İZMİR KÂTİP ÇELEBİ ÜNİVERSİTESİ
2010
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

Analysis of Multichannel EEG Signals for the Detection of Major Depressive Disorder

Submitted to the Graduate School of Natural and Applied Sciences
in partial fulfillment of the requirements for the degree of

Master of Science

in Biomedical Technologies

by
Elif İZCİ

ORCID 0000-0003-1148-8374

August 4, 2021

This is to certify that we have read the thesis **Analysis of Multichannel EEG Signals for the Detection of Major Depressive Disorder** submitted by **Elif İzci**, and it has been judged to be successful, in scope and in quality, at the defense exam and accepted by our jury as a MASTER'S THESIS.

APPROVED BY:

Advisor: **Prof. Dr. Aydın AKAN**
Izmir University of Economics

Co-advisor: **Dr. Onan GÜREN**
İzmir Kâtip Çelebi University

Committee Members:

Doç.Dr. Ahmet MERT
Bursa Technical University

Doç. Dr. Utku Kürşat ERCAN
İzmir Kâtip Çelebi University

Date of Defense: 04.08.2021

Declaration of Authorship

I, **Elif İzci**, declare that this thesis titled **Analysis of Multichannel EEG Signals for the Detection of Major Depressive Disorder** and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for the Master's degree at this university.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. This thesis is entirely my own work, with the exception of such quotations.
- I have acknowledged all major sources of assistance.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date: 28.06.2021

Analysis of Multichannel EEG Signals for the Detection of Major Depressive Disorder

Abstract

Major depressive disorder (MDD) is a common mood disorder encountered worldwide. The diagnosis of MDD is based on qualitative methods such as psychometric questionnaires, patient self-report, and the professional's clinical experience. The aim of this thesis is to develop an quantitative method to differentiate MDD patients from healthy controls (HCs). Electroencephalogram (EEG)-based computer-aided (CAD) methods have great attention due to the success of diagnosing MDD. In this thesis, three approaches for identifying MDD by analyzing multichannel EEG signals are proposed.

EEG signals taken from 16 MDD patients and 16 HCs were analyzed in the first proposed method according to the brain region, and time-domain, frequency-domain, and nonlinear features were extracted. The feature sets are classified using four different classification algorithms. As a result of the study, the classification accuracy of 86.7% was achieved using the Support Vector Machine (SVM) with the features extracted from the frontal region EEG channels.

Empirical Mode Decomposition (EMD) is introduced in the second proposed method. In the method, EEG channels were decomposed into Intrinsic Mode Functions (IMFs) and the first two IMFs (IMF1 and IMF2) of a channel were selected to act as a new channel. After that, the selected features were extracted from the channels and their two IMFs. The feature sets are classified using four different classification algorithms. The highest classification of 91.5% was achieved with the combination of EMD and Linear Discriminant Analysis (LDA) classifier for classification of the parietal region of EEG channels.

The third method proposes Deep Learning (DL) approach to identify MDD in EEG topographic images. The proposed approach identifies MDD classes using Convolutional Neural Network (CNN) architecture trained by two-dimensional (2D) EEG topographic images. The images were obtained from the EEG signals which represent the spread of a feature to the whole cortex and used as inputs in CNN architecture to classify MDD and HC. The results show that the classification performance of the proposed approach reaches a validation ACC of 80.09%.

Automatic detection of MDD can be potential with the approaches through EEG signals. In future studies, developing the proposed methods may provide a crucial decision-making system during the diagnosis of MDD for professionals.

Keywords: Major depressive dissorder, electroencephalogram, machine learning, empirical mode decomposition, classification

Majör Depresif Bozukluğunun Tespiti için Çok Kanallı EEG Sinyallerinin Analizi

Öz

Majör depresif bozukluk (MDB), dünya çapında yaygın olarak görülen bir duygudurum bozukluğudur. MDB tanısı, psikometrik anketler, hastanın kendi raporu ve profesyonelin klinik deneyimi gibi öznel yöntemlere dayanır. Bu tezin amacı, MDB hastalarını sağlıklı kişilerden (HC) ayırt etmeyi sağlayan objektif bir yöntem geliştirmektir. Elektroensefalograma (EEG) dayalı bilgisayar destekli (CAD) algoritmalar, MDB tanısındaki başarısı nedeniyle büyük ilgi görmektedir. Bu tezde, çok kanallı EEG sinyallerini analiz ederek MDB hastalığını ayırt etmek için için üç yaklaşım önerilmiştir.

Önerilen ilk yöntemde 16 MDB hastası ve 16 HC'den alınan EEG sinyalleri beyin bölgelerine göre analiz edilerek sinyallerden zaman alanı, frekans alanı ve doğrusal olmayan öznitelikler çıkarılmıştır. Öznitelik kümeleri, dört farklı sınıflandırma algoritması kullanılarak sınıflandırılmıştır. Çalışma sonucunda, frontal bölge EEG kanallarından çıkarılan öznitelikler ile Destek Vektör Makinesi (SVM) kullanılarak %86,7 sınıflandırma doğruluğu elde edilmiştir.

Görgül Kip Ayrışımı (EMD), önerilen ikinci yöntemde tanıtılmıştır. Yöntemde, EEG kanalları İçsel Mod Fonksiyonlarına (IMF'ler) ayrıştırılmış ve bir kanalın ilk iki IMF'si (IMF1 ve IMF2) yeni bir kanal gibi sayılacak şekilde seçilmiştir. Seçilen öznitelikler kanallardan ve kanalların ilk iki IMF'sinden çıkarılmıştır. Oluşan öznitelik kümeleri, dört farklı sınıflandırma algoritması kullanılarak sınıflandırılmıştır. EEG kanallarının paryetal bölgesinin sınıflandırılması için EMD ve Doğrusal Ayırma Analizi (LDA) sınıflandırıcısının kombinasyonu ile %91,5 ile en yüksek sınıflandırma doğruluğu elde edilmiştir.

Üçüncü yöntem, EEG topografik görüntülerinden MDB'yi tanımlamak için Derin Öğrenme (DL) yaklaşımı önerilmiştir. Önerilen yaklaşım, iki boyutlu (2B) EEG topografik görüntülerle eğitilmiş Evrişimsel Sinir Ağı (CNN) mimarisini kullanarak MDB sınıfını tanımlar. Görüntüler, bir özniteliğin tüm kortekse yayılmasını temsil eden ve MDB ve HC'yi sınıflandırmak için CNN mimarisinde girdi olarak kullanılan EEG sinyallerinden elde edilmiştir. Sonuçlar, önerilen yaklaşımın sınıflandırma performansının %80.09'luk bir doğruluğa ulaştığını göstermektedir.

MDB'nin otomatik ve objektif olarak tespit edilmesi, EEG sinyallerinin kullanılmasıyla önerilen yaklaşımlarla oluşturulabilir. Gelecekteki çalışmalarla önerilen yöntemlerin geliştirilmesi, profesyoneller için MDB'nin teşhisi sırasında önemli bir karar verme sistemi sağlayabilir.

Anahtar Kelimeler: Majör depresif bozukluk, elektroensefalogram, makine öğrenmesi, görgül kip ayrışımı, sınıflandırma

First and foremost, I would like to thank Prof. Dr. Aydın AKAN for his guidance during his time as my advisor, and for his advice and constant support throughout my thesis work. It is thanks to him that I wanted to work on signal processing area and his encourement will always be shaping my career in the future.

I also would like to thank Dr. Onan Güren for her guidance, helping, advice during her time as my co-advisor

I would like to thank Prof. Dr. Kemal ARIKAN and Dr. Mehmet Akif OZCOBAN for their efforts in recording and transmitting EEG data to us.

I also would like to thank Research Assistant Mehmet Akif ÖZDEMİR and Murside DEGİRMENCİ for helping me during the thesis period.

I would like to offer greatest thanks my parents, Necibe İZCİ, Süleyman İZCİ and Selim İZCİ for their love, support and always being with me.

Acknowledgment

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) Scientist Support Department (BIDEB) under the 2210/C program number 1649B021906973.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

IKCU	Izmir Kâtip Çelebi University
TUBITAK	Scientific and Technological Research Council of Turkey
ORCID	Open Researcher and Contributor ID
EEG	Electroencephalography
MDD	Major Depressive Disorder
HC	Healthy Control
CAD	Computer-aided design
ML	Machine Learning
CNS	Central Nervous System
PNS	Peripheral Nervous System
DSM	Diagnostic and Statistical Maual of Mental Disorders
WHO	World Health Organization
ECT	Electroconvulsive Therapy
BDI	Beck’s Depression Inventory
HAM-D	Hamilton Depression Rating Scale
SVM	Support Vector Machine
*k*NN	*k*-Nearest Neigbor
CT	Classification Tree
ANN	Artifical Neural Network
DT	Decition Tree
LR	Logistic Regression
NB	Naive-Bayesian

KEFB-SCP	Kernel Eigen-filter-bank Common Spatial Pattern
DL	Deep Learning
CNN	Convolutional Neural Network
EMD	Empirical Mode Decomposition
IMF	Instrictic Mode Function
PSD	Power Spectral Density
FFT	Fast Fourier Transform
LDA	Linear Discriminant Analysis
ReLU	Rectified linear unit
ACC	Accuracy
REC	Recall
SPE	Specificity
PRE	Precision
ROC	Receiver Operating Characteristic
AUC	Area Under the ROC Curve
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

Chapter 1

Introduction

1.1 Motivation

Major depressive disorder (MDD) commonly termed depression is a debilitating mood disorder that seriously decreases the quality of life. It is a social, economic burden, and life-threatening event that causes increased risk of health problems such as cardiovascular diseases [1], diabetes mellitus [2]. Furthermore, MDD can trigger suicidal thoughts. According to the World Health Organization (WHO), more than 300 million people are suffered from depression worldwide [3] and it is estimated that depression will be the leading global burden of diseases by 2030 [4]. Early and accurate diagnosis of depression has great importance for applying effective treatment and preventing severe problems. Recently, the diagnosis of depression is based on qualitative methods which are psychometric questionnaires, patient's self-reporting, and professional's clinical experiences. Due to subjectivity, the accuracy of diagnosis is influenced by various factors including professional proficiency, uncertain and confused patient symptoms (related to other diseases). These factors may cause misdiagnosis, not selected optimal treatment, or delaying patient recovery time. Therefore, developing quantitative and effective method is significantly important to improve diagnostic efficiency and accuracy, assist professionals, and decrease depression diagnosis time. As explained in the section on 'Literature Review' of the thesis, there are several scientific studies for diagnosis of depression using computer-aided designs (CAD) and engineering methods. In recent studies, Electroencephalograph (EEG)-based machine learning (ML) techniques have great attention due to their properties which are noninvasive, reliable, high time resolution, cost, and easy operation. With these effective potentials, the analysis on the field of MDD detection and prediction from EEG signals offers an area open to development.

1.2 Nervous System

The nervous system is a network that coordinates different body functions. It provides communication between the brain and the rest of the body to control many processes such as movement, memories, internal organs activity. The nervous system is divided into two main parts called central nervous system (CNS) and peripheral nervous system (PNS). CNS consists of the brain and the spinal cord and PNS consists of nerves that are connected to sensory organs. Nerves receive the signals from the sensory organs and transmit them to the CNS. The signals are processed in the brain and an appropriate response is formed. The response is transmitted via nerves to the sensory organs [5].

A neuron is the basic unit of the nervous system that specialized to receive and transmit information to different parts of the body. It consists of three main parts, cell body, an axon, and dendrite as represented in Figure 1.1. There are three neuron types according to their roles: sensory neuron, motor neuron and interneurons. Transmission between neurons is provided by electrical and chemical signals [6].

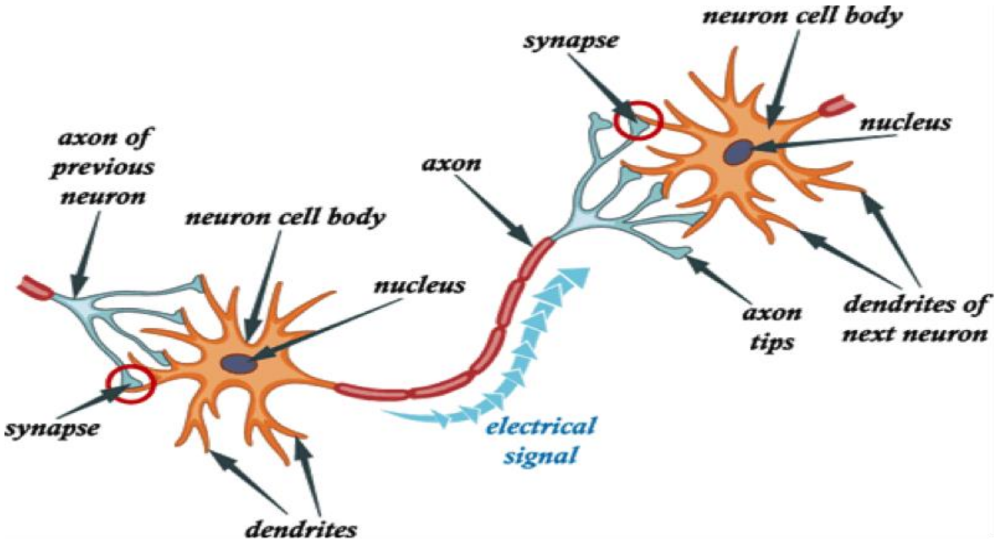

Figure 1.1: Diagram of neurons structure [6]

When a neuron sends information to another neuron, it transmits electrical impulses throughout its axon. Dendrites of another neuron receive impulses from the neuron.

The impulses pass through the cell body and reaches the axon. Electrical impulses are converted to chemical signals at the end of the axon. Chemical signals are transmitted as neurotransmitters from the axon to synapse which is the connection of axon and another neuron cell. The neurotransmitters go through synapse, reach another neuron dendrite and are converted to electrical impulses [5].

Neurological disorders are types of diseases that affect the central and peripheral nervous system include the brain, spinal cord, nerves. They are serious issues for causing severe health problems and leading cause death around the world [7]. Neurological disorders comprise a wide range of diseases as Alzheimer’s disease, dementia, epilepsy, stroke, Parkinson's disease, migraine, tumors.

1.2.1 Brain

The brain is an organ that is located in the head and is protected by the skull. It is most important part of the CNS that performs the management of body functions such as motor activities (movement, speech, walking etc.) and cognitive activities (breathing, thinking, sensations, memory etc.). There are three main parts of the brain that are called as cerebellum, brainstem, and cerebrum. The cerebellum is located at the back and bottom of the brain. It regulates balance, maintains posture, and coordinates body movements. The brainstem is an interface between the brain and spinal cord that directly connects cerebrum to the spinal cord. It regulates automatic functions of the body such as heart rate, breathing, digestion, sleep stages. It is also responsible for reflexes. The cerebrum is the largest part of the brain that has major role for regulating conscious activities. It accepts and processes conscious action, interpreting and creates thought, movement or other conscious action [8].

The cerebrum divided into two hemispheres, right and left hemispheres that are responsible for controlling different functions. Each hemisphere controls the opposite part of the body. If the left hemisphere is damaged, vision, movement, or hearing functions of the right side can be affected. The brain hemispheres subdivided into lobes that associated with different functions. The hemispheres and lobes of the brain are shown as in Figure 1.2. There are four lobes of each hemisphere: frontal, temporal, occipital and parietal lobes. Functions of the lobes are given as follows;

Frontal lobe: It is largest lobe and is located in the front of the brain. It has motor cortex, prefrontal cortex and Broca's area that are provide to control different type of functions. Frontal lobe controls cognitive functions, movement, language, emotional behaviours.

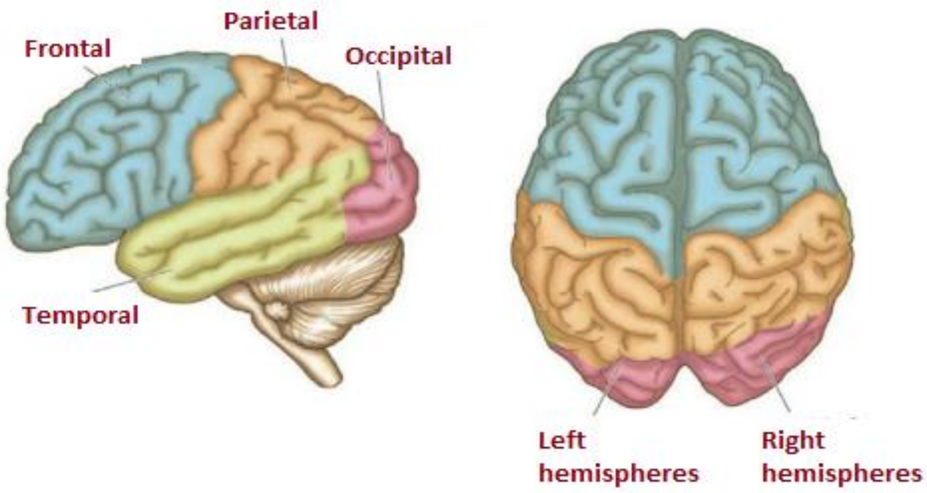

Figure 1.2: Brain hemispheres and lobes [9]

Temporal lobe: It is located below the frontal lobe and near the ears. Temporal lobe controls auditory process, visual and verbal memories, smell, language understanding.

