AI-Driven Multi-Class Diagnosis of Parkinson's Disease: Enhancing Accuracy and Differentiation from Movement Disorders

Shubhika Verma
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
shubhikav02@gmail.com

Jasmeet Kaur
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
jasmeetkaur95460@gmail.com

Deepa Gupta
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
g_deepa@blr.amrita.edu

Aarathi Rajagopalan Nair
Department of Computer Science and Engineering
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
aarathirajagopalan19@gmail.com

***Abstract*—Parkinson's disease is a progressive neurological disorder that is frequently misdiagnosed due to its symptomatic similarities with other movement-related conditions. Ensuring precise and early diagnosis is essential for effective treatment; however, current diagnostic approaches are often subjective and prone to errors. To overcome these limitations, this study introduces a multi-class classification framework designed to differentiate Parkinson's disease from related disorders, including Essential Tremor, Atypical Parkinsonism, Secondary Parkinsonism, Multiple Sclerosis, and Healthy Controls. The classification process is structured into three tasks: distinguishing Parkinson's disease from Healthy Controls, differentiating it from other movement disorders, and performing a multi-class classification that includes Parkinson's disease, Healthy Controls, and additional disorders. Experimental evaluations indicate that XGBoost achieved an F1-score of 93% for Parkinson's disease versus Healthy Controls, SVM attained 85% for Parkinson's disease against other movement disorders, and Logistic Regression yielded 68% accuracy in the multi-class classification. These results highlight the potential of machine learning in providing accurate, scalable, and objective diagnostic solutions for Parkinson's disease.**

***Index Terms*—Parkinson's Disease, Machine Learning, Atypical Parkinsonism, Classification, Essential Tremor**

I. Introduction

Parkinson's disease (PD) is a degenerative neurological disorder caused by the depletion of dopamine-producing neurons in the brain, resulting in motor impairments. The worldwide prevalence of PD increased from 2.5 million in 1990 to 6.1 million in 2016, primarily due to aging populations, longer disease duration, and potential shifts in environmental and social risk factors [1]. As the disease advances, patients' quality of life deteriorates significantly, emphasizing the critical role of promoting good health and well-being through improved diagnostic and treatment strategies. Unlike other movement disorders, the primary method of diagnosing PD remains clinical examination, which can be supported by nuclear imaging techniques. However, these methods are not infallible and can be supplemented by innovative diagnostic approaches.PD impacts more than just the motor fun ction. Common non-motor symptoms (NMS) include cognitive impairment, dementia, hallucinations, depression, apathy, excessive daytime sleepiness, insomnia, and problems with impulse control. The progression of these symptoms is often interrelated and influenced by factors such as age, sex, disease severity, and antiparkinsonian medication—the latter being the sole theoretically changeable component [2]. Symptoms can worsen when accompanied with smoking and drinking alcohol habits. The complexity and interconnection of NMS in PD necessitates a thorough management and treatment strategy.

Early detection and treatment of Parkinson's disease is critical for reducing load on individuals and healthcare systems. However, no neuroprotective or regenerative treatments are currently available [3]. While the symptoms of PD are well-documented, their onset and course vary widely among individuals, complicating the prediction and management of the disease [4] [5]. In the field of medical research, supervised learning techniques are commonly utilized for predictive modelling. Robust evaluation techniques are required to assure the correctness and reliability of these models. For large datasets, a simple train-test split may be sufficient. However, for tiny and imbalanced datasets, cross-validation is important to produce trustworthy results [6]. This approach is especially relevant in the context of PD research, because dataset sizes might vary greatly [7]. The key contributions of this research are:

- This study introduces a novel approach for multi-class classification of Parkinson's disease, covering three dis-

tinct classification tasks: (1) differentiating Parkinson's disease (PD) from Healthy Controls (HC), (2) distinguishing PD from other neurological disorders such as Essential Tremor, Atypical Parkinsonism, Secondary Parkinsonism, and Multiple Sclerosis, and (3) a comprehensive classification of PD, HC, and other disorders.
- To address the significant class imbalance in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) is applied in conjunction with fifteen different classifiers.
- Feature selection and dimensionality reduction techniques are utilized alongside ensemble learning methods like Stacking and Bagging to leverage the strengths of individual classifiers and enhance predictive performance.

