International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Advancing Voice-based Parkinson's Detection with Hybrid XGBoost-LightGBM Models

Rajnish K Ranjan[1], Anupam Agrawal[2], Divyarth Rai[3]

[1,3]LNCT University
[2]IIIT Prayagraj

Abstract

Parkinson's disease (PD) affects millions of individuals across the world and is characterized by gradual degeneration of the nervous system, and leading to both motor and non-motor dysfunctions. Early and reliable identification of the disease is important, as it directly influences therapeutic strategies and long-term patient outcomes. This research work presents a hybrid stacking ensemble technique that integrates Extreme Gradient Boosting (XGBoost) with Light Gradient Boosting Machine (LightGBM) to improve the predictive accuracy of PD diagnosis and disease progression assessment by voice analysis. By leveraging the complementary learning capabilities of both algorithms, the proposed framework aims to deliver strong generalization and high-performance classification, particularly in the early stages of PD detection. The evaluation of proposed work is carried out using two publicly available UCI datasets: the Parkinson's Disease Detection dataset for classification and the Parkinson's Disease Telemonitoring dataset for regression-based prediction of UPDRS scores. On the detection dataset, the model achieved an accuracy of 96.7% with a perfect recall of 100%. On the Telemonitoring dataset, the framework produced an RMSE of 1.70 and an R^2 value of 0.973 for Total UPDRS, and an RMSE of 1.19 with an R^2 of 0.977 for Motor UPDRS. When the regression problem is reframed as binary classification, the model achieved a high accuracy range of 97-98%. Overall, the results show that ensemble-based learning techniques can significantly improve the reliability of predictive models and serve as effective decision support tools for the diagnosis and monitoring of Parkinson's disease.

Keywords: Parkinson's disease, XGBoost, LightGBM, Machine Learning, Dementia.

I. Introduction

Parkinson's Disease (PD) is the second most prevailing neuro-degenerative disorder, after Alzheimer's disease [1][2]. James Parkinson, who narrated PD in 1817, and was originally termed 'Shaking Palsy'. PD is a progressive neurological condition with motor deficits such as muscular rigidity, tremors, and bradykinesia, along with non-motor symptoms including, sleep disturbances, speech difficulties, depression, and cognitive decline.

According to the "World Health Organization (WHO)", it affects over 10 million people globally [3]. Parkinson's disease develops as the nerve cells in the substantia nigra slowly deteriorate and it leads to a significant drop in dopamine level- critical for motor regulation [4]. One of the major challenges in PD diagnosis is the delay in symptom onset, because symptoms appear only when about 60–80% of dopamine-producing neurons have already been destroyed or damaged. In addition, studies spot diagnostic inaccuracies, with up to a 25% mis-diagnosis rates in early stages, especially when symptoms overlap with

other neurological situations. PD primarily reduces quality of life and forces psychological, social, and economic load on caregivers and healthcare systems [5].

Traditionally, PD diagnosis depends on medical screening and rating scales such as the "Unified Parkinson's Disease Rating Scale (UPDRS)" and the Hoehn and Yahr 5-stage scaling system [6], along with neuroimaging techniques including MRI, CT, PET, and DaTSCAN. However, these imaging methods are overpriced and typically detect PD only when the disease has notably progressed. New lab-based tests, such as speech analysis, finger-touch task, and handwriting tests, have been developed to identify early symptoms [7]. While these methods provide semantic facts, these methods are useful only after significant neuronal damage has occurred. At present being, there is no cure for PD, highlighting the crucial need for accurate and reliable diagnostic tools for early detection and better management of the disease.

Among non-motor symptoms, voice related abnormalities come into sight in the early stage of PD and include monotonic speaking, dysarthria, hypophonia, and other articulation problems [8]. Solana-Lavalle et al. [9] noted that such symptoms appear in approximately 90% of PD individuals during early stages, making them valuable indicators for early diagnosis. Recent research has proven that machine learning (ML) and deep learning (DL) approaches hold on to great promise in medical diagnostics, including the detection of neurodegenerative diseases [10][11]. In contrast to conventional movement or imaging-based methods, voice-analysis offers a quick, patient-friendly, and affordable solution for early screening and tracking of PD progression [12]. Ensemble learning methods, which integrate the strengths of multiple classifiers, have shown better performance over individual models, making them promising tools for enhancing early and reliable PD detection.

