Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 26 June 2026
Revised 04 Aug 2026
Accepted 18 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (16s): 1519–1535

KEYWORDS:
Neurological disorders; Physiological sensor data; Machine learning framework; Multimodal signal analysis; Ensemble classifiers; Early disease detection; XGBoost; Clinical decision support systems.

An Intelligent Machine Learning Framework for Multi-Class Detection of Neurological Disorders Using Physiological Signals

Harsimran Bains[1], Khushboo Bansal[2], Yogesh Kumar[3]

[1]Department of Computer Science and Application, Faculty of Engieering Technology and Computing , Desh Bhagat University, Mandi Gobindgarh, Punjab, India, Email- bainsharsimran91@gmail.com
[2]Department of Computer Science and Engieering, Faculty of Engieering Technology and Computing , Desh Bhagat University, Mandi Gobindgarh, Punjab, India, Email- deputy.directorengg@deshbhagatuniversity.in
[3]Department of CSE, School of Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat, India, Email- yogesharora10744@gmail.com

ABSTRACT: Neurological and neurodevelopmental disorders often exhibit overlapping clinical manifestations, making early and accurate diagnosis a significant challenge in healthcare practice. Recent advances in physiological sensing technologies have enabled the continuous acquisition of health-related signals that can support intelligent disease screening systems. This study presents an intelligent machine learning framework for the multi-class detection of neurological disorders using multimodal physiological signals. The proposed framework utilizes heart rate, oxygen saturation (SpO_2), temperature, movement characteristics, infrared sensor measurements, and tri-axial accelerometer data to classify six neurological conditions, including Alzheimer face, Angelman syndrome, MECP2 syndrome, Pitt-Hopkins syndrome, Rett syndrome, and normal cases. To enhance predictive performance, a comprehensive preprocessing pipeline incorporating feature engineering, data normalization, Synthetic Minority Oversampling Technique (SMOTE), mutual information-based feature selection, and Principal Component Analysis (PCA) is employed. Multiple machine learning algorithms, including Logistic Regression, K-Nearest Neighbors, Support Vector Machines, Random Forest, Gradient Boosting, Extra Trees, XGBoost, LightGBM, CatBoost, Voting, and Stacking classifiers, are systematically evaluated. Experimental results demonstrate the superiority of ensemble learning approaches, with LightGBM achieving an accuracy of 99.60%, precision of 99.57%, recall of 99.63%, F1-score of 99.59%, Matthews Correlation Coefficient of 99.52%, and ROC-AUC of 99.99%. These findings highlight the effectiveness of multimodal physiological signal analysis and advanced machine learning techniques for reliable neurological disorder classification, providing a promising decision-support tool for early diagnosis and intelligent healthcare applications.

INTRODUCTION

Neurological diseases are among the major causes of disability and death worldwide and impose a significant social, economic, and healthcare burden. Alzheimer's disease, Rett syndrome, Angelman syndrome, Pitt–Hopkins syndrome, and MECP2 syndrome are examples of disorders, which have complex neurological abnormalities that impact cognitive, behavioral and physiological functions (Sandweiss et al. 2020). These conditions can only be timely managed, treated and optimized if diagnosed early. However, the diagnosis of these disorders remains difficult as the clinical signs of many neurological disorders are similar, especially in the early stages of their course. Typical diagnostic approaches involve clinical observations, neurological assessments, neuro-imaging studies, and genetic testing, all of which may be cumbersome, expensive and impractical in resource-limited healthcare environments (Sadu et al. 2025). With the recent progress in biomedical sensing technology and

wearable healthcare devices, a new opportunity arises for real-time, non-invasive disease assessment by continuous monitoring of physiological signals. Physiological signals including heart rate, oxygen saturation (SpO_2), body temperature, movement patterns, infrared sensor measurements and accelerometer signals serve as indicators of the body's normal activities, which are linked with activities happening at the nervous level and autonomic nervous system Xian (2026).

Physiological signals collected from wearable devices and biomedical sensors contain valuable information about an individual's health status. These signals can detect changes in the body at a level that may not be visible in the clinical picture and may be linked to neurological disease (Kim et al. 2025). This has resulted in an increasing interest in the integration of physiological sensing and intelligent computational techniques that can aid in early disease detection and enable better diagnostics in healthcare. Machine learning has proven to be a valuable tool in this respect due to its ability to detect complex patterns and relationships in large volumes of biomedical data which might not be easily discernible by traditional clinical analysis (Guan 2019; Chen et al. 2024).

In recent years, machine learning has been effectively used in the diagnosis of multiple neurological diseases. These models can be used to make accurate and automatic predictions of physiological measurements and disease conditions, aiding clinical decisions. However, most of the studies are limited to one type of neurological disorder and most often use binary classification methods (Li et al. 2025). These methods may not be applicable to actual clinical settings where physicians and other health professionals must differentiate between various neurological disorders in the presence of several. In addition, physiological data sets are frequently very high-dimensional, feature-redundant (correlated), class-imbalanced, and noisy, which may have adverse impact on model accuracy and generalization performance (Busnatu et al. 2022).

To cope with these challenges, this study presents a machine learning system for multi-class classification of neurological disorders from physiological sensor data. The proposed framework includes physiological data including heart rate, SpO_2, movements, IR sensors, and accelerometers to represent neurological health conditions. The data is preprocessed using a structured data processing pipeline that enhances data quality and predictive performance through feature engineering, normalization, SMOTE class balancing, feature selection, and PCA dimensionality reduction. Clinically relevant metrics are used to evaluate multiple machine learning models. The results demonstrate the efficiency of the proposed system as a reliable decision support system in early detection and diagnosis of neurological diseases.

1.1 Contribution

The major contributions of this research are outlined as follows:

- A machine learning framework is proposed for the multi-class classification of neurological disorders using multimodal physiological sensor data, including heart rate, SpO_2, movement, infrared, temperature, and accelerometer measurements.
- A comprehensive data processing pipeline incorporating class balancing, feature selection, and dimensionality reduction is developed to enhance feature quality and improve classification performance.
- Multiple machine learning, ensemble learning, and gradient boosting algorithms are systematically evaluated to identify the most effective model for neurological disorder detection.
- The presented framework demonstrates the effectiveness of physiological sensor fusion for accurately distinguishing among six neurological conditions using a unified diagnostic model.
- Extensive performance evaluation using clinically relevant metrics highlights the potential of the proposed system as an intelligent decision-support tool for early neurological disorder screening and diagnosis.

1. RELATED WORK

Recent advances in machine learning and deep learning have shown impressive performance in capturing patterns of diseases from physiological signals, neuro-imaging and electro-physiological measurements to facilitate automated and non-invasive diagnosis.