The organization of this paper is as follows: Section II reviews related works, focusing on prior experimental studies involving discharge summaries. Section III describes the dataset used in this research. Section IV elaborates on the proposed methodology. Section V presents the results along with their analysis. Finally, Section VI concludes the paper and explores potential future directions.

II. Related Work

Research on Parkinson's disease has predominantly concentrated on MRI scans, gait assessments, and genetic analysis for predictive modeling. However, there has been limited exploration of symptom-based questionnaire data for early detection. Various modalities, such as hand movement tracking, balance and gait analysis, eye tracking, and voice recordings, have been investigated [8] [9]. For example, [10] implemented an SVM-based approach to predict Parkinson's disease onset in elderly individuals using genetic data. Their improved SVM model achieved an accuracy of 0.9183, surpassing the initial accuracy of 0.889. Their findings also suggested that auditory data provided better classification results than genetic data. Likewise, [11] utilized a Random Forest model to assess Parkinson's disease severity in elderly patients by analyzing keyboard interaction patterns from the UCI telemonitoring dataset. Another study [12] focused on classifying Parkinson's disease patients (PWP) using audio data, though the research was primarily conducted using MATLAB.

Machine learning (ML), particularly deep learning, has significantly contributed to medical research [13], including studies on Parkinson's disease [14]. The study in [15] examined the impact of clinical and biofluid predictors on cognitive outcomes in Parkinson's disease using methods such as conditional inference forests, elastic nets, random forests, and support vector machines [16]. Moreover, GenoML, an open-source ML framework in Python, integrates transcriptomics, genetic, and clinical data to build peri-diagnostic models for Parkinson's disease risk assessment [17]. Despite the advantages of ML models in understanding disease risk, selecting an optimal algorithm and determining the most relevant predictors remains challenging, as different studies employ varying methodologies and predictor sets.

A major issue in Parkinson's disease research is the presence of imbalanced datasets, particularly regarding age and gender distribution, which can limit the generalizability of ML models [18]. Additionally, traditional clinical diagnoses of Parkinson's disease present difficulties. According to [19], approximately 15% of diagnosed patients did not meet strict clinical criteria, whereas 20% of true Parkinson's cases went undetected, highlighting the need for improved diagnostic accuracy. Similarly, [20] pointed out that misdiagnosis is common due to symptom similarities with other neurodegenerative diseases and the absence of definitive biomarkers. These challenges emphasize the necessity for continuous reassessment and expert collaboration to refine diagnostic precision.

Although existing models for Parkinson's disease prediction employ diverse ML techniques and predictor sets, systematic comparisons using identical datasets remain scarce. Additionally, the specific information required for effective risk assessment across various ML methodologies is unclear, especially given the cost and accessibility constraints of different diagnostic tools. Some risk factors, such as cerebrospinal fluid (CSF) biomarkers, require invasive procedures, while others, like PET imaging, are expensive, making data collection complex. Most studies have prioritized voice and speech analysis over questionnaire-based symptom data. Such questionnaires gather details on demographics (age, gender, height, and weight), family history of Parkinson's disease, alcohol's effect on tremors, and responses to 30 yes/no questions regarding Parkinson's-specific non-motor symptoms (PDNMS). Addressing data imbalance is essential, and this study aims to tackle that challenge.

The proposed model seeks to close these research gaps by incorporating class balancing techniques and feature selection to enhance model performance across multiple ML algorithms. Additionally, by focusing on multi-class classification and benchmarking outcomes against recent studies, this work aims to establish a reliable and adaptable framework for Parkinson's disease diagnosis.

III. Dataset Overview

The dataset utilized in this study is sourced from the PhysioNet repository [21] in 2024. It contains data from 469 individuals, each with a unique ID. The dataset includes a variety of data types, such as mobility, time series, and questionnaire responses. However, for this study, only questionnaire responses and patient data were used. The questionnaire data includes responses to Parkinson's Disease Non-Motor Symptom Questions. The dataset was stored in separate files for each individual's questionnaire replies and patient metadata. Patient records include important data such as age, height, weight, and gender. The questionnaire and patient information were combined to create a complete dataset for analysis. The dataset originally had six categories: Parkinson's Disease, Healthy Controls, Other Movement Disorders, Essential Tremor, Atypical Parkinsonism, and Multiple Sclerosis. Futhermore, the dataset was divided into two binary classification problems: Parkinson's Disease vs. Healthy Controls and Parkinson's