Problem Statement

Diagnosis of Parkinson's disease from voice signals is not easy due to the small, noisy, and high-dimensional datasets, leading to overfitting and poor generalization. Existing single-model approaches struggle to capture complex non-linear feature interactions within the data, unable to solve regression and classification tasks simultaneously. Therefore, many existing models require heavy computation and offer limited interpretability, limiting their practical use in clinical deployment.

Objective

This work aims to develop a state-of-the-art hybrid stacking ensemble that integrates XGBoost and LightGBM to capitalize their complementary strengths for reliable PD detection. The framework supports both classification (PD Vs. Healthy) and regression (estimating disease severity) tasks, allowing precise diagnosis and continuous symptom monitoring. By incorporating feature importance analysis, the model ensures interpretability and offers meaningful insights while maintaining strong generalization on small, and noisy datasets.

Paper Organization

Section 1 presents the background of PD, the problem statement, and the objectives. Section 2 reviews the literature on PD detection. Section 3 provides the dataset description in brief. Section 4 details the proposed hybrid ensemble framework, and Section 5 covers the experimental results and discussion. Finally, Section 6 highlights the key conclusions and outlines potential directions for future research.

II. Literature Review

In their research, Govindu and Palwe [13] evaluated classical ML models (SVM, RF, KNN, LR) on the MDVP voice dataset and noted Random Forest as the most effective, achieving 91.83% accuracy and 0.95

sensitivity, supporting telemedicine based diagnosis. While Nahar et al. [14] spotlight feature selection (Boruta, RFE, RF) before classification, achieving 82.35% accuracy using Bagging with RFE and demonstrating that feature reduction can yield competitive performance with lower computational cost.
Kadhim et al. [15] further improved classification by applying Mahala Nobis distance-based feature selection, where KNN achieved 98.31% accuracy, and ensures the benefit of dimensionality reduction for PD diagnosis.
Lamba et al. [16] developed a hybrid speech based model for early Parkinson's disease (PD) detection, integrating feature selection like mutual information gain, genetic algorithm, extra tree, with classification algorithms Naive Bayes, KNN, Random Forest. Using UCI speech dataset and SMOTE to address class imbalance, the genetic algorithm and Random Forest achieved 95.58% accuracy.
Wang et al. [17] focused on early stage Parkinson's disease detection using multimodal premotor indicators from the PPMI dataset, including REM sleep behavior, cerebrospinal fluid biomarkers, olfactory loss, and dopaminergic imaging. Their deep learning (DL) model surpassed twelve machine learning (ML) and ensemble approaches, achieving an average accuracy of 96.45%, and highlighted clinical significance of integrating multiple biomarkers for robust early diagnosis. Whereas, Huda Jasim et al. [18] proposed PD severity prediction using speech recordings and compared with three deep learning architectures: LSTM, 1D CNN, and autoencoder. Reporting that LSTM achieved the highest accuracy of 98%. Together, these studies illustrates deep learning is effective not only for early PD detection but also for the disease staging, supporting more comprehensive patient monitoring and management.
Quan et al. [19] proposed an end to end deep learning (DL) model for Parkinson's disease (PD) detection from speech signals. Combining time distributed 2D-CNNs and 1D-CNNs to capture dynamic temporal features. The model achieved up to 92% accuracy across multiple speech tasks in Chinese and Spanish datasets, effectively identifying hypokinetic dysarthria and emphasizing the clinical importance of low-frequency Mel-spectrogram regions for PD detection.
Al-Tam et al. [20] proposed a machine-learning framework for Parkinson's disease diagnosis with individual and ensemble models (Random Forest, Decision Tree, SVM, Logistic Regression, Gradient Boosting, Stacking, Bagging) on two different benchmark datasets. The Stacking Ensemble are integrating SVM and Gradient Boosting for feature extraction with Logistic Regression for classification, and achieved the best performance with 94.87-96.18% accuracy and 90-96.27% AUC, demonstrating better diagnostic precision.
Omodunbi et al. [21] introduced a robust stacked ensemble learning model for Parkinson's disease prediction using voice-based biomarkers, combining SVM, RF, KNN, and DT as base classifiers with logistic regression as the meta classifier. They incorporated advanced feature selection methods like forward search, gain ratio, Kruskal-Wallis test, handled class imbalance with SMOTE, and used subject-wise validation to prevent data leakage. The model surpassed the performance of individual classifiers, emphasizing the critical role of proper validation and feature optimization for clinically meaningful PD detection.

