



## Application of machine learning models to classify Parkinson disease patients using accelerometer

<sup>1</sup>Ameenu Abdulhameed Rabi, <sup>2</sup>Ismail Ahmad Ibrahim, <sup>3</sup>Buhari Bashir and <sup>4</sup>Salihu Abdullahi Audu

<sup>1</sup>No. 70 Kundilar Gandu, Zoo Road Housing Estate, Kano.

<sup>2</sup>No. 13 Dutsinma Street Tudun Wada, Kaduna.

<sup>3</sup>No. 8 Funtua Crescent K/Kaura New Layout, Katsina.

<sup>4</sup>Department of Computer Science, Nasarawa State University, Nigeria

Corresponding Author's Email: [abdulgiz@nsuk.edu.ng](mailto:abdulgiz@nsuk.edu.ng)

Received on: April, 2025

Revised and Accepted on: May, 2025

Published on: July, 2025

### ABSTRACT

Parkinson's disease (PD) greatly affects mobility, emphasizing the critical need for early detection to optimize treatment outcomes. This study explores the application of advanced machine learning techniques to analyse data collected from specialized motion tracking sensors known as accelerometers. These sensors monitor individual's movements and symptoms associated with Parkinson's disease, with a primary focus on comprehending the unique movement patterns and related symptoms prevalent in those affected by this condition. This research leverages the capabilities of machine learning, employing diverse algorithms such as support vector machines, neural networks, and random forest. Through an integrated approach, the study aims to construct a robust ensemble model capable of synthesizing insights from these techniques. The dataset utilized originates from the University of Zaragoza Hospital, encompassing a diverse range of participants, including individuals both diagnosed and undiagnosed with Parkinson's disease. This diversity ensures a comprehensive exploration of various movement patterns and symptoms among heterogeneous individuals. The primary objective revolves around precise differentiation between individuals affected by Parkinson's disease and those who are not. To achieve this, the study adopts an ensemble model strategically designed to reconcile conflicting predictions from individual classifiers. This methodological approach seeks consensus by aggregating multiple classifier opinions, thereby minimizing uncertainties arising from divergent predictions. The outcomes of this study are particularly significant, with the ensemble model demonstrating superior performance over traditional machine learning models. The best-performing ensemble model, as identified through comparative analysis, achieved an impressive accuracy of 65.69%, specificity of 73.5%, and precision of 71.17%. These metrics underscore the potential of ensemble classifiers in enhancing diagnostic precision, thereby contributing to the early detection and continuous monitoring of parkinson disease.

**Keywords:** Parkinson disease, advanced machine learning, support vector machines, neural networks, random forest

### 1.0 INTRODUCTION

#### 1.1 Background

Parkinson's Disease (PD) is a neurodegenerative disorder that significantly affects millions of people worldwide. It is characterized by a gradual decline in motor functions, evidenced by symptoms such as tremors, stiffness, and slowed movement, which greatly affect an individual's quality of life. Beyond these motor symptoms, PD can also lead to cognitive impairments and mood disorders, adding layers of complexity to its diagnosis and treatment. The variation in symptom manifestation and response to treatment among patients further complicates the clinical approach to PD. Traditionally, PD diagnosis relies on clinical evaluations, including medical history and neurological exams. While these methods are invaluable, their subjective nature can lead to inconsistent diagnoses. This paper is inspired by the promising capabilities of machine learning, particularly

its application to objective data like accelerometer readings. Machine learning offers a sophisticated toolset capable of identifying patterns within movement data that are otherwise difficult for humans to detect. This innovative approach aims to bridge a gap in the early and precise diagnosis of PD, potentially transforming patient care. By providing a more detailed understanding of PD's diverse manifestations, machine learning covers the path for customized treatment plans. This marks a significant stride towards enhancing the well-being and outcomes for those living with PD (Meenakshi and Kishore, 2023).