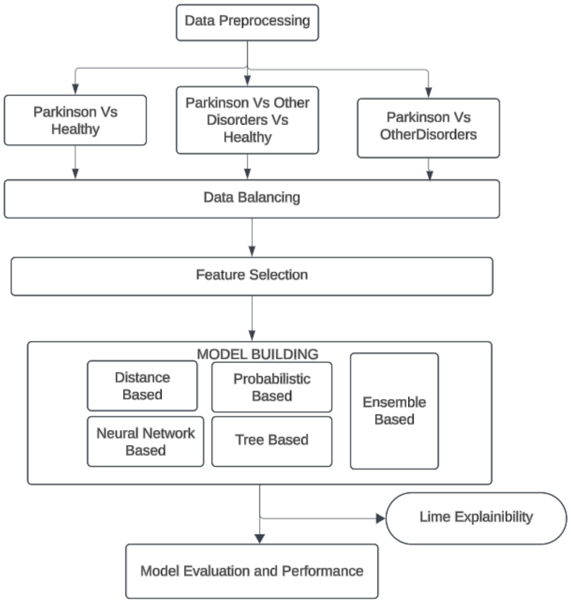

Fig. 1. Proposed architecture for classification

Disease vs. Other Disorder. In addition, a three-class problem was created: Parkinson's Disease, Healthy Controls, and Other Disorders. This approach allowed a more focused study of the conditions.

IV. Methodology

In this work, the proposed architecture represented in Fig 1, includes procedures for data collection, cleaning, visualization, feature selection, and class imbalance management. The following sub-sections provide a detailed explanation of these phases.

A. Data Preprocessing

The dataset was initially recorded in JSON format and later converted to CSV for this study. It comprised two components: (1) self-completion of an electronic questionnaire and (2) an assessment of active movement. This study focused on questionnaire data for symptom classification. The questionnaire gathered information on age, height, weight, gender, familial history of Parkinson's disease (PD), and the effect of alcohol on tremors. Additionally, 30 yes/no responses related to Parkinson's disease-specific non-motor symptoms (PDNMS) were collected using the International Parkinson and Movement Disorder Society's PDNMS questionnaire [22]. Missing and NaN values were addressed, and a data quality report provided statistical insights for each feature. The final dataset contained 469 instances and 51 features.

B. Data Balancing

As depicted in Fig. 2, a substantial class imbalance was evident across all three classification tasks. The distribution for Parkinson's and Healthy Controls was 276:79, while for Parkinson's, Other Disorders, and Healthy Controls, it was 276:114:79. Similarly, the distribution for Parkinson's versus Other Disorders was 276:114. To ensure a balanced analysis, data balancing techniques were necessary. The Synthetic Minority Over-sampling Technique (SMOTE) was utilized to address this issue by generating synthetic samples for the minority class through interpolation of existing data points. ‘

Fig. 2. Class imbalance for Three Types of Classification

C. Feature Selection

Feature selection methods, including ANOVA, Sequential Feature Selection, and Recursive Forward Elimination, were utilized alongside feature reduction techniques such as PCA, ICA, and T-SNE. Model performance was further enhanced through hyperparameter tuning. Additionally, a 10-fold cross-validation approach was implemented to ensure the robustness and reliability of the results.

D. Model Building

This research employs fifteen distinct machine learning models across three classification tasks. The models used include K-Nearest Neighbors (KNN), Naive Bayes, Logistic Regression, Decision Trees, Random Forests, Gradient Boosting, XGBoost, AdaBoost, Extra Trees, LightGBM, Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), CatBoost, Stacking, and Bagging. The dataset undergoes 10-fold cross-validation and is processed using these classifiers. The study also incorporates the Random Forest (RF) Classifier as part of the analysis. To reduce overfitting, the RF model aggregates predictions from multiple decision trees, each trained on different subsets of the data and features. By training each decision tree on a randomly selected set of features and data, the model becomes more robust and resilient to variations in the input data.

E. Model Evaluation

To assess the effectiveness of each model, various performance metrics are utilized, including accuracy, F1-score, precision, recall, confusion matrix, and the Area Under the Receiver Operating Characteristic Curve (ROC AUC). These metrics provide a comprehensive evaluation of model performance, ensuring a balanced assessment of predictive capabilities across different classification tasks.

