

1-1-2022

Classification of parkinson disease based on patient's voice signal using machine learning

Imran Ahmed
Riphah International University

Sultan Aljahdali
Taif University

Muhammad Shakeel Khan
Riphah International University

Sanaa Kaddoura
Zayed University, sanaa.kaddoura@zu.ac.ae

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Computer Sciences Commons](#), and the [Medicine and Health Sciences Commons](#)

Recommended Citation

Ahmed, Imran; Aljahdali, Sultan; Khan, Muhammad Shakeel; and Kaddoura, Sanaa, "Classification of parkinson disease based on patient's voice signal using machine learning" (2022). *All Works*. 4673. <https://zuscholars.zu.ac.ae/works/4673>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact scholars@zu.ac.ae.

Classification of Parkinson Disease Based on Patient's Voice Signal Using Machine Learning

Imran Ahmed¹, Sultan Aljahdali², Muhammad Shakeel Khan¹ and Sanaa Kaddoura^{3,*}

¹Riphah International University, Lahore, 54000, Pakistan

²Department of Computer Science, Taif University, Taif, KSA

³College of Technological Innovation, Zayed University, Abu Dhabi, UAE

*Corresponding Author: Sanaa Kaddoura. Email: Sanaa.kaddoura@zu.ac.ae

Received: 26 July 2021; Accepted: 13 September 2021

Abstract: Parkinson's disease (PD) is a nervous system disorder first described as a neurological condition in 1817. It is one of the more prevalent diseases in the elderly, and Alzheimer's is the second most common neurodegenerative illness. It impacts the patient's movement. Symptoms start gradually with tremors, stiffness in movement, and speech and voice disorders. Researches proved that 89% of patients with Parkinson's has speech disorder including uncertain articulation, hoarse and breathy voice and monotone pitch. The cause behind this voice change is the reduction of dopamine due to damage of neurons in the substantia nigra responsible for dopamine production. In this work, Parkinson's disease is classified with the help of human voice signals. Six different machine learning (ML) algorithms are used in the classification: Stochastic Gradient Descent (SGD) Classifier, Extreme Gradient Boosting (XGB) Classifier, Logistic Regression Classifier, Random Forest Classifier, K-Nearest Neighbour (KNN) Classifier, and Decision Tree (DT) Classifier. This research aims to classify Parkinson's disease using human voice signals and extract essential features to reduce the complexity of the dataset. Then, human voice signals are analyzed to check the voice intensity and spectrum for PD patients. Then, machine learning classifiers are applied to classify the PD patients based on the extracted features. The results show that SGD-Classifier has 91% accuracy, XGB-Classifier has 95% accuracy, Logistic Regression has 91% accuracy, Random Forest shows 97% accuracy, KNN shows 95% accuracy, and Decision Tree has 95% accuracy. Hence, Random Forest has the highest accuracy. The disease can be studied more by looking for more characteristics of PD patients to enhance its proper use in the medical field.

Keywords: SGD-classifier; XGB-classifier; logistic regression; random forest; KNN-classifier; parkinson disease; decision tree classifier

1 Introduction

Substantia nigra is an essential part of the brain which is considered the direct cause of Parkinson's disease. Substantia nigra releases dopamine that acts as a messenger between portions of the nervous



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

system and the brain to coordinate the body's actions. The amount of dopamine in the brain is lowered if these nerve cells die. This indicates that the brain region that controls movement is not working as well as it should, resulting in slow and irregular movements.

Nerve cell death is a gradual process. When roughly 80% of the nerve cells in the substantia nigra have been destroyed, the symptoms of Parkinson's disease begin to appear. Parkinson's disease affects the voices of men, affecting them in mating or perhaps in conversations. PD is only second to Alzheimer's disease in neurological disorders [1]. Expansion is expected in the coming years, and appropriate testing and frameworks are needed along these lines to identify logical remedies. More studies should be done in this area since appropriate measures are essential for detecting and treating the disease [2]. The ideal characteristics of data sets are transferred as a commitment to models, and forecast results are generated. The expected exhibition of substantia nigra (see Fig. 1) is a group in the middle of the brain responsible for producing dopamine.

