

Early Detection of Parkinson's Disease Using Speech Biomarkers and Interpretable Machine Learning Models

Sai Siva Naga Chandra Pathuru¹, S. Sri Harsha²

¹MTech, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (K L E F), Vaddeswaram, Guntur, Andhra Pradesh, India

Email: 2401050035cse@gmail.com

²Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (K L E F), Vaddeswaram, Guntur, Andhra Pradesh, India

Email: sharsha@kluniversity.in

Abstract— Parkinson's disease (PD) is a progressive neurological disorder that manifests with motor dysfunctions, including bradykinesia, rigidity and tremor. Speech problems, such as dysphonia, are common and often precede PD and thus are important early indicators. The Parkinson's illness dataset from UCI repository was used for the predictive modelling, consisting of a range of vocal characteristics. Data pre-processing included missing value treatment, label encoding, normalization and KMeans-SMOTE resampling to overcome class imbalance. Recursive Feature Elimination with Logistic Regression (RFE-LR) was used to select features. Different classifiers were trained and evaluated such as Support Vector Machine (SVM), CatBoost, K-Nearest Neighbor (KNN), Logistic Regression, Random Forest (RF), XGBoost, Multilayer Perceptron (MLP), and PCA with LR. Performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrices, and the XRFILR model achieved an accuracy of 96.46%. The accuracy, precision, recall, F1-score and confusion matrices were used to assess the performance, and the XRFILR model obtained an accuracy of 96.46%. Further, an ensemble-based Voting Classifier that combined MLP with Bagging and Logistic Regression was used for further improvement of prediction performance and an astounding level of accuracy of 98% was achieved. In addition, LIME and SHAP greatly improved the model's interpretability by providing information on which features are contributing to the model, ensuring transparency in clinical decisions. Furthermore, a user interface designed using Flask was developed for processing the user-supplied inputs with the learned models for making predictions in real-time.

Keywords— *Parkinson's disease, K-MeansSMOTE, recursive feature elimination, logistic regression classification with permutation, eXplainable artificial intelligence*”.

HOW TO CITE THIS ARTICLE: SAI SIVA NAGA CHANDRA PATHURU¹ S. SRI HARSHA². EARLY DETECTION OF PARKINSON'S DISEASE USING SPEECH BIOMARKERS AND INTERPRETABLE MACHINE LEARNING MODELS. INT J DRUG DELIV TECHNOL. 2026;16(84s):583-589. DOI: 10.25258/IJDDT.16.84s.62.

I. INTRODUCTION

PD is a progressive degenerative disease that affects the brain and leads to speech, tremor, rigidity and motor dysfunction, the primary cause of which is the death of dopamine-producing cells in the brain. It is a difficult condition to diagnose in its early stages, and is believed to be the second most prevalent neurological condition after Alzheimer's disease [1]. Typical diagnostic approaches rely on imaging techniques such as SPECT and MRI. Alternative, non-invasive tests have been explored in order to detect along with imaging, including HWMA [2], gait signals [3] and biosignals such as EEG [2], EMG [3] and FoG symptoms [4].

According to the World Health Organization, there were more than 8.5 million people with PD in the world in 2019 [5]. According to epidemiological research, men are 1.5 times more likely than women to get PD, especially if they are older than 50. Approximately 60,000 new cases are found annually in the US. Dysphonia is one of the most common symptoms, occurring in around 70–90% of patients. Low voice pitch, sudden bursts, strained voice quality and long pauses are well recognized as the early warning signs [6].

AI and Machine Learning (ML) have transformed the medical diagnostics process by providing highly accurate predictions based on complex data. For pattern identification and decision assistance, these technologies have been widely used in the healthcare industry [7]. Although sampling techniques such as oversampling and undersampling are important in boosting predictive accuracy because the imbalance in medical information gives rise to classification bias [8]. Similarly, high dimensional feature sets pose

computational challenges and efficient feature selection is needed to decrease the dimensionality and boost the classification performance due to the redundancy among features [9].

