is a very time-consuming and tedious task, even more as the number of EEG electrodes increases. Automatic seizure detection may aid epileptologists in evaluating long-term EEG recordings and may also be used in closed-loop therapeutic systems such as implantable electrical stimulation devices.[7].

Some of the information contained in the EEG signal can be extracted as features and used to automatically detect epileptic seizures through the use of classification methods, as illustrated in Figure 1.1.

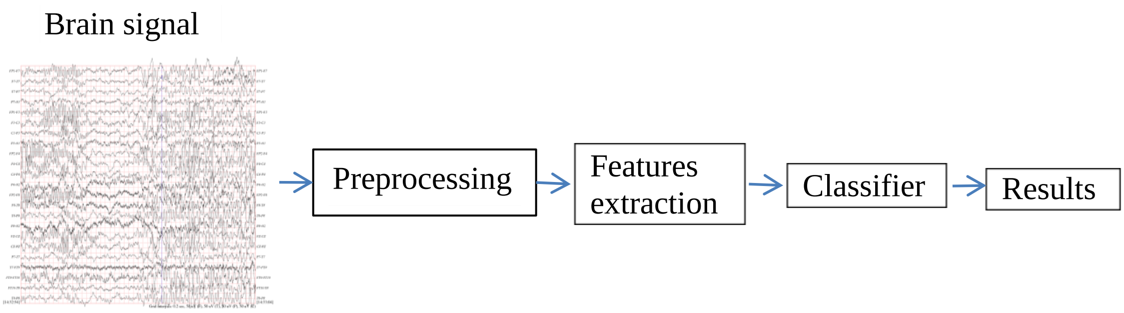


Figure 1.1. The block diagram for the system that is used for automatic seizure detection.

In state of art, many studies tried to find which features that should be extracted and which classification method is employed to give high accuracy. This thesis focused on the accuracy and the time required for automatic seizure detection.

In this thesis, two datasets have been used :

- The first dataset is iEEG (internal EEG) taken from Bonn dataset which is measured through an interracial electrode, as shown in figure 1.2. Depth electrodes were symmetrically implanted into the tops of the hippocampal formations. Segments from sets C, D and E were taken from all of the depth electrode's contacts. Strip electrodes were implanted into the neocortical lateral and basal regions, as well as the middle and bottom.

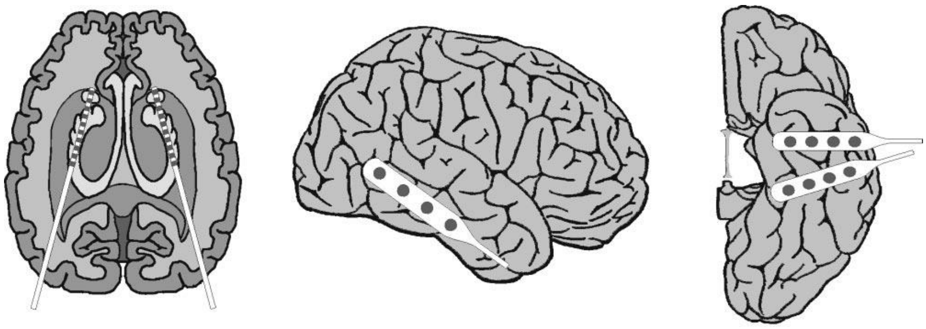


Figure 1.2. . Schematic representation of intracranial electrodes implanted in epilepsy patients prior to surgery[3].

- The other dataset of brain signal that is used in this thesis was EEG signals that are measured from the scalp. This dataset was measured with 23 electrodes and recorded and saved in a digital file figure 1.3.

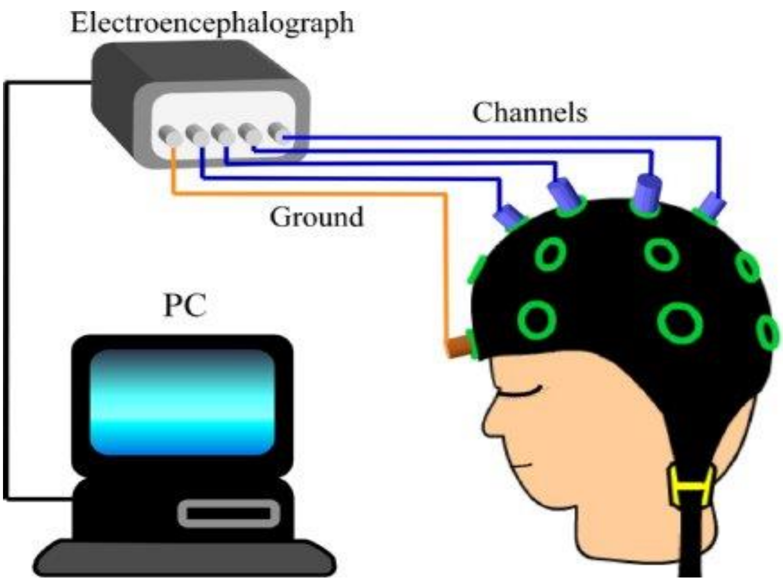


Figure 1.3. Schematic representation of scalp electrodes [8].

To prepare the EEG signl and reduce the noise and the artifact, preprocessing is required. The preprocessing that used was windowing for the data prepare and filter for reduce the noise and the artifact signals. After preprocessing step two types of feature extraction have been used :

- Feature extraction method that extracts feature without any delay (no windowing required). The features values dependent on the current sample or the current and previous samples.
- Feature extraction method that extracts feature with time delay (with window length). The features values dependent on the many samples in window.

The Artificial Neural Network (ANN), adaptive-network-based fuzzy inference system (ANFIS), Decision Tree (DT) and random forest (RF) are used as binary classifiers where one denoted for seizure state and zero denoted for the normal state.

1.2.Literature review

Machine learning has become more interesting for many researchers in recent years. Many researchers have used machine learning in medical applications, and one of these applications is automatic epileptic seizure detection. Many strategies and studies have been used in this field to improve the accuracy of seizer detection at a suitable time.

Some of the published related works for automatic epileptic seizure detection are summarized as follows:

K. Samiee, et al. in 2013 [9]**:** The authors consulted an EEG database made available online by the University of Bonn. The Discrete Short Time Fourier

Transform (DSTFT) technique is used to extract features from EEG data. The classifiers used are Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Multilayer Perceptron (MLP). After a series of numerical experiments, the optimal classifier was determined to be the feedforward MLP where the maximum accuracy was 99.8%.

L. Wang et al. in 2017[10]**:** The Discrete Wavelet Transform (DWT) applied on Bonn dataset and wavelet threshold is used for denoising, and signal analyzed in the time, frequency, and time-frequency domains and then the features are extracted from the analyzed signal, also Empirical Mode Decomposition – Phase Space Representation (EMD–PSR) and entropies used for nonlinear analysis. The total features extracted were 83 features. To improve classification performance and reduce the complexity the researchers applied some techniques such as Principle Component Analysis (PCA) for reducing features and select the optimal features subset from all features extracted. Five classifiers (KNN, LDA, NB, LR, SVM) were used for the classification purpose. The maximum accuracy in this study achieved was 99.25%.

E. Alickovic, et al. in 2017[11]**:** The authors developed a model segment-based for the classification of EEG signals. Large two EEG database applied Freiburg and Children's Hospital Boston- Massachusetts Institute of Technology (CHB-MIT) datasets, and Wavelet Packet Decomposition (WPD), DWT and EMD used for feature extraction where Mean, Average power, Standard deviation, Ratio of absolute mean values, Skewness and Kurtosis applied in every sub-band, and four classifiers which are Random forest (RF), K-Nearest neighbour (k-NN), Multilayer perceptron (MLP) and Support vector machine (SVM) used to test the efficiency of three proposed feature extraction methods. The maximum accuracy in this study achieved was 100%.

M. Mursalin, et al. in 2017 [12]**:** The authors used EEG data from Bonn University to display and extract time domain and entropy-based features from the EEG, as well as the maximum, minimum, mean, and standard deviation from the DWT components. Random Forest Classifier is used as a classifier, improved correlation based feature selection which developed by authors is used to select the most effective features for classification. The maximum accuracy in this study achieved was 100%.

A. Narang, et al. in 2018 [13]**:** The DWT, statistical values, Shannon Entropy, Line Length, Approximate Entropy and Dominant Frequency are used for feature extraction from the EEG data downloaded from University of Bonn, Germany. ANN and Support Vector Machine (SVM) are used for the classification process. From the comparison of the simulation results from the two classifiers, the results from the ANN classifier are better than the results when the SVM classifier was used. The accuracy was 99.6% achieved in this study.

S. Ramakrishnan, et al. in 2018 [14]**:** The authors designed a method that depends on the fuzzy rules applied for sub-band specific features and for the EEG signal classification process layered directed acyclic graph support vector machine (LDAG-SVM) is used. Two EEG datasets used Bonn and CHB-MIT used for testing and evaluate the designed method. The wavelet transformation used for decomposed EEG signal into sub-bands. After the decomposition process, three features are extracted from each sub-band which are Approximate Entropy (AE), Largest Lyapunov Exponent (LLE) and Correlation Dimension (CD). The maximum accuracy achieved was 100%.

M. Ravi, et al. in 2018 [15]: The Bern-Barcelona EEG and the Bonn University EEG databases are used to investigate EEG-based epileptic seizure detection in this study. Decomposition, feature extraction, and classification are the three primary processes in the proposed technique. Decomposition utilizing variational mode decomposition initially provides efficient frequency localization. Following decomposition, semantic feature extraction is performed using differential entropy and the peak magnitude of the root mean square ratio to get optimal feature subsets and to eliminate irrelevant and redundant features. Following the discovery of feature information, a superior classifier known as the RF is used to categorize the data. The maximum accuracy for this model was 94.1%.

C. Mahjoub, et al. in 2019 [16]: The authors used many features, including linear and nonlinear measures associated with both time and frequency domains extracted from data available at the Epileptology Department of Bonn University. The features consist of five statistical measures, the zero-crossing rate, three Hilbert transform attributes, the spectral entropy (SE) and the Permutation Entropy (PE) with different windows size (1000, 2000, 3000 and 4000 samples). These features are extracted from the raw data directly or from its decomposition using Tunable-Q Wavelet Transform (TQWT) and Multivariate Empirical Mode Decomposition (MEMD). For avoiding the potential over-fitting issue, the feature subset selection step implemented by using the ranking criterion is computed for all applied features. The SVM classifier is used as classifier. The maximum accuracy for this model was 99.81%.

M. Madhusmita, et al. in 2019 [17]**:** The seizure EEG signal that was used recorded from cardiac patients with epilepsy (6 patients) in the NIMHANS centre in Bangalore, India. The DWT used for feature extraction the ANFIS used as a classifier with Modified Genetic Algorithm (MGA) to select the best parameters for the classifier. The higher accuracy obtained was 99.35%

W. Hussain, et al. in 2019 [4]**:** The authors focused on the methodology that used one feature extraction from EEG, Freiburg EEG database, signals which is Permutation Fuzzy Entropy (PFEN) and apply it to six different classifiers (ANN, KNN, SVM, Decision Tree (DT), Ensemble Classifier and KNN) and compare the accuracy results from all six classifiers output. The higher accuracy was 98.26%.

W. Mardini et al. in 2020 [18]**:** The authors suggested a method that extracted features from the DWT analysis of the data from Bonn University. These features include Standard Deviation (SD), Mean Absolute Value (MAV), Average Power (AVP), Max, Min, Mean, Skewness, Shannon Entropy, Kurtosis, Normalized SD, and Energy. A Genetic Algorithm (GA) is used to minimized those extracted features where GA select the most relevant features. Finally, four classifiers (SVM, Naive Bayes (NB) Classifier, KNN and ANN) were used for the EEG dataset classification to two cases of an epileptic seizure or not. Accuracy of 100% was achieved in this study.

O. Kaziha et al. in 2020 [19]**:** The authors used dataset "CHB-MIT" and proposed a method a neural network based on a convolutional neural network (CNN). Seizure detection hasn't completely explored deep learning, instead of relying on traditional machine learning algorithms that require feature extraction. Deep learning should be investigated since it eliminates laborious feature extraction and allows for real-time detection of raw signals. The higher accuracy was 96.7%.

In this thesis the two datasets are used iEEG and scalp EEG and the features are extracgted from time domain and from frequency domain, new approach fast and simple component method (SDFT) used to convert the iEEG and EEG time domain signals to frequency domain. Two groups of features used the first group is extracted the features dependent on the current sample or calculated frorm the current and previous samples, in this cases the seizure detection will be instansully while the second method is extracted the features many samples in uncrossing windows for this resean this method need time dalay for features calculation (one and two sec are applied). Finally the two type of black-box classifiers (ANN and ANFIS) and non-black-box classifiers (DT and RF) are used for classification.

1.3. Problem Statements

Basically, the first challenge EEG signals suffer from noise and its amplitude is low that is making the embedded in noise, especially in the low frequencies. Also EEG is complexity and non-stationary signal this leads to difficult to classify the signal into two categories (normal state and seizure state) with less time resolution and high accuracy. To obtain the reusable classification needed system with a high performance of classification depended on the features extracted from the EEG signal.

The other challenge, EEG signals are long time series signals recorded that need large memory space for saving and for feature calculation.

1.4.Thesis Objectives

Two main objectives in this research. They are summarized as follows:

1- Studying and analysing extracted features that have important information that are used to detect the seizure.

2- Comparing and finding the best classifier(which has a high accuracy) that is used for seizure detection.

1.5.Expected Contributions

-Firstly, choosing appropriate statistical features for iEEG and EEG signals is an important first step in processing raw datasets.

-Secondly, while there are several classifiers available, not all of them are appropriate for every dataset. As a result, the focus of this work is on identifying an appropriate classifier for epileptic seizure detection that has a decent performance metric.

-The third major contribution is based on the idea of quickly detecting a seizure while preserving accuracy. For this, we analyse and test different features that extracted from current samples or the current and preivous samples and compare the results with other features that need many samples for its calculation.

-The last and final contribution is reduce the calculation cost by select the the active features that effect on classification.

1.6.Outlines

This thesis is divided into five chapters, and a detailed description for each chapter is as follows:

- **Chapter two:** Brief explanation of type seizure and type of EEG and the methods used for seizure detection by using EEG signals. Also describes the data that used in this thesis and the preprocessing, feature extraction methods, classifiers, the parameters that are used to measure the performance and the methods that are used to measure the feature importance and feature selection.
- **Chapter three:** This chapter shows the design of methods and the machine learning models and the classifiers that are used in each dataset.
- **Chapter four:** This chapter presents the results in each step with the different models when applying the datasets and the result from the measuring methods used to measure the feature importance for each feature used in this thesis and resuls from the select the features without effect on the accuracy.
- **Chapter five:** The simulation results extract is clarified also provides a proposal for some educational points for future work.

CHAPTER TWO

Seizure Definition and Automatic Detection of Epileptic Seizure

2.1. Introduction:

This chapter presents the seizure and their types also would be explained how the brain activity measure by EEG and iEEG system and how to be used these measured data to detect seizures automatically by using machine learning techniques. The first part of the system that is used for seizure detection is preprocessing stage that is used for noise and artifacts reduction and some other processing on information EEG and iEEG signals such as windowing the signals. The second part is feature extraction. In this section, the features extracted from the input signal trying to give more information about normal and seizure signals for the classifier. There are many features that could be extracted from EEG and iEEG signals; Basically, this chapter focused only on features used in this thesis. There are many classifiers used for seizure detection in this chapter that focus on classifiers used in our study, which are artificial neural network (ANN), adaptive-network-based fuzzy inference system (ANFIS), decision tree (DT) and random forest (RF), also explain the parameters that are used for evaluating the detection methods. Finally in this chapter will explain the ways used for measuring the importance of the features that are used and select the features without effect on the system performance.

2.2.Types of seizure

Epilepsy is a term that refers to a variety of abnormalities that have a variety of causes, symptoms, and clinical manifestations. Individuals may have one or more types of seizures. Neurologists classify epileptic seizures into two broad categories: partial seizures, which affect only a portion of the brain, and generalized seizures, which affect the entire brain. Figure 2.1 illustrates the various seizure types.

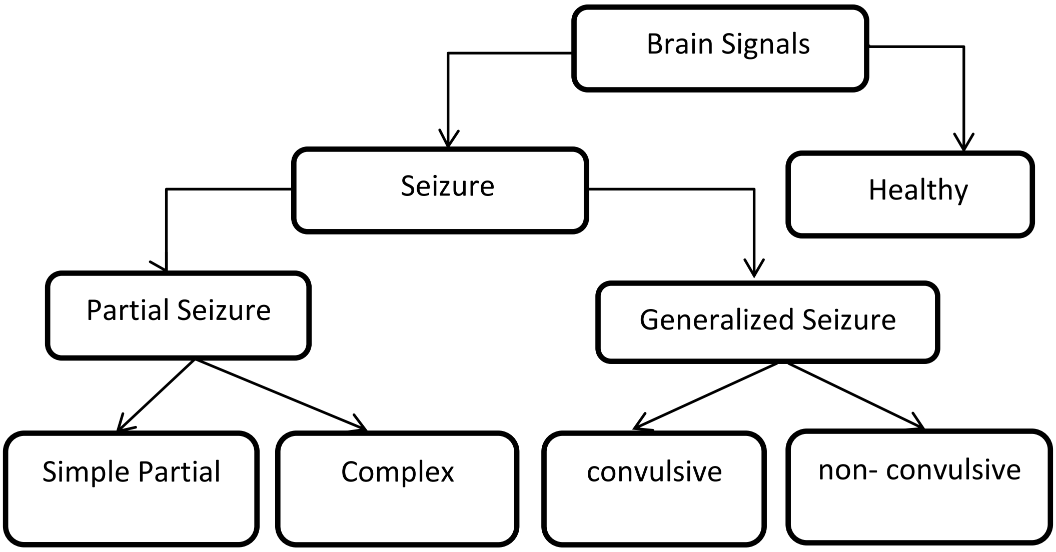


Figure 2.1. The illustrated diagram explain the types of seizures.

1.Partial seizure: This is also known as a focal seizure that occurs in a specific region in the brain; that would affect a specific region of the brain. The abnormal seizure activity begins in one region of the brain and then moves to another, or it may remain stationary in one area. In general patients are aware of their surroundings and maintain the same level of consciousness during this type of seizure. Additionally, partial seizures are classified into two types.

Simple Partial Seizure: Patients experiencing a simple partial seizure are aware of their surroundings during the seizure event. However, the patient will be unable to speak or move during the seizure. A simple partial seizure affects a specific area of the brain, and the patient may be able to move uncontrollably. In comparison to simple partial seizures, complex partial seizures affect a larger portion of the brain and some time affect consciousness. Although complex partial seizures can occur in any region of the brain, they typically occur in a specific region – the two temporal lobes.

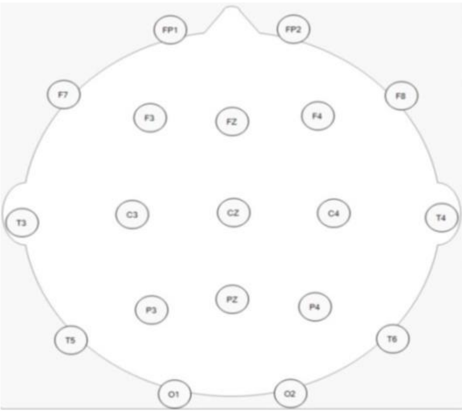
2.Generalized seizure: covers the whole region of the brain due to abnormal activity happening in both hemispheres (left and right) of the brain, and the patient loses the consciousness. A generalized seizure is of two types, 'generalized convulsive' and 'generalized non- convulsive' [20].

2.3.Electroencephalogram (EEG and iEEG)

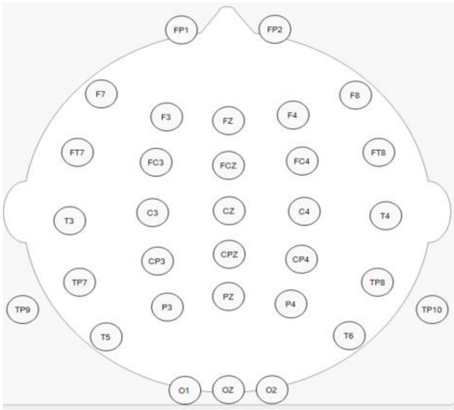
Generally, there are two types of the brain signals illustrated as:

2.3.1. Electroencephalogram (EEG) :

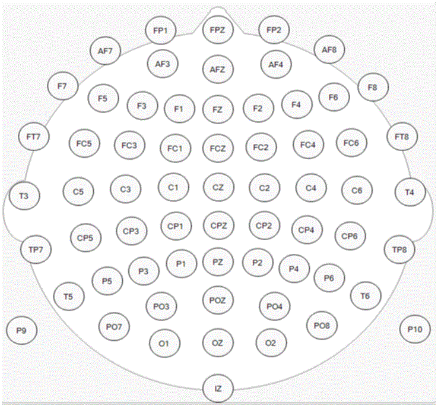
In this type, the signals are recorded using electrodes placed on the scalp to detect differences in electric potential. Hans Berger, a German psychiatrist, made the first EEG recording of a human brain's electrical signals in Jena in 1924 [21]. There are numerous electrode montages available, depending upon the number of electrodes used and their plan distribution on the scalp. Figure 2.2 illustrates several standard montages. All of these electrodes determine the potential difference between them and a reference electrode. This reference electrode is placed near the earlobes, nose, mastoid, chin, neck, or scalp center in close proximity to the other electrodes. Prior to performing any source imaging, the data acquired using this method has to be transformed to a common average reference montage by subtracting the average activity of all remaining channels from each channel. A set of electrodes, a differential amplifier, an analogue to digital converter (ADC), and a recording computer are required to record EEG [22]. To avoid aliasing, a low pass filter is required, and a high pass filter to eliminate low-frequency artifacts. The ADC is required to digitize the EEG signals. Figure 2.3 depicts the recording setup.



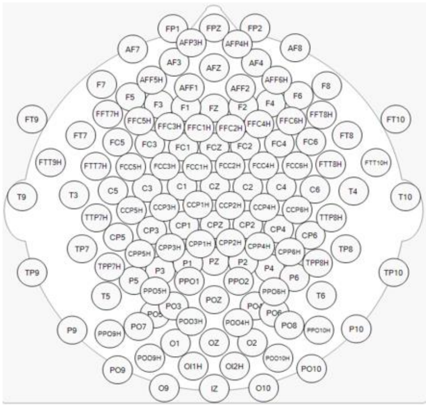
a) The montage with 19 electrode



b) The montage with 32 electrode



c) The montage with 64 electrode



d) The montage with 256 electrode

Figure 2.2. Examples of montages that use for EEG measurement (a) the montage that used 19 electrodes(b)the montage with 32 electrodes (c) with 64 electrodes (d) with 256 electrodes that putting on the scalp.

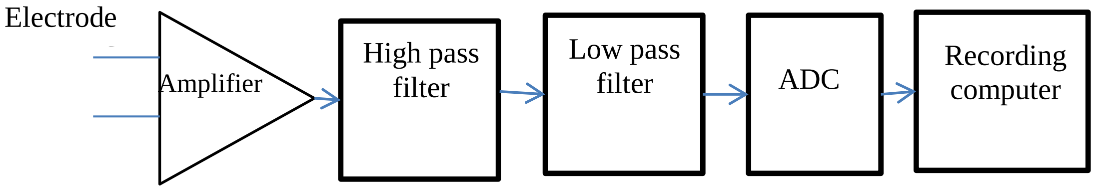


Figure 2.3. The block diagram of the EEG recording system[23]

Electroencephalography (EEG) is extremely useful in diagnosing a variety of neurological disorders. An EEG is a non-invasive system that detects signals via multiple electrodes/channels, resulting in the recording of millions of cortical neurons. The EEG is used for visualizing the brain's electrical signal, which aids in the diagnosis of brain diseases. The electrodes are symmetrically implanted on the brain's surface. The 10-20 systems [24] are typically used to record spontaneous EEG signals.

2.3.2. Internal Electroencephalogram (iEEG):

Another technique for recording brain signals is the internal electrocorticogram (iEEG). The primary distinction from EEG is that the channels or electrodes are invasively placed on the cortex. In this case, invasive means that the brain have to be opened in order to implant the grid. However, the beauty is that the cortex sustains no physical damage [25]. The disadvantage of this method is that if proper precautions are not taken, the infection may occur.

