


Article

Ensemble Machine Learning Approach for Parkinson's Disease Detection Using Speech Signals

Syed Nisar Hussain Bukhari ^{1,*} and Kingsley A. Ogudo ² ¹ National Institute of Electronics and Information Technology (NIELIT), Srinagar 191132, India² Department of Electrical & Electronics Engineering, Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg 0524, South Africa; kingsleyo@uj.ac.za

* Correspondence: nisar@nielit.gov.in

Abstract: The detection of Parkinson's disease (PD) is vital as it affects the population worldwide and decreases the quality of life. The disability and death rate due to PD is increasing at an unprecedented rate, more than any other neurological disorder. To this date, no diagnostic procedures exist for this disease. However, several computational approaches have proven successful in detecting PD at early stages, overcoming the disadvantages of traditional methods of diagnosis. In this study, a machine learning (ML) detection system based on the voice signals of PD patients is proposed. The AdaBoost classifier has been utilized to construct the model and trained on a dataset obtained from the machine learning repository of the University of California, Irvine (UCI). This dataset includes voice attributes such as time-frequency features, Mel frequency cepstral coefficients, wavelet transform features, vocal fold features, and tremor waveform quality time. The model demonstrated promising performance, achieving high accuracy, precision, recall, F1 score, and AUC score of 0.96, 0.98, 0.93, 0.95, and 0.99, respectively. Furthermore, the robustness of the proposed model is rigorously assessed through cross-validation, revealing consistent performance across all iterations. The overarching objective of this study is to contribute to the scientific community by furnishing a robust system for the detection of PD.

Keywords: machine learning; Parkinson's disease; speech signals; boosting; ensemble learning

MSC: 68T05



Citation: Bukhari, S.N.H.; Ogudo, K.A. Ensemble Machine Learning Approach for Parkinson's Disease Detection Using Speech Signals. *Mathematics* **2024**, *12*, 1575. <https://doi.org/10.3390/math12101575>

Academic Editor: Andrea Scozzari

Received: 25 April 2024

Revised: 13 May 2024

Accepted: 15 May 2024

Published: 18 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Parkinson's disease (PD) is a degenerative disease that affects the nervous system and other parts of the body controlled by the nervous system [1]. Second only to Alzheimer's disease in prevalence among neurodegenerative conditions, PD imposes a substantial global burden, with over 10 million individuals estimated to be grappling with its ramifications worldwide [2]. The financial burden corresponding to the direct and indirect treatment of the disease is high. According to estimates, more than USD 2000 is spent annually on the diagnosis and treatment of PD per person in India alone. Timely and accurate detection and therapy can help delay the progress of this disorder in affected patients. The cause and diagnosis of PD remain unknown to this date [3]. However, clinical observations and symptomatology serve as cornerstones for disease identification and characterization. The advancement of symptoms varies in different patients. Patients affected by PD show mainly two types of symptoms: motor symptoms and non-motor symptoms [4]. Motor symptoms can range from tremors in the hand, slowed movement, rigid muscles, and changes in writing to impaired posture and balance and loss of involuntary movements [5]. However, motor symptoms appear at later stages of the disease, resulting in late detection of the disease. Non-motor problems, including speech disorders, sleep behavior disorders, depression, and loss of olfactory sense, can detect PD at its early stage.

PD causes the loss of neurons in the part of the brain called *Substantia Nigra*, which is responsible for the production of *dopamine* [6]. *Dopamine* facilitates the coordination between the nervous system and other parts of the body. The loss of *dopaminergic* neurons can lead to speech disorder in patients with PD, as the generation of speech is a result of various coordination activities. The analysis of speech characteristics is essential in the early detection of PD, as its effect on speech and phonation are clear indicators of PD [7]. The deviation in speech signals indicates PD. For instance, the voice signal of a person with PD has a lack of intensity and monotony in pitch and loudness. Speech disorder in patients with PD is characterized by a slowed speech rate, voice tremor, sunken tone, and unclear speech caused by impaired pharyngeal, oral, and jaw muscles.

The convergence of burgeoning research in neurodegenerative disease, advances in machine learning (ML), and the wealth of diagnostic information embedded within speech signals holds immense promise for the detection and characterization of Parkinson's disease. By using ML, the relevant voice features can be extracted, combined, and employed for the assessment of this disease [8]. Depending on the features that can be extracted from speech recordings of the affected patients, voice signals can be of various types, including temporal, spectral, articulatory, prosodic, statistical, and vocal [9]. Among these, temporal, spectral, statistical, and vocal signals are mostly used in the speech analysis of patients with PD [10]. The aim and focus of this study is to develop an effective ML model for the detection of PD utilizing a diverse array of attributes and their corresponding statistical summaries extracted from speech signals. By leveraging ML algorithms, we seek to identify subtle patterns and biomarkers present in speech recordings that can serve as early indicators of PD onset. The proposed model shall act as a non-invasive and cost-effective diagnostic tool that can aid healthcare professionals in detecting PD at its initial stages, enabling timely intervention and improved patient outcomes.

The existing literature on PD detection often highlights the limitations of current diagnostic methods, which primarily rely on the assessment of motor symptoms and may not detect the disease until its later stages. There is a clear gap in the literature regarding the development of sensitive and reliable diagnostic tools for PD detection, particularly those utilizing non-motor symptoms such as speech impairments. Our study addresses this gap by exploring the potential of ML algorithms to analyze speech signals and extract relevant features for PD diagnosis, thus contributing to the advancement of detection techniques.

Previous studies in the field of PD detection have faced several challenges, including limitations related to dataset quality, model performance, and the lack of robust diagnostic tools. Many existing approaches rely on small or imbalanced datasets, which may lead to biased or unreliable results. Additionally, the performance of some ML models in distinguishing between PD patients and healthy individuals may be suboptimal, highlighting the need for more advanced techniques. This study addresses these challenges by employing state-of-the-art pre-processing methods, leveraging a comprehensive dataset, and utilizing ensemble-learning techniques to improve model accuracy and reliability.

1.1. Contributions

The novel contributions of this paper are summarized as follows:

- In this study, we employ a comprehensive dataset comprising voice recordings from PD patients and healthy individuals, allowing for a detailed analysis of speech characteristics and the extraction of relevant features. This dataset encompasses a wide range of clinically significant attributes, including time-frequency fading (TFF) features, Mel frequency cepstral coefficients (MFCCs), vocal fold features (VFF), and tremor waveform quality time (TWQT), among others. By using these diverse features, the proposed model can capture subtle variations in speech patterns associated with PD pathology, thereby enhancing diagnostic accuracy.
- Our study utilizes advanced pre-processing techniques, including synthetic minority oversampling technique (SMOTE) and principal component analysis (PCA), to address imbalanced datasets and reduce dimensionality while preserving essential information.

These techniques enhance the robustness and efficiency of our model, allowing for more accurate PD detection.

- The study adopts an ensemble learning approach, specifically AdaBoost, to construct a predictive model for PD detection. By combining multiple weak learners, AdaBoost improves the model's performance and generalization ability, making it well suited for handling complex relationships in speech datasets. This novel application of ensemble learning techniques to PD detection represents a significant advancement in the field, offering a more effective and reliable diagnostic tool.
- In contrast to conventional diagnostic approaches that often overlook non-motor symptoms, this study explores the diagnostic potential of such symptoms. By incorporating non-motor manifestations, the model aims to enrich the diagnostic process for PD.
- By facilitating the detection of PD, this proposed study strives to mitigate the substantial economic burden associated with delayed diagnosis. Timely identification of the disease can potentially reduce both direct healthcare costs and indirect economic burdens on individuals and healthcare systems.

1.2. Motivations

Parkinson's disease (PD) presents a significant global health challenge, affecting individuals' quality of life and posing a growing burden on healthcare systems. With escalating disability and mortality rates surpassing those of many other neurological disorders, the imperative for early PD detection is undeniable. Yet, conventional diagnostic methods lack the efficacy to meet this demand. Fortunately, computational approaches offer promising avenues for PD detection, circumventing the limitations of traditional diagnostic modalities.

