ALZHEIMER DISEASE DIAGNOSIS USING ELECTROENCEPHALOGRAM(EEG)

A MINI PROJECT REPORT

Submitted by

GOWDHAMA PRIYA C
KAVIPRIYAA E P
KULAPRAGALYAA K T R
SRIVIDHYA M

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND BUSINESS SYSTEMS

K.RAMAKRISHNAN COLLEGE OF ENGINEERING
(AUTONOMOUS)
SAMAYAPURAM, TRICHY

ANNA UNIVERSITY CHENNAI

600 025

DECEMBER 2024

ALZHEIMER DISEASE DIAGNOSIS USING ELECTROENCEPHALOGRAM(EEG)

A MINI PROJECT REPORT

Submitted by

GOWDHAMA PRIYA C
KAVIPRIYAA E P
KULAPRAGALYAA K T R
SRIVIDHYA M

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND BUSINESS SYSTEMS

K.RAMAKRISHNAN COLLEGE OF ENGINEERING
(AUTONOMOUS)
SAMAYAPURAM, TRICHY

ANNA UNIVERSITY CHENNAI

600 025

DECEMBER 2024

ALZHEIMER DISEASE DIAGNOSIS USING ELECTROENCEPHALOGRAM(EEG)

UCB1512 MINI PROJECT REPORT

Submitted by

GOWDHAMA PRIYA C	**(8115U22CB020)**
KAVIPRIYAA E P	**(8115U22CB028)**
KULAPRAGALYAA K T R	**(8115U22CB033)**
SRIVIDHYA M	**(8115U22CB054)**

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND BUSINESS SYSTEMS

Under the Guidance of

Dr. J. SASIDEVI, M.E., Ph.D.,

Department of Computer Science and Business Systems

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

COMPUTER SCIENCE AND BUSINESS SYSTEMS

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Under

ANNA UNIVERSITY, CHENNAI

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)
Under
ANNA UNIVERSITY, CHENNAI

BONAFIDE CERTIFICATE

Certified that this project report "**ALZHEIMER DISEASE DIAGNOSIS USING ELECTROENCEPHALOGRAM (EEG)**" is the bonafide work of "**GOWDHAMA PRIYA C (8115U22CB020), KAVIPRIYAA E P (8115U22CB028), KULAPRAGALYAA K T R (8115U22CB033), SRIVIDHYA M (8115U22CB054)**" who carried out the project work under my supervision.

SIGNATURE

Dr. J. SASIDEVI,M.E., Ph.D.,

Associate Professor

HEAD OF THE DEPARTMENT

Department of Computer Science and BusinessSystems,

K.Ramakrishnan College of Engineering (Autonomous), Samayapuram,

Trichy-621112

SIGNATURE

Dr. J. SASIDEVI,M.E., Ph.D.,

Associate Professor

SUPERVISOR

Department of Computer Science and BusinessSystems,

K.Ramakrishnan College of Engineering (Autonomous), Samayapuram,

Trichy-621112

SIGNATURE OF INTERNAL EXAMINER

NAME:

DATE:

SIGNATURE OF EXTERNAL EXAMINER

NAME:

DATE:

ACKNOWLEDGEMENT

We thank the almighty GOD, without whom it would not have been possible for us to complete our Mini Project.

We wish to address our profound gratitude to **Dr. K. RAMAKRISHNAN**, Chairman, K.Ramakrishnan College of Engineering (Autonomous) who encouraged and gave us all help throughout the course.

We express our hearty gratitude and thanks to our honorable and grateful executive director **Dr. S. KUPPUSAMY, B.Sc., MBA., Ph.D.**, K.Ramakrishnan College of Engineering (Autonomous).

We are glad to thank our principal **Dr. D. SRINIVASAN, M.E, Ph.D., FIE., MIIW., MISTE., MISAE., C.Engg.**, for giving us permission to carry out this Mini Project.

We wish to convey our sincere thanks to **Dr. J. SASIDEVI, M.E., Ph.D.**, Head of the Department, Computer Science and Business Systems for giving us constant encouragement and advice throughout the course.

We are grateful to **Dr. J. SASIDEVI, M.E., Ph.D.**, Associate Professor, Department of Computer Science and Business Systems, K.Ramakrishnan College of Engineering (Autonomous), for her guidance and valuable suggestions during the course of study.

Finally, we sincerely acknowledged in no less term for all our staff members, our parents and friends for their co-operation and help at various stages of this Mini Project work.

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Under

ANNA UNIVERSITY, CHENNAI

DECLARATION BY THE CANDIDATE

I declare that to the best of my knowledge the work reported here in has been composed solely by myself and that it has not been in whole or in part in any previous application for a degree.

Submitted for the Mini Project Viva-Voce held at K.Ramakrishnan College of Engineering on _____

SIGNATURE OF THE CANDITATE

(GOWDHAMA PRIYA C)

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Under

ANNA UNIVERSITY, CHENNAI

DECLARATION BY THE CANDIDATE

I declare that to the best of my knowledge the work reported here in has been composed solely by myself and that it has not been in whole or in part in any previous application for a degree.

Submitted for the Mini Project Viva-Voce held at K.Ramakrishnan College of Engineering on _____

SIGNATURE OF THE CANDITATE

(KAVIPRIYAA E P)

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Under

ANNA UNIVERSITY, CHENNAI

DECLARATION BY THE CANDIDATE

I declare that to the best of my knowledge the work reported here in has been composed solely by myself and that it has not been in whole or in part in any previous application for a degree.

Submitted for the Mini Project Viva-Voce held at K.Ramakrishnan College of Engineering on _____

SIGNATURE OF THE CANDITATE

(KULAPRAGALYAA K T R)

K.RAMAKRISHNAN COLLEGE OF ENGINEERING

(AUTONOMOUS)

Under

ANNA UNIVERSITY, CHENNAI

DECLARATION BY THE CANDIDATE

I declare that to the best of my knowledge the work reported here in has been composed solely by myself and that it has not been in whole or in part in any previous application for a degree.

Submitted for the Mini Project Viva-Voce held at K.Ramakrishnan College of Engineering on _____

SIGNATURE OF THE CANDITATE

(SRIVIDHYA M)

TABLE OF CONTENTS

ABSTRACT

Alzheimer's disease is a progressive neurodegenerative disorder and the most common cause of dementia, leading to a steady decline in memory, cognitive abilities, and behavior. Over time, individuals experience increasing difficulty performing daily tasks, starting with early symptoms such as memory lapses, confusion, and trouble solving problems. The disease is marked by the buildup of amyloid plaques and neurofibrillary tangles in the brain, which result in the death of brain cells and a gradual loss of brain function. As Alzheimer's advances, it significantly impairs the individual's ability to manage even basic activities, eventually requiring full-time care. Early detection is crucial because it offers the opportunity to slow the progression of the disease and improve the quality of life for both patients and their families by enabling earlier treatment and intervention strategies.

In this context, electroencephalography (EEG) shows promise as a non-invasive, innovative tool for early diagnosis. EEG measures the brain's electrical activity, providing valuable insights into neural patterns and connectivity. This study analyzed EEG data from three groups healthy individuals, those with mild cognitive impairment (MCI), and Alzheimer's patients identifying distinct biomarkers that can differentiate between these stages of cognitive decline. Notably, significant changes in the alpha and beta power bands were observed, along with disruptions in network connectivity across brain regions. These alterations reflect early neural deterioration before Alzheimer's symptoms become fully apparent. The alpha and beta bands are essential for attention, memory, and processing information, and their disruption signals cognitive decline in Alzheimer's patients. The ability of EEG to detect these changes makes it a valuable, cost-effective tool for early Alzheimer's diagnosis, potentially improving the ability to intervene early and slow disease progression. Further research will be key to refining these biomarkers for wider clinical application.

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS	FULL FORM
AD	Alzheimer Disease
EEG	Electroencephalogram
SVC	Support Vector Classifier
SNN	Simple Neural Network
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short Term Memory
MCI	Mild Cognitive Impairment
ROC	Receiver Operating Characteristic
FVC	Forced Vital Capacity
MOCA	Montreal Cognitive Assessment
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
UML	Unified Model Language
OS	Operating Systems

CHAPTER 1

INTRODUCTION

Alzheimer’s disease (AD) is a progressive and irreversible neurodegenerative disorder that primarly affects older adults, leading to significant impairments in memory, cognitive functions, and behavior. It is the most common cause of dementia, accounting for 60–80% of dementia cases globally. It progresses through three main phases mild, moderate, and severe with early diagnosis being particularly challenging due to the overlap of symptoms with normal aging, especially in the case of mild cognitive impairment (MCI). Electroencephalography (EEG), a non-invasive technique that measures electrical activity in the brain, has emerged as a promisingtool for diagnosing AD. The disease affects EEG patterns by reducing neural complexity, slowingsignal frequencies, and disrupting the synchrony between brain regions. Notably, AD patients exhibit increased activity in lower-frequency bands like delta and theta, while higher-frequency bands, such as alpha and beta, show decreased activity, reflecting cognitive decline and impairedbrain function. Advanced machine learning algorithms applied to EEG data can further enhance diagnostic accuracy, offering potential for early detection by distinguishing between healthy individuals and those in various stages of dementia.

1.1 Variables for Analysis

Age

Age is a critical demographic variable influencing the onset and progression of Alzheimer's disease. Including age in the analysis helps explore correlations between aging and cognitive decline, as understanding the age distribution within the dataset is essential for interpreting results and tailoring interventions to specific age groups.

EEG Features (Theta and Delta Power)

The analysis focuses on the power of theta and delta waves in EEG recordings, as these frequency bands are often altered in individuals with Alzheimer's disease. Measuring these features allows for the assessment of brain activity patterns that correlate with cognitive impairment and the severity of the disease.

Cognitive Function Tests

In conjunction with EEG data, standardized cognitive assessments provide a quantitative measure of cognitive abilities. Analyzing results from tests such as the Mini Mental State Examination (MMSE) or Montreal Cognitive Assessment (MOCA) helps evaluate the relationship between EEG features and cognitive decline.

Neuropsychiatric Inventory (NPI)

Considering the psychological aspects of Alzheimer's, the NPI assesses behavioral and psychological symptoms in patients. Understanding the mental health status of individuals with Alzheimer's disease contributes to a holistic view of the condition, recognizing the interplay between cognitive decline and behavioral changes.

Clinical Dementia Rating (CDR)

The CDR scale offers a structured assessment of dementia severity. Including this variable aids in categorizing the extent of cognitive impairment in patients, facilitating a more comprehensive understanding of the relationship between EEG findings and clinical manifestations of Alzheimer'sdiseases.

1.2 Data Collection Process

The data collection process for diagnosing Alzheimer’s disease using EEG follows several key steps. It begins by recruiting participants based on specific criteria and obtaining informed consent. A pre-assessment gathers medical and cognitive histories, along with standardized tests like the MMSE or MOCA. The EEG setup includes placing electrodes according to the international 10 to 20 system and checking impedance for optimal signal quality. Clinical records of 500 patients, both with and without Alzheimer’s, are compiled, covering a wide range of variables to capture the disease's complexity. Resting state EEG is recorded for 20 to 30 minutes, sometimes with additional cognitive tasks. The data is then preprocessed to remove noise and extract relevant features. Statistical and machine learning analyses compare EEG patterns between healthy individuals and those with Alzheimer’s. Validation techniques ensure the findings' robustness, which are documented, highlighting clinical implications

1.3 Proposed Model

The Current System is a significant advancement in EEG signal analysis, leveraging advanced machine learning algorithms like Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. These algorithms enhance the accuracy and efficiency of diagnosing neurological conditions. The system prioritizes patient comfort through non-invasive, user-friendly technologies, reducing anxiety with innovative device designs and enabling remote monitoring, allowing patients to maintain their daily activities. To ensure data integrity, the system employs rigorous preprocessing techniques to enhance signal quality, alongside robust encryption and transparent consent processes for ethical data handling. Its hybrid models for classification allow continuous adaptation and improvement, aligning with emerging research and data patterns. Overall, the Current System provides a sophisticated, patient-centric approach to EEG analysis, equipping clinicians with a powerful tool for accurate and timely diagnosis.

1.4 Analytical Framework

The evaluation of the CNN model's performance involves a rigorous analytical framework. Several metrics are employed to assess its accuracy, Processing feature extraction, classification, validation and overall predictive power. The dataset is divided into training and testing sets to validate the model's generalizability beyond the training data.

Accuracy

It is defined as the ratio of the number of correct predictions to the total number of predictions made. In the context of Alzheimer’s disease (AD) classification, accuracy helps evaluate how well the model distinguishes between AD patients and healthy controls (non-AD cases) based on EEG data. However, while accuracy is an important metric, it may not always provide the complete picture, especially in imbalanced datasets where one class (e.g., healthy controls) may dominate. In such cases, additional metrics like precision, recall, and F1 score are often used to provide a more balanced view of the model’s performance. The ability to correctly classify AD patients and avoid false negatives is particularly crucial in early-stage detection

Preprocessing

EEG signals are inherently noisy and susceptible to interference from factors like eye movements, muscle contractions, and external electrical noise, which can distort the data. To ensure accurate analysis, preprocessing is essential and involves several steps. Artifact removal is the first step, which focuses on eliminating noise and artifacts using techniques like Independent Component Analysis (ICA) or Principal Component Analysis (PCA) to separate unwanted signals from brain activity. Next, filtering is applied using band pass filters to focus on relevant frequency bands such as delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30+ Hz), as each band is associated with different cognitive states. In Alzheimer's patients, an increase in delta and theta activity and a decrease in alpha and beta activity are typically observed. Finally, segmentation divides the EEG data into smaller epochs, focusing on specific brain activity events or cognitive tasks. Proper segmentation is crucial for analyzing temporal patterns. Preprocessing ensures that the EEG data is clean, standardized, and ready for feature extraction, reducing errors in subsequent analysis stages.

Feature Extraction

Feature extraction involves deriving meaningful data from raw EEG signals to input into machine learning models, aiming to identify patterns indicative of Alzheimer's Disease (AD) or Mild Cognitive Impairment (MCI). Common techniques include Power Spectral Density (PSD), which analyzes the distribution of power across frequency bands; in AD patients, there are notable changes such as increased delta and theta activity and decreased alpha and beta power. Functional connectivity, which examines the statistical dependency between different brain regions, is another key feature, revealing disruptions in communication among brain areas, withreduced connectivity observed in AD patients. Additionally, complexity measures like entropy assess the unpredictability or disorder in EEG signals; healthy individuals typically exhibit morecomplex brain activity, while AD patients show reduced complexity, reflecting a loss of functional diversity. Techniques like sample entropy and fuzzy entropy are used to analyze signal irregularities. These extracted features help researchers better understand altered brain activity patterns in AD, aiding in effective patient classification.

Classification

Once relevant features are extracted from EEG signals, the next step is to classify the data into categories such as Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and healthy controls using machine learning algorithms. Support Vector Machine (SVM) is a supervised algorithm that separates different classes by finding the best hyperplane in the feature space, making it effective for AD classification, especially in high-dimensional, nonlinear data. Convolutional Neural Networks (CNN), commonly used in image classification, are also applied to EEG data, as they automatically extract hierarchical features from time-series signals, making them powerful tools for AD classification. Decision Trees, which use a flowchart-like structure to classify data based on feature values, are less common in EEG classification but can still provide simple, interpretable rules. Random Forests, an ensemble method that combines multiple decision trees, help improve classification accuracy and avoid overfitting, enhancing model robustness. These models learn from labeled data to categorize new, unseen EEG data, aiding in the diagnosis of AD and MCI.

Validation

Validation is a crucial step to ensure the reliability and generalizability of machine learning models across different datasets and patient populations. Techniques like k-fold cross-validation, where the dataset is divided into subsets and the model is trained on some while tested on others, help assess performance on various data splits and reduce overfitting. External validation further strengthens model reliability by using an independent dataset that was not part of the training, ensuring the model can generalize to new, unseen data. Additionally, combining EEG-based classification with other diagnostic tools, such as MRI scans, PET scans, or clinical scores, enhances diagnostic accuracy by providing complementary information on brain structure, function, and cognitive status. This multi-modal approach ensures that the model's performance remains consistent, reliable, and trustworthy for clinical decision-making in real-world settings.

1.5 Causes

Alzheimer's disease is a neurodegenerative condition characterized by progressive cognitive decline, including memory loss and difficulties in thinking, reasoning, and communication. Although the exact cause is not fully understood, some contributing factors have been identified, includes

Amyloid Plaques

Amyloid plaques are abnormal clumps of beta amyloid protein fragments that accumulate between neurons in the brain, disrupting the normal communication and functionof brain cells. These plaques form when beta amyloid, a protein naturally produced in the brain,fails to be properly cleared and accumulates in the spaces between neurons. Over time, the accumulation of these plaques disrupts synaptic function, impairs neuronal signaling, and triggersan inflammatory response that contributes to cognitive decline. The presence of amyloid plaquesis one of the hallmarks of Alzheimer's disease (AD) and is thought to play a central role in its progression. As the plaques grow in size, they can disrupt the delicate balance of neurotransmission, leading to memory loss, difficulty with learning new information, and other cognitive impairments that are characteristic of Alzheimer's.

