2019

HUMAN ELECTROENCEPHALOGRAM BASED BIOMARKERS FOR DETECTION OF ALZHEIMER'S DISEASE

Al-Nuaimi, Ali H. Husseen

UNIVERSITY OF PLYMOUTH

HUMAN ELECTROENCEPHALOGRAM BASED BIOMARKERS FOR DETECTION OF ALZHEIMER'S DISEASE

BY

ALI H. HUSSEEN AL-NUAIMI

A thesis submitted to the University of Plymouth
in partial fulfilment of the degree of

DOCTOR OF PHILOSOPHY

School of Engineering, Computing and Mathematics

July 2018

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee. Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed with the aid of a scholarship form the Iraq Ministry of Higher Education and Scientific Research (MoHESR).

Word count of main body of thesis: 52,186.

Publications:

1. A.H.H. Al-Nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Tsallis entropy as a biomarker for detection of Alzheimer's disease, in: Eng. Med. Biol. Soc. (EMBC), 2015 37th Annu. Int. Conf. IEEE, IEEE, 2015: pp. 4166–4169.

DOI: 10.1109/EMBC.2015.7319312

PEARL (OA): http://hdl.handle.net/10026.1/13830

2. A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Changes in the EEG amplitude as a biomarker for early detection of Alzheimer's disease, in: Eng. Med. Biol. Soc. (EMBC), 2016 IEEE 38th Annu. Int. Conf., IEEE, 2016: pp. 993–996.

DOI: http://dx.doi.org/10.1109/EMBC.2016.7590869

PEARL (OA): http://hdl.handle.net/10026.1/13251

3. A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Higuchi Fractal Dimension of the Electroencephalogram as a Biomarker for Early Detection of Alzheimer's

Disease, in: Eng. Med. Biol. Soc. (EMBC), 2017 IEEE 39th Annu. Int. Conf. IEEE, IEEE, 2017: pp. 2320–2324.

DOI: http://dx.doi.org/10.1109/EMBC.2017.8037320

PEARL (OA): http://hdl.handle.net/10026.1/10505

4. A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Changes in the Electroencephalogram as a Biomarker of Alzheimer ' s Disease, in:, International workshop on biosensors for dementia , 13 - 14 June 2017, 2017: p. 40.

DOI: https://www.plymouth.ac.uk/whats-on/biosense-dementia

5. A.H.H. Al-Nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease, Complexity. 2018 (2018) 1–12.

DOI: http://dx.doi.org/10.1155/2018/8915079

PEARL (OA): http://hdl.handle.net/10026.1/12567

Signed .....
Date .....

Acknowledgements

My entire deep thanks to Massive God - ALLAH, Who blessed me with the strength, confidence and determination needed for the completion of my research project.

This thesis would not have been possible without the support and guidance of many people.

First, it is my pleasure to express sincere thanks to my first supervisor and director of studies Professor Emmanuel Ifeachor for his professional guidance, encouragement, patience and support throughout the present research project, for the benefit of his wide knowledge and vision for this project and for the tremendous amount of time and efforts he has spent to ensure the high quality of my papers and this thesis. I would also like to thank Dr Lingfen Sun and Dr Emmanuel Jammeh for their support and guidance during this project.

I would like to acknowledge the Iraq Ministry of Higher Education and Scientific Research (MoHESR) for the award of scholarship that provided the financial support for this research project.

I would like to express my sincere gratitude to DTC Administrator Mrs Carole Watson, and the Administrator Staff especially Julie Platt and Karen Ansell.

Finally, I am very grateful for the encouragement and support of my Mother. I am extremely grateful to my family, especially my wife Shaymma Al-Juboori and my children Hussein, Abdulrahman and Lujain for their lovely smiles that can relieve any kind of tiredness.

Human Electroencephalogram Based Biomarkers for Detection of Alzheimer's Disease

Ali H. Husseen Al-nuaimi

Abstract

Alzheimer's disease (AD) is a progressive disorder that affects cognitive brain functions and develops many years before there are any clinical manifestations. A biomarker that provides a quantitative measure of changes in the brain in the early stages of AD would therefore be useful for early diagnosis. However, this would involve dealing with large numbers of people because up to 50% of dementia sufferers do not receive a formal diagnosis. Thus, there is a need for accurate, low-cost, robust, and easy to use biomarkers that can be used to detect AD in its early stages.

Recent guidelines promote the use of biochemical and neuroimaging biomarkers to improve the diagnosis of AD. Cerebral spinal fluid (CSF) testing for AD is not widely used in clinical practice because it involves an invasive lumbar puncture procedure. Neuroimaging (e.g., positron emission tomography-PET), on the other hand, is expensive, available only in specialist centres, and may be unsuitable for patients with pacemakers or certain other implants. Blood-based biomarkers have shown promising results in terms of AD diagnosis, but these are not yet fully developed and low-cost biosensors to detect such biomarkers do not yet exist.

However, electroencephalogram (EEG) based biomarkers can potentially fulfil these needs and play a vital role in the early diagnosis of AD. AD causes changes in EEGs that are thought to be associated with functional disconnections among cortical areas due to the death of brain cells. EEG analysis may therefore provide valuable information about brain dynamics in AD. Potentially, the EEG could be used to detect changes in brain signals even in the preclinical stages of the disease. This means it

could be used as a first line decision-support tool in AD diagnosis and complement other AD biomarkers.

This thesis describes research into the development of EEG biomarkers that detect AD based on analysis of changes in the EEG. The most characteristic features in AD are slowing of the EEG activities, a decrease in coherence, and a reduction in complexity. These changes can be quantified as a biomarker of AD. In this study, we identified characteristic EEG features that have a significant association with AD. The most promising EEG features were then used to develop EEG biomarkers that can exhibit high diagnostic performance.

Four measures of complexity were investigated and evaluated for their suitability as the basis for EEG-based biomarkers of AD: Tsallis entropy, Higuchi Fractal dimension, Lempel-Ziv complexity, and approximation entropy. Two EEG slowing measures were also investigated and evaluated: changes in zero-crossing intervals, and changes in the power spectrum of EEG. In addition, a new approach to quantifying the slowing of EEGs based on analysing changes in EEG amplitudes was developed and evaluated. The coherence of connections among cortical regions of the brain was also investigated to analyse EEG connectivity.

A new biomarker was developed based on analysing changes in EEG amplitude (ΔEEG_A). This is a marker for the subsequent rate of cognitive and functional decline in AD patients and provides high diagnostic performance. The performance of ΔEEG_A measured 100% and 88.88% for sensitivity and specificity, respectively. Our results therefore show that EEG-based measures can potentially be a good biomarker for AD.

An important contribution of the thesis is the development of a method to derive robust biomarkers from the EEG through selective band filtering and by combining key biomarkers. Thus, this study provides a framework for constructing robust EEG biomarkers that can be used to detect AD with high diagnostic performance (e.g., in terms of sensitivity and specificity).

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Glossary

ACC	Accuracy
AD	Alzheimer's Disease
ADHD	Attention Deficit Hyperactivity Disorder
ApEn	Approximation Entropy
APOE	Apolipoprotein E
APP	Amyloid Peptide Precursor Protein
AUC	Area Under Curve
Aβ	Amyloid Beta
C	Central
CSF	Cerebral Spinal Fluid
CT	Computed Tomography
EEG	Electroencephalogram
ERP	Event-Related Potential
F	Frontal
FFT	Fast Fourier Transform
FN	False Negative
FNR	False Negative Rate
FP	False Positive
Fp	Frontopolar
FPR	False Positive Rate
HFD	Higuchi Fractal Dimension
Hz	Hertz
KNN	K-Nearest Neighbour
LDA	Linear Discriminant Analysis
LZC	Lempel Ziv Complexity
MCC	Matthew's Correlation Coefficient
MCI	Mild Cognitive Impairment
MEG	Magnetoencephalogram
MRI	Magnetic Resonance Imaging
NFTs	Neurofibrillary Tangles
NPV	Negative Predictive Value
O	Occipital
P	Parietal
PET	Positron Emission Tomography
Ph.D.	Doctor Of Philosophy
POST	Positive Occipital Sharp Transient
PPV	Positive Predictive Value
P-tau	Phosphorylated Tau
ROC	Receiver Operating Characteristic
Sen	Sensitivity

Spec	Specificity
SPECT	Single Photon Emission Computed Tomography
SVM	Support Vector Machine
T	Temporal
Tcdf	Student's T Cumulative Distribution
TN	True Negative
TP	True Positive
TsEn	Tsallis Entropy
T-tau	Total Tau
UK	United Kingdom
US	United States
ZCI	Zero-Crossing Intervals
ΔEEG_A	Changes in The EEG Amplitude
ΔPS	Changes in the Power Spectrum

List of Symbols

$	American Dollar
α	Alpha
β	Beta
γ	Gamma
δ)	Delta
θ	Theta
κ	Kappa
λ	Lambda
μ	Micro
ρ	Rho
σ	Sigma
τ	Tau
φ	Phi
χ	Chi

Chapter 1. Introduction

1.1. Motivations

Alzheimer's disease is an age-related, progressive and neurodegenerative disorder characterised by loss of memory and cognitive decline [1][2]. It represents 70%-80% of all dementias [3][4][5], is the main cause of disability among older people [6] and the sixth leading cause of death in the US [7]. The rapid increase in the number of people living with AD and other forms of dementia is due to an ageing population and represents a major challenge for health and social care systems worldwide [8]. Currently, there are over 46.8 million individuals living with dementia in the world and the annual cost of care is estimated at US$818 billion [9][10]. This number is projected to reach 74.7 million by 2030 at an annual cost of US$ 2 trillion [11]. The number of individuals with dementia worldwide is expected to exceed 131 million by 2050, the economic impact of which will be enormous [9]. However, many dementia sufferers do not receive an early diagnosis [9][12]. In fact, it is estimated that up to 50% of people living with dementia may not have received a formal diagnosis [12][13]. For example, in 2011, 28 million dementia sufferers out of a total of 36 million worldwide were as yet undiagnosed [14].

The degeneration of brain cells due to AD begins many years prior to any clinical manifestations [8][15][16][17][18][19]. An early diagnosis of AD will therefore contribute to the development of effective treatments that could slow, stop, or prevent significant cognitive decline [18][20][21]. An early diagnosis of AD could also be useful in identifying dementia sufferers who have not received a formal early diagnosis and provide them with an opportunity to access appropriate health care services and would facilitate the development of new therapies [22][23][24][25][26][27][28].

Age is the main risk factor for AD [29]. Loss of recent memory is one of the first symptoms of AD (early stage), followed by mild cognitive impairment (MCI), and then

severe AD, which is the advanced stage [30]. MCI usually denotes a transitional cognitive state between normal ageing and dementia [29][31], although only 80% of MCI cases go on to develop dementia [29][32]. This is sufficient, however, for MCI to help identify people at high risk of dementia [31].

A biomarker that can measure the degeneration of brain cells caused by AD would help facilitate an early diagnosis [2][33][34][35]. Given the large number of people involved, there is a need for simple, non-invasive, low-cost, and reliable biomarkers that can be accessed in clinical practice for early diagnosis. There is therefore a need for simple, non-invasive, low-cost, and reliable biomarkers that can be accessed in clinical practice for early diagnosis [8][36][37]. Recent guidelines have promoted the use of biochemical and neuroimaging biomarkers to improve the diagnosis of AD. Among these, cerebral spinal fluid (CSF) testing for AD is not widely used in clinical practice because it involves an invasive lumbar puncture procedure [2][38][39]. Neuroimaging, on the other hand, is expensive, available only in specialist centres [40], and may not be suitable for patients with pacemakers or certain other implants [4]. Blood-based biomarkers have shown promising results in AD diagnosis but are not yet fully developed and low-cost biosensors to detect AD do not yet exist [2][36][41].

Potentially, the electroencephalogram (EEG) can play a valuable role in the early diagnosis of AD [15][23][24][35][42][43][44]. EEG is non-invasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics in AD [23][24][43][45][46]. The fundamental utility of EEG in detecting changes in brain signals, even in the preclinical stage of the disease, has been widely demonstrated [43][47][48]. Thus, EEG biomarkers may be used as a first line decision-support tool in AD diagnosis [15][45] and could complement other AD biomarkers [37].

In summary, there is an urgent need to develop low cost, non-invasive, robust biomarkers that can be used to detect AD and monitor its progression.

1.2. Aim and objectives

The aim of this project is to contribute to the investigation and development of automated EEG analysis techniques that can be used to detect Alzheimer's disease in its pre-clinical stages. The degeneration of brain cells due to AD begins many years prior to any clinical manifestations. An early diagnosis of AD will contribute to the development of effective treatments that could slow, stop, or prevent significant cognitive decline. It could also be useful in identifying dementia sufferers who have not received a formal early diagnosis and provide an opportunity for them to access appropriate health care services. To achieve this, it is necessary to develop low-cost, easy to use, accurate, reliable, and robust biomarkers that can be used to quantify and detect changes in the EEG that are attributable to pre-clinical AD. This is a substantial challenge because the EEG is a complex, non-stationary signal that varies between people and is affected by their condition (age, gender, wakefulness, AD stage, other diseases, and so on). It is also affected by stimuli such as light in the eye, sounds applied to the ear, or the sensation of pain. The objectives of this study are therefore as follows:

1. Review available techniques for EEG analysis for AD and identify those which are most promising.
2. Investigate and evaluate EEG measures to detect AD.
3. Identify promising EEG features related to AD that may lead to high diagnostic performance.
4. Investigate the development of novel biomarkers for detecting AD based on the EEG measures.

5. Create a framework for the development of robust EEG based biomarkers that can be used in the detection of AD.
6. Determine the smallest set of EEG biomarkers to detect AD with clinically acceptable performance.

1.3. Contributions of the thesis

This thesis therefore makes the following contributions to knowledge:

1. It provides a detailed understanding of the three main EEG analysis techniques used for AD detection: slowing of EEG, decrease in EEG complexity, and reduction in EEG coherence. This includes the identification and evaluation of key methods to detect AD from the EEG.
2. Development of novel EEG-based biomarkers based on analysing changes in EEG amplitude (ΔEEG_A). The results show that ΔEEG_A is a promising nonlinear EEG marker that can be used to quantify changes in EEG. It measures gradual changes in EEG amplitude. Its performance was found to be 100% and 88.88% for sensitivity and specificity, respectively.
3. Development of a new approach to enhance complexity-based EEG biomarkers used to detect AD. Nonlinear analysis methods based on EEG complexity have shown promising results in detecting changes in the EEG that are thought to be attributable to AD. We found that AD patients have significantly lower complexity values than normal people.
4. Creation of a new framework for the development of robust EEG-based biomarkers for the detection of AD. This has resulted in the development of biomarkers with high diagnostic performance (sensitivity and specificity close to 100%).
5. The development of new software tools for EEG analysis and AD detection.

The majority of the work reported in this thesis has been published at three EMBC conferences [8][45][49], one short paper presented at the international workshop on biosensors for dementia [35], and one journal paper [50]. A final paper on the creation of robust EEG-based biomarkers has been submitted to the IEEE Transactions on Biomedical Engineering journal.

1.4. Thesis outline

This thesis comprises eight chapters. Chapter one (this chapter) provides an introduction that clarifies the motivations, aims and objectives, and contributions of the thesis. Chapter 2 provides important background to AD, the nature of the EEG signal, the computational and signal processing methods that were investigated in this study (e.g., Tsallis entropy, Higuchi fractal dimension, Lempel-Ziv complexity, changes in the EEG amplitude, zero-crossing interval, approximation entropy, EEG coherence, and changes in the power spectrum of EEG signal), machine learning approaches that were used in the study (e.g., support vector machine, K-nearest neighbours, and linear discriminant analysis), and the diagnostic performance measures used in AD classifications. Chapter 3 describes the materials and methods used in the investigations. These include a description of the EEG datasets used in the study together with the computational programs employed (e.g., MATLAB toolbox and the MATLAB functions), the pre-processing techniques for splitting the EEG signal into five EEG frequency bands (e.g., bandpass filters for the delta, theta, alpha, beta, and gamma bands), and an overview of EEG based biomarkers (e.g., slowing of EEG, EEG complexity, and EEG coherence). In Chapter 4, the outcome of investigations of novel EEG biomarkers for detection of Alzheimer's disease (based on Tsallis entropy, changes in the EEG amplitude, and Higuchi fractal dimension) are presented. In Chapter 5, three complexity measures are investigated: Tsallis entropy, Higuchi fractal dimension, and Lempel-Ziv complexity. These are used to

derive the biomarkers from the EEG frequency bands and from the entire EEG signal. The results show that EEG biomarkers extracted from EEG frequency bands exhibit greater performance than the biomarkers from the entire EEG record. In Chapter 6, the three main techniques used to analyse the EEG signal are investigated to create robust EEG biomarkers. These techniques are: slowing of EEG, reduction in EEG complexity, and EEG coherence. For EEG slowing, changes in EEG amplitude method, zero crossing interval, and changes in the power spectrum of EEG are investigated. For the reduction in EEG complexity, Tsallis entropy, Higuchi fractal dimension, Lempel-Ziv complexity, and approximation entropy are investigated. To analyse EEG connectivity, the reduction in EEG coherence among cortical regions is also investigated. The results show that EEG features can be used to construct robust EEG biomarkers by combining key features of the EEG. In Chapter 7, further investigations are carried out to identify the smallest subset of biomarkers to detect AD with clinically acceptable performance. A new EEG dataset was used in the investigation to avoid bias and to increase the sample size. Chapter 8, reviews the findings, presents final conclusions, and considers the potential for future work in this area.

Chapter 2. Background

2.1. Alzheimer's disease

Alzheimer's disease is an irreversible, neurodegenerative¸ and fatal disease of the brain. It is a progressive disease that leads to dementia [30][5] and is characterised by gradual cognitive impairments accompanied by abnormal behaviour, loss of memory, and personality changes[2][30][51]. The two main neuropathologic hallmarks of AD are extracellular amyloid beta (Aβ) plaques and intracellular neurofibrillary tangles (NFTs). The production of Aβ, which represents a vital stage in AD pathogenesis, is the result of an aberrant cleavage of the amyloid peptide precursor protein (APP) which is overexpressed in AD [52][53][54]. In histopathological terms, AD is characterised by the accumulation of senile plaques and neurofibrillary tangles [55]. The senile plaques consist mainly of β-amyloid peptides, while the fibrillary tangles consist of abnormal hyperphosphorylated insoluble forms of the TAU-protein [29][56]. AD is the most common form of dementia (others include vascular, Lewy body, or frontotemporal dementia) and accounts for between 60% to 80% of all dementias worldwide [4][5].

Age is the main risk factor for AD [29]. Loss of recent memory is one of the first symptoms of AD (early stage), followed by mild cognitive impairment (MCI), and then severe AD, which is the advanced stage [30]. MCI describes a transitional cognitive state between normal ageing and dementia [29][31] and has been proposed as a disease describing elderly people with mild cognitive impairment but not dementia [17][31]. However, only 80% of MCI cases go on develop dementia [29][32], although this is sufficient to identify people at high risk of dementia [31]. At the macroscopic level, severe atrophy of the brain is the most common characteristic of AD patients, leading to enlargement of the ventricular system and shrinkage of cortical sulci [57]. In the preclinical stages of AD, the atrophy primarily affects medial temporal parts of

the brain, including the hippocampal formation [57]. Hippocampal atrophy is associated with more severe memory impairment in AD [58]. Such atrophy could therefore be used as a marker of disease development in clinical trials [4].

The rapid increase in the number of people living with AD and other forms of dementia is due to an ageing population and presents a major challenge to health and social care services worldwide [8]. Currently, over 46.8 million individuals worldwide live with dementia and the annual cost of care is estimated to be US$818 billion [9][10]. The number of dementia patients is projected to reach 74.7 million by 2030 at an annual cost of US$ 2 trillion [11]. Furthermore, this number is then expected to exceed 131 million by 2050 which will have a huge economic impact across the globe [9]. It is widely accepted that a preclinical AD diagnosis enables patients to gain access to appropriate health care services and facilitates the development of new therapies. However, many dementia sufferers do not receive an early formal diagnosis [9][12] and up to 50% of people living with dementia may never have received a formal diagnosis [12][13]. For example, in 2011, 28 million dementia sufferers out of a total of 36 million worldwide were undiagnosed [14].

The neurodegeneration of brain cells caused by AD begins many years before any clinical manifestations [8][15][16][17][18]. Such neurodegeneration is estimated to occur 20–30 years before clinical onset; thus, the identification of biological markers for pre-clinical stages is the major aim of researchers concerned with AD diagnosis [29][38]. This is because diagnosis at the pre-clinical stage AD enables patients to gain access to appropriate health care services and facilitates the development of new therapies. Furthermore, the subsequent treatment of preclinical stage AD would be an effective way to prevent the ultimate cognitive decline caused by AD [59]. An early diagnosis may also lead to the development of effective treatments that could slow, stop or prevent cognitive decline [18][20][21]. It could also be useful for

identifying dementia sufferers who have not received a formal diagnosis, thus providing an opportunity for them to access appropriate health care services [22][23][24]. To be clinically useful, the performance of markers should exceed 80% for sensitivity and specificity, respectively [60].

Figure 2-1 presents cross-sections of a healthy brain (left) and a brain with extensive atrophy in the late stages of AD (right) while Figure 2-2 compares the brain of a normal person (left) with the brain of an AD patient (right)

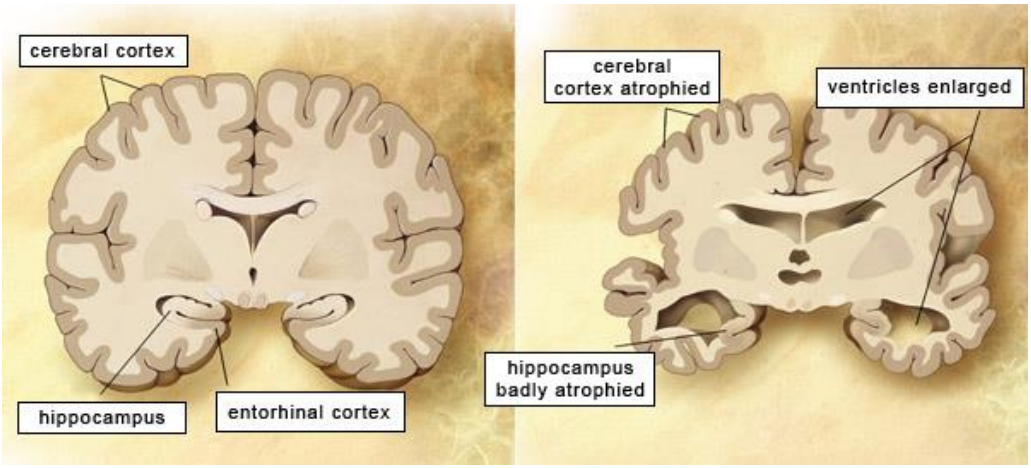

Figure 2-1: Cross-sections of a healthy brain (left) and a brain with extensive atrophy in the late stages of AD (right) [61]

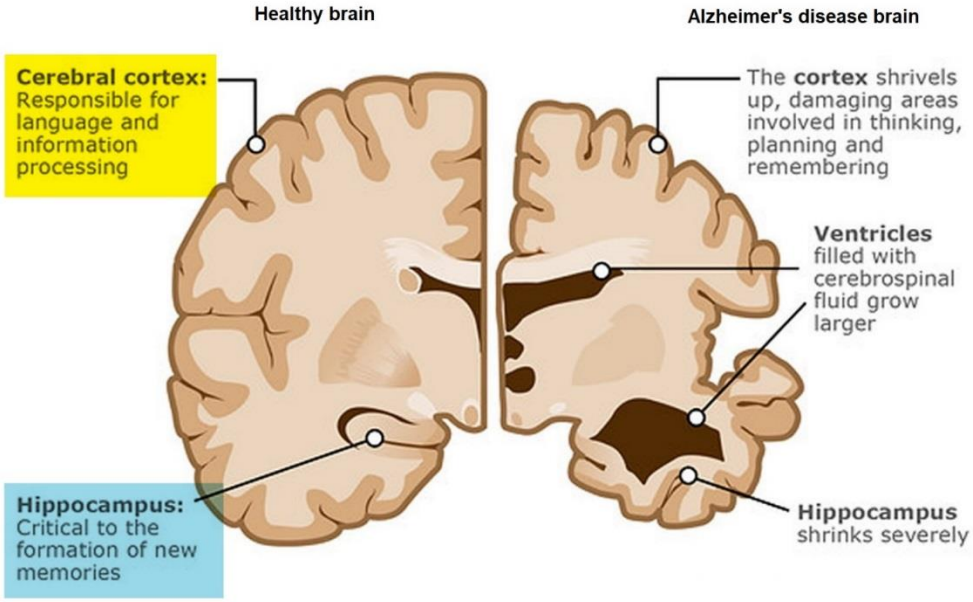

Figure 2-2: A healthy brain and an Alzheimer's disease affected brain [62]

Figure 2-3 provides a schematic representation of the two main neuropathologic hallmarks of AD: extracellular amyloid β (Aβ) plaques and intracellular neurofibrillary tangles.

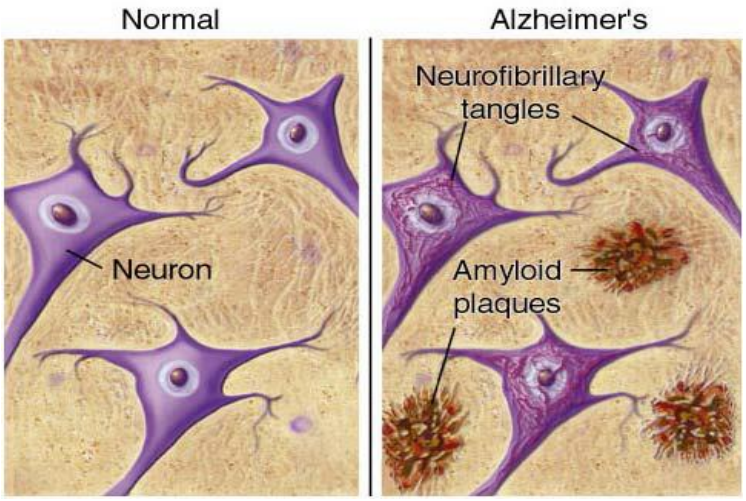

Figure 2-3: Illustration of the extracellular amyloid β (Aβ) plaques and intracellular neurofibrillary tangles (NFTs) in AD patients and in normal people

Figure 2-4 presents the evaluative stages from healthy through to normal age-related memory loss (yellow line) or AD dementia (red line). The blue line represents the mild cognitive impairment (MCI) stage which typically affects memory domains while other cognitive domains are preserved.

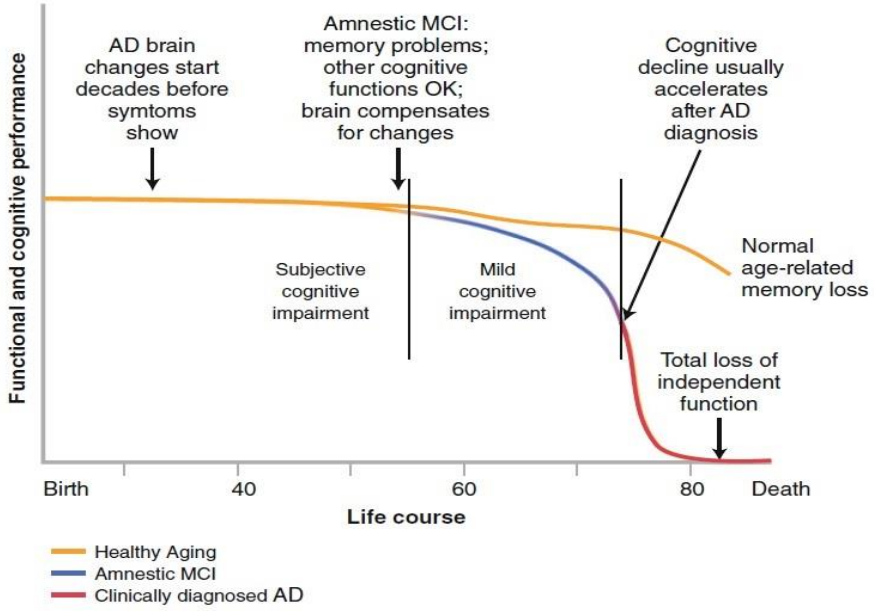

Figure 2-4: AD development stages from normal – MCI – AD [63]

2.2. Biomarkers of Alzheimer's disease

A biomarker is an indicator used to detect a specific biological state, such as functional and structural variations in organs and cells [2][33][34]. Biomarkers for AD are important tools that support a clinical diagnosis [17].

The ideal characteristics of a diagnostic biomarker for AD are as follows:

1. The biomarker should detect an essential feature of the pathophysiologic processes active in AD.
2. The biomarker should be validated in neuropathologically confirmed cases of AD.
3. The biomarker should be accurate, i.e., able to detect AD in its early stages and distinguish it from other types of dementia.
4. The measurement of the biomarker must be reliable, and the process must be non-invasive, easy to use, and low-cost [29][53][60].

Biomarkers for the detection of AD can be divided into three main types: genetic, neuroimaging, and biochemical biomarkers [64].

2.2.1. Genetic markers

Molecular genetic investigations of neurodegenerative disorders have been valuable in specifying genes that may be associated with specific diseases such as AD and in revealing the functional mechanisms supported by the products of those genes that may be markers of early detection. Genetics provides an assessment of who is at risk, but does not provide complete information in the case of rare genetic mutations [64]. The deposition rate and accumulation of b-amyloid protein (Aβ) in different parts of the brain represent the neuropathological key of AD [65][66]. Therefore, genes that reverse or inhibit these processes are an attractive candidate for the detection of AD [67]. The apolipoprotein E (APOE) gene is the only genetic risk factor that has thus far been linked to a risk of late-onset AD [68]. It represents the best example of a genetic association that increases the risk of developing late-onset AD. Genetic

studies provide the rationale for ongoing studies that evaluate the use of cerebrospinal fluid (CSF) markers of protein folding or accumulation as tools to detect early stage AD [64].

2.2.2. Neuroimaging biomarkers

There are three different types of neuroimaging techniques: structural imaging, functional imaging, and molecular imaging. Structural imaging such as magnetic resonance imaging – MRI, and computed tomography – CT, provides measures of the shape, position, or volume of the brain. Functional imaging, such as positron emission tomography – PET and functional MRI – fMRI, provides information about the brain cells working in different regions by screening how actively cells use sugar or oxygen. Molecular imaging, such as single photon emission computed tomography – SPECT), uses highly targeted radiotracers to detect cellular or chemical changes [29][63]. Functional brain imaging provides potential insights into the pathological features of AD such as neuronal loss, tangle deposition, cholinergic depletion, and amyloid plaques, and also facilitates measurement of the neurophysiological correlates of disease-related changes in the brain [29]. Neuroimaging techniques have yielded good results for both preclinical and early diagnosis of AD [64].

2.2.3. Biochemical markers

Biochemical markers of AD include CSF levels of total tau (T-tau), phosphorylated tau-181 (P-tau), and the 42-amino acid fragment of amyloid peptide ($A\beta_{1-42}$) [38][69][70]. Those markers reflect the essential neuropathology of the development of AD [55][69][70][71]. T-tau is a marker of cortical axonal decline, P-tau reflects neurofibrillary pathology, and $A\beta_{1-42}$ is a marker of plaque pathology which is reflected in the inverse relationship between plaque counts and $A\beta_{1-42}$ levels in CSF [72]. Measuring amyloid-β accumulation and tau in CSF provide the earliest pathological signature of AD [73][74]. For example, levels of CSF $A\beta_{1-42}$ are significantly lower in

the AD while levels of CSF tau are significantly higher [70][75][76]. Therefore, CSF measures have potential clinical utility as biomarkers of AD [75].

2.3. Electroencephalogram

2.3.1. Introduction

A biological signal describes the spontaneous activity of the human body [77]. An electroencephalogram (EEG) measures the electrical activity of the brain [77] by recording the activity generated by a large number of neurons [78]. The EEG plays a useful role in clinical diagnosis and the electrophysiological analysis of brain functions [77]. It records using electrodes that are attached to the scalp and have a range of 10 to 100 μV [78]. EEG is non-invasive, low-cost, and has a temporal resolution that makes it more popular in the analysis of brain dynamics [15]. However, various conditions can affect the EEG (e.g., age, gender, disease, and so on).

EEG analysis is very important in different clinical applications such as the detection and monitoring of brain injury, and in detecting abnormal brain states such as epilepsy and different types of dementia [78].

The concept of neurophysiology was first established by Carlo Matteucci (1811–1868) and Emil Du Bois-Reymond (1818–1896). They were the first people to register the electrical signals emitted from muscle nerves using a galvanometer. Later, Ernst Fleischl von Marxow (1845–1891) observed the cerebral electrical activity taking place in the visual cortex of different species of animals. Hans Berger (1873–1941) then became the first person to record human EEG signals [79].

2.3.2. Electrode montage

Montage refers to the process of arranging the locations of EEG channels in a specific order on the scalp [1][2][3][4]. The international 10–20 system refers to the standard arrangement of electrodes placed on the scalp [5]. This system provides a consistent and replicable method of EEG recording with 21 electrodes located at a relative

distance 10% or 20% from the nasion to the inion and the head circumference, hence the name “10-20” [2][5][6]. Alternative montages are also used, such as a 10-10 system for the placement of up to 74 channels, and a 10-5 system for the placement of up to 128-channel system based on the 10-20 system [4]. The standard 10-20 system has now been modified and is known as the Maudsley system. Although this is similar to the 10-20 system, the outer electrodes are slightly lowered to capture signals more effectively. This provides more extensive coverage of the lower part of the cerebral convexity [79].

Figure 2-5 shows the electrodes placed on the scalp based on the standard 10-20 system [79].

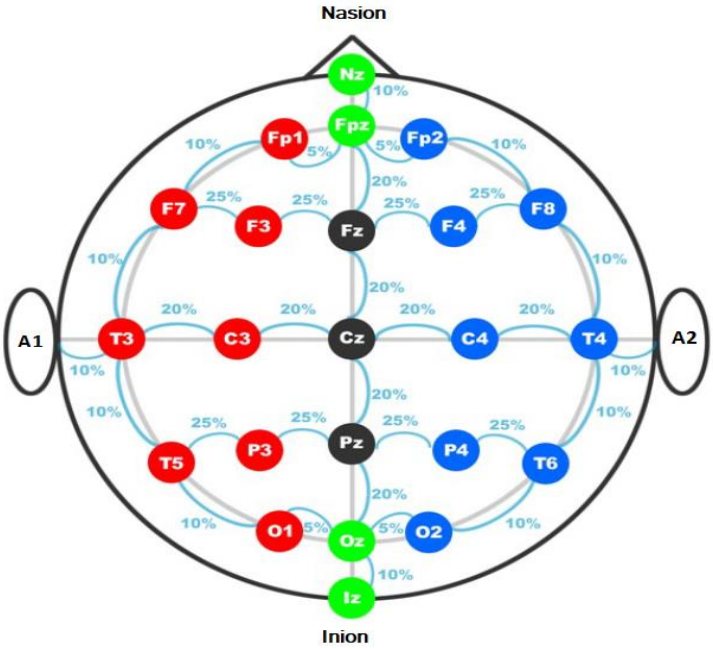

Figure 2-5: Electrode locations on the scalp based on the standard 10-20 system

Different areas of the brain are identified as Fp (frontopolar), F (frontal), C (central), P (parietal), O (occipital), and T (temporal). Odd numbers denote the left side, even numbers the right side, and Z the midline placements. ‘A’ signifies an ear channel (A1 for left ear, A2 for right) [77][80][81].

Figures 2-6 and 2-7 show the frontal and side views of electrode placement based on the 10-20 system [79].

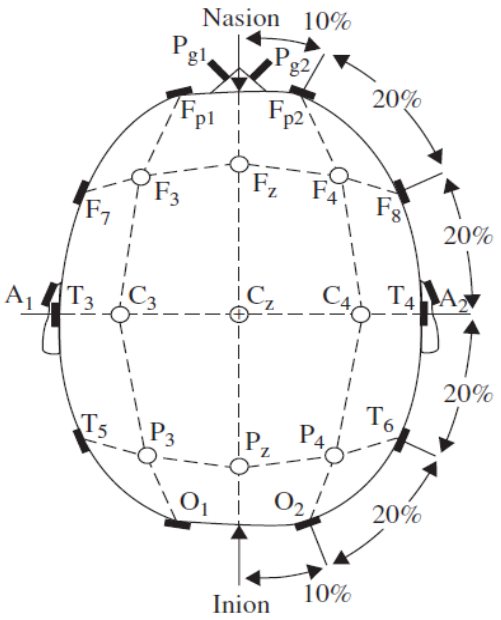

Figure 2-6: Frontal view of electrode placement based on the 10-20 system

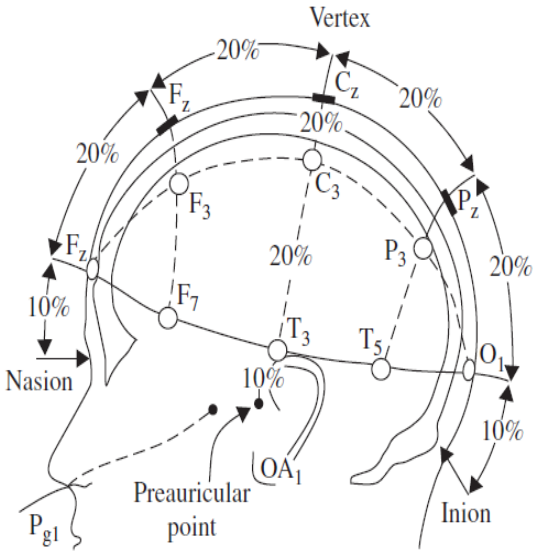

Figure 2-7: Side view of electrode placement based on the10-20 system.

The common referencing methods used in the montage are:

1. Bipolar: Measurements made between selected pairs of electrodes.
2. Common reference: Measurements taken between electrodes and a chosen reference point that is least affected by interference, such as the ear lobes (Al and A2).

3. Common average reference: Measurements taken between electrodes and the mean of the other electrodes used.

2.3.3. Artefacts

EEG measurement can be influenced by many factors including artefacts [78][82]. An ertefact refers to non-cerebral interference with the EEG [82]. The main artefacts can be divided into internal (patient-related) and external (system-related) artefacts [79]. Internal artefacts thus relate to body movement (e.g., eye movements-blinks, muscle movements), while external artefacts relate to the device (e.g., power supply interference, impedance fluctuation, cable defects, electrical noise from the electronic components) [78][79]. Artefacts present significant challenges in EEG analysis.

2.3.4. Interpretation of Human EEGs

Hans Berger (1873–1941) was the first person to measure human EEG signals. EEG analysis plays an important role in clinical diagnosis and the electrophysiological analysis of brain functions. It can be used to discover different brain disorders (e.g., AD, epilepsy, and attention deficit hyperactivity disorder-ADHD). However, it can also be used in non-medical applications (e.g., security, gaming and entertainment, and education). Since Berger's discovery, many methods have been developed to analyse the EEG signal. For example, the specialist analyses the EEG based on visual interpretation (the graphic elements), whereas biomedical engineers use mathematical or geometrical analyses to interpret the EEG signal.

Biomedical engineers therefore adhere to the following steps when interpreting the EEG signal:

1. Pre-processing: This step removes the artefacts from the EEG signal (any unwanted part of the EEG). Also, filtering the EEG signal may help to detect EEG features that are significantly associated with specific disorders or

behaviours. Filtering technique has therefore been used to divide the EEG signal into several bands, each of which may have a specific function.

2. Feature extraction: Some important features are extracted from the EEG signal to be used in the next step. These include mean, standard deviation, and so on.
3. Decision making: In this step, the biomedical engineers compare features for two or more groups to differentiate them. Machine learning techniques can be used to classify the groups.

Several brain disorders can be diagnosed through the visual interpretation of EEG signals. Clinicians specialising in brain disorder are familiar with the manifestation of brain rhythms in EEG signals. The amplitudes and frequencies of such signals differ from one state to another, such as sleep and wakefulness. Furthermore, the characteristics of waves also changes with age.

There are five major brain waves (rhythms) that can be discriminated by their different frequency ranges [79][80][81]:

1. Delta *(δ)*: This band lies in the 0-4 Hz range and is associated with deep sleep, although it may be present in the waking state. It is seen in temporal regions during wakefulness, and in a generalised distribution, maximal anteriorly, during drowsiness.
2. Theta *(θ)*: This band lies in the 4-8 Hz range and is primarily observed when consciousness slips towards drowsiness. It is often present in children and young adults during wakefulness, whereas in adults it occurs predominantly during drowsiness. Theta activity may occur in the temporal regions in normal elderly adults during wakefulness.
3. Alpha (*α*): This band lies in the 8-12 Hz range and is found in posterior regions (occipital and parietal). It varies with age. During wakefulness, the alpha

rhythm is present across posterior regions of the head and is maximal with the person relaxed and their eyes closed.

4. Beta (β): This band lies in the 12-30 Hz range and is usually most prominent anteriorly. It often increases during drowsiness.
5. Gamma (γ): This band lies in ranges over 30Hz and is clearly apparent following visual stimuli or just before a movement task. Gamma, along with other activities in the above bands, can be observed at approximately the same time after performing a movement task, such as finger movement, movement of the right toe, or rather broad and bilateral areas of tongue movement.

Figure 2-8 depicts normal brain rhythms of the EEG signal.

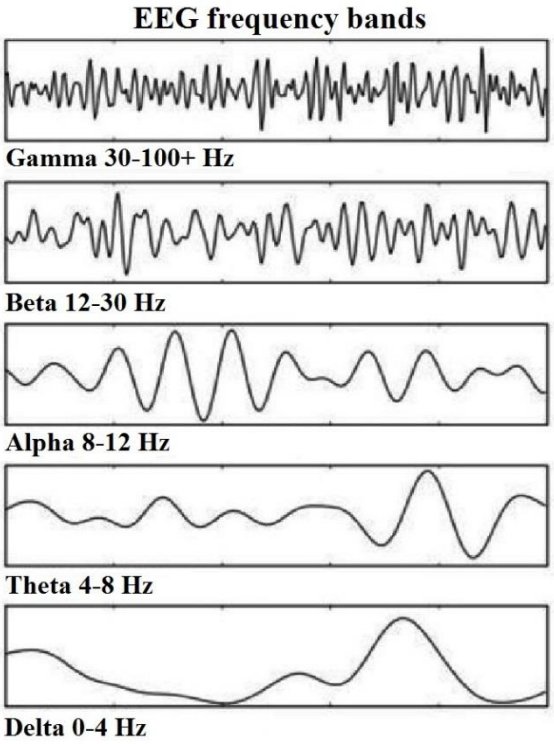

Figure 2-8: Normal brain rhythms of the EEG signal

The above rhythms may persist if the state of the person does not change; therefore, they are approximately cyclic in nature. On the other hand, other brain waveforms exist, which may:

1. Have a wide frequency range or appear as spiky-type signals, such as K-complexes, vertex waves (which occur during sleep), some seizure signals, or a breach rhythm (an alpha-type rhythm arising from a cranial bone defect) which does not respond to movement and is found mainly over the midtemporal region (under electrodes T3 or T4).
2. Be transient, such as an event-related potential (ERP) and contain positive occipital sharp transient (POST) signals (also called rho (ρ) waves).
3. Originate from defective regions of the brain, such as temporal brain lesions.
4. Be spatially localised and ostensibly cyclic in nature yet easily blocked by physical movement such as mu rhythms. Mu denotes motor and is strongly related to the motor cortex. Rolandic (central) mu is related to posterior alpha in terms of amplitude and frequency. However, its topography and physiological significance are quite different. From the mu rhythm, it is possible to investigate cortical functioning and changes in EEG brain (mostly bilateral) activities subject to physical and imaginary movements. The mu rhythm has also been used in feedback training for a range of purposes such as the treatment of epileptic seizure disorder.
5. Phi (ϕ) rhythm (less than 4Hz), which occurs within two seconds of eye closure.
6. Kappa (κ) rhythm, which is an anterior temporal alpha-like rhythm. It is believed to result from discrete lateral oscillations of the eyeballs and is considered an artefact signal.
7. Sleep spindles (also called the sigma (σ) activity) within the 11–15 Hz frequency range.
8. Tau (τ) rhythm, which represents the alpha activity in the temporal region. Eyelid fluttering with closed eyes gives rise to frontal artefacts in the alpha band.

9. Chi (χ) rhythm is a mu-like activity believed to consist of a specific rolandic pattern of 11–17 Hz. This wave has been observed during Hatha Yoga exercises.
10. Lambda (λ) waves are most prominent in waking patients but are not very common. They are sharp transients occurring across the occipital region of the head of walking participants during visual exploration. They are positive and time-locked to saccadic eye movement with varying amplitude, generally below 90 μV.

2.4. Computational and signal processing methods

2.4.1. Introduction

AD is characterized by loss of memory and cognitive decline resulting from damage to brain cells which influence brain activity [48]. AD causes changes in the features of the EEG [45][48][83] and so EEG analysis may provide valuable information about brain dynamics due to AD [23][24][43][45]. The most characteristic features in EEG caused by AD are slowing of EEG, a decrease in EEG coherence, and reduction in EEG complexity [43][44][45][47][48][84][85]. These changes in the EEG can be quantified as a biomarker of AD. A variety of linear and nonlinear methods are being developed to quantify changes in EEG as AD biomarkers [86][87]. AD biomarkers based on the slowing in EEG and a decrease in EEG coherence are often derived using linear analysis methods (i.e., spectral analysis of the EEG signal) [47][88][89]. While biomarkers extracted by analysing the complexity of the EEG are based on nonlinear approaches (e.g., entropy methods, fractal dimension, and Lemple-Ziv complexity). The EEG complexity approaches have shown promising results in AD diagnosis [15][45][49] and appear to be appropriate for AD diagnosis [48][90][91]. Complexity is a measure of the extent to which the dynamic behaviour of a given sequence resembles a random one [92]. The cortical areas of the brain fire spontaneously and this dynamic behavior of the brain is complex [93][94]. AD causes

a reduction in neuronal activity of the brain [95] resulting in decreased capability of the brain to process information [96][97][98] and this may be reflected in the EEG signals [95]. EEG complexity can potentially be a good biomarker for AD diagnosis [48] as AD patients have a significant reduction in EEG complexity [48][84][85][95][99][100].

A number of entropy measures have also been used to derive EEG biomarkers, including Tsallis entropy [101][102], Shannon entropy [99][101], permutation entropy [103][104], and Kolmogorov entropy [105]. In particular, TsEn approach has been shown to be one of the most promising information theoretic methods for quantifying changes in the EEG [102][106]. It has also been shown to be a reliable analysis tool to use with working memory tasks. As its computation is fast, it can serve as a basis for a real-time decision support tool for dementia diagnosis by both specialists and non-specialists [107]. Sneddon et al. [108] investigated TsEn of the EEG and was able to detect mild dementia due to AD with a sensitivity of 88% and specificity of 94%. Bock et al. [102] found TsEn of the EEG to be a highly promising potential diagnostic tool for mild cognitive impairment (MCI) and early dementia with a sensitivity and specificity of 82% and 73%, respectively. Using TsEn approach, Al-nuaimi et al. [45] detected AD from normal subjects with a sensitivity and specificity of 85.8% and 70.9%, respectively. Garn et al. [101], investigated the use of TsEn and Shannon entropy to diagnose AD based on EEG analysis and achieved a p-value<0.0036 for channels T7 and T8 in discriminating between AD patients and normal subjects. McBride et al. [99], found there is a decrease in Shannon entropy in the alpha band in the left parietal region of AD patients. However, Shannon entropy has serious limitations when the time series under consideration is short and noisy such as EEG signal [109]. Deng et al. [103] found that the permutation entropy of AD patients are decreased in contrast with the normal subjects, especially in the theta

band. Mammon et al. [104], suggested the permutation entropy significantly increased in the delta and theta bands specifically in MCI subjects who converted to AD. Hamadicharef et al. [105] used Kolmogorov entropy to detect AD patients with a performance of 82% and 50% for sensitivity and specificity, respectively.

LZC is a nonparametric, non-linear measure of complexity for finite length sequences [110]. It is a simple and powerful method which has been used in several biomedical applications [111]. LZC depends on coarse-grain processing of the measurements [112] and can be applied directly to physiologic signals without pre-processing [113]. LZC has been applied extensively in analysing biomedical signals to obtain a measure of complexity of discrete-time physiological signals [110]. Furthermore, it has been used to analyse brain functions, brain information transmission and EEG complexity in patients with AD [86]. The LZC approach produces a good biomarker for AD detection [113][114]. Hornero et al. [115] used LZC to analyse EEG and magnetoencephalogram (MEG) in AD patients. They found that LZC provides a good insight into the EEG background activity characteristics and the changes associated with AD. Hornero et al. [116] found that LZC values were lower in AD patients and suggested that the most relevant differences are in the posterior region. In addition, they suggested that the MEG activity from AD patients is characterized by a lower degree of irregularity and complexity and that the LZC measures can be used to detect AD with a sensitivity and specificity values of 65% and 76.2%, respectively. McBride et al. [99] analysed EEG complexity based on the LZC method to discriminate between patients with early mild cognitive impairment (MCI), AD patients, and normal subjects. They found EEG complexity features for specific EEG frequency bands with regional electrical activity provide promising results in discriminating between MCI, AD, and normal subjects. Fernandez et al. [117] analysed MEG complexity for MCI patients, AD patients, and normal subjects based

on LZC method for discriminating between the three groups. They found that a combination of age and posterior LZC scores allowed them to distinguish between AD patients and MCI patients with 94.4% sensitivity and specificity.

EEG biomarkers can also be derived using fractal dimension methods, including HFD, and Adapted Box algorithms [118]. Henderson et al [117] found that the Adapted Box algorithms are the most consistent when used in the early detection of AD. HFD is a fast computational method for obtaining the fractal dimension of time series signals [119][120][121] even when very few data points are available [119]. It can track changes in a biosignal from a measure of its complexity [119][120] and it is suited to capturing region-specific neural changes due to AD [51][122]. In addition, HFD provides a more accurate measure of the complexity of signals compared to other methods [119][123][124] and it has been shown to be an efficient method for discriminating between AD patients and normal subjects [41][125]. HFD of the EEG is potentially a good biomarker of AD diagnosis as it is significantly lower in AD patients than in normal subjects [49][126][125]. Smits et al. found that HFD is sensitive to neural changes selectively related to AD patients and normal subjects. Al-nuaimi et al. [49] investigated HFD of EEG for AD diagnosis and found that HFD is a promising EEG biomarker that captures changes in the regions of the brain thought to be affected first by AD and it could be used to detect AD with a sensitivity and specificity values of 100% and 80%, respectively.

2.4.2. Changes in the EEG amplitude (ΔEEG_A)

Changes in the EEG amplitude (ΔEEG_A): ΔEEG_A [8] is used as a measure of the slowing of the EEG. In particular, ΔEEG_A is the sum of the differences between adjacent amplitudes of EEG values per second. The ΔEEG_A calculation of an N-samples data sequence x(1), x(2), ..., x(N) is obtained as,

$$\Delta EEG_A = \frac{\sum \Delta x}{\Delta t} \tag{2.1}$$

where Δx represents the difference between adjacent amplitudes of the EEG in one second and Δt denotes the time interval:

$$\Delta x = x_{i+1} - x_i \tag{2.2}$$

$$\Delta t = t_{i+1} - t_i \tag{2.3}$$

where x_i and x_{i+1} are the current and next EEG amplitude values, respectively, and t_i and t_{i+1} represent the corresponding times i.

ΔEEG_A is first computed using Equation 2.4 for each EEG channel. The mean ΔEEG_A for the channel is then computed as,

$$M_C = \frac{\sum_{i=1}^{N} \Delta EEG_A}{N} \tag{2.4}$$

where M_C is the mean value of ΔEEG_A, and N is the number of samples for the EEG signal.

2.4.3. Zero-crossing intervals (ZCI)

Zero-crossing intervals (ZCI): ZCI [15][126][127][128] is defined, in this context, as the time interval between a positive to negative voltage transition to the next positive to negative voltage transition. It is based on finding a set of instances when the waveform intersects with the time axis. The ZCI calculation for the N-samples EEG signal is obtained as,

$$inst_t = \begin{cases} x_t > 0 \\ x_{t+1} < 0 \end{cases} \tag{2.5}$$

where x_t and x_{t+1} are the times that EEG amplitude changed from positive to a negative value, respectively, and *inst* is the vector that contains the time instances when the amplitude changed from positive to negative value.

$$ZCI = \sum_{i=1}^{K-1} \Delta t \tag{2.6}$$

$$\Delta t = t_{i+1} - t_i \tag{2.7}$$

ZCI is the zero-crossing interval value, k is the indicator for the number of instances, and t_i and t_{i+1} represent the predecessor and successor corresponding to the instances.

2.4.4. Changes in the power spectrum (ΔPS) of EEG signal

Changes in the power spectrum (PS): ΔPS [129][130] biomarker computation is based on the magnitude square of the Fast Fourier transform (FFT) of an N-sample EEG data sequence x(1), x(2), …, x(N)

$$Power_{X(N)} = [|FFT(X(N)|]^2 \tag{2.8}$$

2.4.5. EEG coherence

EEG Coherence (EEG-Coh): Coherence [131][132][133] biomarker computation of an N-sample EEG data sequence x(1), x(2), ..., x(N), is based on the coherence between each two EEG channels and was calculated as,

$$Coh(a,b) = \frac{\left|P_{a,b}\right|^2}{PSD_a * PSD_b} \tag{2.9}$$

where a and b are EEG channels, PSD_a , and PDS_b are the power spectral density for EEG channels a and b, $|P_{a,b}|^2$ is the square cross-spectral density of the channels *a*, and *b*. The EEG coherence is a value between 0 and 1 calculated using the Welch's power spectral density. It represents the functional relationships between the two brain regions.