In a review of deep learning methods for diagnosis of neurological disorders via magnetic resonance imaging (MRI), Noor et al. (2020) showed that convolutional neural networks (CNN) have been shown to be effective for Alzheimer's disease, Parkinson's disease and schizophrenia diagnosis. Likewise, Lima et al. (2022) conducted a detailed review on the detection and

classification of neurological disorders using machine learning techniques, highlighting important challenges such as data heterogeneity, limited data availability and model interpretability. Recently, Mohsen (2025) reviewed the machine learning and deep learning methods for the diagnosis of Alzheimer's disease and highlighted the increasing contribution of AI in the diagnostic system. Moreover, Rosca & Stancu (2025) reviewed recent developments in Neuro-ML technologies and Mehmood et al. (2025) presented studies on the application of AI to analyze EEG signals for neurological and oculomotor conditions, which also demonstrated the potential of intelligent decision-support systems for neurological healthcare. One of the important research trends is neuroimaging and EEG-based approaches to the diagnosis of neurological disorders. A multi-class classification method for EEG analysis of neurodegenerative diseases was proposed by Falaschetti et al. (2024) that used lightweight Long Short-Term Memory (LSTM) network to capture the temporal neural patterns while keeping computational efficiency. In a similar manner, Goyal et al. (2025) proposed ensemble deep learning-based approach for classification of neurodegenerative disease using classifiers and found that the predictive performance has improved by the fusion of classifiers. While these studies show impressive results on the detection of neurological diseases using advanced deep learning architectures, the requirement of these modalities (EEG or neuroimaging) for specialized acquisition systems and physician analysis may hinder their wider clinical application. More recently, there has also been increased attention to using wearable sensing technologies and physiology monitoring solutions to assess neurological domains. Hope et al. (2021) demonstrated the feasibility of combining a wearable IMU-based physiological vibration acceleration sensor with machine learning models to identify, quantify, and monitor physiological impairments associated with concussion injuries. To identify various motor abnormalities associated with PD, Xiang et al. (2025) presented a multidimensional machine learning framework, demonstrating that motion characteristics derived from wearable movement sensors can be used to detect PD. Similarly, Khalil et al. (2025) used wearable sensors and machine learning methods to differentiate Parkinson's disease from other types of parkinsonism, demonstrating the potential benefit of continuous physiological monitoring and sensor-based diagnostic systems. Apart from investigations specific to the identified disease, there is some research that has focused on multi-class classification approaches to neurological disease. Altıntop (2025) proposed a feature-driven machine learning based framework which aims to increase the diagnosis accuracy of the multiple neurological diseases by performing feature selection and classification optimizations. The approach of using an aggregated pattern classification method was employed to improve neural disorder stage detection by Anjum et al. (2024) and highlighted the advantages of ensemble-based learning techniques with complex neurological data.

Advantages of feature engineering and ensemble learning for the classification of neurological disorders have been shown in previous studies. Most studies, however, focus on a specific disease, such as Alzheimer's or Parkinson's and use expensive neuroimaging or EEG information. Furthermore, few studies have focused on incorporating multiple physiological signals and the problems of class imbalance, feature redundancy, high dimensionality and model generalization have not been adequately studied.

This study aims at addressing these limitations by proposing a machine learning framework to detect multiple-classes of neurological disorders based on multi-modal physiological sensor data collected from the human body, such as heart rate, SpO_2, temperature, infrared signals, movement, and accelerometer measurements. It adopts SMOTE-based class balancing, feature selection and dimensionality reduction in the framework to improve the classification performance. The proposed approach enables the benchmarking of various machine learning and ensemble models to produce an accurate, scalable and clinically relevant decision support system for neurological disorder screening and diagnosis.

2. MATERIALS AND METHODOLOGY

This study presents a comprehensive machine learning-based framework for the classification and analysis of physiological sensor data. The proposed methodology integrates multiple stages, as shown in Figure 1 to classify neurological disorders. Physiological parameters such as heart rate (BPM), blood oxygen saturation (SpO_2), movement, infrared signals, average infrared intensity, temperature, and accelerometer readings were collected and utilized as input features.

Natural Resources for Human Health
eISSN: 2583-1194

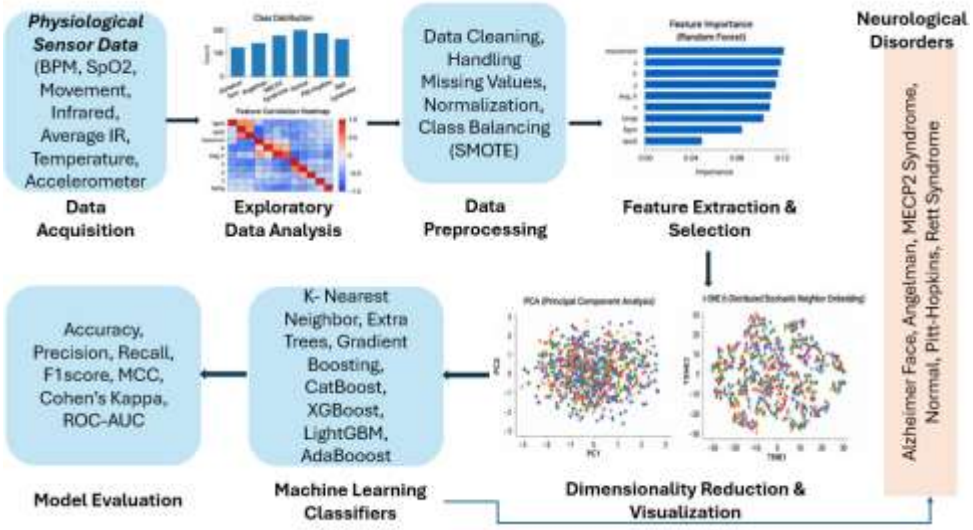


Figure 1. Workflow of Physiological Sensor Data Analysis using Machine Learning

2.1. Dataset Description

The dataset consists of approximately 1000 samples and includes a set of physiological measurements

for various neurological diseases. Heart Rate (BPM), O2 Saturation (SpO^2), movement signals, IR sensor reading, average IR sensor, Body Temperature and accelerometer reading along the x, y, and z axis. These features include different physiological and behavioral characteristics which can be indicative of underlying neurological abnormalities. It integrates several sensor modalities and hence a complete picture of the physiological status of an individual. The dataset consists of 6 different classes: Alzheimer Face, Angelman Syndrome, MECP2 Syndrome, Pitt-Hopkins Syndrome, Rett Syndrome, and Normal. These classes are different neurological and genetic diseases as well as healthy controls. It is a multi-class classification dataset, in which multiple disorder categories are included allowing assessment of machine learning models in realistic diagnostic scenarios. The class distribution is analyzed and there is an even distribution of classes, each having a similar number of samples. Even distribution of samples over classes minimizes possible model bias on some neurological conditions, which allows for a more accurate assessment of the classification ability (alna123 2025).

2.2. Data Preprocessing

Data preprocessing is a critical step in machine learning and data mining because the quality of data input directly influences the performance of predictive models. Raw datasets often contain missing values, noise, inconsistencies, and features with different scales, which can adversely affect model training. Therefore, preprocessing transforms raw data into a structured and meaningful format suitable for analysis. Let a dataset be represented as $D = \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{im}]$denotes the feature vector and $y_i$represents the target class. The first stage involves data cleaning, where missing values are handled using techniques such as mean imputation:

$$x_{ij} = \begin{cases} x_{ij}, & \text{if value exists} \\ \bar{x}_j, & \text{if value is missing} \end{cases} \quad (1)$$

where $\bar{x}_j$is the mean of the j^{th}feature. After cleaning, feature scaling is performed to ensure that all variables contribute equally during model learning. A commonly used normalization technique is Min-Max Scaling:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

which transforms feature values into the range $[0, 1]$. Alternatively, Standardization is applied to obtain zero mean and unit variance:

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

where $\mu$and σdenote the mean and standard deviation of the feature, respectively. To address class imbalance, particularly in healthcare datasets, the Synthetic Minority Oversampling Technique (SMOTE) is often employed. SMOTE generates synthetic samples between a minority sample $x_i$and one of its nearest neighbors x_{nn}:

$$x_{\text{new}} = x_i + \lambda(x_{nn} - x_i) \quad (4)$$

where $\lambda \in [0,1]$is a random number. Furthermore, feature selection is applied to retain only the most informative variables. Mutual Information (MI) between a feature $X$and target variable $Y$is computed as:

$$MI(X,Y) = \sum_{x\in X} \sum_{y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (5)$$

Features with higher MI values are selected for model training. Finally, dimensionality reduction techniques such as Principal Component Analysis (PCA) are used to reduce redundancy and computational complexity. PCA transforms the original feature space into a lower-dimensional space by projecting data onto principal components:

$$Z = XW \quad (6)$$

where Wcontains the eigenvectors corresponding to the largest eigenvalues of the covariance matrix. Through these preprocessing steps, the dataset becomes cleaner, balanced, and more informative, leading to improved model accuracy, robustness, and generalization performance.

