ADVANCED PARKINSONS DISEASE DETECTION WITH COMPREHENSIVE VOICE FEATURE ANALYSIS

Sumit Upadhyay, Shrenik Nandre, Georgy S Pulayath, Prabhujeet Sahu, Gaurav Srivastava

Guided By Dr. Sukrit Gupta, IIT Ropar

ABSTRACT

This study aims to develop an advanced model for Parkinson's disease (PD) detection using diverse voice recordings, including sustained vowels, continuous speech, and other audio samples. By analyzing multiple acoustic features—such as jitter, shimmer, harmonics-to-noise ratio, and MFCCs the study seeks to identify key vocal characteristics that distinguish PD patients from healthy controls. Supervised classification algorithms, including neural networks, were employed in this research, achieving a peak accuracy of 79.83 %, which surpasses the average clinical diagnosis accuracy of non-experts (73.8%). Using a publicly available dataset, this study aspires to create an accessible, non-invasive, speech-based diagnostic tool for early diagnosis and ongoing monitoring of Parkinson's through voice analysis.

Index Terms— Parkinson's Disease (PD), Acoustic Features, Non-Invasive Diagnosis, Machine Learning Model, Speech-Based Diagnostic Tool, Sustained Vowels, Running Speech

1. INTRODUCTION

Parkinson's Disease (PD) represents a significant difficulty within neurological disorders. A degenerative movement illness impacting the neurological system, Parkinson's Disease is characterized by debilitating symptoms including tremors, stiffness, bradykinesia, and postural instability. Despite prominent individuals such as Muhammad Ali, Michael J. Fox, and Pope John Paul II contending with Parkinson's Disease, a definitive solution for this condition remains elusive. Due to the unavailability of exact diagnostic tests and [1] the invasiveness of several current procedures, diagnosing PD creates complex issues. Blood testing, laboratory assessments, and brain imaging are employed to exclude alternative potential ailments, although these techniques may involve strenuous procedures that could exacerbate the discomfort of Parkinson's disease patients. Using speech sample data processing, [2] this research explores a non-invasive diagnosis technique. The idea is to discriminate between those who have PD and those who do not by employing certain features in the acoustic aspects of their voices.

2. MOTIVATION FOR THE WORK

Deep learning techniques, such as convolutional neural networks (CNNs) utilized for spectrographic pictures, are gaining prominence in speech analysis; nonetheless, they are constrained by their deficiency in explicit feature extraction and interpretability. Conversely, conventional machine learning (ML) techniques adhere to a more defined pipeline—from audio recording to pre-processing, feature extraction, and classification—resulting in more interpretable outcomes. Feature selection can enhance machine learning models, especially when a multitude of features is present. Considering that Parkinson's Disease (PD) [3] exhibits symptoms similar to those of illnesses such as REM Sleep Behavior Disorder (RBD), hypokinetic dysarthria, and tremors generated by PD, it is crucial to identify relevant acoustic aspects to discern probable patterns. Frequently utilized elements, such as Fundamental Frequency (F0), jitter, shimmer, MFCCs, and diadochokinetic (DDK) measurements, represent the physical characteristics of the vocal folds and evaluate the capacity to generate rapid, alternating sounds—crucial metrics in examining the effects of Parkinson's Disease on vocalization.

3. IMPORTANCE OF THE RESEARCH IN PARKINSON'S DISEASE

More study is needed to better understand, describe, and identify aspects of Parkinson's disease at its preclinical phase. Finding biological identifiers, or biomarkers, of these early phases is a top aim in order to identify patients who are at high risk of moving to the clinical phase of Parkinson's disease. Therapeutics or other interventions may become available in the future to prevent or delay the start of the clinical phase of Parkinson's disease in persons at high risk. Our study demonstrates the efficacy of machine learning and deep learning techniques in properly recognizing Parkinson's disease from voice signals, highlighting the potential for these approaches to greatly contribute to early diagnosis and intervention strategies for Parkinson's disease.

4. OVERVIEW OF THE STRUCTURE OF THE REPORT.

The report opens with an abstract that summarizes the major themes covered. The introduction then provides the under-

lying framework by stating the key goals and scope. The methodology section describes the research methods, data collection procedures, and analytical approaches used. This is followed by a results section that includes findings, such as key data, observations, and interpretations in relation to the original objectives.