2.4.6. Tsallis entropy (TsEn)

Tsallis entropy (TsEn): The computation of the TsEn [134] biomarker of an N-samples EEG data sequence x(1), x(2), ..., x(N), is based on the generalised measure of entropy, due to Tsallis:

$$TsEn_q = (\sum_{i=1}^{k} P_i - P_i^q)/(q-1) \tag{2.10}$$

where $TsEn_q$ is the Tsallis entropy value, k is the number of states that the amplitudes of the EEG are quantized into, P_i is a probability associated with the ith state, and q is Tsallis parameter (k=2200, and q=0.5).

2.4.7. Higuchi Fractal Dimension (HFD)

Higuchi fractal dimension (HFD): To compute HFD biomarker [49][119][121][135] of an N-sample EEG data sequence x(1), x(2), ..., x(N), the data is first divided into a k-length sub-data set as,

$$\chi_k^m : x(m), x(m+k), x(m+2k), ..., x(m+\left[\frac{N-m}{k}\right]\cdot k) \tag{2.11}$$

where [] is Gauss' notation, k is constant, and $m=1, 2, \ldots, k$. The length $L_m(k)$ for each sub-data set is then computed as,

$$L_m(k) = \left\{ \left[\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |x(m+ik) - x(m+(i-1)\cdot k)| \right] \frac{N-1}{\left[\frac{N-m}{k}\right]\cdot k} \right\} / k \tag{2.12}$$

The mean of $L_m(k)$ is then computed to find the HFD for the data as,

$$HFD = \frac{1}{K}\sum_{M=1}^{K} L_m(k) \tag{2.13}$$

2.4.8. Approximation Entropy (ApEn)

Approximation entropy (ApEn): ApEn [136][90][137][135] calculation of an N-samples EEG signal, two input constants (m, and r) must be identified to calculate ApEn that

referred as ApEn (m, r, N), where m is the run length and r is the tolerance window. To calculate the ApEn:

Initialise the vector sequences $x_1, x_2, \ldots, x_{(N-m+1)}$, where $xi=[x_i, x_{i+1}, \ldots. , x_{i+m-1}]$, i=1, …, N-m+1. These vectors represent *m* successive x values beginning with the i^{th} point. Then, the distance is defined between x_i and x_j as the maximum differences between successive scalar values. For the x_i, the $N_{m(i)}$ refers the number of j (j=1, …, N-m+1, j≠i) therefore, $d[x_i,x_j] \leq r$. Therefore, for i = 1, …, N – m + 1:

$$C_r^m(i) = \frac{N^m(i)}{N - m + 1} \tag{2.14}$$

where C_r^m values compute the regularity within a tolerance r to the specified window m. Then, compute the average natural logarithm of each C_r^m over i:

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_r^m(i) \tag{2.15}$$

The dimension is increased to *m* + 1 and the previous steps will be repeated to get $C_r^{m+1}$and C_r^{m+1}. The final step the ApEn is defined as,

$$ApEn(m,r,N) = \phi^m(r) - \phi^{m+1}(r) \tag{2.16}$$

The average value of ApEn values was computed for each channel as,

$$Avg(ApEn)_a = \frac{1}{s} \sum_{1}^{s} ApEn(m,r,N) \tag{2.17}$$

where *a* is the channel number, and *S* is the number of subjects in their group (AD or normal).

2.4.9. Lempel Ziv Complexity (LZC)

Lempel-Ziv complexity (LZC): To compute the LZC [92][110][111][86][113] biomarker of an N-sample EEG data sequence x(1), x(2), ..., x(N), the EEG signal is first converted into binary string as,

$$x(i) = \begin{cases} 0 & if \quad EEG(i) < M \\ 1 & if \quad EEG(i) \geq M \end{cases} \tag{2.18}$$

where *x(i)* is the equivalent binary value of EEG(*i*), *i* is the index of all values in the EEG signal, and *M* is the median value of each EEG channel. The median value is used to manage the outliers.

The binary string is then scanned from left to right until the end to produce new substrings. A complexity counter *c(N)* is the number of new substrings. The upper bound of *c(N)* is used to normalise *c(N)* to get an independent value from the sequence of length *N*. The upper bound of *c(N)* is $N/\log_2(N)$. *c(N)* is then normalised by *b(N)* as,

$$C(N) = \frac{c(N)}{b(N)} \tag{2.19}$$

where *C(N)* is the normalised value of the LZC, and *b(N)* is the upper bound of the *c(N)*.

2.5. Main machine learning methodologies

Machine learning is a part of artificial intelligence. It involves programming computers to optimize a performance criterion using example data or past experience [138]. Machine Learning is generally taken to encompass automatic computing procedures based on logical or binary operations, that learn a task from a series of examples [139]. In particular, we define machine learning as a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty [140]. We have a model with defined parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or experience. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data, or both.

Machine learning is usually divided into two main types - supervised and unsupervised learning. Also, there is a third type of machine learning, known as reinforcement learning [140].

Supervised learning involves learning from examples and it is described in more detail in the next section as it is used in a number of studies in this thesis. Unsupervised learning attempts to learn patterns and associations from a set of objects that do not have attached class labels [141]. There is no such supervisor and we only have input data. The aim is to find the regularities in the input. There is a structure to the input space such that certain patterns occur more often than others, and we want to see what generally happens and what does not [142].

Reinforcement learning is the most general form of learning. It tackles the issue of how to learn a sequence of actions called a control strategy from indirect and delayed reward information (reinforcement) [141]. Reinforcement learning fills the gap between supervised learning, where the algorithm is trained on the correct answers given in the target data, and unsupervised learning, where the algorithm can only exploit similarities in the data to cluster it [143].

2.5.1. Supervised learning approach

Supervised learning deals with learning a target function from labelled examples [141]. The objective is to learn a model from a given dataset, $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, and then based on the learned model, to make accurate predictions of y for future values of x [144]. The goal is to learn a mapping from inputs x to outputs y, given a labelled set of input-output pairs

$$D = \{(x_i, y_i)\}_{i=1}^{N} \tag{2.20}$$

where D is called the training set, and N is the number of training examples [140].

In supervised learning, the aim is to learn a mapping from the input to an output whose correct values are provided by a supervisor [142].

In the simplest setting, each training input x_i is a D-dimensional vector of numbers, representing, say, the height and weight of a person. These are called features, attributes or covariates. In general, however, x_i could be a complex structured object, such as an image, a sentence, an email message, a time series, a molecular shape, a graph, etc.

Similarly, the form of the output or response variable can in principle be anything, but most methods assume that y_i is a categorical or nominal variable from some finite set, $y_i \in \{1, \ldots, C\}$ (such as male or female), or that y_i is a real-valued scalar (such as income level). When y_i is categorical, the problem is known as classification or pattern recognition, and when y_i is real-valued, the problem is known as regression. Another variant, known as ordinal regression, occurs where label space Y has some natural ordering, such as grades A–F [140].

Neural networks, fuzzy systems, and support vector machines are typical nonparametric classifiers. Through training using input-output pairs, classifiers acquire decision functions that classify an input datum into one of the given classes [145].

2.5.1.1. Support Vector Machine

Support vector machine (SVM) is a set of related supervised learning methods used for classification and regression. SVM is a classification and regression prediction tool that uses machine-learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data. SVM can be defined as systems which use hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory [146]. Figure 2-9 shows the idea of SVM

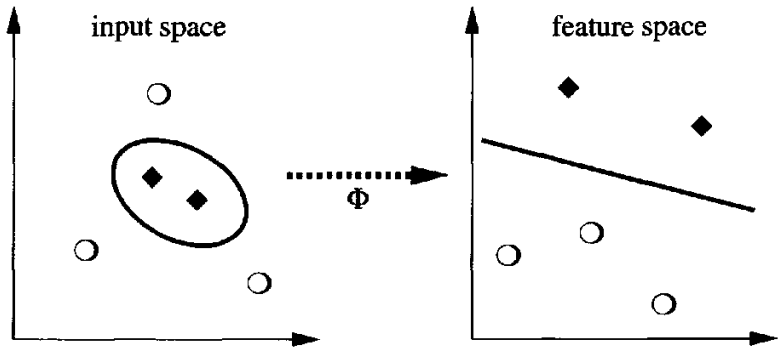

Figure 2-9: The idea of SVM

Figure 2-9 shows the idea of SVM which map the training data into a higher-dimensional feature space via ϕ, and construct a separating hyperplane with maximum margin there. This yields a nonlinear decision boundary in input space[147]. The goals of SVM are separating the data with hyperplane and extend this to non-linear boundaries using kernel trick [148]. For calculating the SVM we see that the goal is to correctly classify all the data [146].

To compute the linearly separable data [149], consider the problem of binary classification. Training data are given as,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l), X \in R^n, Y \in \{+1, -1\} \quad (2.21)$$

For reasons of visualization only, we will consider the case of two-dimensional input space, i.e., $X \in R^2$. Data are linearly separable and there are many different hyperplanes that can perform separation, as shown in figure 2-10. For $X \in R^2$, the separation is performed by,

$$w_1x_1 + w_2x_2 + b = 0 \quad (2.22)$$

In other words, the decision boundary, i.e., the separation line in input space is defined by the equation

$$w_1x_1 + w_2x_2 + b = 0 \quad (2.23)$$

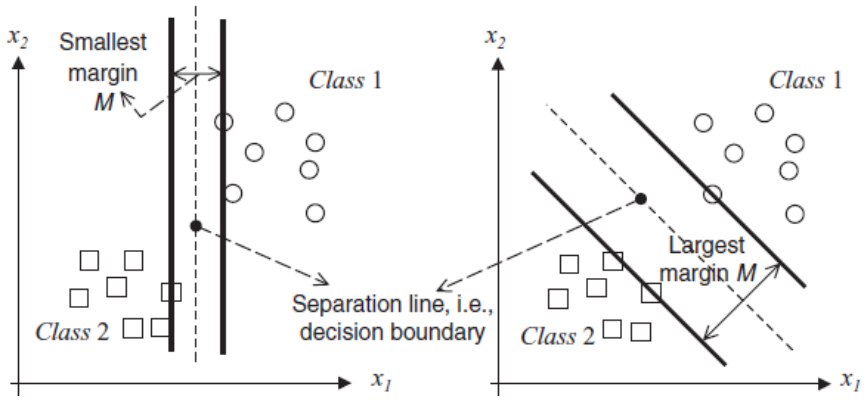

Figure 2-10: Two-out-of-many separating lines: a good one with a large margin (right)

As shown in Figure 2-10, the dashed separation line shown in the right graph seems to promise good classification. It seems to probably be better in generalization than the dashed decision boundary having smaller margin shown in the left graph.

By using given training examples, during the learning stage, our machine finds parameters $W = [w_1 w_2 \ldots w_n]^T$ and b of a discriminant or decision function d(x,w,b) given as

$$d(x,w,b) = w^T x + b = \sum_{i=1}^{n} w_i \, x_{i+b} \tag{2.24}$$

where $X,W \in R^n$, and the scalar b is called a bias. (Note that the dashed separation lines in figure 2-10 represent the line that follows from d(x,w,b) = 0). After the successful training stage, by using the weights obtained, the learning machine, given previously unseen pattern x_p, produces output 0 according to an indicator function (i_F) given as

$$i_F = o = \text{sign}(d(xp,w,b)) \tag{2.25}$$

where o is the standard notation for the output from the learning machine. In other words the decision rule is:

if d(xp,w, b) > 0, the pattern x_p belongs to a class 1 (i.e., o = y1 = +1) , and

if, d(xp,w, b) < 0 the pattern x_p belongs to a class 2 (i.e., o = y2 = −1) [150].

2.5.1.2. Linear Discriminant Analysis

Linear discriminant analysis (LDA) [142] is a supervised method for dimensionality reduction for classification problems. We start with the case where there are two classes, then generalize to K > 2 classes. Given samples from two classes Cl and C2, we want to find the direction, as defined by a vector w, such that when the data are projected onto w, the examples from the two classes are as well separated as possible when,

$$Z = W^T X \tag{2.26}$$

is the projection of X onto W and thus is a dimensionality reduction from d to 1. Figure 2-11 shows the two-dimensional, two-class data projected on W.

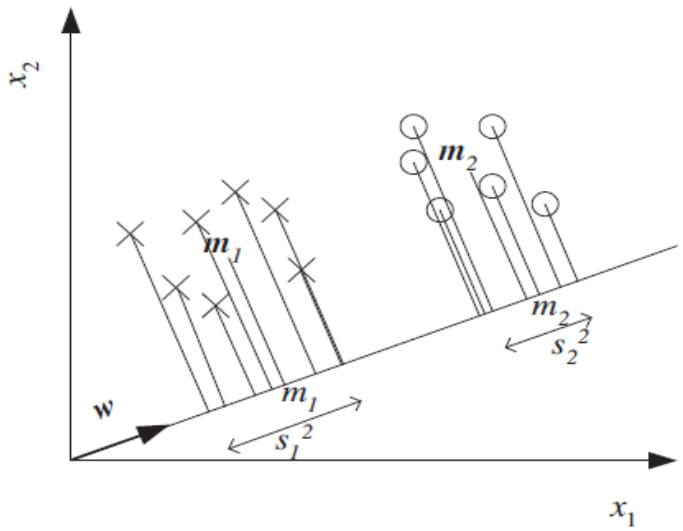

Figure 2-11: Two-dimensional, two-class data projected on W

m_1 and m_1 are the means of samples from C1 before and after projection, respectively. Note that $m_1 \in d$ and $m1 \in R^d$. We are given a sample $X = \{x^t, r^t\}$ such that $r^t = 1$ if $x^t \in C1$ and $r^t = 0$ if $x^t \in C2$. Where,

$$m1 = \frac{\sum_t w^T x^t r^t}{\sum_t r^t} = w^T m_1 \tag{2.27}$$

$$m2 = \frac{\sum_t w^T x^t (1-r^t)}{\sum_t (1-r^t)} = w^T m_2 \tag{2.28}$$

2.5.1.3. K-nearest neighbours

The K-nearest neighbour (KNN) algorithm is a supervised learning algorithm which classifies a new instance based upon some distance formula (e.g., Euclidean). The new instance is classified to a category relative to some majority of K-nearest neighbours. Traditionally, the KNN algorithm is viewed as a machine learner, rather than a method for dividing training and test data [141].

Nearest neighbour classification is quite straightforward; examples are classified based on the class of their nearest neighbours. It is often useful to take more than one neighbour into account so the technique is more commonly referred to as k-Nearest Neighbour (KNN) Classification where k nearest neighbours are used in determining the class. KNN classification has two stages; the first is the determination of the nearest neighbours and the second is the determination of the class using those neighbours. The k-nearest neighbours are selected based on this distance metric. KNN classification has two stages; the first is the determination of the nearest neighbours and the second is the determination of the class using those neighbours [151], as shown in Figure 2-12.

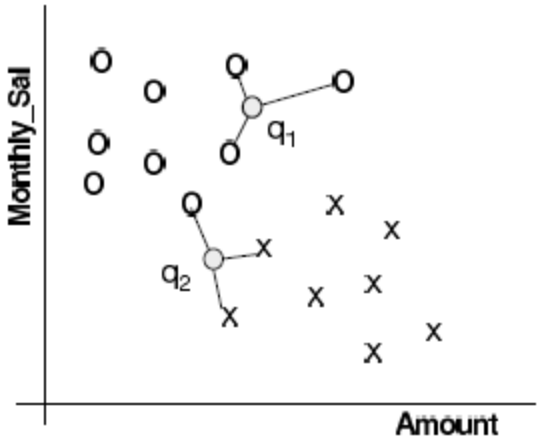

Figure 2-12: Example of 3-nearest neighbour classification

Let us assume that we have a training dataset D made up of $(x_i)_{I \in [1,|D|]}$ training samples. The examples are described by a set of features F and any numeric features have been normalised to the range [0,1]. Each training example is labelled with a

class label $y_j \in Y$. Our objective is to classify an unknown example q. For each $x_i \in D$ we can calculate the distance between q and x_i as,

$$d(q, x_i) = \sum_{f \in F} w_f \delta(q_f, x_{if}) \quad (2.29)$$

2.5.2. Some basic concepts in machine learning

Pattern classification is to classify some object into one of the given categories called classes. For a specific pattern classification problem, a classifier, which is computer software, is developed so that objects are classified correctly with reasonably good accuracy. Inputs to the classifier are called features because they are determined so that they represent each class well or so that data belonging to different classes are well separated in the input space [145].

The complexity of any classifier depends on the number of inputs. This determines both the time and space complexity and the necessary number of training examples to train such a classifier [152].

2.5.2.1. Dimensionality reduction

When dealing with high dimensional data, it is often useful to reduce the dimensionality by projecting the data to a lower dimensional subspace which captures the “essence” of the data. This is called dimensionality reduction [140].

There are two main methods for reducing dimensionality: feature selection and feature extraction. In feature selection, we are interested in finding n of the d dimensions that give us the most information and we discard the other (d - n) dimensions [152]. Feature selection, also known as variable selection, is the process of choosing relevant components of the vector X for use in model construction [153]. While, in feature extraction, we are interested in finding a new set of n dimensions that are the combination of the original d dimensions. These methods may be supervised or unsupervised depending on whether or not they use the output information. The best known and most widely used feature extraction methods are

Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA), which are both linear projection methods, unsupervised and supervised respectively [152].

2.5.2.2. Training, testing, and validation sets

Three sets of data (i.e. for training, testing, and validation) are needed, as shown in Figure 2-13. The training set is used to train the algorithm [143] with the goal of finding the training parameters that result in the best performance [154]. The validation set to keep track of how well it is doing as it learns [143]. It is a set of samples used to find the best training parameters [154], and the test set to produce the final results, and each algorithm is going to need some reasonable amount of data to learn [143].

The dataset is randomly partitioned into K subsets, and one subset is used as a validation set, while the algorithm is trained on all of the others. A different subset is then left out and a new model is trained on that subset, repeating the same process for all of the different subsets. Finally, the model that produced the lowest validation error is tested and used [143].

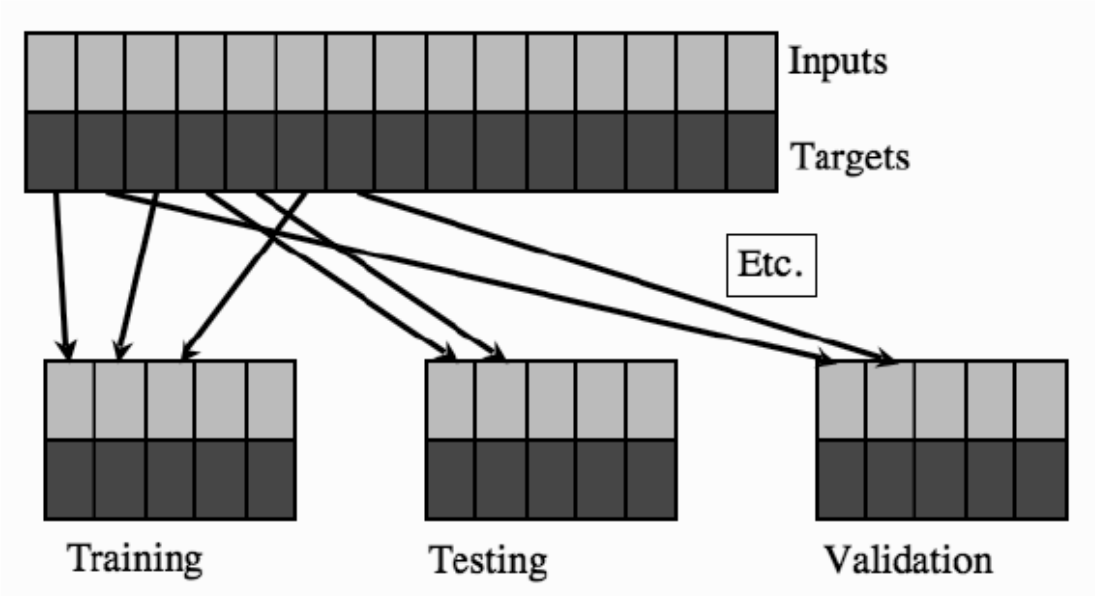

Figure 2-13: The dataset is split into different sets, some for training, and some for validation

2.5.2.3. K-Fold Cross-Validation

K-fold cross validation is a resampling method for validating a model. For this technique, data is partitioned into k-classes, and k models are constructed with each of the k-classes rotated into the test set. K-fold cross validation addresses the issue of data distribution between training and test sets but does not consider the difficulty in modelling the training data [141].

In K-fold cross-validation, the dataset X is divided randomly into K equal sized parts, X_i, i = 1, ... , K. To generate each pair, we keep one of the K parts out as the validation set, and combine the remaining K - 1 parts to form the training set. Doing this K times, each time leaving out another one of the K parts out, we get K pairs [152]:

$V_1=X_1$, $T1=X_2, X_3, X_4 \ldots, X_k$

$V_2=X_2$, $T2=X_1, X_3, X_4 \ldots, X_k$

$V_3=X_3$, $T1=X_1, X_2, X_4 \ldots, X_k$

2.6. Diagnostic performance measures

This thesis aims to evaluate and develop EEG based biomarkers to detect AD in its early stages. The evaluation of current biomarkers and the development of novel biomarkers means their respective performance needs to be measured. Such measurement comprises a set of technical specifications (e.g., mathematical or statistical measurements) that define several important factors pertaining to the quality of diagnosis [155]. This enhances the quality of decisions made by biomarkers, provides feedback to clinicians, and represents an indicator of the quality of the diagnosis [156]. To facilitate a high quality diagnosis, multiple performance measures can be computed to provide an extensive and complete view of overall performance [156][157]. Finally, quantifying performance is an important step in improving the quality of healthcare [158].

Clinically, diagnostic biomarkers are evaluated according to the following factors; sensitivity, specificity, p-value, accuracy, positive predictive value, negative predictive value, f-measure, Matthew's correlation coefficient, receiver operating characteristic, and area under the curve. The terms positive and negative are traditionally defined in that a person is "positive" if they have the disease in question and "negative" if they do not, and a "true" diagnosis is a correct diagnosis while a "false" diagnosis is incorrect. Moreover, true positive (TP) denotes the number of AD patients who were diagnosed correctly, false positive (FP) denotes the number of AD patients who were diagnosed incorrectly, true negative (TN) denotes the number of normal people who were diagnosed correctly, and false negative (FN) denotes the number of normal people who were diagnosed incorrectly.

2.6.1. Accuracy

Accuracy refers to the ratio between the correct number of classified cases and the total number of all cases [122]. It measures how close the classification is to the

correct classification [159]. Accuracy (correct classification) as a performance measure is insufficient because it depends on the true classification and disregards false classifications [122]. Accuracy can be computed as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.30}$$

2.6.2. Sensitivity and specificity

To resolve the drawbacks associated with using accuracy as a performance measure, two measurements can be used to describe performance; sensitivity and specificity. Sensitivity denotes the probability of correctly predicting positive cases, while specificity denotes the probability of correctly predicting negative cases [160]. Sensitivity and specificity represent how well a measure recognises a true positive and true negative, respectively [122][161]. Sensitivity and specificity can be computed as follows:

$$Sensitivity = \frac{TP}{TP+FN} \tag{2.31}$$

$$Specificity = \frac{TN}{TN+FP} \tag{2.32}$$

2.6.3. False positive rate and false negative rate

False positive rate (FPR) and false negative rate (FNR) indicate the rates at which these errors occur in the classification. FPR indicates the rates with which the classifier makes an error by classifying negative cases as positive (e.g., normal but classified as AD). While, FNR indicates the rates with which the classifier makes an error by classifying positive cases as negative (e.g., AD but classified as normal) [162][163][164]. FPR and FNR can be computed as,

$$FPR = \frac{FP}{FP+TN} \tag{2.33}$$

$$FNR = \frac{FN}{FN+TP} \tag{2.34}$$

2.6.4. Positive predictive value (precision), and negative predictive value

The main advantage of a diagnostic test is that it can be utilised in diagnosis. Sensitivity and specificity do not provide clinicians with the probability of the disease being present in a patient. The probability of the test provides this information, and thus facilitates successful diagnoses. Therefore, there is a need to direct the diagnostic test towards predictive values.

The predictive values of the test rely on the prevalence of abnormality in the patients being tested. This can be defined as the probability that a patient has the disease before the test is performed. The positive and negative predictive values (PPV, NPV) are purified estimates of the same probability that people will be given the correct diagnosis [165][166].

PPV represents the probability that a person has a disease and is correctly classified. Mathematically, it is the ratio between the number of true positives and the number of true positives plus the number of false positives. Conversely, NPV represents the probability that a person does not have the disease and is correctly classified. Mathematically, it is the ratio between the number of true negatives and the number of true negatives plus the number of false negatives [167]. PPv and NPV can therefore be calculated as follows:

$$PPV = \frac{TP}{TP+FP} \tag{2.35}$$

$$NPV = \frac{TN}{TN+FN} \tag{2.36}$$

2.6.5. F-measure and Matthew's correlation coefficient

F-measure is a harmonic mean of positive predictive value (precision-P) and sensitivity (recall-R) [168][169]. The two diagnostic measures (P and R) are combined to construct their harmonic mean [170]. It therefore represents a standard balance between precision and recall when evaluating the performance of classifiers. However, the f-measure is an alternative version of the area under curve (AUC) [171]

and provides a better way to evaluate classifiers because it can assign errors for positive and negative results [169]. The F-measure can be computed as follows:

$$F - measure = \frac{2PR}{P+R} \tag{2.37}$$

Matthew's correlation coefficient (MCC) is a correlation coefficient between the actual and predicted results in the binary classification with values ranging from -1 to 1. A correlation coefficient of 1 indicates that the prediction was completely correct while a value of -1 indicates that a completely opposite prediction was made [172]. MCC is used to measure the quality of the binary classification in machine learning by comparing the actual and predicted results [173][174]. MCC can be computed as follows:

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)*(FP+FN)*(TN+FP)*(TN+FN)}} \tag{2.38}$$

2.6.6. Receiver operating characteristic and area under the curve

Receiver Operating Characteristic (ROC) analysis provides tools to distinguish classifiers that are optimal under some class and cost distributions from classifiers that are always sub-optimal, and to select the optimal classifier once the cost parameters are known [175]. ROC graphs are a useful technique for organizing classifiers and visualizing their performance. ROC graphs are commonly used in medical decision making, and in recent years have been increasingly adopted in the machine learning and data mining research communities [176]. One of the main advantages of ROC analysis is that it is threshold-agnostic; i.e. the performance of a predictor is estimated without a specific threshold and also gives a criteria to choose an optimal threshold based on a certain cost function or objective. Typically, an ROC analysis shows how sensitivity (true positive rate) changes with varying specificity (true negative rate or 1−false positive rate). Analyses also typically weigh false positives and false negatives equally [177]. ROC analysis for two classes is based on plotting the true positive rate (TPR) on the y-axis and the false-positive rate (FPR) on the x-axis. This gives a point for each classifier. A curve is obtained because, given

two classifiers, we can obtain as many derived classifiers as we want along the segment that connects them, just by voting them with different weights. Consequently, any point "below" that segment will have greater cost for any class distribution and cost matrix, because it has lower TPR and/or higher FPR. According to that property, given several classifiers, one can discard the classifiers that fall under the convex hull formed by the points representing the classifiers and the points (0,0) and (1,1), which represent the default classifiers always predicting negative and positive, respectively [175].

Figure 2-14 shows an example of a ROC "curve" on a test set of twenty instances created by thresholding a test set. It shows the corresponding ROC curve with each point labelled by the threshold that produces it. Table 2-1 shows ten positive and ten negative instances and the score assigned to each by a scoring classifier. Any ROC curve generated from a finite set of instances is actually a step function, which approaches a true curve as the number of instances approaches infinity [176].

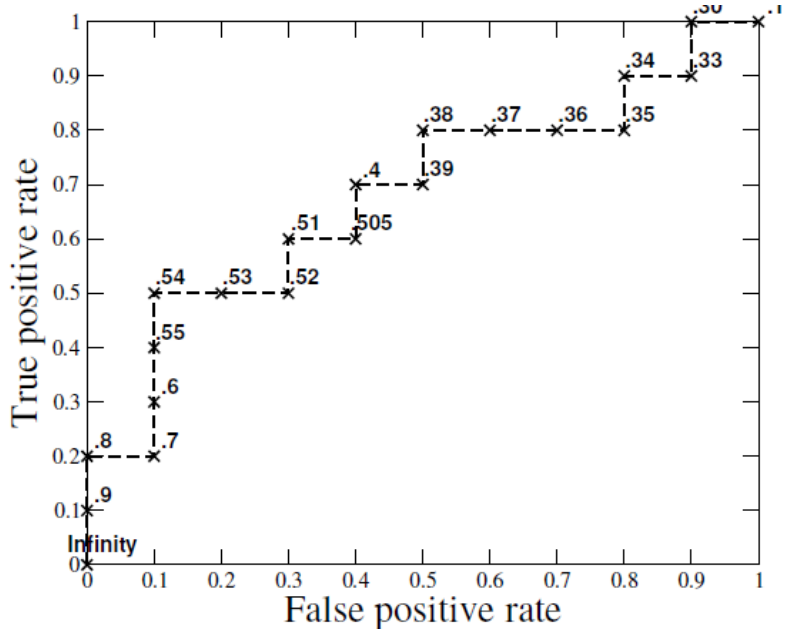

Figure 2-14: The ROC "curve" created by thresholding a test set

Table 2-1: Shows twenty data and the score assigned to each by a scoring classifier

Inst#	Class	Score	Inst#	Class	Score
1	p	0.9	11	p	0.4
2	p	0.8	12	n	0.39
3	n	0.7	13	p	0.38
4	p	0.6	14	n	0.37
5	p	0.55	15	n	0.36
6	p	0.54	16	n	0.35
7	n	0.53	17	p	0.34
8	n	0.52	18	n	0.33
9	p	0.51	19	p	0.3
10	n	0.505	20	n	1

In ROC analyses, the predictive performance is commonly summarized by the area under the curve (AUC), which can be found by integrating areas under the line segments [177]. The AUC, estimated simply from the ranks of the predicted scores, can remain unchanged even when the predicted scores change. This can lead to a loss of useful information, and may produce suboptimal results [178]. A ROC curve is a two-dimensional depiction of classifier performance. To compare classifiers we may want to reduce ROC performance to a single scalar value representing expected performance. A common method is to calculate the area under the ROC curve, abbreviated AUC. Since the AUC is a portion of the area of the unit square, its value will always be between 0 and 1.0. However, because random guessing produces the diagonal line between (0, 0) and (1, 1), which has an area of 0.5, no realistic classifier should have an AUC less than 0.5. The AUC has an important statistical property: the AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [176].

$V_k=X_k$, $T1=X_1, X_2, X_4\ldots, X_{k-1}$

2.7. Null and alternative hypotheses

The null hypothesis, denoted by H_0, is the hypothesis that is to be tested. The alternative hypothesis, denoted by H_1, is the hypothesis that in some sense contradicts the null hypothesis [179][180].

The null hypothesis of this study is that EEG based biomarkers can not be used to detect AD patients with high performance. *The alternative hypothesis* of the study is that EEG based biomarkers can be used to detect AD with high performance. P-value was used in this study to test the null hypothesis. It was computed between the EEG biomarkers for AD patients and normal subjects for different analysis methods were used in this study e.g., between the Tsallis entropy values for AD patients and normal subjects.

P-values

P-value is a common statistical measure that is used in biomedical research to facilitate decisions regarding the independence of two different states [181]. It refers to the probability of obtaining results which show the null hypothesis is true [182]. The null hypothesis specifies that no association exists between the factors under investigation [181][122][183]. The p-value ranges between zero and one [184]. A p-value that satisfies the threshold of significance is less than or equal to 0.05 [182][185].

P-values are computed using t-tests [186][187][188]. A t-test can also be computed using Welch's t-test rather than the Student's t-test [184][189]. To compute the p-value between two data samples (e.g., X1, and X2), the following equation is used:

$$P = 2 * Tcdf(-|t|, V) \quad (2.39)$$

where *Tcdf* is the Student's T cumulative distribution with *V* degrees of freedom of the values in X, and *t* is the Welch's *t*-test which is calculated as shown in Equation 2.40.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} - \frac{s_2^2}{n_2}}} \quad (2.40)$$

where n_1, and n_2 are the size of X_1 and X_2, while, S_1, and S_2 are the variances of X_1, and X_2, respectively.

The formula to calculate the degrees of freedom is known as the Welch-Sattertwaite-equation and is computed as follows:

$$V = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1) + \left(s_2^2/n_2\right)^2/(n_2-1)} \quad (2.41)$$

Welch's t-test divides the variance in the group by its size. Conversely, the Student's t-test uses the standard deviations of each group [187].

2.8. Multiple comparisons problem

When multiple tests are performed on a dataset, it is likely to reject the null hypothesis when it is true. This is a consequence of the logic of hypothesis testing; reject the null hypothesis if we find a rare event. But the larger the number of tests, the easier it is to find rare events and therefore the easier it is to make the mistake of thinking that there is an effect when there is none [190][191]. When rejecting a null hypothesis because a p-value is less than your critical value, it is possible this is wrong; the null hypothesis might really be true, and your significant result might be due to chance. A p-value of 0.05 means that there is a 5% chance of getting your observed result if the null hypothesis were true. It does not mean that there is a 5% chance that the null hypothesis is true [192].

Bonferroni correction

There are several approaches for addressing multiple comparisons concerns. Bonferroni correction is a standard method to deal with multiple comparisons problems in statistics. Bonferroni correction involves dividing the p-value threshold

by the number of comparisons [193]. The Bonferroni correction is an adjustment made to p-values when several dependent or independent statistical tests are being performed simultaneously on a single data set [194]. The p-values so obtained can then be corrected for multiple comparisons [195].

The classic approach to the multiple comparison problem is to control the familywise error rate. Instead of setting the critical P-level for significance, or alpha, to 0.05, a lower alpha is used. If the null hypothesis is true for all of the tests, the probability of getting one result that is significant at this new; lower-alpha level is 0.05. In other words, if the null hypotheses are true, the probability that the family of tests includes one or more false positives due to chance is 0.05. The most common way to control the familywise error rate is with the Bonferroni correction. The significance level (alpha) for an individual test is found by dividing the familywise error rate (usually 0.05) by the number of tests. Thus if 100 statistical tests are being performed, the alpha level for an individual test would be 0.05/100=0.0005, and only individual tests with P<0.0005 would be significant [192].

The following steps summarise the computation of the Bonferroni-corrected P-values

1. Let N be the number of tests
2. Compute the critical p-value

   This is found by dividing the familywise error rate (usually 0.05) by the number of tests, i.e. 0.05/N
3. If (P-values* number of tests)< 1 Then

   Bonferroni-corrected P-values=P-values* number of tests

   Else

   Bonferroni-corrected P-values=1
4. P-value is significant if it is less than the critical p-value

The following example highlights the Bonferroni correction computation for N = 11.

Critical p-value= 0.0045 (0.05/11)

Table 2-2: Bonferroni-corrected P-value computation

P-values	Bonferroni-corrected significance	Bonferroni-corrected P-value
0.001	Significant	0.0110
0.008	Not significant	0.0880
0.039	Not significant	0.4290
0.041	Not significant	0.4510
0.042	Not significant	0.4620
0.002	Significant	0.0220
0.074	Not significant	0.8140
0.205	Not significant	1.0000
0.003	Significant	0.0330
0.216	Not significant	1.0000
0.004	Significant	0.0440

Chapter 3. Materials and Methods

3.1. Electroencephalogram Datasets

In this section, we provide detailed descriptions of each dataset including information about the cohort, recording information and so on. Detailed descriptions of the recording format which is specific to each dataset are also given. Four EEG datasets were used in this study e.g., B, C, and D.

3.1.1. Dataset A

This data set consists of 11 cases: nine normal subjects, and three AD patients.

3.1.1.1. Data origin

Data were recorded in Derriford Hospital and a memory clinic in Plymouth, UK as part of an existing protocol. All subjects are from UK.

3.1.1.2. Diagnostic criteria and cohort information

All patients were referred to the EEG department in the hospital from a specialist memory clinic. Prior to referral, all patients underwent a battery of psychometric tests (including the MMSE [196], The Rey Auditory Verbal Learning Test [197], Benton Visual Retention Test [198] and memory recall tests [199]). All normal subjects were healthy volunteers and had normal EEGs, which were reviewed and confirmed by the consultant clinical neurophysiologist. Two of the normal subjects were young subjects (aged < 10 years). Each of these subjects had three EEG recordings. The remaining were age-matched (> 65 years old). MRI data was not recorded because this facility was not available at the hospital at that time. The ethnicities of subjects were unknown. Details of the cohort are summarised in Table 3-1.

Table 3-1: Cohort information for Dataset A

No.	Subject	Age (yrs)	Diagnosis	Clinical Notes	Comments/ Notes
1	AD1	-	Prob. AD	Probable AD	-
2	AD2	-	Prob. AD	Probable AD	-
3	AD3	-	Prob. AD	Probable AD	-
4	Vol1	> 65	Normal	Normal who went on to develop Alzheimer's	Used as a pre-AD case
5	Vol2		Normal	-	-
6	Vol3		Normal	-	-
7	Vol4		Normal	-	-
8	Vol5		Normal	-	-
9	Vol6		Normal	-	-
10	Vol7		Normal	-	-
11	Vol8		Normal	-	-

3.1.1.3. Recording information

All subjects were recorded at rest comprising of various states: awake, drowsy, alert with periods of eyes closed and open. The EEGs were obtained using the 10-20 system. Details of the recording are given in Table 3-2 recording information for dataset A.

Table 3-2: Recording information for Dataset A

Description	Details
Montage	Common Reference Montage
Sampling Frequency (Hz)	256, downsampled to 128 for storage reasons
Number of Channels	21
Channel Labels	Fp1 = 1, Fp2 = 2, F7 = 3, F3 = 4, FZ = 5, F4 = 6, F8 = 7, A1 = 8, T3 = 9, C3 = 10, CZ = 11, C4 = 12, T4 = 13, A2 = 14, T5 = 15, P3 = 16, PZ = 17, P4 = 18, T6 = 19, O1 = 20, O2 = 21
Duration (sec)	240
Equipment Settings	Not Applicable

3.1.2. Dataset B

This data set consists of 41 cases: 21 normal subjects, and 17 mild AD patients.

3.1.2.1. Data origin

As with Dataset A, the data in this dataset were recorded in Derriford Hospital and a memory clinic in Plymouth, UK as part of an existing protocol. All subjects were from UK.

3.1.2.2. Diagnostic criteria and cohort information

All patients were referred to the EEG department in the hospital from a specialist memory clinic. A battery of psychometric tests (including the MMSE [196], The Rey Auditory Verbal Learning Test [197], Benton Visual Retention Test [198] and memory recall tests [199]) were performed on all patients at the memory clinic. The classification of subjects with dementia was based on the working diagnosis provided by the specialist memory clinic. All healthy volunteers and AD patients had their EEG confirmed by a consultant clinical neurophysiologist at the hospital as normal and probable mild AD respectively. MRI data was not recorded because this facility was not available at the hospital at the time. The ethnicities of subjects were unknown. Details of the cohort are summarised in Table 3-3.

Table 3-3: Cohort information for Dataset B

No.	Subject	Age (yrs)	Gender	Diagnosis	Clinical Notes
1	AD1	77.6 ± 10, max.= 93,min. = 50	9M, 8F	Prob. AD	Probable AD
2	AD2			Prob. AD	Probable AD
3	AD3			Prob. AD	Probable AD
4	AD4			Prob. AD	Probable AD
5	AD5			Prob. AD	Probable AD
6	AD6			Prob. AD	Probable AD
7	AD7			Prob. AD	Probable AD
8	AD8			Prob. AD	Probable AD
9	AD9			Prob. AD	Probable AD
10	AD10			Prob. AD	Probable AD
11	AD11			Prob. AD	Probable AD
12	AD12			Prob. AD	Probable AD
13	AD13			Prob. AD	Probable AD
14	AD14			Prob. AD	Probable AD
15	AD15			Prob. AD	Probable AD
16	AD16			Prob. AD	Probable AD
17	AD17			Prob. AD	Probable AD
18	NORM1	69.4 ± 11.5,	10M, 14F	Normal	-
19	NORM2			Normal	-
20	NORM3	max. = 84, min. = 40		Normal	-
21	NORM4			Normal	-
22	NORM5			Normal	-
23	NORM6			Normal	-
24	NORM7			Normal	-
25	NORM8			Normal	-
26	NORM9			Normal	-
27	NORM10			Normal	-
28	NORM11			Normal	-
29	NORM12			Normal	-
30	NORM13			Normal	-
31	NORM14			Normal	-
32	NORM15			Normal	-
33	NORM16			Normal	-
34	NORM17			Normal	-
35	NORM18			Normal	-
36	NORM19			Normal	-
37	NORM20			Normal	-
38	NORM21			Normal	-
39	NORM22			Normal	-
40	NORM23			Normal	-
41	NORM24			Normal	-

3.1.2.3. Recording information

All subjects were recorded at rest comprising of various states: awake, drowsy, alert with periods of eyes closed and open. The EEGs were obtained using the 10-20 system. Details of the recording are given in Table 3-4.

Table 3-4: Recording information for Dataset A

Description	Details
Montage	Common Reference Montage
Sampling Frequency (Hz)	256, downsampled to 128 for storage reasons
Number of Channels	21
Channel Labels	Fp1 = 1, Fp2 = 2, F7 = 3, F3 = 4, FZ = 5, F4 = 6, F8 = 7, A1 = 8, T3 = 9, C3 = 10, CZ = 11, C4 = 12, T4 = 13, A2 = 14, T5 = 15, P3 = 16, PZ = 17, P4 = 18, T6 = 19, O1 = 20, O2 = 21
Duration (sec)	240
Equipment Settings	Not Applicable

3.1.3. Dataset C

This dataset consists of 20 cases: 10 normal subjects, and 10 mild AD patients.

3.1.3.1. Data origin

The data in this dataset were recorded in Rome and were provided by La Sapienza, University of Rome. Local institutional ethics committees approved the study and all experiments were performed with the informed and overt consent of each participant or caregiver, in line with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by the Author's Institutional Review Board. All subjects were from Italy.

3.1.3.2. Diagnostic criteria and cohort information

The data were collected from 10 mild AD patients (78.3 + 4.03 years) and 10 normal age-matched healthy old subjects (78 + 4.24 years). Probable AD was diagnosed according to NINCDS-ADRDA [200] and DSM IV criteria. The recruited AD patients underwent general medical, neurological and psychiatric assessments. Patients were also rated with a number of standardized diagnostic and severity instruments that included MMSE [196], Clinical Dementia Rating Scale (CDRS) [201], Geriatric Depression Scale (GDS) [202], Hachinski Ischemic Scale (HIS) [203], and

Instrumental Activities of Daily Living scale (IADL) [204]. Neuroimaging diagnostic procedures (CT or MRI) and complete laboratory analyses were carried out to exclude other causes of progressive or reversible dementias, in order to have a homogenous mild AD patient sample. Exclusion criteria included, in particular, any evidence of (i) fronto-temporal dementia, (ii) VAD (i.e. the VaD was also diagnosed according to NINDS-AIREN criteria; [205], (iii) extra-pyramidal syndromes, (iv) reversible dementias (including pseudo-dementia of depression), and (v) fluctuations in cognitive performance (suggestive of a possible Lewy body dementia). The normal control subjects were recruited mainly among patients' spouses. All normal subjects underwent physical and neurological examinations as well as cognitive screening (including MMSE). Subjects affected by chronic systemic illnesses (i.e. diabetes mellitus or organ failure) were excluded, as were subjects receiving psychoactive drugs. Subjects with a history of present or previous neurological or psychiatric disease were also excluded. All normal subjects had a GDS score lower than 14. The ethnicities of the subjects were unknown. Details of the cohort are summarised in Table 3-5.

Table 3-5: Cohort information for Dataset C

No.	Subject	Age (yrs)	Gender	Education (yrs)	MMSE	Diagnosis
1	AD-s01	78	F	2	21	Prob. AD
2	AD-s02	77	M	9	27	Prob. AD
3	AD-s03	78	F	5	21.7	Prob. AD
4	AD-s04	83	M	13	23.1	Prob. AD
5	AD-s05	78	F	5	18.7	Prob. AD
6	AD-s06	72	F	5	23.3	Prob. AD
7	AD-s07	77	F	5	21.7	Prob. AD
8	AD-s08	80	M	13	22.1	Prob. AD
9	AD-s09	74	M	18	25.7	Prob. AD
10	AD-s10	86	M	8	23.8	Prob. AD
11	Nold-s01	85	F	5	30	Normal
12	Nold-s02	77	M	13	30	Normal
13	Nold -s03	78	M	5	28.7	Normal
14	Nold -s04	71	F	8	28.4	Normal
15	Nold -s05	72	M	5	29.3	Normal
16	Nold -s06	83	M	13	29.1	Normal
17	Nold -s07	78	F	5	28.7	Normal
18	Nold -s08	78	F	5	28.7	Normal
19	Nold -s09	79	F	5	29	Normal
20	Nold -s10	79	M	18	30	Normal

3.1.3.3. Recording Information

All subjects were recorded at rest with eyes closed. The EEGs were obtained using the 10-20 system. Details of the recordings are given in Table 3-6.

Table 3-6: Recording information for Dataset C

Description	Details
Montage	Common Reference Montage
Sampling Frequency (Hz)	128
Number of Channels	19
Channel Labels/Order	Fp1 = 1, Fp2 = 2, F7 = 3, F3 = 4, FZ = 5, F4 = 6, F8 = 7, T3 = 8, C3 = 9, CZ = 10, C4 = 11, T4 = 12, T5 = 13, P3 = 14, PZ = 15, P4 = 16, T6 = 17, O1 = 18, O2 = 19. Channels A1 and A2 were not used.
Duration (sec)	60
Equipment Settings	Not Applicable

3.1.4. Dataset D

This dataset consists of 40 cases: 20 normal subjects, and 20 mild AD patients.

3.1.4.1. Data origin

As with Dataset C, this dataset was recorded in Rome and was provided by La Sapienza, University of Rome. The study was approved by the local institutional ethics committees and all experiments were performed with the informed and overt consent of each participant or caregiver, in line with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by the Author's Institutional Review Board. All subjects were from Italy.

3.1.4.2. Diagnostic criteria and cohort information

The data were collected from 20 normal age-matched healthy old subjects (78.5 + 3.85 years), 20 MCI patients (74.1 + 5.47 years) and 20 mild AD patients (77.8 + 5.50 years). Probable AD was diagnosed according to NINCDS-ADRDA [200] and DSM IV criteria. The recruited AD patients underwent general medical, neurological and psychiatric assessments. Patients were also rated with a number of standardized diagnostic and severity instruments that included MMSE [196], Clinical Dementia Rating Scale (CDRS) [201], Geriatric Depression Scale (GDS) [202], Hachinski Ischemic Scale (HIS) [203], and Instrumental Activities of Daily Living scale (IADL) [204]. Neuroimaging diagnostic procedures (CT or MRI) and complete laboratory analyses were carried out to exclude other causes of progressive or reversible dementias, in order to have a homogenous mild AD patient sample. Exclusion criteria included, in particular, any evidence of (i) fronto-temporal dementia, (ii) VAD (i.e. the VaD was also diagnosed according to NINDS-AIREN criteria; [205], (iii) extra-pyramidal syndromes, (iv) reversible dementias (including pseudo-dementia of depression), and (v) fluctuations in cognitive performance (suggestive of a possible Lewy body dementia). The normal control subjects were recruited mainly among patients' spouses. All normal subjects underwent physical and neurological examinations as well as cognitive screening (including MMSE). Subjects affected by chronic systemic illnesses (i.e. diabetes mellitus or organ failure) were excluded, as

were subjects receiving psychoactive drugs. Subjects with a history of present or previous neurological or psychiatric disease were also excluded. All normal subjects had a GDS score lower than 14. The ethnicities of the subjects were unknown. Details of the cohort are summarised in Table 3-7.

Table 3-7: Cohort information for Dataset E

No.	Subject	Age (yrs)	Gender	Education (yrs)	MMSE	Diagnosis
1	AD-s01	77	M	9	27	Prob. AD
2	AD-s02	78	F	5	18.7	Prob. AD
3	AD-s03	72	F	5	23.3	Prob. AD
4	AD-s04	74	M	18	25.7	Prob. AD
5	AD-s05	86	M	8	23.8	Prob. AD
6	AD-s06	67	M	18	25.2	Prob. AD
7	AD-s07	79	M	13	25	Prob. AD
8	AD-s08	82	F	5	18.4	Prob. AD
9	AD-s09	79	F	8	20	Prob. AD
10	AD-s10	81	F	8	19	Prob. AD
11	AD–s11	77	F	5	21.7	Prob. AD
12	AD–s12	71	F	5	23	Prob. AD
13	AD–s13	80	F	8	21	Prob. AD
14	AD–s14	83	F	5	19.4	Prob. AD
15	AD–s15	83	F	13	17.1	Prob. AD
16	AD–s16	68	M	3	18.4	Prob. AD
17	AD–s17	74	M	4	21.7	Prob. AD
18	AD–s18	81	M	17	22	Prob. AD
19	AD–s19	87	M	3	17	Prob. AD
20	AD–s20	77	M	5	16.7	Prob. AD
21	Nold -s01	72	F	5	27.3	Normal
22	Nold -s02	77	M	13	30	Normal
23	Nold -s03	78	M	5	28.7	Normal
24	Nold -s04	71	F	8	28.4	Normal
25	Nold -s05	72	M	5	29.3	Normal
26	Nold -s06	83	M	13	29.1	Normal
27	Nold -s07	78	F	5	28.7	Normal
28	Nold -s08	78	F	5	28.7	Normal
29	Nold -s09	79	F	5	29	Normal
30	Nold -s10	79	M	18	30	Normal
31	Nold –s11	73	F	5	29.3	Normal
32	Nold –s12	73	F	17	26.7	Normal
33	Nold –s13	83	M	10	28	Normal
34	Nold –s14	70	F	5	27.3	Normal
35	Nold –s15	79	F	5	27.7	Normal
36	Nold –s16	74	M	10	28.4	Normal
37	Nold –s17	77	F	13	28.3	Normal
38	Nold –s18	72	M	5	28.3	Normal
39	Nold –s19	73	M	6	26.3	Normal
40	Nold –s20	74	M	10	27.4	Normal

3.1.4.3. Recording Information

All subjects were recorded at rest with eyes closed. The EEGs were obtained using the 10-20 system. Details of the recordings are given in Table 3-8.

Table 3-8: Recording information for Dataset D

Description	Details
Montage	Common Reference Montage
Sampling Frequency (Hz)	128
Number of Channels	19
Channel Labels/Order	Fp1 = 1, Fp2 = 2, F7 = 3, F3 = 4, FZ = 5, F4 = 6, F8 = 7, T3 = 8, C3 = 9, CZ = 10, C4 = 11, T4 = 12, T5 = 13, P3 = 14, PZ = 15, P4 = 16, T6 = 17, O1 = 18, O2 = 19. Channels A1 and A2 were not used.
Duration (sec)	60
Equipment Settings	Not Applicable

3.2. Pre-processing

The EEG data was used in two different ways in this study. In the first way, the entire EEG record was analysed to compute EEG based biomarkers. In the second way, instead of using the entire EEG record, EEG data was filtered into the traditional five EEG frequency bands i.e., delta, theta, alpha, beta, and gamma. The low computational infinite impulse response (IIR) Chebyshev-II bandpass filter was used to retain computational efficiency in extracting the biomarkers [206].

The normal specifications of a bandpass filter H(s) as shown in Figure 3-5 are the cutoff frequencies ω_1 and ω_2, the maximum value of the magnitude in the passband between the cutoff frequencies, the maximum attenuation in this passband or the minimum magnitude at the cutoff frequencies ω1 and ω2, and a frequency ωs (= ω3 or ω4) in the stopband at which the minimum attenuation or the maximum magnitude are specified. The type of passband response required may be a Butterworth or Chebyshev response.