Medication such as levodopa and syndopa remains the primary form of treatment for PD and dosage adjusted depending on the severity of symptoms. These drugs make the symptoms worse, but do not prevent illness from worsening [10]. Scalable, non-invasive and clinically relevant approaches for early diagnosis of PD are of increasing importance due to their ability to provide clinically useful tools, with ML-based predictive frameworks particularly exploiting easily available features, such as speech.

II. RELATED WORK

The rapid advances in AI and ML have inspired a wide range of researchers to build more complex models for early detection and prediction of PD, particularly with the help of non-invasive biomarkers such as speech and voice data. Several recent works have provided valuable techniques that can enhance the accuracy of the classifications, feature selection, data balancing, and interpretability, and have provided a solid basis for future research.

In order to improve PD prediction, an ensemble of closest neighbor boosting method was proposed by Shastry [11]. It focused on making the classifications more reliable in the presence of complex data sets by combining the best aspects of the nearest neighbor classification with boosting. This approach showed that for non-linear distributions, the adaptive weights of the weak learners could be used to reduce

misclassifications and boost the accuracy of ensemble learning.

Ali, Salim and Saeed [12] proposed a hybrid approach that uses filter feature selection along with genetic algorithm and ensemble learning. Their research was often directed to the challenges that are inherent in voice data from high dimensional feature sets, characteristic of PD voice data. The framework improved the computing efficiency while preserving the most relevant features using ensemble learning for effective classification and evolutionary algorithms for effective searching. The significance of tackling both computational complexity and prediction performance in medical datasets was brought to light by this work.

Bukhari & Ogudo [13] developed an ensemble based ML model particularly for PD detection of voice inputs. The authors used PCA for dimensionality reduction and the SMOTE for handling the class imbalance. By using these preprocessing techniques in conjunction with ensemble classifiers, their model showed excellent accuracy of 96%, precision of 98% and AUC of 0.99. They realised that it was successful, but had problems with hyperparameter tuning and reliance on small data sets to employ it more broadly.

Shyamala and Navamani [14] have introduced an interpretable feature ranking XGBoost (IFRX) model for PD prediction using speech features. They integrated some preprocessing methods, including SVM SMOTE (class balance) and RFE using XGBoost (feature selection), as well as SHAP (interpretability). A disadvantage of the study was the use of a limited data set. A strength was the ability to strike a balance between clinical transparency and predictive accuracy. The motivation for this work comes from the challenges of accurate classification and explainability in healthcare decision support systems that are needed at the same time.

According to the report by the WHO in 2024 [15], the prevalence of PD has also grown substantially since 2000 and is a growing public health issue around the world. Deaths associated with PD have more than doubled in the previous 20 years, the research says, underlining the essential need for early detection methods. The statistical evidence shows the need for ML in healthcare for early, non-invasive, accurate and large scale detection.

For the early detection of PD, Wang, Lee, Harrou, and Sun [16] investigated both DL and conventional ML techniques. They compared the performance of their deep learning models with traditional methods, and found that DL models, specifically convolutional neural networks, outperform the traditional methods in learning complex speech patterns. However, they noted that deep models would require more training time and a larger data set, rendering them impractical in some medical contexts. This study provides a clear understanding of the balance of complexity and applicability between the choice of predictive models.

Hoq, Uddin, and Park [17] proposed an AI model for PD identification, which is based on voice feature extraction. Their technique concentrated on examining minute speech differences that are frequently challenging for medical professionals to manually recognize. For non-invasive diagnostics, encouraging results were obtained from the model by extracting discriminative speech cues and adding them to ML classifiers. In this study, speech was pointed as

one of the most significant indicators for early detection of PD.