In general iEEG signals have a higher signal to noise ratio compared with EEG signals where the electrical brain active signals attenuate by the skull in case EEG signals [26].

2.4. Automatic seizure detection

Automatic seizure detection is critical from the expert's perspective, as is searching for seizures in a lengthy signal. The most frequently used method for automatically detecting seizures is through the use of an EEG (Electroencephalogram) signal, which contains a wealth of information about brain activity at a high resolution [27]. The monitored brain signals that are non-stationary, complex, and time-variant and the dataset is largely due to the hourly recording of the brain signals[12]. However, it is difficult to accurately detect it in a short period of time, and there is a high risk of missing an actual seizure [28].

Despite of these efforts, there is still a need to learn how to efficiently recognize a seizure based upon statistical data, as well as how to detect seizures utilizing strategies that aid in speedy seizure detection, allowing us to take the appropriate remedial activities [29]. Numerous automated methods for diagnosing epilepsy have been proposed recently [15]-[30]. The majority of automatic seizure detection algorithms made use of features extracted from EEG and iEEG signals. The features extracted from time-series EEG or iEEG signals are fed into the classifier, and the classifier outputs two values to classify the input signals as seizure or non-seizure, as illustrated in Figure 2.4.

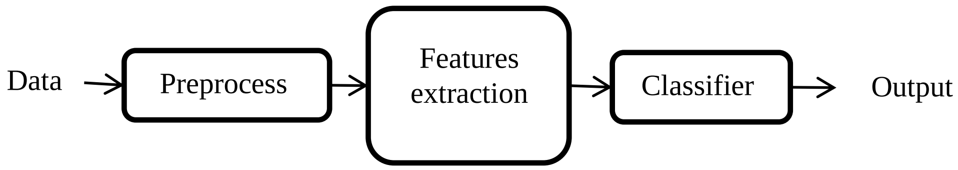


Figure 2.4. The block diagram of the system that used for automatic seizure detection.

2.4.1. Datasets

Two publicly available datasets were used in this study for experimental purposes. The first set (Bonn dataset) of data that was analyzed in this study consists of five sets (A, B, C, D and E) This data was downloaded from the website of the University of California, Irvine's School of Information and Computer Sciences, and the original signal was obtained from the University Hospital of Bonn. This data is used by a large number of researchers [14]-[31]. Each set of data contains 100 signals; each signal is divided into 23 chunks, resulting in a total of 11500 chunks; each chunk contains 179 values; 178 represents the EEG signal value; thus, the datasets contain 2047000 values; and the final value in position (179) denotes the chunk type. The datasets A and B collected data from the surface of the scalp of healthy individuals with their eyes open (A) and closed (B). C, D, and E measured actual and interactive elliptic activities via interracial electrodes, while C recorded EEG activity from a healthy brain area. D was recorded in the area where the tumor was discovered, and E was recorded during seizure activity [3]. the summary of the Bonn dataset is shown in table 2.1.

Table 2.1. The summary of the Bonn dataset that used in this thesis

	Five healthy subjects		**Five epileptic subjects**		
	Set A	Set B	Set C	Set D	Set E
Patient state	Eyes open	Eyes close	Inter ictal (seizure free)	Inter ictal (seizure free)	Ictal (seizure activity)
Electrode type	Surface (Scalp)	Surface (Scalp)	Intracranial	Intracranial	Intracranial
Electrode Position	International 10-20	International 10-20	Healthy area	Tumor area	Epileptogenic zone

The sample time series signals of all type of Bonn dataset shown in figure 2.5.

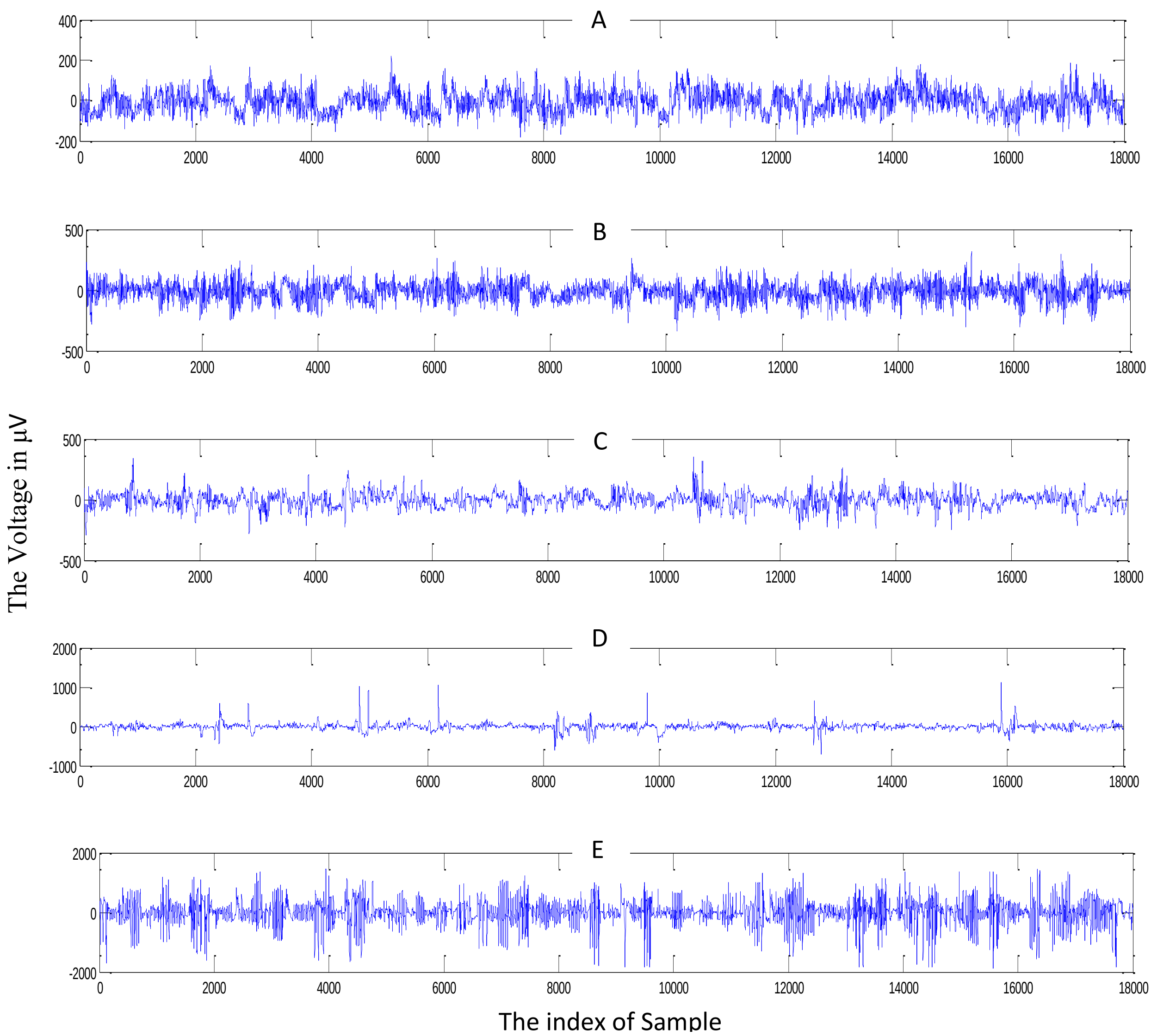


Figure 2.5. The sample time series signals for Bonn dataset for all cases (A, B, C, D and E)

The second set of data was extracted from the CHB-MIT Scalp. EEG Database recordings from 22 subjects (5 males, ages 3–22; and 17 females, ages 1.5–19) were grouped into 23 cases. (Case chb21 was obtained from the same female subject 1.5 years after case chb01.

Each case (chb01, chb02, and so on) contains between nine and forty-two continuous time-series signals files type European Data Format (edf) pertaining to a single subject. In state of arts many researcher observed that hardware has limitations, gaps between consecutively numbered edf files were created during which no signals were recorded; the gaps are typically 10 seconds or less but can occasionally be much longer. To safeguard the subjects' privacy, all protected health information (PHI) is contained in the original .edf files have been replaced with surrogate data in the files provided online. The edf files typically contain exactly one hour of digitized EEG signals, although those for case chb10 are two hours long and those for cases chb04, chb06, chb07, chb09, and chb23 are four hours long; files containing seizures are occasionally shorter.

All signals were sampled at a rate of 256 samples per second and with a resolution of 16 bits. The majority of files contain 23 electroencephalograms (EEG) signals (24 or 26 in a few cases). These recordings were made using the International 10-20 system of EEG electrode positions and nomenclature. Other signals, such as an ECG signal in the last 36 files belonging to case chb04 and a vagal nerve stimulus (VNS) signal in the last 18 files belonging to case chb09, are also recorded in a few records. In some cases, up to five "random" signals (designated "-") were interspersed between the EEG signals to create an easily readable display format; these dummy signals can be ignored [32]. The summary of the second dataset is shown in table 2.2.

Table 2.2. The summary of the CHB-MIT dataset that used in this thesis

Characteristic	**Details**
Number of patients	22 (5 males and 17 female)
Ages	ages 3–22
Recorded system	International 10-20
Sampling rate	256 samples per second
Number of electrodes	23, 28 and 32

2.5.Preprocessing

The EEG is designed to record cerebral activity, it can also record physiologic or extra-physiologic artifacts [33]. During the preprocessing stage, the Noise and artifacts present in the EEG signals identified. This stage is important to minimize their impact of carputed noise and the artifacts on the feature extraction stage that affected the classification result. In some studies, filters are used to reduce the noise and artifacts effect [34]. Additionally, the EEG signals were windowed in this step to prepare data for the next step which is extract features that need many samples for calculate the feature values[35].

2.5.1. Feature extraction

In general, many types of features are used; for instance, statistical values are extracted from the time domain of EEG signals [14]. Other studies have performed Fourier spectral analysis to derive EEG signals [36]. A typical short-time Fourier transform (STFT) method has been developed; in STFT, a window can change in time to measure the spectral density of EEG signals [37] [38]. For EEG signal processing, wavelet transform approaches for time-frequency estimation is typically desirable. For example, a discrete wavelet transform (DWT) technique is a classical method of time-frequency analysis similar to short-time Fourier transform, and it has been used to derive features from EEG signals [11] [39]. Some studies applied the DWT-based EEG de-noising method and feature extraction to identify seizures and diagnose epilepsy [40].

In general the features can be categorized as shown in figure 2.6.

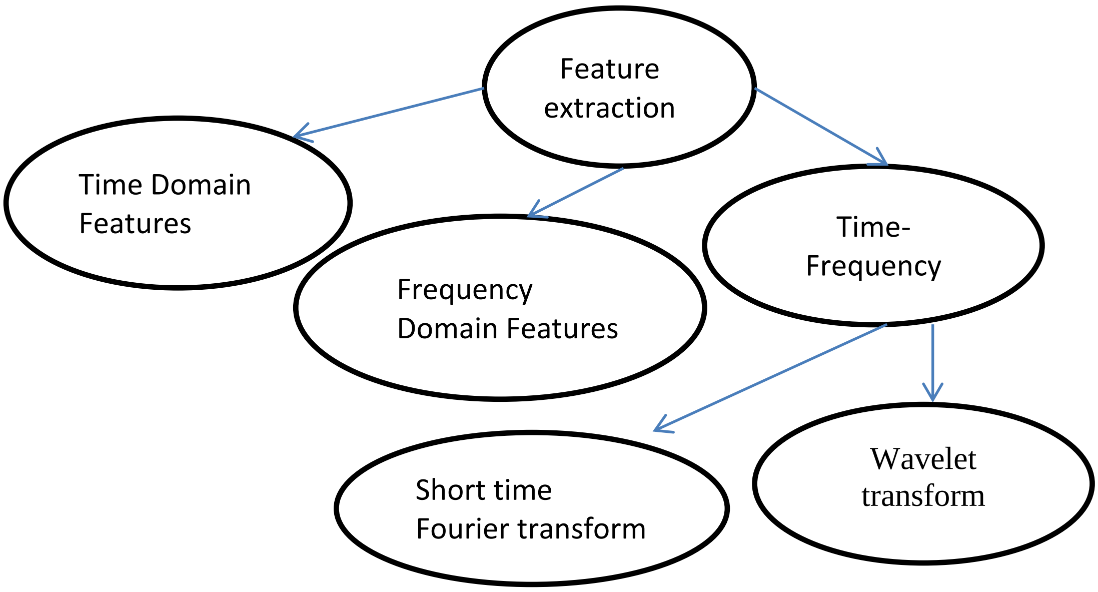


Figure 2.6. The general types of features extraction

Following describe the feature that used in this thesis.

A- Sliding Discrete Fourier Transform(SDFT):

SDFT was introduced by Jacobsen and Lyons as a DSP method that requires fewer computations for real-time spectral analysis and produces results on a sample-by-sample basis with the spectral bin output rate equal to the input data rate [41].

SDFT is applied to a fixed-length window of the signal in the SDFT scenario. Considering a complex input signal x (n), where n equals 0,1,2,..., which will be partitioned into M overlapping windows. Let k be the frequency-domain index in the range $0 \leq k < M$. Then, at time index n, the kth bin of an M-point DFT is computed as follows [42].

$$X_n(k) = \sum_{m=1}^{M-1} x(\hat{n}+m)\, W_M^{-km} \quad (2.1)$$

Where m is the index, n is the time domain index, M is the length of the window, $\hat{n} = n - M + 1$, $W_M^{-km} = e^{-j2pikm/M}$

Equation 1 can be rewritten according to the circular shift property as following:

$$X_n(k) = W_M^{-km}(X_{n-1}(k) + x(n) - x(n-M)) \quad (2.2)$$

Where $W_M^{k+M} = W_M^k$ because it is a periodic signal.

From equation 2.2, the outputs of SDFT will update after each input sample dependent on the previous value of output from the SDFT and the present value of time series; this technique can be built by Infinite Impulse Response (IIR) figure 2.7.

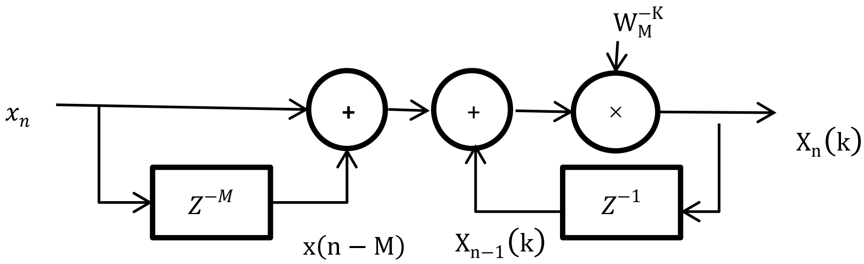


Figure 2.7. The structure of the IIR filter that can use to calculating SDFT from the EEG signals

B- Time-domain feature extraction:

1- **The first derivative of the EEG signals**: For discrete signal, the derivative is approximately equal to the difference between the current sample value with the previous sample value

$$x'(n) = x(n) - x(n-1) \quad (2.3)$$

2-**The second derivative of the EEG signals:** is the derivative of the first derivative

$$x''(n) = x'(n) - x'(n-1) \quad (2.4)$$

3- **The mean of the first, second derivative of the EEG signals with time window**.

4- **Zero crossing of the first and second derivatives in the duration window**: It is the number of the first, second derivative of data that crossing with zero.

5- **Number of local min and local max in the duration window**: It is the summation of positive peaks for the signal and for signals multiply by -1

6- **Line Length or Curve Length**: It is calculated by adding the lengths of consecutive data samples taken from a given signal. It is widely used for detecting and determining unequivocal seizure onset by observing changes in the EEG's amplitude, frequency, and dimensionality [1]

$$L(n) = \frac{1}{k}\sum_{K=n-N}^{n} abs[x(k-1) - x(k)] \tag{2.5}$$

8- **Hjorth parameters (Activity, Mobility, Complexity)**: The Hjorth parameter is one way of expressing a signal's statistical properties in the time domain. It consists of three types of parameters: Activity, Mobility and complicity. These three parameters provide information about a signal's frequency spectrum, and they also aid in the analysis of signals in the time domain. Additionally, their use results in a reduction in computational complexity [43].

$$\text{Activity} = \text{var}\big(y(t)\big) \tag{2.6}$$

$$\text{Mobility} = \sqrt{\frac{\text{var}\big(y'(t)\big)}{\text{var}\big(y(t)\big)}} \tag{2.7}$$

$$\text{Complexity} = \frac{\text{mobility}\big(y'(t)\big)}{\text{mobility}\big(y(t)\big)} \tag{2.8}$$

9- **Maximum and minimum**: The two values that are calculated for each window by taken the maximum and minimum values of samples in the window.

10- **The energy duration of the window**: It is the summation of the square of each value in the window divided by the length of the window.

$$En = \frac{\sum_{n=1}^{k} x^2(n)}{K} \quad (2.9)$$

2.5.2. Machine learning

Machine learning is the method that can be learned from the data. Machine learning can be used when in need to solve complex problems where no or very difficult solutions are available with a traditional approach. Machine learning can be categorized into four types dependent on human supervision during the training[44] as shown in figure 2.8.

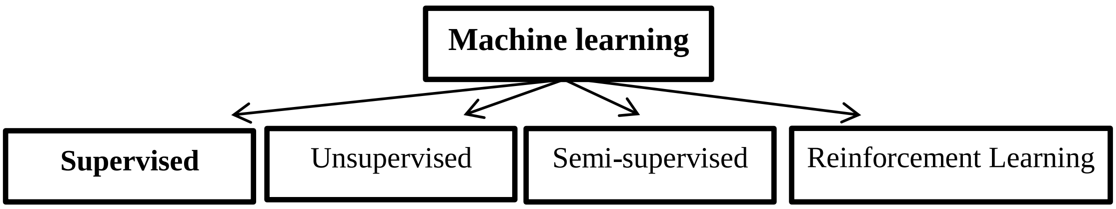


Figure 2.8. The machine learning categories and its required

Supervised Learning: is a prediction learning model, given an unforeseen input instance. To learn the model, a supervised learning algorithm uses a known set of input datasets and targets. The model is then trained with a learning algorithm to generate a forecast for the reaction to new data.

Unsupervised Learningthe: The goal in unsupervised learning the is to find the regularities in the input such that certain pattern occur more often than others.

Semi-supervised Learning: is a combination of supervised and unsupervised learning that uses the labeled knowledge set to classify certain unlabeled data. It's employed when there aren't enough labeled data for a particular application.

Reinforcement Learning: Interacting with the environment is a necessary part of reinforcement learning. Reinforcement learning was developed to answer the problem of how an autonomous agent that detects and acts in its environment may learn to select the best actions to accomplish its objectives. The activities taken by an agent in the environment are used to reward its behavior. It considers the consequences of its actions and takes optimal steps further.

In this thesis the supervised method used for signals classification. Due to the fact that there is no single classification method that is applicable to all subjects and applications. It is advantageous to test ensemble classification methods. Classifiers are typically used to refer to the algorithms used to construct the training model for predictive analysis. These are further classified into two broad categories: 'black-box' [45] systems and 'non-black-box ' systems [46]. The term 'black-box' refers to classifiers that are incapable of providing human-comprehensible reasoning (e.g., logic rules) for their classification method. They generate ambiguous and difficult-to-understand logic rules or patterns. Neural network (NN) is examples of black-box classifiers. The neural network contain three major layers input, hidden and output layer. For simple NN contain one hidden layer as shown in figure 2.9.

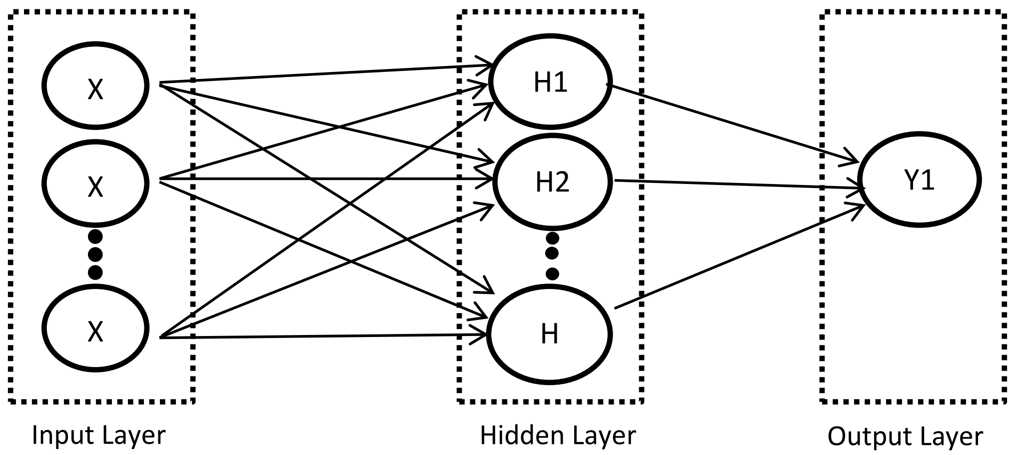


Figure 2.9. Structure of a feedforward NN classifier .

Deep neural network (DNN) is a neural network that has more than one hidden layer.

Logic rules in non-black-box classifiers, on the other hand, are straightforward and easy to comprehend because they are based on if-else conditions. As a result, they are referred to as 'non-black-box classification algorithms. Decision trees and Random Forests are examples of non-black-box classifiers [47].

In this study, four classifiers are used: a artificial neural network (ANN) with Levenberg-Marquardt optimization was used as the first classifier, an adaptive network-based fuzzy inference system (ANFIS) was used as the second classifier, a decision tree classifier was used as the third classifier and fourth classifier used was Random forest (RF).

A- Artifical Neural network

The artifical neural network (ANN) with Levenberg-Marquardt optimization: The neural network (feedforward net) is used as a classifier with three layers, the hidden layer contain 20 nodes.

In feedforward each, the output from node j in the hidden layer is a function of sums the multiplying its input x_j signals with weight w_{ij} [1].

$$output_j = s\left(\sum w_{ij}\, x_j\right) \qquad (2.10)$$

Where s is a sigmoid activation function.

$$s(y_i) = \frac{1}{1+e^{-y_i}} \qquad (2.11)$$

Livenberg–Marquardt optimization method used when required minimization of non-linear function that is used in the training schemes. The result from the system is presented in two values: zero for normal state and one for seizure state. The methods build by using MATLAB R2012b. The main disadvantage of ANN is that it is a 'black-box' this mean the internal processing is not understandable to data analyst.

B- Adaptive network-based fuzzy inference system (ANFIS)

An adaptive network-based fuzzy inference system (ANFIS) is a network with five layers of a feedforward neural network with a supervised learning capability [17]. ANFIS is deep learning and its structure depend on fuzzy and adaptive control figure. 2.10.

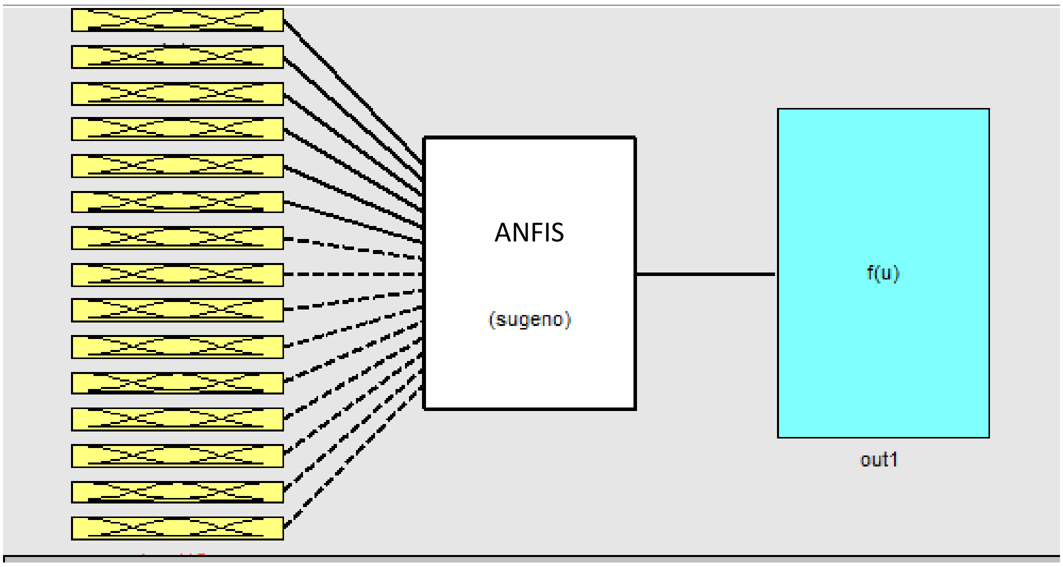


Figure 2.10. ANFIS model structure when fifteen features extracted from the EEG are input to ANFIS

In ANFIS first, the fuzzification method applies by using the membership for the input. ANFIS is dependent on the Sugeno model where fuzzy if-then rules generate the stipulated input-output pairs. The output from the classifier is calculated by summation the results from fuzzification and relations steps as follows :

$$y = \frac{\sum_{i=1}^{2} w_i f_i}{w_1 + w_2} \tag{2.12}$$

where w_i is the result from the rules stage and f_i is the adaptive parameter dependent on the input and the membership.

