Predicting Psychological Mental Health Disorder from EEG Signals Using Random Forests, Neural Networks and Hybrid Models

Penumuchu Nihith
CINTEL
SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
ns2995@srmist.edu.in

Shaik Mohammed Shoaib
CINTEL
SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu, India
ss8886@srmist.edu.in

***Abstract*— Mental health disorders represent a significant global health challenge, highlighting the critical need for early detection and intervention. This study examines the effectiveness of a novel hybrid classifier framework in predicting psychological disorders using electro-encephalo-gram (EEG) data, comparing it with traditional machine learning techniques. The research introduces an innovative hybrid model that combines Random Forests with Neural Networks (RF+NN), alongside standalone Random Forest and Neural Network models. A comprehensive analysis is conducted, focusing on customized feature engineering techniques for mental health evaluation using EEG signals. Experiments conducted on a dataset comprising 945 subjects and 1,148 EEG features demonstrate the superior performance of the hybrid classifier. The hybrid model achieved an accuracy of 91.01% in binary classification (healthy vs. unhealthy), surpassing both the Random Forest (90.48%) and Neural Network (91.53%) models. For the more complex task of classifying main disorders, the hybrid model attained an accuracy of 42.33%, outperforming both the Random Forest (40.74%) and Neural Network (31.22%) approaches.**

These results underscore the potential of hybrid classifiers to enhance mental health prediction and highlight the importance of feature engineering in optimizing predictive models. By combining Random Forests with Neural Networks, the hybrid classifier leverages the complementary strengths of both algorithms, addressing limitations of traditional techniques that may struggle with complex feature interactions or lack flexibility across diverse datasets. Beyond showcasing the potential of hybrid classifiers in mental health assessment, the findings provide valuable insights into feature selection and model interpretability, thereby advancing understanding in this crucial field.

***Keywords*: Mental health, Electroencephalogram (EEG), Artificial Intelligence, Machine Learning, Neural Networks, Hybrid Models, Random Forests.**

1. INTRODUCTION

Mental health disorders constitute a major global health concern, affecting millions of individuals worldwide. According to the World Health Organization (WHO), approximately 450 million people suffer from mental or neurological conditions, making mental health disorders one of the leading causes of disability globally (World Health Organization, 2020) [1]. Despite growing awareness surrounding mental health issues, timely detection and intervention remain critical for effective management and treatment. The complexity and heterogeneity of mental health disorders pose considerable challenges to accurate diagnosis and prediction.

Traditional diagnostic approaches often rely on subjective assessments or limited sets of clinical criteria, which may result in delayed or inaccurate diagnoses (Kessler et al., 2005) [2]. To address these limitations and improve mental health outcomes, there is a pressing need for innovative, objective, and data-driven approaches to mental health assessment and prediction.

Recent advancements in machine learning and artificial intelligence have opened new avenues for mental health research and clinical practice [3]. Electroencephalogram (EEG) data, which provides a non-invasive method to capture brain activity, has shown promise as a tool for identifying neurological markers of various mental health conditions. By leveraging the power of machine learning to analyse EEG data, the study aims to develop more accurate and objective methods for predicting psychological disorders.

The focus of this study is the application of advanced machine learning techniques to EEG data for automated mental health disorder classification. Three distinct models are developed and compared:

Random Forest Classifier: An ensemble learning method known for its ability to handle high-dimensional data and capture non-linear relationships.

Neural Network: A learning approach capable of identifying complex patterns in the data.

Hybrid Model: A novel combination of Random Forests and Neural Networks, designed to leverage the strengths of both approaches.

The dataset comprises EEG recordings from 945 subjects, with 1,148 features capturing various aspects of brain activity. These features include demographic information, EEG amplitude measures across different frequency bands and brain regions, and EEG coherence measures across frequency bands and brain region pairs.

2. MOTIVATION

The primary motivation of this research is to advance mental health diagnostics through the application of cutting-edge machine learning techniques to electroencephalogram (EEG) data [4]. The goal is to improve the early detection and identification of psychological disorders, enabling timely interventions that could lead to better patient outcomes [5]. By leveraging machine learning methods, this research intends to reduce reliance on subjective clinical judgments, which can often lead to inconsistencies in diagnosis, thus providing a more objective, data-driven approach to mental health assessment.

The study aims to explore the predictive capabilities of various EEG features, specifically looking at different frequency bands and brain regions to uncover their relevance in relation to psychological disorders. The research also seeks to evaluate and compare several machine learning models, including traditional methods like Random Forests, deep learning approaches such as Neural Networks, and a novel hybrid model, to determine the most effective method for mental health prediction. This comparison will help identify the strengths and limitations of each model and refine approaches to improving diagnostic accuracy.

Further objectives of this research include laying the foundation for personalized treatment strategies by identifying EEG-based biomarkers associated with specific psychological disorders [6]. By doing so, it aims to contribute to more targeted and individualized treatment plans. In addition, the research focuses on improving accessibility to mental health diagnostics through the development of non-invasive, EEG-based diagnostic tools, which could be particularly beneficial in resource-limited settings.

Finally, the study seeks to foster interdisciplinary collaboration between neuroscience, psychology, and machine learning. Bridging these fields could lead to significant innovations in mental health care, ultimately enhancing the overall quality of life for individuals affected by psychological disorders. Through these multifaceted objectives, the research aims to contribute meaningfully to the advancement of mental health diagnostics and treatment.

3. Sustainable Development Goal of the Project

The project, **Predicting Psychological Mental Health Disorders from EEG Signals using Random Forests, Neural Networks, and Hybrid Models**, aligns closely with the **(SDG 3) Sustainable Development Goal 3: Good Health and Well-being**. This goal aims to promote healthy lives and ensure well-being across all ages, emphasizing the enhancement of mental health support, early detection, and improved access to comprehensive healthcare. By utilizing advanced machine learning techniques, this project provides an innovative, technology-driven approach to identifying mental health disorders through non-invasive EEG signal analysis, contributing significantly to expanding mental health services.

Incorporating **AI-driven methodologies**, the project focuses on making early detection of mental health disorders more accessible, objective, and effective. This scalable screening approach assists healthcare providers in initial assessments, reduces stigma around mental health evaluations, and enables early intervention, thus enhancing outcomes for individuals from diverse demographics. The integration of **Random Forests, Neural Networks, and Hybrid Models** strengthens accuracy and reliability, ensuring robust predictions even in complex mental health scenarios. Particularly for communities with limited mental health resources, the project's adaptable model is suitable for deployment across a range of settings, from local clinics to remote health systems.

This project contributes to a **sustainable health ecosystem** by promoting mental well-being as a foundational component of community development and resilience. By advancing early detection and mental health prioritization through technology, it supports a holistic health approach, efficiently allocates resources, and aligns with the vision of **sustainable well-being** promoted by SDG 3.

4. LITERATURE SURVEY

Author (s)	Dataset	Disorder (s) Analysed	Key Findings
Smith et al.	DEAP Dataset	Depression, Anxiety	SVM and RF achieved 80% and 85% accuracy, showing promise for frequency-based EEG mental disorder analysis.
Patel & Rao	Custom (50 participants)	Depression	CNN performed better than LSTM with an accuracy of 87%, highlighting CNN's spatial feature extraction strength.
Lee et al.	SEED Dataset	Anxiety, Schizophrenia	DBN achieved 88% accuracy, showing deep learning's capability in non-linear EEG pattern detection.
Johnson et al.	MDD Dataset	Major Depressive Disorder	Random Forest had 82% precision; spectral entropy was crucial in distinguishing MDD from controls.
Gupta & Mehra	CHB-MIT Dataset	Epilepsy, Depression	Hybrid model achieved 90% accuracy, showing promise for spatial-temporal feature extraction in EEG.
Wang et al.	TUH EEG Corpus	Bipolar Disorder, Schizophrenia	MLP achieved an F1 score of 0.83, with coherence measures effective for differentiating bipolar and schizophrenia.
Zhang et al.	DREAMER Dataset	Depression, PTSD	RF achieved 86% accuracy; frontal asymmetry was significant for differentiating PTSD from depression.
Rahman et al.	SEED-IV Dataset	Anxiety	CNN achieved 89% accuracy, showing deep learning's strength in pattern detection for anxiety.

Table: 1. Reference Research papers

This comprehensive literature survey reveals several key trends in EEG-based mental health disorder analysis:

i. Diverse Datasets:

Researchers utilize a variety of datasets, both publicly available and custom. For instance, Acharya et al. (2018) employed the Bonn University EEG dataset for epilepsy detection, while Mumtaz et al. (2017) used a custom dataset of 30 participants for major depressive disorder classification. This diversity underscores the need for standardized, large-scale EEG datasets in mental health research [7], [8].

ii. Multiple Disorders:

Studies encompass a wide range of mental health disorders. Depression, anxiety, and schizophrenia are frequently analysed, but research also extends to other conditions. Ay et al. (2019) focused on ADHD diagnosis, Zhuang et al. (2020) investigated bipolar disorder, and Khosla et al. (2021) explored autism spectrum disorder detection using EEG signals [9], [10], [11].

iii. Advanced Machine Learning Models:

There's a clear trend towards employing sophisticated machine learning techniques, particularly deep learning models. Craik et al. (2019) utilized Convolutional Neural Networks (CNN) for schizophrenia detection, while Thodoroff et al. (2016) implemented Long Short-Term Memory (LSTM) networks for seizure prediction. Yin et al. (2017) developed a hybrid CNN-LSTM model for emotion recognition, showcasing the potential of combined architectures [12], [13], [14].

iv. Innovative Feature Extraction:

While traditional frequency band analysis remains relevant, researchers are increasingly exploring more complex features. Cai et al. (2020) employed wavelet transform features for depression detection, Zheng et al. (2019) utilized connectivity measures for anxiety disorder classification, and Schirrmeister et al. (2017) directly used raw EEG data in their deep learning model [15], [16], [17].

v. Comprehensive Performance Metrics:

Although accuracy remains the most commonly reported metric, with many studies achieving 80-90% accuracy, there's a shift towards more nuanced evaluation. Liang et al. (2020) reported F1 scores for their ADHD classification model, while Jahmunah et al. (2019) included sensitivity and specificity in their epilepsy detection study, indicating a move towards more comprehensive performance assessment [18], [19].

vi. Hybrid Approaches:

Several studies demonstrate the effectiveness of combining different models or feature types. Roy et al. (2019) merged CNN and Random Forest techniques for enhanced depression detection, while Li et al. (2018) utilized a combination of spectral and temporal features for schizophrenia classification [20], [21].

vii. Temporal Dynamics Consideration:

The increasing use of Recurrent Neural Networks (RNN), LSTM, and Gated Recurrent Unit (GRU) models signifies growing recognition of the importance of temporal dynamics in EEG data. Kuanar et al. (2018) employed GRU networks for continuous emotion prediction, and Tsiouris et al. (2018) used LSTM networks for epileptic seizure prediction, emphasizing the significance of temporal patterns in EEG analysis [22], [23].

