International Journal of Multidisciplinary Research and Growth Evaluation.

Feature-Based Voice Analysis for Parkinson's Prediction Using ML Classifiers

Shaik Zubair Ahmed [1*], Mohammed Uzair Khan [2]
[1-2] Student of CSE, Neil Gogte Institute of Technology, Osmania University, Hyderabad, India

* Corresponding Author: **Shaik Zubair Ahmed**

Article Info

ISSN (online): 2582-7138
Volume: 06
Issue: 01
January-February 2025
Received: 06-11-2024
Accepted: 07-12-2024
Page No: 1597-1605

Abstract
The possibility of disease called Parkinson's pre diagnosis allows radical change in course of this sickness by introducing, in a timely manner, adequate therapies. Using machine learning algorithms, the presented research improves initial screening of Parkinson's disease on the examination of specific vocal characteristics. This ML technique applies by using a very large dataset from UC Irvine ML Repository with 197 distinct cases and 22 unique attributes. The accuracy of the KNN classifier turns out to be very accurate with an accuracy of 85%. Other than the KNN classifier, this study will look into other ML algorithms like Support Vector Machine (SVM), Random Forest Classifier, Decision Tree Classifier, and Extra Trees Classifier. Preprocessing steps like SMOTE remove redundant features and further balance the classes to improve the performance of the classifier. LIME analysis around the critical findings sh ows that vocal characteristics such as Spread2, RPDE, and MDVP (Hz) are of utmost importance in predicting Parkinson's disease. These results are extremely important for preliminary diagnosis of disease because it can totally change patient care and afford possibilities for more special and effective treatment options. This might also be through the use of voice analysis tools by patients themselves, possibly at home, feeding data to the phone-based system for the ML algorithms mapping disease course or response to treatment. Such really is the essence of this research—that it truly typifies the latest machine learning methods for forecasting Parkinson's disease and hence opening the way toward early therapeutic interventions for improved patient health outcomes.

DOI: https://doi.org/10.54660/.IJMRGE.2025.6.1.1597-1605

Keywords: Machine Learning; Parkinson disease, KNN Classifier, Vocal Characteristics, LIME Analysis, Data Preprocessing

1. Introduction

A dangerous neurological disorder that mostly affects the elderly is Parkinson's disease (PD). It is a neuro concern among degenerative diseases as its frequency approaches that of Alzheimer's. PD often affects people over 60 years, and usually, it presents with various symptoms, including tremors, rigidity, dysphonia, delayed movement (bradykinesia), and balance problems [1]. Interestingly, the initial manifestations of the disease are voice anomalies in many cases. Recently, as a non-invasive, early diagnosis method, sound analysis in Parkinson's disease diagnosis has become more popular. This method is especially of benefit because vocal recording tests may easily be conducted [2].

However, diagnosis of Parkinson's disease is complex because patients' symptoms may be so varied. Diagnosis relies on the identification of motor symptoms, including bradykinesia, resting tremor, and stiffness [3]. These may be subtle at the beginning and can overlap with other diseases, making early and accurate diagnosis difficult. Moreover, non-motor symptoms, often underestimated as belonging to Parkinson's disease, such as mood disorders, sleep disturbances, and cognitive dysfunction, may precede motor symptoms [4]. Variations in the course of the disease also exist; some patients rapidly decline, while in others, the course is milder. This heterogeneity increases the difficulty of developing a homogeneous scheme of PD diagnosis. PD also can't be diagnosed by a single test; rather, a history taking and physical exam are used in combination with imaging and laboratory testing to rule out other conditions.

Recent developments in artificial intelligence have shown promise in supporting the diagnosis of parkinsonism through analysing

patterns in speech, activity, and writing skills. The huge variation in disease presentations makes it very important to provide data sets that are comprehensive and diverse for training machine learning models to ensure high efficiency. Because it reduces the use of diagnostic instruments, coping with a lack of diversity in PD studies can yield more difficulties. Based on the findings, quite a lot of patients present audible abnormalities even in early stages of PD [5, 6, 7]. Consequently, early detection by using voice data for PD detection is non-invasive and reasonably priced. In this paper, speech data is analyzed using machine learning algorithms so that Parkinson's disease can be detected early. This study's objective is to construct robust and interpretable models that can discriminate people with PD from healthy individuals with high accuracy using a dataset from UC Irvine Machine Learning Repository.

Proposed method extracts speech data analysis through machine learning algorithms for initial screening of Parkinson disease. To begin with, a large data set was extracted from the ML Repository at UC Irvine. Recording of voices came under 22 features of that data set in which 197 instances were considered. In data pre-processing numerous steps were taking place to ameliorate model performance. This included feature scaling for best results, removal of redundant features, and handling class imbalance using the Synthetic Minority Over-sampling Technique (SMOTE). The dataset was then divided into a train set and a test set using an 80-20 split. To determine the best machine learning model for the detection of PD, several models were trained and compared. Among the selected ML models are AdaBoost, K-Nearest Neighbors (KNN), Gaussian Naive Bayes (GNB), Random Forest (RF), Random Forest (SVM), Random Forest (RF), Gradient Boosting (GB), XGBoost, and Logistic Regression (LR). To enhance model performance, hyperparameters tuning was performed. In addition, model interpretation methods like LIME also applied for more explainable and trustable.

The following are the study's main contributions:

- Using speech data, powerful machine learning models were developed that achieve as high as 85% in accuracy for early Parkinson's disease detection.
- Critical voice biomarkers were discovered which are needed to classify Parkinson's patients vs. healthy subjects.
- LIME and permutation importance approaches were utilized with a view to ensure the predictability and interpretability of the model to facilitate clinical use.