III. Dataset Description

This study utilizes two publicly available voice based Parkinson's datasets (PD) from the UCI Machine Learning Repository, designed for classification and regression tasks:

A. Parkinson's Disease Detection Dataset: Classification

This dataset is intended for binary classification of Parkinson's disease V/s healthy controls. It comprises

197 voice recordings from 31 subjects, each represented by 22 biomedical voice features including fundamental frequency measures, jitter, shimmer, noise-to-harmonics ratios, and nonlinear dynamic measures. The target variable, status, indicates PD (1) or healthy (0), with 147 PD and 48 healthy samples. This dataset is widely used as a benchmark for voice-based PD detection (Fig. 1 shows the class distribution).

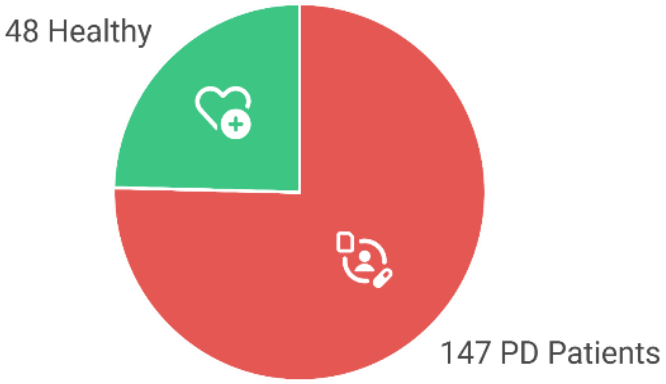


Fig. 1. Distribution of PD patients and healthy controls in the Parkinson's Disease Detection dataset.

B. Parkinson's Disease Telemonitoring Dataset: Regression

This dataset is used for predicting PD severity via regression. It contains 5,875 voice recordings from 42 PD patients, collected over a six-month clinical trial. Each record includes 16 biomedical voice features (e.g., jitter, shimmer, NHR, HNR, RPDE, DFA, PPE) along with demographic details such as subject ID, age, sex, and test time. The two target outputs, motor UPDRS and total UPDRS, are standard clinical measures used to monitor Parkinson's progression.

IV. Proposed Methodology

This study proposes a hybrid stacking ensemble framework that integrates two powerful gradient boosting algorithms, XGBoost and LightGBM, as base learners, and employs a task-specific meta-learner to generate the final predictions. The

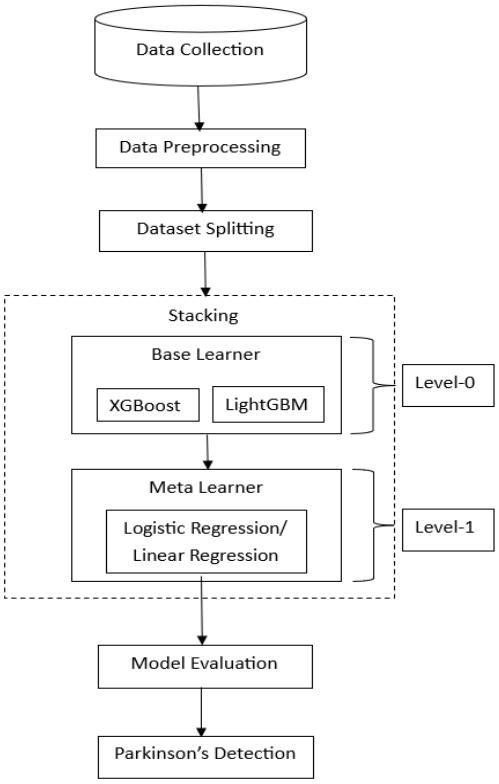


Fig. 2. Working flow diagram for the detection of Parkinson's Disease

framework is applied on two Parkinson's datasets: one formulated as a binary classification task and the other as a regression task. The block diagram of the proposed hybrid stacking ensemble–based framework for the detection of Parkinson's Disease is shown in Fig. 2.