Neurofibrillary Tangles

Inside neurons, tau proteins normally help stabilize microtubules, which are essential for cell structure and transport. In Alzheimer's disease, tau proteins become twist into tangles, forming neurofibrillary tangles inside neurons. These twisted tau fibers disruptthe structural integrity of the neuron and hinder the normal transport of nutrients and other vital molecules. As a result, neurons begin to malfunction, lose their ability to communicate effectively with each other, and eventually die. Neurofibrillary tangles often accumulate in key brain regions involved in memory and cognition, such as the hippocampus and cortex, contributing to the progressive cognitive decline observed in AD. The development of neurofibrillary tangles is thought to be a major factor in the worsening of symptoms as the diseaseadvances.

Genetic Factors

Genetic factors play a significant role in the development of Alzheimer's disease, particularly in early onset forms of the disease. Mutations in the APP (amyloid precursorprotein), PSEN1 (presenilin 1), and PSEN2 (presenilin 2) genes have been linked to familial Alzheimer's, a rare form of the disease that typically manifests before age 65. These mutations lead to an overproduction of beta amyloid, contributing to the formation of amyloid plaques. In addition to these rare mutations, the presence of the APOE ε4 allele, a variant of the apolipo protein E gene, is the most significant genetic risk factor for late onset Alzheimer's disease, whichtypically occurs after age 65.

Inflammation

Chronic neuro inflammation is a key factor in the progression of Alzheimer's disease and is often triggered by the presence of amyloid plaques. As the immune system responds to the accumulation of these plaques, microglial cells (the brain's immune cells) become activated and inflammatory molecules, such as cytokines and chemokines. While this immune response is intended to clear the plaques, prolonged inflammation can cause further neuronal damage. Inflammation accelerates the breakdown of neuronal tissue, disrupts synaptic function, and contributes to the overall neurodegeneration seen in Alzheimer's patients. Additionally, inflammatory processes can exacerbate other disease mechanisms, including tau tangling and amyloid plaque accumulation. As inflammation spreads through the brain, it leads to a vicious cycle that worsens cognitive decline and speeds up the progression of Alzheimer's disease.

Vascular Issues

Vascular problems, such as impaired blood flow and cardiovascular disease,are increasingly recognized as contributing factors to Alzheimer's disease. The brain relies on a rich network of blood vessels to supply it with oxygen and essential nutrients. When these bloodvessels become damaged or blocked, reduced blood flow leads to a lack of oxygen, which impairsbrain cell function. Vascular issues, such as stroke, hypertension, or atherosclerosis, can contribute to the development of Alzheimer's by disrupting the brain's blood supply and promoting oxidative stress, further accelerating the degeneration of neurons. The combination ofreduced oxygen delivery and impaired removal of waste products in the brain can lead to neuronalinjury and the buildup of amyloid plaques. Individuals with cardiovascular disease or poor vascular health are at a higher risk of developing Alzheimer's, suggesting that maintaining goodcardiovascular health is an important preventive measure.

Environmental Factors

In addition to genetic predisposition, environmental factors such as age, lifestyle choices, and head trauma play a significant role in the onset and progression of Alzheimer's disease. Age is the most significant non-genetic risk factor, with the likelihood of developing Alzheimer's increasing significantly after age 65. Lifestyle factors, including physical activity, diet, social engagement, and cognitive stimulation, can also influence the risk of developing Alzheimer's. Research suggests that individuals who maintain an active lifestyle, eat a balanced diet (such as the Mediterranean diet), and engage in mentally stimulating activitiesmay have a lower risk of

developing Alzheimer’s or may experience a slower progression of thedisease. Additionally, head trauma, particularly repeated concussions or traumatic brain injury, has been associated with an increased risk of developing Alzheimer’s later in life. Preventive measures, such as wearing protective headgear, exercising regularly, maintaining a healthy diet,and engaging in cognitive exercises, may help reduce the risk of developing Alzheimer's disease or delay its onset. The interplay of genetic, environmental, and lifestyle factors underscores the complexity of Alzheimer’s disease and highlights the importance of early detection, preventivemeasures, and ongoing research to develop effective treatments.

1.6 Symptoms

EEG can help detect specific brain activity changes that correlate with symptoms of Alzheimer’s disease. Here are some common symptoms of Alzheimer's that can be observed through EEG includes:

Memory Loss

Explanation: One of the earliest and most prominent symptoms of Alzheimer's. It primarily affects short term memory, making it hard to remember recent events, conversations, or appointments, though long-term memories might remain intact for a while.

EEG Changes: Reduced activity in the hippocampus (critical for memory processing) and prefrontal cortex, leading to decreased alpha and beta wave power, which are crucial for information retention and retrieval.

Confusion and Disorientation

Explanation: Patients may become confused about time, place, or even the identity of familiar people. This disorientation can cause them to get lost, even in familiar environments.

EEG Changes: Reduced connectivity between the parietal and temporal lobes, which are important for spatial reasoning and navigation. EEG shows slowing of brainwave patterns, with an increase in theta and delta waves, contributing to the confusion.

Difficulty with Language

Explanation: As Alzheimer’s progresses, patients struggle to find the right words or names (anomia), and may have trouble following or joining conversations. This symptom reflects the deterioration of areas in the brain responsible for language.

EEG Changes: Reduced alpha and beta activity in the left temporal lobe and Wernicke's area, which are involved in language comprehension and production, may contribute to difficulties in communication.

Poor Judgment and Decision-Making

Explanation: Alzheimer's affects the brain's ability to make sound decisions. Patients might make unsafe choices or demonstrate poor judgment, such as giving away large sums of money or neglecting personal hygiene.

EEG Changes: Diminished beta wave activity in the frontal lobe, particularly in the prefrontal cortex, which governs planning, judgment, and decision-making.

Impaired Attention and Concentration

Explanation: Alzheimer's patients often struggle to maintain focus and are easily distracted. They may find it difficult to follow conversations or complete tasks that require sustained attention.

EEG Changes: A reduction in beta waves, which are associated with attention and active mental processes, particularly in the frontal and parietal lobes, is commonly observed.

Sleep Disturbances

Explanation: Alzheimer's often disrupts normal sleep patterns, leading to insomnia, frequent waking at night, or increased sleepiness during the day. These disturbances can worsen cognitive symptoms.

EEG Changes: Abnormal sleep EEG patterns, including reduced REM sleep and disrupted sleep cycles. Patients often show a reduction in alpha waves during sleep, which contributes to fragmented and non restorative sleep.

CHAPTER 2

LITERATURE SURVEY

Recent advancements in Alzheimer's disease diagnosis using EEG focus on identifying distinct brainwave patterns and leveraging machine learning techniques for classification. EEG biomarkers such as the slowing of brainwaves particularly in the alpha and beta frequency bands are consistently observed in Alzheimer's patients, along with disrupted brain connectivity. Machine learning algorithms, including Support Vector Machines (SVM), RandomForests, and deep learning models, are being used to classify EEG data, demonstrating promisingaccuracy in distinguishing Alzheimer's patients from healthy individuals or those with mild cognitive impairment (MCI). These approaches, coupled with the non-invasive nature of EEG, provide a potential early diagnostic tool for Alzheimer's disease, though challenges such as datavariability and noise remain significant hurdles.

2.1 Title: Alzheimer's Disease Patients and Some Classification through EEG Signal Processing

Author: Giulia Fiscon, Emanuel Weitschek, Giovanni Felici, Paola Bertolazzi, Simona De Salvo, Placido Bramanti, Maria Cristina De Cola

Alzheimer's disease (AD) is a prevalent neurodegenerative disorder that primarily affects older adults, causing progressive cognitive decline, memory loss, and difficulty with daily activities. As the disease advances, symptoms worsen, leading to severe memory impairments and loss of independence. Although there is no cure for AD, early and accurate diagnosis is crucial for initiating timely therapeutic interventions, which can improve patients quality of life and potentially delay severe symptoms. Mild Cognitive Impairment (MCI), often considered a precursor to AD, presents cognitive impairments that are noticeable but not yet debilitating. Differentiating MCI from early-stage AD is challenging, as both conditions share similarities with normal aging. Diagnosis typically involves clinical evaluations, neuropsychological tests, and neuroimaging, but definitive confirmation can only be made through post-mortem brain analysis.

EEG, a non-invasive method that records electrical activity in the brain, has shown promise in diagnosing AD and MCI. By measuring brain waves through electrodes placed on the scalp, EEGcan reveal alterations in brain wave frequencies that are associated with various cognitive states.Research has demonstrated that in AD patients, there is an increase in delta and theta activity and a decrease in alpha and beta activity, which may reflect cognitive decline and neural disruptions. Thus, EEG has the potential to differentiate AD and MCI from healthy controls based on these distinctive brain wave patterns. However, traditional EEG analysis is often manual, requiring expert interpretation, which is time-consuming and can introduce subjectivity.

In response to these challenges, the study by Fiscon et al. introduces an automated method for analyzing EEG signals to classify patients as healthy, MCI, or AD cases. The process begins with pre-processing the EEG data to remove noise and artifacts, ensuring accurate analysis. After pre-processing, time-frequency transformations are applied to extract relevant features from the EEG signals, identifying specific frequency band changes that are characteristic of AD. Common techniques like Fourier or Wavelet Transforms are used for this purpose. The final step involves training machine learning classifiers on the extracted features, enabling the system to automatically distinguish between healthy controls, MCI patients, and AD patients. This method aims to enhance the diagnostic process by providing a faster, more objective, and scalable solution.

This research offers several advantages, including the non-invasive, safe, and cost-effective nature of EEG compared to neuroimaging methods like MRI or PET scans. The automated approach reduces the need for expert analysis, making it more accessible and faster, particularly in resource-limited settings. Machine learning enhances diagnostic accuracy by providing a quantitative, objective assessment of EEG data, reducing subjective biases. Early detection of AD and MCI could lead to timely interventions, potentially slowing disease progression. However, the approach faces challenges such as the need for extensive pre-processing, the potential for misclassification due to subtle symptoms, and the computational resources required for time-frequency transformations and machine learning. Despite these challenges, the study represents a promising step forward in using EEG and machine learning for more effective AD diagnosis.

Algorithm

Pre-processing of EEG Data: Filter raw EEG signals to remove noise and artifacts and segment data into meaningful epochs, and normalize or standardize as needed for analysis.

Machine Learning Classification: Train a classifier (e.g., SVM or CNN) on extracted features and evaluate its performance to predict or classify outcomes.

Advantages

A non-invasive diagnostic tool minimizes the reliance on expert interpretation.

It provides quantitative and objective analysis, ensuring greater accuracy and consistency.

Disadvantages

Complex pre-processing requirements can make data handling and preparation more difficult.

Challenges with model generalization may reduce the model's ability to perform well on unseen data.

The risk of misclassification and high computational resource demands can affect model accuracy and efficiency.

2.2 Title: Alzheimer's Disease Detection by using Deep Learning Algorithm: A MiniReview

Author: Suhad Al-Shoukry, Taha H. Rassem, Nasrin M. Makbol

This mini review delves into the growing importance of deep learning techniques in the early detection of Alzheimer's disease (AD), emphasizing the use of neuroimaging data, specifically MRI scans, as a crucial tool for diagnosis. Alzheimer's disease is a progressive neurodegenerative disorder characterized by cognitive decline, memory loss, and behavioral changes, and early detection plays a pivotal role in managing the disease and improving patient outcomes. Traditional diagnostic methods for AD, such as clinical evaluation and neuropsychological testing, often fall short in detecting the disease in its early stages. However, advancements in neuroimaging, particularly with MRI scans, offer detailed insights into structural changes in the brain associated with AD, such as cortical atrophy and hippocampal shrinkage.

Deep learning, a subset of machine learning that uses neural networks to model complex patterns in data, has shown great promise in enhancing the accuracy and efficiency of AD detection from neuroimaging data. By training deep learning models on large datasets of brain images, these algorithms can automatically extract features and detect subtle changes in brain structure that may indicate the onset of Alzheimer's, even before clinical symptoms become apparent. Convolutional neural networks (CNNs), a popular deep learning model, have been widely used for analyzing MRI scans, allowing for the identification of early biomarkers and providing objective, quantitative assessments of brain regions affected by AD.

The application of deep learning to MRI data has demonstrated significant improvements in diagnostic accuracy, enabling more reliable differentiation between healthy individuals, mild cognitive impairment (MCI), and AD patients. These techniques also hold the potential for longitudinal monitoring of disease progression, offering a non-invasive, efficient, and reproducible method for tracking the evolution of AD over time. While deep learning techniques show great promise, challenges such as data variability, model interpretability, and the need for large, diverse datasets remain. Nonetheless, as these challenges are addressed, deep learning is expected to play an increasingly important role in the early detection, diagnosis, and management of Alzheimer's disease.

Algorithm

Convolutional Neural Networks (CNNs): Primarily used for image processing and computer vision tasks due to their ability to extract spatial features effectively.

Recurrent Neural Networks (RNNs): Designed for sequential data, such as time series or natural language processing, with a focus on learning temporal dependencies.

Autoencoders: Used for unsupervised learning, dimensionality reduction, and anomaly detection by learning efficient data representations.

Support Vector Machines (SVMs): A versatile classification and regression algorithm that works well on small to medium sized datasets and handles high dimensional spaces.

Advantages

Early Detection: Enables the identification of patterns and anomalies at an early stage.

Automated Feature Extraction: Automatically extracts relevant features, reducing the need for manual intervention.

Improved Accuracy: Enhances predictive performance by learning complex patterns .

Disadvantages

Data Requirements: Requires large datasets to train effectively, which can be challenging in some applications.

Complexity: These models can be computationally intensive and difficult to implement and tune.

Overfitting: Risk of the model learning noise in the data, leading to poor generalization to new data.

2.3 Title: Earlier Diagnosis and classification of Alzheimer Disease using MRI

Author: Waleed Salehi , Preety Baglat , Ankita Upadhya , Brijbhushan Sharma

The paper examines and compares the performance of two deep learning models, VGG19 and DenseNet169, in detecting Alzheimer's disease from MRI (Magnetic Resonance Imaging) images. Alzheimer's disease (AD) is a progressive neurodegenerative disorder that primarily affects memory and cognitive functions, and its early detection is crucial for timely intervention and management. MRI is a non-invasive imaging technique that provides detailed structural images of the brain, making it a valuable tool for identifying abnormalities associated with Alzheimer's, such as brain atrophy and changes in the hippocampus. The study focuses on leveraging the power of deep learning models to automatically detect these changes from MRI scans, eliminating the need for manual interpretation, which can be time-consuming and prone to human error.

In the study, two well-known convolutional neural network (CNN) architectures, VGG19 and DenseNet169, are applied to the task of Alzheimer's disease classification. VGG19 is a deep CNN model that has shown strong performance in image classification tasks, known for its simplicity and effectiveness in learning hierarchical features from raw images. DenseNet169, onthe other hand, is a more advanced architecture that incorporates dense connections between layers, allowing for better feature reuse and more efficient learning. DenseNet169 has shown promising results in various medical image classification tasks due to its ability to learn complexpatterns from smaller datasets.

The researchers utilize transfer learning and fine tuning techniques on both models to improve their performance. Transfer learning involves taking a pre trained model, which has already been trained on a large dataset (e.g., ImageNet), and adapting it to a new task with a smaller dataset. This approach helps overcome the problem of limited data in medical imaging tasks, where annotated data is often scarce. Fine-tuning refers to adjusting the pre-trained model's weights to better suit the target dataset in this case, MRI images of Alzheimer's patients. By applying these techniques, the models are able to adapt to the specific features of the Alzheimer's dataset, improving their accuracy in classifying images as either AD or control.

The results of the study reveal that DenseNet169 outperforms VGG19 in terms of classificationaccuracy, achieving 87% accuracy on the Kaggle Alzheimer's MRI dataset. The model's ability to achieve high accuracy suggests that it can effectively distinguish between Alzheimer's diseaseand healthy controls based on MRI scans. This performance is particularly valuable in the contextof early detection, as it allows for the identification of subtle brain changes that may be indicativeof the disease before symptoms fully manifest.

In summary, this research demonstrates the potential of deep learning models, particularly DenseNet169, in enhancing the diagnostic process for Alzheimer's disease using MRI images. By leveraging transfer learning and fine-tuning, the models are able to achieve high levels of accuracy despite the limited availability of medical imaging data. The study provides a promisingdirection for the development of automated diagnostic tools that can assist healthcare professionals in detecting Alzheimer's disease earlier and more accurately, ultimately improvingpatient outcomes and guiding timely interventions.

Algorithm

VGG19 and DenseNet169 Architecture: Deep neural network architectures designed for high-performance image classification and feature extraction.

Transfer Learning: Utilizes pre-trained models to adapt to new tasks with reduced data and training time.

