AJCSE ASIAN JOURNAL OF COMPUTER SCIENCE ENGINEERING

Available Online at www.ajcse.info
Asian Journal of Computer Science Engineering 2025;10(3)

RESEARCH ARTICLE

Machine Learning-Based Disease Classification Models for Parkinson's Based on MRI Images

Dr. Pradeep Laxkar*

Associate Professor
Department of Computer Science and Engineering ITM (SLS) University Vadodara, Gujarat

Received on: 19-07-2025; Revised on: 11-08-2025; Accepted on: 22-09-2025

***Abstract*—Parkinson's disease (PD) is a slowly advancing neurological problem of the central nervous system that is manifested by shaking, rigidity, and slowness of movement. Effective early diagnosis is a must; usually, it includes detailed physical tests and analysis of medical history. This study presents an early-stage Parkinson's disease prediction system based on biological voice characteristics and machine learning (ML). In the study, the researcher will use a publicly accessible dataset that is on Kaggle to discriminate between healthy and affected people using advanced classification methods. Exploratory data analysis (EDA) shows feature correlations and class imbalance, making it possible to advance a systematic data processing pipeline that involves cleaning data, identifying outliers, and standardizing data. This was done in order to improve model performance by removing some features that are not important using feature selection, which reduces dimensionality and computational complexity. They created and assessed two models: Logistic Regression (LR) and Extreme Gradient Boosting (XGBoost), utilizing the ROC curve, F1-score, accuracy, precision, recall, and confusion matrix. The experimental results demonstrated that the XGBoost model outperformed the LR and could be used to make an early diagnosis of Parkinson's disease, with an F1-score of 98.3, an accuracy rate of 97.4, and an AUC of 0.9833. These results demonstrate that XGBoost is a useful diagnostic tool that can assist medical professionals in early Parkinson disease detection.**

Keywords—Parkinson's Disease (PD), Early Diagnosis, Voice Recordings, Neurogenerative Disorder, Medical Diagnosis, Clinical Decision Support.

INTRODUCTION

Alzheimer's disease is the most prevalent neurological condition, followed by Parkinson's. Parkinson's disease (PD)-specific symptoms include bradykinesia, resting tremor, hypokinetic movement disorder, muscle stiffness, and unstable posture and steps [1][2][3]. Besides, non-motor characteristics, including dementia, depression and dysautonomia, were outlined. The general disturbances of the motor system on PD are known as Parkinsonism. It is noteworthy that Parkinsonism is primarily linked to PD, but other disorders, including AD-related and PD-related diseases, have identical characteristics.

To effectively intervene, PD must be identified early and manage the disease in time because such identification enables instilling of the right treatments and interventions, which are capable of improving the outcomes of the patients [4][5]. The conventional PD diagnostic tools are by clinical observation and subjective assessment which can only be used to misdiagnose a condition and postponed treatment. The benefits of biosensors, which are easy to build, inexpensive, ready available, and simple to interpret and read, have provided it the potential of becoming an alternative and more promising method of early detection of PD [6][7]. However traditional biosensors have some drawbacks such as limited sensitivity, difficulty in detecting the target molecule in low concentration and low anti-interference ability.

The most common diagnostic method for detecting Parkinson's disease (PD) early on is the examination of brain magnetic resonance imaging (MRI) data. The brain's subcortical structures are shown anatomically in the MRI images, which are then examined to ensure that no aneurysms are present. This information is also thought to be helpful in the early detection of specific disease types. However, because the MRI is a three-dimensional structure, using the human eye to explore the nuances and various features of subcortical areas can be challenging [8][9]. Thus, by utilizing multidimensional healthcare data, computer-aided detection systems have demonstrated remarkable efficacy in illness analysis and diagnosis as intelligent technologies have advanced.

The latest developments in deep learning (DL) and machine learning (ML), two branches of artificial intelligence (AI), are helping doctors diagnose diseases early. As a result, recent studies have used a range of AI and ML algorithms to automatically detect Parkinson's disease from MRI data [10][11]. DL has been used to detect many different diseases and conditions, and the results often surpass conventional benchmarks. DL algorithms are very powerful and often used for image categorization tasks. Because they can recognize intricate patterns and characteristics from pictures, they outperform the outdated ML techniques in terms of accuracy.

Motivation and Contribution of the Study

The motor system is impacted by Parkinson disease (PD), a degenerative neurological condition. Early diagnosis is essential for managing it and enhancing quality of life. Conventional diagnostic methods, however, typically rely on subjective assessments and physical observations, which can be time-consuming to establish. Automated, non-invasive, and effective diagnostic techniques might become more feasible as ML advances and biological voice data becomes more accessible. The project is driven by the need to use these technologies to investigate voice biomarkers for Parkinson's disease (PD) in order to allow early identification, which often fluctuates throughout the disease's early stages. The study uses modern ML algorithms, including XGBoost and LR on voice-based features to improve diagnostic accuracy, do away with manual analysis, and contribute to the development of reliable, data-driven healthcare proposals.

The study's primary contributions are as follows:

- Utilized a Parkinson's Disease dataset from **Kaggle**, enhancing the practical relevance and applicability of the findings.
- Implemented a robust preprocessing pipeline including data cleaning, outlier detection, and standardization of continuous variables to improve data quality and model performance.
- Implemented **XGBoost** and **LR** classifiers to determine the most effective model for diagnosis.
- In order to manage and treat Parkinson's disease early, the proposed study employs machine learning to diagnose the condition.
- Measured the performance of evaluated models with standard classification metrics, Precision, Recall, Accuracy and F1-Score to guarantee robustness and reliability.