F. Model Explainability using AI Tools

The Explainable AI model provides comprehensive explanations for its predictions, including the relevance of characteristics and decision-making process. The dataset uses Explainable AI (XAI) approaches, namely Local Interpretable Model-agnostic Explanations (LIME) and Shapely Additive Explanations (SHAP). They provide explanations for machine learning models' predictions and decisions.

V. Results and Discussion

A total of fifteen different machine learning models were evaluated for classifying the multi-class Parkinson's disease dataset. Among these, CatBoost emerged as the most effective model for distinguishing between Parkinson's disease and Healthy Controls, demonstrating superior performance across various evaluation metrics, including precision, recall, F1-score, and overall accuracy. To further enhance the performance of the CatBoost model, feature reduction techniques such as PCA and ICA were applied. Additionally, feature selection methods, including ANOVA, Sequential Forward Selection, and Recursive Feature Elimination, were utilized to refine the model and improve its classification accuracy.

1) PD Vs HC: As shown in Table I, the results of the five best-performing configurations are summarized below, along with confusion matrices to visualize the classification performance across the two classes: Parkinson's and Healthy. The number of PCA components was reduced to 5, and the number of ICA components was also limited to 5. These dimensionality reduction techniques resulted in significant feature reduction while preserving the integrity of the original features. Additionally, the evaluation metrics are provided after applying SMOTE with a sampling strategy of 0.5, which was used to address the high class imbalance in the dataset.

2) PD Vs OTHER: The results in Table II highlight the importance of class balancing and the application of the ANOVA feature selection technique in enhancing model performance by selecting the top seven features. Although Logistic regression and AdaBoost performed consistently well across all parameters, Gradient Boosting, SVM, and CatBoost emerged as the top three models when SMOTE and feature selection techniques were applied.

3) PD, HD AND OTHER: The trials in Table III shows that AdaBoost performed well on baseline conditions, with the best F1 Score and precision. However, there was a trade-off between predictions and class balancing, as all classifiers faced slight declines in performance when class imbalance was dealt with with SMOTE. Along with SMOTE, backward feature selection helped reduce this drop, which was particularly useful for simpler models such as logistic regression.

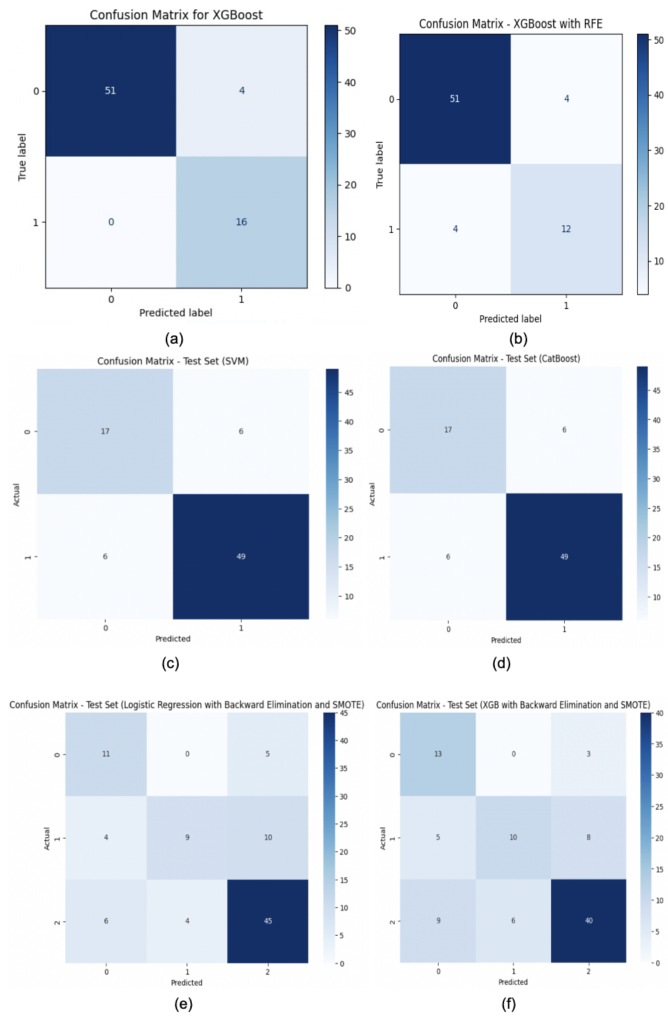

Fig. 3. (a) Confusion matrix for Model 1: XGBoost with standard evaluation. (b) Confusion matrix for Model 2: XGBoost with Recursive Feature Elimination. (c) Confusion matrix for Model 3: SVM with ANNOVA and SMOTE. (d) Confusion matrix for Model 4: CatBoost with ANNOVA and SMOTE. (e) Confusion matrix for Model 5: Logistic Regression with Backward Elimination and SMOTE. (f) Confusion matrix for Model 6: XGBoost with Backward Elimination and SMOTE.