5. METHODOLOGY

5.1. Dataset

This work utilizes the 'Parkinson Speech Dataset with Multiple Types of Sound Recordings' comprising speech samples from both Parkinson's disease sufferers and healthy individuals. Every participant documented 26 speech samples, comprising words, phrases, prolonged vowels, and numerals.

5.1.1. Dataset Description

The dataset was prepared for analysis through loading and pre-processing. It features 1039 entries and 29 columns with varied acoustic elements derived from voice recordings for each entry. We excluded the Unified Parkinson's Disease Rating Scale (UPDRS) scores. Since UPDRS is determined by physicians through medical examination, incorporating it would render the model's classification redundant. Our focus was to leverage voice features solely, ensuring the model's utility in predicting disease severity without prior clinical scoring, thus enhancing its standalone diagnostic capability.

Column No.	Measurement category	Description
1	Subject identifier	This number identifies a study subject
2	Jitter	Jitter in %
3	Jitter	Absolute jitter in microseconds
4	Jitter	Jitter as relative amplitude perturbation (r.a.p.)
5	Jitter	Jitter as 5-point period perturbation quotient (p.p.q.5)
6	Jitter	Jitter as average absolute difference of differences between jitter cycles (d.d.p.)
7	Shimmer	Shimmer in %
8	Shimmer	Absolute shimmer in decibels (dB)
9	Shimmer	Shimmer as 3-point amplitude perturbation quotient (a.p.q.3)
10	Shimmer	Shimmer as 5-point amplitude perturbation quotient (a.p.q.5)
11	Shimmer	Shimmer as 11-point amplitude perturbation quotient (a.p.q.11)
12	Shimmer	Shimmer as average absolute differences between consecutive differences between the amplitudes of shimmer cycles (d.d.a.)
13	Harmonicity	Autocorrelation between NHR and HNR
14	Harmonicity	Noise-to-Harmonic ratio (NHR)
15	Harmonicity	Harmonic-to-Noise ratio (HNR)
16	Pitch	Median pitch
17	Pitch	Mean pitch
18	Pitch	Standard deviation of pitch
19	Pitch	Minimum pitch
20	Pitch	Maximum pitch
21	Pulse	Number of pulses
22	Pulse	Number of periods
23	Pulse	Mean period
24	Pulse	Standard deviation of period
25	Voice	Fraction of unvoiced frames
26	Voice	Number of voice breaks
27	Voice	Degree of voice breaks
28	UPDRS	The Unified Parkinson's Disease Rating Scale (UPDRS) score that is assigned to the subject by a physician via a medical examination to determine the severity and progression of Parkinson's disease.
29	PD indicator	Value "1" indicates a subject suffering from PD. Value "0" indicates a healthy subject.

Fig. 1. Features

5.1.2. Data Division

In this dataset, the categorical column PD indication indicates whether an individual has Parkinson's disease or not. As a result, we began by dividing the dataset into two groups: those without Parkinson's disease and those who had it. The chart demonstrates that in this sample set, the proportion of people with and without Parkinson's disease is nearly equal.

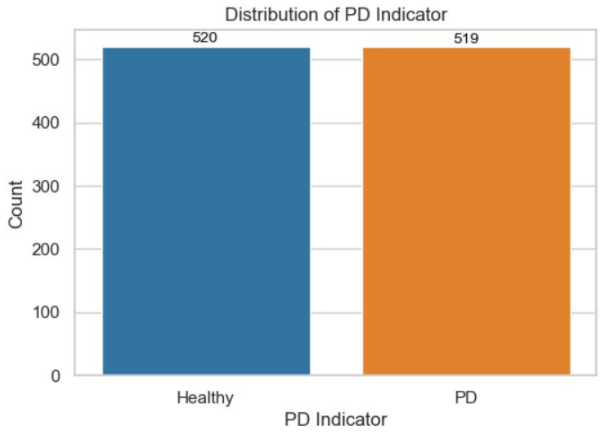

Fig. 2. With or Without PD

5.1.3. Descriptive Analysis

We create side-by-side histograms to examine the distribution of numerical features in healthy versus ill individuals. This image makes it easy to see differences in feature distributions between the two groups.