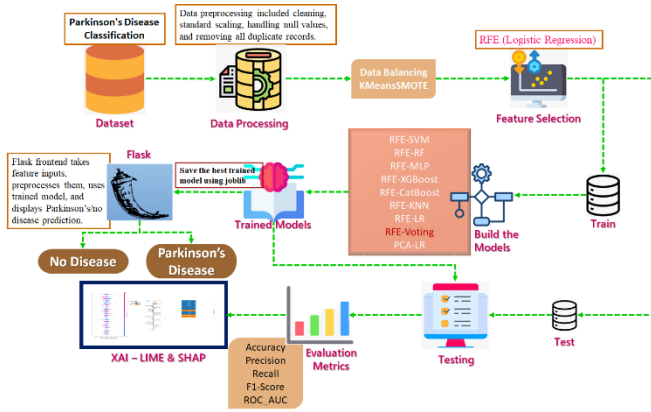
Soumaya et al., Taoufiq et al., Benayad et al., Yunus et al., and Abdelkrim et al. [18] employed a SVM classifier and a genetic algorithm to identify PD. The optimal feature subsets were obtained by using the genetic algorithm, which removed redundant features and enhanced the classification accuracy. The SVM proved to be a strong classification model, being able to handle the non-linear relationships of the data. While cost of computation was known to be an issue, their approach demonstrated the strength of evolutionary algorithms in combination with classic ML models.

Liu et al. [19] presented local Discriminant Preservation Projection (LDPP), which is based on ensemble learning and is a dimensionality reduction method. They had developed a method that would preserve the most informative parts of speech data by incorporating ensemble learning into dimensionality reduction. The researchers made better use of LDPP in speech sounds that are associated with Parkinson's disease, to improve the efficiency of computing and classification accuracy. In this study, the problem of managing a large number of voices and preventing information loss in the preprocessing stage was addressed.

Finally, Goyal, Khandnor and Aseri [20] conducted a comprehensive study and comparison of several ML classifiers used for the categorization of PD using dysphonia. Their study evaluated classifiers like RF, DT, SVM and Neural Networks using voice datasets to show the pros and cons of every type of classifier. The results indicate that the DL and ensemble techniques outperformed simpler methods with respect to accuracy and resilience. This comparison method is useful for the researchers to select a suitable algorithm depending on the size, complexity and therapeutic requirements of their datasets.

III. MATERIALS AND METHODS

The proposed method is based on speech analysis, which is considered a frontline diagnostic method for early detection of PD. Comprehensive voice characteristics of both healthy and afflicted individuals from the UCI repository of Parkinson's disease data are used. The pre-processing procedures in the pipeline first resolve the missing values, then encode the categorical variables, standardize the numerical attributes, and apply class balancing technique KMeans-SMOTE. The foundation of the XRFILR architecture is RFE-LR, which is used for feature selection. ML methods taught and tested include SVM, CatBoost, KNN, LR, RF, XGBoost, MLP, and PCA with LR. To boost the predictive performance an enhanced VC is presented by using MLP and Bagging along with LR. Furthermore, a Flask-based interface is developed for real-time delivery of the features and for their prediction, and XAI techniques such as LIME and SHAP are used to ensure prediction interpretability.



"Fig.1 System Architecture"

Description

A) Dataset Collection:

The UCI ML Repository is the source of the Parkinson's Dataset, which has 756 entries with 754 features, including 5 integer columns and 749 numerical attributes. The speech and vocal characteristics of people classified as either Parkinson's disease or healthy are provided in the dataset. Duplicate records, missing values, and removing the unique identifier put the data set into a state where it could be analyzed. The classes are imbalanced, and there are 192 Parkinson's instances and 563 healthy instances.

	gender	PPE	DFA	RPDE	numPulses	numPeriodsPulses	meanPeriodPulses	stdDevPe
0	1	0.85247	0.71826	0.57227	240	239	0.008064	
1	1	0.76686	0.69481	0.53966	234	233	0.008258	
2	1	0.85083	0.67604	0.58982	232	231	0.008340	
3	0	0.41121	0.79672	0.59257	178	177	0.010858	
4	0	0.32790	0.79782	0.53028	236	235	0.008162	