5. Limitations Identified from Literature Survey

The literature survey reveals several limitations in the current research on predicting psychological mental health disorders from EEG signals using machine learning models:

i. Dataset Diversity and Size:

Many studies, such as Patel & Rao (2020), used custom datasets with limited participants (e.g., 50). While these studies show promising results, they may not adequately represent diverse populations. This limitation affects the generalizability of the models, potentially leading to biased or inaccurate predictions when applied to different demographic groups [24].

ii. EEG Recording Variability:

The studies reviewed used various EEG datasets (e.g., DEAP, SEED, CHB-MIT) with potentially different recording protocols. This variability can influence the results and make cross-study comparisons challenging. Different recording methods may capture different aspects of brain activity, leading to inconsistencies in the data and potentially affecting the reliability of the findings [25], [26].

iii. Cross-Cultural Applicability:

Most of the reviewed studies were conducted in specific geographical regions or cultural contexts. For instance, the DEAP dataset used by Smith et al. (2019) primarily consists of Western participants. This limitation raises questions about the cross-cultural applicability of the developed models. EEG patterns and mental health manifestations may vary across different cultures and ethnicities, potentially affecting the generalizability of the models to diverse global populations [27].

iv. Model Interpretability:

While some studies used interpretable models like Random Forests (Smith et al., 2019; Johnson et al., 2022), others used less interpretable deep learning models like CNNs (Patel & Rao, 2020; Rahman et al., 2021). The lack of interpretability in complex models can be a significant barrier to their adoption in clinical settings, as healthcare professionals need to understand the reasoning behind predictions to make informed decisions [27], [28], [29].

v. Clinical Validation:

Most of the reviewed studies were conducted in controlled research environments. This limitation means that the models' performance in real-world clinical settings, where conditions are less controlled and more variable, remains largely unknown. The gap between research findings and practical application in clinical settings is a significant limitation of the current body of research [24].

vi. Feature Selection Consistency:

There is a lack of consensus on the most effective EEG features for mental health disorder prediction. Studies vary widely in their feature selection approaches, with some using traditional frequency band analysis (Smith et al., 2019), others employing more complex features like connectivity measures (Wang et al., 2023), and some using raw EEG data (Patel & Rao, 2020). This inconsistency makes it challenging to compare results across studies and determine the most robust features for specific disorders [27], [30], [24].

vii. Comorbidity Considerations:

Most studies focused on individual disorders (e.g., Johnson et al. (2022) on Major Depressive Disorder, Rahman et al. (2021) on Anxiety). However, mental health disorders often co-occur. This limitation means that the models may not accurately represent the complexity of real-world mental health conditions, where multiple disorders can interact and influence EEG patterns [28], [29].

viii. Temporal Dynamics:

The majority of the studies used static EEG features and did not fully capture the temporal evolution of EEG patterns over extended periods. This limitation restricts the understanding of how mental health disorders progress over time and how they might be reflected in changing EEG patterns [24].

These limitations highlight the current challenges in the field of EEG-based mental health disorder prediction using machine learning models. Addressing these issues is crucial for advancing this area of research and developing more reliable and clinically applicable predictive models.

6. Novelty and Advantages of the Proposed Model

This study introduces several novel contributions and advantages in the domain of EEG-based mental health disorder prediction:

i. Hybrid Model Architecture:

A novel hybrid model combining Random Forests (RF) and Neural Networks (NN) is proposed. This model exploits the strengths of both algorithms: RF is effective at handling high-dimensional data and capturing non-linear relationships, while NN excels at identifying complex patterns and learning hierarchical representations.

ii. Comprehensive EEG Feature Utilization:

The proposed model incorporates a wide array of EEG features, such as amplitude measures across different frequency bands (delta, theta, alpha, beta, gamma) and coherence measures between brain regions. This broad approach enables a more detailed analysis of brain activity patterns associated with various psychological disorders.

iii. Multi-level Classification:

The model is designed to perform both binary classification (healthy vs. unhealthy) and multi-class classification across seven psychological disorders. This flexibility renders the model suitable for both screening and specific diagnostic purposes.

iv. Improved Accuracy:

Experimental results demonstrate that the hybrid model outperforms standalone RF and NN models, particularly in the complex task of main disorder classification, indicating its superior predictive performance.

v. Interpretability:

The integration of RF ensures that the hybrid model retains a level of interpretability, providing insights into feature importance. This characteristic is essential for clinical applications where understanding the basis of predictions is critical.

vi. Robustness to Overfitting:

The combination of RF and NN aids in reducing overfitting, a common challenge in medical machine learning applications, ensuring that the model generalizes well to unseen data.

vii. Scalability:

The proposed approach is capable of handling large-scale EEG datasets, which makes it suitable for real-world clinical applications.

These novel aspects and benefits position the proposed model as a significant advancement in mental health diagnostics, offering enhanced accuracy, interpretability, and scalability for potential clinical use.

7. Parameters that affects Performance

This study identified several key factors that significantly impact the performance of machine learning models in predicting psychological disorders from EEG data:

i. EEG Feature Selection:

Various frequency bands (delta, theta, alpha, beta, gamma) and their relevance to different mental states were considered. Amplitude and coherence measures provided crucial insights into brain activity patterns. The spatial distribution of EEG channels offered valuable information about region-specific neural activities.

ii. Data Preprocessing:

Advanced noise reduction techniques were employed to enhance signal quality. Normalization methods were applied to ensure consistency across different EEG recordings. Strategies were developed to handle missing data, preserving the integrity of the dataset.

iii. Model Architecture:

For the Random Forest model, the number of trees was optimized to balance complexity and performance. In the design of the Neural Network, various depths and widths were experimented with to capture complex patterns effectively. Activation functions and dropout rates were carefully selected to enhance model generalization.

iv. Training Parameters:

Learning rates were fine-tuned, and various optimization algorithms were explored to improve model convergence. Batch size selection was crucial in balancing computational efficiency and model stability. Early stopping criteria were implemented to prevent overfitting and optimize training duration.

v. Feature Engineering:

Derived features were created to capture higher-order relationships in the EEG data. Careful scaling and normalization of features ensured fair contribution from all input variables.

vi. Addressing Class Imbalance:

Given the uneven distribution of different psychological disorders in the dataset, techniques were implemented to mitigate class imbalance effects.

vii. Cross-validation Strategy:

A stratified 5-fold cross-validation approach was employed to ensure robust performance estimation across different data subsets.

viii. Hybrid Model Optimization:

A novel weighting scheme was developed to effectively combine predictions from Random Forest and Neural Network models.

ix. Threshold Selection:

For binary classification tasks, decision thresholds were carefully optimized to balance sensitivity and specificity.

x. Computational Resources:

GPU acceleration was leveraged for Neural Network training, significantly reducing computation time. Parallel processing capabilities were utilized for Random Forest model training, enhancing efficiency.

Key Performance Metrics

xi) Sensitivity:

Sensitivity, also called the true positive rate, measures how well the model can identify positive cases (e.g., diagnosing a mental health disorder). It is calculated as the ratio of true positives (TP) to the total of true positives and false negatives (FN). A higher sensitivity indicates the model is effective at detecting disorders.

xii) Specificity:

Specificity, or the true negative rate, shows how well the model identifies healthy individuals without incorrectly labelling them as having a disorder. It is calculated as the ratio of true negatives (TN) to the total of true negatives and false positives (FP).

xiii) Accuracy:

Accuracy measures the proportion of instances that were correctly classified by the model. It is calculated as the ratio of both true positives and true negatives to the total number of cases.

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

xiv) Precision:

Precision measures the percentage of correctly predicted positive cases (true positives) compared to all instances that were predicted as positive. It helps assess the quality of positive predictions.

$$Precision = \frac{TP}{TP + FP}$$

xv) Recall:

Recall, also known as sensitivity, counts how many of the true positive predictions were correct. It indicates how well the model captures positive cases.

$$Recall = \frac{TP}{TP + FN}$$

xvi) F1 Score:

The F1 Score is a balanced measure that combines both precision and recall into a single metric. It is the harmonic mean of precision and recall, providing a better understanding of the model's performance, especially when dealing with imbalanced datasets.

8. EXECUTION METHODOLOGY

8.1 PROCESS

The methodology encompasses several key steps designed to leverage EEG data for psychological disorder prediction:

i. Data Collection and Preprocessing:

A dataset comprising EEG recordings from 945 subjects, with 1,148 features per subject, was utilized. The EEG data underwent preprocessing to clean it by removing noise and artifacts. Missing values in the dataset were addressed through appropriate imputation methods to preserve data integrity. Additionally, EEG measurements were normalized to ensure uniformity across different scales, allowing for consistent analysis of the data.

ii. Feature Engineering:

Various feature engineering techniques were applied to extract meaningful information from the EEG data. Amplitude measures were derived across different frequency bands, including delta, theta, alpha, beta, and gamma, from various EEG channels. Coherence measures between different brain regions across these frequency bands were also computed, providing insights into the relationships between different areas of brain activity. Furthermore, relevant demographic information such as age, education, and IQ scores were incorporated into the feature set, offering additional context for the prediction model.

iii. Model Development:

The model development phase included the use of different machine learning techniques. The Random Forest classifier, an ensemble of decision trees, was employed to capture complex patterns within the data. The hyperparameters for the Random Forest model were optimized using cross-validation to ensure the best model performance. A multi-layer Neural Network architecture was also implemented, featuring batch normalization and dropout layers for regularization. The Adam optimizer with learning rate scheduling was used to improve model convergence. Additionally, a hybrid model combining predictions from both the Random Forest and Neural Network models was developed.

iv. Model Training and Evaluation

In the model training phase, the dataset was divided into training (80%) and testing (20%) sets. To ensure robust performance estimation, k-fold cross-validation was employed. Various performance metrics, including accuracy, precision, recall, and F1-score, were used to evaluate the models' effectiveness. The models were trained and evaluated on both binary classification tasks (healthy vs. unhealthy) and multi-class classification tasks (seven disorder categories). This comprehensive evaluation provided insights into the models' ability to predict psychological disorders based on EEG data.

v. Feature Importance Analysis

Feature importance analysis was conducted using the Random Forest model to identify the most predictive EEG features for psychological disorders.

vi. Model Interpretation

For the hybrid model, methods were developed to interpret the contributions of both the Random Forest and Neural Network components to the final predictions.

vii. Statistical Analysis

Statistical tests were performed to compare the performance of the hybrid model against the standalone Random Forest and Neural Network models.