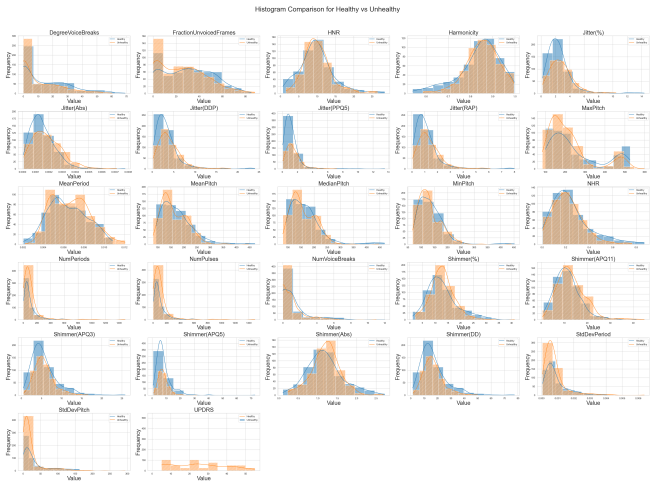

Fig. 3. Histogram Comparison of With or Without PD

Histograms show the distribution characteristics of various features in the dataset. For example, the properties associated with jitter measurements, such as 'Jitter(%)', 'Jitter(Abs)', and 'Jitter(RAP)', have positively skewed distri-

butions, indicating longer tails toward higher values. The 'Shimmer' features, which measure voice variability, show varied degrees of spread and skewness, particularly 'Shimmer(APQ5)'. 'Harmonicity' has a mean of around 0.85 and a standard deviation of 0.09, indicating a pretty consistent distribution. The 'NHR' (noise-to-harmonics ratio) has a positively skewed distribution, indicating that data skews toward higher values. Features such as 'NumPulses' and 'NumPeriods' have a strong positive skew, indicating potential outliers or variability. Meanwhile, metrics like 'MeanPeriod' and 'StdDevPeriod' have distributions centered near zero, indicating limited variance. The distribution of 'UPDRS', which could represent the severity of Parkinson's disease, is skewed to the right.

5.2. Artificial Neural Networks

Figure 2 depicts artificial neural networks, a technology based on brain and nervous system studies. These networks look like biological neural networks, however they only use a subset of the principles found in biological brain systems. More specifically, ANN models [4] imitate the electrical activity of the brain and nervous system. Processing elements (also known as neurodes or perceptrons) are linked to other processing elements. Typically, the neurodes are arranged in a layer or vector, with the output of one layer feeding into the next and possibly further layers. A neurode may be connected to all or a subset of the neurodes in the next layer, imitating the brain's synaptic connections. Weighted data signals entering a neurode replicate the electrical activation of a nerve cell, and hence the flow of information throughout the network or brain. The input values to a processing element, in, are multiplied by a connection weight, wn,m, which represents neural pathway strengthening in the brain. ANNs simulate learning by adjusting the connection strengths or weights.

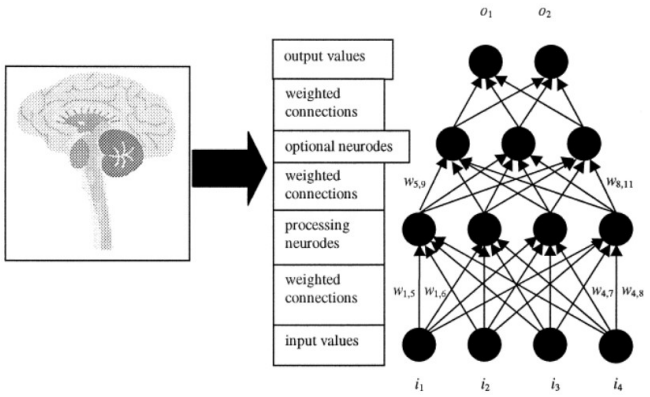

Fig. 4. ANN Architecture

5.3. SHAP (SHapley Additive exPlanation) Values

The Shapley value (SHAP) concept was first created to assess the importance of a single person in a collaborative team. [5] This notion attempted to distribute the overall benefit or payment among participants based on their respective value to the game's final conclusion. Shapley values are a solution for assigning a fair or appropriate reward to each player and provide a unique result characterized by the following natural qualities or axioms: local accuracy (additivity), consistency (symmetry), and nonexistence (null effect).

