DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

DETECTION OF EPILEPTIC SEIZURES FROM EEG SIGNALS BY USING DAUBECHIES WAVELETS

by

Şeyma YOL

September, 2022

İZMİR

DETECTION
OF
EPILEPTIC SEIZURES FROM EEG SIGNALS
BY USING
DAUBECHIES WAVELETS

A Thesis Submitted to the Graduate School of Natural and Applied Sciences of Dokuz Eylül University in Partial Fulfillment of the Requirements for Master of Science in Biomedical Technologies

by

Şeyma YOL

September, 2022

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled "**DETECTION OF EPILEPTIC SEIZURES FROM EEG SIGNALS BY USING DAUBECHIES WAVELETS**" completed by **Şeyma YOL** under supervision of **PROF. DR. GÜLAY TOHUMOĞLU** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Gülay TOHUMOĞLU

Supervisor

Prof. Dr. Aydın AKAN

(Jury Member)

Assoc. Prof. Dr. Yavuz ŞENOL

(Jury Member)

Prof. Dr. Okan FISTIKOĞLU

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Prof. Dr. Gülay TOHUMOGLU, for her patience, encouragement, and guidance throughout my master thesis. I would also like to thank my committee, Prof. Dr. Aydın AKAN and Assoc. Prof. Dr. Yavuz ŞENOL, for their valuable time.

I deeply appreciate my family for their kind and emotional support in all my experience. I know is that without my family could not have come this far. Nothing can be compared to their love.

Most importantly I would like to pay my sincere gratitude to Furkan KABA for him faith in my capabilities, understanding, consistent support me to be a happy, strong and hard-working person.

Last but not least, I would like to thank my colleagues, RA Osman Furkan KARAKUŞ, RA Nursac KURT, RA Yazım Beril ULUER, RA Zehra TUFEKCI, RA Aysenur ESER, RA Ezgi YILDIRIM ARSLAN for helping me in all stages of this project.

Seyma YOL

DETECTION OF EPILEPTIC SEIZURES FROM EEG SIGNALS BY USING DAUBECHIES WAVELETS

ABSTRACT

A seizure is a sudden and abnormal activity of brain, caused by disorderly electrical discharge of cerebral neurons. One of the most known seizures is recurrent attacks namely epileptic seizures which result from uncontrolled discharges of nerve cells. A traditional electroencephalogram (EEG) is often helpful to detect this uncontrolled process that indicates epilepsy. The visual detection of the seizures is required noticeable effort and time, especially in the long recordings, therefore alternative methods are needed. In the literature, there have been various epilepsy detection studies using different approaches. However, most of them utilizes patient-specific classifiers.

In this study, wavelet-based algorithm is adopted to different classification methods for detection of epileptic seizures. To detect the effective frequency intervals of epileptic seizure, it is important to reach subbands of signal. Thus, choosing the suitable mother wavelet that resembles epileptic seizure is challenging. Consequently, ideal order Daubechies (db) wavelet becomes a critical point to achieve best performances. Here, it is selected Daubechies wavelets from db2 to db10 which have strong correlation with seizures to decompose the original EEG signal. The algorithm is applied on two different datasets. Eventually, the most appropriate eight features are used for classification model. Various machine learning algorithms namely, Decision Trees, Discriminant Analysis, Naive Bayes, Support Vector Machine (SVM), and k-Nearest Neighbor are used to differentiate epileptic and non-epileptic groups. The results show that db10 gives the best ACC performance of 95.83-100% in beta band using SVM classifier. This method can be a good alternative to the traditional approaches.

Keywords: Seizure, epilepsy, electroencephalogram, Daubechies wavelet, machine learning algorithms

DAUBECHIES DALGACIKLARI KULLANARAK EEG SİNYALLERİNDEN EPİLEPTİK NÖBETLERİN BELİRLENMESİ

ÖZ

Nöbet, serebral nöronların düzensiz elektriksel boşalmasının neden olduğu, beynin ani ve anormal aktivitesidir. En çok bilinen nöbetlerden biri, sinir hücrelerinin kontrolsüz boşalmasından kaynaklanan epileptik nöbet olarak adlandırılan tekrarlı ataklardır. Geleneksel bir EEG, epilepsiye işaret eden bu kontrolsüz süreci tespit etmek için sıklıkla yardımcı olur. Nöbetlerin görsel tespiti, özellikle uzun kayıtlarda, gözle görülür bir çaba ve zaman gerektirir, bu nedenle epileptik nöbetleri tespit etmek için alternatif yöntemlere ihtiyaç duyulmaktadır. Literatürde farklı yaklaşımların kullanıldığı çok sayıda epilepsi tespit çalışması yapılmıştır. Ancak bunların birçoğu hastaya özel sınıflandırıcılar kullanmaktadır.

Bu çalışmada, epileptik nöbetlerin tespiti için dalgacık tabanlı algoritma farklı sınıflandırma yöntemlerine uyarlanmıştır. Epileptik nöbetin etkin frekans aralıklarını tespit etmek için sinyalin alt bantlarına ulaşmak önemlidir. Bu nedenle epileptik nöbetlere benzeyen bir ana dalgacık seçimi zordur. Dolayısıyla, ideal mertebeden bir Daubachies (db) dalgacığı, en iyi performansları elde edebilmek için kritik bir nokta haline gelir. Burada, orijinal EEG sinyalini ayrıştırmak için nöbetlerle güçlü korelasyona sahip olan db2'den db10'a Daubachies dalgacıkları seçilmiştir. Algoritma iki farklı veri kümesine uygulanır. Son olarak, sınıflandırma modeli için en uygun 8 öznitelik kullanılmıştır. Epileptik ve sağlıklı grupları ayırt etmek için Karar Ağaçları, Diskriminant Analizi, Naive Bayes, Destek Vektör Makinesi (DVM) ve k-En Yakın Komşu (k-NN) olmak üzere çeşitli makine öğrenme algoritmaları kullanıldı. Sonuçlar, db10'un DVM sınıflandırıcısını kullanarak beta bandında %95.83-100'lük en iyi ACC performansını verdiğini göstermektedir. Bu yöntem geleneksel yaklaşımlara iyi bir alternatif olabilir.

Anahtar kelimeler: Nöbet, epilepsi, elektroensefalogram, Daubechies dalgacığı, makine öğrenimi algoritmaları

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS

EEG	: Electroencephalogram
db	: Daubechies
SVM	: Support Vector Machine (SVM)
ES	: Epileptic Seizures
ILAE	: The International League Against Epilepsy
CNS	: Central Nervous System
PNS	: Peripheral Nervous System
MRI	: Magnetic Resonance Imaging
CT	: Computerized Tomography
SPECT	: Single Photon Emission Computed Tomography
PET	: Positron Emission Tomography
fMRI	: functional Magnetic Resonance Imaging
ECG	: Electrocardiography
SNR	: signal-to-noise ratio
FT	: Fourier transform
EMD	: Empirical Mode Decomposition
IMFs	: Intrinsic Mode Functions
EEMD	: Ensemble Empirical Mode Decomposition
CEEMDAN	: Complete Ensemble Empirical Mode Decomposition
LMD	: Local Mean Decomposition
DPL	: Dictionary Pair Learning
WT	: Wavelet Transform
ANN	: Artificial Neural Network
DT	: Decision Tree
k-NN	: k-Nearest Neighbor
DWT	: Discrete Wavelet Transform
CWT	: Continuous Wavelet Transform
WPD	: Wavelet Packet Decomposition

CHB-MIT	: Children's Hospital Boston–Massachusetts Institute of Technology
ApEn	: Approximate Entropy
MAD	: Mean Absolute Deviation
RMS	: Root Mean Square
LDA	: Linear Discriminant Analysis
QLDA	: Quadratic Linear Discriminant Analysis
NB	: Naive Bayes
SVM	: The Support Vector Machine
k-NN	: k-Nearest Neighbor
ED	: Euclidean Distance
TP	: True Positives
FN	: False Negatives
TN	: True Negatives
FN	: False Negatives
SEN	: Sensitivity
SPE	: Specificity
ACC	: Accuracy
ROC	: Receiver Operator Characteristic
AUC	: Area Under Curve

LIST OF SYMBOLS

δ	: Delta
θ	: Theta
α	: Alpha
β	: Beta
γ	: Gamma
Ψ_f^Ψ	: Continuous Wavelet Transform
$\Psi(t)$	: Mother Wavelet
γ_2	: Kurtosis
ρ	: Moment
S	: Standard deviation
z	: Mean of the distribution
x_i	: Each point of data
y_k	: Output labels
m	: Mobility
c	: Complexity
ω	: Weight vector

CHAPTER 1

INTRODUCTION

1.1 Motivation

The seizures happen spontaneously in other words, they can occur at any time without any kind of aura. All mammals, including rats, dogs, and cats, are susceptible to this disease, therefore the idea that only people are affected is a common mistake. It offers no clues as to the cause, a specific gender, age, or period of time for epilepsy. It has been estimated that different types of epileptic seizures (ES) are affecting around 70 million people worldwide. These seizures occasionally induce cognitive disorders which can give rise to physical injury to the patient. On the other hand, the people who diagnosed has been with epilepsy sometimes experience emotional distress because of embarrassment and concern about lack of social status. Another risk factor for epilepsy is brain malfunction, which can result in the development of other conditions such as brain tumors, Alzheimer's disease, depression, heart disease, sleep disorders, migraines, cognitive impairment, and mental decline among other conditions. Epilepsy is not just a neurological disorder; it brings other associated diseases with it (Supriya, Siuly, Wang, & Zhang, 2020). Moreover, sudden unexpected death is 24-fold more likely in people with epilepsy as compared to general.

The emergence of machine learning algorithms makes a great contribution of the development to automated epilepsy detection techniques. New algorithms can shorten detection time and improve the accuracy rate. The process of manual visual inspection and interpretation of EEGs requires tremendous time (hours to a day) and high level of technical and analytical skill. Unfortunately, this traditional method is vulnerable to the high observer error rate due to being subject to inter-observer variability and high prevalence (Yol & TohumoĞlu, 2020). More frequency changes are observed in the epileptic EEG segments. Thus, visual examination of these differences is a difficult task due to nonstationary characteristics. In this thesis, it is adopted machine learning algorithms for wavelet-based techniques by using a sequence of Daubachies wavelets to search the best one for all commonly used classification algorithms.

In the literature, there are many research using different wavelet types for the epileptic seizure classification. The choice of the appropriate wavelet type is a very critical point and enhances the performance of the classification. Different sequences of Daubechies are proposed with advantages and disadvantages in these studies. However, there is no performance comparison of db2, db4, db8 and db10 at the same time. That is why we are interested in the following research question: Can we identify the most effective Daubechies wavelet among these types, subband, and machine learning classifier for detection of epileptic and non-epileptic seizures based on patient-free approach. Through study, the performances of chosen Daubechies type, classifiers and sub bands have been compared, it has been explored that the beta band, Daubachies-10 wavelet and SVM outperform other classifiers.

1.2 Nervous System

Interaction between the physical world and the human body is carried out through the nervous system. The Central Nervous System (CNS) and the Peripheral Nervous System (PNS) are the two primary parts of the human nervous system. The brain is one of the crucial components of the central nervous system which controls and regulates all physiological processes as body temperature, blood pressure, breathing, thinking, reasoning etc. The human brain is the most advanced of all species. It is composed of four primary parts: cerebellum, diencephalon, brain stem, and cerebrum which is the largest part (Purves et al., 2001). as shown in Figure 1.1. Higher thought and the special senses are controlled by the cerebrum, smooth body movement is coordinated by the cerebellum, and respiration and heart rate are regulated by the brain stem.

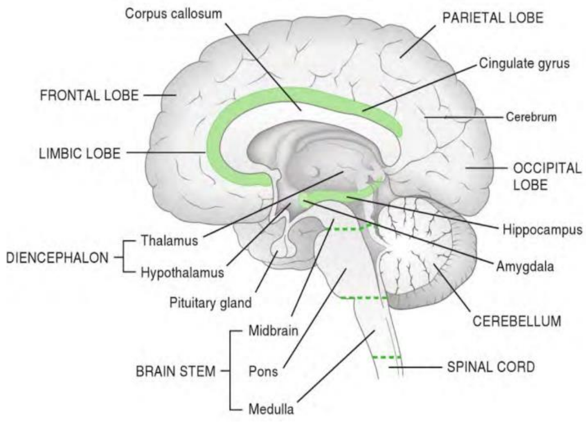

Figure 1.1 Parts of the human brain (Cech, 2011)

The brain has countless nerve cells (neurons) which responsible for carrying electrical impulses named as action potential throughout the body. This electrochemical activity of neurons connects them together for receiving or sending out information. Intersections called synapses provide this connection.

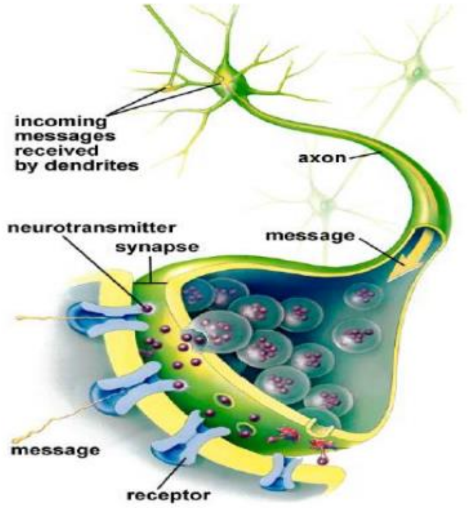

Figure 1.2 Representation of the communication among two neurons by synapse (Darbas & Lohrengel, 2019)

The above Figure represents how information flows from one neuron to another with the help of voltage-gated-sodium-channels. Each nerve cells consist of three basic parts: cell body (soma), axon and dendrites. The dendrite is a short part that carry impulses towards cell body. The metabolic activity of the neuron occurs in the cell body. The axon has a single thread-like structure that carries the impulse away from the cell body to other cells. The chemical impulses known as neurotransmitters are converted to electrical signals. Any disruption in this communication can result in several brain disorders.

1.3 Seizures

Seizure is excessive electrical discharges of nerve cells that interrupt normal function of CNS. They are associated with transient alteration in motor activity, consciousness, or sensation and uncontrolled activity of the brain (Bowman, Dudek, & Spitz, 2001). The type of seizure depends on part of the brain in which the seizure occurs.

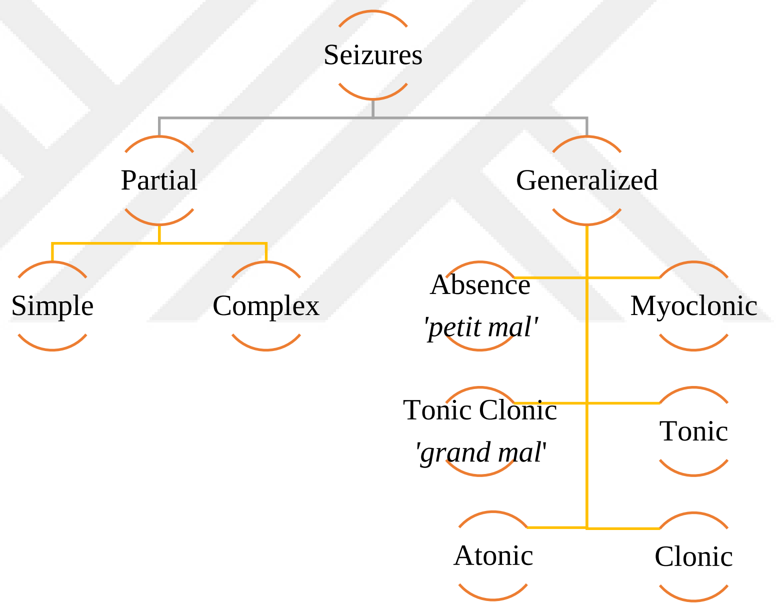

Figure 1.3 Classification of seizures according to ILAE (Angeles, 1981)

Seizures are classified into 2 main groups; first one is partial seizures which originate in one part of the brain. Second one is generalized seizures which are involved the whole brain (Davis & Pirio Richardson, 2015) as presented in Figure 1.3. In this study, we deal with epileptic seizures using EEG signals.

1.4 Epilepsy

Epilepsy word comes from Greek Word ‘epilambanein’ which means “to seize or attack” (Sharmila, Aman Raj, Shashank, & Mahalakshmi, 2018). The uncontrolled excessive activity of the CNS is knowns as epilepsy that affects people of all ages. It is a well-known chronic neurological condition characterized by recurrent epileptic seizures resulting from uncontrolled neuronal discharges in the CNS. The brain has millions of nerve cells that control the way we act, think, or feel, by passing electrical signals throughout your whole body. If these signals are disrupted, it causes an epileptic seizure (Zandi, Javidan, Dumont, & Tafreshi, 2010). All seizures start in the brain, and people with epilepsy have repetitive electrical brain activity disturbances. About 70 million people suffer from epilepsy in the world (Yaffe et al., 2015; Zandi et al., 2010).

A wide variety of clinical signs have been observed, they may be a brief lack of awareness, while there may also be a major motor convulsion(Tohumoglu & Yol, 2021). The uncontrollable bouts of epilepsy are a defining characteristic. Epileptic attacks are accompanied by loss of consciousness, disturbances of autonomic, mental functioning, muscle spasms and movement disorders (Pathak, Waheed, & Mirza, 2016). Epilepsy may be considered incurable; however, it can be controlled with medications. If the patients with epilepsy do not respond to the drugs, the surgery can be considered as an option. Epilepsy has no sexual, geographical, or social boundaries. All ages can be affected but it is generally diagnosed in infancy, childhood, adolescence, and old ages. Unknown genetic or biological predisposition, drugs, trauma, or tumor can cause epilepsy.

1.4.1 Diagnosis of Epilepsy

One important aspect of epilepsy research is that the earlier diagnosis is made, the better result is obtained. Because many patients experience critical morbidity from the incorrect diagnosis such as unnecessary parenteral medications, taking these drugs for years, and changing their life choices according to wrong diagnosis. All of them lead to reduce in quality of life and survival rate. Unfortunately, it is impossible to predict when an epileptic seizure will occur, and we still don't completely understand how they process. Therefore, it is vital to recognize and correctly diagnose the kind of epileptic seizure in order to effectively treat the condition. Diagnostic tests comprise variable methods: accurate pharmacological diagnosis, genetic screening methods, blood tests, physical examination are among these options (Fernández, Loddenkemper, Gaínza-Lein, Sheidley, & Poduri, 2019; Malmgren, Reuber, & Appleton, 2012; Sutton et al., 2020).

Although epilepsy is easy to diagnose, doubts emerge in routine clinical practice. EEG is one of the most widely used techniques to identify the nature of seizure and confirm a diagnosis. If the initial drugs are insufficient to control seizures, a Magnetic Resonance Imaging (MRI) examination may be recommended, or a Computerized Tomography (CT) scan if patient is an older adult. There are various neuroimaging techniques to monitor brain functions such as EEG, Single Photon Emission Computed Tomography (SPECT), Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI) with good spatial resolution. In contrast, EEG has relatively poor spatial resolution. However, EEG is very desirable, informative approach and the most popular method of them is EEG with the high temporal resolution (He, Baxter, Edelman, Cline, & Wenjing, 2015).

Signal processing of EEG is an important tool in the diagnosis of epilepsy. On the other hand, visual interpretation is very subjective, making it conceivable for different people to have different interpretations of the same recordings. For this reason, computer-aided extraction and analysis of EEG signal characteristics are extremely helpful in diagnosis (Akbari & Esmaili, 2020; Y. Wang et al., 2021).

1.5 Electroencephalogram

The Electroencephalogram (EEG) is an essential tool for correct diagnosis of epilepsy. Interpretation of EEG helps determine the type of seizure which is examined by neurologists to detect and classify the patterns of the disease. Since routine EEG is economical, harmless, easily accessible, and applicable, it is often preferred diagnostic method for various neurological and psychiatric diseases apart from epilepsy. It may also play a role in choice of antiepileptic drugs, documentation of seizures that the patient is unaware of, examination of response to treatment, and initial assessment of invasive other treatments (Perucca & Tomson, 2011; Wirrell, 2010).

The EEG is painless and noninvasive tool to read electrical activity of the brain. The voltage range is between 3-100 μV which is 100 times weaker than Electrocardiography (ECG) signals (Kaur & Singh, 2012). The EEG represents the electrical activity of the brain and gives information about any abnormalities. The frequencies of these signals range from 0.5 to 100 Hz. Spikes, sharp and polyspike waves is generally interpreted as epileptic waveform. EEG have non-stationary characteristic and poor signal-to-noise ratio (SNR) (Yan et al., 2016).

In 1924, it is used in humans by Hans Berger who discovered alpha waves having 10 Hz (Zeidman, Stone, & Kondziella, 2014). The EEG uses the 10-20 international system of electrode placement that gives a brief description of the scalp electrode placement. The 10 and 20 numbers represent the distance (approximately 10% or 20%) between electrodes as seen in Figure 1.4. The EEG represents the electrical activity of the brain and give information about any abnormalities.

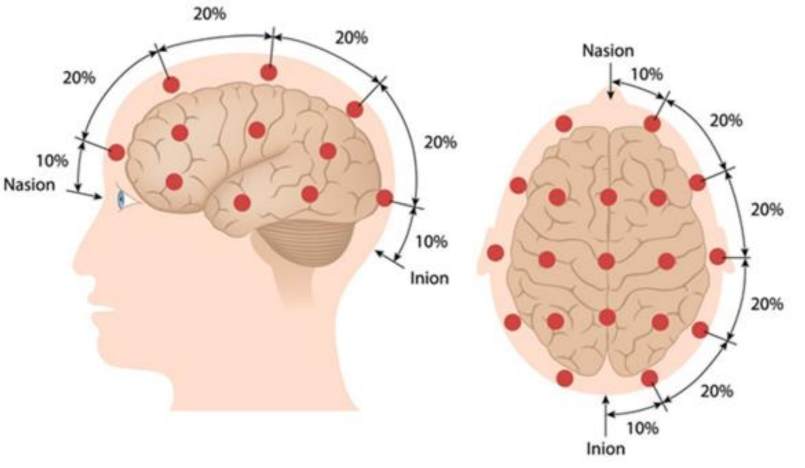

Figure 1.4 Left side image and right side images represent the left side of the head and top view of head, respectively in the international 10-20 system (Graimann, Allison, & Pfurtscheller, 2010)

The brain highly sophisticated system that includes billions of interconnected neurons. To comprehend the characteristics and dynamics of these neurons, signal processing methods are needed. In the biomedical research area, automatic EEG processing techniques have been utilized since the early 1960s. Therefore, automatic detection methods have a substantial advantage including saving of time, rapid diagnosis, remote and continuous monitoring, and reduced cost of medical treatment over the traditional manual. The EEG analysis methods are divided into time domain, frequency domain, time-frequency domain, and nonlinear methods. The ideal method is changed according to characteristic of signal. The signal analysis generally consists of same steps such as consists of data acquisition, signal pre-processing, feature extraction, feature selection, and classification. The most commonly used data is electroencephalogram in the data acquisition step.