The top-performing models for both binary and multi-class classification tasks, including Parkinson's disease (PD), healthy controls (HC), and other disorders (OTHER), are visualized in the confusion matrices shown in Figure 3. Subfigures (c) and (d) indicate that SVM and CatBoost excel in the PD vs. OTHER classification, while subfigures (a) and (b) demonstrate the strong performance of the XGBoost model in the PD vs. HC classification. Subfigures (e) and (f) show that

TABLE I
PERFORMANCE METRICS FOR DIFFERENT CLASSIFIERS - PD VS HC

	Macro Average			Weighted Average		
Classifier	**Precision**	**Recall**	**F1 Score**	**Precision**	**Recall**	**F1 Score**
Logistic Regression	0.91	0.91	0.91	0.92	0.91	0.91
SVM	0.88	0.88	0.88	0.89	0.88	0.88
CatBoost	0.91	0.91	0.91	0.92	0.91	0.91
XGBoost	0.88	0.88	0.88	0.89	0.88	0.88
Random Forest	0.91	0.91	0.91	0.92	0.91	0.91
SMOTE						
Logistic Regression	0.93	0.95	0.93	0.93	0.95	0.93
SVM	0.93	0.93	0.93	0.93	0.94	0.93
CatBoost	0.96	0.94	0.93	0.94	0.95	0.95
XGBoost	0.93	0.93	0.93	0.93	0.94	0.93
Random Forest	0.93	0.93	0.93	0.93	0.94	0.93
RFE + SMOTE						
Logistic Regression	0.88	0.95	0.91	0.95	0.93	0.93
SVM	0.89	0.93	0.93	0.94	0.93	0.91
CatBoost	0.81	0.85	0.83	0.88	0.87	0.88
XGBoost	0.86	0.95	0.89	0.94	0.92	0.92
Random Forest	0.85	0.94	0.88	0.93	0.90	0.91

TABLE II
PERFORMANCE METRICS FOR DIFFERENT CLASSIFIERS - PD VS OTHER

	Macro Average			Weighted Average		
Classifier	**Precision**	**Recall**	**F1 Score**	**Precision**	**Recall**	**F1 Score**
Logistic Regression	0.80	0.75	0.76	0.82	0.82	0.81
AdaBoost	0.80	0.75	0.76	0.82	0.82	0.81
Gradient	0.71	0.70	0.70	0.75	0.76	0.75
SVM	0.67	0.63	0.64	0.71	0.73	0.71
CatBoost	0.65	0.62	0.63	0.70	0.72	0.70
SMOTE						
Logistic Regression	0.78	0.78	0.78	0.82	0.82	0.82
AdaBoost	0.77	0.76	0.77	0.81	0.81	0.81
Gradient	0.72	0.72	0.72	0.77	0.77	0.77
SVM	0.67	0.63	0.64	0.71	0.73	0.71
CatBoost	0.71	0.68	0.69	0.74	0.76	0.75
Annova + SMOTE						
Logistic Regression	0.76	0.73	0.74	0.79	0.79	0.79
AdaBoost	0.72	0.70	0.71	0.76	0.77	0.76
Gradient	0.80	0.81	0.80	0.84	0.83	0.83
SVM	0.82	0.82	0.82	0.85	0.85	0.85
CatBoost	0.82	0.82	0.82	0.85	0.85	0.85

Logistic Regression outperforms XGBoost in the multiclass problem, with results that are more evenly distributed across the three classes.

As depicted in Fig. 3, the matrices represent model performance for each class. In Fig. 3(a), the XGBoost configuration achieved the highest performance, with precision, recall, and F1-score of 0.94 for the weighted average. The rows of the matrix represent the actual classes, while the columns correspond to the predicted classes. The diagonal elements indicate correct predictions, while off-diagonal values represent misclassifications. In Fig. 3(b), the application of Recursive Feature Elimination (RFE) marginally improved the XGBoost model's performance, as evidenced by a slight increase in correct predictions for class 1.