In the context of activity predictions, Shapley values can be justified as a fair or reasonable allocation of feature importance given a specific model result. Shapley values allow for the fact that features contribute to the model's output or prediction in varying degrees and signs. Shapley values thus reflect estimates of both feature relevance (magnitude of contribution) and direction (sign). characteristics with a positive sign predict activity, while characteristics with a negative sign predict inaction.

$$\phi_i = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|! \left(|N| - |S| - 1\right)! \left[f\left(S \cup \{i\}\right) - f\left(S\right)\right]$$

Fig. 5. Shapley value defines the relevance of a feature i

The output f(S) the ML model can be explained using a comprehensive set of S features. Feature i (ϕ_i) final contribution or Shapley value is calculated by averaging its contributions over all feature set permutations. As a result, individual features are added to the set, and the change in model output indicates their significance. Importantly, this approach takes into account feature orderings, which influence reported changes in a model's output in the presence of correlated features.

6. MODEL EVALUATION

To determine the ANN model's maximum performance, we used iterative training to optimize accuracy while avoiding overfitting, using SHAP for feature selection to focus on the most relevant inputs. To provide a thorough comparison, we assessed model performance using a variety of metrics, including the ROC-AUC curve, confusion matrix, accuracy, precision, recall, and F1 score. The equations for these metrics, provided in equations 1-3, use True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) to provide precise information about the model's effectiveness.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

We obtained a peak accuracy of 79.83%, a sensitivity of 80.39%, and a specificity of 78.30% to evaluate our optimized ANN model for Parkinson's disease diagnosis. With an AUC-ROC score of 89.63%, the model showed strong predictive performance, therefore proving its capacity to clearly separate impacted from non-affected cases.

The performance metrics specified in Equations 1-3 were chosen for their ability to provide a balanced perspective of the model's success across various dimensions:

1. **Accuracy** (Eq. 1): The ratio of accurately predicted observations to total observations.

2. **Precision** (Eq. 2): The proportion of true positive forecasts to all positive predictions.

3. **Recall (Sensitivity)** (Eq. 3): The proportion of genuine positive forecasts to all real positives.

The AUC-ROC curve (Fig. 6) provides a comprehensive perspective of model performance across multiple thresholds, whilst the SHAP analysis (Fig. 7) shows the contribution of each feature, guiding the selection of relevant inputs and reducing overfitting. These findings were critical in establishing the model's high accuracy while maintaining generalizability.