1.5.1 Frequency Subbands of EEG

In epileptic research studies, original the EEG is commonly decomposed into five sub-bands: *gamma* (30-60 Hz), *beta* (13-30 Hz), *alpha* (8-12 Hz), *theta* (4-8 Hz), and *delta* (0-4 Hz) as shown in Figure 1.5 and 1.6 where L represent decomposition level. We extracted features using the wavelet transform because to its superior resolution in the time-frequency domain. The subbands provide more specific details about the underlying neural processes of the EEG. Due to the fact that certain alterations cannot be seen in the original full-spectrum EEG, each sub-band is analyzed independently.

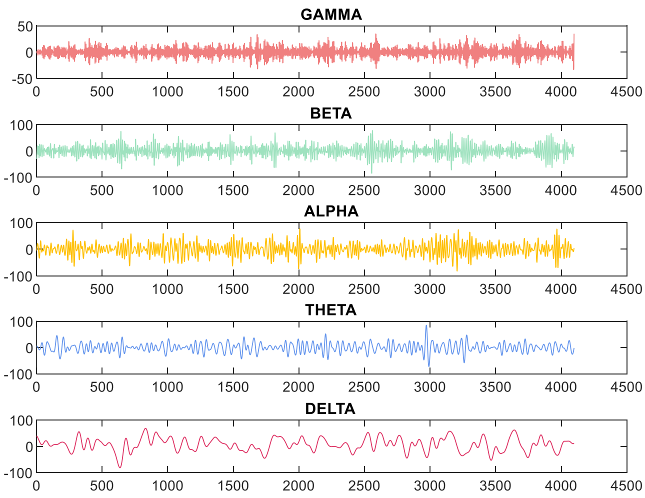

Figure 1.5 Decomposition of a seizure free EEG with 5 levels of details (L=5, db=4)

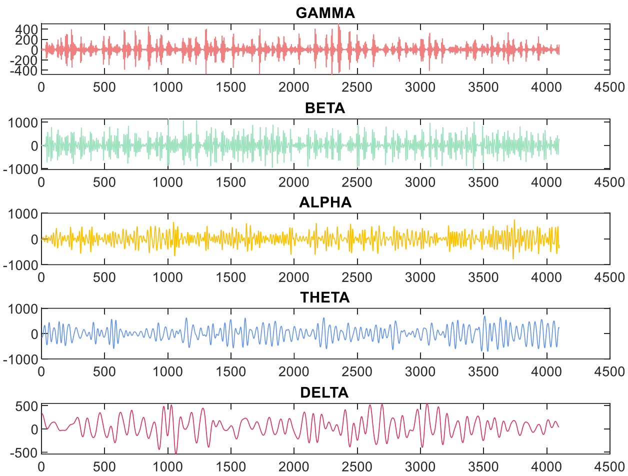

Figure 1.6 Decomposition of an epileptic EEG with 5 levels of details (L=5, db=4)

Here, EEG signal subband frequencies are defined as;

- Delta (δ) are quite slow rhythms, lower than 4 Hz. They are primarily characterized by deep sleep, and pathological conditions such as coma or cancers.

- Theta (θ) rhythms: These are oscillations in the 4-8 frequency band, slightly faster waves. The rhythms are visualized mainly during drowsiness and in children's temporal and parietal regions.

- Alpha (α) waves are approximately 8-12 Hz and appear mainly in the occipital and parietal lobes, the back of the brain. They are generally observed during relaxed states of wakefulness or closed eyes. The asynchronism of the alpha waves is generally interpreted as increased availability of networks or cortical sensory input to the motor command. Several studies on alpha source localization showed that alpha oscillations are obtained from thalamocortical neurons (Da Silva, Vos, Mooibroek, & Van Rotterdam, 1980; Hughes et al., 2004; Lőrincz, Kékesi, Juhász, Crunelli, & Hughes, 2009).

- Beta (β): These waves are relatively fast, generally observed in 13-30 frequency band. These oscillations are generated by the parietal and frontal part of the cortex, associated with e performance of movements, in the motor areas (Pfurtscheller & Neuper, 2001). The most of epileptic seizure activities occur at a frequency ranging from 3 to 25 Hz referred to as Beta activity. Thus, analysis of beta subband is one of the most used methods in the epileptic detection studies.

- Gamma rhythms (γ) consist of all frequencies above 30 Hz. These waves are related to various cognitive and motor functions.

1.6 Review on Classification of Epileptic Seizure Using EEG Signal

A comprehensive literature review of epileptic detection methods will be presented in this section. In the literature, there are number of studies related with epileptic signal analysis based on time-domain, frequency-domain, time-frequency domain. Processing of frequency-based signals is not effective by using simple methods, such as Fourier transform (FT) that produces spectral features. However, FT assumes that the analyzed signal is stationary, so the Fourier transform succeeds in stationary signals, the same success cannot be achieved in non-stationary signals. Another disadvantage is that FT does not contain any time information. Different strategies exist for addressing the non-stationary EEG feature. Over the past decades, different methods have resorted to detection and classification epileptic signals.

Empirical Mode Decomposition (EMD) is used to classify epileptic seizures with the help of spectral energy, spectral peaks and spectral entropy (Martis et al., 2012). Original signal is decomposed into oscillatory components named as Intrinsic Mode Functions (IMFs) by using EMD method which is a time series-based transformation, successfully gives outcomes to decompose the nonlinear and non-stationary data. Despite widespread use of the methods, there are numerous limitations, including the need for more mathematical knowledge and the finding of extremal points and mode mixing problems. This problem of EMD which causes different oscillations in the same mode or the same oscillations in different modes result from the signal intermittency. The improved version of EMD is Ensemble Empirical Mode

Decomposition (EEMD)(Bizopoulos, Tsalikakis, Tzallas, Koutsouris, & Fotiadis, 2013), Complete Ensemble Empirical Mode Decomposition (CEEMDAN) with statistical features (Hassan & Haque, 2015), Local Mean Decomposition (LMD)(Zhang & Chen, 2016), which is proposed to alleviate this shortcoming. Noise assisted sort of EMD is EEMD approach that consists of adding low amplitude white noise to signal to provide a true decompositions frame in the time–frequency domain (Al-Subari et al., 2015). In the several automatic detection methods, these methods are very popular. In their work, researchers of (Yu, Li, Yuan, & Zhou, 2018) propose different approaches combining novel LMD and Dictionary Pair Learning (DPL). Their approach suggests that their method can be used as potential tool in the clinic with sensitivity of 95.89% with specificity of 95.10%. The authors of (Bari & Fattah, 2020) extract statistical and spectral features from normalized IMFs using the CEEMDAN with the reported 100% accuracy. It is combined the Ensemble Empirical Mode Decomposition and Least Squares Support Vector Machine to classify the seizure and non-seizure EEG with an accuracy of 94.7% (Torse & Khanai, 2021).

The performance of the gamma band in an EEG signal is analyzed using another version of the Fourier Transform, namely the Short Term Fourier Transform, which analyzes local features using a variable window size (Sameer, Gupta, Chakraborty, & Gupta, 2019). The disadvantage of this method is that the used window to obtain frequency information is not robust. This drawback is non-negligible issue in terms of resolution. While a narrow window provides better time resolution but worse frequency resolution, a broader window provides better time resolution but worse frequency resolution. Wavelet Transform (WT) plays a vital tool in analysis of non-stationary signals. Because of transient characteristics of EEG in nature, capture of small details and sudden changes in the signals get hard. The flexible analysis of WT overcomes these drawbacks. In the wavelet transform, time-domain EEG signals are converted to time-frequency localization by means of variable window size.

One of the most extensively utilized methods for epilepsy diagnosis is the Wavelet Transform. In studies of (Adeli, Zhou, & Dadmehr, 2003), wavelets are likened to a mathematical microscope. It is considered that the lower order wavelets of the family are too rough to show EEG spikes appropriately. Otherwise, more oscillations are observed in the higher order ones and the spiky form of the absence of epileptic seizure cannot be represented in their research.

In the time-frequency analysis of epileptic seizure signals, the different types of wavelet functions namely Haar, Discrete Meyer, Daubechies, Coiflets, Reverse biorthogonal, Symlets, and Biorthogonal are most widely used. One of these wavelets is Daubechies wavelet which has been considered as main wavelet to extract time-frequency features. It is recommended that the best suitable wavelet function is Daubechies of order 4 for epileptic EEG signal analysis (Guo, Rivero, Dorado, Munteanu, & Pazos, 2011).

The comparative study of wavelet families-Haar, Daubechies (orders 2–10), Coiflets (orders 1–10), and Biorthogonal (orders 1.1, 2.4, 3.5, and 4.4) is achieved to identify the most appropriate wavelet function for EEG analysis (Gandhi, Panigrahi, & Anand, 2011). The features obtained from subbands are fed into Probabilistic Neural Network. They notice that the Coiflet of order 1 is the most suitable member of the current wavelet families, offering both higher classification accuracy and lower computing cost in comparison to Daubechies 2/3.

In some studies, Daubechies 4 is one of the most successful wavelets owing to filter lengths. All mother wavelets have different filter lengths. It is important that selection of longer wavelet filters' length should take into account because of higher the computational cost (Abbate, DeCusatis, & Das, 2012). On the other hand, different wavelet families such as Symlet, Daubechies, and Coiflet wavelets are proposed to classify and detect epileptic seizures. There are several elements that determine the characteristics of these families, including regularity or symmetry, the number of vanishing points, and the length of a wavelet's support. (Ngui, Leong, Hee, & Abdelrhman, 2013).

Using Daubechies 2 (db2), Daubechies 4 (db4), and Daubechies 8 (db8), the wavelet coefficients are computed in the study of (Orosco, Correa, & Laciar, 2013).

Daubechies sequence is mostly used due to better stability and more flexible options for weighing boundary problems. The analysis of EEG signals using Daubechies Wavelets is introduced in many studies (Ayyad, Saleh, & Labib, 2019; Faust, Acharya, Adeli, & Adeli, 2015; Gazalba & Reza, 2017; Gu, Yan, Zhang, Li, & Yu, 2018; Nabil, Benali, & Reguig, 2020; L. Wang et al., 2017). One of these studies is achieved by (Juárez-Guerra, Alarcon-Aquino, & Gomez-Gil, 2015) In the feature extraction step, they use the two types of wavelet transform that is DWT and the Maximal Overlap Discrete Wavelet Transform with the Haar, db2 and db4. The computed features as mean, absolute median and variance from the Alpha and Delta subbands fed into Feed-Forward Artificial Neural Networks classifier. Accuracy rate of 93.23 % is achieved using whole segments whereas accuracy of 99.26 % is obtained using sub-segments for training.

In some research articles, because of the spike-wave pattern characteristic of EEG, it is thought that Daubechies-2 wavelet function is appropriate (Fathima, Bedeeuzzaman, Farooq, & Khan, 2011; Tong, Aliyu, & Lim, 2018). New family of wavelets is constructed with generalized Laguerre polynomial wavelets (Peachap & Tchiotsop, 2019). They obtain a scalogram of the signal and CWT coefficients with the Laguerre 1-5 wavelets and dimension of features is reduced with the help of Principal Component Analysis to extract significant information from data. Dimensionality reduced features are classified by using SVM and Artificial Neural Network classifiers. The classification rate changes between 95%-100%. The researchers of (Ahmad, Singh, & Khan, 2019) achieve the sequential segmentation of EEG signals to detect epileptic seizures using Machine Learning. The approximation and detail coefficients are obtained by db4 and decomposed to frequency sub-bands. They compare the result of the different classifiers including Simple Decision Trees, Bagged Trees, Quadratic Discriminant, Subspace k-Nearest Neighbor, Medium Gaussian SVM. They observe that Simple Decision Tree has high sensitivity and specificity rate in percentage.

The most suitable wavelet function is identified among the wavelet families (second, fourth, sixth tenth and twelfth order of Daubechies) in (Anila Glory, Vigneswaran, & Shankar Sriram, 2020). In their classification methodology, the entropy-based features (Renyi, Sample, Shannon and Permutation) are extracted by employing Bonn University EEG Dataset. Minimum entropy criterion and reconstruction criterion are utilized to select the appropriate feature. These relevant features are classified using Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Tree (DT) and k-Nearest Neighbor (k-NN) and the result are compared. It is found that db10 is optimal basic wavelet function with the highest accuracy, sensitivity, specificity rate and improve epileptic detection performance.

It is compared that the performance of discrete wavelet transform (DWT) and wavelet packet decomposition (WPD) to classify normal, generalized epilepsy and focal epilepsy (Sairamya, Premkumar, George, & Subathra, 2021). Their attempt to combinate of three factors; wavelet coefficient features, decomposition level, and selection of mother wavelet. Using DWT and WPD, the acquired from the Karunya EEG database are divided into sub-bands. Nine statistical features acquired from each subbands are classified using SVM. They use 54 different wavelet types (Bior, Coif, Rbio, Haar, db, Sym, Dmey and derivatives). In each situation, the highest classification accuracy is attained by the Haar, reverse biorthogonal and wavelet member biorthogonal. According to their findings, the best approach for analyzing EEG data to diagnose epilepsy is DWT. Furthermore, the best wavelet type for signal classification is the mother wavelet rbio1.1, which has the best classification rate.

The authors of (Omidvar, Zahedi, & Bakhshi, 2021) use 5-level decomposition, Daubechies 4 (db4) that combine with different statistical and entropy-based features (9-11 main features). They use only db4 type of wavelet and did not compare the other wavelet types. The classification is performed by SVM and ANN classifiers which contain genetic algorithms to select the more suitable features. Two classifiers give the 100% and 98.7% accuracy rates between seizure and normal classes, and seizure, normal, and seizure-free classes.

The authors of (El-Gindy et al., 2021) develop a method for predicting epileptic seizures based on several wavelet transform families (Haar, Daubechies (db4, and db8), Symlets (Sym4), and Coifets (Coif)). They use the amplitude, local mean, local median, local variance, derivative, and entropy of the wavelet-transformed signals. Their results show that db4 gives the high sensitivity of 100% with low average False Prediction Rate of 0.0818 h−1 and a high average Prediction Time of 38.1676 min.

As mentioned above, the different wavelet types are applied with various classification algorithms in the literature. In order to more clearly compare its effectiveness, we chose the type of Daubechies most frequently used in other studies. Selected wavelets namely db2, db4, db8 and db10 that are combined with 8 features is tested with 5 different classification methods. In addition to above studies, there are many epileptic detection studies that are used Bonn University and Children's Hospital Boston–Massachusetts Institute of Technology (CHB-MIT) EEG Database as shown in Table 1.1 and Table 1.2.

Table 1.1 Comparison of different classifier performances for Bonn University Dataset

Authors	Features	Classifier	Performance metrics (%)
(Shoeibi et al., 2022)	Fuzzy entropy features	ANFIS classifier Autoencoders	ACC: 99.74-99.46
(Qureshi, Afzaal, Qureshi, & Fayaz, 2021)	Temporal-Spectral Features	Fuzzy Classifiers	ACC: 98.70-99.81
(Anuragi, Sisodia, & Pachori, 2021)	line-length, log-energy-entropy, and norm-entropy	LS-SVM SVM k-NN Ensemble Bagged Tree	ACC: 100
(Akyol, 2020)	Not performed	Deep Neural Networks Stacking ensemble approach	ACC: 97.17 SEN: 93.11
(Jiang, Chen, & Li, 2020)	symplectic geometry decomposition-based features	SVM	ACC: 99-100
(Atal & Singh, 2020)	ECT+MGT+NPT+FD+GLCM	Random Forest	ACC: 99-98
(Amin, Yusoff, & Ahmad, 2020)	Discrete Wavelet +Arithmetic coding	k-NN, Naive Bayes, MLP, and SVM	ACC: 100
(Mahjoub, Jeannès, Lajnef, & Kachouri, 2020)	MEMD + Temporal, Non-linear features	SVM	ACC: 99.54-98.61
(Ravi Kumar & Srinivasa Rao, 2019)	Differential entropy Peak magnitude to Root Mean Square ratio	Random Forrest Classifier	SEN: 93.33 SPE: 96.67
(V. Gupta & Pachori, 2019)	FBSE + WMRPE	LS-SVM	ACC: 99.5-97.5
(Acharya, Oh, Hagiwara, Tan, & Adeli, 2018)	Not performed	CNN	SEN: 95 SPE: 90
(Ullah, Hussain, & Aboalsamh, 2018)	Not performed	CNN	SEN: 98 SPE: 98
(A. Gupta, Singh, & Karlekar, 2018)	Hurst Exponent and ARMA Parameter	SVM	SEN: 95 SPE: 94 ACC: 94
(Mingyang Li, Chen, & Zhang, 2017)	Mean Energy Standard deviation Max value	NNE	ACC: 98.78

[1]enhanced curve let transform (ECT), Modified Graph Theory (MGT), Novel Pattern Transformation (NPT), fractal dimension (FD) and Gray Level Co-occurrence Matrix (GLCM), Multilayer Perceptron (MLP), multivariate empirical mode decomposition (MEMD), weighted multiscale Renyi permutation entropy (WMRPE), Fourier–Bessel series expansion (FBSE), least square support vector machine (LS-SVM), Neural Network Ensemble (NNE)

Table 1.2 Comparison of different classifier performances for CHB MIT Dataset

Authors	Features	Classifier	Performance metrics (%)
(Qureshi et al., 2021)	Temporal-Spectral Features	Fuzzy Classifiers	ACC: 92.79-99.38
(Anuragi et al., 2021)	line-length, log-energy-entrop, and norm-entropy	LS-SVM SVM k-NN Ensemble Bagged Tree	ACC: 99.84
(Gómez et al., 2020)	Not performed	Fully Convolutional Neural Networks	ACC: 98.0 SPE: 98.3
(Jiang et al., 2020)	symplectic geometry decomposition-based features	SVM	ACC: 97.17-99.72
(Choi et al., 2019)	Multi-scale 3D-CNN	Deep Neural Network	SEN: 89.4
(Tăuţan et al., 2019)	Amplitude, skewness, kurtosis, entropy, maxPSD, maxF, mean Gamma, mean Beta, mean Theta, mean Delta, varPSD	SVM, RF	ACC: 94
(Kaleem, Gurve, Guergachi, & Krishnan, 2018)	signal-derived EMD-based dictionary approach	SVM	SEN: 94.27 SPE: 91.55 ACC: 92.91
(Tsiouris, Markoula, Konitsiotis, Koutsouris, & Fotiadis, 2018)	Spectral analysis, variation in EEG energy distribution over the delta, theta, and alpha rhythms	SSM	SEN: 88
(Bhattacharyya & Pachori, 2017)	Empirical Wavelet transform, þ joint instantaneous amplitudes and frequencies	Random Forest	SEN: 97.91 SPE: 99.57 ACC: 99.41
(Janjarasjitt, 2017)	Mean, Standard Deviation	SVM	SEN: 72.99 SPE: 98.13 ACC: 96.87
(Van Esbroeck, Smith, Syed, Singh, & Karam, 2016)	Temporal variability information	SVM	SEN: 100
(Orosco, Correa, Diez, & Laciar, 2016)	Spectral and energy features	LDA Pattern Recognition Neural Network	SEN: 87.5 SPE: 99.5
(Bugeja, Garg, & Audu, 2016)	Magnitude, spectral energy variation, and relevance frequency	SVM ELM	SVM: - SEN: 97.98 SPE: 89.90 ELM: - SEN: 99.48 SPE: 81.39

[2]Linear Discriminant Analysis (LDA), Extreme Learning Machine (ELM)

1.7 Contribution

Researchers showed that different wavelet types are the most appropriate wavelet for use in seizure detection. However, the potential advantages of each wavelet and detailed results aren’t compared in detail. Thus, this creates a gap in this area having a research importance in the literature. This is the motivation for this study.

The purpose of this study is to find the best effective Daubechies types and sub-band type to classify the epileptic and nonepileptic EEG signal and understand the general perspective of different Machine Learning Algorithms that can be associated with feature extraction process. Though Daubechies has shown promising results, it is still an open question regarding which wavelet has the differential feature, and also which is the best suitable for the classification of epileptic and non-epileptic seizures.

Contextually,

1. The EEG signal is decomposed in 5-levels decomposition by using db2, db4, db8, and db10-WT to extract specific information and eliminate redundant data.
2. Accurate epilepsy diagnosis by integration of time-frequency domain characteristics
3. To combine signal processing algorithms and machine learning techniques. SVM, K-NN, Decision Tree, Discriminant Analysis, and Naive Bayes are among the classifiers used to extract features and classifying them into several categories.
4. Employing different machine learning algorithms and the performance of each algorithm is compared in terms of accuracy, sensitivity, specificity
5. The experimental results propose low computational burden. It is built highly fast, sensible, robust, cost-effective epilepsy detection algorithm that achieves the results with high accuracy rate without patient-dependent classification. Our algorithm offers more precise detection of epileptic EEGs from nonepileptic ones.

CHAPTER 2

MATERIALS AND METHODS

In Figure 2.1, to explain the whole process the block diagram is shown. The procedure covers the pre-processing, feature extraction and training-test steps of machine learning algorithms and classification of EEG signal for epileptic seizures. In the preprocessing, the Wavelet Transform is applied to the signal with the Daubechies mother wavelet. The signal subbands are obtained for 5-level. In the feature extraction step, eight features are applied to all five subbands. After all features are computed, the feature matrix is fed into classifiers.

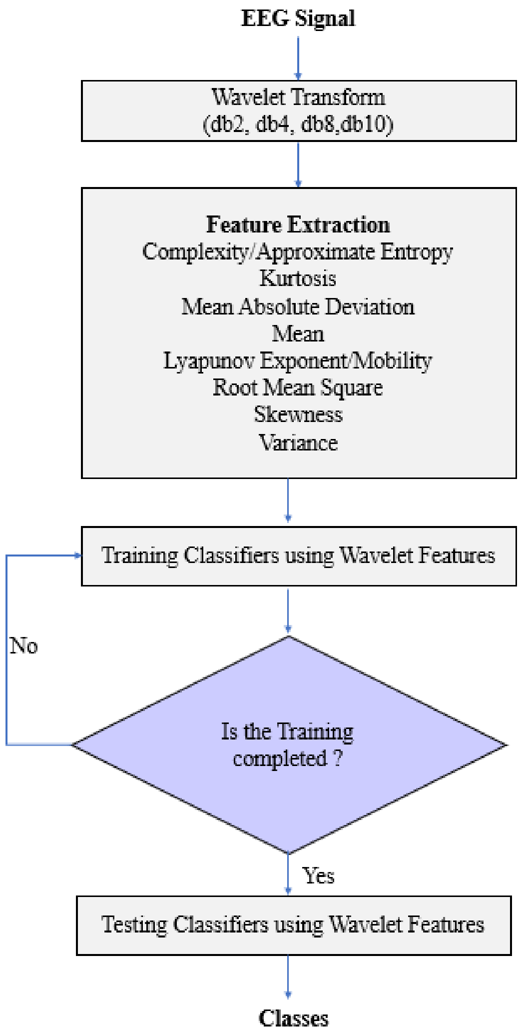

Figure 2.1 Overall system architecture for EEG signal classification

2.1 EEG Database

To evaluate the performance of our proposed technique, we apply it to two distinct EEG databases: the Bonn University EEG Dataset and the CHB-MIT EEG dataset. Unlike The Bonn Dataset, which includes normal, interictal, and ictal EEG recordings, the CHB MIT dataset includes long interictal and ictal EEG recordings.