The accuracy of the Logistic Regression model, shown in Fig. 3(e), is 69%, with higher diagonal values reflecting strong performance in identifying the majority of classes. However, certain misclassifications, such as class "1" being mistaken for class "2," can still be observed. The XGBoost model in Fig. 3(f) shows slightly lower accuracy at 67%, with comparable patterns of correct predictions and misclassifications.

As demonstrated in Fig. 4, the feature contributions for a test instance were analyzed using the LIME technique. The LIME results provide valuable insights into the decision-making process of the model for the binary classification problem. The XGBoost classifier was used to visualize the explanations provided by LIME for the data.

Whereas in Fig 5, For a test instance, the model focuses on symptoms such as dribbling, urgency, and constipation to identify cases like Other Disorders. SVM classifier is implemented to visualize the data.

In Fig 6, the model labeled the test case as Parkinson. Taste/smelling and Nocturia predicted Other Disorders, but Dribbling, Urgency, and Constipation dropped the risk of Other Disorders, correlating more closely with Parkinson. Random Forest is the classifier used to implement data vi-

TABLE III
PERFORMANCE METRICS FOR DIFFERENT CLASSIFIERS - PD, HC AND OTHER

	Macro Average			Weighted Average		
Classifier	**Precision**	**Recall**	**F1 Score**	**Precision**	**Recall**	**F1 Score**
Logistic Regression	0.61	0.59	0.58	0.65	0.67	0.65
AdaBoost	0.67	0.65	0.63	0.72	0.69	0.69
XGBoost	0.60	0.58	0.58	0.65	0.66	0.64
CatBoost	0.68	0.63	0.63	0.71	0.71	0.69
SMOTE						
Logistic Regression	0.61	0.63	0.61	0.66	0.67	0.66
AdaBoost	0.58	0.59	0.56	0.66	0.63	0.62
XGBoost	0.63	0.63	0.61	0.64	0.64	0.63
CatBoost	0.63	0.63	0.61	0.68	0.68	0.67
Backward Elimination + SMOTE						
Logistic Regression	0.66	0.63	0.63	0.70	0.69	0.68
AdaBoost	0.60	0.60	0.56	0.68	0.64	0.63
XGBoost	0.63	0.66	0.62	0.69	0.67	0.67
CatBoost	0.64	0.62	0.59	0.69	0.67	0.67

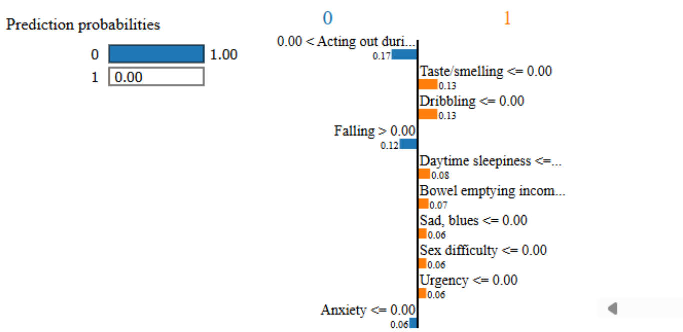

Fig. 4. Prediction probabilism used by machine learning model to classify PD vs HC

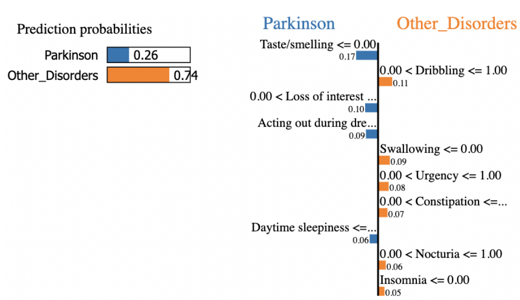

Fig. 5. Prediction probabilism used by machine learning model to classify PD vs Other Disorders

sualization.