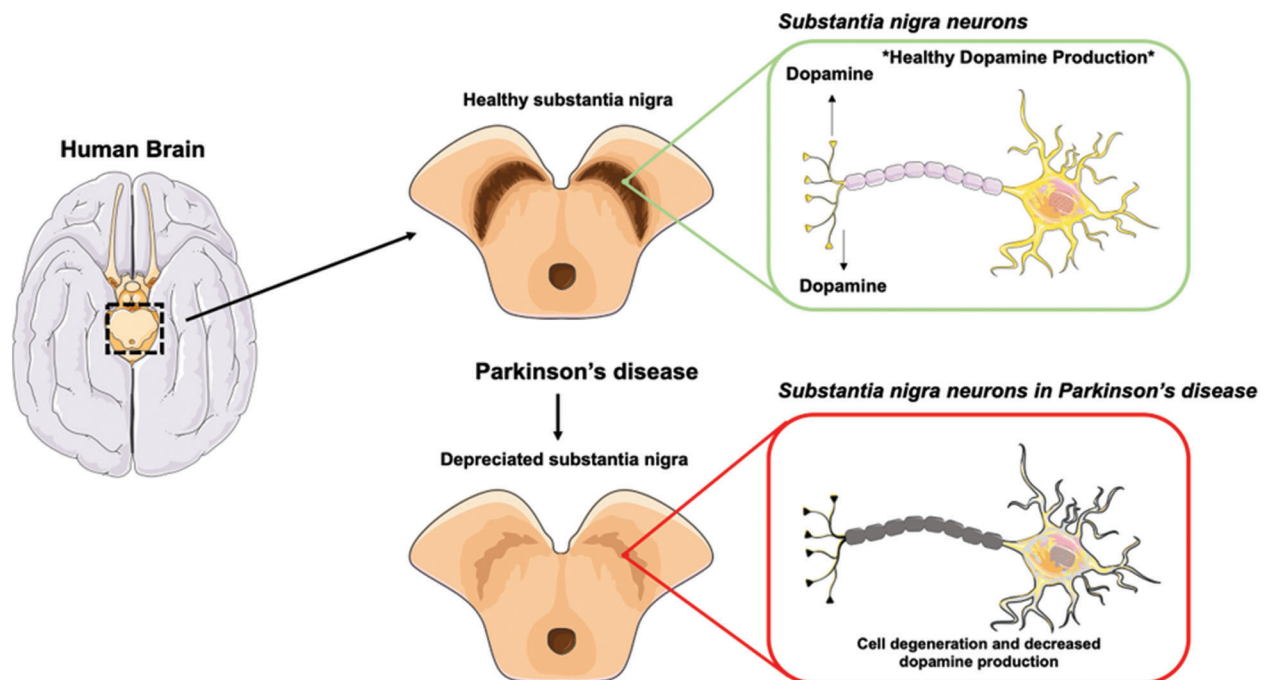


Figure 1: Substantia nigra neurons in Parkinson's disease

The six different machine learning models give better results than previous work in literature. The algorithms used in this work are:

- SGD-Classifer
- XGB-Classifer
- Logistic Regression Classifier
- Random Forest Classifier
- KNN-Classifer
- Decision Tree Classifier

Gradient Descent is the most used machine learning optimization approach. It's a first-order optimization procedure. It means that while updating the parameters, it only takes into consideration the first derivative. The parameters are in the opposite direction of the objective function $J(w)$ gradient with

respect to the parameters on each iteration, where the gradient indicates the steepest ascent direction. The learning rate determines the amount of the step we take each iteration to reach the local minimum. As a result, we descend in the direction of the slope until we hit a local minimum [3,4]. XGBoost is part of a family of boosting algorithms that help weak learners improve their performance. Random guessing is marginally better than a slow learner. Boosting is a sequential process in which trees are grown one after the other, using information from a previously developed tree. This method slowly learns from the data in successive cycles and improves its forecast [5]. A random forest comprises a vast number of individual decision trees that work together as an ensemble.

Each tree in the random forest produces a class prediction, and the class with the most votes becomes the prediction of our model. The KNN is a non-parametric technique proposed by the Thomas cover used in assembly and relapse in the confirmation design. In both applications, the information consists of the nearest K setting forms in the component space. The yield depends on whether KNN is used for regulation or relapse. The component space is grouped by the majority of the votes, indicated by the nearest-distance-neighbor [5]. A Decision Tree is a non-parametric learning strategy used in characterization and regression. The aim is to develop a model that predicts objective variable points by introducing the essential criteria for the selection collected from the information covered [6].

2 Related Works

2.1 Classification of Parkinson Disease

The project in [6] aims to classify PD patients by analyzing their gait results. The data used was derived from the database of Physio net Gait Analysis. The database includes data on 93 idiopathic PD patients and 73 stable control subjects. It consists of the subjects' Vertical Ground Reaction Force as they walk for about two minutes. Each subject has 16 sensors. Eight sensors were placed on the bottom of each foot, which measures Newton's force as a function of time. Some mathematical functions were used to improve data accuracy and decrease the necessary computing capacity. This reduced data in the range of $[-1, 1]$ while maintaining the data singularity unchanged. The new data reduction was then used to create machine learning models using machine learning algorithms. The highest accuracy of KNN is 93.08%.

As compared to previous research, PD was established differently. MATLAB graphical user interface was used to incorporate the machine learning method in the application. Also, a new program called ParkDet 2.0 was developed. Then, to increase the accuracy of the classifier, different combinations of ML were applied using ParkDet. The goal is to enable clinician technicians to use the ParkDet software to diagnose PD patients and not force them to use advanced engineering systems such as MATLAB and automated detection methods.

Several variations of ML processes, such as PCA and factor analysis (FA), were introduced via the ParkDet software. Additionally, seven classifiers were established: Support Vector Machine (SVM), Boosting, KNN, Naive Bayes (NB), Decision Tree, Linear Discriminant Analysis (LDA), and Quadratic Discriminant Analysis (QDA) with 2-, 5- and 10-fold cross-validations and an additional blind examination. Since the classification accuracy hits 90.1%, the findings are promising [2].