2.1.1 CHB-MIT Dataset

This study utilizes a dataset contributed by Children's Hospital Boston–Massachusetts Institute of Technology (CHB MIT). Extracranial EEG signals from 23 children with an intractable seizure are obtained using a standard 10-20 electrode placement scheme. Males ranges in age from 3 to 22 years, while girls' range in age from 1.5 to 19 years. To detect their seizures, they are observed for many days following the cessation of anticonvulsant medication. Continuous scalp recordings are collected with a sampling rate of 256 Hz per second and 16-bit resolution. To detect their seizures, they are followed for many days following the cessation of a nticonvulsant medication. Continuous scalp recordings are collected with a sampling rate of 256 Hz per second and 16-bit resolution. In total, each EEG segments consist of 91600 sampling points and lasted around 1 hour.

The EEG database uses in this study is accessible at the Physionet website; http://physionet.org/physiobank/database/chbmit/ (Shoeb & Guttag, 2010). In the experiment, all patients have experienced 2-14 seizures and the total seizure number is 198 seizure events from 23 patients that are listed in Table 2.1. When anti-seizure medication is stopped, they are followed for a few days and the continuous scalp recordings are taken at a sampling rate of 256 Hz per second and 16-bit resolution to identify their seizures.

Table 2.1 CHB-MIT dataset for 23 subjects

Cases	Gender	Age (Years)	Total number of records	Number of seizures
chb01	F	11	42	7
chb02	M	11	36	3
chb03	F	14	38	7
chb04	M	22	42	3
chb05	F	7	39	2
chb06	F	1.5	18	7
chb07	F	14.5	19	3
chb08	M	3.5	20	5
chb09	F	10	19	3
chb10	M	3	25	7
chb11	F	12	35	3
chb12	F	2	24	13
chb13	F	3	33	8
chb14	F	9	26	7
chb15	M	16	40	14
chb16	M	7	19	5
chb17	M	12	21	3
chb18	M	18	36	6
chb19	M	19	30	3
chb20	M	6	29	6
chb21	M	13	33	4
chb22	M	9	31	3
chb23	M	6	09	3
chb24	F	-	22	12

2.1.2 Bonn University Dataset

EEG signals are acquired from an online database available from Epilepsy Center of the Bonn University Hospital of Freiburg [5]. There are five groups within this database. (A, B, C, D, E). They are single-channel group with 100 samples which are 23.6 seconds. Each data sample contains 4096 data points, and the sampling frequency is 173.61 Hz. In addition, the data are filtered between 0.5-85 Hz. A and B data set is obtained from five healthy volunteers. The eyes of the volunteers in A are open while the eyes of the volunteers in B are closed. C, D, E are records of five epilepsy patients that are described in Table 2.2. During the EEG signal measurement of the C class subjects, the electrodes placed on the skull are taken from the moment when they are not in a seizure state with the electrode placed across the epileptogenic region. Class D subjects are taken with electrodes placed in the epileptogenic region when they are not in a seizure state. Class E samples are taken while the patients are having seizures.

Briefly;

Table 2.2 Clinical data brief of Bonn University

Data Set	A set (Z data)	B set (O data)	C set (N data)	D set (F data)	E set (S data)
Subjects	Health individual	Health individual	Patient with epilepsy	Patient with epilepsy	Patient with epilepsy
Electrode type	Surface	Surface	Intracranial	Intracranial	Intracranial
Electrode placement	International 10-20 systems	International 10-20 systems	Within epileptogenic zone	Within epileptogenic zone	Within epileptogenic zone
State of patient	Awake and eyes open	Awake and eyes open	Seizure-free (Interictal)	Seizure-free (Interictal)	Seizure activity (Ictal)
Number of segments	100	100	100	100	100
Segment duration	23.6	23.6	23.6	23.6	23.6

2.2 Pre-Processing

The processing step is started with raw signal entered into band-pass filter to remove artifacts. The boundary of filter is between 0.3 and 60 Hz (CHB MIT dataset). 0-40 Hz frequency band range gives more significant information for epileptic seizure detection. The Bonn dataset is filtered using 40 Hz low pass filter, so it is not applied the filter for this dataset.

2.2.1 Channel Selection

The rhythmic activity that occurs during epileptic seizures is not prominent in all channels, and the channel varies significantly from patient to patient. Therefore, knowing which channel we will work with in feature extraction is an important point. In order to identify significant channel that carries rhythmic activity and reduce system complexity, using smaller number of channels is contributive step for multichannel data.

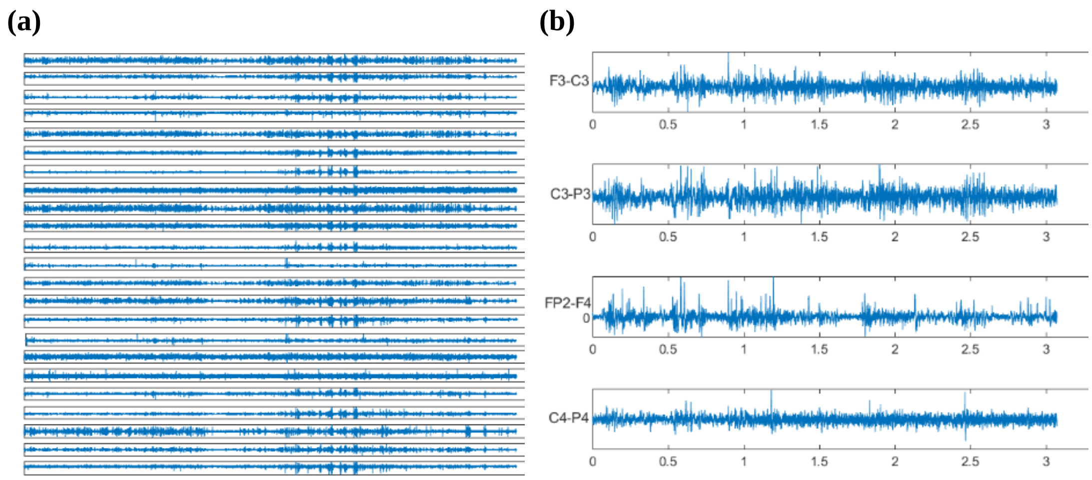

Figure 2.2 a. All channels of CHB MIT Dataset, b.Selected four channels

As indicated in Figure 2.2, we have chosen four channels, Channels (F3-C3), (C3-P3), (FP2-F4), and (C4-P4), each of which manifest rhythmic activity. Considering these, the following equation is used to average of most effective four channels;

$$Averaged_{EEG}[n] = \frac{1}{c}\left(\sum_{i=1}^{c} x_i[n]\right) \tag{2.1}$$

This process is only implemented on CHB MIT database because of 23 channels, Bonn Database has the only one channel that have been used.

2.3 Signal Processing

In the signal processing step, original signal is decomposed into detailed coefficients with the help db2, db4, db8 and db10.

2.3.1 Wavelet Transforms

In the signal analysis, Fast Fourier Transforms, Wavelet Transform, Power Spectral Density are the most applied techniques for time-frequency analysis. In the recent years, Wavelet Transform is used in EEG signal processing to analyze EEG signal due to the rapidly changing spectral content.

To distinguish between different frequencies, a wavelet represented by Ψ must be oscillatory in some way. Wavelet Transform enables the accurate decomposition of complicated information content into elementary form at various scales and positions, and subsequent reconstruction. To analyze the time domain $f(t)$ signal, Continuous Wavelet Transform (Ψ_f^{Ψ}) is used and expressed as Equation 2.2;

$$\Psi_f^{\Psi} = \frac{1}{\sqrt{|S|}} \int_{-\infty}^{+\infty} f(t) \Psi^* \left(\frac{t-\tau}{s} \right) dt \qquad (2.2)$$

where $\Psi(t)$ is referred to as the mother wavelet, τ and s correspondingly indicate the translation and scaling parameters. When a complex wavelet is present, the complex conjugate is employed, as indicated by the sign *. If translation and scaling are discretized, the wavelet transform is called as Discrete Wavelet Transform.

In this study, it is introduced the wavelet analysis-based feature extraction technique using Daubechies family. Frequency domain signals of EEG are more apparent compared with time-domain signals (Ren & Wu, 2014). As a result of high time-frequency resolution, wavelet types are one of the techniques frequently utilized in EEG analysis. They are divided into 7 main groups that briefly represented in Figure 2.3.

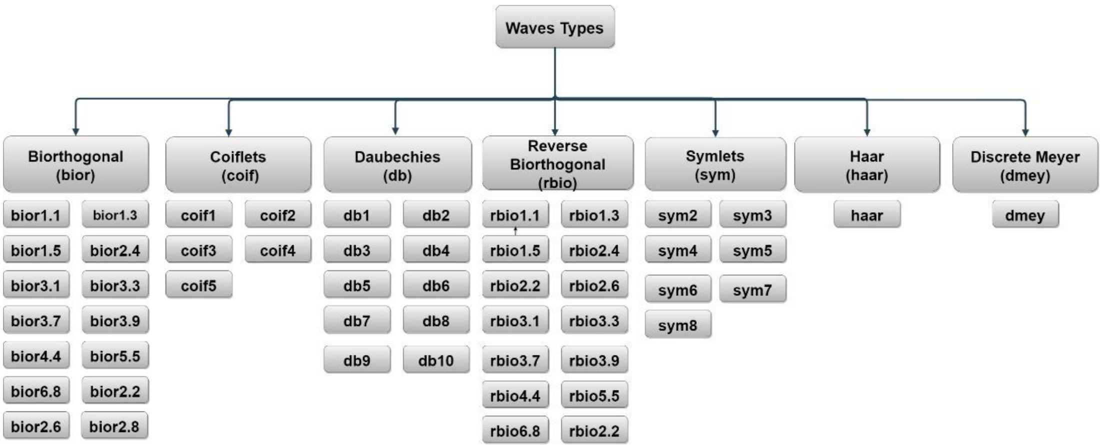

Figure 2.3 Wavelet family representation (Al-Qerem, Kharbat, Nashwan, Ashraf, & Blaou, 2020)

2.3.1.1 Daubechies Wavelets

Daubechies Wavelets are proposed by Belgian physicist Ingrid Daubechies. dbN format is used to abbreviate the Daubechies wavelets, where N is order. The order increases its regularity. This wavelet is known for non-symmetric and orthogonal attributes, efficient filter implementation. They give remarkable results in epileptic detection algorithms due to the previous properties. The maximum number of vanishing moments is the most significant characteristic. Daubechies wavelet has different forms (e.g., db1 to db10). The vanishing moments are determined by these forms. Overlapping windows are utilized in these wavelets, thus all high frequency changes are represented by the high frequency coefficient spectrum. Thus it is applied to compress and remove noise audio signal processing (Mahmoud, Dessouky, Deyab, & Elfouly, 2007). Daubechies wavelets are better adopted natural signals than a flat Haar wavelet. Because Haar wavelets cannot capture the high frequencies of the epileptic signals effectively. Daubechies wavelet is also known for detecting the change in frequency.

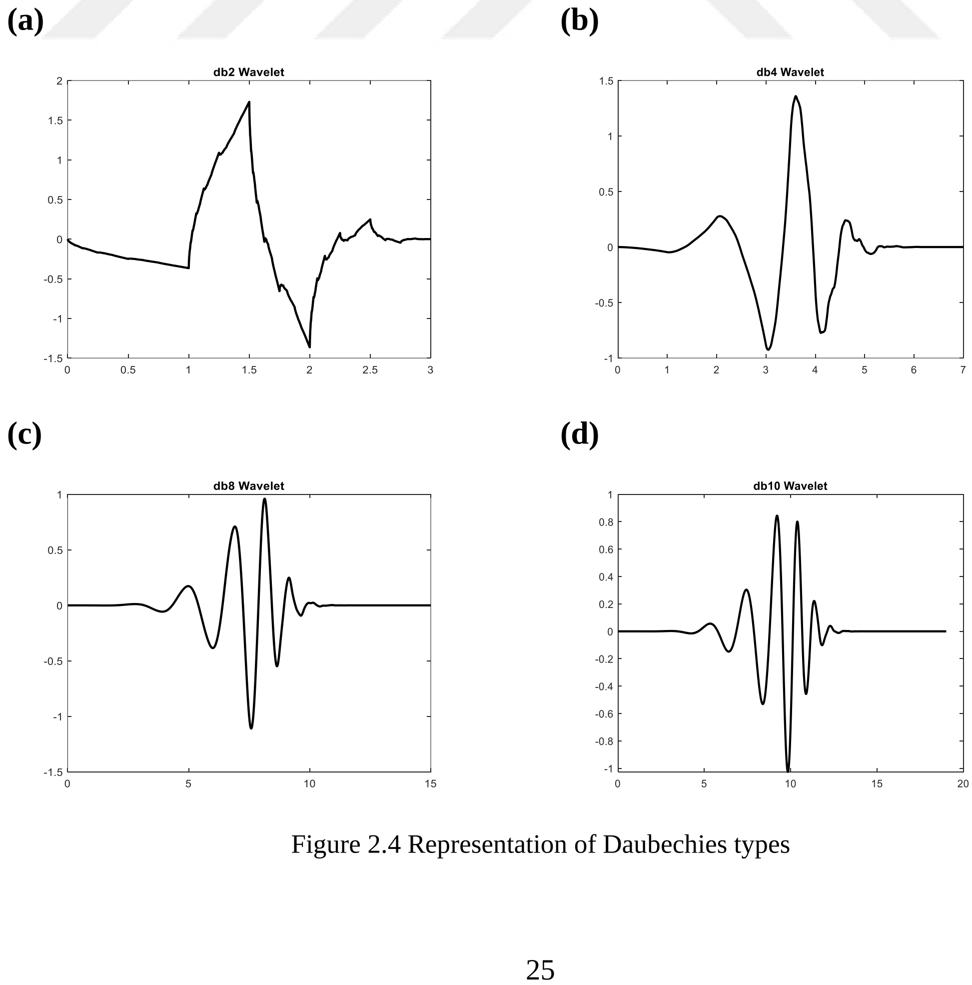

Figure 2.4 Representation of Daubechies types

In the figure, a, b, c and d are the representation of 2nd of Daubechies wavelet, 4th of Daubechies wavelet, 8th order of Daubechies wavelet, 10th of Daubechies wavelet. In order to obtain the optimal features which are associated with classifier to acquire the highest classification rate, the original EEG signal is decomposed in time-frequency domain. The original EEG signal is decomposed into frequency sub-bands such as alpha, beta, theta, gamma and delta by employing DWT. The wavelet coefficients are obtained using second order Daubechies (db2) wavelet and a fourth order Daubechies (db4) wavelet, eighth order Daubechies (db8) and tenth order Daubechies (db10) as seen in Figure 2.4. The extracted features of each channel include mean, kurtosis, skewness, approximate entropy/complexity, root mean squared, variance, Lyapunov Exponent /mobility, Mean Absolute Deviation.

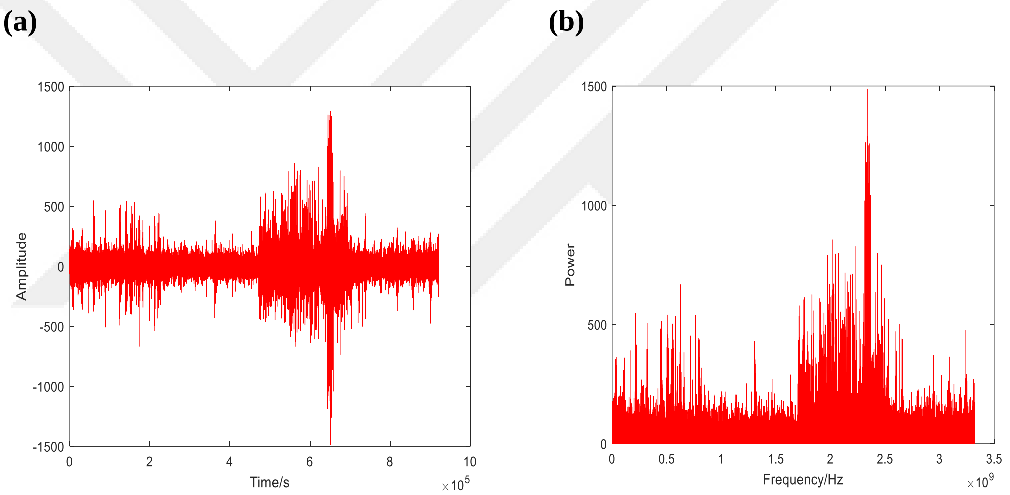

Figure 2.5 a. Recordings in the time domain, b. Recordings in the frequency domain

For each channel, the time domain signals are converted to frequency domain signals using the Wavelet Transform described as Figure 2.5.

2.3.2 Wavelet Decomposition

The performance of Wavelet Transform (WT) is affected directly by four main factors, as follows: decomposition level, frequency band, mother wavelet and selecting features by using the WT coefficients. Therefore, it should be noted that the selection of an appropriate number of decomposition levels is an important step. In order to identify the number of decomposition level, it is needed to handle dominant frequency.

If the wavelet coefficient levels and the useful frequency range of the EEG signal are well correlated, they can be selected in the next processing steps. In the 0-30 Hz of frequency spectrum, EEG signals give meaningful information. Oscillations called rhythms of EEG signals in which seen in EEG waves at frequencies lower than 30 Hz. At that point, decomposition of the signal such that it optimally correlates within significant frequency band. If the decomposition levels are less than five, the differentiation of lower rhythmic activities gets hard, this theta and delta rhythms may be disappearing. In contrast, the decomposition level that is higher than 5, it does not give significant results. By taking into account sampling frequency of dataset and above discussion, the most suitable choice for decomposition is five-level decomposition in this algorithm. Using 5-level-decomposition transform and Daubechies wavelets family of order 2 (db2), order 4 (db4), order 8 (db8) and order10 (db10), it is obtained the detailed coefficients from D1 to D5 and approximation coefficients (A5) at the lower subband. This wavelet decomposition of EEG signal is shown in Figure 2.5.

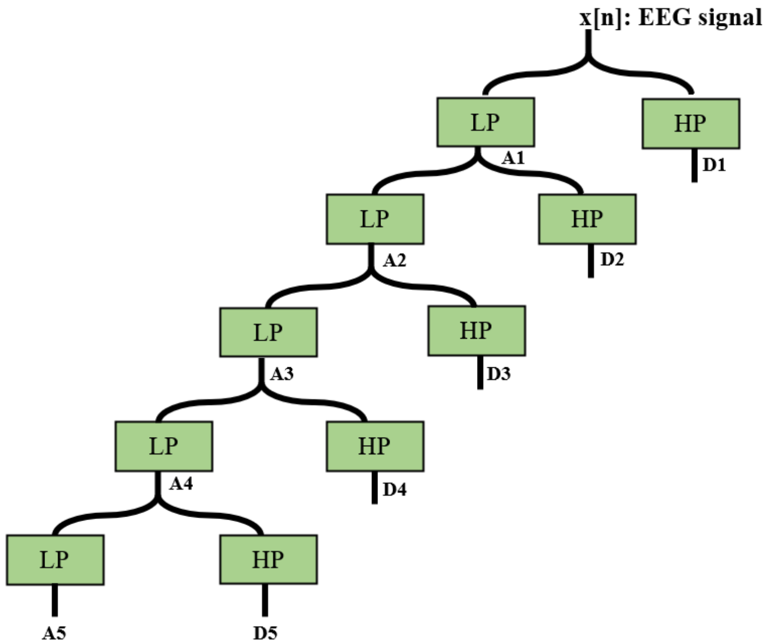

Figure 2.6 Schematic representation of the 5th level wavelet decomposition

The following is a step-by-step process to Wavelet Decomposition:

- This technique utilizes two digital filters and two downsamplers at each level. In the first step, LP represents low pass filter whereas HP is the high pass filter. The downsampled outputs of the first high-pass and low-pass filters produce, respectively, the level of detail D1 and the approximation A1.
- In the second level, the output of low pass filter (A2) is fed into the filters of second step, and this process for the 1-D signal is repeated until level 5 and the outputs of the filters are ordered as A5, D5, D4, D3, D2, D1.

2.4 Feature Extraction and Selection

The features define some critical properties. In the signal analysis, the feature extraction of original signal important. After computing all sub-bands, then it is calculated all statistical features including mean value, kurtosis, skewness, approximate entropy, root mean squared, variance, Lyapunov exponent, mean absolute deviation. In the proposed algorithm, all features are calculated from four detailed wavelet coefficients as D1, D2, D3, D4, D5 and one approximation coefficient (A5) only. These parameters represent the location and variability of the EEG data. In the usage of the wavelet transform with the 5-level decomposition and Daubechies mother wavelet, all subband coefficients are computed.

96 samples are randomly chosen out of 120 samples of each feature and used for training classifiers, the remaining 24 samples from each feature are employed for the testing process. Consequently, for each class feature vectors are 120x8. All these data set is obtained from feature extraction are normalized before the classification procedure. All normalized data are subjected to cross validation process. A 10-fold cross validation is also employed to prevent overfitting during this phase of training. Each column of the input data matrix is one feature (totally 8 columns are listed as Table 2.3.), and each row represented one observation. To demonstrate the effectiveness of the suggested method, two different EEG datasets are compared.

Table 2.3 The eight-dimension features for CHB MIT dataset

EEG Signal	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7	Feature 8
Gamma	Complexity	Kurtosis	MAD	Mean	Mobility	RMS	Skewness	Variance
Beta	Complexity	Kurtosis	MAD	Mean	Mobility	RMS	Skewness	Variance
Alpha	Complexity	Kurtosis	MAD	Mean	Mobility	RMS	Skewness	Variance
Theta	Complexity	Kurtosis	MAD	Mean	Mobility	RMS	Skewness	Variance
Delta	Complexity	Kurtosis	MAD	Mean	Mobility	RMS	Skewness	Variance

Table 2.4 The eight-dimension features for Bonn University

EEG Signal	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7	Feature 8
Gamma	Entropy	Kurtosis	MAD	Mean	LYP	RMS	Skewness	Variance
Beta	Entropy	Kurtosis	MAD	Mean	LYP	RMS	Skewness	Variance
Alpha	Entropy	Kurtosis	MAD	Mean	LYP	RMS	Skewness	Variance
Theta	Entropy	Kurtosis	MAD	Mean	LYP	RMS	Skewness	Variance
Delta	Entropy	Kurtosis	MAD	Mean	LYP	RMS	Skewness	Variance

This is convenient to the fact that the approximate entropy and Lyapunov Exponent features are found to form time consuming set of features and ineffective for CHB MIT dataset. Both features are eliminated to improve efficiency of the proposed method by using feature ranking algorithm. Mobility and Complexity are selected instead of using these features in CHB MIT database, list of used features is represented in Table 2.4. These selected components are processed by Decision Trees, Discriminant Analysis, Naive Bayes, SVM and k-NN.