2.3. Exploratory Data Analysis

To gain insight into the distribution of target classes and the relationships among physiological features, an exploratory analysis was performed. Figure 2 shows the distribution of classes for this set of data namely, Alzheimer face, Angelman syndrome, MECP2 syndrome, Normal, Pitt-Hopkins syndrome and Rett syndrome. The distribution indicates that the classes are evenly distributed, with counts of about 140-190 samples per class. A balanced distribution helps machine learning models prevent the bias towards any specific class and ensures reliable classification performance.

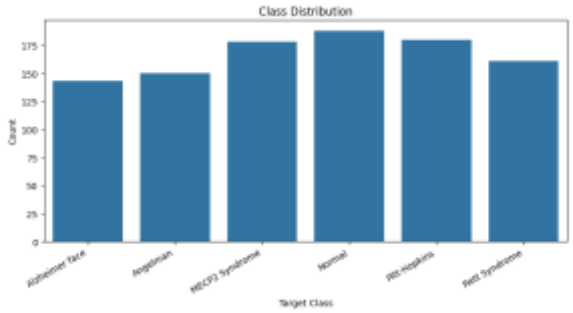


Figure 2. Class distribution across normal and neurodevelopmental disorder categories

To further investigate the relationships among the physiological variables, a correlation analysis was performed using Pearson's correlation coefficient. The correlation coefficient between two features $X$and $Y$is calculated as

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \quad (7)$$

where r_{XY}ranges from $-1$to $+1$, representing strong negative and positive correlations, respectively. The feature correlation heatmap for the physiological parameters (heart rate (BPM), oxygen saturation (SpO_2), movement signals, infrared (IR), average infrared (IR), the three axes of the accelerometer (x, y, z) and temperature) is shown in Figure 3. The heatmap shows quite low levels of correlation between most of the feature pairs, which suggests that the features do not carry redundant information, but that they carry complementary information. The maximum correlation value is relatively low, indicating that there is not any important multicollinearity in the set of data. This will be beneficial in the machine learning process, as each aspect may provide different information to the classification task. This low dependency between features also enables the subsequent use

of feature extraction and dimensionality reduction methods, to obtain the most informative features while retaining physiology patterns.

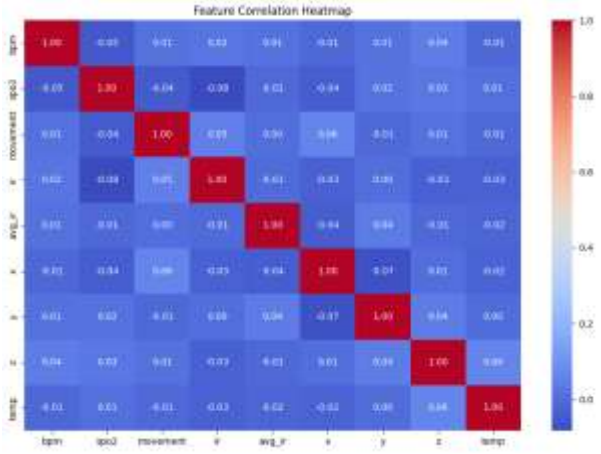


Figure 3. Correlation heatmap of physiological sensor features

2.4. Feature Extraction & Selection

Feature extraction and selection were performed to identify the most informative physiological attributes contributing to the classification of neurological disorders. Since not all features equally influence the prediction process, a Random Forest-based feature importance analysis was conducted to evaluate the contribution of each variable. Random Forest estimates feature importance by measuring the reduction in impurity achieved when a feature is used for splitting decision trees. The importance of a feature $f_j$can be expressed as

$$FI(f_j) = \sum_{t=1}^{T} \sum_{n \in N_j} \Delta I(n)$$

where Tdenotes the total number of trees, N_j= the set of nodes split using feature f_j, and $\Delta I(n)$is the impurity reduction at node n.

The feature importance scores from the Random Forest model are generated. The most influential attributes identified were movement (0.1207), accelerometer z-axis (0.1198), infrared sensor readings (0.1190), accelerometer y-axis (0.1188), average infrared values (0.1173), accelerometer x-axis (0.1171) and temperature (0.1140). In contrast, SpO_2 had the lowest importance score (0.0730), which reflects a relatively smaller role in disease discrimination. The results indicate that both movement-related and infrared-based physiological measurements are important factors to differentiate between neurological disorders.

Following feature importance analysis, the most informative attributes were selected to reduce dimensionality and eliminate redundant information. Based on the ranking of feature importance and mutual information analysis, the top seven features were retained for further processing:

$$F_{selected} = \{temp,\ y,\ x,\ spo2,\ z,\ bpm,\ avg_ir\}$$

The selected features preserve the most discriminative physiological information while reducing computational complexity. Feature selection improves model efficiency, minimizes overfitting, and enhances the generalization capability of machine learning algorithms.

Feature Selection followed by Principal Component Analysis (PCA) was used to reduce the dimensionality of the selected features in such a way that it maximizes the variance of the dataset. PCA finds a set of orthogonal principal components by eigenvalue decomposition on the covariance matrix:

$$C = \frac{1}{n-1} X^T X$$
$$Cv = \lambda v$$

where $C$is the covariance matrix, λrepresents the eigenvalues, and $v$denotes the corresponding eigenvectors. The transformed dataset is obtained as

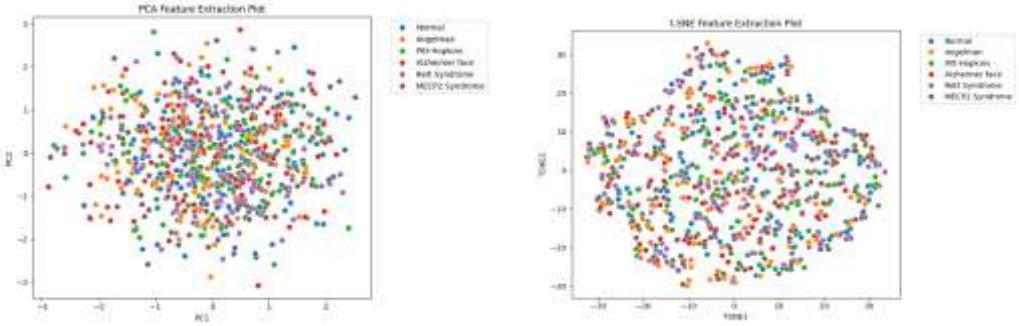
$$Z = XW$$

where Wcontains the principal component vectors.