C- Decision Tree :

For diverse domain datasets, decision trees provide logic rules that aid in the discovery of knowledge (Fayyad et al., 1996).

In this case, the classification processes are fairly clear and simple. These categorization processes are given in the form of a flow chart structure based on an if-then condition, which makes the findings human-readable and allows for the study of logic rules/patterns inherent in a data collection. Not only that, but a decision tree can handle both continuous and categorical data with high dimensions (Han et al., 2011). The decision tree structure is made up of three main elements: nodes, leaves, and lines. As illustrated in Figure 2.11.

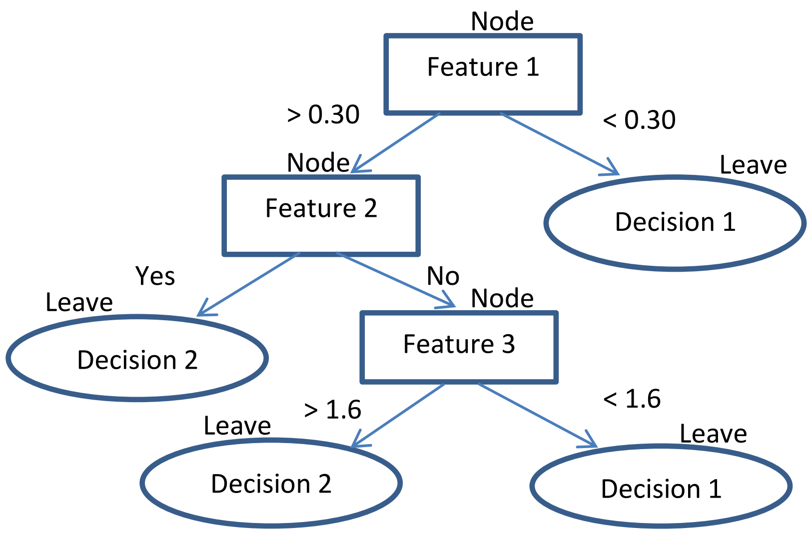


Figure 2.11. The example of decision tree structure

Several induction techniques are developed to form a decision tree, each of which is based on a different criterion for measuring characteristics such as impurity, distance, and so on. Information gain, gain ratio, and Gini index is among the impurity measurements considered (Adnan, 2017, Han et al., 2011). Shannon's 'information theory, often known as Entropy, is the foundation for these metrics (Han et al., 2011). In mathematic, the Gini impurity can be calculated as the following equation

$$G = 1 - \sum {p_i}^2 \tag{2.13}$$

where p_i is the probability of each state (normal and seizure)

d- Random Forest:

A Random Forest is a machine learning technique structured from multiple trees. Each tree used subset random samples from all input set samples with the same distribution. Each tree calculates the result independent of the other tree.

The final result is dependent on the results calculated from each tree where the popular class result is chosen by the vote for all results from the trees as the final RF result. [48].

2.6. Prediction of Performance Indices

Three statistical parameters, classification accuracy, sensitivity and specificity, have been calculated to measure the performance of the methods used. The definitions of these parameters are :

$$\textbf{Accuracy (AC)} = \frac{\textbf{correct classification patterns}}{\textbf{total patterns}} \times \mathbf{100}\% \tag{2.14}$$

$$\textbf{Sensitivity(SEN)} = \frac{\textbf{true positives}}{\textbf{true positives+false negatives}} \times \mathbf{100}\% \tag{2.15}$$

$$\textbf{Specificity (SPE)} = \frac{\textbf{true negatives}}{\textbf{true negatives+false positives}} \times \mathbf{100\%} \tag{2.16}$$

The values of true positive, true negative, false positive and false negative are dependent on the automatic system detection and the target as shown in table 2.3.

Table 2.3. The stat of the output from the classifier compared with the doctor's notes.

	Diseased	**Non- Diseased**
Test positive	True positive	False-positive
Test negative	False-negative	True negative

2.7.Feature importance

In machine learning, it is important to explain why a given model behaves the way that it does. One way to explain the behaviour of a model is to describe the importance or ranking for each features used in the classification. The ranking with identify a level can be used for features selection. There are many methods for feature ranking, but in this study, we will focus only on the two types of ranking from the models which are Permutation Importance and Mean Decrease in Impurity(MDI).

2.7.1. Permutation Importance

With the Permutation Importance feature selection method, the performance of a model is tested after removing each individual feature and replacing that feature with random noise. In this way, the importance of individual features can be directly compared, and a quantitative threshold can be used to determine feature inclusion [48].

2.7.2. Mean Decrease in Impurity(MDI)

It is another method for calculating the importance of each feature that is used for classification. With (MDI), the feature importance is calculated by calculating the sum of the number of splits on each tree in RF dependent on this feature, proportionally to the total number of samples that need splits. [49].

2.8.Feature Selection

In the feature selection method the features that are most relevant to the classification process are selected. This algorithm is important to reduce the dimensions of data used in classification. This lead to improving the computational efficiency by removing irrelevant features [50]. There are three different types of feature selection approaches filtering, wrapping and embedding. In the filter approach there is no learning algorithm used. The filter approach selects the features with the greatest ranks among them, and the subset can then be prepared for any classification algorithm. Because of its speed and scalability, the filter method is a common feature selection approach. By using a classifier algorithm as a black box, the wrapper technique calculates scores for feature sets that rely on estimated power. Testing and training on a given dataset are used to evaluate a subset. The space of all subset characteristics is gained by the wrapped search technique around the classifier. In embedded technique, feature selection is combined with the classifier's structure. The advantage of the embedded method is that it interacts with its categorization model and does not require complex calculation[51]. In this thesis the Sequential Feature Selection (SFS) algorithms are used which is a part of the wrapper methods where it adds and removes features from the dataset sequentially.

CHAPTER THREE

Methodology

3.1.Introduction:

In this study, two datasets were used. The first dataset was downloaded from the website of the School of Information and the Computer Sciences University of California, Irvine[52]. The original signal is taken from the University Hospital of Bonn, which contains EEG and iEEG and the second dataset was CHB-MIT from Hospital Boston, which contains scalp EEG signals. We will apply these two data in the designed models for automatic seizure detection, as shown in figure 3.1.

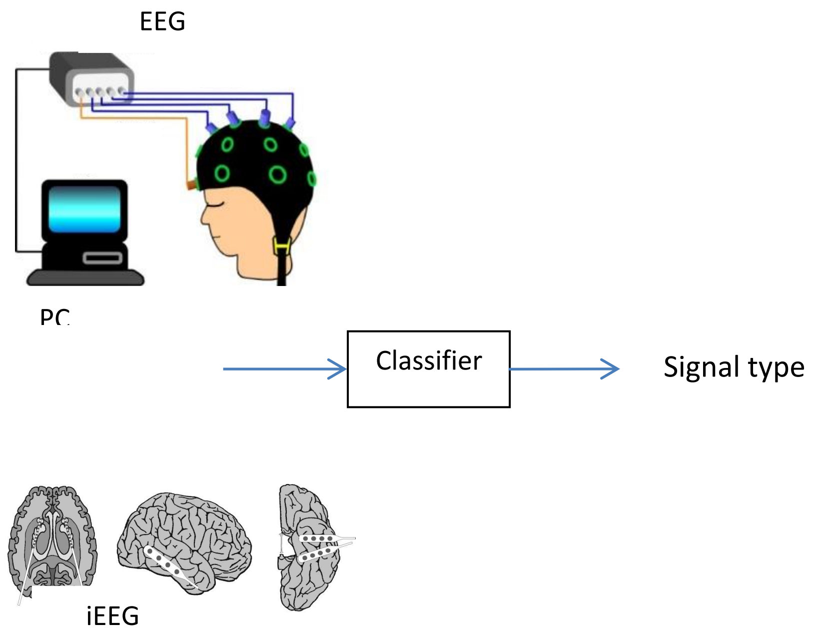


Figure 3.1. The general idea for automatic seizure detection.

3.2. Datasets

3.2.1. Bonn dataset

1. **Prepare the data:**

The file that downloaded for the first dataset was a CSV file containing 11500 rows for different measured types (A, B, C, D and E) each one has 100 measured and each measured contains 23 chunks. Each row in the CSV file contains 179 values. The first 178 values contain the data (2047000 value) of time series signals recorded, and the last value indicates the type of signal measured. First, each measured type of signal separated dependent on the type value on the file to five types (A, B, C, D, E) each type contains 409400 samples.

After separating, the target generated depended on the type of signals, normal state A, B, C, and D targeted value with zero's and seizure state E targeted value one Table 3.1.

Table 3.1. The five types in the data and the type number in CSV file in position 179, all types target with zero except type E target with one.

Type	**Type value in the file**	**Target**
A	5	0
B	4	0
C	3	0
D	2	0
E	1	1

Five different cases generated from separated types for study, as shown in Table 3.2.

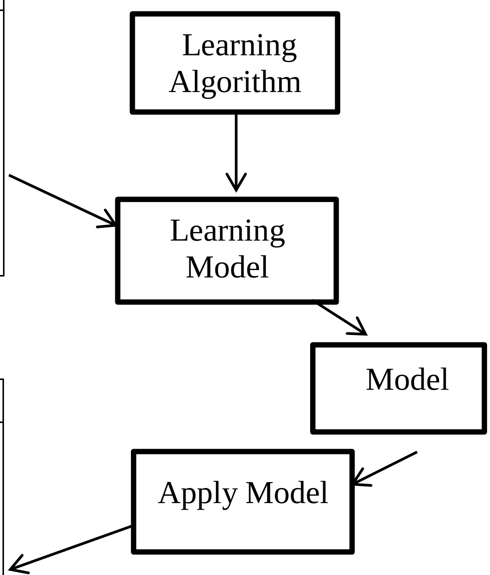
Table 3.2. The cases extracted from the dataset that used for classification.

Case	Description	No. of samples
A-E	Detection seizure from signals contain seizure and normal signal with open eyes	818800
ABCD-E	Detection seizure from signals contain all states	2047000
C-E	Detection seizure from signals contains C normal and E seizure.	818800
D-E	Detection seizure from signals contain D and E states	818800
CD-E	Detection seizure from signals contain C and E states	1228200

For each case from the five cases, 50% from the values and its targets used for training the classifier in the model for seizure detection and the other 50% used for testing the model[8][53], as shown in figure 3.2. the target for the testing set is used with the classifier output to evaluate the machine learning for seizure detection.

Training Set

No. of Samples Values	Target
178 Values	**0**
178 Values	**0**
178 Values	**1**
178 Values	**1**
178 Values	**1**
178 Values	**1**
………..	……



Testing Set

No. of Samples Values	Target
178 Values	**?**
178 Values	**?**
178 Values	**?**
178 Values	**?**
178 Values	**?**
178 Values	**?**
………..	……

Figure 3.2. General diagram for building a classification model.

2. Features that used with Bonn dataset:

There are two types of features used with this dataset as shown in table 3.3, The first type was the feature that needs no delay for calculating its value. The results values are equal to the number of samples in the dataset. The six features used in this thesis are extracted from the dataset in figure 3.3.

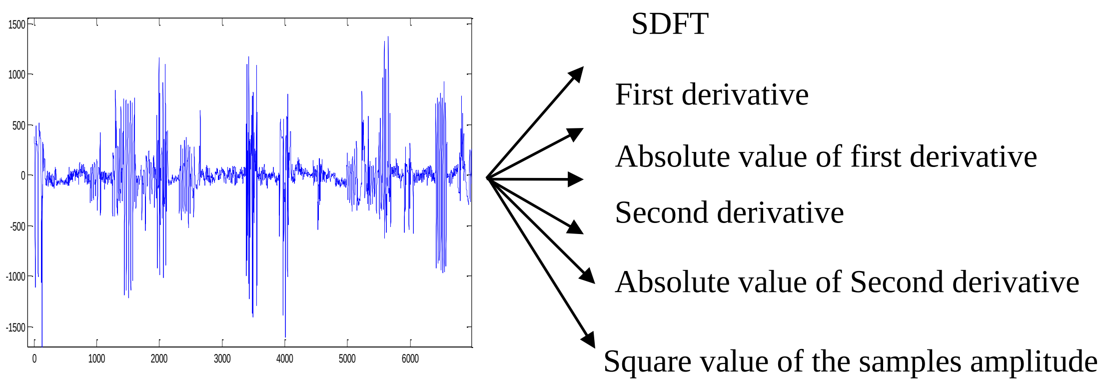


Figure 3.3. The six features extracted from the signal without delay

While the second type of feature was the feature that needs time delay for calculation, to get the one feature value many samples needed for this purpose window is used to select many samples for calculating the feature value for the samples inside the window, for this reason, the time delay needed depends on the length of the window. For this dataset, two windows length used the first one 178 samples per window and the second state was a window with 356 samples and each window state is applied for seizure detection.

Table 3.3. The features used with Bonn Dataset in all cases.

The Features	**Type of Features**
SDFT	No delay required
First Derivative	No delay required
Second Derivative	No delay required
Absolute of First Derivative	No delay required
Absolute of Second Derivative	No delay required
Squared of samples value	No delay required
Zero-Crossing for First Derivative	delay required
Zero-Crossing for Second Derivative	delay required
The No. of Peaks	delay required
Line Length	delay required
Activity	delay required
Mobility	delay required
Complexity	delay required
Maximum Value	delay required
Minimum Value	delay required
Energy	delay required
Mean of SDFT	delay required
Mean of First Derivative	delay required
Mean of Second Derivative	delay required
Mean of the absolute of First Derivative	delay required
Mean of the absolute of Second Derivative	delay required

3. Classifiers:

In all cases, four types of classifiers used ANN with Levenberg-Marquardt optimization with 20 nodes in the hidden layer and the second classifier was ANFIS, DT used as the third classifier and the fourth classifier was RF.

4. Methods :

Two methods with four classifiers used in each method that are applied to this dataset for all cases, as shown in figure 3.4.

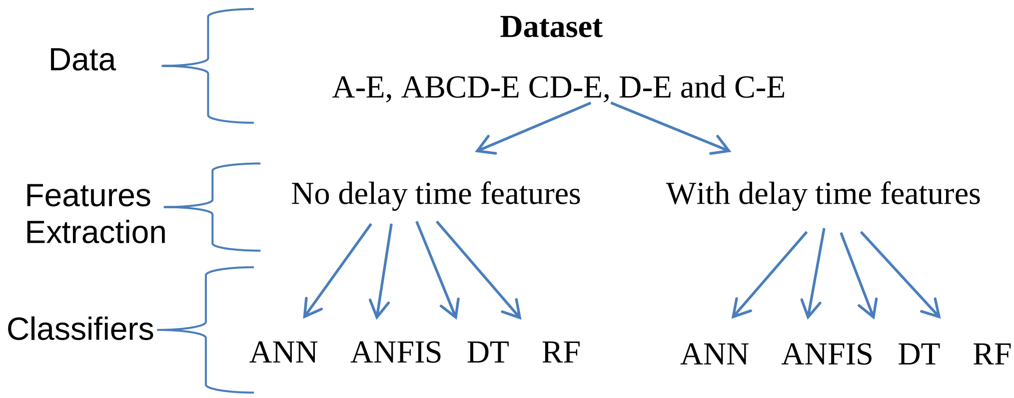


Figure 3.4. The features and classifier used with all cases.

- **The first method:**

In this method, the feature extraction from the time series of cases is calculated with no delay and then, these features are used as input to classifiers. The training dataset for the artificial neural network is 50% from the dataset figure 3.5 (a), the remaining dataset, which is 50% is used for testing the system figure 3.5 (b).

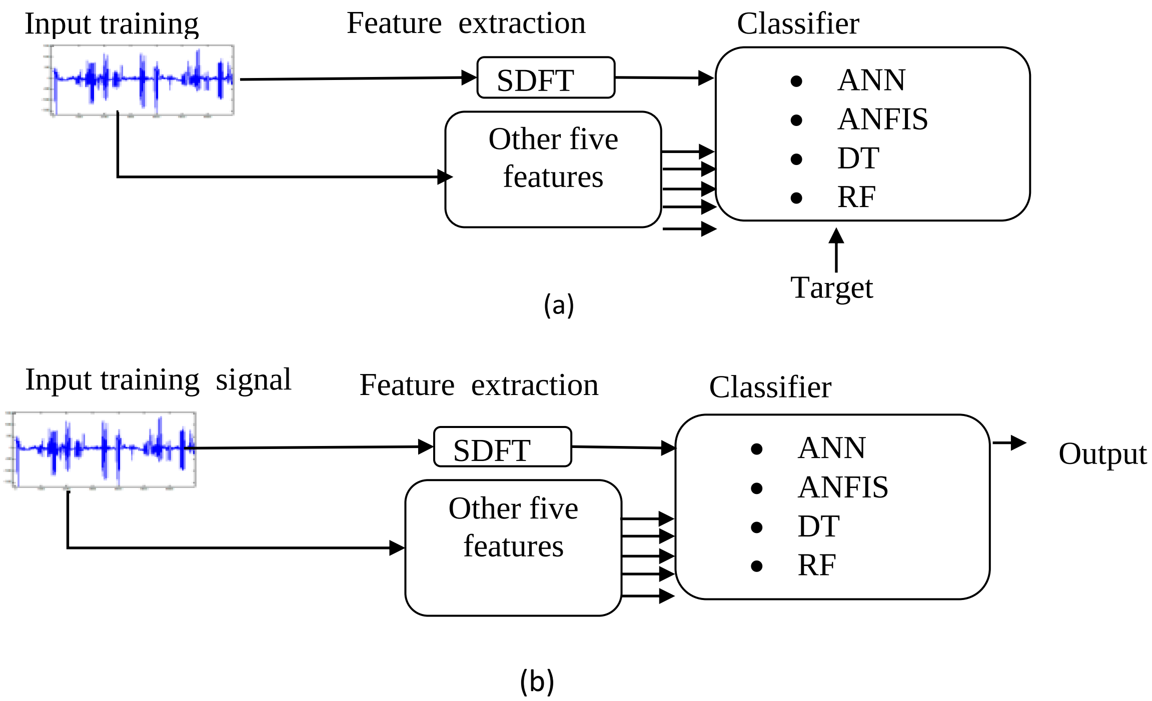


Figure 3.5. The first method structure used with the Bonn dataset is (a) the structure of the training phase (b) the testing phase.

The flow chart of method one explain in figure 3.6.

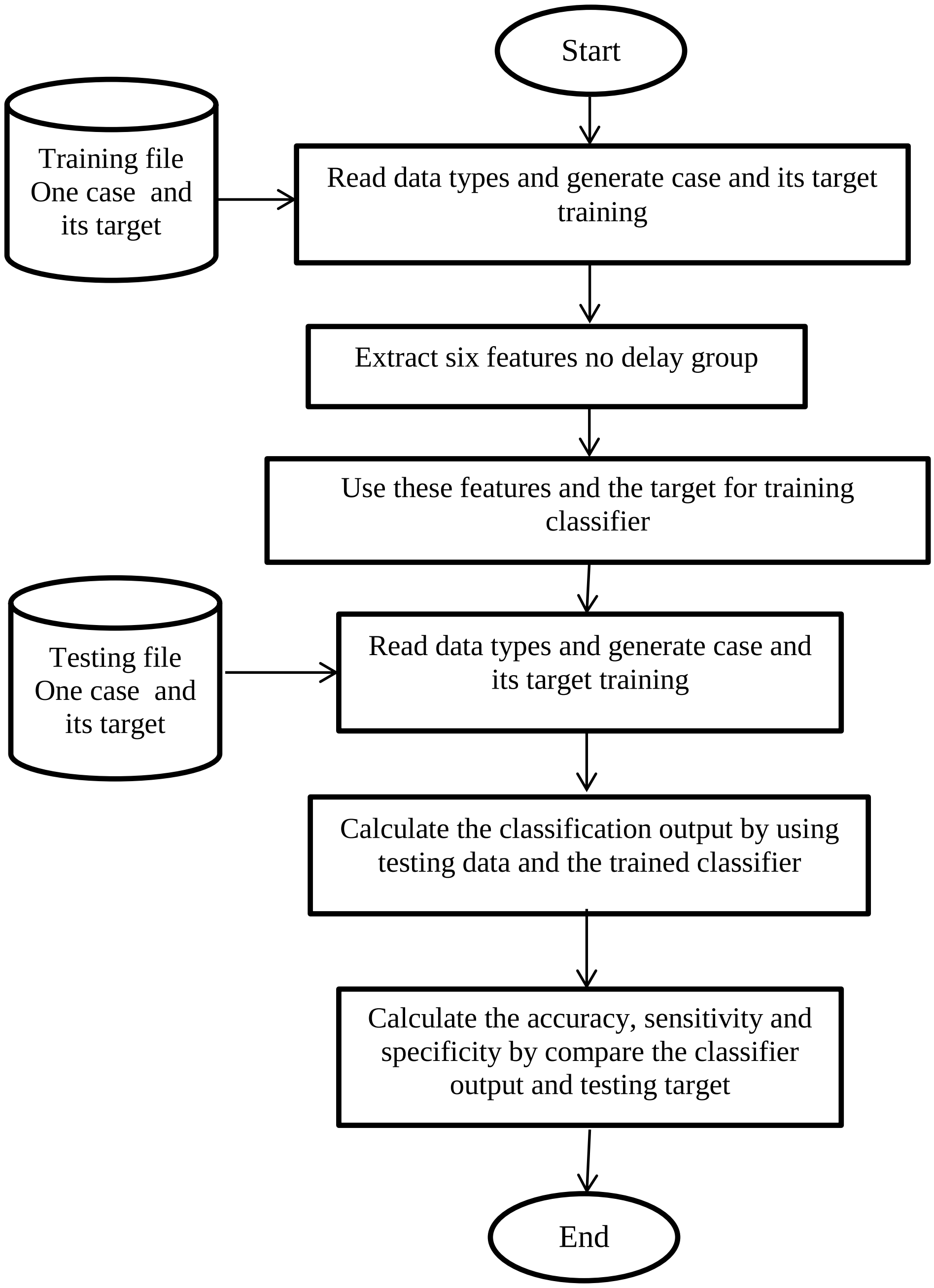
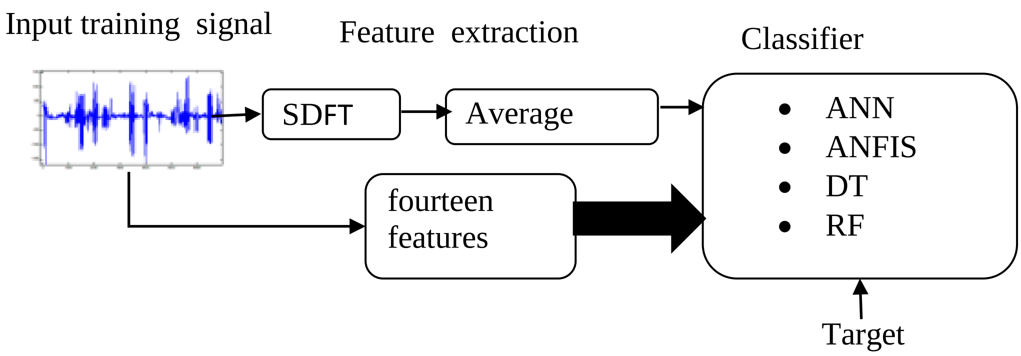


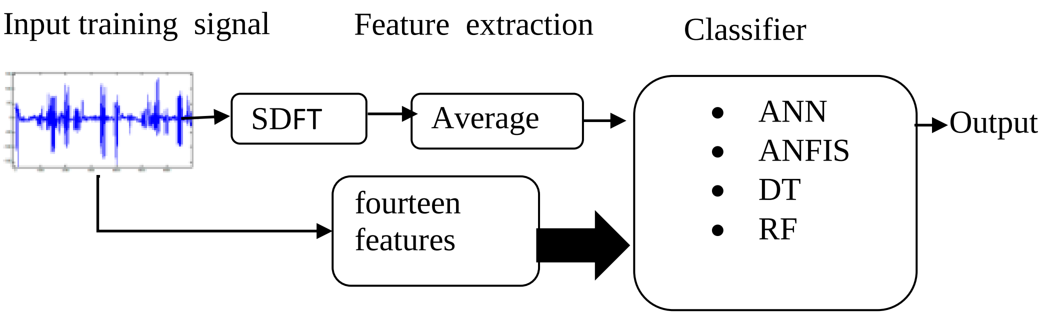
Figure 3.6. The flow chart of seizure detection from Bonn dataset when no delay feature extraction used.