This study is motivated by the pressing need to advance PD diagnosis through the application of ML techniques, specifically focusing on speech signals as a diagnostic modality. Speech analysis holds potential as a non-invasive, cost-effective means of detecting PD, offering insights into the subtle vocal impairments characteristic of the disease. Leveraging the extensive dataset from the University of California Irvine (UCI) Machine Learning Repository, encompassing diverse voice attributes such as time-frequency features, Mel frequency cepstral coefficients, and vocal fold features, this research endeavors to develop a robust ML detection system. The utilization of the AdaBoost classifier represents a deliberate choice, guided by its proven efficacy in handling complex classification tasks. Through meticulous model development and rigorous evaluation, this study aims to achieve superior diagnostic performance, as evidenced by the attained accuracy, precision, recall, F1 score, and AUC score metrics. Additionally, the robustness of the proposed model is systematically validated through cross-validation, ensuring consistent performance across diverse datasets. By providing the scientific community with a reliable, validated ML-based detection system, this research seeks to catalyze advancements in PD diagnosis. Ultimately, the overarching goal is to equip clinicians and researchers with a potent tool for PD detection, thereby enhancing patient outcomes and alleviating the societal burden imposed by this debilitating disease.

The rest of the paper is structured in the following manner. Section 2 contains an illustration of related works in this field. Section 3 provides a detailed description of data collection, pre-processing, feature selection, and model building. Section 4 summarizes the results obtained by the use of various metrics. Section 5 concludes the study with possible future improvements in this field.

2. Related Work

Several statistical methods and data-driven techniques have been developed to detect PD. Due to its data-driven techniques, machine learning (ML), which is an emerging technique in the field of medical research, has contributed sufficiently to the diagnostic process of this disease. A detection model was introduced based on the waves collected from the smartwatches of patients with PD [11]. The model effectively detected the symptoms

of PD with high classification accuracy. The model effectively detected the symptoms of PD with high classification accuracy. A study conducted by [12] employed a pre-trained CNN (Inception V3) with transfer learning to analyze spectrograms of voice recordings of patients with PD. Their deep learning model demonstrated superior performance in classifying patients with PD, with an AUC of 0.97 for colored spectrograms and 0.96 for grayscale spectrograms. In an attempt to diagnose PD at its early stage, a study by [13] utilized several ML algorithms and a three-layer deep neural network (DNN2), which outperformed all, reaching an accuracy of 95.41%. The study suggested that DNNs exhibit superior performance compared to traditional ML methods in categorizing patients with PD based on speech signals. With the aim of developing a soft diagnosis tool for identifying PD through voice signal characteristics, a research investigation carried out by [14] evaluated various ML models. The study demonstrated that the SVM-based model surpassed other ML models in classifying patients with PD with a notable accuracy of 96%. To promote the integration of ML in telemedicine, a study by [15] trained four ML models (SVM, random forest, KNN, and logistic regression) utilizing MDVP (multi-dimensional voice program) audio data from 30 individuals with PD and healthy subjects. The results declared the random forest classifier as the optimal technique, with a detection accuracy of 91.83% and a sensitivity of 0.95. In another study, conducted by [16], feature elimination methods were combined with ML classifiers to diagnose PD using voice disorders. The findings from their research illustrated that the combination of Random Forest with t-SNE (t-distributed stochastic neighbor embedding) and MLP (multi-level perceptron) with the PCA (principal component analysis) exhibited superior performance with an accuracy of 97% and 98%, respectively. Aiming to highlight the importance of using acoustic features in the detection of PD at its early and mid-advance stages, ref. [17] integrated feature selection methods with ML classifiers. In addition to detecting early and mid-advanced stages of PD with an accuracy of 95.4%, their hybrid model also detected stage 3 and stage 4 of PD with an accuracy of 89.48% and 86.62%, respectively. Following a similar approach, ref. [18] experimented with multiple combinations of feature selection methods and classifiers to detect PD with the help of speech signal features. The results concluded that random forest combined with the genetic algorithm outperformed the rest of the combinations with an accuracy of 95.58%. Owing to its capacity to handle large datasets, deep learning techniques have also been applied to the effective diagnosis of PD [19]. By utilizing voice signal characteristics and data balancing techniques, a hybrid LSTM-GRU model proposed by [20] achieved a noteworthy accuracy of 100%. Research conducted by [21] proposed a hybrid CNN-LSTM model that works in various stages, including noise removal, feature extraction, and the final classification stage. The proposed hybrid model demonstrated an enhanced accuracy of 93.51% compared to neural network, CART, SVM, and XGBoost models with accuracies of 72.69%, 84.21%, 73.51%, and 90.81%, respectively. A hybrid method suggested by [22] combined resonance based sparse signal decomposition (RSSD) and a time-frequency (TF) algorithm to extract features from voice signals in order to diagnose PD. The proposed hybrid model also integrated CNN for classification and achieved a validation accuracy of 99.37%. An attempt towards building a non-invasive detection model based on customized CNN was made by [23] using spirals drawn by patients. Using the DL approach to assess the severity of PD, ref. [24] attained the highest accuracy of 99.5% with a hybrid model consisting of CNN and weighted RF. Another hybrid approach recommended by [25] combined the MIRFE feature selection method with an XGBoost classifier. Subsequently, the model achieved a substantial feature reduction ratio of 94.69%, an accuracy of 93.88%, and an AUC of 0.978. Further research was conducted by [26] to explore the diagnosis of PD using a dual-branch deep learning model and gait signals. Their approach involved combining CNN and Bi-LSTM architectures for each foot's gait signal and extracting features from each gait cycle. A study by [27] utilizes CNN to assist in diagnosing PD by analyzing abnormal motor patterns in patient-drawn spiral exercises. The convolutional classifier achieved 91.67% accuracy in distinguishing PD patients from controls based on spiral drawings. Another study by [28] presents a

method for automating PD diagnosis using EEG signals decomposed into sub-bands and analyzed with machine learning models. The VKF method, introduced for the first time in this context, coupled with SVM, achieves nearly perfect classification accuracy, marking a significant advancement in PD diagnostic systems. These findings suggest promising potential for accurate and early PD detection compared to recent research efforts. The study by C. Tran et al. addresses the critical need for improved PD diagnostics and treatment by exploring retinal fundus imaging as a potential screening tool [29]. Through systematic evaluation of machine learning and deep learning techniques, it achieves 68% accuracy in distinguishing PD individuals from healthy subjects, offering promising advancements in early detection and intervention.

3. Materials and Methods

This study follows a series of steps, which include acquiring a suitable dataset, pre-processing the data, visualization, model building, and evaluating the model. The collected data have been pre-processed first and then utilized for classification by the model. Figure 1 displays the experimental framework that has been followed in this study.