VI. CONCLUSION

This project focused on multi-class classification to differentiate Parkinson's Disease (PD) from other movement disorders, including Healthy Controls, Essential Tremor, Atypical Parkinsonism, and Multiple Sclerosis. The primary aim was to tackle the challenge of distinguishing PD from these disorders, which share overlapping symptoms, using various machine learning algorithms. The study demonstrated that machine learning methods can significantly enhance the accuracy of PD diagnosis and effectively differentiate it from other

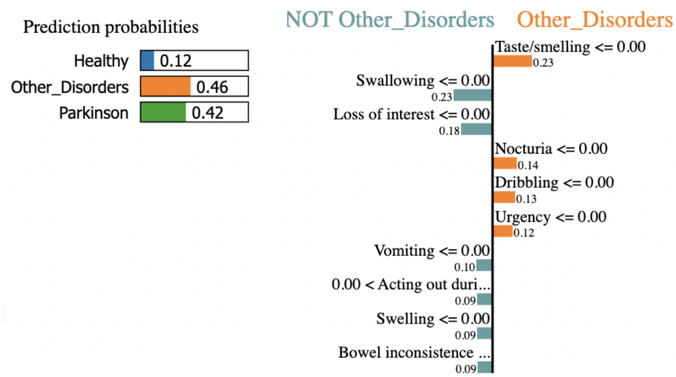

Fig. 6. Prediction probabilism used by machine learning model to classify the three classes

movement disorders. The proposed multi-class classification approach successfully addressed the issue of symptom overlap, achieving an F1-score of 93% for distinguishing PD from Healthy Controls using XGBoost, 85% for differentiating PD from other disorders using SVM, and 68% for multi-class classification with Logistic Regression. For future work, this research can be expanded by exploring the application of deep learning techniques for classification, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).

REFERENCES

[1] Z. Ou, J. Pan, S. Tang, D. Duan, D. Yu, H. Nong, and Z. Wang,"Global trends in the incidence, prevalence, and years lived with disability of Parkinson's disease in 204 countries/territories from 1990 to 2019," Frontiers in public health, vol. 9, p. 776847, 2021, doi:10.3389/fpubh.2021.776847.

[2] K. Nanda, S. Maurya and D. Gupta,"An Explainable AI Approach to Classify Smoker and Drinker Profiles via Medical Parameters," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-7, doi: 10.1109/I2CT61223.2024.10543849.

[3] R. B. Postuma, "Prodromal Parkinson disease: do we miss the signs?," Nature Reviews Neurology, vol. 15, no. 8, pp. 437-438, 2019, doi:10.1038/s41582-019-0215-z.

[4] J. Zhang, L. Wei, X. Hu, B. Xie, Y. Zhang, G.-R. Wu, and J. Wang,“Akinetic-rigid and tremor-dominant Parkinson’s disease patients show different patterns of intrinsic brain activity," Parkinsonism & Related Disorders, vol. 21, no. 1, pp. 23-30, 2015, doi:10.1016/j.parkreldis.2014.10.017.

[5] G. Rizzo, M. Copetti, S. Arcuti, D. Martino, A. Fontana, and G. Logroscino, “Accuracy of clinical diagnosis of Parkinson disease: a systematic review and meta-analysis," Neurology, vol. 86, no. 6, pp. 566-576, 2016, doi:10.1212/WNL.0000000000002350.

[6] S. Vekkot, N. N. V. Sai Prakash, T. S. Eswar Reddy, S. Reddy Sripathi, S. Lalitha, D. Gupta, M. Zakariah, and Y. A. Alotaibi, “Dementia Speech Dataset Creation and Analysis in Indic Languages—A Pilot Study," IEEE Access, vol. 11, pp. 130697-130718, 2023, doi: 10.1109/ACCESS.2023.3334790.

[7] S. Nematzadeh, F. Kiani, M. Torkamanian-Afshar, and N. Aydin, “Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases," Computational Biology and Chemistry, vol. 97, 2022, doi: 10.1016/j.compbiolchem.2021.107619.

[8] K. Duvvuri, H. Kanisettypalli, T. N. Masabattula, S. Vekkot, D. Gupta, and M. Zakariah, “Unravelling stress levels in continuous speech through optimal feature selection and deep learning," Procedia Computer Science, vol. 235, pp. 1722-1731, 2024, doi: 10.1016/j.procs.2024.04.163.

[9] A. M. Abubakar, D. Gupta, and S. Parida, “A reinforcement learning approach for intelligent conversational chatbot for enhancing mental health therapy," Procedia Computer Science, vol. 235, pp. 916-925, 2024, doi: 10.1016/j.procs.2024.04.087.

[10] S. Moradi, L. Tapak, and S. Afshar, “Identification of novel noninvasive diagnostics biomarkers in the Parkinson’s diseases and improving the disease classification using support vector machine," BioMed Research International, vol. 2022, p. 5009892, 2022, doi:10.1155/2022/5009892.