- **The second method**:

In this method, the feature extraction from the time series of cases is calculated with delay and then these features are used as input to classifiers. The training dataset for the artificial neural network is 50% from the dataset, as shown in figure 3.7. (a), the remaining dataset, which is 50% is used for testing the system, as shown in figure 3.7 (b).



(a)



(b)

Figure 3.7. The second method structure that used with the first dataset case A-E and case ABCD-E (a) the structure of training phase (b) testing phase.

The flow chart of the second method is shown in figure 3.8

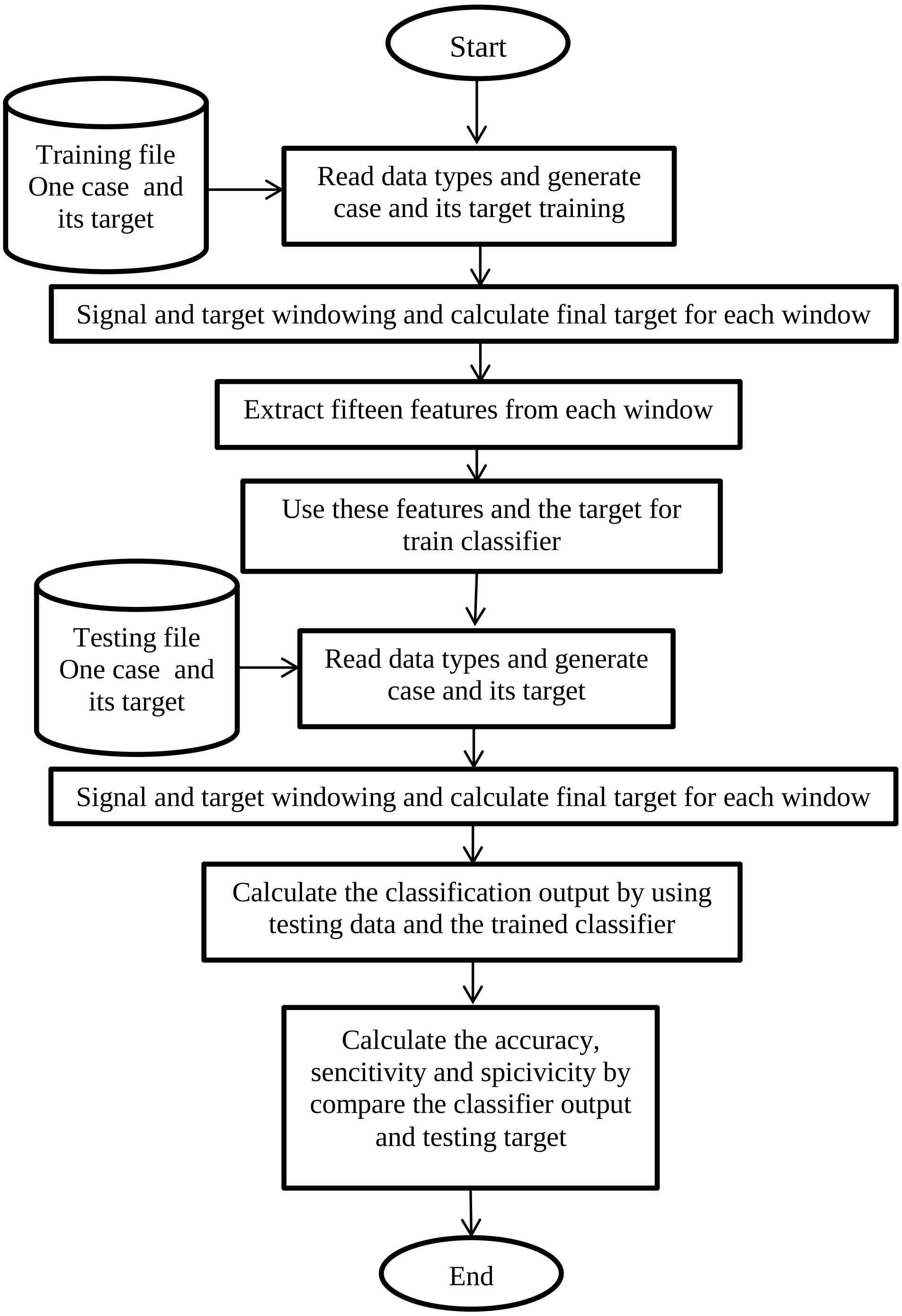


Figure 3.8. The flow chart of seizure detection from Bonn dataset when delay feature extraction used.

To choose the best value of M, which is the delay for the samples input to the IIR for SDFT calculation, many values of M (form 1 to 100) are applied for the SDFT computation step, and ANN is used as a classifier, then values of accuracies calculated for each value of M as shown in figure 3.9.

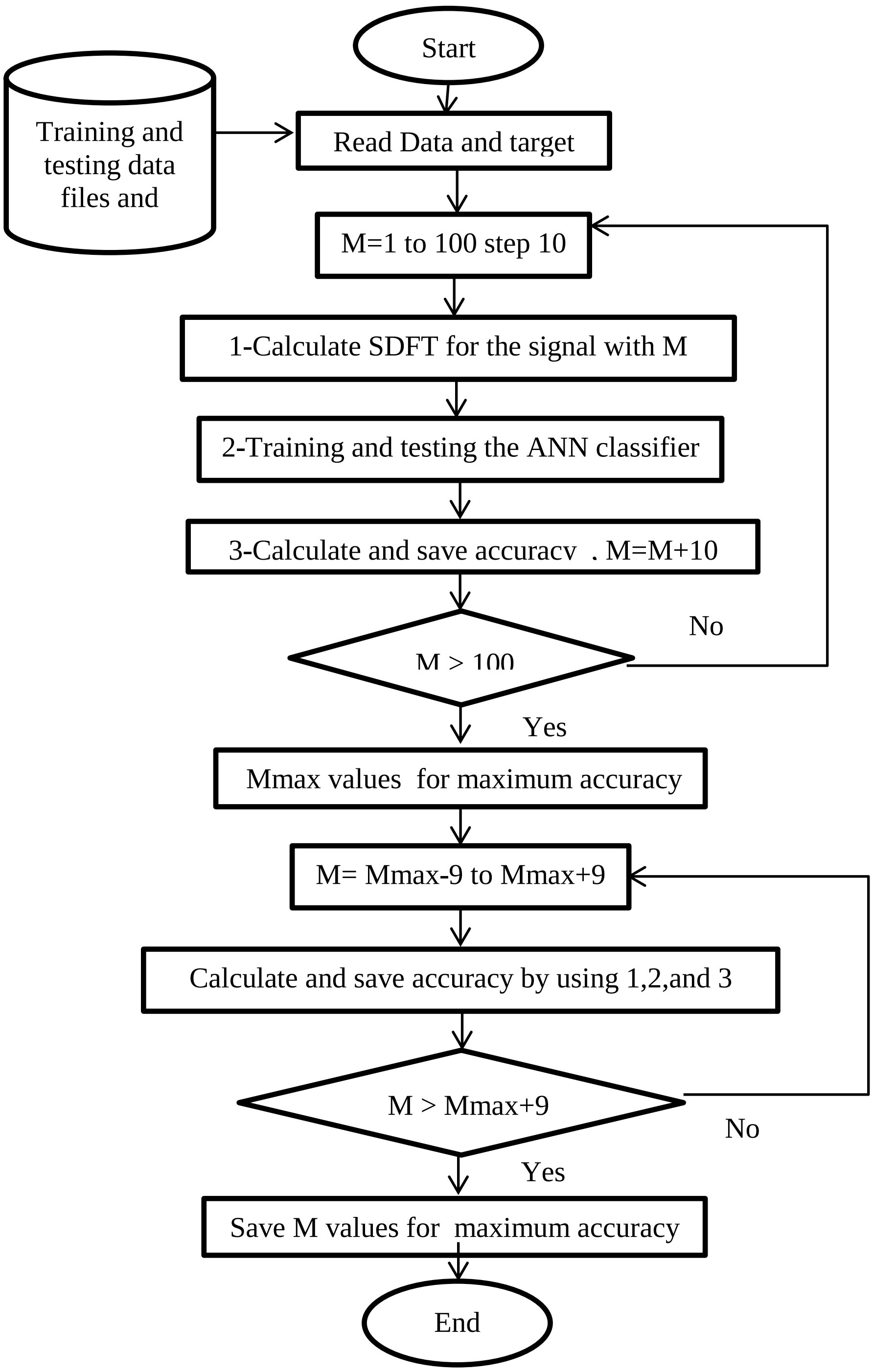


Figure 3.9. The flow chart of a method that used for calculating the best value of M

When applying the method explained in figure 3.6 the maximum value of M was 35 for case A-E as shown in figure 3.10.

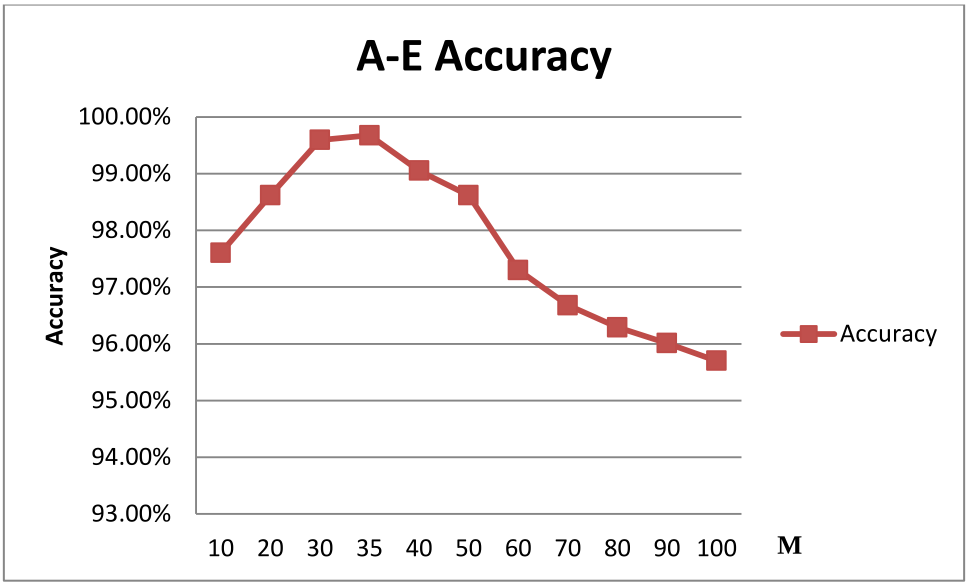


Figure 3.10. A–E case classification accuracies when SDFT is applied to EEG signals before ANN with different values of M, which is the forward delay in the IIR filter.

5. Feature importance

There are many methods that are used to measure the importance of each feature. This measure is important to know which features have more effect in the classification and also can be used for feature selection. In this study the two methods of selecting from the model used are Permutation Importance and Mean Decrease in Impurity (MDI) applied to measure the feature importance for each feature extracted. First, these methods are applied to the first group which are the six features to measure importance for each feature. Also, these two methods were applied for the second group the fifteen features to calculate the importance for each feature in this group.

6. Feature Selection

Feature selection is used to select only the feature that has more effect in classification and delete the features that have no effect or the features that reduce the accuracy of classification. The sequential feature selection (SFS) types forward are used in this thesis for this purpose. This method is a greedy procedure where, at each iteration, the SFS method choice is the best new feature to add to the selected features based on a cross-validation score.

3.2.2. CHB-MIT Dataset Prepare :

The second dataset used in this thesis are gotten from the link https//physionet.org/content/chbmit/1.0.0. The data described in section 2.4.1 from which ten patients have chosen five patients for the training phase and the other five patients for the testing phase, as shown in Table 3.4

Table 3.4. The information about the CHB-MIT dataset that used in this study.

Patient No.	Age	Dataset	Duration sec	Seizure duration sec	No. of channel	Purpose
chb02 M	11	16+ 16	3600 960	81 82	23 23	Train
chb05 F	7	6 17	3600 3600	115 120	23 23	
chb07F	14. 5	12 13	14400 3720	86 96	23 23	
chb18 F	18	35 36	3600 3600	68 46	28 28	
chb21 F	13	19 20	3600 3000	56 50	28 28	
chb01 F	11	3 4	3600 3600	40 27	23 23	Test
chb03 F	14	2 3	3600 3600	65 69	23	
chb11 F	12	92 99	3600 2210	32 752	28 28	
chb19 F	19	29 30	3600 3334	77 81	28 28	

chb23 F	6	9	14426	71 62 27 82	23	

From table 3.4, some EEG measured with montage 23 channel and other with montage 28 channel, the example of EEG for patients number two and dataset 16 from the time 00:02:00 to 00:02:40 shown in figure 3.11 the seizure started at 00:02:10 The EEG signal drawing by using the website link https://physionet.org/lightwave/?db=chbmit/1.0.0, first step the EEG signal with montage 28 channels convert to montage 23 channels by delete five channels and rearranged the remain channels.

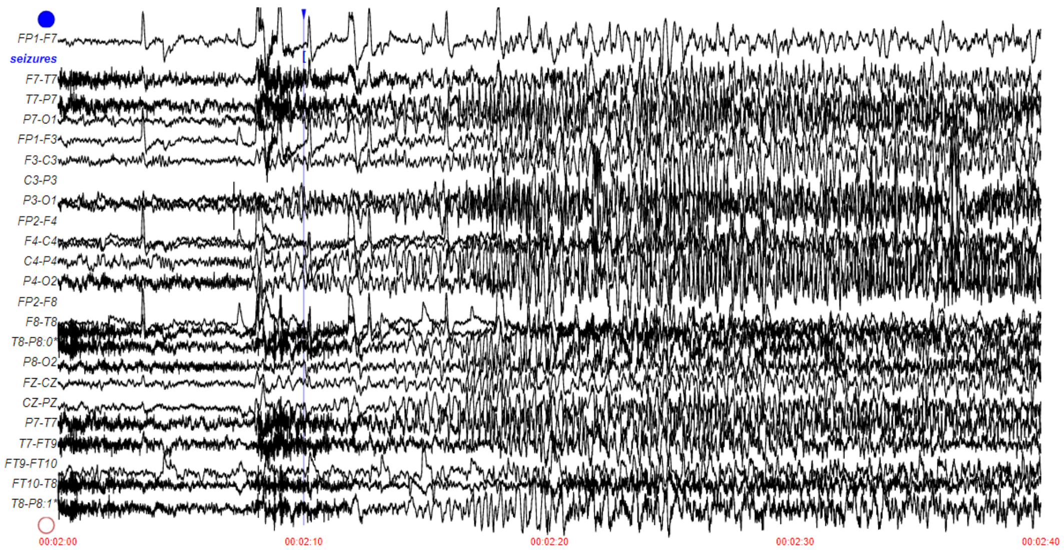


Figure 3.11. One example of part EEG signals measured from patient number two is the seizure start after 2 minutes and 10 seconds.

In the second step, two types of Butterworth filters applied bandpass filter 3 to 32 Hz more information about the seizure in this range [30] and notch filter for 60 Hz. After this step, the data for training was created by damage EEG data for five patients and in the same way for the testing phase from the other five patients. The target generated by zero array length equal to the EEG duration and then replaced some of them by value one dependent on seizure position and duration in information file about EEG recorded for each patient.

- **Methods used with CHB-MIT Data**:

There are two methods structure that are used, and in each method, three classifiers used TD, RF and ANN (with adam optimization). These two methods are :

1. **The first method:**

In the first method, the time series EEG signals that were prepared for training were converted to the frequency domain by using SDFT, then the result from SDFT was used as input to the classifier in the training phase with its target. In the test phase, the EEG signal that prepares to test is converted to the frequency domain by using SDFT and then the result from SDFT is used as input to a classifier, as shown in figure 3.12.

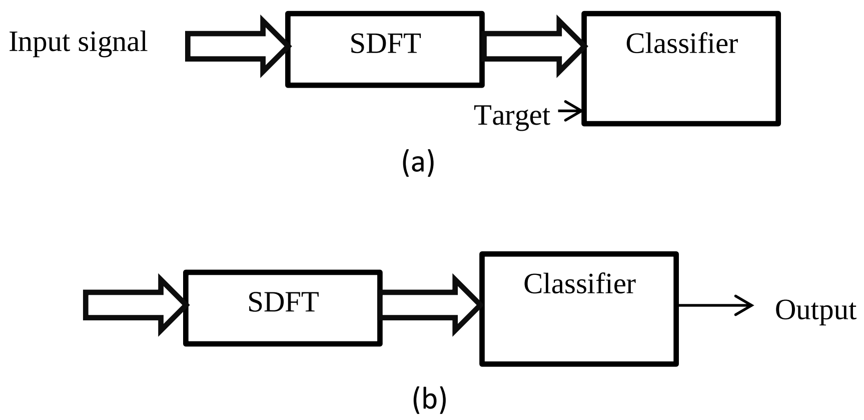


Figure 3.12. The first method structure used with a second dataset (a) the structure of the training phase (b) testing phase.

To calculate the best value of M (the delay of input in IIR filter), which is used for SDFT computation, many values of M choose from one to one hundred, and the output from SDFT inputted to the classifier and then calculated the accuracy in each value of M for tree decision classifier figure 3.13 the result shown the best value of accuracy was 85% achieved when the value of M was M=20.

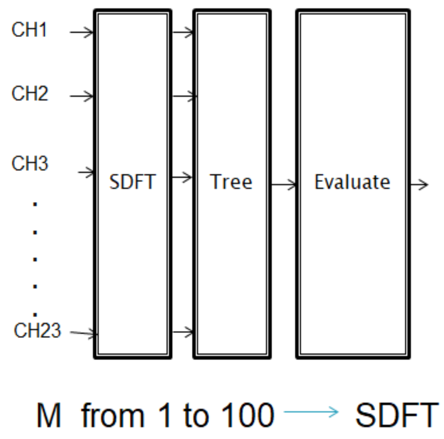


Figure 3.13. The method used to calculate the best value of M (the forward delay in the IIR filter) the value of M apply to all channels.

Another way used to compute the best value of M by calculating the best value of M in each channel of 23 channels separately, as shown in figure 3.14

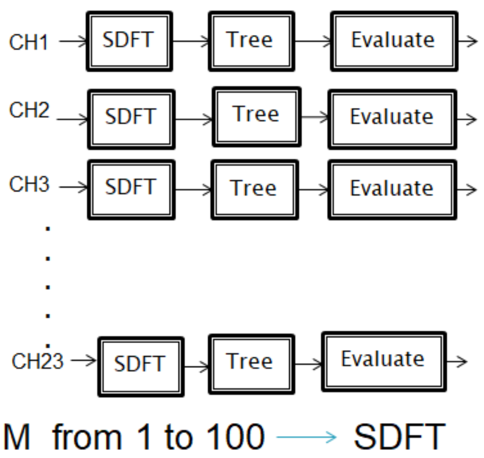


Figure 3.14. The method used to calculate the best value of M (the forward delay in the IIR filter) for each channel separately.

The maximum accuracy achieved was 84% when using the M value for each channel, as shown in table 3.5.

Table 3.5. Best value of M for each channel

Channel	Best M value
Ch1	1
Ch2	2
Ch3	16
Ch4	20
Ch5	18
Ch6	20
Ch7	20
Ch8	20
Ch9	20
Ch10	20
Ch11	20
Ch12	20
Ch13	20
Ch14	20
Ch15	20
Ch16	20
Ch17	30
Ch18	20
Ch19	20
Ch20	20
Ch21	22
Ch22	20
Ch23	20

From the results of accuracy, the first way of calculating accuracy was better than the second way.

2. **The second method:**

The second method is the same first method, but an additional step added after the classifier to calculate the average of one second the results in the testing phase, as shown in figure 3.15.

Figure 3.15. The second method structure that used with a second dataset.

The flow chart of the two methods that used with CHB-MIT dataset shown in figure 3.16.

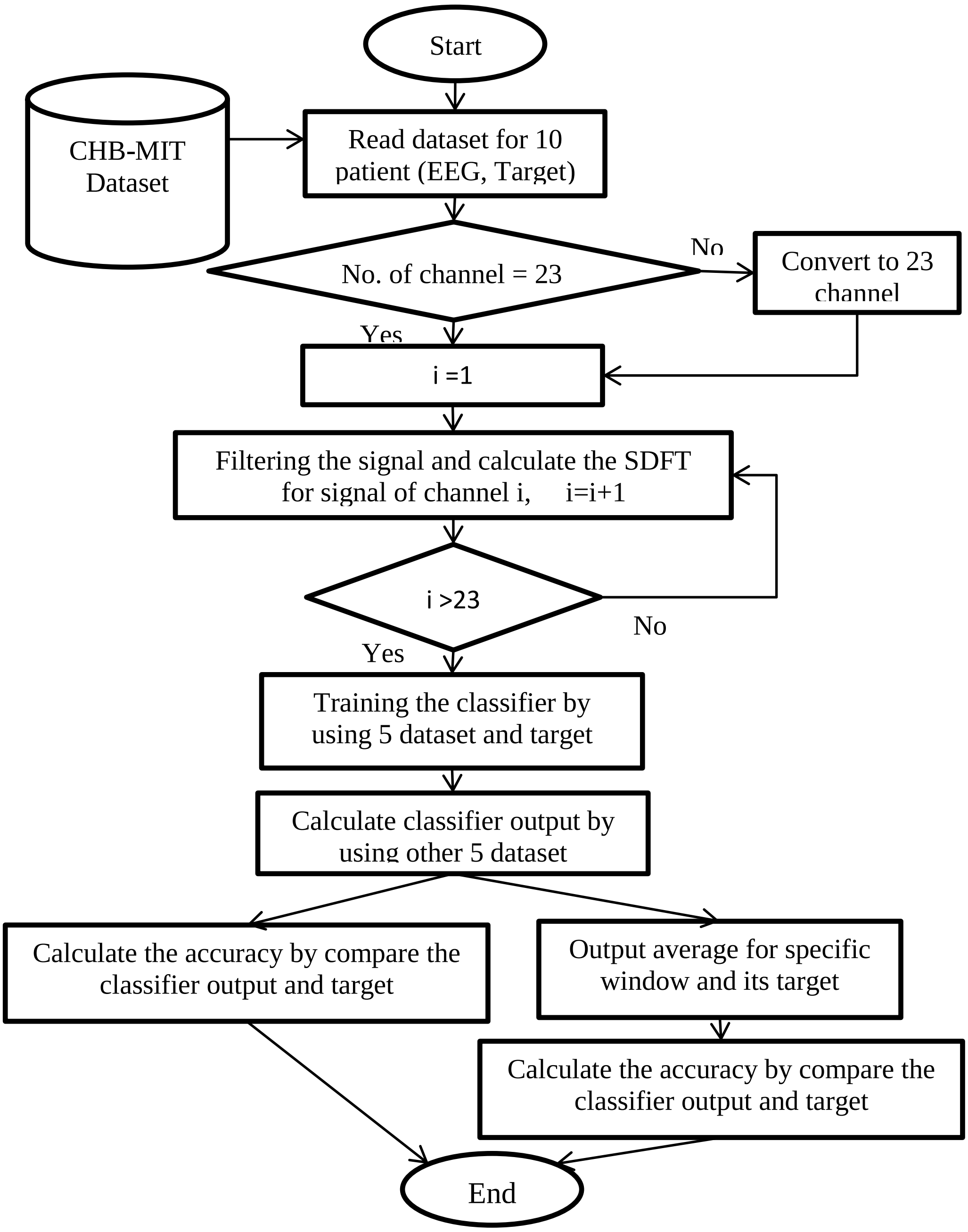


Figure 3.16. The flow chart of methods of seizure detection that used with CHB-MIT dataset.