2. Literature Review

In the recent years, machine learning (ML) has witnessed tremendous development in the diagnosis of Parkinson's disease (PD) by voice analysis. To improve the accuracy and reliability of PD diagnosis, various machine-learning schemes and diverse datasets have been suggested. Little *et al*. [8] quantified dysphonia on the voice recordings of 31 individuals with vowel sound "a" to investigate Parkinson's disease. They showed the voice analysis potential for PD diagnosis in their work by the classification of extracted dysphonia traits from these recordings by an SVM model. Their study was, however, limited by the short dataset size and the need for further validation. Ali *et al*. [9] worked with advanced fusion of L1 regularized SVM and Deep Neural Network (DNN) methods on datasets provided by Max Little and Sarkar. Their approach showed very good accuracy, at 97.5% with k-fold cross-validation and 100% accuracy with Leave-One-Subject-Out validation—stunning results. Although the study by had high accuracy, it too had some shortcomings: incomplete patient records and differences in the length and severity of disease.

Utilizing data from figshare and the Irvine ML repository, Neto *et al*. [10] used algorithms like DNN, RF, GB, and SVM as well as ensemble methods like the Ensemble Stacked Model (ESM) and Ensemble Voting Model (EVM). The accuracy rates for DNN and ESM were 72.1% and 84.49%, respectively, hence validating the need for fusing multiple data sources and using diverse datasets for robust PD identification. In the work of Lizbeth *et al*. [11], Bayesian analysis with Gibb's sampling was applied for duplicated voice recordings and obtained an accuracy of 86.2%. The study did contain some flaws, including not enough validation and thin comprehensive data, even though it had responded well to intra-subject variability and improved interpretability of the results.

With an accuracy of 95.89%, Pahuja*et al*. [12] used various machine learning techniques, including multilayer perceptron (MLP), SVM, K-NN, and ANN, to analyze recorded speech signals from the MA and NCVS datasets. The study had problems such as class imbalance and noise in the dataset, although it was good at the classification of PD. Ali *et al*. [13] used GA, NN and LDA to speech sample recordings by Sarkar *et al*. Although they had the disadvantages of an unbalanced testing database and no details about feature extraction, they could attain 100% accuracy on the testing database and 95% accuracy on the training data.

Based on Parkinson's Progression Markers Initiative dataset, Wang *et al*. [14] implemented DNNs to score an accuracy rate of 96.45%. Despite some challenges such as deep learning's black-box characteristics and independent dataset verification, it brought a new approach to the study focused on early diagnosis. Gunduz*et al*. [15], they used UCI ML repository and applied CNN, and they achieved an accuracy equal to 86.9%. This ". Though the study has used deep features, it showed the promise of using multi-modal data despite the constraints in exploring deep learning models and only focusing on one type of feature.

Nizamuddin *et al*.[16] proposes an advanced AI model combining Multi-Layer Perceptron and Convolutional Neural Network to enhance prediction accuracy, framework emphasizes real-time adaptability, personalized healthcare interventions, and reduces false positives, setting a benchmark for AI applications in health monitoring and predictive care.

These results indicate the promising potential of using machine learning methods in speech analysis for PD identification. They also demonstrated the need for larger and more diverse datasets, and the application of advanced techniques for model interpretation to deal with potential challenges related to data biases and class imbalances. The proposed research methodology is presented in Figure 1. This study, therefore, goes further than previous work to increase the precision and dependability of PD diagnosis by using a full range of machine learning algorithms combined with stringent data pre-processing, feature engineering, and sophisticated interpretability methodologies.

3. Methodology

This is also a body paragraph at the first level. Font Size: 10, Font Type: Palatino Linotype Alignment: Appropriate; no spacing; no indentation with special (first line 0.75 cm). The source of Fig. 1 is [1].

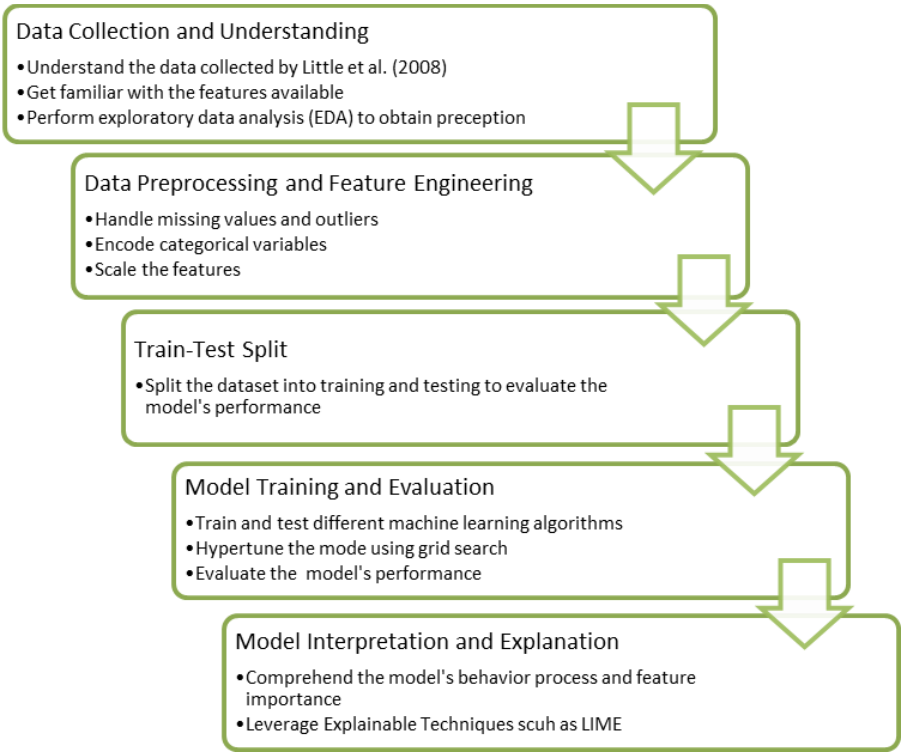

Fig 1: Suggested Approach.