2.4.1 Mean

In statistical computation, the **mean** value is the measure of the arithmetic average of all values of data. It indicates the sum of all values divided by total number of values. Mean is not complicated value which has good estimate of data values. It is expected that the mean of seizure data is higher than the seizure free data because of high amplitudes. The mean formula is as given in Eq. 2.3;

$$E = \frac{\sum_{i=1}^{N} x_i^2}{N} \tag{2.3}$$

2.4.2 Kurtosis

The Kurtosis word derives from the Greek Word 'Kurtos' that means the data dispersion. It is defined as measure of the peakedness and gives information about the shape of frequency distribution. The most characteristic of epileptic EEG includes spikes, sharps, spike and slow wave complex. Higher values are signature of an existing seizure activity whereas lower values refer to normal data without seizure. Kurtosis (γ_2) is the fourth standardized moment and is defined as;

$$\gamma_2 = {}^{\rho_4}/_{S^4} \tag{2.4}$$

where ρ_4 is the fourth moment about the mean and S is the standard deviation.

2.4.3 Skewness

Skewness is a quantify of asymmetry. The skewness is given by;

$$\gamma_1 = {}^{\rho_3}/_{S^3} \tag{2.5}$$

where ρ_3 the third moment about the mean and S^3 is the standard deviation. If the data is more spread out on the right side of the graph, it means that skewness is a positive value. On the contrary, the negative values cause spreading out to the left of the mean. It is expected that in optimal symmetric distribution about mean, the skewness has zero value.

2.4.4 Approximate Entropy

Approximate Entropy (ApEn) is a complexity index which can measure the complexity or irregularity of a specific time series (Ocak, 2009). It is one of the most helpful tools to extract information from the biological signal. It has always positive value for any time series. Higher ApEn value indicates higher degree of complexity or irregularity. It is expected that the regularity of the epileptogenic data set patterns higher than seizure-free EEG. The EEG in those with epilepsy is more regular, predictable, and simplified than that of healthy individuals. The regularity of the epileptic EEG is higher than normal EEG data. Formally, if a time series contains N data points, $y(n) = [y(1), y(2), \ldots\ldots., y(N)]$, the steps for calculating ApEn are as follows:

1. The definitions of the form m vectors Y(1), Y(2), Y(3),...,Y(N-m+1) are as follows:

$$Y(i) = [y(i), y(i+1), \ldots\ldots., y(i+m-1)]; \quad i = 1,2, \ldots., N-m+1$$

2. Distance between Y(i) and Y(j) is defined by d[Y(i),Y(j)] which indicates represents the highest difference between their scalar components in absolute terms

$$d[Y(i), Y(j)] = {}_{k=1,2,\ldots e}^{max} |y(i+k+1) - y(j+k-1)|$$

3. Determine the number of (j=1,…….,N-m+1, j≠i) such that d[Y(i), Y(j)] ≤ k, represented as $N^m(i)$. Then for i=1,2,…..N-m+1

$$C_k^m(i) = \frac{N^m(i)}{N-m+1}, for\ i = 1 \ldots.. N-m+1$$

4. Figure out the natural algorithm of $C_k^m(i)$ and the find mean value over

$$\emptyset^m(k) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} lnC_k^m(i)$$

5. As another steps, increase the dimension to m+1. Repeat steps between 1–4 to obtain $\emptyset^{m+1}(k)$ and $C_k^{m+1}(i)$

6. Finally, ApEn is computed based on the following formula;

$$ApEn(m, k, N) = \emptyset^m(k) - \emptyset^{m+1}(k) \tag{2.6}$$

Before computing the value of ApEn, two parameters must be known where these are m and k, the embedding dimension, a tolerance window, respectively. (Upadhyay, Padhy, & Kankar, 2016).

2.4.5 Mean Absolute Deviation

Mean Absolute Deviation (MAD) is one of the parameters to measure statistical dispersion. It is the mean of the absolute deviations of dataset, mathematically is expressed as;

$$MAD = \frac{1}{n}\sum_{i=1}^{p}|x_i - z| \tag{2.7}$$

where z is the mean of the distribution with n sample points. It is comprehensible feature that is more effective for distributions except ideal normal. It gives better results compared to standard deviation in realistic conditions where some of the determinations are in error (Gorard, 2005).

2.4.6 Lyapunov Exponent

Lyapunov Exponent estimates chaos and complexity of nonlinear dynamic systems and spatiotemporal dynamics in epileptic EEG time are characterized. When a system becomes unpredictable, it can be mentioned that the system is chaotic and have aperiodic dynamics. If Lyapunov Exponent has a positive value, the system becomes chaotic. It is possible to detect seizures using chaotic metrics with 90% sensitivity (Kannathal, Chee, Er, Lim, & Tat, 2014).

2.4.7 Hjorth's Parameters

The Hjorth parameters is a set of statistical parameters of a time series proposed by Bo Hjorth (1970) that are extensively used in signal processing. These parameters are calculated that contain activity, mobility, and complexity in time domain (Türk, Şeker, Akpolat, & Özerdem, 2017). They are originally introduced for several online EEG analyses (e.g., sleep staging). Otherwise, they are commonly used to detect epileptic seizures (Kaushik, Gaur, Sharma, & Pachori, 2022; Tanveer, Pachori, & Angami, 2018).

The first Hjorth parameter**, activity**, defined as the variance of the EEG amplitudes, is the degree of statistical distribution. This value of seizure data is generally higher than healthy data. The variance defines as;

$$Activity = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \tag{2.8}$$

where x_i is the signal, μ represents mean of the signal and N is length of the signal.

Mobility (m) is the second Hjorth parameter that means approximation of frequency. This can be represented in mathematical terms as square root of the ratio of the activity of the first derivative of the signal divided by the activity of EEG amplitudes and defined by the function;

$$m = \sqrt{\frac{Activity\left(\frac{dx_i}{dt}\right)}{Activity(x_i)}} \tag{2.9}$$

The next last parameter **complexity (c)** is defined as an approximation of the signal bandwidth. This term can be used to determine the closeness of a time series to original sine wave. It is defined as mobility of the first derivative of the signal is divided by the mobility of the signal itself;

$$c = \frac{Mobility\left(\frac{dx_i}{dt}\right)}{Mobility(x_i)} \tag{2.10}$$

2.4.8 Root Mean Square

One of the time domain features is Root Mean Square (RMS) which is widely utilized for analyzing the EEG signal. It has been considered as good signal power estimator in frequency bands of biosignals, measures magnitude of the variable quantity (Abdul-Latif, Cosic, Kumar, Polus, & Da Costa, 2004; Patel, Chua, Fau, & Bleakley, 2009). The mathematical equation of RMS is described in Eq. (2.10);

$$RMS = \sqrt{\frac{1}{K}\sum_{i=1}^{K} x_i^2} \tag{2.11}$$

2.5 Classification

The selected favorable features are utilized for 5 classification methods using Decision Trees, Naive Bayes, Discriminant Analysis, SVM, k-NN. These five classifiers run out for different algorithms which are given in Table 2.5. The data is trained and tested using randomly selected for defined features. The average accuracy percentage of all classifiers is presented in Result and Discussion part. The average ACC percentage of each classifier is obtained by dividing summation of all algorithms of method results to total number of algorithms of method. The all sensitivity, specificity and ACC values of classifiers are represented in Tables 3.5 and 3.6.

Table 2.5 Classification methods with different algorithms used in this study.

Classifier	**Algorithm**
Decision Tree	Fine Tree Medium Tree Coarse Tree
Discriminant Analysis	Linear Discriminant Quadratic Discriminant
Naive Bayes	Gaussian Naive Bayes Kernel Naive Bayes
Support Vector Machine	Linear SVM Quadratic SVM Cubic SVM
k-Nearest Neighbor classifiers	Fine k-NN Medium k-NN Cosine k-NN Cubic k-NN Weighted k-NN

In the following subsections, all these classifier methods are explained one by one.

2.5.1 Decision Trees

Decision Tree (DT) is a learning algorithm that is one of the most widely used basic classification and regression techniques in signal processing. This provides both nominal and numerical features. The classification model is similar to tree structure that consists of root node, branches, and leaves (Han, Pei, & Kamber, 2011). Internal nodes evaluate an attribute while branches store outcomes of each occurrence and the classification is taken place by leaves. First of all, the dataset is divided into smaller and smaller subsets. Secondly, classify the data completely based on selection of each attribute. With the high interpretability, flexibility, and easy debugging decision tree has the potential to prediction of heart disease problems (Thenmozhi & Deepika, 2014), chronic kidney diseases (Chaurasia, Pal, & Tiwari, 2018), sleep apneas (Rohan & Kumari, 2021), breast cancer (Venkatesan & Velmurugan, 2015) and epilepsy classification (Gifu, 2021; Martis et al., 2012). Even if data is incomplete, it gives effective results without scaling and normalization.

Decision tree can have different names according to the number of splits; fine tree, medium tree, coarse tree. The maximum number of branches on the medium tree is 20, and it has a medium number of leaves for making fine differences between the classes. Fewer leaves are used to separate the classes by coarse tree. Accordingly, the maximum number of splits is determined to be four. In order to draw clearer differences between the classes, a fine tree is regarded when there are numerous leaves.

2.5.2 Discriminant Analysis

In the thesis, the two types of Discriminant Analysis namely Linear Discriminant Analysis, and Quadratic Linear Discriminant Analysis is used.

2.5.2.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is supervised algorithm that compares the variety of the test data with the defined dataset of LDA. The linear discriminant features is computed by maximizing distance between classes and minimizing distance between classes (Subasi & Gursoy, 2010). The main purpose of this procedure, it is desired that every class may have a standard submission related with discriminant ratings. These ratings demonstrate that it is determined with discriminant complete that has the shape;

$$D = f_1 C_1 + f_2 C_2 + \cdots \ldots . . + f_p C_p \tag{2.12}$$

2.5.2.2 Quadratic Linear Discriminant Analysis

Quadratic Linear Discriminant Analysis (QLDA) is closely related to LDA except the covariance matrix. It has assumption that the measurements from each class are distributed normally whereas QLDA does not suppose that the covariance of each class is identical (Eva & Lazar, 2015). The covariance matrix $\sum_i$ is classified for each class, i=1, 2,…….I.

Quadratic discriminant function:

$$\delta_i = -\frac{1}{2}\log\left|\sum_i\right| - \frac{1}{2}(x-\mu_i)^T \sum_i^{-1} (x-\mu_i) + \log \pi i \tag{2.13}$$

2.5.3 Naive Bayes

Naive Bayes (NB) is a probabilistic classifier that is derived on Bayesian theory. Maximum likelihood estimates the occurrence or particular absence of Naive Bayesian algorithm (Fielding, 2006). The occurrence probability of an effect C is associated with the certainty of an event B that happens before and given using Eq.

$$P(C\backslash X) = \frac{P(C) * P(X\backslash C)}{P(C)} \tag{2.14}$$

Where $P(C)$ is the probability of C occuring, $P(X)$ is the probability of X occuring, $P(C\backslash X)$is the probability of C given X, $P(X\backslash C)$is the probability of X given C.

In multiclass and binary classification, NB is useful, and less training data is required. It gives the best results with the datasets including functionality-dependent features or completely independent features. The NB is preferred in many medical studies for prediction of heart diseases (Saritas & Yasar, 2019), diabetes diagnosis (Choudhury & Gupta, 2019) and classification of human emotion (Oktavia, Wibawa, Pane, & Purnomo, 2019) and epileptic EEG (Sameer & Gupta, 2021).

In Naive Bayes, there are different kinds of models based on the flexibility of the model. The parameters of Gaussian Naive Bayes cannot be altered to regulate the model's flexibility. With spite of this, in Kernel Naive Bayes, the parameters may be modified to see how the classifier predicts predictor distributions.

2.5.4 Support Vector Machine

The Support Vector Machine (SVM) is kernel-based classifier that is designed by Cortes and Vapnik in 1995. This method is originally proposed to overcome binary classification problems. SVM is more widely used for epileptic seizure detection with its accuracy and capability to overcome many predictors. Depending on the flexibility of the model, there are six versions of SVM. Although linear SVM only generates basic differences within classes, it is not as flexible for multiple model parameters. The model flexibility of quadratic and cubic SVM is medium compared to linear SVM. This classification is used in many different research areas apart from epilepsy. Some of these studies include performance evaluation of dementia prediction (Battineni, Chintalapudi, & Amenta, 2019), detection of Alzheimer’s disease (Rabeh, Benzarti, & Amiri, 2016) and cervical cancer (Jia, Li, & Zhang, 2020), leukemia diagnosis (Vogado, Veras, Araujo, Silva, & Aires, 2018).

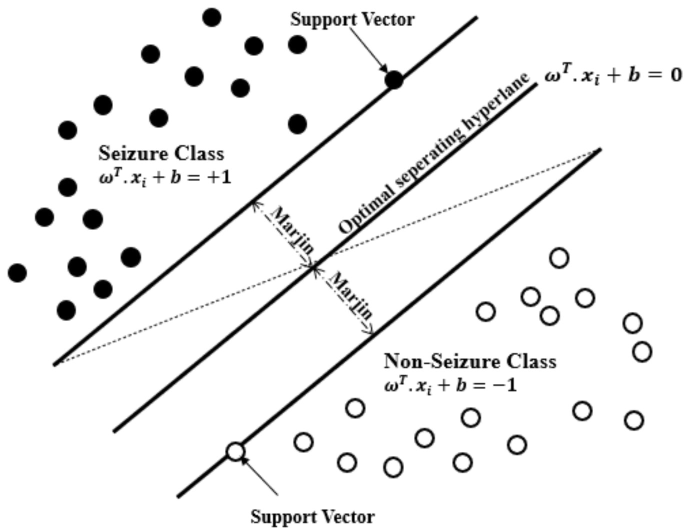

Figure 2.7 Optimal linear seperating hyperplane

The main concept of SVM is to develop an optimal hyperplane that separates two classes describes as Figure 2.7 in the data. If classes can be separated linearly, hyperplanes that have maximum margins might be used to classify them. Otherwise, if data cannot be divided linearly, the bigger space (i.e., feature space) is utilized for separating them linearly. This conversion is named as kernel function. The Kernel approach is utilized by SVM in order to map training data from an input space to a feature space that has a higher dimension. For the two-class linear classification process (class = -1 for non-epileptic and class = +1 for epileptic), the equations for the separating hyperplane are given by;

$$If \qquad \omega^T.x_i + b \geq +1 \quad , \qquad y = +1 \tag{2.15}$$

$$If \qquad \omega^T.x_i + b \leq +1 \quad , \qquad y = -1 \tag{2.16}$$

where each point of data is assigned to x_i represents, $y_k \in \{+1; -1\}$ represents output labels, ω represents the weight vector and b is constant value (Nkengfack, Tchiotsop, Atangana, Louis-Door, & Wolf, 2020).

2.5.5 k- Nearest Neighbor

k-Nearest Neighbor (k-NN) is a member of nonlinear classifiers, derived from nearest neighbor method which aims to handle both classification and regression problems. In the classification algorithm, the distance metric between different feature values is calculated and the distance metric (e.g., Euclidean distance) determines the similarity between two data points as seen in Figure 2.8. It saves all accessible data samples and classifies the new data according to previously known data samples based on measured similarity among training and test set. Varieties of k-NN exist based on the number of neighbors and nature of class differences. Here, six different model types are examined. The number of neighbors is assumed to be one while using Fine k-NN, and fine differentiation is provided between the different classes. The medium differentiation of classes is achieved by medium K-NN using 10 neighbors. The other classifier is Cosine k-NN that utilizes a cosine distance metric with a medium degree of separation between the 10 neighboring classes. If the distance metric is cubic, Cubic k-NN is another type of k-NN which is used, the number of neighbors is defined to 10, and the differentiation between classes is medium. It is referred to as weighted k-NN when the k-NN algorithm uses the distance weight between each pair of neighbors as a parameter, selects 10 neighbors, and selects those neighbors with medium differences.

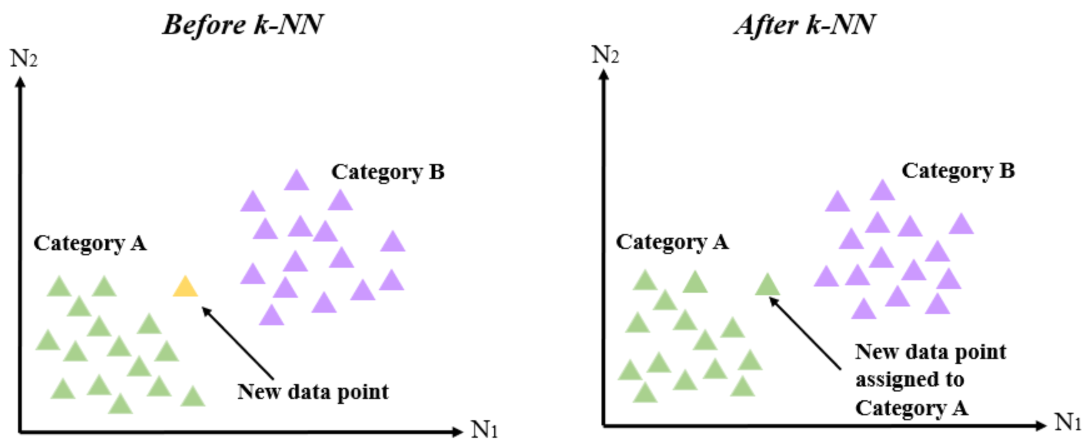

Figure 2.8 A sample classification of k-NN algorithm

The Euclidean Distance (ED) between two points, a and b, with k dimensions is as follows;

$$\sqrt{\sum_{j=1}^{k}\left(a_j - b_j\right)^2} \tag{2.17}$$

where

$a_j = a_1, a_2, \ldots \ldots a_k$ and $b_j = b_1, b_2, \ldots \ldots b_k$

In our experiments, we set k=5 to find nearest neighbors. k-NN is also effective for noisy and large data. On the other hand, simple interpretation attitude and multiclass capability are among other advantages. The k-NN is very popular in EEG analysis that includes major depressive disorders (Saeedi, Saeedi, & Maghsoudi, 2020), emotion recognition (Mi Li, Xu, Liu, & Lu, 2018), and epileptic prediction (Savadkoohi, Oladunni, & Thompson, 2020) and detection (Choubey & Pandey, 2021). The most important feature of this algorithm is that not required to crate model, so fast results are obtained.

2.6 Performance Evaluating

Where *True Positives (TP)* is the number of epileptic cases that are predicted as epileptic, *False Positives (FP)* is the number of epileptic cases, that are predicted as non-epileptic, *True Negatives (TN)* is the number of non-epileptic cases which are predicted as non-epileptic and *False Negatives (FN)* is the number of non-epileptic cases that are classified as epileptic by the system. If it is required to show all these terms in a single table, they are represented as Table 2.6.

Table 2.6 Confusion matrix for detection of epileptic seizure

		Predicted	
		Epileptic	**Non-Epileptic**
True	**Epileptic**	True Positive (TP)	False Positive (FP)
	Non-Epileptic	False Negative (FN)	True Negative (TN)

Accuracy, sensitivity, and specificity all play a role in the evaluation of classifications' performance. The meanings of these metrics are given below; Sensitivity is defined as the ratio of the total number of positive cases that are correctly classified, also called the true positive rate, as given by the formula;

$$Sensitivity\ (SEN) = TPR = \frac{TP}{FN + TP} \tag{2.18}$$

On the other hand, another performance metric is specificity that also called the true negative ratio, is described below;

$$Specificity(SPE) = TNR = \frac{TN}{TN + FP} \tag{2.19}$$

The last term ACC is the ratio of number of correct classifications (summation of TP and TN) to Total Samples;

$$Accuracy(ACC) = \frac{TP + TN}{Total\ Samples} \tag{2.20}$$

2.6.1 The Receiver Operator Characteristic Curve

To assess the performance of the suggested approach, it is applied to confusion matrix to obtain accuracy, sensitivity and specificity, The Receiver Operator Characteristic (ROC) curve. It is computed the region under the ROC curve to estimate the classifier's ability to discriminate between classes. The probability curve illustrates the True Positive Rate (sensitivity) vs the False Positive Rate (specificity) at different threshold settings. It is considered as the signal is removed from the noise with the help of ROC curve.

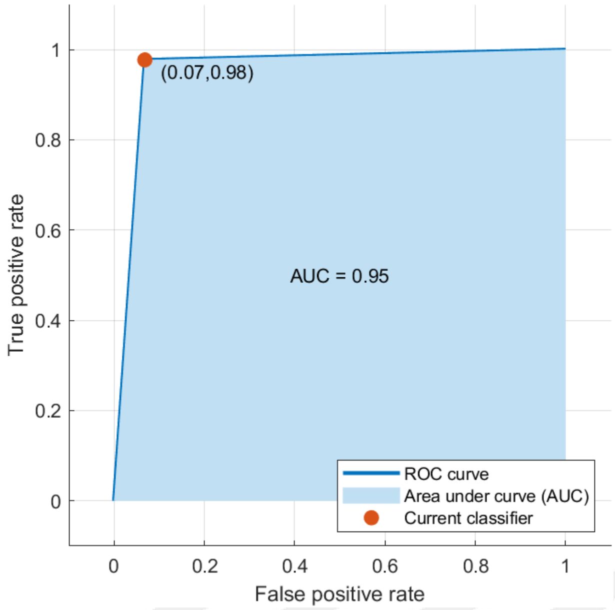

Figure 2.9 ROC Curve of Decision Tress classifier that associated with db10

When the Area Under Curve (AUC) is higher, it indicates that the model shows better performance when it comes to differentiating between positive and negative classes as shown in Figure 2.9.

2.6.2 k-Fold Validation

The robustness of the suggested method is evaluated by 10-fold cross validation during the training and testing process. This is the most decent way for estimating the performance of a machine learning algorithm on a given dataset.