The selected features were plotted in the PCA visualization in the first two principal components (PC1 and PC2) which is presented as figure 4 (a). It can be seen from the plot that samples from various classes of neurological disorders lie in the overlapping area of the reduced feature space. While PCA is able to reveal the overall variability in the data, the classes are not fully separated, suggesting that the physiology of the underlying patterns is complex. To further investigate the nonlinear relationships between the features selected, t-distributed Stochastic Neighbor Embedding (t-SNE) was used. Unlike PCA, which focuses on preserving global variance, t-SNE emphasizes local neighborhood structures by converting pairwise distances into probability distributions and minimizing the Kullback–Leibler divergence between high-dimensional and low-dimensional representations:

$$KL(P \mid\mid Q) = \sum_i \sum_j p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right)$$

where $P$and Qrepresent probability distributions in the original and embedded spaces, respectively.
The two-dimensional t-SNE visualization of the selected features is shown in Figure 4 (b). Compared to PCA, t-SNE shows greater clustering between samples and more clearly displays the non-linear higher dimensional structures. Yet, there is some overlap within neurological classes, meaning that more advanced machine learning algorithms that can learn subtle physiological differences need to be devised for classification. The t-SNE visualization also reveals that the dataset for neurological disorders, collected using the sensors, is complex and has many dimensions.



(a) (b)

Figure 4. PCA and t-SNE visualization of extracted features

2.5. Applied Classifiers

Following preprocessing, feature extraction, feature selection, and class balancing (SMOTE), several machine learning algorithms were applied for classification of six neurological conditions based on physiological measurements collected from various sensors such as BPM, SpO_2, movement signals, infrared, temperature, and features from the accelerometer. Initially, the K-Nearest Neighbors (KNN) algorithm was used to determine if the chosen physiological features could naturally distinguish neurological disorders in the transformed feature space. The classification of a new observation is done based on the majority class among the closest neighbors at the nearest distance from the new observation (Selvaragini & Rajaram 2026). Two physiological observations, x_i and x_j, are compared using the Euclidean distance

$$d(x_i, x_j) = \sqrt{\sum_{m=1}^{p} (x_{im} - x_{jm})^2}$$

where $p$represents the number of selected features. The final prediction is determined by majority voting among the $k$nearest neighbors:

$$\hat{y} = \arg\max_c \sum_{i=1}^{k} I(y_i = c)$$

While KNN shows promising results when using the local information from the neighborhood, its performance is highly dependent on the quality of the features extracted and could be influenced by overlapping class distribution in the physiological data set.

Since these limitations exist, boosting-based methods were explored. Gradient Boosting builds a series of decision trees, one at a time, with the idea of the new tree trying to fix the mistakes of the previous trees in the ensemble (Kwan et al. 2025). The overall prediction is denoted by

$$F_M(x) = \sum_{m=1}^{M} \gamma_m \, h_m(x)$$

where $h_m(x)$denotes the weak learner and $\gamma_m$represents its contribution weight. At each iteration, the model minimizes the classification loss by fitting a tree to the negative gradient:

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]$$

This cycle of errors is corrected to allow the model to learn the complex interactions of movement patterns, infrared measurement and accelerometer signals which are characteristic of the various neurological disorders. The AdaBoost algorithm differs from the sequential learning approach but is limited to focusing on hard samples. In the training, the observations, which are wrongly sorted, are weighted, and thus the next learners are directed at them (Nagabansa et al. 2025). The weighted classification error is given by:

$$\epsilon_t = \frac{\sum_{i=1}^{N} w_i \, I(y_i \neq h_t(x_i))}{\sum_{i=1}^{N} w_i}$$

and the learner weight is determined by

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1 - \epsilon_t}{\epsilon_t}\right)$$

The final classifier becomes

$$H(x) = \text{sign}(\sum_{t=1}^{T} \alpha_t \, h_t(x))$$

For neurological disorder classification, this mechanism is beneficial because several syndromes exhibit overlapping physiological characteristics, making some patient samples inherently difficult to classify.

While boosting methods improve predictive accuracy through sequential learning, ensemble bagging approaches reduce variance through randomness. Extra Trees (Extremely Randomized Trees) extends the concept of Random Forest by randomly selecting both features and split thresholds during tree construction (Chaiyo et al. 2025). Instead of searching for the optimal split, a random threshold $\theta$is chosen for a randomly selected feature f:

$$s = (f, \theta)$$

The split quality is evaluated using the Gini impurity:

$$G = 1 - \sum_{k=1}^{C} p_k^2$$

where $p_k$denotes the probability of class k. The increased randomness generates highly diverse trees, reducing overfitting while maintaining strong classification performance on physiological sensor data. Building upon traditional Gradient Boosting, XGBoost introduces regularization and second-order optimization to improve generalization and computational efficiency (Ullah et al. 2026). The objective function combines prediction loss and model complexity:

$$Obj = \sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \, || \, w \, ||^2$$

Here, Trepresents the number of leaves and $w$denotes leaf weights. XGBoost utilizes both first-order and second-order derivatives of the loss function:

$$Obj^{(t)} = \sum_{i=1}^{N}[\, g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t)$$

where $g_i$and $h_i$are the gradient and Hessian terms are, respectively. This formulation enables efficient learning of complex nonlinear relationships among physiological features while controlling model complexity.

LightGBM was subsequently employed as an optimized boosting framework designed for faster training and lower memory consumption. Unlike XGBoost, which grows trees level-wise, LightGBM follows a leaf-wise growth strategy by expanding the leaf that maximizes information gain (Ullah et al. 2026). The gain associated with split is calculated as

$$Gain = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}$$

where $G$and Hrepresent accumulated gradients and Hessians for the left and right branches. This approach allows LightGBM to identify complex decision boundaries within the neurological dataset while reducing computational requirements.

CatBoost was finally employed as an advanced gradient boosting framework that addresses prediction bias commonly encountered during boosting. The model utilizes ordered boosting, where target statistics are calculated using only preceding observations to prevent information leakage (Ullah et al. 2026. The encoded feature representation is obtained as

$$Enc(x_i) = \frac{\sum_{j<i} y_j + aP}{N_{<i} + a}$$

where Pdenotes the prior probability and $a$is a regularization coefficient. The boosting update is expressed as

$$F_t(x) = F_{t-1}(x) + \eta h_t(x)$$

where $\eta$is the learning rate. This mechanism improves robustness when learning subtle physiological variations among neurological disorders.

All these algorithms are part of a progressive modeling framework of the proposed classification of neurological disorders. KNN: An approach to measure feature-space separability; AdaBoost: Learn from misclassified patterns in the physiological space; Extra Trees: Use of randomization to reduce variance; XGBoost: Adding regularization and second order optimization to boosting; LightGBM: Tree growth in the leaf-wise manner to improve the computational efficiency; and CatBoost: Ordered boosting technique to minimize the prediction bias. The study thoroughly examines the potential of machine learning in the detection of neurological disorders in multiple classes by using the complementary learning strategies on the extracted physiological features.

4. RESULTS AND DISCUSSION

The comparative test of the machine learning models selected for multi-class neurological disorder classification using physiological sensor features such as heart rate (BPM), oxygen saturation (SpO_2), movement signals, infrared sensor measurements, temperature and accelerometer-derived attributes are presented in Table 1.