2.2 Symptoms of Parkinson Disease

The neuropathological condition of PD or other types of Parkinsonism may result in Parkinson's disease. About 90% of the cases are intermittent with no specific etiology for neuropathologic PD; a further 10% have a hereditary cause, and at least 11 separate lineages have been reported with six gene mutations. Genetic variants of PD in young-onset PD are observed. Medicines, sequelae to central nervous system inflammation, toxins, or vascular/metabolic conditions may be responsible for secondary forms of Parkinsonism [7-9]. This paper's participants underwent baseline in-clinic assessments, including the

Unified Parkinson's Disease Rating Scale (UPDRS). They were provided smartphones with an Android operating system that contained a smartphone application to assess voice, posture, gait, finger tapping, and response time. Twenty participants performed an average of 2.7 tests per day (68.9% adherence) for the study duration (average of 34.4 days) in a home and community setting. The analyses of the five tasks differed between those with Parkinson's disease and those without it. In discriminating participants with PD from controls, the mean sensitivity was 96.2% (standard deviation 2%), and mean specificity was 96.9% (standard deviation 1.9%).

2.3 Voice Frequencies of Parkinson Disease Affected Patients with Parkinson Disease

Some research focuses on the tonal level, i.e., the fundamental frequency (F0) variations produced by the laryngeal activity. The significant improvements in F0 phonation in subjects with sensorineural hearing loss (SNHL) and PD using power spectral analysis of F0 have been studied in [8–10]. For the two disorders, F0's rhythm was distinct in frequency and amplitude. This research also revealed that the F0 analysis could become a valuable instrument for the two neurological disorders to be tested. By using the autocorrelation function algorithm, fundamental frequencies of voice signals were obtained. They were then transformed to a series of cents after linear interpolation and re-sampling to a 50 Hz scale [8–10]. The hypothesis states that the Parkinsonian dysprosody has often been described as a mere neuro-motor disturbance. To verify the hypothesis, both the comprehension and the production of pitch parameters in a group of Parkinsonian patients were tested. The influence of traditional anti-parkinsonian treatments, L-DOPA which is an amino acid precursor of dopamine and is the most effective and commonly used drug in the treatment of PD, was also studied to determine whether improved performance occur [11].

Deep learning is used in [12] to evaluate the patient's voice data and divide it into "severe" or "not severe" classes. The two Unified Parkinson's Disease Ranking Scale (UPDRS) scores are the assessment measures used in this analysis. On a scale of 0–108, the motor UPDRS tests the patient's motor capacity, while total UPDRS offers a wider variety of scores from 0–1766. The 16 biomedical voice functions are the input dataset, and the output vector is the Motor UPDRS score. The classification accuracy achieved for train and test datasets is 83.367% and 81.6667%, respectively. Since they tested their model on the same set of UCI Parkinson's Telemonitoring Voice Results, they contrasted their result with that of Nilashi et al. in forecasting the progression of Parkinson's disease, and they used adaptive neuro-fuzzy interference system and support vector regression. Their work provided an average accuracy of 47.2 percent for the overall UPDRS ranking [7–12].

2.4 Causes of Parkinson Disease

This research demonstrated that pesticides affect PD production. Priyadarshi et al. [13] conducted a meta-analysis to analyze multiple studies examining environmental risk factors to produce PD. Findings show a greater risk of PD for individuals living in rural areas using wells as their water source. However, the environmental hazards causing PD could not be identified. Indeed, as applied to rural and urban areas, the findings yielded an insignificant association between risk factors. Due to the differences between the results of different trials, improved data collection techniques tend to be required to understand better possible causes of PD [13,14]. [Tab. 1](#) summarizes the literature review.

3 Proposed Work

Previous research stated that Parkinson's disease had been spread all over the world. Doctors use the most extensive equipment for physical diagnosis, which is a very time-consuming and non-accurate process. So, we will classify Parkinson's disease using human voice signal frequencies and explore it in three phases, as shown in [Fig. 2](#). Firstly, we are extracting some essential features to classify for

understanding. Secondly, we apply data mining techniques to classify the healthy/affected patients based on some voice features to generate the results in graphs and tables, accuracy-score. Thirdly, we are going to make a comparison of all machine learning algorithms to find out the best accuracy result algorithm.

Table 1: Tabulated literature review

Reference	Features/Factors	Contribution
[7]	Self-selected speed-walking	Artificial Neural Network and SVM showed perfect ability to distinguish gait patterns using three parameters: spatiotemporal, kinematic, and kinetic.
[2]	Variations of dysphonia, gait, or motor skills	Using ML methods on vocal features, clinicians use a blind test interface for PD detection.
[10]	Smartphone monitoring voice, posture, gait, finger tapping	Analyzing tests results for participants with/without PD shows the feasibility of diagnosing PD.
[11]	Gait patterns, gait initiation, and freezing of gait	Using behavioral factors, detecting and classifying tremors shows high efficiency in distinguishing healthy people from PD using voice data.
[12]	Rapid Eye Movement and olfactory loss	Using non-motor features, automated diagnosis models from ML techniques show efficiency in the early prediction of PD.
[15]	Tele-monitoring voice data	Using “TensorFlow” deep neural networks, prediction of PD severity shows high accuracy values.