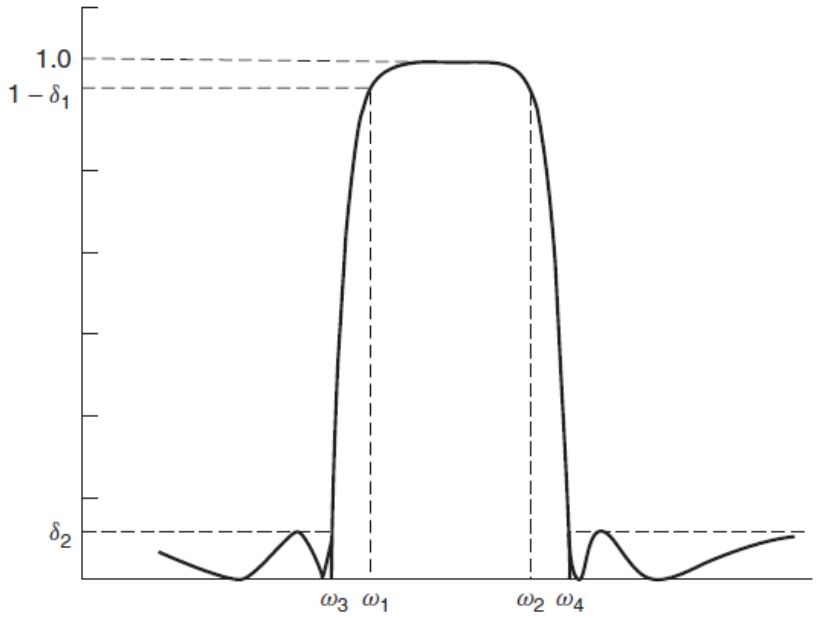

Figure 3-1: Typical specifications of a bandpass filter

The lowpass-bandpass (LP-BP) frequency transformation p = g(s) that is used for the design of a specified bandpass filter is

$$p = \frac{1}{B}\left(\frac{s^2+\omega_0^2}{s}\right) \tag{3.1}$$

where B = ω2 − ω1 is the bandwidth of the filter and ω0 =$\sqrt{\omega_1\omega_2}$ is the geometric mean frequency of the bandpass filter. A frequency s = jΩk in the bandpass filter is mapped to a frequency p = jk under this transformation, which is obtained by

$$\mathrm{j\Omega k} = \frac{j}{B}\left(\frac{\omega_0^2-\omega_k^2}{\omega_k}\right) \tag{3.2}$$

$$= \frac{j_{\omega 0}}{B}\left(\frac{\omega_k}{\omega_0}-\frac{\omega_0}{\omega_k}\right) \tag{3.3}$$

Therefore the frequencies ω1 and ω2 map to Ω =∓1, and the frequencies −ω1 and −ω2 map to Ω =±1. Similarly, the positive value of the stopband frequency Ωs to which the frequency ωs maps is calculated from

$$\Omega_{S=} \left| \frac{1}{B} \left(\frac{\omega_0^2 - \omega_S^2}{\omega_S} \right) \right| \quad (3.4)$$

The magnitude or the attenuation at the frequencies Ω = 1and Ω_s for the prototype filter are the same as those at the corresponding frequencies of the bandpass filter. From the specification of the lowpass prototype filter, we obtain its transfer function H(p), following the appropriate design procedure discussed earlier. Then we substitute (3.1) in H(p) to get the transfer function H(s) of the bandpass filter specified [207].

Chebyshev II filter minimizes the absolute difference between the ideal and the actual frequency response over the entire stopband by using an equal ripple in the stopband. Passband response is maximally flat. The stopband does not approach zero as quickly as the type I filter (and does not approach zero at all for even-valued filter order n). The absence of ripple in the passband, however, is often an important advantage.

A Chebyshev-II filter is an equiripple filter that has the ripples in the stopband rather than the passband [208].

Fdatool of MATLAB R2017b was used to design the (IIR) Chebyshev-II bandpass filters for the delta, theta, alpha, beta, and gamma bands, as shown in the Figures 3-2 to 3-6. The fdatool is a MATLAB toolbox for filters development. Filter design by using fdatool needs to detect the specification of the required filter.

The specifications for the bandpass filters are summarised in Table 3-9.

Table 3-9: Specifications for the bandpass filter for delta, theta, alpha, beta, and gamma bands

Filter parameters	Delta	Theta	Alpha	Beta	Gamma
Sampling Frequency, Fs (Hz)	128	128	128	128	128
First Stopband Frequency, Fstop1 (Hz)	0.1	3.5	7.5	11.5	29.5
First passband Frequency, Fpass1 (Hz)	0.5	4	8	12	30
Second passband Frequency, Fpass1 (Hz)	4	8	12	30	45
Second Stopband Frequency, Fstop2 (Hz)	4.5	8.5	12.5	30.5	45.5
Stopband attenuation, Astop1 (dB)	60	60	60	60	60
Stopband attenuation, Astop 2 (dB)	80	80	80	80	80

The choice of filter parameters are determined by the requirements of the pre-processing. For example, for the delta filter, there is often a need to remove the low frequency baseline and this necessitates a low stop band frequency (0.1 Hz). The second passband frequencies are determined by the widely accepted band edge frequencies for the EEG activities. After designing the required filters, the MATLAB fdatool is used to generate the corresponding MATLAB function for the designed filters. The functions can then be used to filter the EEG signal, as shown in Section 3.3.2.

In Sections 3.2.1 to 3.2.5, the design of the filters is presented.

3.2.1. Chebyshev-II bandpass filter design for delta band

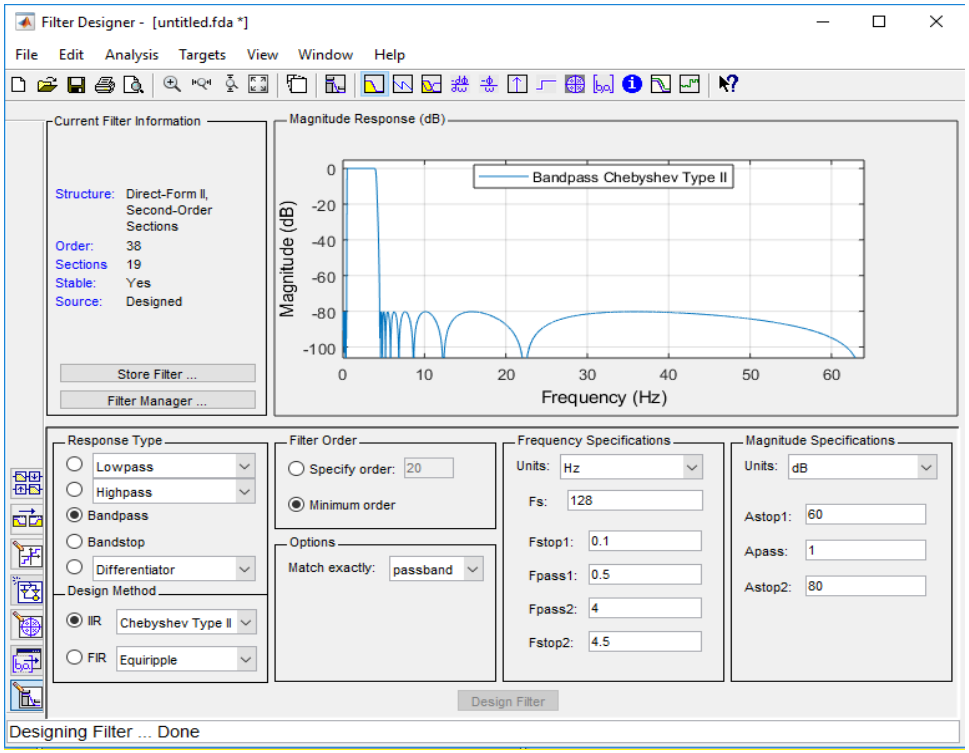

Figure 3-2: Magnitude response of (IIR) Chebyshev-II bandpass filters for delta band from 0Hz to 4Hz

3.2.2. Chebyshev-II bandpass filter design for theta band

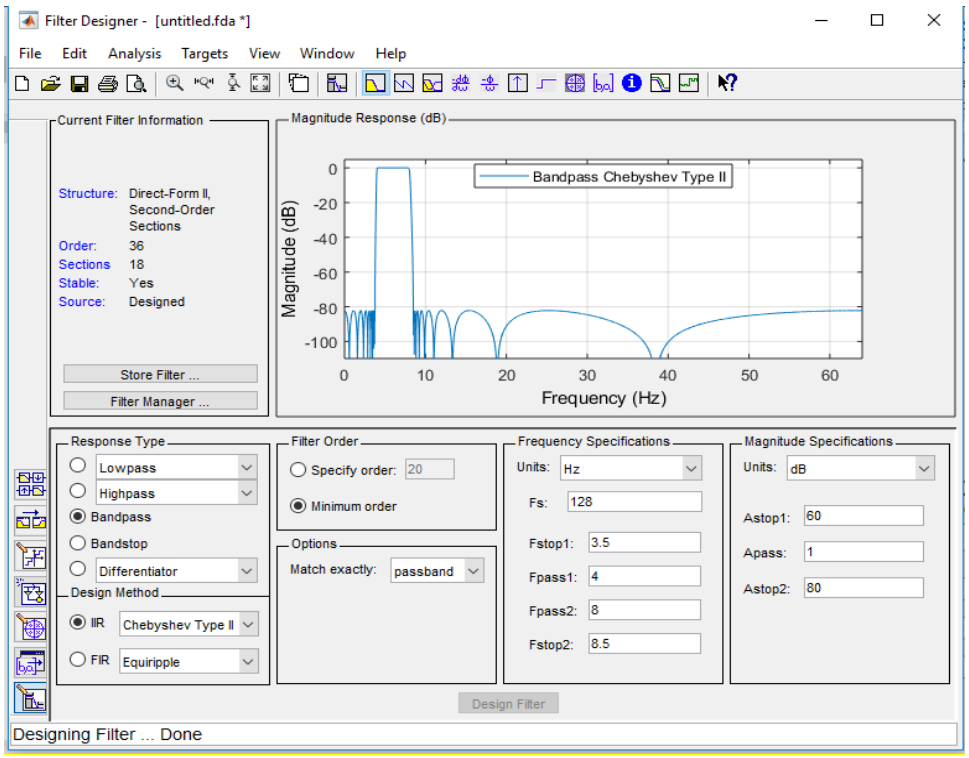

Figure 3-3: Magnitude response of (IIR) Chebyshev-II bandpass filters for theta band from 4Hz to 8Hz

3.2.3. Chebyshev-II bandpass filter design for alpha band

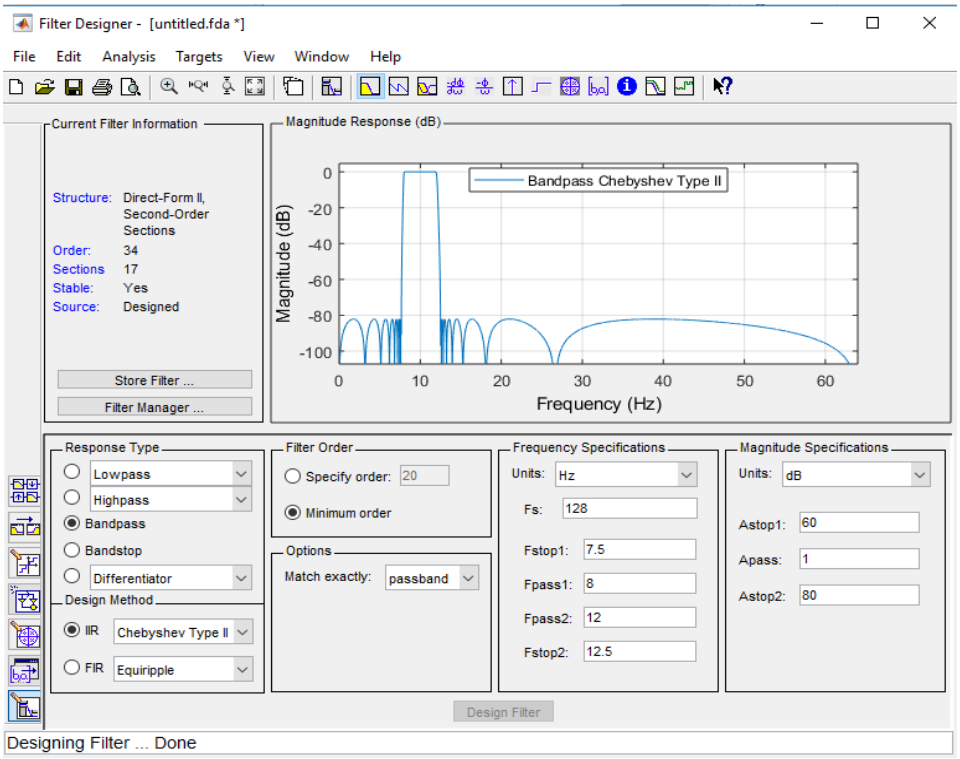

Figure 3-4: Magnitude response of (IIR) Chebyshev-II bandpass filters for alpha band from 8Hz to 12Hz

3.2.4. Chebyshev-II bandpass filter design for beta band

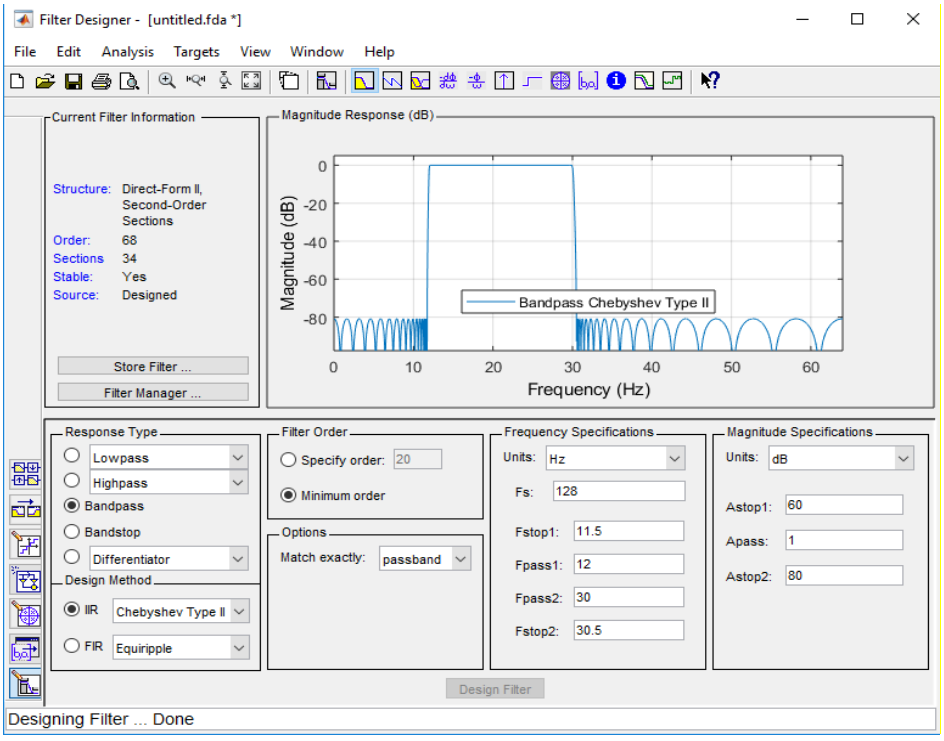

Figure 3-5: Magnitude response of (IIR) Chebyshev-II bandpass filters for beta band from 12Hz to 30Hz

3.2.5. Chebyshev-II bandpass filter design for gamma band

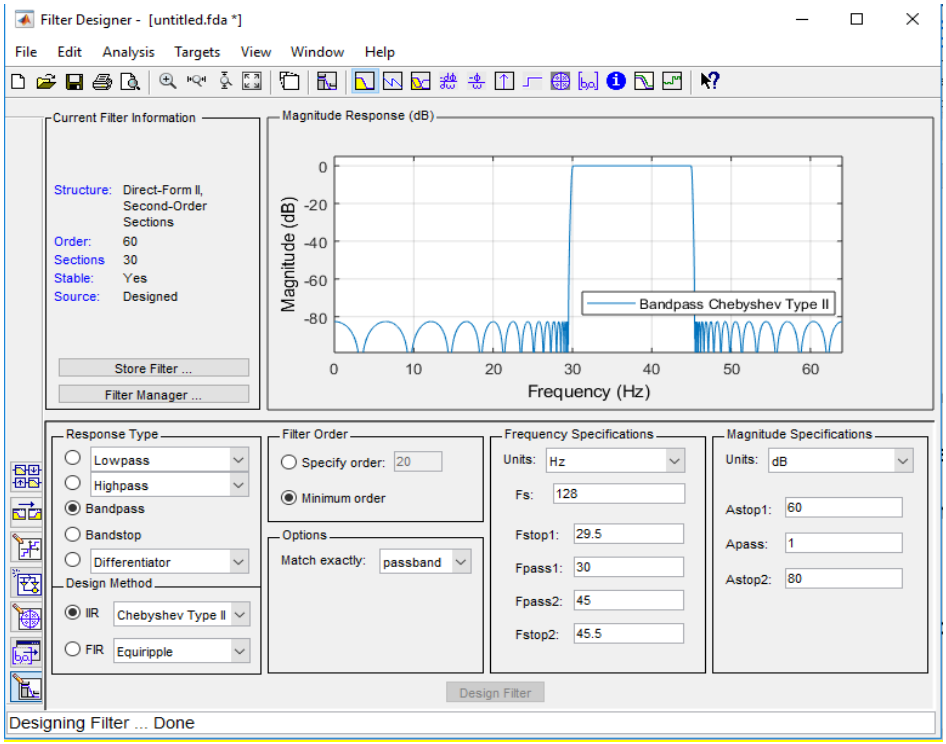

Figure 3-6: Magnitude response of (IIR) Chebyshev-II bandpass filters for gamma band from 30Hz to 45Hz

3.3. Computational programs

3.3.1. Introduction

The mathematical computation, signal processing and analysis, statistical analysis, and machine learning techniques used in this study were carried out using MATLAB toolboxes (i.e., the signal processing toolbox, and machine learning toolbox). Different versions of MATLAB were used throughout (MATLAB® 2015 to MATLAB® 2018).

MATLAB was selected because most biomedical engineering analysis for EEG signal is based on MATLAB. For example, EEGLAB was developed using MATLAB and this means it is likely to be compatible with other biomedical applications. Furthermore, it is an interactive system with a high-performance programming language for technical computing that integrates computation, visualisation, and programming, and is easy-

to-use. It is therefore the tool of choice for high-productivity research, development, and analysis.

A biomedical engineering toolbox was then developed during the study that can be used to detect AD based on EEG analysis.

The following figures show the toolbox that was developed.

Figure 3-7 shows the main menu for the biomedical engineering toolbox and the toolbar that contains the main tasks.

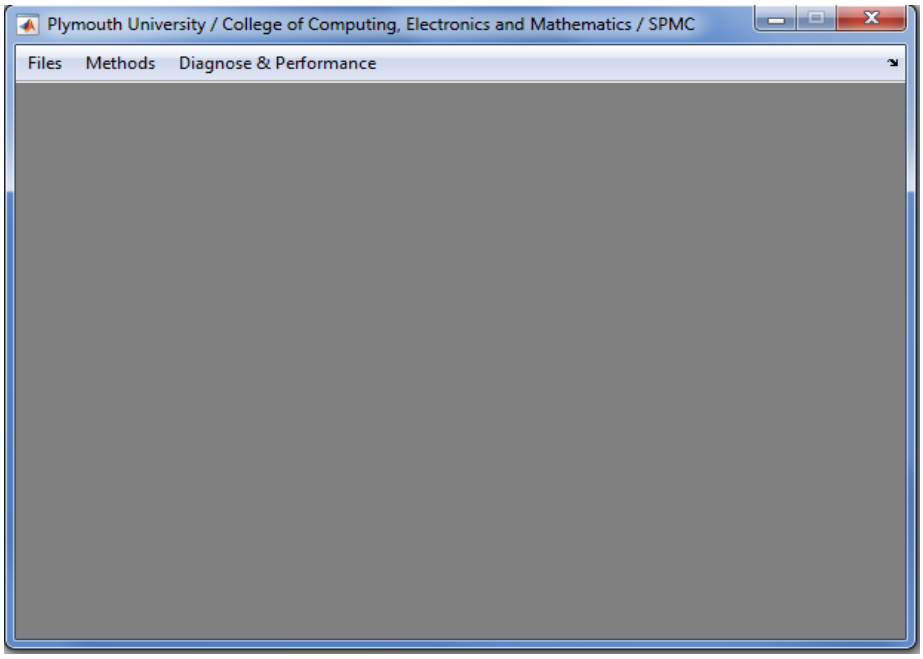

Figure 3-7: Main menu for the biomedical engineering toolbox

Figure 3-8 shows the files submenu which contains operations related to file processing.

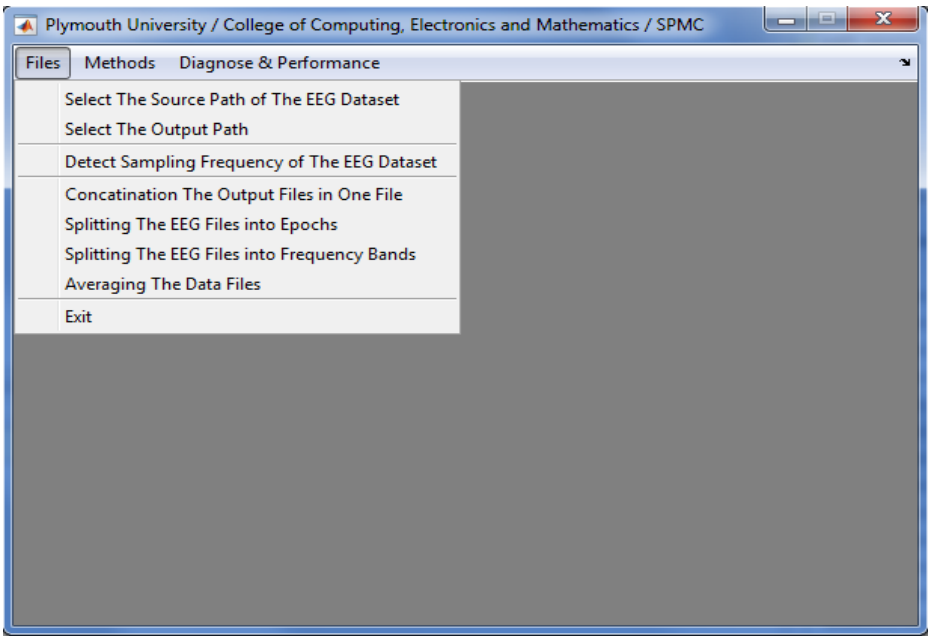

Figure 3-8: The files submenu for the biomedical engineering toolbox

Figure 3-9 shows the methods submenu which contains operations related to computation methods.

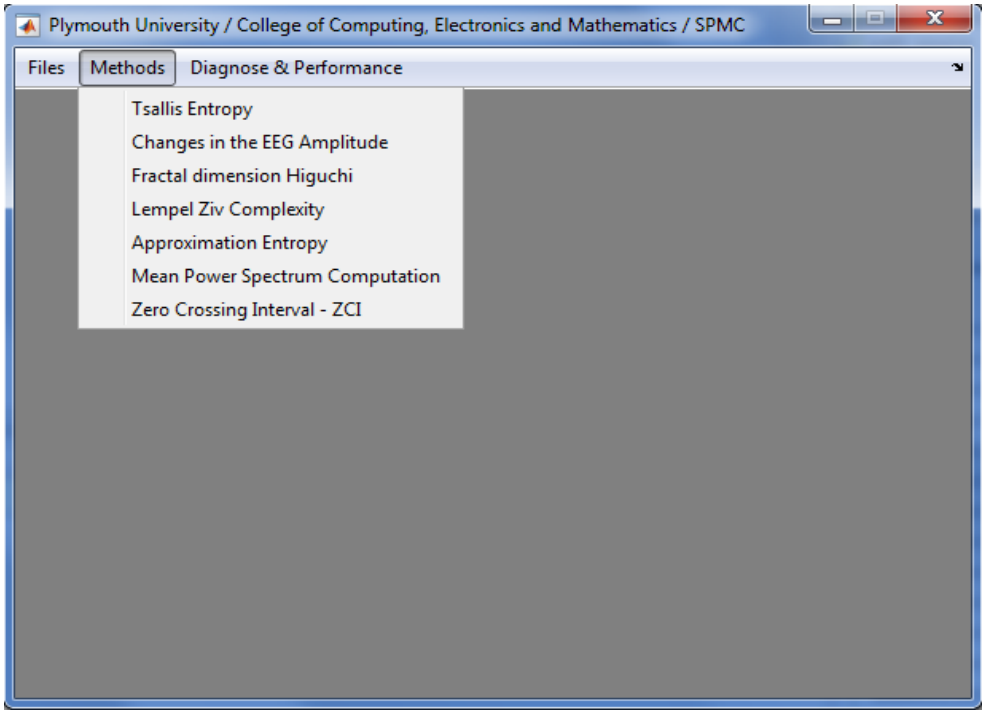

Figure 3-9: The methods submenu for the biomedical engineering toolbox

Figure 3-10 shows the diagnosis and performance analysis submenu which contains operations related to diagnosis and performance computations.

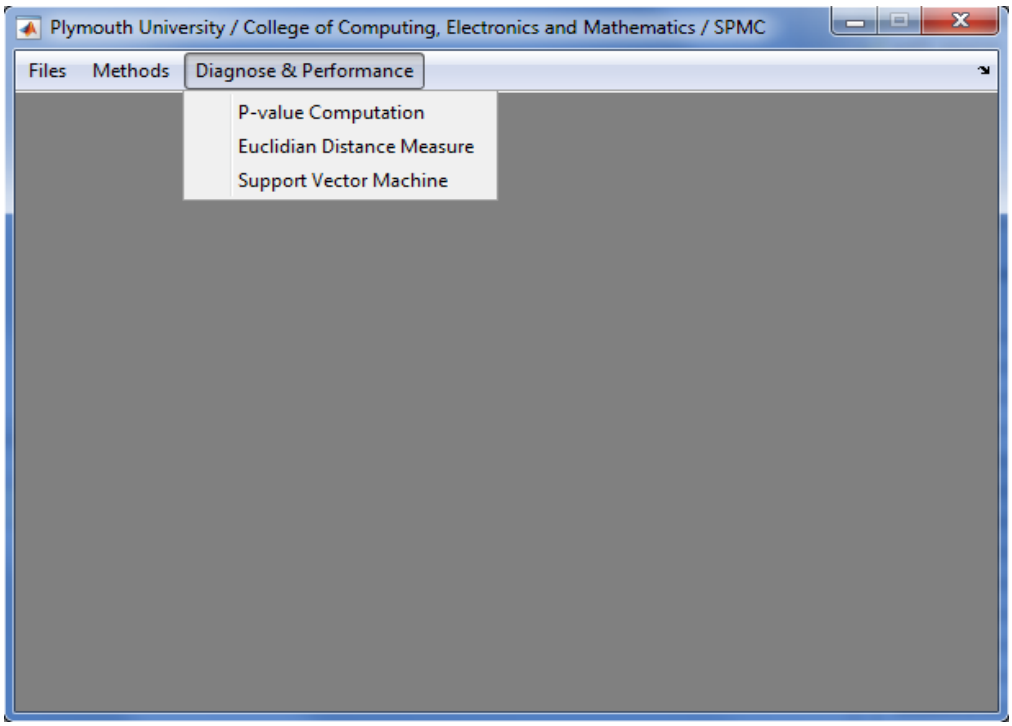

Figure 3-10: The diagnosis and performance analysis submenu containing operations related to diagnosis and performance analysis

Figure 3-10 shows the diagnosis and performance analysis submenu which contains operations related to diagnosis and performance computations.

As shown in Figures 3-7 to 3-10, there are several menus and each menu has multiple selections, and there is a specific function associated with each selection. These functions are not standard MATLAB functions but were created by using MATLAB codes. The MATLAB codes developed for the functions are given in the Appendixes 1, 2 and 3 for convenience.

For example, to compute the ZCI values for a given input EEG dataset, the zci_function can be used.

The developed functions are explained in Sections 3.3.2, 3.3.3, and 3.3.4 below.

3.3.2. MATLAB functions for filtering EEG signal

EEG signal was filtered into five EEG frequency bands i.e., delta, theta, alpha, beta, and gamma. For each band, a MATLAB function was developed using the fdatool, as

shown in Sections 3.2.1 to 3.2.5. The whole body of each function was placed in Appendix 1.

Figure 3-11 demonstrates the effect of using Chebyshev-II bandpass filter for filtering real EEG signal into five traditional EEG frequency bands e.g., delta, theta, alpha, beta, and gamma. The MATLAB codes were used to generate Figure 3-11 are given in Appendix 1.

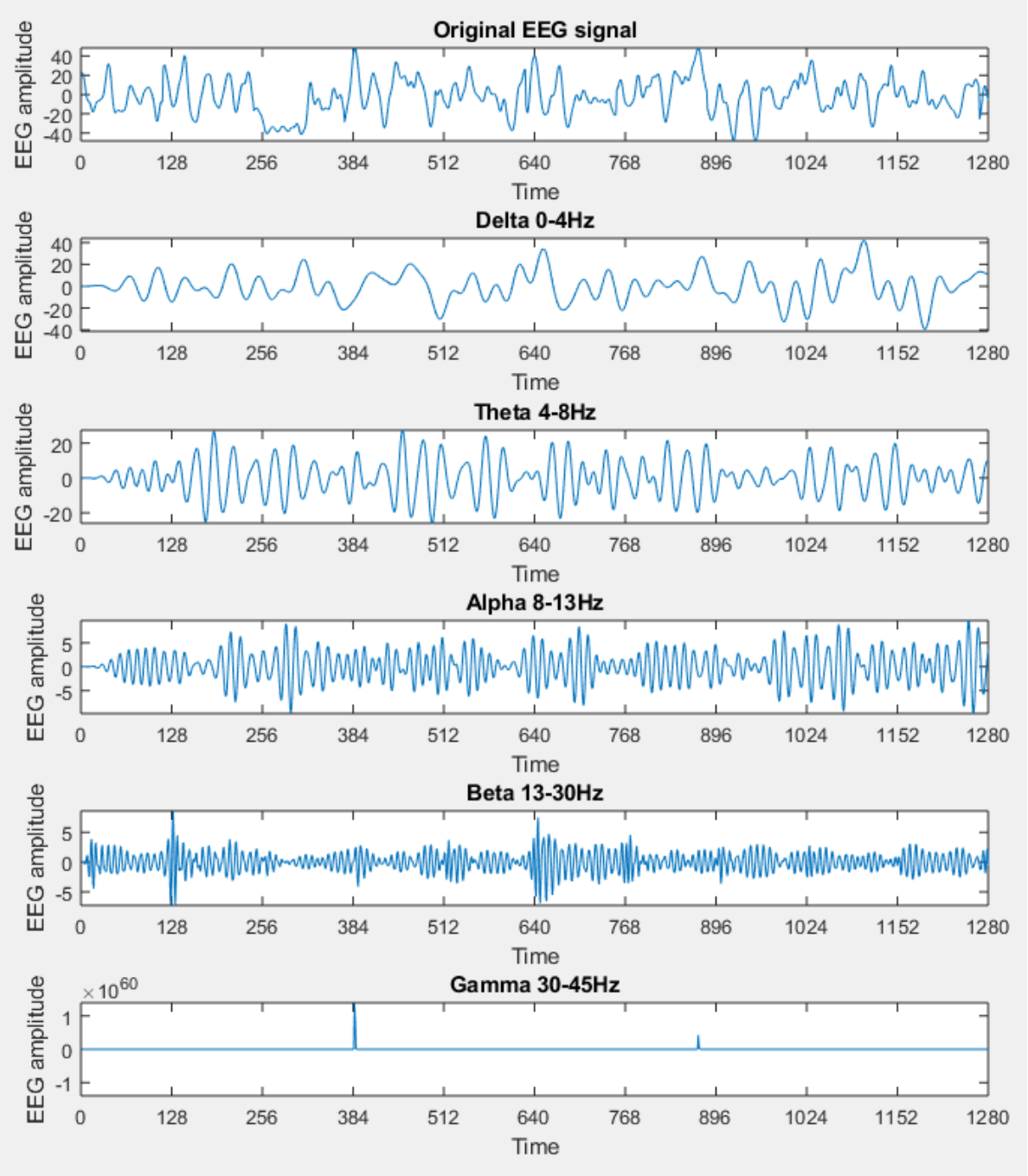

Figure 3-11: An original EEG signal and its delta, theta, alpha, beta, and gamma bands

3.3.2.1. Chebyshev-II bandpass filter function for delta band

Purpose: Filter EEG signal into delta band.

Syntax: [output] = delta (data, Fs)

Description: The function delta was used to filter the input signal into delta band. data is the input EEG signal, Fs is the sampling frequency for the input EEG signal, and output is the filtered signal.

Example: output = delta (data,128);

An illustrative example for data is given below.

3.3.2.2. Chebyshev-II bandpass filter function for theta band

Purpose: Filter EEG signal into theta band.

Syntax: [output] = theta (data, Fs)

Description: The function theta was used to filter the input signal into theta band. data is the input EEG signal, Fs is the sampling frequency for the input EEG signal, and output is the filtered signal.

Example: output = theta (data,128);

3.3.2.3. Chebyshev-II bandpass filters for alpha band

Purpose: Filter EEG signal into alpha band.

Syntax: [output] = alpha (data, Fs)

Description: The function alpha was used to filter the input signal into alpha band. data is the input EEG signal, Fs is the sampling frequency for the input EEG signal, and output is the filtered signal.

Example: output = alpha (data,128);

3.3.2.4. Chebyshev-II bandpass filters for beta band

Purpose: Filter EEG signal into beta band.

Syntax: [output] = beta (data, Fs)

Description: The function beta was used to filter the input signal into beta band. data is the input EEG signal, Fs is the sampling frequency for input EEG signal, and output is the filtered signal.

Example: output = beta (data,128);

3.3.2.5. Chebyshev-II bandpass filters for gamma band

Purpose: Filter EEG signal into gamma band.

Syntax: [output] = gamma (data, Fs)

Description: The function gamma was used to filter the input signal into gamma band. data is the input dataset of EEG signal, Fs is the sampling frequency for the input EEG signal, and output is the filtered signal.

Example: output = gamma (data,128);

3.3.3. MATLAB code for computing EEG signal processing methods

The EEG based biomarkers were investigated based on eight analysis methods. These methods include ΔEEG_A, ΔPS, ZCI, TsEn, HFD, ApEn, LZC, and EEG Coherence. For each analysis method, a MATLAB function was developed. All the functions were used in this study are not standard MATLAB functions.

The whole body of each function was placed in Appendix 2.

To illustrate the execution of the functions, suppose we have an array e.g., data(12,10). As shown in Table 3-10, it contains 10 columns and 12 arrows. This means we have an EEG dataset for one subject and it has 10 EEG channels, and each EEG signal has 12 samples. Suppose the sampling frequency is 10Hz.

Table 3-10: Sample EEG dataset was used to illustrate the execution of the functions

4.41	7.93	-3.59	2.34	1.57	3.51	2.67	-3.66	0.00	-0.51
4.31	13.95	-8.84	4.13	2.33	5.89	2.65	-5.22	0.06	0.61
3.94	17.58	-11.74	5.77	3.02	7.72	0.83	-5.64	0.40	1.66
3.28	18.32	-12.22	6.77	3.41	8.68	-1.57	-5.34	1.18	2.31
2.46	16.81	-11.07	6.87	3.38	8.61	-2.70	-4.91	2.29	2.36
1.73	14.45	-9.47	6.15	2.92	7.62	-1.18	-4.69	3.35	1.85
1.31	12.65	-8.44	4.95	2.19	6.05	3.02	-4.65	3.89	1.01
1.27	12.20	-8.46	3.69	1.35	4.45	8.32	-4.41	3.60	0.22
1.39	12.92	-9.44	2.67	0.50	3.24	12.20	-3.66	2.61	-0.20
1.26	13.83	-10.96	1.98	-0.38	2.59	12.48	-2.41	1.43	-0.13
0.40	13.72	-12.58	1.48	-1.42	2.31	8.37	-1.09	0.75	0.35
4.41	7.93	-3.59	2.34	1.57	3.51	2.67	-3.66	0.00	-0.51

3.3.3.1. Changes in the EEG amplitude (ΔEEG_A) computation

Purpose: Compute the ΔEEG_A for EEG signal.

Syntax: [output] = EEG_amplitude(data,Fs)

Description: The function EEG_amplitude was used to compute changes in the EEG amplitude values for the input signal. 'data' is the input dataset of EEG signal, Fs is the sampling frequency for input EEG signal, and output is the resulted values from applying EEG_amplitude function. Computation of that function was based on applying Equations 2.1 – 2.4 as mentioned in Section 2.4.2.

Example: output = EEG_amplitude(data,10);

Result:

22.24	50.74	42.00	15.19	19.40	79.65	33.93	64.24	29.16	36.63

Each value in the result represents the ΔEEG_A value for each EEG channel.

3.3.3.2. Zero-crossing intervals (ZCI) computation

Purpose: Compute the ZCI for EEG signal.

Syntax: [output] = zci_function (data,Fs)

Description: The function zci_function was used to compute zero-crossing intervals values for the input signal. 'data' is the input dataset of EEG signal, Fs is the sampling frequency for the input EEG signal, and output is the resulted values from applying the zci_function function. Computation of that function was based on applying Equations 2.5 – 2.7 as mentioned in Section 2.4.3.

Example: output = zci_function(data,10);

Result:

0.83	1.81	3.05	0.69	0.00	-4.61	2.77	2.97	-0.99	0.00

Each value in the result represents the ZCI value for each EEG channel.

3.3.3.3. Changes in the power spectrum (ΔPS) of EEG signal computation

Purpose: Compute the ΔPS for EEG signal.

Syntax: [output] = power_spectrum(data)

Description: The function power_spectrum was used to compute changes in the power spectrum values of the input signal. data is the input dataset of EEG signal, and output is the resulted values from applying power_spectrum function. Computation of that function was based on applying Equations 2.8 – 2.9 as mentioned in Section 2.4.4.

Example: output = power_spectrum (data,10);

Result:

0.83	1.81	3.05	0.69	0.00	-4.61	2.77	2.97	-0.99	0.00

Each value in the result represents the ΔPS value for each EEG channel.

3.3.3.4. EEG coherence computation

Purpose: Compute the EEG coherence for EEG signal.

Syntax: [output] = coherence_function (data,Fs)

Description: The function coherence_function was used to compute coherence values of the input signal. data is the input dataset of EEG signal, Fs is the sampling frequency for the input EEG signal, Ch is the EEG channel numbers e.g. [1 2 3], and output is the resulted values from applying the coherence_function function. Computation of that function was based on applying Equation 2.9 as mentioned in Section 2.4.5.

Example: output = coherence_function(data,[1 2 3 4 5],10);

Result:

0.83	1.81	3.05	0.69	0.00	-4.61	2.77	2.97	-0.99	0.00

Each value in the result represents the coherence value for each EEG channel.

3.3.3.5. Tsallis entropy (TsEn) computation function

Purpose: Compute the TsEn for EEG signal.

Syntax: [output] = Tsallis_entro (data,q)

Description: The function Tsallis_entro was used to compute Tsallis entropy values of the input signal. q is a Tsallis factor, data is the input dataset of EEG signal, and output is the resulted values from applying Tsallis_entro

function. Computation of that function was based on applying Equation 2.10 as mentioned in Section 2.4.6.

Example: output = Tsallis_entro (data,0.5);

Result:

-23.08	29.19	-6.17	91.90	67.77	31.45	26.79	7.38	29.42	16.91

Each value in the result represents the TsEn value for each EEG channel.

3.3.3.6. Higuchi Fractal Dimension (HFD) computation function

Purpose: Compute the HFD for EEG signal.

Syntax: [output] = Higuchi_function (data)

Description: The function Higuchi_function was used to compute HFD values of the input signal. data is the input dataset of EEG signal, and output is the resulted values from applying Higuchi_function function. Computation of that function was based on applying Equations 2.11 - 2.13 as mentioned in Section 2.4.7.

Example: output = Higuchi_function (data);

Result:

-0.45	-0.42	-0.27	-0.42	-1.21	0.32	-0.63	-0.35	-0.82	-0.57

Each value in the result represents the HFD value for each EEG channel.

3.3.3.7. Approximation Entropy (ApEn) computation function

Syntax: [output] = ApEn(dim, r, data)

Description: The function ApEn was used to compute ApEn values of the input signal. dim is the embedded dimension, r is tolerance (typically 0.2 * std), data is the input dataset of EEG signal, and output is the resulted values from

applying ApEn function. Computation of that function was based on applying Equations 2.14 - 2.16 as mentioned in Section 2.4.8.

Example: output = ApEn(2, 5.12, data)

Result:

0.29	0.61	0.37	0.03	0.08	-0.15	-0.18	-0.22	0.07	-0.41

Each value in the result represents the ApEn value for each EEG channel.

3.3.3.8. Lempel Ziv Complexity (LZC) computation function

Purpose: Compute the LZC for EEG signal.

Syntax: [output] = LZC (data)

Description: The function LZC was used to compute LZC values of the input signal. data is the input dataset of EEG signal, and output is the resulted values from applying LZC function. Computation of that function was based on applying Equations 2.18 - 2.19 as mentioned in Section 2.4.9.

Example: output = LZC (data);

Result:

1.49	1.79	1.49	1.79	1.79	1.19	1.49	1.19	1.49	1.19

Each value in the result represents the LZC value for each EEG channel.

3.3.4. MATLAB machine learning approaches

The EEG based biomarkers were investigated to assess their performance. Machine learning model was developed for each biomarker using the MATLAB machine learning toolbox (classification learner). The biomarkers were extracted for eight analysis methods. These methods include ΔEEG_A, ΔPS, ZCI, TsEn, HFD, ApEn, LZC, and EEG Coherence. 10-fold cross-validation was used to learn the developed models. For each machine learning approach, a MATLAB function was developed.

All functions were used in this study are not standard MATLAB functions. The whole body of each function was placed in Appendix 3.

The following steps explain using MATLAB functions for uploading EEG data into MATLAB and using the developed machine learning functions e.g., SVM, LDA, and KNN.

1. Prepare CSV file contains the biomarker for AD patients and normal subjects as shown in Table 3-11. Suppose we have a data file saved in Excel format e.g., data.xlsx as shown in Table 3-11.

Table 3-11: CSV EEG data file for uploading in MATLAB toolbox (classification learner)

VarName1	VarName2	VarName3	VarName4	VarName5	VarName6	Stat
10.716	0.018364	0.638427	0.152862	0.068752	0.856114	AD
10.733	0.018318	0.639364	0.153254	0.068966	0.856175	AD
10.745	0.018243	0.637837	0.15297	0.069534	0.856224	AD
10.732	0.018086	0.637822	0.153039	0.068979	0.854451	AD
10.723	0.0177	0.635625	0.152946	0.068819	0.855088	AD
10.75	0.017454	0.635193	0.152995	0.069432	0.856945	AD
10.745	0.017056	0.635536	0.152566	0.07119	0.85299	AD
10.747	0.016569	0.636369	0.152796	0.072434	0.85044	AD
0.008587	0.342673	0.000735	0.03965	0.035008	0.016558	Norm
0.008772	0.233128	0.000752	0.039545	0.034953	0.016595	Norm
0.008593	0.120405	0.000736	0.039161	0.035027	0.016589	Norm
0.007657	0.087793	0.000648	0.039442	0.034994	0.016666	Norm
0.008925	0.051825	0.000759	0.039408	0.035011	0.016543	Norm
0.007679	0.034675	0.000647	0.038812	0.034979	0.016685	Norm
0.008645	0.025307	0.000734	0.039181	0.035008	0.016582	Norm

The variables in Table 3-10 refers to the biomarkers e.g., Var1 for TsEn(Alpha/theta(T6)), Var2 for ZCI(Alpha/theta(P3)), Var3 for ZCI(Delta/alpha(P3)), and so on.

2. Uploading the CSV file into the MATLAB by using MATLAB function importfile. The developed importfile function was used to import the EEG data into MATLAB, as shown:

Purpose: Upload data file into MATLAB.

Syntax: tableout = importfile(workbookFile,sheetName,startRow,endRow)

Description: The function importfile was used to read the data from the worksheet in the Microsoft Excel spreadsheet file named workbookFile. It then returns the data as a table e.g., tableout into MATLAB toolbox (classification learner), sheetName is worksheet name e.g., 'Sheet1'. Specify startRow and endRow as a pair of scalars or vectors of matching size for dis-contiguous row intervals. To read to the end of the file specify an endRow.

Example: Table = importfile ("data.xlsx", "Sheet1",1,15);

3. The uploaded EEG data can be classified using the developed machine learnings functions e.g., SVM, LDA, and KNN.

The following Sections 3.3.4.1 – 3.3.4.3 explain using the developed machine learning approaches to detect AD patients.

3.3.4.1. Support Vector Machine (SVM)

Function trainClassifier_SVM was used to develop an SVM classification model using 10-fold cross-validation. The developed model was used to detect AD patients as shown in Section 3.3.4.4.

Function trainClassifier_SVM returns a trained classifier and its accuracy.

Purpose: Develop an SVM classification model using 10-fold cross-validation.

Syntax: [trainedClassifier, validationAccuracy] = trainClassifier_SVM(trainingData)

Description: The function trainClassifier_SVM was used to develop an SVM classification model using 10-fold cross-validation, trainingData is EEG dataset table was used to develop the classification model, trainedClassifier is a struct containing the trained classifier, and validationAccuracy is the accuracy of the developed model.

Example: [trainedClassifier, validationAccuracy] = trainClassifier_SVM(Table)

3.3.4.2. Linear Discriminant Analysis computation

Function trainClassifier_LDA was used to develop an LDA classification model using 10-fold cross-validation. The developed model was used to detect AD patients as shown in Section 3.3.4.4.

Function trainClassifier_ LDA returns a trained classifier and its accuracy.

Purpose: Develop an LDA classification model using 10-fold cross-validation.

Syntax: [trainedClassifier, validationAccuracy] = trainClassifier_ LDA (trainingData)

Description: The function trainClassifier_LDA was used to develop an LDA classification model using 10-fold cross-validation, trainingData is EEG dataset table was used to develop the classification model, trainedClassifier is a struct containing the trained classifier, and validationAccuracy is the accuracy of the developed model.

Example: [trainedClassifier, validationAccuracy] = trainClassifier_ LDA (Table)

3.3.4.3. K-nearest neighbour (KNN) computation

Function trainClassifier_ KNN was used to develop an KNN classification model using 10-fold cross-validation. The developed model was used to detect AD patients as shown in Section 3.3.4.4.

Function trainClassifier_KNN returns a trained classifier and its accuracy.

Purpose: Develop an KNN classification model using 10-fold cross-validation.

Syntax: [trainedClassifier, validationAccuracy] = trainClassifier_ KNN (trainingData)

Description: The function trainClassifier_KNN was used to develop an KNN classification model using 10-fold cross-validation, trainingData is EEG

dataset table was used to develop the classification model, trainedClassifier is a struct containing the trained classifier, and validationAccuracy is the accuracy of the developed model.

Example: [trainedClassifier, validationAccuracy] = trainClassifier_KNN (Table)

3.3.4.4. Model validation

The developed model can be validated by using another EEG dataset. The developed models e.g., trainedClassifier for the SVM, LDA, and KNN were used for validation using different EEG dataset. The function predictFcn was used to make a prediction using a new dataset (unseen EEG dataset e.g., New_EEG_data).

For example, to validate the SVM developed model using the New_EEG_data the function would be used as shown:

Yfit_SVM = trainedModel_SVM.predictFcn(New_EEG_data);

the output (results) in yfit_SVM as shown,

```
yfit_SVM =
  20×1 categorical array
     AD
     AD
     AD
     AD
     AD
     AD
     AD
     AD
     AD
     AD
     AD
     Norm
     Norm
     Norm
     Norm
     Norm
     Norm
     Norm
```

Norm

Norm

Norm

3.4. Overview of EEG Based Biomarkers

3.4.1. Introduction

A biomarker is an indicator used to detect a specific biological state, along with functional and structural variations in organs and cells [2][33][34]. AD is characterised by loss of memory and cognitive decline resulting from damage to brain cells which affects brain activity [48]. Damage to nerve cells/pathways in the brain due to AD causes changes in the information-processing activity of the brain. These changes are thought to be reflected in the informational content of the EEG [23]. The electroencephalogram (EEG) can therefore play a potentially valuable role in the early detection of AD [15][23][24][42][43][44][50]. EEG is non-invasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics related to AD [23][24][43][45]. EEG exhibits high sensitivity when discriminating between AD patients and normal people [49][50][209] . The fundamental utility of EEG in detecting changes in brain signals, even in the preclinical stage of the disease, has been widely demonstrated [47]. EEG based biomarkers have proved to be of great value in the identification of preclinical AD [41]. EEG based biomarkers can thus be used as a first line decision-support tool for AD diagnosis in clinical practice.

The most characteristic features evident in an EEG that detects AD are slowing of the EEG, a decrease in EEG coherence, and a reduction in EEG complexity [43][45][47][48][84][85]. These changes can thus be quantified as biomarkers of AD. EEG based biomarkers can therefore be divided into three main categories: the slowing of the EEG, reduction in complexity, and a decrease in the coherence between cortical regions [43][44][45][47][48][50][84][85].

3.4.2. Slowing of EEG

The slowing of the EEG is one of the most consistent features relating to the detection of AD. Slowing may therefore be quantified as a biomarker of AD [15][48][210]. It can be measured in several ways such as changes in EEG amplitude (ΔEEG_A), zero-crossing intervals (ZCI) [15], and changes in the power spectrum (ΔPS) of the EEG signal [15][43][44][48][85][211][212][213][214][215][216][217][218]. Al-nuaimi et al. [8] quantified slowing in EEG by measuring the ΔEEG_A. Their results showed that ΔEEG_A is a promising nonlinear EEG marker in the time domain. It can be measured through changes in EEG amplitude and can track changes in the EEG over time [8]. Their results showed that a gradual change in EEG amplitude is a marker for the subsequent rate of cognitive and functional decline in AD patients [8]. The reduction of ZCI of an EEG signal has also been shown to be a promising biomarker of AD [15] [126].The slowing of the EEG can also be quantified by the power of the EEG signal in different frequency bands (i.e., delta, theta, alpha, beta, and gamma) where slowing is manifest in a decrease in power of high frequency bands (alpha and beta) and an increase in power of low frequency bands (delta and theta). These changes can be used to distinguish AD patients from those with other types of dementia [15][43][44][48][85]. An increase in the power ratio of the alpha/middle alpha bands is an indicator of mild cognitive impairment (MCI) in people who may go on to develop AD [211]. Conversely, an increase in the power ratio of theta/gamma bands has been associated with MCI patients who may not develop AD [219]. This increase was related to a decline in memory and can therefore be used to identify MCI patients in a cohort of normal people [47]. Numerous studies have shown that power changes in the EEG frequency bands are promising markers of AD [211][212][213][214][215][216][217] [218].

3.4.3. Reduction in EEG complexity

Complexity is a measure of the extent to which the dynamic behaviour of a given sequence resembles one that is random [92]. The cortical areas of the brain fire spontaneously and such dynamic behaviour is complex [93][94]. AD causes a reduction in the neuronal activity of the brain [95] resulting in a decreased ability to process information [96][97][98] which may be reflected in EEG signals [95]. EEG complexity can potentially be a good biomarker for AD diagnosis [48] as AD patients exhibit a significant reduction in EEG complexity [48][50][84][85][95][99][100]. This reduction can be measured using Tsallis entropy (TsEn) [45][220][102], Higuchi Fractal Dimension (HFD) [125], Approximation Entropy (ApEn) [136][221][222][115], and Lempel Ziv Complexity (LZC) [86][223]. TsEn is one of the most promising information theoretic methods for quantifying EEG complexity [45][220][102] [105][106]. Its capacity for rapid computation may serve as the basis for real-time decision support tools for diagnosing AD [45][220][105][224][107]. HFD is a fast, nonlinear computational method for obtaining the fractal dimension of time series signals [119][120][121], even when very few data points are available [119]. The fractal dimension of EEG signals is significantly lower in AD patients than in normal people [126][125]. HFD of the EEG is therefore a potentially good biomarker of AD [49][125]. ApEn is a powerful nonlinear powerful method that can be used to analyse the dynamic behaviour of the brain based on complex biological signals such as the EEG [90][225]. ApEn is less affected by noise, robust to outliers, and offers good reproducibility with medium-sized datasets [225][226][137]. The reduction in ApEn of EEG is thus a good indicator of AD [136][221][115][90][135] . LZC is a nonparametric and nonlinear method that provides a powerful method for quantifying the complexity of finite length sequences [110][111]. It has previously been used to analyse EEG complexity in patients with AD [86][112]. The reduction of LZC values may therefore be a good biomarker for AD [86][112][113].

3.4.4. Decrease in EEG coherence

AD causes changes in the cortical electrical activity of the brain [131] which affects the coherence of structural connectivity among cortical regions of the brain [48]. These changes may then be reflected in EEG coherence, which can be quantified by assessing the functional coupling between brain regions [48][227]. Coherence measures depend on the channel location and frequency bands of the EEG signal [131][228][229][132][230]. AD patients show a significant reduction of coherence in the alpha band, especially in the temporo-parieto-occipital regions, and an increase in the coherence of the delta band. AD patients show a reduction in both the left temporal alpha band and interhemispheric theta band [132]. Analysis of left temporal alpha coherence may enhance the usefulness of an EEG in AD diagnosis [132]. Furthermore, a positive association has been found between EEG coherence in the frontal region for delta and beta bands, and incoherence between frontal and posterior in the theta band [229]. EEG coherence is therefore a sensitive and selective method for assessing the integrity of structural connections between brain areas in AD patients [230]. As such, it is a useful marker for AD diagnosis [48][133].

Chapter 4. Investigation of the Novel EEG Biomarker for Detection of Alzheimer's Disease

4.1. Introduction

Many techniques exist for deriving AD biomarkers from the EEG [87]. However, time domain-based approaches are potentially one of the most reliable ways to derive robust EEG biomarkers for AD [15][220][231]. In this chapter, new and emerging EEG biomarkers were investigated and developed.

AD biomarkers based on analysis of the complexity and slowing of EEG signal provided promising results [99][106][232]. Two analysis of complexity methods in EEG signal such as Tsallis entropy (TsEn) and Higuchi fractal dimension (HFD) were investigated and we present a new approach to quantify the slowing of the EEG in the time domain by measuring changes in the EEG amplitudes.

TsEn was selected form the other types of entropies e.g., Shannon entropy, permutation entropy, and Kolmogorov entropy because our experiments showed the performance of Tsallis entropy in AD detection was better than the others types of entropies as shown in Table 4-1.

The changes in the amplitudes over time may be viewed as the mean velocity of the EEG [233]. The approach is easy to implement and is computationally efficient.