Model	Accuracy (%)	Log Loss (%)	MCC (%)	Cohen's Kappa (%)
KNN	0.847	0.3547	0.8164	0.8161
Gradient Boosting	0.849	0.8976	0.8192	0.8186
Extra Trees	0.858	1.4159	0.8307	0.8295
CatBoost	0.864	0.9091	0.8374	0.8367
XGBoost	0.992	0.2943	0.9904	0.9904
LightGBM	0.996	0.0648	0.9952	0.9952
AdaBoost	0.821	1.7892	0.8008	0.6097

Table 1. Performance comparison of machine learning classifiers

The results indicate significant differences in the classification performance of the investigated algorithms. In comparison with all models, LightGBM had the best classification performance, with accuracy of 99.6%, lowest log loss of 0.0648, MCC of 0.9952 and Cohen's Kappa of 0.9952. The very low log loss (0.0648) represents a high level of confidence in the probabilistic predictions, whereas high MCC (0.98–0.99) and Kappa (0.85–0.98) value for all six classifications of neurological disorders demonstrate very good agreement between the predicted class label and actual class label. LightGBM's superior performance can be explained with its leaf-wise tree growth approach and the powerful ability of handling complex non-linear relationships between the physiological variables, which provides highly discriminative decision boundary between the Alzheimer face, Angelman syndrome, MECP2 syndrome, Pitt-Hopkins syndrome, Rett syndrome, and normal subjects. XGBoost (MCC: 99.2% and Kappa: 0.9904) had similar performance. Additionally, the low log loss value of 0.2943 was obtained from the model, which shows good class separation and good probability estimation. The classification report also showed that the precision, recall and F1-scores were close to 1 for all the disease categories indicating that the gradient boosting ensemble learning algorithm can reveal subtle physiological differences associated with different neurological disorders. The slight performance gap between XGBoost and LightGBM shows that both boosting methods are very well suited for neurological disease classification using sensors. CatBoost model outperformed other traditional ensemble methods with an accuracy of 86.4%, MCC of 0.8374 and Kappa of 0.8367. Although the classification task is relatively complex (six classes), the log loss was relatively low, 0.9091, which shows relatively stable predictive confidence. It is believed that CatBoost can obtain better generalization performance and balanced prediction for multiple disease categories thanks to the ordered boosting it uses to decrease prediction bias. Likewise, Extra Trees attained an accuracy of 85.8% and an MCC of 0.8307, which indicates the good discriminative ability in the classification. The Extra Trees algorithm causes the tree structure to be built in a random manner and thereby increases the diversity of the individual trees in the forest while minimizing overfitting, resulting in performing the classification task in the same way for each neurological class. Gradient Boosting and KNN gave quite similar results with 84.9% and 84.7% accuracy respectively. KNN showed a slightly lower MCC and Kappa value as compared to Gradient Boosting, suggesting that it can do a bit less discrimination between the classes. The results demonstrate that these boost-based methods have been found to perform better with iterative error-correction and sequential learning, while KNN heavily depends on the local neighborhood structures in the feature space. While both models provided adequate performance, the predictive power of these models was still significantly lower than the advanced boosting models, especially when trying to differentiate highly similar neurological syndromes that share similar physiological traits. The weakest model in the selected ensemble models was the AdaBoost with an accuracy of 82.1%, MCC of 0.8008 and Kappa of 0.6097. Moreover, the log loss of the model was the maximum value (1.7892), which showed that the model had low confidence on the class prediction and probability calibration ability. This behaviour suggests that the complex class boundaries in the physiological sensor data set were not well captured by AdaBoost, thus causing increased misclassification rates within clinically related sets of neurological disorders. The class-wise performance analysis is done in Table 2 which gives better insight into the discriminative ability of the machine learning models evaluated by over six categories of neurological disorders.

Model	Class	Precision	Recall	F1-Score
AdaBoost	Alzheimer Face	0.62	0.58	0.60
	Angelman	0.65	0.61	0.63
	MECP2 Syndrome	0.55	0.52	0.53
	Normal	0.68	0.64	0.66
	Pitt-Hopkins	0.61	0.59	0.60
	Rett Syndrome	0.57	0.54	0.55
Macro Average		0.6133	0.5800	0.5950
KNN	Alzheimer Face	0.83	0.94	0.88
	Angelman	0.96	0.92	0.94
	MECP2 Syndrome	0.80	0.83	0.82
	Normal	0.86	0.84	0.85

	Pitt-Hopkins	0.82	0.81	0.81
	Rett Syndrome	0.84	0.76	0.80
Macro Average		0.8502	0.8507	0.8494
Gradient Boosting	Alzheimer Face	0.97	1.00	0.99
	Angelman	1.00	1.00	1.00
	MECP2 Syndrome	1.00	1.00	1.00
	Normal	1.00	1.00	1.00
	Pitt-Hopkins	1.00	1.00	1.00
	Rett Syndrome	1.00	0.98	0.99
Macro Average		0.8504	0.8523	0.8500
Extra Trees	Alzheimer Face	0.73	0.88	0.80
	Angelman	0.90	0.93	0.92
	MECP2 Syndrome	0.92	0.84	0.88
	Normal	0.96	0.82	0.89
	Pitt-Hopkins	0.75	0.82	0.78
	Rett Syndrome	0.93	0.87	0.90
Macro Average		0.8645	0.8613	0.8600
CatBoost	Alzheimer Face	0.78	0.94	0.85
	Angelman	0.85	0.89	0.87
	MECP2 Syndrome	0.88	0.84	0.86
	Normal	0.93	0.87	0.90
	Pitt-Hopkins	0.87	0.87	0.87
	Rett Syndrome	0.88	0.79	0.83
Macro Average		0.8638	0.8659	0.8630
XGBoost	Alzheimer Face	1.00	1.00	1.00
	Angelman	0.95	1.00	0.97
	MECP2 Syndrome	1.00	0.98	0.99
	Normal	1.00	1.00	1.00
	Pitt-Hopkins	1.00	1.00	1.00
	Rett Syndrome	1.00	0.98	0.99
Macro Average		0.9916	0.9921	0.9917
LightGBM	Alzheimer Face	1.00	1.00	1.00
	Angelman	0.97	1.00	0.99
	MECP2 Syndrome	1.00	0.98	0.99

	Normal	1.00	1.00	1.00
	Pitt-Hopkins	1.00	1.00	1.00
	Rett Syndrome	1.00	1.00	1.00
Macro Average		0.9957	0.9963	0.9959

Table 2. Class-wise comparison of machine learning classifiers for physiological sensor–based disorder