The paper reviews a variety of ML techniques applied to create a diagnostic system for PD. It searches through into using individual models such as ANN, KNN, SVM, and Naive Bayesian Classifiers for predicting PD (Anisha and Saranya, 2022). It also outlines different approaches and introduces a



universal framework for PD diagnosis based on voice signals and specifically explores various machine learning models employed in PD prediction, including ensemble classifiers (Anisha and Saranya, 2022). The paper underscores the precision and resilience of ensemble classifiers when compared to individual ones. It shortly explains diverse ML algorithms and their characteristics. Stacking, or stacked generalization, is a method that brings together information from different basic learners, making the whole system more reliable and complete (Anisha and Saranya, 2022). They work by combining different models, giving more solid and precise results for diagnosing PD. Another study by has indicated that ML algorithms can be proficiently employed to evaluate changes in an individual's vocal patterns to establish whether they have PD. The algorithm, utilizing parameter extracts and XGBoost classifiers, has demonstrated a significant level of precision in ascertaining the health condition of individuals diagnosed with PD Significantly, the XGBoost classifier demonstrates notable effectiveness in accurately detecting the presence of PD Divya et al. (2023). The algorithm employs feature extraction parameters and XGBoost classifiers to analyse vocal patterns and identify the presence of PD. The study utilizes an audio dataset from the UCI dataset database as input for the algorithm. Notably, the algorithm attains a high level of precision in PD detection, with XGBoost classifiers demonstrating the highest accuracy among the employed methods. Furthermore, Early identification of vocal cord issues, leading to speech difficulties, serves as a vital early sign of PD. As no specific tests for early PD detection currently exist, scientists are working on developing generalized data based decision systems. The study suggests a combined approach of ML boosting classifiers, like randomforest, AdaBoost, and XGBoost, with assigned weights for automatic pd identification. To ensure accuracy, the proposed system incorporates outlier rejection and attribute selection as necessary initial steps. The combined model exhibits superior performance compared to individual classifiers, presenting a more efficient tool for healthcare professionals in diagnosing PD (Ahmad et al., 2023).

By applying ML, we can teach the algorithms to recognize patterns and trends in the UPDRS data, helping us to make more accurate predictions and assessments about the progression of the disease. This integration allows us to gain deeper insights into how PD affects individuals, potentially leading to more personalized and precise treatment plans tailored to each patient's specific needs. The scores are vital for assessing how serious PD is, especially when it comes to predicting how the illness might progress. In

addition, factors like Jitter percentage and Shimmer percentage are considered separate elements in regression models to anticipate UPDRS scores. This helps doctors understand how these specific factors might influence the overall severity of the disease and its potential trajectory. Another study by Shastry (2023) utilized four different types of regression techniques - multilinear, ridge, robust, and UPDRS regression to anticipate UPDRS scores based on these specific factors. multilinear regression models relationships with multiple independent variables, ridge regression prevents overfitting with a regularization term, robust regression is less influenced by outliers, and UPDRS regression aids in feature selection by shrinking some coefficients to zero. After an examination, they determined that ridge regression proved to be the most precise model for foreseeing UPDRS scores in the investigation. This finding suggests that ridge regression is the most reliable method for estimating the severity of PD based on the given parameters (Shastry, 2023). Furthermore, this study explores the use of force plate data during quiet standing to predict motor symptom severity in PD patients. Employing ML, specifically the Random Forest algorithm, it identifies key features correlated with the centre of pressure range, crucial for predicting PD symptoms. The study highlights the potential of quiet standing as an additional diagnostic tool for PD, enabling the identification of disease progression and providing specialists with valuable insights for accurate diagnosis and monitoring (Exley et al., 2022).

### 1.2 Aim and Objectives

The Primary aim of the research is to employ ML techniques to categorize accelerometer data from patients with PD. The proposed approach is expected to provide accuracy and faster processing time than existing techniques.

The following are the objectives of this research work.

- i To conduct a comprehensive review of existing literature on the application of ML in the classification of accelerometer data from PD patients and identify the strengths and limitations of previous studies in this domain.
- ii To develop a robust and accurate ML models that can effectively differentiate between accelerometer data from PD patients and controls and implement various machine learning algorithms and techniques, such as support vector machines, neural networks, and random forests, for the creation of the classification model.
- iii To assess the classification model's effectiveness and reliability in accurately distinguishing PD patients from non-PD patients based on their accelerometer data and to measure the



model's classification accuracy, precision, recall, specificity, sensitivity, ROC AUC, F1 score, MCC and determine its capacity in aiding the early detection and monitoring of PD.

### 1.3 The Hypothesis

Application of ensemble classifiers to accelerometer data from PD patients will result in a more accurate and robust classification of motor symptoms compared to individual classifiers, providing a promising approach for the early detection and monitoring of PD.