Occipital lobe: It is posterior part of the brain and is associated with vision process. It contains visual cortex that allows to receive and analyze visual information.

Parietal lobe: It is middle part of the brain that is located behind the frontal lobe. It is involved in interpreting sensory information from different parts of the brain such as pain, temperature, touch.

1.3 Psychiatric Disorders

Psychiatric disorders cause disability of behavior, emotion, and/or cognitive functions. According to literature, 30% of the world population are affected by mental disorders in a year [10]. Mental disorders impair quality of life due to change a person's thinking, feeling, moods. Changing of these factors is not be easily controlled by the person and causes distress or disability in social, environmental, occupational areas. The special causes of mental disorders are not entirely clear, different factors can contribute to developing the disorders such as genetic, environmental, social factors, stress, chemical imbalances in the brain, health problems as trauma, brain, injury, acute/chronic diseases as diabetes, cancer. These factors may affect the nervous system negatively to disturbances in the neurotransmitter system, dysfunction of the brain [11]. Almost 300 mental disorders are listed in Diagnostic and Statistical Manual of Mental Disorders (DSM-5) which is a handbook to guide identifying, classifying, and diagnosing mental disorders [12]. Most common types are schizophrenia, mood, personality, eating, anxiety disorders.

Mood disorders are types of psychiatric disorders characterized by disturbance in mood and psychomotor activity. Different symptoms can be seen in a person depending on severity of mood disorder. Most pervasive symptoms are ongoing sad, angry, irritable mood, irregular sleep or eating stage, slowing of thought, speech, and motion or feeling worthlessness and guilt. Mood disorders are threating factor on nervous system due to cause expansion of neurological disorders [13]. Some types of mood disorders are listed as below:

- Major depressive disorder
- Bipolar disorder
- Dysthymia
- Cyclothymic disorder
- Seasonal affective disorder
- Mood disorder related to different disease

1.3.1 Major Depressive Disorder

MDD is a potentially life-threatening mood disorder that affects millions of people aound the world. It can develop with devastating, debilitating impacts at any age of a person. It is classified as mild, moderate, and severe related to the severity of symptoms. Various symptoms are seen with MDD such as feeling of sadness, impairment of concentration and thinking, loss of pleasure in daily activity. Furthermore, MDD is a trigger suicide factor. According to WHO, more than 300 million people suffered from MDD, and nearly 800,000 people are dying from it in a year [3]. In studies, women are more affected by MDD than men [14]. MDD may contribute to developing other diseases such as cardiovascular disease [1], diabetes mellitus [15], or stroke [16].

1.3.1.1 Etiology

Even though there are many theories related to the pathophysiology of depression, it is not clearly explained what causes it and the mechanism of MDD development. The cause of the disorder is associated with various factors such as genetic factors, stressful environment, imbalance of hormonal activity, trauma, cognitive impairment, other medical problems [17]. Some of them are explained as follows;

Genetics: There is no single gene for causing depression but different genetic factors can susceptible to the development of MDD. According to twin genetic studies, the heritability of MDD has been estimated at 37% [18]. Family studies have indicated that first-degree relatives of MDD patients demonstrate a threefold increased risk of MDD compared with first-degree relatives of control ones [19].

Environmental factors: Childhood trauma or stressful living areas could be a significant risk factor for MDD. Early life stress reasons by exposure to trauma, grind, violence, or any form of mistreatment causes a two-fold increased risk of MDD in adults [20].

Other health problems: Studies have been shown that MDD occured more frequently with other medical ilnesses. The prevalence of MDD three-fold higher with type 1 diabetes and two-fold higher with type 2 diabetes compare to control ones [21]. Aproximately 20% of patients with heart disease are also have MDD [22].

Chemical imbalance of brain: Some brain regions are responsible to mood regulation. Researchers have been revealed that imbalance between neurotransmitters may contribute to developing MDD. Three neurotransmitter types which are serotonin, dopamine and norepinephrine are thought to be involved in MDD [23]. Studies have been found that suidical thought is associated with low serotonin metabolites levels [24]. Dopamine regulates positive feelings and lower dopamine may increased risk of MDD [25].

1.3.1.2 Treatment

There are various approaches for treating MDD depending on the severity of the disorder. The treatment model can be form one of following options or a combination of these;

Psychotherapy: It is also referred to as talking therapy that including cognitive-behavioral, interpersonal and family-focused therapies.

Pharmacotherapy: Many types of antidepressants are approved to treated MDD. Most of them are used for modulating brain chemistry. Selective serotonin reuptake inhibitors, monoamine oxidase inhibitors, serotonin and norepinephrine reuptake inhibitors, atypical and tricyclic antidepressants are major antidepressant types that increased the avaibility of neurotransmitters [26].

Electroconvulsive Therapy (ECT): It is a somatic therapy that is used when other treatments are not effective or acute phase of MDD such as suicidality. It is used for relieving depression by applying electrical current to the brain to improve function of neurotransmitters [27].

1.3.1.3 Diagnosis

Diagnostic evaluation has critical point to prepare a patient-centered and effective treatment model. Due to the most of symptoms of MDD are intangible, diagnosis is based on qualitative methods. The most widely used methods for depression diagnosis are patient self-report, clinical interview based on doctor's experience, psychiatric questionnaires with accepted diagnostic criteria such as DSM-5, Beck’s Depression Inventory (BDI) [28], or Hamilton Depression Rating Scale (HAM-D) [29]. These criteria based on investigation of patient behavior during a specific period by asking some questionnaires.

Diagnostic criteria for MDD are defined by DSM-5 that five or more of following nine symptoms must be present in a person during 2 weeks or more [12]:

1. Depressed mood nearly everyday

2. Remarkable decrease of interest and pleasure
3. Significantly change of body weight or appetite
4. Insomnia or hypersomnia
5. Disruption of psychomotor activity
6. Fatigue or loss of energy
7. Feelings of worthlessness or guilty
8. Inefficient concentration
9. Recurrent suicidal thought

Recently, there is no certain quantitative method to assessment of MDD. Due to utilizing qualitative methods, the accuracy of diagnosis often affected by various factors such as a complex spectrum of symptoms, doctor's proficiency, patient's assistance. Furthermore, these factors may prevent selecting appropriate treatment and reaching sufficient recovery time. If the diagnosis period does not proceed correctly the disorder may reach a severe and irreversible situation. Therefore, developing an quantitative and effective method for the detection of depression has great importance in terms of providing an early and accurate diagnosis, contribute to doctor's findings, reducing the recovery time.

Recent studies indicate that mental disorders are addressed as brain disorders associated with disruption of function, structure, or chemical balance of the neural networks [30]. Based on this, neuroimaging techniques have been gaining attention for quantitative diagnosis of depression due to analyzing the brain activity. As detailed explained below, EEG is more preferred than other neuroimaging techniques because of its noninvasive, easy to use, high temporal resolution, low cost. In addition to this, many studies have demonstrated that as discussed in 'Literatur Review' section EEG based CAD methods have achieved high success to discriminate MDD patients from HCs.

1.4 EEG

EEG is a type of electrophysiological technique that records the electrical activity of the brain. It provides to obtain information about different brain region's activity by using many electrodes on the scalp. Nerve cells generate electrical potential during communicating with each other and this potential is recorded by the EEG device as oscillations. Any disturbance of neural activity in the brain is expected to reflect on the bioelectrical activity of the brain. Abnormalities in the bioelectrical activity can be detected with the recordings that can be taken from the EEG device [31]. Compared to other techniques, EEG has some advantages like easy to use, high temporal resolution, low cost. It is widely preferred by doctors and scientists to understand the brain activity of a person or diagnose neurological diseases such as epilepsy [32], Alzheimer disease [33], Parkison disease [34], and cognitive disorders [35, 36].

The principle of EEG based on recording voltage differences between different parts of the brain by multiple metal electrodes attached to the scalp. The spontaneous and rhythmic EEG signal is recorded and it represents the electrical activity of many billions of cortical neurons which usually ranges from −100 to +100 μV [37]. The electrical activity is based on summation of postsynaptic potentials of cortical neurons. Postsynaptic potentials are the electrical potential that obtained during neurotransmitters are bind to the postsynaptic cell [38].

1.4.1 EEG Electrode Position

EEG is usually recorded by considering The International 10-20 system as recommended by the International Federation of Societies for EEG and Clinical Neurophysiology [39]. As shown in Figure 1.3, the system standardized location of electrodes on the surface of the scalp. Four landmarks are used to divide the head into proportional positions: nasion, inion, and two preauricular points.

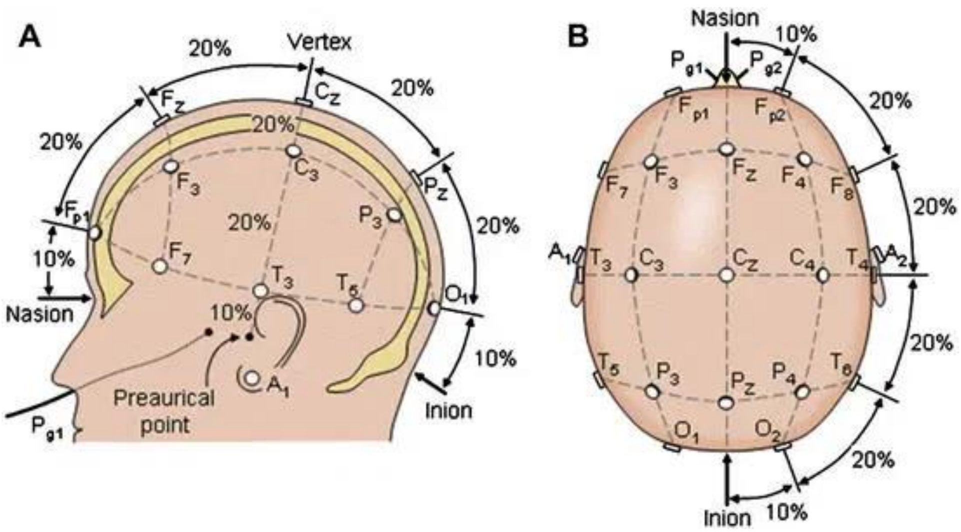

Figure 1.3: The International 10-20 system of electrode placement [40]

The "10" and "20" refer to the distance between neighbor electrodes either 10% or 20% intervals. Electrodes are labeled according to the relationship between their position and the brain regions as F: frontal, C: central, T: temporal, P: posterior, and O: occipital. After the brain region is labeled, the exact location of electrodes is indicated as numbers or 'z' which refers to the midline electrode position. Odd numbers represent left hemisphere of the brain, and even numbers represent right hemisphere of the brain.

1.4.2 Frequency Bands of EEG

The raw EEG signal is a complex and non-stationary signal that includes various frequency oscillations. Brain waves are mainly subdivided into 5 frequency bands: delta (<4 Hz), theta (4–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (>35 Hz) [41]. Their waveforms are shown in Figure 1.4. The frequency bands are related to different psychological behaviors [42]. Delta wave is the slowest wave and highest amplitude formation that associated with deep sleep. It occurs frequently infants up to one year. Theta waves are classified as slow activity that appears at very relaxed status. It is common wave in young children. The wave is associated with deep meditation and mental imagery.

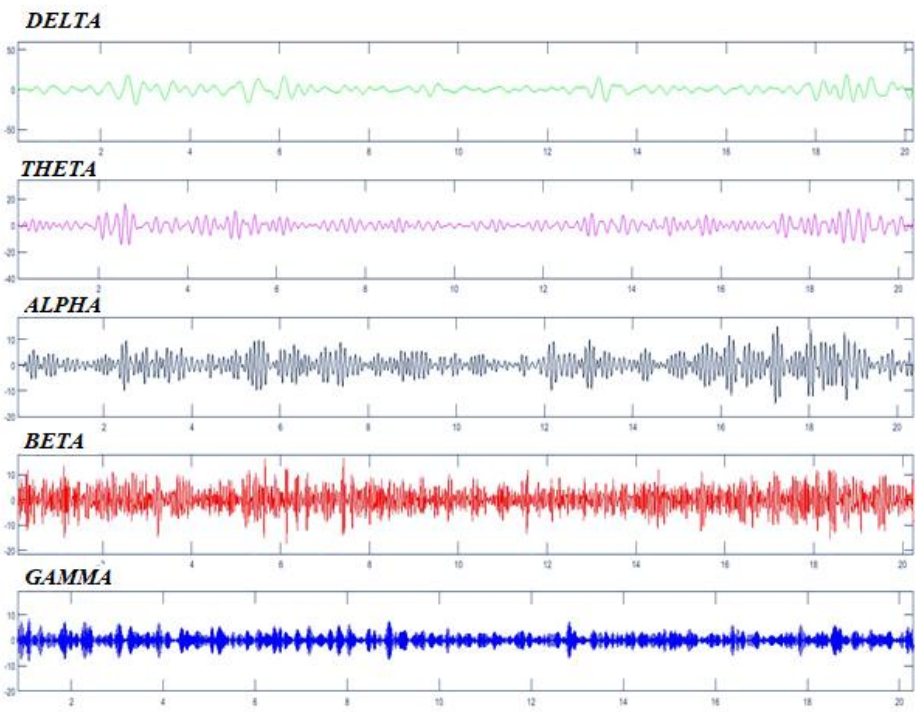

Figure 1.4: The frequency bands of an EEG signal

Alpha wave appears posterior regions of head and mainly present during eyes close and resting status. It is associated with mental activity and physical relaxation. Beta wave has higher frequency, lower amplitude formation and a symmetrical distribution of activity. It is dominant in unrest status such as concentration, logical thinking, alert positions. Gamma wave is fastest brain wave that related to higher mental activity as perception and consciousness [42].

1.5 Literature Review

Numerous studies have been indicated that electrophysiological differences between MDD patients and HCs can be detected by examining the EEG signals. Lee et al. [43] suggested that high-alpha power originating from the central areas of the left hemisphere can be used for differentiating depression patient from euthymic peers. Gollan et al. [44] focused on alpha EEG asymmetry differences between depressed and healthy persons and they reported that depressed patient has significantly higher alpha EEG asymmetry than healthy ones. In a study found that more alpha and distributed beta activity, and less distributed delta activity are observed in depression patients compared to non-depressed persons [45]. In addition to these studies, there are many studies that used a combination of EEG and CAD system to automatically classified depressed patients and non-depressed ones as listed in Table 1.1.

Cai et al. explored a depression detection method using prefrontal lobe channels as Fp1, Fp2, and Fpz [46]. In their study, resting-state EEG signals under sound stimulation were recorded from 92 depressed patients and 121 HCs. 270 features as composed of linear and nonlinear features were extracted from the EEG signals and minimal-redundancy-maximal-relevance technique was used for feature selection process. Four different classification algorithms as Support Vector Machine (SVM), *k*-Nearest Neighbor (*k*NN), Classification Trees (CT), and Artificial Neural Network (ANN) were applied in classification process. As a result of the classification, the highest success was achieved by *k*NN algorithm with accuracy of 79.27%. The study also suggested that it is a relationship between the power of theta wave and depression.

Table 1.1: List of recent studies related to detection of depression using EEG signals

Authours	Year	Database	Method	Classifier	Performance (%)		
					Accuracy	***Sensitivity***	***Specificity***
Zhang et al. **[47]**	2021	24 MDD 24 HC	Frequency Bands	Brain Functional Network	93.31	95.76	90.85
Liu et al. **[48]**	2020	20 MDD 19 HC	Phase-Lag Index Graph Theory Frequency Bands	SVM	89.7	89.4	89.9
Cai et al. **[49]**	2020	86 Depressed 92 HC	Linear, Nonlinear Features Feature-Level Fusion	*k*NN	86.98	–	–
Peng et al. **[50]**	2019	27 MDD 28 HC	Phase-Lag Index Frequency Bands	SVM	92.73	–	–
Aydemir et. al. **[51]**	2021	34 MDD 30 HC	Melamine Pattern	Weighted *k*NN	99.11	98.40	99.87
Li et al. **[52]**	2019	14 Depressed 14 HC	PSD Ensemble model	SVM	89.02	–	–
Sandheep et al. **[53]**	2019	30 Depressed 30 HC	–	CNN	99.31	–	–
Mao et al. **[54]**	2018	17 Depressed 17 HC	Non-distance Projection	CNN	77.20	–	–

Mahato et al. [55] developed a classification method to differentiate MDD patients by utilizing frequency band powers and theta asymmetry. They used 5 min eye close (EC) resting-state EEG signals recorded from 34 MDD patients and 20 HCs. Multi-Cluster Feature Selection technique was used for feature selection. SVM, Logistic Regression (LR), Naive-Bayesian (NB), and Decision Tree (DT) algorithms were used as classifiers in the study. It was found that accuracy of 88.33% was obtained in SVM with a combination of alpha power and theta asymmetry features.

Liao et al. [56] proposed a method based on EEG and kernel eigen-filter-bank common spatial pattern (KEFB-CSP) which is a feature extraction method. The method was testd using 12 MDD patients and 12 HCs with recording resting-state EEG. The study obtained accuracy of 81.23% in SVM by using features extracted from temporal lobe channels.

