AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

Automated Multiclass EEG Signal Classification for Early Diagnosis of Psychiatric Disorders via Hybrid Graph Neural Networks

C. Bagath Basha[1*], Badanagari Shivapriya[2], Kaparthy Shiva Pranav[2], Pasham Shivani[2]
[1]Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering
[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.
*Correspondence: C. Bagath Basha (basha@kpritech.ac.in)

ABSTRACT

Neurological and psychiatric disorders are emerging as major global health concerns, making early and accurate diagnosis critical for effective treatment and improved patient outcomes. Electroencephalography (EEG) signals offer a non-invasive and highly informative method for studying brain activity. However, due to their high dimensionality, noise, and complex nonlinear patterns, EEG-based automated classification remains a challenging task. Many existing diagnostic models struggle with issues such as limited generalization across subjects, inability to effectively capture nonlinear relationships, and reduced classification accuracy, which impacts their clinical reliability. To overcome these limitations, this study proposes a multiclass machine learning framework for the simultaneous classification of epilepsy, migraine, and schizophrenia using EEG data. The framework begins with a structured preprocessing pipeline that includes data cleaning, missing value imputation, normalization, and label encoding to ensure data quality and consistency. The dataset is then divided using stratified train–test splitting to maintain class balance and prevent bias during model training. Several machine learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), and Decision Tree (DT), are implemented and evaluated. In addition, a Hybrid Graph Neural Network (HGNN) approach is introduced, where a Tree-based Generalized Additive Model (TreeGAM) is utilized to capture complex nonlinear interactions among EEG features more effectively.

Keywords: Machine Learning, Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Hybrid Graph Neural Network (HGNN), Tree-based Generalized Additive Model (TreeGAM).

1. INTRODUCTION

Neurological and psychiatric conditions such as schizophrenia, epilepsy, and migraine pose a major burden on global healthcare systems due to their high prevalence and long-term effects on patients' quality of life. Epidemiological studies suggest that schizophrenia impacts about 1% of the population, migraine disorders affect hundreds of millions of people worldwide, and epilepsy has a significant lifetime incidence rate. These statistics highlight the pressing need for diagnostic approaches that are accurate, scalable, and widely accessible. While advanced neuroimaging modalities such as magnetic resonance imaging (MRI), functional MRI (fMRI), and positron emission tomography (PET) are commonly used in clinical settings, they are often expensive, less accessible in low-resource regions, and limited in capturing rapid temporal changes in brain activity. In comparison, electroencephalography (EEG) offers a more practical alternative due to its non-invasive nature, low cost, portability, and high temporal resolution. These characteristics make EEG particularly effective for monitoring real-time brain dynamics associated with neurological and psychiatric disorders. Furthermore, EEG-derived features play a crucial role in enabling machine learning-based approaches for automated detection and classification of such conditions. The significance of these features in supporting data-driven diagnostic models is illustrated in Figure 1.