The aim of this chapter is to investigate the promising EEG analysis methods used to detect AD. Therefore, new and emerging EEG biomarkers were investigated and developed. The research question in this chapter: is it possible to develop EEG based biomarkers using a minimal number of EEG channels. The hypothesis of this study is that it is possible to identify the candidate EEG channels that can be used to detect AD. The whole EEG record and all 19 EEG channels were investigated. The analysis of EEG signal was achieved in time domain. P-value was used as a criterion to detect EEG channels that can be used to develop a diagnostic model for AD detection. KNN

was used to develop a diagnostic model for each analysis method. The performance of each biomarker was assessed by its sensitivity, specificity.

Table 4-1 shows the performance of TsEn entropy compared to other types of entropy e.g., Shannon entropy, permutation entropy, and Kolmogorov entropy. EEG datasets A, and B were selected in the experiments. Euclidian distance measure was used to compute the distance between AD patients and normal subjects. The results suggested that the TsEn was better than the other entropies for the detection of AD.

Table 4-1: Performance analysis of TsEn entropy compared to types of entropies

Analysis method	Sensitivity	Specificity	Accuracy	F-measure
Tsallis entropy	65.38	88.46	76.92	73.91
Shannon entropy	53.85	76.92	65.38	60.87
permutation entropy	50.00	71.43	61.54	54.55
Kolmogorov entropy	48.15	72.00	59.62	55.32

4.2. Methodology

The EEG data used in this chapter consists of two datasets (A, and B). Dataset A includes 11 age-matched subjects (3 AD patients, and 8 normal subjects). Dataset B includes 41 subjects that were not perfectly age-matched with 24 normal subjects (10 males and 14 females) and 17 were probable AD patients (9 males and 8 females).

The EEG dataset was split into training and testing data (60% for training and 40% for testing) with subjects selected at random. More details are given in Chapter 3.

In our approach, the process of deriving the EEG biomarker from the TsEn, ΔEEG_A, and HFD methods was divided into training and testing phases. Dataset B was used for training (developing the KNN classification model) because of its larger size which meant it has more diversity and covered most of the problem space and dataset A with the remaining of dataset B were used for testing (testing the developed KNN classification model). The training dataset has 20 normal subjects and 12 AD patients

were selected at random from dataset B. The testing dataset has 12 normal subjects and 8 AD patients; 5 AD patients and 4 normal subjects were selected from dataset B that combined with the 3 AD patients and 8 normal subjects of dataset A. 10-fold cross validation was used for developing the KNN classification model. The EEG biomarkers were computed for the whole EEG record in time domain for all 19 EEG channels.

The following steps outline the procedure for deriving the biomarkers

1. Compute TsEn, ΔEEG_A, and HFD values for each the 19 EEG channel for the EEG dataset (i.e., Fp1, Fp2, F7, F3, FZ, F4, F8, T3, C3, CZ, C4, T4, T5, P3, PZ, P4, T6, O1, and O2).
2. Determine which features have a significant statistical association with AD. P-value between AD patients and normal subject for each of the TsEn, ΔEEG_A, and HFD method was computed.
3. For each of the candidate method, identify the EEG channels that show a significant separation between AD patients and normal subjects based on p-value analysis. EEG channels with the smallest p-values between AD patients and normal subject are selected to detect AD.
4. Select identified EEG channels for each of the candidate methods to construct an EEG panel as a feature vector that can be used in the classification between AD patients and normal subjects.
5. Use machine learning techniques to develop diagnostic models. The selected EEG channels from each method were used to develop a KNN classification model for each of the candidate methods. Therefore, three KNN classification models were developed.
6. The performance of each model was assessed based on its sensitivity, specificity, accuracy, f-measure, PPV, NPV, and ROC.

4.2.1. Tsallis entropy (TsEn) computation

TsEn values were computed for the 19 EEG channels of EEG dataset. Figure 4-1 shows the TsEn values for one AD patient and one normal subject for the 19 EEG channels. The mean values of TsEn were computed for AD patients and normal subjects as shown in Figure 4-2.

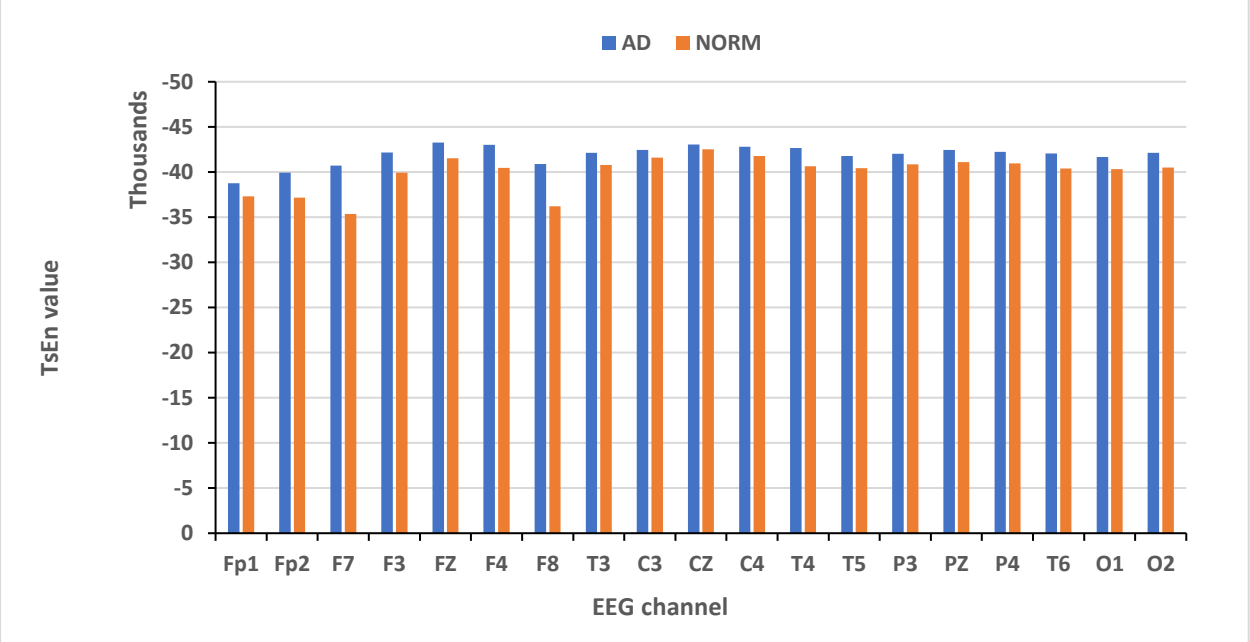

Figure 4-1: TsEn values for one AD patient and one normal subject

Figure 4-1 shows the actual TsEn values for one AD patient and one normal subject.

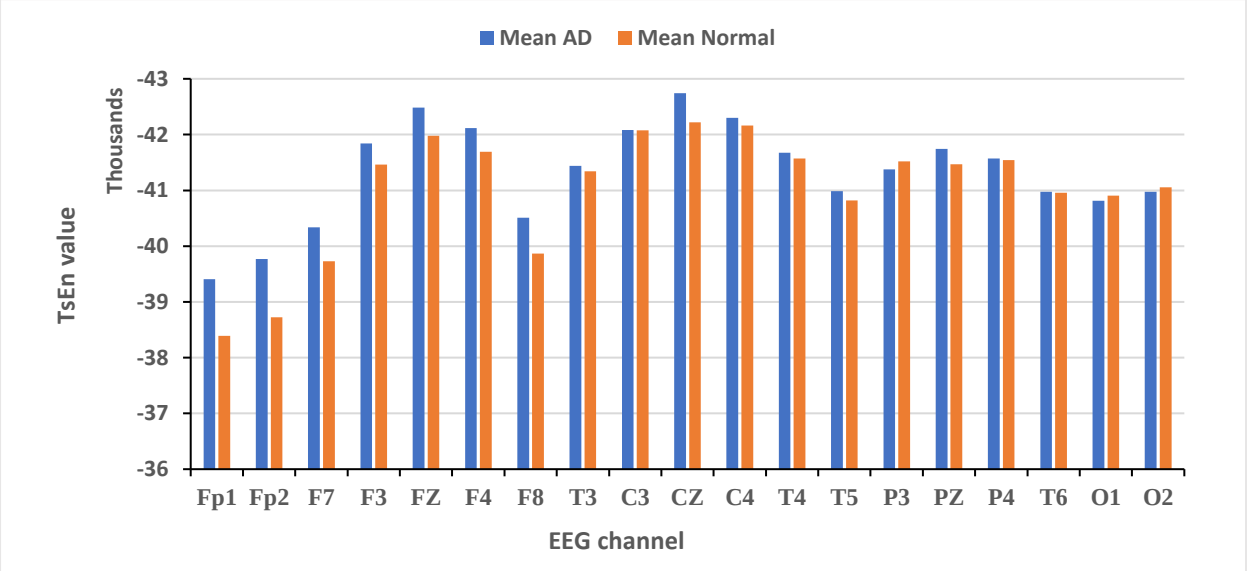

Figure 4-2: Mean TsEn for AD patients and normal subjects

Figure 4-2 shows the mean TsEn values for AD patients and normal subjects.

4.2.2. Changes in the EEG amplitude (ΔEEG_A) computation

The ΔEEG_A values were computed for the 19 EEG channels of EEG dataset. Figure 4-3 shows the ΔEEG_A values for one AD patient and one normal subject for the 19 EEG channels. The mean values of ΔEEG_A were computed for AD patients and normal subjects as shown in Figure 4-4.

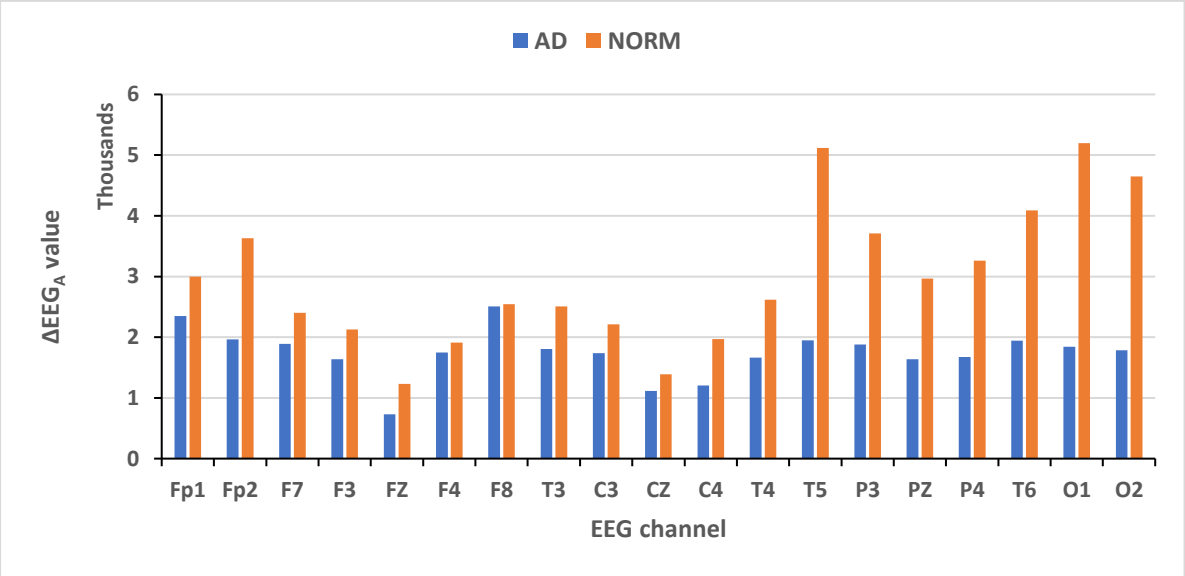

Figure 4-3: ΔEEG_A values for one AD patient and one normal subject

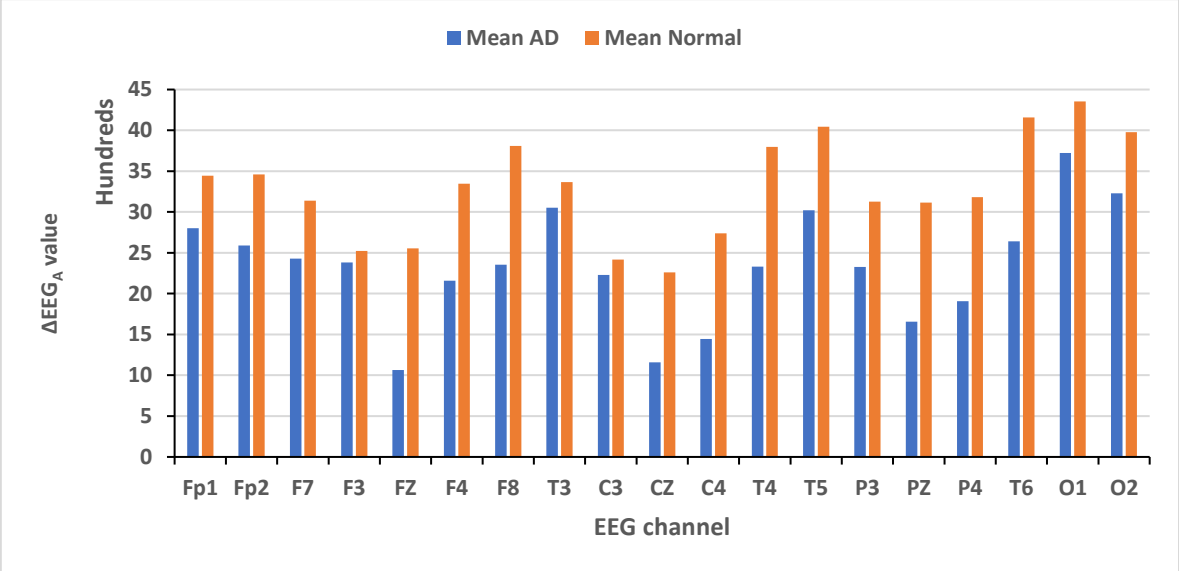

Figure 4-4: Mean ΔEEG_A for AD patients and normal subjects

4.2.3. Higuchi fractal dimension (HFD) computation

The HFD values were computed for the 19 EEG channel of EEG dataset. Figure 4-5 shows the HFD values for one AD patient and one normal subject for the 19 EEG channels.. The mean values of HFD were computed for AD patients and normal subjects as shown in Figure 4-6.

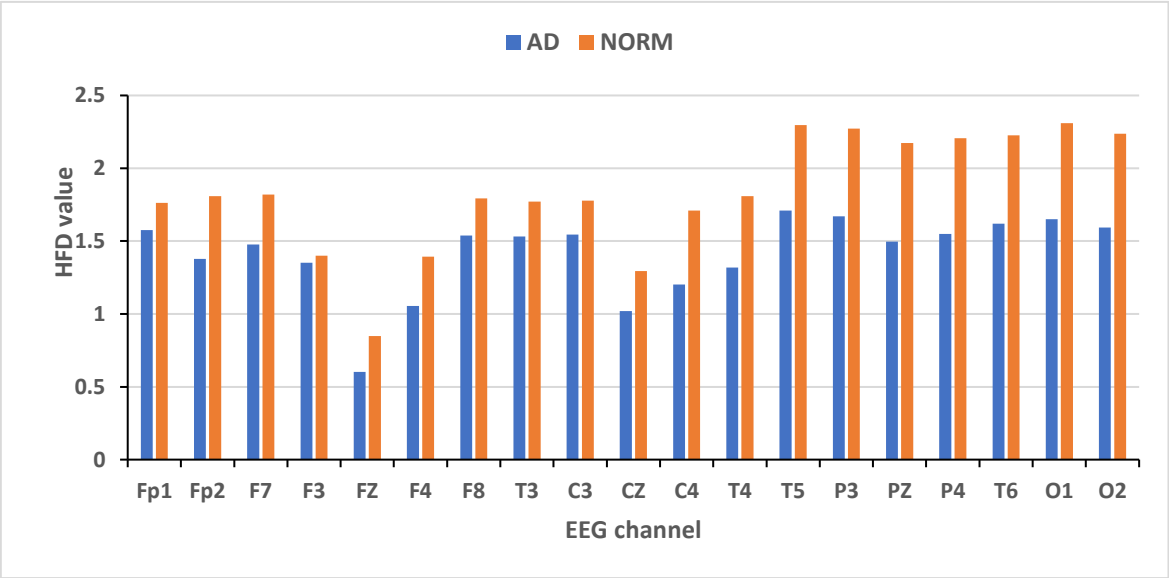

Figure 4-5: HFD values for one AD patient and one normal subject

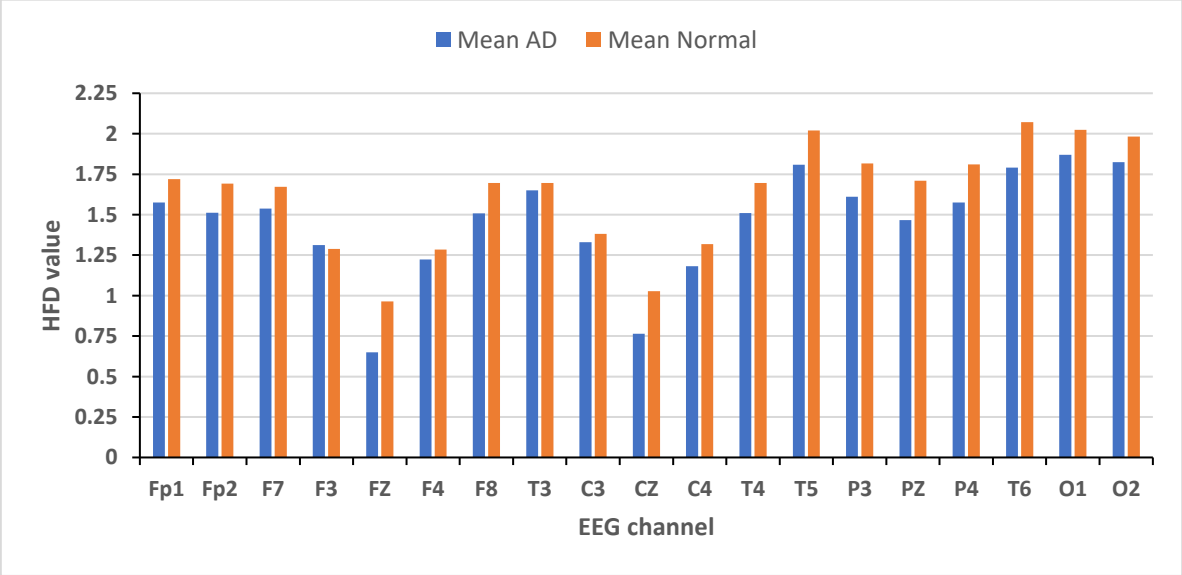

Figure 4-6: Mean HFD for AD patients and normal subjects

4.2.4. Biomarker selection

Figures 4-1, 4-3, and 4-5 show TsEn, ΔEEG_A, and HFD values for all 19 EEG channels in the training dataset. It can be seen that although there is significant overlap between

TsEn, ΔEEG_A, and HFD values between AD patients and normal subjects in some of EEG channels, not for all the channels. EEG channels have less overlap may be used to discriminate between AD patients and normal subjects.

Figures 4-2, 4-4, and 4-6 show the mean TsEn, ΔEEG_A, and HFD values of AD patients are lower than those in normal subjects. This reduction in TsEn, ΔEEG_A, and HFD values of EEG is because AD causes a reduction in neuronal activity of the brain and this may be reflected in the content of EEG signals [215][234]. It can be seen from those figures that there is little or no difference in mean of TsEn, ΔEEG_A, and HFD values for all 19 channels. This is because the changes in the EEG of AD are qualitatively similar to those in normal subjects [235]. It may therefore be difficult to develop a diagnosis model based on the 19 EEG channels. For example, the mean TsEn values for EEG channels T3, C3, and C4 show no discernible difference, but channels Fp1, Fp2, Fz, F8, Cz, and Pz showed a difference between AD patients and normal subjects as shown in Figure 4-2. Figure 4-4 shows ΔEEG_A values of EEG for all 19 channels, the mean ΔEEG_A values for channels F7, F8, and T3 show less discernible difference, but channels Fz, F4, Cz, and Pz show significant differences between AD patients and normal subjects. Of all the 19 EEG channels, the mean HFD values for channels F3, F4, and C3 showed the least discernible difference between AD patients and normal subjects whereas the mean values for channels Fp1, Fp, Fz, Cz, Pz, and T6 showed the most difference (see Figure 4-6). This suggests that it may therefore be possible to discriminate between AD patients and normal subjects based on the analysis of the biomarker values for the EEG channels for the candidate methods.

P-values between AD patients and normal subjects were analysed for all 19 EEG channels to identify the channels have a significant separation between AD patients and normal subjects. EEG channels have a minimum p-value were selected as a

feature vector that can be used to discriminate between AD patients and normal subjects. EEG channels that have the lowest p-value compared to other channels were selected. A threshold of 0.09 was set for the p-values for this based on the analysis of p-values between AD patients and normal subjects.

Figure 4-7 shows the p-values between AD patients and normal subjects for the training dataset for all 19 EEG channel of TsEn method.

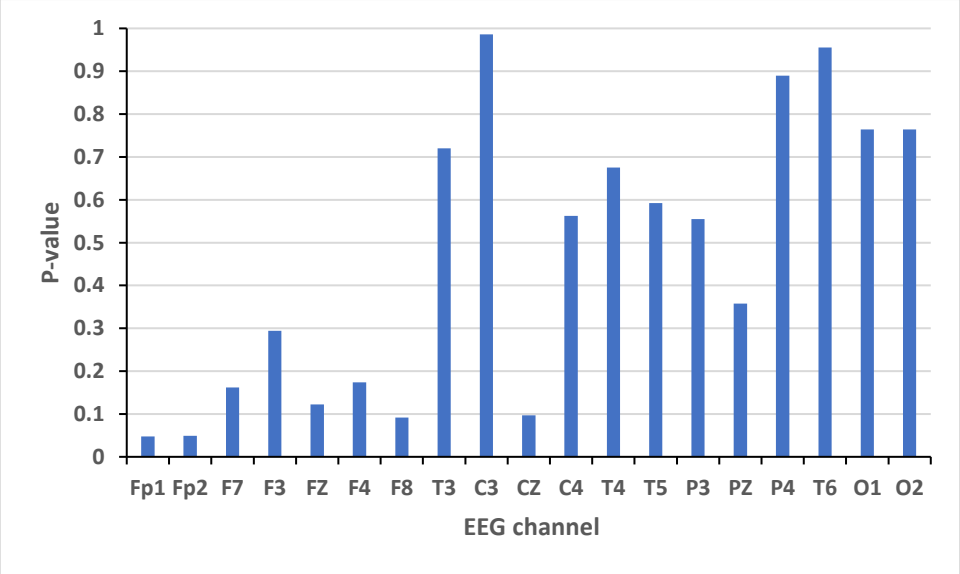

Figure 4-7: P-values between AD patients and normal subjects for all 19 EEG channels of TsEn method

Figure 4-7 shows EEG channels Fp1, Fp2, F8, and Cz have the minimum p-values between the two groups e.g., AD and normal, but other channels such as T3, and C3 have the maximum p-values.

Figure 4-8 shows the p-values between AD patients and normal subjects for training dataset for all 19 EEG channel of ΔEEG_A method.

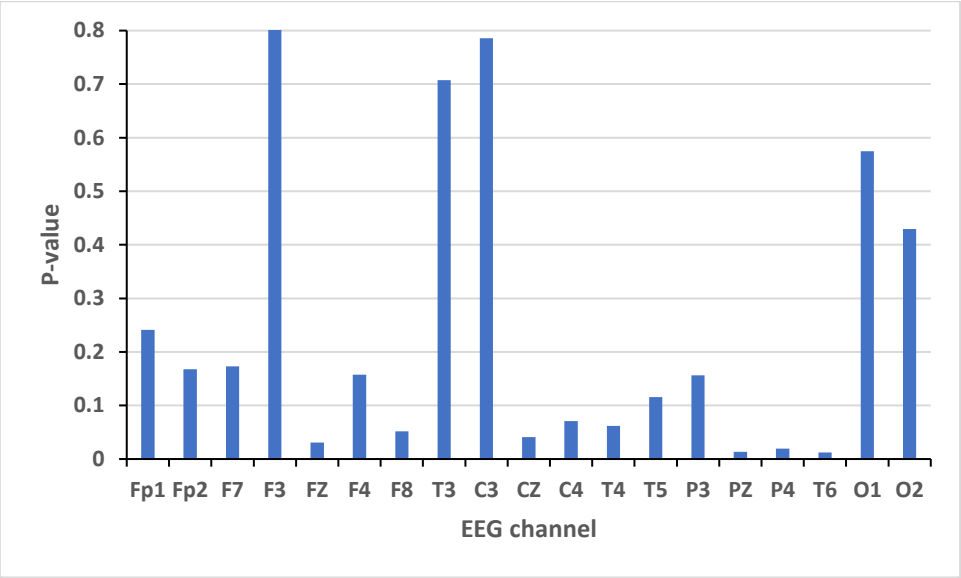

Figure 4-8: P-values between AD patients and normal subjects for all 19 EEG channels of ΔEEG_A method

Figure 4-8 shows channels Fz, F8, Cz, C4, T4, Pz, P4, and T6 have the minimum p-values between the two groups e.g., AD and normal, but other channels such as F3, T3, and C3 have the maximum p-values.

Figure 4-9 shows the p-values between AD patients and normal subjects for training dataset for all 19 EEG channels for the HFD method.

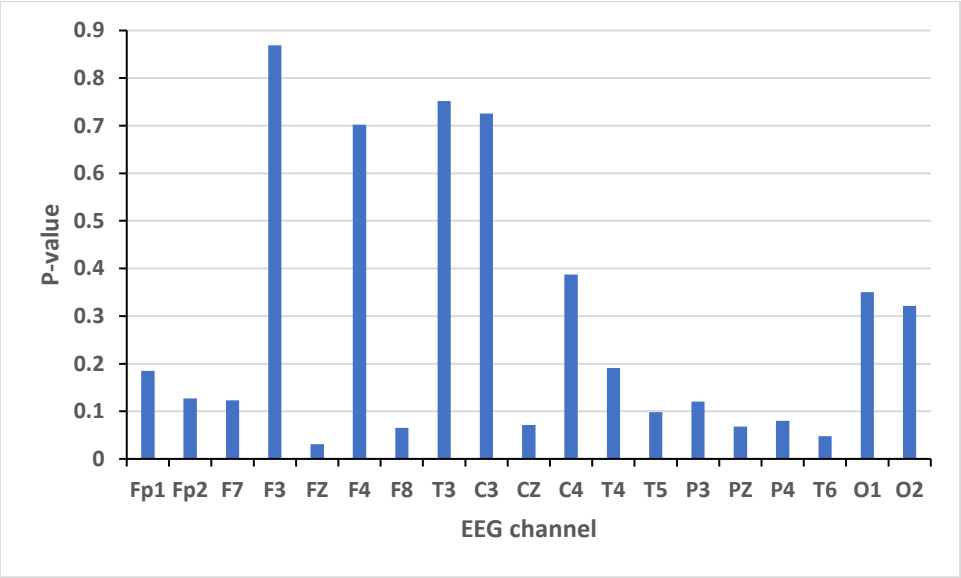

Figure 4-9: P-values between AD patients and normal subjects for all 19 EEG channels of HFD Method

Figure 4-9 shows channels Fz, F8, Cz, T5, Pz, P4, and T6 have the minimum p-values between the two groups e.g., AD and normal, but other channels such as F3, T4, and T3 have the maximum p-values.

4.3. Results

4.3.1. Tsallis Entropy (TsEn)

The EEG channels: Fp1, Fp2, F8, and Cz were selected in discriminating between AD patients and normal subjects e.g., TsEn(Fp1), TsEn(Fp2), TsEn(F8), and TsEn(Cz). Table 4-2 shows the performance of the KNN-based classification model of TsEn biomarkers for the whole EEG record for the 19 EEG channels. In this case, the sensitivity and specificity were 72.7% and 100%, respectively.

Table 4-2: Performance results of TsEn biomarker

Sensitivity	72.7%
Specificity	100%
Accuracy	85%
F-measure	84%
PPV	100%
NPV	75%

Figure 4-10 shows the ROC and AUC for the performance of the KNN classification model for TsEn biomarkers.

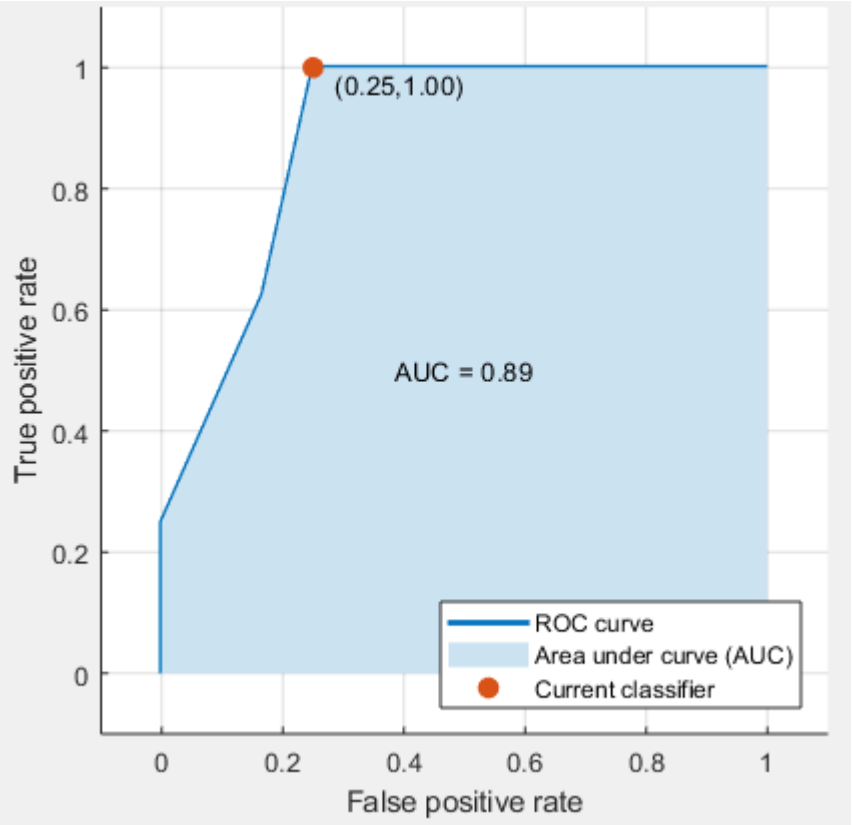

Figure 4-10: ROC and AUC for the performance of TsEn biomarkers

4.3.2. Changes in the EEG Amplitude (ΔEEG_A)

The EEG channels Fz, F8, Cz, C4, T4, Pz, P4, and T6 were selected to discriminate between AD patients and normal subjects e.g., $\Delta EEG_A(Fz)$, $\Delta EEG_A(F8)$, $\Delta EEG_A(Cz)$, $\Delta EEG_A(C4)$, $\Delta EEG_A(T4)$, $\Delta EEG_A(Pz)$, $\Delta EEG_A(P4)$, and $\Delta EEG_A(T6)$. Table 4-3 shows the performance of the KNN-based classification model of ΔEEG_A biomarkers for the whole EEG record for the 19 EEG channels. In this case, the Sensitivity and Specificity were 87.5% and 91.6%, respectively.

Table 4-3: Performance results of ΔEEG_A biomarker

Sensitivity	87.5%
Specificity	91.6%
Accuracy	90.0%
F-measure	87.5%
PPV	87.5%
NPV	91.6%

Figure 4-11 shows the ROC and AUC for the performance of the KNN classification model for ΔEEG_A biomarkers.

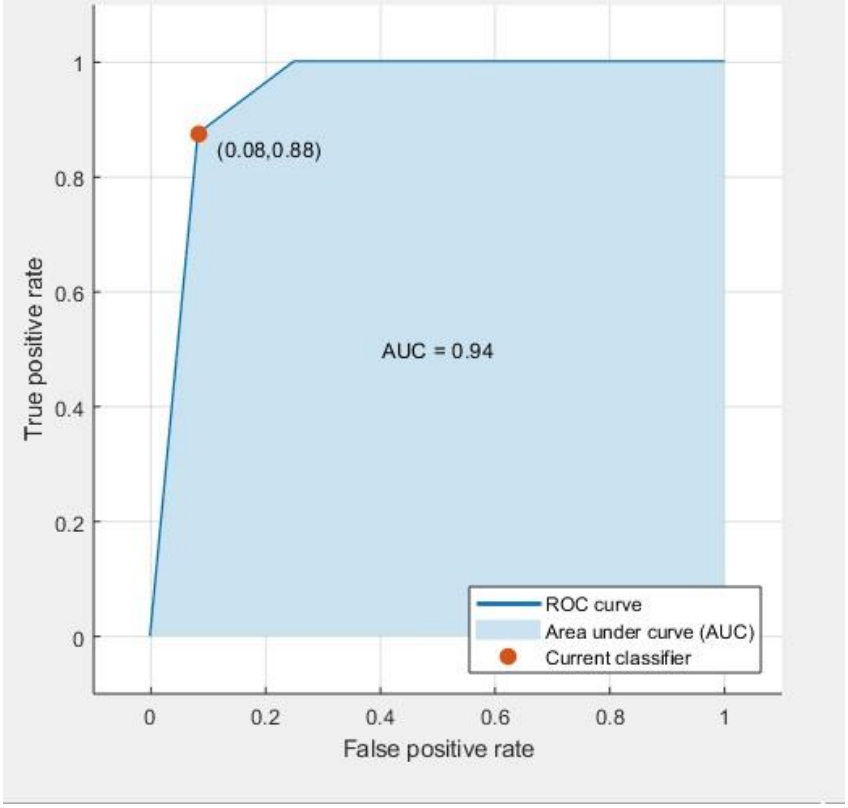

Figure 4-11: ROC and AUC for the performance of ΔEEG_A biomarkers

4.3.3. Higuchi Fractal Dimension (HFD)

The EEG channels Fz, F8, Cz, T5, Pz, P4, and T6 were selected to discriminate between AD patients and normal subjects e.g., HFD(Fz), HFD(F8), HFD(Cz), HFD(T5), HFD(Pz), HFD(P4), and HFD(T6). Table 4-4 shows the performance of the KNN-based classification model of ΔEEG_A biomarkers for the whole EEG record for the 19 EEG channels. In this case, the sensitivity and specificity were 77.7% and 90.9%, respectively.

Table 4-4: Performance results of HFD biomarker

Sensitivity	77.7%
Specificity	90.9%
Accuracy	85%
F-measure	82.3%
PPV	87.5%
NPV	83.3%

Figure 4-12 shows the ROC and AUC for the performance of the KNN classification model for HFD biomarkers.

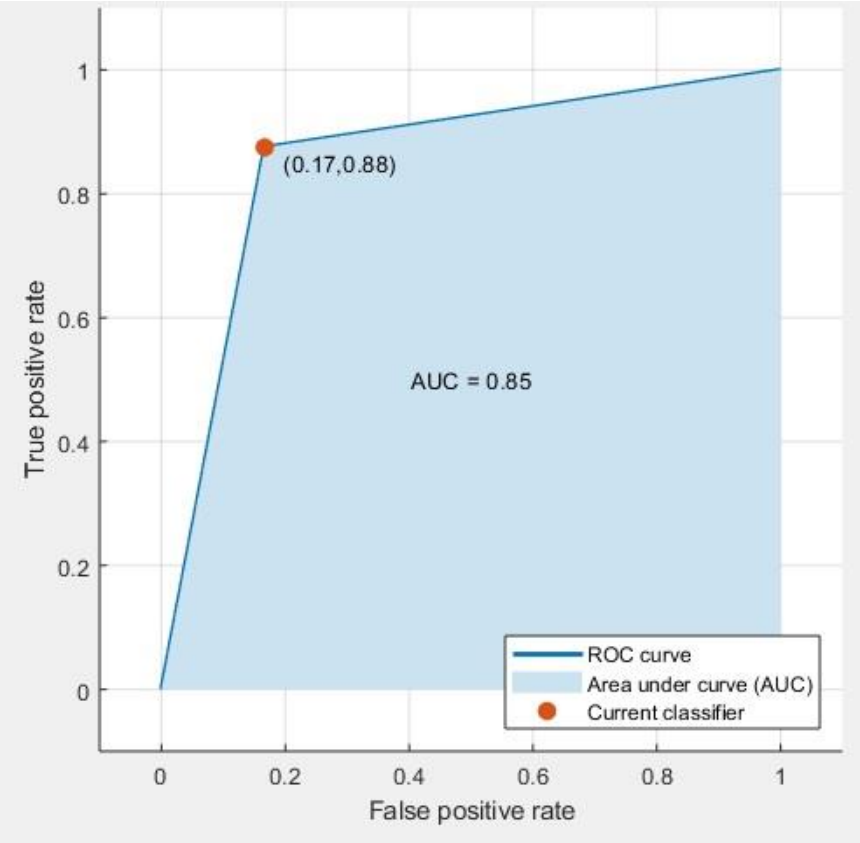

Figure 4-12: ROC and AUC for the performance of HFD biomarkers

4.4. Discussions

AD causes a reduction in the neuronal activity of the brain [95] resulting in a decreased ability to process information [96][97][98] which may be reflected in EEG signals [95]. The EEG signal was investigated in AD detection based on the analysis of the

complexity and slowing in EEG. The slowing of the EEG is one of the most consistent features which is relating to the detection of AD. Slowing may therefore be quantified as a biomarker of AD [15][48][210]. Furthermore, AD patients exhibit a significant reduction in EEG complexity [48][50][84][85][95][99][100]. In this chapter, the whole EEG record was investigated in the time domain. TsEn and HFD were used to compute the complexity and ΔEEG_A was used to compute the slowing in EEG. TsEn, ΔEEG_A, and HFD were computed for all 19 EEG channels. P-value was used between AD patients and normal subjects for all 19 EEG channels and it was used as criteria to identify the EEG channels that have a significant association with AD. For TsEn, four EEG channels were selected i.e., Fp1, Fp2, F8, and Cz. For ΔEEG_A, eight EEG channels were selected i.e., Fz, F8, Cz, C4, T4, Pz, P4, and T6. For HFD, seven EEG channels were selected i.e., Fz, F8, Cz, T5, Pz, P4, and T6. Most of the channels are located at the back part of the brain and this is compatible with the findings in other studies that demonstrated that the gradual slowing of the brain wave activity caused by AD starts from the back of the brain and then spreads out to other parts over time [210][236][237]. The candidate EEG channels from the 19 EEG channels were four for TsEn, seven for ΔEEG_A, and eight for HFD have proved the hypothesis that it is possible to identify the EEG channels that can be used to detect AD. The selected channels were then used to construct a KNN classification model for each method. The results suggest that AD can be detected with high performance using a few EEG channels. The performance for TsEn values of EEG was 72.7% for sensitivity and 100% for specificity, for ΔEEG_A it was 87.5% for sensitivity and 91.6% for specificity, and for HFD it was 77.7% for sensitivity and 90.9% for specificity. These results are consistent with those of others and show that EEG based biomarkers can play a potentially valuable role in the early detection of AD [15][23][24][42][43][44][50].

4.5. Summary

Three promising EEG analysis methods were investigated to detect AD based on the analysis of EEG signal. The results suggest that the EEG signal could be used to detect AD with a small number of channels which have a significant association with AD. Also, the results suggest that Tsallis entropy-based biomarkers, ΔEEG_A based biomarkers, and HFD-based biomarkers are suitable for the detection of dementia. At this level of performance, these approaches may serve as a basis for a first line of a decision support tool for detection of dementia. In addition, the results suggest that TsEn, ΔEEG_A , and HFD biomarkers are promising biomarkers that identify the EEG channels thought to be affected first by AD and so could be used to detect subjects at high risk of dementia. As AD subjects have significantly lower TsEn, ΔEEG_A, and HFD values, this provides an effective way to discriminate between AD patients and normal subjects. Potentially, the complexity and slowing of the EEG could be markers for the subsequent rate of cognitive and functional decline in AD patients.

Chapter 5. Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease

5.1. Introduction

Several studies have investigated EEG complexity as a potential AD biomarker using whole EEG record with the objective of achieving a high performance. Given the association of EEG activities (e.g. alpha, delta activities) with AD, we hypothesised that complexity measures based on the EEG frequency bands would provide better results than those derived directly from the whole EEG record. The aim was to enhance the performance of the complexity measures and to demonstrate their usefulness in quantifying changes in EEG due to AD.

This is a cross-sectional study aimed at demonstrating the usefulness of EEG-based complexity measures to detect AD. In this chapter, we investigated an important class of complexity measures, information theoretic methods, which offers a potentially powerful approach for quantifying changes in the EEG due to AD [220]. Information theoretic methods (i.e*.,* TsEn and LZC) have emerged as a potentially useful complexity-based approach to derive robust EEG biomarkers of AD [220][102][105] [90][238][239]. They are attractive because of the potential natural link between information theory-based biomarkers and changes in the brain caused by AD [220]. Conceptually, information processing activities in the brain are thought to be reflected in the information content of the EEG.

It is widely accepted that AD causes a decrease in the power of high frequencies (alpha, beta, and gamma) and an increase in the power of low frequencies (delta and theta) [15][43][44][48][85].

Digital filters were used to extract the five EEG frequency bands (i.e., delta, theta, alpha, beta, and gamma). Complexity measures were then obtained for each of the

five EEG frequency bands and for each channel using each of the three methods of computing complexity measures (TsEn, HFD, and LZC).

For each method, we computed a panel of 114 biomarkers (i.e., 19 biomarkers for the whole EEG record, and 19 biomarkers for each of the five EEG frequency bands). The performance measures for each biomarker was computed (including the sensitivity and specificity).

5.2. Methodology

The following steps outline the procedure that was used to derive the biomarkers for the three complexity methods (i.e., TsEn, HFD, and LZC)

1. The EEG signal was filtered using infinite impulse response (IIR) Chebyshev-II band pass filter into five frequency bands (i.e., delta 0-4Hz, theta 4-8Hz, alpha 8-12Hz, beta 12-30Hz, and gamma 30-45Hz). A low computational IIR filter was used to retain the computational efficiency of the derived complexity-based biomarkers [206].
2. The biomarkers were then derived first from the whole EEG record and then for each of EEG frequency bands for each of the three EEG complexity methods.
3. For each biomarker of the EEG complexity methods (i.e., TsEn, HFD, and LZC), p-values were computed between AD patients and normal subjects using student t-test.
4. The performance of each complexity measure to detect AD is then assessed. For each complexity measure, a classification model, based on the support vector machine (SVM), was used to detect AD.

A panel of 114 biomarkers was computed (19 biomarkers for the whole EEG record and 19 biomarkers for each of EEG frequency band (i.e., delta, theta, alpha, beta, and gamma). To determine which features have a significant statistical association with AD, we computed p-values between AD patients and normal subjects using

student t-test. This allowed us to identify significant features that may be useful to discriminate between AD patients and normal subjects. The EEG data used in this chapter consists of two datasets (A and B). Dataset A includes 11 age-matched subjects over 65 years old (3 AD patients, and 8 normal subjects). Dataset B includes 41 subjects that were not perfectly age-matched with 24 normal subjects (10 males and 14 females) and 17 were probable AD patients (9 males and 8 females). The dataset was split into training and testing data (60% for training and 40% for testing) with subjects selected at random. For the training dataset, 1-AD patients from dataset A and 3-AD patients from dataset B, and 2-normal subjects from dataset A and 10-normal subjects from sub dataset B were selected. But for the testing dataset, 2-AD patients from dataset A and 10-AD patients from dataset B, and 6-normal subjects from dataset A and 14-normal subjects from dataset B were selected. P-values were computed using the training EEG dataset. Machine learning techniques were used to develop models based on biomarkers. As a classifier, we used support vector machine (SVM) to model biomarkers extracted using TsEn, HFD, and LZC methods. SVM classifier was

used because it is widely used in machine learning and has found application in dementia diagnosis. It has shown better performance in biomedical data analysis and in automatic AD diagnosis compared to other conventional classifiers (e.g., Euclidean distance classifier), good capability to learn from experimental data [240][241], and it has a stable classification performance [242]. It has also been shown to outperform other machine learning techniques (e.g., Naive Bayes, Multilayer Perceptron, Bayes Network, egging, Logistic Regression, Random Forest,) in diagnosis of MCI and dementia [243]. We used the testing EEG dataset to test the performance of the models. For each complexity method, six performance tables were created (whole EEG record, and table for each EEG frequency band).

The performance of the TsEn, HFD, and LZC biomarkers for AD diagnosis was assessed in terms of Sen, Spec, ACC, F-measure, FPR, FNR, PPV, and NPV. MCC was computed to measure the quality of the binary classification (AD and normal) between the actual and predicted results [173][174].

5.3. Results

We analysed the performance of the three different complexity measures in quantifying changes in EEG due to AD. For this purpose, we examined the differences between the values of the complexity measures derived from EEG signals of AD subjects and those of normal subjects. Biomarkers that do not show significant differences between AD patients and normal subjects may not be suitable for quantifying changes in EEG due to AD as they may not be capable of being used to discriminate between AD and normal subjects.

We found that complexity measures derived from the EEG frequency bands for AD patients were significantly different to those of normal subjects compared to complexity measures derived from the whole EEG record. This suggests that they may be better suited to quantifying changes in the EEG due to AD and potentially may provide better results in AD diagnosis.

Figure 5-1 shows the EEG biomarkers derived from whole EEG record (i.e. unfiltered) and those derived from the five EEG bands (delta, theta, alpha, beta, and gamma bands) using the TsEn method. The results show that TsEn values for AD patients are lower than those for normal subjects for the whole EEG record. This is consistent with the findings in other studies [45][220][102][107]. Figure 5-1 also shows that the differences between the TsEn values for AD patients and for normal subjects for the EEG bands (delta and theta bands in particular) are larger than those for whole EEG record. This is a desirable feature in a biomarker as it suggests that TsEn biomarkers

derived the EEG bands may provide better performance in detecting AD than those from whole EEG record.

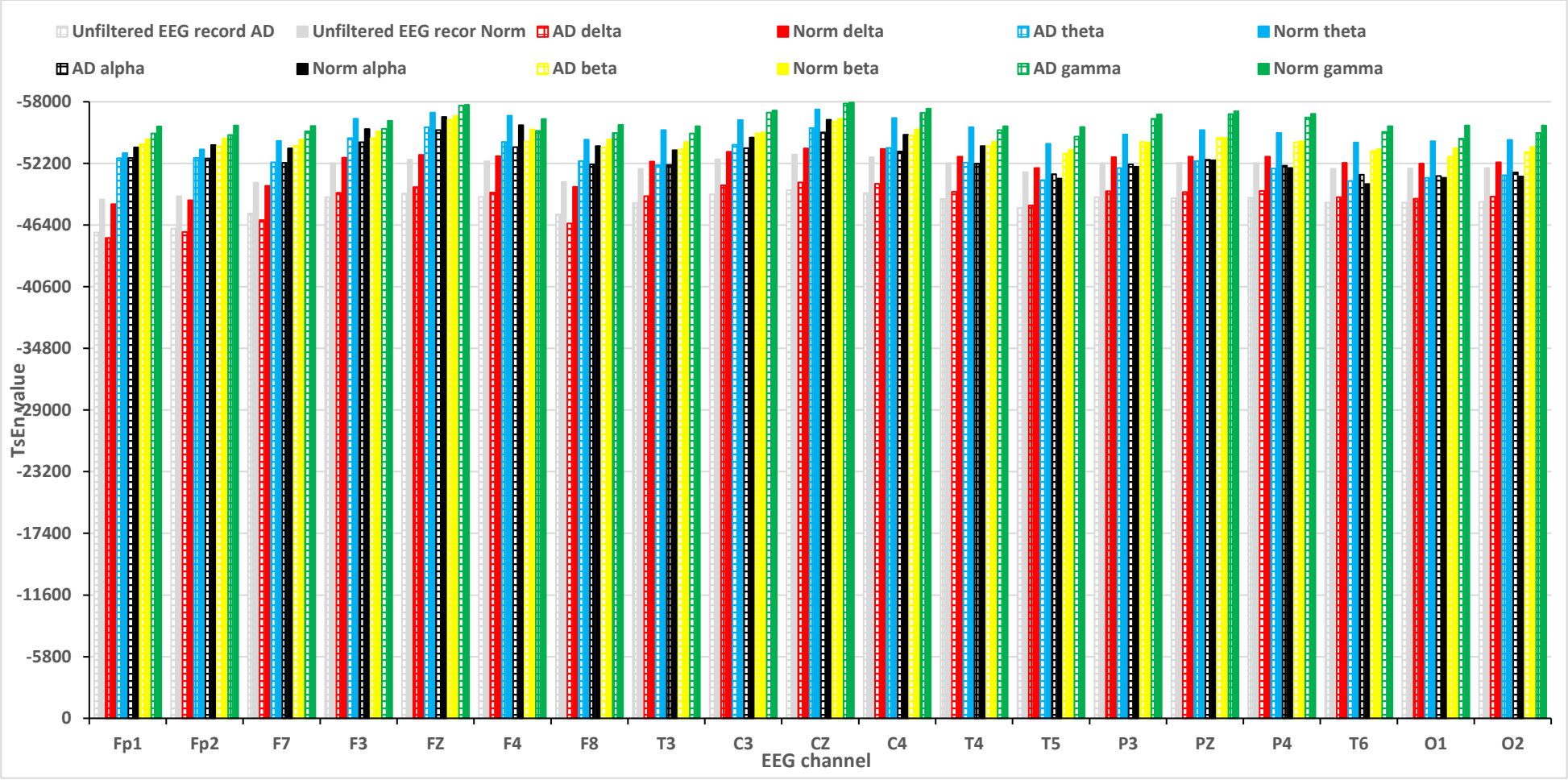

Figure 5-1: EEG biomarkers for TsEn

Figure 5-2 shows the EEG biomarkers derived from whole EEG record and those derived from the five EEG bands (delta, theta, alpha, beta, and gamma bands) using the HFD method. In this case, the results show that HFD values for AD patients are lower than those for normal subjects. This result is consistent with the finding in other studies [49][125]. As with the TsEn, the differences between HFD biomarkers for AD patients and normal subjects for the five EEG frequency bands (delta, theta, and alpha bands) were larger than those for the whole EEG record suggesting that the use of biomarkers derived from the frequency bands would be better at detecting AD than the use of whole EEG record.

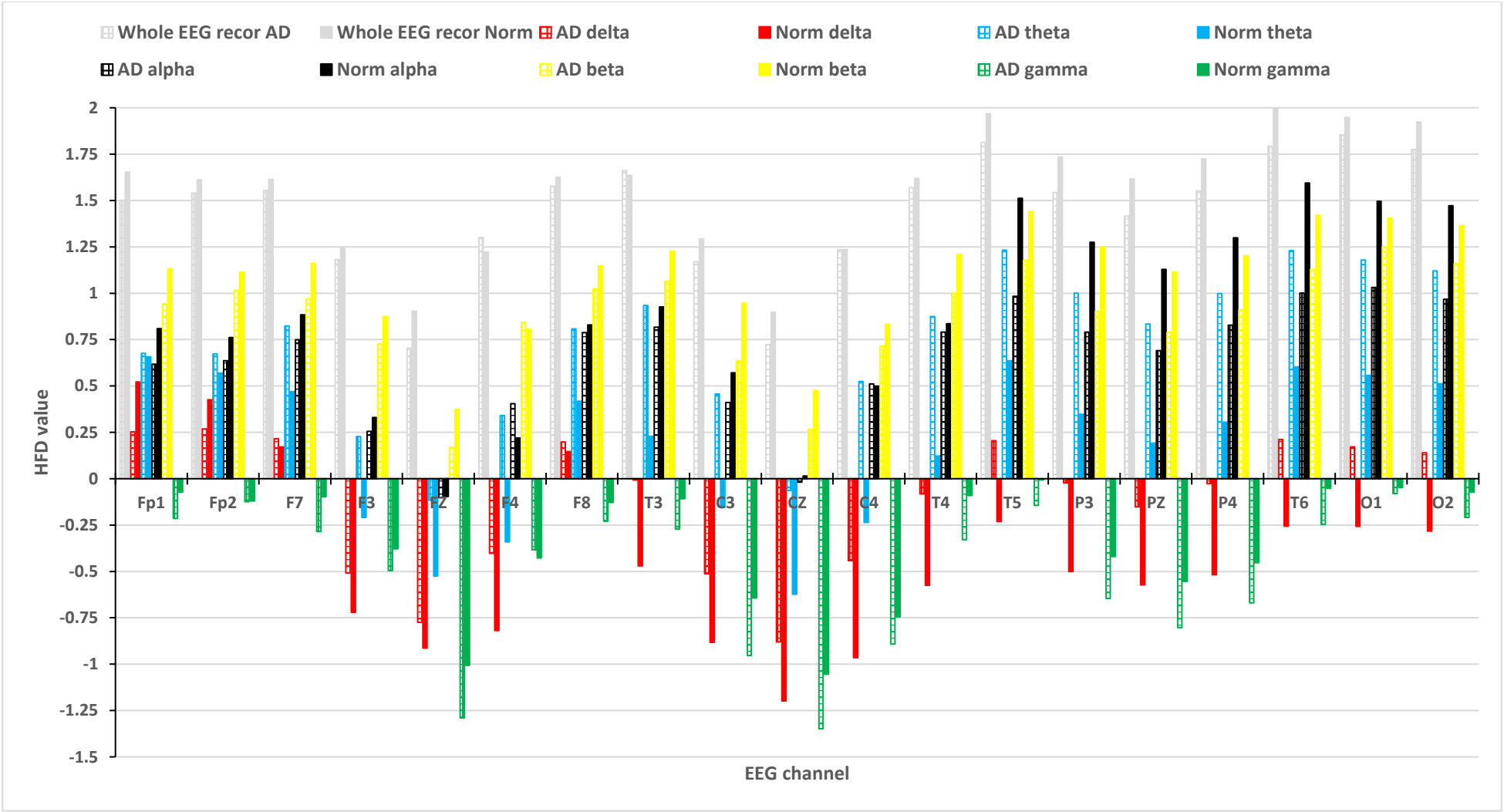

Figure 5-2: EEG biomarkers for HFD

Figure 5-3 shows similar results for the LZC method. In this case, the results show that LZC values for AD patients were lower than those for normal subjects and these are consistent with the finding in other studies [86][223]. Again, the differences between the LZC biomarkers for AD patients and normal subjects for the five EEG frequency bands (the theta, beta and gamma bands, in particular) were larger those for the whole EEG record, suggesting that the use of biomarkers derived from the frequency bands would be better at detecting AD than the use of whole EEG record.