Mumtaz et al. [35] proposed a ML algorithm for differentiating MDD patients from HCs by analyzing resting-state EEG signals taken from 22 MDD patients and 30 HCs. EEG frequency bands and interhemispheric alpha asymmetry was used to feature extraction and LR, SVM, and NB algorithms were used for classification. Combination of frequency band powers and interhemispheric alpha asymmetry features reached the highest success, accuracy of 98.4 % by using SVM classifier.

Acharya et al. [57] developed a deep learning (DL) method to classified depressed patients and HCs by using Convolutional Neural Network (CNN) model. 15 depressed patient and 15 healhty controls were chosen for testing the model. Feature extraction and selection processes are eliminated with using the DL model. The algorithm achive highest success with using EEG signals obtained from the right hemisphere as accuracy of 93.5% and left hemisphere as accuracy of 96.0, respectively.

Uyulan et al. [58] built up a diagnosis model for MDD based on DL and EEG. 53 MDD patients and 43 HCs were used in the study. EEG signals were analyzed by different CNN structures as ResNet-50, Mobile-Net, and Inception-v3. EEG signals were converted image formation to using the CNN structures as an input. The study reached 92.66% accuracy with right hemisphere by using MobileNet.

1.6 Contribution of The Thesis

The aim of the thesis is to develop an quantitative MDD detection method based on a computer-based model by analyzing EEG signals. Multichannel EEG signals were recorded from voluntary subjects which are categorized as MDD patient and HC. Empirical Mode Decomposition (EMD) is introduced as a technique to decompose the signal without leaving the time domain. Moreover, a novel DL classification model is presented in the thesis with a new form of input that converts one-dimensional (1D) EEG signal into two-dimensional (2D) image formation. For this, three different methods are proposed in the thesis to differentiate MDD patients from HCs.

The first method is ML classification algorithm by analyzing multichannel EEG signals taken from different brain regions. In the method, various features are extracted from the EEG signals, and the features are classified by different ML algorithms.

EMD is proposed in the second method. EEG signals are decomposed into a finite number of oscillations called intrinsic mode functions (IMFs) for use as an EEG signal. The remaining process is similar to the first method with feature extraction and classification process. These are applied to the EEG signals and their selective IMFs.

The third method is based on the classification of topographic images separated from EEG signals using a DL model. Topographic images are obtained by positioned frequency band features considering EEG channels. The classification process is performed with given topographic images as input to the DL model.

The proposed methods were analyzed and evaluated by using MATLAB software (version 2018b and 2021b). The classification performances of the proposed models are compared with current state of art methods in this field.

1.7 Organization

This chapter describes the nervous system, mental disorders, MDD, related studies based on the relation between EEG and MDD, and the contribution of the thesis. The remaining thesis is organized as follows:

In Chapter 2, the datasets used in the thesis are explained. Filtering processes are also explained for applying to remove noise on EEG signals. The steps of the proposed three methods are described in detail, the EMD method is introduced, and theoretical information about the used features and classification algorithms is given in this section.

In Chapter 3 represented as the Result section, the classification performances of the three proposed methods are presented.

In Chapter 4, the results of the proposed methods are compared to recent studies, and the positive aspects, limitations of the proposed methods are discussed. Besides, the conclusion of the thesis is presented in this section.

Chapter 2

Material and Method

In this section, the proposed methods for analyzing EEG signals and the database which are used in these methods are explained. The schematic diagram as demonstrated in Figure 2.1 briefly decribes the three methods.

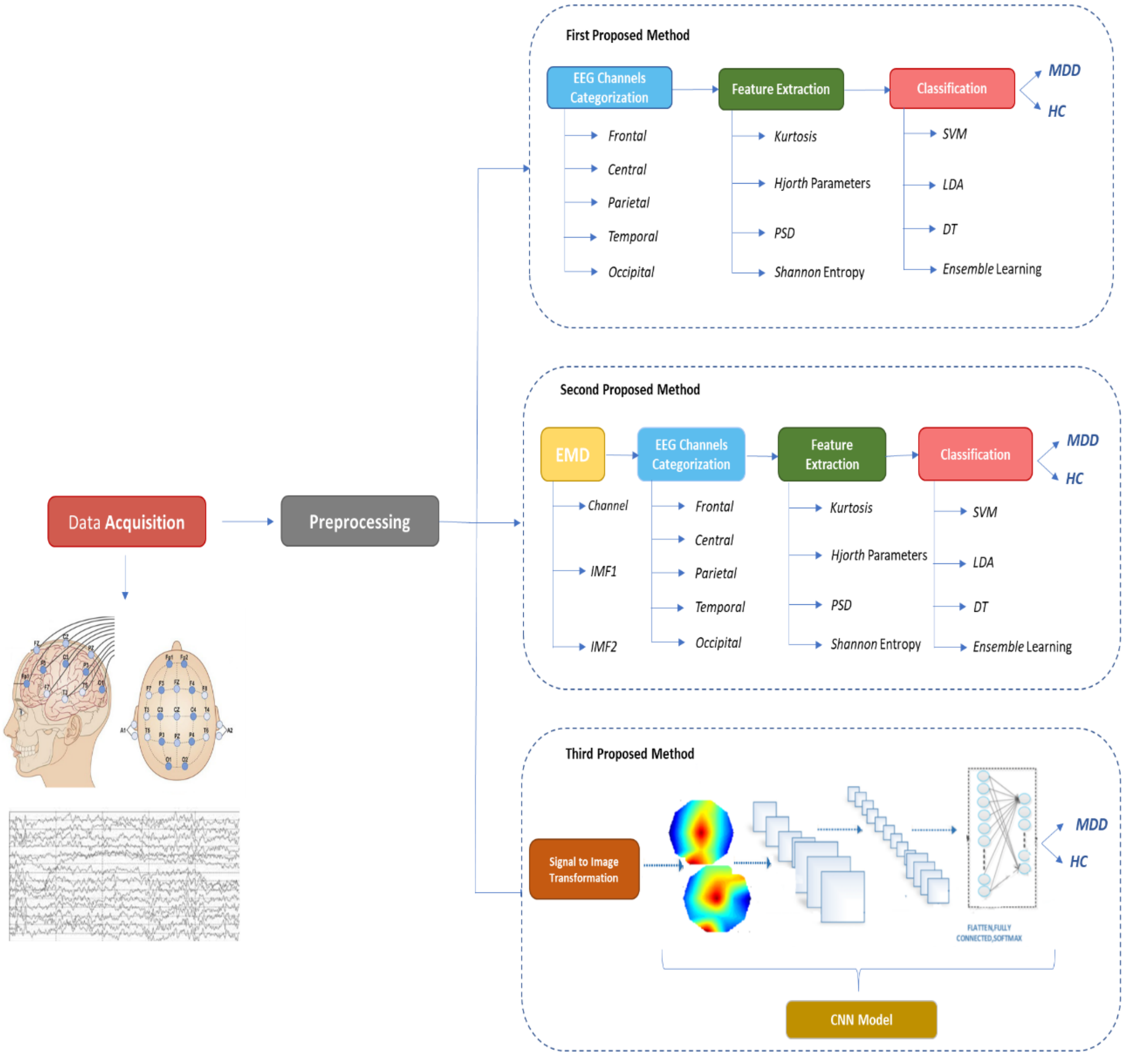

Figure 2.1: Schematic diagram of the three proposed methods in this thesis

2.1 Data Acquisition and Preprocessing

Multichannel EEG signals recorded from 16 MDD patients and 16 HCs were used in the study. The signals were obtained from the outpatient clinic database of Kemal Arikan Psychiatry Center. Voluntary subjects were selected according to the diagnosis of DSM-5 criteria and doctor's assessment. The experimental data were recorded in a quiet dim room where has mild temperature, and humidity environment. During the experiment period, the subjects sat comfortably on a chair with relax, awake, eyes closed position. Simultaneously, the lowest head, body, and eye movements were required for avoiding noise of body movement. The resting-state EEG signals were recorded for approximately 7 minutes by using the Neuron-Spectrum-4/P device with 19 electrodes (Fp1, Fp2, F3, F4, Fz, C3, C4, Cz, P3, P4, Pz, O1, O2, F7, F8, T3, T4, T5, T6) arranged based on the international 10–20 system as seen in Figure 2.2. Linked mastoid electrodes (A1-A2) were chosen as reference electrodes during data collection and the sampling rate is 500 samples per second.

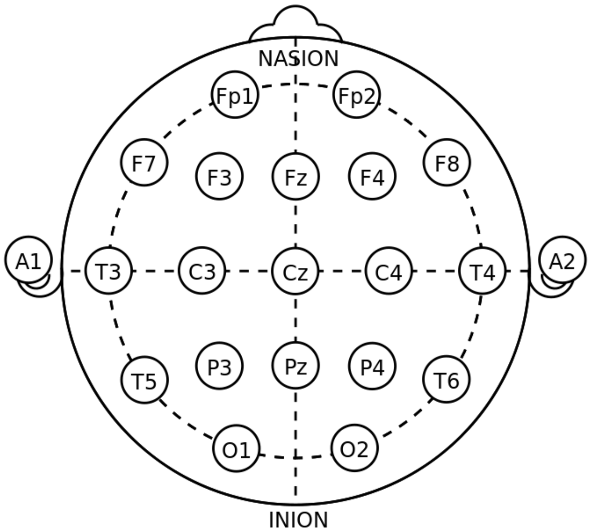

Figure 2.2: 19 electrodes position according to international 10–20 system

In the study, all the signal processing has been done using MATLAB software (version 2018b and 2021b). When an EEG signal is recorded it is contaminated by different noise types or artifacts. Eye blinks, muscular activity, the voluntary or involuntary movement of the patient, electrical power lines are types of noises that may change the characteristics of the EEG signal. Consequently, it is difficult to extract effective features from the signal. These noise types may disrupt the EEG signal, and make the feature extraction and classification steps less accurate. In order to obtain proper signal processing and classification, it is necessary to preprocess the EEG signal which includes removing noises from the EEG signals. Muscle artifacts were eliminated from raw EEG data manually for each subject by the responsible technician. The filtering process is mainly used for preprocessing of the signal. The raw EEG data is passed through a band-pass filter and notch filter to get the clean EEG data. Band-pass filtered with cutoff frequencies at 0.15–70 Hz was applied to eliminate the artifacts from raw EEG data. Power line interference is eliminated with an applied Notch filter of 50 Hz.

In order to standardize all of the EEG recording values, the Z-Score normalization technique [59] was applied to the signals. This method is among the statistical data transformation techniques and uses the mean and standard deviation values of the data. The purpose of the normalization process is to transform the data into normalized values into appropriate intervals, to enable faster processing and to obtain easily interpretable results.

2.2 Signal Processing

EEG channels contain different information about brain activity due to matching different regions of the brain. Based on this, the multichannel EEG signals were analyzed and classified considering the brain region. Subgroups were created which are frontal, central, temporal, parietal, and occipital by categorized the EEG channels according to the brain region. The subgroups and their representative channels that are used in the study are listed in Table 2.1. 6 min EEG signals were used in this study. As ML algorithms a large number of datasets are required, each selected channel in the EEG signal divided into epochs with durations of 30 seconds, and 12 epochs were obtained from a channel that each of them has 15000 samples. In this way, a large number of samples were obtained to avoid underfitting.

Table 2.1: The brain region subgroups and EEG signal channels used in the study

Brain Region	Channels	Total Number of Dataset (Person x Epoch x Channel)
Frontal	Fp1, Fp2, F3, F4, Fz	$32\ x12x5 = 1920$
Central	C3, C4	$32\ x12x2 = 768$
Temporal	T3, T4, T5, T6	$32\ x12x2 = 768$
Parietal	P3, P4	$32\ x12x4 = 1536$
Occipital	O1, O2	$32\ x12x2 = 768$

The number of the database for each subgroup are given in Table 2.1. Due to selected channel numbers are not equal for subgroups, the number of datasets shows a slight difference. Feature vectors were obtained from each subgroup in the feature extraction process and these feature vectors were classified with various classification algorithms. Besides this, EMD based method was developed to differentiate MDD patients from HCs. Each EEG channel was decomposed into IMFs by using the EMD. Feature vectors were created by the combination of signal channels and their selective IMFs and then the classification process was applied. Classification performances of the developed methods were compared with evaluating performance measure parameters.

2.2.1 Empirical Mode Decomposition

EEG signal is a non-linear and non-stationary signal that its statistical parameters such as mean, variance, amplitude are change with respect to time. Various methods are developed to process, analyze and, evaluate EEG signals depending on the time domain, frequency domain, and time-frequency domain. EMD is a data-driven signal processing method proposed by Huang et al. [60] for analyzing non-linear and/or non-stationary signals. It decomposes a signal without leaving the time domain and no need any information about the stationarity and linearity of the signal. It differs from other methods as Fourier Transform, Wavelet Transform with this property. EMD method composed of decomposition of any complex signal into a finite and small number of oscillations named IMFs which are elementary AM-FM-type components.

The decomposition is based on local characteristics of the complex signal and the IMFs are obtained by an iterative procedure called a sifting process. An EEG signal and its IMFs are demonstrated in Figure 2.3. Each IMFs should satisfy two conditions;

- In the signal, the number of extrema and zero crossings must be equal to each other or different by one at most,
- The mean value of the upper envelope defined by local minima and the lower envelope defined by the local maxima should be zero

The process of the EMD method to decompose a signal into its IMFs is explained as below:

1) Find the local maxima and local minima of the original signal $x(t)$.

2) Use cubic spline interpolation and generate the upper and lower envelopes $e_{min}(t), e_{max}(t)$, respectively.

3) Calculate the average of the upper and lower envelopes.

$$m(t) = \frac{e_{min}(t) + e_{max}(t)}{2} \tag{2.1}$$

4) Subtract the average number from the original signal

$$c_1(\mathrm{t}) = x(t) - m(t) \tag{2.2}$$

5) Check the result $c_1(\mathrm{t})$ to satisfied IMF conditions. IMF1 is obtained if the $c_1(\mathrm{t})$ provides these conditions. Else, repeat the procedure from step 1.

6) Subtract IMF1 from the original signal and find residual.

$$r_1(\mathrm{t}) = x(t) - c_1(\mathrm{t}) \tag{2.3}$$

7) Then use the residual as the original signal and repeat the process from step 1 to step 5 to produce next IMF.

This procedure is repeated for created other IMFs until the final residue signal is a monotonic function and does not sufficient to produce a new IMF. At the end of the decomposition, IMFs and final residual are obtained. The original signal can be represented as the summation of IMFs and final residue. This procedure is symbolized as;

$$x(t) = \sum_{i=1}^{n} c_i(\mathrm{t}) + \mathrm{r}(t) \tag{2.4}$$

where $c_i(\mathrm{t})$ is number of IMFs and $\mathrm{r}(t)$ is the final residual.

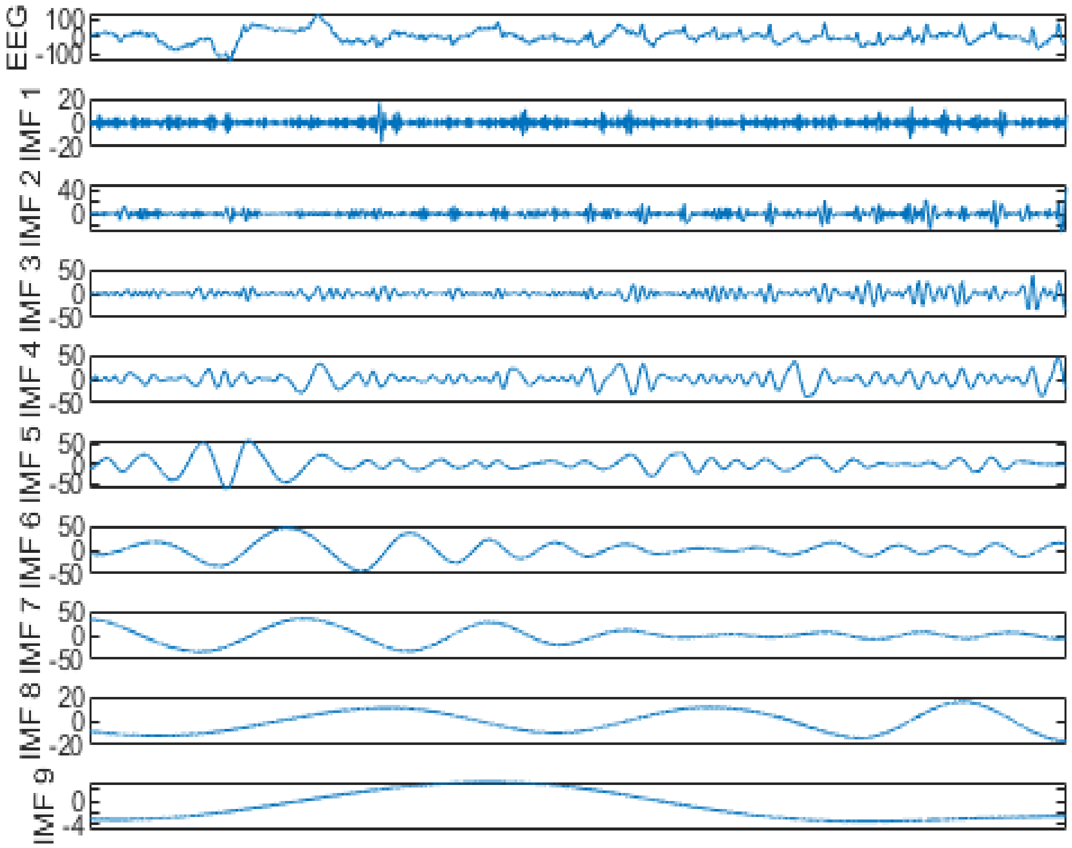

Figure 2.3: An EEG signal and its IMFs

The EMD method decomposes signals into narrow-band components with decreasing frequency. Therefore lower-order IMFs carry high-frequency components of the original signal and the order of the IMFs increases, frequency decreases. IMF1and IMF2 have the highest frequencies between the IMFs. The EEG signals and their first two IMFs were used for extracting certain features as explained next section.