Novelty and Justification of the Study

The proposed study is novel because it uses a holistic method of detecting PD based on voice attributes by using ensemble learning (XGBoost) and classical statistical analysis LR benchmarks to measure overall performance. Compared to the previous works where a single model or a small number of features can be used, this study employs a wide variety of vocal biomarkers producing delicate patterns related to PD based on biomedical voice measurements. The use of advanced preprocessing, feature selection, and cross-validation techniques ensures robust model training and generalization. The justification for this study stems from the urgent need for accurate, non-invasive, and early-stage diagnostic tools, as current clinical assessments are prone to delays and subjectivity. By comparing and validating multiple ML models, this study offers important new information on the predictive power of voice characteristics, supporting the development of scalable, real-time diagnostic applications in clinical settings.

Structure of the Paper

The following is the structure of the paper: Section II examines pertinent studies on Parkinson's disease early diagnosis, Section III describes the technique, Section IV displays the findings and model comparisons, and Section V offers conclusions and suggestions for further study.

LITERATURE REVIEW

The material currently available on the early diagnosis of PD is reviewed in this section. The majority of studies emphasize the use of diverse algorithms to enhance the efficiency of task scheduling in cloud environments. Common themes emerging from the reviewed literature include:

Jain and Srivastava (2025) proposed neurological disorders, the use of MRI and CT images as input data in DL models is becoming increasingly widespread. In this study, MRI images from the "Alzheimer Parkinson 3 Class Data Set" available on the Kaggle platform were used for the diagnosis of Alzheimer's and PD. The dataset includes three classes: 2,561 Alzheimer's, 906 Parkinson's, and 3,010 Control (Normal) images. In this work, the Alzheimer, Parkinson, and Normal classes were trained using ResNet-18, VGG-16, and ConvNext architectures, yielding accuracy rates of 96.2%, 95.4%, and 98.9%, respectively. Additionally, Alzheimer and Parkinson diseases were tested against the normal class using binary classifiers. For the Alzheimer-Normal and Parkinson-Normal classes, the models achieved the following results: ResNet-18 with accuracy rates of 82.0% and 96.1%, VGG-16 with 95.4% and 89.4%, and ConvNext with 99.4% and 99.5%, respectively [12].

Nawal, Habib and Barua (2025), stated an approach combining Histogram of Oriented Gradients (HOG) with It is suggested to use a customized CNN for early Parkinson's disease diagnosis. Preprocessing methods were used to improve the consistency and quality of a medical image collection. The CNN extracts key features while HOG provides edge orientation information, and their fusion creates a robust feature map. An integrated attention mechanism further refines focus on crucial regions. Evaluation demonstrates a balanced performance in terms of accuracy (99%) and parameter (0.8M) requirement. Visualization tools like Grad CAM offer insights into model decisions, aiding interpretability. This approach offers an accurate PD detection, potentially transforming diagnosis and improving patient outcomes [13].

Mehta and Khurana (2024) aimed to determine whether DBNs are suitable for detecting PD early since they can assess complicated and high-dimensional medical information. During the DBN modelling, the data used were trained and tested using publicly available datasets, and the accuracy level recorded was 92%. In comparison, the sensitivity was 90%, and the specificity was 94%. The ROC-AVC of the timing of task execution was calculated to be 95% in the diagnostic capacity, which indeed indicates the high level had been maintained. According to the above results, the DBN model provided superior performance to other diagnostic methods, which include a low FNR. Traditional techniques, where the diagnosis depends on a doctor's assignment and imaging techniques, are usually less accurate and take more time to detect diseases early [14].

Vats and Mehta (2024) suggested deploying a deep belief network (DBN) method, considered a highly advanced ML algorithm, which is more of a memory structure capable of DL and hierarchically. Their study implied the use of a DBM model for a diverse data set of 500 PD subjects suspected to have disease in its early stage. The dataset contains medical records, speaking analysis, audio recordings of subjects, and biometric monitoring. The model was trained using a two-phase training approach. The first phase is an unsupervised pre-training process to learn general characteristics. The DBN model's accuracy of 93%, sensitivity of 90%, specificity of 93%, and AVC of 0. 7. 97 were all extremely positive outcomes. With an accuracy of 85%, sensitivity of 80%, specificity of 85%, and AVC of 0.85%, these measurements perform better than standard diagnostic techniques [15].

Tesfai (2023) focused on the development of a speech and audio-based ML pipeline for PD diagnosis. Two voice recording datasets are assembled using data augmentation techniques. Paired with traditional ML models, acoustic features yield 99.21% accuracy, while Log-Mel spectrograms with CNN's achieve 99.71% accuracy. The highest accuracy of 99.82% is attained through an ensemble model that combines both spectrogram and acoustic models. These outcomes provide compelling evidence for the effectiveness of multimodal ensemble models in PD diagnosis, offering promising prospects for non-invasive early detection [16].

Lyu and Guo (2023) Brain Graph Convolutional Networks (BGCN) is a unified framework designed to integrate brain functional connectivity based on the non-Euclidean heuristic into a DL model (GCN) based on graphs for diagnosing Parkinson's illness. In order to preserve the spatial dependency between the EEG channels and make it easier to formulate the functional connectivity building issue, the graph format of EEG data is used. It used the GCN to simulate the flow of brain information between nodes using convolutions along functional connectivity. Functional connection was achieved in this study by using a heuristic search technique to solve an MST (Minimum Spanning Tree) issue. The resulting functional connectivity in terms of the afflicted areas and hub shift was in line with previous MRI investigations. The effectiveness of the suggested framework was assessed by contrasting random/uniform connectivity produced by k-NN with the heuristic functional connectivity speculation. Both learning robustness and accuracy (95.59%) have been attained by the suggested system [17].

Chang et al. (2022) proposed that bradykinesia, rest tremor, and stiffness are the three primary motor symptoms of PD. Among neurodegenerative movement disorders, PD is the most common. Using a high-speed camera system, the accuracy of a novel algorithm approach created to recognize each motor evaluation on the Unified Parkinson's Disease Rating Scale (UPDRS) has been confirmed. The three categories of detection parameters that comprise this system are the angle, time-frequency, and trajectory parameters. With IMU, the average detection accuracy is 87%, 90%, and 95%, respectively. There are some disparities in the movement characteristics between the 17 patients and the 20-year-old youth controls, according to the results of the trial tests. The typical control rotation speed for 3.6 pronation and supination can be double that of the patient, and A typical control's amplitude deviation is 5 degrees, whereas the patients can exceed 45 degrees [18].