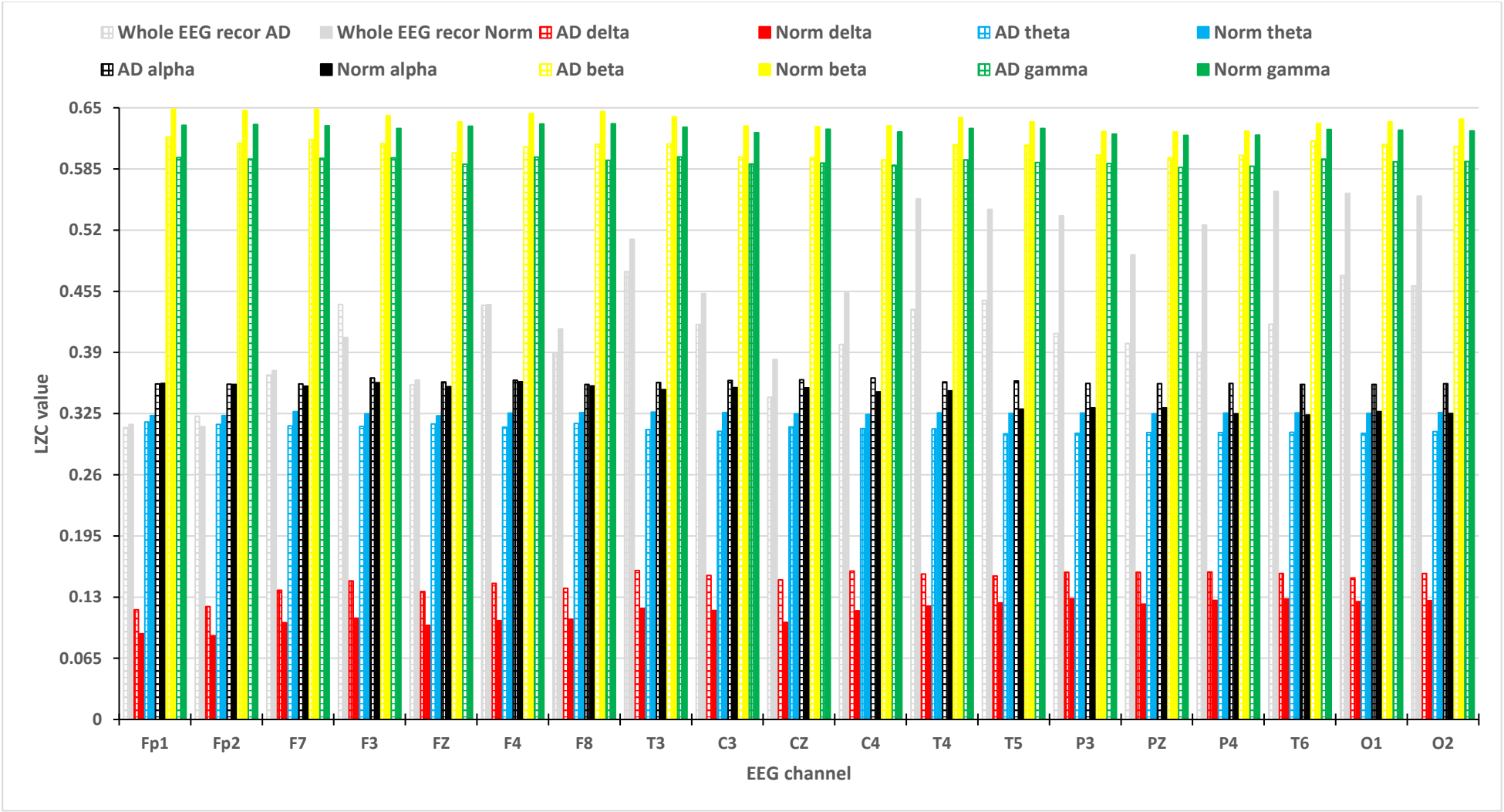

Figure 5-3: EEG biomarkers for LZC

We analysed the complexity measures using p-values to determine the statistical significance in detecting AD

Figure 5-4 shows p-values of the differences in TsEn measures between AD patients and normal subjects for the whole EEG record and those from the EEG frequency bands. The results show that TsEn biomarkers that were extracted from theta bands have the smallest p-values while the TsEn biomarkers derived from gamma band have the maximum p-value between AD patients and normal subjects. This suggests that biomarkers that are extracted from theta band may provide the best performance in AD diagnosis. Figure 5-4 also shows that biomarkers that were extracted from EEG frequency bands may have a more significant association with AD than the EEG biomarkers that are derived from whole EEG record based on p-value analysis. Therefore, the complexity measures derived from the EEG frequency band may provide better results in the classification between AD patients and normal subjects.

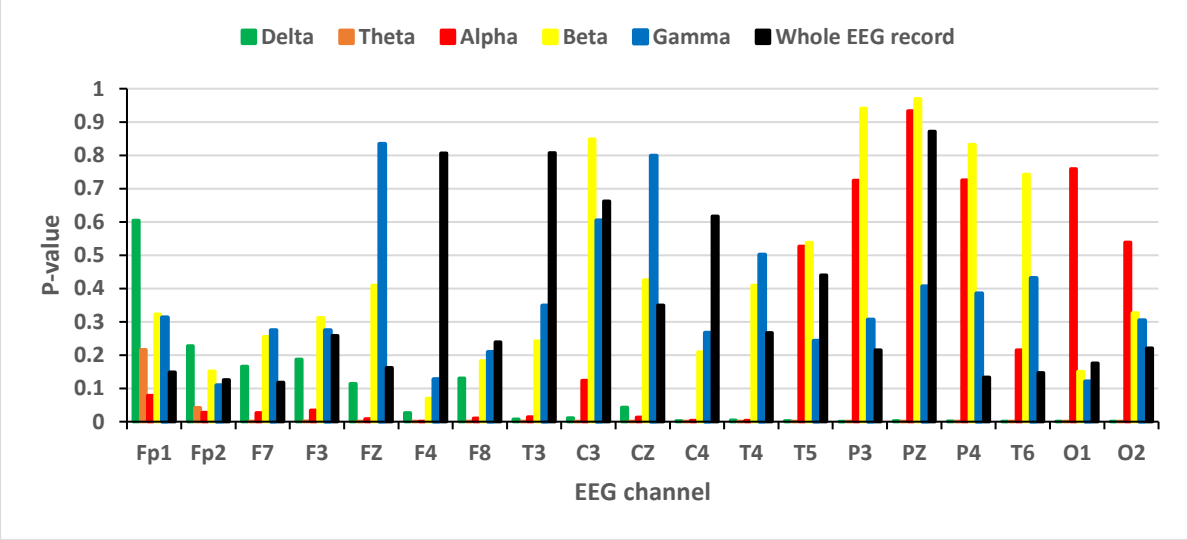

Figure 5-4: P-values for TsEn between AD patients and normal subjects of the training EEG dataset

Figures 5-5 and 5-6 depict the results of similar p-value analysis for HFD and LZC measures, respectively. The results show that in both HFD and LZC methods, the complexity measures derived from the EEG frequency bands theta band have

significantly smaller p-values compared to those of measures derived from the whole EEG record. In both methods, complexity measures derived from the theta band gave the smallest p-value. This implies that biomarkers derived from the frequency bands, the theta band in particular, may provide the best possible performance in AD diagnosis using the HFD and LZC methods.

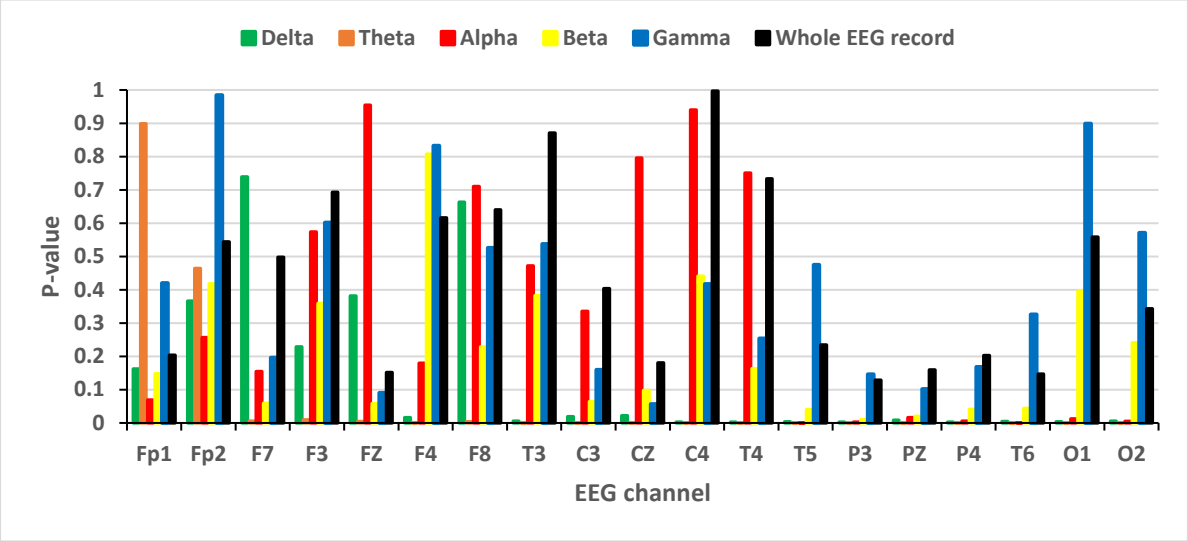

Figure 5-5: P-values for HFD between AD patients and normal subjects of the training EEG dataset

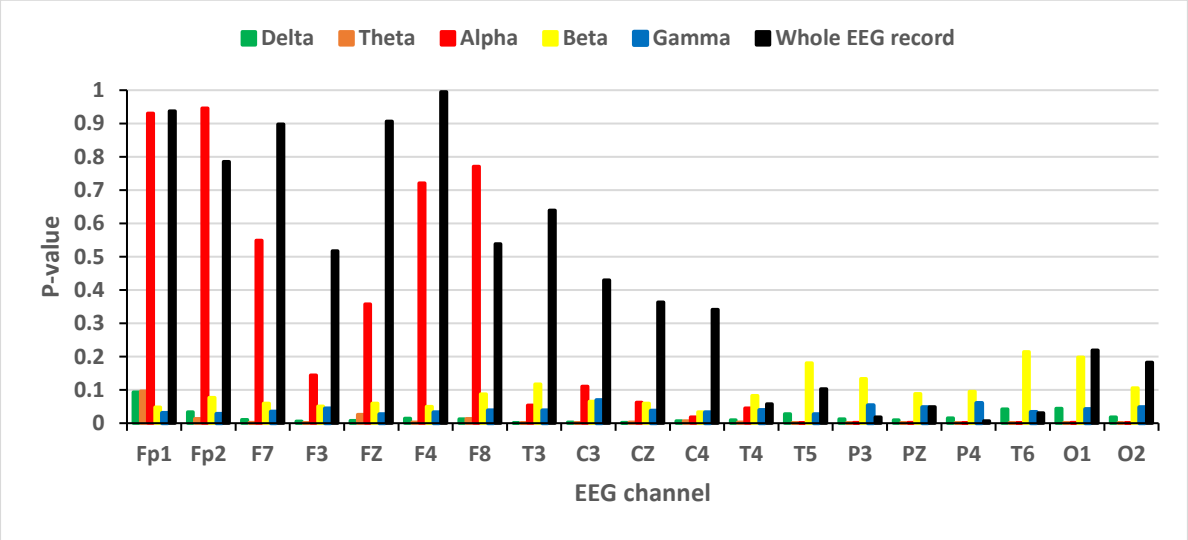

Figure 5-6: P-values for LZC between AD patients and normal subjects of the training EEG dataset

Looking across all the results (Figure 5-4, Figure 5-5, and Figure 5-6), the theta band has a minimum p-value between AD patients and normal subjects for all three

complexity methods (i.e., TsEn, HFD, and LZC). Thus, EEG biomarkers derived from EEG frequency bands are better than the biomarkers were extracted from whole EEG record. The biomarkers derived from theta band may provide the best performance in AD diagnosis across all three methods.

5.3.1. The performance of the EEG complexity-based measures

Table 5-1 shows the performance of the SVM-based classification model using TsEn biomarkers for whole EEG record for the 19 EEG channels. In this case, the best Sensitivity and Specificity were 46.67% and 80%, respectively, for Fp2 and F7 EEG channels.

Similar performance indices were computed for each of the five EEG bands using the TsEn. As an example, Table 5-2 shows the performance indices for TsEn biomarkers for the delta band for the 19 EEG channels. The best Sensitivity and Specificity were 85.71%, and 84.62%, respectively, for T4, O1, and O2 EEG channels.

Similar performance indices were computed for each of the five EEG bands using HFD and LZC methods. Table 5-3 summarises the best performance indices for the three complexity measures.

Table 5-1: TsEn performance for whole EEG record

EEG channel	Sen. %	Spec. %	Acc. %	F-measure %	MCC	FPR %	FNR %	PPV %	NPV %
Fp1	43.75	75.00	50.00	58.33	0.153	25.00	56.25	87.50	25.00
Fp2	**46.67**	**80.00**	**55.00**	**60.87**	**0.236**	**20.00**	**53.33**	**87.50**	**33.33**
F7	**46.67**	**80.00**	**55.00**	**60.87**	**0.236**	**20.00**	**53.33**	**87.50**	**33.33**
F3	43.75	75.00	50.00	58.33	0.153	25.00	56.25	87.50	25.00
FZ	44.44	100.00	50.00	61.54	0.272	0.00	55.56	100.00	16.67
F4	44.44	100.00	50.00	61.54	0.272	0.00	55.56	100.00	16.67
F8	44.44	100.00	50.00	61.54	0.272	0.00	55.56	100.00	16.67
T3	37.50	50.00	40.00	50.00	-0.102	50.00	62.50	75.00	16.67
C3	35.71	50.00	40.00	45.45	-0.134	50.00	64.29	62.50	25.00
CZ	42.11	100.00	45.00	59.26	0.187	0.00	57.89	100.00	8.33
C4	44.44	100.00	50.00	61.54	0.272	0.00	55.56	100.00	16.67
T4	35.29	33.33	35.00	48.00	-0.229	66.67	64.71	75.00	8.33
T5	33.33	50.00	40.00	40.00	-0.167	50.00	66.67	50.00	33.33
P3	28.57	33.33	30.00	36.36	-0.356	66.67	71.43	50.00	16.67
PZ	37.50	50.00	40.00	50.00	-0.102	50.00	62.50	75.00	16.67
P4	35.71	50.00	40.00	45.45	-0.134	50.00	64.29	62.50	25.00
T6	26.67	20.00	25.00	34.78	-0.471	80.00	73.33	50.00	8.33
O1	27.27	44.44	35.00	31.58	-0.287	55.56	72.73	37.50	33.33
O2	30.00	50.00	40.00	33.33	-0.204	50.00	70.00	37.50	41.67

Table 5-2: TsEn Performance for delta band of the EEG signal

EEG channel	Sen. %	Spec. %	Acc. %	F-measure %	MCC	FPR %	FNR %	PPV %	NPV %
Fp1	50.00	66.67	60.00	50.00	0.167	33.33	50.00	50.00	66.67
Fp2	50.00	62.50	60.00	33.33	0.102	37.50	50.00	25.00	83.33
F7	55.56	72.73	65.00	58.82	0.287	27.27	44.44	62.50	66.67
F3	80.00	73.33	75.00	61.54	0.471	26.67	20.00	50.00	91.67
FZ	50.00	62.50	60.00	33.33	0.102	37.50	50.00	25.00	83.33
F4	50.00	61.11	60.00	20.00	0.068	38.89	50.00	12.50	91.67
F8	57.14	69.23	65.00	53.33	0.257	30.77	42.86	50.00	75.00
T3	71.43	76.92	75.00	66.67	0.471	23.08	28.57	62.50	83.33
C3	60.00	66.67	65.00	46.15	0.236	33.33	40.00	37.50	83.33
CZ	100.00	63.16	65.00	22.22	0.281	36.84	0.00	12.50	100.00
C4	71.43	76.92	75.00	66.67	0.471	23.08	28.57	62.50	83.33
T4	**85.71**	**84.62**	**85.00**	**80.00**	**0.685**	**15.38**	**14.29**	**75.00**	**91.67**
T5	80.00	73.33	75.00	61.54	0.471	26.67	20.00	50.00	91.67
P3	75.00	83.33	80.00	75.00	0.583	16.67	25.00	75.00	83.33
PZ	100.00	75.00	80.00	66.67	0.612	25.00	0.00	50.00	100.00
P4	83.33	78.57	80.00	71.43	0.579	21.43	16.67	62.50	91.67
T6	83.33	78.57	80.00	71.43	0.579	21.43	16.67	62.50	91.67
O1	**85.71**	**84.62**	**85.00**	**80.00**	**0.685**	15.38	14.29	75.00	91.67
O2	**85.71**	**84.62**	**85.00**	**80.00**	**0.685**	**15.38**	**14.29**	**75.00**	**91.67**

Table 5-3: Summary of the best performance indices for the three complexity measures

Method	TsEn		HFD		LZC
Feature	Delta	Theta	Theta	Alpha	Theta
EEG channel	T4, O1, O2	F4	C4	T5, P3	C3
Sen. %	85.71	85.71	66.67	66.67	100
Spec. %	84.62	84.62	100	100	92.31
Acc. %	85	85	80	80	95
F-measure %	80	80	80	80	93.33
MCC	0.685	0.685	0.667	0.667	0.9
FPR %	15.38	15.38	0	0	7.69
FNR %	14.29	14.29	33.33	33.33	0
PPV%	75	75	100	100	87.5
NPV%	91.67	91.67	66.67	66.67	100

Figures 5-7, 5-8, and 5-9 summarise the performance indices of the TsEn, HFD and LZC methods.

The results show that TsEn, HFD and LZC EEG biomarkers derived from the EEG frequency bands provide better performance than EEG biomarkers derived from the whole EEG record.

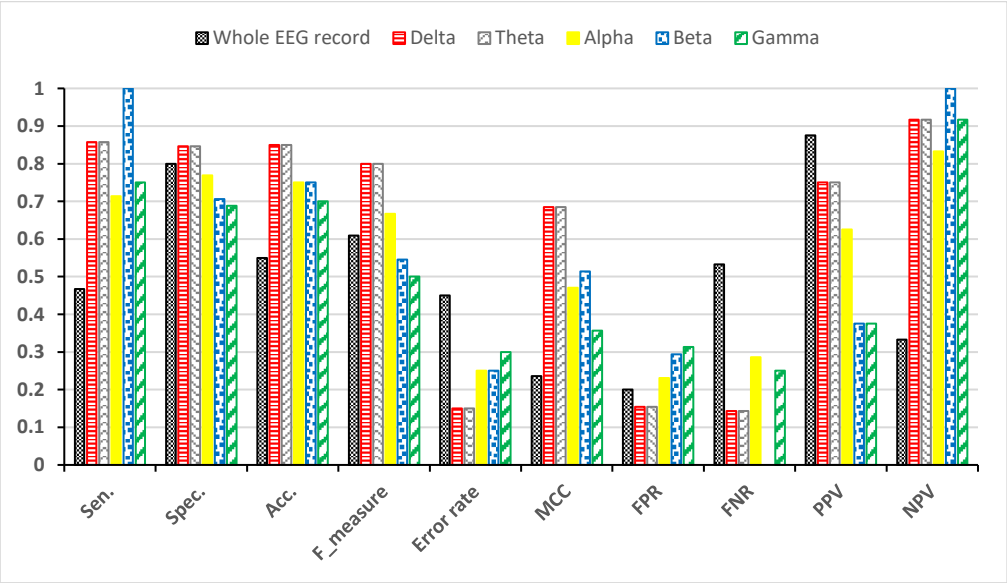

Figure 5-7: TsEn performance

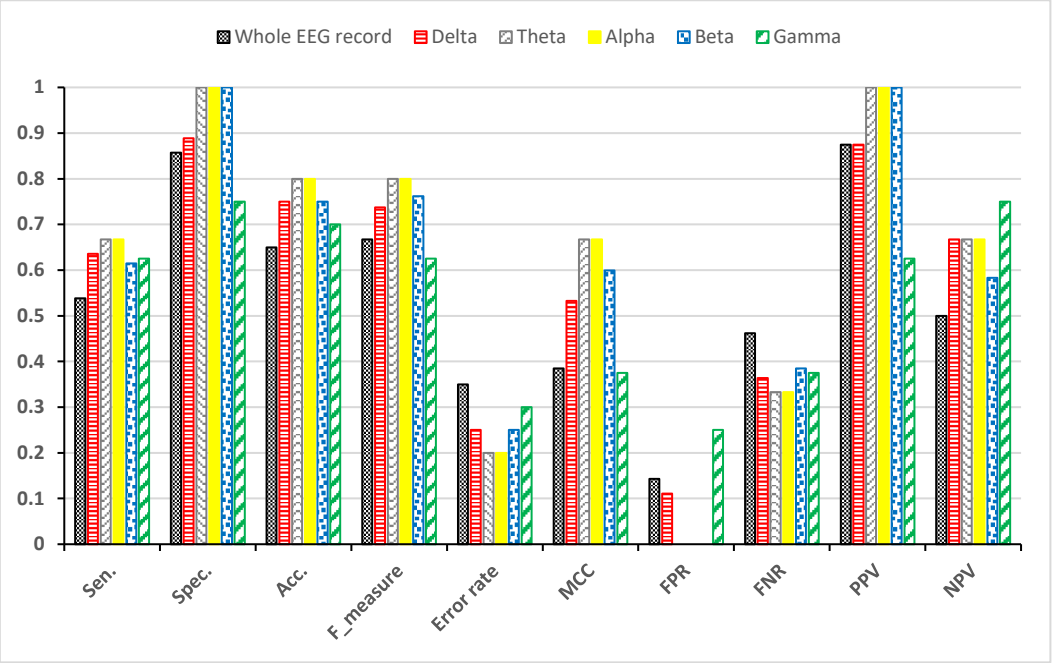

Figure 5-8: HFD performance

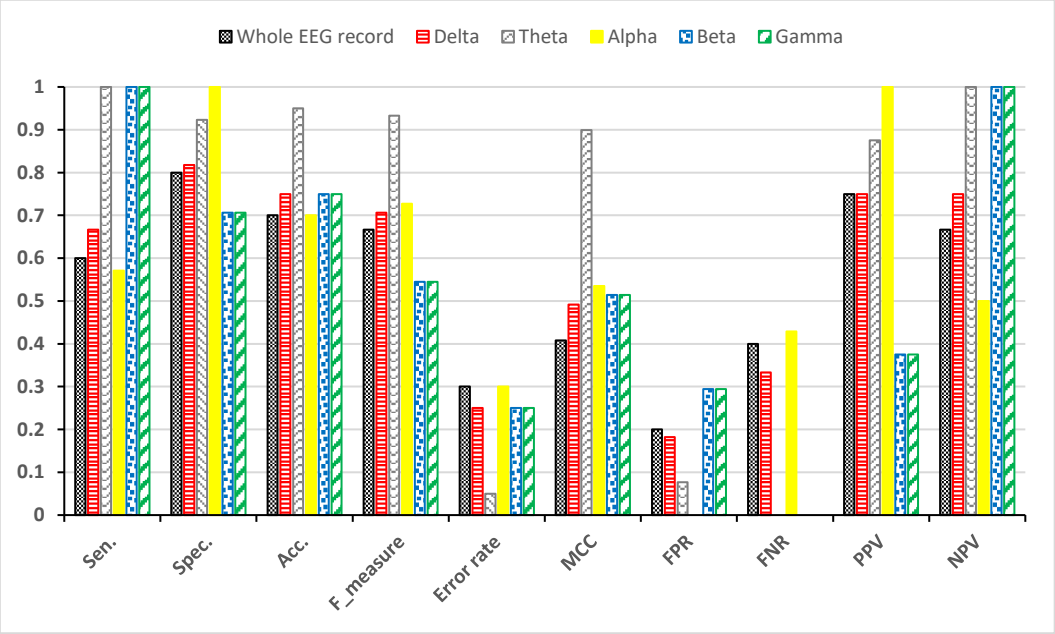

Figure 5-9: LZC performance

5.4. Discussions

The results show that EEG complexity-based measures provide a potentially useful way to detect AD. This is consistent with other studies [46][105][106][108].

For all three methods of computing complexity measures (i.e., TsEn, HFD, and LZC), we found that the complexity measures derived from the EEG frequency bands

provide significantly better performance in detecting AD than measures derived from whole EEG records. This comes from the greater differences between the complexity measures for AD and normal subjects when they are derived from the frequency bands compared to when they are derived from whole record which is a desirable property of a good biomarker.

In particular, we found that for the TsEn and HFD complexity measures derived from the delta and theta bands gave the best performance. For the delta band, three EEG channels (T4, O1, and O2) gave the best performance. For the theta band, F4 gave the best performance.

Similar results were obtained for the LZC complexity measures, except that the best EEG channel was C3 for the theta band. This is consistent with the findings of other studies which suggested that AD starts from the back of the brain and then spreads gradually to other parts of the brain [8][210][49][237][236]. This implies that it may be possible to use only a small number of EEG channels to detect AD.

Table 5-3 shows the best performance indices for the three complexity measures i.e., TsEn, HFD, and LZC was for theta band. Therefore, analysis of the theta band based on EEG in AD diagnosis has high impact and this is compatible with the findings of others that found the theta analysis is most closely related to AD severity [101].

Our findings in AD diagnosis were based on EEG complexity measures and could be used with other types of EEG analysis *e.g.,* slowing, and coherence in EEG signal. In addition, our results show that EEG complexity methods can be used to diagnose AD patients with low cost biomarkers (only one EEG channel was used in AD diagnosis).

Our study has a number of limitations. At present, our methods have been applied only to the detection of AD, the most common form of dementia. A more detailed study is necessary to evaluate the methods using much larger and diverse EEG

datasets. This includes using the methods to differentiate between normal, MCI and AD subjects [100][106][116].

This study shows that the abnormalities caused by AD can be detected by the complexity measures. However, similar changes may be caused by other neurodegenerative diseases, such as other types of dementia. To enhance the diagnostic usefulness of the methods, it may be necessary to enhance them further to differentiate between dementias.

5.5. Summary

AD causes changes in the EEG due to loss of memory and cognitive decline and these changes are thought to be associated with functional disconnections among cortical areas resulting from the death of brain cells. Therefore, EEG analysis may provide valuable information about brain dynamics in AD. AD causes a reduction in neuronal activity of the brain and this may be reflected in EEG signals. Non-linear methods based on EEG complexity approaches have shown promising results in detected changes in the EEG thought to be due to AD. Therefore, EEG complexity can potentially be a good biomarker for AD diagnosis. We investigated three complexity measures, TsEn, HFD, and LZC methods, derived from EEG frequency bands. We found that AD subjects have significantly lower TsEn, HFD, and LZC values in specific EEG frequency bands and specific EEG channels compared to normal patients. This may provide an effective way to discriminate between AD patients and normal subjects.

Chapter 6. Robust EEG Based Biomarkers to Detect Alzheimer's Disease in its Early Stages

6.1. Introduction

Although a large number of EEG-based biomarkers for AD exists, no single biomarker is robust enough for use in clinical practice or provides a clear-cut detection of AD in the early stages [2][209][113][244][245][246][247]. A robust biomarker of AD should be consistent and have high detection performance (e.g. in terms of sensitivity and specificity)[2][51][248]. Potentially, a combination of biomarkers could provide the required level of robustness necessary for clinical use [2][8][37][36]. Few studies have investigated how to combine different EEG biomarkers to exploit their strengths. Hamadicharef et al. [105] developed a logistic regression model to combine AD biomarkers, however the study did not investigate factors that determine the optimal set of biomarkers. In addition, only a small number of specific biomarkers were considered. Houmani et al. [249] developed an EEG diagnostic model for AD detection based on two EEG features, namely epoch-based entropy (a measure of signal complexity) and bump modelling (a measure of synchrony). They found these two features are sufficient for discriminating among subjective cognitive impairment (SCI) patients, MCI patients, possible AD patients, and patients with other pathologies groups. A classification accuracy of 91.6% (specificity = 100%, sensitivity = 87.8%) was obtained when discriminating SCI patients from possible AD patients and 81.8% to 88.8% accuracy was obtained for the 3-class classification of SCI, possible AD and other patients. They did not investigate other key characteristics of dementia EEG (i.e. slowing of the EEG, reduction in EEG complexity and reduction in EEG coherence). The development of robust EEG biomarkers requires a comprehensive study of key EEG features with significant association with AD and the integration of the most appropriate biomarkers into a composite biomarker. The aim of this chapter is to provide a methodological

framework for the development of robust EEG biomarkers to detect AD with clinically acceptable performance by exploiting the combined strengths of different EEG based biomarkers. A large number of biomarkers were investigated based on the three key characteristics of dementia EEG (slowing of the EEG, reduction in EEG coherence and reduction in EEG complexity). Furthermore, the five EEG bands (delta, theta, alpha, beta, and gamma), channel locations, and all possible combinations of the biomarkers were investigated. This is the first such study and this approach has made it possible to achieve high performance (close to 100% for sensitivity and specificity) and should facilitate acceptance of EEG biomarkers.

6.2. Methodology

Figure 6-1 provides an overview of the methodological framework for the development of robust EEG biomarkers to detect AD with clinically acceptable performance by exploiting the combined strengths of different EEG based biomarkers. It supports the investigation, development, integration and assessment of the performance of new and promising biomarkers based on the three main characteristics of the dementia EEG (slowing of the EEG, reduction in complexity and reduction in coherence). The emphasis is on finding the best possible combination of biomarkers to detect AD accurately. The development of robust and composite EEG biomarkers requires the identification of EEG features which have a significant association with AD. The features were extracted from the five traditional EEG bands (i.e., delta, theta, alpha, beta, and gamma). Previous studies have shown that such an approach may enhance the performance of biomarkers (i.e., using EEG frequency bands instead of using the entire EEG record) [38]. The features were then used to derive the EEG biomarkers and the best performing biomarkers found using support vector machine (SVM) because of its robustness with high-dimensional data and because it has been shown to perform well in automated AD studies compared to

other machine learning methods. A linear discriminant analysis (LDA) model was then used to combine the selected biomarkers to create the final model to detect AD. The biomarkers investigated included those based on ΔEEG_A, ΔPS, ZCI, TsEn, HFD, ApEn, LZC, and EEG Coherence which have shown promise.

The following steps outline the procedure for deriving the biomarkers

1. Filter the EEG signals into five frequency bands (i.e., delta 0-4Hz, theta 4-8Hz, alpha 8-12Hz, beta 12-30Hz, and gamma 30-45Hz). Previous studies have shown that biomarkers derived from the EEG frequency bands (instead of the entire EEG) have enhanced performance [50]. For this step, a low computational infinite impulse response (IIR) Chebyshev-II bandpass filters were used to retain computational efficiency in extracting the biomarkers [206].
2. Identify EEG features and compute the EEG indices based on the slowing of EEG (i.e., ΔPS, ΔEEG_A, and ZCI), reduction in EEG complexity (i.e., TsEn, HFD, LZC, and ApEn) and reduction in coherence for each of the five EEG frequency bands and each EEG channel.
3. Compute the EEG biomarkers from the indices in Step 2 above (e.g., TsEn(delta)/TsEn(theta), TsEn(delta)/TsEn(alpha), …, etc), and select the biomarkers and EEG channels that have a significant association with AD (in terms of the p-values).
4. Construct panels of biomarkers from the selected biomarkers.
5. For each panel of biomarkers, develop a classification model using SVM and assess its performance in detecting AD.
6. Select EEG biomarkers panels with sensitivity and specificity values above a specified threshold (80% in this case).

7. Develop a diagnostic model to detect AD by combining the selected biomarkers. In the study, a linear discriminant analysis model was used to combine the biomarkers and to produce the diagnostic model [250].
8. Validate the diagnostic model using unseen dataset.

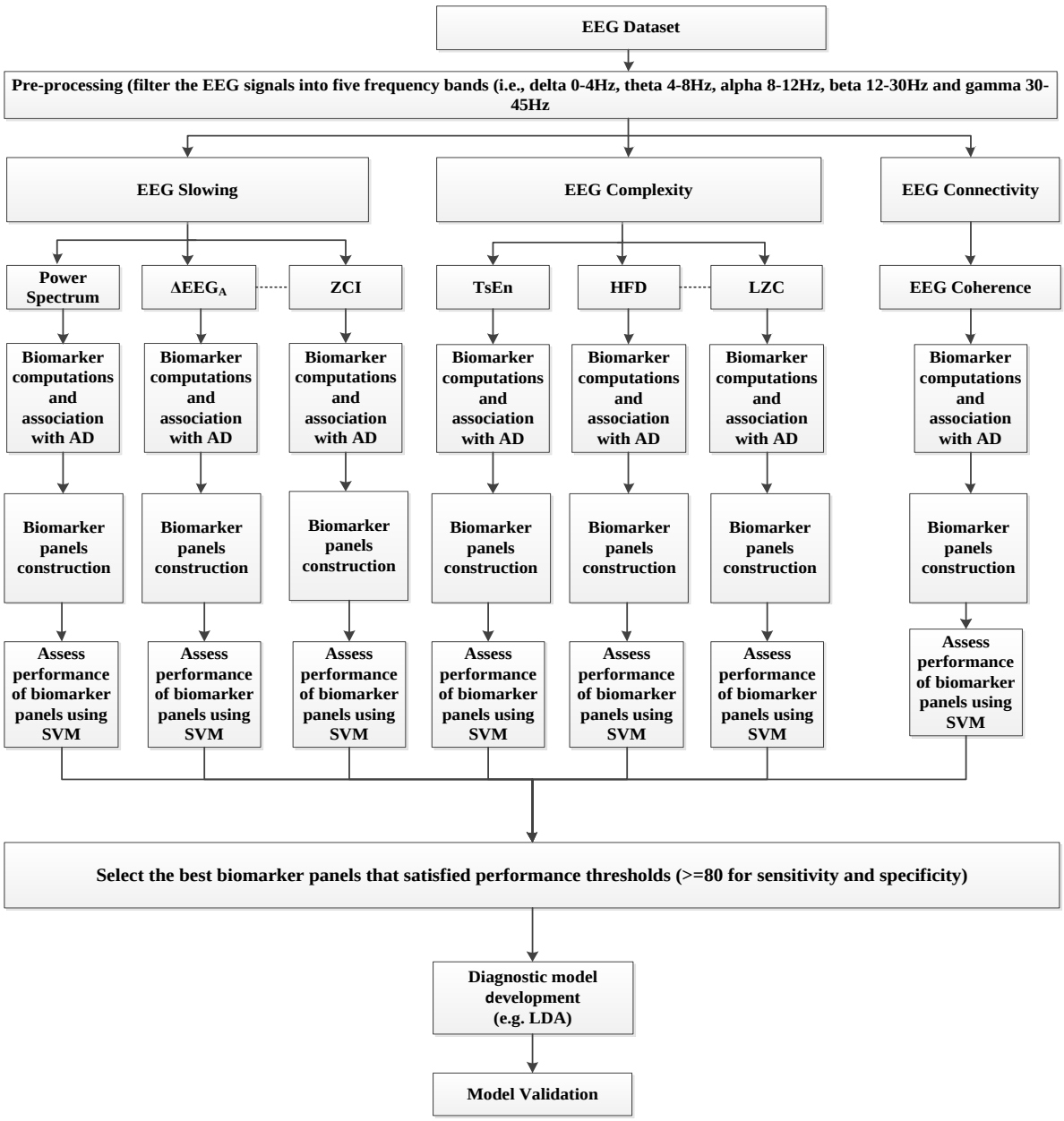

Figure 6-1: A framework for developing robust EEG based biomarker

The EEG dataset that was used in this chapter consists of two datasets (A and B). Dataset A includes 11 age-matched subjects (3 AD patients and 8 normal subjects). Dataset B consists of 41 subjects (24 normal subjects and 17 probable AD patients). More details are given in Chapter 3.

6.2.1. Identification of EEG features and Computation of Biomarkers (Steps 2 and 3)

The EEG signals are first filtered into the five EEG frequency bands using a low computational IIR filter as described in chapter 3 and the main features identified. The main EEG features in dementia are those associated with the slowing of the EEG (e.g. shifts in the EEG power to the lower frequencies), reduction in EEG complexity (reduction in complexity measures) and reduction in EEG coherence. These features are quantified by computing appropriate indices in each of the five EEG frequency bands - for EEG slowing the indices are ΔPS, ΔEEG_A, and ZCI; for reduction in EEG complexity, the indices are TsEn, HFD, LZC, and ApEn; for coherence, the indices are the coherence values between the channels.

6.2.2. Biomarker selection and biomarker panels

The methodological framework involves a thorough consideration of all possible biomarkers. This creates a large number of biomarkers at the biomarker computation stage and so it is necessary to select biomarkers with significant statistical association with AD as these may be useful in discriminating between AD patients and normal subjects. P-values are used here as a statistical criterion to select the biomarkers. Dataset B was split 60% for training and 40% for testing for this purpose, and 10 fold cross-validation was used for training the developed model. Biomarkers with p-values no greater than 0.001 are selected as having significant association with AD. Similarly, EEG channels for which the biomarkers have significant association with AD (p-values <= 0.001) are taken to have a significant association with AD and was also selected to use in the classification. Figure 6-2 shows the procedure used to select the EEG biomarkers of AD.

The selected biomarkers are then used to construct panels of biomarkers for AD detection. A machine learning model is then developed for each panel and the performance of each in detecting AD assessed. This allowed us to determine the best

performing biomarker panels. Support vector machine (SVM) is used to develop the models.

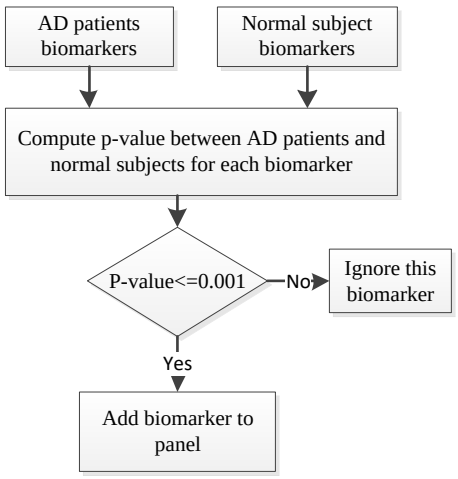

Figure 6-2: Construct panels of biomarkers for AD detection

6.2.3. Diagnostic model to detect AD

The outcome of biomarker selection is a set of panels of biomarkers from the different methods of computing biomarkers. The next step is to find the best possible combination of biomarkers to allow the detection of AD with high accuracy. To do this, we first select biomarker panels that satisfied the threshold for sensitivity and specificity of at least 80% for inclusion in the final model. Then, the biomarkers with the highest performance are analysed. Biomarkers with the fewest number of EEG channels are selected for inclusion in the development of the diagnostic model. These are then combined into to produce the robust diagnostic model using LDA. LDA is a supervised method which is commonly used in classification and provides optimal separation in classification [251][138][141]. The performance of the LDA model is then assessed using unseen data.

6.3. Result

6.3.1. Biomarker computations

For each of the seven methods (i.e., ΔPS, ΔEEG_A, ZCI, TsEn, HFD, ApEn, and LZC) and for each channel, 25 biomarkers were computed, five biomarkers for the five EEG frequency bands and 20 for ratios between bands. Examples of the biomarkers are ZCI(alpha), TsEn(delta), TsEn(alpha/beta), ΔEEG_A(theta/beta), …. etc. Thus, 475 biomarkers were computed and analysed for each of the seven methods (25 biomarkers x 19 EEG channels for each method). Thus after biomarker computations, there was a total of 3325 biomarkers for the seven methods (475 x 7).

For the coherence based approach, there were 171 coherence values between pairs of EEG channels (e.g., Fp1-Fp2, Fp1-F7, …, O1-O2) for each of the traditional EEG band and 20 ratios. The total number of biomarkers computed for the method was 4275 (171 biomarkers for each of the 25 biomarkers). Thus, a total of 7600 (3325 + 4275 biomarkers) candidate biomarkers were used to determine a panel of robust EEG biomarkers based on the promising EEG methods (i.e., ΔPS, ΔEEG_A, ZCI, TsEn, HFD, ApEn, LZC, and EEG coherence).

6.3.2. Biomarker selection

P-values were computed for each of the 7600 extracted biomarkers. For example, Table 6-1 shows theta/alpha ratio for TsEn method. There are two EEG channels (T5, and T6) satisfied the criterion. The TsEn method showed two EEG channels that satisfied the p-value criterion for the theta/alpha ratio (as shown in Table 6-1).

To assess the relative effects of the EEG bands and EEG channels in the detection of AD, the probability distribution ratios were computed by using Equation 6.1:

$$P_X = \frac{x_i}{\sum_{i=1}^{N} x_i} \tag{6.1}$$

where P_X is the probability distribution ratio for an N-sample data sequence x(1), x(2), ..., x(N).

Bonferroni correction [192] was used for adjusting p-values as shown in Table 6-1. The corrected p-values show the two EEG channels T5, and T6 for theta/alpha ration of TsEn are the two significant channels and this is consistent with our analysis.

Table 6-1: P-value and corrected p-values for theta/alpha band for TsEn method

Ch No.	EEG channel	P-value	Bonferroni-corrected P-value	Bonferroni-corrected significance
1	Fp1	0.0264	0.5016	Not Significant
2	Fp2	0.0394	0.7486	Not Significant
3	F7	0.4883	1	Not Significant
4	F3	0.5511	1	Not Significant
5	FZ	0.3612	1	Not Significant
6	F4	0.1582	1	Not Significant
7	F8	0.3105	1	Not Significant
8	T3	0.1352	1	Not Significant
9	C3	0.2859	1	Not Significant
10	CZ	0.8901	1	Not Significant
11	C4	0.5549	1	Not Significant
12	T4	0.2418	1	Not Significant
13	**T5**	**0.0010**	0.019	**Significant**
14	P3	0.0048	0.0912	Not Significant
15	PZ	0.0207	0.3933	Not Significant
16	P4	0.0036	0.0684	Not Significant
17	**T6**	**0.0002**	0.0038	**Significant**
18	O1	0.0028	0.0532	Not Significant
19	O2	0.0035	0.0665	Not Significant

The critical p-value was 0.00263 (e.g., 005/19).

The probability distribution ratio and Bonferroni correction of p-value for each biomarker may reflect their significance in AD detection. The Probability distribution ratio was computed for the 25 biomarkers for each of the eight methods (i.e., ΔPS, ΔEEG_A, ZCI, TsEn, HFD, ApEn, LZC, and EEG coherence). It was also computed for the 19 EEG channels for all biomarkers.

Table 6-2 summarises the results for the 25 EEG biomarkers that were analysed for the eight methods. For each biomarker, 19 EEG channels were analysed by computing the p-value to determine the biomarkers that may have significant

association with AD. The EEG biomarkers have significant separation between AD patients and normal subjects were selected.

Table 6-2: Probability distribution ratio for all 25 features for each method

EEG bands	Number of EEG channel for the								Sum.	Probability distribution ratio
	ApEn	LZC	HFD	TsEn	ΔPS	ΔEEG_A	ZCI	Coh		
Theta/alpha	12	5	9	2	18	12	15	1	74	15.579
Alpha/theta	12	7	0	2	17	15	15	0	68	14.316
Alpha/delta	0	0	0	0	18	11	8	0	37	7.789
Beta/theta	13	0	0	0	2	10	9	1	35	7.368
Theta/beta	12	0	3	0	4	5	9	0	33	6.947
Alpha	2	0	4	0	19	3	4	0	32	6.737
Delta/alpha	1	0	0	0	18	4	7	0	30	6.316
Delta	0	0	0	0	19	0	0	3	22	4.632
Theta	0	0	0	0	18	0	0	3	19	4.000
Theta/delta	0	0	0	0	17	0	0	0	18	3.789
Delta/theta	8	0	0	1	0	4	4	0	17	3.579
Gamma/theta	0	0	0	0	16	0	0	0	17	3.579
Beta/delta	0	0	0	0	2	7	3	2	14	2.947
Beta	2	0	4	0	1	1	3	2	13	2.737
Theta/gamma	6	0	0	1	0	0	4	0	11	2.316
Alpha/beta	0	4	0	0	4	0	0	0	8	1.684
Gamma/delta	3	0	0	0	1	3	1	0	8	1.684
Delta/beta	0	0	0	0	2	0	2	0	4	0.842
Alpha/gamma	0	4	0	0	0	0	0	0	4	0.842
Gamma	0	2	0	0	1	0	0	2	3	0.632
Gamma/alpha	0	0	1	0	2	0	0	0	3	0.632
Gamma/beta	0	1	0	0	1	0	0	0	3	0.632
Beta/alpha	0	0	0	0	1	0	0	0	2	0.421
Delta/gamma	0	0	0	0	0	0	0	0	0	0.000
Beta/gamma	0	0	0	0	0	0	0	0	0	0.000
Summation	71	23	21	6	181	75	84	14	475	100

Table 6-2 shows the number of EEG channels (out of 19 channels) that satisfied the p-value threshold (less than or equal to 0.001) for each biomarker. This means the

theta/alpha ratio has the highest number of channels that have a p-value which is less than or equal to 0.001. While delta/gamma and beta/gamma biomarkers have no EEG channels that satisfied the threshold of p-value. To rank the biomarkers, the probability distribution ratio was computed for each biomarker, as shown in Figure 6-3.

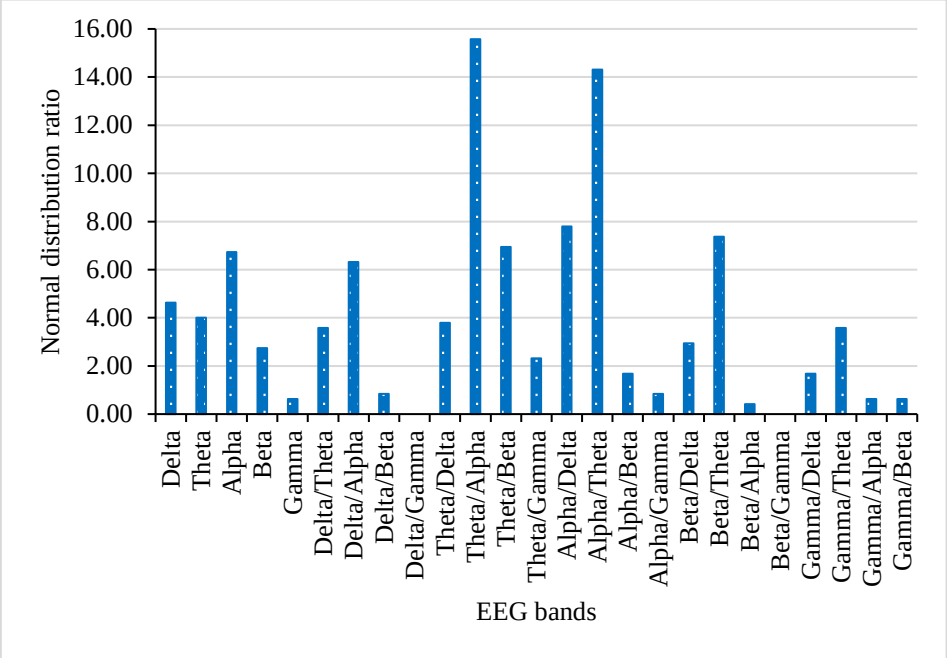

Figure 6-3: Probability distribution ratio for all 25 biomarkers

Figure 6-3 shows the probability distribution ratio for all EEG biomarkers. The maximum probability distribution ratio was 15.579 for the theta/alpha ratio. It was computed by dividing the probability distribution ratio of theta/alpha i.e., 74 by the total probability distribution ratio of all bands i.e., 475, and multiply the resulted value by 100.

To determine the EEG biomarkers that could be used in discriminating between AD patients and normal subjects with the best performance, we selected the biomarkers that occupied 80% of the normal distribution ratios to all biomarkers. Therefore, 11 biomarkers were selected that have a more significant in AD diagnosis. Table 6-3 shows the EEG biomarkers have a more significant with AD and can be used to diagnose AD with the best performance.

Table 6-3: The selected features that could be used in the classification

Seq.	Features
1	Theta/Alpha
2	Alpha/Theta
3	Alpha/Delta
4	Beta/Theta
5	Theta/Beta
6	Alpha
7	Delta/Alpha
8	Delta
9	Theta
10	Theta/Delta
11	Delta/Theta

These 11 biomarkers were used to develop a classification model for AD diagnosis.

To determine the EEG channels that have a more significant association with AD, we selected channels that met the p-value threshold of less than or equal to 0.001. Then, the probability distribution ratio was computed for all the 19 EEG channels. The probability distribution ratio was computed to identify the EEG channels that are considered promising in terms of having a significant association with AD. The EEG channels are ranked in Table 6-4 in terms of probability distribution ratio.

Table 6-4: Probability distribution ratio for all 25 EEG features and for all 19 EEG channels

EEG Channels	Number of biomarkers							Sum.	Probability distribution ratio
	ApEn	LZC	HFD	TsEn	ΔPS	ΔEEG_A	ZCI		
P4	7	5	4	0	10	10	12	48	10.412
P3	9	2	4	0	9	8	13	45	9.761
PZ	8	2	3	2	11	9	10	45	9.761
T6	3	6	2	2	9	7	7	36	7.809
T5	2	3	4	2	9	6	7	33	7.158
C4	6	0	0	0	11	6	4	27	5.857
T4	4	2	1	0	8	7	3	25	5.423
T3	4	0	1	0	14	3	2	24	5.206
C3	4	0	0	0	12	4	4	24	5.206
CZ	6	0	0	0	9	3	6	24	5.206
O2	3	3	1	0	10	3	4	24	5.206
O1	4	0	1	0	8	3	4	20	4.338
F8	2	0	0	0	12	1	2	17	3.688
F7	3	0	0	0	9	2	2	16	3.471
F3	3	0	0	0	9	1	2	15	3.254
FZ	3	0	0	0	7	2	2	14	3.037
Fp1	0	0	0	0	9	0	0	9	1.952
F4	0	0	0	0	9	0	0	9	1.952
Fp2	0	0	0	0	6	0	0	6	1.302
Sum.	71	23	21	6	181	75	84	461	100

Table 6-4 shows the number of EEG channels for the methods (i.e., ΔPS, ΔEEG_A, ZCI, TsEn, HFD, ApEn, and LZC) that met the p-value threshold and the probability distribution ratio for each channel. Figure 6-4 shows the probability distribution ratio for all EEG channels. The maximum probability distribution ratio is for P3, PZ and P4.

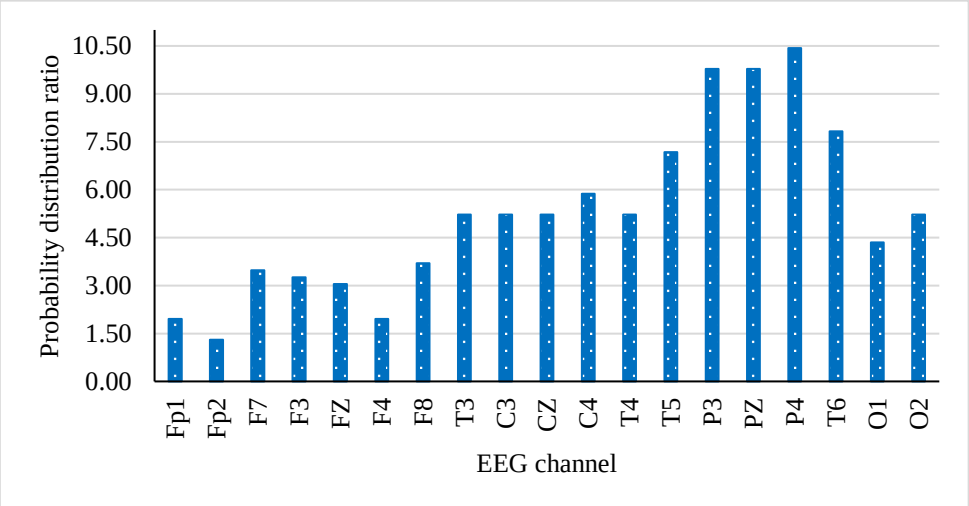

Figure 6-4: Probability distribution ratio for all 19 EEG channels and for all eight methods

The rank of each EEG channel in terms of the number of biomarkers is shown in Table 6-4 and Figure 6-.5. Channel P4 was the top-ranked channel among the 25 biomarkers, with channels P3 and PZ the second highest ranked channels. Whilst channel Fp2 is the lowest ranked channel. This is consistent with findings in the literature where it is thought that AD starts from the back of the brain and then spread to other parts [126][120][110][111][86][113].

Based on the analysis of the 25 biomarkers for each EEG channel (see Tables 6-2 and 6-4, and Figures 6-4 and 6-5), 12 EEG channels (P4, P3, PZ, T6, T5, C4, T3, C3, CZ, T4, O2, and O1) were selected in the classification as having a significant association with AD. These channels accounted for more than 80% in the probability distribution ratios of all EEG channels.