3.1. Dataset Description

The multivariate dataset, taken from the "UC Irvine ML Repository [8]," contains 197 instances with 22 features. It contains several biological voice measurements and data for 31 people, 23 of which suffered from parkinsonism. 195 vocal recordings of the individuals in row of pool of data and is labelled in the "name" column. Each column in the dataset represents a unique speech metric. In the binary-coded "status" column, 0 represents a healthy subject, while 1 represents a person suffering from parkinsonism. Main purpose of the collection of data is make it easier to distinguish between people who are well and those who have the illness.

Table 1: Names and descriptions of features.

Feature Name	Description
name	The name or identifier of the individual
MDVP:Fo(Hz)	Average vocal fundamental frequency
MDVP:Fhi(Hz)	Maximum vocal fundamental frequency
MDVP:Flo(Hz)	Minimum vocal fundamental frequency
MDVP:Jitter(%)	variation in the duration of the voice cycles, measured as a percentage
MDVP: Jitter(Abs)	variation in the duration of the voice cycles, measured in absolute values
MDVP: RAP	Relative Amplitude Perturbation, a measure of the variation in the amplitude of the voice
MDVP: PPQ	Pitch Period Perturbation Quotient, a measure of the variation in the amplitude of the voice
Jitter: DDP	Average Absolute Difference of Differences, a measure of the variation in the duration of the voice cycles
MDVP: Shimmer	Shimmer (local). The variation in the amplitude of the voice, related to the roughness of the voice
MDVP: Shimmer(dB)	Shimmer (decibels). The variation in the amplitude of the voice, measured in decibels
Shimmer: APQ3	Shimmer (three-point amplitude perturbation quotient), admeasure of the variation in the amplitude of the voice
Shimmer: APQ5	Shimmer (five-point amplitude perturbation quotient), a measure of the variation in the amplitude of the voice
MDVP: APQ	Shimmer (amplitude perturbation quotient), a measure of the variation in the amplitude of the voice
Shimmer: DDA	Shimmer (dB-by-delta amplitude), a measure of the variation in the amplitude of the voice
NHR	Noise-to-Harmonics Ratio, a measure of the amount of noise in the voice signal
HNR	Harmonics-to-Noise Ratio, a measure of the amount of noise in the voice signal
status	Health status of the patient where a 0 = subject is healthy and 1 = subject has Parkinson's disease
RPDE	Recurrence Period Density Entropy, a measure of the complexity of the voice signal
DFA	Detrended Fluctuation Analysis, a measure of the long-range dependence in the voice signal
spread1	The first spectral moment, statistical measure of the voice signal's frequency spread
spread2	The second spectral moment, a statistical measure of the voice signal'sfrequency spread
D2	A nonlinear dynamic parameter, statistical measure of the complexity of the voice signal
PPE	Pitch Period Entropy, a measure of the variation in fundamental frequency

3.2. Data pre-processing

Data preprocessing are among the key steps toward developing machine learning models. Therefore, extensive cleaning of the data is necessary to ensure its reliability by handling any missing values and outliers. It applies the SMOTE to the training data so that synthetic samples are created in the minority classes to balance both classes.

Pearsons correlation coefficient was applied to identify features that exhibit high correlations in order to achieve model efficiency while minimizing computational load. A

highly correlated pattern is evident from the correlation matrix as shown in Figure 2. This underlines the fact that dimensionality reduction is a must to avoid over-fitting. High correlation was defined as having a Pearson co-efficient greater than 0.8. Fig 3. shows matrix after removing the highly linked characteristics.