A comparative analysis of the background study, based on its methodology, Dataset/Environment, Problem Addressed, Performance, and Future Work/Limitations, is provided in Table I.

TABLE I. REVIEW OF LITERATURE ON EARLY DIAGNOSIS OF PARKINSON'S DISEASE

Author	Methodology	Environment	Problem Addressed	Performance	Future Work / Limitation
Jain and Srivastava (2025)	MRI image deep learning with ResNet-18, VGG-16, and ConvNext	Alzheimer Parkinson 3 Class Data Set" (Kaggle)	MRI imaging for Alzheimer's and Parkinson's disease diagnosis	Multi-class: 96.2% (ResNet-18), 95.4% (VGG-16), 98.9% (ConvNext); Binary: Up to 99.5% (ConvNext)	Focused on classification; could explore lightweight models for real-time or mobile deployment
Nawal, Habib, and Barua (2025)	HOG + custom CNN with attention mechanism; Grad CAM visualization	Curated medical image dataset	Early Parkinson's detection through hybrid feature learning	Accuracy: 99%, Parameters: 0.8M	Limited details on dataset diversity and generalizability; clinical validation needed
Mehta and Khurana (2024)	Deep Belief Network (DBN) on public PD datasets	Public datasets	High-dimensional medical data analysis for early PD detection	Accuracy: 92%, Sensitivity: 90%, Specificity: 94%, ROC-AUC: 95%	Lacks multimodal data usage; focused only on DBN architecture
Vats and Mehta (2024)	DBN with unsupervised pretraining on multimodal data (voice, biometric, medical)	Diverse dataset with 500 PD subjects	Early-stage PD detection with various physiological and biometric indicators	Accuracy: 93%, Sensitivity: 90%, Specificity: 93%, AUC: 97%	AVC reported inconsistently; real-world deployment readiness not assessed
Tesfai (2023)	Traditional ML with acoustic features and CNNs with Log-Mel spectrograms; ensemble model	Speech and audio datasets + data augmentation	PD diagnosis through non-invasive speech signals	ML: 99.21%, CNN: 99.71%, Ensemble: 99.82%	Real-time application and language/accent variation unaddressed
Lyu and Guo (2023)	Brain Graph Convolutional Networks (BGCN) using EEG functional connectivity via MST heuristic	EEG data + graph-based deep learning	EEG-based PD diagnosis preserving spatial interdependence	Precision: 95.59%, Robust learning performance	Heuristic connectivity may vary across individuals; needs clinical validation and real-time efficiency review
Chang et al. (2022)	Wearable IMU system with UPDRS motor exam analysis (trajectory, time-frequency, angle)	IMU + CMOS chip + high-speed camera validation	Objective quantification of PD motor symptoms	Accuracy: 87%-95% depending on metric; Power: 0.3713mW; Area: 4.2mm × 4.2mm	Small subject pool (17 patients); generalization and long-term use unassessed

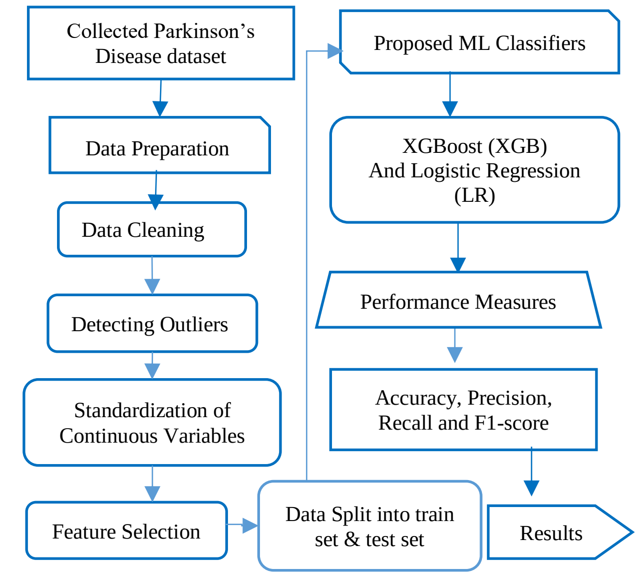

Fig. 1. Flowchart of Early Diagnosis of Parkinson's Disease

METHODOLOGY

The symptoms of PD, a complex, progressive neurological disease that causes tremor, rigidity, and bradykinesia. As the illness progresses, some people may have postural instability. This section illustrates how to use ML to make an early diagnosis of PD. The PD dataset is gathered from Kaggle to start the procedure. The second step also includes data preparation extensively (data cleaning, the identification of outliers and the normalization of continuous variables). This is followed by the feature selection process so as to keep the most pertinent attributes of classification. From this cleaner dataset, the training and testing datasets are further segregated. LR and XGBoost (XGB), two ML classifiers, are used to build predictive models. These classifiers' performance is commonly assessed using metrics like F1-score, recall, accuracy, and precision. The models' ability to diagnose PD is then determined by looking at the evaluation results in Figure 1.