## 2.0 MATERIALS AND METHOD(S)

### 2.1 Data Collection

This study initiated an in depth data collection process, making use of wrist worn accelerometers to gather detailed movement information from two primary groups: those diagnosed with PD and a healthy control group for comparison. Importantly, within the PD group, individuals were categorized based on whether they were receiving medication, acknowledging the significant effects that medication can have on the symptoms of the disease. This data gathering was conducted with the expertise of medical professional which are the clinicians from the University of Zaragoza Hospital in Spain, ensuring the process was thorough and aimed to cover a broad range of daily activities. This careful and detailed approach to collecting data guarantees that the dataset not only accurately captures the complex motor difficulties that PD patients experience but also includes comparative data from healthy individuals, providing a rich basis for analysis.

### 2.2 Data Set Description

The dataset derived from accelerometer data of individuals with PD, tracks their movements using a specialized device called an accelerometer. This device captures the complexities of their actions over time, offering a detailed map of their motions through various coordinate features. Features were extracted from the collected data, resulting in a dataset with 192 features and 82 rows. This dataset plays a crucial role in enabling ML techniques to identify patterns and categorize different movements associated with PD. However, challenges arise from imbalances in the dataset, with fewer instances of disease related movements impacting the accuracy of the models. Additionally, Time series complexities such as temporal dependencies and irregular intervals also affect the precision of modeling efforts.

### 2.3 Software Environment

The implementation of the ML models and data pre-processing techniques described in this study was

conducted using the Python programming language. Specifically, Python version 3.12 was utilized for the development and execution of the code

### 2.4 Libraries and Frameworks

The following libraries and frameworks were employed for various tasks within the paper:

- Scikit-learn: A comprehensive machine learning library in Python used for data preprocessing, model training, and evaluation (Fabian, 2011).
- TensorFlow: The study of Smilkov et al. (2019) shows tensorflow as an open-source machine learning framework developed by Google for building and training neural networks.
- XGBoost: A scalable and efficient implementation of gradient boosting for classification and regression tasks.
- Pandas: A powerful data manipulation library in Python used for data handling and analysis.
- NumPy: A fundamental package for scientific computing with Python, providing support for large, multi-dimensional arrays and matrices.
- Matplotlib: A plotting library for the Python programming language and its numerical mathematics extension, NumPy (Sial et al., 2021).
- Seaborn: A data visualization library based on Matplotlib, providing a high-level interface for drawing attractive statistical graphics (Sial et al., 2021).

### 2.5 Evaluation and Comparative Analysis

In the evaluation and comparative analysis, Comprehensive set of metrics were adopted to carefully assess the performance of the ML models utilized in our study. This evaluative approach went beyond the conventional measure of accuracy, precision, recall, and the MCC among others. This multidimensional framework for assessment allowed for a refined understanding of each model's effectiveness. Through this detailed evaluation, I was able to clearly identify the superior capabilities of ensemble methods in handling the complexities of PD data, which often presents complex patterns and variations. Moreover, The analysis underscored the critical need for careful model selection, tailored to the unique characteristics of the dataset at hand and the specific requirements of the diagnostic challenge being addressed.

This methodical comparison not only showcases the distinct advantages of different modeling approaches but also highlights our commitment to advancing the precision and reliability of PD diagnosis. By carefully assessing these models against a broad spectrum of performance metrics, I aim to contribute to the



development of more effective diagnostic tools, ultimately facilitating early detection and more personalized treatment plans for individuals with PD (Nilashi et al., 2022). The logistic regression model with carefully tuned hyperparameters, showed a balanced performance with an accuracy of 68.6% and a notable recall of 72.5%, indicating its strength in correctly identifying true positive cases. The neural network, optimized for a combination of layer sizes and other parameters in table 3.2, also performed well, particularly in precision at 77.5%, suggesting its efficacy in minimizing false positives.