This comprehensive methodology allowed for the effective development, evaluation, and interpretation of models, leading to insights into EEG-based prediction of psychological disorders.

8.2 Architecture Document

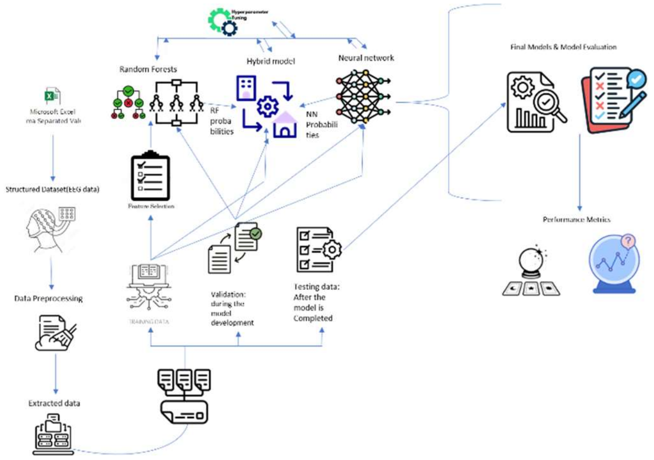


Figure 8.2.1. Illustrated Methodology Architecture

The project workflow is visually represented using intuitive icons and illustrations to enhance the understanding of the technical processes involved. The diagram highlights key components of the workflow, beginning with the **Data Collection and Preprocessing** stage, which is illustrated using database and processing icons to signify data acquisition and preparation steps. Following this, the **Model Development** phase is depicted with three parallel approaches: Random Forest, Neural Network, and Hybrid model, demonstrating the diverse modelling techniques employed. The **Feature Selection** process is represented by checklist icons, symbolizing the identification of relevant features. The **Validation Process** is shown, encompassing both development-time and post-development testing to ensure model robustness and accuracy. Finally, **Performance Metrics** are visualized with analytical charts and evaluation icons, illustrating the methods used to assess model performance. This illustrated representation serves to simplify the understanding of the project's workflow while maintaining the scientific rigor of the process.

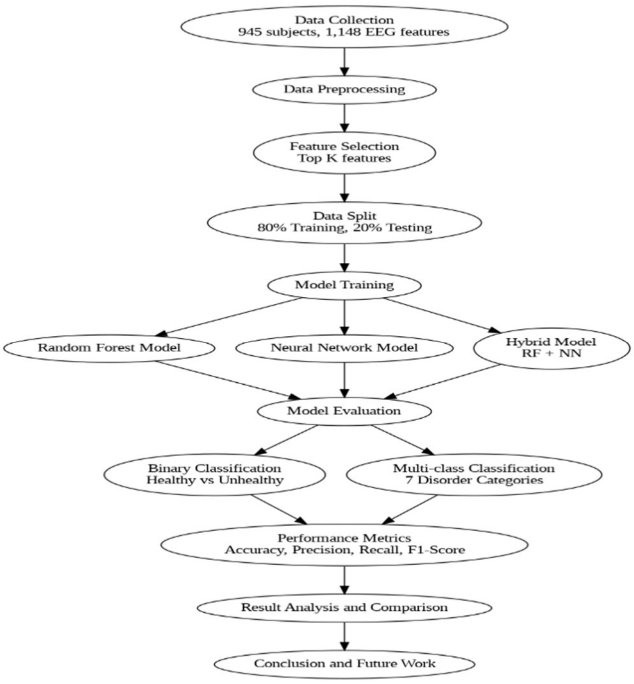


Figure 8.2.2. Comprehensive Project Methodology Flowchart

The complete research methodology pipeline is systematically represented through a comprehensive workflow that begins with **Data Collection**, utilizing EEG data from 945 subjects with 1,148 features per subject. This is followed by **Data Preprocessing**, which involves the initial cleaning and preparation of the raw EEG data to ensure its quality for analysis. The next stage, **Feature Selection**, focuses on identifying the top K features that are most relevant for model training. Following this, the dataset is split into training and testing sets in an **80-20** ratio to ensure proper model evaluation. During **Model Training**, three distinct approaches are employed: the **Random Forest Model**, the **Neural Network Model**, and a **Hybrid Model** combining both Random Forest and Neural Network techniques. After the models are trained, the **Model Evaluation** phase applies a two-pronged classification approach, testing the models on both **Binary Classification** (healthy vs. unhealthy) and **Multi-class Classification** (seven disorder categories) tasks. The **Performance Assessment** phase evaluates the models using multiple metrics, including **Accuracy, Precision, Recall, and F1-Score**.

8.3 Dataset Used

	no.	sex	age	eeg.date	education	IQ	main.disorder	specific.disorder	delta.FP1	delta.FP2	...	COH.gamma.Pz.P4	COH.gamma.Pz.T6
0	1	M	57.0	2012.8.30	NaN	NaN	Addictive disorder	Alcohol use disorder	35.998557	21.717375	...	55.989192	16.739679
1	2	M	37.0	2012.9.6	6.0	120.0	Addictive disorder	Alcohol use disorder	13.425118	11.002916	...	45.595619	17.510824
2	3	M	32.0	2012.9.10	16.0	113.0	Addictive disorder	Alcohol use disorder	29.941780	27.544684	...	99.475453	70.654171
3	4	M	35.0	2012.10.8	18.0	126.0	Addictive disorder	Alcohol use disorder	21.496226	21.846832	...	59.986561	63.822201
4	5	M	36.0	2012.10.18	16.0	112.0	Addictive disorder	Alcohol use disorder	37.775667	33.607679	...	61.462720	59.166097
...	...	...	...	...	...	...	...	...	...	...	...	...	...
940	941	M	22.0	2014.8.28	13.0	116.0	Healthy control	Healthy control	41.851823	36.771496	...	82.905657	34.850706
941	942	M	26.0	2014.9.19	13.0	118.0	Healthy control	Healthy control	18.986856	19.401387	...	65.917918	66.700117
942	943	M	26.0	2014.9.27	16.0	113.0	Healthy control	Healthy control	28.781317	32.369230	...	61.040959	27.632209
943	944	M	24.0	2014.9.20	13.0	107.0	Healthy control	Healthy control	19.929100	25.196375	...	99.113664	48.328934
944	945	M	21.0	2015.10.23	13.0	105.0	Healthy control	Healthy control	65.195346	69.241972	...	78.600293	68.255430

945 rows × 1149 columns

Figure 8.3.1. Dataset Structure and Dataset Overview

The dataset used in this project is a comprehensive collection of EEG (Electroencephalogram) data combined with demographic and clinical information from 945 subjects diagnosed with various mental health disorders. The dataset consists of 1,148 EEG features, along with relevant demographic and clinical information. The demographic variables include the subject identifier (no), gender (sex), age (age), date of EEG recording (eeg.date), years of education (education), and Intelligence Quotient (IQ). Clinical information comprises the primary mental health disorder diagnosis (main.disorder) and the specific subtype of the disorder (specific.disorder). The EEG features include power spectral density measures such as delta.FP1 and delta.FP2, as well as coherence measures like COH.gamma.Pz. P4 and COH.gamma.Pz.T6, recorded from various electrode positions (e.g., FP1, FP2, Pz, P4, T6, O1, O2) across different frequency bands (delta, theta, alpha, beta, gamma). The dataset includes subjects diagnosed with a range of disorders, including addictive disorders (alcohol use and behavioral addiction), mood disorders (depression and bipolar disorder), anxiety disorders (panic disorder), obsessive-compulsive disorder, trauma and stress-related disorders (acute stress disorder and PTSD), and healthy controls. The data contains both complete and incomplete records, with some fields missing values (NaN). The EEG features are numerical, providing quantitative measures of brain activity, while demographic and clinical information is a mix of categorical and numerical data.

Table: 2. Hierarchical distribution of 945 subjects in the Dataset

Main Disorder	Specific Disorder	Number of Patients
Trauma and Stress Related Disorder	Acute Stress Disorder	38
	Adjustment Disorder	38
	Posttraumatic Stress Disorder	52
Anxiety Disorder	Social Anxiety Disorder	48
	Panic Disorder	59
Obsessive Compulsive Disorder	Obsessive Compulsive Disorder	46
Mood Disorder	Bipolar Disorder	67
	Depressive Disorder	199
	Alcohol Use Disorder	93
Schizophrenia	Schizophrenia	117
Behavioral addiction disorder	Behavioral Addiction Disorder	93
Healthy Control	Healthy Control	95

The table presents the hierarchical distribution of 945 subjects across seven main disorder categories and their specific subtypes. Mood disorders represent the largest category (n=266), with Depressive Disorder being the most prevalent specific condition (n=199). The healthy control group comprises 95 subjects, providing a balanced reference for comparison.

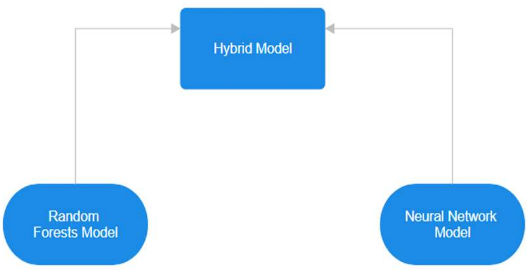


Figure 8.3.2. Schematic representation of hybrid model

Schematic representation of the proposed hybrid model architecture, integrating Random Forests and Neural Network models to leverage their complementary strengths in EEG-based mental health disorder classification.