Each step of the flowchart is explained in the section below:

Data Collection

In this study, the Parkinson's Disease dataset, which was acquired via Kaggle, was used. There are 31 people in this collection, 23 of whom have Parkinson's disease, and a variety of biological voice metrics are included. The index is the "name" column in the database, and each row corresponds to a voice measure, and each column to one of the 195 voice recordings of these people. The "status" column is set to 0 for healthy and 1 for PD in order to distinguish between those with PD and those in excellent health. This is the primary goal of the information. Some EDA graphs are given in this section below:

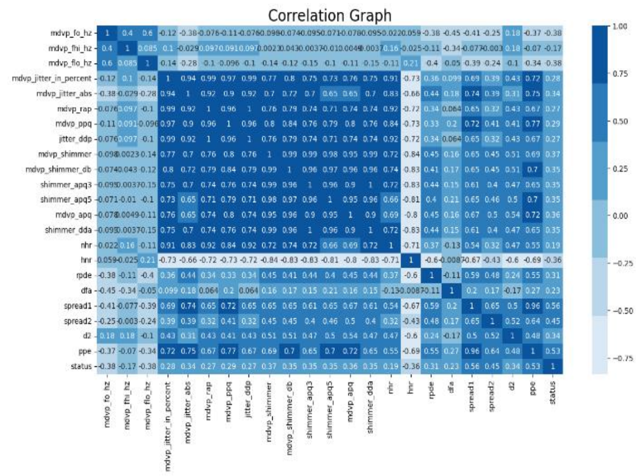

Fig. 2. Correlation between features

Figure 2 visualizes the pairwise relationships between features in the dataset used to identify Parkinson's illness. The Pearson correlation coefficient between two attributes is shown in each cell of the heatmap; Perfect negative correlation (value -1) and perfect positive correlation (value +1) are the two extremes. Lighter blue hues and values close to 0 signify weak or nonexistent linear associations, whereas darker blue hues suggest higher positive correlations. The status variable, representing the disease state, shows moderate correlation with certain acoustic features, indicating their predictive relevance. This heatmap aids in identifying multicollinearity, guiding dimensionality reduction and feature selection strategies in the model development process.

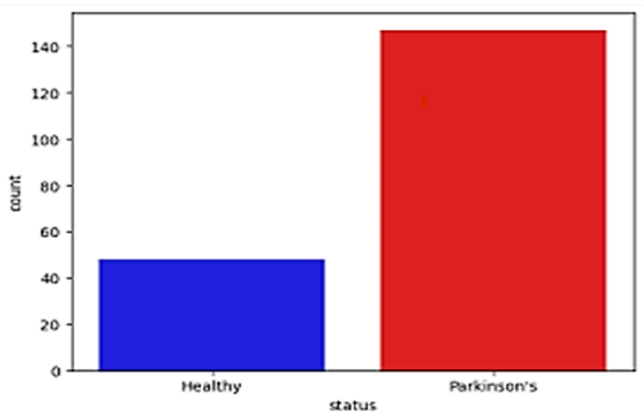

Fig. 3. Plot between Healthy & Parkinson's from the dataset

Figure 3 displays the distribution of individuals based on their health status, categorized into Healthy and Parkinson's. The y-axis shows the overall number of people, while the x-axis shows the current situation. With a noticeably higher percentage of individuals with PD than healthy individuals, the graph clearly illustrates the dataset's imbalance. Specifically, there are approximately 50 healthy individuals (represented by the blue bar) and around 145 individuals with Parkinson's (represented by the red bar), indicating that the dataset is imbalanced towards the Parkinson's class.

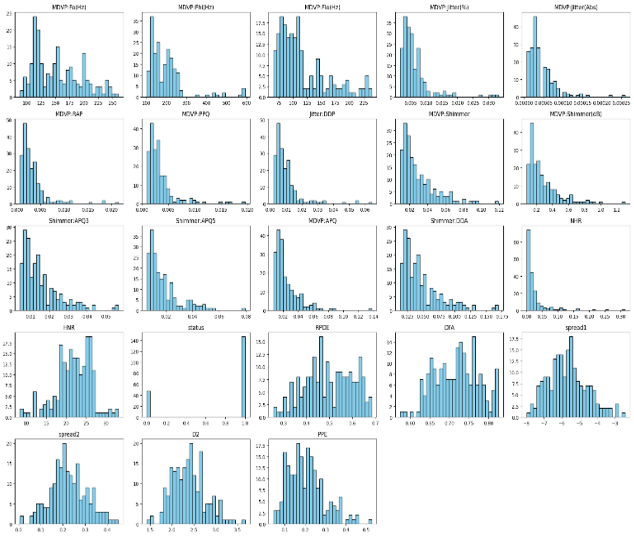

Fig. 4. Analyzing the Data Attributes

Figure 4 displays a grid of 23 histograms, each representing the distribution of a different feature. All histograms are blue on a white background, consistent with a standard plotting style, and appear to have similar scales on their y-axes (representing frequency or count), though the x-axis scales vary for each feature. Many histograms frequently exhibit a skewed distribution that extends towards higher values with a lengthy tail and a high frequency of values concentrated at the lower end. This indicates that most features are not normally distributed but rather exhibit a positive skew, meaning there are more instances of lower values and fewer instances of higher values.

Data Preprocessing

This part used a variety of preprocessing techniques to improve the data's quality while keeping the original characteristics for further examination. The preprocessing involves data cleaning, outlier detection, standardization of continuous variables, and feature selection techniques, which are discussed below:

1) Data Cleaning

Single-value and missing-value columns were eliminated prior to preprocessing and analysis [19]. In order to provide more dependable and significant findings, effective data cleansing makes that the information is trustworthy, consistent, and suitable for ML or analysis.

2) Detecting Outliers

The mode, median, and mean are all at the same location, indicating that the data are symmetrical [20]. A longer or fatter tail distribution to the right indicates positive skewness in the data, meaning that the mode is lower than the mean and median.