Support vector machines tailored with the optimal choice of kernel and regularization parameters, demonstrated a moderate performance with a precision of 62.5% and a recall of 60%, reflecting its cautious approach to classification. The ensemble methods, including stacking, random forest, and XGBoost, brought diverse strategies into play. Stacking, which leverages the predictions of multiple models, achieved a precision of 70.8% and a recall of 60%, showing its ability to balance false positives and negatives effectively. XGBoost which is also known for its gradient boosting framework, and the random forest model both highlighted their strengths through a combination of high precision, recall, and specificity, indicating their robustness in handling both classes effectively. XGBoost, with its fine tuned learning rate and tree specific parameters, and random forest, optimized for its ensemble strategy, stood out for their ability to maintain a good balance across all evaluated metrics

Comparing these models, ensemble classifiers like

XGBoost and random forest generally provided a stronger balance between sensitivity and specificity, indicating a solid ability to distinguish between the classes effectively. Their advanced algorithms that aggregate multiple decision processes help in reducing variance and bias, contributing to their overall superior performance in this context. In contrast, traditional ML models like logistic regression and support vector machines, while effective in certain metrics, demonstrated the importance of model selection based on the specific needs of the dataset and the prediction task at hand. This comparative analysis underlines the significance of evaluating a range of models and metrics to identify the most suitable approach for a given classification problem, emphasizing that the choice of model can greatly impact the effectiveness of predictions and the insights drawn from the data

### 3.0 RESULTS AND DISCUSSION

#### 3.1 Results Presentation

Figure 3 The assessment of different ML models for detecting PD has been carried out with attention to detail, and the results are laid out clearly and methodically. We have placed the performance indicators Ahmed et al. (2018) such as accuracy, precision, recall, F1 score, ROC AUC, sensitivity, specificity, and the MCC next to each other for every model. This setup allows for a direct comparison. This method highlights each model's unique strengths and provides insights into how effectively they can identify PD. Organizing the results this way helps to clearly see where each model excels and where it might fall short, aiding those involved in PD diagnosis to choose the most suitable model based on solid evidence.

**Table 1: Model Performance Metrics**

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	ROC AUC (%)	Sensitivity (%)	Specificity (%)	MCC
Logistic Regression	68.6±18.5	68.1±25.8	72.5±27.5	68.2±22.4	68.5±18.8	72.5±27.5	64.5±28.8	0.40±0.38
Neural Network	68.5±15.1	77.5±25.7	60.0±24.2	63.8±18.8	68.0±15.3	60.0±24.2	76.0±28.1	0.41±0.32
Support Vector Machine	63.9±25.7	62.5±29.2	60.0±33.7	59.8±30.1	63.5±26.0	60.0±33.7	67.0±28.5	0.28±0.53
Stacking	64.4±15.2	70.8±22.5	60.0±12.9	62.9±11.9	64.3±15.2	60.0±12.9	68.5±29.3	0.31±0.32
XGBoost	65.7±14.4	71.2±22.8	57.5±16.9	62.0±15.4	65.5±14.4	57.5±16.9	73.5±22.1	0.33±0.31
Random Forest	66.0±18.9	68.5±24.4	62.5±21.2	64.0±20.0	65.8±19.0	62.5±21.2	69.0±26.1	0.33±0.39

Table 1 presents the performance metrics of various ML models used for classifying individuals with PD who are not taking medication. Each row corresponds to a different model, including Logistic Regression, Neural Network, SVM, Stacking, XGBoost, and Random Forest. The columns represent different evaluation metrics, including Accuracy, Precision, Recall, F1 Score, ROC AUC, Sensitivity, Specificity, and MCC. These metrics measure different aspects of

the models performance, such as overall correctness, accuracy in predicting positive outcomes, ability to identify actual positive cases, balance between precision and recall, capability to distinguish between classes, ability to identify actual negatives, and a balanced measure considering all four quadrants of the confusion matrix. The values in each cell indicate the percentage achieved by each model for the corresponding metric. These metrics provide insights

into the strengths and weaknesses of each model in accurately classifying individuals with PD who are not taking medication.