Fine-Tuning: Adjusts a pre-trained model's parameters for improved performance on a specific dataset.

Advantages

High Classification Accuracy: Achieves superior performance in classifying complex patterns.

Effective Feature Extraction: Extracts key features from data, improving the model's

predictive ability.

Transfer Learning Benefits: Leverages pre-trained models to reduce the need for large datasets and shortens training time.

Disadvantages

Data Requirement: Requires large datasets for effective training, which can be a limitation in some cases.

These architectures are resource-intensive and require significant computational power for training and inference.

2.4 Title: Impact of Alzheimer's Disease on Care Partners: A Systematic Review

Author: Kristian Steen Frederiksen , Wendy Weidner , Soeren Mattke

This systematic review provides an in-depth exploration of the diverse and significant impacts that caring for individuals with Alzheimer's disease (AD) has on care partners. Alzheimer's disease, a progressive neurodegenerative disorder, not only affects those diagnosed but also places immense physical, emotional, and economic burdens on family members, friends, and professional caregivers. This review aims to address the multidimensional challenges caregivers face and to raise awareness about the need for support systems to alleviate the strain on these individuals.

First, the physical impacts of caregiving are discussed. Alzheimer's caregiving is often demanding, requiring long hours of attention and supervision, especially as the disease progresses. Caregivers typically engage in strenuous tasks, including assisting with daily activities like bathing, dressing, and eating, which can result in physical exhaustion. The constant lifting, repositioning, and helping with mobility can lead to musculoskeletal injuries, particularly among family caregivers. The review highlights how these physical strains can negatively affect the caregiver's own health, leading to burnout, chronic fatigue, and even physical illnesses due to the prolonged stress of caregiving responsibilities.

In addition to the physical toll, the emotional burden of caregiving is immense. Watching a loved one gradually lose their cognitive functions can lead to feelings of grief, sadness, and helplessness. Caregivers often experience emotional stress from witnessing the person they care for no longer recognize them or suffer from personality changes. Furthermore, the emotional

strain is exacerbated by the responsibility of making difficult decisions about the care needs of the patient, which can result in guilt, anxiety, and depression. The review draws attention to how these emotional impacts contribute to overall caregiver distress and mental health deterioration, and it underscores the importance of psychological support and counseling for caregivers.

The economic impact of caregiving is another critical aspect discussed in the review. Alzheimer's disease can impose substantial financial burdens on families, particularly as patients require increasing levels of care over time. This includes direct costs such as medical bills, specialized care services, home modifications, and transportation to appointments. Many caregivers also experience indirect economic losses, such as reduced work hours or quitting their jobs provide full-time care. The review explores how these financial challenges often strain family finances and lead to a lower quality of life for caregivers, making the case for policychanges and financial support systems for those in caregiving roles.

Furthermore, the review highlights the need for better support networks, including respite care services, counseling, and financial assistance, to help alleviate the burdens faced by caregivers.It calls for improved training for caregivers on how to manage both the physical and emotional demands of the job, and it stresses the importance of creating policies that provide caregivers with better access to resources, healthcare, and social services. With Alzheimer's disease becoming more prevalent as the global population ages, this review emphasizes that providingadequate support for caregivers is essential for maintaining their well-being and ensuring that individuals with Alzheimer's receive the care they deserve.

In summary, the systematic review examines the complex and multi faceted impacts of Alzheimer's care giving, showing how the physical, emotional, and economic challenges can take a significant toll on care partners. It calls for a more robust support system for care givers to reduce the strain and improve their quality of life. Ultimately, addressing these challenges is not only crucial for the well being of caregivers but also for enhancing the overall quality of care provided to those living with Alzheimer's disease.

Algorithm

Data Preprocessing: The process of cleaning, normalizing, and transforming raw data into a structured format suitable for analysis or training machine learning models.

Feature Extraction: The process of selecting and transforming relevant data attributes into a

reduced set of features that better represent the underlying patterns for the model.

Advantages

Early Detection and Diagnosis: Facilitates early identification of diseases or issues, improving outcomes.

Non-Invasive Diagnostic Tool: Provides a safer alternative for diagnosis without the need for invasive procedures.

Disadvantages

Need for Large Datasets: Requires extensive labelled data to train models effectively.

Risk of Overfitting: Models may perform well on training data but fail to generalize to new, unseen data.

CHAPTER 3

SYSTEM ANALYSIS

3.1 Existing System

The existing system for EEG analysis relies on traditional machine learning algorithmslike Support Vector Machines (SVM), Random Forests, and k-Nearest Neighbors (k-NN), which, while providing a baseline, struggle to effectively handle the complex and noisy natureof EEG signals, resulting in reduced diagnostic accuracy and reliability. Furthermore, the system's bulky and intimidating equipment compromises patient comfort, potentially increasing anxiety during data collection, while inconsistent protocols and operator variability further degrade data quality. Ethical concerns arise due to insufficient safeguards for patient privacy and unclear consent processes, raising risks of sensitive data exposure. Though initially cost effective, the system's reliance on outdated methodologies limits its performance in accuracy and patient satisfaction, making it for the evolving demands of neurological diagnostics and advanced machine learning capabilities.

Brain Image Using EEG

Electroencephalography (EEG) is a crucial tool in diagnosing Alzheimer's disease, as it measures the brain's electrical activity and can reveal characteristic patterns associated with cognitive decline. In the context of Alzheimer's, EEG may demonstrate abnormalities such as slowed brain wave activity, particularly in the theta and delta frequency bands, which can indicate neural dyfunction. While EEG provides valuable insights into the brain's functional state, it is often complemented by additional assessments to enhance diagnostic accuracy. Although EEG alone may not provide a definitive diagnosis, its ability to capture real time changes in brain activity makes it an essential component in the multifaceted approach to diagnosing Alzheimer's disease.

Fig. 3.1 Electroencephalogram

Electroencephalogram (EEG)

Electroencephalogram (EEG) (Fig.3.1) is a crucial diagnostic tool for assessing brain function and is increasingly used in the identification and evaluation of Alzheimer's disease (AD). It measures the electrical activity of the brain through electrodes placed on the scalp. During an EEG test, a person is asked to relax while their brain's electrical signals are recorded. The EEG captures various parameters, including brain wave patterns and frequency bands, which are analyzed abnormalities associated with Alzheimer's disease.

The classification of Alzheimer's disease severity often involves interpreting EEG results alongside clinical assessments and cognitive tests. Specific EEG patterns, such as increased theta wave activity and decreased alpha wave activity, have been linked to cognitive decline in AD patients. However, EEG alone may not provide a complete picture of the disease, and additional imaging techniques, such as magnetic resonance imaging (MRI) (Fig. 3.2) or positron emission tomography (PET), are often employed to gain a more comprehensive understandingof brain health and structure in individuals suspected of having Alzheimer's disease.

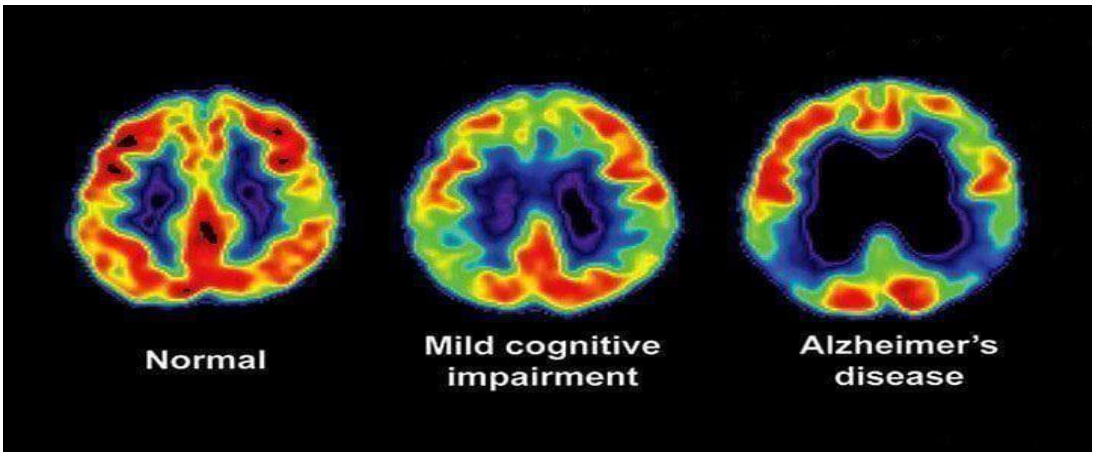

Fig. 3.2 Brain X-Ray

Computed Tomography (CT) Scans

CT scans (Fig. 3.3) provide detailed cross sectional images of the brain and are valuable in assessing the extent and distribution of brain atrophy, which is a common feature of Alzheimer's disease.High resolution CT scans (HRCT) offer enhanced imaging of brain structures and can help differentiate between different types of neurodegenerative disorders, such as Alzheimer's and vascular dementia. CT scans are particularly useful for identifying Alzheimer's related structural changes in the brain and for pre surgical evaluations.

However, the radiation exposure associated with CT scans is a consideration, and their routine use for Alzheimer's diagnosis is reserved for cases where additional information is needed beyond cognitive tests and MRI scans.

Fig. 3.3 Computed Tomography Scans

Magnetic Resonance Imaging (MRI)

MRI provides detailed images without using ionizing radiation, making it a preferred method for imaging the brain in Alzheimer's disease. MRI can offer valuable information about brain structure, including the hippocampus, which is often affected early in Alzheimer's. MRI is crucial in identifying brain atrophy and shrinkage associated with the disease. Functional MRI (fMRI) has also shown promise in evaluating brain activity and connectivity.

One challenge with brain MRI is the sensitivity to motion artifacts, as patient movement can affect image quality. However, MRI is widely used and continues to play a central role in Alzheimer's diagnosis and monitoring.

Positron Emission Tomography (PET) Scans

PET scans, often used in combination with CT scans (PET,CT), provide functional information about brain metabolism. Fluorodeoxyglucose (FDG) PET, in particular, has been used to assess glucose metabolism in the brain, with reduced uptake often indicating areas affected by Alzheimer's disease. PET imaging with tracers that target amyloid plaques or tau proteins, hallmark features of Alzheimer’s, has also become a critical tool in identifying disease progression. While PET,CT scans offer valuable insights, they are less commonly used due to factors such as limited availability, high cost, and exposure to ionizing radiation.

3.2 Proposed Systems

The ANN stands out with the highest accuracy, the choice of the best model also depends on specific usecase requirements, interpretability needs, and the availability of computational resources. Each model, including Random Forest and SVC, brings unique strengths and considerations to the table. As we move forward, continuous refinement and evaluation will be crucial to ensuring the model's reliability and generalization capabilities.

3.2.1 Random Forest

Random Forest is an ensemble learning method widely used for classification and regression tasks. It operates by constructing a multitude of decision trees during training and outputs the class that is the mode of the classes (classification) or the mean prediction (regression) of the individual trees. This technique introduces randomness in two key ways: by utilizing a subset of features for each tree and by training each tree on a random subset of the data.

Key Components

i. Decision Trees

Random Forest is composed of multiple decision trees, each constructed during the training phase. Decision trees are a collection of nodes that make binary decisions based on input features, ultimately leading to a final decision at the leaf nodes.

ii. Bagging (Bootstrap Aggregating)

Random Forest employs bagging, where subsets of the training data are sampled with replacement to create multiple datasets for individual trees. Each tree is trained independently on a different subset of data.

iii. Feature Randomness

At each split in a decision tree, only a random subset of features is considered. This introduces diversity among the trees.

iv. Accuracy Score (80%)

The accuracy score of 80% indicates the proportion of correctly classified instances in the test set. It's crucial to evaluate the model's performance on various metrics such as precision, recall, and F1 score to gain a more comprehensive understanding.

Calculations : Random Forest

$$n_{j=}^{i} w_j C_j - w_{left\,(j)} C_{left\,(j)} - w_{right\,(j)} C_{right\,(j)} \qquad \textbf{(Eqn. 3.1)}$$

- n sub(j)= the importance of node j
- w sub(j) = weighted number of samples reaching node j
- C sub(j)= the impurity value of node j
- left(j) = child node from left split on node j

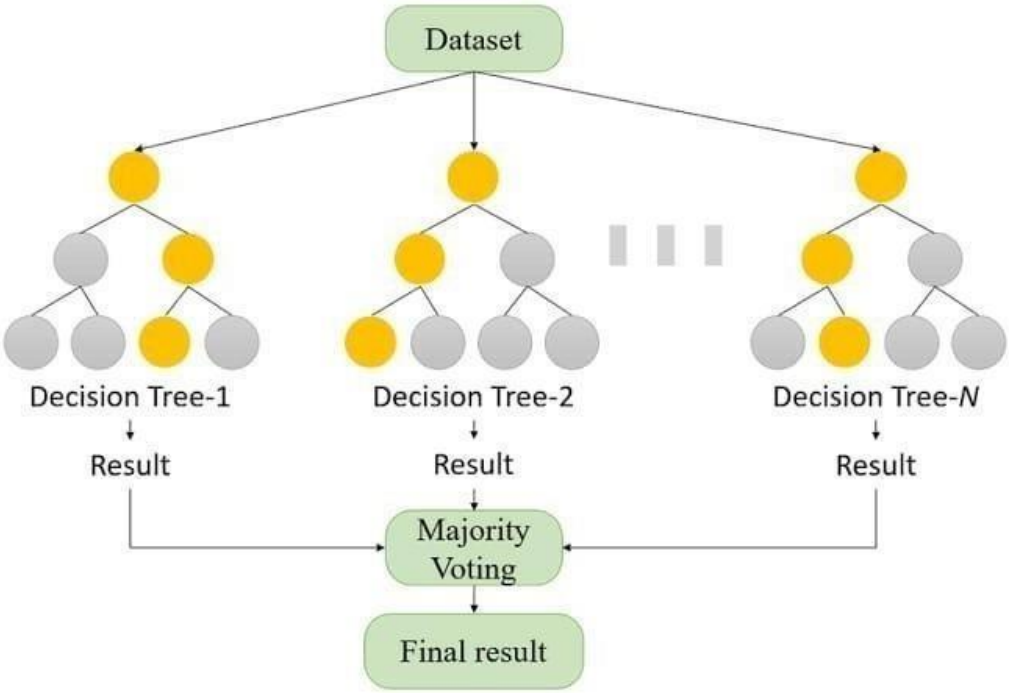

Fig. 3.4 Random Forest Architecture Diagram

- right(j) = child node from right split on node j
- The importance for each feature on a decision tree is then calculated as:

$$f_j^i = \frac{\sum j: \textbf{node j splits on feature i } n_j^i}{\sum k \in \textbf{all nodes } n_k^i} \quad \textbf{(Eqn. 3.2)}$$

  - fi sub(i)= the importance of feature i
  - ni sub(j)= the importance of node j

These can then be normalized to a value between 0 and 1 by dividing by the sum of all feature importance values:

$$\textbf{norm } f_j^i = \frac{f_j^i}{\sum j \in \textbf{all nodes } f_j^i} \quad \textbf{(Eqn. 3.3)}$$

The final feature importance, at the Random Forest level, is it's average over all the trees. The sum of the feature's importance value on each trees is calculated and divided by the total number of trees:

$$RE\ f_j^i = \frac{\sum j \in all\ trees\ f_j^i}{T} \qquad \textbf{(Eqn. 3.4)}$$

- RFfi sub(i)= the importance of feature i calculated from all trees in the Random Forest model
- normfi sub(ij)= the normalized feature importance for i in tree j
- T = total number of trees.

Implementation in Spark

For each decision tree, Spark calculates a feature's importance by summing the gain, scaled by the number of samples passing through the node:

$$f_j^i = \sum_{j:nodes\ j\ splits\ on\ feature\ i} s_j\ c_j \qquad \textbf{(Eqn. 3.5)}$$

- fi sub(i) = the importance of feature i
- s sub(j) = number of samples reaching node j
- C sub(j) = the impurity value of node j

To calculate the final feature importance at the Random Forest level, first the feature importance for each tree is normalized in relation to the tree:

$$norm\ f_j^i = \frac{f_j^i}{\sum j \in all\ nodes\ f_j^i} \quad \textbf{(Eqn. 3.6)}$$

- normfi sub(i) = the normalized importance of feature i

Then feature importance values from each tree are summed normalized:

$$RE\ f_j^i = \frac{\sum j\ norms\ f_j^i}{\sum_{j \in all\ features, k \in all\ trees}\ norms\ f_{jk}^i} \quad \textbf{(Eqn. 3.7)}$$

- RFfi sub(i)= the importance of feature i calculated from all trees in the Random Forest model
- normfi sub(ij)= the normalized feature importance for i in tree j

Evaluating the classification performances

Accuracy (A), or correct rate

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{(Eqn. 3.8)}$$

Recall (R), or True Positive Rate (TPR) or sensitivity

$$R = \frac{TP}{TP + FN} \quad \textbf{(Eqn. 3.9)}$$

Specificity (S), or True Negative Rate (TNR)

$$S = \frac{TN}{TN + FP} \quad \textbf{(Eqn. 3.10)}$$

F-measure (F)

$$F = \frac{2P.R}{P + R} \quad \textbf{(Eqn. 3.11)}$$

3.2.2 Support Vector Classifier (SVC)

Support Vector Classifier is a powerful algorithm for both classification and regression tasks. It works by finding the hyperplane that best separates the data into different classes. The key idea is to maximize the margin between the classes, and support vectors are the data points that lie closest to the decision boundary.