3) Standardization of Continuous Variables

The standardization approach was used to make sure that all of the data had a uniform format because the dataset derived from the earlier phases included continuous variables [21]. The dataset was standardized using Equation (1), where the mean of each characteristic is taken out of split by its value and the data's standard deviation.

$$\text{stand} = \frac{\text{x} - \text{mean}}{\text{Std Dev}} \quad (1)$$

4) Feature Selection

A crucial step before using classification algorithms is feature selection, which lowers the algorithms' complexity and computation time while also improving overall classification performance [22]. The following describes the feature selection: The aim of feature selection is to find the optimal subset Q', where $Q' \subset Q$ and has a size of n^', where $(n' < n)$, such that the following equivalence is assured in eq. Given an evaluation function E_{val} and a feature set $Q = q_1, q_2, \dots, q_n$ of size n, where n is the total number of Equation (2):

$$\mathrm{E_{val}}(\mathrm{Q'}) = \underset{\mathrm{M} \subset \mathrm{Q}}{\operatorname{argmin}} \mathrm{E_{val}}(\mathrm{M}) \quad (2)$$

In this case, $\left||\mathrm{M}|\right| = n'$, where n' is a user-defined number or dictated by the selection criteria.

Data Splitting

The data splitting, which comprises separating the dataset into subsets for testing and training, typically 30% for testing and 70% for training, is a crucial step in DL.

Proposed Machine Learning (ML) Classifiers

The ML models are described in this section:

5) XGBoost Classifier

XGBoost is a classifier that uses the gradient boosting (GB) technique, which is based on DT. Its speed, effectiveness, and scalability have led to its usage [23]. The following is a general explanation of GB and XGBoost. Using D=[x,y] to characterize a dataset with n observations, where x is the feature (an independent variable) and y is the dependent variable [24].

The scores for each leaf may then be added together to determine the final forecast for a specific sample x_i, as shown in Equation (3).

$$\hat{y}_i = \sum_{b=1}^{B} f_b(x_i) \quad (3)$$

A tree construction q is indicated by f_b , and leaf j has a weight score w_j. If boosting is k in GB, use a B function to anticipate the outcome using $\hat{y}_i$ as the prediction for the i-th sample at the b-th boost.

6) Logistic Regression

The majority of early 20th-century biological research and applications employed logistic regression. When dealing with categorical target variables, one of the most used ML techniques is logistic regression (LR). Lately, LR has gained popularity as a technique for binary classification issues [25][26]. Additionally, a discrete binary product between 0 and 1 is shown. Using the underlying logistic function, logistic regression evaluates probabilities (p) to calculate the connection between the feature variables [27]. In the initial phase of the analysis, LR, a widely-used method for predictive analytics and classification tasks, was applied to transform, which is the probability of success divided by the probability of failure, the logit formula was employed as shown in Equation (4):

$$Logit(p) = \frac{1}{1+\exp(-p)} \quad (4)$$

The function of $Logit(p)$ in LR is to transform the odds of success to a linear scale, facilitating binary classification by modeling the probability of the outcome as given in Equation (5):

$$\ln\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k \quad (5)$$

where $X_1 \dots X_k$ are predictor variables, p is the probability of an occurrence, and $\beta_0, \beta_1, \dots, \beta_k$ are coefficients that determine each predictor variable's proportional relevance.

Performance Matrix

The suggested model's performance was evaluated using the four commonly used evaluation metrics of recall, accuracy, precision, and F1-score. The predictive ability of the model was demonstrated by comparing its predictions with the test dataset's actual class labels using a confusion matrix. This matrix summarizes the right and wrong classifications in a simple way, giving you a better idea of how well the model worked. It also serves as a basis for calculating key performance indicators that reflect the model's classification effectiveness. The confusion matrix's essential elements include:

- **True Positives (TP):** The proportion of PD patients who the algorithm correctly forecasts will have the condition.
- **False Positives (FP):** The quantity of instances in which a patient is misdiagnosed with PD by the model when they do not.
- **True Negatives (TN):** The frequency with which the model accurately predicts that a patient is actually healthy and does not have PD.
- **False Negatives (FN):** The frequency with which the model predicts a patient to be healthy while in fact they have PD.

7) Accuracy

Evaluates the overall diagnostic precision of the model for both Parkinson's disease patients and those without the condition. The accuracy is calculated for the overall model using Equation (6):

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})} \quad (6)$$

8) Precision

Evaluates the model's capacity to identify Parkinson's disease in authentic situations. High recall is crucial for early diagnosis to avoid missed cases. The precision is calculated in Equation (7):

$$\text{Precision} = \frac{\text{TP}}{(\text{TP}+\text{FP})} \quad (7)$$

9) Recall

The percentage of TP evaluations that the model accurately detects. An elevated recall signifies that the model can detect the vast majority of TP emotions. The recall is mathematically depicted in Equation (8):

$$\text{Recall} = \frac{\text{TP}}{(\text{TP}+\text{FN})} \quad (8)$$

10) F1-Score

A single performance metric that balances the importance of both detecting true Parkinson's cases and avoiding FP. In situations when there is an unequal distribution of classes, it is invaluable. The F1-score is formulated in Equation (9):

$$\text{F}_1 - \text{Score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad (9)$$

11) ROC-AUC

The classification problem's performance is measured using the ROC curve. The x-axis displays the FPR, while the y-axis displays the TPR. The AUC and ROC, is a separability statistic that shows how well a model can differentiate between classes. The model predicts classes more correctly when the AUC is larger.

Results And Discussion

The system used for this study is equipped with a 6th Generation Intel Core i5 processor, supported by 12 GB of RAM to ensure smooth multitasking and efficient data handling. It also has a dedicated 4 GB GPU to make computations faster especially those related to ML. The ML models for Parkinson's disease prediction are compared in Table II according to important performance characteristics such as F1-score, recall, accuracy, and precision. The XGBoost model's maximum accuracy of 97.4%, precision of 99.9%, recall of 96.6%, and F1-score of 98.3% show how effectively the model can distinguish between favorable and unfavorable situations. Comparatively, the LR model performs a little bit lower with the accuracy standing at 92.3%, precision at 93.5%, recall at 96.6% and an accuracy of 95.0%. Such findings underscore the efficiency and stability of XGBoost in predicting the presence of PD in its initial stages, which were better than LR with regard to all the measured variables.

