Hybrid Method of Detecting Mental Health Disorder Through Neural Networks & Random Forest

B Vidhya[1], Arun A[2], Nithish Kumar W[3]
Dept of Computational Intelligence SRM Institute of Science and Technology, KTR Campus Chennai, India

***Abstract*—Mental health disorders pose a major global challenge, necessitating early detection for effective intervention. This approach proposes a hybrid classification framework integrating Random Forest and Neural Network (RF+NN) to enhance psychological disorder prediction using EEG data. Additionally, this technique aims to establish a foundation for personalized treatment strategies by identifying EEG-based biomarkers associated with specific mental health conditions. Evaluations on a dataset of 945 subjects with 1,148 EEG features demonstrate the hybrid model's superior performance, achieving 91.01% accuracy in binary classification (healthy vs. unhealthy) and 42.33% in multi-class classification, surpassing individual RF and NN models. The findings emphasize the advantages of hybrid classifiers in mental health assessment, particularly in feature selection and model interpretability, highlighting their potential for improving diagnostic accuracy.**

***Index Terms*—Mental health, EEG, Machine Learning, Neural Networks, Hybrid Models, Random Forests**

I. INTRODUCTION

Mental health disorders are a significant global concern, impacting millions of individuals and contributing to widespread disability. The World Health Organization (WHO) estimates that around 450 million people suffer from mental or neurological conditions, making early detection and intervention is crucial for effective management. However, the complexity and variability of these disorders pose challenges in accurate diagnosis and prediction.

Traditional diagnostic methods often rely on subjective assessments and predefined clinical criteria, which may lead to inconsistencies or delays in diagnosis. To overcome these limitations, there is a need for data-driven, objective approaches to mental health evaluation. Recent advancements in artificial intelligence(AI) and machine learning (ML) have enabled new possibilities in the field, particularly through the analysis of electroencephalogram (EEG) data. EEG, a non- invasive technique for monitoring brain activity, has shown promise in identifying neurological markers linked to mental health conditions.

The proposed approach investigates the application of advanced ML techniques for automated classification of psychological disorders using EEG data. Three models were developed and compared.

- Random Forest (RF) - An ensemble learning method known for its capability to handle high-dimensional data and identify complex relationships.
- Neural Network (NN) - A deep learning model designed to recognize intricate patterns within data.
- Hybrid Model (RF+NN) - A novel approach that integrates Random Forests and Neural Networks to combine their respective strengths.

This research aims to enhance mental health diagnostics by utilizing advanced machine learning techniques applied to EEG data. The primary objective is to improve early detection and identification of psychological disorders, enabling timely interventions that lead to better patient outcomes. By leveraging machine learning, this approach seeks to minimize reliance on subjective clinical assessments, ensuring a more objective and data-driven approach to diagnosing mental health conditions.

The research focuses on analyzing EEG features across various frequency bands and brain regions to determine their significance in psychological disorder prediction. It compares different machine learning models, including traditional Random Forests, deep learning-based Neural Networks, and a novel hybrid model, to identify the most effective classification approach. This comparative analysis helps uncover the strengths and limitations of each method, contributing to the refinement of diagnostic models.

This project aligns with Sustainable Development Goal 3 (SDG 3) - Good Health and Well-being, which emphasizes improved healthcare access, early detection, and mental health support. By integrating machine learning techniques, the project offers a technology-driven, non-invasive approach for identifying psychological disorders through EEG signal analysis, contributing to the expansion of mental health services.

The use of AI-based methodologies enhances accessibility, objectivity, and efficiency in diagnosing mental health disorders. This scalable model aids healthcare professionals in early assessments, reduces stigma associated with mental health evaluations, and enables timely interventions, thereby improving patient outcomes. The combination of Random Forests and Neural Networks as a Hybrid Model ensures greater predictive accuracy and reliability, particularly in complex cases.

Furthermore, this approach is adaptable for deployment in various healthcare settings, including underserved communities with limited mental health resources. By prioritizing early detection and promoting mental well-being through technology, the study contributes to a sustainable healthcare system, ensuring efficient resource allocation and supporting the broader vision of SDG 3 for global health improvement.