CHAPTER FOUR

Results and Discussion

4.1.Introduction

In this chapter, the results achieved from different models of machine learning are explained and discussed the effect of features and the time delay on the accuracy results for a different classifier, using two different EEG datasets Bonn and CHB-MIT datasets. At first, the results from models that used Bonn datasets were explained and discussed and then explain the results that are achieved and discussed what we achieved from models when the CHB-MIT dataset used. Reading the digital files that contain the iEEG and EEG dataset, prepare the data, preprocessing filter, feature extraction and classification models is programed with MATLAB R2012b and python 3.6.

4.2.Bonn Dataset

In this dataset, there are five cases of signals CD-E, C-E, D-E, A-E, and ABCD-E used for seizure detection from these cases, the information about these cases descripted in chapter 3 section 3.2.1. These signals are shown in figure 4.1 for the first ten second which is of 17800 samples of the signals from all cases ABCD-E, A-E, CD-E, C-E and D-E.

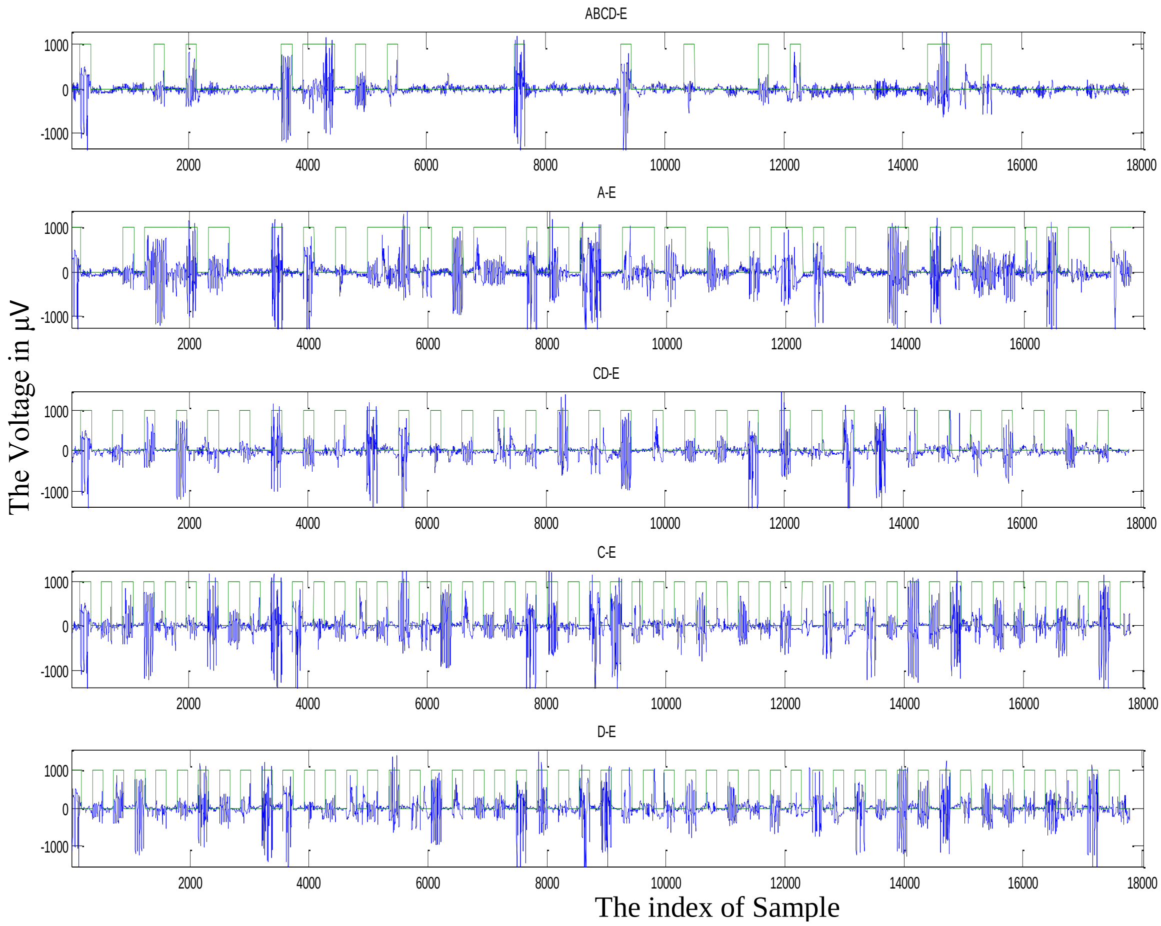


Figure 4.1. The first ten second (17800 samples) of time-series signals in each case ABCD-E, A-E, CD-E, C-E and D-E.

From figure 4.1 obviously the signal oscillated between 1000 to - 1000 µv and it still has noise in some interval and clearly there are differences between healthy and non-healthy cases. When these cases are used with different models and classifiers. The results for each step are explained in the following subsections.

4.2.1. Artificial Neural Network Classifier

First, the type of no delay features extraction is applied to these cases. The results are shown in figure 4.2 for case CD-E. These features are used as input to the classifier.

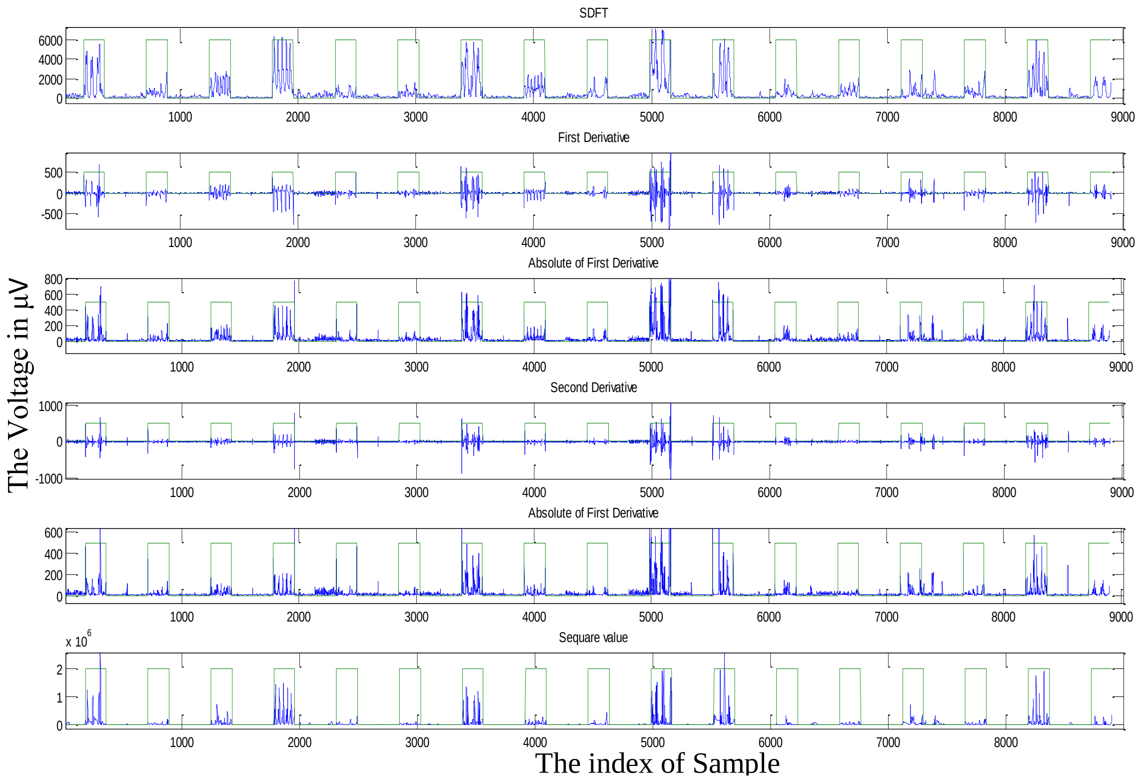


Figure 4.2. The first five-second for the results when six no delays feature extracted applied for case CD-E.

When used these six features with ANN with the structure six node in input layer , 20 node in hidden layer and one output as shown in figure 4.3

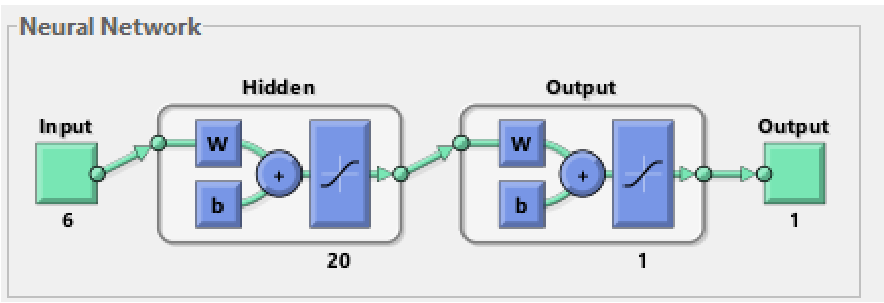


Figure 4.3. the structure of ANN that used six input and twenty hidden nodes with one output

In the training phase, the train dataset 50% from the dataset of the case CD-E used with 1000 epoch. The mean square error characteristic for each epoch of training steps is shown in figure 4.4.

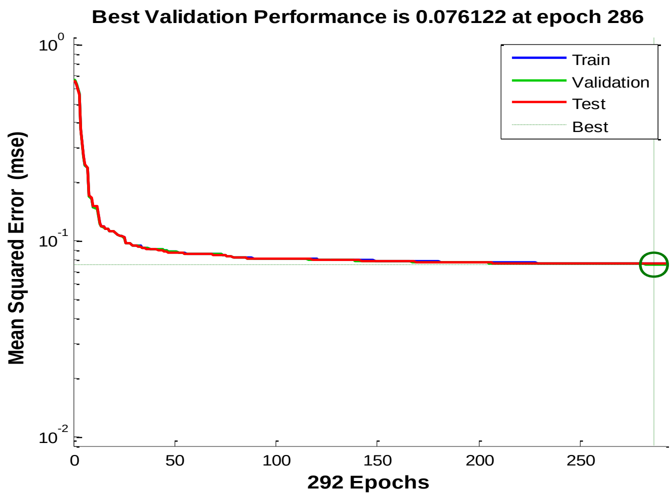


Figure 4.4. The mean squared error for each epoch in the training phase when no delay features are used as input to the classifier in case CD-E.

From figure 4.4 obviously the mean square error (MSE) reach to point that his value less that 0.01 that means the training process going well and all signals are memorized in neural network as weight. When the other 50% of the data is set as input to trained ANN, the output from ANN is shown in figure 4.5.

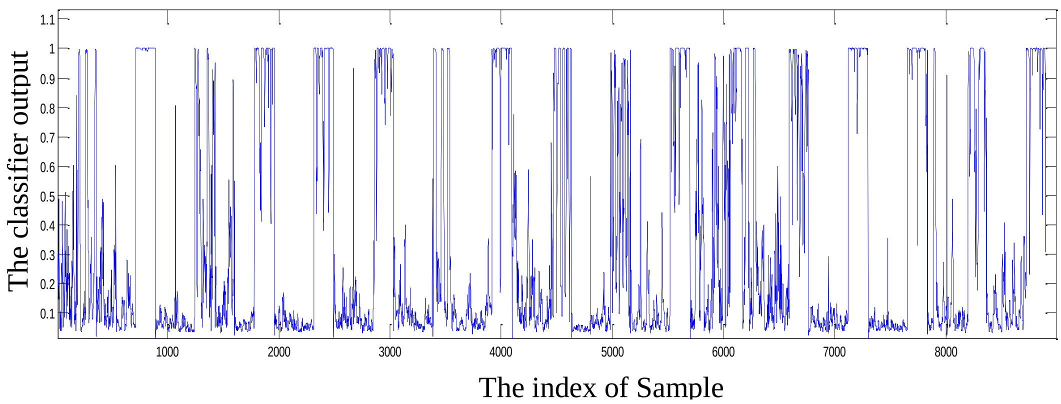


Figure 4.5. The first five seconds of the output from trained ANN when no delay features are used as input to the classifier in case CD-E.

To evaluate the classifiers, the true positive, true negative, false positive and false negative are used to calculate the accuracy, sensitivity and specificity calculated using equations in section 2.6 and the result shown in figure 4.7.

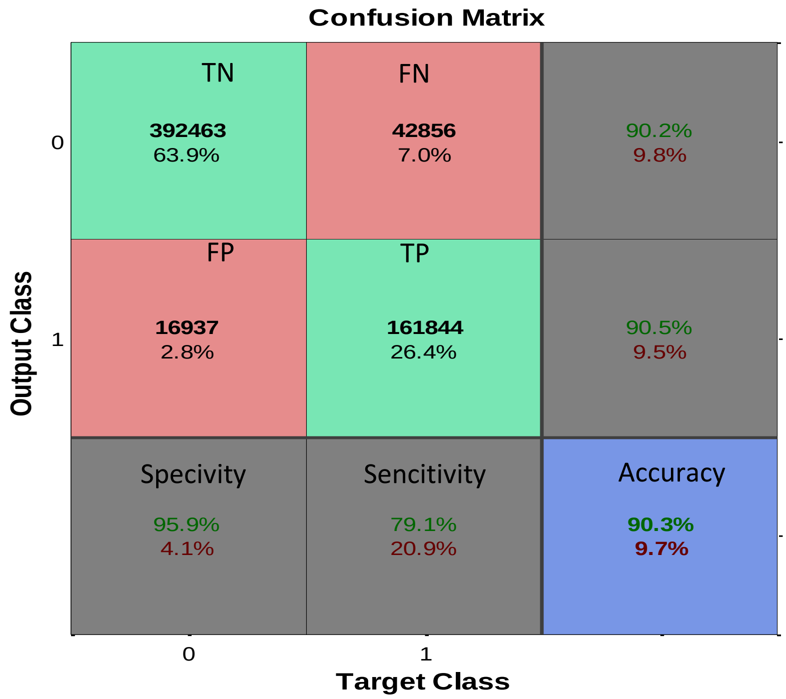


Figure 4.6. The values of true positive, true negative, false positive and false negative, accuracy, sensitivity and specificity for case CD-E and no delay features used with the ANN classifier.

And when using the second group of features that need time delay with one second delay to calculate the features value (15 features used that explained in section 3.2.1). The features that are achieved shown in figure 4.7. In figure 4.7 the blue color is the features value and the red color is the target that generated dependent on the digital files that contain the type of the signals.

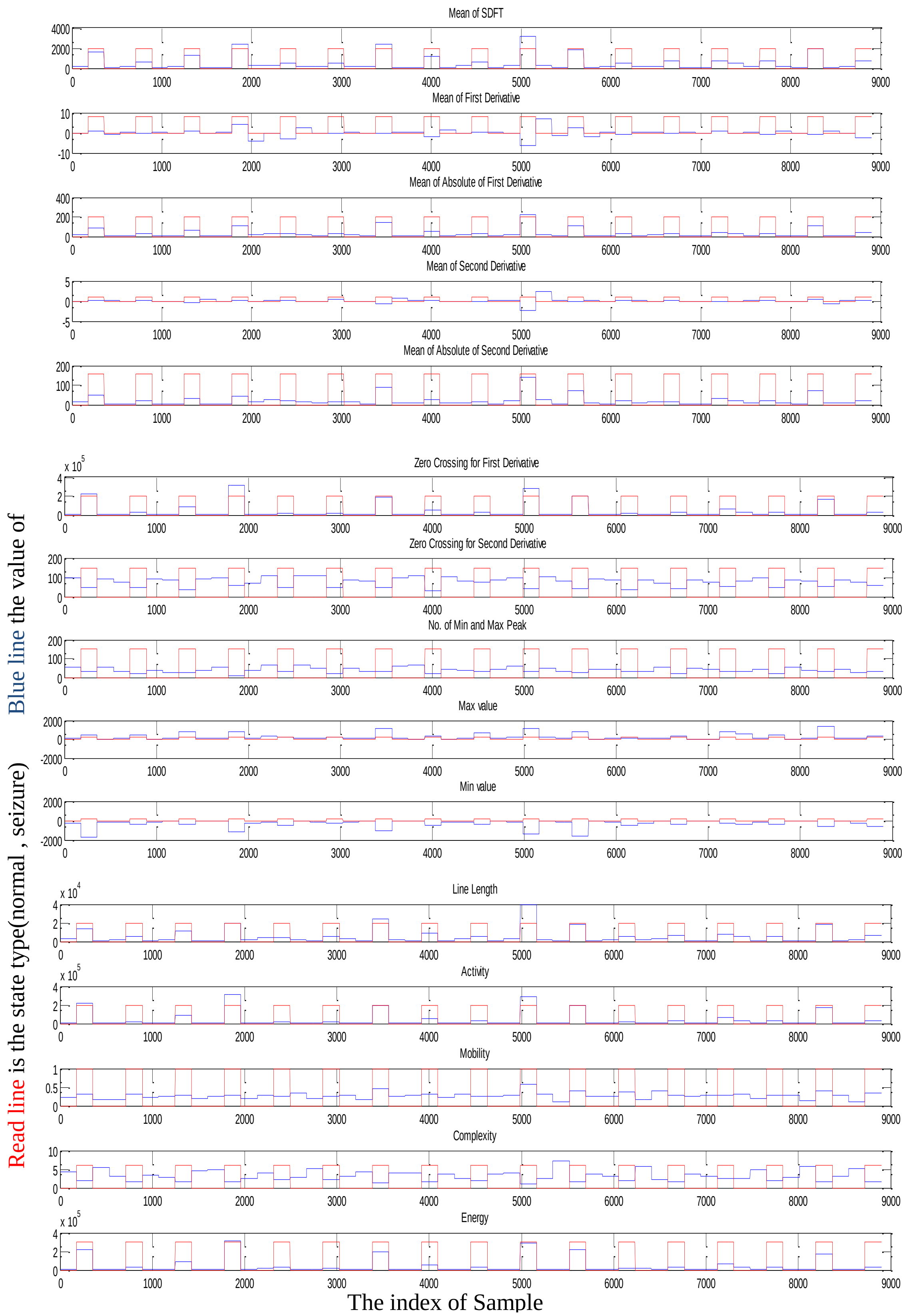


Figure 4.7. the first five-second for the results when fifteen features with one second delay applied for case CD-E

When used fifteen with one-second delay features with ANN, the structure of ANN will be as shown in figure 4.8.

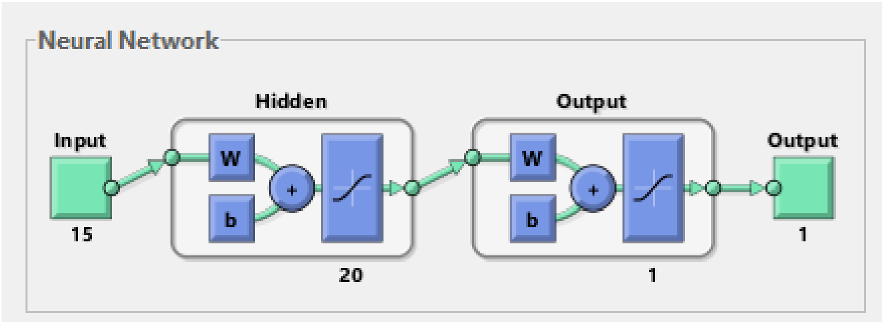


Figure 4.8. The structure of ANN used fifteen inputs and twenty hidden nodes with one output.

In the training phase, the train dataset of the case CD-E was used with 1000 epoch. The mean square error characteristic of training steps is shown in figure 4.9.

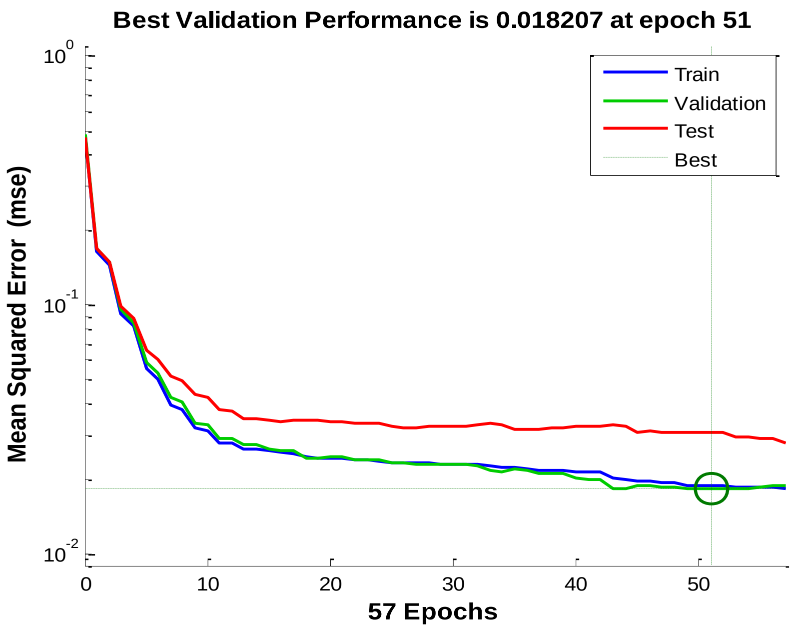


Figure 4.9. The mean squared error for each epoch in the training phase when one-second delay features are used as input to ANN classifier in case CD-E.

When used the other 50% dataset as input to trained ANN for testing the output from ANN shown in figure 4.10.

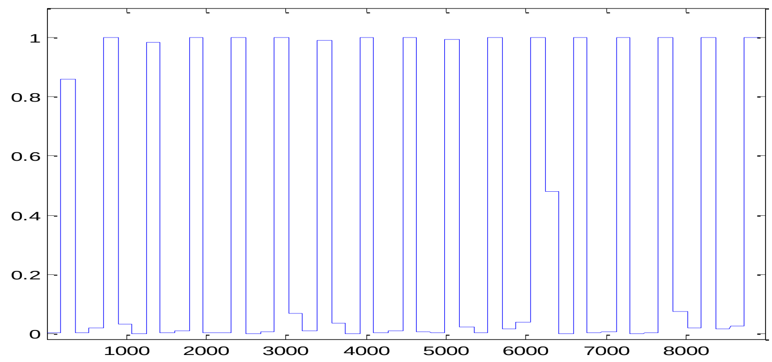


Figure 4.10. the first five seconds of the output from trained ANN when one-second delay features are used as input to the classifier in case CD-E.

To evaluate the classifier, the true positive, true negative, false positive and false negative calculated as shown in figure 4.11. where these values use to calculate the accuracy, sensitivity and specificity.

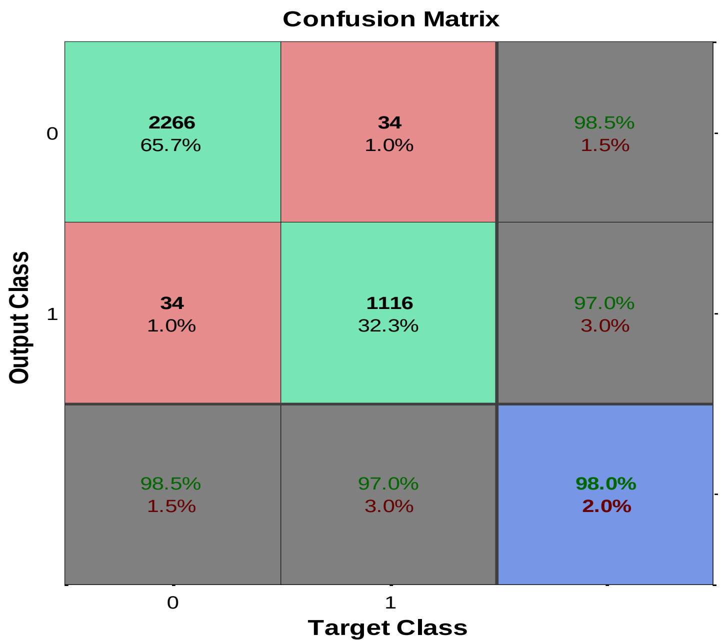


Figure 4.11. The values of true positive, true negative, false positive and false negative, accuracy, sensitivity and specificity for case CD-E with one-second delay features used with ANN classifiers.