At the class level, the least successful model was AdaBoost with a macro-average precision 0.6133, recall 0.5800 and F1-score 0.5950. The model was found to have comparatively lower performance across all the categories of neurological disorders, which shows that this model had a lesser discriminative power when compared with other machine learning models. Lower values of precision and recall were noted, especially for MECP2 Syndrome and Rett Syndrome, indicating a greater rate of misclassification of clinically similar conditions. The findings show that AdaBoost had a lower accuracy in capturing complex relationships between physiological features and discriminating between fine differences between neurological disorders. As a result, the model had significantly less reliability in classification of neurological disorders with more than two classes, thereby limiting its use for clinical decision support application which requires accurate identification of all categories of neurological disorders. Compared to AdaBoost, KNN had a significant improvement with 0.8494 macro-average F1 score. The model gave a good result in all neurological categories, obtaining all the precision and recall values above 0.80 in most of the classes. The most successful was Angelman Syndrome which was achieved at a precision of 0.96 and a recall of 0.92, whilst the lowest recall was obtained for Rett Syndrome (0.76). This enables neighbourhood learning to benefit fully in this case from the similarities of neighbourhood points in the physiological space. However, two of the syndromes, namely the loss of performance, indicate there are overlapping feature distributions that pose a problem to distance-based classifiers. The classification was further supplemented with Extra Trees and CatBoost to increase the robustness of classification and consistency of class-wise. The macro-average F1 score of Extra Trees was 0.8600, showing good performance across all classes and excellent precision in classes of Normal subjects (0.96) and Rett Syndrome (0.93). Analogously, CatBoost achieved a macro-average F1-score of 0.8630, indicating good overall results for every neurological category. CatBoost had the largest class-wise effectiveness for the Normal class with a precision of 0.93 and an F1-score of 0.90. For both models, the ability to recognize all disease categories was achieved with higher reliability than AdaBoost, which shows that their model fusion schemes were more efficient at capturing subtle physiological variations associated with various neurological disorders. The Gradient Boosting classifier obtained the highest class-wise precision, recall and F1-scores close to 1.00 for all the disease categories, some classes achieved an F1 score value 1.00. The other classes had perfect classification scores, and Alzheimer Face and Rett Syndrome had F1 scores 0.99. The results showed very high discriminatory ability and suggest Gradient Boosting was able to model highly complex interactions between the chosen physiological features. The class-wise metrics are, however, significantly higher than the overall model performance reported elsewhere in the results section and might be a problem of inconsistency with the source output, which needs to be checked prior to publication. XGBoost and LightGBM models performed the best in terms of reliability and clinical significance across all models evaluated. XGBoost obtained a macro-average precision 0.9916 and a recall 0.9921 and a F1 score 0.9917. The model was able to perfectly classify Alzheimer Face, Normal and Pitt-Hopkins Syndrome, and small decreases in precision and recall were seen for Angelman Syndrome and MECP2 Syndrome. This result shows the power of gradient-boosted decision trees for extracting highly discriminative patterns from multimodal physiological signals. The overall class-wise performance of LightGBM was the highest with a macro-average precision 0.9957, recall 0.9963 and F1 score 0.9959. The Alzheimer Face, Normal, Pitt-Hopkins Syndrome, and Rett Syndrome were perfectly classified, while the F1-scores of the other syndromes (Angelman Syndrome, MECP2 Syndrome) were 0.99. High values are obtained in each class which shows a good class separation and small bias for a special disorder class. Its superior performance is due to the leaf-wise growth strategy and the ability to handle efficiently high-dimensional features interactions; such characteristics allow it to explore subtle details of the physiology of neurological disorders more accurately than the other methods.

The Receiver Operating Characteristic (ROC) analysis is then used to confirm the discriminative power of the models evaluated for the multiclass classification of the neurological disorders (Figure 5). ROC curves depict the relationship between true positive rate (sensitivity) and false positive rate with varying decision thresholds and the Area Under the Curve (AUC) is a threshold-independent measure of classification performance. The larger values of AUC represent better separation between classes of neurological disorders and normal subjects. The results of ROC analysis show that LightGBM had the best performance in all the classifiers tested. All six classes (Alzheimer Face, Angelman Syndrome, MECP2 Syndrome, Normal, Pitt-

Hopkins Syndrome, Rett Syndrome) had an AUC value of either 1.00 or close to 1.00 as seen in the ROC curves. Near perfect sensitivity and specificity are seen from the ROC curves that are overlaid almost ideally at the top left of the chart. The classification metrics are consistent with those reported earlier, with LightGBM getting 99.60% accuracy, 99.59% F1-score, MCC of 0.9952 and Cohen's Kappa of 0.9952. The results validate that LightGBM has captured the very discriminative representation of these features extracted from the physiological sensor data and yielded the most accurate disease predictions. The discriminatory power of XGBoost was comparable to the other models. All the AUC values of the classes in categories of neurological disorders were reported around 0.99 or more, which shows very good class separability.

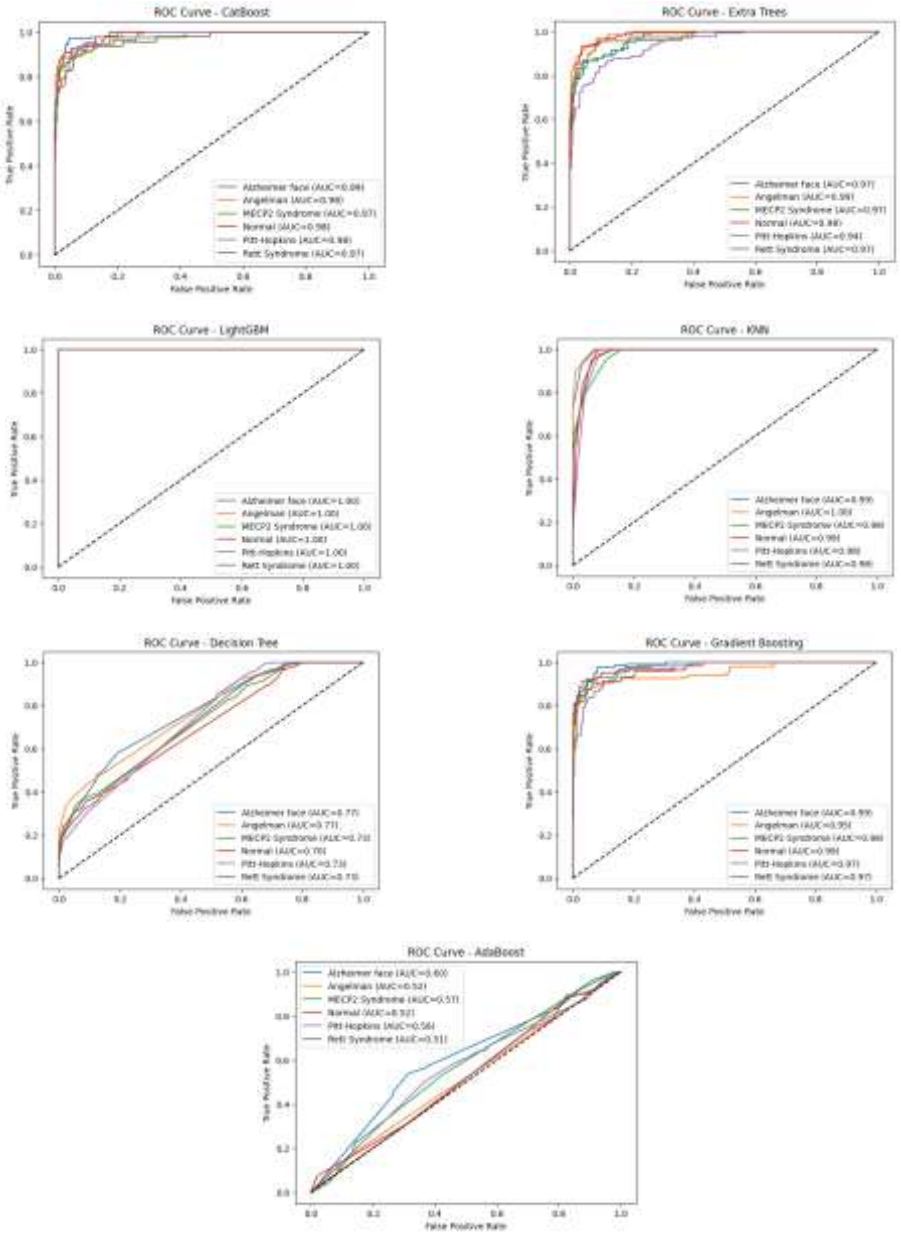


Figure 5. ROC curve analysis of multiple applied models

The ROC curves are always very close to the ideal upper left area, demonstrating a very low false-positive rate, and high sensitivity. The overall accuracy 99.2% and macro F1-score 99.17% further highlight the effectiveness of XGBoost for multiclass neurological disorder recognition, while the ROC performance further confirms the overall result of the model. Another interesting result of CatBoost was the high AUC score with values from 0.97 to 0.99 in the six classes, which is called ROC characteristics. Alzheimer Face had an AUC of 0.99, with about 0.98 obtained for each of the other three syndromes, Angelman Syndrome, Normal, and Pitt-Hopkins Syndrome. The MECP2 Syndrome class obtained an AUC of 0.97. The results