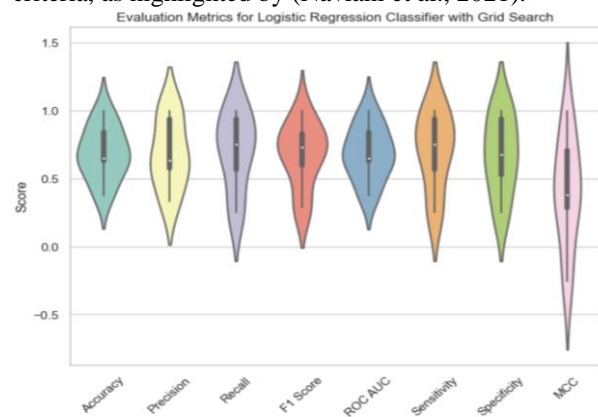
### 3.2 Interpretation of Results

The results of the classification models offer valuable insights into their performance. Starting with logistic regression, which demonstrated a moderate accuracy of 68.6%, the model exhibited balanced precision and recall, indicating its ability to provide reliable predictions without bias towards false positives and false negatives. Moving to neural network models, despite showcasing promising precision at 77.5%, the slightly lower recall at 60.0% suggests a potential for accurate identification of positive cases while maintaining a fair balance between precision and recall. The SVM model displayed a moderate accuracy of 63.9% along with balanced sensitivity and specificity, showcasing its capability to classify both positive and negative instances effectively. Ensemble methods, such as stacking and XGBoost, achieved competitive accuracies at 64.4% and 65.7%, respectively. These methods demonstrated balanced precision and recall, indicating robust classification capabilities across various evaluation metrics. Similarly, random forest models exhibited comparable performance to ensemble methods, further emphasizing their effectiveness in handling complex datasets. Lastly, the findings suggest that ensemble methods and neural networks present promising approaches for accurately classifying individuals with PD who are not on medication, offering potential advancements in disease diagnosis and management.

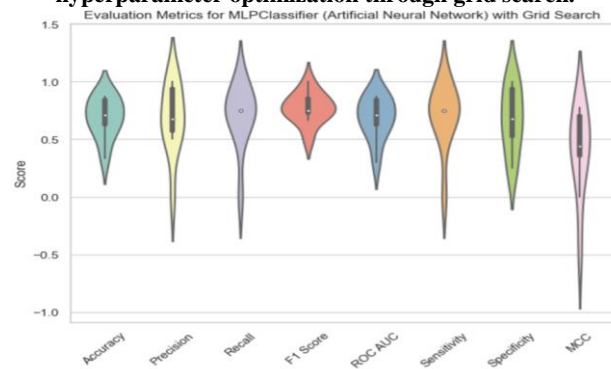
### 3.3 Violin Plots Explanation

The explanation of violin plots in the context of evaluating ML models for PD detection is both relevant and insightful. Violin plots serve as a powerful visual tool to illustrate the distribution and central tendencies of various performance metrics such as accuracy, precision, recall, F1 score, ROC AUC, sensitivity, specificity, and the MCC for each model. These plots illustrate the density of metric values, effectively highlighting where most data points cluster, thereby indicating the consistency and variability in model performance. For instance, a violin plot for the Logistic Regression model shows a wider section around the middle for recall and MCC metrics in figure 4.1, suggesting a significant concentration of data points near the mean, which denotes high reliability in these areas. Also, the Neural Network model notable for its precision have a violin plot with a narrower middle section at high precision values as shown in figure 4.2, indicating a strong ability to consistently minimize false positives. Similarly, In Support Vector Machine model moderate variability across the metrics is

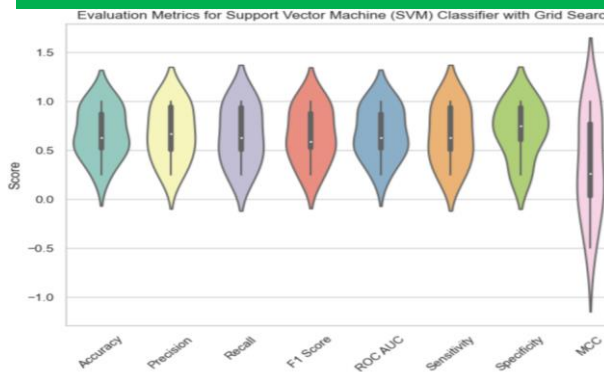
observed, indicating relatively consistent performance across different measures shown in figure 4.3. Ensemble methods such as Stacking, XGBoost and Random Forest in figure 4.4, 4.5 and 4.6, praised for their balanced performance across a range of metrics, exhibit violin plots with less variability, marked by more uniform shapes that signal steady performance across different measures. These visual representations are invaluable not only for highlighting the performance strengths and variabilities of each model but also for enabling a direct, visual comparison. By doing so, they aid significantly in the process of refining and choosing the most appropriate model for this classification, based on specific performance criteria, as highlighted by (Navlani et al., 2021).