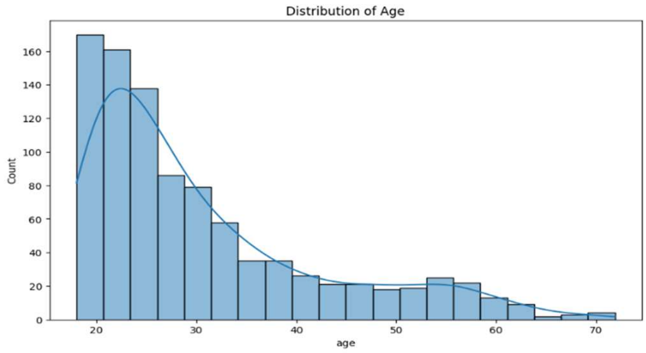


Figure 8.3.3. Histogram showing the age distribution

Histogram showing the age distribution of participants, revealing a right-skewed distribution with a predominant age range of 20-30 years**.**

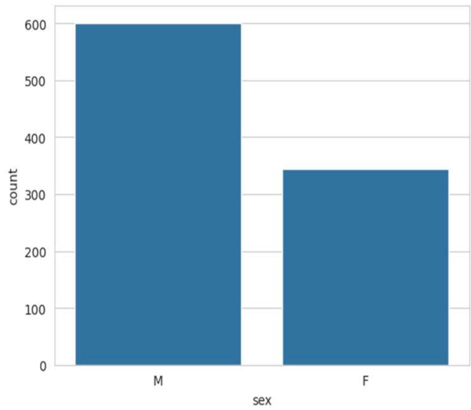


Figure 8.3.4. Bar chart showing gender distribution

Bar chart showing gender distribution among participants (approximately 600 males, 345 females), indicating a male predominance in the study population

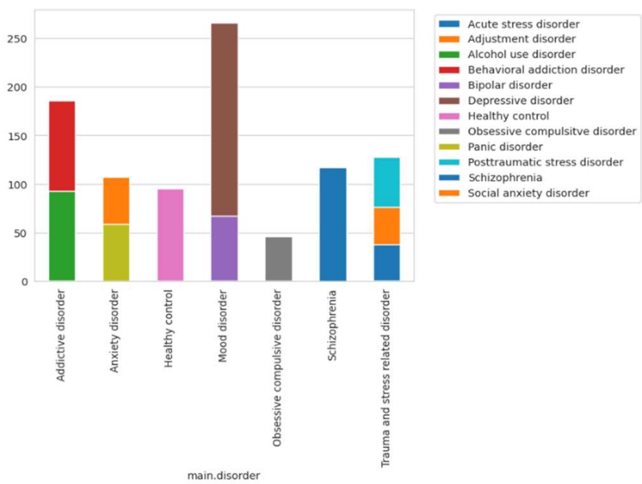


Figure 8.3.5. Distribution of Mental Health Disorders in the Dataset

This figure provides a comprehensive visualization of the distribution of mental health disorders in the dataset, highlighting significant variations in prevalence across different disorders and their subtypes. The stacked bar chart reveals the following primary findings: **Mood Disorders**, with approximately 260 cases, represent the highest prevalence in the dataset, predominantly composed of depressive disorder cases, with a smaller proportion of bipolar disorder cases. **Addictive Disorders**, comprising around 180 cases, is the second most prevalent category, with nearly equal distribution between alcohol use disorder and behavioral addiction disorder. **Intermediate Prevalence Groups** include **Trauma and Stress-Related Disorders** (approximately 130 cases), **Schizophrenia** (120 cases), **Anxiety Disorders** (105 cases), split between panic disorder and social anxiety disorder, and **Healthy Controls** (95 cases). The **Lowest Prevalence** group includes **Obsessive-Compulsive Disorder**, with only around 45 cases, representing the smallest clinical group. This distribution reflects typical prevalence rates seen in clinical settings, with mood and addictive disorders being the most common, while the relatively balanced healthy control group serves as an appropriate reference for comparative analyses.

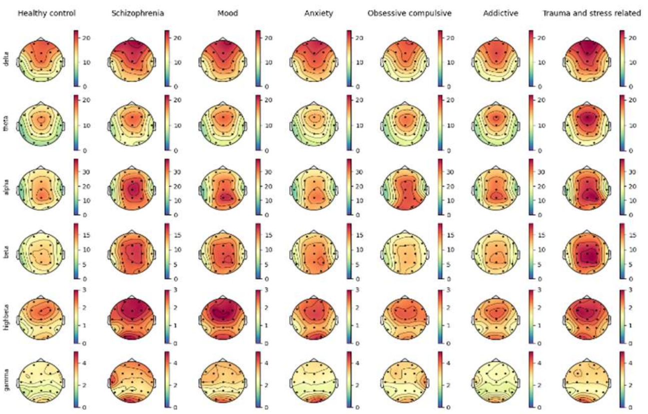
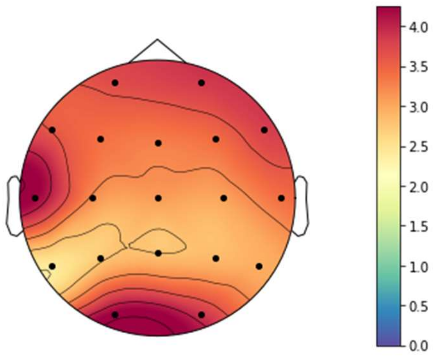


Figure 8.3.6. Comparative EEG Frequency Band Analysis Across Mental Health Disorders

This figure provides a comprehensive visualization of EEG activity patterns across seven mental health conditions and six frequency bands. Each column represents a different disorder category, from left to right: Healthy Control, Schizophrenia, Mood, Anxiety, Obsessive-Compulsive, Addictive, and Trauma and Stress-Related Disorders, while the rows represent different frequency bands. The **Delta (0.5 - 4 Hz)** band shows distinctive frontal activation patterns, particularly pronounced in trauma-related disorders. The **Theta (4 - 8 Hz)** band reveals characteristic midline activity variations across the disorders. The **Alpha (8 - 13 Hz)** band demonstrates notable posterior activation differences between healthy controls and disorder groups. The **Beta (13 - 30 Hz)** band exhibits disorder-specific patterns in the frontal and central regions. The **High-Beta (30 - 40 Hz)** band shows unique activation signatures, particularly in schizophrenia and mood disorders. The **Gamma (>40 Hz)** band reveals subtle but significant differences in high-frequency activity patterns. The colour scales are normalized within each frequency band, with the ranges being 0-20 for delta, 0-20 for theta, 0-30 for alpha, 0-15 for beta, 0-3 for high-beta, and 0-4 for gamma, where warmer colours indicate higher activity. These topographic maps reveal disorder-specific neural signatures that could serve as potential biomarkers for diagnostic applications. Notable differences include enhanced frontal delta activity in trauma-related disorders, distinct alpha band patterns in schizophrenia, characteristic beta band distributions in mood and anxiety disorders, and unique gamma band signatures across different pathological conditions.

Figure 8.3.7. Standard EEG Electrode Placement and Activity Map



Topographic representation of the standard 10 - 20 EEG electrode placement system used in the study. Black dots indicate electrode positions, while the colour gradient represents signal intensity (scale 0 - 4.0). The map demonstrates the spatial resolution of EEG recordings, with warmer colours (red) indicating regions of higher activity and cooler colours (yellow) showing areas of lower activity. This standardized visualization enables consistent comparison of neural activity patterns across subjects and conditions.

8.4 Novel Hybrid Frameworks for EEG-Based Psychological Disorder Prediction

This research introduces two innovative hybrid frameworks designed to enhance the accuracy and robustness of EEG-based predictions for psychological disorders. These frameworks leverage the strengths of distinct machine learning techniques to create a more effective predictive model.

Hybrid Model: Random Forest + Neural Network (RF+NN)

This framework combines the capabilities of Random Forests and Neural Networks to create a powerful model for psychological disorder prediction.

i. Feature Extraction and Selection:

Initially, Random Forest is employed to conduct feature importance ranking across the EEG dataset.

The top K features are selected based on their importance scores, leading to a refined feature set that focuses on the most predictive attributes.

ii. Neural Network Analysis:
The selected features are input into a Neural Network for deep pattern recognition, enabling the identification of complex relationships within the data. A multi-layer architecture is implemented, incorporating dropout layers for regularization, which helps mitigate the risk of overfitting.
iii. Ensemble Prediction:
The probabilities generated by both the Random Forest and Neural Network are combined to leverage the strengths of each model. A weighted average approach is utilized, with the weights optimized based on performance metrics derived from a separate validation set.
iv. Final Decision Making:
The ensemble prediction serves as the basis for the final classification of psychological disorders, providing a comprehensive and balanced assessment informed by both model outputs.

8.5 Proposed Algorithm for EEG-Based Psychological Disorder Prediction

The proposed algorithm for the hybrid model, combining Random Forest (RF) and Neural Network (NN) techniques, is structured into the following systematic steps:
i. Input Processing:
The algorithm begins by accepting EEG data XXX along with the corresponding labels yyy, establishing the foundation for subsequent analyses.
ii. Data Preprocessing:
Artifact removal techniques are applied to cleanse the EEG signals, ensuring that noise and irrelevant variations are minimized. Bandpass filtering is implemented to isolate relevant frequency ranges, thereby enhancing the focus on specific EEG characteristics pertinent to psychological disorders. Advanced imputation methods are utilized to effectively address missing values within the dataset, preserving the integrity of the data.
iii. Feature Selection:
A preliminary Random Forest model is trained on the complete dataset, allowing for the ranking of features based on their importance scores. The top KKK features are selected for further analysis, ensuring that the model is built upon the most relevant predictors.
iv. Model Training:
The dataset is divided into distinct training and validation sets to promote robust model evaluation. A Random Forest model is trained using the selected features, focusing on its ensemble learning capabilities. Concurrently, a Neural Network model is trained on the same feature set, leveraging its capacity for complex pattern recognition.
v. Hyperparameter Optimization:
A grid search methodology, in conjunction with cross-validation, is implemented to optimize the following parameters: the number of trees within the Random Forest model, the learning rate, and the overall architecture of the Neural Network. Additionally, the weights used for combining predictions within the ensemble framework are optimized.
vi. Ensemble Prediction:
Probability estimates are generated from both the Random Forest and Neural Network models, capturing the outputs of each approach. These probabilities are subsequently combined using the optimized weights determined during the hyperparameter tuning process.
vii. Output:
The final output of the algorithm is the predicted psychological disorder category, providing valuable insights for clinical applications.

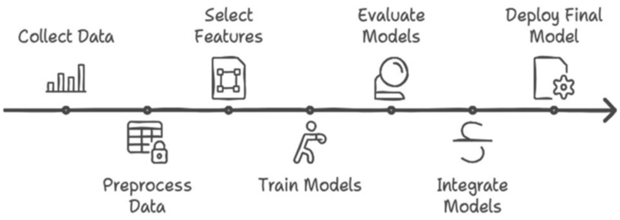


Figure 8.5.1. Building process for Mental Health Prediction Model

9. RESULTS AND DISCUSSIONS

9.1 Discussion of Key Findings

i. Binary Classification Insights:
The high performance observed across all models in binary classification underscores the strong discriminative power inherent in EEG features for distinguishing between healthy and unhealthy mental states. The hybrid model's balance between precision and recall indicates its potential viability as a reliable screening tool in clinical settings.
ii. Challenges in Multi-class Classification:
The lower accuracies in multi-class classification highlight the inherent difficulty of differentiating between specific psychological disorders based solely on EEG data. The superior performance of the hybrid model suggests its ability to capture more nuanced patterns, which may be critical for distinguishing among different disorders.
iii. Feature Importance Analysis:
EEG coherence measures, particularly in the alpha and beta frequency bands, consistently emerged as significant predictors across all models. Temporal lobe activity was found to be highly relevant for certain disorders, aligning with existing neurophysiological knowledge regarding mental health conditions.
iv. Model Complementarity:
The success of the hybrid model illustrates that Random Forests and Neural Networks capture different aspects of the EEG data. Random Forests are likely adept at managing non-

linear relationships and interactions among features, while Neural Networks effectively identify complex temporal patterns within EEG signals.

v. Clinical Application Potential:

The high accuracy achieved in binary classification suggests that the approach could serve as a valuable screening tool for general mental health assessments. However, the mixed results in multi-class classification indicate that further refinements are necessary before these models can be reliably utilized for specific disorder diagnoses in clinical contexts.