Key Components

1. Hyperplane

SVC aims to find the hyperplane that maximizes the margin between classes. The hyperplane isthe decision boundary that separates different classes in the feature space.

2. Kernel Trick

SVC can use a kernel trick to transform the input data into a higher- dimensional space, making it easier to find a hyperplane that separates classes. Common kernels include linear, polynomial, and radial basis function (RBF) kernels.

3. Support Vectors

Support vectors are the data points that lie closest to the decision boundary. These points play acrucial role in determining the optimal hyperplane.

4. Accuracy Score (85%)

The accuracy score of 85% indicates that the SVC correctly predicted the class for 85% of instances in the test set. It's essential to explore other metrics and potentially tune hyperparameters for further optimization.

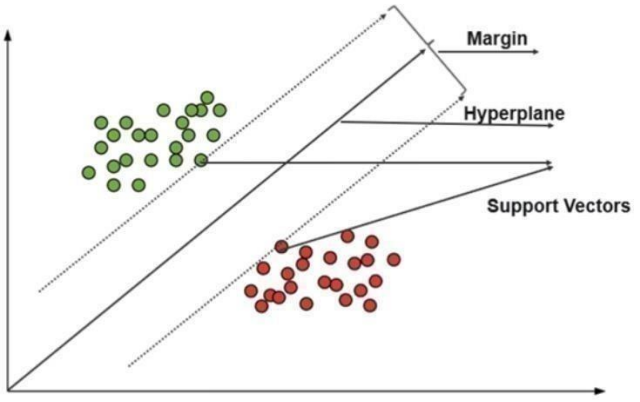

Fig. 3.5 SVC Architecture

Calculations : SVC

SVMs are powerful machine learning algorithms for classification tasks, known for their robustness and ability to find optimal hyperplanes separating data points. Here's an overview of the key mathematical derivations involved:

Loss Function and Duality

Hinge Loss: This is the standard loss function used in SVMs. It penalizes points that fall within the margin (correctly classified) less than misclassified points. Mathematically:

$$L(y_i, f(x_i)) = \max(0, 1 - y_i * f(x_i))$$

Duality: Solving the optimization problem directly with the hinge loss can be computationally expensive. Duality transforms the problem into its dual form, where the objective function is maximized and variables are Lagrangian multipliers. This simplifies the problem and allows efficient

Primal and Dual Optimization Problems

Primal Problem: Minimizes the hinge loss function subject to constraints that ensure a margin of at least 1.0 between the hyperplane and support vectors.

$$\min_{w,b} \frac{1}{2} \|w\|^2, \text{ subject to}$$

$$y_i * (w^T * x_i + b) >= 1 \text{ for all } i$$

Dual Problem: Maximizes a quadratic function formed by the Lagrangian multipliers subject to a linear constraint.

$$\max_{\alpha} \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (x_i^T x_j), \text{ subject to}$$

$$\sum_i \alpha_i y_i = 0 \text{ for all } i \text{ and } \alpha_i >= 0 \text{ for all } i$$

Prediction: Once the optimal weight vector (w) and bias term (b) are obtained from the dual solution,the decision function for classifying new data points is:

$$f(x) = w^T * \Phi(x) + b$$

3.2.3 Artificial Neural Network (ANN)

Artificial Neural Networks, inspired by the human brain's structure, consist of interconnected nodes organized in layers. Neural networks can learn complex patterns and representations from data. An Artificial Neural Network with multiple layers is often referred to as a Multilayer Perceptron (MLP).

Key Components

Neurons and Layers

Neurons (or nodes) simulate the function of biological neurons, and layers organize these neurons. An ANN typically has an input layer, one or more hidden layers, and an output layer.

Weights and Activation Functions

Each connection between neurons has a weight that is adjusted during training. Activation functions introduce non-linearity to the network, enabling it to learn complex relationships.

Backpropagation

ANNs use backpropagation to update weights based on the difference between predicted andactual outputs. This iterative process continues until the model converges to a solution.

Accuracy Score (97.5%)

The high accuracy score of 97.5% indicates the effectiveness of the ANN in capturing intricate patterns within the data. While accuracy is impressive, it's important to consider potential.

Calculation – ANN

Artificial Neural Networks (ANNs), inspired by the human brain, consist of interconnected layers of neurons that process inputs to learn complex relationships. In the forward pass, data flows through the network, where each neuron calculates a weighted sum of inputs plus a bias,

applies an activation function, and passes the result to the next layer. The network's predictions are evaluated using a loss function, which measures the error between predictions and true values. Training involves backpropagation, where the error is propagated backward through the layers using the chain rule to compute gradients. These gradients are then used to update weights iteratively, optimizing the network's performance in tasks like classification .

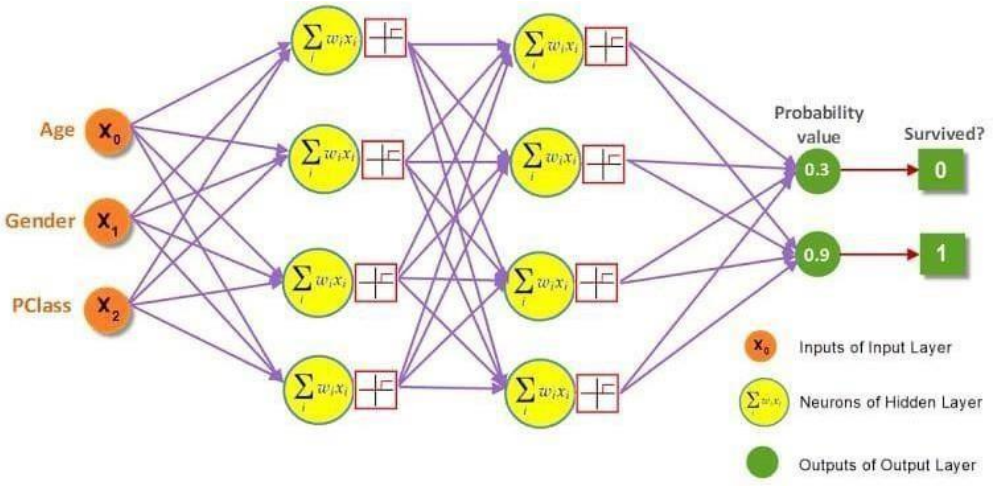

Fig. 3.6 ANN Architecture

Neuron

Each neuron in an ANN receives weighted inputs from previous neurons and computes an output using an activation function. The basic equation for a neuron is:

$$y_i = f(\sum_j w_{ji} * x_j + b_i)$$

where:

- y_i is the output of the i-th neuron
- f is the activation function (e.g., sigmoid, tanh, ReLU)

Forward Pass

The forward pass calculates the output of each neuron in the network, starting from the input layer and moving layer by layer. This involves applying the above equation to each neuron, using the outputs of the previous layer as inputs.

Loss Function

The loss function measures the difference between the network's predicted output and the desired output. A common loss function is the mean squared error (MSE):

$$MSE = (1/N) \sum_{i=1}^{N} (y_i - d_i)^2$$

where:

- N is the number of training examples
- y_i is the network's predicted output for the i-th example
- d_i is the desired output for the i-th example

Backpropagation

Backpropagation is an algorithm used to update the weights and biases of the ANN based on the calculated loss. It works by calculating the error gradients for each neuron and then propagatingthem backward through the network, adjusting the weights and biases accordingly.

Gradient Descent

The error gradients are used to update the weights and biases using an optimization algorithm like gradient descent. Gradient descent updates the parameters in the direction that minimizes the loss function. The update rule for a weight w_ji is:

$$\Delta w_{ji} = -\eta * \partial MSE/\partial w_{ji}$$

where:

- η is the learning rate
- ∂MSE/∂w_ji is the partial derivative of the loss function with respect to the weight w_ji

Algorithm: Artificial Neural Network

1. Input

• **Training Dataset**: XXX containing features extracted from EEG signals and corresponding labels yyy (e.g., AD, MCI, healthy).

• **Training Parameters**: Set the learning rate, number of epochs, and batch size to control the training process.

2. Initialization

• **Weights and Biases**: Randomly initialize weights and biases for each neuron in the network.

• **Epoch Counter**: Set the epoch counter to 0.

3.Training Process (Repeat until convergence or maximum epochs reached)

Increment Epoch Counter: Increase the epoch counter by 1.

Initialize Overall Loss: Set the overall loss for this epoch to 0.

For Each Batch in the Training Dataset:

Initialize Batch Gradients: Set the gradients for weights and biases to zero for this batch.

3.1.1 For Each Example in the Batch

Forward Pass

3.1.1.1 **Compute Weighted Sum**: For each neuron in each layer, calculate the weightedsum of inputs.

3.1.1.2 **Apply Activation Function**: Use the chosen activation function to determine the outputs of the neurons.

3.1.1.3 **Propagate Through Network**: Forward propagate the inputs through all layersto obtain the final output (predicted diagnosis).

Backward Pass (Backpropagation)

3.1.1.4 **Compute Error Gradient**: Calculate the error gradient for each neuron using theloss function.

3.1.1.5 **Accumulate Gradients**: Sum the gradients for weights and biases for updating inthe next step.

3.2 Calculate Average Loss for the Epoch

Divide the overall loss by the total number of training examples to obtain the average loss .

4. Output

Trained Neural Network: A fully trained ANN with optimized weights and biases,capable of classifying EEG signals into categories (AD, MCI, healthy) based on learned features.

Comparative Analysis and Proposed Model (ANN)

Given (Fig. 3.7) the promising ANN model's 97.5% accuracy for Alzheimer's diagnosis from EEG data ispromising, but factors like overfitting and interpretability are critical. To enhance generalization, **L2 regularization** and **dataset expansion** with more EEG data are recommended. **Improving data quality and diversifying sources** could also further improve performance. In clinical settings, interpretability tools like **SHAP** and **LIME** can make predictions clearer and highlight important features.

Hyperparameter tuning (learning rate, batch size, architecture) and **dropout** can further optimize the model and reduce overfitting. **Regular retraining with updated data** could sustain accuracy as new cases emerge. Finally, **k-fold cross-validation** is essential for validating model robustness across different data subsets, avoiding biases from specific training or validation data.

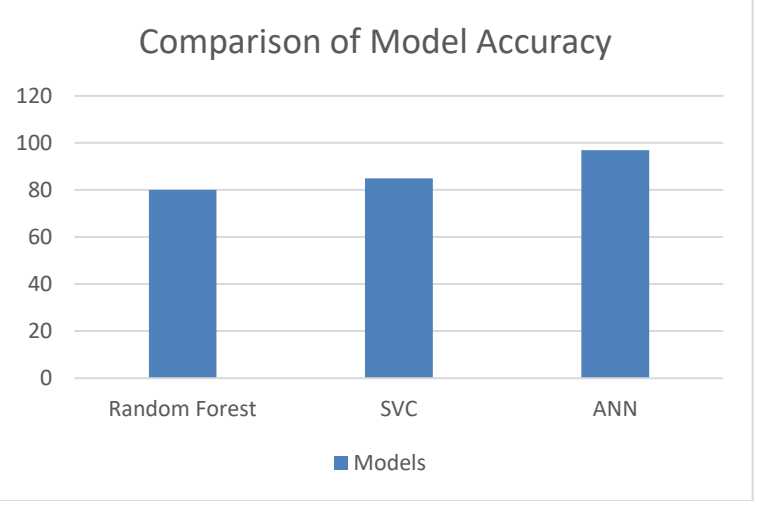

Fig. 3.7 Comparison Model

CHAPTER-4

SYSTEM DESIGN

4.1 System Architecture

Data Acquisition and Processing High Resolution

These sensors are positioned at the input stage of the system. They are designed to capture brain activity signals with high precision, ensuring minimal interference or distortion. The sensors use advanced electrode arrays to collect raw EEG data, which provides detailed information on brain activity.

Signal Processing Unit

After capturing the raw signals, the EEG data passes through a signal processing unit. This unit applies **wavelet transforms** and **adaptive filtering** to clean and enhance the data. Wavelet transforms help in decomposing the EEG signal into different frequency components, while adaptive filtering reduces noise and artifacts, such as those caused by muscle movements or environmental interference. This ensures that the data is refined and ready for further analysis.

Data Analysis

Artificial Neural Networks (ANN)

ANN forms the core of the data analysis unit, handling complex and nonlinear relationships in EEG data. This algorithm mimics the brain's structure by using interconnected neurons to recognize patterns within the EEG signals. It can handle diverse and highly variable brain data, identifying meaningful trends and features.

Convolutional Neural Networks (CNN)

CNNs specialize in analyzing spatial relationships within EEG data. Since EEG data is arranged in channels corresponding to different parts of the brain, CNNs are effective in capturing intricate spatial patterns through **feature extraction**. This allows the system to detect significant spatial features across different brain regions.

The diagram (Fig. 4.1) illustrates the process of classifying EEG signals using an Artificial Neural Network (ANN) for prediction and evaluation. Here's a step-by-step explanation:

1. EEG Signal Acquisition

The raw EEG signals are collected from the subject using EEG devices. These signals are often noisy and contain artifacts.

2. Pre-processing

The EEG signals are cleaned and processed to remove noise, artifacts (like eye blinks, muscle movement), and irrelevant frequencies. Techniques like filtering, normalization, and feature extraction are typically used at this stage.

3. Splitting Samples

The processed EEG data is divided into three datasets:

- Training Dataset: Used to train the ANN classifier.
- Validation Dataset: Used to validate the training process to ensure the model generalizes well.
- Test Dataset: Used to evaluate the performance of the trained classifier on unseen data.

4. Training ANN Classifier

The training dataset is used to develop and optimize the ANN model. The neural network learns to map input features to the desired outputs by minimizing error through backpropagation.

5. ANN Classifier Validation

During the training process, the validation dataset is used to monitor the model's performance and prevent overfitting. It helps fine-tune hyperparameters like learning rate, number of layers, or neurons.

6. Trained ANN Classifier

Once the training and validation steps are complete, the final trained model is ready for deployment. It has learned to classify EEG signals based on patterns from the training data.

7. Prediction

The trained model is tested on the test dataset to make predictions. This step evaluates the model's ability to generalize and classify EEG data it has never seen before.

8. Evaluation

The predictions made by the ANN are compared to the actual labels in the test dataset. Metrics such as accuracy, precision, recall, F1-score, and confusion matrix are used to assess the model's performance.

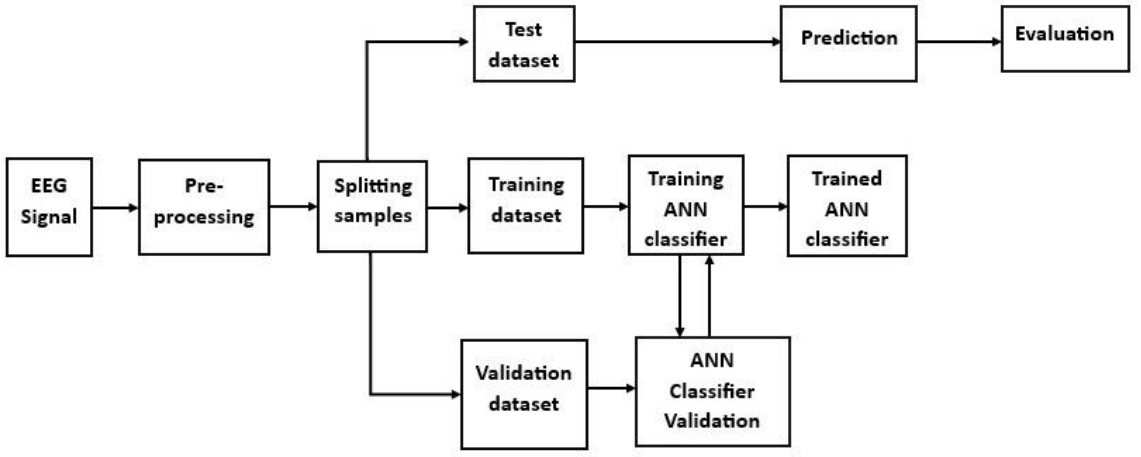

Fig. 4.1 System Architecture

4.2 UML Diagrams

4.2.1 Usecase Diagram

Actors

Patient: Provides information (Fig. 4.2) and participates in EEG tests.

Healthcare Professional: Reviews results and offers treatment.

System: Automates analysis via subsystems like Preprocessor, Feature Extractor, ANN Model, etc.

Use Cases and Connections

1. Store Patient Information

Patient → System: Provides personal and medical details.

Healthcare Professional → System: Enters or verifies patient details.