5 rows × 754 columns

"Fig.2 Parkinson's Dataset"

B) EDA:

Data visualization and EDA were conducted prior to building the model to better understand the data. It was determined that all the data types, null values, possible duplicates and dimensions of the datasets were appropriate during the initial inspection and cleaned accordingly. The difference between the Parkinson's and the healthy patients was demonstrated using a pie chart of classes which was then balanced with KMeansSMOTE. Variance plots and correlation heatmaps were used to show feature selection findings, making it possible to identify the most pertinent features. For reliable model training, these EDA procedures guaranty cleaner, balanced, and well-structured data.

C) Data Processing:

Initially, the data set is verified for duplicate data and missing values for data integrity. `drop_duplicates()` will drop any duplicates. `dropna` is the method that will drop any missing values. This data set has no categorical columns; however, they are recognized and validated. `StandardScaler` is used to scale the numerical features to the same scale as the other features, which have a mean of 0 and a standard deviation of 1. This pre-processing helps the ML algorithms to converge faster and to improve the performance of ML models and ensures all features have the same weightage during the model training.

D) Class Balancing:

The data is suffering from class imbalance problem as the healthy data points are more than the Parkinson's disease data

points. The KMeans-SMOTE approach addresses this problem by using oversampling to generate synthetic samples of the minority class and clustering. This balanced dataset is used to minimize the bias during the training period in the model that leads to higher accuracy in the diagnosis of the Parkinson's disease by the classifier. For predictive models we resample the data, and present a bar chart to validate that the data is evenly distributed (balanced) between healthy and impacted samples.

E) Feature Selection:

The most important features of the scaled dataset are selected using RFE with Logistic Regression. In RFE, the least significant features are ranked according to their contribution to the prediction performance of the model, and are deleted one by one. In this case, the 30 optimal features were those most relevant voice biomarkers for identifying Parkinson's disease. This selection helps to preserve the maximum classification capacity of the original data set, minimizes dimensionality, computing costs and makes the model easy to read.

F) Training and Testing:

The selected and scaled features are divided into 80% training and 20% testing sets and then the model performance is tested. Several ML models are trained using the training data and the performance of the models in predicting the test data is analyzed using the metrics: accuracy, precision, recall, F1-score, and AUC. The evaluation method is based on the computation of confusion matrices and the storing of the results on performances for comparison. This helps in the selection of the most reliable classifier to predict Parkinson's disease and assures model validation.

G) Algorithms:

SVM (SVM): SVM can work with high dimensional and non-linear data, which allows for more accurate diagnosis, reduces misclassified instances in the diagnosis of speech-based Parkinson's disease and also categorizes vocal features by finding the optimal hyper plane between the classes of healthy and Parkinson's patients.

$$\text{minimize } \frac{1}{2} ||W||^2 + C \sum_{i=1}^n \xi_i \quad (1)$$

CatBoost: CatBoost knows how to model complex feature interactions and categorical features very well, and has a way to constantly learn from the misclassified data to improve the classification accuracy. It has showed potential for medical applications such as voice characteristics for identifying Parkinson's illness because of its interpretability and strength.

$$F(x) = F_0(x) + \sum_{m=1}^M \sum_{i=1}^N f_m(x_i) \quad (2)$$

KNN (KNN): KNN compares patterns of the speech of sick people with healthy people and classifies the pattern based on how close it is to the existing patterns. Because of its neighborhood based, non-parametric approach, forecasts are easy to understand and explain and do not need to be complicated to be trained.

$$\text{distance}(x, X_i) = \sqrt{\sum_{j=1}^d (x_j - X_{ij})^2} \quad (3)$$

Logistic Regression: LR is a model that predicts the probability of belonging to one of the classes and is easy to interpret for identification of Parkinson's disease. In

conjunction with feature selection, it discovers important predictors, which aids in transparency and well-informed clinical decision making.