The dataset consists of EEG recordings from various individuals, with 1,148 extracted features representing brain activity, demographic details, and coherence measures. The objective of this research is to enhance the accuracy and efficiency of mental health disorder classification through machine learning.

II. LITERATURE REVIEW

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

A. Studies of Different Disorders

Various mental health issues have been the focus of recent studies. The World Health Organization [1] reported that mental illness and substance use affect many people around the world. Kessler et al. [2] found that mental health problems often begin during the teenage years. Kitsios et al. [3] mentioned that AI is being used more often in health care to help detect and manage conditions like depression, anxiety, and bipolar disorder. These studies show that technology is playing a growing role in understanding and supporting mental health.

B. Application of AI and EEG in Mental Health

Recent research highlights how artificial intelligence (AI) is being used to support mental health care. Chen et al. [4] showed that AI tools can help identify mental health conditions and assist in treatment planning. Jin et al. [5] explained how AI is useful in analyzing mental health data and could support doctors in making better decisions. Yasin et al.

[6] reviewed studies that use brain signals (EEG) with AI models like neural networks to detect mental disorders such as depression and bipolar disorder. These findings show that combining EEG with AI has strong potential for improving early diagnosis.

C. Diverse Datasets

Researchers utilize a variety of datasets, both publicly available and custom. For instance, Acharya et al., employed the Bonn University EEG dataset for epilepsy detection, while Mumtaz et al., 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].

D. 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. Ozdemir et al., focused on ADHD diagnosis, Zhuang et al., investigated bipolar disorder, and Khosla et al., explored autism spectrum disorder detection using EEG signals [9-11].

E. Advanced Machine Learning Models

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

F. Innovative Feature Extraction

While traditional frequency band analysis remains relevant, researchers are increasingly exploring more complex features. Wavelet transform features were employed for depression detection [15]. Connectivity measures were utilized for anxiety disorder classification [16] and raw EEG data were directly used in their deep learning model [17].

G. 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., reported F1 scores for their ADHD classification model [18], while Jahmunah et al., included sensitivity and specificity in their epilepsy detection study, indicating a move towards more comprehensive performance assessment [19].

III. PROPOSED METHODOLOGY

The project workflow is visually represented using intuitive icons and illustrations to enhance the understanding of the technical processes involved. The Figure 1. highlights the 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 modeling techniques employed.

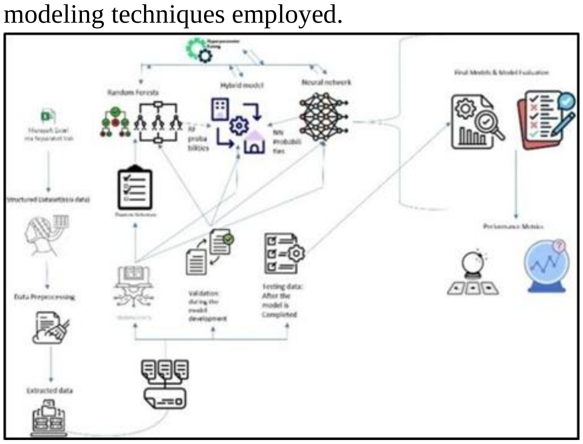

Fig. 1. Proposed Project Architecture

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.

The complete research methodology pipeline is systematically represented through a comprehensive workflow as shown in Figure 2. It begins with Data Collection, utilizing EEG data from various 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.

Fig. 2. Comprehensive Flowchart of Project Methodology

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 at wo- 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.

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

A. 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 relationship, while NN excels at identifying complex patterns and learning hierarchical representations

B. 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.

C. 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.

D. 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.

E. 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.

F. 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 generalize well to unseen data.

G. 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.

IV. RESULTS

The outcomes of the existing models in terms of performance evaluation, comparisons and testing results are discussed in this section.

A. Random Forest

This model achieved an accuracy of 40.74%. While it showed strong performance for Mood Disorder (F1-score: 0.56) and Addictive Disorder (F1-score: 0.53), but it failed to classify Anxiety, Schizophrenia, and Trauma and Stress- Related Disorders (F1-score: 0.00). The overall macro- average F1-score was 0.22, indicating significant class imbalance effects.