Algorithm 1: Hybrid Stacking Ensemble for PD Diagnosis

Input: Dataset D, Base Learners B = {XGBoost, LightGBM}, Meta Learner M

Output: Final Trained Ensemble Model M_{final}

1: Preprocess D and split into D_{train}, D_{val}, D_{test}
2: For each $B_i \in$ B do
3: Tune hyperparameters of B_i on D_{train}
4: Train B_i on D_{train}
5: End For
6: Generate $F_{meta} \leftarrow$ predictions of $\{B_1(x), B_2(x)\}$ on D_{val}
7: Train M on F_{meta} using true labels y_{val}
8: For each $x \in D_{test}$ do
9: Obtain base predictions P $\leftarrow [B_1(x), B_2(x)]$
10: Predict final output $y_{pred} \leftarrow M(P)$
11: End For
12: Evaluate the predictive performance of the ensemble
13: Compute average feature importance FI_{avg} across base learners
14: Return $M_{final} \leftarrow$ M

A. Data Preprocessing

Prior to model development, both datasets underwent Exploratory Data Analysis (EDA) to examine its

structure, detect anomalies, and identify any missing values. No missing entries were found. Irrelevant attributes such as subject identifiers were removed during cleaning phase. All features were normalized using StandardScaler to ensure a uniform distribution and prevent features with larger ranges from dominating model training. And then, the dataset partitioned into three subsets 70%-training, 15%-validation, and 15%-testing. The training set were used to fit the models, the validation set for hyperparameter tuning, and the test set for unbiased evaluation of the final model.

B. Base Learners(Level-0)

At the very first level of the stacking ensemble, two gradient boosting models: XGBoost and LightGBM are employed as base learner. Given an input dataset $D = \{(x_i, y_i)\}_{i=1}^{n}$, where $x_i \in R^d$ denotes the feature vector and y_i represents the target variable, the prediction functions of the two learners are defined as:

$$f^{(1)}(x_i) = XGBoost(x_i; \theta_1), \quad f^{(2)}(x_i) = LightGBM(x_i; \theta_2) \quad (1)$$

Where θ_1 and θ_2 are the trainable parameters of XGBoost and LightGBM, respectively. Eq. (1) defines the parametric mapping functions of the base learners prior to model training, illustrating how input x_i would be transformed by each model.

XGBoost is a powerful and scalable implementation for gradient boosting, that constructs an ensemble of decision trees sequentially to reduce the overall loss. Its overall objective function is defined as Eq. (2):

$$L_{XGB} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k),$$

$$\Omega(f) = \gamma T + \tfrac{1}{2}\lambda||w||^2 \quad (2)$$

Here, $l(y_i, \hat{y}_i)$ represent the differentiable convex loss function, and $\Omega(f)$ is the regularization term that manage model complexity and helps to prevent overfitting.

LightGBM, on the other hand, is optimized for large-scale applications and improves computational efficiency using histogram-based feature binning and leaf-wise tree growth. Its objective function is given as Eq. (3):

$$L_{LGBM} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \lambda||w||^2 \quad (3)$$

In Eq. (3), the first component denotes the loss function and the second introduces an L_2 regularization penalty applied to leaf weights, which improves generalization performance. Once both base learners are trained on dataset D, their individual predictions for each input sample can be expressed as:

$$\hat{y}_i^{(1)} = f^{(1)}(x_i), \quad \hat{y}_i^{(2)} = f^{(2)}(x_i) \quad (4)$$

where $\hat{y}_i^{(1)}$ and $\hat{y}_i^{(2)}$ correspond to the outputs produced by the optimized XGBoost and LightGBM models, respectively. Eq. (4) thus captures the prediction phase post-training. These outputs serve as intermediate features and are concatenated to construct the meta-feature vector, and then fed into the Level-1 meta-learner for final prediction.

C. Meta Feature Construction

The predictions obtained from the base learners are concatenated to form a new feature vector. For each input sample x_i, the stacked representation is Eq (5):

$$Z_i = [\hat{y}_i^{(1)}, \hat{y}_i^{(2)}] \in R^2 \quad (5)$$

This stacked vector Z_i captures complementary information from both gradient boosting models, as each base learner may focus on different feature relationships and decision boundaries. Merging these outputs into the unified meta space, the framework enhances the richness and diversity of the feature representation. These meta features are used as input to the Level-1 meta learner, allowing it to exploit the strengths of both base models and refine the final prediction.