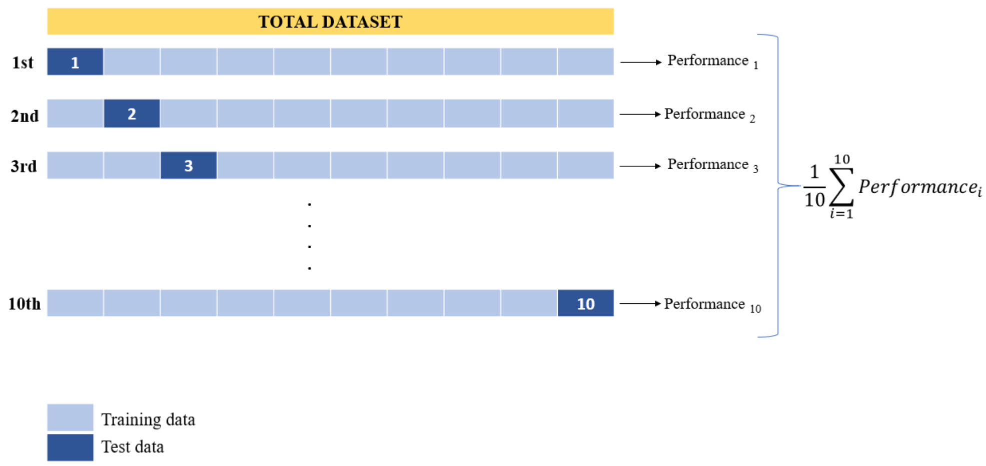

Figure 2.10 Representation of 10-fold cross validation

In this algorithm, the feature matrix is randomly divided into k equal size subsets. In testing, one subset is chosen, whereas in training, the other subsets are selected. (R. Kumar & Indrayan, 2011). This process is iterated for k times leaving one-fold for computation each time that illustrated in Figure 2.10. Overfitting can be avoided by performing rotation estimation. In our proposed method, 10-fold cross validation is computed for better approximation error.

In the following chapter, all these classification algorithms run out and their performances are given. The results are also compared with the previous studies in the literature and discussion is given.

CHAPTER 3

RESULT and DISCUSSION

3.1 Experimental Result

The proposed algorithm is applied to Bonn University and Children's Hospital Boston–Massachusetts Institute of Technology datasets in this thesis. Eighty percent of samples of each group (i.e., 48 samples of healthy groups and 48 samples of epileptic group) are randomly formed as training data. The remaining %20 of samples (i.e., 15 samples of healthy groups and 15 samples of epileptic group) are adopted as test data. First of all, raw EEG signals are filtered by using a Band-Pass filter and divided into five sub-bands using fifth level decomposition. Secondly, the features are extracted from wavelet coefficients. These features are used as input to machine learning algorithms for the classification of seizures and non-seizure events.

In this section, it is presented the classification performances of the proposed algorithm and comparison of the results for a sequence of Daubechies filters (ie.from db2 to db10), 5 subbands (gamma, beta, alpha, theta, delta), and 5 different classifiers (Decision Trees, Discriminant Analysis, Naive Bayes, SVM, k-NN) utilized for the detection of epileptic seizures. The feature matrix is classified by MATLAB classification learner box. The average ACC rates provided by the two different datasets are given in the below tables. The yellow bold lines represent the highest average ACC rates.

3.1.1 Application of Classification Algorithm on CHB MIT Dataset

In this part, the performance of 5 different classifiers such that Decision Trees, Discriminant Analysis, Naive Bayes, SVM and k-NN are shown in Table 3.1 for CHB MIT dataset in the form of percent average ACC performance.

Table 3.1 Average ACC rates of db filters for 5 subbands on classifiers (in %)

	Gamma	Beta	Alpha	Theta	Delta	Average
			DECISION TREES			
db2	100.00	95.50	95.50	90.90	95.50	**95.48**
db4	95.50	95.50	90.90	81.80	95.50	91.84
db8	86.40	100.00	90.90	77.30	77.30	86.38
db10	90.90	100.00	95.50	95.50	95.50	**95.48**
			DISCRIMINANT ANALYSIS			
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	97.75	97.75	90.95	97.75	97.75	96.39
db4	97.75	100.00	95.50	95.50	95.50	96.85
db8	95.45	97.75	93.20	93.20	100.00	95.92
db10	100.00	97.75	100.00	95.45	100.00	**98.64**
			NAIVE BAYES			
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	100.00	100.00	90.95	100.00	95.50	97.29
db4	100.00	100.00	95.50	95.50	95.50	97.30
db8	100.00	95.50	90.90	100.00	100.00	97.28
db10	100.00	100.00	100.00	95.50	97.75	**98.65**
			SVM			
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	89.43	100.00	93.97	100.00	100.00	96.68
db4	100.00	98.50	100.00	100.00	100.00	**99.70**
db8	100.00	93.83	95.50	100.00	100.00	97.87
db10	100.00	100.00	100.00	96.97	100.00	**99.39**
			k-NN			
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	95.50	100.00	91.82	100.00	100.00	97.46
db4	98.20	100.00	100.00	100.00	100.00	**99.64**
db8	96.40	95.50	96.40	97.30	100.00	97.12
db10	100.00	99.10	100.00	99.10	100.00	**99.64**

In the above tables, we want to see the characteristic of each type of Daubechies along the signal with the average values. Combination of Decision Trees and db10/db2 give the best performance as it maintained an ACC of 95.48% according to the Table 3.1. For the Discriminant Analysis, Naive Bayes, the best performances (98.64%, 98,65%, respectively) are obtained from db10. The accuracy rates of other Daubechies types are slightly lower than db10 (97.29%, 97.30%, and 97.28%, respectively) in Naive Bayes classifier. In comparison with the ACC performances of other wavelet types, Table 3.1 shows that db4 has the higher ACC performance of 99.70% for SVM. The second highest average ACC values of 99.39% is obtained from db10 as other classifiers. In the k-NN classifier, db4 and db10 resulted in the best performance with an average ACC of 99.64%. Interestingly, the lowest average ACC rates are achieved by db8 in all classes except SVM. Even if the best average performance of Decision Trees is 95.48%, the combination of db8 and theta-delta subbands associated with Decision Trees has the lowest ACC of 77.30%. In the Discriminant Analysis and Naive Bayes classifier, combination of db2 and alpha subband has the lowest ACC rate of 90.95%. Similarly, this combination with the poor ACC performance of 91.82% is seen in the k-NN classifier. Consequently, we could say that combination of db2 and alpha subband failed to classify epileptic seizures. In the SVM classifier, the lowest average ACC rate is associated with combining of db2 and Gamma subbands.

Table 3.2 Average ACC performance of subbands with Daubechies wavelets (in %)

	db2	***db4***	***db8***	***db10***
Gamma	96.09	98.20	95.47	98.18
Beta	98.80	98.80	96.37	**99.40**
Alpha	**92.75**	97.28	93.96	99.10
Theta	97.88	95.46	93.65	96.99
Delta	98.20	98.20	95.46	98.80

In this table, we want to observe the comparison of all subbands versus Daubechies wavelets from db2 to db10 in the classification algorithms. It is inferred that db10 gives the best result in Beta subband with an average ACC of 99.40%. We can see that the second highest ACC rate of 99.10% is achieved by db10 in alpha band. Another important point is that the combination of alpha and db2 has the lowest ACC of 92.75%. On the other hand, we should note that the performances of all Daubechies types in beta band is above the ACC rate of 96%.

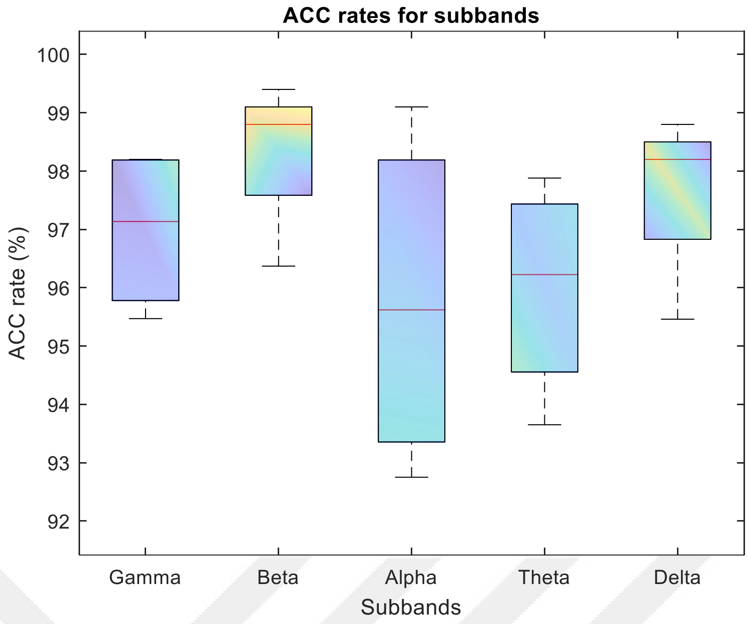

Figure 3.1 Illustration of ACC performances versus 5-subbands

The above figure, it is seen that beta band has the highest accuracy rate while alpha has the lowest rate. Specifically, max. and min. accuracy are from 92.75 to 99.10% for Alpha subband whereas the similar observation of Beta band as from 96.37 to 99.40%. The underlying reason is that beta band can capture epileptic characteristics of signal compared to other subbands.

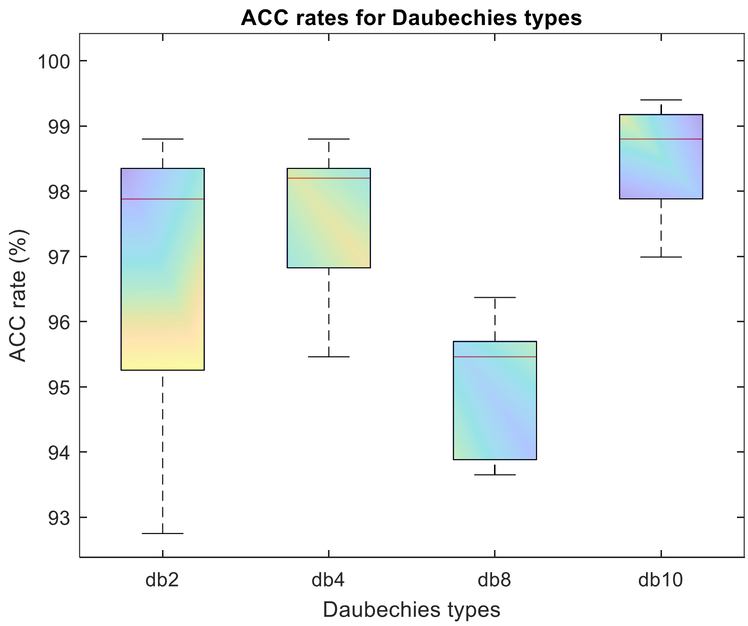

Figure 3.2 Distribution of the ACC performances achieved by db2, db4, db8 and db10

Each Daubechies wavelet from db2 to db10 and correspondingly computed accuracy rates are shown in Figure 3.2. It is seen that db2 has a broader range from 92.75% to 98.20% than the other Daubechies types. This wider range is interpreted as unpredictability and also is not preferred for classification. The narrowest range is evaluated for db 10 which shows the best performance from 96.99% to 99.40%.

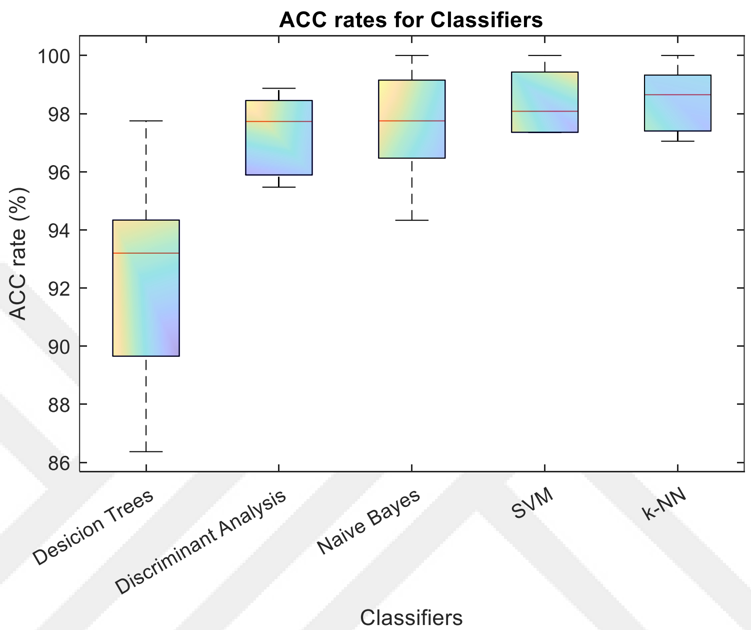

Figure 3.3 Distribution of ACC performances achieved by 5 classifiers

Figure 3.3 clearly portrayed that SVM is the most consistent classifier with narrow ACC rates among other classifiers. Its ACC values range from 97.35% to 100%. Decision Trees classifier is insufficient with the wider ACC values from 86.37% to 97.75%.

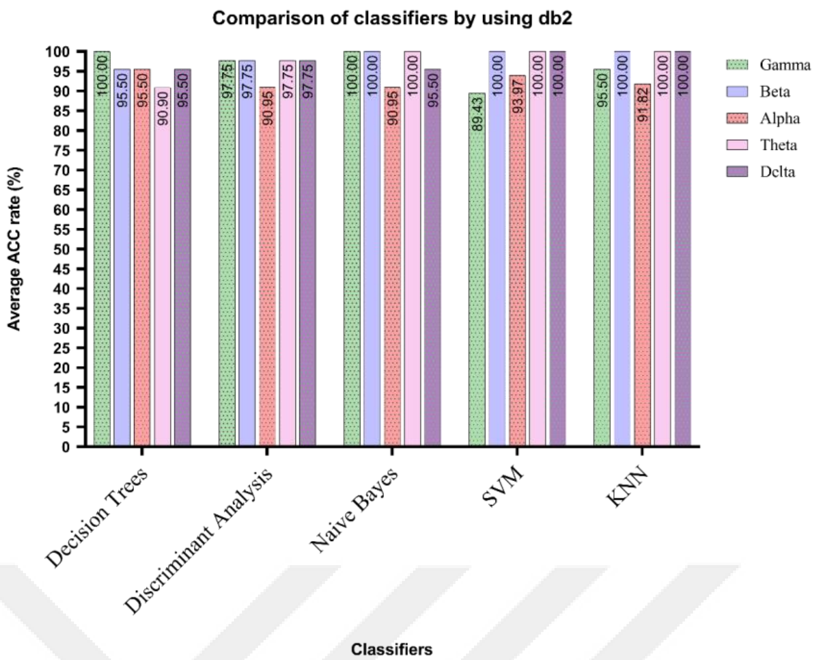

Figure 3.4 Comparison between average ACC rates for 5 classifiers in the subbands

Figure 3.4 gives average ACC rates of each subband in one classifier. In this part, we analyze difference between highest accuracy and lowest accuracy of classifiers associated with subbands. This difference has also impact algorithm stability. When it is obtained high differences, we can say that the classifier is unstable in EEG signal. An important other observation is that ACC difference between subbands in limited range which can be assumed to be an advantage of proposed method. The lowest difference between max.-min. accuracy rate is obtained from Discriminant Analysis (highest-lowest average ACC rates are 90.9%, 97.7%, respectively) whereas SVM has the highest difference (highest-lowest average ACC rates are 89.43%, 100%, respectively) for db2.

When Decision Trees are used, the highest accuracy of 100% is obtained from gamma band. The performance of beta band is lower than gamma band in this classifier. While using Discriminant Analysis classifier, the lowest accuracy (90.9 %) is obtained from alpha band. For the Naive Bayes classifier, gamma, beta, and theta band have the highest accuracy rate of 100%. SVM and K-NN have the same highest accuracy from beta, theta, and delta. The common subband is the beta which has the highest accuracy in the whole classifier except Decision Trees classifiers. It is obvious that alpha band has the insufficient ACC performance with Discriminant Analysis, Naive Bayes and k-NN.

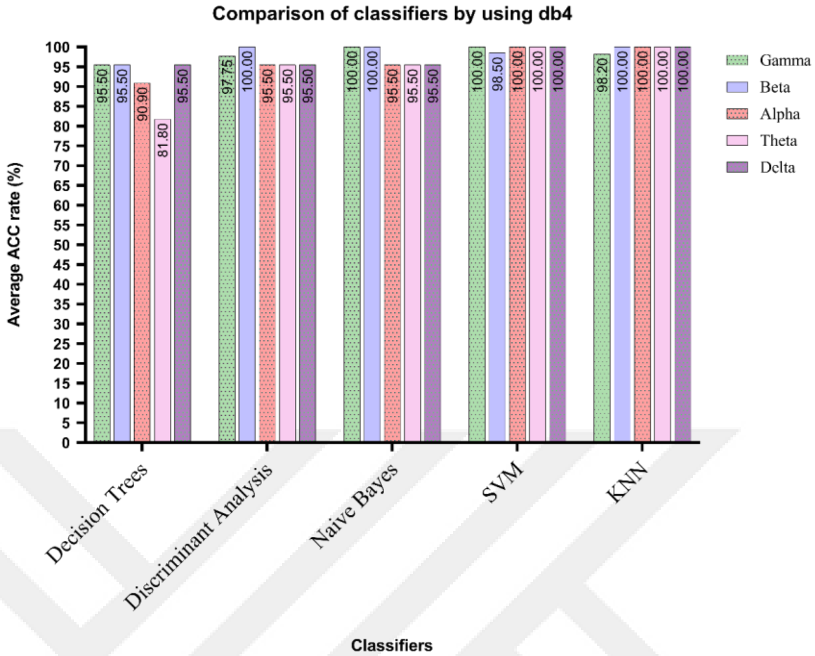

Figure 3.5 Comparison between average ACC rates for 5 classifiers in the subbands

The highest difference is 13.7% from combination of Decision Tree classifier and theta subband. As mentioned in Figure 3.5, beta band is regarded as best subband for Daubechies 4 with 95.50, 98.50,100% ACC values. At the same time, SVM gives worst ACC rate (98.50%) in beta band by using db4.

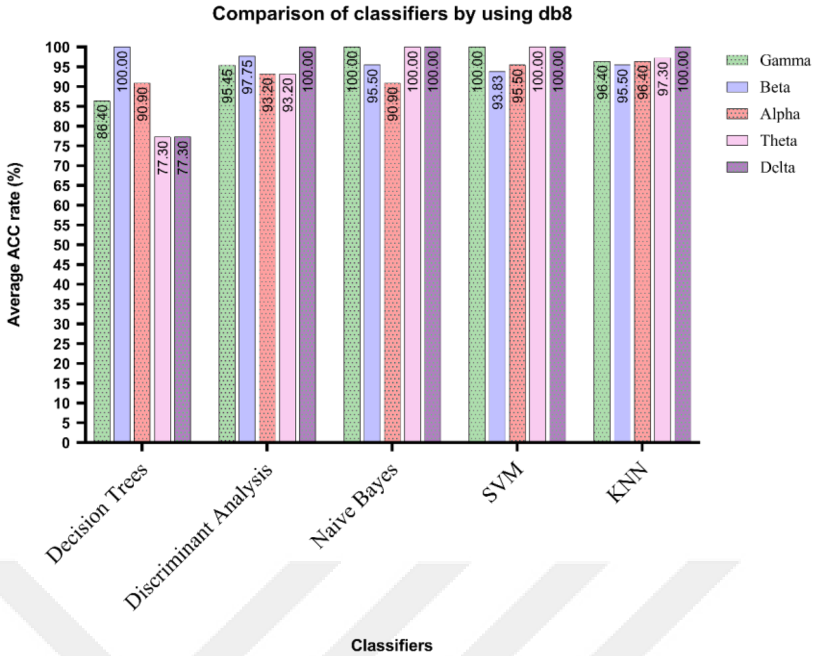

Figure 3.6 Comparison between average ACC rates for 5 classifiers in the subbands

As can be seen from the table 3.6, The Decision Tree has the highest difference between max. and min. ACC rate (77.3% and 100% respectively). Beta band shows the most significant difference in this classifier. We also investigate that *k*-NN have the effective ACC performance and the lowest difference between min.-max. ACC at the same time. For the CHB MIT dataset, the lowest accuracy rate is obtained from Decision Trees classifier with an ACC of 77.3% using db8. Another important point is that Decision Trees gives best ACC performances of 100% in the beta band.

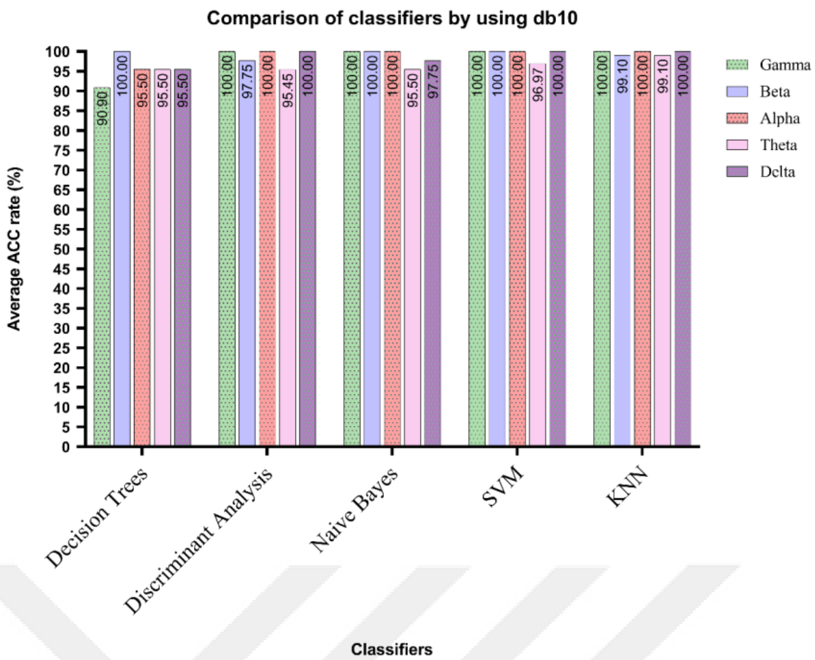

Figure 3.7 Comparison between average ACC rates for 5 classifiers in the subbands

It is shown in Figure 3.7 that Decision Tree indicates instability and broader average ACC range between 90.9% and 100% while k-NN outperforms the other classifiers with an average ACC of between 99.1-100% for all subbands. Although the combination of K-NN-db10 seems perfect, obtaining an ACC between 99.1-100% from all subbands is not reliable. Because frequency oscillations caused by epileptic seizures are uncommon in all frequency ranges. SVM with 3.1% difference can be accepted as best classifier for db10.

3.1.2 Application of Classification Algorithm on Bonn University Dataset

In this part, the performance of 5 different classifiers such that Decision Trees, Discriminant Analysis, Naive Bayes, SVM and k-NN are shown in Table 3.3 for Bonn University dataset in the form of percent average ACC performance as Table 3.1.