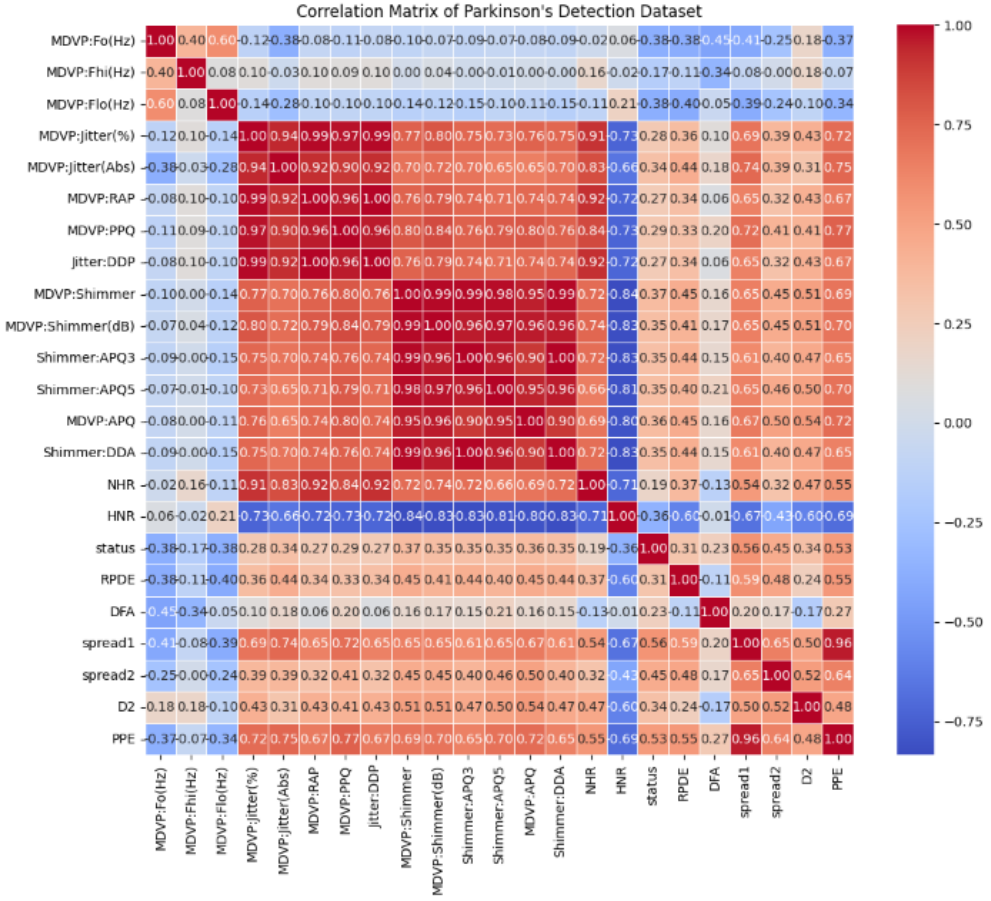

Fig 2: Following removal of highly correlated characteristics, the correlation matrix is shown

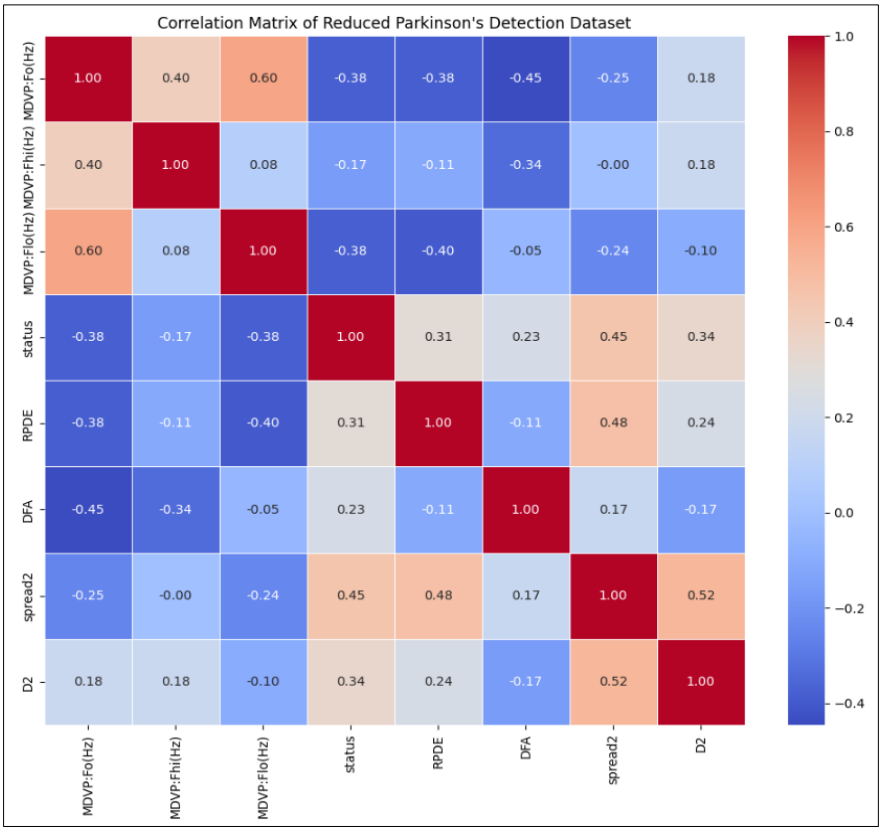

Fig 3: Following the removal of highly linked characteristics, the correlation matrix is shown

3.3. Model Training

The data set is split in an 80:20 ratio, with 80% going into training and the rest for testing. It is, therefore, essential to optimize their performance in the development of the ML model since the hyperparameter tuning optimizes model performance of ML models. Grid search is applied to help the model find patterns in the data, which it then uses in making anticipations on data not seen before.

Those are the models:

1) LR: In LR, there is a binary dependent variable that is modeled by a logistic function. One of the statistical techniques used in estimating the probability of different experimental results is LR [17]. It assumes a linear relationship between the features and the log-odds of the result, hence providing a simple and interpretable way of detecting the presence of PD. Mathematical Description The LR models are mathematically described by Equation 1, which describes the probability P of the binary outcome—PD or not—as a function of input features X: where the coefficients for the input features X_i are denoted by β_i and the intercept term by β_0.