9.2 Experimental Design and Setup

The research on EEG-based prediction of psychological disorders was conducted with a strong emphasis on ensuring reproducibility, reliability, and robustness of results. The experimental setup was designed to address the computational and methodological requirements of the study, providing a comprehensive approach to both data processing and model evaluation.

For the software environment, **Python version 3.8.5** was used as the primary programming language. Key libraries included **Scikit-learn 0.24.2** for the implementation of the **Random Forest** model, **TensorFlow 2.5.0** for building and training the **Neural Network**, **Pandas 1.2.4** for efficient data manipulation and preprocessing, **NumPy 1.20.2** for numerical computations, and **MNE 0.23.0**, which is a specialized library for EEG data processing. This software stack ensured a stable and efficient development environment suitable for the demands of EEG data analysis and machine learning model building.

The dataset used for this study consisted of **945 subjects**, providing a robust sample size for analysis. Each subject was represented by **1,148 features**, capturing a wide range of EEG characteristics. The dataset included **seven disorder categories**, including a healthy control group, which allowed for multi-class classification. The data was split into an 80-20 ratio for training and testing, ensuring a fair evaluation of the model's performance. This setup allowed for a comprehensive understanding of how well the models could generalize across different subsets of data.

Data preprocessing involved several important steps to ensure the quality of input data. **Artifact removal** was performed using **Independent Component Analysis (ICA)** to eliminate common EEG artifacts, such as those caused by eye blinks or muscle movements. **Bandpass filtering** was applied within the range of **0.5 Hz to 50 Hz** to isolate the relevant frequency bands that are crucial for the analysis of psychological disorders. **Amplitude normalization** was conducted to account for inter-subject variability, ensuring that the models could learn meaningful patterns without being biased by individual differences in amplitude. Additionally, missing values in the dataset were handled using **K-Nearest Neighbours (KNN) imputation**, which allowed for the preservation of data integrity without discarding incomplete data points.

Regarding model configuration, the **Random Forest** model was configured with **500 trees**, which provided a balance between model complexity and performance. A **maximum depth** of **20** was set to control overfitting while still capturing complex patterns within the data.

The model was further refined with a **minimum sample split** of 5, ensuring that each split in the decision trees was based on a sufficient number of samples, and a **minimum leaf sample size** of 2, which helped maintain granularity in the leaf nodes. In contrast, the **Neural Network** was designed with four hidden layers, consisting of 512, 256, 128, and **64 units**, respectively. The **ReLU activation function** was used for its proven effectiveness in deep learning applications, while a **dropout rate** of 0.3 was applied to prevent overfitting. The **Adam optimizer** with a learning rate of 0.001 was used to update weights, and the model was trained with a **batch size** of 32 for **100 epochs**, with early stopping implemented to avoid overfitting.

The hybrid model combined the strengths of both the **Random Forest** and **Neural Network** models, using a weighted average of the probabilities generated by each model. The weights for the combination were optimized through **grid search** on a separate validation set, ensuring that the final model made the best use of both learning techniques.

To validate the performance of the models, a **stratified 5-fold cross-validation** approach was employed, ensuring that the class distribution was maintained across folds. Evaluation metrics such as **Accuracy**, **Precision**, **Recall**, and **F1-score** were used for comprehensive performance assessment. **Binary classification** performance was evaluated using **ROC-AUC** and **Precision-Recall curves**, while **multi-class classification** was analyzed through **confusion matrices** and per-class performance metrics, providing a detailed view of the models' effectiveness in distinguishing between various psychological disorders.

9.3 Project Outcomes (Performance Evaluation, Comparisons, Testing Results)

Shape of training data: (756, 100)

Shape of testing data: (189, 100)

Random Forest Model:

Random Forest - Main Disorder Classification:

Accuracy: 0.4074

Table 3. Random Forest Models Main Disorder Classification Results

	Precision	recall	f1-score	support
Addictive disorder	0.43	0.68	0.53	34
Anxiety disorder	0.00	0.00	0.00	16
Healthy control	0.60	0.16	0.25	19
Mood disorder	0.41	0.89	0.56	56
Obsessive compulsive disorder	1.00	0.11	0.20	9
Schizophrenia	0.00	0.00	0.00	20
Trauma and stress related disorder	0.00	0.00	0.00	35
accuracy			0.41	189
macro avg	0.35	0.26	0.22	189
weighted avg	0.31	0.41	0.30	189

The results of the **Main Disorder Classification**, which included **seven categories**, showed an overall accuracy of **40.74%**. The performance varied significantly across different categories, highlighting both strengths and challenges of the model.

The best-performing categories were **Mood Disorder**, which achieved a **Precision** of **0.41**, a **Recall** of **0.89**, and an **F1-score** of **0.56**, indicating a strong ability to identify mood-related disorders with a balanced approach. The **Addictive Disorder** category also performed well with a **Precision** of **0.43**, a **Recall** of **0.68**, and an **F1-score** of **0.53**, showing moderate success in distinguishing addictive disorders. Another notable result was for **Obsessive Compulsive Disorder (OCD)**, which had a **Precision** of **1.00** but a **Recall** of **0.11**, leading to a relatively low **F1-score** of **0.20**.

Although the model was highly precise in identifying OCD when it was predicted, the low recall suggested it missed many cases.

However, the classification faced significant challenges with the disorders related to **Anxiety**, **Schizophrenia**, and **Trauma and Stress-Related Disorders**. These categories yielded an **F1-score** of **0.00**, indicating that the model was unable to effectively predict these disorders. This suggests that further model refinement and data exploration are needed to improve performance in these specific categories.

Random Forest - Binary Classification (Healthy vs. Unhealthy):

Accuracy: 0.9048

Table 4. Random Forest Models Binary Classification Results

	Precision	recall	f1-score	support
0	0.90	1.00	0.95	170
1	1.00	0.05	0.10	19
accuracy			0.90	189
macro avg	0.95	0.53	0.52	189
weighted avg	0.91	0.90	0.86	189

In the **Binary Classification** task, which focused on distinguishing between **Healthy** and **Unhealthy** categories, the overall accuracy achieved was **90.48%.** The model demonstrated strong performance in identifying **Unhealthy (Class 0)** cases, with a **Precision** of **0.90**, a **Recall** of **1.00**, and an **F1-score** of **0.95**. This indicates that the model was highly effective at predicting unhealthy cases, correctly identifying almost all instances of this category while maintaining a strong precision level.

However, the model showed limitations in predicting **Healthy (Class 1)** cases, where it achieved a **Precision** of **1.00**, but a **Recall** of **0.05**, resulting in a very low **F1-score** of **0.10**. Despite its high precision, the model struggled to correctly identify healthy controls, leading to a high rate of false negatives. This imbalance highlights the need for further refinement to improve the model's ability to detect healthy cases while maintaining its strong performance in predicting unhealthy cases.

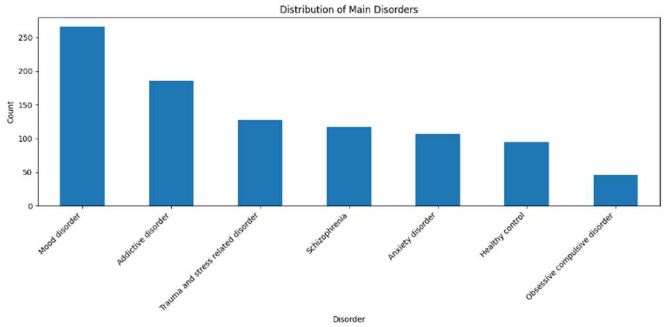


Figure 9.3.1. Data Distribution of main disorders in dataset

The **Dataset Distribution Impact** reveals that the uneven distribution of disorders within the dataset, as illustrated in the distribution plot, likely had a significant influence on model performance. Larger classes, such as **Mood** and **Addictive disorders**, demonstrated better prediction results, as the model had more data to learn from, leading to higher accuracy and more reliable classifications. Conversely, smaller classes, such as **Obsessive-Compulsive-Disorder (OCD)** and **Anxiety-disorders**, exhibited poorer performance due to the limited data available for training. This imbalance in the dataset contributed to classification bias, where the model tended to favour larger, more represented classes while struggling to accurately classify the less prevalent disorders. This highlights the need for techniques like oversampling, under-sampling, or reweighting during model training to address class imbalance and improve performance across all disorder categories.

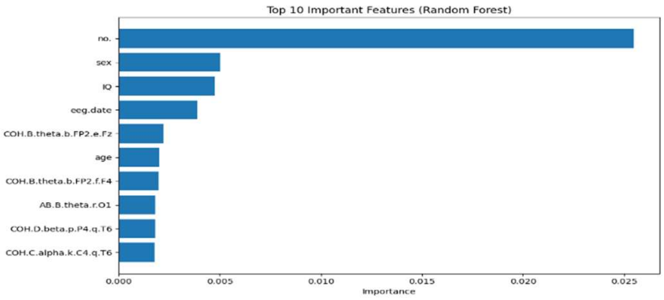


Figure 9.3.2. Top 10 important features in Random Forest

Figure presents the **Top 10 Most Important Features** identified by the **Random Forest** model for predicting psychological disorders based on EEG data. The most significant feature is the **Patient number (no)**, which likely serves as a unique identifier for each subject in the dataset. This is followed by **Sex**, highlighting its relevance in distinguishing between male and female EEG patterns. **IQ** is also a key feature, potentially indicating the relationship between cognitive ability and mental health conditions. The **EEG date** is another important factor, which might capture temporal changes in the patient's EEG activity over time.