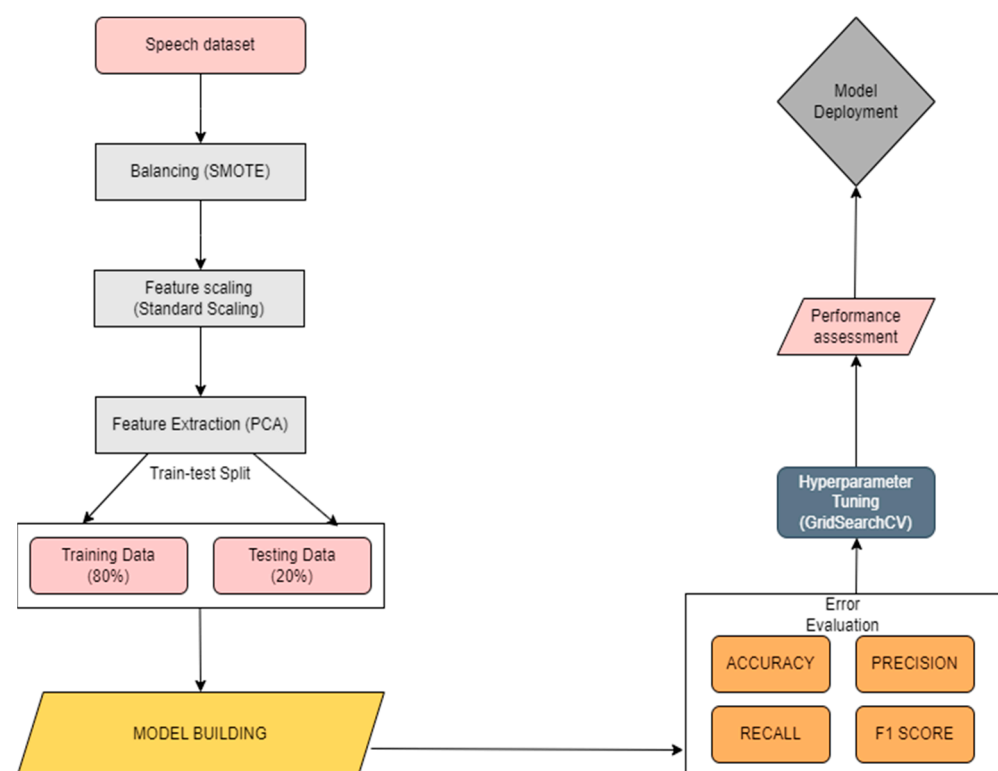


Figure 1. Experimental framework of the proposed methodology.

3.1. Retrieval of Dataset

The selection of an appropriate dataset holds paramount importance in the development of predictive models, as their efficacy is contingent upon the quality and relevance of the data used for training [30]. It underscores the necessity of meticulously choosing a dataset that ensures accuracy and reliability in predictions. In this study, the dataset is sourced from the UCI ML repository, comprising data collected from 188 patients diagnosed with PD, of whom 107 are male and 81 are female, spanning ages from 33 to 87 [31]. The dataset encompasses a total of 756 instances, each characterized by 754 clinically significant attributes crucial for PD detection. The dataset primarily consists of various voice signal attributes obtained from voice recordings of participants, including 188 PD patients and 64 healthy individuals who were instructed to sustain the phonation of the vowel /a/. Data collection involved configuring the microphone to a sampling rate of 44.1 KHz, with

each subject contributing three repetitions of the sustained /a/ vowel sound. To facilitate the training of a model for PD detection, the dataset incorporates a comprehensive array of the following features and associated metrics. These features serve as vital inputs for the model, enabling it to discern subtle patterns indicative of PD pathology.

- Time frequency fading (TFF) features: These features combine time and frequency components of voice signals.
- Mel frequency cepstral coefficients (MFCCs): MFCCs capture vital information about the spectral data of speech signals and are obtained using a series of steps like framing, windowing, and Fourier transform.
- Wavelet transform-based features (WTF): These features include pitch, formants, energy, and spectral density of speech signals obtained from wavelet transform functions.
- Vocal fold features (VFF): VFFs characterize the human vocal cord using certain parameters like jitter, shimmer, and fundamental frequency.
- Tremor waveform quality time (TWQT): It measures the quality of the oscillatory movement of voice signals.

The dataset also includes these additional noise-related features aimed at enhancing the model's robustness and ensuring optimal performance [32], even in the presence of noise in the speech signals:

- Entropy: Entropy provides information about the disorder in the signal and can help the model identify and adapt to noisy conditions, as in the presence of noise, the signal's entropy may change [33].
- Detrended fluctuation analysis (DFA) features: To enhance the model's resilience to short-term noise, DFA features, which are effective for evaluating long-term correlations in the signal, have been leveraged. This approach enables the model to prioritize the underlying structure of the speech signal.
- Quantile, skewness, and kurtosis features: Quantile, skewness, and kurtosis features capture deviations in the voice signals caused by noise, aiding the model in identifying and handling noisy segments.
- Signal-to-noise and noise-to-signal ratios: Integrating these features during model training allows the model to understand the proportion of noise in the signal and distinguish between the actual signal and surrounding noise, promoting robustness in the presence of varying noise levels [34].

3.2. Data Pre-Processing

In ML, data pre-processing holds significant importance because the quality of the data and the valuable insights that can be extracted from it using data pre-processing techniques have a direct impact on the learning ability of our model. Hence, it is crucial to preprocess the data before inputting it into our model. The subsequent divisions include the pre-processing techniques implemented in this study to address imbalanced and raw data, converting it into a structure appropriate for training the model.

3.2.1. Feature Weighting

Imbalancing refers to a situation where the distribution of target class labels in a dataset is not equal. For binary classification problems, one class label has significantly more or less observations than the other class label in an imbalanced situation. To ensure accurate predictions, it is essential to address imbalancing, as it can result in misclassification. In this study, the synthetic minority oversampling technique (SMOTE) is employed as a method to tackle the issue of the imbalanced dataset [35]. The SMOTE technique involves generating synthetic samples for the minority class specifically designed for that class through oversampling. This is accomplished by linear interpolation and creating new instances of the minority class rather than simply duplicating existing instances. The workings of SMOTE can be summarized in the following steps:

- For the given minority class set X , the k -nearest neighbors of each sample a in X are determined by computing the Euclidean distance between a and all other samples in set X .
- The sampling rate N is determined based on the degree of class imbalance. For every sample in the minority class set X , N instances (a_1, a_2, \dots, a_n) are chosen randomly from their k -nearest neighbors to form the set X_1 .
- A new example is generated for each instance a_k in X_1 (where $k = 1, 2, 3, \dots, N$) using Equation (1).

$$a' = a + \text{rand}(0, 1) * |a - a_k| \quad (1)$$

In its original form, the UCI dataset consisted of 564 instances classified as positive and 192 instances classified as negative. The graphs in Figures 2 and 3 depict the distribution of two classes in the dataset before and after balancing the minority class. With SMOTE, the number of instances in the minority class has been increased to 564, resulting in a balanced dataset.

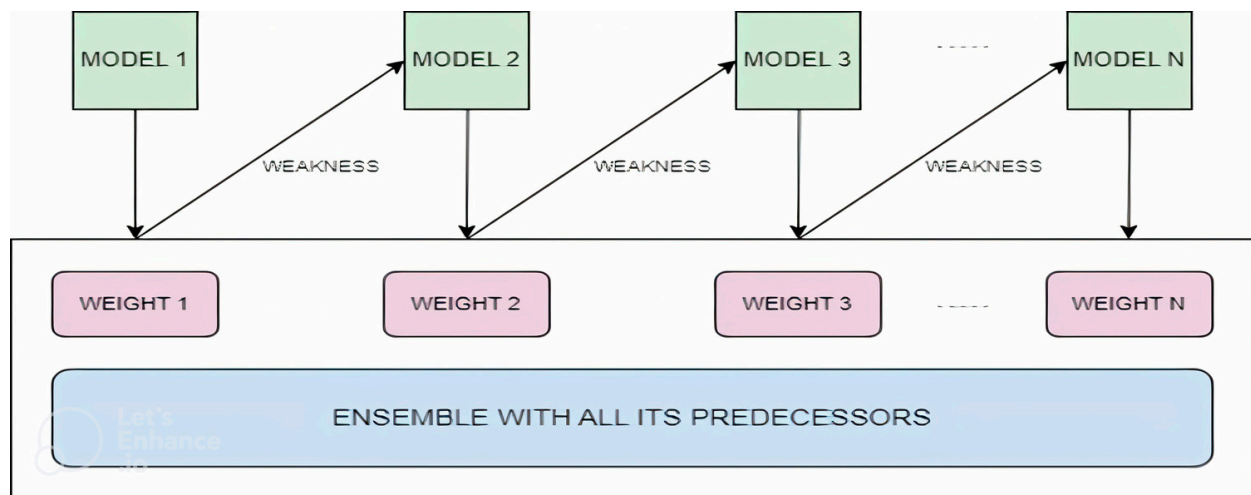


Figure 2. Architecture of the AdaBoost ensemble classifier.