TABLE I. RESULTS OF RANDOM FOREST MODEL- BINARY CLASSIFICATION

	Precision	Recall	F1-score
Healthy (0)	0.90	1.00	0.95
Unhealthy (1)	1.00	0.05	0.10

Binary Classification (Healthy vs. Unhealthy) - Accuracy reached 90.48%, with an F1-score of 0.95 for Unhealthy (Class 0) which is listed in Table I. However, the model struggled with Healthy (Class 1) cases (Recall: 0.05, F1-score: 0.10), revealing a severe imbalance in predictions. Model Insights are given below.

- The ROC-AUC score was 0.90, showing excellent binary discrimination.
- The learning curve indicated overfitting, with training accuracy near 1.0 and validation accuracy plateauing at 0.35.
- Feature importance analysis highlighted Patient ID, Sex, IQ, and EEG coherence metrics as top contributors.

B. Neural Network

It achieved an overall accuracy of 31.22%, with the best performance on Healthy Control (F1-score: 0.49) and balanced but low metrics across other classes. This model handled class imbalance better, maintaining a more distributed confusion matrix and consistent performance.

Binary Classification - The model reached 91.53% accuracy, outperforming Random Forest slightly. It yielded strong results for Unhealthy cases (F1-score: 0.95), and improved performance for Healthy (F1-score: 0.53) that is shown in Table II. Model Insights are as follows

- The ROC-AUC was 0.81, indicating good binary classification capabilities.
- The neural network showed lower class-specific bias and better generalization, though with a lower multi- class accuracy than Random Forest.

TABLE II. RESULTS OF NEURAL NETWORK - BINARY CLASSIFICATION

	Precision	Recall	F1-score
Healthy (0)	0.94	0.96	0.95
Unhealthy (1)	0.60	0.47	0.53

C. Hybrid Model

The hybrid model achieved 42.33% accuracy, the highest among all models. Notable performance was observed for Addictive Disorders (F1-score: 0.60) and Mood Disorders (F1-score: 0.53). However, it failed to predict Anxiety, OCD, Schizophrenia, and Trauma/Stress Disorders (F1-score: 0.00).

Binary Classification is shown in Table III., With 91.01% accuracy, the hybrid model balanced precision and recall better than individual models. It showed improved recall and F1-score for the Healthy class (F1-score: 0.60), suggesting enhanced robustness. Model Insights are

- Combines the strengths of Random Forest and Neural Networks.
- Better performance in binary classification, but still challenged by minority disorder classes in multi- class tasks.

TABLE III. RESULTS OF HYBRID MODEL - BINARY CLASSIFICATION

	Precision	Recall	F1-score
Healthy (0)	0.96	0.94	0.95
Unhealthy (1)	0.54	0.68	0.60

Performance comparison of the three models is illustrated in the Figure 3. From this, it is obvious that the proposed hybrid method outperforms the Random Forest and Neural network models.

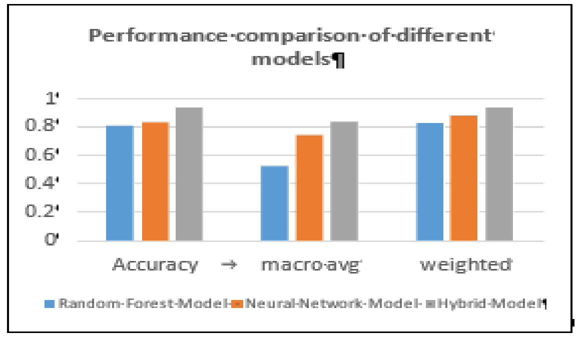

Fig. 3. Performance comparison of different models

V. CONCLUSION

This proposed technique 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 of individual subjects, with 1,148 EEG features per subject, split into training (80%) and testing (20%) sets.

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.

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.

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.

The development of non-invasive, EEG-based diagnostic tools could enhance accessibility to mental health assessments, particularly in resource-constrained settings.

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. [2005Jun]; 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, [1October2023], 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. ttps://doi.org/10.1109/ACCESS.2020.3006404

[16] Zheng, W., Liu, Y., & He, H. [2019]. Connectivity measures for anxiety disorder

classification from EEG signals. Neuro computing ,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] https://www.kaggle.com/datasets/shashwatwork/eeg-psychiatric- disorders-dataset