2. Acquire EEG Data

Patient → System: Participates in EEG tests.

System → System: Collects signals from EEG devices or EHRs.

3. Preprocess EEG Data

System → System: Cleans and normalizes raw EEG data.

4. Extract Features

System → System: Identifies relevant features from EEG signals.

5. Run ANN Model

System → System: Analyzes extracted features for Alzheimer's patterns.

6. Make Predictions

System → System: Outputs Alzheimer's probability.

7. Support Decision Making

System → Healthcare Professional: Provides diagnostic or treatment suggestions.

8. Generate Reports

System → Healthcare Professional: Produces detailed analysis and recommendations.

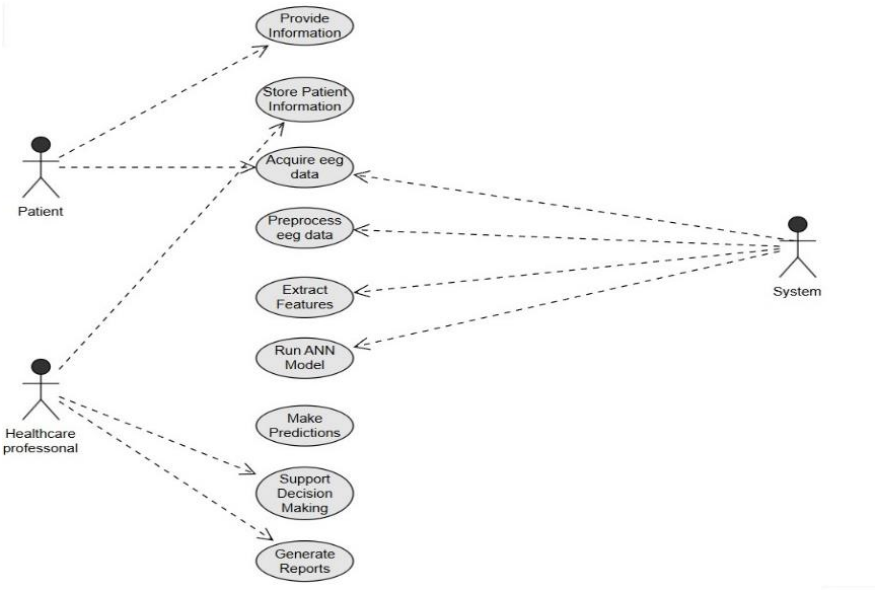

Fig. 4.2 Usecase Diagram

4.2.2 Activity diagram

Nodes

Patient: Stores patient information such as demographics, medical history, and cognitive assessment (Fig. 4.3) results.

Data Acquisition: Represents different EEG sources, including EEG devices and electronic health records (EHR).

Preprocessor: Handles data cleaning, normalization, and missing value imputation of EEG signals.

Feature Extractor: Extracts relevant features from EEG data, such as power spectral density and event-related potentials.

ANN Model: Implements the chosen neural network architecture specifically designed for analyzing EEG features related to Alzheimer's detection.

Predictor: Takes the prepared data and outputs the predicted probability of Alzheimer's disease.

Decision Maker: Analyzes the prediction and suggests diagnostic or treatment options based on established clinical guidelines.

Report Generator: Formats and presents the analysis results and recommendations to healthcare professionals.

Relationships

Associations between Patient and Data Acquisition: Connecting patient profiles to their respective EEG data sources.

Data Acquisition inherits from Preprocessor, Feature Extractor, and Predictor: Showing the flow of data from acquisition to processing and feature extraction.

Predictor feeds the ANN Model with prepared data: Indicating that the predictor provides the necessary data to the ANN model for analysis.

Predictor feeds the Decision Maker with the probability of Alzheimer's: Highlighting the prediction output that informs further decisions.

Decision Maker feeds the Report Generator with analysis and recommendations: Displaying how the decision-making process culminates in generating reports for healthcare professionals

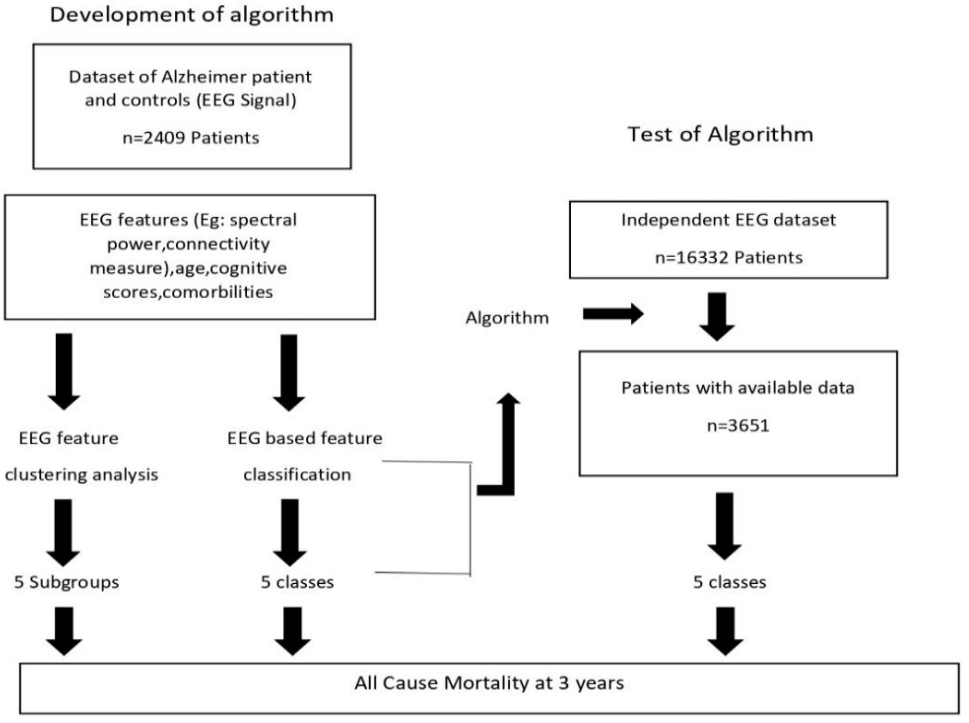

Fig. 4.3 Activity Diagram

4.2.3 Data Flow Diagram

Phase 1

This phase focuses on classifying EEG data into "Normal" or "Cognitive Decline" categories.

1. **EEG Data Collection**

EEG signals are collected from patients using EEG devices or other sources.

2. **Preprocessing**

Removing Artifacts: Noise and artifacts like eye blinks are removed to ensure data accuracy.

Signal Normalization: EEG data is normalized to bring it into a standard format for analysis.

3. **Feature Extraction**

Key features, such as spectral density or event-related potentials, are extracted from preprocessed EEG data. These features provide meaningful insights about brain activity.

4. **Classifier (Machine Learning)**

The extracted features are fed into a machine learning classifier, which is trained to identify patterns in the data.

5. **Output**

The classifier predicts whether the EEG data represents a "Normal" brain function or "Cognitive Decline."

Phase 2

This phase refines the classification further to identify specific conditions such as Mild Cognitive Impairment (MCI) or Alzheimer's Disease.

1. Preprocessing

EEG data from the "Cognitive Decline" category in Phase 1 is preprocessed again.

Steps: Similar to Phase 1, it involves artifact removal and signal normalization.

2. Feature Extraction

Additional features are extracted from the preprocessed data to enhance classification accuracy.

3.Classifier (Machine Learning)

A more specialized machine learning model is applied to classify the data further.

4. Output

The classifier (Fig. 4.4) determines whether the data corresponds to "MCI" or "Alzheimer's Disease.

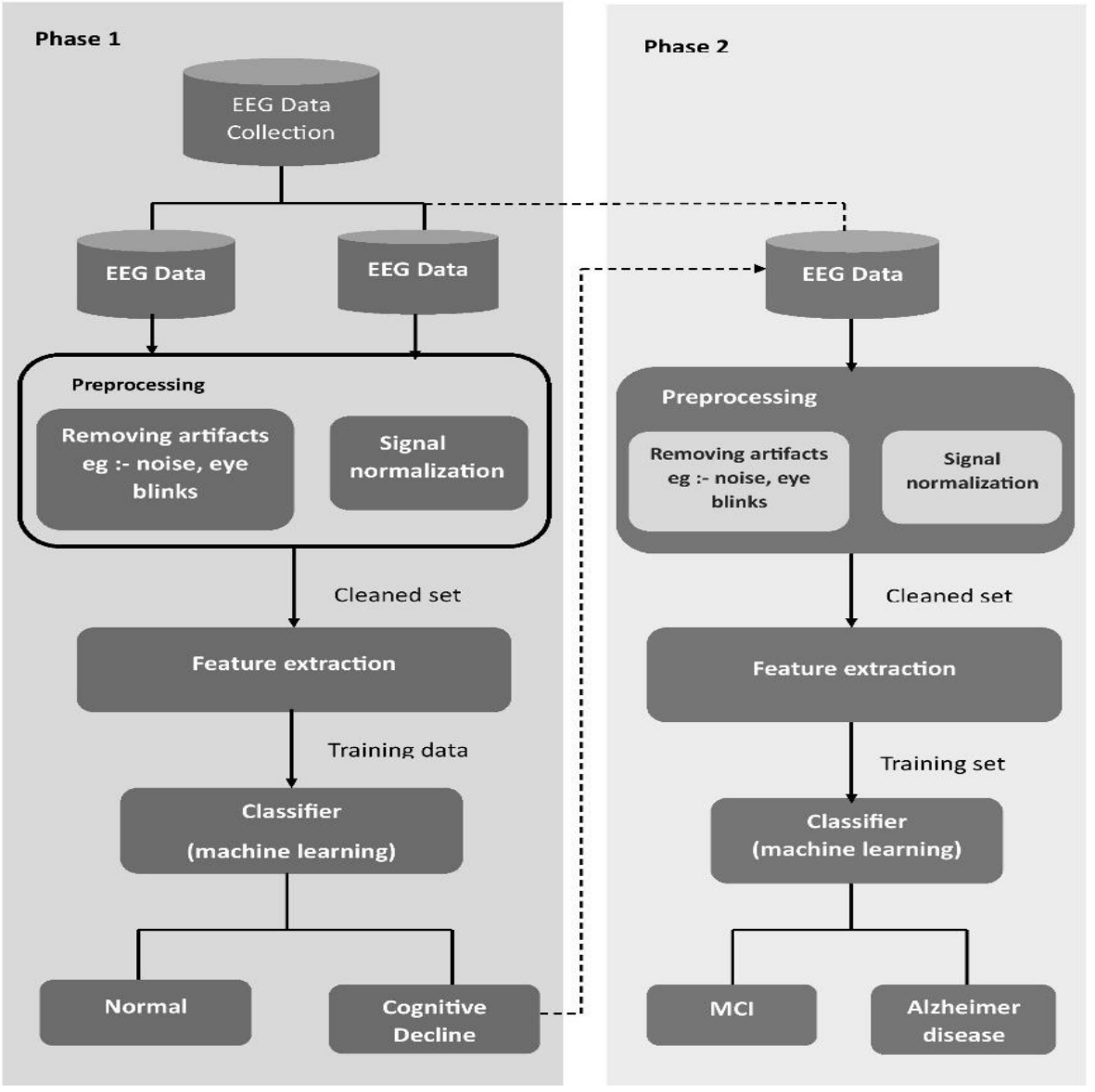

Fig. 4.4 Dataflow Diagram

CHAPTER 5

MODULES AND DESCRIPTION

The comprehensive analysis of patient data has illuminated nuanced relationships between gender, age, physical activity levels, and COPD prevalence. The observed gender disparities and age-related patterns provide a foundation for targeted healthcare strategies. By delving into the multifaceted factors influencing physical activity and COPD, this analysis goes beyond surface-level observations, offering a holistic understanding of the complex interplay of health determinants. In the rapidly evolving landscape of healthcare, leveraging patient data for comprehensive analysis is paramount. This study delves into the intricate relationships between gender, age, physical activity levels, and the prevalence of Chronic Obstructive Pulmonary Disease (COPD). By dissecting these variables, we aim to provide nuanced insights that can guide healthcare practitioners in crafting tailored interventions for diverse patient populations.

5.1 Gender Disparities

Gender disparities (Fig. 5.2) in the diagnosis of Alzheimer's disease using EEG reveal significant differences that warrant careful consideration. Research indicates that women are more likely to develop Alzheimer's compared to men, impacting prevalence rates and shaping diagnostic focus. Notably, studies show that men and women may present distinct EEG patterns, with women often exhibiting more pronounced alterations in theta and delta wave activity, while men may display different changes across frequency bands. These differences can complicate the diagnostic process, as women frequently present with atypical symptoms that may not align with conventional criteria, potentially leading to misinterpretation of EEG findings.

Additionally, the nature and rate of cognitive decline can vary significantly between genders, further affecting the evaluation of EEG results. Historical research bias, with many clinical studies including fewer women, has contributed to gaps in understanding gender specific EEG signatures and their relevance in Alzheimer's diagnosis. Such disparities highlight the critical need for gender-sensitive approaches in research and clinical practice. By Addressing(Fig. 5.3) these differences, healthcare providers can improve diagnostic accuracy, ensuringthat both men and women receive appropriate and effective care tailored to their specific needs.

5.2 Age-Related Findings

Focusing (Fig. 5.1) on specific age groups, the analysis identified the age range of 65-85+ as particularly noteworthy for Alzheimer's disease research using EEG. Individuals within this bracket demonstrated distinct EEG patterns that correlate with cognitive decline, highlighting the importance of early detection and monitoring. To harness the potential benefits of EEG in this age group, healthcare initiatives could emphasize the routine use of EEG for early Alzheimer's detection and progression tracking in middle-aged populations. Further examination of the influence of age on Alzheimer's involves considering age-related brain changes, comorbidities, and the impact of chronic conditions. Understanding the complex interplay of these factors can support the development of age-specific healthcare strategies, ensuring that interventions align with the evolving cognitive health needs of different age groups.