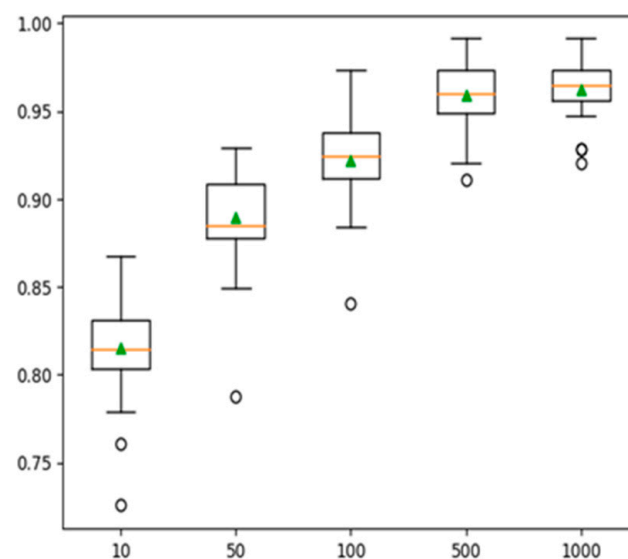


Figure 3. Boxplot depicting accuracies on a varying number of DTs.

3.2.2. Feature Scaling

Feature scaling (FS) [36] refers to the process of transforming the values in a dataset to a specific scale (range), preventing features with larger scales from dominating the learning process. This is typically performed by converting the values to a range that is easier to handle, such as scaling all values to a range between 0 and 1 or standardizing them to have a mean of 0 and a standard deviation of 1. In this study, standard scaling has been employed as a scaling technique, bringing the raw data onto a scale ranging from 0 to 1. For any variable X in a dataset, the standard value can be calculated using Equation (2).

$$X_{stand} = \frac{X - \text{mean}(X)}{\text{Standard deviation}(X)} \quad (2)$$

By scaling the features, the issues that arise with features having vast ranges can be avoided, ultimately improving the performance of ML algorithms.

3.3. Feature Extraction

Feature extraction (FE) is a technique for detecting and extracting significant and meaningful attributes from the original data [37]. These attributes depict the dataset relationships effectively and are useful in constructing a predictive model. Typically, the feature set obtained after extraction is more concise than the original data, which facilitates more precise and effective analysis and modeling. FE involves a wide variety of methods, such as statistical analysis, signal processing, and dimensionality reduction. In the current study, dimensionality reduction has been utilized as a technique for extracting features through principal component analysis (PCA).

Principal Component Analysis

PCA [38] is an unsupervised ML algorithm mainly used for FE. It is a linear transformation method that aims at discovering the directions in high-dimensional data exhibiting maximum variance and then mapping the data onto a fresh subspace with the same or fewer dimensions than the initial space. By identifying significant correlations in the data, PCA alters the data, evaluates the relevance of these correlations, and retains the most essential ones while discarding the rest. This way, it allows us to extract and retain the most critical information from the data. The working method of PCA can be divided into 4 distinct steps:

- **Standardizing the data:** To ensure that each variable has the same influence in the analysis, it is important to standardize the data. This involves adjusting the data so that each variable has a mean of zero (0) and a variance of one (1). Without standardization, variables with larger values can have a disproportionate impact on the analysis, leading to biased results.
- **Constructing the covariance matrix:** The covariance matrix is an $N \times N$ matrix that displays the covariance between each pair of variables in the dataset. Two features can have positive covariance, i.e., they tend to vary in the same direction, or negative covariance, i.e., with an increase in one feature, the other feature tends to decrease.
- **Decompose the covariance matrix:** The covariance matrix is decomposed into Eigen vectors and Eigen values that represent the principal components and their magnitudes, respectively.
- **Projecting the data onto a new sub-space:** In the last step, the most important k eigenvectors are selected based on their respective eigenvalues (ones with the largest Eigen values are chosen), and those are used to transform the data and project it onto the new subspace.

The obtained data have fewer dimensions, with the new attributes being linear combinations of the original ones. A total of 6 principal components were retained from the original dataset to capture the fundamental correlations in the data while reducing its

dimensionality, providing a more concise view of the features, and facilitating the smooth training of the model.

3.4. Model Selection and Training

The objective of this study is to create a predictive model that can aid in the detection of PD. The fundamental stages involved in constructing the proposed model are as follows:

- **Choosing an appropriate ML model:** The process of selecting an ML model involves choosing one model from several potential candidates to learn the complexities in a dataset. In this study, an ensemble modeling approach has been adopted to process the speech dataset for its effectiveness in handling non-linear relationships often exhibited by the speech datasets. Ensemble methods are also recognized for their ability to mitigate the impact of irrelevant information (noise) in the training data, consequently enhancing the proposed model's robustness against speech data affected by noise.
- **Fitting the ML model:** Fitting an ML model entails evaluating how effectively the model can extrapolate to comparable data to that for which it was trained. In this study, the dataset was partitioned into training and testing subsets with an 80:20 ratio. Post the train-test split, the ensemble model was fitted to the training subset, allowing it to capture the underlying intricacies of the data. Subsequently, the model's predictive capability was assessed on the testing subset, which remained unseen during the training phase. Generalizing the model to unseen data is essential for ensuring its ability to make predictions.

Model Building

The proposed ensemble learning model [39] is designed to enhance prediction accuracy and minimize errors by consolidating predictions from multiple models, with boosting selected as the method of choice. Boosting, an iterative technique, aims to elevate weak learners such as basic decision trees (DTs) into robust models by aggregating their predictions using techniques like voting or averaging. Notably, boosting encompasses four primary categories of algorithms in machine learning: AdaBoost, CatBoost, gradient boosting, and light gradient boosting. In this study, AdaBoost is employed as the boosting algorithm due to its robustness against overfitting and its efficacy in handling datasets comprising speech signals with a large number of samples and features.

AdaBoost is a meta-estimator that works by initially fitting a classifier on the original dataset, where each data instance has an equal weight [40]. Then, it fits more copies of the classifier on the same dataset while increasing the weights of inaccurately classified instances and decreasing the weights of correctly classified instances, thus enabling subsequent classifiers to concentrate more on complicated cases. This process continues for a defined number of iterations or until the desired results are obtained. This can be understood by Figure 2. By default, the estimator used by AdaBoost is a DT with one split-level, also called a decision stump (DS) [41,42]. The pseudocode of AdaBoost is presented in Algorithm 1. The detailed explanation of the AdaBoost algorithm is given through the following points:

- **Initialization:** Initially, each sample in the training dataset is assigned an equal weight.
- **Base learning algorithm:** A base learning algorithm, typically a decision tree with a single level (also known as a decision stump), is trained on the weighted dataset. The decision stump aims to find the simplest rule that can classify the data.
- **Weight update:** The misclassified samples are assigned higher weights, while the correctly classified samples are assigned lower weights, emphasizing the importance of the former in subsequent iterations.
- **Iterative process:** AdaBoost repeats the process of training a new base learner on the weighted dataset and updating the sample weights iteratively. Each subsequent learner focuses more on the instances that were misclassified by the previous learners.

- Weighted voting: After all iterations, the final model combines the predictions of all base learners through a weighted voting mechanism, where the weight of each learner's prediction is determined by its accuracy on the training data.
- Output: The final prediction is made by aggregating the weighted predictions of all base learners, resulting in a strong classifier capable of accurately classifying instances, even in the presence of noise or complex patterns.