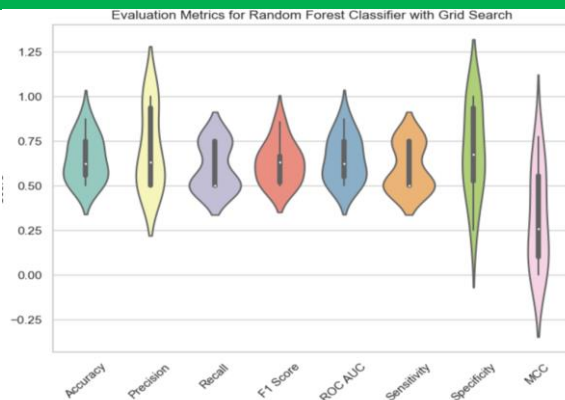
**Figure 1: Violin Plot Illustration of Evaluation Metrics for Logistic Regression.** This plot demonstrates the distribution and range of various performance scores including accuracy, precision, recall, F1 score, ROC AUC, sensitivity, specificity, and MCC, after hyperparameter optimization through grid search.



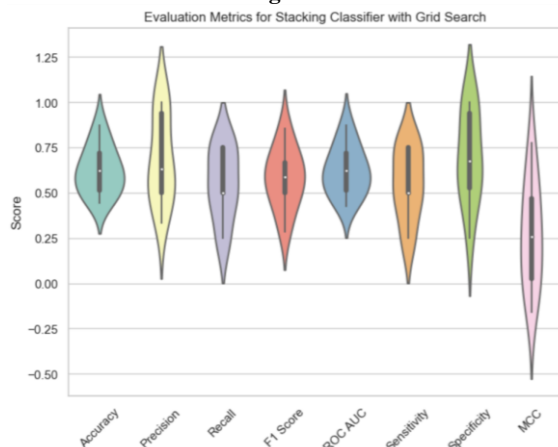
**Figure 2: Distribution of Performance Metrics for Neural Networks as Depicted by Violin Plots.** Showcasing the variability and central tendencies of the metrics, these plots highlight the effects of hyperparameter tuning on the model's ability to classify Parkinson's Disease.



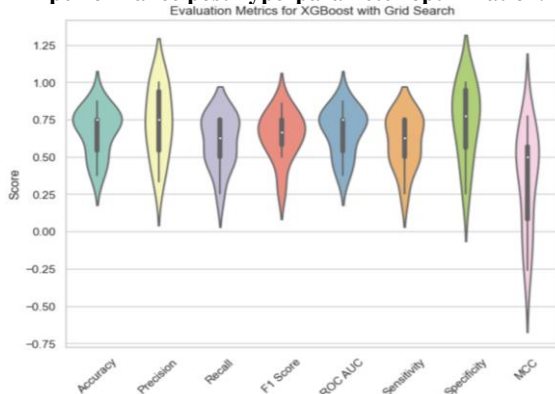
**Figure 3: Evaluation Metrics for SVM Classifier with Grid Search.** This violin plot visualizes the distribution and density of performance metrics accuracy, precision, recall, F1 score, ROC AUC, sensitivity, specificity, and MCC following hyperparameter tuning via grid search.



**Figure 6: Random Forest Classifier Evaluation Metrics Represented in Violin Plots.** These plots offer a visual summary of the model's performance across key metrics, revealing the range and concentration of scores achieved after tuning with a grid search approach.



**Figure 4: Performance Distribution for Stacking Classifier Using Violin Plots.** Each violin plot depicts the variability and median scores across several evaluation metrics, showcasing the classifier's performance post hyperparameter optimization.



**Figure 5: XGBoost Classifier Performance Metrics Illustrated with Violin Plots.** The plots highlight the results of the model's evaluation metrics obtained after systematic hyperparameter tuning, providing a visual comparison across the different scores.