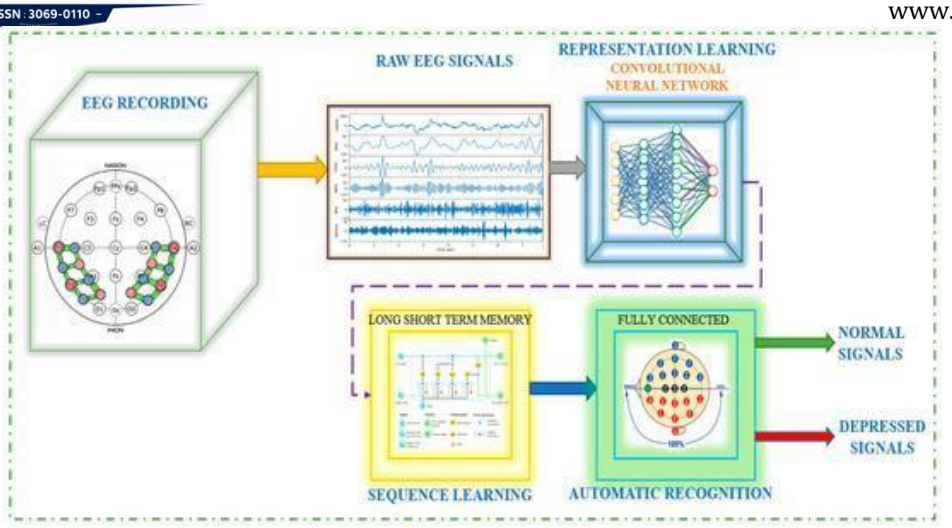


Figure 1: EEG features based detection using ML

The growing utilization of EEG in both clinical and research domains is largely driven by these advantages, especially its capability to capture rapid neural fluctuations. EEG signals are recorded through electrodes placed on the scalp and reflect brain activity across multiple frequency bands, including delta, theta, alpha, beta, and gamma, each associated with specific cognitive and physiological states. This rich temporal information makes EEG highly compatible with artificial intelligence (AI)-based analytical approaches. Recent advancements in computational neuroscience have significantly enhanced the role of EEG in diagnostic applications, where machine learning and deep learning techniques are employed to identify complex patterns within brain signals. These approaches have demonstrated promising results in detecting and classifying psychiatric conditions such as depression and schizophrenia. Furthermore, the integration of EEG data with additional clinical or demographic features, particularly through hybrid deep learning architectures like convolutional neural networks, has further improved diagnostic accuracy and robustness.

2. LITERATURE SURVEY

Li et al. [1] introduced a comprehensive framework that integrates empirical mode decomposition (EMD) with support vector machines (SVMs), achieving high performance with sensitivity of 98.00% and specificity of 99.40% through detailed analysis of intrinsic mode functions (IMFs) and their statistical characteristics. Expanding on advanced architectures, Satapathy et al. [2] demonstrated the capability of transformer-based models for EEG classification, reporting an accuracy of 99.5% using wavelet-transformed signals. Similarly, Manali et al. [3] proposed a structured approach for mental task classification by applying variational mode decomposition (VMD) on single-channel EEG signals. Their methodology involved signal decomposition, computation of variational mode energy ratios, and classification using adaptive boosting techniques.

Feature optimization remains a critical aspect in EEG-based learning systems, as explored by Conrado et al. [4], who applied artificial neural networks (ANNs) for mental task classification with emphasis on feature reduction. Convolutional neural networks (CNNs) have also been widely adopted, with Pallavi et al. [5] utilizing scalogram representations of EEG signals for emotion classification, achieving subject-independent performance across multiple datasets. Jinru et al. [6] further advanced this domain by employing Bidirectional Long Short-Term Memory (BiLSTM) networks for emotion recognition from EEG data. Additionally, EEGNet, a specialized CNN-based architecture for brain–computer interfaces (BCIs), introduced depth-wise and separable convolutions and demonstrated effectiveness across multiple paradigms including P300, ERN, MRCP, and SMR.

Madhuri et al. [7] developed a hierarchical classification framework using optimized neural networks for tasks such as hand movement and word generation. Earlier work by Xiu et al. [8] applied deep learning techniques to motor imagery tasks, specifically focusing on left and right hand movement

imagination. Saadat et al. [9] combined Back Propagation Neural Networks (BPNN) with Hidden Markov Models (HMM) for mental task classification, while Jose et al. [10] explored online learning mechanisms in BCIs to stabilize learning rates in patients with neurological conditions. Debarshi et al. [11] compared subject-dependent and subject-independent EEG models, concluding that conventional machine learning approaches are more effective for subject-independent scenarios. Chen et al. [12] evaluated multiple ML classifiers and identified least squares SVM (LS-SVM) as the most effective, achieving 99.5% accuracy for three-class classification tasks.

In migraine-related studies, Yin et al. [13] developed a decision-support system to distinguish migraine from tension-type headaches using K-nearest neighbors (KNN), achieving 90% accuracy. Akben et al. [14] analyzed EEG responses under photic stimulation at varying frequencies and applied ANN classification to reach 85% accuracy. Subasi et al. [15] enhanced this approach using discrete wavelet transform (DWT) for feature extraction combined with Random Forest (RF), achieving 85.95% accuracy. Frid et al. [16] employed radial basis function SVM to differentiate migraine subtypes, achieving 84.62% accuracy.