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \quad (1)$$

2) SVM: SVM is a supervised learning approach used in classification tasks. It works by finding out the best hyperplane that maximally separates data points from different categories [18]. The biological voice measurements are used as bases for the SVM to separate the healthy people from the PD-affected ones. In using kernels, SVM handles non-linear relationships in the data, resulting in a robust model that separates classes through complex feature patterns. SVM seeks a hyperplane in Equation (2) that optimally separates the two classes (PD or not).

$$f(\mathbf{X}) = \mathbf{w} \cdot \mathbf{X} + b \quad (2)$$

3) RF: An RF Classifier creates a lot of DTs during training and generates a class that is the average of the classes (classification) of the individual trees. To identify Parkinson's disease and create a robust model, which lowers overfitting and increases accuracy, it uses biological voice measures. By averaging the output of a multitude of DTs, it improves the prediction performance and stability, allowing it to distinguish between people with PD and healthy controls. Equation (3) shows how all the forecasts from individual trees are aggregated to obtain the final prediction, usually by majority vote. Individual DTs in the forest are represented by T_i.

$$\hat{Y} = \text{mode}\left(T_1(\boldsymbol{X}), T_2(\boldsymbol{X}), \ldots, T_m(\boldsymbol{X})\right) \quad (3)$$

4) GB: A GB Classifier is an ensemble learning method that iteratively constructs multiple weak learners. The model corrects the PD identification errors of past trees by biological voice measures and then iteratively builds a new tree. This increases precision and makes the model more robust. Its final prediction is represented by equation (4). where h_m are the weak learners and α_mare the learning rates.

$$\hat{Y} = \sum_{m=1}^{M} \alpha_m h_m(\boldsymbol{X}) \quad (4)$$

5) Extreme GB, popularly known as XGBoost, is a distributed GB library optimized for maximum portability, efficiency, and versatility. It is an advanced variant of the GB method, and it is fast becoming one of the most popular machine learning techniques for regression and classification [19]. It constructs a series of DTs, training each tree to correct errors of its predecessors. With added regularization terms, in order to avoid overfitting, the prediction is akin to Equation (4) in reputation for scalability, speed, and performance. where the complexity of the trees is controlled by the use of regularization.

6) KNN: This is a very strong algorithm in machine learning. Since it makes it possible to compare a data point with its nearest neighbors in order to make a precise prediction, it is quite useful for classification tasks [20]. The model is implemented with k=3 for the Parkinson's dataset, meaning it uses the three nearest data points for classification. It calculates the distance between each of the training instances and the test instance, usually Euclidean. Then, it assigns the class with the majority of the neighbors to the test instance. Assuming that data points with similar attributes—voice measures and other features—belong to the same class—Patison's or healthy, a new instance is classified by KNN. The decision rule is presented in equation (5).

$$\hat{Y} = \text{mode}\left(Y_{NN_1}, Y_{NN_2}, \ldots, Y_{NN_k}\right) \quad (5)$$

7) GNB: This classifier makes its prediction using probabilistic calculations and assumptions of feature distribution and independence. It examines the characteristics of each class—healthy or Parkinson's—to decide the probability of every incident in the Parkinson's dataset. The model assigns the class with the highest computed probability to the instance. This technique works quite fine when dealing with small datasets and data with several dimensions. Its simplicity and efficiency in handling the various attributes gleaned from voice measurements make it particularly useful in this situation. GNB assumes feature independence and a Gaussian distribution when calculating the posterior probability using Bayes' theorem (Equations 6 and 7). where the mean and standard deviation of feature X_ifor class c are denoted by u_ic and σ_icare.

$$P(Y = c|X) = \frac{P(Y = c)\prod_{i=1}^{n} P(X_i|Y = c)}{P(X)} \quad (6)$$

$$P(X_i|Y = c) = \frac{1}{\sqrt{2\pi\sigma_{ic}^2}} \exp\left(-\frac{(X_i - \mu_{ic})^2}{2\sigma_{ic}^2}\right) \quad (7)$$

8) DT: This model classifies data by placing it into subsets based on feature values. For the Parkinson's dataset, it creates a model reliably used to detect parkinsonism. In making its predictions, the model takes into account voice measures along with a number of other characteristics. It splits the data at each node by using criteria like information gain or Gini impurity; a path from the root to any leaf shows the classification criteria, model is, therefore, quite due to its simplicity in understanding and visualization. Based on the feature values, a DT moves from the root to a leaf node to make a prediction. On example X, the prediction by (8) is:

$$\hat{Y} = T(\boldsymbol{X}) \quad (8)$$

9) DT: Adaptive Boosting, or AdaBoost, DTs are trained sequentially on the Parkinson's dataset, and the weights of misclassified instances are updated so that more attention is given to the harder instances. Training before tree bugs

rectified by each subsequent tree is trained to correct the mistakes of the earlier trees. The final model uses a weighted majority vote to combine the predictions from all of the trees. Repeating this process has an effect of improving overall classification accuracy, making it more reliable at differentiating between subjects with PD and healthy subjects. The final prediction can be represented using Equation 9. Here, α_mare weights assigned to each and weak learner h_m.

$$\hat{Y} = \text{sign}\left(\sum_{m=1}^{M} \alpha_m h_m(\boldsymbol{X})\right) \quad (9)$$

3.4. Interpretation and explanation of the model

AI techniques are necessary for ML model interpretation. Feature importance comes in handy to identify important features that are used in a decision to quantify how much an individual input feature contributes toward the prediction made by the model. Permutation importance works by analyzing the performance of a model with regard to feature values and understanding the dependencies on such a feature through a random rearrangement. Mathematically, if (fˆ) is the trained model and (L(fˆ)) is the loss function, then the permutation importance of feature (j) is given by "Importance" (j)=L(fˆ)-L((f_π(j))ˆ). where the model's prediction after shuffling feature (j) is denoted by ((f_π(j))ˆ. Another effective way to explain predictions for individual instances is to approximate the model locally with an interpretable model-say, a linear model-using LIME. To that end, LIME generates perturbed samples around the instance of interest and then fits a simple, interpretable model: yˆ=β_0+β_1 x_1+β_2 x_2+⋯+β_n x_n, where yˆ represents the predicted value, βi are coefficients, and xi denotes feature values.