TABLE II. Evaluation of Machine Learning Models on Early Diagnosis of Parkinson's Disease

Model	Accuracy	Precision	Recall	F1-Score
XGBoost	97.4	99.9	96.6	98.3
LR	92.3	93.5	96.6	95.0

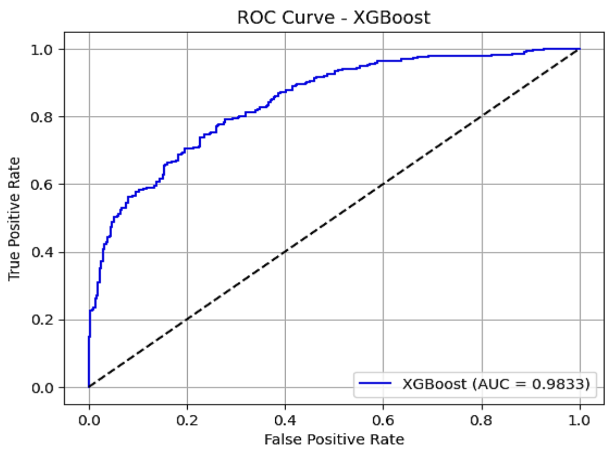

Fig. 5. ROC Graph of XGBoost Model

The ROC curve for the XGBoost model is displayed in Figure 5. There are problems with both the FPR on the x-axis and the TPR on the y-axis. The blue solid line displays the XGBoost model's performance, whereas the black dashed line indicates a random classifier (where AUC = 0.5). As seen in the legend, the XGBoost model's curve performs admirably, maintaining a significant margin above the random classifier line and attaining an AUC score of 0.9833. The XGBoost model appears to have outstanding discriminating power, successfully differentiating between positive and negative classes, based on its high AUC value.

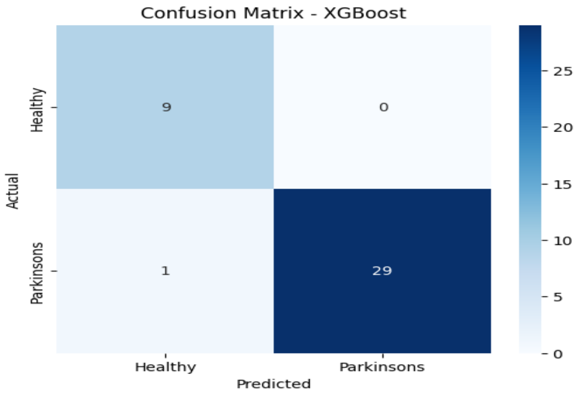

Fig. 6. Confusion Matrix of the XGBoost Model

The XGBoost model's confusion matrix is shown in Figure 6, showing strong classification performance. All 9 healthy individuals were correctly identified TN, with no false positives. Among Parkinson's cases, 29 were correctly classified TP, and only 1 was misclassified FN. The darker blue shades emphasize the high number of accurate predictions.

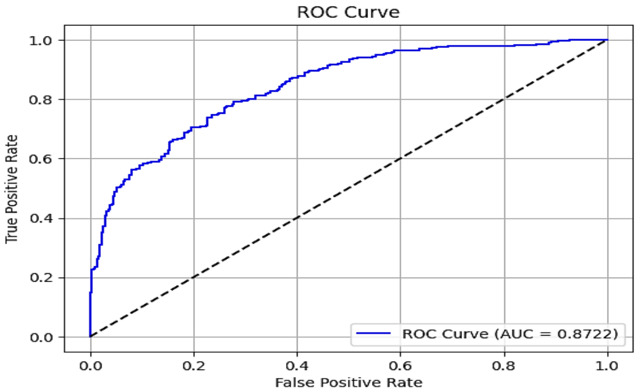

Fig. 7. ROC Graph of the LR Model

An LR, ROC curve, which shows how effectively a binary classifier system can identify issues when its discriminating threshold is altered, is shown in Figure 7. The TPR (sensitivity) is shown on the y-axis, while the FPR (specificity) is shown on the x-axis. The blue solid line shows the model's ROC curve, while a random classifier is shown by the black dashed line. According to the description, this ROC curve's AUC is 0.8722.

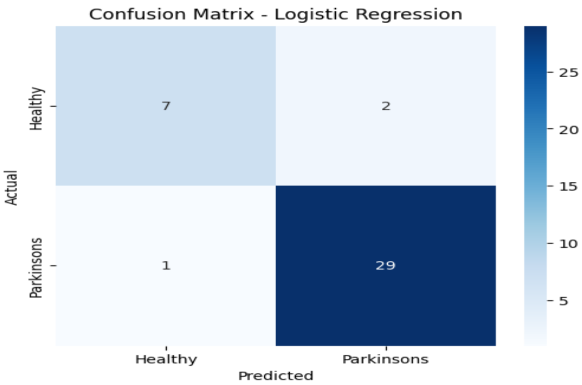

Fig. 8. Confusion Matrix of the Logistic Regression (LR) Model

In Figure 8, the classification performance of an LR model is shown graphically as a confusion matrix. The matrix, labeled "Confusion Matrix - LR," has 'Actual' classes (Healthy and Parkinsons) on the y-axis and 'Predicted' classes (Healthy and Parkinsons) on the x-axis. According to the matrix, 7 individuals who were actually 'Healthy' were correctly predicted as 'Healthy' (true negatives). The false positive results were 2 cases of 'Healthy' being erroneously classified as 'Parkinsons'. Among people with real cases of having 'Parkinsons' 1 was falsely classified as being healthy FN and 29 as being 'Parkinsons' TP.

Comparative Analysis

In this section, a comparative statement is provided to compare the proposed XGBoost and Logistic Regression (LR) models with the current ML techniques, Decision Tree (DT), and Support Vector Machine (SVM). Table III shows that the XGBoost model has the highest accuracy of 97.4%, indicating that it has great predictive ability. Another model, LR, works quite well and achieves 92.3% accuracy, and the last model is Random Forest (RF)[28] with an accuracy of 91.01%. Bagging [29] produces a moderately high accuracy of 88.2 per cent whereas SVM [30] and DT [31] have lesser accuracies of 76.32 per cent and 60.7 per cent respectively. Such results highlight the high level of precision and diagnostic efficiency of the suggested XGBoost model in comparison with both standard and ensemble-based methods of ML.