This approach enables AdaBoost to iteratively improve the model's performance by focusing on the instances that are challenging to classify, ultimately leading to a robust and accurate predictive model for the detection of PD.

Algorithm 1: AdaBoost for the detection of Parkinson's Disease

Input : A training dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_{n-1}, y_{n-1}), (x_n, y_n)\}$;

Base algorithm = L ; Number of iterations = m .

Process:

```

[1]   for  $i \in \{1, 2, \dots, n\}$  do
[2]        $w_i(i) \leftarrow \frac{1}{n}$ 
[3]   end for
[4]        $L \leftarrow \emptyset$ 
[5]   for  $m = 1, 2, \dots, M$  do
[6]        $l_m \leftarrow \arg \min_{l \in L} P_{i \sim w_i}(l(x_i) \neq y_i)$ 
[7]       set
[8]        $err_m \leftarrow P_{i \sim w_i}(l(x_i) \neq y_i)$ 
[9]        $\theta_m \leftarrow \frac{1}{2} \ln \left( \frac{1 - err_m}{err_m} \right)$ 
[10]       $L \leftarrow L \cup \{(\theta_m, l_m)\}$ 
[11]   for  $i \in \{1, 2, \dots, n\}$  do
[12]        $w_{m+1}(i) \leftarrow \frac{w_m(i) e^{-\theta_m y_i l_m(x_i)}}{\sum_{j=1}^n w_m(j) e^{-\theta_m y_j l_m(x_j)}}$ 
[13]   end for
[14]   end for
[15]   Return  $L$ ;
[16]   Output :  $L(x) = \text{sign} \left( \sum_{m=1}^M \theta_m l_m(x) \right)$ 

```

The functioning of AdaBoost can be explained subsequently:

- Initializing weights: For each sample in the dataset, initially the weight is set to N , where N is the total number of instances in the dataset. Sample the dataset using weights.
- Calculating Gini impurity (GI) for each feature variable: GI for each node can be calculated using Equation (3).

$$\text{Gini Impurity} = 1 - (\text{probability of true})^2 - (\text{the probability of false})^2 \quad (3)$$

- The GI for a feature can be computed by taking the weighted average of the impurities at each node. The feature variable with the lowest GI is used to create the first DS.
- Determine the amount of say for the newly created DS: This is performed by calculating the total error, which is equal to the sum of the weights of the misclassified samples, as shown in Equation (4).

$$\text{Amount of say} = \frac{1}{2} \log \frac{1 - \text{total error}}{\text{total error}} \quad (4)$$

- Calculate sample weights for the next DS: In this step, the sample weight for misclassified data points is increased and that of correctly classified instances is decreased by the use of Equations (5) and (6).

$$\text{Sample weight for incorrectly classified data points} = \text{old sample weight} * e^{\text{amount of say}} \quad (5)$$

$$\text{Sample weight for correctly classified data points} = \text{old sample weight} * e^{-\text{amount of say}} \quad (6)$$

The DS focuses more on misclassified data points in the new dataset created from the samples with new weights. The process is continued for a selected number of iterations or until the desired results are acquired [43]. The training process included fitting the AdaBoost model to the training dataset using a few selected hyperparameters, including the number of decision trees ($n_estimators$), the maximum depth of the decision trees (max_depth), and the learning rate ($learning_rate$). Initially, the model was trained on the default values of these hyperparameters, which were later fine-tuned to achieve better results. The best hyperparameter values were identified using grid search cross-validation, in which various combinations of hyperparameter values are systematically explored to finalize the optimum values.

4. Results and Discussion

4.1. Performance Analysis

The results obtained from the proposed model stem from the meticulous methods employed during model creation and testing. Initially, the dataset used for testing encompassed a broad array of features crucial for Parkinson's disease (PD) detection, as evidenced by prior research [25]. Various state-of-the-art methodologies such as SMOTE and PCA were applied during the pre-processing stages, drawing on their proven effectiveness in the previous literature [14–16]. The experimental design incorporated a well-established algorithm known for yielding favorable outcomes in earlier studies [44]. Through the integration of these systematic approaches and the utilization of a carefully curated dataset, this research ensures the statistical relevance of the achieved results. Furthermore, to offer a comprehensive assessment of the proposed model, a detailed description of various performance metrics is discussed in this section.

The AdaBoost classifier has been repeatedly trained on training data, and prediction results have been summarized on the testing data using several confusion matrix based metrics, namely accuracy, precision, recall, F1 score, and false negative rate (FNR) [16,45,46]. Mathematically, the metrics can be represented by Equations (7)–(11), respectively.

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$\text{Precision (P)} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{Recall (R)} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{F1 score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

$$\text{FNR} = \frac{FN}{FN + TP} \quad (11)$$

where TP, FP, TN, and FN are the number of true positives, the number of false positives, the number of true negatives, and the number of false negatives, respectively. The measure of accuracy assesses the percentage of accurately classified instances. Precision (or positive prediction value) refers to the proportion of correct positive predictions out of all positive predictions made by the model, while recall value (or true positive rate or sensitivity) and FNR pertain to the ability of the trained model to accurately identify all the positive instances. The F1 score combines recall and precision and gives the number of times a correct prediction was made by the model [47,48].

In addition to these four metrics, the model's performance was also assessed using the area under the curve of the receiver operating characteristic (AUROC) score. The AUC score measures the area under the ROC curve and represents the ability of the model to distinguish between two classes. Since the research is carried out on a biological dataset, the consequences of FPs and FNs can be severe. Hence, the AUC score helps to assess the performance of diagnostic models better than other metrics as it summarizes the ability to classify positives and negatives correctly. Additionally, the AUC score is robust to class imbalances commonly encountered in biological datasets.

From Figure 3, it can be observed that the accuracy score has significantly increased with the increase in the number of DTs employed in the training of AdaBoost. The increasing trend in accuracy implies that the model can benefit from a larger ensemble of DTs, as a result of which the final model was trained on an ensemble of 500 DTs. Figure 4 depicts that the proposed model is robust to learning rate (LR). The accuracy of the classifier is 80% on an LR of 0.1, with a minor difference of 1% when trained on an LR of 1.0. Consequently, an LR of 1.0 was chosen for the final model to achieve faster convergence while ensuring that each DT in the ensemble contributes fully to the final predictions. Figure 5 shows the increasing trend in performance metrics with respect to the maximum depth of DT chosen for training the model. This upward trend in accuracy aligns seamlessly with the feature-rich nature of the speech dataset. Including deeper DTs in the training of AdaBoost substantiates the demands of the employed dataset. Tables 1–3 display the numerical values of accuracy and various other performance metrics that guided this research towards selecting the optimal values for various hyperparameters. Table 1 demonstrates that other performance metrics except FNR exhibit a positive correlation with respect to the number of DTs in AdaBoost. As these metrics increase with the increasing number of DTs, it is important to highlight that FNR gradually decreases, which is an additional indication of the model's ability to predict positive instances accurately. An analogous trend of increasing accuracy, precision, recall, and F1 score and decreasing FNR can be observed in Table 2, listing the mathematical value of these metrics against different depths of DTs in AdaBoost. However, like the accuracy of AdaBoost with respect to different values of LR, other metrics also remain more or less constant, indicating the model is not sensitive to LR as depicted in Table 3. The results achieved by the tuned model in terms of the various metrics employed in this study are depicted in Table 4. The AUROC curve of the tuned model is shown in Figure 6. The tuned model performed well with an accuracy of 96%, precision of 98%, recall of 93%, F1 score of 95%, FNR of 0.07%, and an AUC score of 0.99.