3. PROPOSED METHODOLOGY

The proposed system presents an end-to-end EEG classification framework that combines data preprocessing, multiple machine learning models, and secure web-based deployment within a unified architecture. The process begins with structured ingestion of EEG data from CSV files, followed by preprocessing steps such as removal of irrelevant attributes and mean-based imputation to handle missing values and ensure data consistency. After preprocessing, the dataset is split into features and target variables, and label encoding is applied to make the data compatible with supervised learning algorithms. The framework integrates several classification models, including Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), and Decision Tree (DT), along with a Hybrid Graph Neural Network (HGNN) component to enhance predictive performance and robustness, as shown in Figure 2. A centralized training engine is responsible for model training, evaluation, and storage, enabling efficient reuse and comparison of trained models.

1. Data Input and Preprocessing

The system begins by loading EEG data from a structured CSV file, where irrelevant or unnamed columns are removed to maintain consistency. Missing numerical values are handled using mean imputation to avoid data loss and bias. This step ensures that the dataset is clean, structured, and ready for effective analysis by machine learning models.

2. Feature–Target Separation and Encoding

In this stage, the dataset is divided into input features and target variables representing disorder classifications. The categorical target labels are transformed into numerical values using label encoding, enabling compatibility with supervised learning algorithms. This transformation allows models to effectively learn patterns from structured data. Model performance is evaluated using standard metrics such as accuracy, precision, recall, F1-score, and confusion matrix, allowing comprehensive assessment across different algorithms.

3. Model Training Engine

The processed data is used to train multiple classifiers, including RF, SVM, LR, and along with a hybrid HGNN model. Each model is trained using predefined configurations and stored for future reuse, ensuring efficiency and consistency across multiple executions.

AMERICAN JOURNAL OF MANAGEMENT AND IOT MEDICAL COMPUTING
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3069-0110
Vol.5, No.2(1) (2026)
www.ajmimc.com

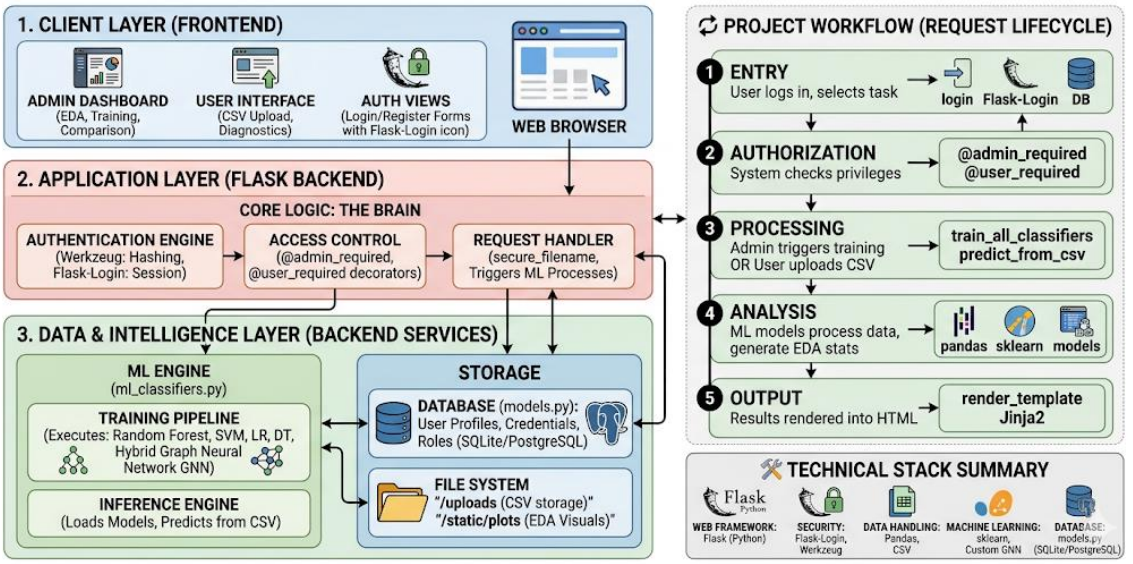


Figure 2: Proposed system architecture

4. Performance Evaluation and Comparison

After training, each model is evaluated using standard performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix. These metrics provide a detailed understanding of model performance and enable comparison across different classifiers to identify the most effective model.

5. Role-Based System Deployment

The system is deployed using a Flask-based web application integrated with secure authentication mechanisms. It supports role-based access control with separate admin and user logins, where administrators manage training and evaluation, while users access prediction functionalities in a controlled environment.