TABLE III. COMPARATIVE ANALYSIS OF ML MODELS ON EARLY DIAGNOSIS PARKINSON'S DISEASE

Model	Accuracy	Precision	Recall	F1-Score
Bagging	88.2	67.6	94.8	-
SVM	76.32	86.0	81.0	84.0
DT	60.7	58.4	60.7	59.5
RF	91.01	89.25	93.26	91.21
XGBoost	97.4	99.9	96.6	98.3
LR	92.3	93.5	96.6	95.0

The suggested XGBoost and LR models are excellent for early Parkinson disease detection because they are reliable, generalizable, and able to handle intricate data patterns. The XGBoost is a successful ensemble learning technique, is a reliable algorithm because it is useful in characterizing non-linear feature relationships and interactions, and thus it should be useful in biomedical tasks of classification. The fact that it has internal regularization and optimization helps in increasing model stability and minimizing overfitting. On one hand, LR is praised due to its simplicity, interpretability, and effectiveness of processing linearly separable data which is quite crucial in the medical diagnosis when transparency and explain ability are vital. These models, when combined, outperform traditional ML techniques in several ways: they produce more accurate and reliable predictions, offer better classification, and can identify individuals in good health and those with Parkinson's disease, enabling more effective early intervention and treatment planning.

CONCLUSION AND FUTURE SCOPE

The neurological degenerative disorder known as Parkinson disease (PD) can cause both motor and non-motor symptoms. Nonmotor symptoms include sleep difficulties, depression, and irregularities in cognition, whereas motor symptoms including bradykinesia, tremors, and stiffness have been linked to striatal dopamine deficit. There are currently no reliable tests to identify PD, however, identifying illnesses that have characteristics with the Parkinson's syndrome is a crucial first step in the diagnosing process. Finally, a novel and effective method for early Parkinson's disease detection may be the suggested strategy, which combines NLP with ML methods like XGBoost and LR. Regarding precision, accuracy, recall, and F1-score, the suggested paradigm shows encouraging results for practical clinical use. Better patient outcomes and early intervention can be facilitated by this automated and scalable method.

For future scope, the model can be extended by incorporating DL techniques, larger and more diverse datasets, and multilingual clinical records. Additionally, expanding the pipeline to detect other neurological disorders or integrating it into a real-time diagnostic support tool could further enhance its utility and impact in the medical field. "

REFERENCES

[1] E. Paccosi and L. Proietti-De-Santis, "Parkinson's Disease: From Genetics and Epigenetics to Treatment, a miRNA-Based Strategy," *Int. J. Mol. Sci.*, vol. 24, no. 11, May 2023, doi: 10.3390/ijms24119547.

[2] P. P. Jalaja, D. Kommineni, A. Mishra, R. Tumati, C. A. Joseph, and R. V. S. S. B. Rupavath, "Predictors of Mortality in Acute Myocardial Infarction: Insights From the Healthcare Cost and Utilization Project (HCUP) Nationwide Readmission Database," *Cureus*, vol. 17, no. 5, May 2025, doi: 10.7759/cureus.83675.

[3] N. Patel, "Quantum Cryptography In Healthcare Information Systems: Enhancing Security In Medical Data Storage And Communication," *J. Emerg. Technol. Innov. Res.*, vol. 9, no. 8, pp. 193–202, 2022.

[4] P. Jiang, N. Gao, G. Chang, and Y. Wu, "Biosensors for Early Detection of Parkinson's Disease: Principles, Applications, and Future Prospects," *Biosensors*, vol. 15, no. 5, Apr. 2025, doi: 10.3390/bios15050280.

[5] S. Pahune and N. Rewatkar, "Large Language Models and Generative AI ' s Expanding Role in Healthcare Large Language Models and Generative AI ' s Expanding Role in Healthcare," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 8, pp. 2288–2302, 2024, doi: 10.13140/RG.2.2.20109.72168.

[6] A. Mobed, S. Razavi, A. Ahmadalipour, S. K. Shakouri, and G. Koohkan, "Biosensors in Parkinson's disease," *Clin. Chim. Acta*, vol. 518, pp. 51–58, Jul. 2021, doi: 10.1016/j.cca.2021.03.009.

[7] S. Singamsetty, "Neurofusion: Advancing Alzheimer's Diagnosis with Deep Learning and Multimodal Feature Integration," *Int. J. Educ. Appl. Sci. Res.*, vol. 8, no. 1, pp. 23–32, 2021, doi: 10.5281/14889013.

[8] S. Chakraborty, S. Aich, and H. C. Kim, "Detection of Parkinson's Disease from 3T T1 Weighted MRI Scans Using 3D Convolutional Neural Network," *Diagnostics*, vol. 10, no. 6, Jun. 2020, doi: 10.3390/diagnostics10060402.

[9] M. A. Mostafiz, "Machine Learning for Early Cancer Detection and Classification: AI- Based Medical Imaging Analysis in Healthcare," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 3, pp. 251–260, 2025.

[10] B. Majhi *et al.*, "An improved method for diagnosis of Parkinson's disease using deep learning models enhanced with metaheuristic algorithm," *BMC Med. Imaging*, vol. 24, no. 1, Jun. 2024, doi: 10.1186/s12880-024-01335-z.

[11] S. S. S. Neeli, "Heart Disease Prediction For A Cloud-Based Smart Healthcare Monitoring System Using Gans And Ant Colony Optimization," *Int. J. Med. Public Heal.*, vol. 14, no. 4, pp. 1219–1229, 2024, doi: 10.70034/ijmedph.2024.4.223.