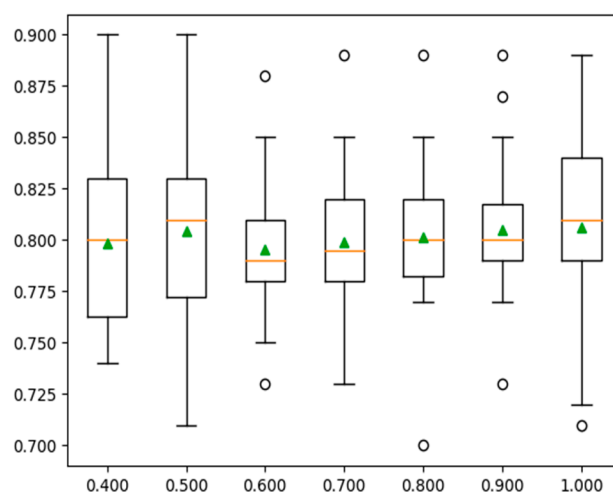


Figure 4. Boxplot depicting accuracies on varying values of LR.

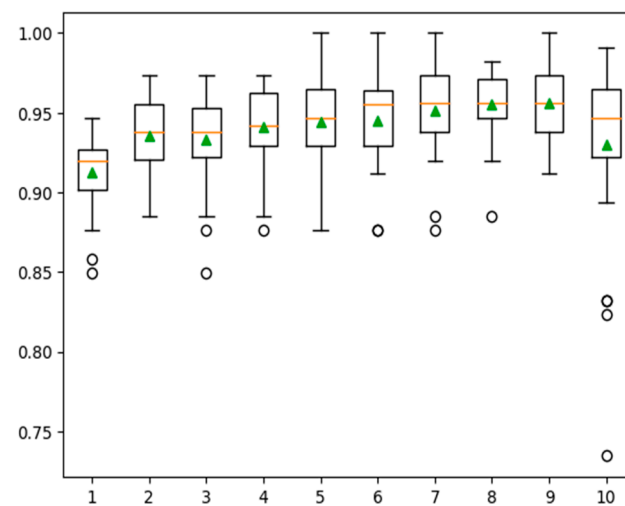


Figure 5. Boxplot depicting accuracies at varying tree depths.

Table 1. Accuracy metrics with respect to different numbers of DTs.

No. of DT	Acc	Precision	Recall	F1 Score	FNR
10	0.81	0.83	0.78	0.80	0.22
50	0.88	0.90	0.85	0.97	0.15
100	0.92	0.94	0.89	0.91	0.11
500	0.96	0.98	0.93	0.95	0.07

Table 2. Accuracy metrics with respect to different tree depths.

Tree Depth	Acc	Precision	Recall	F1 Score	FNR
3	0.93	0.95	0.90	0.92	0.10
5	0.94	0.96	0.91	0.93	0.09
7	0.96	0.98	0.93	0.95	0.07
9	0.96	0.98	0.93	0.95	0.07

Table 3. Accuracy metrics with respect to different values of LR.

LR	Acc	Precision	Recall	F1 Score	FNR
0.1	0.80	0.82	0.77	0.79	0.23
0.5	0.81	0.83	0.78	0.80	0.22
1.0	0.81	0.83	0.78	0.80	0.22

Table 4. Accuracy metrics of the tuned model.

Performance Metrics	Values
Acc	0.96
Precision	0.98
Recall	0.93
F1 score	0.95
FNR	0.07
AUC score	0.99

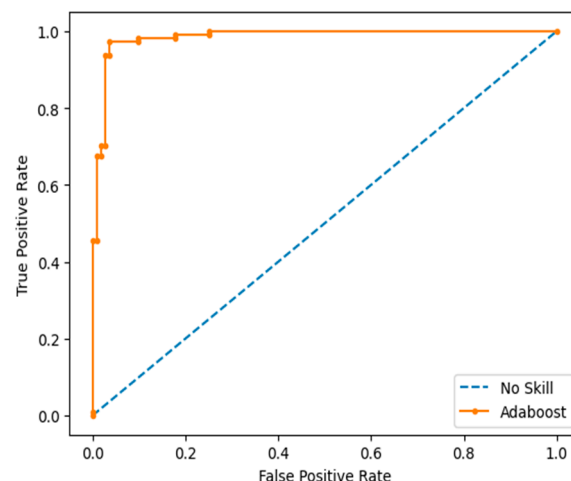


Figure 6. AUROC curve of the tuned model.

4.2. Comparative Analysis

Numerous prior research endeavors have utilized speech-derived characteristics to build a PD detection model. To claim the novelty of the proposed model, a comprehensive comparative analysis has been conducted, and the results are summarized in Table 5. A notable strength of the proposed work lies in the fact that superior results, compared to previous works utilizing deep learning (DL) techniques, have been achieved using an ML model. Using a pre-trained CNN, Inception V3 [12] evaluated their model on AUC, which did not exceed the value of 0.97. Another deep learning based model, DNN2, was proposed by [13], achieving an accuracy of 95.41% and an AUC score of 0.96. Furthermore, the proposed model outperformed existing ML models, as displayed in Table 5. In addition to benchmarking the proposed model with existing ML and DL approaches, it has also been rigorously compared with hybrid approaches in the existing literature. A DL hybrid model that combines CNN and LSTM, as suggested by [21], attained an accuracy of 93.51%. With the aim of combining feature extraction methods with ML classifiers, two hybrid models introduced by [17,18] achieved an accuracy of 95.48% and 95.58%, respectively. With a superior accuracy of 96% and an AUC score of 0.99, the proposed model thrives across well-known performance metrics, qualifying as an effective and innovative solution for the detection of PD using speech signals. These accomplishments have been made possible through the implementation of state-of-the-art pre-processing techniques and leveraging a feature-rich dataset.

In our comprehensive evaluation of the proposed algorithm for PD detection, we have meticulously examined its effectiveness and performance. Based upon the results presented and discussed previously, next we provide a closer examination of the algorithm's strengths and weaknesses, shedding light on both its merits and demerits to offer a comprehensive understanding of its performance and potential limitations.

The proposed algorithm exhibits several noteworthy strengths in the realm of PD detection. Firstly, it demonstrates a commendable level of accuracy, showcasing its efficiency in identifying the PD with high precision. Through the integration of pre-processing techniques such as synthetic minority oversampling technique (SMOTE) and principal component analysis (PCA), the algorithm displays robustness to noise commonly present in clinical data, ensuring reliable performance even in imperfect conditions. Furthermore, its capability to handle large datasets, characterized by a multitude of features, underscores its scalability and potential for application in diverse clinical settings. One of its most significant advantages lies in its non-invasive nature, leveraging speech signals for disease detection. This approach not only simplifies the diagnostic process but also enhances patient comfort and compliance, compared to invasive and time-consuming traditional diagnostic methods.

Table 5. Comparison of the proposed model with recent works.

Study	Feature Set	Pre-Processing Method	Classification Algorithm	Accuracy	AUC Score
Iyer et al., 2023 [12]	Acoustic signals, Spectrogram	Resultant wav files were filtered with floor and ceiling values of 75 decibels (dB) and 300 dB, respectively, for males and 100 dB and 600 dB for females, along with a scaling range of $[-1, 1]$.	Inception v3	NA (Only AUC has been used)	0.97
Rahman et al., 2023 [13]	MDVP, NHR, HNR, RPDE, DFA, PPE, etc.	Random over sampler, PCA	DNN2	95.41%	0.96
Lihore et al., 2023 [21]	MDVP, NHR, HNR, RPDE, DFA, PPE, etc.	dynamic feature breakdown using CNN and LSTM	CNN, LSTM	93.51%	1.0
Mondol et al., 2023 [17]	Frequency, jitter shimmer, HNR, MFCC, etc.	SMOTE	MLP	95.48%	0.98
Lamba et al., 2023 [18]	MDVP, NHR, HNR, RPDE, DFA, PPE, etc.	Genetic algorithm	RF	95.58%	0.98
Proposed model	MFCC, TFF, NHR, HNR, DFA, WTF, VFF, etc.	SMOTE, PCA	AdaBoost	96%	0.98

Despite its promising performance, the proposed algorithm is not without limitations. Prominent among them is its computational complexity, particularly during training phases involving large datasets or intricate pre-processing techniques. This computational demand could pose challenges in resource-constrained environments, limiting the algorithm's practical applicability. Additionally, the algorithm may exhibit sensitivity to certain hyper-parameters, necessitating careful tuning to achieve optimal performance. Without proper regularization techniques, there is also a risk of overfitting, where the model learns noise or irrelevant patterns from the training data, potentially compromising its generalization performance on unseen data. Accordingly, these limitations highlight areas for further research and improvement to enhance the algorithm's effectiveness and practical utility in clinical settings.