## 4.0 DISCUSSION

### 4.1 Analysis of Results

The discussion effectively shows the comparative evaluation of ML models, including logistic regression, neural networks, SVM, stacking, XGBoost, and random forests, emphasizing the analysis of their performance through metrics such as accuracy, precision, recall, F1 score, ROC AUC, sensitivity, specificity, and the MCC. A key insight from the evaluation is the noticeable performance gap between single models (logistic regression, SVM, neural networks) and ensemble methods (stacking, XGBoost, random forest), with the recent demonstrating superior precision, sensitivity, and MCC. This superiority arises from the ensemble methods enhanced capability to model the complex, non-linear relationships inherent in PD symptoms and the effects of medication, alongside their robustness against overfitting, a common challenge for single models. Among the ensemble techniques, XGBoost stood out for its balanced performance across multiple metrics, notably in specificity and MCC, Examining its proficiency at accurately distinguishing PD cases without extreme bias towards either class. These findings align with existing literature, highlighting the effectiveness of ensemble methods in analyzing complex health data. Furthermore, our results, particularly the specificity of 74% and an MCC of 0.30 achieved by our XGBoost model, vibrate with the outcomes of Gattam (2024) research on deep learning for PD detection. This consistency with prior studies underscores the relevance of our research in advancing the application of ML in PD diagnostics.



#### 4.2 Project Complexities

This complexities are the challenges encountered in this paper. Firstly, accurately categorizing individuals with PD is difficult because of the need to distinguish between those taking medication and those not, considering how medication might affect or hide symptoms. Secondly, ensuring the dataset is reliable and balanced, especially in terms of class distribution and handling missing data, is a major complication. Furthermore, selecting the right ML model from many options requires careful thought about factors like how complex the model is and how easy it is to understand. Ethical concerns around automating disease detection make these challenges even more difficult, as issues like patient privacy and potential biases in the algorithm need to be addressed. Lastly, the changing nature of PD symptoms, influenced by medication and disease progression, adds another layer of complexity. To tackle these challenges, an approach is needed to combine technical skills with ethical considerations to create effective diagnostic tools for PD.

#### 4.3 Limitations and Challenges

One limitation lies in the size and balance of the dataset. While described as balanced, the overall dataset size may limit the models ability to generalize to a broader population. Smaller datasets might fail to capture the full spectrum of PD symptoms, especially considering the variability introduced by medication. Another critical aspect is the generalizability of the models to diverse populations or settings. Demographic variability, disease progression stages, and different medication types could significantly impact model performance. Additionally, handling missing values, particularly in the medication feature, presents a challenge. While imputing the most frequent value is practical, it may introduce bias underlying patterns related to medication effects on PD symptoms. These limitations underscore the need for careful consideration and mitigation strategies throughout the research.

#### 4.4 Key Findings

The key discoveries from this paper offer valuable insights into PD diagnosis. Firstly, it emphasizes the crucial role of medication status in accurately identifying PD cases. This means that whether someone is on medication significantly impacts how accurately we can diagnose PD. This highlights the complexity of PD symptoms and how they can change depending on whether someone is taking medication. Secondly, the study tells that ensemble methods like XGBoost, stacking, and random forests perform better than individual models such as logistic regression and SVM. These ensemble methods are more effective at capturing the detailed symptoms of PD, showing

promise for diagnosis. Among them, XGBoost stands out for its balanced performance across different metrics, suggesting it could be a reliable tool for PD diagnosis. These findings stress the importance of refining diagnostic models to include medication status and using ensemble methods to enhance diagnostic accuracy for PD.

#### 4.5 Implications for PD Diagnosis and Patient Care

The implications of this work for PD diagnosis and patient care are significant. By incorporating medication status into the diagnostic process, we can achieve more accurate and refined classifications of PD cases. This means that considering whether someone is on medication or not can lead to more precise diagnoses, which is crucial for early detection and the development of tailored treatment plans. Early and accurate diagnosis can have a profound impact on patient outcomes, enabling more personalized and effective management of PD. Moreover, by refining diagnostic models to account for factors like medication status, we can move closer to the goal of personalized medicine. This approach ensures that treatment strategies are tailored to each individual's specific condition and needs, ultimately improving the quality of care and outcomes for PD patients.