D. Meta-Learner (Level-1)

A task specific meta learner is used, which adjusts according to the prediction task- a logistic regression model is applied for classification tasks, whereas a linear regression model is used for regression. The meta learner takes the stacked outputs of the base learners as input and generates the final prediction $\hat{y}_i$ as define in Eq. (6):

$$\hat{y}_i = MetaLearner(Z_i; \theta_m) \quad (6)$$

Where θ_m denotes the parameters of the meta learner. This two level hybrid framework ensure that the model benefits from the predictive strength of gradient boosting learner while leveraging generalization ability of linear learners. The performance of the proposed ensemble model is validated using multiple evaluation metrics, along with confusion matrix and ROC-AUC analysis. And confirming its robustness and reliability. Furthermore, feature importance averaged from the XGBoost and LightGBM, highlights the most influential predictors and adding interpretability to the model.

V. Results and Discussion

A. Performance Metrics

To comprehensively test the effectiveness of this proposed models, different performance metrics were employed for both the classification and regression tasks. The use of various evaluation measures ensures robust comparison of baseline models with the proposed hybrid-stacking ensemble model.

1) Classification Metrics: The following metrics were used to assess model performance for classification tasks-

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$F1-score = \frac{2\times Recall\times Precision}{Recall+Precision} \quad (10)$$

$$AUC-ROC = \int_0^1 TPR(FPR^{-1}(t))\,dt \quad (11)$$

Where "TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives", respectively.

2) Regression Metrics: Regression performance were evaluated using the following metrics-

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (12)$$

$$RMSE = \sqrt{MSE} \quad (13)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (15)$$

Where, y_i and $\hat{y}_i$ are actual value and predicted value, and $\bar{y}$ is the mean of observed value.

B. Results on Parkinson's Disease Detection Dataset

The performance of XGBoost, LightGBM, and the proposed Hybrid-Stacking Ensemble model were tested using various classification metrics. Table 1 summarizes the results below. Performance of various models are also visualized in fig. 2. Performance of hybrid stacking ensemble model is demonstrated through ROC-AUC curve in fig. 3.

TABLE I. PERFORMANCE COMPARISON OF MODELS ON PARKINSON'S DISEASE DATASET

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
XGBoost	93.33	92.86	100.00	96.30	94.23
LightGBM	96.67	98.40	96.15	98.04	100.00
Hybrid	96.67	96.30	100.00	98.11	98.08

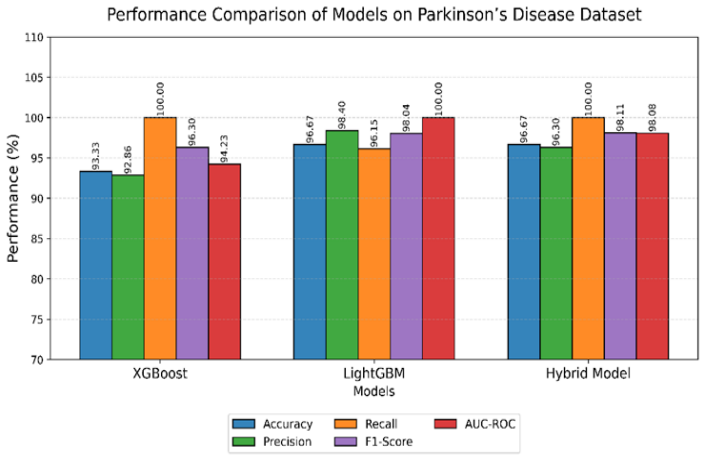


Fig. 2. Performance comparison of XGBoost, LightGBM, and Hybrid Model on Parkinson's Disease dataset.

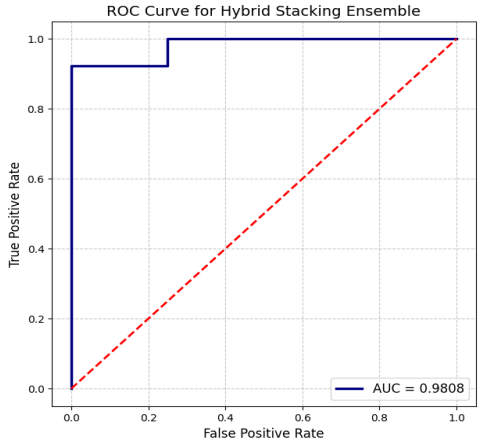


Fig. 3. ROC-AUC curve validating the effectiveness of the proposed Hybrid Model.