5. Conclusions

With the advance of ML in neuroscience, it is possible to not only address the complex challenges of the medical domain but also to solve them. The integration of ML and neuroscience, especially in the context of PD detection, is a significant leap forward towards unraveling the complexities associated with medical data and diagnosis. Identification of PD is not only crucial for gaining deeper insights into the underlying causes of the disease but can also serve as a foundation for initiating timely therapeutic measures and creating suitable remedies. Subsequently, through targeted and personalized treatment, the course of the disease progression can potentially be altered, improving the quality of life for those affected by this disease. However, the conventional assessment of PD through various medical tests is time-consuming and costly for patients. Although there exists no definitive test for the identification of PD, health professionals opt for a combination of tests like MRI, CT, PET, SPECT, DaT scan, blood tests, and clinical evaluation to diagnose this disease. Most of these advanced diagnostic procedures, especially the imaging studies, are expensive and may not always be accessible to all patients. Additionally, the traditional diagnostic practices are reliant on the assessment of motor symptoms like tremors, bradykinesia, and rigidity, causing a delay in the detection process as these appear at later stages of PD. On the other hand, non-motor symptoms that precede the emergence of motor symptoms and

manifest in various forms, including speech impairment, sleep disturbances, autonomic dysfunction, and mood disorders, are often overlooked [22]. The expense of care for PD is high and increases gradually as the disease progresses [49]. In light of the challenges posed by the time-consuming and expensive nature of traditional diagnostic practices, there arises a compelling need for the development of an ML model that leverages non-motor symptoms to enable the detection of PD. The proposed disease detection system not only addresses the challenge of detecting the disease but also serves as a potential avenue to enhance the accessibility and affordability of PD diagnosis. It can also cut the cost of building and maintaining sophisticated laboratories and installing the heavy imaging machinery required in traditional diagnosis. With the availability of large speech-related datasets in future, the model can be further evolved to reduce error as larger datasets with a greater number of features effectively enhance the prediction capacity of ML-based detection systems. Inclusion of other biomarkers, such as genetic information relevant to this disease, into the model can lead to a more precise diagnosis. By incorporating features such as user-friendly interfaces and ensuring compatibility with various healthcare settings, this non-invasive detection system holds the potential to revolutionize the field of clinical diagnostics.

Author Contributions: Conceptualization, S.N.H.B.; methodology, S.N.H.B.; software, S.N.H.B.; validation, K.A.O., S.N.H.B.; formal analysis, S.N.H.B.; investigation, K.A.O.; resources, K.A.O.; data curation, S.N.H.B.; writing—original draft preparation, S.N.H.B.; writing—review and editing, K.A.O.; visualization, K.A.O.; supervision, K.A.O.; project administration, K.A.O.; funding acquisition, K.A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Johannesburg’s University Research Committee (URC) grant for K.A.O_2019, and the Department of Electrical and Electronic Engineering Technology KA_Ogudo Research cost center. and The APC was funded by a grant from the University of Johannesburg Library Research Funds (UJ).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, S.N.H.B., upon reasonable request.

Acknowledgments: The authors acknowledge the support and express gratitude to the Department of Electrical and Electronic Engineering Technology, University of Johannesburg.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Parkinson, J. An essay on the shaking palsy. *J. Neuropsychiatry Clin. Neurosci.* **2002**, *14*, 223–236. [[CrossRef](#)] [[PubMed](#)]
2. Lombardo, J.M.; Lopez, M.A.; Miron, F.; López, M.; León, M.; Arambarri, J.; Álvarez, D. MOBEEZE. Natural interaction technologies, virtual reality and artificial intelligence for gait disorders analysis. *Int. J. Interact. Multimed. Artif. Intell.* **2019**, *5*, 54.
3. Adam, H.; Gopinath, S.C.; Md Arshad, M.K.; Adam, T.; Parmin, N.A.; Husein, I.; Hashim, U. An update on pathogenesis and clinical scenario for Parkinson’s disease: Diagnosis and treatment. *3 Biotech* **2023**, *13*, 142. [[CrossRef](#)]
4. Church, F.C. Treatment options for motor and non-motor symptoms of parkinson’s disease. *Biomolecules* **2021**, *11*, 612. [[CrossRef](#)] [[PubMed](#)]
5. Park, Y.H.; Suh, J.H.; Kim, Y.W.; Kang, D.R.; Shin, J.; Yang, S.N.; Yoon, S.Y. Machine learning based risk prediction for Parkinson’s disease with nationwide health screening data. *Sci. Rep.* **2022**, *12*, 19499. [[CrossRef](#)] [[PubMed](#)]
6. Saeed, F.; Al-Sarem, M.; Al-Mohaimed, M.; Emara, A.; Boulila, W.; Alasli, M.; Ghabban, F. Enhancing Parkinson’s disease prediction using machine learning and feature selection methods. *Comput. Mater. Contin.* **2022**, *71*, 5639–5658. [[CrossRef](#)]
7. Pramanik, A.; Sarker, A. Parkinson’s disease detection from voice and speech data using machine learning. In *International Joint Conference on Advances in Computational Intelligence*; Springer: Singapore, 2021.
8. Cherubini, A.; Nistico, R.; Novellino, F.; Salsone, M.; Nigro, S.; Donzuso, G.; Quattrone, A. Magnetic resonance support vector machine discriminates essential tremor with rest tremor from tremor-dominant Parkinson disease. *Mov. Disord.* **2014**, *29*, 1216–1219. [[CrossRef](#)] [[PubMed](#)]
9. Moro-Velazquez, L.; Garcia, J.A.G.; Arias-Londono, J.D.; Dehak, N.; Godino-Llorente, J.I. Advances in Parkinson’s disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomed. Signal Process. Control.* **2021**, *66*, 102418. [[CrossRef](#)]