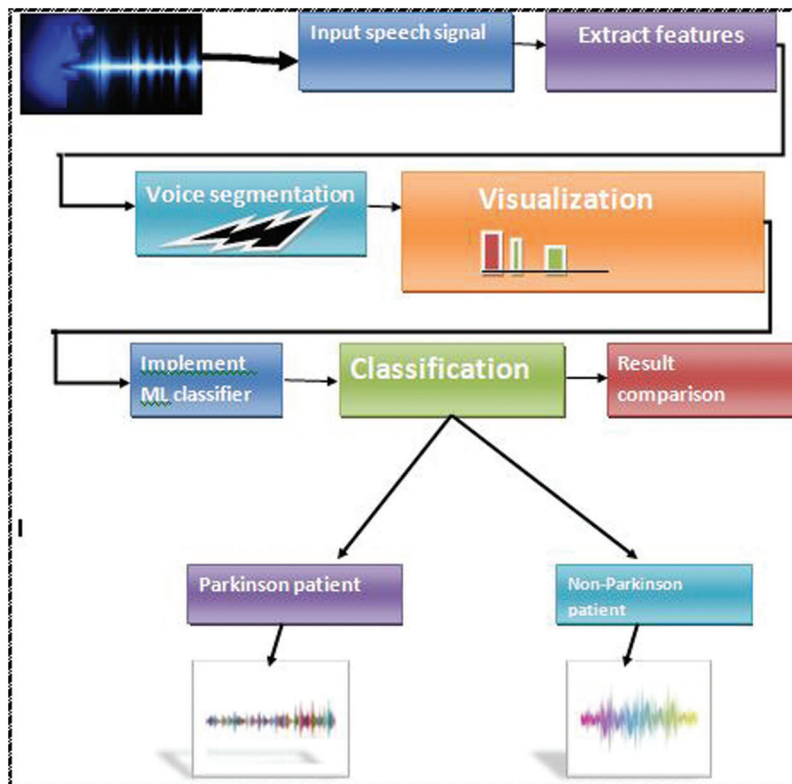


Figure 2: Proposed workflow

3.1 Place of Work and Facility

Our research is based on qualitative analysis rather than quantitative. We will use the qualitative data analysis online tool. All research will be done at the university and in our personal space.

3.2 Feature Extraction

Feature extraction increases the accuracy of learned models by extracting features from input data. This phase reduces the dimensionality of data by removing the redundant records. Of course, it enhances the classification speed. Feature extraction helps get the best feature from those big data sets by selecting and combining variables into features, thus, effectively reducing the amount of data. These features are easy to process but still able to describe the actual data set with accuracy and originality.

3.3 Feature Selection

Feature selection is an essential approach for reducing the dimension of high-dimensional data. In recent years, many feature selection algorithms have been proposed. However, most of them only exploit information from the data space. They often neglect helpful information contained in the feature space and typically do not exploit information about the underlying geometry of the data. To overcome these problems, we introduce new unsupervised feature selection methods based on feature selection.

3.4 Univariate Selection

Statistical tests can select those features with the most robust relationship with the output variable. The scikit-learn library provides the SelectKBest class for this purpose.

3.5 Voice Segmentation

In this phase, we classify the voice in segments and find out the relationship of features so used in which seven extract features for classification of the human voice. Based on these seven features, we identify the ranges of frequencies and compare them with patients' health status.

3.6 Visualization

According to the voice segments classes, we classify frequencies with visualization. Data visualization helps to analyze the measurement reports. Visualization is a method for transmitting a message by generating pictures or animations. Visualization is any technique for creating images, diagram plots, charts, and animations. In these analyses, we analyzed the human voice frequencies into the graph box plot and tabular form and compared these all frequencies one by one with patients' health status and identified the ranges of all frequencies, and measured it with health status.

4 Findings and Results

4.1 Input Dataset

The voice signal dataset [15] display some rows and columns of data in Fig. 3, showing head (5) of the dataset, which displayed all attributes of the data and values in float. This dataset comprises a range of biomedical tone of voice estimations from 31 individuals, of which 23 have PD. Every section in the table is a specific voice measure, and each line compares to one of the voice chronicles from these people ("name" segment). The "status" segment is set to 0 for non-patient and 1 for PD [16].

4.2 Average Local Fundamental Frequencies (fo)

In this condition, most people with Parkinson's disease will change vocabulary, voice, and swallowing. The exact effects of Parkinson's disease that appear in the human body's muscles, tremor, stiffness, slow

motion, and slow speech, can occur in the muscles used to speak and swallow. In the medical treatment of reading the human voice with a machine those machines, read in different frequencies in which one of the most critical frequencies of fundamental vocal frequencies.

	name	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP	MDVP:Shimmer	...	S
0	phon_R01_S01_1	119.992	157.302	74.997	0.00784	0.00007	0.00370	0.00554	0.01109	0.04374	...	
1	phon_R01_S01_2	122.400	148.650	113.819	0.00968	0.00008	0.00465	0.00696	0.01394	0.06134	...	
2	phon_R01_S01_3	116.682	131.111	111.555	0.01050	0.00009	0.00544	0.00781	0.01633	0.05233	...	
3	phon_R01_S01_4	116.676	137.871	111.366	0.00997	0.00009	0.00502	0.00698	0.01505	0.05492	...	
4	phon_R01_S01_5	116.014	141.781	110.655	0.01284	0.00011	0.00655	0.00908	0.01966	0.06425	...	

5 rows x 24 columns

Figure 3: Head of 5 elements of the dataset

Fig. 4 shows the fundamental vocal frequencies (Fo) of the human voice. The ranges of frequencies 100–250 define the human voice. This range increases when the person is affected with PD. Fig. 5 shows the maximum vocal fundamental frequency (Fhi(Hz)). The ranges are higher than the frequencies shown in Fig. 4 because these frequencies measure the vocal voice length. Fig. 6 shows the minimum vocal fundamental frequency (Flo). It visualizes the ranges of frequencies signal of human voices.