$$P(y = 1 | X) = \frac{1}{1 + e^{-(W^T x + b)}} \quad (4)$$

Random Forest (RF): The RF is a combination of several decision trees to take account of multiple decision trees' complex relations for reliable classification of audio features. By identifying important biomarkers, its feature importance ranking enhances interpretability and boosts confidence in Parkinson's disease diagnostic results.

$$Gini = 1 - \sum_{i=1}^C (P_i)^2 \quad (5)$$

XGBoost: XGBoost is a Gradient Boosting algorithm that has very good ability to learn tree, and can capture the interaction between the speech features to a certain extent. Regularization can reduce overfitting to produce more robust and generalizable predictions that can be employed in challenging medical speech classification tasks.

$$\hat{y}_i = \sigma \left(\sum_{k=1}^K f_k(x_i) \right), f_k \in F \quad (6)$$

Multilayer Perceptron (MLP): Non-linear correlations in vocal characteristics are learned by MLP neural networks, thus enabling automatic learning of subtle speech patterns that are associated with PD. Its DL capabilities allow for more accurate and early diagnosis of complex voice-based biomarkers.

$$\hat{y} = f(W^L f(W^{L-1} \dots f(W^1 X + b^1) + b^{(L-1)}) + b^L) \quad (7)$$

PCA with Logistic Regression: The high dimensional vocal characteristics are reduced to lower dimensional using PCA and then classified using logistic regression. This combination not only decreases noise but also increases the computational efficiency and keeps the important biomarkers without loss, which makes it more effective in predicting Parkinson's disease.

Voting Classifier (MLP + Bagging with Logistic Regression): The Voting Classifier has an advantage over the other classifiers due to the incorporation of the advantage of DL and ensemble learning (ensemble learning LT with Bagging and MLP). Based on voice characteristics, it ensures accurate and comprehensible classification of Parkinson's disease, increases accuracy and stabilizes forecasts.

$$\hat{y} = \operatorname{argmax}_c \left(\sum_{i=1}^n II(\hat{y}_i = c) \right) \quad (8)$$

H) Integration of XAI and Flask Framework:

When coupled with the XAI classification system, the Flask framework aids deployment and understanding of the system. The application of Explainable AI techniques, including SHAP and LIME, illustrates how these approaches can be used to evaluate the performance of the hybrid model in the context of a Parkinson's disease diagnosis and offers insights into the relative importance of features and the decision-making process. LIME generates local explanations with visualizations like waterfall plots, whereas SHAP highlights the importance of global features with summary charts. Furthermore, the Flask framework guarantees the model's practical utility and trustworthiness, enabling seamless deployment of the trained model and handling of

input data and their predictions with a user-friendly web interface for processing, scaling and delivery.

IV. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a test is its ability to correctly identify healthy people as healthy and patients as patients. The proportion of true positive and true negative for each tested case should be calculated, in order to estimate a test's accuracy. This can be expressed mathematically as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} (1)$$

Precision: Precision is the percentage of samples or incidents that are correctly classified as positive. Hence, the following formula can be used to find out the precision:

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} (2)$$

Recall: Recall measures, in ML, a model's ability to identify all relevant observations of a certain class. It indicates the accuracy at which a model predicts positive observations from the population and is the proportion of positive observations predicted correctly to all positive observations in the population.

$$Recall = \frac{TP}{TP + FN} (3)$$

F1-Score: The F1 score is a ML metric used for evaluating the accuracy of a model. Combines a model's Recall and Precision. The accuracy data is the number of times a model predicted an entire data set correctly.

$$F1 \text{ Score} = 2 * \frac{Recall * Precision}{Recall + Precision} * 100(1)$$

AUC-ROC Curve: One of the metrics used for classification problems for various threshold values is AUC-ROC Curve. ROC is the graph of the True Positive Rate against the False Positive Rate. AUC is the overall ability of the model to distinguish between the classes; the higher the value of AUC, the more accurate the model's functioning.