C. Results on Parkinson's Disease Telemonitoring Dataset

Telemonitoring dataset was evaluated for both regression 'predicting UPDRS scores' and classification 'disease severity classification based on UPDRS thresholds'. Accuracy comparison of three models are illustrated in fig. 4(a) and fig. 4(b) in terms of R^2 score and percentage accuracy for Total_UPDRS and Motor_UPDRS respectively. Model reliability is illustrated through fig. 5(a) and fig. 5(b).

Model performance results are also explained through various table. Performance of our proposed ensemble model ((XGBoost, LightGBM and Hybrid) for regression analysis on Total_UPDRS is tabularised in Table II, where Table III is for classification. Similarly, performance measure on Motor_UPDRS for regression analysis is given in Table IV, and for the classification, performance of all three ensemble models (XGBoost, LightGBM and Hybrid) is tabulated in Table V.

TABLE II. PERFORMANCE COMPARISON OF MODELS ON TOTAL UPDRS FOR REGRESSION ANALYSIS

Model	MSE	RMSE	MAE	R^2 SCORE
XGBoost	3.1401	1.7720	0.9997	0.9711
LightGBM	3.7605	1.9392	1.1200	0.9654
Hybrid	2.8962	1.7018	0.9636	0.9734

TABLE III. PERFORMANCE COMPARISON OF MODELS ON TOTAL UPDRS FOR CLASSIFICATION

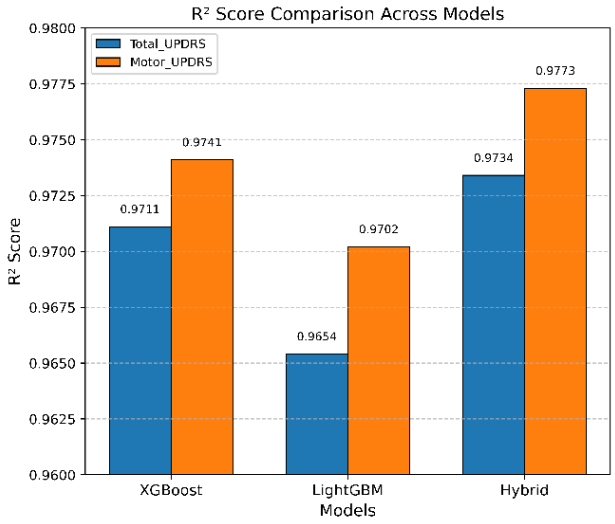
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
XGBoost	97.17	97.25	99.12	97.18
LightGBM	96.49	96.82	98.67	97.74
Hybrid	97.17	97.25	99.12	98.18

TABLE IV. PERFORMANCE COMPARISON OF MODELS ON MOTOR UPDRS FOR REGRESSION ANALYSIS

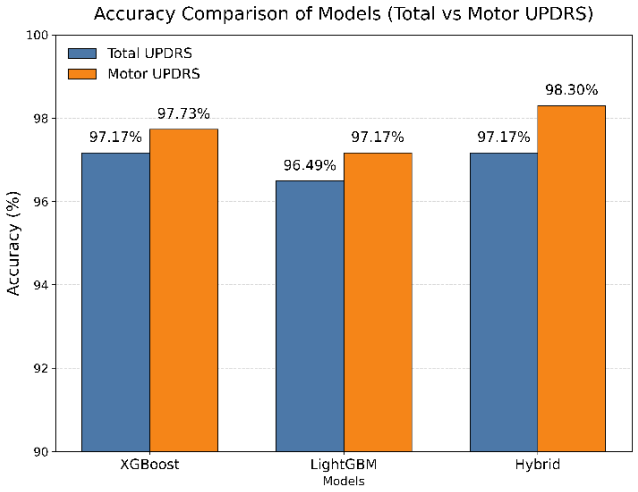
Model	MSE	RMSE	MAE	R^2 SCORE
XGBoost	1.6341	1.2783	0.7561	0.9741
LightGBM	1.8806	1.3714	0.8613	0.9702
Hybrid	1.4333	1.1972	0.202	0.9773

TABLE V. PERFORMANCE COMPARISON OF MODELS ON MOTOR UPDRS FOR CLASSIFICATION

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
XGBoost	97.73	97.18	98.46	97.82
LightGBM	97.17	96.34	98.24	97.28
Hybrid	98.30	97.83	98.90	98.36