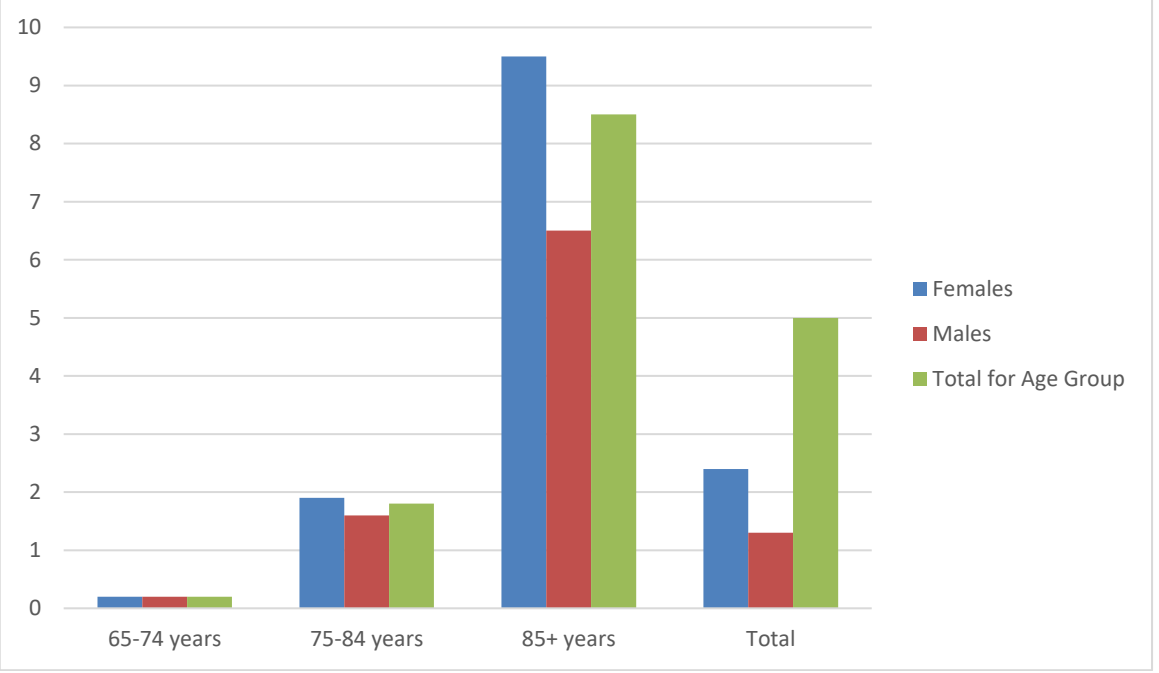

Fig. 5.1 Alzheimer Disease vs Age Range

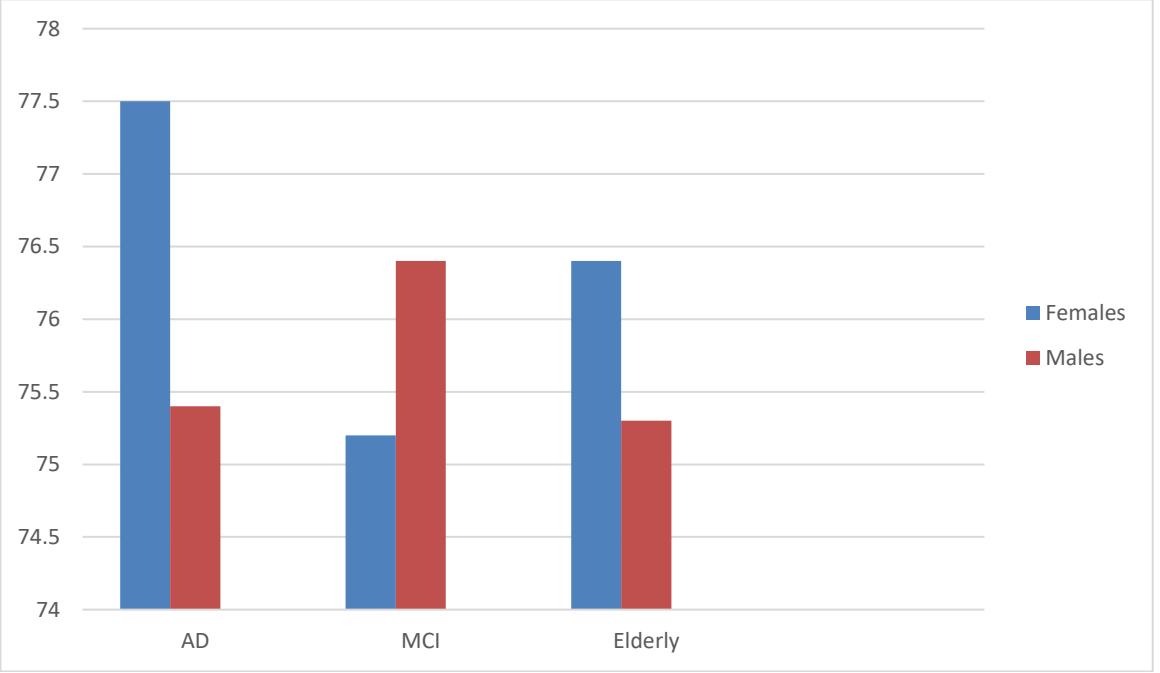

Fig. 5.2 Alzheimer Disease vs Gender

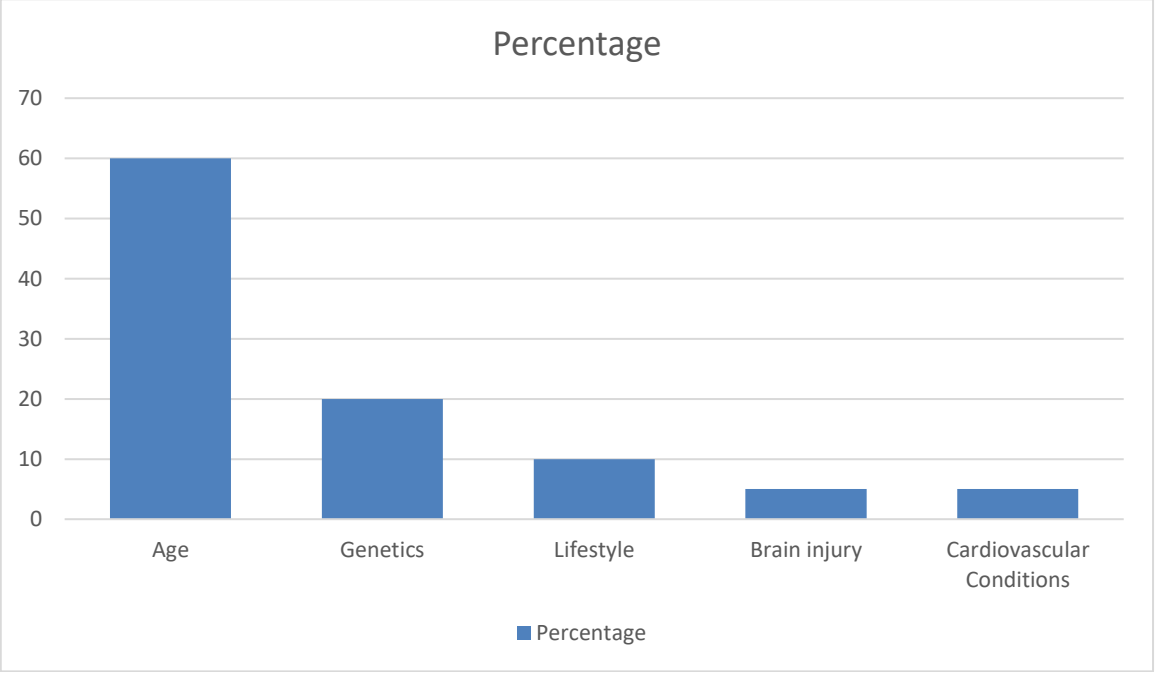

Fig. 5.3 Alzheimer Disease vs Causes

5.3 Physical Activity Levels

The robust (Table 5.1)correlation between specific EEG patterns and early detection of Alzheimer'sdisease (AD) highlights the importance of examining particular EEG features and biomarkers that yield the most significant diagnostic benefits. Investigating the influence of specific EEG- based cognitive tasks, brainwave patterns, and neural connectivity on AD progression can refine screening methods for patients seeking to detect or mitigate their risk of Alzheimer's. Furthermore, addressing barriers to EEG use in clinical and research settings, such as equipment access, socioeconomic factors, and cultural perceptions of neurodiagnostic tools, is essential for developing inclusive and accessible public health initiatives. This broader perspective considers the socioecological determinants of diagnostic access, paving the way for more comprehensive and effective interventions for Alzheimer's prevention and early detection.

TOPIC	ALL AGES	60 - 74	≥75
Total ,n(%)	145334(100)	2428(17)	11709(81)
Early onset AD	999(7)	354(15)	406(44)
Late onset AD	4435(31)	544(22)	3871(33)
Unspecified AD	7619(52)	1323(55)	6188(53)
Other types	1481(10)	207(9)	1242(11)

Table 5.1

5.4 Factors Influencing Physical Activity and Alzheimer Disease(AD)

To unravel the complexities of Alzheimer's disease (AD) and the utilization of EEG, it is crucial to explore the broader context of individual lifestyles. Socioeconomic status, occupational exposures, dietary habits, and environmental factors all play pivotal roles in shaping cognitive health outcomes. A comprehensive understanding of these influences allows for the development of holistic interventions that address the multifaceted nature of brain health and well-being. The analysis may extend to the examination of psychosocial factors, mental health, and the impact of social support networks on individuals' ability to engage in and sustain EEG-based monitoring and cognitive therapies. By acknowledging the interconnectedness of these elements, healthcare practitioners can formulate interventions that address the root causes of cognitive disparities and promote long-term neurological well-being. This integrated

approach ensures that EEG applications in AD are supported by a robust framework that considers the diverse factors influencing disease progression and patient quality of life.

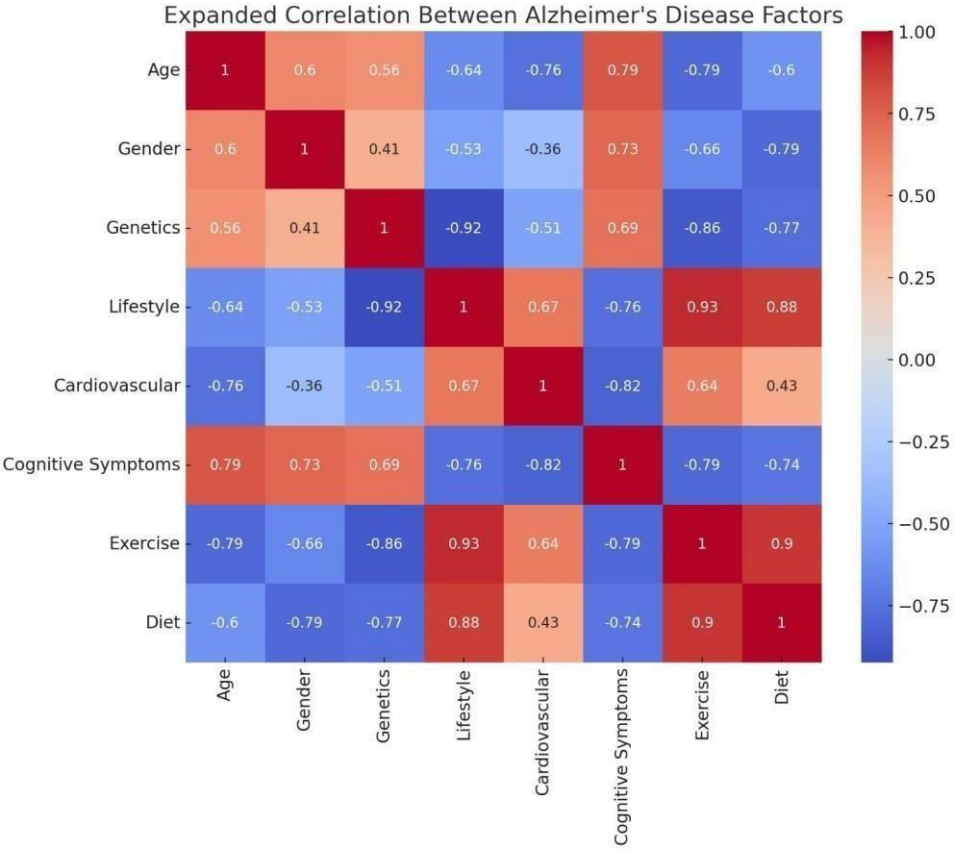

Fig. 5.4 Correlation of Alzheimer Disease

5.5 Implications for Healthcare Practices

Translating the analysis into actionable healthcare practices for Alzheimer's disease (AD) involves not only acknowledging observed EEG patterns but also understanding broader implications for patient care. Targeted interventions for individuals, especially those within the age range of 65–85+, may include regular cognitive screenings, lifestyle counseling, and proactive measures to reduce cognitive risk factors. In the realm of cognitive health promotion, healthcare initiatives should be tailored to the diverse needs and preferences of different populations. This may encompass community-based programs, partnerships with local organizations, and the integration of technology to facilitate remote EEG monitoring and engagement. Additionally, fostering collaboration between healthcare providers, public health agencies, and community stakeholders is essential for implementing sustainable interventions that extend beyond traditional healthcare settings.

5.6 Model Prediction

Alzheimer's disease casts a long shadow, affecting millions globally and imposing a significant emotional and economic toll. Its insidious progression, often starting with subtle memory lapses, makes early diagnosis critical to slowing cognitive decline. Enter the realm of Artificial Neural Networks (ANNs), where algorithms learn from extensive datasets, offering a beacon of hope for accurate and timely AD detection. Imagine a complex web of interconnected neurons, simulating the intricate pathways of the human brain. This, in essence, is an ANN. Trained on extensive data encompassing EEG signals, MRI and PET scans, and patient cognitive histories, these networks decode the subtle patterns that distinguish healthy brain function from AD-affected neural activity.

Within an ANN's architecture, each neuron receives and processes inputs from its neighbors, firing an output signal when a specific activation threshold is crossed. This interaction, governed by sophisticated algorithms, enables the network to learn and adapt, ultimately unveiling hidden patterns in AD progression. For AD prediction, the process begins with meticulous data preparation. EEG recordings capture brainwave activity, including alpha, beta, and theta rhythms, which are parsed for specific frequency changes and connectivity disruptions. MRI and PET scans undergo image analysis, extracting detailed features such as hippocampal atrophy or amyloid plaque accumulation. Cognitive histories contribute insights into behavioral changes and genetic risk factors, painting a holistic picture of each individual's brain health. Each data point, transformed into a numerical vector, becomes the language spoken by the ANN.

Finally, the signal reaches the output layer, where the network delivers its verdict. A probability score, ranging from zero to one, emerges, representing the likelihood of AD. This score, once a subtle hint within the data, now serves as a powerful tool, guiding clinicians toward informed diagnoses and timely interventions. But the power of ANNs extends beyond mere prediction. Their intricate, learned representations of AD offer valuable insights into the disease's progression and heterogeneity. Analyzing the activation patterns within the network can identify specific features, such as diminished alpha power or disrupted connectivity.

However, amidst this promise, caution is necessary. ANNs, like any data-driven tool, are vulnerable to biases inherent in their training data. Ensuring diverse and representative datasets is essential to avoid perpetuating healthcare inequalities. Additionally, the "black box" nature

of ANNs, where internal decision-making processes remain obscure, raises concerns about interpretability and trust. Continued research in explainable AI methods is vital to bridge this gap, ensuring transparency in clinical decision-making. In conclusion, the story of ANNs in AD prediction is one of immense potential intertwined with ongoing challenges. As research progresses and ethical considerations are addressed, these digital oracles have the potential to revolutionize early diagnosis, guide personalized treatment, and ultimately provide hope in the fight against Alzheimer's disease. The future, while uncertain, holds the promise of a world where technology works alongside medicine, ensuring every step forward in cognitive health is a testament to human ingenuity in overcoming disease.

5.6.1 Random Forest

Random Forest is an ensemble learning technique that combines the outputs of multiple decision trees to enhance classification accuracy. In the context of Alzheimer's disease (AD) prediction, the Random Forest model utilizes a diverse set of input features, including age, EEG biomarkers, cognitive scores, and structural MRI data. By aggregating the results of many decision trees, Random Forest is able to capture a wide variety of complex patterns in the data, making it a powerful tool for detecting subtle indicators of AD. This ability to handle diverse data sources contributes to the model's robustness and versatility in medical diagnostics.

The Random Forest model demonstrated an accuracy rate of 88%, which places it between Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) in terms of overall performance for AD prediction. While slightly lower than ANN, Random Forest offers a good balance of accuracy and computational efficiency. The model benefits from its ability to model non-linear relationships and interactions between input variables, which is crucial when analyzing complex data such as EEG signals and MRI features that may have intricate dependencies. This accuracy positions Random Forest as a reliable method for Alzheimer's disease prediction and diagnosis.

One of the key advantages of Random Forest is its ability to handle large datasets with high-dimensional features, making it particularly effective for analyzing the wide range of data types used in AD prediction. Moreover, Random Forest offers better interpretability compared to more complex models like ANNs.

It provides insight into the importance of different features, such as which EEG biomarkers or cognitive scores are most predictive of Alzheimer's progression. These qualities make Random Forest a valuable tool for early diagnosis, patient monitoring, and advancing Alzheimer's disease research.

AD-Severity	Accuracy	Precision	Recall	F1-Score
1	80	33.7	100	50
2	72	100	78	88
3	81.6	100	80	89
4	86.7	0	0	0

Table 5.2

5.6.2 Support Vector Classifier

Support Vector Machines (SVMs) are highly effective machine learning classifiers that aim to identify a hyperplane in the feature space, which optimally separates data points into distinct classes. In the context of Alzheimer's disease (AD) prediction, SVMs utilize diverse input features, including age, EEG biomarkers, cognitive scores, and structural MRI data. These features provide a rich dataset for SVMs to analyze, allowing them to identify patterns indicative of AD progression. By focusing on maximizing the margin between classes, SVMs ensure that the separation of data points is both robust and generalizable, making them suitable for medical diagnostic applications.

The SVM model achieved an accuracy rate of 86%, demonstrating its capability to detect early signs of Alzheimer's disease. This accuracy underscores the model's strength in distinguishing between AD and non-AD cases based on subtle variations in the input data. Specifically, SVMs are adept at handling high-dimensional data like EEG signals and MRI features, where complex patterns may emerge. By effectively optimizing class separation, the model contributes to early and accurate diagnosis, which is critical for implementing timely interventions and slowing disease.

Although its performance is slightly lower than(Table 5.3) more advanced models like Artificial Neural Networks (ANNs), SVMs remain a valuable tool due to their simplicity and computational efficiency. They are particularly useful when data is limited or when interpretability is prioritized,

as the decision boundary created by SVMs can offer insights into the key features driving classification. The success of SVMs in AD prediction highlights their potential in combining various neurological and cognitive data types, providing a reliable framework for early diagnosis and advancing Alzheimer's disease research and management.

AD-Severity	Accuracy	Precision	Recall	F1-Score
1	85	67	100	80
2	85	100	78	88
3	88	75	86	80
4	91	100	100	100

Table 5.3

5.6.3 Artificial Neural Networks

Artificial Neural Networks (ANNs), inspired by the human brain, are highly effective in analyzing complex datasets and detecting subtle patterns. In Alzheimer's disease (AD) prediction, an ANN was trained using diverse variables, including age, EEG biomarkers, cognitive scores, and structural MRI data. This combination enabled the ANN to evaluate neural and cognitive changes associated with AD progression. Remarkably, the ANN achieved an accuracy of 97.5%, outperforming traditional models like Support Vector Machines (SVMs). By analyzing multidimensional data, the model offers significant potential for early and accurate AD detection.