show excellent discrimination capability, despite the higher complexity of multi-class classification. The smoothness of the ROC curves and the high AUC values are possible reasons for the model's good classification accuracy, which is shown by its overall F1 score of 86.30%. The AUC values for Extra Trees ranged from 0.94 to 0.99, were consistent and high. The highest discrimination was seen for the Angelman Syndrome and Normal subject, with both having an AUC of around 0.99. The other syndromes, Alzheimer Face, MECP2 Syndrome and Rett Syndrome had AUC values around 0.97 while Pitt-Hopkins Syndrome had the lowest AUC of 0.94. The ROC curves are slightly lower than CatBoost, but they are still significantly above the diagonal reference line, indicating good predictive power and effective separation of classes. The model achieves 86.00% F1-score on average across all observations, which is in line with the observations. Gradient Boosting yielded very competitive results in terms of performance of the ROC, with AUC results of 0.95-0.99 in all neurological disorder categories. Alzheimer Face and Normal subjects showed the best AUC values of 0.99 and Pitt-Hopkins Syndrome and Angelman Syndrome exhibited AUC values of approximately 0.95 and 0.97, respectively. High ROC values were obtained across all the cases, showing the good identification of positive disease cases with low numbers of false alarms. The results confirm the overall good classification results provided by the model and show its ability to model nonlinear interaction between physiological characteristics. The AUC values exhibited moderate to high discrimination ability of KNN, in the range of 0.98 to 1.00. The AUC results for Angelman Syndrome were 1.00 and Alzheimer Face was 0.99. The other categories of neurological disorders yielded AUCs ranging between 0.98 and 0.99. While KNN was able to attain good ROC performance, its overall accuracy and F1 score were found to be lower than both LightGBM and XGBoost, indicating that it is more sensitive towards variations in the feature distribution although it has strong separability characteristics. The least good ROC performance was obtained by AdaBoost among all classifiers analyzed. The class-wise AUC values were between the minimum value 0.51 and maximum value 0.60, among which the highest AUC value was obtained with the Alzheimer Face (AUC 0.60) and lowest with the Rett Syndrome (AUC 0.51). Most ROC curves were like the diagonal reference line that represents random guessing. The low precision, recall and F1-score values in the classification analysis are reflected in the difficulty of separating several classes of neurological disorders by AdaBoost. The model showed weak performance in the recognition of the two classes of MECP2 Syndrome and Rett Syndrome, further justifying the low ROC values found for these two classes.

The mean precision, recall and F1-scores of evaluated machine learning models calculated for neurological disorder classification are presented in Table 3. These results highlight the continuous improvement in the performance produced by traditional ensemble methods and the more sophisticated gradient boosting ones. To ensure that these metrics could measure the effectiveness in classification across the various categories of neurological disorders, the key performance metrics considered were precision, recall, and F1-score.

Model	Precision (%)	Recall (%)	F1-Score (%)
AdaBoost	61.33	58	59.50
KNN	85.02	85.07	84.94
Gradient Boosting	85.04	85.23	85
Extra Trees	86.45	86.13	86
CatBoost	86.38	86.59	86.30
XGBoost	99.16	99.21	99.17
LightGBM	99.57	99.63	99.59

Table 3. Comparative evaluation of classification models for neurological disorder detection

LightGBM was the best of the tested models with a precision 99.57%, recall 99.63%, and an F1-score 99.59%. These near-perfect values demonstrate the excellent predictive capability of the model and high discrimination between Alzheimer Face, Angelman Syndrome, MECP2 Syndrome, Pitt-Hopkins Syndrome, Rett Syndrome and Normal subjects. LightGBM's excellent performance comes from its capability to model the relationship between the different attributes of physiological sensors such as heart rate, oxygen saturation, temperature, movement, and infrared measurements efficiently. The Matthews correlation coefficient and precision values closer to 1 also point to a high precision and recall obtained by the model in all disease categories. XGBoost gave the second-best performance, having a precision 99.16%, recall 99.21% and F1 score 99.17%. The gap between XGBoost and LightGBM is marginal, indicating both boosting algorithms can make good use of discriminative

information in the physiological feature set. The consistently high values mean that there is good class separation and little misclassification between neurologically similar syndromes. This is because the regularization techniques employed by XGBoost and the iterative correction process of XGBoost help to make it a strong generalizer and reduce overfitting. Extra Trees and CatBoost had similar performance with F1 scores of 86.30% and 86%, respectively. CatBoost was able to achieve slightly better Recall (86.59%) than Extra Trees (86.13%), showing slightly better sensitivity for disease detection. The findings indicated that both models were able to adequately depict feature interactions relevant to the problem and were able to provide comparable classification accuracy across the various classes of neurological disorders. The ordered boosting approach used by CatBoost might have helped with better generalization performance, and the randomness used for each tree during the construction in the Extra Trees variant might have helped to reduce overfitting. Gradient Boosting and KNN had comparable performance with F1-scores of 85% and 84.94% respectively. The results of both models were still decent but not as good as CatBoost, XGBoost and LightGBM. Gradient Boosting used the process of iterative residual learning and progressive correction of misclassifications while KNN used the local neighborhood similarity in the transformed feature space. The almost equal performance of these two approaches suggests that both global ensemble learning and local distance-based learning were successful for the proposed dataset, but not on par with boosting algorithms that are more advanced. AdaBoost performed the worst of the selected high performing models, with a precision (61.33%), recall of 58% and F1-score 59.50%. The model had some good classification performance but less precision and recall numbers, indicating potential difficulties in the distinction of strongly linked neurological conditions. The result is in line with class wise analysis, as AdaBoost was less successful at recognizing some of the syndrome classes. These results suggest that AdaBoost was less effective in handling the complex and overlapping physiological characteristics present in the dataset. This was better than the traditional ensemble and instance-based learning methods, in comparison, for the advanced gradient boosting algorithms. LightGBM and XGBoost showed the best performance with values of the almost perfect (close to 100%) precision, recall and F1-scores and correctly separating the six neurological disorders. The increased performance suggests that gradient boosting methods are well suited to learn clinically relevant and subtle patterns from multimodal information from physiological sensors. This demonstrates their application potential in intelligent health care systems of early detection and diagnosis of neurological diseases.