Table 3.3 Average ACC rates of db filters for 5 subbands on classifiers (in %)

DESICION TREES						
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	75.00	100.00	100.00	100.00	100.00	95.00
db4	100.00	100.00	100.00	91.70	100.00	98.34
db8	100.00	91.30	100.00	100.00	100.00	98.26
db10	100.00	100.00	100.00	100.00	100.00	**100.00**
DISCRIMINANT ANALYSIS						
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	91.70	100.00	95.85	100.00	91.65	95.84
db4	87.50	95.85	95.85	95.85	95.85	94.18
db8	97.85	95.70	93.50	95.65	91.30	94.80
db10	89.60	100.00	97.90	97.90	97.90	**96.66**
NAIVE NAYES						
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	91.70	100.00	100.00	100.00	100.00	**98.34**
db4	91.70	100.00	100.00	91.70	100.00	96.68
db8	91.35	95.70	95.70	100.00	100.00	96.55
db10	91.65	95.80	100.00	100.00	100.00	**97.49**
SVM						
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	94.47	100.00	91.70	97.23	91.70	95.02
db4	88.90	100.00	100.00	91.70	91.70	94.46
db8	92.77	97.13	94.23	94.20	100.00	95.67
db10	98.60	95.83	97.20	98.60	97.20	**97.49**
k-NN						
	Gamma	**Beta**	**Alpha**	**Theta**	**Delta**	**Average**
db2	83.32	98.34	88.32	88.34	90.02	89.67
db4	83.30	95.02	90.00	90.02	83.30	88.33
db8	94.82	99.14	95.70	86.98	84.36	92.20
db10	90.86	100.00	95.80	89.18	86.66	**92.50**

As indicated in above Figure, db10 has an improved performances in all classifiers except Naive Bayes. Unlike in other cases, the highest ACC performance of 98.34% is achieved by db2 in the Naive Bayes classifier. However, the second highest ACC performance of 97.49% is obtained from db10. Alike SVM classifier performance on the CHB MIT Dataset, the combination of db2 and gamma subband obviously give the lowest performance with the Decision Trees on the Bonn University Dataset. The combination of db4 and Gamma subband give the poorest average ACC performance in Discriminant Analysis, SVM and k-NN. (87.50%, 88.90%, 83.30%, respectively.) The all-average ACC values of Naive Bayes classifiers are above 91.35% compared to other classifiers.

Table 3.4 Average ACC performance of subbands with Daubechies wavelets (in %)

	db2	***db4***	***db8***	***db10***
Gamma	86.12	89.44	95.39	94.17
Beta	95.56	97.79	96.25	**98.61**
Alpha	93.89	96.11	95.97	97.76
Theta	95.56	91.69	93.92	95.83
Delta	93.90	85.94	93.63	94.71

When comparing the CHB MIT database, the higher average accuracy rate (98.61%) is also obtained from beta in the Bonn database as represented in Table 3.4. Similarly, it can be clearly observed that db10 also gives the best results, with an average ACC above 94.17%. The combination of db4 and delta subband fails to classify the epileptic seizures with an ACC of 85.94%. In both datasets, the best Daubechies type is db10 which has the highest average ACC. Thus, db10 is the most effective wavelet type for the detection of epileptic seizures according to the Table 3.1 and Table 3.3.

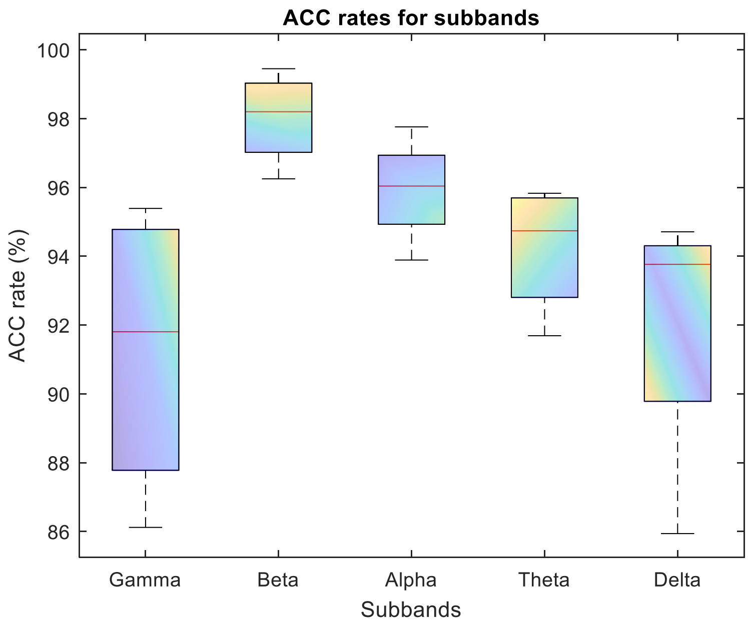

Figure 3.8 Illustration of ACC performances versus 5-subbands

Based on Figure 3.8, the corresponding boxplots are obtained. The boxplots of all subbands shown as Fig. that beta subband has maximum performance as compared to other subbands for Bonn Dataset. It can be understood that the features pertaining to epileptic seizures are most distinguished in beta subband. It proves information that seizures in recorded EEG are significant in the frequency range 3-25 Hz (Ahmad et al., 2019). Interestingly, the Beta band has been attained a high classification accuracy for both databases used. In this dataset, the results of delta and gamma band are slightly poor in terms of ACC of 85.94%.

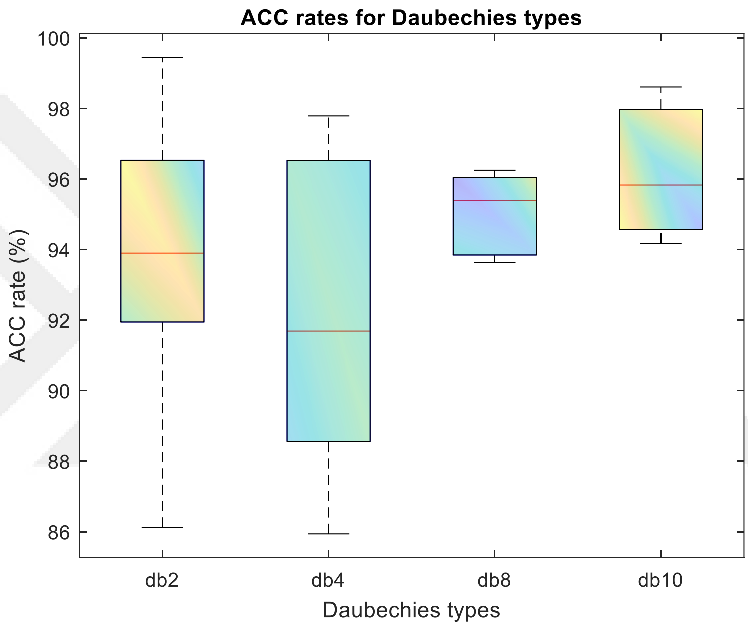

Figure 3.9 Distribution of ACC performances achieved by db2, db4, db8 and db10

Comparing the results obtained in Figure 3.9, it is obvious that the classification accuracy obtained by Daubechies 10 could reach 98.61% which is higher than all other wavelet types. Daubechies 2 is associated with insufficient performance, as the accuracy range widely from 86.12% to 99.45%. Even though db8 has lower performance than db10, it is relatively consistent.

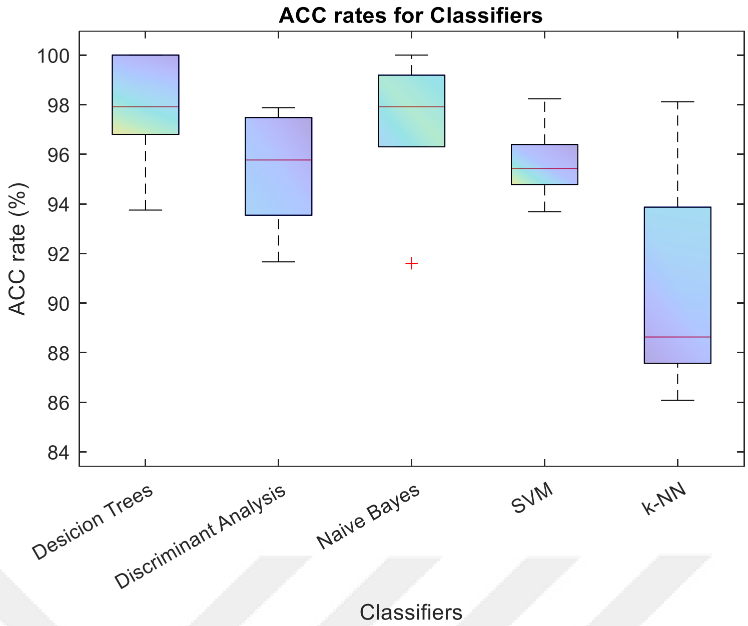

Figure 3.10 Distribution of ACC performances for classifiers

Based on all classifiers as shown in the Figure 3.10, k-NN has the wider ACC rates from 86.08% to 98.12%. The most consistent values are achieved by SVM classifiers compared to other classifier types. Although it seems that the Naive Bayes has higher values than SVM, the wider range of ACC values is seen from 91.60% to 100%

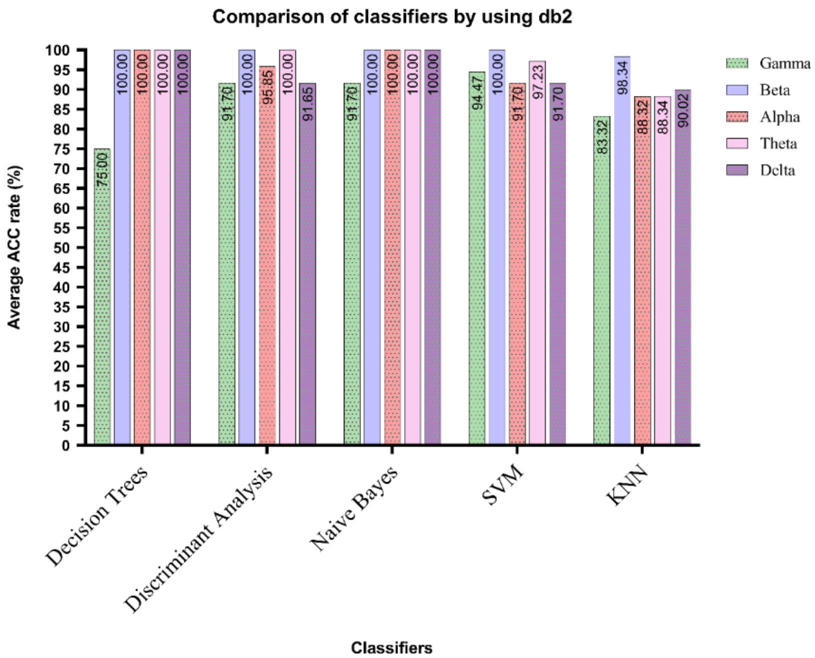

Figure 3.11 Comparison between average ACC rates for 5 classifiers in the subbands

Compared to the difference between max. and min. ACC performance seen as Figure 3.11, Decision Trees is in broader range with 75%-100%. Naive Bayes and SVM attain the same ACC difference. Their classification scores are in the range of 91.7% to 100%, which shows the potential of this method for epilepsy detection.

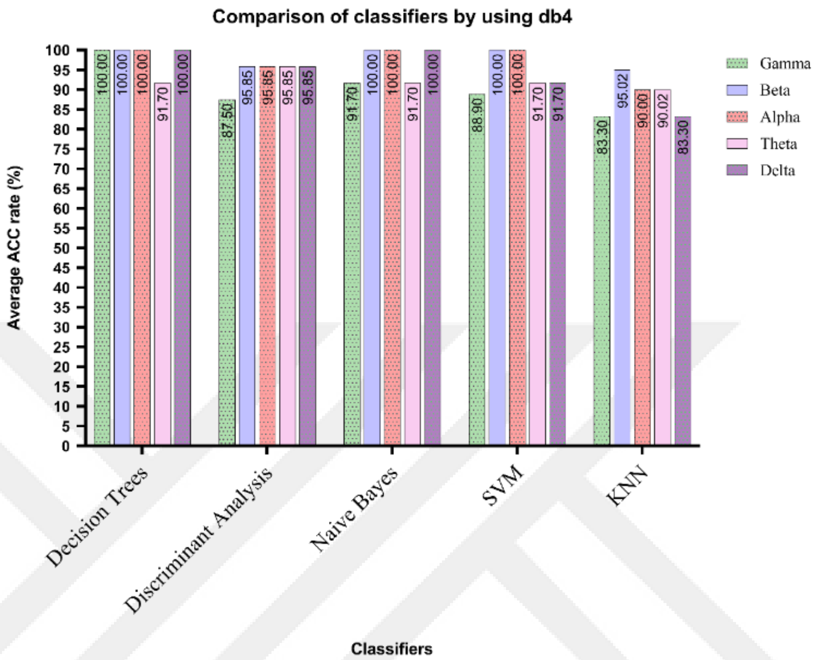

Figure 3.12 Comparison between average ACC rates for 5 classifiers in the subbands

The comparison of the all classifiers in Figure 3.12 shows that k-NN classifier has a significant differentiation according to its lowest rate of 83.30% average ACC. In the gamma and delta subbands, the lowest rate is found in the k-NN-db4 combination, which has an average ACC of 83.3%. Although Decision Tress and Naive Bayes have the different ACC performances, they have the same differences between max. and min. ACC from 91.7% to 100%.

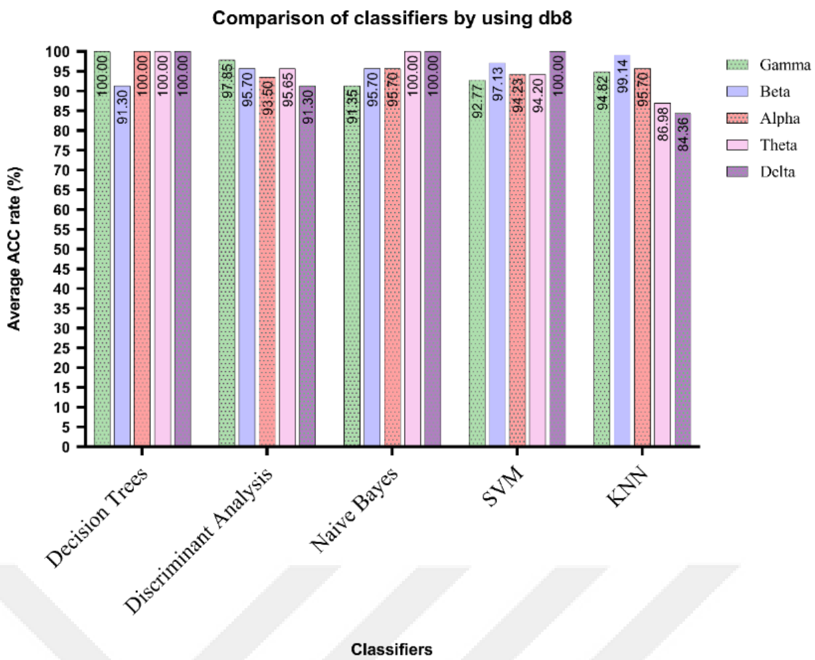

Figure 3.13 Comparison between average ACC rates for 5 classifiers in the subbands

According to the Figure 3.13, the significant difference is obtained from k-NN with 84.36%-99.14%. Although beta band gives best performance compared to other subbands, combination of beta subband-Decision Trees classifier is prominently lower with an ACC performance of 91.3%.

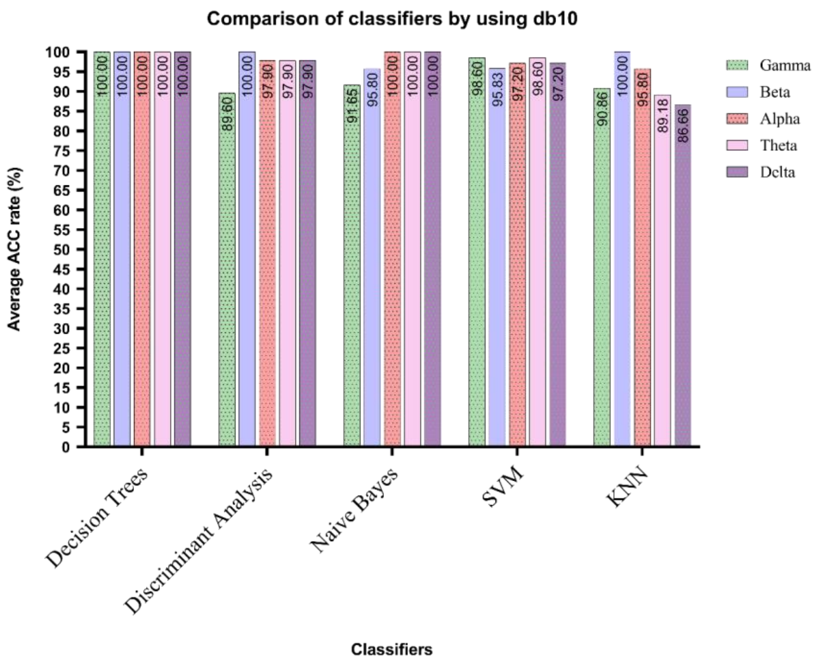

Figure 3.14 Comparison between average ACC rates for 5 classifiers in the subbands

Even though the k-NN is resulted with an accuracy of 100% for Beta subband, the difference between the max. and min. accuracy rates is quite high as shown in Figure 3.14. In this algorithm, Decision Tree is not discriminative for Bonn database because the accuracy rates of subbands are 100 %. These results are not desired because epileptic seizures do not affect all subbands equally.

As a result of the classification on the two datasets, receiver operating characteristic curve and confusion matrix are obtained. The SVM classifier performance of Bonn Dataset in beta band is represented by the ROC Curve shown in 3.15. According to Figure 3.15-3.16, SVM has the ideal value for classification of seizures. Figure 3.17 and 3.18. shows SVM classifier performance of ROC Curve and confusion matrix in beta band for CHB MIT dataset. AUC value is 0.97 that is close to ideal value.

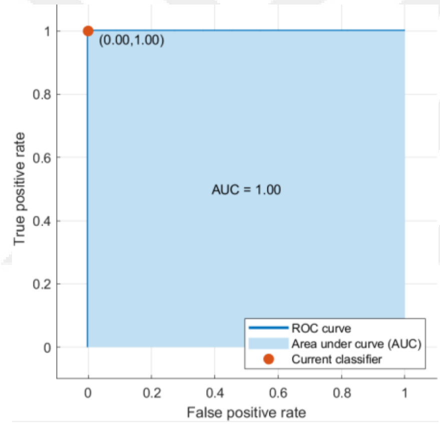

Figure 3.15 ROC Curve of SVM on Bonn EEG database

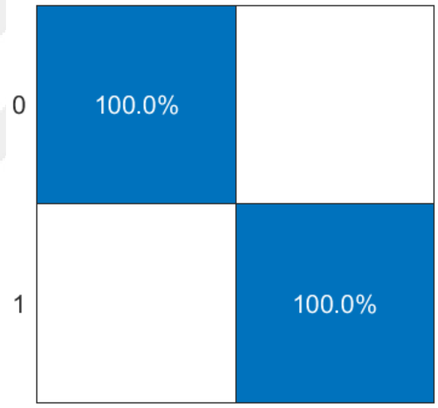

Figure 3.16 Classification performance of SVM on Bonn EEG Database

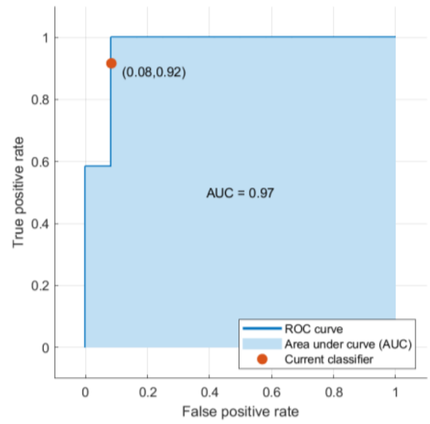

Figure 3.17 ROC Curve of SVM on CHB MIT EEG Database

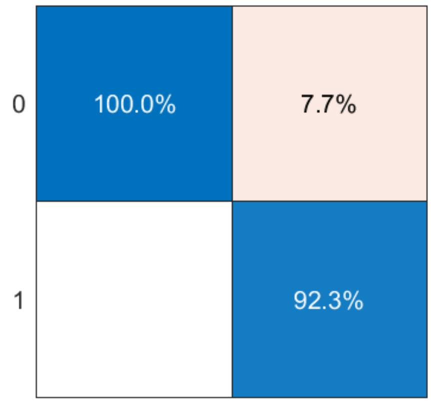

Figure 3.18 Classification performance of SVM on CHB MIT database

According to the results given in all the tables, db10 wavelet gives its best performance when it is associated with SVM which has 97.69 % accuracy. If db10 wavelet is used with beta subband and when it is classified with SVM, an average classification accuracy of 95.83% is achieved along with an average sensitivity of 97.27% and an average specificity of 94.67% in the Bonn dataset. If db10 wavelet is used with beta subband and when it is classified with SVM, an average classification accuracy of 100% is achieved along with an average sensitivity of 100% and an average specificity of 100% in the CHB MIT dataset. It is observed that SVM classifier is ideal classifier with a stable and higher average ACC rate for the two datasets.

It is applied 10-fold cross validation to avoid overlap between test and training dataset. Eighty percent of dataset is randomly selected for training set and twenty percent of remaining dataset is used to test classifier. This procedure is repeated five times to ensure obtaining approximate values without overlap. In this process, it is compared only successive results of beta band combination with different classifiers and wavelet types.

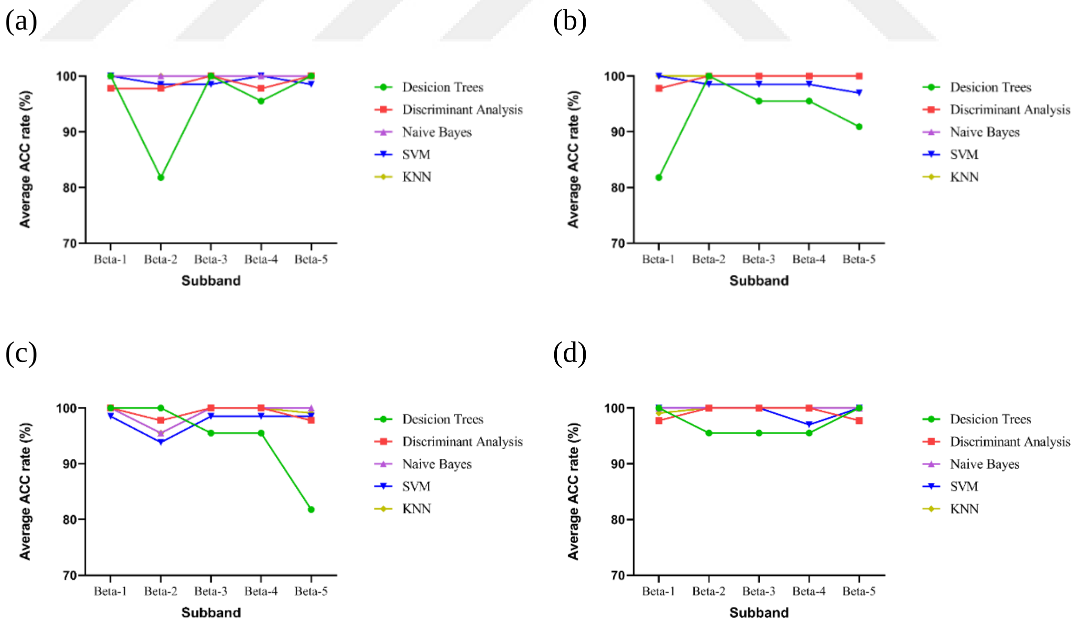

Figure 3.19 Combination of Daubechies wavelets and training rate of 20%

In the figure, option a, b, c and d represent combination of db2 and training rate of 20%, db4 and training rate of 20%, db8 and training rate of 20%, db10 and training rate of 20%, respectively on CBH MIT dataset. Beta-1 represents average ACC values obtained from first iteration. Beta-2 represents average ACC values obtained from second iteration. Beta-3 represents average ACC values obtained from third iteration. Beta-4 represents average ACC values obtained from forth iteration as shown in Figure 3.19.