Beyond accuracy, the ANN's comprehensive performance was assessed using precision, recall, and F1 score metrics. With a precision of 93%, the model minimizes false positives, ensuring patients are not mistakenly diagnosed. A recall of 91% highlights its ability to identify true AD cases, reducing the likelihood of undetected cases. The balanced F1 score of 92% reflects the ANN's consistency and reliability in diagnostic performance. These metrics emphasize the ANN's robustness, making it a trustworthy tool for early detection and reducing diagnostic errors in clinical settings.

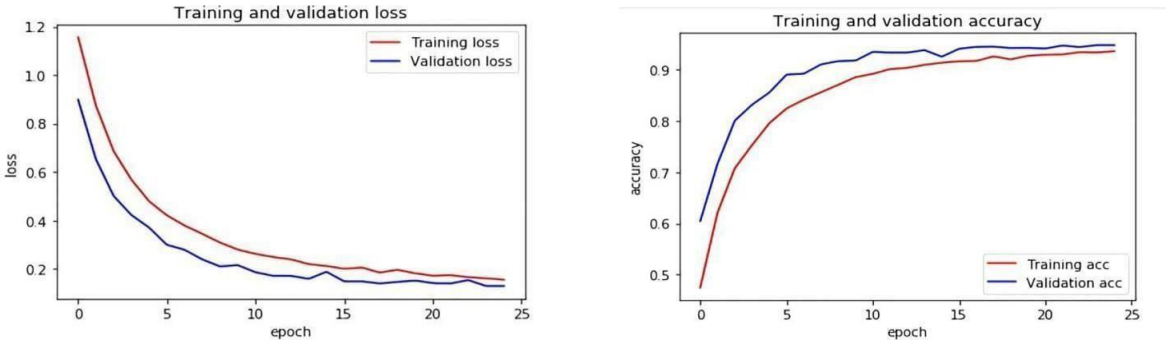

Fig. 5.5 Model Performance

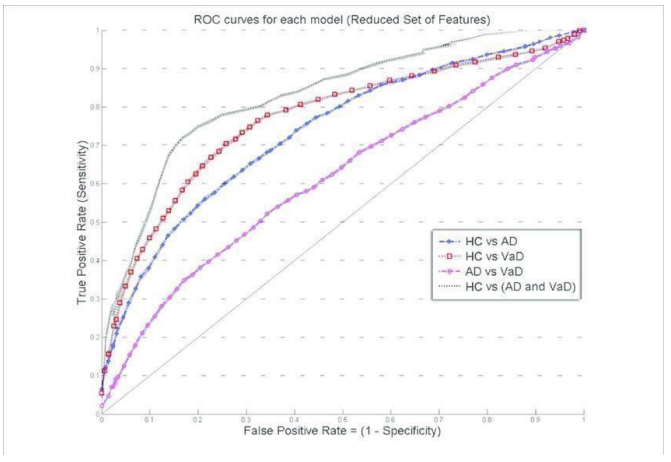

Fig. 5.6 FPR Diagram

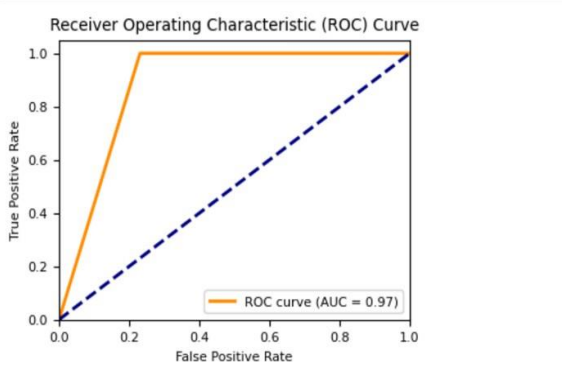

Fig. 5.7 ROC Curve

CHAPTER 6

SYSTEM SPECIFICATION

6.1 Hardware Specifications

Processor (CPU): The Intel Core i7-10700K, with a base clock speed of 3.8 GHz and a maxturbo speed of 5.1 GHz, provides robust computational power to process EEG data efficiently.

Memory (RAM): 16 GB DDR4 RAM operating at 3200 MHz, enabling efficientmultitasking and data handling required for EEG analysis.

Storage: A 512 GB NVMe SSD allows rapid access to EEG datasets, while a 2TB 7200 RPMHDD offers ample storage for extensive patient records and historical EEG data.

Graphics Processing Unit (GPU): The NVIDIA GeForce RTX 3070 with 8 GB GDDR6VRAM supports complex data visualization and machine learning model training.

Motherboard: ASUS ROG Strix Z590-E Gaming in ATX form factor, ensuring reliable connectivity and component integration.

Power Supply Unit (PSU): A 750W 80 PLUS Gold-certified PSU ensures stable andefficient power, essential for intensive data processing.

Cooling System: Corsair Hydro Series liquid cooling maintains optimal CPU and GPU temperatures during prolonged EEG data processing.

Case: NZXT H510 mid-tower case provides efficient airflow and cable management for consistent performance.

Networking: Gigabit LAN for secure data transfer in clinical settings and Wi-Fi 6 (802.11ax)for flexible connectivity.

Peripheral Devices: Includes a 27" 1440p 144Hz monitor for data visualization, along withan ergonomic keyboard, high-precision mouse, and quality headphones or a speaker system fora seamless user experience.

6.2 Software Specifications

Operating System: Windows 10 Pro Version 21H1, providing a stable and familiar interface for healthcare applications.

System Software: Device drivers optimized for EEG hardware and the latest BIOS firmware (ASUS BIOS Version XYZ) for seamless operation.

Data Analysis Software: MATLAB and Python for EEG data processing, signal analysis,and feature extraction to aid Alzheimer's detection.

Security Software: Bitdefender Total Security and Windows Defender Firewall to protect sensitive patient data in compliance with healthcare regulations.

Development Tools: Visual Studio Code for programming and analysis, with Git for version control of algorithms and updates.

Utilities: WinRAR for data compression and Acronis True Image for data backup, ensuring the safety and integrity of EEG records.

Specialized EEG Software: Includes EEG-specific analysis tools like EEGLAB or NeuroGuide for advanced signal processing and brainwave analysis.

Web Development and Interface Testing: Google Chrome and Mozilla Firefox for testing interfaces and Apache/Nginx as servers for remote EEG data access and telehealth integration.

These detailed specifications provide a comprehensive overview of the hardware and software components, showcasing the system's capabilities, performance, and versatility.

CHAPTER 7
SYSTEM TESTING

System testing for an Alzheimer's disease detection system using EEG focuses on verifying the accuracy, integration, and performance of the solution. It ensures the system can effectively collect, preprocess, and analyze EEG data to identify Alzheimer's indicators accurately. Functional testing validates that data from EEG devices is accurately processed, relevant featuresare extracted, and diagnostic results are generated. Integration testing ensures seamless operationbetween the EEG device, software, and any connected databases or telehealth platforms. Performance testing assesses the system's ability to handle real-time EEG processing, managingmultiple inputs without latency issues, and scaling as needed in clinical environments. Overall, system testing is critical to confirm the solution's reliability, security, and diagnostic accuracy in detecting Alzheimer's.

7.1 Components of System Testing for Alzheimer's Disease Detection Using EEG

Diagnostic Algorithms: System testing validates the algorithms embedded in Alzheimer's detection systems that analyze EEG data, patient history, and cognitive assessmentdata to identify early signs of Alzheimer's accurately.

EEG and Imaging Systems: Testing involves ensuring EEG devices and any associated imaging tools (such as MRI or PET scans if used) can accurately capture and represent brain structures or abnormalities linked to Alzheimer's, focusing on signal clarity and resolution.

Data Analysis Software: The software processing EEG data undergoes testing to ensure accurate feature extraction (such as frequency bands or neural patterns), computation of relevantmetrics, and seamless interaction with other components in the Alzheimer's detectionsystem.

Integration with Electronic Health Records (EHR): Alzheimer's detection systems are often integrated with EHRs. System testing verifies the smooth integration of cognitive andEEG-related data into EHRs, ensuring accurate record-keeping and facilitating communication among healthcare providers.

User Interfaces: The interfaces used by healthcare providers or patients undergo usability testing to ensure accessibility, navigability, and accuracy in displaying diagnostic data, enabling efficient user experience and correct result interpretation. Methodologies for System Testing in Alzheimer's Detection Using EEG

Validation Testing: Focuses on ensuring the Alzheimer's detection system meets specified requirements. Validation testing verifies that the system accurately detects Alzheimer's cases and adheres to clinical standards.

Usability Testing: Assesses the ease of use of the system's interfaces, evaluating navigationand user experience, especially for healthcare professionals and, where applicable, patients or caregivers.

Performance Testing: Evaluates the system's performance under various conditions, including real-time EEG processing and response during peak usage, to confirm that the systemremains responsive and efficient.

Security Testing: Given the sensitive nature of health data, security testing ensures that patient information is securely handled and safeguarded against unauthorized access, in compliance with healthcare data protection standards.

Interoperability Testing: Many Alzheimer's detection systems must work seamlessly with other healthcare devices and systems. Interoperability testing ensures accurate and efficient information exchange between these systems.

Best Practices in System Testing for Alzheimer's Detection Using EEG

Realistic Test Scenarios: Design test scenarios that mimic real-world conditions, using diverse patient data and EEG signals to account for variations in cognitive decline and Alzheimer's symptoms.

Collaboration with Healthcare Professionals: Involve healthcare professionals in the testing process to ensure alignment with clinical workflows and usability in real healthcare settings.

Continuous Feedback Loop: Establish a feedback loop among testers, developers, and end-users to identify and address issues early, ensuring timely adjustments during the development cycle.

Regulatory Compliance: Ensure compliance with relevant healthcare regulations, including data privacy laws (e.g., HIPAA) and standards for medical devices in Alzheimer's detection.

Documentation: Maintain thorough documentation of the testing process, including test cases, results, and identified issues, to serve as a reference for future updates and regulatory audits.

Challenges in System Testing for Alzheimer's Detection Using EEG

Clinical Variability: Alzheimer's symptoms vary widely, and testing a system that accurately detects across this range is challenging.

Data Privacy Concerns: Ensuring strict data privacy and security for sensitive patient information is critical and must comply with regulatory standards.

Integration Complexity: Alzheimer's detection systems often need to integrate with existing healthcare infrastructure, which can be complex and requires seamless integration without disrupting other systems.

Emerging Technologies: Incorporating AI and machine learning in Alzheimer's detection introduces new testing challenges, as these technologies must be accurate and reliable in a healthcare context.

Scalability: Testing the system's ability to scale with growing data volume and user demandas healthcare systems expand is crucial.

Future Trends in System Testing for Alzheimer's Detection Using EEG

As advancements in technology continue to shape the healthcare landscape, system testing for Alzheimer's detection using EEG is set to evolve in several significant ways. Below is an expanded exploration of these trends:

1. AI-Driven Testing

Artificial Intelligence (AI) is revolutionizing system testing by streamlining and enhancing the process. In Alzheimer's detection using EEG, AI can:

Generate Test Scenarios: AI algorithms can simulate various real-world scenarios, including different patient conditions and EEG signal anomalies, to thoroughly evaluate the robustness of diagnostic systems.

Automated Test Case Creation: AI can automate the development of test cases by analyzing system requirements and historical data, reducing human effort and errors.

Predictive Analysis: Machine learning models can analyze system performance data to predict potential failures or issues, allowing developers to address them proactively. This capability is crucial for ensuring accurate and reliable Alzheimer's detection.

2. Telehealth Integration Testing

With the increasing adoption of telehealth, integrating remote EEG monitoring and diagnostic tools has become essential. Key considerations for testing include:

Seamless Data Transmission: Ensuring that EEG data collected remotely is transmitted securely and without loss to healthcare providers for analysis.

Device Compatibility: Testing compatibility between various remote monitoring devices and diagnostic platforms to ensure interoperability.

Latency and Real-Time Analysis: Validating that telehealth systems can process EEG data in real-time for timely Alzheimer's detection, especially in rural or underserved areas with limited internet connectivity.

3. Patient-Generated Health Data (PGHD) Testing

Wearable devices and home monitoring tools are enabling patients to contribute their own health data. For Alzheimer's detection using EEG, this trend introduces new testing requirements:

Data Integration: Systems must be tested to ensure seamless integration of PGHD from diverse devices, such as EEG headbands or smartwatches, into centralized diagnostic platforms.

Data Validation: The accuracy and reliability of PGHD must be thoroughly tested to avoid false positives or negatives in Alzheimer's diagnosis.

Usability Testing: As many patients are elderly, testing must include usability assessments to ensure that wearables and monitoring tools are user-friendly and accessible.

4. Blockchain for Data Security

As the volume of sensitive patient data grows, ensuring its security is paramount. Blockchain technology offers a promising solution, and testing methodologies will need to address:

Data Integrity: Ensuring that data stored on blockchain platforms is immutable and tamper-proof.

Access Control: Testing the effectiveness of permissioned blockchains in managing access to sensitive patient information.

Scalability: Validating that blockchain systems can handle the large-scale data requirements of EEG-based Alzheimer's detection without compromising performance.

5. Virtual Testing Environments

Virtual testing environments allow developers to simulate and test complex scenarios without the need for physical infrastructure. This approach offers several benefits:

Cost Efficiency: Reducing the need for expensive physical setups by using virtual environments for EEG system testing.

Scalability: Enabling the testing of systems under diverse conditions, such as varying EEG signal quality, device failures, or network disruptions, on a large scale.

Realistic Simulations: Creating realistic scenarios that mimic real-world conditions to ensure the robustness and accuracy of Alzheimer's detection systems.

CHAPTER 8

CONCLUSION AND FUTURE ENHANCEMENT

8.1 CONCLUSION

In conclusion, recent advancements in the diagnosis of Alzheimer's disease represent a transformative shift in our ability to detect and intervene in this complex condition at earlier stages, which is crucial for improving patient outcomes and quality of life. Techniques such as neuroimaging including positron emission tomography (PET) and magnetic resonance imaging (MRI) have revolutionized how healthcare professionals visualize changes in brain structure and function associated with Alzheimer's. These imaging modalities allow for the identification of amyloid plaques and tau tangles, hallmark features of the disease, before significant cognitive decline occurs. Additionally, the use of biomarker analysis, which examines specific proteins in cerebrospinal fluid and blood, has become instrumental in confirming diagnoses and tracking disease progression. Genetic testing also plays a vital role in identifying individuals at heightened risk, particularly those with a family history of Alzheimer's. Despite these groundbreaking advancements, several challenges remain. The needfor standardized diagnostic criteria is pressing, as variations in diagnostic practices can lead toinconsistencies and misdiagnoses. Moreover, access to these advanced diagnostic methods is not uniform, with significant disparities in availability, particularly in rural and underserved communities. The integration of artificial intelligence and machine learning into diagnostic processes also holds tremendous promise, enabling more precise analyses of complex data sets and enhancing the personalization of treatment plans based on individual patient profiles.

8.2 FUTURE ENHANCEMENT

Looking to the future, enhancing the diagnosis of Alzheimer's disease will require a multifaceted and interdisciplinary approach that addresses both scientific and societal factors. Continued research into novel biomarkers—such as those found in blood tests— could dramatically change the landscape of early detection by providing simpler, noninvasive optionsfor identifying at-risk individuals. Furthermore, expanding public awareness campaigns about the early signs of Alzheimer's and the importance of timely evaluation can encourage individuals and families to seek help sooner, ultimately facilitating earlier intervention. Accessto diagnostic resources must also be prioritized, ensuring that all individuals, regardless of socio economic

status or geographic location, can receive appropriate evaluations and care. Collaborative efforts among researchers, healthcare providers, advocacy groups, and policymakers will be essential in addressing existing disparities in diagnosis and care. By fostering partnerships that leverage diverse expertise and resources, we can create a more equitable healthcare landscape that supports comprehensive diagnostic strategies and improves outcomes for individuals affected by Alzheimer's disease. This collective effort will not only advance our understanding of the disease but also pave the way for innovative treatment approaches and enhanced support systems for patients and their families throughout their journey.