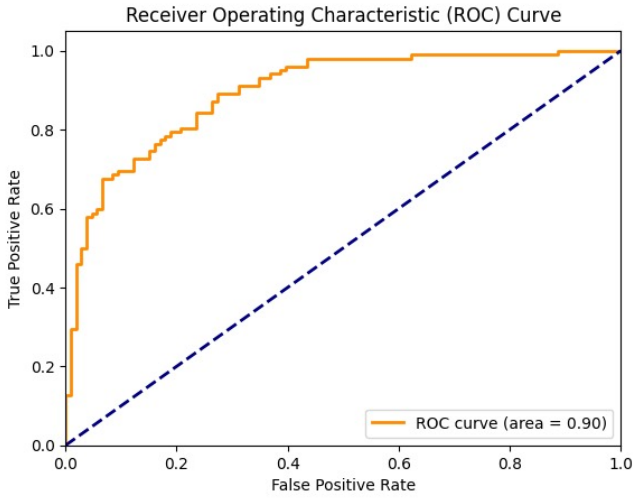

Fig. 6. AUC-ROC curve

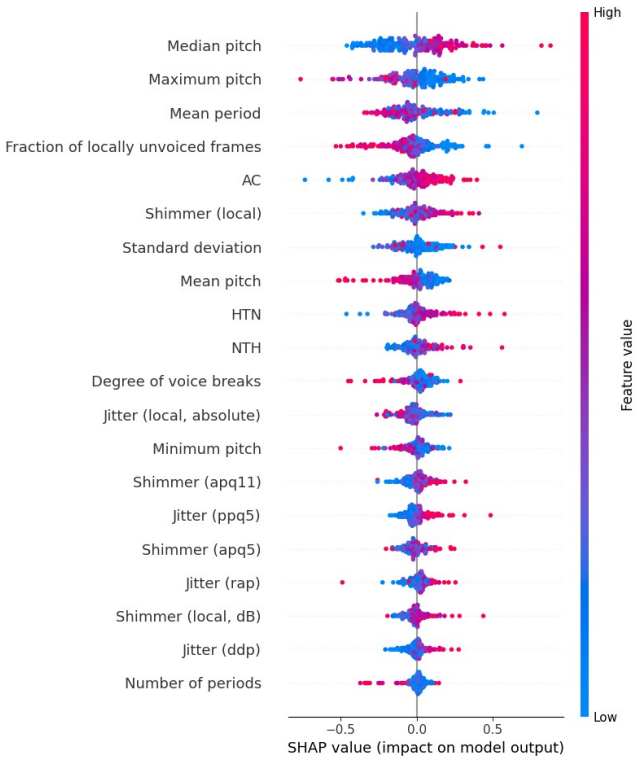

Fig. 7. Shapley value relevance of the features

7. RESULTS

We created an improved Artificial Neural Network (ANN) model to identify Parkinson's illness by analyzing voice parameters such as jitter, shimmer, harmonics-to-noise ratio, and MFCCs. Our technique included iterative training and feature selection with SHAP analysis, which assisted in identifying the most relevant voice aspects and reducing overfitting. This meticulous selection contributed to the model's strong performance on a variety of evaluation metrics.

The ANN model had a high accuracy of 79.83%, a sensitivity of 80.39%, and a specificity of 78.30%, showing a balanced capacity to properly identify both Parkinson's-affected and unaffected patients. The AUC-ROC score was 89.63%, indicating the model's high discriminative potential. As shown in Fig. 6, the AUC-ROC curve demonstrates the model's ability to maintain high performance across various threshold settings, while Fig. 7 depicts feature relevance based on SHAP values, demonstrating the importance of specific voice characteristics in Parkinson's disease detection. These findings highlight the model's effectiveness and reliability in detecting Parkinson's disease, indicating its potential use in diagnostic situations.

8. ACKNOWLEDGMENTS

We would like to thank Dr. Sukrit Gupta sir for his persistent support and exceptional teaching throughout the course. His advice has been quite helpful in navigating the complexity of this study topic. We also like to thank Pritam Sunil Mahajan for his persistent efforts in organizing project review meetings, providing constructive feedback, and assuring the smooth flow of our work. Their contributions have substantially enhanced our research experience and established a collaborative environment that has aided the effective completion of this project.

9. REFERENCES

[1] K. Prabhavathi and S. Patil, "Tremors and Bradykinesia, in Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation", S.P. Arjunan and D.K. Kumar, Eds. Series in BioEngineering. Singapore: Springer, 2022, pp. 135–149. doi: https://doi.org/10.1007/978-981-16-3056-9_9.

[2] R. Alkhatib, M. O. Diab, C. Corbier, and M. E. Badaoui, "Machine Learning Algorithm for Gait Analysis and Classification on Early Detection of Parkinson, IEEE Sensors Letters", vol. 4, no. 6, pp. 1–4, Jun. 2020, Art no. 6000604. doi: https://doi.org/10.1109/LSENS.2020.2994938.

[3] F. Cordella, A. Paffi, and A. Pallotti, "Classification-based screening of Parkinson's disease patients through voice signal, in Proc. IEEE Int. Symp. Medical Measurements and Applications (MeMeA)", Lausanne, Switzerland, 2021, pp. 1–6. doi: https://doi.org/10.1109/MeMeA52024.2021.9478683.

[4] S. Walczak and N. Cerpa, "Artificial Neural Networks, in Encyclopedia of Physical Science and Technology", 3rd ed., R. A. Meyers, Ed. Amsterdam, The Netherlands: Academic Press, 2003, pp. 631–645. doi: https://doi.org/10.1016/B0-12-227410-5/00837-1.

[5] M. Louhichi, R. Nesmaoui, M. Mbarek, and M. Lazaar, "Shapley Values for Explaining the Black Box Nature of Machine Learning Model Clustering, Procedia Computer Science", vol. 220, pp. 806–811, 2023. doi: https://doi.org/10.1016/j.procs.2023.03.107.