Figure 4: MDVP: Fo(Hz) Average Vocal fundamental frequencies of the human voice

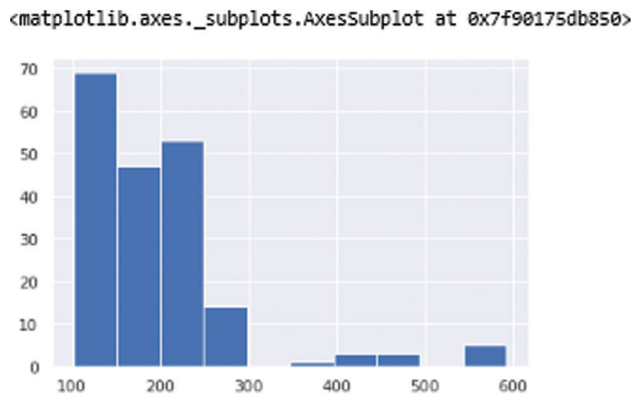


Figure 5: MDVP: Fhi(Hz) in these vocal frequencies

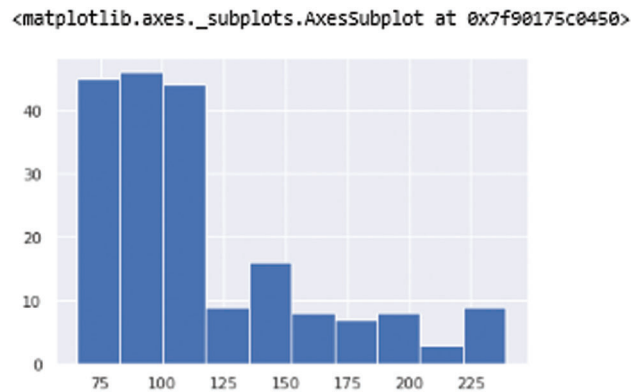


Figure 6: MDVP: Flo(Hz) the minimum vocal fundamental frequency of vocal frequencies

The fundamental vocal frequencies are classified to check the patient health status and interpret the frequency. Fig. 7 shows the results of all vocal frequencies. Patient health status in 0 represents a non-Parkinson, whereas 1 is the Parkinson's patient. The vocal frequencies of non-patient is 0.006, which is considered low compared to the affected patient which highest vocal frequencies range is 0.0010.

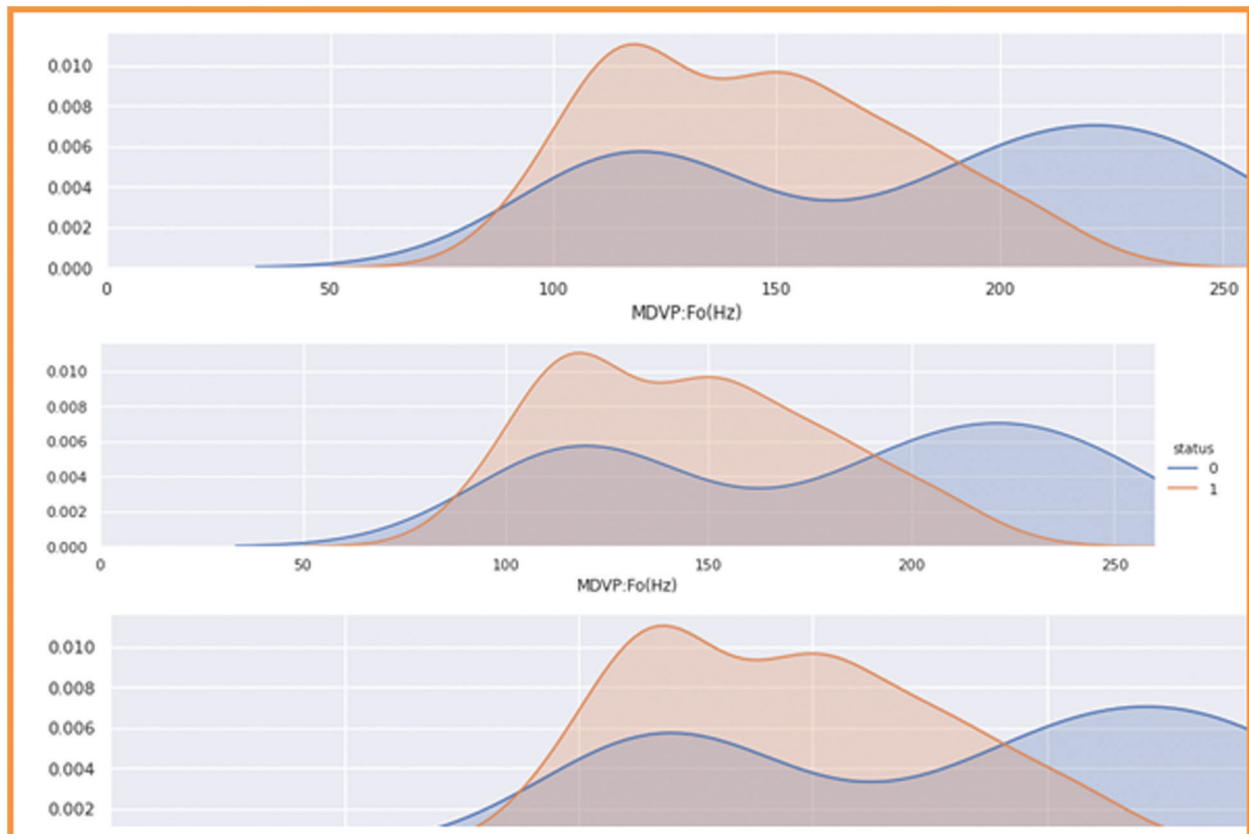


Figure 7: Results of all focal frequencies

4.3 Harmonic Parameter Frequencies

In these frequencies in which harmonic-noise ratio and harmonic noise ratio of two types of frequencies are available Kinematic changes in PD gait are well documented; however, there is less comprehension of upper body dynamics. The rhythm of trunk accelerations is measured by harmonic ratios (HRs) and can be studied in longitudinal, anterior-posterior, and mediolateral planes, providing an indicator of global walking stability (lower HR indicates poorer stability) (see Figs. 8 and 9).