For each of the 11 EEG biomarkers that were selected, 4082 combinations were investigated for the 12-EEG channels (all combinations from length 1 to length 10 were constructed for the 12-EEG channels). For each method, we investigated 44902 biomarkers (11- biomarkers x 4082-combinations).

Based on analysis of the probability distribution ratio for each of the 25 biomarkers, 11 biomarkers and 12 EEG channels were selected in discriminating between AD patients and normal subjects as shown in Tables 6-1 and 6-3, and Figure 6-5.

For EEG coherence, 10 pairs of EEG channels that satisfied the p-value threshold (less than or equal to 0.001) were selected for discriminating between AD patients and normal subjects as shown in Table 6-5.

Table 6-5: Probability distribution ratio for all 25 EEG features and for all 19 EEG channels

EEG Channel	No. of biomarkers	Probability distribution ratio
F4-F8	3	21.4286
Fz-F8	2	14.2857
T3-T4	2	14.2857
Fp2-F4	1	7.1429
F4-T3	1	7.1429
F4-T4	1	7.1429
F8-P4	1	7.1429
T3-P4	1	7.1429
T3-T6	1	7.1429
T4-P3	1	7.1429
Summation	14	100

6.3.3. Performance analysis

For seven of the methods investigated (i.e., ΔPS, ΔEEG_A, ZCI, TsEn, HFD, ApEn, and LZC), 11 EEG bands and 12 EEG channels for each band were selected to develop the SVM classification models using the testing dataset. For EEG coherence, 11 EEG bands and 10 pairs of EEG channels for each band were also selected to develop the SVM classification models.

Each method was assessed by computing its performance in discriminating between AD patients and normal subjects. For each biomarker, a support vector machine (SVM) model was developed based on the training dataset and assessed by using the testing dataset. The best biomarkers for each method were determined based on their performance (in terms of sensitivity, specificity, and the number of EEG channels used). Sensitivity and specificity thresholds of at least 80% in the detection of AD were set for the biomarkers. Table 6-6 shows the performance of ApEn method in AD diagnosis.

Table 6-6: Performance of the ApEn method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C4)**
Theta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C4)**
Alpha/delta	75.00	88.89	82.35	80.00	0.65	85.71	80.00	(P3)
Beta/theta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(T3)
Alpha	**100.00**	**83.33**	**88.24**	**83.33**	**0.77**	**71.43**	**100.00**	**(P3, T6, O1)**
Delta/alpha	75.00	88.89	82.35	80.00	0.65	85.71	80.00	(C3, P3)
Theta/beta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(T3)
Theta/delta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(T3, CZ, P3,P4), and (C3, T5, PZ, P4)
Delta	50.00	71.43	58.82	58.82	0.21	71.43	50.00	(T3, P3), and (T5, P3)
Theta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(C4)
Delta/theta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(T3, C3,CZ, T5, P3, P4), and (C3, CZ, T4, T5, P3, P4)

The best performance for ApEn method was for the biomarkers ApEn(alpha/theta(C4)), ApEn(theta/alpha(C4)), and ApEn(alpha(P3, T6, O1)). Table 6-7 shows the performance of LZC method in the detection of AD.

Table 6-7: Performance of LZC method for all the 11 biomarkers

EEG bands	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(T3, T6)**
Theta/alpha	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(T3, T6)**
Alpha/delta	**100.00**	**83.33**	**88.24**	**83.33**	**0.77**	**71.43**	**100.00**	**(CZ, C4, P3, P4), (CZ, T4, P3, P4), and (CZ)**
Beta/theta	**83.33**	**81.82**	**82.35**	**76.92**	**0.63**	**71.43**	**90.00**	**(C4, P3, T6, O1), and (C4, P3, T6, O2)**
Alpha	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(C3, CZ, PZ, T6, O1)**
Delta/alpha	**100.00**	**83.33**	**88.24**	**83.33**	**0.77**	**71.43**	**100.00**	**(CZ, T4, P3, P4), (CZ, T4, P3, T6), and (CZ, T4, P3, O2)**
Theta/beta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(C4, T5, PZ, P4, O1, O2)
Theta/delta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(CZ)
Delta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(CZ, P3), and (CZ, P4)
Theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3, C3, P3)**
Delta/theta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(CZ)

The best performance for LZC method was for the biomarkers LZC(alpha/theta(T3, T6)), LZC(theta/alpha(T3, T6)), LZC(alpha/delta(CZ, C4, P3, P4)), …. , and LZC(theta(T3, C3, P3)). Table 6-8 shows the performance of HFD method in AD diagnosis.

Table 6-8: Performance of HFD method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(CZ, C4, T4, P4)
Theta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (T5)**
Alpha/delta	**83.33**	**81.82**	**82.35**	**76.92**	**0.63**	**71.43**	**90.00**	**(CZ, PZ, O2)**
Beta/theta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(CZ, C4, T4, P4)
Alpha	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(T5, PZ), (PZ, T6), (T6, O1), and (T6, O2)**
Delta/alpha	80.00	75.00	76.47	66.67	0.51	57.14	90.00	(O1)
Theta/beta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(C3, CZ, C4)
Theta/delta	60.00	85.71	70.59	70.59	0.46	85.71	60.00	(CZ, C4,T4, PZ, O1,O2)
Delta	62.50	77.78	70.59	66.67	0.41	71.43	70.00	(T3, C3, CZ, T5, P3)
Theta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(T3, C4), and (C3, C4)
Delta/theta	70.00	100.00	82.35	82.35	0.70	100.00	70.00	(CZ, C4, T4, P4)

The best performance for HFD method was for the biomarkers HFD(theta/alpha(T3)), HFD(alpha/delta(CZ, PZ, O2)), , …… , and HFD(alpha(T6, O2)). Table 6-9 shows the performance of TsEn method in AD diagnosis.

Table 6-9: Performance of TsEn method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**85.71**	**90.00**	**88.24**	**85.71**	**0.76**	**85.71**	**90.00**	**(T3, CZ, P3, T6), and (C3, CZ, O1, O2)**
Theta/alpha	**85.71**	**90.00**	**88.24**	**85.71**	**0.76**	**85.71**	**90.00**	**(T3, CZ, C4, T6, O1), and (T3, CZ, T5, P4, O1)**
Alpha/delta	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(C3, T6)**
Beta/theta	60.00	66.67	64.71	50.00	0.25	42.86	80.00	(T3, P3, PZ, O1), (C3, C4, T4, O1), (C3, T4, PZ, O1), and (P3, PZ, O1, O2)
Alpha	75.00	88.89	82.35	80.00	0.65	85.71	80.00	(T3, T6, O1), and (T3, T6, O2)
Delta/alpha	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(C3, T6)**
Theta/beta	57.14	70.00	64.71	57.14	0.27	57.14	70.00	(T5, O1)
Theta/delta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(PZ)
Delta	100.00	76.92	82.35	72.73	0.66	57.14	100.00	(T3, C4), (T3, P4), (T3, O1), (T3, O2), (T5, P3), (T5, O1), and (T5, O2)
Theta	62.50	77.78	70.59	66.67	0.41	71.43	70.00	(C3)
Delta/theta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(PZ)

The best performance for TsEn method was for the biomarkers TsEn(alpha/theta(T3, CZ, P3, T6)), TsEn(alpha/theta(C3, CZ,O1, O2)) , …, and TsEn(delta/alpha(C3, T6)).

Table 6-10 shows the performance for ΔPS method in AD diagnosis.

Table 6-10: Performance of ΔPS method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (C3), (T4), (P3), (PZ), and (P4)**
Theta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C3), (CZ), (C4), (P4), and (O1)**
Alpha/delta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3, T4, P3, P4, O1)**
Beta/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C3)**
Alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (C3), (CZ), (C4), (T4), (T5), (P3), (PZ), (P4), (T6), (O1), and (O2)**
Delta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (C3), and (CZ)**
Theta/beta	**100.00**	**83.33**	**88.24**	**83.33**	**0.77**	**71.43**	**100.00**	**(C4)**
Theta/delta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(P3, T6)**
Delta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C3), (CZ), (C4), (T4), (T5), (P3), (PZ), (P4), (T6), (O1), and (O2)**
Theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(CZ), (P4), and (O2)**
Delta/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (T4), (P3), (PZ), (P4), and (T6)**

The best performance for ΔPS method was for the biomarkers ΔPS(alpha/theta(T3)), ΔPS(alpha/theta(C3)) , …, and ΔPS(delta/theta (T6)). Table 6-11 shows the performance of ΔEEG_A method in AD diagnosis.

Table 6-11: Performance of ΔEEG_A method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F- measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3)**
Theta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T3), (C3), (T5), (P3), (PZ), and (P4)**
Alpha/delta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(T6)**
Beta/theta	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(C4), and (P4)**
Alpha	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(T6)**
Delta/alpha	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(P4, T6), (T6, O1), and (T6, O2)**
Theta/beta	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(T3, CZ, T4, O1, O2)**
Theta/delta	75.00	88.89	82.35	80.00	0.65	85.71	80.00	(CZ, C4, T5, O1), (CZ, T5, PZ, O1), and (CZ, T5, O1, O2)
Delta	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(P3)
Theta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(C3)
Delta/theta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(T3, C4), (C4, T5), and (C4, P3)

The best performance for ΔEEG_A method was the biomarkers ΔEEG_A(alpha/theta(T3)), ΔEEG_A(theta/alpha (T3)), …, and ΔEEG_A(theta/beta (T3, CZ, T4, O1, O2)). Table 6-12 shows the performance of ZCI method in AD diagnosis.

Table 6-12: Performance of ZCI method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C3), (P3), (PZ), and (P4)**
Theta/alpha	**100.00**	**100.00**	**100.00**	**100.00**	**1.00**	**100.00**	**100.00**	**(C3), (P3), and (P4)**
Alpha/delta	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(C3, P4, T6, O1)**
Beta/theta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(C3, T5, PZ), (C3, P3, PZ), (C3, P3, P4), (C3, PZ, P4), (C4, T4, PZ), (T4, P3, PZ), (T4, P3, O1), (T4, P3, O2), and (T4, P4, O1)
Alpha	**87.50**	**100.00**	**94.12**	**93.33**	**0.89**	**100.00**	**90.00**	**(T6, O1), and (T6, O2)**
Delta/alpha	**100.00**	**90.91**	**94.12**	**92.31**	**0.88**	**85.71**	**100.00**	**(C3, P3, O1)**
Theta/beta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(T3, P4, O1), (C3, T5, PZ), (C3, P3, P4), (C3, PZ, P4), (C4, T4, P3), (T4, P3, O1), (T4, P3, O2), and (T4, P4, O2)
Theta/delta	62.50	77.78	70.59	66.67	0.41	71.43	70.00	(T3, C3, C4, PZ)
Delta	50.00	80.00	58.82	63.16	0.28	85.71	40.00	(PZ)
Theta	66.67	87.50	76.47	75.00	0.55	85.71	70.00	(C4)
Delta/theta	62.50	77.78	70.59	66.67	0.41	71.43	70.00	(T3, C4, P3, PZ), and (T3, C4, PZ, P4)

The best performance for ZCI method was the biomarkers ZCI(alpha/theta(C3)), ZCI(alpha/theta(P3)), … , and ZCI(delta/alpha (C3, P3, O1)). Table 6-13 shows the performance of coherence method in AD diagnosis.

Table 6-13: Performance of coherence method for all the 11 biomarkers

EEG Biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %	EEG channel
Alpha/theta	**85.71**	**90.00**	**88.24**	**85.71**	**0.76**	**85.71**	**90.00**	**(Fp2-F4, Fz-F8, F4-T4, F8-P4, T3-T4, T3-P4)**
Theta/alpha	63.64	100.00	76.47	77.78	0.62	100.00	60.00	(F8-P4, T3-P4)
Alpha/delta	75.00	88.89	82.35	80.00	0.65	85.71	80.00	(F4-T3, F4-T4, T3-T4, T3-P4)
Beta/theta	**85.71**	**90.00**	**88.24**	**85.71**	**0.76**	**85.71**	**90.00**	**(Fp2-F4, F4-F8, F4-T3, F4-T4)**
Alpha	71.43	80.00	76.47	71.43	0.51	71.43	80.00	(F8-P4, T4-P3)
Delta/alpha	53.85	100.00	64.71	70.00	0.46	100.00	40.00	(T3-P4)
Theta/beta	**83.33**	**81.82**	**82.35**	**76.92**	**0.63**	**71.43**	**90.00**	**(F4-F8, F4-T3, F4-T4)**
Theta/delta	53.85	100.00	64.71	70.00	0.46	100.00	40.00	(Fp2-F4, F4-T3, F4-T4, T3-T4), and (Fp2-F4, F4-T3, F8-P4, T3-T4)
Delta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(F4-T4, T3-T4)
Theta	77.78	100.00	88.24	87.50	0.79	100.00	80.00	(T3-T4)

The best performance for coherence method was for the biomarkers Coh(alpha/theta (Fp2-F4, Fz-F8, F4-T4, F8-P4, T3-T4, T3-P4)), Coh(beta/theta(Fp2-F4, F4-F8, F4-T3, F4-T4)), and Coh(theta/beta(F4-F8, F4-T3, F4-T4)). Tables 6-6 to 6-13 showed the highest performance were selected of each EEG frequency band for the eight methods investigated. A few of the single biomarkers could provide with a high performance (e.g., ApEn(Alpha/theta(C4)) and ΔEEG_A(alpha/theta(T3))) because the size of the used EEG dataset was not large enough to demonstrate the actual performance to each biomarker.

6.3.4. Diagnostic model to detect AD

The key goal in the study is to find the best combination of EEG based biomarkers to detect AD patients with high sensitivity and specificity. Tables 6-6 to 6-13 summarise the best performing EEG biomarkers for the eight methods investigated. Biomarkers that satisfied the threshold for sensitivity and specificity of at least 80% were selected for inclusion in the final model. Then, the biomarkers with the highest performance

were analysed. Biomarker panels emerge high sensitivity and specificity that include the fewest number of EEG channels were selected for inclusion in the development of the diagnostic model. Table 6-14 shows the best 69 EEG biomarkers panel selected (69 out of 325567). The 69 EEG biomarkers were then combined in a model using LDA. For example, ApEn(Alpha(P3, T6, O1)), ApEn(Alpha/theta(C4)), …. , and ΔPS(Alpha/theta(T3, C3, T4, P3, PZ, P4) were computed and used to develop the LDA model. The training and testing EEG datasets were used in the model development. Dataset A was used to validate the model and the performance found to be 100% for sensitivity and specificity.

Table 6-14 shows the EEG based biomarkers selected and used in the diagnostic model.

Table 6-14: Panel of robust EEG biomarkers

Seq.	Method	EEG Biomarker	EEG channel	Category
1	ApEn	Alpha	P3, T6, O1	EEG complexity
2	ApEn	Alpha/theta	C4	
3	HFD	Alpha	T5, PZ, T6, O1, O2	
4	HFD	Theta/alpha	T3, T5	
5	LZC	Alpha	C3, CZ, PZ, T6, O1	
6	LZC	Alpha/delta	CZ, C4, T4, P3, P4, T6	
7	TsEn	Alpha/delta	C3, T6	
8	TsEn	Alpha/theta	T3, C3, CZ, P3, T6, O1	
9	ZCI	Alpha	T6, O1, O2	EEG slowing
10	ZCI	Alpha/theta	C3, P3, PZ, P4	
11	ZCI	Delta/alpha	C3, P3, O1	
12	ΔEEG_A	Alpha/delta	T6	
13	ΔEEG_A	Theta/alpha	T3, C3, T5, P3, PZ, P4	
14	PS	Alpha	T3, C3, P3, P4, T6	
15	PS	Alpha/delta	T3, T4, P3, P4, O1	
16	PS	Alpha/theta	T3, C3, T4, P3, PZ, P4	
17	Coh	Alpha/theta	Fp2-F4, Fz-F8, F4-T4, F8-P4, T3-T4, T3-P4	EEG connectivity

For example, the performance for ApEn(Alpha(P3, T6, O1)) was 100% and 88% for sensitivity and specificity, respectively as shown in Table 6-6, and the performance for Coh (Alpha/theta (Fp2-F4, Fz-F8, F4-T4, F8-P4, T3-T4, T3-P4)) was 86% and

90% for sensitivity and specificity, respectively as shown in Table 6-13. We can conclude, the performance of inclusion multiple biomarkers in one diagnostic model guided us to high diagnostic performance.

6.4. Discussions

AD affects cognitive memory and brain functions [1][2], each EEG frequency band has been shown to be associated with specific brain functions [1][244]. Furthermore, each lobe of the brain has specific tasks [23][24][42][245][246][22] and does not generate the same brainwave frequency simultaneously [247]. Each EEG band is related to specific functions (e.g., drowsiness, wakefulness, and deep sleep) [80][81][79]. Thus, decline in brain functions may be reflected in the EEG activities and EEG frequency bands [23][252][253]. Consequently, deriving EEG biomarkers from frequency bands is thought to provide a better performance in detecting AD compared to EEG biomarkers derived from the whole EEG record [50]. Furthermore, EEG measurement can be influenced by many artefacts (e.g. muscle and eye movements) [78][82]. Filtering EEG signal into frequency bands could reduce these artefacts. For example, the peak frequency of the masseter muscle around 50–60Hz, 40–80Hz for temporal muscles, posterior head muscles sternocleidomastoids, splenius capitus, and trapezius higher peak frequencies around 100Hz, and the extraocular muscles that contain both striate and smooth fibres and control saccadic eye movements produce activity that peaks around 65Hz [254].

The development of robust EEG based biomarkers requires a thorough investigation of all the possible factors that affect AD detection. These factors may include the basis for the biomarkers (e.g. EEG slowing, reduction in complexity, reduction in coherence), the EEG features used and hence the methods used to derive the biomarkers, and the EEG channels from which the biomarker is derived. Considering these factors and integrating biomarkers into a composite biomarker should lead to

robust EEG based biomarkers. Our methodological framework makes it possible to do so and our results support this.

We selected EEG biomarkers and channels with a significant association with AD detection as having high performance in distinguishing between AD patients and normal subjects. These biomarkers and EEG channels were then used to find the best combination of the EEG biomarkers that is robust enough to be used to detect AD.

The total number of EEG biomarkers were used to detect AD was 325567 biomarkers.

From 325567 biomarkers, 69 biomarkers were determined that may have a more significant with AD as shown in Table 6-14 with performance greater than or equal to 80% for sensitivity and specificity.

In this study, we found that the robust EEG biomarkers that can be used in AD diagnosis with performance close to 100% for sensitivity and specificity. Table 6-15 shows the probability distribution ratio for the selected EEG biomarkers, but Table 6-16 shows the probability distribution ratio for the selected EEG channels. These biomarkers and channels were evaluated to develop a robust diagnostic model for AD.

Table 6-15: EEG biomarkers that may have a more significant association with AD

Biomarker	Total	Probability distribution ratio
Alpha	5	29.41
Alpha/Theta	5	29.41
Alpha/Delta	4	23.53
Theta/Alpha	2	11.76
Delta/Alpha	1	5.88
Summation	17	100

Table 6-16: EEG channels that may have a more significant association with AD

EEG channel	Total	Probability distribution ratio
P3	11	12.50
T6	11	12.50
T3	9	10.23
C3	9	10.23
P4	9	10.23
O1	8	9.09
T4	7	7.95
PZ	7	7.95
CZ	6	6.82
O2	5	5.68
C4	3	3.41
T5	3	3.41
Summation	88	100

To illustrate the effect of AD on EEG signal. The changes in EEG due to AD can be shown for the 69 robust EEG biomarkers that may have a more significant association with AD, as shown in Table 6-14. Table 6-17 shows the changes in EEG signal due to AD for the 17 EEG features (e.g., ApEn(Alpha), ApEn(Alpha/Theta)) that may have a more significant association with AD for the three key characteristics of dementia EEG (i.e. slowing of the EEG, reduction in EEG complexity and reduction in EEG coherence). The shaded boxes in Table 6-17 referred to the decrease in EEG biomarkers due to AD, otherwise it means the increase. Table 6-17 shows the changes in EEG characteristics due to AD.

Table 6-17: Changes in EEG signal due to AD for the 17 robust EEG biomarker panels

Category	Method	Biomarker	AD	Norm
EEG complexity	ApEn	Alpha	■	
	ApEn	Alpha/Theta	■	
	HFD	Alpha	■	
	HFD	Theta/Alpha		■
	LZC	Alpha		■
	LZC	Alpha/Delta	■	
	TsEn	Alpha/Delta	■	
	TsEn	Alpha/Theta		■
EEG slowing	Power	Alpha	■	
	Power	Alpha/Delta	■	
	Power	Alpha/Theta	■	
	ZCI	Alpha	■	
	ZCI	Alpha/Theta	■	
	ZCI	Delta/Alpha		■
	ΔEEG_A	Alpha/Delta	■	
	ΔEEG_A	Theta/Alpha		■
EEG coherence	Coherence	Alpha/Theta	■	

We analysed the performance of the main three approaches on which EEG biomarkers are based (i.e., reduction in EEG complexity, slowing of the EEG signal, and decrease in EEG coherence) in quantifying changes in EEG due to AD. Therefore, we investigated the differences between the values of the slowing, complexity, and coherence measures derived from EEG signals of AD subjects to those from normal subjects. From Table 6-17, we can conclude AD cause changes in the characteristics of EEG. These characteristics include decrease in EEG complexity, slowing in EEG and decrease in EEG coherence among cortical regions. This finding is consistent with the finding in other studies that referred that AD leads to a decrease in the performance of information processing activity of the brain and these changes are reflected in the content of EEG signal [43][50][44][45][47][84][48][85][255].

6.5. Summary

AD causes a reduction in neuronal activity of the brain and this may be reflected in EEG signals, which may be used to detect AD accurately. The changes in the EEG may provide useful information about AD development which may be used to detect AD with high performance (e.g., close to 100%). We analysed eight promising methods for quantifying changes in EEG based on three approaches to EEG analysis (i.e., the decrease in EEG complexity, slowing in EEG, and decrease in EEG coherence among cortical regions). For the slowing of EEG, we analysed three promising methods (i.e., changes in the power of EEG, changes in EEG amplitude, and zero crossing interval). For the EEG complexity, four promising methods were analysed (i.e., approximation entropy, Tsallis entropy, Lempel Ziv Complexity, and Higuchi Fractal Dimension). For EEG connectivity, EEG coherence was analysed. We analysed EEG changes based on the five EEG frequency bands (i.e., delta, theta, alpha, beta, and gamma) and the 20 ratios between features in traditional EEG frequency bands. The results showed that AD detection with high performance could be achieved by combining multiple biomarkers to create a robust biomarker and this is consistent with the results in other studies; Mao et al. [256] suggested the combination of markers of AD can improve the identification effect for auxiliary diagnosis of AD. This study provides a framework for constructing robust biomarkers that could be used to detect AD with high performance for sensitivity and specificity closed to 100%. We develop a robust diagnostic model from 69 EEG biomarkers that have shown promise in AD diagnosis. We believe that the combination of these biomarkers could be used to provide a reliable model for AD diagnosis.

However, although the diagnostic model gave good results there are significant limitations in the study, the most important of which is the low sample sizes of the datasets for AD patients and normal subjects (37 AD and 35 normals). Given the

number of biomarkers involved, there is a potential for over-fitting of the model. Further investigation is necessary to identify the smallest subset of biomarkers that might give similar performance and if possible to increase the sample sizes.

Chapter 7. Optimisation of Robust EEG Based Biomarkers

7.1. Introduction

In Chapter 6, 69 robust EEG based biomarkers were identified (see Table 6-14) and combined into one diagnostic model. The combined biomarker consisted of 30 biomarkers from analysis of the reduction in EEG complexity, 33 biomarkers from analysis of the slowing of the EEG, and 6 biomarkers from the analysis of the reduction in EEG coherence. Although the results were good, this may not be the smallest possible subset of EEG biomarkers.

Further investigation is necessary to identify the smallest subset of EEG biomarkers from the 69 biomarkers. A new EEG dataset was used in the investigation to avoid bias if the same datasets were used. Using a new EEG dataset would also help to assess how the method would perform in different clinical settings. The new EEG datasets (C and D) were recorded in Rome and were provided by La Sapienza, University of Rome.

7.2. Methodology

10-fold cross-validation was used for developing the model to help mitigate against over-fitting. The selection of the training and testing datasets was performed randomly.

The EEG dataset used in this chapter consists of four datasets (A, B, C, and D). EEG dataset D was used for training and cross-validation. Datasets A, B and C were used for subsequent testing.

Dataset A includes 11 age-matched subjects over 65 years old (3 AD patients, and 8 normal subjects). Dataset B includes 41 subjects that were not perfectly age-matched with 24 normal subjects (10 males and 14 females) and 17 were probable AD patients (9 males and 8 females). Dataset C includes 10 mild AD patients and 10 normal age-matched healthy old subjects. Dataset D includes 20 normal age-matched healthy

old subjects and 20 mild AD patients. Altogether, a total of 112 subjects were involved, 57 AD and 55 normals. More details are given in Chapter 3.

The following steps outline the procedure for extracting the best combination of biomarkers from the 69 robust EEG based biomarkers.

1. Following the methodology in Chapter 6, the 69 biomarkers (see Table 6-14) were computed using dataset D: Steps 1 – 6 in Chapter 6, Section 6.2
2. Create panels of biomarkers from the 69 biomarkers (i.e. by combining one or more biomarkers). The maximum panel size is limited to 4 biomarkers to avoid the excessive number of possible panels and given that the goal is to find the smallest subset of biomarkers (i.e. panels with the fewest number of biomarkers) to give satisfactory performance compared to the use of all 69 biomarkers.
3. Develop and test a model for each panel to detect AD. In the study, an SVM model was used to combine the biomarkers in each panel [250].
4. Select panels and hence the subset of biomarkers that meet the diagnostic criteria and develop a diagnostic model from these.
5. Test the diagnostic model using unseen datasets (datasets A, B and C).

7.2.1. Biomarker selection

The methodological framework involves a thorough consideration of all 69 extracted biomarkers. Starting with the 69 biomarkers, finding the smallest subset of biomarkers involved creating panels of biomarkers by combining the biomarkers. The maximum number of biomarkers in a panel was limited to 4 because the goal is to find the smallest subset of biomarkers to detect AD with acceptable performance and the need to avoid the exponential increase in the number of possible biomarker panels as panel size increases. Limiting the panel size to 4 still yielded 919310 biomarker panels (see Table 7-1).

A machine learning model was developed for each biomarker panel. The performance of each model was assessed based on its sensitivity and specificity for detecting AD patients. SVM was used to develop the models. Biomarker panels with performance more than or equal to 80% for sensitivity and specificity were selected . This allowed us to determine the best performing biomarker panels.

Table 7-1: Number and distribution of panels with one, two, three and four biomarkers

No. of biomarkers in a panel	No. of panels
1	69
2	2346
3	52394
4	864501
Total	919310

7.2.2. Diagnostic model to detect AD

The next step is to select the smallest subset of biomarkers with acceptable performance. The selected biomarkers are then combined into one model to produce the best diagnostic model. To achieve this, first we select biomarker panels that satisfied the selection criterion and from these a subset of biomarkers was selected based on their performance and then combined into one biomarker panel. The new biomarker panel was used to develop a new SVM model. The performance of the SVM model is then assessed using unseen EEG dataset.

7.3. Results

7.3.1. Biomarker computations and selections

A new EEG dataset (dataset D) was used to compute the 919310 EEG biomarker panels. For each of the 919310 biomarker panels, an SVM diagnostic model was developed. To minimise overfitting, a 10-fold cross validation was used in model development [257]. As a result, 919310 SVM models were developed. Each biomarker panel of 919310 was assessed based on its performance and the number

of biomarkers in the panel. Biomarker panels with high performance and fewest number of biomarkers were selected to develop the fina model.

Table 7-2 summarises the performance of the single-biomarker panels in order of performance. Only the first six biomarkers in Table 7.1 have satisfactory performance (sensitivity and specificity equal to or greater than 80%).

Table 7-2: Performance of the 69 single-biomarker panels

EEG biomarker	Sen. %	Spec. %	Acc. %	F-measure %	MCC	PPV %	NPV %
TsEn(Alpha/theta(T6))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ZCI(Alpha/theta(P3))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ZCI(Delta/alpha(P3))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ΔEEG_A(Alpha/delta(T6))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ΔEEG_A(Theta/alpha(T3))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ΔEEG_A(Theta/alpha(T5))	100.00	100.00	100.00	100.00	1.00	100.00	100.00
ApEn(Alpha/theta(C4))	66.67	100.00	75.00	80.00	0.58	100.00	50.00
ΔPS(Alpha(P3))	69.23	85.71	75.00	78.26	0.52	90.00	60.00
ΔPS(Alpha(T6))	62.50	100.00	70.00	76.92	0.50	100.00	40.00
ΔPS(Alpha/delta(P4))	62.50	100.00	70.00	76.92	0.50	100.00	40.00
ΔPS(Alpha/delta(O1))	62.50	100.00	70.00	76.92	0.50	100.00	40.00
ΔPS(Alpha/theta(T3))	62.50	100.00	70.00	76.92	0.50	100.00	40.00
ΔPS(Alpha/theta(C3))	62.50	100.00	70.00	76.92	0.50	100.00	40.00
Coh(Alpha/theta(F4-T4))	72.73	77.78	75.00	76.19	0.50	80.00	70.00
Coh(Alpha/theta(Fz-F8))	64.29	83.33	70.00	75.00	0.44	90.00	50.00
LZC(Alpha/delta(T6))	55.56	100.00	60.00	71.43	0.33	100.00	20.00
ΔPS(Alpha/theta(P3))	61.54	71.43	65.00	69.57	0.31	80.00	50.00
LZC(Alpha/delta(P3))	56.25	75.00	60.00	69.23	0.25	90.00	30.00
LZC(Alpha/delta(C4))	52.63	100.00	55.00	68.97	0.23	100.00	10.00
HFD(Alpha(PZ))	52.94	66.67	55.00	66.67	0.14	90.00	20.00
HFD(Alpha(O1))	52.94	66.67	55.00	66.67	0.14	90.00	20.00
LZC(Alpha(O1))	52.94	66.67	55.00	66.67	0.14	90.00	20.00
LZC(Alpha/delta(T4))	52.94	66.67	55.00	66.67	0.14	90.00	20.00
ΔPS(Alpha/theta(P4	52.94	66.67	55.00	66.67	0.14	90.00	20.00
ZCI(Alpha(O2))	50.00	0.00	50.00	66.67	0.00	100.00	0.00
LZC(Alpha/delta(CZ))	50.00	50.00	50.00	64.29	0.00	90.00	10.00
ΔEEG_A(Theta/alpha(PZ))	60.00	60.00	60.00	60.00	0.20	60.00	60.00
ΔPS(Alpha/theta(PZ))	60.00	60.00	60.00	60.00	0.20	60.00	60.00
Coh(Alpha/theta(Fp2-F4))	71.43	61.54	65.00	58.82	0.31	50.00	80.00
ApEn(Alpha(O1))	80.00	60.00	65.00	53.33	0.35	40.00	90.00
ΔPS(Alpha(P4))	80.00	60.00	65.00	53.33	0.35	40.00	90.00
ΔPS(Alpha/delta(T4))	80.00	60.00	65.00	53.33	0.35	40.00	90.00
LZC(Alpha(C3))	55.56	54.55	55.00	52.63	0.10	50.00	60.00
HFD(Theta/alpha(T5))	42.86	33.33	40.00	50.00	-0.22	60.00	20.00

LZC(Alpha(PZ))	66.67	57.14	60.00	50.00	0.22	40.00	80.00
ΔPS(Alpha/theta(T4))	66.67	57.14	60.00	50.00	0.22	40.00	80.00
Coh(Alpha/theta(T3-P4))	42.86	33.33	40.00	50.00	-0.22	60.00	20.00
LZC(Alpha(T6))	57.14	53.85	55.00	47.06	0.10	40.00	70.00
ΔEEG_A(Theta/alpha(P3))	57.14	53.85	55.00	47.06	0.10	40.00	70.00
ΔEEG_A(Theta/alpha(P4))	57.14	53.85	55.00	47.06	0.10	40.00	70.00
ΔPS(Alpha/delta(P3))	57.14	53.85	55.00	47.06	0.10	40.00	70.00
Coh(Alpha/theta(F8-P4))	37.50	0.00	30.00	46.15	-0.50	60.00	0.00
ApEn(Alpha(P3))	50.00	50.00	50.00	44.44	0.00	40.00	60.00
Coh(Alpha/theta(T3-T4))	38.46	28.57	35.00	43.48	-0.31	50.00	20.00
ΔPS(Alpha(C3))	42.86	46.15	45.00	35.29	-0.10	30.00	60.00
HFD(Alpha(O2))	100.00	55.56	60.00	33.33	0.33	20.00	100.00
LZC(Alpha(CZ))	50.00	50.00	50.00	28.57	0.00	20.00	80.00
ZCI(Alpha/theta(PZ))	23.08	0.00	15.00	26.09	-0.73	30.00	0.00
TsEn(Alpha/delta(T6))	33.33	42.86	40.00	25.00	-0.22	20.00	60.00
TsEn(Alpha/theta(T3))	33.33	42.86	40.00	25.00	-0.22	20.00	60.00
ΔPS(Alpha(T3))	33.33	42.86	40.00	25.00	-0.22	20.00	60.00
ΔPS(Alpha/delta(T3))	50.00	50.00	50.00	16.67	0.00	10.00	90.00
ApEn(Alpha(T6))	33.33	47.06	45.00	15.38	-0.14	10.00	80.00
TsEn(Alpha/delta(C3))	33.33	47.06	45.00	15.38	-0.14	10.00	80.00
TsEn(Alpha/theta(CZ))	16.67	35.71	30.00	12.50	-0.44	10.00	50.00
TsEn(Alpha/theta(P3))	16.67	35.71	30.00	12.50	-0.44	10.00	50.00
HFD(Alpha(T5))	0.00	47.37	45.00	0.00	-0.23	0.00	90.00
HFD(Alpha(T6))	0.00	47.37	45.00	0.00	-0.23	0.00	90.00
HFD(Theta/alpha(T3))	0.00	44.44	40.00	0.00	-0.33	0.00	80.00
LZC(Alpha/delta(P4))	0.00	41.18	35.00	0.00	-0.42	0.00	70.00
TsEn(Alpha/theta(C3))	0.00	44.44	40.00	0.00	-0.33	0.00	80.00
TsEn(Alpha/theta(O1))	0.00	50.00	50.00	0.00	0.00	0.00	100.00
ZCI(Alpha(T6))	0.00	0.00	0.00	0.00	-1.00	0.00	0.00
ZCI(Alpha(O1))	0.00	50.00	50.00	0.00	0.00	0.00	100.00
ZCI(Alpha/theta(C3))	0.00	28.57	20.00	0.00	-0.65	0.00	40.00
ZCI(Alpha/theta(P4))	0.00	37.50	30.00	0.00	-0.50	0.00	60.00
ZCI(Delta/alpha(C3))	0.00	50.00	50.00	0.00	0.00	0.00	100.00
ZCI(Delta/alpha(O1))	0.00	50.00	50.00	0.00	0.00	0.00	100.00
ΔEEG_A(Theta/alpha(C3))	0.00	50.00	50.00	0.00	0.00	0.00	100.00

As shown in Table 7-2 the performance of biomarkers started at 100% for sensitivity and specificity for the first six biomarkers and then droped significantly to 66.67% for sensitivity for the biomarker ApEn(Alpha/theta(C4)). Our criterion is to select the biomarkers with the highest performace and so the first six biomarkers were selected from single-biomarker. This table shows that the biomarkers [TsEn(Alpha/theta(T6))], [ZCI(Alpha/theta(P3))], [ZCI(Delta/alpha(P3))], [ΔEEG_A(Alpha/delta(T6))],

[ΔEEG_A(Theta/alpha(T3))], and Δ[EEG_A(Theta/alpha(T5))] have the highest performance, and that biomarkers [ZCI(Delta/alpha(C3))], [ZCI(Delta/alpha(O1))], and [ΔEEG_A(Theta/alpha(C3))] have the lowest performance.

Table 7-3 summarises the performance of the best two-biomarker panels (out of the 2346 two-biomarker panels in Table 7-1), with sensitivity and specificity greater than or equal to 80%.

Table 7-3: Summary of the performance of the best two-biomarker panels

Biomarker 1	Biomarker 2	Sen. %	Spec. %	Acc. %
LZC(Alpha(C3))	**ZCI(Alpha/theta(P3))**	100.00	100.00	100.00
LZC(Alpha/delta(P3))	**ZCI(Delta/alpha(P3))**	100.00	100.00	100.00
TsEn(Alpha/theta(CZ))	**ΔEEG_A(Theta/alpha(T3))**	100.00	100.00	100.00
ΔEEG_A(Theta/alpha(T5))	Coh(Alpha/theta(T3-P4))	100.00	100.00	100.00
LZC(Alpha(O1))	**TsEn(Alpha/theta(T6))**	90.91	100.00	95.00
TsEn(Alpha/theta(T6))	ΔPS(Alpha(T3))	90.91	100.00	95.00
ZCI(Alpha/theta(P3))	**ΔPS(Alpha/theta(T3))**	90.91	100.00	95.00
ΔEEG_A(Alpha/delta(T6))	ΔPS(Alpha(T6))	90.91	100.00	95.00
ZCI(Alpha/theta(P3))	Coh(Alpha/theta(F4-T4))	83.33	100.00	90.00
ZCI(Alpha/theta(P3))	ΔEEG_A(Theta/alpha(P4))	81.82	88.89	85.00

As shown in Table 7-3 the performance of biomarkers was started at 100% for sensitivity and specificity for the first four biomarkers and then reduced gradually to 81.82% for sensitivity and 88.89% for specificity for biomarker the [ZCI(Alpha/theta(P3)), ΔEEGA(Theta/alpha(P4))].

Table 7-4 summarises the performance of the best three-biomarker panels (out of 52394 three-biomarker panels in Table 7-1), with sensitivity and specificity greater than or equal to 80%.

Table 7-4: Summary of the performance of the best three-biomarker panels

Biomarker 1	Biomarker 2	Biomarker 3	Sen. %	Spec. %	Acc. %
ApEn(Alpha(P3))	ApEn(Alpha(T6))	**TsEn(Alpha/theta(T6))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha(O1))	**ZCI(Delta/alpha(P3))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha/theta(C4))	**ZCI(Alpha/theta(P3))**	100.00	100.00	100.00
ApEn(Alpha(P3))	HFD(Alpha(PZ))	**ΔEEG_A(Theta/alpha(T3))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ZCI(Alpha/theta(P3))	**ΔEEG_A(Theta/alpha(P3))**	90.91	100.00	95.00
LZC(Alpha(O1))	**ZCI(Delta/alpha(P3))**	Coh(Alpha/theta(T3-T4))	90.91	100.00	95.00
HFD(Alpha(PZ))	LZC(Alpha(O1))	**ΔEEG_A(Theta/alpha(T5))**	100.00	90.91	95.00
LZC(Alpha(O1))	**ΔEEG_A(Alpha/delta(T6))**	ΔPS(Alpha/theta(C3))	100.00	90.91	95.00
ΔEEG_A(Theta/alpha(T5))	ΔPS(Alpha/delta(T3))	ΔPS(Alpha/theta(P4))	100.00	90.91	95.00
ΔEEG_A(Alpha/delta(T6))	Coh(Alpha/theta(F4-T4))	Coh(Alpha/theta(T3-P4))	100.00	83.33	90.00

As shown in Table 7-4 the performance of biomarkers started at 100% for sensitivity and specificity for the first four biomarkers and then reduced gradually to 100% for sensitivity and 83.33% for specificity for the biomarker [ΔEEG_A(Alpha/delta(T6)) Coh(Alpha/theta(F4-T4)) Coh(Alpha/theta(T3-P4))] .

Table 7-5 summarises the performance of the best four-biomarker panels (out of the 864501 four-biomarker panels in Table 7-1), with sensitivity and specificity greater than or equal to 80%.

Table 7-5: Summary of the performance of the best four-biomarker panels

Biomarker 1	Biomarker 2	Biomarker 3	Biomarker 4	Sen. %	Spec. %	Acc. %
ApEn(Alpha(P3))	ApEn(Alpha(T6))	ApEn(Alpha(O1))	**TsEn(Alpha/theta(T6))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	LZC(Alpha(CZ))	**ZCI(Delta/alpha(P3))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	LZC(Alpha/delta(T6))	**ZCI(Delta/alpha(P3))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	ZCI(Alpha(O1))	**ΔEEG_A(Theta/alpha(T5))**	100.00	100.00	100.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	LZC(Alpha/delta(CZ))	**ZCI(Alpha/theta(P3))**	90.91	100.00	95.00
ApEn(Alpha(P3))	**ΔEEG_A(Theta/alpha(T3))**	**TsEn(Alpha/theta(T6))**	ΔPS(Alpha/delta(P3))	90.91	100.00	95.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	**ΔEEG_A(Alpha/delta(T6))**	ΔPS(Alpha/theta(PZ))	100.00	90.91	95.00
ApEn(Alpha(P3))	HFD(Theta/alpha(T5))	LZC(Alpha(O1))	**ZCI(Alpha/theta(P3))**	100.00	90.91	95.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	**ΔEEG_A(Alpha/delta(T6))**	ΔPS(Alpha(P3))	83.33	100.00	90.00
ApEn(Alpha(P3))	ApEn(Alpha(T6))	ZCI(Alpha/theta(P3))	**ΔEEG_A(Theta/alpha(P4))**	81.82	88.89	85.00

As shown in Table 7-5 the performance of biomarkers started at 100% for sensitivity and specificity for the first four biomarkers and reduced gradually to 81.82% for sensitivity and 88.89% for specificity for biomarker [ApEn(Alpha(P3)) ApEn(Alpha(T6)) ZCI(Alpha/theta(P3)) ΔEEG_A(Theta/alpha(P4))].

7.3.2. Diagnostic model to detect AD

The key goal of this study is to find the smallest subset of biomarkers that may give a similar or close performance to that of the robust 69 biomarkers in Chapter 6. Tables 7-2 to 7-5 summarise the performance of the best biomarker panels for single, two, three, and four biomarker panels, respectively.

To select the best biomarkers from Tables 7-2 to 7-5, we focused on the smallest subset of biomarkers that have a high performance in AD detection. Based on that criterion, we found from analysis of the biomarker panels in Tables 7-2 to 7-5 that six biomarkers are shared between the best performing biomarker panels in Tables 7-2 to 7-5. To illustrate the single-biomarker panel TsEn(Alpha/theta(T6)) in Table 7-2 appears twice in Table 7-3, once in Table 7-4, and twice in Table 7-5. Therefore, the single-biomarker panel TsEn(Alpha/theta(T6)) was selected as it has a high performance and it has only one parameter. Table 7-6 summarises the first six single-biomarker panels in Table 7-2 and their occurrence in Tables 7-3, 7-4, and 7-5. Table 7-6 summarises the smallest subset of biomarkers that have a high performance in AD detection (6 out of 919310). These six biomarkers represent the smallest subset of biomarkers that have a high performance in AD detection. The six biomarkers were combined into one SVM model. Dataset D using 10-fold cross-validation was used in the development of the final SVM model.

Table 7-6: The smallest subset of biomarkers that have a high performance in AD detection and their occurrence in Tables 7-3 to 7- 5

No.	EEG biomarker	Table 7-3	Table 7-4	Table 7-5	Total
1	TsEn(Alpha/theta(T6))	2	1	2	5
2	ZCI(Alpha/theta(P3))	4	2	3	9
3	ZCI(Delta/alpha(P3))	1	2	2	5
4	ΔEEG_A(Alpha/delta(T6))	1	3	2	6
5	ΔEEG_A(Theta/alpha(T3))	1	1	1	3
6	ΔEEG_A(Theta/alpha(T5))	1	1	1	3

The resulting model was then tested using unseen datasets. In particular, datasets A, B, and C were used to validate the developed model.

The performance of the final model using dataset C was 100% for sensitivity and specificity, respectively as shown in Figure 7-1. Also, the performance of the final diagnostic model using dataset A was also 100% for sensitivity and specificity, respectively.

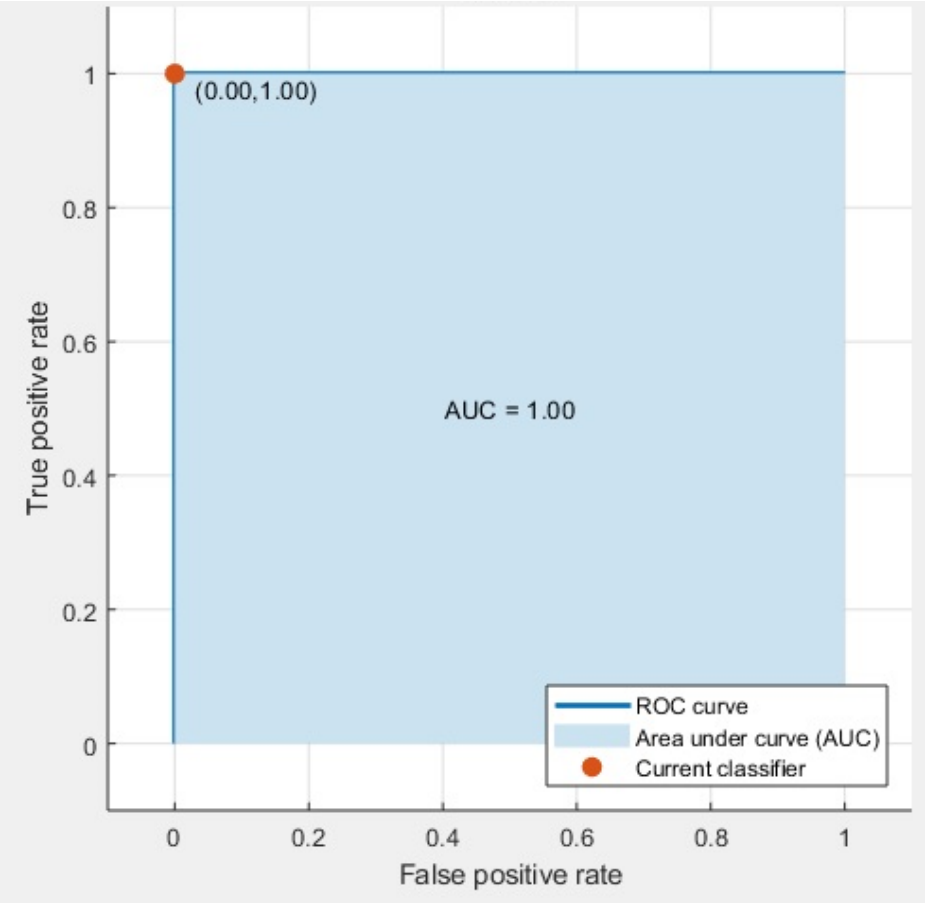

Figure 7-1: ROC and AUC of the final diagnostic model based on a subset of six biomarkers using dataset C

The performance of final model using dataset B was 85% for sensitivity and 100% for specificity. Figure 7-2 shows the ROC and AUC for the final diagnostic model based on dataset B.

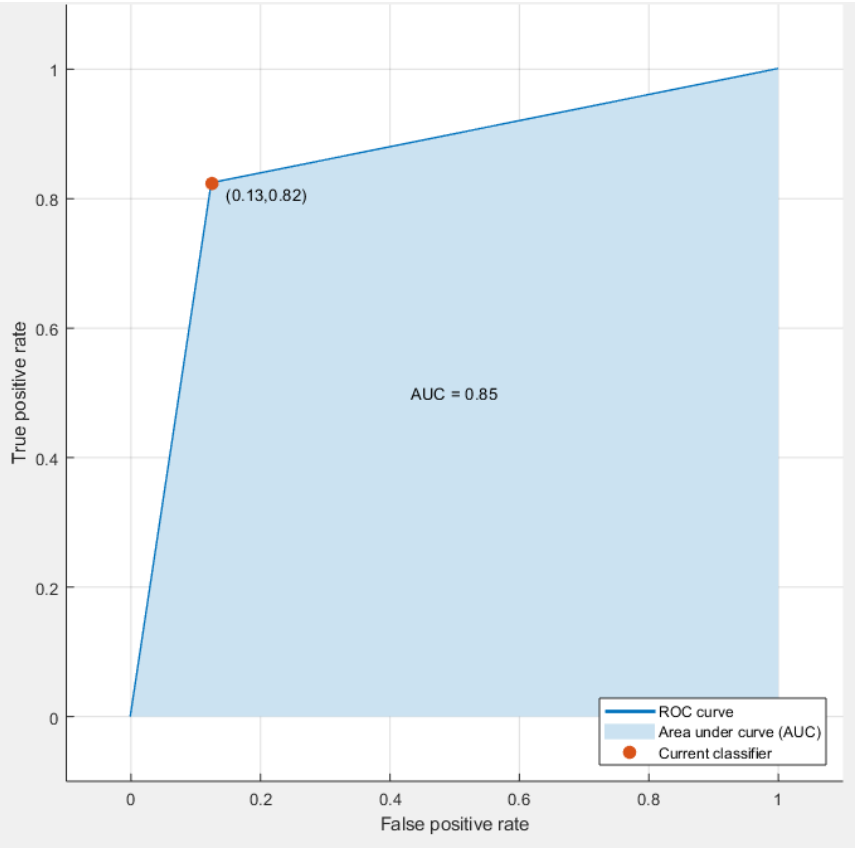

Figure 7-2: ROC and AUC of the final diagnostic model based on a subset of six biomarkers using dataset B

7.4. Discussions

Identification of the best possible subset of EEG biomarkers for AD detection requires a detailed investigation of the subsets of the 69 robust EEG biomarkers found in Chapter 6. Starting with 69 biomarkers, the number of possible subsets (or panels) of the 69 biomarkers grows exponentially with the number of biomarkers in the subsets. To keep the number of possible biomarker panels (or subsets of biomarkers) manageable and given that the goal is to find the smallest subset of biomarkers to give satisfactory performance, the number of biomarkers in a panel was limited to four. The best performing 1, 2, 3 and 4 biomarker panels were identified and analysed. The biomarkers common in the best performing biomarker panels and which also have high performance were selected. The outcome of this is the six biomarkers as shown in Table 7-6. The selected biomarkers were then used to

develop the final diagnostic model which is then tested using unseen EEG datasets i.e. datasets A, B and C (see Figures 7-1, and 7-2). To mitigate over-fitting, a10-fold cross validation was used in model development [257].

The selected biomarkers included one biomarker based on the reduction in EEG complexity i.e., TsEn(Alpha/theta(T6)), and 5 biomarkers based on the slowing of the EEG i.e., ZCI(Alpha/theta(P3)), ZCI(Delta/alpha(P3)), ΔEEG_A(Alpha/delta(T6)), ΔEEG_A(Theta/alpha(T3)), and ΔEEG_A(Theta/alpha(T5)). Furthermore, analysis of the results show that the EEG channels in the temporal and parietal lobes (i.e. T6, T3, T5, and P3) gave better results compared to other channels that relate to other lobes e.g., the frontal lobe.

Although the performance of the diagnostic model is good, there are a number of significant limitations of the study. First, the size of the dataset used in the model development and testing of the model is small. In the study, 40 cases were used for training the model (i.e. 20 AD patients and 20 normal subjects (dataset D)), and 72 cases for testing I (i.e. 30 AD patients and 42 normal subjects (datasets A, B, and C)). Although this compares well with the size of the dataset in other studies (e.g., Amezquita-Sanchez et al. [258] used 74 cases, 37 MCI and 37 AD patients with an accuracy of 90.3%. Chai et al. [259] used 20 cases, 10 AD patients and 10 healthy people with AUC reaching 0.89), the number of subjects is still quite low and there is a risk of over-fitting. For clinical acceptance, there is a need for further study using larger datasets from different settings to assess the full potential of the model and methodology. Finding the smallest subset of biomarkers from 69 biomarkers involves examining excessively large number of biomarker panels. As a result, the maximum number of biomarkers in a panel was limited to four and the determination of the smallest subset was based on these. Thus, the subset we used to develop the final diagnostic model may not be the optimal subset (in terms of performance, number

and type of biomarkers). In future, it should be possible to examine all the possible subsets of biomarkers if the computational resources are available.

7.5. Summary

In this chapter, we have developed a method to find the smallest set of biomarkers from the 69 biomarkers developed in Chapter 6. This led to the selection of six biomarkers which featured in the best performing 1, 2, 3 and 4 biomarker panels. A diagnostic model based on the six biomarkers were then developed and tested with unseen datasets. A new dataset was used to develop the model. 10 fold cross-validation was used throughout in the model development to mitigate against over-fitting. The performance of the final model compares well with other studies, but further studies is necessary using a much larger dataset to develop and fully evaluate the diagnostic model.

Chapter 8. Review, Conclusions and Future Work

8.1. Review

Alzheimer's disease is a progressive disorder that affects cognitive brain functions and begins many years prior to any clinical manifestations. A biomarker that provides a quantitative measure of early stage changes in the brain due to AD would therefore be useful for the early diagnosis of AD. However, because up to 50% of dementia sufferers do not receive a formal diagnosis, this would involve dealing with large numbers of people. Thus, there is a need for accurate, low-cost, robust, and easy to use biomarkers that can be used to detect AD in its early stages.

Degeneration of brain cells due to AD begins many years before any clinical issues arise [8][15][16][17][18][19]. Early diagnosis of AD will therefore contribute to the development of effective treatments that could slow, stop or prevent significant cognitive decline [18][20][21]. Early diagnosis of AD could also be useful for identifying dementia sufferers who have not received a formal early diagnosis and will provide them with an opportunity to access appropriate health care services [22][23][24].