10. Narendra, N.P.; Schuller, B.; Alku, P. The detection of Parkinson's disease from speech using voice source information. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2021**, *29*, 1925–1936. [\[CrossRef\]](#)
11. Almahadin, G.; Lotfi, A.; Carthy, M.M.; Breedon, P. Task-oriented intelligent solution to measure Parkinson's disease tremor severity. *J. Healthc. Eng.* **2021**, *4*, 9624386. [\[CrossRef\]](#)
12. Iyer, A.; Kemp, A.; Rahmatallah, Y.; Pillai, L.; Glover, A.; Prior, F.; Larson-Prior, L.; Virmani, T. A machine learning method to process voice samples for identification of Parkinson's disease. *Sci. Rep.* **2023**, *13*, 20615. [\[CrossRef\]](#)
13. Rahman, S.; Hasan, M.; Sarkar, A.K.; Khan, F. Classification of Parkinson's disease using speech signal with machine learning and deep learning approaches. *Eur. J. Electr. Eng. Comput. Sci.* **2023**, *7*, 20–27. [\[CrossRef\]](#)
14. Alshammri, R.; Alharbi, G.; Alharbi, E.; Almubark, I. Machine learning approaches to identify Parkinson's disease using voice signal features. *Front. Artif. Intell.* **2023**, *6*, 1084001. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Govindu, A.; Palwe, S.A. Early detection of Parkinson's disease using machine learning. *Preced. Comput. Sci.* **2023**, *218*, 249–261. [\[CrossRef\]](#)
16. Alalayah, K.M.; Senan, E.M.; Altam, H.F.; Ahmed, I.A.; Shatnawi, H.S.A. Automatic and early detection of Parkinson's disease by analyzing acoustic signals using classification algorithms based on recursive feature elimination method. *Diagnostics* **2023**, *13*, 1924. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Mondol, S.R.; Kim, R.; Lee, S. Hybrid machine learning framework for multistage Parkinson's disease classification using acoustic features of sustained Korean vowels. *Bioengineering* **2023**, *10*, 984. [\[CrossRef\]](#)
18. Lamba, R.; Gulati, T.; Alharbi, H.; Jain, A. A hybrid system for Parkinson's disease diagnosis using machine learning techniques. *Int. J. Speech Technol.* **2022**, *25*, 583–593. [\[CrossRef\]](#)
19. Das, R. A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Syst. Appl.* **2010**, *37*, 1568–1572. [\[CrossRef\]](#)
20. Rehman, A.; Saba, T.; Mujahid, M.; Alamri, F.; ElHakim, N. Parkinson's disease detection using hybrid LSTM-GRU deep learning model. *Electronics* **2023**, *12*, 2856. [\[CrossRef\]](#)
21. Lihore, U.K.; Dalal, S.; Faujdar, N.; Margala, M. Hybrid CNN-LSTM model with efficient hyperparameter tuning for prediction of Parkinson's disease. *Sci. Rep.* **2023**, *13*, 14605. [\[CrossRef\]](#)
22. Goyal, J.; Khandnor, P.; Aseri, T.C. A Hybrid Approach for Parkinson's Disease diagnosis with resonance and time-frequency based features from speech signals. *Expert Syst. Appl.* **2021**, *182*, 115283. [\[CrossRef\]](#)
23. Chowdhary, C.L.; Srivatsan, R. Non-invasive detection of Parkinson's disease using deep learning. *Int. J. Image Graph. Signal Process.* **2022**, *14*, 38–46. [\[CrossRef\]](#)
24. Asuroglu, T.; Ogul, H. A deep learning approach for parkinson's disease severity assessment. *Health Technol.* **2022**, *12*, 943–953. [\[CrossRef\]](#)
25. Lamba, R.; Gulati, T.; Jain, A. An intelligent system for Parkinson's diagnosis using hybrid feature selection approach. *Int. J. Softw. Innov.* **2022**, *10*, 1–13. [\[CrossRef\]](#)
26. Liu, X.; Li, W.; Liu, Z.; Du, F.; Zou, Q. A dual-branch model for diagnosis of Parkinson's disease based on the independent and joint features of the left and right gait. *Appl. Intell.* **2021**, *51*, 7221–7232. [\[CrossRef\]](#)
27. Shabu, S.J.; Sivapriya, V.; Refonaa, J.; Dhamodaran, S.; Poornima. Parkinson's Disease Detection using Machine Learning Algorithm. In Proceedings of the 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 1–3 June 2023; pp. 990–997. [\[CrossRef\]](#)
28. Latifoğlu, F.; Penekli, S.; Orhanbulucu, F.; Chowdhury, M.E.H. A novel approach for Parkinson's disease detection using Vold-Kalman order filtering and machine learning algorithms. *Neural Comput. Appl.* **2024**, *36*, 9297–9311. [\[CrossRef\]](#)
29. Tran, C.; Shen, K.; Liu, K.; Ashok, A.; Ramirez-Zamora, A.; Chen, J.; Li, Y.; Fang, R. Deep learning predicts prevalent and incident Parkinson's disease from UK Biobank fundus imaging. *Sci. Rep.* **2024**, *14*, 3637. [\[CrossRef\]](#)
30. Fenza, G.; Gallo, M.; Loia, V.; Orciuoli, F.; Herrera-Viedma, E. Data set quality in machine learning: Consistency measure based on Group Decision Making. *Appl. Soft Comput.* **2021**, *106*, 107366. [\[CrossRef\]](#)
31. Şakar, C.O.; Serbes, G.; Gunduz, A.; Tunc, H.C. A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Appl. Soft Comput.* **2019**, *74*, 255–263. [\[CrossRef\]](#)
32. Yin, S.; Liu, C.; Zhang, Z.; Lin, Y.; Wang, D.; Tejedor, J.; Zheng, T.F.; Li, Y. Noisy training for deep neural networks in speech recognition. *EURASIP J. Audio Speech Music. Process.* **2015**, *2015*, 2. [\[CrossRef\]](#)
33. Toh, M.; Togneri, R.; Nordholm, S. Spectral entropy as speech features for speech recognition. *Proc. PEECS* **2005**, *1*, 92.
34. Kumalija, E.; Nakamoto, Y. Performance evaluation of automatic speech recognition systems on integrated noise-network distorted speech. *Front. Signal Process.* **2022**, *2*, 999457. [\[CrossRef\]](#)
35. Chawla, N.V.; Bowyer, K.; Lawrence, H.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
36. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
37. Hubel, D.H.; Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **1962**, *160*, 106–154. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 498–520. [\[CrossRef\]](#)

39. Dietterich, T.G. Ensemble methods in machine learning. In *International Workshop on Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2000.
40. Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. In *Proceedings of the International Conference on Machine Learning*, Bari, Italy, 3–6 July 1996.
41. Bukhari, S.N.H.; Webber, J.; Mehbodniya, A. Decision tree based ensemble machine learning model for the prediction of Zika virus T-cell epitopes as potential vaccine candidates. *Sci. Rep.* **2022**, *12*, 7810. [[CrossRef](#)] [[PubMed](#)]
42. Şentürk, Z.K. Early diagnosis of Parkinson's disease using machine learning algorithms. *Med. Hypotheses* **2020**, *138*, 109603. [[CrossRef](#)] [[PubMed](#)]
43. Bukhari, S.N.H.; Jain, A.; Haq, E.; Mehbodniya, A.; Webber, J. Ensemble machine learning model to predict SARS-CoV-2 t-cell epitopes as potential vaccine targets. *Diagnostics* **2021**, *11*, 1990. [[CrossRef](#)]
44. Nour, M.; Şentürk, Ü.K.; Polat, K. Diagnosis and classification of Parkinson's disease using ensemble learning and 1D-PDCovNN. *Comput. Biol. Med.* **2023**, *161*, 107031. [[CrossRef](#)]
45. Anoruo, C.M.; Bukhari, S.N.H.; Nwofor, O.K. Modeling and spatial characterization of aerosols at Middle East AERONET stations. *Theor. Appl. Climatol.* **2023**, *152*, 617–625. [[CrossRef](#)]
46. Masoodi, F.; Quasim, M.; Bukhari, S.N.H.; Dixit, S.; Alam, S. *Applications of Machine Learning and Deep Learning on Biological Data*; CRC Press: Boca Raton, FL, USA, 2023.
47. Bukhari, S.N.H.; Masoodi, F.; Dar, M.A.; Wani, N.I.; Sajad, A.; Hussain, G. Prediction of erythemato-squamous diseases using machine learning. In *Applications of Machine Learning and Deep Learning on Biological Data*; CRC Press and Taylor & Francis: Boca Raton, FL, USA, 2023; pp. 87–96.
48. Bukhari, S.N.H.; Jain, A.; Haq, E. A Novel Ensemble Machine Learning Model for Prediction of Zika Virus T-Cell Epitopes. In *Proceedings of Data Analytics and Management*; Gupta, D., Polkowski, Z., Khanna, A., Bhattacharyya, S., Castillo, O., Eds.; Springer: Singapore, 2022; pp. 275–292.
49. Ola, V.; Puri, I.; Goswami, D.; Vibha, D.; Shukla, G.; Goyal, V.; Srivastava, A.; Behari, M. Annual cost of care of Parkinson's Disease and its determinants in North India—A cost of illness study with patient perspective. *Ann. Indian Acad. Neurol.* **2022**, *25*, 660. [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.