[12] S. Jain and R. Srivastava, "Neurodegenerative Disease Alzheimer's and Parkinson's Classification with Deep Learning," in *2025 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, IEEE, Apr. 2025, pp. 798–803. doi: 10.1109/InCACCT65424.2025.11011366.

[13] R. Nawal, N. Habib, and S. Barua, "A Deep Learning Based Non-Invasive Framework for Neurodegenerative Parkinson's Disease Diagnosis from Template-Less Handwritten Shapes," in *2025 4th International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, IEEE, Jan. 2025, pp. 65–70. doi: 10.1109/ICREST63960.2025.10914445.

[14] S. Mehta and S. Khurana, "From Data to Diagnosis: Utilizing Deep Belief Networks for Early Parkinson's Disease Detection," in *2024 International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit)*, IEEE, Sep. 2024, pp. 347–351. doi: 10.1109/GlobalAISummit62156.2024.10947917.

[15] S. Vats and S. Mehta, "The Impact of Deep Belief Networks on the Early Diagnosis of Parkinson's Disease," in *2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, Aug. 2024, pp. 980–984. doi: 10.1109/ICESC60852.2024.10690058.

[16] S. Tesfai, "Multimodal Ensemble Models for Parkinson's Disease Diagnosis Using Log-Mel Spectrograms and Acoustic Features," in *2023 IEEE MIT Undergraduate Research Technology Conference (URTC)*, IEEE, Oct. 2023, pp. 1–5. doi: 10.1109/URTC60662.2023.10534982.

[17] T. Lyu and H. Guo, "BGCN: An EEG-based Graphical Classification Method for Parkinson's Disease Diagnosis with Heuristic Functional Connectivity Speculation," in *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)*, IEEE, Apr. 2023, pp. 1–4. doi: 10.1109/NER52421.2023.10123796.

[18] W. H. Chang, K. Du Liou, Y. T. Liu, and K. A. Wen, "Precise Motor Function Monitor for Parkinson Disease using Low Power and Wearable IMU Body Area Network," in *2022 14th Biomedical Engineering International Conference (BMEiCON)*, IEEE, Nov. 2022, pp. 1–5. doi: 10.1109/BMEiCON56653.2022.10012114.

[19] S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani, and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," in *2024 9th International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Dec. 2024, pp. 1149–1154. doi: 10.1109/ICCES63552.2024.10859381.

[20] J. Wang, Z. Zhou, Z. Li, and S. Du, "A Novel Fault Detection Scheme Based on Mutual k-Nearest Neighbor Method: Application on the Industrial Processes with Outliers," *Processes*, vol. 10, no. 3, Mar. 2022, doi: 10.3390/pr10030497.

[21] K. H. Chang, C. H. Wang, B. G. Hsu, and J. P. Tsai, "Serum Osteopontin Level Is Positively Associated with Aortic Stiffness in Patients with Peritoneal Dialysis," *Life*, vol. 12, no. 3, Mar. 2022, doi: 10.3390/life12030397.

[22] N. Khoury, F. Attal, Y. Amirat, L. Oukhellou, and S. Mohammed, "Data-Driven Based Approach to Aid Parkinson's Disease Diagnosis," *Sensors*, vol. 19, no. 2, Jan. 2019, doi: 10.3390/s19020242.

[23] R. P. Mahajan, "Transfer Learning for MRI image reconstruction: Enhancing model performance with pretrained networks," *Int. J. Sci. Res. Arch.*, vol. 15, no. 1, pp. 298–309, Apr. 2025, doi: 10.30574/ijsra.2025.15.1.0939.

[24] G. Abdurrahman and M. Sintawati, "Implementation of XGboost for Classification of Parkinson's Disease," *J. Phys. Conf. Ser.*, vol. 1538, no. 1, pp. 1–8, May 2020, doi: 10.1088/1742-6596/1538/1/012024.

[25] C. K. Gomathy, "The Parkinson ' S Disease Detection Using Machine Learning," *Int. Res. J. Eng. Technol.*, vol. 8, no. 10, pp. 440–444, 2021.

[26] R. Dattangire, D. Biradar, R. Vaidya, and A. Joon, "A Comprehensive Analysis of Cholera Disease Prediction Using Machine Learning," in *Congress on Intelligent Systems*, 2025, pp. 555–568. doi: 10.1007/978-981-96-2697-7_39.

[27] S. B. Shah, "Artificial Intelligence (AI) for Brain Tumor Detection: Automating MRI Image Analysis for Enhanced Accuracy," *Int. J. Curr. Eng. Technol.*, vol. 14, no. 06, pp. 320–327, Dec. 2024, doi: 10.14741/ijcet/v.14.5.5.

[28] X. Yang, Q. Ye, G. Cai, Y. Wang, and G. Cai, "PD-ResNet for Classification of Parkinson's Disease From Gait," *IEEE J. Transl. Eng. Heal. Med.*, vol. 10, pp. 1–11, 2022, doi: 10.1109/JTEHM.2022.3180933.

[29] O. Barukab, A. Ahmad, T. Khan, and M. R. T. Kunhumuhammed, "Analysis of Parkinson's Disease Using an Imbalanced-Speech Dataset by Employing Decision Tree Ensemble Methods," *Diagnostics*, vol. 12, no. 12, Nov. 2022, doi: 10.3390/diagnostics12123000.

[30] M. A. Hossain and F. Amenta, "Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech

Biomarkers," *J. Park. Dis.*, vol. 14, no. 1, pp. 95–109, Jan. 2024, doi: 10.3233/JPD-230002.

[31] H. Ali, E. Hashmi, S. Y. Yildirim, and S. Shaikh, "Analyzing Amazon Products Sentiment: A Comparative Study of Machine and Deep Learning, and Transformer-Based Techniques," *Electronics*, vol. 13, no. 7, Mar. 2024, doi: 10.3390/electronics13071305.