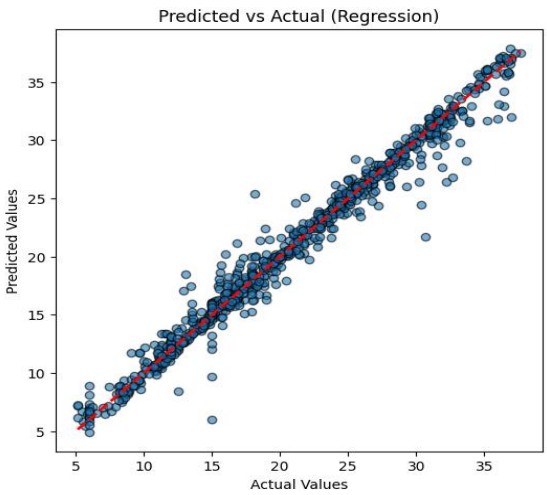


(a)

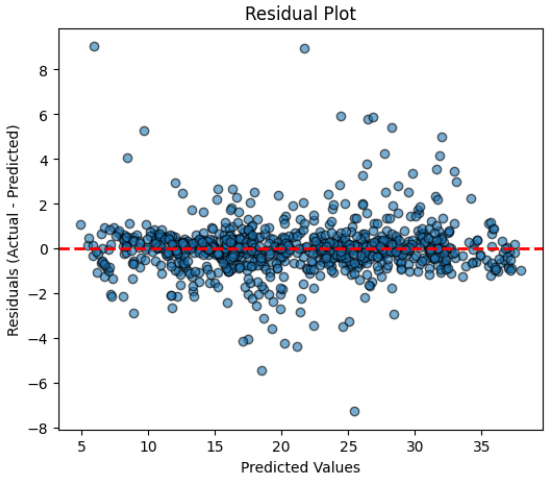


(b)

Fig. 4. (a) Comparison of R² scores, (b) Comparison of prediction accuracies of models for Total and Motor UPDRS.



(a)



(b)

Fig. 5. (a) Predicted vs. actual values showing high accuracy. (b) Residual plot with random errors around zero, confirming model reliability.

D. Discussion

The outcomes of the suggested work show how well the hybrid-stacking ensemble model works for Parkinson's disease monitoring and detection across several datasets. The model successfully capitalized on complementary strengths by combining XGBoost and LightGBM as base learners with a meta-learner, capturing intricate non-linear patterns while preserving generalization. Importantly, the model's high recall emphasizes its clinical utility by reducing false negatives, which are important for accurate and timely diagnosis.

The key biomarkers that play a major role in the accurate diagnosis of Parkinson's disease were identified. For this, the feature-importance scores obtained from both XGBoost and LightGBM models were analyzed by taking an average. This process provided useful and practical information for clinical evaluation and early intervention. An interesting observation was that when additional preprocessing techniques such as correlation-based feature selection, outlier removal or SelectKBest were used, the performance of the model either declined or showed no improvement. This makes it clear that the basic feature distributions in medical datasets already contain the necessary predictive information, and changing them excessively can weaken the model's ability. Compared to the single-model approaches used in earlier researches, the ensemble model of this study turned out to be better in both recall and generalization. High recall is especially beneficial for neurologists as it reduces the chances of misdiagnosis and enables them to start treatment on time. Overall, the results obtained show that this hybrid stacking model can really help in clinical decision-making. It remains stable across different datasets, maintaining accuracy, and is also useful in monitoring disease progression. Because of this, this model can become a reliable and practically useful option for clinical use.

VI. Conclusion and Future Scope

In summary, the hybrid stacking framework presented in this research - which works with combined strength of XGBoost and LightGBM, provides a reliable and practical way to detect and track the progression of Parkinson's disease. By combining the different characteristics of the two models, this framework gives improved generalization across different datasets and also makes clinical assessment more accurate, as it maintains critical feature-relationships without impairment.

Future research will focus on integrating multimodal data to make this system more robustness, using explainable AI techniques to enhance interpretability, and conducting comprehensive validation studies so that this model can be equally effective in real-world clinical deployment.