And when using the second group for feature selection with a time delay of two seconds, the feature extracted, these features shown in figure 4.12.

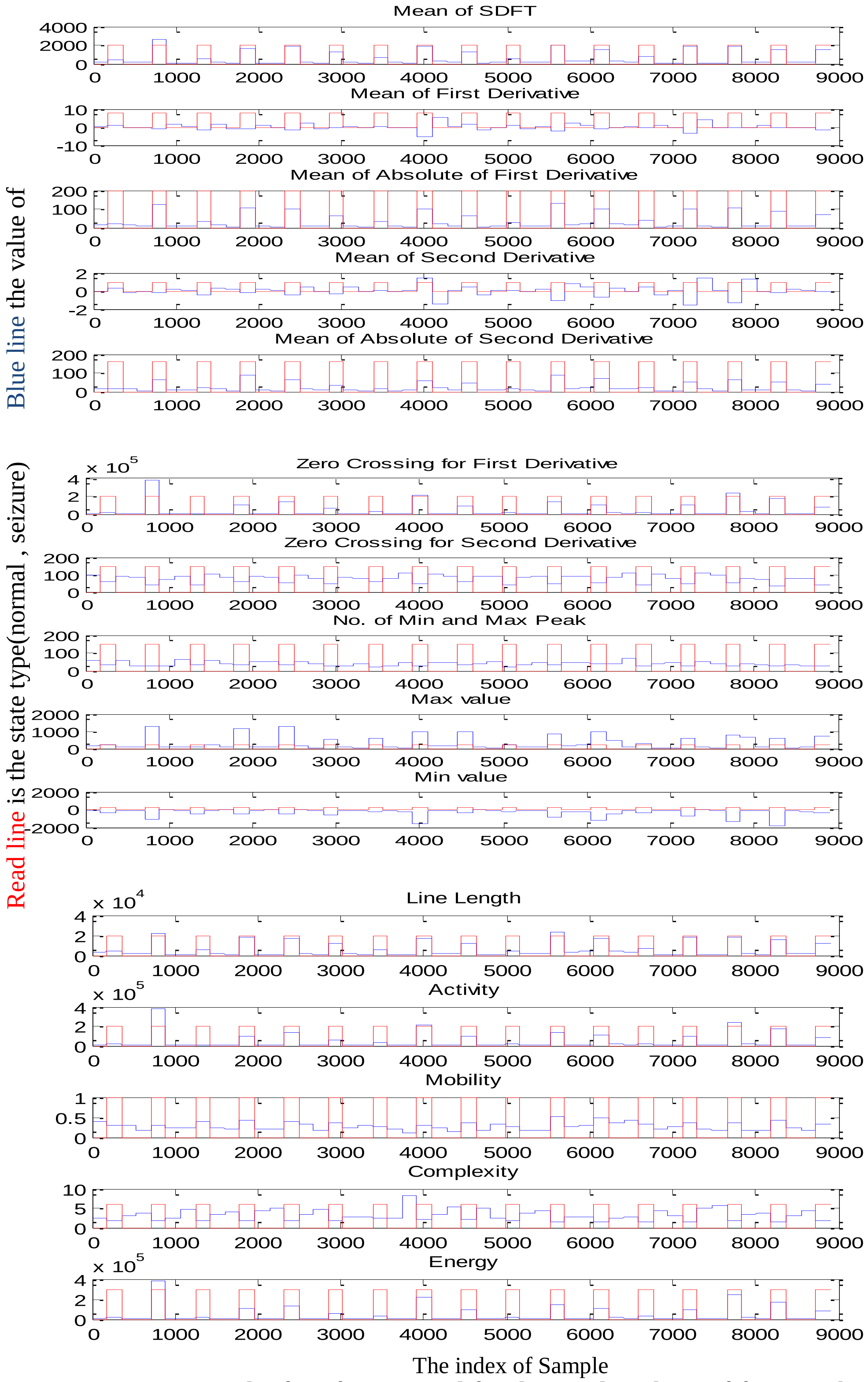


Figure 4.12. the first five-second for the results when a fifteen with two-second delay applied for case CD-E.

These features are used as input to ANN with the structure shown in figure 4.10. In the training phase, the train dataset of the case CD-E was used with 1000 epoch. The mean square error characteristic of training steps is shown in figure 4.13.

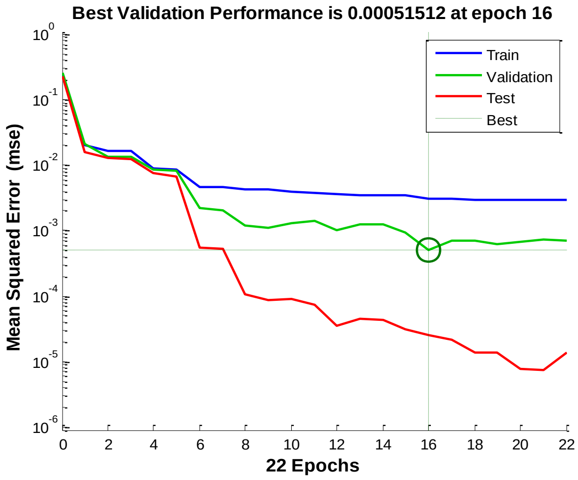


Figure 4.13. The mean squared error for each epoch in the training phase when two-second delay features are used as input to ANN classifier in case CD-E.

When used the other 50% dataset as input to trained NN for testing the output from NN shown in figure 4.14.

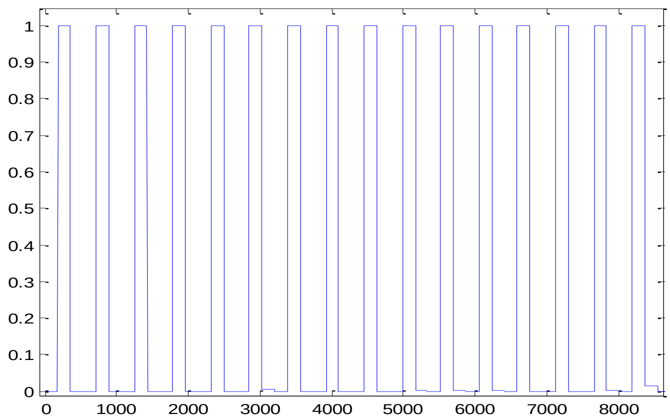


Figure 4.14. the first five seconds of the output from trained ANN when two-second delay features are used as input to the classifier in case CD-E.

The true positive, true negative, false positive and false negative values when comparing the output from the classifier and the target are shown in figure 4.15.

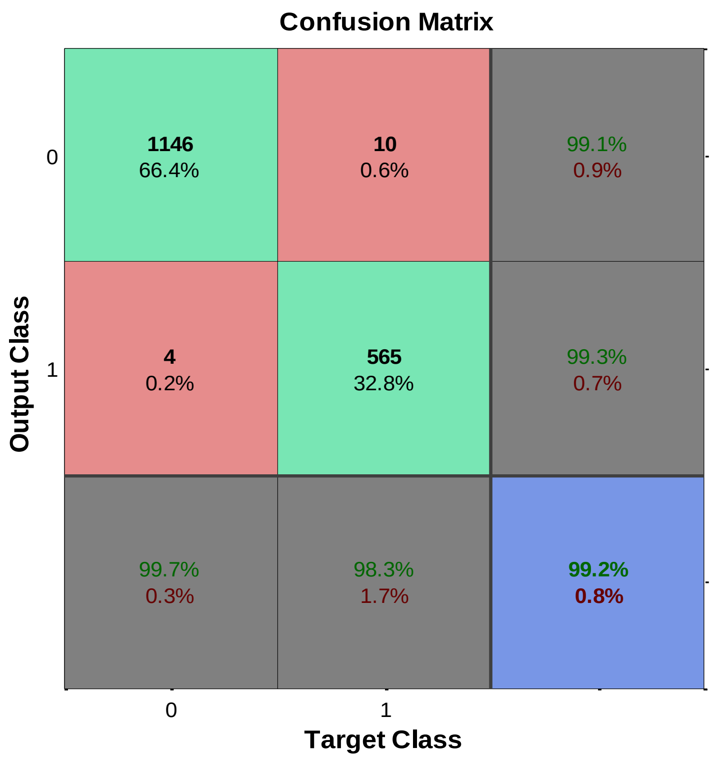


Figure 4.15. The values of true positive, true negative, false positive and false negative, accuracy, sensitivity and specificity for case CD-E with two-second delay features used with ANN classifiers.

The same steps that used with case CD-E used with the other cases (ABCD-E, A-E, C-E and D-E) where these cases input to machine system that shown in figure 4.8 with 50% of data used in training and the other 50% used to calculate the output and compare this output with the target for evaluating the automatic seizure detection .Two types of features no delay required group and the other group that need time delay for calculating the features were used. The second group used two-time values one second and two seconds and ANN used as classifiers the results for all cases shown in table 4.1.

Table 4.1. The result of all cases with features extracted in three ways and ANN used as the classifier.

Feature extraction type	Case	Accuracy%	Sensitivity%	Specificity%
Six features no delay	CD-E	90.2633	79.06	95.86
	C-E	91	87.33	94.66
	D-E	84.6	78.95	90.25
	ABCD-E	94.5374	79.76	98.12
	A-E	94.77	91.53	98.02
Fifteen features with one-second delay	CD-E	98	97.04	98.52
	C-E	99.0435	98.96	99.13
	D-E	95.6522	94.78	96.52
	ABCD-E	97.98	93.85	98.98
	A-E	99.74	99.57	99.91
Fifteen features with a two-second delay	CD-E	99.2	98.3	99.7
	C-E	99.75	99.5	99.84
	D-E	98.5	97.74	99.13
	ABCD-E	99.3	97.57	99.74
	A-E	100	100	100

From Table 4.1, less accuracy is achieved when the first way for the six feature extraction without delay is used. The accuracy will be increased when the second method of feature extraction is used. Also, the accuracy increased when the time delay increased.

4.2.2. ANFIS Classifier

The second classifier used was ANFIS that shown in section 2.5.2. The structure of the classifier is shown in figure 4.16 when the six features are input to ANFIS.

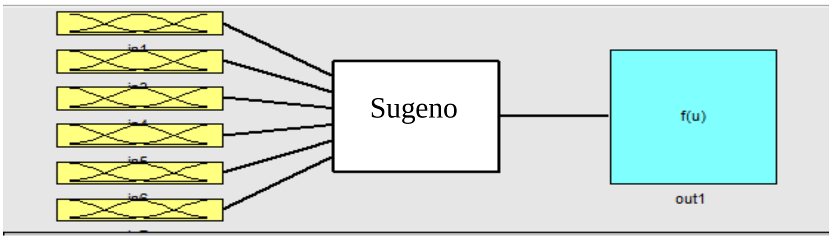


Figure 4.16. The structure of ANFIS with six inputs which are the first-way feature extracted (no delay required) used.

The Gaussian memberships that are used for fuzzification of the input in the first layer of ANFIS is shown in figure 4.17 in case CD-E.

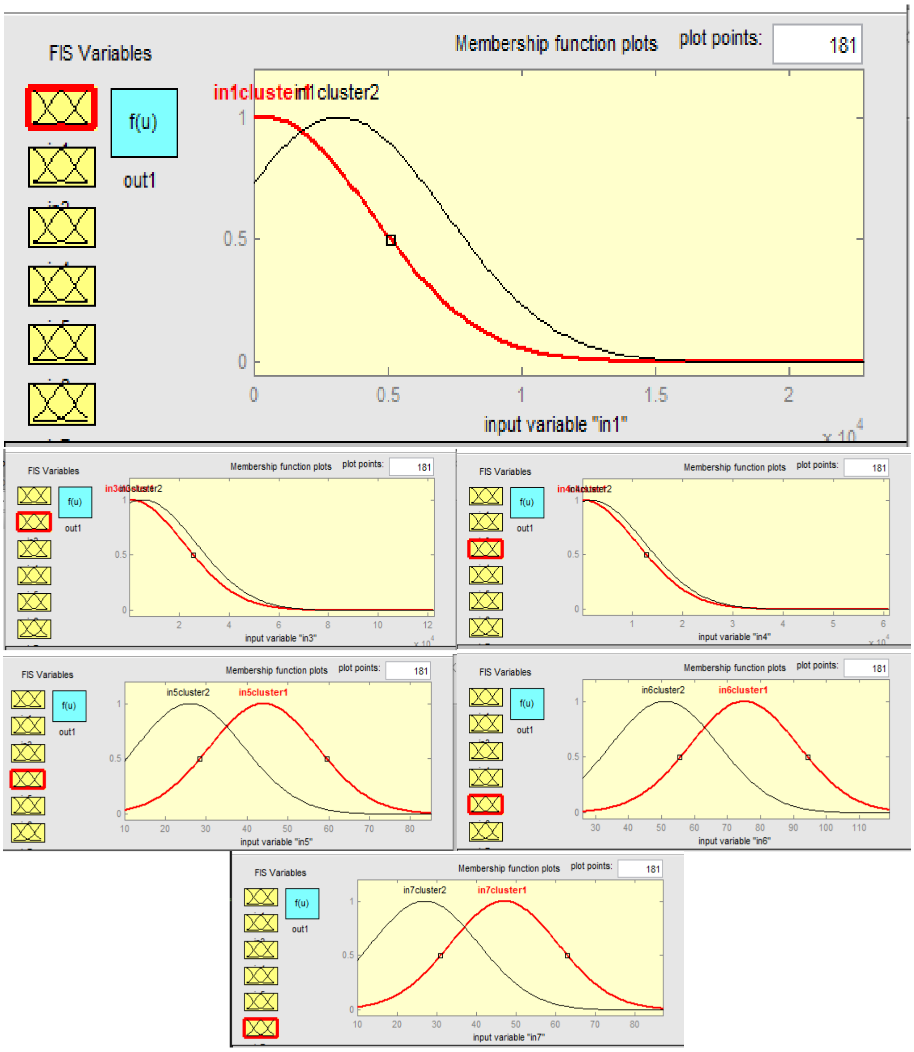


Figure 4.17. The Gaussian memberships for fuzzification the six inputs to the ANFIS in case CD-E.

The true positive, true negative, false positive and false negative values that are achieved when comparing the output from the classifier and the target are shown in figure 4.18 in case CD-E.

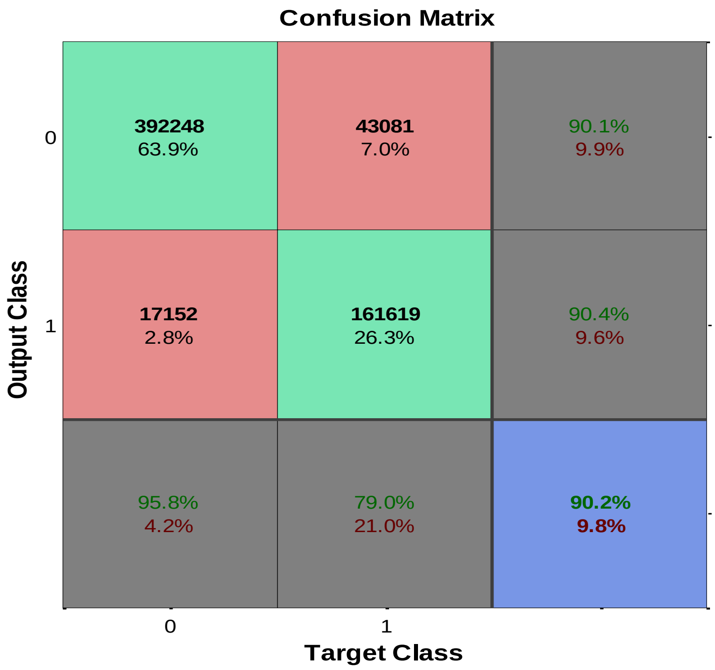


Figure 4.18. The values of true positive, true negative, false positive and false negative, accuracy, sensitivity and specificity for case CD-E with first-way feature extraction used with ANFIS classifier.

When using an ANFIS classifier with fifteen features input, the structure of the classifier will be as shown in figure 4.19.

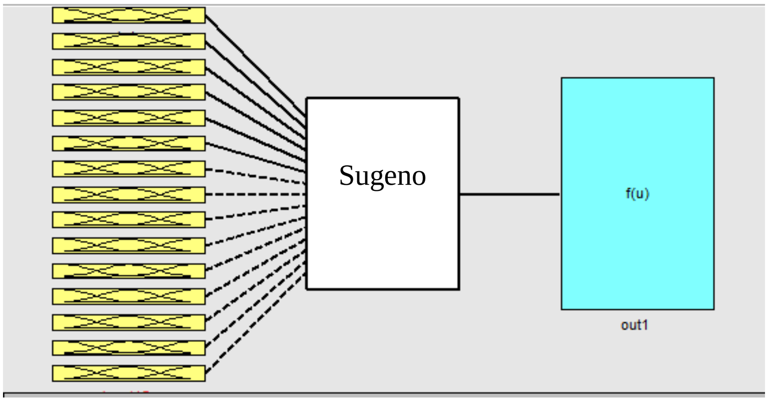


Figure 4.19. the relationship for fuzzification for the fifteen inputs to the ANFIS in case CD-E.

And the result value for the evaluation parameter when using the ANFIS classifier with fifteen inputs is shown in figure 4.20 for case CD-E with a one-second delay.

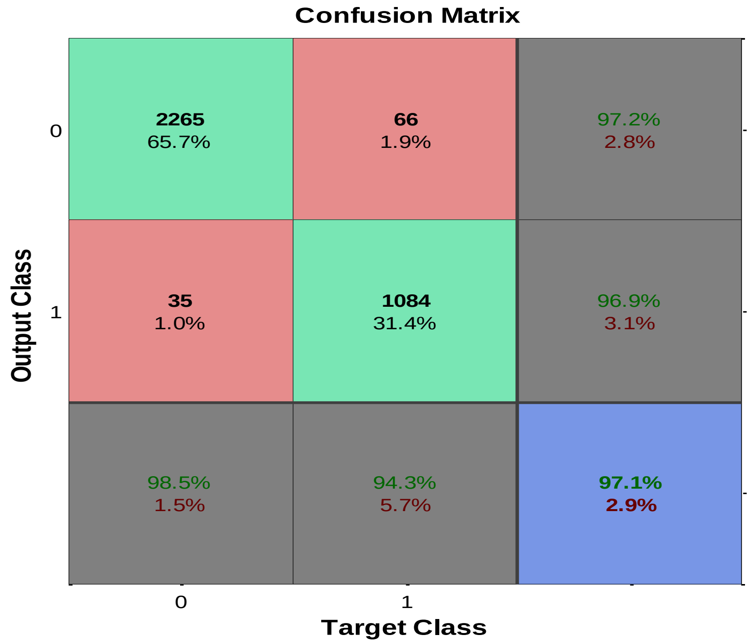


Figure 4.20. The values for evaluation parameters for case CD-E with second-way feature extraction used with ANFIS classifiers.

Figure 4.21 showed the result when fifteen features were used with a two-second delay.

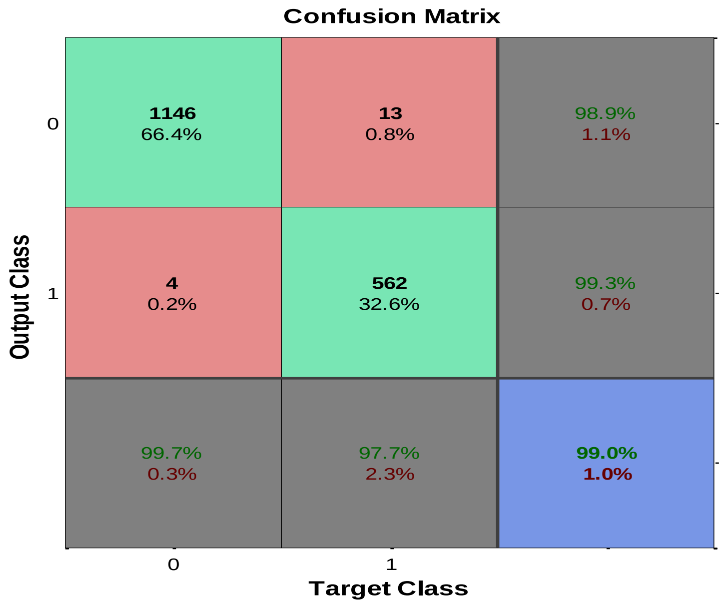


Figure 4.21. The values for evaluation parameters for case CD-E with Third way feature extraction used with ANFIS classifier.

The results for all cases (CD-E, C-E, D-E, ABCD-E and A-E) by extract features using no delay, one second and two-second delay with ANFIS classifier are shown in table 4.2.

Table 4.2. The result of all cases with features extracted in three ways and ANFIS classifier.

Feature extraction type	**Case**	**Accuracy%**	**Sensitivity%**	**Specificity%**
Six features no delay	CD-E	90.1917	78.95	9581
	C-E	90.1715	86.39	93.95
	D-E	86.7670	82.55	90.98
	ABCD-E	93.8133	77.32	97.94
	A-E	94.2059	90.24	98.18
Fifteen features with one-second delay	CD-E	97.0725	94.26	9848
	C-E	99.0435	99.13	9896
	D-E	95.4783	9443	9652
	ABCD-E	97.8261	93.22	98.98
	A-E	99.67	99.39	100
Fifteen features with a two-second delay	CD-E	99.0145	97.74	9965
	C-E	99.6522	99.48	99.83
	D-E	98.3478	97.74	9896
	ABCD-E	98.47	96.17	99.04
	A-E	99.31	98.61	100

From Table 4.2, less accuracy is achieved when the first way for the six feature extraction without delay is used. The accuracy will be increased when the second method of feature extraction is used. Also, the accuracy increased when the time delay increased.

4.2.3. Decision Tree Classifier

The third classifier used was DT described in the chapter two section five part two with no delay features and other features that need time to calculate the values. The structure of DT when the input was the features that need a two-second delay used in the training phase is shown in figure 4.22.

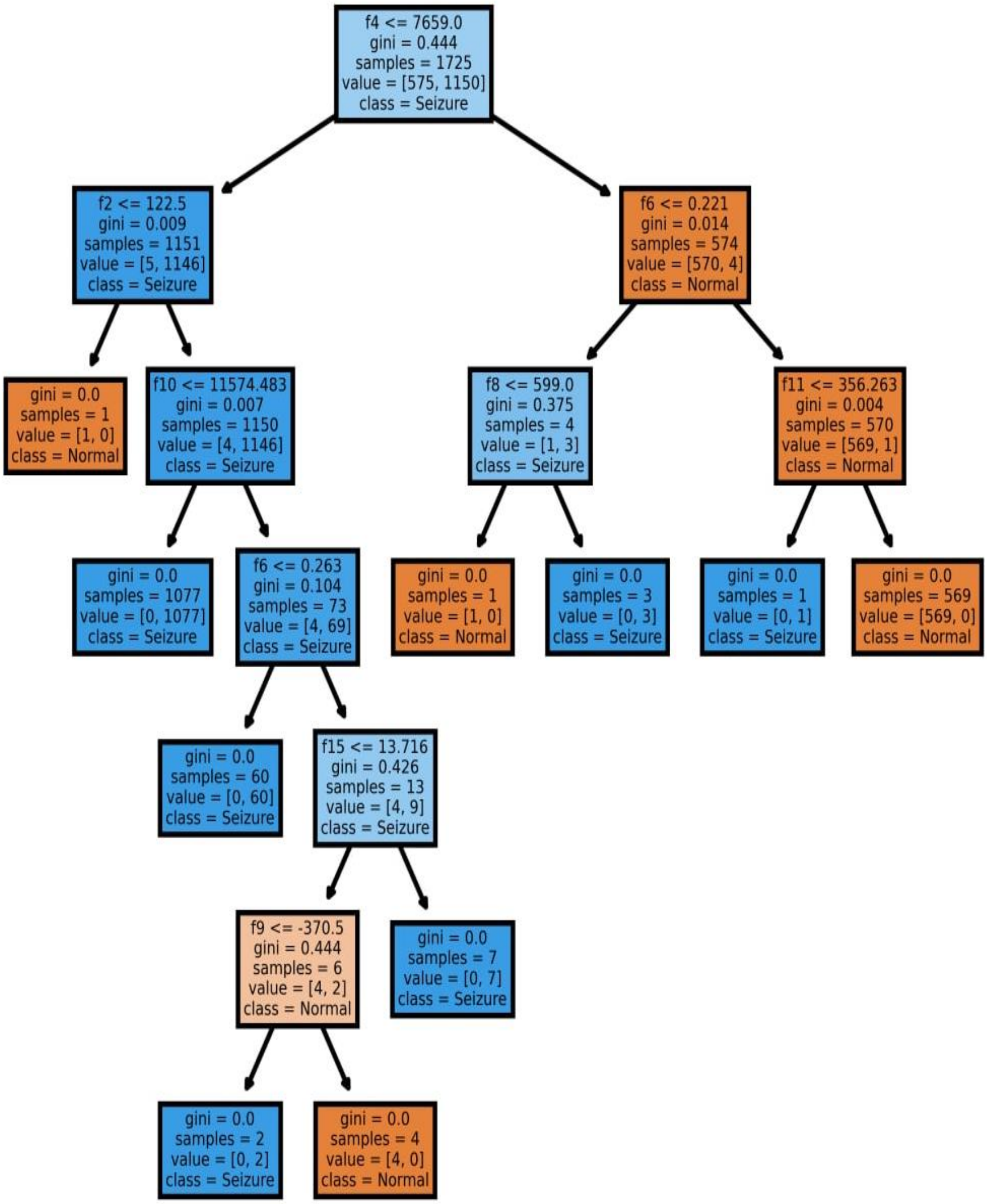


Figure 4.22. The structure of DT when 50 % of CD-E is used with the third way for feature calculation.