Potentially, the electroencephalogram (EEG) can play a valuable role in the early diagnosis of AD [15][23][24][42][43][44] as it is non-invasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics [23][24][43][45][46]. The fundamental utility of EEG in detecting changes in brain signals, even in the preclinical stage of the disease, has been widely demonstrated [43][47][48]. Thus, EEG biomarkers may be used as a first line decision-support tool in AD diagnosis [15][45] and could complement other AD biomarkers [37].

In this thesis, we described research into the development of EEG biomarkers to detect AD based on an analysis of changes in the EEG. The most characteristic features of an EEG that detects AD are the slowing of the EEG, a decrease in EEG coherence, and a reduction in EEG complexity. These changes can be quantified as

biomarkers of AD. In this study, we focused on detecting EEG features that have a more significant association with AD. Promising EEG features were extracted from the main three techniques of EEG analysis and used to develop EEG biomarkers with high diagnostic performance. A cross-sectional EEG dataset was used in this study. Four measures of complexity were investigated: Tsallis entropy, Higuchi fractal dimension, Lempel-Ziv complexity, and approximation entropy. Two slowing measures were also investigated: zero-crossing intervals, and changes in the power spectrum of EEG. We also proposed a new approach to quantify the slowing of the EEG that was based on analysing changes in EEG amplitude. EEG connectivity was investigated by analysing the coherence of electrical structural connectivity among cortical regions of the brain.

A thorough investigation was performed for the eight methods selected. The EEG features were extracted in both time and frequency domains. In the frequency domain, traditional EEG frequency bands and the ratios between those bands were explored. The results showed that EEG biomarkers extracted from the frequency bands were better at detecting AD than EEG biomarkers extracted from the entire EEG record. This is because each EEG frequency band is associated with specific brain functions, each lobe of the brain carries specific tasks, and bands do not emit the same brain wave frequency simultaneously. Thus, any decline in brain function may be reflected in the content of EEG frequency bands. This complements our results showing that slowing of the EEG, a decrease in EEG coherence, and a reduction in EEG complexity are potentially good biomarkers for AD diagnosis.

This study thus provides a framework for constructing robust EEG biomarkers that can be used to detect AD with high diagnostic performance.

8.2. Conclusions

AD causes changes in the EEG and these are thought to be associated with functional disconnections among cortical areas resulting from the death of brain cells. Therefore, EEG analysis may provide valuable information about brain dynamics in patients with AD. AD causes a reduction in neuronal activity in the brain and this may be reflected in EEG signals. Potentially, the EEG may be able to detect changes in brain signals even in the early stages of the disease. Thus, EEG may be used as a first line decision-support tool in AD diagnosis and could complement other AD biomarkers.

EEG biomarkers can therefore be used to detect AD in its early stages and have exhibited high performance in AD detection. Our results showed that EEG biomarkers are stable because their performance remained high when different evaluation methods were used in both time and frequency domains. These were based on traditional EEG frequency bands and the ratio between those bands. Thus, the EEG signal can be used to develop robust EEG biomarkers for AD detection.

The results showed that EEG analysis based on specific frequency bands is more valuable than an analysis of the entire EEG signal because each EEG frequency band has specific functions. Furthermore, we found that the analysis of a specific EEG channel may provide better results in terms of AD detection because the pathological analysis of AD shows that it spreads from specific parts of the brain to others.

In our analysis, we found that combining of multiple EEG biomarkers provides a high level of performance. This may be because each method exploits different characteristics of the EEG signal; combining those characteristics into one biomarker therefore means the inclusion of more characteristics associated with AD.

However, this approach creates a large number of EEG biomarkers and it can be difficult to find the best combination of EEG biomarkers to use in the diagnostic model.

There is also the possibility of over-fitting in the model. However, our study suggests that combining the best biomarkers of EEG can lead to the development of an EEG diagnostic model with a high diagnostic performance.

However, large datasets of EEG will be neededto fully evaluate the usefulness of the methodology.

Future work

The outcomes of this study suggest several areas that can be addressed in future work. These include the following:

1. Evaluate the techniques using a larger EEG dataset

   As noted previously, there are over 46.8 million individuals living with dementia worldwide and a rapid increase in the number of people living with AD and other forms of dementia due to the ageing population presents a major challenge to health and social care systems. For these reasons, it is necessary to include a large EEG dataset because this will permit exploration of a wide variety of AD cases from the entire AD population and may also help identify reasons for the development of AD.

2. Evaluate the use of EEG biomarkers to detect changes in EEG signals in different stages of AD using a longitudinal EEG dataset (normal, MCI, and AD).

   AD progression occurs in different stages ranging from moderate to severe [260][261]. The treatment of AD at an early stage is more effective than at a later stage. Therefore, it is important to analyse the progression of AD as early as possible. This will also help to determine the extent of AD progression at each stage.

3. Evaluate the use of EEG biomarkers in detecting different types of dementia, such as AD, vascular dementia, dementia with Lewy bodies (DLB), and mixed dementia

because of potential overlap between AD and other types of dementia [262][263]. Therefore, biomarkers that can distinguish between different types of dementia may help to detect AD in its early stages and contribute to the development of effective treatments that could slow, stop, or prevent significant cognitive decline [18][20][21].

4. Collect different types of AD dataset, such as MRI, EEG, amyloid-β, and CSF, at different times (maybe every four months) from the same patient (i.e., construct a bio profile for each patient) as this will help to benchmark results and to provide new insights.

   Studies investigating different biomarkers of AD from the same patient are relatively rare yet may help counter the rapid increase in AD cases. There is currently no certain clinical biomarker of AD that can be used in clinical diagnosis [264]. However, a certain diagnosis of AD can be made following an autopsy of patients' brains [265][266]. We suggest that analysing different types of biomarker for the same individual may therefore provide a more precise diagnosis of AD. Furthermore, this may help in detecting the progression of AD long before its symptoms become manifest.

5. Investigate different types of EEG analysis

   This will help in the evaluation of available EEG biomarkers and may contribute to the development of new methods of analysis.

Appendix 1: MATLAB functions for filter EEG signal

1. (IIR) Chebyshev-II bandpass filter for theta band

```
function [out] = delta(x,Fs)
% MATLAB Code
% Generated by MATLAB(R) 9.3 and DSP System Toolbox 9.5.
% Chebyshev Type II Bandpass filter designed using FDESIGN.BANDPASS.
% All frequency values are in Hz.
% x is the input signal

% Fs is Sampling Frequency
Fstop1 = 0.1;         % First Stopband Frequency
Fpass1 = 0.5;         % First Passband Frequency
Fpass2 = 4;           % Second Passband Frequency
Fstop2 = 4.5;         % Second Stopband Frequency
Astop1 = 60;          % First Stopband Attenuation (dB)
Apass  = 1;           % Passband Ripple (dB)
Astop2 = 80;          % Second Stopband Attenuation (dB)
match  = 'passband';  % Band to match exactly

% Construct an FDESIGN object and call its CHEBY2 method.
h  = fdesign.bandpass(Fstop1, Fpass1, Fpass2, Fstop2, Astop1, Apass,
...
    Astop2, Fs);
Hd = design(h, 'cheby2', 'MatchExactly', match);
out=filter(Hd,x)
% [EOF]
```

2. (IIR) Chebyshev-II bandpass filter for theta band

```
function [out] = theta(x,Fs)

% MATLAB Code
% Generated by MATLAB(R) 9.3 and DSP System Toolbox 9.5.
% Chebyshev Type II Bandpass filter designed using FDESIGN.BANDPASS.
% All frequency values are in Hz.
% x is the input signal

% Fs is Sampling Frequency
Fstop1 = 3.5;         % First Stopband Frequency
Fpass1 = 4;           % First Passband Frequency
Fpass2 = 8;           % Second Passband Frequency
Fstop2 = 8.5;         % Second Stopband Frequency
Astop1 = 60;          % First Stopband Attenuation (dB)
Apass  = 1;           % Passband Ripple (dB)
Astop2 = 80;          % Second Stopband Attenuation (dB)
match  = 'passband';  % Band to match exactly
% Construct an FDESIGN object and call its CHEBY2 method.
h  = fdesign.bandpass(Fstop1, Fpass1, Fpass2, Fstop2, Astop1, Apass,
...
                      Astop2, Fs);
Hd = design(h, 'cheby2', 'MatchExactly', match);
out=filter(Hd,x)
% [EOF]
```

3. (IIR) Chebyshev-II bandpass filter for alpha band

```
function [out] = alpha(x,Fs)

% MATLAB Code
% Generated by MATLAB(R) 9.3 and DSP System Toolbox 9.5.
% Chebyshev Type II Bandpass filter designed using FDESIGN.BANDPASS.
% All frequency values are in Hz.
% x is the input signal

% Fs is Sampling Frequency

Fstop1 = 7.5;         % First Stopband Frequency
Fpass1 = 8;           % First Passband Frequency
Fpass2 = 12;          % Second Passband Frequency
Fstop2 = 12.5;        % Second Stopband Frequency
Astop1 = 60;          % First Stopband Attenuation (dB)
Apass  = 1;           % Passband Ripple (dB)
Astop2 = 80;          % Second Stopband Attenuation (dB)
match  = 'passband';  % Band to match exactly

% Construct an FDESIGN object and call its CHEBY2 method.
h  = fdesign.bandpass(Fstop1, Fpass1, Fpass2, Fstop2, Astop1, Apass,
...
                      Astop2, Fs);
Hd = design(h, 'cheby2', 'MatchExactly', match);
out=filter(Hd,x)

% [EOF]
```

4. (IIR) Chebyshev-II bandpass filter for beta band

```
function [out] = beta(x,Fs)

% MATLAB Code
% Generated by MATLAB(R) 9.3 and DSP System Toolbox 9.5.
% Chebyshev Type II Bandpass filter designed using FDESIGN.BANDPASS.
% All frequency values are in Hz.
% x is the input signal

% Fs is Sampling Frequency

Fstop1 = 11.5;        % First Stopband Frequency
Fpass1 = 12;          % First Passband Frequency
Fpass2 = 30;          % Second Passband Frequency
Fstop2 = 30.5;        % Second Stopband Frequency
Astop1 = 60;          % First Stopband Attenuation (dB)
Apass  = 1;           % Passband Ripple (dB)
Astop2 = 80;          % Second Stopband Attenuation (dB)
match  = 'passband';  % Band to match exactly

% Construct an FDESIGN object and call its CHEBY2 method.
h  = fdesign.bandpass(Fstop1, Fpass1, Fpass2, Fstop2, Astop1, Apass,
...
                      Astop2, Fs);
Hd = design(h, 'cheby2', 'MatchExactly', match);
out=filter(Hd,x)

% [EOF]
```

5. (IIR) Chebyshev-II bandpass filter for gamma band

Suppose the EEG data was saved in excel file e.g., Sample_EEG.xlsx.

```
function [out] = gamma(x,Fs)

% MATLAB Code
% Generated by MATLAB(R) 9.3 and DSP System Toolbox 9.5.
% Chebyshev Type II Bandpass filter designed using FDESIGN.BANDPASS.

% All frequency values are in Hz.
% x is the input signal

% Fs is Sampling Frequency

Fstop1 = 29.5;        % First Stopband Frequency
Fpass1 = 30;          % First Passband Frequency
Fpass2 = 45;          % Second Passband Frequency
Fstop2 = 45.5;        % Second Stopband Frequency
Astop1 = 60;          % First Stopband Attenuation (dB)
Apass  = 1;           % Passband Ripple (dB)
Astop2 = 80;          % Second Stopband Attenuation (dB)
match  = 'passband';  % Band to match exactly

% Construct an FDESIGN object and call its CHEBY2 method.
h  = fdesign.bandpass(Fstop1, Fpass1, Fpass2, Fstop2, Astop1, Apass,
...
                      Astop2, Fs);
Hd = design(h, 'cheby2', 'MatchExactly', match);
out=filter(Hd,x)

% [EOF]
```

6. Applying (IIR) Chebyshev-II bandpass filter into a real EEG signal

```
num= xlsread('Sample_EEG.xlsx');
% Sample_EEG.xlsx is an EEG signal
x1=num(1:1280);

d1=delta(x1,128);
% delta is the Chebyshev Type II Bandpass filter

t1=theta(x1,128);
% theta is the Chebyshev Type II Bandpass filter

a1=alpha(x1,128);
% alpha is the Chebyshev Type II Bandpass filter

b1=beta(x1,128);
% beta is the Chebyshev Type II Bandpass filter

g1=gamma(x1,128);
```

```
% gaMMA is the Chebyshev Type II Bandpass filter

lim1=0:128:1280;
figure
subplot(6,1,1)
plot(x1)
xlim([0 1280]);
min1=min(x1);
max1=max(x1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Original EEG signal');


subplot(6,1,2)
plot(d1)
xlim([0 1280]);
min1=min(d1);
max1=max(d1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Delta 0-4Hz');



subplot(6,1,3)
plot(t1)
xlim([0 1280]);
min1=min(a1);
max1=max(a1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Theta 4-8Hz');

subplot(6,1,4)
plot(a1)
xlim([0 1280]);
min1=min(a1);
max1=max(a1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Alpha 8-13Hz');

subplot(6,1,5)
plot(b1)
```

```
xlim([0 1280]);
min1=min(b1);
max1=max(b1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Beta 13-30Hz');

subplot(6,1,6)
plot(g1)
xlim([0 1280]);
min1=min(g1);
max1=max(g1);
ylim([min1 max1]);
xticks([lim1]);
xlabel('Time');
ylabel('EEG amplitude');
title('Gamma 30-45Hz');
```

Appendix 2: MATLAB functions for computing EEG signal processing methods

1. Changes in the EEG amplitude (ΔEEGA) computation function

```
function [ x ] = EEG_amplitude(x2,Fs1)

% X2 is EEG signal
% Fs1 is sampling frequency

[r1 c1]=size(x2);
Seconds1=(r1/Fs1);
for m1=1:c1
    Y2=x2(:,m1);
    [t1 t2]=size(Y2);
    Y2=reshape(Y2,t2,t1);
    i=1;
    for k=1:Seconds1
        for j=1:Fs1
            times(k,j)=Y2(1,i);
            i=i+1;
        end;
    end;
    avg1(m1)=mean(times(:));
end;
mast_avg=reshape(avg1,1,m1);
x= mast_avg ;
end
```

2. Zero-crossing intervals (ZCI) computation function

```
function zci = zci_function( x,fs )

% This function to compute the ZCI method for one EEG file
% X is EEG signal
% fs is sampling frequency


[m n]=size(x);
data3(m,n)=0;
t=((0:m-1)/fs)';
dt=0.5;
frac_dim(1:n)=0;
for j=1:n
    data2(1:m,1:6)=0;
    data=x(:,j);
    zero_array=[];
    k=0;
    for i=2:m
          if (data(i,1)<0 && data(i-1,1)>0)
            data(i,2)=data(i,1);
            k=k+1;
            zero_array(k,1)=i;
        end;
    end
    data2(1:m1,1)=zero_array;
    temp=0;
```

```
    p1=1;
    p2=1;
    for i=2:m1
        temp=(data2(i,2)-data2(p1,2));
        if temp>dt
           data2(i-1,3)=(data2(i-1,2)-data2(p1,2));
           data2(i-1,4)=p2;
        else
           p2=p2+1;
           p1=i;
        end
    end
    inst1=data2(find(data2(:,4)>0),4);
    N=length(inst1);
    %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
    maxstep=length(inst1);
    xx=1:maxstep;
    yy=inst1;
    xx=xx*log(2);
    yy=log2(yy);

    % To estimate the fractal dimension, the slope of the line
    % "closest" (in the least square sense) to the points (x,y) is
calculated:
    A = zeros(maxstep, 2); % set up the matrix
    A(:, 1) = xx; % the first column
    A(:, 2) = ones(maxstep,1); % the second column
    [Q,R] = qr(A); % compute the QR-decomposition
    c = Q'*yy; % solution step 1
    param = R\c; % solution step 2
    frac_dim(1,j) = abs(log(inst1(1,1))-(param(2))); % the estimated
fractal dimension
end;
zci=frac_dim;
```

3. Changes in the power spectrum (ΔPS) of EEG signal computation function

```
function [ Z1 ] = power_spectrum(sig1)
%UNTITLED2 Summary of this function goes here
%   Detailed explanation goes here
% Fs is the frequency of the signal
% sig1 is the data samples of the EEG signal

      data=sig1;
      [m n]=size(data);

      %% Transform and visualise in frequency domain

      fdata=abs(fft(data,m));
      Mag = abs(fdata);              %Compute magnitude complex array
      PS = Mag.^2;                   % Take power sepctrum as
magnitude squared

      Z1=PS;
end
```

4. EEG coherence computation function

```
function [ coh ] = coherence_function( x,ch,Fs )
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%               x is the EEG data file                  %
%              ch is the channels e.g. [1 2 3 ...]      %
%               Fs is the sampling frequency            %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

l=length(ch);
c=1;
ch_pairs=[];
for i=1:l
    for j=(i+1):l
        if i~=j
            ch_pairs(c,1)=ch(i);
            ch_pairs(c,2)=ch(j);
            c=c+1;
        end
    end
end

for i=1:k

temp1=mscohere(x(:,ch_pairs(i,1)),x(:,ch_pairs(i,2)),kaiser(win_size)
,[],[],Fs);
    ch_pairs(i,3)=mean(temp1);
end
coh=ch_pairs;
end
```

5. Tsallis entropy (TsEn) computation function

```
function y=Tsallis_entro(x,q)

% x is EEG signal
% q is Tsallis factor

[M N]=size(x);
y=zeros(1,N);
for l=1:N
    sum1=sum(x(:,l)-(x(:,l).^q));
    sum2=sum1/(q-1);
    y(1,l)=sum2;
end;
```

6. Higuchi Fractal Dimension (HFD) computation function

```
function [ z1 ] = Higuchi_function(d )

% d is EEG signal

[sr yr]=size(d);
for r1=1:yr
    x=d(:,r1);
```

```
    [N,m] = size(x);
    width = max(size(x));    % largest size of the box
    p = log(width)/log(2);
    % remap the array if the sizes are not all equal,
    % or if they are not power of two
    % (this slows down the computation!)

    n=zeros(1,p); % pre-allocate the number of box of size r4
    for i=1:N/p:N
        i;
    end;

    k=3;
    in1=1;
    L=(1:k);
    L(:)=0;
    L1=(1:k);
    for m=1:k
        j=0;
        for i=m:k:(m+fix((N-m)/k)*k)
            j=j+1;
            x1(m,j)=x(i);
        end;
    end;

    [N1 m1]=size(x1);
    for j=2:m1
        t1=0;
        for i=1:N1
            t1=t1+abs(x1(i,j-1)-x1(i,j));
        end;
        x2(j-1)=(t1/k)/k;
    end;

    xx2(r1,:)=x2*(N-1)/(((N-m1)/k)*k)/k;

    xx3(r1,:)=xx2(r1,:); % This array contains the fractal Higuchi
distance
    xx4(r1,:)=log(xx3(r1,:)); % This array contains the LOG of
fractal Higuchi distance
end;
z1=xx4;
```

7. Approximation Entropy (ApEn) computation function

```
function apen = ApEn( dim, r, data)
%ApEn
%   dim : embedded dimension
%   r : tolerance (typically 0.2 * std)
%   data : time-series data

N = length(data);
result = zeros(1,2);

for j = 1:2
    m = dim+j-1;
    phi = zeros(1,N-m+1);
    dataMat = zeros(m,N-m+1);
```

```
    % setting up data matrix
    for i = 1:m
        dataMat(i,:) = data(i:N-m+i);
    end

    % counting similar patterns using distance calculation
    for i = 1:N-m+1
        tempMat = abs(dataMat - repmat(dataMat(:,i),1,N-m+1));
        boolMat = any( (tempMat > r),1);
        phi(i) = sum(~boolMat)/(N-m+1);
    end

    % summing over the counts
    result(j) = sum(log(phi))/(N-m+1);
end

apen = result(1)-result(2);

end
```

8. Lempel Ziv Complexity (LZC) computation function

```
function [ z ] = LZC(d)

% d is EEG signal

[sr yr]=size(d);
for r1=1:yr
    data1=d(:,r1);

    m1=median(data1);
    [x y]=size(data1);
    data2 = zeros(x,y);
    for i=1:x
        for j=1:y
            if data1(i,j)>=m1
                data2(i,j)=1;
            end
        end
    end
    [C, H] = calc_lz_complexity(data2, 'exhaustive', true);
    [x y]=size(H);
    mast = zeros(x,y);

    for i=1:y
        b=double(H{i});
        %b1=binary_seq_to_string(b)
        [x1 y1]=size(b);
        for i1=1:x1
            for j1=1:y1
                mast(i,j1)=b(i1,j1);
            end;
        end;
    end;
    mast_dec=bi2de(mast)';

    Cn(r1)=C;

end
z=Cn;
```

```
end

%CALC_LZ_COMPLEXITY Lempel-Ziv measure of binary sequence complexity.
%   This function calculates the complexity of a finite binary
sequence,
%   according to the algorithm published by Abraham Lempel and Jacob
Ziv in
%   the paper "On the Complexity of Finite Sequences", published in
%   "IEEE Transactions on Information Theory", Vol. IT-22, no. 1,
January
%   1976.  From that perspective, the algorithm could be referred to
as
%   "LZ76".
%
%   Usage: [C, H] = calc_lz_complexity(S, type, normalize)
%
%   INPUTS:
%
%   S:
%   A vector consisting of a binary sequence whose complexity is to
be
%   analyzed and calculated.  Numeric values will be converted to
logical
%   values depending on whether (0) or not (1) they are equal to 0.
%
%   type:
%   The type of complexity to evaluate as a string, which is one of:
%       - 'exhaustive': complexity measurement is based on
decomposing S
%       into an exhaustive production process.
%       - 'primitive': complexity measurement is based on decomposing
S
%       into a primitive production process.
%   Exhaustive complexity can be considered a lower limit of the
complexity
%   measurement approach proposed in LZ76, and primitive complexity
an
%   upper limit.
%
%   normalize:
%   A logical value (true or false), used to specify whether or not
the
%   complexity value returned is normalized or not.
%   Where normalization is applied, the normalized complexity is
%   calculated from the un-normalized complexity, C_raw, as:
%       C = C_raw / (n / log2(n))
%   where n is the length of the sequence S.
%
%   OUTPUTS:
%
%   C:
%   The Lempel-Ziv complexity value of the sequence S.
%
%   H:
%   A cell array consisting of the history components that were found
in
%   the sequence S, whilst calculating C.  Each element in H consists
of a
%   vector of logical values (true, false), and represents
%   a history component.
%
%   gs:
```

```
%   A vector containing the corresponding eigenfunction that was
calculated
%   which corresponds with S.
%
%
%
%   Author: Quang Thai (qlthai@gmail.com)
%   Copyright (C) Quang Thai 2012

function [C, H, gs] = calc_lz_complexity(S, type, normalize)

%% Some parameter-checking.

% Make sure S is a vector.
if ~isvector(S)
    error('''S'' must be a vector');
end

% Make sure 'normalize' is a scalar.
if ~isscalar(normalize)
    error('''normalize'' must be a scalar');
end

% Make sure 'type' is valid.
if ~(strcmpi(type, 'exhaustive') || strcmpi(type, 'primitive'))
    error(['''type'' parameter is not valid, must be either ' ...
        '''exhaustive'' or ''primitive''']);
end



%% Some parameter 'conditioning'.
S = logical(S);
normalize = logical(normalize);


%% ANALYSIS

% NOTE: Many of these comments will refer to the paper "On the
Complexity
% of Finite Sequences" by Lempel and Ziv, so to follow this code, it
may
% be useful to have the manuscript in front of you!


% Allocate memory for eigenfunction (vector of eigenvalues).
% The first value of this vector corresponds with gs(0), and is
always
% equal to 0.
% Please note that, since MATLAB array indices start at 1, gs(n) in
MATLAB
% actually holds gs(n-1) as defined in the paper.
n = length(S);
gs = zeros(1, n + 1);
gs(1) = 0;  % gs(0) = 0 from the paper



% The approach we will use to find the eigenfunction values at each
% successive prefix of S is as follows:
% - We wish to find gs(n), where 1 <= n <= l(S) (l(S) = length of S)
% - Lemma 4 says:
%       k(S(1,n-1)) <= k(S(1,n))
```

```
%           equivalently
%       gs(n-1) <= gs(n)
%   In other words, the eigenfunction is a non-decreasing function of
n.
% - Theorem 6 provides the expression that defines the
eigenvocabulary of
%   a sequence:
%       e(S(1,n)) = {S(i,n) | 1 <= i <= k(S(1,n))}
%           equivalently
%       e(S(1,n)) = {S(i,n) | 1 <= i <= gs(n)}
%   Note that we do not know what gs(n) is at this point - it's what
we're
%   trying to find!!!
% - Remember that the definition of the eigenvocabulary of a sequence
S(1,n),
%   e(S(1,n)), is the subset of the vocabulary of S(1,n) containing
words
%   that are not in the vocabulary of any proper prefix of S(1,n),
and the
%   eigenvalue of S(1,n) is the subset's cardinality: gs(n) =
|e(S(1,n))|
%   (p 76, 79)
% - Given this, a corollary to Theorem 6 is:
%       For each S(m,n) | gs(n) < m <= n, S(m,n) is NOT a member of
%       the eigenvocabulary e(S(1,n)).
%       By definition, this means that S(m,n) is in the vocabulary of
at
%       least one proper prefix of S(1,n).
% - Also note that from Lemma 1: if a word is in the vocabulary of a
%   sequence S, and S is a proper prefix of S+, then the word is also
%   in the vocabulary of S+.
%
% As a result of the above discussion, the algorithm can be expressed
in
% pseudocode as follows:
%
% For a given n, whose corresponding eigenfunction value, gs(n) we
wish to
% find:
% - gs(0) = 0
% - Let m be defined on the interval: gs(n-1) <= m <= n
% - for each m
%       check if S(m,n) is in the vocabulary of S(1,n-1)
%       if it isn't, then gs(n) = m
%       end if
%   end for
%
% An observation: searching linearly along the interval
% gs(n-1) <= m <= n will tend to favour either very complex sequences
% (starting from n and working down), or very un-complex sequences
% (starting from gs(n-1) and working up).  This implementation will
% attempt to balance these outcomes by alternately searching from
either
% end and working inward - a 'meet-in-the-middle' search.
%
% Note that:
% - When searching from the upper end downwards, we are seeking
% the value of m such that S(m,n) IS NOT in the vocabulary of S(1,n-
1).
% The eigenfunction value is then m.
% - When searching from the lower end upwards, we are seeking the
value
% of m such that S(m,n) IS in the vocabulary of S(1,n-1).  The
```

```
% eigenfunction value is then m-1, since it is the MAXIMAL value of m
% whereby S(m,n) IS NOT in the vocabulary of S(1,n-1)


%% Calculate eigenfunction, gs(n)

% Convert to string form - aids the searching process!
S_string = binary_seq_to_string(S);
gs(2) = 1;  % By definition.  Remember: gs(2) in MATLAB is actually
gs(1)
            % due to the first element of the gs array holding the
            % eigenvalue for n = 0.

for n = 2:length(S)

    eigenvalue_found = false;

    % The search space gs(n-1) <= m <= n.
    % Remember: gs(n) in MATLAB is actually gs(n-1).
    % Note that we start searching at (gs(n-1) + 1) at the lower end,
since
    % if it passes the lower-end search criterion, then we subtract 1
    % to get the eigenvalue.
    idx_list = (gs(n)+1):n;
    for k = 1:ceil(length(idx_list)/2);

        % Check value at upper end of interval
        m_upper = idx_list(end - k + 1);
        if ~numel(strfind(S_string(1:(n-1)), S_string(m_upper:n)))
            % We've found the eigenvalue!
            gs(n+1) = m_upper;    % Remember:
                                  % gs(n+1) in MATLAB is actually
gs(n)
            eigenvalue_found = true;
            break;
        end

        % Check value at lower end of interval.
        %
        % Note that the search at this end is slightly more
complicated,
        % in the sense that we have to find the first value of m
where the
        % substring is FOUND, and then subtract 1.  However, this is
        % complicated by the 'meet-in-the-middle' search adopted, as
        % described below...
        m_lower = idx_list(k);
        if numel(strfind(S_string(1:(n-1)), S_string(m_lower:n)))
            % We've found the eigenvalue!
            gs(n+1) = m_lower-1;    % Remember:
                                    % gs(n+1) in MATLAB is actually
gs(n)
            eigenvalue_found = true;
            break;
        elseif (m_upper == m_lower + 1)
            % If we've made it here, then we know that:
            % - The search for substring S(m,n) from the upper end
had a
            %   FOUND result
            % - The search for substring S(m,n) from the lower end
had a
            %   NOT FOUND result
```

```
            % - The value of m used in the upper end search is one
more
            %   than the value of m used in this lower end search
            %
            % However, when searching from the lower end, we need a
FOUND
            % result and then subtract 1 from the corresponding m.
            % The problem with this 'meet-in-the-middle' searching is
that
            % it's possible that the actual eigenfunction value
actually
            % does occur in the middle, such that the loop would
terminate
            % before the lower-end search can reach a FOUND result
and the
            % upper-end search can reach a NOT FOUND result.
            %
            % This branch detects precisely this condition, whereby
            % the two searches use adjacent values of m in the
middle,
            % the upper-end search has the FOUND result that the
lower-end
            % search normally requires, and the lower-end search has
the
            % NOT FOUND result that the upper-end search normally
requires.

            % We've found the eigenvalue!
            gs(n+1) = m_lower;      % Remember:
                                    % gs(n+1) in MATLAB is actually
gs(n)
            eigenvalue_found = true;
            break;
        end

    end

    if ~eigenvalue_found
        % Raise an error - something is not right!
        error('Internal error: could not find eigenvalue');
    end

end

%% Calculate the terminal points for the required production
sequence.

% Histories are composed by decomposing the sequence S into the
following
% sequence of words:
%       H(S) = S(1,h_1)S(h_1 + 1,h_2)S(h_2 + 1,h_3)...S(h_m-1 +
1,h_m)
% The indices {h_1, h_2, h_3, ..., h_m-1, h_m} that characterise a
history
% make up the set of 'terminals'.
%
% Alternatively, for consistency, we will specify the history as:
%       H(S) = ...
%           S(h_0 + 1,h_1)S(h_1 + 1,h_2)S(h_2 + 1,h_3)...S(h_m-1 +
1,h_m)
% Where, by definition, h_0 = 0.
```

```
% Efficiency measure: we don't know how long the histories will be
(and
% hence, how many terminals we need).  As a result, we will allocate
an
% array of length equal to the eigenfunction vector length.  We will
also
% keep a 'length' counter, so that we know how much of this array we
are
% actually using.  This avoids us using an array that needs to be
resized
% iteratively!
% Note that h_i(1) in MATLAB holds h_0, h_i(2) holds h_1, etc., since
% MATLAB array indices must start at 1.
h_i = zeros(1, length(gs));
h_i_length = 1;     % Since h_0 is already present as the first
value!

if strcmpi(type, 'exhaustive')

    % - From Theorem 8, for an exhaustive history, the terminal
points h_i,
    % 1 <= i <= m-1, are defined by:
    %       h_i = min{h | gs(h) > h_m-1}
    % - We know that h_0 = 0, so this definition basically bootstraps
our
    % search process, allowing us to find h_1, then h_2, etc.

    h_prev = 0;     % Points to h_0 initially
    k = 1;
    while ~isempty(k)
        % Remember that gs(1) in MATLAB holds the value of gs(0).
        % Therefore, the index h_prev needs to be incremented by 1
        % to be used as an index into the gs vector.
        k = find(gs((h_prev+1+1):end) > h_prev, 1);

        if ~isempty(k)
            h_i_length = h_i_length + 1;

            % Remember that gs(1) in MATLAB holds the value of gs(0).
            % Therefore, the index h_prev needs to be decremented by
1
            % to be used as an index into the original sequence S.
            h_prev = h_prev + k;
            h_i(h_i_length) = h_prev;
        end
    end

    % Once we break out of the above loop, we've found all of the
    % exhaustive production components.
else

    % Sequence type is 'primitive'

    % Find all unique eigenfunction values, where they FIRST occur.

    % - From Theorem 8, for a primitive history, the terminal points
h_i,
    % 1 <= i <= m-1, are defined by:
    %       h_i = min{h | gs(h) > gs(h_i-1)}
    % - From Lemma 4, we know that the eigenfunction, gs(n), is
```

```
    % monotonically non-decreasing.
    % - Therefore, the following call to unique() locates the first
    % occurrance of each unique eigenfunction value, as well as the
values
    % of n where the eigenfunction increases from the previous value.
    % Hence, this is also an indicator for the terminal points h_i.

    [~, n] = unique(gs, 'first');

    % The terminals h_i, 1 <= i <= m-1, is ultimately obtained from n
by
    % subtracting 1 from each value (since gs(1) in MATLAB actually
    % corresponds with gs(0) in the paper)
    h_i_length = length(n);
    h_i(1:h_i_length) = n - 1;
end

% Now we have to deal with the final production component - which may
or
% may not be exhaustive or primitive, but can still be a part of an
% exhaustive or primitive process.
%
% If the last component is not exhaustive or primitive, we add it
here
% explicitly.
%
% - From Theorem 8, for a primitive history, this simply enforces
% the requirement that:
%       h_m = l(S)
if h_i(h_i_length) ~= length(S)
    h_i_length = h_i_length + 1;
    h_i(h_i_length) = length(S);
end

% Some final sanity checks - as indicated by Theorem 8.
% Raise an error if these checks fail!
% Also remember that gs(1) in the MATLAB code corresponds with gs(0).
if strcmpi(type, 'exhaustive')
    % Theorem 8 - check that gs(h_m - 1) <= h_m-1
    if gs(h_i(h_i_length) - 1 + 1) > h_i(h_i_length-1)
        error(['Check failed for exhaustive sequence: ' ...
            'Require: gs(h_m - 1) <= h_m-1']);
    end
else
    % Sequence type is 'primitive'

    % Theorem 8 - check that gs(h_m - 1) = gs(h_m-1)
    if gs(h_i(h_i_length) - 1 + 1) ~= gs(h_i(h_i_length-1) + 1)
        error(['Check failed for primitive sequence: ' ...
            'Require: gs(h_m - 1) = gs(h_m-1)']);
    end
end


%% Use the terminal points to construct the production sequence.

% Note the first value in h_i is h_0, so its length is one more than
the
% length of the production history.
H = cell([1 (h_i_length-1)]);
for k = 1:(h_i_length-1)
    H{k} = S((h_i(k)+1):h_i(k+1));
```

```
end

%% Hence calculate the complexity.
if normalize
    % Normalized complexity
    C = length(H) / (n / log2(n));
else
    % Un-normalized complexity
    C = length(H);
end

%% Eigenfunction is returned.
% The (redundant) first value (gs(0) = 0) is removed first.
gs = gs(2:end);
```

Appendix 3: MATLAB machine learning functions

1. The following function will be generated to import the EEG data into MATLAB toolbox (classification learner)

```
function tableout =
importfile(workbookFile,sheetName,startRow,endRow)
%IMPORTFILE Import data from a spreadsheet
%   DATA = IMPORTFILE(FILE) reads data from the first worksheet in
the
%   Microsoft Excel spreadsheet file named FILE and returns the data
as a
%   table.
%
%   DATA = IMPORTFILE(FILE,SHEET) reads from the specified worksheet.
%
%   DATA = IMPORTFILE(FILE,SHEET,STARTROW,ENDROW) reads from the
specified
%   worksheet for the specified row interval(s). Specify STARTROW and
%   ENDROW as a pair of scalars or vectors of matching size for
%   dis-contiguous row intervals. To read to the end of the file
specify an
%   ENDROW of inf.%
% Example:
%   data = importfile('data.xlsx','Sheet1',2,16);
%
%   See also XLSREAD.


%% Input handling

% If no sheet is specified, read first sheet
if nargin == 1 || isempty(sheetName)
    sheetName = 1;
end

% If row start and end points are not specified, define defaults
if nargin <= 3
    startRow = 2;
    endRow = 16;
end

%% Import the data
[~, ~, raw] = xlsread(workbookFile, sheetName,
sprintf('A%d:G%d',startRow(1),endRow(1)));
for block=2:length(startRow)
    [~, ~, tmpRawBlock] = xlsread(workbookFile, sheetName,
sprintf('A%d:G%d',startRow(block),endRow(block)));
    raw = [raw;tmpRawBlock]; %#ok<AGROW>
end
stringVectors = string(raw(:,7));
stringVectors(ismissing(stringVectors)) = '';
raw = raw(:,[1,2,3,4,5,6]);

%% Create output variable
I = cellfun(@(x) ischar(x), raw);
raw(I) = {NaN};
```

```
data = reshape([raw{:}],size(raw));

%% Create table
tableout = table;

%% Allocate imported array to column variable names
tableout.VarName1 = data(:,1);
tableout.VarName2 = data(:,2);
tableout.VarName3 = data(:,3);
tableout.VarName4 = data(:,4);
tableout.VarName5 = data(:,5);
tableout.VarName6 = data(:,6);
tableout.Stat = categorical(stringVectors(:,1));
```

2. Support Vector Machine (SVM) function

```
function [trainedClassifier, validationAccuracy] =
trainClassifier_SVM(trainingData)
% [trainedClassifier, validationAccuracy] =
trainClassifier_SVM(trainingData)
% returns a trained classifier and its accuracy. This code recreates
the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data,
or to
% learn how to programmatically train models.
%
%  Input:
%      trainingData: a table containing the same predictor and
response
%       columns as imported into the app.
%
%  Output:
%      trainedClassifier: a struct containing the trained classifier.
The
%       struct contains various fields with information about the
trained
%       classifier.
%
%      trainedClassifier.predictFcn: a function to make predictions
on new
%       data.
%
%      validationAccuracy: a double containing the accuracy in
percent. In
%       the app, the History list displays this overall accuracy
score for
%       each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your
original
% data or new data as the input argument trainingData.
%
% For example, to retrain a classifier trained with the original data
set
% T, enter:
%   [trainedClassifier, validationAccuracy] = trainClassifier(T)
```

```
%
% To make predictions with the returned 'trainedClassifier' on new
data T2,
% use
%   yfit = trainedClassifier.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns
as used
% during training. For details, enter:
%   trainedClassifier.HowToPredict


% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.Stat;
isCategoricalPredictor = [false, false, false, false, false, false];

% Train a classifier
% This code specifies all the classifier options and trains the
classifier.
template = templateSVM(...
    'KernelFunction', 'linear', ...
    'PolynomialOrder', [], ...
    'KernelScale', 'auto', ...
    'BoxConstraint', 1, ...
    'Standardize', true);
classificationSVM = fitcecoc(...
    predictors, ...
    response, ...
    'Learners', template, ...
    'Coding', 'onevsone', ...
    'ClassNames', categorical({'AD'; 'Norm'; 'Stat'}));

% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
svmPredictFcn = @(x) predict(classificationSVM, x);
trainedClassifier.predictFcn = @(x)
svmPredictFcn(predictorExtractionFcn(x));

% Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'VarName1', 'VarName2',
'VarName3', 'VarName4', 'VarName5', 'VarName6'};
trainedClassifier.ClassificationSVM = classificationSVM;
trainedClassifier.About = 'This struct is a trained model exported
from Classification Learner R2017b.';
trainedClassifier.HowToPredict = sprintf('To make predictions on a
new table, T, use: \n  yfit = c.predictFcn(T) \nreplacing ''c'' with
the name of the variable that is this struct, e.g. ''trainedModel''.
\n \nThe table, T, must contain the variables returned by: \n
c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype)
must match the original training data. \nAdditional variables are
ignored. \n \nFor more information, see <a
href="matlab:helpview(fullfile(docroot, ''stats'', ''stats.map''),
''appclassification_exportmodeltoworkspace'')">How to predict using
an exported model</a>.');

% Extract predictors and response
% This code processes the data into the right shape for training the
```

```
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.Stat;
isCategoricalPredictor = [false, false, false, false, false, false];

% Perform cross-validation
partitionedModel = crossval(trainedClassifier.ClassificationSVM,
'KFold', 10);

% Compute validation predictions
[validationPredictions, validationScores] =
kfoldPredict(partitionedModel);

% Compute validation accuracy
validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun',
'ClassifError');
```

3. Linear Discriminant Analysis (LDA) function

```
function [trainedClassifier, validationAccuracy] =
trainClassifier_LDA(trainingData)
% [trainedClassifier, validationAccuracy] =
trainClassifier(trainingData)
% returns a trained classifier and its accuracy. This code recreates
the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data,
or to
% learn how to programmatically train models.
%
%  Input:
%      trainingData: a table containing the same predictor and
response
%       columns as imported into the app.
%
%  Output:
%      trainedClassifier: a struct containing the trained classifier.
The
%       struct contains various fields with information about the
trained
%       classifier.
%
%      trainedClassifier.predictFcn: a function to make predictions
on new
%       data.
%
%      validationAccuracy: a double containing the accuracy in
percent. In
%       the app, the History list displays this overall accuracy
score for
%       each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your
original
% data or new data as the input argument trainingData.
%
```

```
% For example, to retrain a classifier trained with the original data
set
% T, enter:
%   [trainedClassifier, validationAccuracy] = trainClassifier(T)
%
% To make predictions with the returned 'trainedClassifier' on new
data T2,
% use
%   yfit = trainedClassifier.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns
as used
% during training. For details, enter:
%   trainedClassifier.HowToPredict

% Auto-generated by MATLAB on 27-Jul-2019 12:32:25


% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.AD;
isCategoricalPredictor = [false, false, false, false, false, false];

% Train a classifier
% This code specifies all the classifier options and trains the
classifier.
classificationDiscriminant = fitcdiscr(...
    predictors, ...
    response, ...
    'DiscrimType', 'linear', ...
    'Gamma', 0, ...
    'FillCoeffs', 'off', ...
    'ClassNames', categorical({'AD'; 'Norm'}));

% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
discriminantPredictFcn = @(x) predict(classificationDiscriminant, x);
trainedClassifier.predictFcn = @(x)
discriminantPredictFcn(predictorExtractionFcn(x));

% Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'VarName1', 'VarName2',
'VarName3', 'VarName4', 'VarName5', 'VarName6'};
trainedClassifier.ClassificationDiscriminant =
classificationDiscriminant;
trainedClassifier.About = 'This struct is a trained model exported
from Classification Learner R2017b.';
trainedClassifier.HowToPredict = sprintf('To make predictions on a
new table, T, use: \n  yfit = c.predictFcn(T) \nreplacing ''c'' with
the name of the variable that is this struct, e.g. ''trainedModel''.
\n \nThe table, T, must contain the variables returned by: \n
c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype)
must match the original training data. \nAdditional variables are
ignored. \n \nFor more information, see <a
href="matlab:helpview(fullfile(docroot, ''stats'', ''stats.map''),
''appclassification_exportmodeltoworkspace'')">How to predict using
an exported model</a>.');
```

```
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.AD;
isCategoricalPredictor = [false, false, false, false, false, false];

% Perform cross-validation
partitionedModel =
crossval(trainedClassifier.ClassificationDiscriminant, 'KFold', 10);

% Compute validation predictions
[validationPredictions, validationScores] =
kfoldPredict(partitionedModel);

% Compute validation accuracy
validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun',
'ClassifError');
```

4. K-nearest neighbour (KNN) function

```
function [trainedClassifier, validationAccuracy] =
trainClassifier_KNN(trainingData)
% [trainedClassifier, validationAccuracy] =
trainClassifier(trainingData)
% returns a trained classifier and its accuracy. This code recreates
the
% classification model trained in Classification Learner app. Use the
% generated code to automate training the same model with new data,
or to
% learn how to programmatically train models.
%
%  Input:
%      trainingData: a table containing the same predictor and
response
%       columns as imported into the app.
%
%  Output:
%      trainedClassifier: a struct containing the trained classifier.
The
%       struct contains various fields with information about the
trained
%       classifier.
%
%      trainedClassifier.predictFcn: a function to make predictions
on new
%       data.
%
%      validationAccuracy: a double containing the accuracy in
percent. In
%       the app, the History list displays this overall accuracy
score for
%       each model.
%
% Use the code to train the model with new data. To retrain your
% classifier, call the function from the command line with your
original
% data or new data as the input argument trainingData.
%
```

```
% For example, to retrain a classifier trained with the original data
set
% T, enter:
%   [trainedClassifier, validationAccuracy] = trainClassifier(T)
%
% To make predictions with the returned 'trainedClassifier' on new
data T2,
% use
%   yfit = trainedClassifier.predictFcn(T2)
%
% T2 must be a table containing at least the same predictor columns
as used
% during training. For details, enter:
%   trainedClassifier.HowToPredict


% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.AD;
isCategoricalPredictor = [false, false, false, false, false, false];

% Train a classifier
% This code specifies all the classifier options and trains the
classifier.
classificationKNN = fitcknn(...
    predictors, ...
    response, ...
    'Distance', 'Euclidean', ...
    'Exponent', [], ...
    'NumNeighbors', 1, ...
    'DistanceWeight', 'Equal', ...
    'Standardize', true, ...
    'ClassNames', categorical({'AD'; 'Norm'}));

% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
knnPredictFcn = @(x) predict(classificationKNN, x);
trainedClassifier.predictFcn = @(x)
knnPredictFcn(predictorExtractionFcn(x));

% Add additional fields to the result struct
trainedClassifier.RequiredVariables = {'VarName1', 'VarName2',
'VarName3', 'VarName4', 'VarName5', 'VarName6'};
trainedClassifier.ClassificationKNN = classificationKNN;
trainedClassifier.About = 'This struct is a trained model exported
from Classification Learner R2017b.';
trainedClassifier.HowToPredict = sprintf('To make predictions on a
new table, T, use: \n  yfit = c.predictFcn(T) \nreplacing ''c'' with
the name of the variable that is this struct, e.g. ''trainedModel''.
\n \nThe table, T, must contain the variables returned by: \n
c.RequiredVariables \nVariable formats (e.g. matrix/vector, datatype)
must match the original training data. \nAdditional variables are
ignored. \n \nFor more information, see <a
href="matlab:helpview(fullfile(docroot, ''stats'', ''stats.map''),
''appclassification_exportmodeltoworkspace'')">How to predict using
an exported model</a>.');

% Extract predictors and response
```

```
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'VarName1', 'VarName2', 'VarName3', 'VarName4',
'VarName5', 'VarName6'};
predictors = inputTable(:, predictorNames);
response = inputTable.AD;
isCategoricalPredictor = [false, false, false, false, false, false];

% Perform cross-validation
partitionedModel = crossval(trainedClassifier.ClassificationKNN,
'KFold', 10);

% Compute validation predictions
[validationPredictions, validationScores] =
kfoldPredict(partitionedModel);

% Compute validation accuracy
validationAccuracy = 1 - kfoldLoss(partitionedModel, 'LossFun',
'ClassifError');
```

Appendix 4: Summary of the biomarkers combination from length 1-4

The following table shows a summary of the combinations EEG biomarkers from 1 – 4. The actual table consists of 919310 combinations.

No	Biomarker 1	Biomarker 2	Biomarker 3	Biomarker 4
1	ApEn(Alpha(O1))			
2	Coh(Alpha/theta(F4-T4))			
3	HFD(Alpha(O1))			
4	LZC(Alpha(C3))			
5	TsEn(Alpha/delta(C3))			
6	ΔEEGA(Alpha/delta(T6))			
7	ΔPS(Alpha(P4))			
8	ApEn(Alpha(O1))	Coh(Alpha/theta(F4-T4))		
9	HFD(Alpha(O1))	Coh(Alpha/theta(F4-T4))		
10	LZC(Alpha(C3))	Coh(Alpha/theta(F4-T4))		
11	TsEn(Alpha/delta(C3))	Coh(Alpha/theta(F4-T4))		
12	ΔEEGA(Alpha/delta(T6))	Coh(Alpha/theta(F4-T4))		
13	ΔPS(Alpha(P4))	Coh(Alpha/theta(F4-T4))		
14	ApEn(Alpha(O1))	HFD(Alpha(O2))		
15	HFD(Alpha(O1))	HFD(Alpha(O2))		
16	ApEn(Alpha(O1))	LZC(Alpha(CZ))		
17	HFD(Alpha(O1))	LZC(Alpha(CZ))		
18	LZC(Alpha(C3))	LZC(Alpha(CZ))		
19	ApEn(Alpha(O1))	TsEn(Alpha/theta(P3))		
20	HFD(Alpha(O1))	TsEn(Alpha/theta(P3))		
21	LZC(Alpha(C3))	TsEn(Alpha/theta(P3))		
22	TsEn(Alpha/delta(C3))	TsEn(Alpha/theta(P3))		
23	ApEn(Alpha(O1))	ΔEEGA(Theta/alpha(T5))		
24	HFD(Alpha(O1))	ΔEEGA(Theta/alpha(T5))		
25	LZC(Alpha(C3))	ΔEEGA(Theta/alpha(T5))		
26	TsEn(Alpha/delta(C3))	ΔEEGA(Theta/alpha(T5))		
27	ΔEEGA(Alpha/delta(T6))	ΔEEGA(Theta/alpha(T5))		
28	ApEn(Alpha(O1))	ΔPS(Alpha/theta(T3))		
29	HFD(Alpha(O1))	ΔPS(Alpha/theta(T3))		
30	LZC(Alpha(C3))	ΔPS(Alpha/theta(T3))		
31	TsEn(Alpha/delta(C3))	ΔPS(Alpha/theta(T3))		
32	ΔEEGA(Alpha/delta(T6))	ΔPS(Alpha/theta(T3))		
33	ΔPS(Alpha(P4))	ΔPS(Alpha/theta(T3))		
34	ApEn(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	
35	HFD(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	
36	ApEn(Alpha(O1))	HFD(Alpha(O2))	LZC(Alpha(T6))	
37	HFD(Alpha(O1))	HFD(Alpha(O2))	LZC(Alpha(T6))	
38	ApEn(Alpha(O1))	LZC(Alpha(CZ))	LZC(Alpha(T6))	
39	HFD(Alpha(O1))	LZC(Alpha(CZ))	LZC(Alpha(T6))	

No	Biomarker 1	Biomarker 2	Biomarker 3	Biomarker 4
40	LZC(Alpha(C3))	LZC(Alpha(CZ))	LZC(Alpha(T6))	
41	ApEn(Alpha(O1))	HFD(Alpha(O2))	ΔEEGA(Theta/alpha(T5))	
42	HFD(Alpha(O1))	HFD(Alpha(O2))	ΔEEGA(Theta/alpha(T5))	
43	ApEn(Alpha(O1))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	
44	HFD(Alpha(O1))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	
45	LZC(Alpha(C3))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	
46	ApEn(Alpha(O1))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	
47	HFD(Alpha(O1))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	
48	LZC(Alpha(C3))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	
49	TsEn(Alpha/delta(C3))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	
50	ApEn(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	LZC(Alpha(C3))
51	HFD(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	LZC(Alpha(C3))
52	ApEn(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	LZC(Alpha(T6))
53	HFD(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	LZC(Alpha(T6))
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
919295	ApEn(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	ΔPS(Alpha/theta(T3))
919296	HFD(Alpha(O1))	HFD(Alpha(O2))	HFD(Theta/alpha(T3))	ΔPS(Alpha/theta(T3))
919297	ApEn(Alpha(O1))	HFD(Alpha(O2))	LZC(Alpha(T6))	ΔPS(Alpha/theta(T3))
919298	HFD(Alpha(O1))	HFD(Alpha(O2))	LZC(Alpha(T6))	ΔPS(Alpha/theta(T3))
919299	ApEn(Alpha(O1))	LZC(Alpha(CZ))	LZC(Alpha(T6))	ΔPS(Alpha/theta(T3))
919300	HFD(Alpha(O1))	LZC(Alpha(CZ))	LZC(Alpha(T6))	ΔPS(Alpha/theta(T3))
919301	LZC(Alpha(C3))	LZC(Alpha(CZ))	LZC(Alpha(T6))	ΔPS(Alpha/theta(T3))
919302	ApEn(Alpha(O1))	HFD(Alpha(O2))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919303	HFD(Alpha(O1))	HFD(Alpha(O2))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919304	ApEn(Alpha(O1))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
901905	HFD(Alpha(O1))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919306	LZC(Alpha(C3))	LZC(Alpha(CZ))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919307	ApEn(Alpha(O1))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919308	HFD(Alpha(O1))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919309	LZC(Alpha(C3))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))
919310	TsEn(Alpha/delta(C3))	TsEn(Alpha/theta(P3))	ΔEEGA(Theta/alpha(T5))	ΔPS(Alpha/theta(T3))

References

[1] T. Harmony, T. Fernández, J. Gersenowies, L. Galán, A. Fernández-Bouzas, E. Aubert, L. Díaz Comas, Specific EEG frequencies signal general common cognitive processes as well as specific task processes in man, Int. J. Psychophysiol. 53 (2004) 207–216.

[2] C. Humpel, Identifying and validating biomarkers for Alzheimer's disease, Trends Biotechnol. 29 (2011) 26–32.

[3] A. Alzheimer's, 2015 Alzheimer's disease facts and figures., Alzheimer's Dement. J. Alzheimer's Assoc. 11 (2015) 332.

[4] M. Weiner, Z. Khachaturian, The use of MRI and PET for clinical diagnosis of dementia and investigation of cognitive impairment: a consensus report, Alzheimer's Assoc. Chicago,. (2005).

[5] J.K. Silver, S. Parangi, R. Phitayakorn, Biographies of Disease, Ser. Ed. (2011).

[6] W.H. Organization, WHO Mental Health Gap Action Programme (mhGAP) 2008, (2014).

[7] ADNI, Alzheimer's Disease Neuroimaging Initiative, (n.d.). http://www.adni-info.org/Home.html.