2.2.2 Feature Extraction

Feature extraction is a process that has importance for the recognition, classification, or detection of a large set of data. In the process, new data matrices are created by extracting features from the original dataset. Significant, informative, and independent features should be selected to obtain less computational time and high success for the classification process [61]. As mention 'Literature Review' section, the depressed EEG signal demonstrate different characteristic property in various parameters like power, complexity, frequency, etc. compare to healthy ones. Considering the differences, the following features were selected in this study and this section explains the details of the features. The selected features can be categorized as below according to their domains.

• *Time domain features*: Kurtosis, Hjorth Parameters (mobility, and complexity)

• *Frequency domain features*: Power Spectral Density of different frequency bands

• *Non-linear features*: Shannon Entropy

2.2.2.1 Kurtosis

Kurtosis (β) is a statistical parameter that characterized the peaked distribution of random variables by measures the peakedness of the time-series signal [62]. It is the fourth-order moment that the value of kurtosis gives information about peak distribution as sharp, flat, or Gaussian distribution. A higher kurtosis value ((β >3) indicates that a signal has a sharper peak distribution. Signal with lower kurtosis value (β <3) has relatively flat peaks. If the kurtosis value equal to three in a signal, its distribution is represented as Gaussian distribution. With respect to this, the kurtosis feature may include useful information about the complexity of EEG signals. Kurtosis is calculated with following formula;

$$\beta = \frac{\sum(x - \breve{x})^4}{N(s)^4} \tag{2.5}$$

where x represent peaks in the data over the interval of time, $\breve{x}$ is mean value, N is size of the data and s is standard deviation.

2.2.2.2 Hjorth Parameters

Hjorth parameters composed of three statistical parameters which are activity, mobility, and complexity introduced by Hjorth in 1970 in EEG signal analysis [63]. They used for indicating characteristics of a signal in the time domain and widely preferred to analyze the biological signals [64]. Besides these parameters are calculated in the time domain, they also include information about the frequency spectrum [65]. Activity refers to the variance of signal amplitude and gives information about signal power. Mobility is the square root of the function which is the variance of the first derivative of the signal divided by variance of the signal. It represents the mean frequency of the signal. Complexity indicates the change in frequency and defined as the ratio between the mobility of the derivative of the signal and mobility of the signal amplitude. The equations of all parameters are given as follows;

$$\text{Activity} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \check{x})^2 \tag{2.6}$$

where x_n represent the signal and $\check{x}$ is mean of the signal and N is length of the signal.

$$\text{Mobility} = \sqrt{\frac{Activity(\frac{dx}{dt})}{Activity(x)}} \tag{2.7}$$

$$\text{Complexity} = \frac{Mobility(\frac{dx}{dt})}{Mobility(x)} \tag{2.8}$$

In this study, two Hjorth parameters which are mobility and complexity are used as features and extracted from EEG signal channels and their first two IMFs.

2.2.2.3 Shannon Entropy

Shannon Entropy is firstly defined by Shannon in 1948 and it measures the uncertainty of a random signal [66]. Shannon entropy is calculated to find the predictability of future amplitudes of a signal by determining the probability density function of the signal. The higher value of Shannon entropy indicates more uncertainty and randomness [49].

In this study, Shannon entropy was chosen as a feature because it can be used to distinguish the different distributions of EEG signals. Shannon entropy is calculated as follows:

$$\mathrm{SE} = -\sum_{i=1}^{n} p_i - \log p_i \qquad (2.9)$$

where i is the amplitude range of the signal and p_i is the probability that the signal is at the given amplitude.

2.2.2.4 Power Spectral Density

Power spectral density (PSD) is a type of frequency domain analysis of a signal which is defined as the signal power distribution over the frequency. The method preferred to characterize the signal which has broad bandwidth. It gives information about different frequency ranges related to variations of power distribution. From this, the idea is performed about which frequency range is effective for selecting as a feature. This method can be obtained by using Fast Fourier Transform (FFT) of the autoregressive model of the signal.

In this study, each channel and their IMFs are divided into frequency bands by filtering method and total power of delta, theta, alpha bands were calculated to used as features. FFT and PSD of a signal are evaluated as follows [46];

The FFT of a signal is calculated following equation;

$$X(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)e^{-j(\frac{2k\pi n}{N})} \qquad k = 0,1,\dots,N-1. \qquad (2.10)$$

The PSD of a signal is calculated by using FFT of the signal as follows,

$$s(k) = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^2 \qquad (2.11)$$

2.2.2.5 Feature Matrices

In the study, multichannel EEG signals were analyzed taken from 16 MDD patients and 16 HCs. All EEG signals divided into epochs which are 12 epochs per channel for increasing dataset. The detailed explanation of matrices is listed in Table 2.2.The EEG signals were classified separately according to the brain region subgroup.All subgroups were separated as training and test set for classification process which is 80% training set and 20% test set described as Figure 2.4. 7 features which are kurtosis, Hjorth parameters (mobility, complexity), power of delta, theta and alpha frequency bands and Shannon entropy were extracted from both training and test set and feature matrices were obtained.

Table 2.2: The number of database and feature matrices used in the study

Brain Region	Total Number of Dataset (Person x Epoch x Channel)	Feature Matrix (Dataset x Feature)
Frontal	$32\ x12x5 = 1920$	Train $1536x7$ Test $384x7$
Central	$32\ x12x2 = 768$	Train $614x7$ Test $154x7$
Parietal	$32\ x12x2 = 768$	Train $614x7$ Test $154x7$
Temporal	$32\ x12x4 = 1536$	Train $1228x7$ Test $308x7$
Occipital	$32\ x12x2 = 768$	Train $614x7$ Test $154x7$

Besides these, each EEG channel was decomposed into IMFs by using the EMD as mentioned before. Feature matrices were obtained by the combination of features and the signal channels and their first two IMFs. So, the total number of datasets per subgroup is rearranged by 3 x the number of the total dataset of a subgroup given in Table 2.2. The EEG signals divided into subgroup brain regions and the selective channels were decomposed into IMFs. The dataset obtained by the channels and their first two IMFs were separated training and test set as same procedure of without EMD process. 7 features were extracted from the dataset of each subgroup.

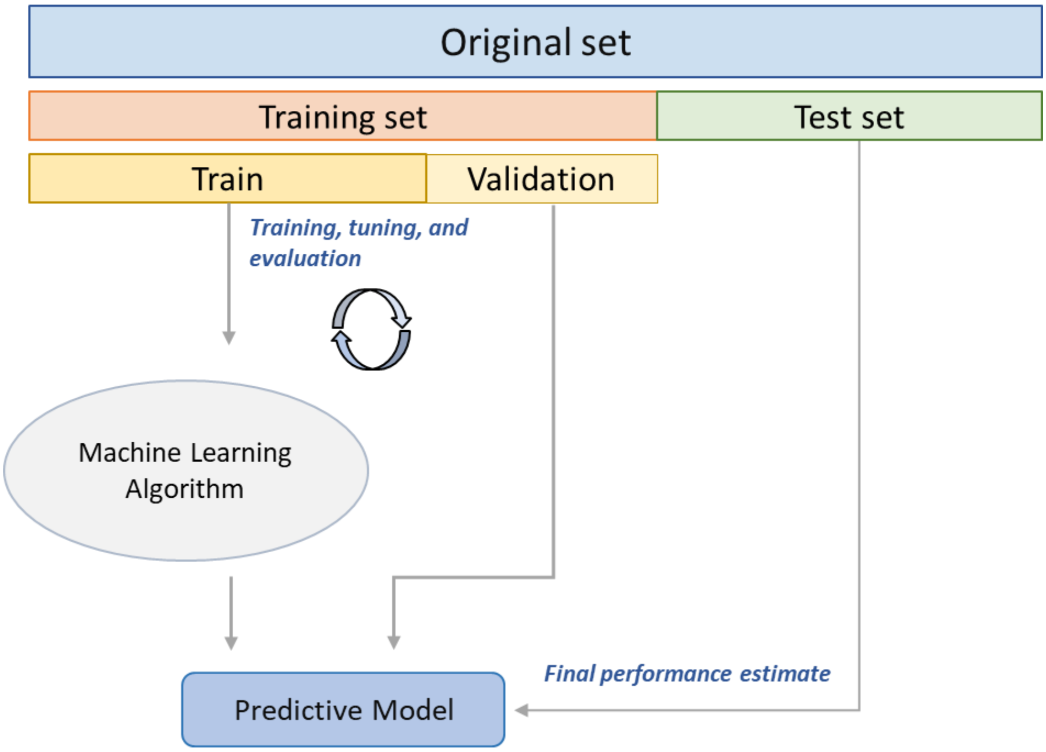

Figure 2.4: Shematic diagram of training and test set from original dataset

Each row of the feature matrix in the training set was labeled 0 or 1, depending on the group it represents. 0 represents HC and 1 represents MDD patient.

2.2.3 Signal to Image Transformation

EEG contains multiple time series corresponding to measurements at different spatial locations on the cortex. Conventional EEG analysis approaches generally based on the representation of information in time and frequency domains while ignoring information about the location of cortical regions in the spatial domain. Applying topographic mapping to an EEG signal provides to transform the information obtained from various scalp electrodes into 2D images. EEG signals could be analyzed in the DL approach by 2D topographic images to be used as inputs [67]. In this way, the spatial, spectral, and temporal qualities and structure of the information are preserved. In order to analyze the input data in the DL model, 2D topographic images should be generated from the EEG time series data.

EEG electrodes are positioned on the scalp in a 3D space. To determine 2D EEG electrode locations, 2D projection of the the locations from 3D space is required. Together with this, such a projection must also protect the relative distance between adjacent electrodes. For this purpose, 3D space was transformed into a 2D space which is referred to as an image plane. The projection based on the transforms of the scalp 3D cartesian coordinate (x, y, z) into 2D cartesian coordinate (y-z) on the image plane. Inspired by the study [68], 2D cartesian coordinates were found by assuming the endpoints of each electrode spreads on a circle in the image plane.

After localized the points of the electrodes, feature mapping was created by extracted a feature from all EEG signal channels and assigning the feature value to the represented electrode location on the image plane. All EEG signal channels were segmented into 5 s epochs and each epoch divided into frequency bands total power of delta, theta, alpha bands were calculated to use as a feature. Frequency band power was selected as a feature for creating topographic images due to having significant differences between MDD patients and HC as mention in previous studies [35]. Delta, theta, and alpha powers were assigned separately to the represented 2D electrode location and so the frequency band power of each channel is mapped onto the 2D image plane. A natural 2D neighbor interpolation method was applied between the electrode location to produce a smooth 2D colorful image (RGB color representation) of the topographic images [68]. The generated 2D topographic image has represented the distribution of the frequency band power onto the whole cortex. Figure 2.5 shows topographic images of MDD patients and HC generated from theta power feature. When the figure is examined, it can be observed that differences between the power distribution of MDD patients and HC.

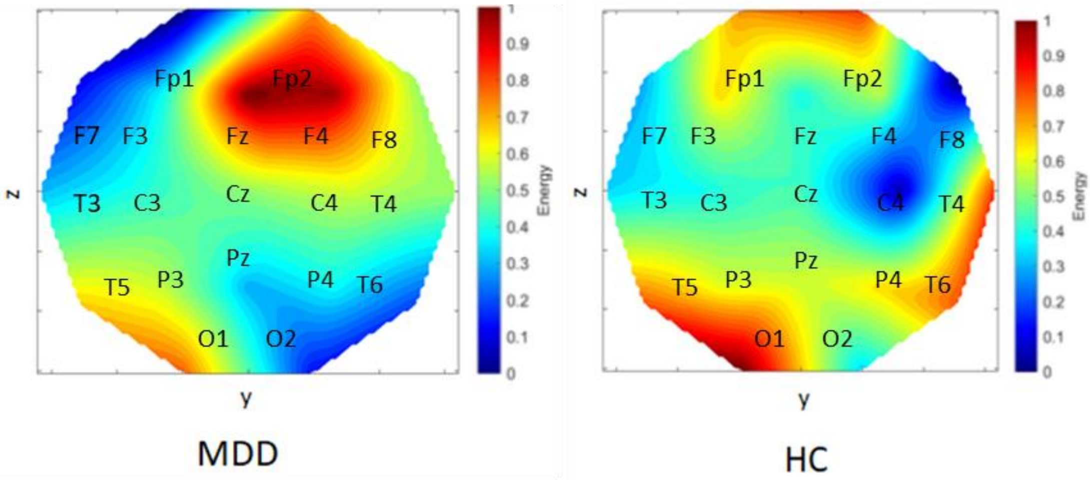

Figure 2.5: Topographic images generated from EEG signals

With the proposed 2D image transformation, all EEG channel information is preserved into a single image. Three topographic image types obtained from delta, theta, and alpha powers were used as input in the DL model.

2.2.4 Classification

Classification is a common data mining process that is the prediction of the target class to which a pattern belongs, by evaluating the characteristics of that pattern. The purpose of the process is to create a decision mechanism and match the patterns whose class information is unknown with the closest classes with the least error. [61] Different ML algorithms have been developed to precisely predict the target class. These algorithms are divided into two groups based on the application learn the algorithm for make predictions namely supervised and unsupervised learning.

Supervised learning is a technique that is based on the data whose target classes are known [69]. For this technique, a function that matches between inputs and expected outputs (labeled data) is produced which is represented as training data. The technique learns the information given from the training data and makes a decision by evaluating the new information accordingly.

In the unsupervised learning method, there is no labeled information before. This method is using a function to predict an unknown structure through unlabelled data [70]. In this method, it is unclear to which class the input data belongs. It works by establishing a connection between previously untrained and unlabeled data and classifying data that are close to each other. It enables inferences about the data by using the distances of the data samples from each other and/or the neighborhood relations.

In this study, four different ML algorithms were used to classified the feature vectors which are explained in ‘Feature Extraction’ section. These algorithms are SVM, DT, Linear Discriminant Analysis (LDA), Ensemble Learning that are types of supervised learning. In addition, a DL model was also developed in this study based on 2D topographical image classification to differentiate MDD patients from HCs. Detailed information about the classifiers is given below.

2.2.4.1 Support Vector Machine

Support Vector Machines (SVM) is a ML algorithm proposed by Vapnik et al. that based on statistical learning theory [71]. The basis of the SVM algorithm is the process of classifying data by separating it with a plane or hyperplane. The working principle is finding the most appropriate decision function with determining the most proper hyperplane that can separate the data belonging to the two classes. The mathematical algorithm of SVM was originally designed for the linear classification of two-class data. Then generalized for the classification of nonlinearly separable or multi-class data and it has been widely used in the solution of these problems [72].

There can be many hyperplanes separating linear data belonging to different classes from each other as shown in Figure 2.6. SVM aims to find the optimal hyperplane that maximizes the distance between support vectors belonging to different classes demonstrated in Figure 2.6. Support vectors are represented as examples of both classes nearest to the separation hyperplane. These support vectors define the boundary of the class to which they belong and are located on a plane parallel to the separation hyperplane.

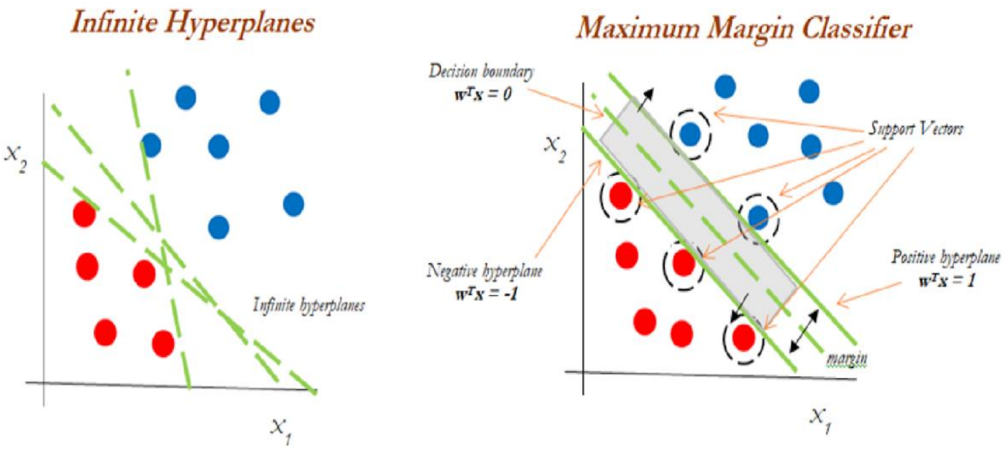

Figure 2.6: The hyperplanes that separated two classes

In the two-class linear classification process, the equations for the optimum hyperplane are as follows;

$$wx_i + b \geq +1 \quad for\ y = +1 \tag{2.12}$$

$$wx_i + b \leq +1 \quad for\ y = -1 \tag{2.13}$$

where x_i is each point of data, $y \in \{-1, +1\}$ is class labels or outputs, w is the weight vector (normal of the hyperplane) and b is the constant value [72].