REFERENCES

1. Mamun M, Mahmud MI, Hossain MI, Islam AM, Ahammed MS, Uddin MM. Vocal feature guided detection of parkinson's disease using machine learning algorithms. In2022 IEEE 13th annual ubiquitous computing, electronics & mobile communication conference (UEMCON) 2022 Oct 26 (pp. 0566-0572). IEEE.
2. Ranjan, Rajnish K., and Divyarth Rai. "Alzheimer's Disease Image Diagnosis by Active Contour Segmentation & Bootstrap Bagging Learning Model." Journal of Computational Analysis & Applications 33.5 (2024).
3. Oh SL, Hagiwara Y, Raghavendra U, Yuvaraj R, Arunkumar N, Murugappan M, Acharya UR. A deep learning approach for Parkinson's disease diagnosis from EEG signals. Neural Computing and Applications. 2020 Aug;32(15):10927-33.
4. Prashanth R, Roy SD, Mandal PK, Ghosh S. High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. International journal of medical informatics. 2016 Jun 1;90:13-21.
5. Dhivyaa CR, Nithya K, Anbukkarasi S. Enhancing Parkinson's disease detection and diagnosis: a survey of integrative approaches across diverse modalities. IEEE Access. 2024 Oct 28.
6. Templeton JM, Poellabauer C, Schneider S. Classification of Parkinson's disease and its stages using machine learning. Scientific reports. 2022 Aug 18;12(1):14036.
7. Dentamaro V, Impedovo D, Musti L, Pirlo G, Taurisano P. Enhancing early Parkinson's disease detection through multimodal deep learning and explainable AI: insights from the PPMI database. Scientific Reports. 2024 Sep 9;14(1):20941.
8. Suppa A, Costantini G, Asci F, Di Leo P, Al-Wardat MS, Di Lazzaro G, Scalise S, Pisani A, Saggio G. Voice in Parkinson's disease: a machine learning study. Frontiers in neurology. 2022 Feb 15;13:831428.
9. Solana-Lavalle G ,Galán-Hernández JC, Rosas-Romero R. Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. Biocybernetics and Biomedical Engineering. 2020 Jan 1;40(1):505-16.
10. Ranjan, Rajnish K., and Divyarth Rai. "A Statistical Study of Lewy Body and Other Types of Dementia." 2024 IEEE International Conference on Intelligent Signal Processing and Effective Communication Technologies (INSPECT). IEEE, 2024.
11. Ranjan, Rajnish K., and Divyarth Rai. "Dementia and Its Subtypes: A Statistical Review." International Conference on Soft Computing: Theories and Applications. Singapore: Springer Nature Singapore, 2024.
12. Di Cesare MG, Perpetuini D, Cardone D, Merla A. Machine learning-assisted speech analysis for early detection of Parkinson's disease: a study on speaker diarization and classification techniques. Sensors. 2024 Feb 26;24(5):1499.
13. Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Computer Science. 2023 Jan 1;218:249-61.

14. Nahar N, Ara F, Neloy MA, Biswas A, Hossain MS, Andersson K. Feature selection based machine learning to improve prediction of Parkinson disease. InInternational Conference on Brain Informatics 2021 Sep 15 (pp. 496-508). Cham: Springer International Publishing.
15. Kadhim MN, Al-Shammary D, Mahdi AM, Ibaida A. Feature selection based on Mahalanobis distance for early Parkinson disease classification. Computer Methods and Programs in Biomedicine Update. 2025 Jan 1;7:100177.
16. Lamba R, Gulati T, Alharbi HF, Jain A. A hybrid system for Parkinson's disease diagnosis using machine learning techniques. International Journal of Speech Technology. 2022 Sep;25(3):583-93.
17. Wang W, Lee J, Harrou F, Sun Y. Early detection of Parkinson's disease using deep learning and machine learning. IEEE access. 2020 Aug 12;8:147635-46.
18. Jasim H, Alshakarchy ND. Predicting the Severity of Parkinson's Disease Based on Voice Analysis Using Deep Learning. In2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA) 2024 Jul 12 (pp. 61-66). IEEE.
19. Quan C, Ren K, Luo Z, Chen Z, Ling Y. End-to-end deep learning approach for Parkinson's disease detection from speech signals. Biocybernetics and Biomedical Engineering.
20. Al-Tam RM, Hashim FA, Maqsood S, Abualigah L, Alwhaibi RM. Enhancing Parkinson's disease diagnosis through stacking ensemble-based machine learning approach. IEEE Access. 2024 Jun 3;12:79549-67.
21. Omodunbi BA, Olawade DB, Awe OF, Soladoye AA, Aderinto N, Ovsepian SV, Boussios S. Stacked Ensemble Learning for Classification of Parkinson's Disease Using Telemonitoring Vocal Features. Diagnostics. 2025 Jun 9;15(12):1467.