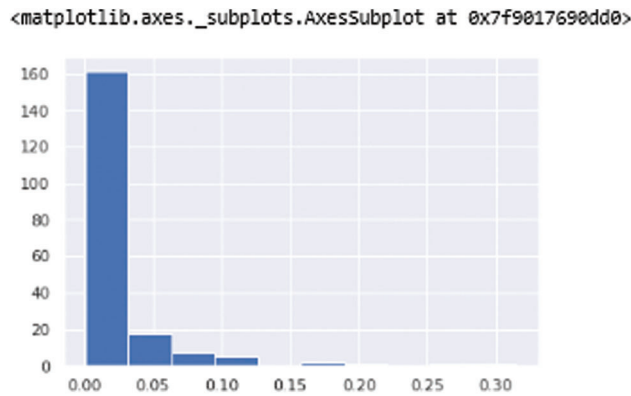


Figure 8: NHR—NHR (Noise Harmonic ratio)

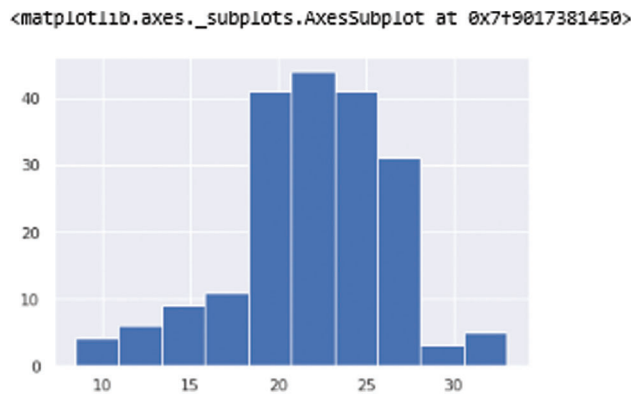


Figure 9: HNR—The HNR (Harmonic-noise ratio)

Now here is the comparison of the harmonic parameter ratio of the human voice. The affected patients and non-affected patients are classified with these two frequencies of ratio. Fig. 10 displays these frequencies ratio. The blue color is healthy person whereas orange color is for patients.

The Harmonic ratio of PD was compared with the patient health status, clearly showing 0 is non-patient, and 1 is positively affected with Parkinson’s person. Hence, in the first figure, the Harmonic noise ratio of the non-affected person is high because the non-Parkinson person can quickly speak, and the noise ratio is high, so there is no gap in this voice. In non-patient, the harmonic ratio is low because its speaking ratio is low, creating gaps in the voice. Thus, the ratio goes down in the second figure.

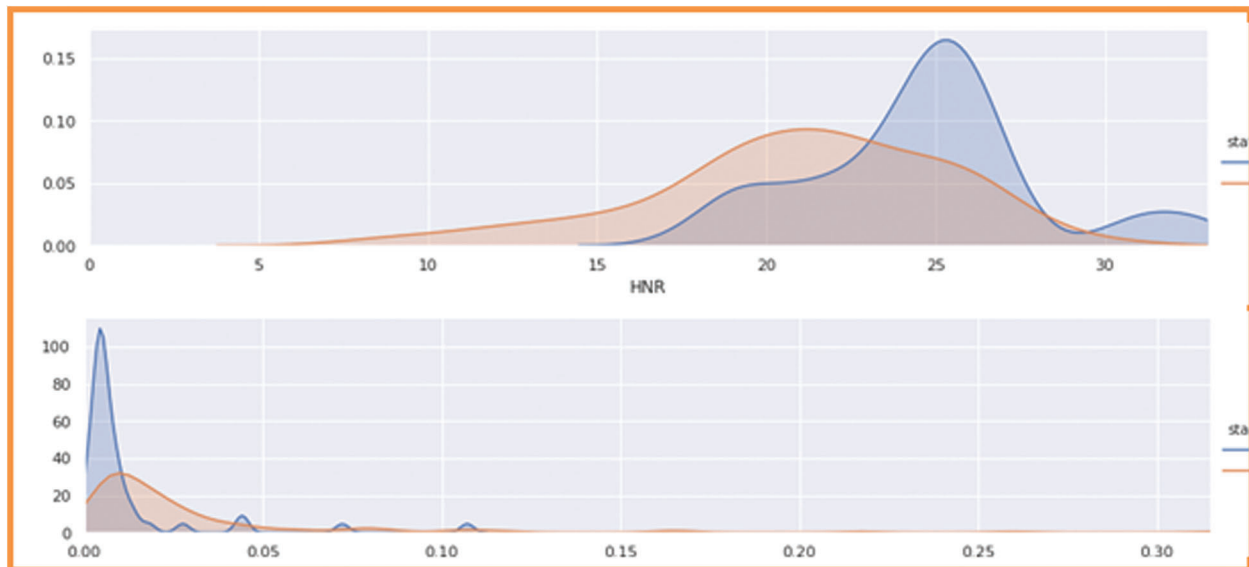


Figure 10: Harmonic parameter ratio with patient health status

4.4 Non-linear Dynamic Complexity Measure

DFA (Detrended fluctuation analysis) (shown in Fig. 11) and PPE (pitch period of entropy) (shown in Fig. 12) show the Parkinson's effect on the human voice. The frequencies are break down because of the slow speaking flow of the human and sibling in the voice when speaking.