In the classification of nonlinearly separable data, SVM carries the original data to a higher dimensional space with a mapping method, where it tries to find the linear optimum hyperplane that can classify the data.

2.2.4.2 Decision Tree

DTs are one of the tree-based supervised learning algorithms used in classification and regression problems. It can be used on complex datasets. DT is a recursive process and a flowchart like tree structure is used in the algorithm [73]. For classification, decision tree structure aims to create a model that predicts classes of input variables by learning decision mechanism inferred from the data feature. A tree structure is created by starting with a single node and branching out to new results demonstrated in Figure 2.7. The first nodes of decision trees are called root nodes. Internal nodes are found underneath of root nodes. The complexity of the model increases as the number of internal nodes increases. The leaf nodes at the bottom of the decision tree. Each interval node represents a test on a variable, each branch is a result of the test and each leaf node represents a class label.

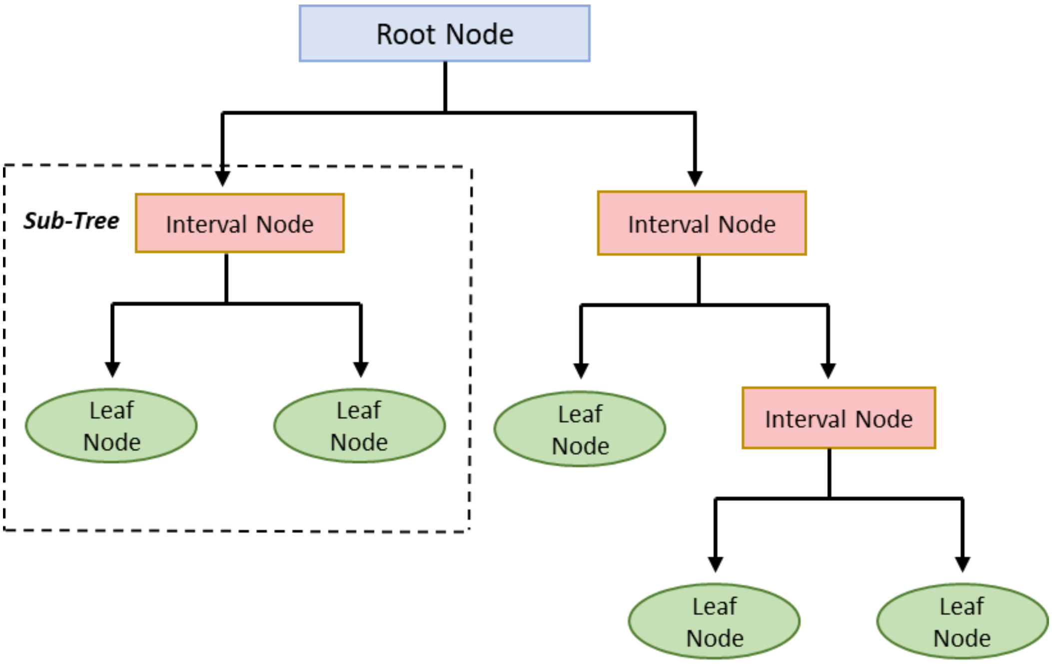

Figure 2.7: Flowchart of a DT structure

The method of splitting in DT algorithms is a factor affecting the tree's accuracy. DTs use multiple methods to decide to split a node into two or more subnodes. Algorithm selection is based on the type of target variable. The most commonly used methods in decision trees are Entropy, information gain, Gini index, Gain Ratio. In this study, Gini index [74] was used as split criterion and its equation is given as below;

$$G = 1 - \sum_{i=1}^{C} (p_{i,k})^2 \tag{2.14}$$

where C is the list of classes, p_i is probability of class i at node k.

2.2.4.3 Linear Discriminant Analysis

LDA is one of the most widely used classification methods developed by Fisher in 1936. This method generates a new variable with a combination of existing data [75]. It aims to maximize the differences between two classes according to the new variable. For maximizing of separability of the classes, the data is projected onto a lower dimension. In LDA, discriminant function is calculated to determine the differences between the classes. Based on this function, the class of unlabeled test data is estimated. The discriminant function is represented as given equation;

$$g(x) = w^t x + w_0 \tag{2.15}$$

where w is weight vector, x is input variable and w_0 is bias.

2.2.4.4 Ensemble Learning

Ensemble Learning is a paradigm that consists of combining multiple ML algorithms to improve the classification performance of a model. In this technique, the ML algorithms are trained to classify input data and the test samples are assigned to the class after combining each output of single learning algorithms based on some criteria [76]. The final predictions are made according to different strategies like combined learning algorithms considering statistical parameters or learning which algorithms are more accurate to predict. In general, better classification performance is obtained by using Ensemble Learning compared to the single classifier. Boosting, bagging and stacking are main types of Ensemble Learning. Shematic diagram of an Ensemble Learning model is given in Figure 2.8.

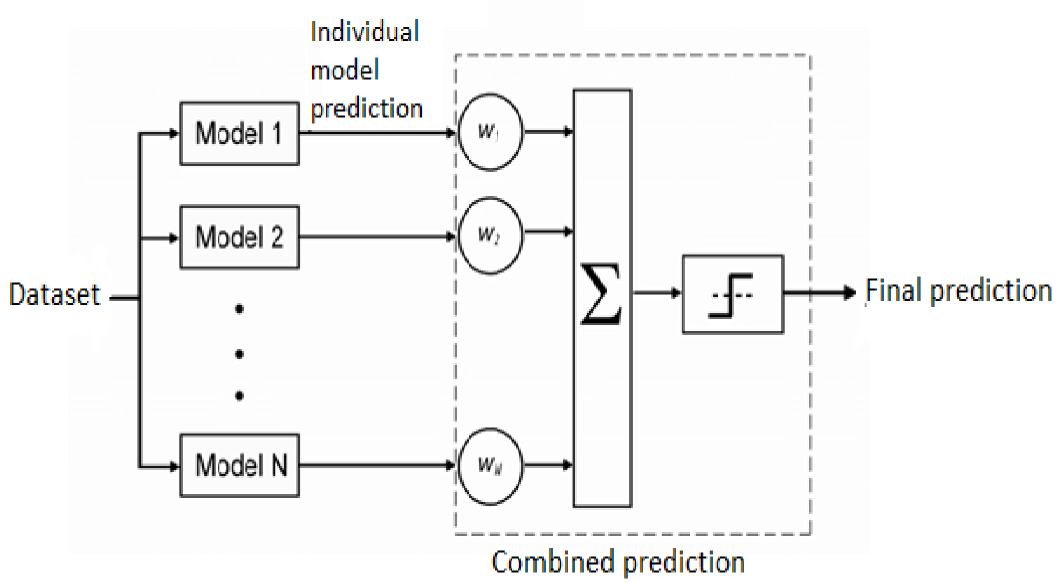

Figure 2.8: Framework of an Ensemble Learning model

2.2.4.5 Convolution Neural Network

DL is a ML technique derived from artificial neural networks that automatically extract features from raw data and classified them. In most cases, unlike conventional ML techniques, it does not require data preprocessing. The basic structure of DL is inspired by the working mechanism of the human brain, which shows a multi-layered learning structure. In recent years, one of the most successful DL techniques in pattern recognition is CNN. CNN is frequently used in the analysis of 2D data due to its structure [77]. It is used in many fields such as computer vision, speech recognition, object recognition and segmentation, image processing as well as in the classification of biological signals [68]. 1D biological signals like EEG can be transformed into two-2D images using different techniques and classified in CNN structure [31].

Unlike conventional ML techniques, the CNN structure consists of an input layer, three or more hidden layers, and an output layer [77]. The input layer collects data that is fed into the model. Hidden layers learn and save the features of the data from the raw data without any preprocessing [78]. These features are then sent to the output layer and assigned to target classes as model outputs. Learning is a repetitive and periodic process based on detecting implicit and complex information on data and adjusting the weights of the connections between artificial neurons embedded in the network. Artificial neurons are provided to detect features from the input data and to transmit the features to other neurons in the network [77]. An artificial neuron structure is demonstrated in Figure 2.9.

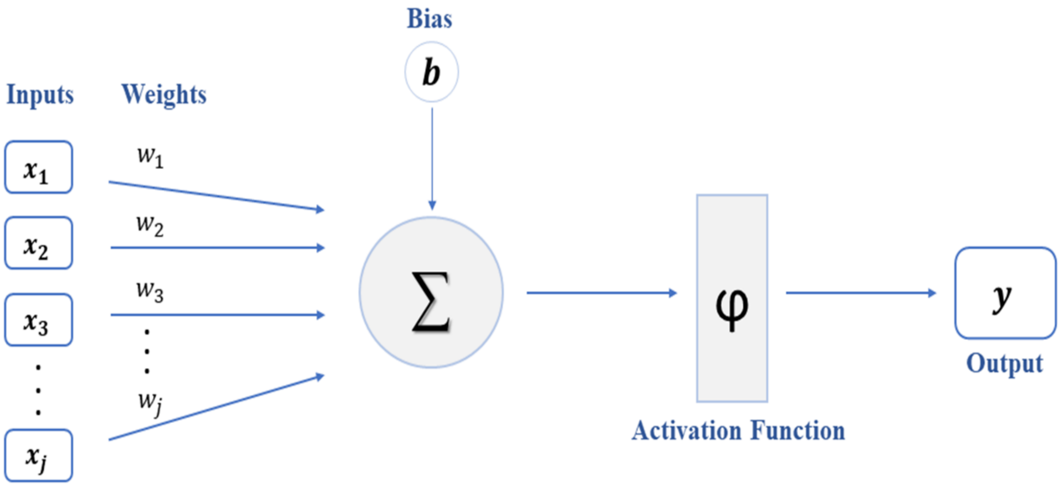

Figure 2.9: Shematic diagram of an artificial neuron structure

The equation of an artificial neuron is given as;

$$y_j = f_j(x) = \varphi(w_j x, b) \tag{2.16}$$

where x is input, w_j is weight, φ is activation function and b is bias.

CNN has three types of layers: convolutional layer, pooling layer, and fully-connected layer. A schematic diagram of a CNN model is demonstrated in Figure 2.10. The convolution layer is the main structure of CNN. It is responsible for recognizing the features from the input image [77]. In this layer, input image is convolved with a specific kernel.

$$M * N(i,j) = \sum_{k,l} M(k,l)N(i+k, j+l)$$

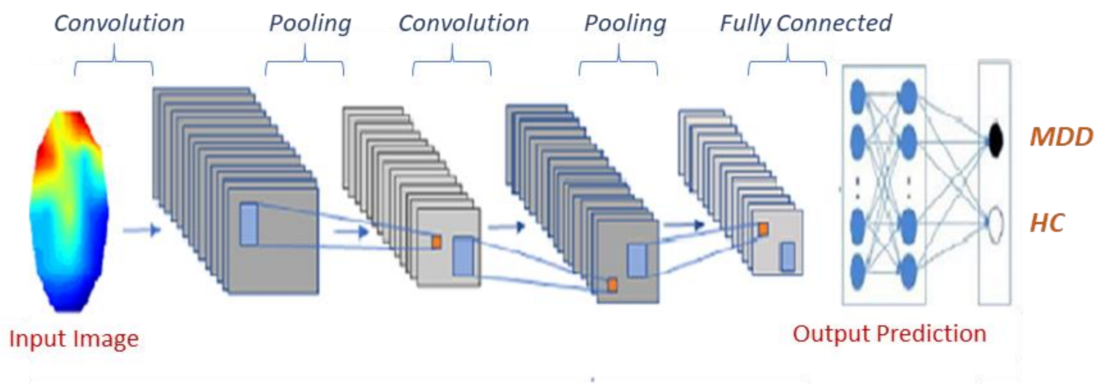

Figure 2.10: A schematic diagram of a CNN model.

The equation of convolutional function for 2D inputs is defined as;

$$M * N(i,j) = \sum_{k,l} M(k,l)N(i+k, j+l) \tag{2.17}$$

where M is convolutional kernel, N is 2D input.

A kernel smaller than the original image size moves over the image and finds certain features from the images by creating a feature map as shown in Figure 2.11. There are three structures in the convolution layer: kernel size, kernel number, and stride. Kernel size determines the output of each neuron in the convolution layer. Kernel sizes can be different such as 2x2, 3x3, 5x5

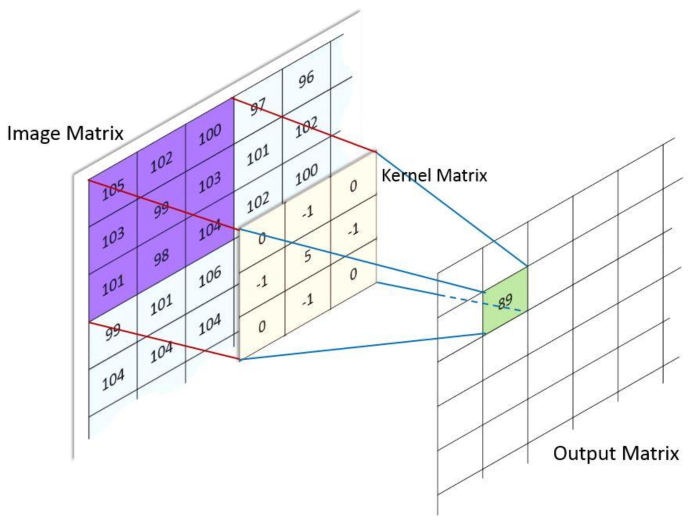

Figure 2.11 An example of image kernel and convolution

. Kernel values represent the weights of the model. When the kernel is moved over the input image, it creates a feature map. Stride value is a value that can be changed as a parameter in the CNN model it determines how many pixels the filter will slide over the input image. Filtering images result in a smaller output size than the original size. To prevent this, padding is applied.

The pooling layer reduces the spatial size of the output of the convolution layer to reduce the number of parameters and calculations in the network [78]. It also controls overfitting. The pooling layer is usually used between two convolution layers. There are two popular pooling methods as maximum pooling and average pooling. For a given input window, the average pooling takes the average values in the window while the max-pooling takes the maximum values within the window.

Convolution and poolling layers produce rectangular shaped outputs. This matrix outputs are converted into vector matrix in flattened layer. Fully connected layer is the last layer of CNN. In fully connected layer, the output data are taken from flattened layer. The outputs of the last layer are classified through the softmax function over the fully connected layer with a certain number of neurons. Some popular CNN architectures used today are AlexNet, VGGNet, GoogleNet, ResNet, MobileNet.

In this study, MobileNet architecture was used to classified topographic images obtained from EEG signals of MDD patients and HCs. MobileNet totally consists of 154 layers but in the study, dropout was added before the fully connected layer to prevent overfitting. The first and last 10 layers of the MobileNet arthitecture are listed in Table 2.3. Average-pooling is applied as the pooling layer. To define the output of the kernel in the convolution layer, the activation function is implemented. The rectified linear unit (ReLU) is applied in the architecture as an activation function that eliminates the negative values of the output kernel.

All topographic images represented as power of frequency band features were classified separately. The input images for each band were divided into training and test data, 70%, 30% respectively. These images are randomly split as training or test data. Adam optimizer was used for optimization and the learning rate is chosen as 0.001. Standard backpropagation with different batch sizes was applied for stochastic learning.

Table 2.3: The first and last 10 layers of the CNN arthitecture used in the study

Layers	**Type**	**Input Size**
Layer 1	Image Input	$32x32x3$
Layer 2	Convolution	$16x16x32$
Layer 3	Batch Normalization	$16x16x32$
Layer 4	ReLU	$16x16x32$
Layer 5	Convolution	$16x16x32$
Layer 6	Batch Normalization	$16x16x32$
Layer 7	ReLU	$16x16x32$
Layer 8	Convolution	$16x16x16$
Layer 9	Batch Normalization	$16x16x16$
Layer 10	Convolution	$16x16x96$
Layer 145	ReLU	$1x1x960$
Layer 146	Convolution	$1x1x320$
Layer 147	Batch Normalization	$1x1x320$
Layer 148	Convolution	$1x1x1280$
Layer 149	Batch Normalization	$2x2x1280$
Layer 150	ReLU	$2x2x1280$
Layer 151	Avarage Pooling	$1x1x1280$
Layer 152	Dropout	$1x1x1280$
Layer 153	Fully Connected	$1x1x2$
Layer 154	Softmax	$1x1x2$
Layer 155	Classification Output	---

The weights and biasis are updated according to the equations 2.18, 2.19 respectively.

$$w_l = \left(1 - \frac{m_\tau}{s}\right) w_{l-1} - \frac{m}{x}\frac{\partial c}{\partial w} \tag{2.18}$$

$$b_l = b_{l-1} - \frac{m}{x}\frac{\partial c}{\partial w} \tag{2.19}$$

where w is weight, b is bias, l is layer, m is learning rate, τ is regulation parameter, s is training sample numbers, x is batch size, and c is cost function.