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \quad (5)$$

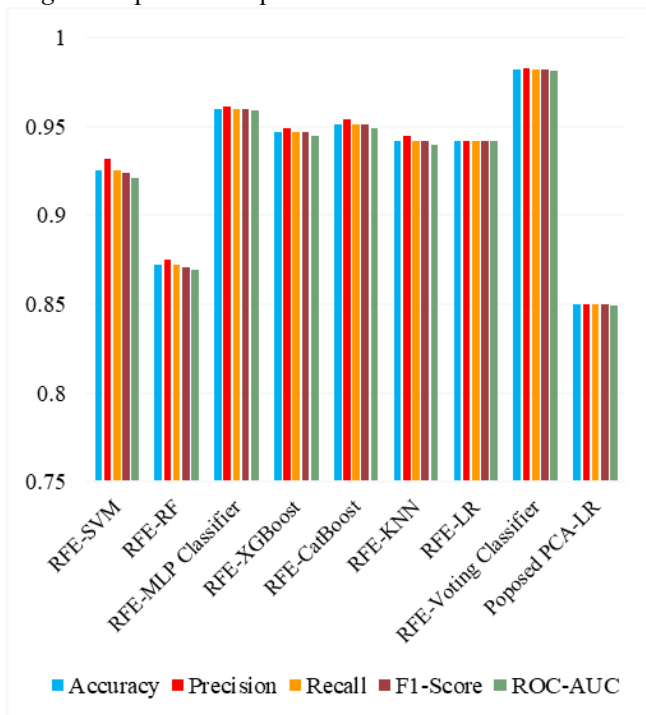
“Table.1 Performance Evaluation Table”

ML Model	Accuracy	Precision	Recall	F1-Score	ROC - AUC
RFE-SVM	0.925	0.932	0.925	0.924	0.921
RFE-RF	0.872	0.875	0.872	0.871	0.869
RFE-MLP Classifier	0.960	0.961	0.960	0.960	0.959
RFE-XGBoost	0.947	0.949	0.947	0.947	0.945
RFE-CatBoost	0.951	0.954	0.951	0.951	0.949
RFE-KNN	0.942	0.945	0.942	0.942	0.940

RFE-LR	0.942	0.942	0.942	0.942	0.942
RFE-Voting Classifier	0.982	0.983	0.982	0.982	0.981
Proposed PCA-LR	0.850	0.850	0.850	0.850	0.849

To compare several classifiers, the accuracy, precision, recall, F1-score and ROC-AUC metrics are used, as shown in Table (1). The RFE-based Voting Classifier, compares favourably with all other classifiers in detecting Parkinson's disease.

“Fig.3 Comparison Graph”



Several classifiers are compared in terms of their performance on five measures as illustrated in Fig. 3. The colour red indicates the accuracy; the orange indicates the precision; the brown indicates the recall; the green indicates F1 score; and the blue indicates ROC-AUC. The RFE-Voting Classifier consistently achieves the best score on all measures and is often the winner of the competitions.

-0.46639

0.0085182

0.32867

3.1541

4.2105

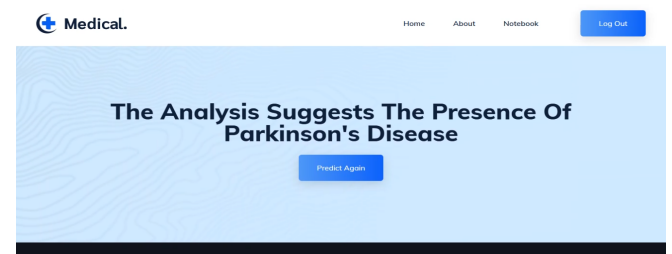
4.4625

Submit Now →

↑

“Fig.4 Upload Input Data”

In Fig. 4. Users provide numerical values to the model as features to make predictions. The system "processes" these inputs and outputs diagnostic data for the prognosis of Parkinson's disease after processing them through the trained model.