#### 5.0 CONCLUSION

The paper represents a significant advancement in using ML to diagnose PD. Through careful analysis of data, we have gained important insights into how PD cases can be classified, especially considering whether individuals are taking medication. Ensemble methods like XGBoost have proven to be essential, showing better accuracy in diagnosis compared to individual models. Incorporating medication status into our diagnostic models has been crucial, allowing for more precise classification of PD and enabling early intervention and personalized treatment plans. These findings are important for patient care, allowing healthcare providers to tailor treatments to each person's needs and improving outcomes for PD patients. Looking ahead, further research and refinement of diagnostic methods will be necessary to enhance our ability to diagnose PD and improve the lives of those affected by the disease.

#### 5.2 Future Work

Several areas for future work present themselves in advancing the application of ML to PD diagnosis. Firstly, further research is warranted to explore more sophisticated modeling techniques, such as deep learning architectures, which may better capture the complex patterns inherent in PD symptomatology.



Additionally, investigating the integration of multimodal data sources, including genetic information, neuroimaging scans, and wearable sensor data, could enhance the predictive power of diagnostic models. Moreover, longitudinal studies are essential to track disease progression over time and evaluate the effectiveness of personalized treatment plans based on diagnostic insights.

Furthermore, addressing the ethical considerations surrounding patient privacy, algorithm bias, and transparency remains crucial for the responsible deployment of ML in healthcare settings. Collaborative efforts between researchers, clinicians, and policymakers will be instrumental in driving these future endeavors to fruition, ultimately improving diagnostic accuracy and patient outcomes in PD care.

## REFERENCES

- Ahmad, M., Uddin, M. S., and Jang, Y. M., editors (2023). *Proceedings of International Conference on Information and Communication Technology for Development: ICICTD 2022*. Springer Singapore, Singapore, 1 edition.
- Ahmed, F. Y., Ali, Y. H., and Shamsuddin, S. M. (2018). Using k-fold cross validation proposed models for spikeprop learning enhancements. *International Journal of Engineering & Technology*, 7(4.11):145–151.
- Anisha, C. and Saranya, K. (2022). A machine learning based parkinson prediction system—a comprehensive.
- Bhuyan, H. K., Ravi, V., Brahma, B., and Kamila, N. K. (2022). Disease analysis using machine learning approaches in healthcare system. *Health and Technology*, 12(5):987–1005.
- Divya, J., Radhakrishnan, P., G, P., Gopatoti, A., Baburao, D., and Krishnamoorthy, R. (2023). Detection of parkinson disease using machine learning. In *2023 International Conference on Inventive Computation Technologies (ICICT)*, pages 53–57.
- Exley, T., Moudy, S., Patterson, R. M., Kim, J., and Albert, M. V. (2022). Predicting updrs motor symptoms in individuals with parkinson’s disease from force plates using machine learning. *IEEE Journal of Biomedical and Health Informatics*, 26(7):3486–3494.
- Fabian, P. (2011). *Scikit-learn: Machine learning in python*. *Journal of machine learning research* 12, page 2825.
- Gattam, S. S. R. (2024). Prediction of protein acetylation sites with machine learning. PhD thesis, Rutgers University-Graduate School-Camden.
- Meenakshi, K. and Kishore, D. (2023). Parkinson’s

disease classification using machine learning techniques. In *2023 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–5.

- Navlani, A., Fandango, A., and Idris, I. (2021). *Python Data Analysis: Perform data collection, data processing, wrangling, visualization, and model building using Python*. Packt Publishing Ltd.
- Nilashi, M., Abumalloh, R. A., Minaei-Bidgoli, B., Samad, S., Ismail, M. Y., Alhargan, A., and Zogaan, W. A. (2022). Predicting parkinson’s disease progression: evaluation of ensemble methods in machine learning. *Journal of healthcare engineering*, 2022.
- Shastri, K. A. (2023). Ensemble machine learning regression model based predictive framework for parkinson’s updrs motor score prediction from speech data. *International Journal of Speech Technology*, 26(2):433–457.
- Sial, A. H., Rashdi, S. Y. S., and Khan, A. H. (2021). Comparative analysis of data visualization libraries matplotlib and seaborn in python. *International Journal*, 10(1):45.
- Smilkov, D., Thorat, N., Assogba, Y., Nicholson, C., Kreeger, N., Yu, P., Cai, S., Nielsen, E., Soegel, D., Bileschi, S., et al. (2019). *Tensorflow.js: Machine learning for the web and beyond*. *Proceedings of Machine Learning and Systems*, 1:309–321.