3.5. Evaluation

We evaluated the performance of each model in binary classification using common assessment metrics, including F1-score, accuracy, precision, and recall. Precision calculates the percentage of true positive predictions out of all positive predictions. Accuracy measures the correctness of the classification as a whole. Recall, also known as sensitivity, is the proportion of true positives that are correctly identified by the model. The F1-score is the harmonic mean of precision and recall, which fairly estimates the performance of the model. These can be calculated using Equation 10-13.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (13)$$

4. Analysis of the experimental setup and results

4.1. Visualization and comparative analysis

Figure 4 shows metrics F1 Score, Accuracy, Precision, and Recall, the performances of the models. Where a series top performances in all four comparison categories are concerned, KNN gives the best results among all models. It has an accuracy similar to those obtained using complex ensemble methods such as RF, GB, XGBoost, and AdaBoost—with great balance. Moreover, KNN presents good F1 Score, Precision, and Recall in all cases, proving to be resilient and dependable in the classification tasks. While the other algorithms—LR, SVM, and ensemble approaches—also yield good results, KNN surpasses them by being consistent across the evaluation criteria. Naive Bayes and DT also show reasonable results for the lower F1 Score and Recall.

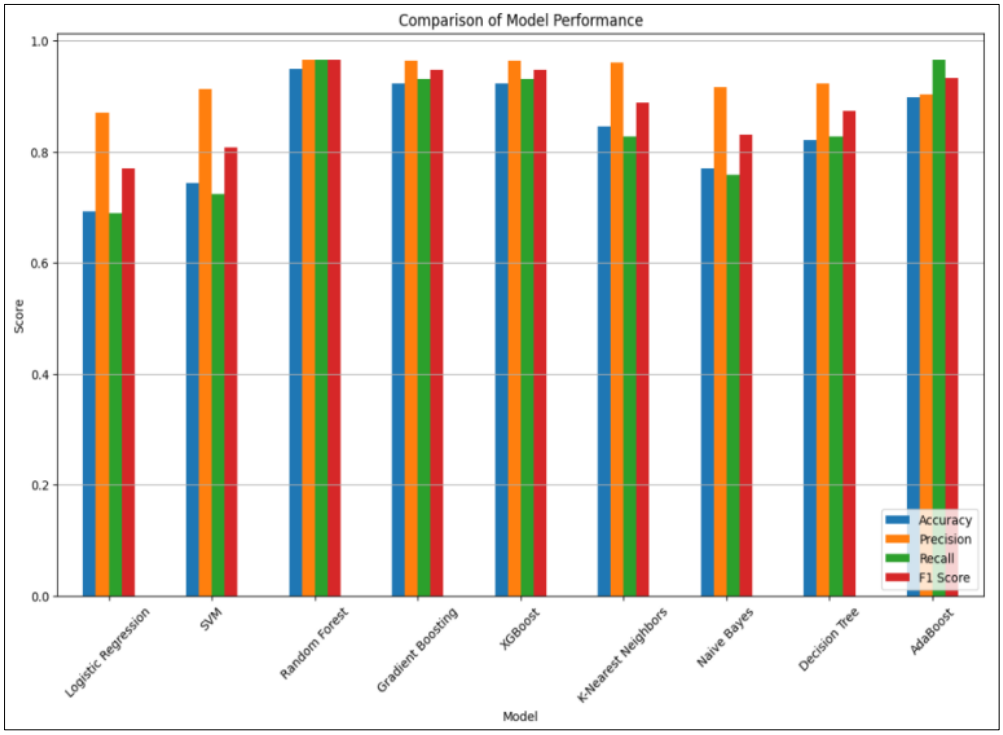

Fig 4: Model Performance Comparison.

Figure 5 reveals the high performance of the KNN model in PD classification as confusion matrix. The precision of the model is very high, at 96%, which means it is very precise in predicting PD—24 true positives and only 1 false positive. It also has a respectable recall, about 82.8%, to correctly detect 82.8% of real PD cases. The overall efficacy of the model can be illustrated using the accuracy value of 84.6%. Five false negatives, however, would mean that some PD patients had been misclassified as healthy. The approximately 88.9% F1 Score indicates a performance balancing between memory and precision.

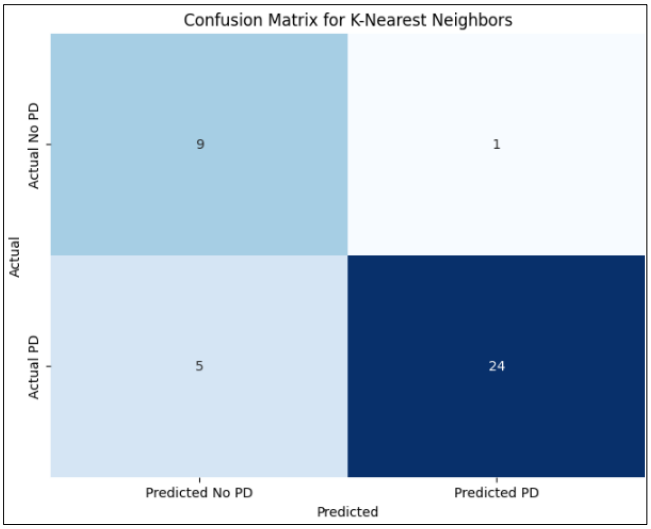

Fig 5: Confusion matrix for KNN.