2.2.4.6 Performance Evaluation Metrics

Performance evaluation criteria are important parameters for evaluating ML algorithms applied on biomedical signal processing [31]. In this study, accuracy (ACC), specificity (SPE), recall (REC), precision (PRE), and F1-Score evaluation parameters were calculated to analyze the classification performance of the models developed for MDD detection objectively. The ACC is indicated as the ratio of correctly classified EEG signals according to the categories in the model to the total dataset. REC indicates the model's ability to be classified as 'diseased' among real patients. It ranges from 0 to 1. The REC value of a diagnostic model is expected to be high. A model's REC value of 1 indicates that that the model can accurately diagnose all patients from input data. SPE indicates the model's ability to be classified as 'healthy' among real healthy contols. A model's SPE value of 1 indicates that the model can accurately detect all healthy individuals. These evaluation parameters are calculated as follows;

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.20}$$

$$SPE = \frac{TN}{TN + FP} \tag{2.21}$$

$$REC = \frac{TP}{TP + FN} \tag{2.22}$$

$$PRE = \frac{TP}{TP + FP} \tag{2.23}$$

$$F1 - Score = \frac{2xRECxPRE}{REC + PRE} \quad (2.24)$$

where TPs, TNs, FPs and FNs are indicated as True Positives (TP) True Negative (TN), False Positives (FN) and False Negatives (FN), respectively. ‘Positive’ and ‘Negative’ are represented different classes of the dataset. In this study, positive is represents MDD class and negative is represents healthy class. TP and TN are the number of samples whose classes are correctly predicted. FP and FN are the number of samples whose classes are not correctly predicted.

The Receiver Operating Characteristic (ROC) [77] is a probability curve and used to evaluate classification performance. Each point on the ROC curve represents a sensitivity/specificity corresponding to a certain decision threshold. A successful classification performance has a ROC curve passing through the upper left corner (100% SEN 100% SPE). Therefore, the closer the ROC curve is to the upper left corner, the higher the overall accuracy of the test. The Area Under the ROC Curve (AUC) represents the degree of separation of different classes. It takes values between 0 and 1 and the ideal value for AUC is 1. The higher area under the curve, the higher the discrimination performance between classes. Figure 2.12 is an example of ROC curve and AUC representation. In the figure, AUC value is found 0.89 that near the ideal value.

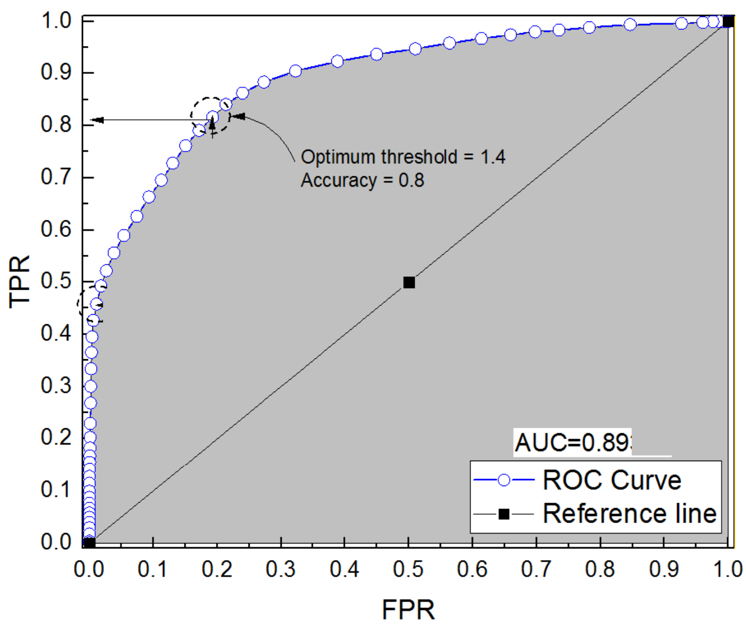

Figure 2.12: An example of ROC curve and AUC

2.2.4.7 *k*-fold Cross-Validation

k-fold cross-validation is a method that uses during the training process. The method ensures that obtaining the consistent result from the overall performance of the classification process on a dataset [79]. In this technique, before the classification process, the training dataset is randomly divided into k equal sub-datasets. Then, each subset is combined, one as a test set, and the other (*k*-1) sub-data set to form a training set. This process is repeated until each of the *k* subsets is used at least once as a test set. The total classification result is obtained by taking the average of the classification results obtained after the *k* steps. A cross-validation process with the *k* parameter set to 5 is shown Figure 2.13. In the data set divided into 5 parts, 4 parts are used as training set and the remaining part is used as test set. This process is repeated 5 times and at each step, a different part is selected as the test set. In this study, 10-fold cross-validation was applied to the training set during the classification process to prevent overfitting.

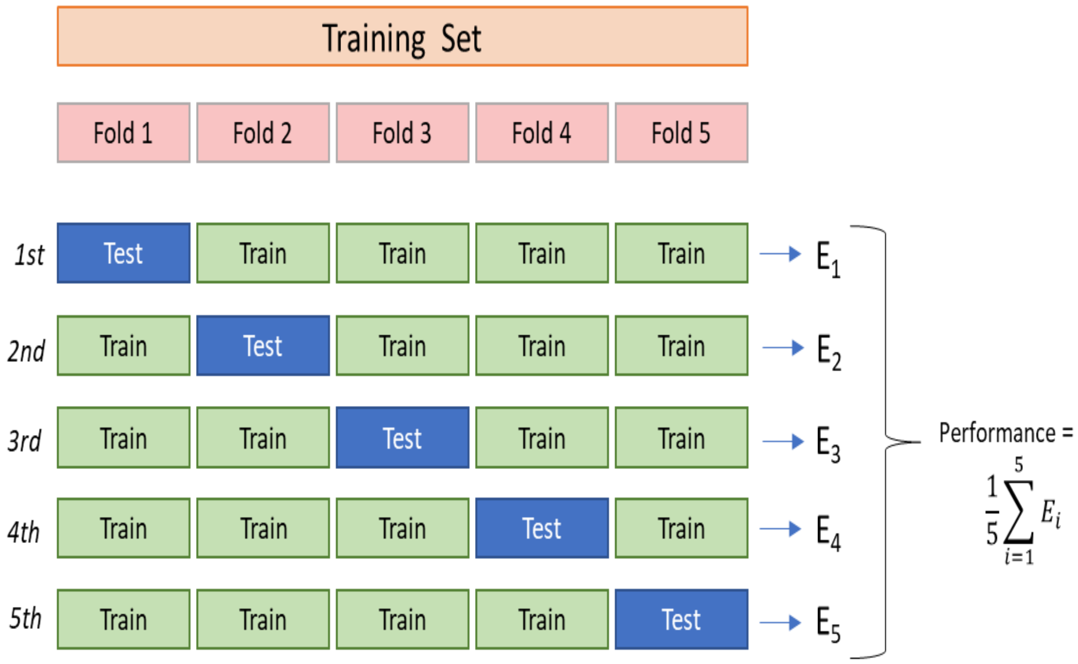

Figure 2.13: 5-fold cross-validation process

Chapter 3

Experimental Result

This thesis contributes to the detection of MDD by providing a computer-based automatic classification method that allows the data to be stored as MDD or healthy. In this thesis, EEG signals recorded from volunteer MDD subjects and HCs in Kemal Arıkan Psychiatry Center were differentiated using three proposed methods. The two methods are based on ML algorithms that composed of preprocessing, feature extraction, and classification parts. The difference between the two methods is using the EMD method before extracting features from EEG signal channels.

The classification performances of the proposed two methods which are ML algorithms with/without EMD are given in the below tables. The tables present the classification performances of four classifiers as LDA, SVM, Ensemble Learning, and DT in terms of the ACC, REC, SPE, PRE, and the F1-Score for the five brain region subgroups. In the tables, the highest classification results provided by the classifier have been emphasized in bold and the highest classification performances in the overall classification models have been emphasized in red bold.

3.1 Classification Performances of The Classifiers

Table 3.1 summarized the classification results based on SVM classifier. According to table, ACC values above 70% were achieved for all brain regions. It can be seen that the highest ACC, 86.7%, has been achieved when the feature vectors obtained from frontal region channels applied to the classifiers as the input. Combining SVM and frontal region channels have been reached the highest classification performances comparing other classification processes except DT in the proposed methods. The second highest classification result was obtained with the temporal region with an ACC of 81.2%.

Table 3.1: Classification performance of the SVM classifier (%)

	ACC	REC	SPE	PRE	F1-Score
Frontal	**86.7**	**94.0**	**81.3**	**78.0**	**85.2**
Central	72.9	77.2	69.0	65.0	70.0
Parietal	76.6	95.0	68.8	55.8	70.0
Temporal	81.2	94.0	74.0	66.2	77.7
Occipital	78.2	94.0	70.74	58.0	71.7

Figure 3.1 represents ROC curve and confusion matrix of SVM classifier for the frontal region. A confusion matrix is a table that used to specify the classification performance of a model on a set of test data for which the actual values are known. The confusion matrix composed of true and predicted classes, where the column is the true class and the row is the predicted classes. According to the confusion matrix of the blue cells represent the number of correct predictions of the classes (TPs and TNs) and the white cells represent the number of incorrect predictions of the classes (FPs and FNs). The classification result of the SVM for the frontal region is successful as observed by the high number of true predictions in the blue cells. As seen in the ROC curve, the AUC value is high and close the ideal value.

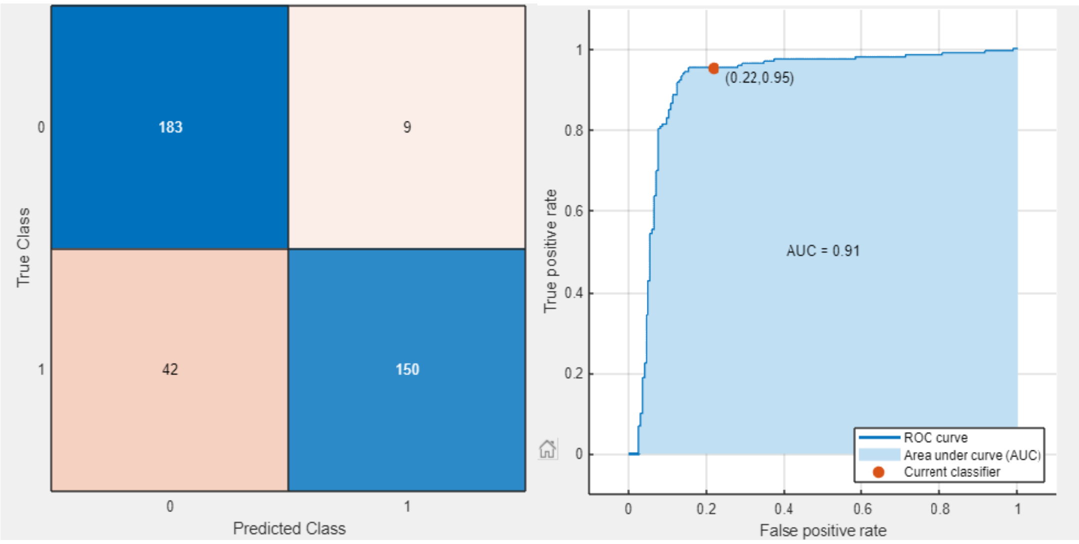

Figure 3.1: ROC curve and confusion matrix of SVM classifier for frontal region

Table 3.2 indicates the results of classification performance based on the LDA classifier for the five subgroups as frontal, central, parietal, temporal, occipital. As can be seen from the table, ACC values above 70% were achieved for all brain regions. The highest classification ACC of 78.2% was obtained with the occipital region. The second highest classification performance was achieved with the parietal region with ACC of 75.3%.

Table 3.2: Classification performance of the LDA classifier (%)

	ACC	**REC**	**SPE**	**PRE**	**F1-Score**
Frontal	74.2	86	68.0	57.9	69.1
Central	72.3	79.6	67.8	60.0	68.0
Parietal	75.3	100	66.0	50.0	66.6
Temporal	73.1	100	64.0	46.0	63.0
Occipital	**78.2**	**95.0**	**70.3**	**58.9**	**72.79**

Table 3.3 presents classification accuracies of brain regions based on the DT classifier. The DT classification accuracy of 86.6% was achieved with the frontal region same as the SVM classification result. The second highest classification performance was obtained with the parietal region with an ACC of 86.4%.

Table 3.3: Classification performance of the DT classifier (%)

	ACC	**REC**	**SPE**	**PRE**	**F1-Score**
Frontal	**86.6**	**95.4**	**69.5**	**77.0**	**85.2**
Central	77.4	72.0	85.0	88.4	79.0
Parietal	86.4	98.2	79.1	74.0	84.0
Temporal	70.0	88.0	66.0	48.0	62.1
Occipital	65.4	88.8	61.6	42.0	57.1

Table 3.4 indicates the results of classification performance based on the Ensemble Learning classifiers for brain regions. As can be seen from the table, the highest classification ACC of 85.7% was achieved with the frontal region. The second highest classification performance was obtained with the parietal region with an ACC of 83.1%.

Table 3.4: Classification performance of the Ensemble Learning classifier (%)

	ACC	**REC**	**SPE**	**PRE**	**F1-Score**
Frontal	**85.7**	**95.9**	**79.1**	**74.4**	**83.8**
Central	82.6	78.0	89.0	88.7	83.0
Parietal	83.1	100	74.1	66.2	79.5
Temporal	77.3	96.0	69.0	56.5	71.1
Occipital	65.4	92.8	59.3	35.0	50.8

Figure 3.2 is a graphical representation of the classification results of the classifiers with brain region. As seen in the Figure, the proposed method is successfully applied in the frontal region and gives the highest results except for LDA compared to other regions.

Figure 3.2: Graphical representation of classification results of the classifiers.

3.2 Classification Performances of The Classifiers with EMD

The second proposed method is based on EMD which is decomposed of the EEG signal channel into IMFs. In the proposed method, after the decomposition, IMF1 and IMF2 of EEG signal channels were selected and the significant features were extracted from the channels and their selected IMFs. The consisted feature vectors were classified with using SVM, LDA, DT and Ensemble Learning classifiers which are the same algorithms as the first method.

Table 3.5 demonstrates the classification performances of the SVM classifier after applying EMD for the brain region. According to the table, ACC values above 70% were achieved for all brain regions. Compared to the first method, the classification success in the central, parietal and occipital regions has increased with the application of the EMD method. The highest success was achieved with the classification in the parietal region. As can be seen from the table, the highest classification ACC of 90.4% was achieved with the parietal region. The second highest classification performance was obtained with the occipital region with an ACC of 82.4%.

Table 3.5: Classification performans of the SVM classifier with EMD (%)

	ACC	**REC**	**SPE**	**PRE**	**F1-Score**
Frontal	72.9	69.6	77.5	81.3	75.1
Central	73.5	72.1	74.8	74.1	75.8
Parietal	**90.4**	**86.9**	**94.7**	**95.0**	**90.0**
Temporal	76.6	77.4	75.7	74.8	76.0
Occipital	82.4	81.5	82.7	83.1	2.3

Table 3.6 summarized the classification results of LDA classifiers after applying EMD for the brain region subgroups. Compared to the first method, the classification success in the parietal and occipital regions has increased with the application of the EMD method. With the ACC of 91.5% in the classification of the parietal region, the highest classification success was obtained compared to the overall classification results of the

first and second methods. The second highest classification performance was obtained with the occipital region with an ACC of 84.4%.

Table 3.6: Classification performance of the LDA classifier with EMD (%)

	ACC	REC	SPE	PRE	F1-Score
Frontal	69.4	67.7	71.2	73.8	70.6
Central	72.0	74.0	70.2	67.7	70.8
Parietal	**91.5**	**96.1**	**87.7**	**86.5**	**91.1**
Temporal	73.1	74.1	72.2	71.2	72.5
Occipital	84.4	84.8	83.9	83.7	84.3

Figure 3.3 represents ROC curve and confusion matrix of LDA classifier after applying EMD for the parietal region. The classification result of the LDA is successful as observed by the high number of true predictions in the blue cells. According to the confusion matrix, TNs and TPs have high values which means that using this algorithm correctly distinguished MDD classes from Healthy classes. As seen in the ROC curve, the AUC value is high with 0.97 and close the ideal value.