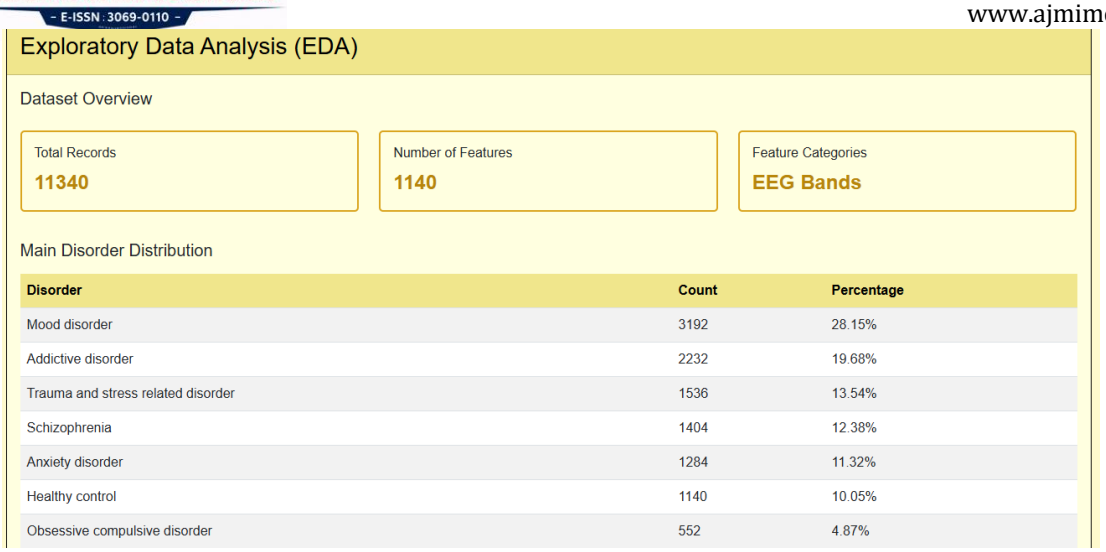
6. Prediction and Result Generation

Users upload new EEG data in CSV format through the interface, which is then preprocessed and passed to the trained models. The system generates predictions and converts them into human-readable labels, providing meaningful diagnostic insights for practical use.

4. RESULTS AND DISCUSSIONS

The implementation of the EEG-based psychiatric disorder identification system begins with setting up the Python environment and required libraries for data processing, machine learning, and web deployment. EEG datasets uploaded by the administrator are loaded, cleaned, and preprocessed by removing irrelevant columns, handling missing values, normalizing features, and encoding disorder labels. Relevant EEG features are selected and the dataset is split into training and testing subsets using stratified sampling. Multiple machine learning classifiers, including RF, SVM, LR, , and HGNN are trained and evaluated using standard performance metrics. The trained models are stored for reuse, and the Flask-based web interface enables real-time prediction by allowing users to upload new EEG data and view accurate, model-driven psychiatric disorder predictions in a user-friendly manner.

Figure 3 depicts the exploratory data analysis of main disorder categories, presenting the distribution of EEG-based classes at a broader level. The dataset contains 11,340 total records, where mood disorder accounts for 3,192 cases (28.15%), addictive disorder includes 2,232 cases (19.68%), trauma and stress-related disorder comprises 1,536 cases (13.54%), schizophrenia represents 1,404 cases (12.38%), and anxiety disorder includes 1,284 cases (11.32%). Additionally, healthy control has 1,140 cases (10.05%), while obsessive compulsive disorder contributes 552 cases (4.87%). This distribution highlights class variability and supports understanding of dataset composition for effective model training.



Exploratory Data Analysis (EDA)