Figure 6 The learning curve, describing the relationship between the number of training instances and the performance of the model, including training and cross-validation scores. While the green curve shows the cross-validation score, representing the model's performance on unknown data, the red curve shows the training score, indicating great performance on the training data. While the cross-validation score is lower, indicating poor generalization, the training score is initially high, showing a good fit to the training data. Both scores rise more training examples, indicating less overfitting and improved generalization. The convergence of the two scores points to better model performance with additional data. Variability is shown by the shaded areas surrounding the curves; smaller areas denote less variability.

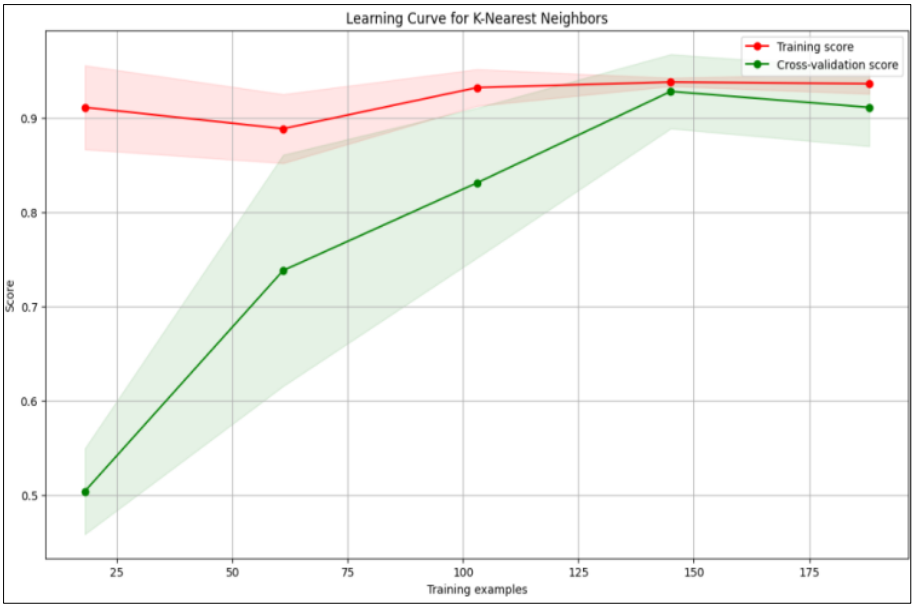

Fig 6: Learning Curve for KNN.

Prediction probabilities for the two classes of "healthy" (0.60) and "parkinson" (0.40) are shown in Figure 7. Features with corresponding weights are listed on the right. Interestingly, "D2 <= -0.70" has by far the largest weight (roughly 0.30) to contribute to a "Parkinson" prediction, but other traits have varying contributions. Positive attributes of Figure 8 are believed to have more influence on the output of the model, while negative variables have less or even a negative effect. For instance, "RPDE" is -1.14, so it is less influential or even negatively influential, while "MDVP:Flo(Hz)" has a value of 0.17, so it is positively influential.

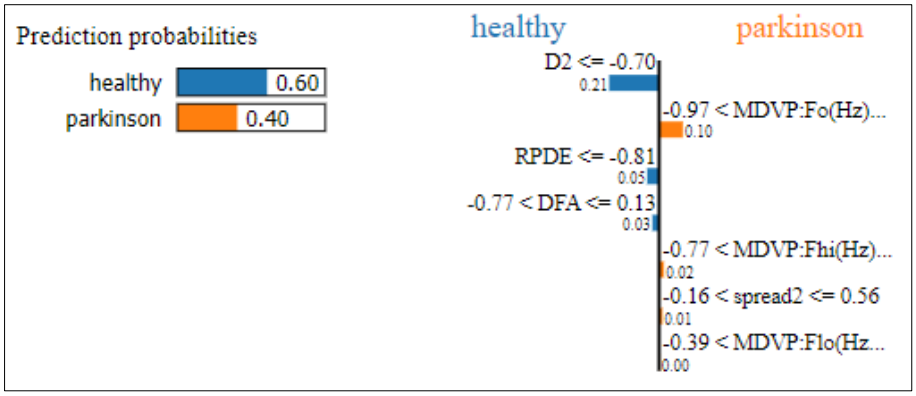

Fig 7: Prediction probabilities of KNN.

Feature	Value
MDVP:Fo(Hz)	-0.27
MDVP:Fhi(Hz)	-0.45
MDVP:Flo(Hz)	0.17
RPDE	-1.14
DFA	0.08
spread2	0.06
D2	-0.77

Fig 8: Feature analysis of the KNN model.

Among the major medical diagnostics, the key progress is that voice biomarkers are being explored for the screening and tracking of Parkinson's disease at initial stages adopting ML models. Current study is using a non-invasive, easily accessible source of data—voice sample analysis—that may eventually make it possible for patients to be diagnosed quicker and in a more convenient manner. Explainable AI is important because it allows for the explanation of models on how decisions are being made, aside from the basic prediction. The complex relation between voice and neurological disorders is depicted by the determination of Spread2, RPDE, and MDVP (Hz) as key vocal biomarkers. Together, MDVP offers an analysis of the basic frequency range, Spread2 analyzes pitch variation, and RPDE looks at the complexity of voice pattern, thus offering a multi-dimensional view of how PD affects speech.