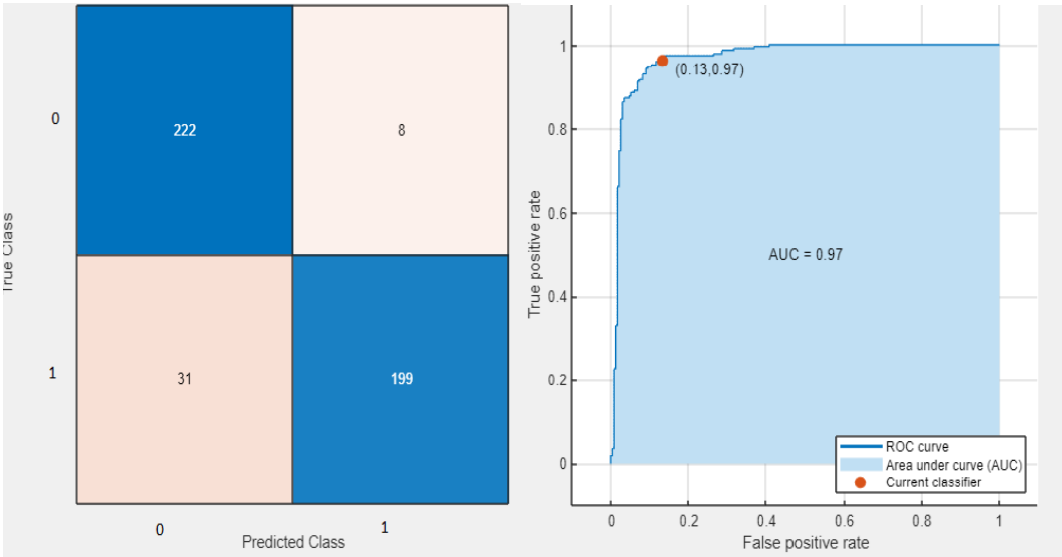

Figure 3.3: ROC curve and confusion matrix of LDA classifier for frontal region with EMD

Table 3.7 presents the classification performances of the DT classifier after applying EMD for the brain region. According to the table, ACC values above 70% were achieved for all brain regions. Compared to the first method, the classification performances in the temporal and occipital regions have increased with the application of the EMD method. The highest success was achieved with the parietal region with an ACC of 84.8% however the result is lower than the classification result of the first method. The second highest classification result was obtained with the frontal region with an ACC of 78.4%.

Table 3.7: Classification performance of the DT classifier with EMD (%)

	ACC	**REC**	**SPE**	**PRE**	**F1-Score**
Frontal	78.4	83.9	74.3	70.1	76.4
Central	76.5	81.3	72.9	68.6	70.7
Parietal	**84.8**	**92.5**	**79.4**	**75.7**	**83.2**
Temporal	71.4	79.6	66.8	57.6	66.8
Occipital	77.9	76.9	79.0	79.9	79.5

Table 3.8 indicates the classification performance of Ensemble Learning after applying EMD for the brain region subgroups. According to the table, ACC values above 70% were obtained for all brain regions. Compared to the first method, only the classification performance in the occipital region has increased with the application of the EMD method. The highest success was achieved with the occipital region with an ACC of 88.6%. The second highest classification result was achieved with the parietal region with an ACC of 81.1%.

Table 3.8: Classification performance of the Ensemble Learning classifier with EMD (%).

	ACC	REC	SPE	PRE	F1-Score
Frontal	80.3	88.0	75.2	70.1	77.9
Central	80.5	82.6	78.5	77.1	79.8
Parietal	81.1	98.6	72.8	63	76.9
Temporal	71.1	77.1	68.3	62.1	68.7
Occipital	**88.6**	**99**	**81.7**	**77.9**	**87.3**

Figure 3.4 is a graphical representation of the classification results of the classifiers after applying EMD with the brain region. It can be seen that the proposed method is successfully applied in the parietal region and gives the highest results except for Ensemble Learning compared to other regions.

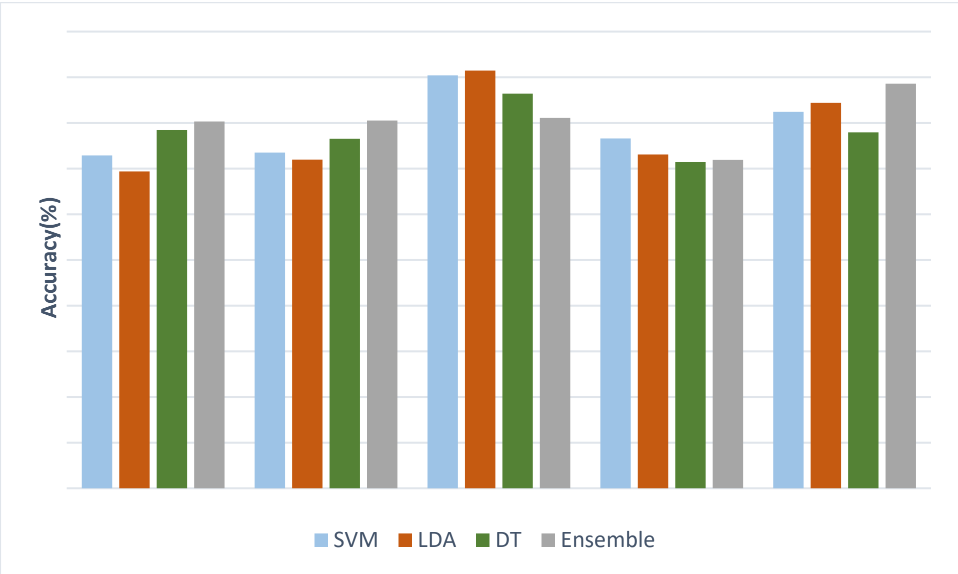

Figure 3.4: Graphical representation of classification results of the classifiers with EMD

3.3 Classification Performance of The CNN Model

In the third proposed method, Deep Network Design toolbox in MATLAB was used to evaluate the CNN model performance. The proposed 2D CNN model based on the classification of 2D topographic images obtained from EEG signals. A topographic image was designed with the aim of showing the spread of a selected feature over the cortex. In the study, the power of frequency bands was selected as a feature obtained from all channels of an EEG signal epoch. 2D topographic images with represent the power of delta, theta, and alpha frequency bands were used as an input in the CNN model. MobilNet model was used in the classification process inspired by the previous study. For finding optimal CNN architecture, different image, epoch, batch sizes, learning rates were tested in the training phase. To adjust the optimum parameters, the certain parameter is changed while the other parameters are constant until the best accuracy is obtained. Initially, the topographic images were resized to different image sizes as (256 × 256 × 3, 128 x 128 x 3, 64 x 64 x 3, 32 x 32 x 3) to feed the CNN architecture and the image size optimized to 32 x 32 x 3. In order to observe the effect of the batch size hyperparameter, the batch size was chosen as 32, 64, 128, respectively, and the batch size optimized to 32. Moreover, different learning rates (0.01, 0.001, 0.0001) were tested to achieve a lower error rate, and it fine-tuned on 0,001. Finally, 32x32x3 image sizes, 30 epoch sizes, 32 batch sizes, and 0.001 learning rate were tuned to be optimum for the proposed CNN model. The best classification performance was achieved with 80.09% ACC in the topographic images with the theta band power. Figure 3.5 shows the training and validation performance of the CNN architecture for topographic images with the theta band power about 0.59 validation loss, and 80.09% validation ACC.

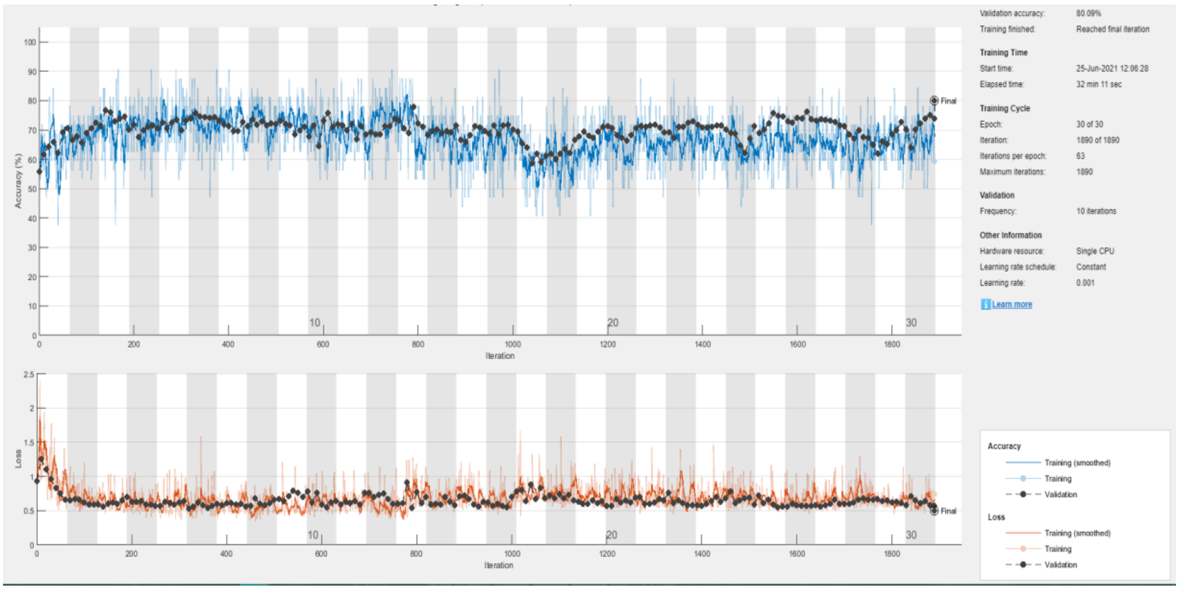

Figure 3.5: Classification performance of the CNN architecture for topographic images with the the theta band power (32x32x3 image sizes)

Figure 3.6 demonstrates training and validation performance of the same CNN architecture except image size to 64x64 with ACC of 74.88%. Increasing the image sizes also reduced the ACC of the proposed model.

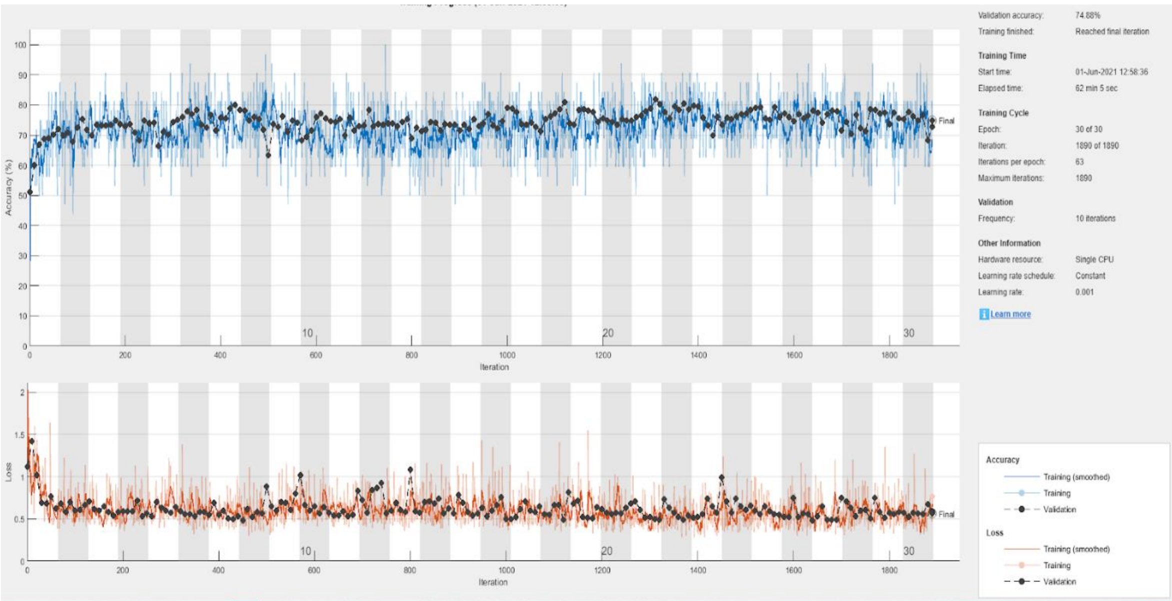

Figure 3.6: Classification performance of the CNN arhitecture for topographic images with theta band power (64x64x3 image sizes)

Chapter 4

Discussion and Conclusion

Detection of any abnormalities in neural activity plays an important role in the early diagnosis of MDD. There are various neuroimaging tools to record the electrophysiologic activity of the brain. Among them, EEG is preferred in this study due to its low cost, non-invasive, and easy-to-operate properties. Besides, it provides high time resolution to follow the very low time precision of the status of the brain activity. Since EEG signals are nonstationary, nonlinear, and complicated, manual analysis for the diagnosis of MDD is a quite difficult issue for professionals. EEG-based CAD algorithms can be automatically differentiated MDD patients from HCs. In literature, there are various studies in this field for the classification of MDD and HC groups.

In this study, three methods that can be used in the diagnosis of MDD have been proposed with the combination of multichannel EEG signals and ML algorithms. The methods were developed to explore the effect of EMD on the classification process and whether 2D image classification using the CNN architecture can be successful to distinguish between the MDD and HC groups. The methods were applied in 19 channel EEG signals taken from 16 MDD patients and 16 HC controls. The success of these methods depends on various factors which are the quality of the signals used, the extraction of significant features, and the development of the appropriate ML algorithms.

The two proposed methods consist of four main steps: preprocessing of the EEG signals, selecting channels from EEG signals to group according to the brain region, feature extraction and classification proceses. The second method follows same algorithm of first method except using EMD based feature extraction process. EMD can be used for analyzing nonlinear and/or nonstationary signals like EEG signals and it decomposes the signals into IMFs that represent finite and simple functions. EMD was introduced as a method to detect MDD by decompose the signal without leaving the time domain. It is applied to remove the high-frequency components and to hold the useful information of EEG signal channels for extract efficient features. With decomposition, the main physical characteristics in the original signal are reflected from its IMFs. In the comparison of the first two methods, EMD based classification process performed better than without EMD. The three performance parameters: ACC, REC, and SPE of EMD based classification got to higher values compared to those without EMD with an ACC of 91.5%. Besides this, the first method also showed high performance in the classification of MDD and HC groups with ACC of 86.7%. Comparison of depression classification performances of the proposed study with previous studies is given in Table 4.1.

Table 4.1: Comparison of previous studies on EEG signal analysis for depression

Authours	**Year**	**Database**	**Method**	**Classifier**	**ACC (%)**
Mahato et al. **[55]**	2020	34 MDD 30 HC	Frequency band power and theta asymmetry Multi-Cluster Feature Selection	SVM	88.3
Shen et al. **[80]**	2019	81 Depressed 89 HC	EMD	SVM	83.2
Bachmann et al. **[81]**	2017	13 Depressed 13 HC	Single channel EEG using 3 linear and 3 nonlinear measures	LR	92.0
Duan et al. **[82]**	2020	16 MDD 16 HC	Interhemispheric asymmetry and cross-correlation	CNN	94.1
The second proposed method	2021	16 MDD 16 HC	EMD	LDA	91.5

It has been stated in studies in the literature that brain regions have different effects for the diagnosis of MDD disease. Considering this, EEG signals were analyzed according to brain regions in the first two proposed methods. The developed classification processes were performed with feature sets obtained from selected channels of five brain regions as frontal, central, parietal, temporal, and occipital, respectively. In the first method, the classification result of SVM for the channels collected from the frontal (Fp1, Fp2, F3, F4, Fz) region achieved better result compared with other brain regions. The result verified previous studies which are emphasized the frontal region could be a potential biomarker to identify MDD [83]. In the second proposed method, the parietal region (P3, P4) classification with the combination of EMD and LDA was demonstrated highest classification performance among all brain regions classification. Using EMD before feature extraction improved the classification process for the parietal region and increased the performance of the method to identify MDD automatically.

In our previous studies, EEG-based emotion recognition [31] and ECG arrhythmia detection [77] problems successfully performed with a combination of CNN and 2D images obtained from EEG and ECG signals. Inspired by our previous studies, the third method presents a DL model for the identification of MDD using 2D topographic images obtained from EEG signals. Based on the robust ability of CNN to classify 2D data, the CNN model was used to classify EEG signals where all EEG channels are represented as a single RGB color 2D image. The images have represented the spread of a certain feature in the whole cortex which is obtained from EEG channels. As to the frequency band powers which are selected the feature for obtaining images, using theta frequency band powers gives better result compared to delta and alpha power for MobilNet architecture. It has been observed that classification result was increased with decreasing image sizes. 32x32 image sizes demonstrated the highest result. The generated topographic images contain information for 19 channels. The lower image size showed a more efficient result for this information to be interpreted. The proposed DL model achieved a classification ACC of 80.09% for identifying MDD from 2D images. The novelty of the proposed CNN model is that it does not require preprocessing and feature extraction processes to classified input images. The model automatically learns the distinctive features and classified them with their many hidden layers. And also, transforming EEG signals to the 2D topographic images preserve the spectral, and temporal qualities and information about the location of cortical regions in the spatial domain.

Despite all of the advantages mentioned above, the developed methods have some limitations. The methods for MDD detection were constructed using a database of 32 subjects (16 MDD and 16 HC). The database is limited in order to develop an quantitative and successful method and it is necessary to increase the database. Due to the EEG signals were analyzed according to brain regions in the first two method the imbalance database were obtained between region groups. For the frontal region, five channels were selected for classification in contrast to the central, parietal, and occipital regions which were selected two channels. The limitation of the third method is the resized of topographic images to different image sizes. Resolution variations in resized topographic images may problems for differentiation in the features obtained and may affect the color intensity of the 2D images.

This thesis aimed to develop an quantitative algorithm that can be used in the field of health and can help professionals during the diagnosis of MDD. Two main approaches were introduced in the study, it was shown that the EMD algorithm can be used in the diagnosis of MDD with improved the algorithms in future and that 1D biomedical data can also be examined in 2D images using DL algorithm. The proposed methods will be able to offer a model that professionals can refer to during the diagnosis of MDD as it gives fast and accurate results. The developed methods provide a simple and robust automatic MDD detection scheme for the classification of EEG signals and appropriate for mobile device-based diagnosis systems. In future studies, the methods can be used with extended for recognizing different stages of MDD.