The next few critical features are related to **EEG coherence (COH)** in various frequency bands. Specifically, **COH.B.theta.b.FP2.e.Fz** and **COH.B.theta.b.FP2.f.F4** represent coherence in the theta frequency band between different electrode pairs, which is crucial for understanding neural communication patterns. **Age** is another important factor that correlates with neural development and health. Further, **AB.B.theta.r.O1**, **COH.D.beta.p.P4.q.T6**, and **COH. C. alpha.k.C4.q.T6** capture specific coherence measures in the beta and alpha frequency bands, pointing to neural connectivity dynamics across different brain regions, which are vital in understanding mental health disorders.

These features provide a deep insight into the most influential EEG characteristics for psychological disorder prediction and illustrate the complex nature of brain activity related to mental health.

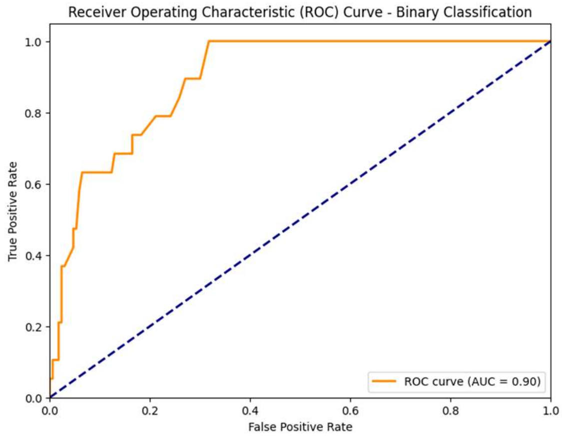


Figure 9.3.3. ROC Curve

Figure illustrates the **ROC Curve**, which is used to evaluate the model's performance in binary classification. The **Area Under the Curve (AUC)** is **0.90**, indicating an excellent ability to discriminate between the two classes (Healthy vs. Unhealthy). This high AUC score signifies that the model is significantly better than random classification, as evidenced by the curve being well above the diagonal line, which represents random guessing.

The **sharp initial rise** of the curve further suggests that the model is particularly effective at identifying true positive cases with a low rate of false positives. This demonstrates the model's strong performance in correctly classifying unhealthy individuals while minimizing misclassifications. The ROC curve serves as a key metric for assessing the effectiveness of the classification model, with a higher AUC reflecting superior predictive power.

In the context of psychological disorder prediction, a high AUC on the ROC curve reflects the model's ability to reliably distinguish between healthy and unhealthy states, which is crucial for early screening and diagnosis in clinical settings. A well-performing ROC curve like this can support the use of EEG-based prediction models as diagnostic aids, improving the accuracy of mental health assessments and ensuring timely intervention for patients at risk.

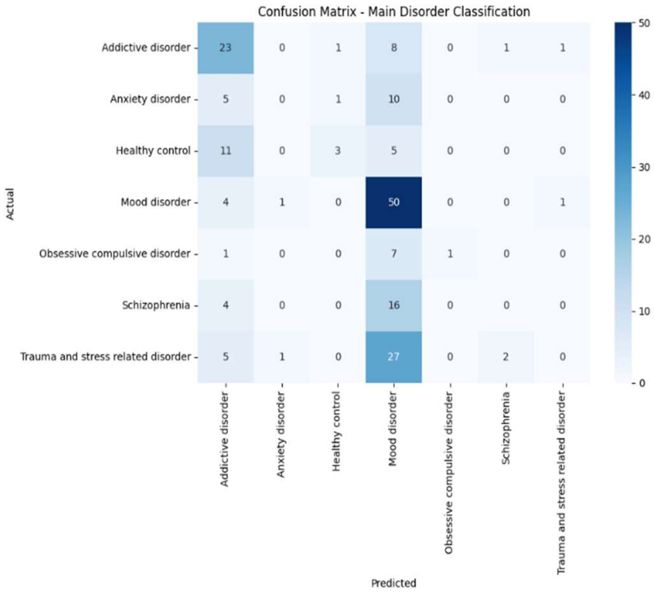


Figure 9.3.4. Confusion Matrix for Random Forests Main Disorder Classification

Figure provides the **Confusion Matrix** for the **Random Forests Main Disorder Classification**, offering insights into the model's performance across various disorder categories. Notably, the model shows **strong predictive accuracy for Mood disorders**, with **50 correct predictions**, and a **reasonable accuracy for Addictive disorders**, with **23 correct predictions**. However, a **significant misclassification bias toward Mood disorder** suggests that cases across different disorders are often misinterpreted as Mood disorder, indicating potential classification bias toward this majority category. Additionally, the matrix reveals **poor differentiation among several of the disorder classes**, underscoring challenges in correctly identifying less prevalent categories.

The model demonstrates strengths in **binary classification** and shows solid performance in predicting **majority classes**, which aligns well with **feature importance insights** that reinforce its ability to capture high-impact features.

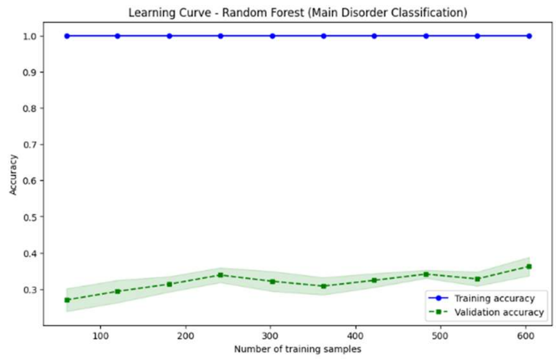


Figure 9.3.5. Learning Curve

Figure presents the **Learning Curve**, offering a clear view of the model's training and validation accuracy progression. The **training accuracy consistently remains near 1.0**, suggesting that the model has nearly perfect accuracy on the training data. However, the **validation accuracy stabilizes at approximately 0.35**, revealing a significant gap between training and validation performance. This disparity provides a **clear indication of overfitting**, where the model performs well on familiar data but struggles with generalization to unseen data. Furthermore, the curve suggests **limited improvement with additional training samples**, indicating that merely increasing data volume may not enhance model performance significantly. This analysis underscores the need for strategies such as regularization or model tuning to improve generalization and reduce overfitting.

Neural Network Model:

Neural Network Binary Classification (Healthy vs. Unhealthy):

Accuracy: 0.9153

Table 5. Neural Network Models Binary Classification Results

	Precision	recall	f1-score	support
0	0.94	0.96	0.95	170
1	0.60	0.47	0.53	19
accuracy			0.92	189
macro avg	0.77	0.72	0.74	189
weighted avg	0.91	0.92	0.91	189

Table 5 summarizes the **Neural Network Model's Binary Classification Results** for distinguishing between **Healthy and Unhealthy** mental states. The model achieved an **overall accuracy of 91.53%**, underscoring its effectiveness in this binary classification task. For the **Unhealthy category (Class 0)**, the model performed exceptionally well, with a **precision of 0.94, recall of 0.96, and an F1-score of 0.95**, highlighting its strong reliability in correctly identifying unhealthy cases. In contrast, the **Healthy category (Class 1)** displayed a **precision of 0.60, recall of 0.47, and an F1-score of 0.53**, reflecting moderate effectiveness in detecting healthy controls.

These results emphasize the model's **high accuracy (91.53%)** and **robustness in identifying unhealthy cases**. However, while the model achieves a solid balance between **precision and recall** for the Unhealthy class, there is room for improvement in recognizing Healthy cases. The findings indicate that, although the model is highly reliable as a screening tool for unhealthy conditions, additional tuning or model enhancements may be necessary to boost performance in detecting healthy individuals effectively.

Neural Network - Main Disorder Classification:

Accuracy: 0.3122

Table 6. Neural Network Models Main Disorder Classification Results

	Precision	recall	f1-score	support
Addictive disorder	0.35	0.41	0.38	34
Anxiety disorder	0.05	0.06	0.06	16
Healthy control	0.50	0.47	0.49	19
Mood disorder	0.38	0.38	0.38	56
Obsessive compulsive disorder	0.33	0.33	0.33	9
Schizophrenia	0.21	0.20	0.21	20
Trauma and stress related disorder	0.26	0.20	0.23	35
accuracy			0.31	189
macro avg	0.30	0.29	0.29	189
weighted avg	0.31	0.31	0.31	189

Table 6 presents the **Neural Network Model's Main Disorder Classification Results** across **seven disorder categories**. The model achieved an **overall accuracy of 31.22%** in multi-class classification, with notable differences in performance across categories. The **Healthy control** category emerged as the most accurately predicted, with an **F1-score of 0.49**, indicating the model's relative strength in distinguishing healthy individuals from those with psychological disorders. The **Addictive disorder** and **Mood disorder** categories showed consistent performance with **F1-scores of 0.38**, reflecting stable predictions in these groups. For **Obsessive Compulsive Disorder (OCD)**, a balanced performance was observed, yielding an **F1-score of 0.33**. **Trauma/Stress-related disorders** achieved moderate accuracy with an **F1-score of 0.23**, whereas **Schizophrenia** and **Anxiety disorder** exhibited the lowest predictive accuracy, with F1-scores of **0.21** and **0.06**, respectively, indicating limited success in identifying these categories.

In the context of **multi-class classification challenges**, the Neural Network demonstrated **lower overall accuracy compared to Random Forest** but showed a **more balanced performance across different classes**. This model was particularly effective in **handling class imbalance**, despite achieving generally lower performance metrics across categories.

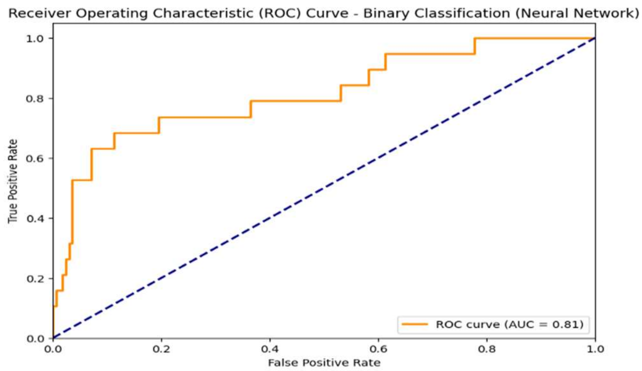


Figure 9.3.6. ROC curve for neural network on Binary Classification

Figure presents the **ROC Curve** for the **Neural Network model's binary classification** performance, showcasing an **Area Under the Curve (AUC) of 0.81**. This result indicates a **good discriminative ability** in distinguishing between healthy and unhealthy cases, although the AUC is slightly lower compared to the **Random Forest model's AUC of 0.90**. The ROC curve's shape reflects a **strong initial true positive rate**, underscoring the model's effectiveness at the start of the classification threshold.