[11] P. Raundale, C. Thosar and S. Rane, “Prediction of Parkinson’s disease and severity of the disease using Machine Learning and Deep Learning algorithm," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-5, doi: 10.1109/INCET51464.2021.9456292.

[12] F. Cordella, A. Paffi and A. Pallotti, “Classification-based screening of Parkinson’s disease patients through voice signal," 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Lausanne, Switzerland, 2021, pp. 1-6, doi: 10.1109/MeMeA52024.2021.9478683.

[13] M. Belić, V. Bobić, M. Badža, N. Šolaja, M. Đurić-Jovičić, and V. S. Kostić, “Artificial intelligence for assisting diagnostics and assessment of Parkinson’s disease—A review," Clinical Neurology and Neurosurgery, vol. 184, 2019, doi: 10.1016/j.clineuro.2019.105442.

[14] A. H. Shahid and M. P. Singh, “A deep learning approach for prediction of Parkinson’s disease progression," Biomedical Engineering Letters, vol. 10, pp. 227-239, 2020, doi:10.1007/s13534-020-00156-7.

[15] A. Boutet, R. Madhavan, G. JB Elias, S. E. Joel, R. Gramer, M. Ranjan, V. Paramanandam, et al., “Predicting optimal deep brain stimulation parameters for Parkinson’s disease using functional MRI and machine learning," Nature communications, vol. 12, no. 1, p. 3043, 2021, doi:10.1038/s41467-021-23311-9.

[16] M. B. Makarious, H. L. Leonard, D. Vitale, H. Iwaki, L. Sargent, A. Dadu, I. Violich, et al., “Multi-modality machine learning predicting Parkinson’s disease," npj Parkinson’s Disease, vol. 8, no. 1, p. 35, 2022, doi:10.1038/s41531-022-00288-w.

[17] J. Harvey, R. A. Reijnders, R. Cavill, A. Duits, S. Köhler, L. Eijssen, B. P. F. Rutten, et al., “Machine learning-based prediction of cognitive outcomes in de novo Parkinson’s disease," npj Parkinson’s Disease, vol. 8, no. 1, p. 150, 2022, doi:10.1038/s41531-022-00409-5.

[18] L. Omberg, E. C. Neto, T. M. Perumal, A. Pratap, A. Tediarjo, J. Adams, B. R. Bloem, et al., “Remote smartphone monitoring of Parkinson’s disease and individual response to therapy," Nature biotechnology, vol. 40, no. 4, pp. 480-487, 2022, doi:10.1038/s41587-021-00974-9.

[19] A. Schrag, Y. Ben-Shlomo, and N. Quinn, “How valid is the clinical diagnosis of Parkinson’s disease in the community?," Journal of Neurology, Neurosurgery & Psychiatry, vol. 73, no. 5, pp. 529-534, 2002, doi:10.1136/jnnp.73.5.529.

[20] S. Priyadharshini, K. Ramkumar, S. V. Vairavasundaram, K. Narasimhan, S. Venkatesh, R. Amirtharajan, and K. Kotecha, “A comprehensive framework for Parkinson’s disease diagnosis using explainable artificial intelligence empowered machine learning techniques," Alexandria Engineering Journal, vol. 107, pp. 568-582, 2024, doi:10.1016/j.aej.2024.07.106.

[21] J. Varghese, A. Brenner, L. Plagwitz, C. van Alen, M. Fujarski, and T. Warnecke, “PADS-Parkinsons Disease Smartwatch dataset," PhysioNet, 2024, doi:10.13026/m0w9-zx22. [Online]. Available: https://www.kaggle.com/datasets/julianvarghese/pads-parkinsons-disease-smartwatch-dataset.

[22] K. R. Chaudhuri, P. Martinez-Martin, A. H. V. Schapira, F. Stocchi, K. Sethi, P. Odin, R. G. Brown, W. Koller, P. Barone, G. MacPhee, L. Kelly, M. Rabey, D. MacMahon, S. Thomas, W. Ondo, D. Rye, A. Forbes, S. Tluk, V. Dhawan, A. Bowron, A. J. Williams, and C. W. Olanow, “International multicenter pilot study of the first comprehensive self-completed nonmotor symptoms questionnaire for Parkinson’s disease: The NMSQuest study," Mov. Disord., vol. 21, pp. 916-923, 2006. [Online]. doi:10.1002/mds.20844.