References

[1] Celano CM, Huffman JC. Depression and cardiac disease: a review. Cardiology in review. 2011;19(3):130-42.

[2] Korczak D, Pereira S, Koulajian K, Matejcek A, Giacca A. Type 1 diabetes mellitus and major depressive disorder: evidence for a biological link. Diabetologia. 2011;54(10):2483-93.

[3] Organization WH. Depression and other common mental disorders: global health estimates. World Health Organization; 2017.

[4] Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine. 2006;3(11):e442.

[5] Purves D. Body and brain: a trophic theory of neural connections: Harvard University Press; 1988.

[6] Miryala R. The Role Of Neural Connections In Physical Behaviors Associated With Dance.

[7] Feigin VL, Vos T, Nichols E, Owolabi MO, Carroll WM, Dichgans M, et al. The global burden of neurological disorders: translating evidence into policy. The Lancet Neurology. 2020;19(3):255-65.

[8] Boldbaatar G. A Neuromarketing Study on Mongolian Consumers'Buying Decision Process. . Proceedings of the Mongolian Academy of Sciences. 2017:97-108.

[9] Lim SH, Nisar H, Thee KW, Yap VV. A novel method for tracking and analysis of EEG activation across brain lobes. Biomedical Signal Processing and Control. 2018;40:488-504.

[10] Group LGMH. Scale up services for mental disorders: a call for action. The Lancet. 2007;370(9594):1241-52.

[11] Sperner-Unterweger B, Kohl C, Fuchs D. Immune changes and neurotransmitters: possible interactions in depression? Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2014;48:268-76.

[12] Association AP. Diagnostic and statistical manual of mental disorders (DSM-5®): American Psychiatric Pub; 2013.

[13] Kanner AM. Is major depression a neurologic disorder with psychiatric symptoms? Epilepsy & behavior. 2004;5(5):636-44.

[14] Seedat S, Scott KM, Angermeyer MC, Berglund P, Bromet EJ, Brugha TS, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Archives of general psychiatry. 2009;66(7):785-95.

[15] Gonzalez JS, Fisher L, Polonsky WH. Depression in diabetes: have we been missing something important? Diabetes care. 2011;34(1):236-9.

[16] Surtees P, Wainwright N, Luben R, Wareham N, Bingham S, Khaw K-T. Psychological distress, major depressive disorder, and risk of stroke. Neurology. 2008;70(10):788-94.

[17] Berton O, Nestler EJ. New approaches to antidepressant drug discovery: beyond monoamines. Nature Reviews Neuroscience. 2006;7(2):137-51.

[18] Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. American Journal of Psychiatry. 2000;157(10):1552-62.

[19] Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. Science. 2015;349(6255):1489-94.

[20] Mandelli L, Petrelli C, Serretti A. The role of specific early trauma in adult depression: a meta-analysis of published literature. Childhood trauma and adult depression. European psychiatry. 2015;30(6):665-80.

[21] Roy T, Lloyd CE. Epidemiology of depression and diabetes: a systematic review. Journal of affective disorders. 2012;142:S8-S21.

[22] Thombs BD, Ziegelstein RC, Whooley MA. Optimizing detection of major depression among patients with coronary artery disease using the patient health questionnaire: data from the heart and soul study. Journal of General Internal Medicine. 2008;23(12):2014-7.

[23] Nutt DJ. Relationship of neurotransmitters to the symptoms of major depressive disorder. J Clin Psychiatry. 2008;69(Suppl E1):4-7.

[24] Malone KM, Waternaux C, Haas GL, Cooper TB, Li S, Mann JJ. Cigarette smoking, suicidal behavior, and serotonin function in major psychiatric disorders. American Journal of Psychiatry. 2003;160(4):773-9.

[25] Meyer JH, Krüger S, Wilson AA, Christensen BK, Goulding VS, Schaffer A, et al. Lower dopamine transporter binding potential in striatum during depression. Neuroreport. 2001;12(18):4121-5.

[26] Hanwella R, de Silva V. Diagnosis and management of depression. The Ceylon medical journal. 2008;53(2):60-2.

[27] Beck AT. The diagnosis and management of depression. 1973.

[28] Richter P, Werner J, Heerlein A, Kraus A, Sauer H. On the validity of the Beck Depression Inventory. Psychopathology. 1998;31(3):160-8.

[29] Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. Archives of general psychiatry. 1988;45(8):742-7.

[30] Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am Psychiatric Assoc; 2010.

[31] Ozdemir MA, Degirmenci M, Izci E, Akan A. EEG-based emotion recognition with deep convolutional neural networks. Biomedical Engineering/Biomedizinische Technik. 2020;1.

[32] Acharya UR, Sree SV, Swapna G, Martis RJ, Suri JS. Automated EEG analysis of epilepsy: a review. Knowledge-Based Systems. 2013;45:147-65.

[33] Ahmadlou M, Adeli H, Adeli A. Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease. Alzheimer Disease & Associated Disorders. 2011;25(1):85-92.

[34] Klassen B, Hentz J, Shill H, Driver-Dunckley E, Evidente V, Sabbagh M, et al. Quantitative EEG as a predictive biomarker for Parkinson disease dementia. Neurology. 2011;77(2):118-24.

[35] Mumtaz W, Xia L, Ali SSA, Yasin MAM, Hussain M, Malik AS. Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD). Biomedical Signal Processing and Control. 2017;31:108-15.

[36] Coutin-Churchman P, Anez Y, Uzcategui M, Alvarez L, Vergara F, Mendez L, et al. Quantitative spectral analysis of EEG in psychiatry revisited: drawing signs out of numbers in a clinical setting. Clinical Neurophysiology. 2003;114(12):2294-306.

[37] Read GL, Innis IJ. Electroencephalography (Eeg). The international encyclopedia of communication research methods. 2017:1-18.

[38] Teplan M. Fundamentals of EEG measurement. Measurement science review. 2002;2(2):1-11.

[39] Jasper HH. The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol. 1958;10:370-5.

[40] Foldvary-Schaefer N, Grigg-Damberger MM. Identifying interictal and ictal epileptic activity in polysomnograms. Sleep Medicine Clinics. 2012;7(1):39-58.

[41] Başar E, Başar-Eroğlu C, Güntekin B, Yener GG. Brain's alpha, beta, gamma, delta, and theta oscillations in neuropsychiatric diseases: proposal for biomarker strategies. Supplements to Clinical neurophysiology. 2013;62:19-54.

[42] Srimaharaj W, Chaisricharoen R, Chaising S, Sittiprapaporn P, editors. Classification of human brain attention focused on meditation, effected by L-

theanine acid in Oolong tea. 2018 International Conference on Digital Arts, Media and Technology (ICDAMT); 2018: IEEE.

[43] Lee PF, Kan DPX, Croarkin P, Phang CK, Doruk D. Neurophysiological correlates of depressive symptoms in young adults: a quantitative EEG study. Journal of Clinical Neuroscience. 2018;47:315-22.

[44] Gollan JK, Hoxha D, Chihade D, Pflieger ME, Rosebrock L, Cacioppo J. Frontal alpha EEG asymmetry before and after behavioral activation treatment for depression. Biological psychology. 2014;99:198-208.

[45] Fingelkurts AA, Fingelkurts AA, Rytsälä H, Suominen K, Isometsä E, Kähkönen S. Composition of brain oscillations in ongoing EEG during major depression disorder. Neuroscience research. 2006;56(2):133-44.

[46] Cai H, Han J, Chen Y, Sha X, Wang Z, Hu B, et al. A pervasive approach to EEG-based depression detection. Complexity. 2018;2018.

[47] Zhang B, Yan G, Yang Z, Su Y, Wang J, Lei T. Brain Functional Networks Based on Resting-State EEG Data for Major Depressive Disorder Analysis and Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2020;29:215-29.

[48] Liu W, Zhang C, Wang X, Xu J, Chang Y, Ristaniemi T, et al. Functional connectivity of major depression disorder using ongoing EEG during music perception. Clinical Neurophysiology. 2020;131(10):2413-22.

[49] Cai H, Qu Z, Li Z, Zhang Y, Hu X, Hu B. Feature-level fusion approaches based on multimodal EEG data for depression recognition. Information Fusion. 2020;59:127-38.

[50] Peng H, Xia C, Wang Z, Zhu J, Zhang X, Sun S, et al. Multivariate pattern analysis of EEG-based functional connectivity: A study on the identification of depression. IEEE Access. 2019;7:92630-41.

[51] Aydemir E, Tuncer T, Dogan S, Gururajan R, Acharya UR. Automated major depressive disorder detection using melamine pattern with EEG signals. Applied Intelligence. 2021:1-18.

[52] Li X, Zhang X, Zhu J, Mao W, Sun S, Wang Z, et al. Depression recognition using machine learning methods with different feature generation strategies. Artificial intelligence in medicine. 2019;99:101696.

[53] Sandheep P, Vineeth S, Poulose M, Subha D, editors. Performance analysis of deep learning CNN in classification of depression EEG signals. TENCON 2019-2019 IEEE Region 10 Conference (TENCON); 2019: IEEE.

[54] Mao W, Zhu J, Li X, Zhang X, Sun S, editors. Resting state EEG based depression recognition research using deep learning method. International Conference on Brain Informatics; 2018: Springer.

[55] Mahato S, Paul S. Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry. Journal of medical systems. 2020;44(1):1-8.

[56] Liao S-C, Wu C-T, Huang H-C, Cheng W-T, Liu Y-H. Major depression detection from EEG signals using kernel eigen-filter-bank common spatial patterns. Sensors. 2017;17(6):1385.

[57] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H, Subha DP. Automated EEG-based screening of depression using deep convolutional neural network. Computer methods and programs in biomedicine. 2018;161:103-13.

[58] Uyulan C, Ergüzel TT, Unubol H, Cebi M, Sayar GH, Nezhad Asad M, et al. Major depressive disorder classification based on different convolutional neural network models: Deep learning approach. Clinical EEG and Neuroscience. 2021;52(1):38-51.

[59] Sulaiman N, Taib MN, Aris SAM, Hamid NHA, Lias S, Murat ZH, editors. Stress features identification from EEG signals using EEG Asymmetry & Spectral Centroids techniques. 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES); 2010: IEEE.

[60] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London Series A: mathematical, physical and engineering sciences. 1998;454(1971):903-95.

[61] Safi MS, Safi SMM. Early detection of Alzheimer's disease from EEG signals using Hjorth parameters. Biomedical Signal Processing and Control. 2021;65:102338.

[62] DeCarlo LT. On the meaning and use of kurtosis. Psychological methods. 1997;2(3):292.

[63] Hjorth B. EEG analysis based on time domain properties. Electroencephalography and clinical neurophysiology. 1970;29(3):306-10.

[64] Oh S-H, Lee Y-R, Kim H-N. A novel EEG feature extraction method using Hjorth parameter. International Journal of Electronics and Electrical Engineering. 2014;2(2):106-10.

[65] Shon D, Im K, Park J-H, Lim D-S, Jang B, Kim J-M. Emotional stress state detection using genetic algorithm-based feature selection on EEG signals. International journal of environmental research and public health. 2018;15(11):2461.

[66] Shannon CE. A mathematical theory of communication. The Bell system technical journal. 1948;27(3):379-423.

[67] Topic A, Russo M. Emotion recognition based on EEG feature maps through deep learning network. Engineering Science and Technology, an International Journal. 2021.

[68] Ozdemir MA, Ozdemir GD, Guren O. Classification of COVID-19 electrocardiograms by using hexaxial feature mapping and deep learning. BMC Medical Informatics and Decision Making. 2021;21(1):1-20.

[69] Liu B. Supervised learning. Web data mining: Springer; 2011. p. 63-132.

[70] Ghahramani Z, editor Unsupervised learning. Summer School on Machine Learning; 2003: Springer.

[71] Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995;20(3):273-97.

[72] Aksehirli ÖY, Ankarali H, Aydin D, Saracli Ö. Tıbbi Tahminde Alternatif Bir Yaklaşım: Destek Vektör Makineleri. Turkiye Klinikleri Journal of Biostatistics. 2013;5(1).

[73] Rutkowski L, Jaworski M, Pietruczuk L, Duda P. The CART decision tree for mining data streams. Information Sciences. 2014;266:1-15.

[74] Raileanu LE, Stoffel K. Theoretical comparison between the gini index and information gain criteria. Annals of Mathematics and Artificial Intelligence. 2004;41(1):77-93.

[75] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert systems with applications. 2010;37(12):8659-66.

[76] Sewell M. Ensemble learning. RN. 2008;11(02):1-34.

[77] Degirmenci M, Ozdemir M, Izci E, Akan A. Arrhythmic heartbeat classification using 2d convolutional neural networks. IRBM. 2021.

[78] Izci E, Ozdemir MA, Degirmenci M, Akan A, editors. Cardiac arrhythmia detection from 2d ecg images by using deep learning technique. 2019 Medical Technologies Congress (TIPTEKNO); 2019: IEEE.

[79] Swami P, Gandhi TK, Panigrahi BK, Tripathi M, Anand S. A novel robust diagnostic model to detect seizures in electroencephalography. Expert Systems with Applications. 2016;56:116-30.

[80] Shen J, Zhang X, Hu B, Wang G, Ding Z. An improved empirical mode decomposition of electroencephalogram signals for depression detection. IEEE Transactions on Affective Computing. 2019.

[81] Bachmann M, Päeske L, Kalev K, Aarma K, Lehtmets A, Ööpik P, et al. Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis. Computer methods and programs in biomedicine. 2018;155:11-7.

[82] Duan L, Duan H, Qiao Y, Sha S, Qi S, Zhang X, et al. Machine learning approaches for MDD detection and emotion decoding using EEG signals. Frontiers in Human Neuroscience. 2020;14.

[83] Allen JJ, Reznik SJ. Frontal EEG asymmetry as a promising marker of depression vulnerability: Summary and methodological considerations. Current opinion in psychology. 2015;4:93-7.

Appendix A

Publications from the Thesis

Conference Papers

1. **İzci, E.,** Özdemir, M.A., Akan, A., Özçoban, M.A., Arıkan, M.K. "Majör Depresif Bozukluğun Tespiti için EEG ve Makine Öğrenmesi Tabanlı Bir Yöntem," IEEE 29. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU-2021, Çevrimiçi, 9-11 Haziran 2021.

Journal Articles

1. Ozdemir, M.A., Degirmenci, M., **Izci, E.,** and Akan, A., "EEG-Based Emotion Recognition with Deep Convolutional Neural Networks" Biomedical Engineering/Biomedizinische Technik (BMT), https://doi.org/10.1515/bmt-2019-0306, Vol. 66. No. 1, pp. 43-57, Feb, 2021. WOS:000621777900005
2. Degirmenci, M., **Izci, E.,** Ozdemir, M.A., and Akan, A., "Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks," Innovation and Research in BioMedical Engineering-IRBM, https://doi.org/10.1016/j.irbm.2021.04.002 June. 2021

Curriculum Vitae

Name Surname : Elif Izci

Education:

2018–2021 İzmir Kâtip Çelebi University, Dept. of Biomedical Technologies

2013–2018 İzmir Kâtip Çelebi University, Dept. of Biomedical Engineering

Work Experience:

2020 – 2021 Turkey Energy, Nuclear, Mine Research Institute - Boron Research Institute (TENMAK-BOREN)

Publications:

1. **İzci, E.,** Özdemir, M.A., Akan, A., Özçoban, M.A., Arıkan, M.K. "Majör Depresif Bozukluğun Tespiti için EEG ve Makine Öğrenmesi Tabanlı Bir Yöntem," IEEE 29. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU-2021, Çevrimiçi, 9-11 Haziran 2021.
2. Degirmenci, M., **Izci, E.,** Ozdemir, M.A., and Akan, A., "Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks," Innovation and Research in BioMedical Engineering-IRBM, https://doi.org/10.1016/j.irbm.2021.04.002 June. 2021.
3. Ozdemir, M.A., Degirmenci, M., **Izci, E.,** and Akan, A., "EEG-Based Emotion Recognition with Deep Convolutional Neural Networks" Biomedical Engineering/Biomedizinische Technik (BMT), https://doi.org/10.1515/bmt-2019-0306, Vol. 66. No. 1, pp. 43-57, Feb, 2021. WOS:000621777900005
4. **İzci, E.,** Değirmenci, M., Özdemir, M. A., Akan, A. "Derin Öğrenme ile EKG Aritmi Tespiti." 28. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, SIU-2020, Çevrimiçi, 5-7 Ekim 2020, 10.1109/SIU49456.2020.9302219
5. **Izci, E.,** Ozdemir, M. A., Degirmenci, M., Akan, A. "Cardiac arrhythmia detection from 2d ecg images by using deep learning technique." 2019 Medical Technologies Congress (TIPTEKNO), Izmir, 3-5 Oct. 2019, 10.1109/TIPTEKNO.2019.8895011

6. **Izci, E.,** Ozdemir, M. A., Sadighzadeh, R., Akan, A. "Arrhythmia detection on ECG signals by using empirical mode decomposition." 2018 Medical Technologies National Congress (TIPTEKNO), Cyprus, 8-10 Nov. 2018, 10.1109/TIPTEKNO.2018.8597094