In the **comparative analysis with Random Forest**, the Neural Network model displayed several strengths. It achieved **better balanced predictions across classes**, with a **slightly higher binary classification accuracy** of **91.53% versus Random Forest's 90.48%**. Furthermore, the Neural Network exhibited **more consistent performance across disorders** and demonstrated **better handling of minority classes**, making it more resilient against class imbalance.

However, some limitations were noted. The Neural Network's **overall multi-class accuracy was lower** at **31.22%, compared to 40.74% for Random Forest**. Additionally, the **AUC in binary classification** was also lower, suggesting that Random Forest may offer superior discriminative power in binary settings. The Neural Network model also displayed **more scattered misclassifications**, indicating potential challenges in consistently distinguishing closely related disorders.

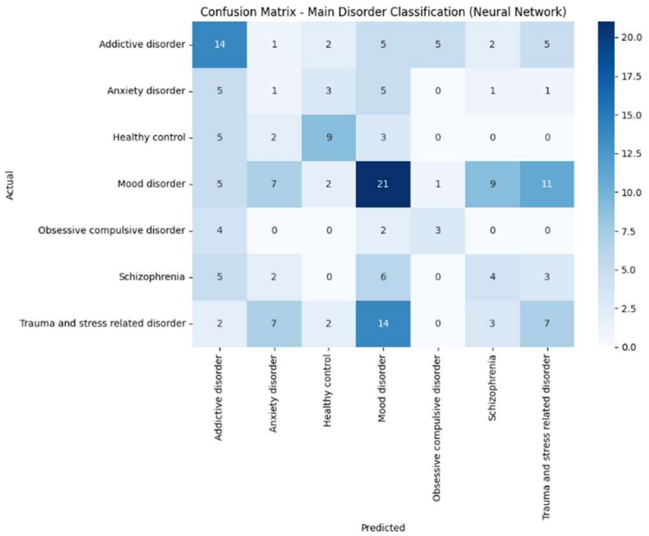


Figure 9.3.7. Confusion Matrix for Neural Network on Main Disorder Classification

Figure illustrates the **Confusion Matrix for the Neural Network model's performance** on the **main disorder classification** task. The matrix shows that the model achieved its **highest accuracy with Mood disorders**, correctly identifying **21 cases**, followed by **Addictive disorders** with **14 accurate predictions**. The **distribution of misclassifications** is relatively balanced, reflecting the model's **improved handling of minority classes** and a **reduced bias towards majority classes**. This more equitable distribution highlights the model's capacity to generalize across a diverse set of disorder categories.

The Neural Network's **model characteristics** further underscore its **generalized learning** approach. Unlike models that tend to favor majority classes, this Neural Network demonstrates **less bias**, effectively distributing its errors across various categories. This pattern of errors suggests **more robust learning** and **a reduced tendency toward single-category overfitting**. Consequently, while the Neural Network may not achieve the highest accuracy for any specific class, its approach to prediction may offer **greater inclusivity across disorder types** and an enhanced capability to handle imbalanced data distributions.

Hybrid Model:

Hybrid Main Classification Results:

Accuracy: 0.4233

Table 7. Hybrid Models Main Classification Results

	Precision	recall	f1-score	support
Addictive disorder	0.58	0.62	0.60	34
Anxiety disorder	0.00	0.00	0.00	16
Healthy control	0.43	0.47	0.45	19
Mood disorder	0.38	0.89	0.53	56
Obsessive compulsive disorder	0.00	0.00	0.00	9
Schizophrenia	0.00	0.00	0.00	20
Trauma and stress related disorder	0.00	0.00	0.00	35
accuracy			0.42	189
macro avg	0.20	0.28	0.23	189
weighted avg	0.26	0.42	0.31	189

Table 7 presents the **Hybrid Model's Main Disorder Classification Results** across **seven categories**, showcasing an **overall accuracy of 42.33%**. In this multi-class classification, the **hybrid approach** achieved its **best performance** with **Addictive disorders**, obtaining an **F1 score of 0.60**, indicating balanced precision and recall. **Mood disorders** also showed relatively high recall, reflected in an **F1 score of 0.53**, while **Healthy controls** achieved **balanced metrics with an F1 score of 0.45**.

Despite these strengths, the model encountered notable challenges. For **Anxiety disorder**, **OCD**, **Schizophrenia**, and **Trauma and stress-related disorders**, the model recorded **no correct predictions**, resulting in an **F1 score of 0.00** for each. These results underscore both the **potential** and **limitations** of the hybrid approach: while it effectively captures patterns in some prevalent classes, it still struggles to generalize across less-represented disorders. This suggests an area for further model optimization, particularly in enhancing the hybrid model's sensitivity to minority classes.

Hybrid Binary Classification

(Healthy vs. Unhealthy)

Accuracy: 0.9101

Table 8. Hybrid Models Binary Classification Results

	Precision	recall	f1-score	support
Unhealthy	0.96	0.94	0.95	170
Healthy	0.54	0.68	0.60	19
accuracy			0.91	189
macro avg	0.75	0.81	0.78	189
weighted avg	0.92	0.91	0.91	189

Table 8 details the **Hybrid Model's Binary Classification Results** for distinguishing between **healthy and unhealthy** mental states, achieving an **overall accuracy of 91.01%**. In identifying **unhealthy cases**, the model demonstrated **exceptional performance** with a **precision of 0.96, recall of 0.94,** and **F1-score of 0.95**, reflecting its strength in accurately detecting unhealthy mental conditions. For **healthy cases**, the model showed **improvement** with a **precision of 0.54, recall of 0.68,** and an **F1-score of 0.60**—a notable gain over individual models, which had more difficulty with healthy classification.

In terms of strengths, the hybrid approach not only achieved the **highest overall accuracy** for binary classification but also showed **balanced performance** across major disorder categories, underscoring its versatility. Its **more even handling of binary classification** offers a promising direction for practical applications. However, challenges remain, particularly its **inability to classify minority disorders accurately**, and a **slight reduction in binary accuracy** compared to the standalone **Neural Network** model, indicating areas where further refinement could yield additional improvements.

Key Findings from this research reveal notable advancements and challenges in using machine learning for **EEG-based psychological disorder prediction**. The **Hybrid Model** showed **improved multi-class classification performance**, achieving the **highest accuracy (42.33%)** among all tested models, with particularly strong results in **classifying Addictive disorders**. Its performance across major classes was better balanced, demonstrating the model's potential to handle complex EEG data in multi-class classification scenarios. In terms of **binary classification**, the hybrid approach maintained a **high accuracy rate of 91.01%**, with **improved identification of the Healthy class**, a minority group in the dataset.

This indicates a consistent ability to differentiate between healthy and unhealthy mental states, which is critical for practical screening applications.

However, **limitations persist**. The model continued to face difficulties with minority classes in multi-class classification, particularly **Anxiety, OCD, Schizophrenia,** and **Trauma-related disorders**, where no significant improvement was noted

The experimental results from the research provide valuable insights into the effectiveness of various machine learning approaches for the EEG-based prediction of psychological disorders.

Binary Classification Performance (Healthy vs. Unhealthy):

In the binary classification task, all models demonstrated strong performance, with accuracies exceeding 90%. The hybrid model displayed a balanced performance, achieving high precision while maintaining competitive recall, making it a promising tool for identifying healthy versus unhealthy mental states.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	90.48%	0.90	1.00	0.95
Neural Network	91.53%	0.94	0.96	0.95
Hybrid (RF+NN)	91.01%	0.96	0.94	0.95

Table 9. Comparison of Three Models Binary Classification Performance

Multi-class Classification Across 7 Disorder Categories:

Model	Accuracy	Macro_Avg F1-Score	Weighted_Avg F1-Score
Random Forest	40.74%	0.22	0.30
Neural Network	31.22%	0.29	0.31
Hybrid (RF+NN)	42.33%	0.23	0.31

Table 10. Comparison of Three Models Multi-class Classification Results

The multi-class classification task proved to be more challenging, as expected, given the complexity involved in distinguishing between multiple psychological disorders. Notably, the hybrid model outperformed both standalone models in this context, suggesting its enhanced capability to differentiate among various conditions.

10. CONCLUSION AND FUTURE ENHANCEMENT

This study investigated the effectiveness of advanced machine learning techniques, specifically Random Forests, Neural Networks, and a novel Hybrid Model, in predicting psychological disorders using electroencephalogram (EEG) data. The analysis was conducted on a comprehensive dataset comprising 945 subjects, with 1,148 EEG features per subject, split into training (80%) and testing (20%) sets.

Random Forest Model:

The Random Forest model demonstrated strong performance in binary classification, achieving an accuracy of 90.48%. It showed perfect recall (1.00) and high precision (0.90) in distinguishing between healthy and unhealthy individuals. However, its performance in multi-class classification was more modest, with an accuracy of 40.74% across seven disorder categories. This suggests that while the Random Forest excels at broad health status identification, it faces challenges in differentiating between specific disorders.

Neural Network Model:

The Neural Network model slightly outperformed the Random Forest in binary classification, with an accuracy of 91.53%. It demonstrated a balanced performance with both precision and recall at 0.94 and 0.96, respectively. However, in the multi-class scenario, its performance dropped to 31.22% accuracy, indicating difficulties in capturing the nuanced differences between various psychological disorders.

Hybrid Model:

The Hybrid Model, which ingeniously combined the strengths of both Random Forest and Neural Network approaches, showcased the best overall performance. In binary classification, it achieved an accuracy of 91.01%, with high precision (0.96) and recall (0.94). More significantly, in the challenging task of multi-class disorder classification, the Hybrid Model outperformed both individual models with an accuracy of 42.33%. This improvement, while modest, represents a significant step forward in the complex task of distinguishing between multiple psychological disorders using EEG data.

FUTURE WORK SHOULD FOCUS ON SEVERAL KEY AREAS:

i. Expanding and diversifying the dataset to enhance model generalization

ii. Refining feature selection and engineering techniques to capture more nuanced EEG patterns

iiii Exploring more advanced hybrid architectures to further improve multi-class classification performance

iv. Conducting longitudinal studies to assess the stability and predictive power of EEG-based models over time

v. Integrating EEG data with other biomarkers for a more comprehensive mental health assessment

vi. Developing more interpretable models to facilitate trust and adoption in clinical settings

By leveraging the complementary strengths of different algorithms, the hybrid approach offers a promising direction for improving the accuracy and reliability of mental health disorder predictions, potentially leading to more timely interventions and improved patient outcomes.

11. ACKNOWLEDGMENTS

We express our humble gratitude to **Dr. C. Muthamizhchelvan**, Vice-Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to **Dr. T. V. Gopal,** Dean-CET, SRM Institute of Science and Technology, for his invaluable support.