<matplotlib.axes._subplots.AxesSubplot at 0x7f9017541b50>

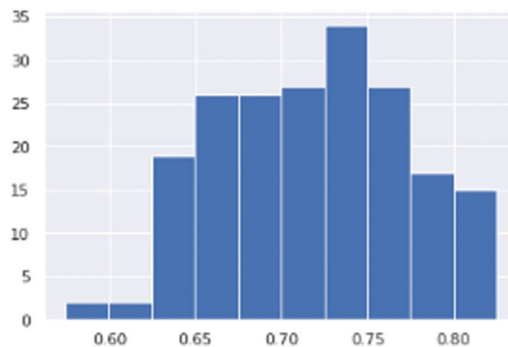


Figure 11: DFA

<matplotlib.axes._subplots.AxesSubplot at 0x7f901e83f6d0>

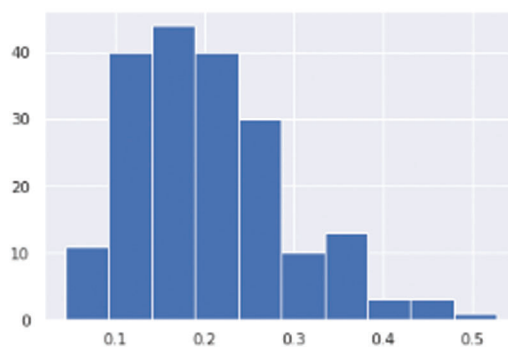


Figure 12: PPE

Fig. 13 shows the results representing the human voice signal. The first chart is DFA in which 0 is non-affected patients. The voice is higher than the affected patient of Parkinson’s disease.

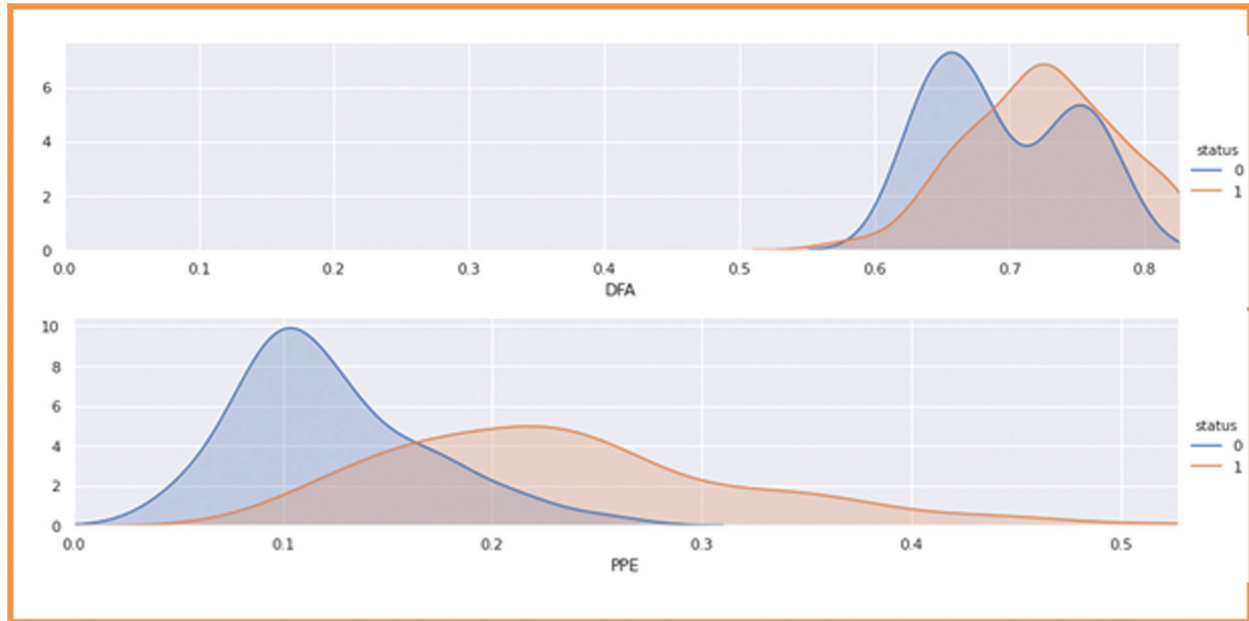


Figure 13: Classified the non-linear dynamical complexity measure with patient health

4.5 Classification of PD with ML Classifier

In this process, we are going to classify Parkinson’s disease with a machine learning classifier. First, we set the target of patient health status and count how many patients are available in this report. After checking the patient’s health status, we visualize the results graphically. The dataset was split into 80% for the training phase whereas 20% for the testing phase. In Fig. 14, 0 represents the healthy people who are 48 in the dataset, whereas 1 represents the patients who are 147 in the dataset.