[8] A.H.A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Changes in the EEG amplitude as a biomarker for early detection of Alzheimer's disease, in: Eng. Med. Biol. Soc. (EMBC), 2016 IEEE 38th Annu. Int. Conf., IEEE, 2016: pp. 993–996. doi:10.1109/EMBC.2016.7590869.

[9] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, M. Karagiannidou, World Alzheimer Report 2016 Improving healthcare for people living with dementia. Coverage, Quality and costs now and in the future, (2016) 1–140. https://www.alz.co.uk/research/world-report-2016.

[10] G. Livingston, A. Sommerlad, V. Orgeta, S.G. Costafreda, J. Huntley, D. Ames, C. Ballard, S. Banerjee, A. Burns, J. Cohen-Mansfield, Dementia prevention, intervention, and care, Lancet. 390 (2017) 2673–2734.

[11] M. Prince, Anders Wimo, Maëlenn Guerchet, G.-C. Ali, Dr Yu-Tzu Wu, Matthew Prina, World Alzheimer Report, 2015. https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf.

[12] E. Jammeh, C. Carroll, S. Pearson, J. Escudero, A. Anastasiou, J. Zajicek, E. Ifeachor, Using NHS primary care data to identify undiagnosed dementia, J. Neurol. Neurosurg. Psychiatry. 86 (2015) e4–e4.

[13] B. Michalowsky, T. Eichler, J.R. Thyrian, J. Hertel, D. Wucherer, W. Hoffmann, S. Flessa, Healthcare resource utilization and cost in dementia: are there differences between patients screened positive for dementia with and those without a formal diagnosis of dementia in primary care in Germany?, Int. Psychogeriatrics. 28 (2016) 359–369.

[14] M. Prince, R. Bryce, C. Ferri, World Alzheimer Report 2011: The benefits of early diagnosis and intervention, Alzheimer's Disease International, 2011.

[15] G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram, S. Wimalaratna, C. Del Percio, F. Vecchio, Development and assessment of methods for detecting dementia using the human electroencephalogram, IEEE Trans. Biomed. Eng. 53 (2006) 1557–1568.

[16] A.L. Sutton., ed., Alzheimer Disease Sourcebook, Fifth edit, Peter E. Ruffner, Detroit: Omnigraphics, 2011.

[17] D. Galimberti, E. Scarpani, eds., Biomarkers for Early Diagnosis of Alzheimer's Disease, Galimberti, Nova Biomedical Books, 2008.

[18] A. Association, 2016 Alzheimer's disease facts and figures, Alzheimer's

Dement. 12 (2017) 459–509.
[19] K.A. Jellinger, B. Janetzky, J. Attems, E. Kienzl, Biomarkers for early diagnosis of Alzheimer disease:'ALZheimer ASsociated gene'–a new blood biomarker?, J. Cell. Mol. Med. 12 (2008) 1094–1117.
[20] K. Ritchie, S. Lovestone, The dementias, Lancet. 360 (2002) 1759–1766.
[21] R.A. Sperling, P.S. Aisen, L.A. Beckett, D.A. Bennett, S. Craft, A.M. Fagan, T. Iwatsubo, C.R. Jack, J. Kaye, T.J. Montine, Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, Alzheimer's Dement. 7 (2011) 280–292.
[22] M.D. Greicius, B. Krasnow, A.L. Reiss, V. Menon, Functional connectivity in the resting brain: a network analysis of the default mode hypothesis, Proc. Natl. Acad. Sci. 100 (2003) 253–258.
[23] A. Gevins, M.E. Smith, L. McEvoy, D. Yu, High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice., Cereb. Cortex. 7 (1997) 374–385.
[24] K. Sasaki, A. Nambu, T. Tsujimoto, R. Matsuzaki, S. Kyuhou, H. Gemba, Studies on integrative functions of the human frontal association cortex with MEG, Cogn. Brain Res. 5 (1996) 165–174.
[25] B.T. Hyman, J.H. Growdon, M.W. Albers, R.L. Buckner, J. Chhatwal, M.T. Gomez-Isla, C. Haass, E. Hudry, C.R. Jack, K.A. Johnson, Massachusetts Alzheimer's disease research center: progress and challenges, Alzheimer's Dement. J. Alzheimer's Assoc. 11 (2015) 1241–1245.
[26] B. Winblad, P. Amouyel, S. Andrieu, C. Ballard, C. Brayne, H. Brodaty, A. Cedazo-Minguez, B. Dubois, D. Edvardsson, H. Feldman, Defeating Alzheimer's disease and other dementias: a priority for European science and society, Lancet Neurol. 15 (2016) 455–532.
[27] N. Coley, A. Gallini, S. Andrieu, Prevention Studies in Alzheimer's Disease: Progress Towards the Development of New Therapeutics, CNS Drugs. 29 (2015) 519–528.
[28] J.D. Grill, J. Karlawish, Study partners should be required in preclinical Alzheimer's disease trials, Alzheimers. Res. Ther. 9 (2017) 93.
[29] V. Garibotto, D. Perani, The Role of Neuroimaging in The Early Diagnosis of Alzheimer's Disease, BioMarkers Early Diagnosis Alzheimer's Dis. (2008) 29–66.
[30] R. Squitti, G. Dal Forno, S. Cesaretti, M. Ventriglia, P.M. Rossini, Copper Studies in Alzheimer's Disease, Top. Alzheimer's Dis. (2006) 129.
[31] I. Guidi, D. Galimberti, Mild Cognitive Impairment, BioMarkers Early Diagnosis Alzheimer's Dis. (2006) 1–9.
[32] R.C. Petersen, G.E. Smith, S.C. Waring, R.J. Ivnik, E. Kokmen, E.G. Tangelos, Aging, memory, and mild cognitive impairment, Int. Psychogeriatrics. 9 (1997) 65–69.
[33] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, Y. Saeys, Robust biomarker identification for cancer diagnosis with ensemble feature selection methods, Bioinformatics. 26 (2010) 392–398.
[34] A. Boucheham, M. Batouche, Robust biomarker discovery for cancer diagnosis based on meta-ensemble feature selection, in: Sci. Inf. Conf. (SAI), 2014, IEEE, 2014: pp. 452–560.
[35] A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Changes in the Electroencephalogram as a Biomarker of Alzheimer ' s Disease, in: Biosense Dement. 2017 - Int. Work. Biosens. Dement. from 13 – 14 June 2017 – Plymouth Univ. Plymouth, UK, International workshop on biosensors for dementia , 13 - 14 June 2017, 2017: p. 40. https://www.plymouth.ac.uk/whats-on/biosense-

dementia.
[36] A. Cedazo-Minguez, B. Winblad, Biomarkers for Alzheimer's disease and other forms of dementia: clinical needs, limitations and future aspects, Exp. Gerontol. 45 (2010) 5–14.
[37] D. Ferreira, V. Jelic, L. Cavallin, A.-R. Oeksengaard, J. Snaedal, P. Høgh, B.B. Andersen, M. Naik, K. Engedal, E. Westman, Electroencephalography is a good complement to currently established dementia biomarkers, Dement. Geriatr. Cogn. Disord. 42 (2016) 80–92.
[38] K. Blennow, H. Hampel, CSF markers for incipient Alzheimer's disease, Lancet Neurol. 2 (2003) 605–613.
[39] S.E. O'Bryant, M.M. Mielke, R.A. Rissman, S. Lista, H. Vanderstichele, H. Zetterberg, P. Lewczuk, H. Posner, J. Hall, L. Johnson, Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic, Alzheimer's Dement. 13 (2017) 45–58.
[40] D. Ferreira, L. Perestelo-Pérez, E. Westman, L.-O. Wahlund, A. Sarría, P. Serrano-Aguilar, Meta-review of CSF core biomarkers in Alzheimer's disease: the state-of-the-art after the new revised diagnostic criteria, Front. Aging Neurosci. 6 (2014) 47.
[41] B. Dubois, H. Hampel, H.H. Feldman, P. Scheltens, P. Aisen, S. Andrieu, H. Bakardjian, H. Benali, L. Bertram, K. Blennow, Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria, Alzheimer's Dement. 12 (2016) 292–323.
[42] H. Jiang, M.P. White, M.D. Greicius, L.C. Waelde, D. Spiegel, Brain activity and functional connectivity associated with hypnosis, Cereb. Cortex. (2016).
[43] M. Signorino, E. Pucci, N. Belardinelli, G. Nolfe, F. Angeleri, EEG spectral analysis in vascular and Alzheimer dementia, Electroencephalogr. Clin. Neurophysiol. 94 (1995) 313–325.
[44] C. Besthorn, R. Zerfass, C. Geiger-Kabisch, H. Sattel, S. Daniel, U. Schreiter-Gasser, H. Förstl, Discrimination of Alzheimer's disease and normal aging by EEG data, Electroencephalogr. Clin. Neurophysiol. 103 (1997) 241–248.
[45] A.H.H.A.H. Al-Nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Tsallis entropy as a biomarker for detection of Alzheimer's disease, in: Eng. Med. Biol. Soc. (EMBC), 2015 37th Annu. Int. Conf. IEEE, IEEE, 2015: pp. 4166–4169. doi:10.1109/EMBC.2015.7319312.
[46] A. Horvath, A. Szucs, G. Csukly, A. Sakovics, G. Stefanics, A. Kamondi, EEG and ERP biomarkers of Alzheimer's disease: a critical review, (n.d.).
[47] D. V Moretti, C. Fracassi, M. Pievani, C. Geroldi, G. Binetti, O. Zanetti, K. Sosta, P.M. Rossini, G.B. Frisoni, Increase of theta/gamma ratio is associated with memory impairment, Clin. Neurophysiol. 120 (2009) 295–303.
[48] J. Jeong, EEG dynamics in patients with Alzheimer's disease, Clin. Neurophysiol. 115 (2004) 1490–1505.
[49] A.H.A.H.H. Al-nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Higuchi Fractal Dimension of the Electroencephalogram as a Biomarker for Early Detection of Alzheimer's Disease, in: Eng. Med. Biol. Soc. (EMBC), 2017 IEEE 39th Annu. Int. Conf. IEEE, IEEE, 2017: pp. 2320–2324. doi:10.1109/EMBC.2017.8037320.
[50] A.H.H. Al-Nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Complexity Measures for Quantifying Changes in Electroencephalogram in Alzheimer's Disease, Complexity. 2018 (2018) 1–12. doi:10.1155/2018/8915079.
[51] T.K. Khan, D.L. Alkon, Peripheral biomarkers of Alzheimer's disease, J. Alzheimer's Dis. 44 (2015) 729–744.
[52] W.S.T. Griffin, Inflammation and neurodegenerative diseases-, Am. J. Clin.

Nutr. 83 (2006) 470S-474S.
[53] E. Venturelli, C. Villa, E. Scarpini, Cerebrospinal Fluid Biomarkers for Alzheimer's Disease, BioMarkers Early Diagnosis Alzheimer's Dis. (2008) 67–79.
[54] S.L. Risacher, A.J. Saykin, Neuroimaging and other biomarkers for Alzheimer's disease: the changing landscape of early detection, Annu. Rev. Clin. Psychol. 9 (2013) 621–648.
[55] C. Mulder, P. Scheltens, J.J. Visser, G.J. van Kamp, R.B.H. Schutgens, Genetic and biochemical markers for Alzheimer's disease: recent developments, Ann. Clin. Biochem. 37 (2000) 593–607.
[56] S. Neurol, Neuroimaging biomarkers of neurodegenerative diseases and dementia, Semin Neurol. 33 (2013) 386–416.
[57] B. Delatour, C. Le Cudennec, N. El Tannir-El Tayara, M. Dhenain, Transgenic models of Alzheimer's pathology: Success and caveat, Prog. Alzheimer's Dis. Res. (2006).
[58] N.M. Hooper, Alzheimer's disease: methods and protocols, Springer Science & Business Media, 2000.
[59] D.M. Holtzman, J.C. Morris, A.M. Goate, Alzheimer's disease: the challenge of the second century, Sci. Transl. Med. 3 (2011) 77sr1-77sr1.
[60] B. Thies, E. Truschke, M. Morrison-Bogorad, R.J. Hodes, Consensus Report of the Working Group on: "Molecular and Biochemical Markers of Alzheimer's Disease," Neurobiol. Aging. 19 (1998) 109–116. doi:10.1016/S0197-4580(98)00022-0.
[61] M.D. Richard J. Hodes, National Institute on Aging, National Institutes of Health (NIH), (n.d.).
[62] Alzheimer's Association, Alzheimer's and Dementia Resources, (n.d.).
[63] M.M. González, Atlas of Biomarkers for Alzheimer's Disease, Springer, 2014.
[64] S.T. DeKosky, K. Marek, Looking backward to move forward: early detection of neurodegenerative disorders, Science (80-.). 302 (2003) 830–834.
[65] T. Harkany, I. Abraham, W. Timmerman, G. Laskay, B. Toth, M. Sasvari, C. Konya, J.B. Sebens, J. Korf, C. Nyakas, β-Amyloid neurotoxicity is mediated by a glutamate-triggered excitotoxic cascade in rat nucleus basalis, Eur. J. Neurosci. 12 (2000) 2735–2745.
[66] R.G. Nagele, M.R. D'Andrea, H. Lee, V. Venkataraman, H.-Y. Wang, Astrocytes accumulate Aβ42 and give rise to astrocytic amyloid plaques in Alzheimer disease brains, Brain Res. 971 (2003) 197–209.
[67] L. Bertram, D. Blacker, K. Mullin, D. Keeney, J. Jones, S. Basu, S. Yhu, M.G. McInnis, R.C.P. Go, K. Vekrellis, Evidence for genetic linkage of Alzheimer's disease to chromosome 10q, Science (80-.). 290 (2000) 2302–2303.
[68] A. Myers, P. Holmans, H. Marshall, J. Kwon, D. Meyer, D. Ramic, S. Shears, J. Booth, F.W. DeVrieze, R. Crook, Susceptibility locus for Alzheimer's disease on chromosome 10, Science (80-.). 290 (2000) 2304–2305.
[69] J.-H. Kang, M. Korecka, J.B. Toledo, J.Q. Trojanowski, L.M. Shaw, Clinical utility and analytical challenges in measurement of cerebrospinal fluid amyloid-β1–42 and τ proteins as Alzheimer disease biomarkers, Clin. Chem. 59 (2013) 903–916.
[70] T. Khan, Biomarkers in Alzheimer's Disease, Academic Press, 2016.
[71] A. Papassotiropoulos, C. Hock, Biochemical markers of Alzheimer's disease: wish and reality, Neurobiol. Aging. 23 (2002) 513–514.
[72] H. Zetterberg, N. Mattsson, L.M. Shaw, K. Blennow, Biochemical markers in Alzheimer's disease clinical trials, Biomark. Med. 4 (2010) 91–98.
[73] A. Nakamura, N. Kaneko, V.L. Villemagne, T. Kato, J. Doecke, V. Doré, C. Fowler, Q.-X. Li, R. Martins, C. Rowe, High performance plasma amyloid-β

biomarkers for Alzheimer's disease, Nature. 554 (2018) 249.
[74] M.W. Weiner, D.P. Veitch, P.S. Aisen, L.A. Beckett, N.J. Cairns, J. Cedarbaum, R.C. Green, D. Harvey, C.R. Jack, W. Jagust, E. Liu, others, 2014 Update of the Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception, Alzheimer's Dement. J. Alzheimer's Assoc. 11 (2015) e1-e120.
[75] T. Sunderland, G. Linker, N. Mirza, K.T. Putnam, D.L. Friedman, L.H. Kimmel, J. Bergeson, G.J. Manetti, M. Zimmermann, B. Tang, Decreased β-amyloid1-42 and increased tau levels in cerebrospinal fluid of patients with Alzheimer disease, Jama. 289 (2003) 2094–2103.
[76] H.M. Schipper, The role of biologic markers in the diagnosis of Alzheimer's disease, Alzheimer's Dement. J. Alzheimer's Assoc. 3 (2007) 325–332.
[77] R. Palaniappan, Biological signal analysis, BookBoon, 2011.
[78] S. Tong, N.V. Thakor, Quantitative EEG analysis methods and clinical applications, Artech House, 2009.
[79] S. Sanei, J.A. Chambers, EEG signal processing, John Wiley & Sons, 2013.
[80] L. Hirsch, R. Brenner, Atlas of EEG in critical care, John Wiley & Sons, 2011.
[81] J.D. Kropotov, Quantitative EEG, event-related potentials and neurotherapy, Academic Press, 2010.
[82] W.O. Tatum IV, Handbook of EEG interpretation, Demos Medical Publishing, 2014.
[83] Y.A.L. Pijnenburg, Y. Vd Made, A.M.V.C. Van Walsum, D.L. Knol, P. Scheltens, C.J. Stam, EEG synchronization likelihood in mild cognitive impairment and Alzheimer's disease during a working memory task, Clin. Neurophysiol. 115 (2004) 1332–1339.
[84] C.-F.V. Latchoumane, E. Ifeachor, N. Hudson, S. Wimalaratna, J. Jeong, Dynamical nonstationarity analysis of resting EEGs in Alzheimer's disease, in: Int. Conf. Neural Inf. Process., Springer, 2007: pp. 921–929.
[85] J. Dauwels, F.-B. Vialatte, A. Cichocki, On the early diagnosis of Alzheimer's disease from EEG signals: a mini-review, in: Adv. Cogn. Neurodynamics, Springer, 2011: pp. 709–716.
[86] D. Abásolo, R. Hornero, C. Gómez, M. García, M. López, Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure, Med. Eng. Phys. 28 (2006) 315–322.
[87] C. Babiloni, G. Binetti, E. Cassetta, G. Dal Forno, C. Del Percio, F. Ferreri, R. Ferri, G. Frisoni, K. Hirata, B. Lanuzza, Sources of cortical rhythms change as a function of cognitive impairment in pathological aging: a multicenter study, Clin. Neurophysiol. 117 (2006) 252–268.
[88] C. Babiloni, C. Del Percio, F. Vecchio, F. Sebastiano, G. Di Gennaro, P.P. Quarato, R. Morace, L. Pavone, A. Soricelli, G. Noce, Alpha, beta and gamma electrocorticographic rhythms in somatosensory, motor, premotor and prefrontal cortical areas differ in movement execution and observation in humans, Clin. Neurophysiol. 127 (2016) 641–654.
[89] G.G. Yener, D.D. Emek-Savaş, R. Lizio, B. Çavuşoğlu, F. Carducci, E. Ada, B. Güntekin, C.C. Babiloni, E. Başar, Frontal delta event-related oscillations relate to frontal volume in mild cognitive impairment and healthy controls, Int. J. Psychophysiol. 103 (2016) 110–117.
[90] D. Abásolo, R. Hornero, P. Espino, J. Poza, C.I. Sánchez, R. de la Rosa, Analysis of regularity in the EEG background activity of Alzheimer's disease patients with Approximate Entropy, Clin. Neurophysiol. 116 (2005) 1826–1834.
[91] H. Kantz, T. Schreiber, Nonlinear time series analysis, Second edi, Cambridge university press, Cambridge , UK, 2004.
[92] A. Lempel, J. Ziv, On the complexity of finite sequences, IEEE Trans. Inf.

Theory. 22 (1976) 75–81.
[93] G. Tononi, G.M. Edelman, O. Sporns, Complexity and coherency: integrating information in the brain, Trends Cogn. Sci. 2 (1998) 474–484.
[94] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, H. Scheingraber, A comparative classification of complexity measures, Chaos, Solitons & Fractals. 4 (1994) 133–173.
[95] C. Besthorn, H. Sattel, C. Geiger-Kabisch, R. Zerfass, H. Förstl, Parameters of EEG dimensional complexity in Alzheimer's disease, Electroencephalogr. Clin. Neurophysiol. 95 (1995) 84–89.
[96] A.P. Burgess, J. Rehman, J.D. Williams, Changes in neural complexity during the perception of 3D images using random dot stereograms, Int. J. Psychophysiol. 48 (2003) 35–42.
[97] F. Gu, X. Meng, E. Shen, Z. Cai, Can we measure consciousness with EEG complexities?, Int. J. Bifurc. Chaos. 13 (2003) 733–742.
[98] H. Adeli, S. Ghosh-Dastidar, Automated EEG-based diagnosis of neurological disorders: Inventing the future of neurology, CRC press, 2010.
[99] J.C. McBride, X. Zhao, N.B. Munro, C.D. Smith, G.A. Jicha, L. Hively, L.S. Broster, F.A. Schmitt, R.J. Kryscio, Y. Jiang, Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease, Comput. Methods Programs Biomed. 114 (2014) 153–163.
[100] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy, T. Musha, F.-B. Vialatte, C. Latchoumane, J. Jeong, A. Cichocki, Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin?, Int. J. Alzheimer's Dis. 2011 (2011).
[101] H. Garn, M. Waser, M. Deistler, T. Benke, P. Dal-Bianco, G. Ransmayr, H. Schmidt, G. Sanin, P. Santer, G. Caravias, Quantitative EEG markers relate to Alzheimer's disease severity in the Prospective Dementia Registry Austria (PRODEM), Clin. Neurophysiol. 126 (2015) 505–513.
[102] T.J. De Bock, S. Das, M. Mohsin, N.B. Munro, L.M. Hively, Y. Jiang, C.D. Smith, D.R. Wekstein, G.A. Jicha, A. Lawson, Early detection of Alzheimer's disease using nonlinear analysis of EEG via Tsallis entropy, in: Biomed. Sci. Eng. Conf. (BSEC), 2010, IEEE, 2010: pp. 1–4.
[103] B. Deng, L. Liang, S. Li, R. Wang, H. Yu, J. Wang, X. Wei, Complexity extraction of electroencephalograms in Alzheimer's disease with weighted-permutation entropy, Chaos An Interdiscip. J. Nonlinear Sci. 25 (2015) 43105.
[104] N. Mammone, L. Bonanno, S. De Salvo, S. Marino, P. Bramanti, A. Bramanti, F.C. Morabito, Permutation disalignment index as an indirect, EEG-based, measure of brain connectivity in MCI and AD patients, Int. J. Neural Syst. 27 (2017) 1750020.
[105] B. Hamadicharef, C. Guan, E. Ifeachor, N. Hudson, S. Wimalaratna, Performance evaluation and fusion of methods for early detection of Alzheimer Disease, in: Biomed. Eng. Informatics, 2008. BMEI 2008. Int. Conf., IEEE, 2008: pp. 347–351.
[106] J. Dauwels, F. Vialatte, A. Cichocki, Diagnosis of Alzheimer's disease from EEG signals: where are we standing?, Curr. Alzheimer Res. 7 (2010) 487–505.
[107] C. Coronel, H. Garn, M. Waser, M. Deistler, T. Benke, P. Dal-Bianco, G. Ransmayr, S. Seiler, D. Grossegger, R. Schmidt, Quantitative EEG Markers of Entropy and Auto Mutual Information in Relation to MMSE Scores of Probable Alzheimer's Disease Patients, Entropy. 19 (2017) 130.
[108] R. Sneddon, W.R. Shankle, J. Hara, A. Rodriquez, D. Hoffman, U. Saha, EEG detection of early Alzheimer's disease using psychophysical tasks, Clin. EEG Neurosci. 36 (2005) 141–150.

[109] N. Nagaraj, K. Balasubramanian, Dynamical complexity of short and noisy time series, Eur. Phys. J. Spec. Top. 226 (2017) 2191–2204.
[110] D. Abásolo, S. Simons, R.M. da Silva, G. Tononi, V. V Vyazovskiy, Lempel-Ziv complexity of cortical activity during sleep and waking in rats, J. Neurophysiol. 113 (2015) 2742–2752.
[111] M.W. Rivolta, M. Migliorini, M. Aktaruzzaman, R. Sassi, A.M. Bianchi, Effects of the series length on Lempel-Ziv Complexity during sleep, in: Eng. Med. Biol. Soc. (EMBC), 2014 36th Annu. Int. Conf. IEEE, IEEE, 2014: pp. 693–696.
[112] X.-S. Zhang, R.J. Roy, E.W. Jensen, EEG complexity as a measure of depth of anesthesia for patients, IEEE Trans. Biomed. Eng. 48 (2001) 1424–1433.
[113] M. Aboy, R. Hornero, D. Abásolo, D. Álvarez, Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis, IEEE Trans. Biomed. Eng. 53 (2006) 2282–2288.
[114] S. Simons, D. Abásolo, Distance-Based Lempel-Ziv Complexity for the Analysis of Electroencephalograms in Patients with Alzheimer's Disease, Entropy. 19 (2017) 129.
[115] R. Hornero, D. Abásolo, J. Escudero, C. Gómez, Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease, Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci. 367 (2009) 317–336.
[116] R. Hornero, J. Escudero, A. Fernández, J. Poza, C. Gómez, Spectral and nonlinear analyses of MEG background activity in patients with Alzheimer's disease, IEEE Trans. Biomed. Eng. 55 (2008) 1658–1665.
[117] A. Fernandez, R. Hornero, C. Gomez, A. Turrero, P. Gil-Gregorio, J. Matías-Santos, T. Ortiz, Complexity analysis of spontaneous brain activity in Alzheimer disease and mild cognitive impairment: an MEG study, Alzheimer Dis. Assoc. Disord. 24 (2010) 182–189.
[118] C. Goh, B. Hamadicharef, G. Henderson, E. Ifeachor, Comparison of fractal dimension algorithms for the computation of EEG biomarkers for dementia, in: 2nd Int. Conf. Comput. Intell. Med. Healthc., 2005.
[119] A. Accardo, M. Affinito, M. Carrozzi, F. Bouquet, Use of the fractal dimension for the analysis of electroencephalographic time series, Biol. Cybern. 77 (1997) 339–350.
[120] H. Preißl, W. Lutzenberger, F. Pulvermüller, N. Birbaumer, Fractal dimensions of short EEG time series in humans, Neurosci. Lett. 225 (1997) 77–80.
[121] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, Phys. D Nonlinear Phenom. 31 (1988) 277–283.
[122] C. Pizzuti, M.D. Ritchie, M. Giacobini, Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics: 9th European Conference, EvoBIO 2011, Torino, Italy, April 27-29, 2011, Proceedings, Springer Science & Business Media, 2011.
[123] R. Esteller, G. Vachtsevanos, J. Echauz, B. Lilt, A comparison of fractal dimension algorithms using synthetic and experimental data, in: Circuits Syst. 1999. ISCAS'99. Proc. 1999 IEEE Int. Symp., IEEE, 1999: pp. 199–202.
[124] C. Gómez, Á. Mediavilla, R. Hornero, D. Abásolo, A. Fernández, Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients, Med. Eng. Phys. 31 (2009) 306–313.
[125] F.M. Smits, C. Porcaro, C. Cottone, A. Cancelli, P.M. Rossini, F. Tecchio, Electroencephalographic fractal dimension in healthy ageing and Alzheimer's disease, PLoS One. 11 (2016) e0149587.
[126] T. Staudinger, R. Polikar, Analysis of complexity based EEG features for the diagnosis of Alzheimer's disease, in: Eng. Med. Biol. Soc. EMBC, 2011 Annu. Int. Conf. IEEE, IEEE, 2011: pp. 2033–2036.

[127] A. Molinaro, Y.D. Sergeyev, An efficient algorithm for the zero crossing detection in digitized measurement signal, Measurement. 30 (2001) 187–196.
[128] A.S. Zandi, R. Tafreshi, M. Javidan, G.A. Dumont, Predicting epileptic seizures in scalp EEG based on a variational Bayesian Gaussian mixture model of zero-crossing intervals, IEEE Trans. Biomed. Eng. 60 (2013) 1401–1413.
[129] J.L. Semmlow, B. Griffel, Biosignal and medical image processing, CRC press, 2014.
[130] D. Brandwood, Fourier transforms in radar and signal processing, Artech House, 2012.
[131] T. Locatelli, M. Cursi, D. Liberati, M. Franceschi, G. Comi, EEG coherence in Alzheimer's disease, Electroencephalogr. Clin. Neurophysiol. 106 (1998) 229–237.
[132] N.K. Al-Qazzaz, S. Brassen, A. Jajcevic, EEG coherence in Alzheimer's dementia, J. Neural Transm. 110 (2003) 1051–1058.
[133] J.J. Dunkin, A.F. Leuchter, T.F. Newton, I.A. Cook, Reduced EEG coherence in dementia: state or trait marker?, Biol. Psychiatry. 35 (1994) 870–879.
[134] M. Gell-Mann, C. Tsallis, Nonextensive entropy: interdisciplinary applications, Oxford University Press, 2004.
[135] C. Gómez, D. Abásolo, J. Poza, A. Fernández, R. Hornero, MEG analysis in Alzheimer's disease computing approximate entropy for different frequency bands, in: Eng. Med. Biol. Soc. (EMBC), 2010 Annu. Int. Conf. IEEE, IEEE, 2010: pp. 2379–2382.
[136] D. Abásolo1, R. Hornero, P. Espino, Approximate entropy of EEG background activity in alzheimer's disease patients, Intell. Autom. Soft Comput. 15 (2009) 591–603.
[137] S.M. Pincus, Approximate entropy as a measure of system complexity., Proc. Natl. Acad. Sci. 88 (1991) 2297–2301.
[138] E. Alpaydin, Introduction to machine learning, MIT press, 2014.
[139] D. Michie, D.J. Spiegelhalter, C.C. Taylor, Machine learning, neural and statistical classification, (1994).
[140] K.P. Murphy, Machine learning : a probabilistic perspective, The MIT Press, London, 2012.
[141] D. Zhang, Advances in machine learning applications in software engineering, IGI Global, 2006.
[142] E. Alpaydin, Introduction to Machine Learning/Ethem Alpaydin, (2010).
[143] S. Marsland, Machine learning: an algorithmic perspective, Chapman and Hall/CRC, 2014.
[144] K.-Z. Huang, H. Yang, I. King, M.R. Lyu, Machine learning: modeling data locally and globally, Springer Science & Business Media, 2008.
[145] S. Abe, Support vector machines for pattern classification, Springer, 2005.
[146] V. Jakkula, Tutorial on support vector machine (svm), Sch. EECS, Washingt. State Univ. 37 (2006).
[147] B. Schölkopf, A.J. Smola, F. Bach, Learning with kernels: support vector machines, regularization, optimization, and beyond, MIT press, 2002.
[148] N. Cristianini, J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods, Cambridge university press, 2000.
[149] V. Kecman, L. Wang, Support vector machines: theory and applications, (2005).
[150] M. Pontil, A. Verri, Properties of support vector machines, Neural Comput. 10 (1998) 955–974.
[151] P. Cunningham, S.J. Delany, k-Nearest neighbour classifiers, Mult. Classif. Syst. 34 (2007) 1–17.
[152] E. Alpaydin, Introduction to machine learning, The MIT Press, Cambridge,

Massachusetts, 2004.
[153] G. Biau, L. Devroye, Lectures on the nearest neighbor method, Springer, 2015.
[154] K.L. Priddy, P.E. Keller, Artificial neural networks: an introduction, SPIE press, 2005.
[155] Margaret E. O'Kane, NCQA Measuring quality. Improving health care, (n.d.). http://www.ncqa.org/hedis-quality-measurement/performance-measurement (accessed April 11, 2018).
[156] P.C. Smith, Performance measurement for health system improvement: experiences, challenges and prospects, Cambridge University Press, 2009.
[157] E.D. Peterson, E.R. DeLong, F.A. Masoudi, S.M. O'brien, P.N. Peterson, J.S. Rumsfeld, D.M. Shahian, R.E. Shaw, ACCF/AHA 2010 position statement on composite measures for healthcare performance assessment: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures (Writing Committee to Develop a Position , Circulation. 121 (2010) 1780–1791.
[158] J.A. Spertus, K.A. Eagle, H.M. Krumholz, K.R. Mitchell, S.-L.T. Normand, A.C. of Cardiology, A.H.A.T.F. on P. Measures, American College of Cardiology and American Heart Association methodology for the selection and creation of performance measures for quantifying the quality of cardiovascular care, J. Am. Coll. Cardiol. 45 (2005) 1147–1156.
[159] M.R. King, N.A. Mody, Numerical and statistical methods for bioengineering: applications in MATLAB, Cambridge University Press, 2010.
[160] P. Baldi, S. Brunak, Bioinformatics: the machine learning approach, MIT press, 2001.
[161] C.P. Friedman, J. Wyatt, Evaluation methods in biomedical informatics, Springer Science & Business Media, 2005.
[162] Á.H.E. Corchado, Computational Intelligence in Security for Information Systems, (n.d.).
[163] R.J. Mural, M.N. Liebman, Biomedical informatics in translational research, Artech House, 2008.
[164] T. Chankong, N. Theera-Umpon, S. Auephanwiriyakul, Cervical cell classification using Fourier transform, in: 13th Int. Conf. Biomed. Eng., Springer, 2009: pp. 476–480.
[165] D.G. Altman, J.M. Bland, Statistics Notes: Diagnostic tests 2: predictive values, Bmj. 309 (1994) 102.
[166] A.K. Akobeng, Understanding diagnostic tests 1: sensitivity, specificity and predictive values, Acta Paediatr. 96 (2007) 338–341.
[167] B. Thies, E. Truschke, M. Morrison-Bogorad, R.J. Hodes, N.I. on A.W. Group, Consensus Report of the Working Group on: "Molecular and Biochemical Markers of Alzheimer's Disease," Neurobiol. Aging. 19 (1998) 109–116. doi:10.1016/S0197-4580(98)00022-0.
[168] Y. Sasaki, The truth of the F-measure, Teach Tutor Mater. 1 (2007).
[169] Z. Liu, M. Tan, F. Jiang, Regularized F-measure maximization for feature selection and classification, Biomed Res. Int. 2009 (2009).
[170] G. Hripcsak, A.S. Rothschild, Agreement, the f-measure, and reliability in information retrieval, J. Am. Med. Informatics Assoc. 12 (2005) 296–298.
[171] J. De Weerdt, M. De Backer, J. Vanthienen, B. Baesens, A robust F-measure for evaluating discovered process models, in: Comput. Intell. Data Min. (CIDM), 2011 IEEE Symp., IEEE, 2011: pp. 148–155.
[172] O. Lund, Immunological bioinformatics, MIT press, 2005.
[173] S. Raschka, An Overview of General Performance Metrics of Binary Classifier Systems, ArXiv Prepr. ArXiv1410.5330. (2014).
[174] D.M.W. Powers, Evaluation: From Precision, Recall and F-Factor to ROC,

Informedness, Markedness & Correlation, School of Informatics and Engineering, Flinders University, Adelaide, Australia, TR SIE-07-001, Journal of Machine Learning Technologies 2: 1 37-63. https://dl-web. dropbox. com/get/Public/201101-Evaluation_JMLT_Postprint-Colour. pdf, 2007.
[175] C. Ferri, P. Flach, J. Hernández-Orallo, Learning decision trees using the area under the ROC curve, in: ICML, 2002: pp. 139–146.
[176] T. Fawcett, ROC graphs: Notes and practical considerations for researchers, Mach. Learn. 31 (2004) 1–38.
[177] J. Muschelli, ROC and AUC with a Binary Predictor: a Potentially Misleading Metric, ArXiv Prepr. ArXiv1903.04881. (2019).
[178] S. Wu, P. Flach, A scored AUC metric for classifier evaluation and selection, in: Second Work. ROC Anal. ML, Bonn, Ger., 2005.
[179] R. Boddy, G. Smith, Statistical methods in practice: for scientists and technologists, John Wiley & Sons, 2009.
[180] B. Rosner, Fundamentals of biostatistics, Nelson Education, 2015.
[181] D.B. Panagiotakos, The value of p-value in biomedical research, Open Cardiovasc. Med. J. 2 (2008) 97.
[182] P. Sedgwick, What is a P value?, BMJ Br. Med. J. 345 (2012).
[183] C. Daraio, L. Simar, Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications, Springer Science & Business Media, 2007.
[184] M.K. Chung, Statistical and computational methods in brain image analysis, CRC press, 2013.
[185] L. Kang, Y. Liu, S. Zeng, Advances in Computation and Intelligence, Springer, 2007.
[186] N.A.C. Cressie, H.J. Whitford, How to Use the Two Sample t-Test, Biometrical J. 28 (1986) 131–148.
[187] M. Delacre, D. Lakens, C. Leys, Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test, Int. Rev. Soc. Psychol. 30 (2017).
[188] R.R. Wilcox, Introduction to robust estimation and hypothesis testing, Academic press, 2011.
[189] R.R. Wilcox, Fundamentals of modern statistical methods: Substantially improving power and accuracy, Springer Science & Business Media, 2010.
[190] H. Abdi, Bonferroni and Sidak corrections for multiple comparisons, Encycl. Meas. Stat. 3 (2007) 103–107.
[191] B. Walsh, Multiple comparisons: Bonferroni corrections and false discovery rates, Lect. Notes EEB. 581 (2004).
[192] J.H. McDonald, Handbook of biological statistics, sparky house publishing Baltimore, MD, 2009.
[193] M.X. Cohen, MATLAB for brain and cognitive scientists, MIT Press, 2017.
[194] M.A. Napierala, What is the Bonferroni correction, AAOS Now. 6 (2012) 40.
[195] M. Stumpf, D.J. Balding, M. Girolami, Handbook of statistical systems biology, John Wiley & Sons, 2011.
[196] M.F. Folstein, S.E. Folstein, P.R. McHugh, "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician, J. Psychiatr. Res. 12 (1975) 189–198.
[197] O. Spreen, C.D. Stretton, A compendium of neuropsychological tests; administration, norms and commentary. (ISBN 0-19-510019-0). New YorL: Oxford University Press, (1998).
[198] A.L. Benton, Revised Visual Retention Test: Clinical and Experimental Applications Psychological Corporation, New York. (1974).
[199] G.A. Talland, M. Ekdahl, Psychological studies of Korsakoff's psychosis: IV. The rate and mode of forgetting narrative material., J. Nerv. Ment. Dis. 129

(1959) 391–404.
[200] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, E.M. Stadlan, Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease, Neurology. 34 (1984) 939.
[201] C.P. Hughes, L. Berg, W. Danziger, L.A. Coben, R.L. Martin, A new clinical scale for the staging of dementia, Br. J. Psychiatry. 140 (1982) 566–572.
[202] J.A. Yesavage, T.L. Brink, T.L. Rose, O. Lum, V. Huang, M. Adey, V.O. Leirer, Development and validation of a geriatric depression screening scale: a preliminary report, J. Psychiatr. Res. 17 (1982) 37–49.
[203] W.G. Rosen, R.D. Terry, P.A. Fuld, R. Katzman, A. Peck, Pathological verification of ischemic score in differentiation of dementias, Ann. Neurol. Off. J. Am. Neurol. Assoc. Child Neurol. Soc. 7 (1980) 486–488.
[204] M.P. Lawton, E.M. Brody, Assessment of older people: self-maintaining and instrumental activities of daily living, Gerontologist. 9 (1969) 179–186.
[205] G.C. Roman, T.K. Tatemichi, T. Erkinjuntti, J.L. Cummings, J.C. Masdeu, J.H. Garcia, L. Amaducci, J.M. Orgogozo, A. Brun, A. Hofman, Diagnostic criteria for research studies: report of the NINDS-AIREN International Workshop, Neurology. 43 (1993) 250–260.
[206] J. Tuqan, P.P. Vaidyanathan, Optimum low cost two channel IIR orthonormal filter bank, in: Acoust. Speech, Signal Process. 1997. ICASSP-97., 1997 IEEE Int. Conf., IEEE, 1997: pp. 2425–2428.
[207] B.A. Shenoi, Introduction to digital signal processing and filter design, John Wiley & Sons, 2005.
[208] R.J. Schilling, S.L. Harris, Fundamentals of digital signal processing using MATLAB, Cengage learning, 2011.
[209] M. Tsolaki, Clinical workout for the early detection of cognitive decline and dementia, Eur. J. Clin. Nutr. 68 (2014) 1186.
[210] C. Babiloni, R. Ferri, G. Binetti, F. Vecchio, G.B. Frisoni, B. Lanuzza, C. Miniussi, F. Nobili, G. Rodriguez, F. Rundo, Directionality of EEG synchronization in Alzheimer's disease subjects, Neurobiol. Aging. 30 (2009) 93–102.
[211] D.V. Moretti, Association of EEG, MRI, and regional blood flow biomarkers is predictive of prodromal Alzheimer's disease, Neuropsychiatr. Dis. Treat. 11 (2015) 2779.
[212] D. V Moretti, G.B. Frisoni, G. Binetti, O. Zanetti, Comparison of the effects of transdermal and oral rivastigmine on cognitive function and EEG markers in patients with Alzheimer's disease, Front. Aging Neurosci. 6 (2014).
[213] P.A.M. Kanda, E.F. Oliveira, F.J. Fraga, EEG epochs with less alpha rhythm improve discrimination of mild Alzheimer's, Comput. Methods Programs Biomed. 138 (2017) 13–22.
[214] C. Babiloni, R. Lizio, C. Del Percio, N. Marzano, A. Soricelli, E. Salvatore, R. Ferri, F.I.I. Cosentino, G. Tedeschi, P. Montella, Cortical sources of resting state EEG rhythms are sensitive to the progression of early stage Alzheimer's disease, J. Alzheimer's Dis. 34 (2013) 1015–1035.
[215] C. Babiloni, R. Lizio, N. Marzano, P. Capotosto, A. Soricelli, A.I. Triggiani, S. Cordone, L. Gesualdo, C. Del Percio, Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms, Int. J. Psychophysiol. 103 (2016) 88–102.
[216] C. Babiloni, C. Del Percio, P. Capotosto, G. Noce, F. Infarinato, C. Muratori, C. Marcotulli, G. Bellagamba, E. Righi, A. Soricelli, Cortical sources of resting state electroencephalographic rhythms differ in relapsing-remitting and secondary progressive multiple sclerosis, Clin. Neurophysiol. 127 (2016) 581–

590.
[217] D.V. Moretti, electroencephalography-driven approach to prodromal Alzheimer's disease diagnosis: from biomarker integration to network-level comprehension, Clin. Interv. Aging. 11 (2016) 897.
[218] W. Klimesch, EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, Brain Res. Rev. 29 (1999) 169–195.
[219] D. V Moretti, D. Paternicò, G. Binetti, O. Zanetti, G.B. Frisoni, EEG markers are associated to gray matter changes in thalamus and basal ganglia in subjects with mild cognitive impairment, Neuroimage. 60 (2012) 489–496.
[220] P. Zhao, P. Van-Eetvelt, C. Goh, N. Hudson, S. Wimalaratna, E. Ifeachor, Characterization of EEGs in Alzheimer's disease using information theoretic methods, in: Eng. Med. Biol. Soc. 2007. EMBS 2007. 29th Annu. Int. Conf. IEEE, IEEE, 2007: pp. 5127–5131.
[221] D. Abásolo, R. Hornero, P. Espino, J. Escudero, C. Gomez, Electroencephalogram background activity characterization with approximate entropy and auto mutual information in Alzheimer's disease patients, in: Eng. Med. Biol. Soc. 2007. EMBS 2007. 29th Annu. Int. Conf. IEEE, IEEE, 2007: pp. 6191–6194.
[222] D. Abásolo, J. Escudero, R. Hornero, C. Gómez, P. Espino, Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer's disease patients, Med. Biol. Eng. Comput. 46 (2008) 1019–1028.
[223] S. Simons, D. Abasolo, M. Hughes, Investigation of Alzheimer's Disease EEG Frequency Components with Lempel-Ziv Complexity, in: 6th Eur. Conf. Int. Fed. Med. Biol. Eng., Springer, 2015: pp. 46–49.
[224] N.K. Al-Qazzaz, S. Ali, S.A. Ahmad, M.S. Islam, J. Escudero, Entropy-Based Markers of EEG Background Activity of Stroke-Related Mild Cognitive Impairment and Vascular Dementia Patients, (n.d.).
[225] R. Alcaraz, J.J. Rieta, A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms, Biomed. Signal Process. Control. 5 (2010) 1–14.
[226] S.M. Pincus, Assessing serial irregularity and its implications for health, Ann. N. Y. Acad. Sci. 954 (2001) 245–267.
[227] L.I. Aftanas, N. V Lotova, V.I. Koshkarov, S.A. Popov, Non-linear dynamical coupling between different brain areas during evoked emotions: an EEG investigation, Biol. Psychol. 48 (1998) 121–138.
[228] S. Gudmundsson, T.P. Runarsson, S. Sigurdsson, G. Eiriksdottir, K. Johnsen, Reliability of quantitative EEG features, Clin. Neurophysiol. 118 (2007) 2162–2171.
[229] J.I. Fleck, J. Kuti, J. Brown, J.R. Mahon, C. Gayda-Chelder, Frontal-posterior coherence and cognitive function in older adults, Int. J. Psychophysiol. 110 (2016) 217–230.
[230] A.F. LEUCHTER, T.F. NEWTON, I.A.N.A. COOK, D.O. WALTER, S. ROSENBERG-THOMPSON, P.A. LACHENBRUCH, Changes in brain functional connectivity in Alzheimer-type and multi-infarct dementia, Brain. 115 (1992) 1543–1561.
[231] E.C. Ifeachor, G.T. Henderson, C. Goh, H.S.K. Wimalaratna, N. Hudson, Biopattern analysis and subject-specific diagnosis and care of dementia, in: Eng. Med. Biol. Soc. 2005. IEEE-EMBS 2005. 27th Annu. Int. Conf., IEEE, 2006: pp. 2490–2493.
[232] C. Laske, H.R. Sohrabi, S.M. Frost, K. López-de-Ipiña, P. Garrard, M. Buscema, J. Dauwels, S.R. Soekadar, S. Mueller, C. Linnemann, Innovative diagnostic tools for early detection of Alzheimer's disease, Alzheimer's Dement. 11 (2015) 561–578.

[233] J.R. Evans, A. Abarbanel, Introduction to quantitative EEG and neurofeedback, 2nd ed., Elsevier, 1999.
[234] E. Capecci, Z.G. Doborjeh, N. Mammone, F. La Foresta, F.C. Morabito, N. Kasabov, Longitudinal study of alzheimer's disease degeneration through EEG data analysis with a NeuCube spiking neural network model, in: 2016 Int. Jt. Conf. Neural Networks, IEEE, 2016: pp. 1360–1366.
[235] G.W. Fenton, Electrophysiology of Alzheimer's disease, Br. Med. Bull. 42 (1986) 29–33.
[236] D. V Moretti, Theta and alpha EEG frequency interplay in subjects with mild cognitive impairment: evidence from EEG, MRI, and SPECT brain modifications, Front. Aging Neurosci. 7 (2015) 31.
[237] U.A. Khan, L. Liu, F.A. Provenzano, D.E. Berman, C.P. Profaci, R. Sloan, R. Mayeux, K.E. Duff, S.A. Small, Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease, Nat. Neurosci. 17 (2014) 304–311.
[238] J. Escudero, D. Abásolo, R. Hornero, P. Espino, M. López, Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy, Physiol. Meas. 27 (2006) 1091.
[239] M. Zanin, L. Zunino, O.A. Rosso, D. Papo, Permutation entropy and its main biomedical and econophysics applications: a review, Entropy. 14 (2012) 1553–1577.
[240] J. Ramírez, J.M. Górriz, D. Salas-Gonzalez, A. Romero, M.M. López, I. Álvarez, M. Gómez-Río, J. Ramirez, J.M. Górriz, D. Salas-Gonzalez, A. Romero, M.M. López, I. Álvarez, M. Gómez-R\'\io, Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features, Inf. Sci. (Ny). 237 (2013) 59–72.
[241] R. Chaves, J. Ramírez, J.M. Górriz, M. López, D. Salas-Gonzalez, I. Alvarez, F. Segovia, SVM-based computer-aided diagnosis of the Alzheimer's disease using t-test NMSE feature selection with feature correlation weighting, Neurosci. Lett. 461 (2009) 293–297.
[242] J.M. Górriz, J. Ramírez, A. Lassl, D. Salas-Gonzalez, E.W. Lang, C.G. Puntonet, I. Álvarez, M. López, M. Gómez-Río, Automatic computer aided diagnosis tool using component-based SVM, in: Nucl. Sci. Symp. Conf. Rec. 2008. NSS'08. IEEE, IEEE, 2008: pp. 4392–4395.
[243] A. So, D. Hooshyar, K.W. Park, H.S. Lim, Early Diagnosis of Dementia from Clinical Data by Machine Learning Techniques, Appl. Sci. 7 (2017) 651.
[244] I. Neuner, J. Arrubla, C.J. Werner, K. Hitz, F. Boers, W. Kawohl, N.J. Shah, The default mode network and EEG regional spectral power: a simultaneous fMRI-EEG study, PLoS One. 9 (2014) e88214.
[245] T. Harmony, T. Fernández, J. Silva, J. Bosch, P. Valdés, A. Fernández - Bouzas, L. Galán, E. Aubert, D. Rodríguez, Do specific EEG frequencies indicate different processes during mental calculation?, Neurosci. Lett. 266 (1999) 25–28.
[246] T. Fernández, T. Harmony, M. Rodriguez, J. Bernal, J. Silva, A. Reyes, E. Marosi, EEG activation patterns during the performance of tasks involving different components of mental calculation, Electroencephalogr. Clin. Neurophysiol. 94 (1995) 175–182.
[247] M. Teplan, others, Fundamentals of EEG measurement, Meas. Sci. Rev. 2 (2002) 1–11.
[248] K.E. Peters, C.C. Walters, J.M. Moldowan, The biomarker guide, Cambridge University Press, 2005.
[249] N. Houmani, F. Vialatte, E. Gallego-Jutglà, G. Dreyfus, V.-H. Nguyen-Michel, J. Mariani, K. Kinugawa, Diagnosis of Alzheimer's disease with

Electroencephalography in a differential framework, PLoS One. 13 (2018) e0193607.
[250] W. Buntine, M. Grobelnik, D. Mladenic, J. Shawe-Taylor, Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2009, Bled, Slovenia, September 7-11, 2009, Proceedings, Springer, 2009.
[251] Y. Pang, S. Wang, Y. Yuan, Learning regularized LDA by clustering, IEEE Trans. Neural Networks Learn. Syst. 25 (2014) 2191–2201.
[252] K. Grill-Spector, R. Henson, A. Martin, Repetition and the brain: neural models of stimulus-specific effects, Trends Cogn. Sci. 10 (2006) 14–23.
[253] W. Klimesch, Memory processes, brain oscillations and EEG synchronization, Int. J. Psychophysiol. 24 (1996) 61–100.
[254] S. Muthukumaraswamy, High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations, Front. Hum. Neurosci. 7 (2013) 138.
[255] Y. Cao, L. Cai, J. Wang, R. Wang, H. Yu, Y. Cao, J. Liu, Characterization of complexity in the electroencephalograph activity of Alzheimer's disease based on fuzzy entropy, Chaos An Interdiscip. J. Nonlinear Sci. 25 (2015) 83116.
[256] N. Mao, Y. Liu, K. Chen, L. Yao, X. Wu, Combinations of Multiple Neuroimaging Markers using Logistic Regression for Auxiliary Diagnosis of Alzheimer Disease and Mild Cognitive Impairment, Neurodegener. Dis. 18 (2018) 91–106.
[257] A. Kassambara, Machine Learning Essentials: Practical Guide in R, sthda, 2018.
[258] J.P. Amezquita-Sanchez, N. Mammone, F.C. Morabito, S. Marino, H. Adeli, A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals, J. Neurosci. Methods. 322 (2019) 88–95.
[259] X. Chai, X. Weng, Z. Zhang, Y. Lu, G. Liu, H. Niu, Quantitative EEG in Mild Cognitive Impairment and Alzheimer's Disease by AR-Spectral and Multi-scale Entropy Analysis, in: World Congr. Med. Phys. Biomed. Eng. 2018, Springer, 2019: pp. 159–163.
[260] Y.-X. Sun, X. Ji, X. Mao, L. Xie, J. Jia, V. Galvan, D.A. Greenberg, K. Jin, Differential activation of mTOR complex 1 signaling in human brain with mild to severe Alzheimer's disease, J. Alzheimer's Dis. 38 (2014) 437–444.
[261] R. Schmidt, E. Hofer, F.H. Bouwman, K. Buerger, C. Cordonnier, T. Fladby, D. Galimberti, J. Georges, M.T. Heneka, J. Hort, EFNS-ENS/EAN Guideline on concomitant use of cholinesterase inhibitors and memantine in moderate to severe Alzheimer's disease, Eur. J. Neurol. 22 (2015) 889–898.
[262] J.T. Coyle, D.L. Price, M.R. Delong, Alzheimer's disease: a disorder of cortical cholinergic innervation, Science (80-.). 219 (1983) 1184–1190.
[263] A.N. Leggett, S. Zarit, A. Taylor, J.E. Galvin, Stress and burden among caregivers of patients with Lewy body dementia, Gerontologist. 51 (2010) 76–85.
[264] A.J. Haes, L. Chang, W.L. Klein, R.P. Van Duyne, Detection of a biomarker for Alzheimer's disease from synthetic and clinical samples using a nanoscale optical biosensor, J. Am. Chem. Soc. 127 (2005) 2264–2271.
[265] S. Engelborghs, K. De Vreese, T. Van de Casteele, H. Vanderstichele, B. Van Everbroeck, P. Cras, J.-J. Martin, E. Vanmechelen, P.P. De Deyn, Diagnostic performance of a CSF-biomarker panel in autopsy-confirmed dementia, Neurobiol. Aging. 29 (2008) 1143–1159.
[266] H. Struyfs, W. Van Hecke, J. Veraart, J. Sijbers, S. Slaets, M. De Belder, L. Wuyts, B. Peters, K. Sleegers, C. Robberecht, Diffusion kurtosis imaging: a possible MRI biomarker for AD diagnosis?, J. Alzheimer's Dis. 48 (2015) 937–948.