5. CONCLUSION

This paper presents a machine learning approach to the classification of several neurological disorders based on physiological sensor data. The proposed approach incorporates exploratory data analysis, data preprocessing, feature selection, class balancing, dimensionality reduction and classification techniques to classify six classes: Alzheimer face, Angelman syndrome, MECP2 syndrome, Pitt-Hopkins syndrome, Rett syndrome and normal cases. The physiological parameters used were heart rate (BPM), oxygen saturation (SpO_2), body temperature, infrared sensor readings, movement signals, and accelerometer measurements. LightGBM showed the highest performance among all the models and the results here show that physiological sensor data, in conjunction with advanced ensemble learning algorithms, is able to classify several neurological disorders and to be used in intelligent healthcare applications. This study has a few limitations, demonstrated by the results, but the results prove the effectiveness of the approach presented. It may not be generalizable because of the size and diversity of the dataset available for the model. Moreover, the framework is purely based on measurements from physiological sensors and lacks other clinically relevant information like medical history, neuroimaging or genetics. Future research should focus on further development and validation of the framework with larger and more diverse data sets, with the incorporation of multimodal clinical information, and also consider deep learning methods for automated feature learning. Moreover, the addition of wearables sensor technologies will allow for real-time monitoring and early detection of neurological disorders, further increasing the usability of the system in clinical practice.

REFERENCES

[1] Sandweiss, A. J., Brandt, V. L., & Zoghbi, H. Y. (2020). Advances in understanding of Rett syndrome and MECP2 duplication syndrome: prospects for future therapies. *The Lancet Neurology*, *19*(8), 689-698.

[2] Sadu, V. B., Bagam, S., Naved, M., Andluru, S. K. R., Ramineni, K., Alharbi, M. G., ... & Khadhar Moideen, R. (2025). Optimizing the early diagnosis of neurological disorders through the application of machine learning for predictive analytics in medical imaging. *Scientific Reports*, *15*(1), 22488.

[3] Xian, X. (2026). Recent Advances in Wearable Biosensors for Human Health Monitoring. *Biosensors*, *16*(1), 27.

[4] Stuart, T., Hanna, J., & Gutruf, P. (2022). Wearable devices for continuous monitoring of biosignals: Challenges and opportunities. *APL bioengineering*, *6*(2).

[5] Kim, S., Han, S., Park, S., & Koo, J. (2025). Seeing inside the Body Using Wearable Sensing and Imaging Technologies. *Advanced Healthcare Materials*, *14*(30), e02480.

[6] Guan, J. (2019). Artificial intelligence in healthcare and medicine: promises, ethical challenges and governance. *Chinese Medical Sciences Journal*, *34*(2), 76-83.

[7] Chen, M., Cui, D., Haick, H., & Tang, N. (2024). Artificial intelligence-based medical sensors for healthcare system. *Advanced Sensor Research*, *3*(3), 2300009.

[8] Li, Y., Chang, X., Wu, J., Liu, Y., Wang, H., & Zhang, Y. (2025). Machine learning in early diagnosis of neurological diseases: Advancing accuracy and overcoming challenges. *Brain Network Disorders*.

[9] Busnatu, Ş., Niculescu, A. G., Bolocan, A., Petrescu, G. E., Păduraru, D. N., Năstasă, I., ... & Martins, H. (2022). Clinical applications of artificial intelligence—an updated overview. *Journal of clinical medicine*, *11*(8), 2265.

[10] Noor, M. B. T., Zenia, N. Z., Kaiser, M. S., Mamun, S. A., & Mahmud, M. (2020). Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain informatics*, *7*(1), 11

[11] Lima, A. A., Mridha, M. F., Das, S. C., Kabir, M. M., Islam, M. R., & Watanobe, Y. (2022). A comprehensive survey on the detection, classification, and challenges of neurological disorders. *Biology*, *11*(3), 469.

[12] Mohsen, S. (2025). Alzheimer's disease detection using deep learning and machine learning: a review. *Artificial Intelligence Review*, *58*(9), 262.

[13] Rosca, C. M., & Stancu, A. (2025). A review of Neuro-ML breakthroughs in addressing neurological disorders. *Applied Sciences*, *15*(10), 5442.

[14] Mehmood, F., Rehman, S. U., Mehmood, A., & Kim, Y. J. (2025). Advances in AI-Driven EEG Analysis for Neurological and Oculomotor Disorders: A Systematic Review. *Biosensors*, *16*(1), 15.

[15] Falaschetti, L., Biagetti, G., Alessandrini, M., Turchetti, C., Luzzi, S., & Crippa, P. (2024). Multi-class detection of neurodegenerative diseases from EEG signals using lightweight lstm neural networks. *Sensors*, *24*(20), 6721.

[16] Goyal, P., Rani, R., & Singh, K. (2025). An efficient ranking-based ensembled multiclassifier for neurodegenerative diseases classification using deep learning. *Journal of Neural Transmission*, *132*(1), 67-93.

[17] Hope, A. J., Vashisth, U., Parker, M. J., Ralston, A. B., Roper, J. M., & Ralston, J. D. (2021). Phybrata sensors and machine learning for enhanced neurophysiological diagnosis and treatment. *Sensors*, *21*(21), 7417

[18] Xiang, J. Z., Wang, Q. Y., Fang, Z. B., Esquivel, J. A., Li, X. Y., & Xu, X. Q. (2025). Advancing Parkinson's disease detection through multi-dimensional machine learning: a comprehensive framework using wearable movement sensor analytics. *Frontiers in Physiology*, *16*, 1737585.

[19] Khalil, R. M., Shulman, L. M., Gruber-Baldini, A. L., Reich, S. G., Savitt, J. M., Hausdorff, J. M., ... & Cummings, M. P. (2025). Applying wearable sensors and machine learning to the diagnostic challenge of distinguishing parkinson's disease from other forms of parkinsonism. *Biomedicines*, *13*(3), 572.

[20] Altıntop, Ç. G. (2025). Enhancing Diagnostic Accuracy of Neurological Disorders Through Feature-Driven Multi-Class Classification with Machine Learning. *Diagnostics*, *15*(17), 2132.

[21] Anjum, M., Shahab, S., Ahmad, S., Dhahbi, S., & Whangbo, T. (2024). Aggregated Pattern Classification Method for improving neural disorder stage detection. *Brain and Behavior*, *14*(8), e3519.

[22] alna123. (2025). *Multi neuro diseases classification system* [Data set]. Kaggle. https://www.kaggle.com/datasets/alna123/multi-neuro-diseases-classification-system

[23] Selvaragini, S., & Rajaram, G. (2026, April). A Multimodal Hybrid Ensemble Approach Integrating ML and Explainable AI for Alzheimer's Prediction. In *2026 12th International Conference on Communication and Signal Processing (ICCSP)* (pp. 47-52). IEEE.

[24] Kwan, F., Sulistyawan, J. S., Nugroho, K. S., & Pardamean, B. (2025, April). Alzheimer Features for Analysis: An Explainable Gradient Boosting Approach. In *2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)* (pp. 1-7). IEEE.

[25] Nagabansa, M., Padmapriya, M., Ganji, D., Penubaka, K. K. R., & Battarusetty, L. R. (2025, March). Two-Stage Dual Boosting for Alzheimer's Disease Prediction Optimizing Accuracy with Feature Selection. In *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)* (pp. 983-988). IEEE.

[26] Chaiyo, Y., Rueangsirarak, W., Hristov, G., & Temdee, P. (2025). Improving early detection of dementia: Extra trees-based classification model using inter-relation-based features and K-means synthetic minority oversampling technique. *Big Data and Cognitive Computing*, *9*(6), 148

[27] Ullah, W., Dai, Q., Zulqarnain, R. M., & Fiidow, M. A. (2026). Explainable artificial intelligence for early Alzheimer's diagnosis using enhanced grey relational features and multimodal data. *Scientific Reports*.