“Fig.5 Predicted Results – Parkinson’s disease”

In Fig. 5. After prediction, the output of the system is displayed, and the analysis result indicates that the Parkinson's disease exists.

-0.65047

0.016816

0.57497

2.115

4.805

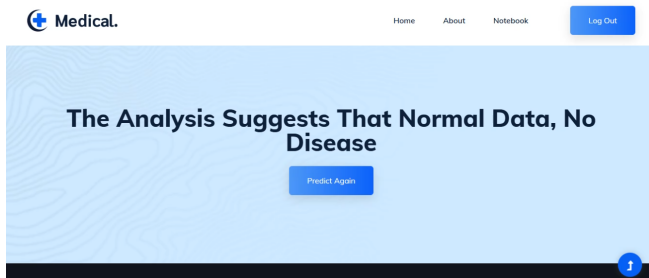
5.8098

Submit Now →

↑

“Fig.6 Upload another Input Data”

In Fig. 6. The numerical values that the user provides are input features for prediction. These variables are fed into the trained model, and the prognosis of Parkinson's disease is obtained as diagnostic results.



“Fig.7 Predicted Results – No Disease”

In Fig. 7. After prediction, the system outputs and the result of analysis is displayed that there is no Parkinson's disease.

V. CONCLUSION

The application of voice biomarkers in an early prediction system for PD is a reliable and accurate way of early detection of this progressive neurological disease. Using the UCI repository speech dataset of Parkinson's disease, it is used to preprocess the speech characteristics in a effective way to handle missing values, label encoding, standardization and class imbalance reduction by applying KMeans-SMOTE resampling. To ensure that the most pertinent voice characteristics were used in model training, RFE-LR was used for feature selection. For evaluating predictive capability, several ML algorithms were trained such as SVM, CatBoost, KNN, LR, RF, XGBoost, MLP, and PCA with Logistic Regression. The XRFE-LR model with the integration of RFE-LR and XAI was validated with a confusion matrix and achieved an accuracy of 96.46%. An advanced Voting Classifier model with the combination of MLP and Bagging with Logistic Regression showed an accuracy of 98% with a confusion matrix. LIME and SHAP provided interpretability and clinical transparency, showcasing the system's effectiveness, reliability, and ability to facilitate well-informed clinical decision-making in early-stage PD identification, while the Flask-based interface enabled real-time predictions.

In order to enhance generalization across populations, this system's future scope includes adding varied and multilingual speech samples to the dataset. Real-time voice recording and analysis can be used to enhance usability in home-based and clinical environments. More complex DL models, such as convolutional or recurrent neural networks, could be explored to more accurately represent temporal speech patterns. Multimodal data (handwriting, gait, genetic information) can be used to improve the accuracy of the diagnosis. A framework deployed on mobile devices, supported by the cloud, would also be more accessible and scalable, and would help healthcare professionals with ongoing monitoring and early detection.