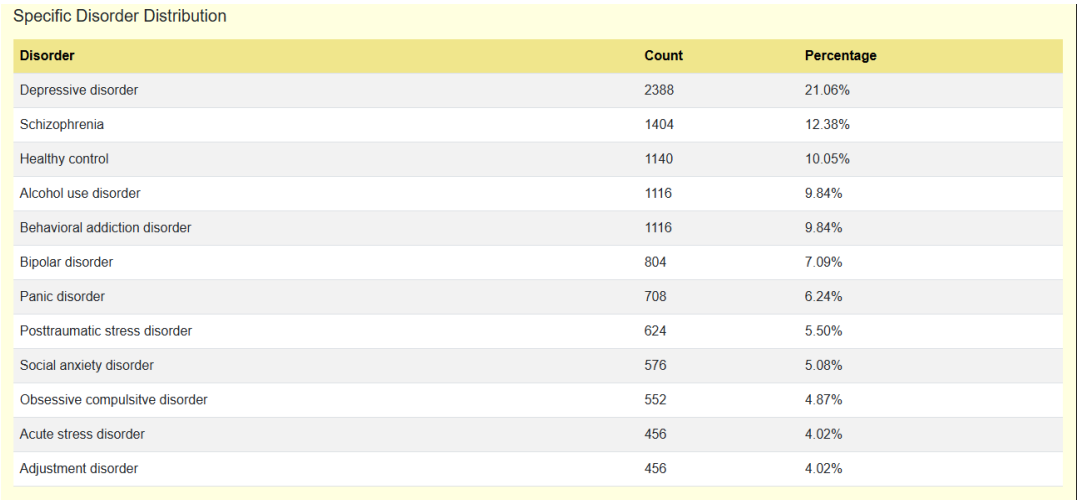
Dataset Overview

Total Records	Number of Features	Feature Categories
11340	1140	EEG Bands

Main Disorder Distribution

Disorder	Count	Percentage
Mood disorder	3192	28.15%
Addictive disorder	2232	19.68%
Trauma and stress related disorder	1536	13.54%
Schizophrenia	1404	12.38%
Anxiety disorder	1284	11.32%
Healthy control	1140	10.05%
Obsessive compulsive disorder	552	4.87%

Figure 3: EDA- Data Distribution of classes main disorder



Specific Disorder Distribution

Disorder	Count	Percentage
Depressive disorder	2388	21.06%
Schizophrenia	1404	12.38%
Healthy control	1140	10.05%
Alcohol use disorder	1116	9.84%
Behavioral addiction disorder	1116	9.84%
Bipolar disorder	804	7.09%
Panic disorder	708	6.24%
Posttraumatic stress disorder	624	5.50%
Social anxiety disorder	576	5.08%
Obsessive compulsitve disorder	552	4.87%
Acute stress disorder	456	4.02%
Adjustment disorder	456	4.02%

Figure 4: EDA- Data Distribution of classes specific disorder

Figure 4 depicts the exploratory data analysis of specific disorder categories, providing a detailed breakdown of class distribution. Depressive disorder accounts for 2,388 cases (21.06%), followed by schizophrenia with 1,404 cases (12.38%) and healthy control with 1,140 cases (10.05%). Alcohol use disorder and behavioral addiction disorder each contribute 1,116 cases (9.84%), while bipolar disorder includes 804 cases (7.09%) and panic disorder has 708 cases (6.24%). Posttraumatic stress disorder consists of 624 cases (5.50%), social anxiety disorder has 576 cases (5.08%), obsessive compulsive disorder includes 552 cases (4.87%), and both acute stress disorder and adjustment disorder have 456 cases each (4.02%). This detailed distribution enables deeper insight into class-level variations for precise classification modeling.

Figure 5 depicts the confusion matrix for the HGNN model, where nearly perfect classification performance is achieved across all classes. Addictive disorder (447), anxiety disorder (257), healthy control (228), mood disorder (638), obsessive compulsive disorder (110), schizophrenia (281), and trauma and stress-related disorder (307) are all correctly classified with minimal to no misclassification. The strong diagonal dominance across all classes indicates the model's superior ability to capture complex nonlinear relationships and effectively distinguish between different psychiatric disorders.

Confusion Matrix (CM)

	Addictive disorder	Anxiety disorder	Healthy control	Mood disorder	Obsessive compulsive disorder	Schizophrenia	Trauma and stress related disorder
Addictive disorder	447	0	0	0	0	0	0
Anxiety disorder	0	257	0	0	0	0	0
Healthy control	0	0	228	0	0	0	0
Mood disorder	0	0	0	638	0	0	0
Obsessive compulsive disorder	0	0	0	0	110	0	0
Schizophrenia	0	0	0	0	0	281	0
Trauma and stress related disorder	0	0	0	0	0	0	307

Figure 5: Confusion matrix obtained for main disorder column HGNN

Confusion Matrix (CM)