According to the results, the differences between average ACC rates obtained from Decision Trees is quite high. The other classifiers have approximated average ACC rates at each time. They are generally between 95% and 100% for db2. The closest values are obtained from db10 that are between 95.50%-100%. The lowest value is 81.8% achieved by Decision Trees from db2, db4, db8.

Unlike previous procedure, seventy five percent of dataset is randomly selected for training set and twenty five percent of remaining dataset is used to test classifier. This procedure is repeated four times.

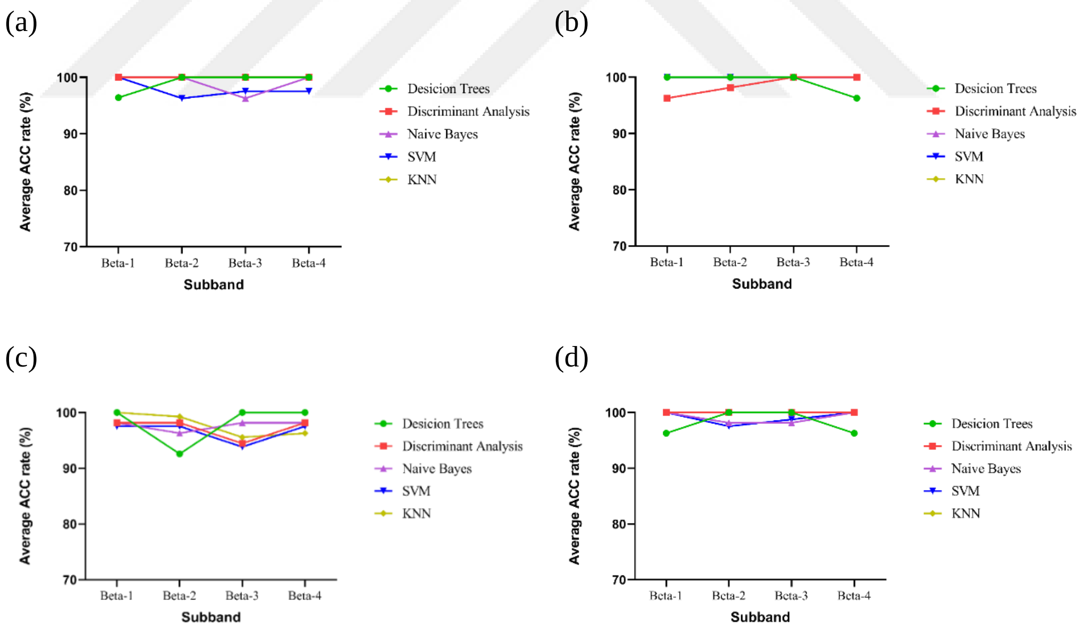

Figure 3.20 Combination of Daubechies wavelets and training rate of 25%

In the above figure, option a, b, c and d represent combination of db2 and training rate of 25%, db4 and training rate of 25%, db8 and training rate of 25%, db10 and training rate of 25%, respectively on CBH MIT dataset.

The most stable values are also obtained from db10 (The ACC ranges of Decision Tress, Discriminant Analysis, Naive Bayes, SVM, k-NN are in 96.30%-100, 100%, 98.15%-100%, 97.53%-100%, 100%, respectively) as described in Table 3.18. The highest average ACC values of 100% are achieved by Discriminant Analysis and k-NN with a db10 in all iteration. The average ACC values of Decision Trees varies in the range [92.60%-100] compared to other classifiers. The lowest ACC value is 92.60 % which is achieved by db8.

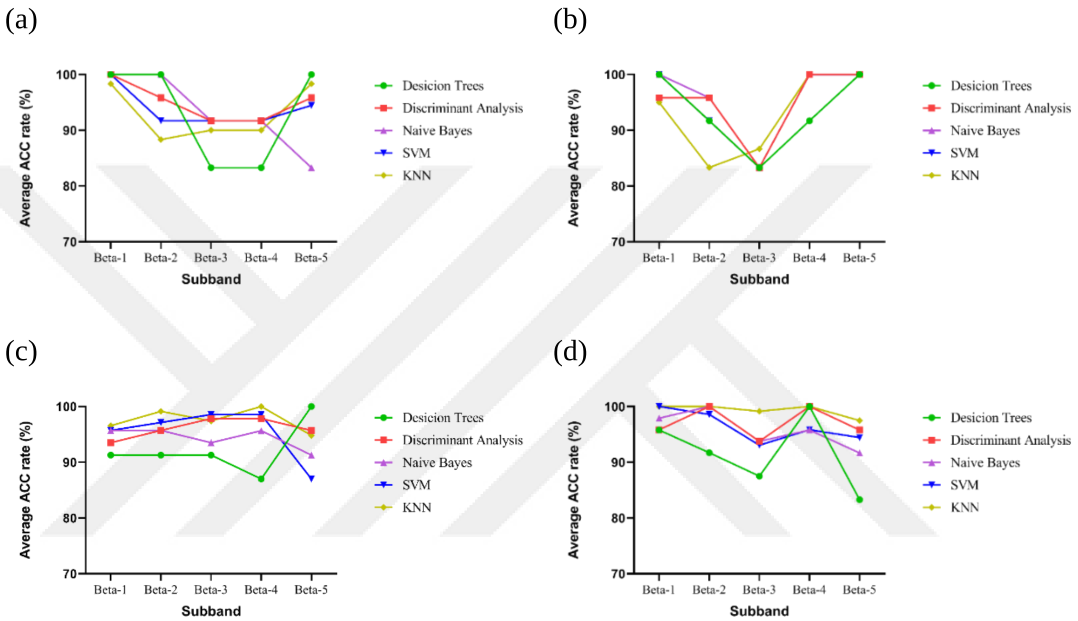

Figure 3.21 Combination of Daubechies wavelets and training rate of 20%

All explanations in Figure 3.19 are available for these Figures. Significant differences are obtained using Bonn University Dataset. According to the Table 3.21, Bonn dataset has wider range from 83.30% to 100%. Decision Trees and k-NN give the lowest average ACC of 83.30%,83.32%, respectively. The average ACC values of db2 are between 83.30% (obtained from Decision Tress with second and third iteration)-100%. The lowest ACC values of db4 are 83.30% achieved by Decision Trees, Discriminant Analysis, Naive Bayes and SVM in the third iteration while all classifiers give average ACC of %100 in the fifth iteration.

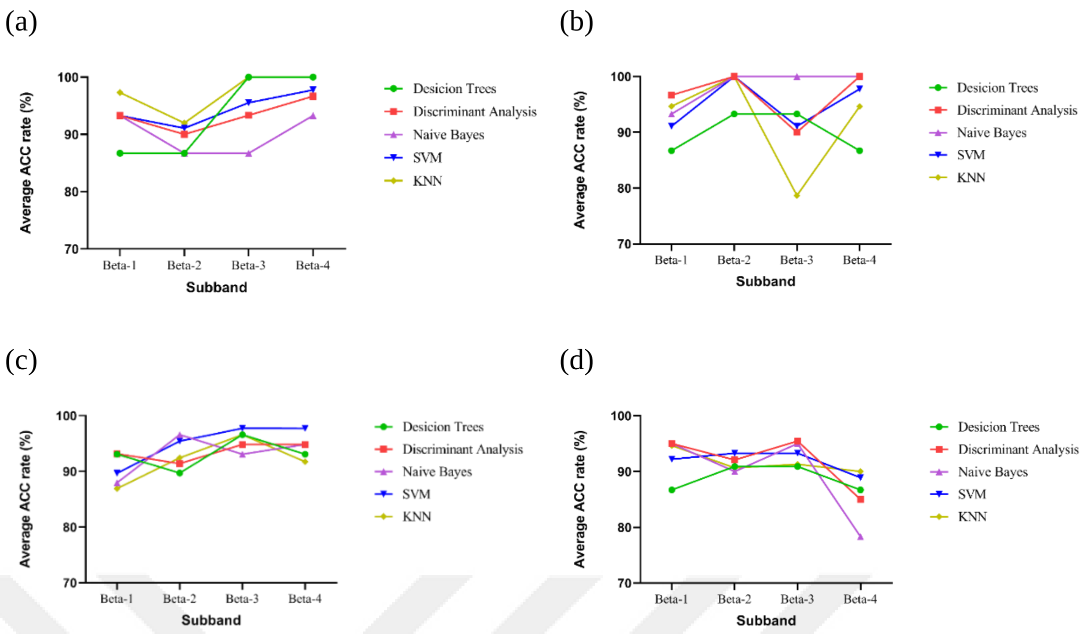

Figure 3.22 Combination of Daubechies wavelets and training rate of 25%

As indicated previous Figure 3.20, all options represent same conditions. The only one difference is these options applied on Bonn University Dataset. Figure 3.22 represents the differences between maximum and minimum average ACC values increase with twenty-five percentage training rate. Although Daubechies 10 gives best results, the lowest ACC rate (78.25%) is achieved by combination of db10 and Naive Bayes classifier. The average ACC range of db2, db4, db8, db10 are 86.70-100%, 78.66-100%, 87.95-100%, 78.35-100%, respectively. These lower ACC ratios are mostly generated by Naive Bayes. In conclusion, with these experiment results of two datasets, we can say that combination of twenty-five percentage and Naive Bayes classifier is not suitable for Bonn Dataset.

Table 3.5 Comparison of ACC, SEN, SPE performances for Bonn University Dataset

db2	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	75.00	80.00	71.40	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Discriminant Analysis	91.70	92.85	92.85	100.00	100.00	100.00	95.85	100.00	92.85	100.00	100.00	100.00	91.65	100.00	87.50
Naive Bayes	91.70	85.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
SVM	94.47	90.47	100.00	100.00	100.00	100.00	91.70	100.00	85.70	97.23	100.00	95.23	91.70	100.00	85.70
KNN	83.32	93.14	79.28	98.34	100.00	97.78	88.32	100.00	82.14	88.34	100.00	81.42	90.02	96.66	85.22

db4	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	91.70	100.00	85.70	100.00	100.00	100.00
Discriminant Analysis	87.50	84.50	91.65	95.85	100.00	92.85	95.85	100.00	92.85	95.85	100.00	92.85	95.85	100.00	92.85
Naive Bayes	91.70	85.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	91.70	100.00	85.70	100.00	100.00	100.00
SVM	88.90	84.90	94.43	100.00	100.00	100.00	100.00	100.00	100.00	91.70	100.00	85.70	91.70	100.00	85.70
KNN	83.30	83.30	83.30	95.02	100.00	91.42	90.00	100.00	84.28	90.02	100.00	83.56	83.30	100.00	75.00

db8	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	100.00	100.00	100.00	91.30	100.00	84.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Discriminant Analysis	97.85	100.00	95.85	95.70	96.15	95.85	93.50	95.85	92.85	95.65	100.00	92.85	91.30	100.00	87.50
Naive Bayes	91.35	86.15	100.00	95.70	92.30	100.00	95.70	91.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00
SVM	92.77	96.67	89.53	97.13	94.87	100.00	94.23	100.00	90.10	94.20	91.83	97.23	100.00	100.00	100.00
KNN	94.82	98.34	91.54	99.14	98.46	100.00	95.70	100.00	92.30	86.98	97.78	80.86	84.36	100.00	77.00

db10	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Discriminant Analysis	89.60	90.00	92.85	100.00	100.00	100.00	97.90	100.00	96.15	97.90	96.15	100.00	97.90	100.00	96.15
Naive Bayes	91.65	87.50	100.00	95.80	96.15	96.15	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
SVM	98.60	100.00	97.43	95.83	97.23	94.67	97.20	100.00	94.87	98.60	100.00	97.43	97.20	100.00	94.87
KNN	90.86	96.52	86.68	100.00	100.00	100.00	95.80	100.00	92.30	89.18	100.00	82.28	86.66	100.00	79.00

Table 3.6 Comparison of ACC, SEN, SPE performances for CHB MIT Dataset

db2	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	100.00	100.00	100.00	95.50	100.00	91.70	95.50	90.90	100.00	90.90	84.60	100.00	95.50	91.70	100.00
Discriminant Analysis	97.75	100.00	95.85	97.75	100.00	95.85	90.95	86.35	95.45	97.75	95.85	100.00	97.75	95.85	100.00
Naive Bayes	100.00	100.00	100.00	100.00	100.00	100.00	90.95	83.90	100.00	100.00	100.00	100.00	95.50	91.70	100.00
SVM	89.43	100.00	84.07	100.00	100.00	100.00	93.97	88.37	100.00	100.00	100.00	100.00	100.00	100.00	100.00
k-NN	95.50	100.00	91.70	100.00	100.00	100.00	91.82	84.82	100.00	100.00	100.00	100.00	100.00	100.00	100.00

db4	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	95.50	100.00	90.90	95.50	100.00	91.70	90.90	90.90	90.90	81.80	88.90	76.90	95.50	100.00	91.70
Discriminant Analysis	97.75	100.00	95.45	100.00	100.00	100.00	95.50	95.85	100.00	95.50	95.85	100.00	95.50	95.85	100.00
Naive Bayes	100.00	100.00	100.00	100.00	100.00	100.00	95.50	91.70	100.00	95.50	91.70	100.00	95.50	91.70	100.00
SVM	100.00	100.00	100.00	98.50	97.23	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
k-NN	98.20	100.00	96.36	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

db8	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	86.40	86.40	83.30	100.00	100.00	100.00	90.90	100.00	84.60	77.30	71.40	87.50	77.30	80.00	75.00
Discriminant Analysis	95.45	100.00	92.30	97.75	95.85	100.00	93.20	95.45	91.30	93.20	92.30	95.85	100.00	100.00	100.00
Naive Bayes	100.00	100.00	100.00	95.50	91.70	100.00	90.90	87.75	95.45	100.00	100.00	100.00	100.00	100.00	100.00
SVM	100.00	100.00	100.00	93.83	91.43	96.97	95.50	100.00	91.70	100.00	100.00	100.00	100.00	100.00	100.00
k-NN	96.40	100.00	93.36	95.50	91.70	100.00	96.40	100.00	93.36	97.30	100.00	95.02	100.00	100.00	100.00

db10	Gamma			Beta			Alpha			Theta			Delta		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
Desicion Trees	90.90	84.60	100.00	100.00	100.00	100.00	95.50	100.00	91.70	95.50	91.70	100.00	95.50	100.00	91.70
Discriminant Analysis	100.00	100.00	100.00	97.75	100.00	95.85	100.00	100.00	100.00	95.45	92.30	100.00	100.00	100.00	100.00
Naive Bayes	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	95.50	91.70	100.00	97.75	95.85	100.00
SVM	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	96.97	100.00	94.87	100.00	100.00	100.00
k-NN	100.00	100.00	100.00	99.10	100.00	98.34	100.00	100.00	100.00	99.10	100.00	98.34	100.00	100.00	100.00

CHAPTER 4

CONCLUSIONS

Epileptic seizure detection using EEG signal are discussed in this thesis to determine the effective algorithm with higher classification accuracy. In EEG signal, there are 5 frequency intervals defined as Gamma (30-60 Hz), Beta (13-30 Hz), Alpha (8-12 Hz), Theta (4-8 Hz), and Delta (0-4 Hz). The epileptic seizures are characterized by beta band. Furthermore, seizures in recorded EEG are more clear in the range 3-25 Hz (Khan & Gotman, 2003). The frequency resolution of signal is associated with the order of Daubechies wavelet. That is, the higher order means corresponds the higher resolution (Luhach, Kosa, Poonia, Gao, & Singh, 2020). Daubechies wavelets are more localized and smoother than Haar wavelet. In our work, db10 is used as a suitable mother wavelet in beta subband besides of db2, db4, db8 and it is appropriate to detect the epileptic seizures in EEG signal (Anila Glory et al., 2020).

Generally, machine learning algorithms are successful and practical in beta subband with Daubechies 10. Specifically, it seen that SVM classifier is more effective to lead to desirable results and improved accuracy of detection. Results are evaluated by average accuracy, sensitivity, and specificity metrics. It gives average ACC of 99.39%, SEN of 100%, SPE of 99.18% on CHB MIT dataset and give average ACC of 97.49%, SEN of 99.12%, SPE of 95.85% on Bonn dataset. Considering these findings, the SVM are more successful with db10 in classification of epileptic signals.

In both datasets, the highest ACC rates are obtained beta subband with an average ACC of 100%, average SEN of 100%, average SPE of 100% for CHB MIT Dataset and an average ACC of 95.83%, average SEN of 97.23%, average SPE of 94.67% for Bonn Dataset. An obvious observation is that the beta band outperforms the other subbands in both datasets. This concludes that the best outcomes are obtained when beta band is used with SVM.

Table 4.1 A comparison of various methods performances for epileptic seizure detection

Authors	Features	Wavelets	Classifier	Performance metrics
(Acharya, Sree, Ang, Yanti, & Suri, 2012)	Entropies Higher Order Spectra based features	db10	Fuzzy	99.7% ACC
(Xie & Krishnan, 2013)	*	Haar	SVM, k-NN, FLD	99 % ACC
(Y. Kumar, Dewal, & Anand, 2014)	Fuzzy approximate entropy	db4	SVM	99.65%-95.85% ACC
(Ahammad, Fathima, & Joseph, 2014)	Maximum, Minimum, Mean, Standard Deviation Entropy, Energy	db2	Linear classifier	84.2% ACC 98.5% SEN
(or Rashid & Ahmad, 2017)	Statistical features	db4	NN	80.0%-100% ACC
(Rajaguru & Prabhakar, 2017)	the mean, standard deviation, minimum approximate coefficient value and maximum approximate coefficient	db2 db4	LDA	95.83% ACC 95.03% ACC
(Nanthini & Santhi, 2017)	mean, median, mode, standard deviation, skewness, kurtosis, maximum, and minimum, gray level co-occurrence matrix, Renyi entropy	db2	SVM	85-100% ACC
(Tzimourta et al., 2017)	energy, entropy, standard deviation, variance and mean	db4	SVM	93% SEN 99% SPE
(El-Gindy et al., 2021)	Local Mean Variance Median Derivative Amplitude Entropy	Haar, db4, db8, Sym4 or Coif4	*	100% SEN average FPR of 0.0818 h−1
(Omidvar et al., 2021)	standard deviation, mean, band power, Hjorth mobility, Hjorth complexity, Shannon entropy, log-energy entropy, maximum value, kurtosis, Skewness and median	db4	ANN, SVM	98.7-100% ACC
This study	Entropy, Kurtosis, LYP, MAD, Mean, Mobility, RMS, Skewness, Variance	db2 db4 db8 db10	Decision Tress Discriminant Analysis K-NN Naive Bayes SVM	95.83- 100 % ACC

In this thesis, the obtained results with the other studies in the literature where have used Daubechies wavelets are summarized in above tables. Generally, all studies use only one Daubechies wavelet type with one or two classifiers. The strong side of this thesis is that the comparison of 5 different classification methods for 5 subbands and 4 different Daubechies wavelets at the same time.

In comparison of results, this study highlights in five different machine learning algorithms as Decision Trees, Discriminant Analysis, Naive Bayes, SVM and k-NN. These are worked out on different dataset, CHB MIT and Bonn University database. The ACC performance above 95% as described in Table 4.1. As a result, this study compared and summarize better performances other than traditional individuals. This allows improved decision making to assist physicians in the detection of epileptic seizures. Additionally, the used algorithms become highly practical, accurate and robust under the light of these performances.

As an extension of this study,

i. The number of features can be increased
ii. Different databases can be experienced
iii. A hybrid classification method can be developed to apply for the PNES and ES classification if enough data is evaluable by studying together with Medical Faculties.

REFERENCES

Abbate, A., DeCusatis, C., & Das, P. K. (2012). *Wavelets and subbands: fundamentals and applications*: Springer Science & Business Media.

Abdul-Latif, A. A., Cosic, I., Kumar, D. K., Polus, B., & Da Costa, C. (2004). *Power changes of EEG signals associated with muscle fatigue: the root mean square analysis of EEG bands.* Paper presented at the Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004.

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine, 100*, 270-278.

Acharya, U. R., Sree, S. V., Ang, P. C. A., Yanti, R., & Suri, J. S. (2012). Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals. *International Journal of Neural Systems, 22*(02), 1250002.

Adeli, H., Zhou, Z., & Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods, 123*(1), 69-87.

Ahammad, N., Fathima, T., & Joseph, P. (2014). Detection of epileptic seizure event and onset using EEG. *BioMed Research International, 2014*.

Ahmad, Z. K., Singh, V., & Khan, Y. U. (2019). *Sequential Segmentation of EEG Signals for Epileptic Seizure Detection using Machine Learning.* Paper presented at the 2019 2nd International Conference on Signal Processing and Communication (ICSPC).

Akbari, H., & Esmaili, S. (2020). A Novel Geometrical Method for Discrimination of Normal, Interictal and Ictal EEG Signals. *Traitement du Signal, 37*(1), 59-68.

Akyol, K. (2020). Stacking ensemble based deep neural networks modeling for effective epileptic seizure detection. *Expert Systems with Applications*, *148*, 113239.

Al-Qerem, A., Kharbat, F., Nashwan, S., Ashraf, S., & Blaou, K. (2020). General model for best feature extraction of EEG using discrete wavelet transform wavelet family and differential evolution. *International Journal of Distributed Sensor Networks, 16*(3), 1550147720911009.

Al-Subari, K., Al-Baddai, S., Tomé, A. M., Volberg, G., Hammwöhner, R., & Lang, E. W. (2015). Ensemble empirical mode decomposition analysis of EEG data collected during a contour integration task. *PLoS ONE, 10*(4), e0119489.

Amin, H. U., Yusoff, M. Z., & Ahmad, R. F. (2020). A novel approach based on wavelet analysis and arithmetic coding for automated detection and diagnosis of epileptic seizure in EEG signals using machine learning techniques. *Biomedical Signal Processing and Control*, *56*, 101707.

Angeles, D. (1981). Proposal for revised clinical and electroencephalographic classification of epileptic seizures. *Epilepsia, 22*(4), 489-501.

Anila Glory, H., Vigneswaran, C., & Shankar Sriram, V. (2020). *Identification of suitable basis wavelet function for epileptic seizure detection using EEG signals.* Paper presented at the First International Conference on Sustainable Technologies for Computational Intelligence.

Anuragi, A., Sisodia, D. S., & Pachori, R. B. (2021). Automated FBSE-EWT based learning framework for detection of epileptic seizures using time-segmented EEG signals. *Computers in Biology and Medicine*, *136*, 104708.