REFERENCES

- [1] Shyamala, K., & Navamani, T. M. (2025). Design of an Optimized Feature Driven Severity Stage Classifier for Parkinson's disease Prediction using DL. *IEEE Access*.
- [2] Kansizoglou, I., Tsintotas, K. A., Bratanov, D., & Gasteratos, A. (2025). Drawing-aware Parkinson's disease detection through hierarchical DL models. *IEEE Access*.
- [3] Khan, M., Moiz, A., Khan, G. N., Wajid, M., Usman, M., & Ali, J. (2025). An FPGA Prototype for Parkinson's disease Detection using ML on Voice Signal. *IEEE Access*.
- [4] Park, H., Youm, C., Cheon, S. M., Kim, B., Choi, H., Hwang, J., & Kim, M. (2025). Using ML to identify Parkinson's disease severity subtypes with multimodal data. *Journal of NeuroEngineering and Rehabilitation*, 22(1), 126.
- [5] Siva, O. V., Prasanna, K. L. N., Sravanthi, V., & Anitha, V. (2025, April). A DL Approach for Predicting Parkinson's Disease Progression. In *2025 International Conference on Inventive Computation Technologies (ICICT)* (pp. 403-409). IEEE.
- [6] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, “A DL approach for Parkinson's disease diagnosis from EEG signals,” *Neural Comput. Appl.*, vol. 32, no. 15, pp. 10927–10933, Aug. 2020.
- [7] C. Loconsole, G. D. Cascarano, A. Brunetti, G. F. Trotta, G. Losavio, V. Bevilacqua, and E. Di Sciascio, “A model-free technique based on computer vision and sEMG for classification in Parkinson's disease by using computer-assisted handwriting analysis,” *Pattern Recognit. Lett.*, vol. 121, pp. 28–36, Apr. 2019.
- [8] Ö. F. Ertuğrul, Y. Kaya, R. Tekin, and M. N. Almalı, “Detection of Parkinson's disease by shifted one dimensional local binary patterns from gait,” *Expert Syst. Appl.*, vol. 56, pp. 156–163, Sep. 2016.
- [9] A. Ma, K. K. Lau, and D. Thyagarajan, “Voice changes in Parkinson's disease: What are they telling us?” *J. Clin. Neurosci.*, vol. 72, pp. 1–7, Feb. 2020.
- [10] S. Saravanan, K. Ramkumar, K. Adalarasu, V. Sivanandam, S. R. Kumar, S. Stalin, and R. Amirtharajan, “A systematic review of artificial intelligence (AI) based approaches for the diagnosis of Parkinson's disease,” *Arch. Comput. Methods Eng.*, vol. 29, no. 6, pp. 3639–3653, Oct. 2022.
- [11] K. A. Shastri, “An ensemble nearest neighbor boosting technique for prediction of Parkinson's disease,” *Healthcare Anal.*, vol. 3, Nov. 2023, Art. no. 100181.
- [12] A. M. Ali, F. Salim, and F. Saeed, “Parkinson's disease detection using filter feature selection and a genetic algorithm with ensemble learning,” *Diagnostics*, vol. 13, no. 17, p. 2816, Aug. 2023.
- [13] S. N. H. Bukhari and K. A. Ogudo, “Ensemble ML approach for Parkinson's disease detection using speech signals,” *Mathematics*, vol. 12, no. 10, p. 1575, May 2024.
- [14] K. Shyamala and T. M. Navamani, “Design of an efficient prediction model for early Parkinson's disease diagnosis,” *IEEE Access*, vol. 12, pp. 137295–137309, 2024.
- [15] World Health Org. (2024). Parkinson Disease. Accessed: Mar. 7, 2024. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/parkinson-disease>
- [16] W. Wang, J. Lee, F. Harrou, and Y. Sun, “Early detection of Parkinson's disease using DL and ML,” *IEEE Access*, vol. 8, pp. 147635–147646, 2020.
- [17] M. Hoq, M. N. Uddin, and S.-B. Park, “Vocal feature extraction-based artificial intelligent model for Parkinson's disease detection,” *Diagnostics*, vol. 11, no. 6, p. 1076, Jun. 2021.

- [18] Z. Soumaya, B. D. Taoufiq, N. Benayad, K. Yunus, and A. Abdelkrim, "The detection of Parkinson disease using the genetic algorithm and SVM classifier," *Appl. Acoust.*, vol. 171, Jan. 2021, Art. no. 107528.
- [19] Y. Liu, Y. Li, X. Tan, P. Wang, and Y. Zhang, "Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of Parkinson's disease," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102165.
- [20] J. Goyal, P. Khandnor, and T. C. Aseri, "A comparative analysis of ML classifiers for dysphonia-based classification of Parkinson's disease," *Int. J. Data Sci. Analytics*, vol. 11, no. 1, pp. 69–83, Jan. 2021.