Lower fundamental frequencies, as measured by MDVP (Hz), in Parkinson's patients may serve as a useful marker for early identification, which may lead to further medical research. In addition to the confirmation of these results, the consistency of this study with previous studies contributes to the body of knowledge and supports the continuing efforts in the fight against Parkinson's disease. That the suggestion should be that future studies incorporate both time-series speech data and a wider range of demographics speaks to at least an attempt at making the models more robust and interpretable. These would bring about more precise and individualized health care solutions. The results of the study are, therefore, likely to spur innovative and creative approaches in health care, while the combination of machine learning with vocal analysis in medical research holds great promise.

5. Conclusion

It should be realized that before these models can be put into real-life settings and integrated into clinical decision-making, much improvement and validation are still required. A number of activities are under way in the pursuit of making the models more robust and applicable in all possible scenarios. This comes about through much research, substantiation of findings using a variety of datasets, and collaboration with clinicians. There is a lot of hope that ML-based diagnostic methods will bring about early detection of PD and individualized treatment. Ultimately, this might bring improved standards of living for patients and better patient outcomes.

6. References

1. Radhakrishnan DM, Goyal V. Parkinson's disease: A review. Neurol India. 2018 Mar;66(7):S26–S35. DOI: 10.4103/0028-3886.226451.
2. Mei J, Desrosiers C, Frasnelli J. Machine learning for the diagnosis of Parkinson's disease: A review of literature. Front Aging Neurosci. 2021 May;13:633752. DOI: 10.3389/fnagi.2021.633752.
3. Senturk ZK. Early diagnosis of Parkinson's disease using machine learning algorithms. Med Hypotheses. 2020 May;138:109603. DOI: 10.1016/j.mehy.2020.109603.
4. Pahuja G, Nagabhushan TN. A comparative study of existing machine learning approaches for Parkinson's disease detection. IETE J Res. 2021;67(1):4–14. DOI: 10.1080/03772063.2018.1531730.
5. Oliveira AM, Coelho L, Carvalho E, Ferreira-Pinto MJ, Vaz R, Aguiar P. Machine learning for adaptive deep brain stimulation in Parkinson's disease: closing the loop. J Neurol. 2023 Nov;270(11):5313–5326. DOI: 10.1007/s00415-023-11873-1.
6. Coelho BFO, Massaranduba ABR, Souza CAS, Viana GG, Brys I, Ramos RP. Parkinson's disease effective biomarkers based on Hjorth features improved by machine learning. Expert Syst Appl. 2023 Feb;212:118772. DOI: 10.1016/j.eswa.2022.118772.
7. Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Comput Sci. 2023 Jan;218:249–261. DOI:

10.1016/j.procs.2023.01.007.
8. Little M, McSharry P, Hunter E, Spielman J, Ramig L. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng. 2009 Apr;56(4):1015. DOI: 10.1109/TBME.2008.2005954.
9. Ali L, Javeed A, Noor A, Rauf HT, Kadry S, Gandomi AH. Parkinson's disease detection based on features refinement through L1 regularized SVM and deep neural network. Sci Rep. 2024 Jan;14(1):1–14. DOI: 10.1038/s41598-024-51600-y.
10. Neto OP. Harnessing voice analysis and machine learning for early diagnosis of Parkinson's disease: A comparative study across three datasets. J Voice. 2024 May. DOI: 10.1016/j.jvoice.2024.04.020.
11. Naranjo L, Perez CJ, Martin J, Campos-Roca Y. A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. Comput Methods Programs Biomed. 2017 Apr;142:147–156. DOI: 10.1016/j.cmpb.2017.02.019.
12. Pahuja G, Nagabhushan TN. A comparative study of existing machine learning approaches for Parkinson's disease detection. IETE J Res. 2021;67(1):4–14. DOI: 10.1080/03772063.2018.1531730.
13. Ali L, Zhu C, Zhang Z, Liu Y. Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network. IEEE J Transl Eng Health Med. 2019;7. DOI: 10.1109/JTEHM.2019.2940900.
14. Wang W, Lee J, Harrou F, Sun Y. Early detection of Parkinson's disease using deep learning and machine learning. IEEE Access. 2020;8:147635–147646. DOI: 10.1109/ACCESS.2020.3016062.
15. Gunduz H. Deep learning-based Parkinson's disease classification using vocal feature sets. IEEE Access. 2019;7:115540–115551. DOI: 10.1109/ACCESS.2019.2936564.
16. Nizamuddin MK, Raziuddin S, Farheen M, Atheeq C, Sultana R. An MPL-CCN model for real-time health monitoring and intervention. Eng Technol Appl Sci Res. 2024;14(4):15553–15558. DOI: 10.48084/etasr.7684.
17. Haque R, Islam MB, Parameshachari BD, Khushbu KG, Rahman S, *et al*. Bengali emotion classification using hybrid deep neural network. In: 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE). 2023 Nov:1–7. DOI: 10.1109/AIKIIE60097.2023.10389834.
18. Liu Y, Wu F, Liu M, Liu B. Abstract sentence classification for scientific papers based on transductive SVM. Comput Inf Sci. 2013 Sep;6(4):125. DOI: 10.5539/cis.v6n4p125.
19. Shaheed K, Abbas Q, Hussain A, Qureshi I. Optimized Xception learning model and XgBoost classifier for detection of multiclass chest disease from X-ray images. Diagnostics. 2023 Aug;13(15):2583. DOI: 10.3390/diagnostics13152583.
20. Çakir M, Yilmaz M, Oral MA, Kazanci HO, Oral O. Accuracy assessment of RFerns, NB, SVM, and kNN machine learning classifiers in aquaculture. J King Saud Univ Sci. 2023 Aug;35(6):102754. DOI: 10.1016/j.jksus.2023.102754.