Atal, D. K., & Singh, M. (2020). A hybrid feature extraction and machine learning approaches for epileptic seizure detection. *Multidimensional Systems and Signal Processing, 31*(2), 503-525.

Ayyad, S. M., Saleh, A. I., & Labib, L. M. (2019). Gene expression cancer classification using modified K-Nearest Neighbors technique. *Biosystems, 176*, 41-51.

Bari, M. F., & Fattah, S. A. (2020). Epileptic seizure detection in EEG signals using normalized IMFs in CEEMDAN domain and quadratic discriminant classifier. *Biomedical Signal Processing and Control, 58*, 101833.

Battineni, G., Chintalapudi, N., & Amenta, F. (2019). Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Informatics in Medicine Unlocked, 16*, 100200.

Bhattacharyya, A., & Pachori, R. B. (2017). A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform. *IEEE Transactions on Biomedical Engineering, 64*(9), 2003-2015.

Bizopoulos, P. A., Tsalikakis, D. G., Tzallas, A. T., Koutsouris, D. D., & Fotiadis, D. I. (2013). *EEG epileptic seizure detection using k-means clustering and marginal spectrum based on ensemble empirical mode decomposition.* Paper presented at the 13th IEEE International Conference on Bioinformatics and Bioengineering.

Bowman, J., Dudek, F. E., & Spitz, M. (2001). Epilepsy. *eLS*.

Bugeja, S., Garg, L., & Audu, E. E. (2016). *A novel method of EEG data acquisition, feature extraction and feature space creation for early detection of epileptic seizures.* Paper presented at the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

Cech, D. J. (2011). *Functional movement development across the life span*: Elsevier Health Sciences.

Chaurasia, V., Pal, S., & Tiwari, B. (2018). Chronic kidney disease: a predictive model using decision tree. *International Journal of engineering Research and Technology, 11*(11), 1781-94.

Choi, G., Park, C., Kim, J., Cho, K., Kim, T.-J., Bae, H., . . . Chong, J. (2019). *A novel multi-scale 3D CNN with deep neural network for epileptic seizure detection.* Paper presented at the 2019 IEEE International Conference on Consumer Electronics (ICCE).

Choubey, H., & Pandey, A. (2021). A combination of statistical parameters for the detection of epilepsy and EEG classification using ANN and KNN classifier. *Signal, Image and Video Processing, 15*(3), 475-483.

Choudhury, A., & Gupta, D. (2019). A survey on medical diagnosis of diabetes using machine learning techniques. In *Recent developments in machine learning and data analytics* (pp. 67-78): Springer.

Da Silva, F. L., Vos, J., Mooibroek, J., & Van Rotterdam, A. (1980). Relative contributions of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analysis. *Electroencephalography and Clinical Neurophysiology, 50*(5-6), 449-456.

Darbas, M., & Lohrengel, S. (2019). Review on mathematical modelling of electroencephalography (EEG). *Jahresbericht der Deutschen Mathematiker-Vereinigung, 121*(1), 3-39.

Davis, L. E., & Pirio Richardson, S. (2015). Seizures and status epilepticus. In *Fundamentals of Neurologic Disease* (pp. 189-198): Springer.

El-Gindy, S. A.-E., Hamad, A., El-Shafai, W., Khalaf, A. A., El-Dolil, S. M., Taha, T. E., . . . El-Samie, F. E. A. (2021). Efficient communication and EEG signal classification in wavelet domain for epilepsy patients. *Journal of Ambient Intelligence and Humanized Computing, 12*(10), 9193-9208.

Eva, O. D., & Lazar, A. M. (2015). Comparison of classifiers and statistical analysis for EEG signals used in brain computer interface motor task paradigm. *International Journal of Advanced Research in Artificial Intelligence (IJARAI), 1*(4), 8-12.

Fathima, T., Bedeeuzzaman, M., Farooq, O., & Khan, Y. U. (2011). Wavelet based features for epileptic seizure detection. *MES Journal of Technology and Management, 2*(1), 108-112.

Faust, O., Acharya, U. R., Adeli, H., & Adeli, A. (2015). Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure, 26*, 56-64.

Fernández, I. S., Loddenkemper, T., Gaínza-Lein, M., Sheidley, B. R., & Poduri, A. (2019). Diagnostic yield of genetic tests in epilepsy: A meta-analysis and cost-effectiveness study. *Neurology, 92*(5), e418-e428.

Fielding, A. H. (2006). *Cluster and classification techniques for the biosciences*: Cambridge University Press.

Gandhi, T., Panigrahi, B. K., & Anand, S. (2011). A comparative study of wavelet families for EEG signal classification. *Neurocomputing, 74*(17), 3051-3057.

Gazalba, I., & Reza, N. G. I. (2017). *Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification.* Paper presented at the 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE).

Gifu, D. (2021). The Use of Decision Trees for Analysis of the Epilepsy. *Procedia Computer Science, 192*, 2844-2853.

Gómez, C., Arbeláez, P., Navarrete, M., Alvarado-Rojas, C., Le Van Quyen, M., & Valderrama, M. (2020). Automatic seizure detection based on imaged-EEG signals through fully convolutional networks. *Scientific reports, 10*(1), 1-13.

Gorard, S. (2005). Revisiting a 90-year-old debate: the advantages of the mean deviation. *British Journal of Educational Studies, 53*(4), 417-430.

Graimann, B., Allison, B. Z., & Pfurtscheller, G. (2010). *Brain-computer interfaces: Revolutionizing human-computer interaction*: Springer Science & Business Media.

Gu, Z., Yan, G., Zhang, J., Li, Y., & Yu, Z. L. (2018). Automatic epilepsy detection based on wavelets constructed from data. *IEEE Access, 6*, 53133-53140.

Guo, L., Rivero, D., Dorado, J., Munteanu, C. R., & Pazos, A. (2011). Automatic feature extraction using genetic programming: An application to epileptic EEG classification. *Expert Systems with Applications, 38*(8), 10425-10436.

Gupta, A., Singh, P., & Karlekar, M. (2018). A novel signal modeling approach for classification of seizure and seizure-free EEG signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(5), 925-935.

Gupta, V., & Pachori, R. B. (2019). Epileptic seizure identification using entropy of FBSE based EEG rhythms. *Biomedical Signal Processing and Control, 53*, 101569.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.

Hassan, A. R., & Haque, M. A. (2015). *Epilepsy and seizure detection using statistical features in the complete ensemble empirical mode decomposition domain.* Paper presented at the TENCON 2015-2015 IEEE region 10 conference.

He, B., Baxter, B., Edelman, B. J., Cline, C. C., & Wenjing, W. Y. (2015). Noninvasive brain-computer interfaces based on sensorimotor rhythms. *Proceedings of the IEEE, 103*(6), 907-925.

Hughes, S. W., Lörincz, M., Cope, D. W., Blethyn, K. L., Kékesi, K. A., Parri, H. R., . . . Crunelli, V. (2004). Synchronized oscillations at α and θ frequencies in the lateral geniculate nucleus. *Neuron, 42*(2), 253-268.

Janjarasjitt, S. (2017). Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and SVM. *Medical & Biological Engineering & Computing, 55*(10), 1743-1761.

Jia, A. D., Li, B. Z., & Zhang, C. C. (2020). Detection of cervical cancer cells based on strong feature CNN-SVM network. *Neurocomputing, 411*, 112-127.

Jiang, Y., Chen, W., & Li, M. (2020). Symplectic geometry decomposition-based features for automatic epileptic seizure detection. *Computers in Biology and Medicine, 116*, 103549.

Juárez-Guerra, E., Alarcon-Aquino, V., & Gomez-Gil, P. (2015). Epilepsy seizure detection in EEG signals using wavelet transforms and neural networks. In *New trends in networking, computing, E-learning, systems sciences, and engineering* (pp. 261-269): Springer.

Kaleem, M., Gurve, D., Guergachi, A., & Krishnan, S. (2018). Patient-specific seizure detection in long-term EEG using signal-derived empirical mode decomposition (EMD)-based dictionary approach. *Journal of Neural Engineering, 15*(5), 056004.

Kannathal, N., Chee, J., Er, K., Lim, K., & Tat, O. H. (2014). *Chaotic analysis of epileptic EEG signals.* Paper presented at the The 15th International Conference on Biomedical Engineering.

Kaur, S., & Singh, M. G. (2012). *Detection of epilepsy disorder by EEG using discrete wavelet transforms.* [Doctoral dissertation]. Thapar University.

Kaushik, G., Gaur, P., Sharma, R. R., & Pachori, R. B. (2022). EEG signal based seizure detection focused on Hjorth parameters from tunable-Q wavelet sub-bands. *Biomedical Signal Processing and Control, 76*, 103645.

Khan, Y., & Gotman, J. (2003). Wavelet based automatic seizure detection in intracerebral electroencephalogram. *Clinical Neurophysiology, 114*(5), 898-908.

Kumar, R., & Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics, 48*(4), 277-287.

Kumar, Y., Dewal, M., & Anand, R. S. (2014). Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine. *Neurocomputing, 133*, 271-279.

Li, M., Chen, W., & Zhang, T. (2017). Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. *Biomedical Signal Processing and Control, 31*, 357-365.

Li, M., Xu, H., Liu, X., & Lu, S. (2018). Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. *Technology and Health Care, 26*(S1), 509-519.

Lőrincz, M. L., Kékesi, K. A., Juhász, G., Crunelli, V., & Hughes, S. W. (2009). Temporal framing of thalamic relay-mode firing by phasic inhibition during the alpha rhythm. *Neuron, 63*(5), 683-696.

Luhach, A. K., Kosa, J. A., Poonia, R. C., Gao, X.-Z., & Singh, D. (2020). *First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019*: Springer.

Mahjoub, C., Jeannès, R. L. B., Lajnef, T., & Kachouri, A. (2020). Epileptic seizure detection on EEG signals using machine learning techniques and advanced preprocessing methods. *Biomedical Engineering/Biomedizinische Technik, 65*(1), 33-50.

Mahmoud, M. I., Dessouky, M. I., Deyab, S., & Elfouly, F. H. (2007). Comparison between haar and daubechies wavelet transformions on FPGA technology. *World Academy of Science, Engineering and Technology, 26*(1), 68-72.

Malmgren, K., Reuber, M., & Appleton, R. (2012). Differential diagnosis of epilepsy. *Oxford textbook of epilepsy and epileptic seizures*, 81-94.

Martis, R. J., Acharya, U. R., Tan, J. H., Petznick, A., Yanti, R., Chua, C. K., . . . Tong, L. (2012). Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals. *International Journal of Neural Systems, 22*(06), 1250027.

Nabil, D., Benali, R., & Reguig, F. B. (2020). Epileptic seizure recognition using EEG wavelet decomposition based on nonlinear and statistical features with support vector machine classification. *Biomedical Engineering/Biomedizinische Technik, 65*(2), 133-148.

Nanthini, B. S., & Santhi, B. (2017). Electroencephalogram signal classification for automated epileptic seizure detection using genetic algorithm. *Journal of Natural Science, Biology, and Medicine, 8*(2), 159.

Ngui, W. K., Leong, M. S., Hee, L. M., & Abdelrhman, A. M. (2013). Wavelet analysis: mother wavelet selection methods. *Applied mechanics and materials, 393*, 953-958.

Nkengfack, L. C. D., Tchiotsop, D., Atangana, R., Louis-Door, V., & Wolf, D. (2020). EEG signals analysis for epileptic seizures detection using polynomial transforms, linear discriminant analysis and support vector machines. *Biomedical Signal Processing and Control, 62*, 102141.

Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications, 36*(2), 2027-2036.

Oktavia, N. Y., Wibawa, A. D., Pane, E. S., & Purnomo, M. H. (2019). *Human emotion classification based on EEG signals using Naïve bayes method.* Paper presented at the 2019 International Seminar on Application for Technology of Information and Communication (iSemantic).

Omidvar, M., Zahedi, A., & Bakhshi, H. (2021). EEG signal processing for epilepsy seizure detection using 5-level Db4 discrete wavelet transform, GA-based feature selection and ANN/SVM classifiers. *Journal of Ambient Intelligence and Humanized Computing, 12*(11), 10395-10403.

or Rashid, M. M., & Ahmad, M. (2017). *Epileptic seizure classification using statistical features of EEG signal.* Paper presented at the 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE).

Orosco, L., Correa, A. G., Diez, P., & Laciar, E. (2016). Patient non-specific algorithm for seizures detection in scalp EEG. *Computers in Biology and Medicine, 71*, 128-134.

Orosco, L., Correa, A. G., & Laciar, E. (2013). A survey of performance and techniques for automatic epilepsy detection. *Journal of Medical and Biological Engineering, 33*(6), 526-537.

Patel, K., Chua, C.-P., Fau, S., & Bleakley, C. J. (2009). *Low power real-time seizure detection for ambulatory EEG.* Paper presented at the 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare.

Pathak, S., Waheed, A., & Mirza, R. (2016). Epilepsy: A brief review. *PharmaTutor, 4*(9), 21-28.

Peachap, A. B., & Tchiotsop, D. (2019). Epileptic seizures detection based on some new Laguerre polynomial wavelets, artificial neural networks and support vector machines. *Informatics in Medicine Unlocked, 16*, 100209.

Perucca, E., & Tomson, T. (2011). The pharmacological treatment of epilepsy in adults. *The Lancet Neurology, 10*(5), 446-456.

Pfurtscheller, G., & Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proceedings of the IEEE, 89*(7), 1123-1134.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). Circuits within the basal ganglia system. In *Neuroscience. 2nd edition*: Sinauer Associates.

Qureshi, M. B., Afzaal, M., Qureshi, M. S., & Fayaz, M. (2021). Machine learning-based EEG signals classification model for epileptic seizure detection. *Multimedia Tools and Applications, 80*(12), 17849-17877.

Rabeh, A. B., Benzarti, F., & Amiri, H. (2016). *Diagnosis of Alzheimer diseases in early step using SVM (support vector machine).* Paper presented at the 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV).

Rajaguru, H., & Prabhakar, S. K. (2017). *Time frequency analysis (dB2 and dB4) for Epilepsy classification with LDA classifier.* Paper presented at the 2017 2nd International Conference on Communication and Electronics Systems (ICCES).

Ravi Kumar, M., & Srinivasa Rao, Y. (2019). Epileptic seizures classification in EEG signal based on semantic features and variational mode decomposition. *Cluster Computing, 22*(6), 13521-13531.

Ren, Y., & Wu, Y. (2014). *Convolutional deep belief networks for feature extraction of EEG signal.* Paper presented at the 2014 International Joint Conference on Neural Networks (IJCNN).

Rohan, R., & Kumari, L. R. (2021). *Classification of sleep apneas using decision tree classifier.* Paper presented at the 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).

Saeedi, M., Saeedi, A., & Maghsoudi, A. (2020). Major depressive disorder assessment via enhanced k-nearest neighbor method and EEG signals. *Physical and Engineering Sciences in Medicine, 43*(3), 1007-1018.

Sairamya, N., Premkumar, M. J., George, S. T., & Subathra, M. (2021). Performance evaluation of discrete wavelet transform, and wavelet packet decomposition for automated focal and generalized epileptic seizure detection. *IETE Journal of Research, 67*(6), 778-798.

Sameer, M., Gupta, A. K., Chakraborty, C., & Gupta, B. (2019). *Epileptical seizure detection: Performance analysis of gamma band in EEG signal using short-time Fourier transform.* Paper presented at the 2019 22nd International Symposium on Wireless Personal Multimedia Communications (WPMC).

Sameer, M., & Gupta, B. (2021). ROC analysis of EEG subbands for epileptic seizure detection using naïve bayes classifier. *Journal of Mobile Multimedia*, 299–310-299–310.

Saritas, M. M., & Yasar, A. (2019). Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *International Journal of Intelligent Systems and Applications in Engineering, 7*(2), 88-91.

Savadkoohi, M., Oladunni, T., & Thompson, L. (2020). A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. *Biocybernetics and Biomedical Engineering, 40*(3), 1328-1341.

Sharmila, A., Aman Raj, S., Shashank, P., & Mahalakshmi, P. (2018). Epileptic seizure detection using DWT-based approximate entropy, Shannon entropy and support vector machine: a case study. *Journal of Medical Engineering & Technology, 42*(1), 1-8.

Shoeb, A. H., & Guttag, J. V. (2010). *Application of machine learning to epileptic seizure detection.* Paper presented at the ICML.

Shoeibi, A., Ghassemi, N., Khodatars, M., Moridian, P., Alizadehsani, R., Zare, A., . . . Gorriz, J. M. (2022). Detection of epileptic seizures on EEG signals using ANFIS classifier, autoencoders and fuzzy entropies. *Biomedical Signal Processing and Control, 73*, 103417.

Subasi, A., & Gursoy, M. I. (2010). EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications*, *37*(12), 8659-8666.

Supriya, S., Siuly, S., Wang, H., & Zhang, Y. (2020). Automated epilepsy detection techniques from electroencephalogram signals: a review study. *Health Information Science and Systems, 8*(1), 1-15.

Sutton, F., Barca, D., Komoltsev, I., Craiu, D., Guekht, A., von Oertzen, T., & Cock, H. R. (2020). Testing blood and CSF in people with epilepsy: a practical guide. *Epileptic Disorders, 22*(4), 381-398.

Tanveer, M., Pachori, R. B., & Angami, N. (2018). *Classification of seizure and seizure-free EEG signals using Hjorth parameters.* Paper presented at the 2018 IEEE Symposium Series on Computational Intelligence (SSCI).

Tăuţan, A.-M., Mândruţă, I., Băjenaru, O.-A., Strungaru, R., Ţarălungă, D., Hurezeanu, B., & Mihaela Neagu Ungureanu, G. (2019). *The Automatic Detection of Epileptic Seizures Based on EEG Signals Processing: Investigation of Different Features and Classification Algorithms.* Paper presented at the World Congress on Medical Physics and Biomedical Engineering 2018.

Thenmozhi, K., & Deepika, P. (2014). Heart disease prediction using classification with different decision tree techniques. *International Journal of Engineering Research and General Science*, *2*(6), 6-11.

Tohumoglu, G., & Yol, Ş. (2021). *A Short Tutorial: From Neuroscience to Detection of Pseudo Epileptic Seizure.* Paper presented at the 2021 Medical Technologies Congress (TIPTEKNO).

Tong, Y., Aliyu, I., & Lim, C.-G. (2018). Analysis of dimensionality reduction methods through epileptic EEG feature selection for machine learning in BCI. *The Journal of the Korea Institute of Electronic Communication Sciences, 13*(6), 1333-1342.

Torse, D. A., & Khanai, R. (2021). *Classification of Epileptic Seizures using Ensemble Empirical Mode Decomposition and Least Squares Support Vector Machine.* Paper presented at the 2021 International Conference on Computer Communication and Informatics (ICCCI).

Tsiouris, K. M., Markoula, S., Konitsiotis, S., Koutsouris, D. D., & Fotiadis, D. I. (2018). A robust unsupervised epileptic seizure detection methodology to accelerate large EEG database evaluation. *Biomedical Signal Processing and Control, 40*, 275-285.

Türk, Ö., Şeker, M., Akpolat, V., & Özerdem, M. S. (2017). *Classification of mental task EEG records using Hjorth parameters.* Paper presented at the 2017 25th Signal Processing and Communications Applications Conference (SIU).

Tzimourta, K., Tzallas, A., Giannakeas, N., Astrakas, L., Tsalikakis, D., & Tsipouras, M. (2017). *Epileptic seizures classification based on long-term EEG signal wavelet analysis.* Paper presented at the International Conference on Biomedical and Health Informatics.

Ullah, I., Hussain, M., & Aboalsamh, H. (2018). An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Systems with Applications, 107*, 61-71.

Upadhyay, R., Padhy, P., & Kankar, P. K. (2016). A comparative study of feature ranking techniques for epileptic seizure detection using wavelet transform. *Computers & Electrical Engineering, 53*, 163-176.

Van Esbroeck, A., Smith, L., Syed, Z., Singh, S., & Karam, Z. (2016). Multi-task seizure detection: addressing intra-patient variation in seizure morphologies. *Machine Learning, 102*(3), 309-321.

Venkatesan, E., & Velmurugan, T. (2015). Performance analysis of decision tree algorithms for breast cancer classification. *Indian Journal of Science and Technology, 8*(29), 1-8.

Vogado, L. H., Veras, R. M., Araujo, F. H., Silva, R. R., & Aires, K. R. (2018). Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. *Engineering Applications of Artificial Intelligence, 72,* 415-422.

Wang, L., Xue, W., Li, Y., Luo, M., Huang, J., Cui, W., & Huang, C. (2017). Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis. *Entropy, 19*(6), 222.

Wang, Y., Dai, Y., Liu, Z., Guo, J., Cao, G., Ouyang, M., . . . Zhao, G. (2021). Computer-aided intracranial EEG signal identification method based on a multi-branch deep learning fusion model and clinical validation. *Brain Sciences, 11*(5), 615.

Wirrell, E. C. (2010). Prognostic significance of interictal epileptiform discharges in newly diagnosed seizure disorders. *Journal of Clinical Neurophysiology, 27*(4), 239-248.

Xie, S., & Krishnan, S. (2013). Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. *Medical & Biological Engineering & Computing, 51*(1), 49-60.

Yaffe, R. B., Borger, P., Megevand, P., Groppe, D. M., Kramer, M. A., Chu, C. J., . . . Sarma, S. V. (2015). Physiology of functional and effective networks in epilepsy. *Clinical Neurophysiology, 126*(2), 227-236.

Yan, J., Wang, Y., Ouyang, G., Yu, T., Li, Y., Sik, A., & Li, X. (2016). Analysis of electrocorticogram in epilepsy patients in terms of criticality. *Nonlinear Dynamics, 83*(4), 1909-1917.

Yol, Ş., & TohumoĞlu, G. (2020). *EEG Sinyallerinde Epilepsi ve Psödo (sahte) Nöbetlerinin Belirlenmesi: Kısa Derleme Detection Methods of Pseudo and Epileptic Seizures from EEG signals: A Short Review.* Paper presented at the 2020 Medical Technologies Congress (TIPTEKNO).

Yu, Z., Li, Y., Yuan, Q., & Zhou, W. (2018). *Epileptic seizure detection based on local mean decomposition and dictionary pair learning.* Paper presented at the 2018 International Conference on Information Systems and Computer Aided Education (ICISCAE).

Zandi, A. S., Javidan, M., Dumont, G. A., & Tafreshi, R. (2010). Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform. *IEEE Transactions on Biomedical Engineering, 57*(7), 1639-1651.

Zeidman, L. A., Stone, J., & Kondziella, D. (2014). New revelations about Hans Berger, father of the electroencephalogram (EEG), and his ties to the Third Reich. *Journal of Child Neurology, 29*(7), 1002-1010.

Zhang, T., & Chen, W. (2016). LMD based features for the automatic seizure detection of EEG signals using SVM. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25*(8), 1100-1108.