	Acute stress disorder	Adjustment disorder	Alcohol use disorder	Behavioral addiction disorder	Bipolar disorder	Depressive disorder	Healthy control	Obsessive compulsitve disorder	Panic disorder	Posttraumatic stress disorder	Schizophrenia	Social anxiety disorder
Acute stress disorder	91	0	0	0	0	0	0	0	0	0	0	0
Adjustment disorder	0	91	0	0	0	0	0	0	0	0	0	0
Alcohol use disorder	0	0	223	0	0	0	0	0	0	0	0	0
Behavioral addiction disorder	0	0	0	223	0	0	0	0	0	0	0	0
Bipolar disorder	0	0	0	0	161	0	0	0	0	0	0	0
Depressive disorder	0	0	0	0	0	478	0	0	0	0	0	0
Healthy control	0	0	0	0	0	0	228	0	0	0	0	0
Obsessive compulsitve disorder	0	0	0	0	0	0	0	110	0	0	0	0
Panic disorder	0	0	0	0	0	0	0	0	142	0	0	0
Posttraumatic stress disorder	0	0	0	0	0	0	0	0	0	125	0	0
Schizophrenia	0	0	0	0	0	0	0	0	0	0	281	0
Social anxiety disorder	0	0	0	0	0	0	0	0	0	0	0	115

Figure 6: Confusion matrix obtained for specific disorder column HGNN

Figure 6 depicts the confusion matrix for the HGNN-inspired model, where near-perfect classification is achieved across all specific disorder classes. Acute stress disorder (91), adjustment disorder (91), alcohol use disorder (223), behavioral addiction disorder (223), bipolar disorder (161), depressive disorder (478), healthy control (228), obsessive compulsive disorder (110), panic disorder (142), posttraumatic stress disorder (125), schizophrenia (281), and social anxiety disorder (115) are all correctly classified. The strong diagonal dominance with negligible off-diagonal values highlights the model's superior ability to distinguish between fine-grained disorder categories.

The figure 7 shows main disorder prediction results generated using the proposed HGNN display the model's ability to accurately classify EEG samples into their respective psychiatric disorder categories. Each EEG sample is assigned a predicted main disorder label such as addictive disorder, anxiety disorder, healthy control, mood disorder, or obsessive–compulsive disorder, demonstrating clear differentiation among disorder groups. The consistent and well-structured predictions across multiple samples indicate the effectiveness of the HGNN model in capturing complex EEG feature relationships. This output confirms the suitability of the proposed hybrid approach for reliable and automated main-level psychiatric disorder identification.

Sample #	Predicted main.disorder
1	Addictive disorder
2	Addictive disorder
3	Addictive disorder
4	Anxiety disorder
5	Anxiety disorder
6	Anxiety disorder
7	Anxiety disorder
8	Healthy control
9	Healthy control
10	Healthy control
11	Healthy control
12	Healthy control
13	Mood disorder
14	Mood disorder
15	Mood disorder
16	Obsessive compulsive disorder

Figure 7: Prediction obtained using HGNN on main disorder

The figure 8 shows specific disorder prediction results obtained using the proposed HGNN demonstrate the model's capability to perform fine-grained psychiatric disorder classification from EEG features. Each EEG sample is classified into a precise disorder category such as depressive disorder, alcohol use disorder, social anxiety disorder, bipolar disorder, or healthy control. The consistent and coherent predictions across multiple samples indicate that the HGNN effectively captures subtle EEG signal variations associated with different psychiatric conditions. This output validates the effectiveness of the proposed model for detailed, specific-level psychiatric disorder identification.

Sample #	Predicted specific.disorder
1	Depressive disorder
2	Alcohol use disorder
3	Alcohol use disorder
4	Depressive disorder
5	Alcohol use disorder
6	Depressive disorder
7	Social anxiety disorder
8	Depressive disorder
9	Depressive disorder
10	Depressive disorder
11	Healthy control
12	Depressive disorder
13	Depressive disorder
14	Bipolar disorder
15	Depressive disorder
16	Depressive disorder
17	Depressive disorder
18	Depressive disorder

Figure 8: Prediction obtained using HGNN on specific disorder

5. CONCLUSION

The proposed system provides a structured and efficient framework for EEG-based mental health classification by combining effective preprocessing techniques with multiple machine learning models, including Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), and a Hybrid Graph Neural Network (HGNN). The framework ensures reliable performance through thorough evaluation using standard metrics such as accuracy, precision, recall, and F1-score. Among all implemented methods, the hybrid model shows superior ability in learning and representing complex EEG signal patterns. The system also incorporates model persistence, exploratory data analysis, and a Flask-based role-driven web interface to improve usability, accessibility, and overall system control. This integration allows efficient management of data processing and trained models within a unified environment. In addition, the real-time prediction module supports practical deployment by enabling the classification of newly uploaded EEG data. Overall, the system offers a scalable, user-friendly, and effective solution for real-world EEG-based mental health diagnosis and decision support.