The output from the DT classifier was two values, zero or one. The results of evaluation values when DT is used as a classifier for all cases are shown in table 4.3.

Table 4.3. The result of all cases with features extracted in three ways and DT classifier.

Feature extraction type	Case	Accuracy%	Sensitivity%	Specificity%
Six features no delay	CD-E	85	77	89
	C-E	86	86	86
	D-E	81	80	81
	ABCD-E	91	76	94
	A-E	92	91	92
Fifteen features with a one-second delay	CD-E	98	97	98
	C-E	98	98	99
	D-E	96	96	96
	ABCD-E	98	95	99
	A-E	100	100	100
Fifteen features with a two-second delay	CD-E	99	98	99
	C-E	100	100	100
	D-E	98	99	98
	ABCD-E	99	98	99
	A-E	100	100	100

From Table 4.3, less accuracy is achieved when the first way for the six feature extraction without delay is used. The accuracy will be increased when the second method of feature extraction is used. Also, the accuracy increased when the time delay increased from one second to two seconds.

4.2.4. Random Forest Classifier

Another classifier used is RF that described in 2.5.2. The input to the classifier was the feature that no delay needs for calculation that extracted from training data, and then the accuracy calculated is dependent on the comparison between the classifier output and the target also the accuracy for other states that features need time for calculation used. the structure of five trees in RF with a one-second delay for the case CD-E is shown in figure 4.23.

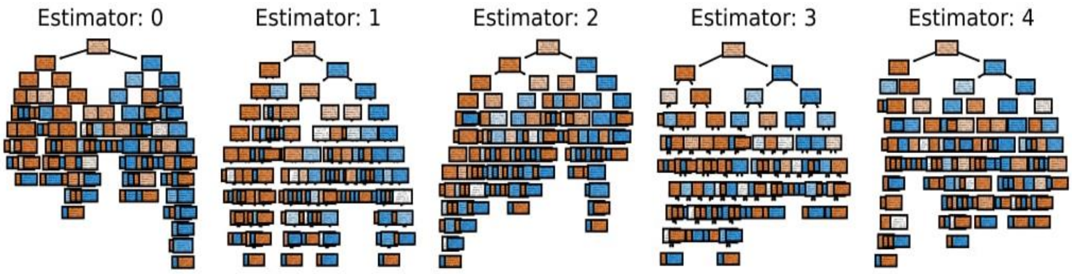


Figure 4.23. The structure of RF when 50 % of the case CD-E was used with the second way for feature calculation.

And the structure of five trees in RF with a Two-second delay when 50 % from the case CD-E data used for the training RF classifier is shown in figure 4.24.

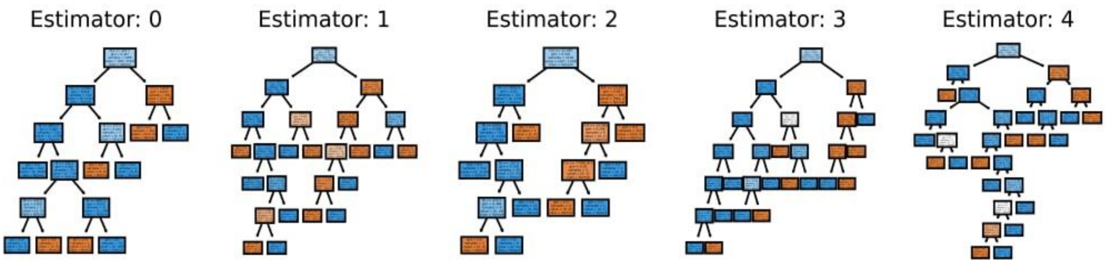


Figure 4.24. The structure of RF when 50 % of CD-E are used with the third way for feature calculation.

From figure 4.24, the RF contains five trees with the voting method used for classification. The first tree structure is shown in figure 4.25 in case the data using for training was CD-E and the features extracted was the third way.

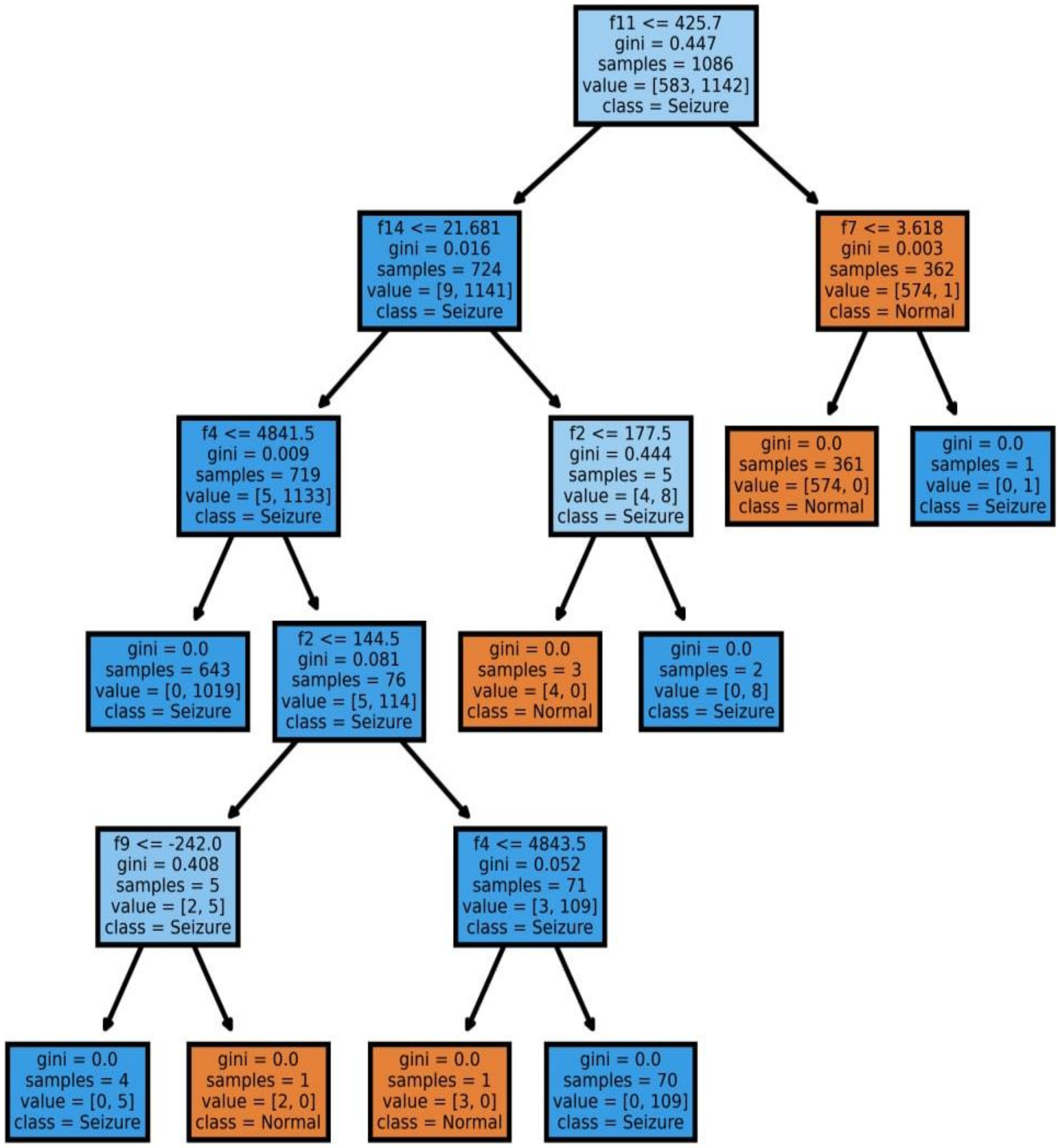


Figure 4.25. The structure of the First tree in the RF classifier when 50 % of case CD-E is used with the third way for feature calculation.

In the testing phase, when used the remain 50% of case CD-E, the output from the RF classifier was zeroes and ones, the same method is used with other cases. The results of evaluation variables values are shown in Table 4.4.

Table 4.4. The result of all cases with features extracted in three ways and RF classifier.

Feature extraction type	Case	Accuracy %	Sensitivity%	Specificity%
Six features no delay	CD-E	90	88	91
	C-E	89	91	88
	D-E	84	85	84
	ABCD-E	93	89	95
	A-E	94	95	92
Fifteen features with a one-second delay	CD-E	99	99	99
	C-E	99	99	99
	D-E	97	98	97
	ABCD-E	99	98	99
	A-E	100	100	100
Fifteen features with a two-second delay	CD-E	99	98	99
	C-E	100	100	100
	D-E	99	99	99
	ABCD-E	99	99	99
	A-E	100	100	100

From Table 4.4, less accuracy is achieved when the first way for the six feature extraction without delay is used. The accuracy will be increased when the second method of feature extraction is used. Also, the accuracy increased when the time delay increased from one second to two seconds.

4.3.Feature importance

To measure the effect for each feature on the classification, the methods with Model used, which are the Permutation on the full model and MDI that explain in section 2.7, the values for effect each six no delay feature that are calculated in a first way are explained in table 4.5.

Table 4.5. The values of effective no delay features on classification for case CD-E.

The features	MDI	Permutation on the full model
SDFT	**0.38064184**	0.20121357
First Derivative	0.07402836	-0.0024658
Second Derivative	0.04767845	-0.00141841
Absolute of First Derivative	0.18356998	0.02785808
Absolute of Second Derivative	0.07943939	-0.00106671
Square Values	0.23464197	0.10247739

From table 4.5, we find the SDFT has more importance in the classification than other features, as shown in figure 4.26 when permutation on the full model is used and figure 4.27 when MDI is used.

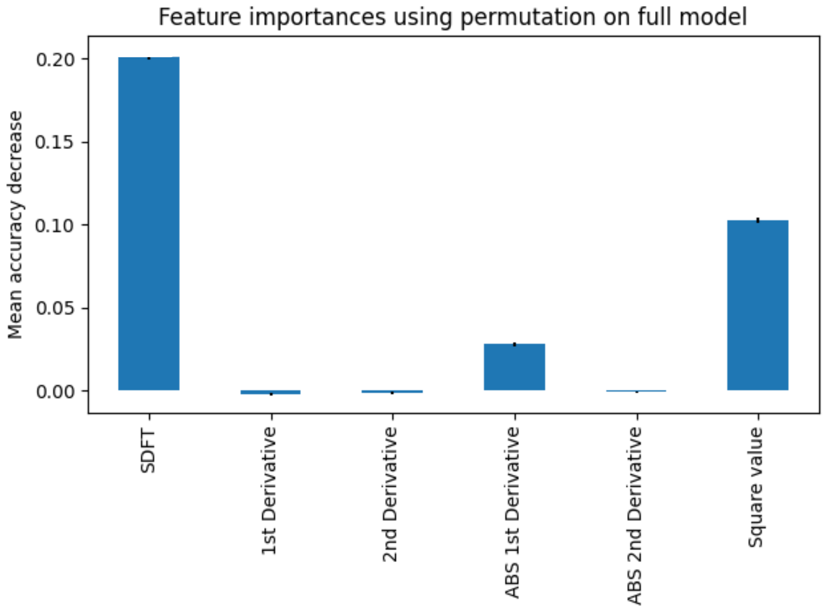


Figure 4.26. The importance of six no delay features measured by Permutation on full model with CD-E case.

From figure 4.26, when useing Permutation on full model we find the SDFT has more importance than the other features in classification. Also, the first and second derivative and the absolute value of the second derivative have the very low value nearly to zero of importance in classification.

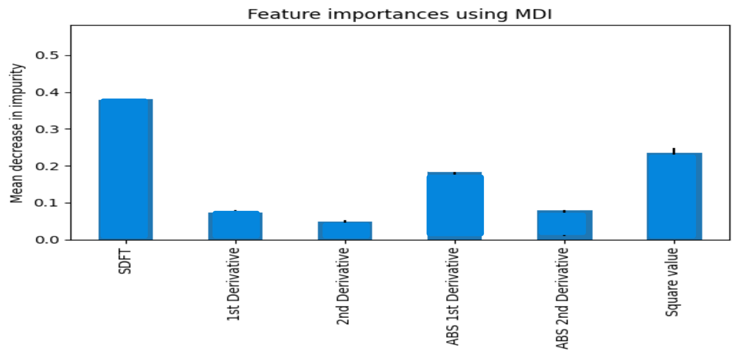


Figure 4.27. The importance of six no delay features measured by MDI with CD-E case

From figure 4.27, also when use MDI on full model we find the SDFT has more importance than the other features in classification and the first , second derivative and the absolute value of the second derivative have the weak value of importance in classification. The values of effect each feature from the fifteen features calculated with one-second delay explain in table 4.6.

Table 4.6. The values of effective features with one-second delay on classification in case CD-E.

The features	MDI	Permutation on the full model
Zero Crossing for First Derivative (zd1)	0.00769375	0.00043478
Zero Crossing for Second Derivative (zd2)	0.03293209	0.00510145
The No. of Peaks (nmm)	0.01992611	0.0006087
Line Length (LL)	0.11182255	0.00255072
Activity	0.11766693	**0.00513043**
Mobility	0.01393434	0.00368116
Complexity	0.0912946	0.00336232
Maximum Value (mxi)	0.04782589	0.00228986
Minimum Value (mni)	0.0506346	0.00118841
Energy	0.12946912	0.00330435
Mean of SDFT	**0.22229255**	0.00113043
Mean of First Derivative (d1mean)	0.00576929	0.00136232
Mean of Second Derivative (d2mean)	0.00598382	0.00313043
Mean of the absolute of First Derivative (absd1mean)	0.11180382	0.00043478
Mean of the absolute of Second Derivative (absd2mean)	0.03095054]	0.00223188

From table 4.6, we find the SDFT has more importance in the classification than other features in case of using MDI and the activity has more importance in the classification than other features in case of using Permutation on the full model, as shown in figure 4.28 when useing permutation on full model and figure 4.29 when using MDI.

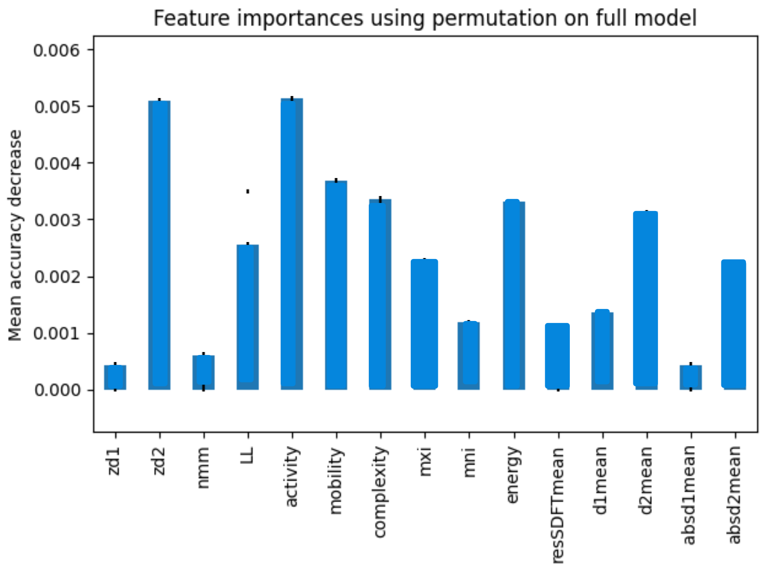


Figure 4.28. The importance of fifteen with one-second delay features measured by Permutation on full model with case CD-E.

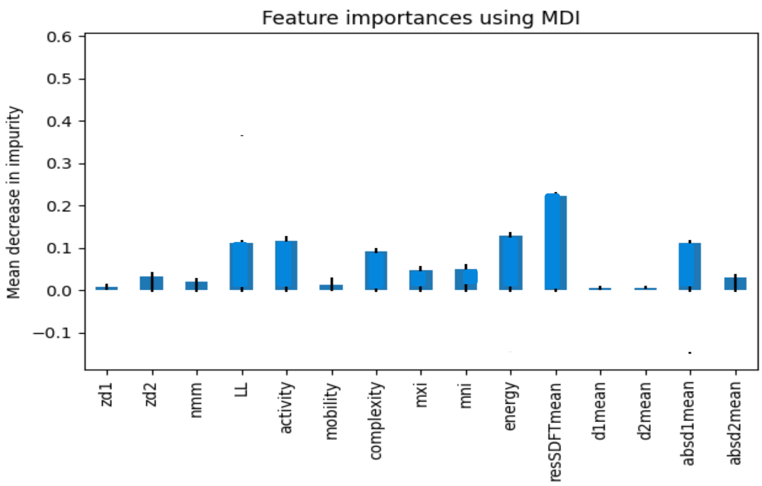


Figure 4.29. The importance of six fifteen with one-second delay features measured by MDI with CD-E case**.**

The values of effect each feature from the fifteen features calculated with two-second delay explained in table 4.7.

Table 4.7. The values of effective features with one-second delay two classifications for case CD-E.

The features	MDI	Permutation on the full model
Zero Crossing for First Derivative (zd1)	0.00178382	0.00017391
Zero Crossing for Second Derivative (zd2)	0.1253836	0.00069565
The No. of Peaks (nmm)	0.01965214	0.00052174 -
Line Length (LL)	0.13036289	0.00104348
Activity	0.07171936	0.00075362
Mobility	0.0054555	0.00104348
Complexity	0.02855865	**0.0013913**
Maximum Value (mxi)	0.01924116	0.00023188
Minimum Value (mni)	0.03630588	0.00046377
Energy	0.09053102	-0.00017391
Mean of SDFT	**0.19596095**	-0.00028986
Mean of First Derivative (d1mean)	0.00071119	-0.00046377
Mean of Second Derivative (d2mean)	0.00051599	0.
Mean of the absolute of First Derivative (absd1mean)	0.15584085	-0.00104348
Mean of the absolute of Second Derivative (absd2mean)	0.11797698	-0.0004058

From table 4.7, we find the SDFT has more importance in the classification than other features in case using MDI, and the complexity has more importance in the classification than other features in case using Permutation on the full model, as shown in figure 4.30 when use permutation on full model and figure 3.31 when using MDI.

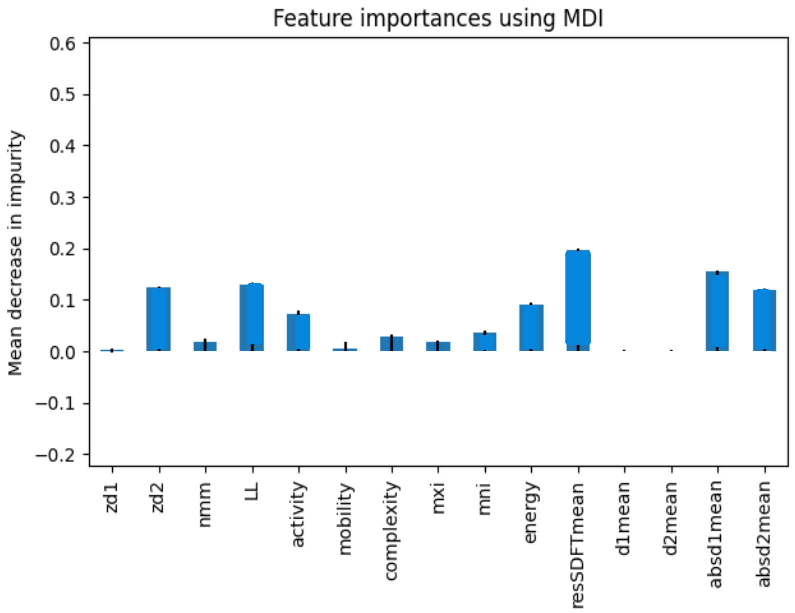


Figure 4.30. The importance of fifteen with two-second delay features measured by Permutation on full model with case CD-E.

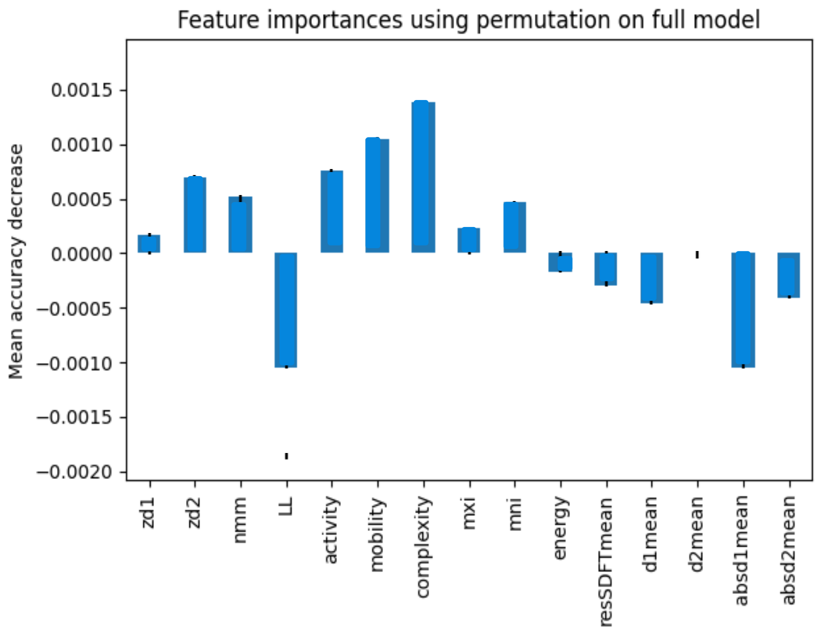


Figure 4.31. The importance of fifteen with two-second delay features measured by MDI with case CD-E.

4.4.Feature selection

When applying the SFS method for the first group of features the results are explained in table 4.8.

Table 4.8. The accuracy for different no. of first group features that are used in SFS

No. of features	The accuracy for training data%	The accuracy for testing data%
2	86.1	86.2
3	87.1	87.1
4	88.3	88.4
5	88.4	88.4
6	88.4	88.4

From table 4.8 we find when using four, five or six have the same accuracy when using testing data from these results we can reduce the feature from six features to four features (SDFT, the second derivative, the absolute of first derivative and the square value). And when applying the SFS method for the second group of features the results that we achieved are explained in table 4.9.

Table 4.9. The accuracy for different no. of second group features that used with SFS

No. of features	The accuracy for training data%	The accuracy for testing data%
2	96	95.8
3	96.98	96.5
4	97.18	96.7
5	97.16	96.8
6	97.45	97.3
7	97.51	97.7
8	97.54	97.3
9	97.59	97.4
10	97.74	97.5
11	97.68	97.5
12	97.71	97.6
13	97.62	97.4
14	97.77	97.7
15	97.85	97.7

From table 4.9 we find when using seven, fourteen and fifteen have the same accuracy 97.7 when using testing data from these results we can reduce the feature from fifteen features to seven features (zero-crossing for the first derivative, the activity, complexity, energy, the mean of SDFT ,the mean of the second derivative and the mean of the absolut of first derivative).