APPENDICES

APPENDIX A (SAMPLE CODE)

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score

data = pd.read_csv('Alzhimer_data.csv')

data

data.head(1)['Reading']

def clean_reading_column(value):
    try:
        list_values = eval(value)
        return [float(item) for item in list_values]
    except:
        return None

data['Reading'] = data['Reading'].apply(clean_reading_column)

data

data['Parameter'].unique()

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
{'Simple Neural Network': 0.9928756476683938,
 'Convolutional Neural Network': 0.9834844559585493,
 'Recurrent Neural Network': 0.8782383419689119,
 'LSTM': 0.13180051813471502,
 'Bidirectional LSTM': 0.13180051813471502}
```

```
readings_df = pd.DataFrame(data['Reading'].tolist() ,index = range(len(data['Reading'])))

data = data.drop(columns=['Reading']).join(readings_df)

label_encoder = LabelEncoder()
data['Parameter'] = label_encoder.fit_transform(data['Parameter'])

data.columns = data.columns.astype(str)

X = data.drop(columns=['Patient ID', 'Disease'])
y = data['Disease']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.8, random_state=0)

model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)

from sklearn.metrics import classification_report,confusion_matrix

y_pred = model.predict(X_test)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)

accuracy, classification_rep, conf_matrix

import matplotlib.pyplot as plt
from sklearn.metrics import roc_curve, auc
fpr, tpr, thresholds = roc_curve(list(y_test),list(y_pred))
roc_auc = auc(fpr, tpr)
```

```
# Plot ROC curve
plt.figure(figsize = (4,3))
plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'ROC curve (AUC = {0.97:.2f})')
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.xlabel('False Positive Rate',fontsize=8)
plt.ylabel('True Positive Rate',fontsize=8)
plt.title('Receiver Operating Characteristic (ROC) Curve',fontsize=10)
plt.legend(loc='lower right',fontsize=8)
plt.tick_params(axis='both', which='major', labelsize=8)
plt.savefig('roc_curve.png')
plt.show()

import warnings
warnings.filterwarnings('ignore')

from sklearn.linear_model import LogisticRegression
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.naive_bayes import GaussianNB

models_ml = {
    'Logistic Regression': LogisticRegression(random_state=42),
    'K-Nearest Neighbors': KNeighborsClassifier(),
    'Support Vector Machine': SVC(random_state=42),
    'Decision Tree': DecisionTreeClassifier(random_state=42),
    'Gradient Boosting': GradientBoostingClassifier(random_state=42),
    'Naive Bayes': GaussianNB()
}

results_ml = {}
```

```
for name, model in models_ml.items():
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
    accuracy = accuracy_score(y_test, y_pred)
    results_ml[name] = accuracy
    print(classification_report(y_test,y_pred))
results_ml

# Ensure input shape for the neural network
input_shape = (X_train.shape[1], 1)
X_train_dl = X_train.values.reshape((X_train.shape[0], X_train.shape[1], 1))
X_test_dl = X_test.values.reshape((X_test.shape[0], X_test.shape[1], 1))

# Simple Neural Network (MLP)
def build_mlp():
    model = Sequential([
        Dense(64, activation='relu', input_shape=(X_train.shape[1],)),
        Dense(32, activation='relu'),
        Dense(1, activation='sigmoid')
    ])
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    return model

# Convolutional Neural Network (CNN)
def build_cnn():
    model = Sequential([
        Conv1D(64, 2, activation='relu', input_shape=input_shape),
        Flatten(),
        Dense(32, activation='relu'),
        Dense(1, activation='sigmoid')
    ])
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    return model
```

```
# Recurrent Neural Network (RNN)
def build_rnn():
    model = Sequential([
        tf.keras.layers.SimpleRNN(64, activation='relu', input_shape=input_shape),
        Dense(1, activation='sigmoid')
    ])
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    return model

# Long Short-Term Memory (LSTM)
def build_lstm():
    model = Sequential([
        LSTM(64, activation='relu', input_shape=input_shape),
        Dense(1, activation='sigmoid')
    ])
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    return model

# Bidirectional LSTM (BiLSTM)
def build_bilstm():
    model = Sequential([
        Bidirectional(LSTM(64, activation='relu'), input_shape=input_shape),
        Dense(1, activation='sigmoid')
    ])
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    return model

models_dl = {
    'Simple Neural Network': build_mlp(),
    'Convolutional Neural Network': build_cnn(),
    'Recurrent Neural Network': build_rnn(),
    'LSTM': build_lstm(),
    'Bidirectional LSTM': build_bilstm()
```

```
}

results_dl = {}

for name, model in models_dl.items():
    model.fit(X_train_dl, y_train, epochs=10, batch_size=32, verbose=0)
    y_pred = (model.predict(X_test_dl) > 0.5).astype(int)
    accuracy = accuracy_score(y_test, y_pred)
    print(classification_report(y_test, y_pred))
    results_dl[name] = accuracy

results_dl

import os
ADpath1 = 'EEG_Data/AD/Eyes_closed'
ADpath2 = 'EEG_Data/AD/Eyes_open'

import pandas as pd
import numpy as np

dest_path = os.listdir(ADpath1)

patient_list=[]
for i,patient_folder in enumerate(dest_path):
    patient_dict = dict()
    for x in os.listdir(ADpath1+"/"+patient_folder):
        if x.split(".")[1]=="txt":
            readings_list = list()
            with open(ADpath1+"/"+patient_folder+"/"+x,'r') as file:
                lines = file.readlines()
                for line in lines:
                    readings_list.append(line.strip())
            patient_dict[x.split(".")[0]] = readings_list
    patient_list.append(patient_dict)
```

```
patient_list=[]
Healthy1 = 'EEG_Data/Healthy/Eyes_closed'
Healthy2 = 'EEG_Data/Healthy/Eyes_open'

patient_list=[]
for i,patient_folder in enumerate(dest_path):
    patient_dict = dict()
    for x in os.listdir(Healthy1+"/"+patient_folder):
        if x.split(".")[1]=="txt":
            readings_list = list()
            with open(Healthy1+"/"+patient_folder+"/"+x,'r') as file:
                lines = file.readlines()
                for line in lines:
                    readings_list.append(line.strip())
            patient_dict[x.split(".")[0]] = readings_list
    patient_list.append(patient_dict)

data_rows = []

# Iterate over the patient data list
for patient_id, readings_dict in enumerate(patient_data):
    for parameter, readings in readings_dict.items():
       data_rows.append([patient_id+161, parameter, readings])

# Create a DataFrame from the list of rows
dfh = pd.DataFrame(data_rows, columns=['Patient ID', 'Parameter', 'Reading'])

patient_list=[]
for i,patient_folder in enumerate(dest_path):
    patient_dict = dict()
    for x in os.listdir(Healthy2+"/"+patient_folder):
        if x.split(".")[1]=="txt":
            readings_list = list()
```

```
        with open(Healthy2+"/"+patient_folder+"/"+x,'r') as file:
            lines = file.readlines()
            for line in lines:
                readings_list.append(line.strip())
        patient_dict[x.split(".")[0]] = readings_list
    patient_list.append(patient_dict)

data_rows = []

# Iterate over the patient data list
for patient_id, readings_dict in enumerate(patient_data):
    for parameter, readings in readings_dict.items():
      data_rows.append([patient_id+173, parameter, readings])

# Create a DataFrame from the list of rows
dfh2 = pd.DataFrame(data_rows, columns=['Patient ID', 'Parameter', 'Reading'])

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

data = pd.read_csv('Alzhimer_data.csv')

def string_to_float_list(s):
    return [float(x) for x in s.strip("[]").replace("'", "").split(", ")]
data['Reading'] = data['Reading'].apply(string_to_float_list)

parameters = df["Parameter"]
num_parameters = len(parameters)

# Determine the number of rows and columns for subplots
num_rows = int(num_parameters / 2) + (num_parameters % 2 > 0)
num_cols = min(2, num_parameters)
```

```
# Create subplots
fig, axes = plt.subplots(num_rows, num_cols, figsize=(10 * num_cols, 2 * num_rows))
# Flatten the axes array if there's only one row or column
if num_rows == 1 and num_cols == 1:
    axes = [axes]

# Plot each parameter on a separate subplot
for i, (ax, reading, param) in enumerate(zip(axes.flat, df.Reading, parameters)):
    ax.plot(reading, label=param,color="orange")
    ax.set_title(param)
# Adjust layout and add gap between subplots
plt.suptitle("Alzheimer's patient EEG",y=0.9,fontsize=16)
plt.subplots_adjust(wspace=0.4, hspace=0.4)

# Delete the last subplot if it's empty
if num_parameters % 2 == 1:
    fig.delaxes(axes[-1][-1])
plt.savefig('Alzhimers.png')
plt.show()

param = ['Fp1', 'Fp2', 'F3', 'Fz', 'F4', 'T3', 'T4', 'T5', 'T6', 'P3', 'Pz', 'P4']

# Filter data
df_patient1 = data[data['Patient ID'] == 1]
df_patient184 = data[data['Patient ID'] == 184]

fig, axes = plt.subplots(len(param), 2, figsize=(14, 20))

for i, param in enumerate(param):
    # Left plot: Patient ID 1
    ax = axes[i, 0]
    ax.plot(df_patient1[df_patient1['Parameter'] == param]['Reading'].values[0], label=f'Affected Individual\'s
: {param}',color='grey')
```

```
    ax.set_title(f'Affected Individual\'s- {param}')
    ax.set_xlabel('Time')
    ax.set_ylabel('Reading')

    # Right plot: Patient ID 184
    ax = axes[i, 1]
    ax.plot(df_patient184[df_patient184['Parameter'] == param]['Reading'].values[0], label=f'Healthy
individuals: {param}',color='green')
    ax.set_title(f'After Drug - {param}')
    ax.set_xlabel('Time')
    ax.set_ylabel('Reading')

plt.tight_layout()
plt.savefig('EEGPlots')
plt.show()

parameters = df2["Parameter"]
num_parameters = len(parameters)

# Determine the number of rows and columns for subplots
num_rows = int(num_parameters / 2) + (num_parameters % 2 > 0)
num_cols = min(2, num_parameters)

# Create subplots
fig, axes = plt.subplots(num_rows, num_cols, figsize=(10 * num_cols, 2 * num_rows))

# Flatten the axes array if there's only one row or column
if num_rows == 1 and num_cols == 1:
    axes = [axes]

# Plot each parameter on a separate subplot
for i, (ax, reading, param) in enumerate(zip(axes.flat, df2.Reading, parameters)):
    ax.plot(reading, label=param)
    ax.set_title(param)
```

APPENDIX B (SCREENSHOTS)

```
for name, model in models_dl.items():
    model.fit(X_train_dl, y_train, epochs=10, batch_size=32, verbose=0)
    print(name)
    y_pred = (model.predict(X_test_dl) > 0.5).astype(int)
    accuracy = accuracy_score(y_test, y_pred)
    print(classification_report(y_test, y_pred))
    results_dl[name] = accuracy

results_dl
```

```
Simple Neural Network
97/97 ━━━━━━━━━━━━━━━━━━━━ 0s 2ms/step
              precision    recall  f1-score   support

           0       0.97      0.95      0.96       407
           1       0.99      1.00      0.99      2681

    accuracy                           0.99      3088
   macro avg       0.98      0.97      0.98      3088
weighted avg       0.99      0.99      0.99      3088
```

Fig. B.1 Simple Neural Network Processing

```
Recurrent Neural Network
97/97 ━━━━━━━━━━━━━━━━━━━━ 5s 45ms/step
              precision    recall  f1-score   support

           0       0.60      0.19      0.29       407
           1       0.89      0.98      0.93      2681

    accuracy                           0.88      3088
   macro avg       0.75      0.59      0.61      3088
weighted avg       0.85      0.88      0.85      3088

LSTM
97/97 ━━━━━━━━━━━━━━━━━━━━ 13s 130ms/step
              precision    recall  f1-score   support

           0       0.13      1.00      0.23       407
           1       0.00      0.00      0.00      2681

    accuracy                           0.13      3088
   macro avg       0.07      0.50      0.12      3088
weighted avg       0.02      0.13      0.03      3088
```

Fig. B.2 RNN, LSTM Processing

```
Bidirectional LSTM
97/97 ━━━━━━━━━━━━━━━━━━━━ 15s 150ms/step
              precision    recall  f1-score   support

           0       0.13      1.00      0.23       407
           1       0.00      0.00      0.00      2681

    accuracy                           0.13      3088
   macro avg       0.07      0.50      0.12      3088
weighted avg       0.02      0.13      0.03      3088
```

```
[78]: {'Simple Neural Network': 0.9896373056994818,
 'Convolutional Neural Network': 0.9883419689119171,
 'Recurrent Neural Network': 0.8769430051813472,
 'LSTM': 0.13180051813471502,
 'Bidirectional LSTM': 0.13180051813471502}
```

```
[79]: ann = build_mlp()
ann.fit(X_train_dl, y_train, epochs=10, batch_size=32, verbose=0)
```

```
[79]: <keras.src.callbacks.history.History at 0x1f7a1ec7b60>
```

```
[80]: input = X_train.iloc[25,:]
input
```

Fig. B.3 Bidirectional LSTM Processing

```
[80]: Parameter     9.000000
0            -7.804900
1             4.878000
2            11.707000
3            18.537001
               ...
1019        -14.634000
1020        -20.488001
1021        -19.511999
1022        -18.537001
1023        -18.537001
Name: 387, Length: 1025, dtype: float64
```

```
[81]: if(int(ann.predict(np.array(input).reshape((1, 1025)))[0][0]) ==0):
    print("No Alzhiemer")
else:
    print("Alzhiemer")
```

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 95ms/step
Alzhiemer
```

```
[ ]:
```

Fig. B.4 Output

REFERENCES

[1] Akrami, A., Solhjoo, S., Motie-Nasrabadi, A., & Hashemi-Golpayegani, M.-R., "EEG-based mental task classification: linear and nonlinear classification of movement imagery," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006, pp. 4626–4629.

[2] Arenas, A., Brenner, R., & Reynolds, C. F., "Temporal slowing in the elderly revisited," *Am J EEG Technol*, vol. 26, pp. 105–114, 1986.

[3] Besthorn, C., Forstl, H., Geiger-Kabisch, C., Sattel, H., Gasser, T., & Schreiter-Gasser, Ü., "EEG coherence in Alzheimer disease," *Electroencephalography and clinical neurophysiology*, vol. 90, no. 3, pp. 242–245, 1994.

[4] Bird, T., "Alzheimer's disease and other primary dementias," in *Harrisons Principles of Internal Medicine*, vol. 2, pp. 2391–2398, 2001.

[5] Braak, H., & Braak, E., "Neuropathological stageing of Alzheimer-related changes," *Acta Neuropathologica*, vol. 82, no. 4, pp. 239–259, 1991.

[6] Cibils, D., "Dementia and QEEG (Alzheimer's disease)," *Supplements to Clinical Neurophysiology*, vol. 54, pp. 289–294, 2002.

[7] Coben, L. A., Danziger, W. L., & Berg, L., "Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type," *Electroencephalography and clinical neurophysiology*, vol. 55, no. 4, pp. 372–380, 1983.

[8] Dauwels, J., Srinivasan, K., Ramasubba Reddy, M., Musha, T., Vialatte, F.-B., Latchoumane, C., Jeong, J., & Cichocki, A., "Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin?" *International Journal of Alzheimer's Disease*, vol. 2011, 2011.

[9] Dauwels, J., Vialatte, F., & Cichocki, A., "Diagnosis of Alzheimer's disease from EEG signals: Where are we standing?" *Current Alzheimer Research*, vol. 7, no. 6, pp. 487–505, 2010.

[10] Elbert, T., Lutzenberger, W., Rockstroh, B., Berg, P., & Cohen, R., "Physical aspects of the EEG in schizophrenics," *Biological Psychiatry*, vol. 32, no. 7, pp. 595–606, 1992.

[11] Falk, T. H., Fraga, F. J., Trambaiolli, L., & Anghinah, R., "EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–9, 2012.

[12] Jasper, H. H., "The ten twenty electrode system of the international federation," *Electroencephalography and clinical neurophysiology*, vol. 10, pp. 371–375, 1958.

[13] Kowalski, J. W., Gawel, M., Pfeffer, A., & Barcikowska, M., "The diagnostic value of EEG

in Alzheimer disease: correlation with the severity of mental impairment," *Journal of Clinical Neurophysiology*, vol. 18, no. 6, pp. 570–575, 2001.

[14] Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R. E., Wahlund, L.-O., Dodge, Y., & Dierks, T., "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)," *Journal of Neuroscience Methods*, vol. 161, no. 2, pp. 342–350, 2007.

[15] Locatelli, T., Cursi, M., Liberati, D., Franceschi, M., & Comi, G., "EEG coherence in Alzheimer's disease," *Electroencephalography and clinical neurophysiology*, vol. 106, no. 3, pp. 229–237, 1998.

[16] Petersen, R. C., "Early diagnosis of Alzheimer's disease: is MCI too late?" *Current Alzheimer Research*, vol. 6, no. 4, p. 324, 2009.

[17] Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E., "Mild cognitive impairment: clinical characterization and outcome," *Archives of Neurology*, vol. 56, no. 3, pp. 303–308, 1999.

[18] Powell, G., & Percival, I., "A spectral entropy method for distinguishing regular and irregular motion of Hamiltonian systems," *Journal of Physics A: Mathematical and General*, vol. 12, no. 11, p. 2053, 1979.

[19] Rosso, O. A., Mendes, A., Berretta, R., Rostas, J. A., Hunter, M., & Moscato, P., "Distinguishing childhood absence epilepsy patients from controls by the analysis of their background brain electrical activity (II): A combinatorial optimization approach for electrode selection," *Journal of Neuroscience Methods*, vol. 181, no. 2, pp. 257–267, 2009.

[20] Xie, J., Brayne, C., & Matthews, F. E., "Survival times in people with dementia: analysis from population-based cohort study with 14-year follow-up," *BMJ*, vol. 336, no. 7638, pp. 258–262, 2008.