REFERENCES

[1]. Li, S.; Zhou, W.; Yuan, Q.; Geng, S.; Cai, D. Feature extraction and recognition of ictal EEG using EMD and SVM. Comput. Biol.Med. 2013,43, 807–816.

[2]. Satapathy, S.K.; Jagadev, A.K.; Dehuri, S. Wei ghted majority voting based ensemble of classifiers using different machine learning techniques for classification of EEG signal to detect epileptic seizure. Informatica 2017,41.

[3]. Manali, M.; Satija, U.; Upadhayay, M.D. Variational Mode Decomposition Based Mental Task Classification from Electroencephalogram. In Proceedings of the 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 10–13 December 2020; pp. 1–7.

[4]. Conrado, C.F.; Sison, L.G. Mental Task Classification Using Artificial Neural Network with Feature Reduction. In Proceedings of the 2020 6th International Conference on Control, Automation and Robotics (ICCAR), Virtual, 20–23 April 2020; pp. 753–757.

[5]. Pandey, P.; Seeja, K.R. Subject independent emotion recognition system for people with facial deformity: An EEG based approach. J. Ambient Intell. Humaniz. Comput. 2021, 12, 2311–2320.

[6]. Jinru, J.; Huang, X.; Wu, H.; Yang, X. EEG-based emotion classification based on Bidirectional Long Short-Term Memory Network. Procedia Comput. Sci. 2020, 174, 491–504.

[7]. Madhuri, M.N.; Bhurchandi, K.M. Classification of Mental Tasks using EEG and Hierarchical Classifier employing Optimised Neural Networks. Int. J. Comput. Appl. 2021, 975, 8887.

[8]. Xiu.; Kuang, D.; Guo, X.; Zhao, Y.; He, L. A deep learning method for classification of EEG data based on motor imagery. In Proceedings of the International Conference on Intelligent Computing, Taiyuan, China, 3–6 August 2022; pp. 203–210.

[9]. Saadat.; Pourghassem, H. Mental task classification based on HMM and BPNN. In Proceedings of the 2013 International Conference on Communication Systems and Network Technologies, Gwalior, India, 6–8 April 2023; pp. 210–214.

[10]. Jose.R. On the need for on-line learning in brain-computer interfaces. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), Budapest, Hungary, 25–29 July 2024; Volume 4, pp. 2877–2882.

[11]. Debarshi.; Singh, M.; Sethia, D.; Kalra, D.; Indu, S. A comparative study of subject-dependent and subject-independent strategies for EEG-based emotion recognition using LSTM network. In Proceedings of the 2020 the 4th International Conference on Compute and Data Analysis, Silicon Valley, CA, USA, 9–12 March 2020; pp. 142–147.

[12]. Chen, S.; Zhang, X.; Chen, L.; Yang, Z. Automatic diagnosis of epileptic seizure in electroencephalography signals using nonlinear dynamics features. IEEE Access 2019, 7, 61046–61056.

[13]. Yin, Z.; Dong, Z.; Lu, X.; Yu, S.; Chen, X.; Duan, H. A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case-based reasoning. J. Headache Pain , 16, 1–9.

[14]. Akben, S.B.; Tuncel, D.; Alkan, A. Classification of multichannel EEG signals for migraine detection. Biomed. Res, 27, 743–748.

[15]. Subasi, A.; Ahmed, A.; Alicković, E.; Hassan, A.R. Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform. Biomed. Signal Process. Control , 49, 231–239.

[16]. Frid, A.; Shor, M.; Shifrin, A.; Yarnitsky, D.; Granovsky, Y. A biomarker for discriminating between migraine with and without aura: Machine learning on functional connectivity on resting-state EEGs. Ann. Biomed. Eng. **2020**, 48, 403–412.