4.5.CHB-MIT Dataset

In this dataset which is described in section 2.4.1, the matrix 23 * 360001 values that are taken from EEG for five patients used for the training phase and another matrix 23 *3400005 values generated from EEG for other five patients used for the testing phase the patients information and signal information explains in table 3.4 these data passed through a bandpass filter and notch filter figure 4.32.

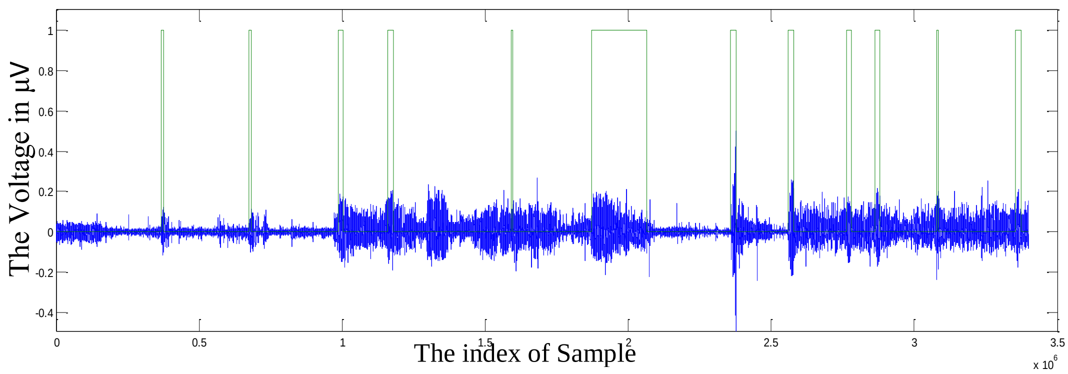


Figure 4.32. The first channel from the dataset with 23 channels used in the testing phase which achieved from five patients.

When applying SDFT, the frequency domain of the time series EEG feature on the training and testing signals with the best value of M is achieved as described in section 3.2.2. When using SDFT with DT the results are shown in table 4.10

Table 4.10. The results when used SDFT as feature extraction with DT classifier

Time duration in sec	Accuracy%	Sensitivity%	Specificity%
0	84	32	94
1	93	96	93
2	95	98	95
3	95	99	95
4	95	99	95
5	95	99	95

From the table 4.10, the accuracy increased to 93 % when the window time duration increased from zero to one and then the accuracy will be constant but the sensitivity will increase and reach to 99%.

When applying the same scenario with the RF classifier, the result is shown in table 4.11.

Table 4.11. The result when used SDFT as feature extraction with RF classifier.

Time in sec	Accuracy%	Sensitivity%	Specificity%
0	92	66	94
1	94	95	94
2	95	98	94
3	95	99	94
4	96	99	95
5	96	99	95

From the table 4.11 also when the time duration increases, the accuracy increased to 96% and then it will be constant for this value but we see the sensitivity increased to reach 99%.

Finally, The ANN classifier with fourtee hidden nodes is used as the third classifier and the results when ANN clssifier used is shown in table 4.12.

Table 4.12. The result when used SDFT as feature extraction with ANN classifier

Time duration in sec	Accuracy%	Sensitivity%	Specificity%
0	94	81	95
1	96	97	96
2	96	99	96
3	96	99	96
4	96	99	96
5	96	99	96

From the table 4.12, when the time duration increases, the accuracy increased to 96% and then it will be constant for this value but we see the sensitivity increased to reach 99%.

From tables 4.10, 4.11 and 4.12, the result from the ANN classifier has better than the RF and DT classifiers for zero sec delay, the accuracy in case DT classifier was 84% while the accuracy was 92% in case RF classifier and 94% in case ANN classifier used. For the other result, also the result from ANN classifier has better than the other classifiers where the accuracy was 96% with ANN and one sec delay and the sensitivity value 99% with two sec delay.

The accuracy is approximately equal for both other classifiers but the accuracy will be better 96% with a 4-sec delay in case RF classifier. From this result we find the ANN classifier has better performance than RF and DT classifiers and the RF classifier better than DT classifier.

Our results were compared with previous findings by using the same datasets as shown in table 4.13.

Table 4.13. Literature of Seizure Detection based on Features and Classifier.

Authors	Dataset	Classification Dataset	Accuracy (%)
K. Samiee, et al. in 2013 [9]	Bonn	A-E B-E C-E D-E ABCD-E	99.8 99.3 98.5 94.9 98.1
L. Wang et al. in 2017 [10]	Bonn	S-FNOZ	99.25
E. Alickovic, et al. in 2017[11]	Freiburg and CHB-MIT	-	100
M. Mursalin, et al. in 2017 [12]	Bonn	A-E B-E C-E D-E ACD-E BCD-E CD-E ABCD-E	100 98.0 99.0 98.5 98.5 97.5 98.67 97.4
A. Narang, et al. in 2018 [13]	Bonn	ABCD-E	99.6
S. Ramakrishnan, et al. in 2018 [14]	Bonn CHB-MIT	A-E ABCD-E A-D-E AB-CD-E A-B-C-D-E -	100 99 98 96 95 98
M. Ravi, et al. in 2018 [15]	Bern-Barcelona Bonn	- -	78.5 94.1
C. Mahjoub, et al. in 2019 [16]	Bonn	A-E B-E C-E D-E CD-E ABCD-E	99.81 99.25 99.54 98.81 98.61 98.78
M. Madhusmita, et al. in 2019 [17]	NIMHANS	-	99.35
W. Hussain, et al. in 2019 [4]	Freiburg	-	98.26
W. Mardini et al. in 2020 [18]	Bonn	A-E B-E C-E D-E CD-E ABCD-E	100 100 96.6 100 95.5 98.6
O. Kaziha et al. in 2020 [19]	CHB-MIT	-	96.7
Our work	Bonn CHB-MIT	A-E C-E D-E CD-E ABCD-E -	100 100 99 99.2 99.3 96

CHAPTER FIVE

Conclusions and Future Work

5.1.Conclusions

Bonn dataset

Several features have applied that are extracted from the time series of the brain signals. the features are divided into two types: The first type is the six features that are calculated with no delay required and the second type is the fifteen features that need delay for calculation. This is required since applying a data classifier immediately will not be particularly effective. From the result, we find the second type of features that need delay are better than the features that are calculated without delay.

Also the ranking methods have applied to measure the importance of each feature that used in this thesis. From the result, we find many features have height importance in classification such as SDFT, LL, complexity and others and this importance will change when the time of feature calculation changed.

For selecting classifiers: The two types of classifiers black-box (ANN, ANFIS) and non-black-box (DT, RF) have applied. In general, each classifier has its own set of advantages and disadvantages. For this, we looked for the best classifier for our brain signal datasets, one that can extract useful information from large datasets. Based on our experiments, We discovered that the ensemble (ANN) is the best classifier and the RF and ANFIS have approximately the same results when used with the first group of features that need no delay for calculation, and RF is the best classifier when used second group features with one-sec delay and finally RF, DT, ANFIS and ANN has the best results when used the second group features with two-sec delay.

CHB-MIT dataset

In this case, according to the high dataset with 23 channels, one feature which is SDFT has applied for all channels. From the result we find the result from ANN is better than RF and DT.

To summarize, ANN, ANFIS, DT and RF classifier systems demonstrate the ability to be given accurate results as better as the Physician doctor and that would be used by coporating with IOT to give earlier response to the clinic center or to the patient himself before her/his case become more complicated.

5.2.Future Work

In this section, we have proposed that this work can be moved in three main directions in the future, as follows:

a- Applying machine learning techniques for seizure prediction.

b- Designing and implementing IOT kit device that could be connected to the cloud serves or to the clinical center to send an alarm when the patient gets the case by localizing his position.

c- Applying Independent Component Analysis (ICA) for artifact reduction in EEG signals.

References:

[1] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal, and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," *J. Neurosci. Methods*, vol. 191, no. 1, pp. 101–109, 2010, doi: 10.1016/j.jneumeth.2010.05.020.

[2] W. Zhou, Y. Liu, Q. Yuan, and X. Li, "Epileptic seizure detection using lacunarity and bayesian linear discriminant analysis in intracranial EEG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3375–3381, 2013, doi: 10.1109/TBME.2013.2254486.

[3] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, vol. 64, no. 6, p. 8, 2001, doi: 10.1103/PhysRevE.64.061907.

[4] W. Hussain *et al.*, "Epileptic Seizure Detection with Permutation Fuzzy Entropy Using Robust Machine Learning Techniques," *IEEE Access*, vol. 7, pp. 182238–182258, 2019, doi: 10.1109/ACCESS.2019.2956865.

[5] A. J. Wang, S. K. Bick, and Z. M. Williams, "Vagus Nerve Stimulation versus Responsive Neurostimulator System in Patients with Temporal Lobe Epilepsy," *Stereotact. Funct. Neurosurg.*, vol. 98, no. 1, pp. 21–29, 2020, doi: 10.1159/000504859.

[6] 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," *Comput. Biol. Med.*, vol. 100, pp. 270–278, 2018, doi: 10.1016/j.compbiomed.2017.09.017.

[7] N. Mahmoodian, A. Boese, M. Friebe, and J. Haddadnia, "Epileptic seizure detection using cross-bispectrum of electroencephalogram signal," *Seizure,* vol. 66, pp. 4–11, 2019, doi: 10.1016/j.seizure.2019.02.001.

[8] M. Akin, M. B. Kurt, N. Sezgin, and M. Bayram, "Estimating vigilance level by using EEG and EMG signals," *Neural Comput. Appl.*, vol. 17, no. 3, pp. 227–236, 2008, doi: 10.1007/s00521-007-0117-7.

[9] K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 541–552, 2015, doi: 10.1109/TBME.2014.2360101.

[10] L. Wang *et al.*, "Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis," *Entropy*, vol. 19, no. 6, pp. 1–17, 2017, doi: 10.3390/e19060222.

[11] 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," *Biomed. Signal Process. Control*, vol. 39, pp. 94–102, 2018, doi: 10.1016/j.bspc.2017.07.022.

[12] M. Mursalin, Y. Zhang, Y. Chen, and N. V. Chawla, "Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier," *Neurocomputing*, vol. 241, pp. 204–214, 2017, doi: 10.1016/j.neucom.2017.02.053.

[13] A. Narang, B. Batra, A. Ahuja, J. Yadav, and N. Pachauri, "Classification of EEG signals for epileptic seizures using Levenberg-Marquardt algorithm based Multilayer Perceptron Neural Network," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3,

pp. 1669–1677, 2018, doi: 10.3233/JIFS-169460.

[14] S. Ramakrishnan and A. S. Muthanantha Murugavel, "Epileptic seizure detection using fuzzy-rules-based sub-band specific features and layered multi-class SVM," *Pattern Anal. Appl.*, vol. 22, no. 3, pp. 1161–1176, 2019, doi: 10.1007/s10044-018-0691-6.

[15] M. Ravi Kumar and Y. Srinivasa Rao, "Epileptic seizures classification in EEG signal based on semantic features and variational mode decomposition," *Cluster Comput.*, vol. 22, pp. 13521–13531, 2019, doi: 10.1007/s10586-018-1995-4.

[16] C. Mahjoub, R. Le Bouquin Jeannès, T. Lajnef, and A. Kachouri, "Epileptic seizure detection on EEG signals using machine learning techniques and advanced preprocessing methods," *Biomed. Tech.*, vol. 65, no. 1, pp. 33–50, 2020, doi: 10.1515/bmt-2019-0001.

[17] M. Madhusmita, B. Mousumi, P. D. Narayan, and M. S. Kumar, *A novel method for epileptic EEG classification using DWT, MGA, and anfis: A real time application to cardiac patients with epilepsy*, vol. 768. Springer Singapore, 2019.

[18] W. Mardini *et al.*, "Enhanced Detection of Epileptic Seizure Using EEG Signals in Combination With Machine Learning Classifiers," *IEEE Access*, vol. 8, pp. 24046–24055, 2020.

[19] O. Kaziha and T. Bonny, "A convolutional neural network for seizure detection," *2020 Adv. Sci. Eng. Technol. Int. Conf. ASET 2020*, 2020, doi: 10.1109/ASET48392.2020.9118362.

[20] J. H. Cross, *Epilepsy and Epileptic Seizures*. 2013.

[21] R. İnce, S. S. Adanır, and F. Sevmez, "The inventor of electroencephalography

(EEG): Hans Berger (1873–1941)," *Child's Nerv. Syst.*, 2020, doi: 10.1007/s00381-020-04564-z.

[22] G. Gargiulo *et al.*, "A new EEG recording system for passive dry electrodes," *Clin. Neurophysiol.*, vol. 121, no. 5, pp. 686–693, 2010, doi: 10.1016/j.clinph.2009.12.025.

[23] F. Group, *Practical Biomedical Signal Analysis*. 2012.

[24] T. D. Lagerlund *et al.*, "Determination of 10-20 system electrode locations using magnetic resonance image scanning with markers," *Electroencephalogr. Clin. Neurophysiol.*, vol. 86, no. 1, pp. 7–14, 1993, doi: 10.1016/0013-4694(93)90062-Z.

[25] T. N. Lal *et al.*, "Methods towards invasive human brain computer interfaces," *Adv. Neural Inf. Process. Syst.*, 2005.

[26] M. J. Kahana, D. Seelig, and J. R. Madsen, "Theta returns," *Curr. Opin. Neurobiol.*, vol. 11, no. 6, pp. 739–744, 2001, doi: 10.1016/S0959-4388(01)00278-1.

[27] D. Selvathi and V. K. Meera, "Realization of epileptic seizure detection in EEG signal using wavelet transform and SVM classifier," *Proc. IEEE Int. Conf. Signal Process. Commun. ICSPC 2017*, vol. 2018-Janua, no. July, pp. 18–22, 2018, doi: 10.1109/CSPC.2017.8305848.

[28] J. Craley, E. Johnson, and A. Venkataraman, *Integrating Convolutional Neural Networks and Probabilistic Graphical Modeling for Epileptic Seizure Detection in Multichannel EEG*, vol. 11492 LNCS. Springer International Publishing, 2019.

[29] C. C. Jouny, P. J. Franaszczuk, and G. K. Bergey, "Improving early seizure detection," *Epilepsy Behav.*, vol. 22, no. SUPPL. 1, pp. S44–S48, 2011, doi:

10.1016/j.yebeh.2011.08.029.

[30] L. S. Vidyaratne and K. M. Iftekharuddin, "Real-Time Epileptic Seizure Detection Using EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 2146–2156, 2017, doi: 10.1109/TNSRE.2017.2697920.

[31] A. S. M. Murugavel and S. Ramakrishnan, "Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification," *Med. Biol. Eng. Comput.*, vol. 54, no. 1, pp. 149–161, 2016, doi: 10.1007/s11517-015-1351-2.

[32] A. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," *Diss. Massachusetts Inst. Technol.*, pp. 157–162, 2009, [Online]. Available: http://dspace.mit.edu/handle/1721.1/54669.

[33] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal - State-of-the-art and guidelines," *J. Neural Eng.*, vol. 12, no. 3, p. 31001, 2015, doi: 10.1088/1741-2560/12/3/031001.

[34] Z. Lasefr, S. Shiva, V. N. R. Ayyalasomayajula, R. R. Ramasani, and K. Elleithy, "An Efficient Automated Technique and Smartphone Application for Epilepsy Seizure Detection Using EEG signals," vol. 45, p. 9995, 2014.

[35] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with Support Vector Machines," *Clin. Neurophysiol.*, vol. 122, no. 3, pp. 464–473, 2011, doi: 10.1016/j.clinph.2010.06.034.

[36] R. B. Pachori and S. Patidar, "Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 494–502, 2014, doi: 10.1016/j.cmpb.2013.11.014.

[37] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection

in EEGs using time-frequency analysis," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 703–710, 2009, doi: 10.1109/TITB.2009.2017939.

[38] K. Fu, J. Qu, Y. Chai, and Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM," *Biomed. Signal Process. Control*, vol. 13, no. 1, pp. 15–22, 2014, doi: 10.1016/j.bspc.2014.03.007.

[39] A. B. Das, M. I. H. Bhuiyan, and S. M. S. Alam, "Classification of EEG signals using normal inverse Gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection," *Signal, Image Video Process.*, vol. 10, no. 2, pp. 259–266, 2016, doi: 10.1007/s11760-014-0736-2.

[40] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015, doi: 10.1016/j.seizure.2015.01.012.

[41] R. Lyons, "dsp tips & tricks - the sliding DFT," *IEEE Signal Process. Mag.*, vol. 20, no. 2, pp. 74–80, 2003, doi: 10.1109/msp.2003.1184347.

[42] C. S. Park, "Fast, Accurate, and Guaranteed Stable Sliding Discrete Fourier Transform [sp Tips&Tricks]," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 145–156, 2015, doi: 10.1109/MSP.2015.2412144.

[43] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, "A Novel EEG Feature Extraction Method Using Hjorth Parameter," *Int. J. Electron. Electr. Eng.*, vol. 2, no. 2, pp. 106–110, 2014, doi: 10.12720/ijeee.2.2.106-110.

[44] V. N. Gudivada and C. R. Rao, "Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications," *Handb. Stat.*, pp. 197–228, 2018.

[45] J. Cepukenas, C. Lin, and D. Sleeman, "Applying rule extraction & rule

refinement techniques to (Blackbox) classifiers," *Proc. 8th Int. Conf. Knowl. Capture, K-CAP 2015*, 2015, doi: 10.1145/2815833.2816950.

[46] M. K. Siddiqui, M. Z. Islam, and M. A. Kabir, "Analyzing performance of classification techniques in detecting epileptic seizure," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10604 LNAI, pp. 386–398, 2017, doi: 10.1007/978-3-319-69179-4_27.

[47] M. N. Adnan and M. Z. Islam, "Optimizing the number of trees in a decision forest to discover a subforest with high ensemble accuracy using a genetic algorithm," *Knowledge-Based Syst.*, vol. 110, pp. 86–97, 2016, doi: 10.1016/j.knosys.2016.07.016.

[48] Y. L. Pavlov, "Random forests," *Random For.*, pp. 1–122, 2019, doi: 10.1201/9780429469275-8.

[49] S. Soltaninejad, A. Basu, and I. Cheng, "Automatic Classification and Monitoring of Denovo Parkinson's Disease by Learning Demographic and Clinical Features," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 3968–3971, 2019, doi: 10.1109/EMBC.2019.8857729.

[50] S. Shafiee, L. M. Lied, I. Burud, J. A. Dieseth, M. Alsheikh, and M. Lillemo, "Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery," *Comput. Electron. Agric.*, vol. 183, no. 1432, p. 106036, 2021, doi: 10.1016/j.compag.2021.106036.

[51] H. Polat, H. Danaei Mehr, and A. Cetin, "Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods," *J. Med. Syst.*, vol. 41, no. 4, 2017, doi: 10.1007/s10916-017-0703-x.

[52] D. Toradmalle, J. Muthukuru, and B. Sathyanarayana, "Certificateless and

provably-secure digital signature scheme based on elliptic curve," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 4, pp. 3228–3231, 2019, doi: 10.11591/ijece.v9i4.ppxx-xx.

[53] Gant, Katie, et al. "EEG-controlled functional electrical stimulation for hand opening and closing in chronic complete cervical spinal cord injury." *Biomedical Physics & Engineering Express*, vol. 4, no. 6, pp. 065005, 2018.

Appendix

MATLAB and python are used for programming the methods that used in this thesis.

MATLAB R2012b is used for reading the csv Bonn dataset and generating the five cases (A-E, C-E, D-E, CD-E, ABCD-E) , also the windowing this dataset and feature extraction and ANN and ANFIS classifiers are programed with matlab. While the DT and RF classifiers and feature ranking and feature selecting are programed by Python 3.6. For reading the CHB-MIT edf dataset files, feature extraction and the three classifiers (ANN ,DT ,RF) programed with python 3.6. many python libraries are installed for model bulding and features ranking and selecting these libraries are:

- scipy is a scientific and technical computing library.
- numpy is library used for working with arrays
- sklearn is incluse classifiers methods as ANN, DT and RF that used in this thesis.
- sklearn.ensemble is used for features ranking by MDI method
- sklearn.inspection is used for features ranking by permutation method
- mlxtend this library is used for Sequential Feature Selection (SFS) algorithm.
- matplotlib this library is used for plot the figures that represent the importance of each feature.

List of Publication

[1] Amal Salman Abdulhussien, and Ahmad T. Abdulsadda. "New automatic EEG epileptic seizure detection approach using sliding discrete Fourier transform and machine learning techniques" The 2nd Asia Conference on Computers and Communications (ACCC 2021) Singaporepublished, with IEEE conference proceedings and included in IEEE Xplore .

[2] Amal Salman Abdulhussien, and Ahmad T. Abdulsadda. " Automatic epileptic seizure detection in EEGs based on an efficient machine learning technique" 1st International Conference on Advances in Engineering Science and Technology (AEST-2021), Iraq, with the STEPS Ltd, and Liverpool John Moores University, UK.

[3] Amal Salman Abdulhussien, and Ahmad T. Abdulsadda. "Automatic seizure detection for different time delays using SDFT and time-domain features extraction" Journal of Biomedical Research (JBR).

الكشف التلقائي لنوبات الصرع من خلال الاشارات الكهربائية للدماغ بالاعتماد على تقنيات تعلم الآلة الفعالة

الخلاصة

يعتبر الصرع من الامراض الخطيرة التي تصيب حوالي 1 % من مختلف الاعمار من السكان. ان العقل البشري يولد اشارات كهربائية في جميع الاوقات. عند حدوث خلل في الدماغ يؤدي الى توليد اشارات غير طبيعية. يمكن الاستفادة من مخطط الكهربائي للراس Electroencephalogram (EEG) والمخطط الدماغ الداخلي internal Electroencephalogram (iEEG) لمراقبة وتتبع سلوك الاشارات الصرع . وبما ان هذه الاشارات هي قليلة القيم وغير ثابتة ومعقدة فسوف تزداد احتمالية عدم الكشف عن نوبة الصرع . على الرغم من هنالك عدة طرق من الكشف عن الصرع باستخدام الاشارات الكهربائية المقاسة الا انه لا تزال هنالك بعض التحديات مثل الكشف بدقة عالية وسرعة الكشف. في هذه الدراسة تم الاستفادة من التقنيات لمعالجة هذه التحديات و دراسة تأثير الكشف السريع على الدقة.

في البداية تم استخدام نوعين من البيانات (Bonn و CHB-MIT) وتم دراسة الميزات المستخرجة من اشارة EEG وتم تطبيق مجموعتين من الميزات (المجموعة الاولى هي الميزات التي ممكن حسابها بدون تأخير والمجموعة الثانية هي الميزات التي تحتاج وقت لحسابها) هذه الميزات تم استخراجها من اشارة ال EEG مباشرة وكذلك بعد تحويل الاشارة من المجال الزمني الى

المجال الترددي باستخدام نهج جديد وهذه الطريقة هي Sliding Discrete Fourier Transform (SDFT). هذه الميزات المستخرجة من الاشارة تساعد في اكتشاف نوبه الصرع.

المرحلة الثانية تم تطبيق عدة مصنفات من نوع آله التعلم وتحديد افضل مصنف والاكثر ملائمة لاكتشاف النوبات من اشارات ال EEG و iEEG من المصنفات التي طبقت .المرحلة الثالثة هي دراسة وتحليل اهمية كل ميزة واختيار الميزات التي لها اكثر اهمية في اكتشاف النوبات .

تظهر هذه الدراسة أن ميزة SDFT لها أهمية كبيرة لاكتشاف النوبات وان المصنفان RF , ANN افضل من DT,A NFIS وتم ملاحظة عندما يزداد الوقت المطلوب لحساب الميزات سوف يزيد من دقة التصنيف.

الكشف التلقائي لنوبات الصرع من خلال الاشارات الكهربائية للدماغ بالاعتماد على تقنيات تعلم الآلة الفعالة

الأطروحة

مقدمة الى قسم هندسة تقنية الاتصالات كجزء من متطلبات نيل درجة

الماجستير

تقدم بها

امال سلمان عبد الحسين

اشراف

أ.د. احمد طه عبد الساده

2021/11