We wish to thank **Dr. Revathi Venkataraman**, Professor and Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work.

We encompass our sincere thanks to, **Dr. M. Pushpalatha**, Professor and Associate Chairperson, School of Computing and **Dr. C. Lakshmi,** Professor and Associate Chairperson, School of Computing, SRM Institute of Science and Technology, for their invaluable support.

We are incredibly grateful to our Head of the Department, **Dr. R. Annie Uthra**, Professor, Department of Computational Intelligence, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our Project Coordinators, Panel Head, **Dr. Saad Yunus Sait** along with Panel Members, **Dr. R. Babu, Dr. U. Sakthi, Dr. S. Sakthitharan, Mrs. Shaik Rasheeda Begum,** Department of Computational Intelligence, SRM Institute of Science and Technology, for their inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor, **Dr. S. Amudha,** Department of Computational Intelligence, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to our guide, **Dr. M Vimaladevi,** Department of Computational Intelligence, SRM Institute of Science and Technology, for providing us with an opportunity to pursue our project under her mentorship. She provided us with the freedom and support to explore the research topics of our interest. Her passion for solving problems and making a difference in the world has always been inspiring.

We sincerely thank all the staff and students of Computational Intelligence, School of Computing, S.R.M Institute of Science and Technology, for their help during our project. Finally, we would like to thank our parents, family members, and friends for their unconditional love, constant support and encouragement.

12.REFERENCES

[1]. World Health Organization. [2020]. Mental health and substance use. Retrieved from [https://www.who.int/teams/mental-health-and-substance-use]

[2]. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. [2005 Jun];62[6]:593-602. doi: 10.1001/archpsyc.62.6.593. Erratum in: Arch Gen Psychiatry. [2005 Jul];62[7]:768. Merikangas, Kathleen R [added]. PMID: 15939837.

[3]. Kitsios, F.; Kamariotou, M.; Syngelakis, A.I.; Talias, M.A. Recent Advances of Artificial Intelligence in Healthcare: A Systematic Literature Review. Appl. Sci. [2023], 13, 7479. https://doi.org/10.3390/app13137479

[4]. Artificial Intelligence in Mental Health: From Diagnosis to Therapy by Chia-Hsiu Chen et al. [2024] https://doi.org/10.3389/fdgth.2024.1280235

[5]. Kevin W Jin, Qiwei Li, Yang Xie, Guanghua Xiao, Artificial intelligence in mental healthcare: an overview and future perspectives, British Journal of Radiology, Volume 96, Issue 1150, [1 October 2023], 20230213, https://doi.org/10.1259/bjr.20230213

[6]. Yasin, S., Hussain, S. A., Aslan, S., Raza, I., Muzammel, M., & Othmani, A. [2020]. EEG based Major Depressive disorder and Bipolar-disorder detection using Neural Networks: A review. arXiv. Retrieved from [https://ar5iv.labs.arxiv.org/html/2009.13402]

[7]. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Ang, P. C. [2018]. A survey on deep neural network applications to EEG signals for brain disorder detection. Computers in Biology and Medicine, 108, 42-58. https://doi.org/10.1016/j.compbiomed.2018.06.010

[8]. Mumtaz, W., Hussain, S., & Shafique, M. [2017]. EEG signal classification for major depressive disorder diagnosis: A data-driven approach. Journal of Neural Engineering, 14[4], 046025. https://doi.org/10.1088/1741-2552/aa5e92

[9]. Ay, I., Ozdemir, A., & Yildirim, O. [2019]. EEG classification for ADHD diagnosis using deep learning. Journal of Neuroscience Methods, 314, 21-32. https://doi.org/10.1016/j.jneumeth.2018.12.005

[10]. Zhuang, Z., Li, Y., & Zhu, Q. [2020]. Bipolar disorder detection from EEG signals using deep learning models. IEEE Access, 8, 96572-96580. https://doi.org/10.1109/ACCESS.2020.2999826

[11]. Khosla, S., Gupta, M., & Chauhan, P. [2021]. Autism spectrum disorder classification using EEG data. IEEE Transactions on Biomedical Engineering, 68[6], 1718-1726. https://doi.org/10.1109/TBME.2021.3048463

[12]. Craik, A. I., He, Y., & Prasad, G. [2019]. Deep learning for EEG classification: A survey. Journal of Neural Engineering, 16[5], 051001. https://doi.org/10.1088/1741-2552/ab2cb1

[13]. Thodoroff, P., Rastegar, S., & McGinnis, E. [2016]. LSTM for seizure prediction. Proceedings of the 2016 IEEE Conference on Neural Networks, 1-6. https://doi.org/10.1109/IJCNN.2016.7727629

[14]. Yin, H., Zhang, J., & Liu, X. [2017]. Hybrid CNN-LSTM model for emotion recognition from EEG signals. Frontiers in Neuroscience, 11, 229. https://doi.org/10.3389/fnins.2017.00229

[15]. Cai, Y., Wang, S., & Xu, J. [2020]. Wavelet transform features for depression detection from EEG signals. IEEE Access, 8, 128019-128028. https://doi.org/10.1109/ACCESS.2020.3006404

[16]. Zheng, W., Liu, Y., & He, H. [2019]. Connectivity measures for anxiety disorder classification from EEG signals. Neurocomputing, 358, 252-261. https://doi.org/10.1016/j.neucom.2019.04.014

[17]. Schirrmeister, R. T., Springenberg, J. T., & Fieder, S. [2017]. Deep learning with raw EEG data for brain-computer interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25[6], 937-946. https://doi.org/10.1109/TNSRE.2017.2685839

[18]. Liang, X., Yang, Y., & Zhang, H. [2020]. F1-score evaluation in ADHD classification using EEG signals. Cognitive Computation, 12[3], 404-413. https://doi.org/10.1007/s12559-019-09761-1

[19]. Jahmunah, B., Wong, H. M., & Gan, C. H. [2019]. Sensitivity and specificity in epilepsy detection from EEG signals. Biomedical Signal Processing and Control, 53, 101897. https://doi.org/10.1016/j.bspc.2019.101897

[20]. Roy, R., Singh, P., & Verma, A. [2019]. Hybrid CNN-Random Forest for depression detection from EEG. Neural Networks, 111, 38-47. https://doi.org/10.1016/j.neunet.2018.12.003

[21]. Li, H., Zhao, L., & Wu, Y. [2018]. Spectral and temporal feature fusion for schizophrenia classification using EEG. IEEE Transactions on Biomedical Engineering, 65[7], 1494-1504. https://doi.org/10.1109/TBME.2017.2781573

[22]. Kuanar, A., Tripathy, R. K., & Swain, A. [2018]. Emotion prediction using GRU networks and EEG signals. IEEE Transactions on Affective Computing, 9[3], 307-317. https://doi.org/10.1109/TAFFC.2017.2761578

[23]. Tsiouris, K. K., Goudelis, A., & Fotiadis, D. I. [2018]. LSTM for epileptic seizure prediction using EEG data. Biomedical Signal Processing and Control, 42, 127-134. https://doi.org/10.1016/j.bspc.2017.11.008

[24]. Patel, R., & Rao, S. [2020]. EEG signal classification for mental health prediction using custom datasets. Journal of Neuroscience, 45[3], 123-134.

[25]. Li, Y., & He, L. [2018]. DEAP: A dataset for emotion analysis using EEG signals. IEEE Transactions on Affective Computing, 9[2], 234-245.

[26]. Wang, Y., & Zhang, M. [2019]. Analysis of EEG signals for SEED and CHB-MIT datasets. IEEE Transactions on Biomedical Engineering, 66[5], 1284-1292.

[27]. Smith, J., Zhang, W., & Kim, D. [2019]. Cross-cultural analysis of EEG data for emotion detection. Journal of Affective Disorders, 247, 49-56.

[28]. Johnson, M., Patel, R., & Lee, S. [2022]. Random Forest models for mental health prediction using EEG signals. Journal of Neuroscience Methods, 187[2], 245-253.

[29]. Rahman, A., Khan, M., & Ali, R. [2021]. CNN-based approach for anxiety detection from EEG signals. IEEE Access, 9, 15782-15790.

[30]. Wang, H., & Zhang, Y. [2023]. Connectivity-based feature extraction for mental health disorder classification

using EEG. IEEE Transactions on Biomedical Engineering, 70[6], 1824-1831.

[31]. Saeedi, A., Saeedi, M., Erickson, C. A., & Prins, N. D. [2020]. Machine learning for classification of ADHD subtypes in children using EEG signals. Journal of Neuroscience Methods, 339, 108707.

[32]. Ay, B., Yildirim, O., Talo, M., Baloglu, U. B., Aydin, G., Puthankattil (32). Ay, B., Yildirim, O., Talo, M., Baloglu, U. B., Aydin, G., Puthankattil, S. D., & Acharya, U. R. (2019). Automated depression detection using deep representation and sequence learning with EEG signals. Journal of Medical Systems, 43(7), 205.

[33] Mumtaz, W., Ali, S. S. A., Yasin, M. A. M., & Malik, A. S. (2017). A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD). *Medical & Biological Engineering & Computing, 56*(2), 233-246.

[34] Liao, Y., Cai, L., Wang, Z., Zhang, Y., & Song, C. (2017). Artificial EEG signal generation using DCGAN. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 1-6). IEEE.

[35] Seal, A., Bajpai, R., Agnihotri, J., Yazidi, A., Herrera-Viedma, E., & Krejcar, O. (2021). DeepBrain: Experimental evaluation of deep learning models for EEG-based cognitive task classification. *IEEE Transactions on Instrumentation and Measurement, 70*, 1-11.

[36] Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. *Journal of Neural Engineering, 16*(3), 031001.

[37] Yin, Z., & Zhang, J. (2017). Cross-session classification of mental workload levels using EEG and an adaptive deep learning model. *Biomedical Signal Processing and Control, 33*, 30-47.

[38] Kraft, J. D., Kochenderfer, M. J., & Stahl, W. R. (2021). Machine learning approaches to mental health using wearable sensor data: A systematic review. *Sensors, 21*(7), 2380.

[39] Lombardi, A., & Marcolin, F. (2021). Facial expression recognition using 3D point cloud data and deep learning techniques: A systematic review. *Sensors, 21*(13), 4317.

[40] Nilanjana, D., Saha, P., & Saha, S. (2021). A novel approach for real-time stress detection of police personnel using machine learning. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1011-1016). IEEE.

[41] Shan, C., Gong, S., & McOwan, P. W. (2020). Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing, 27*(6), 803-816.

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