- Number of Parkinson people in the dataset 147 (75.39%)
- Number of Healthy people in the dataset 48 (24.61%)

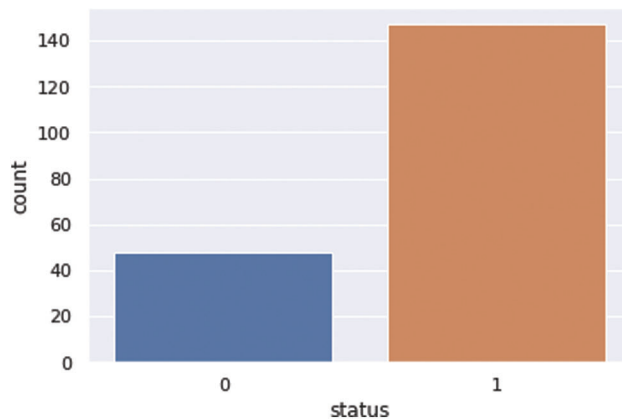


Figure 14: PD patient health status

4.6 Models Building

The machine learning classifiers were implemented. First, we set the target variable, which is a Parkinson's patient health status. After setting up the target, we transform the dataset columns extracted from the dataset to take as input. After setting up the input variables of the dataset, we standardize the dataset columns and transform the dataset with Min and Max Scalar to use in our codes.

4.6.1 SGD-Classifier

The Scikit-learn algorithm provides the SGD-Classifier module for the implementation of the SGD classification. The results are shown in [Tab. 2](#).

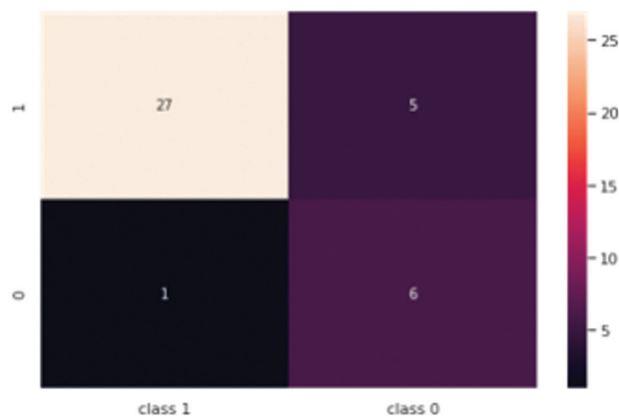
Table 2: SGD classifier results

Name	Results
SGD-classifier	91.666667%
Cross-validation score	66.6667% (+/-24.6581%)
Execution time	0.02922 s

4.6.2 XGB-Classifier Model

XGB-classifier prediction results are shown in [Fig. 15](#).

	precision	recall	f1-score	support
0	0.55	0.86	0.67	7
1	0.96	0.84	0.90	32
accuracy			0.85	39
macro avg	0.75	0.85	0.78	39
weighted avg	0.89	0.85	0.86	39



Score of XGBoost classifier Test Data : 0.8461538461538461
 Score of XGBoost classifier Train Data : 1.0
 Roc AUC score of XGBoost classifier : 0.8504464285714286
 f1 score of XGBoost classifier : 0.8999999999999999

Figure 15: XGB-Classifier results in prediction tables

Fig. 15 and Tab. 3 show the best results of the XGB-classifier producing the precision, recall, F1-score, support, accuracy score. The results show that the testing score is 0.84 and the train score is 1.0.

Table 3: XGB classifier results

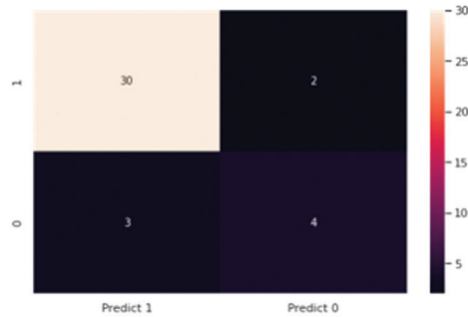
Name	Results
Score of XGB-classifier Test Data	0.8461538461538461
Score of XGB-classifier Train Data	1.0
AUC of XGB-classifier score	0.8504464285714286
F1-score of XGB-classifier	0.8999999999999999
Total accuracy score of XGB-classifier	95.08196721311475

4.6.3 Logistic Regression Classifier

Fig. 16 and Tab. 4 show the results of the logistic regression classifier.

```

Accuracy of Training Data: 0.8260869565217391
Accuracy of Test Data: 0.8717948717948718
D2 Jitter:DDP MDVP:APQ MDVP:Jitter(%) MDVP:Jitter(Abs) MDVP:PPQ
0 0.454476 0.492643 1.22633 -0.929222 0.129734 -0.376542
MDVP:RAP MDVP:Shimmer MDVP:Shimmer(db) RPDE Shimmer:APQ3 \
0 0.494756 0.284885 -0.023437 -0.51382 -0.466563
Shimmer:APQ5 Shimmer:DDA spread1 spread2 intercept
0 0.383768 -0.467944 2.098717 0.722762 1.558423
precision recall f1-score support
0 0.67 0.57 0.62 7
1 0.91 0.94 0.92 32
accuracy 0.87 39
macro avg 0.79 0.75 0.77 39
weighted avg 0.87 0.87 0.87 39
    
```



f1 score 0.923076923076923
AUC Roc Score: 0.7544642857142856

Figure 16: Logistic regression-classifier results

4.6.4 Random Forest Classifier

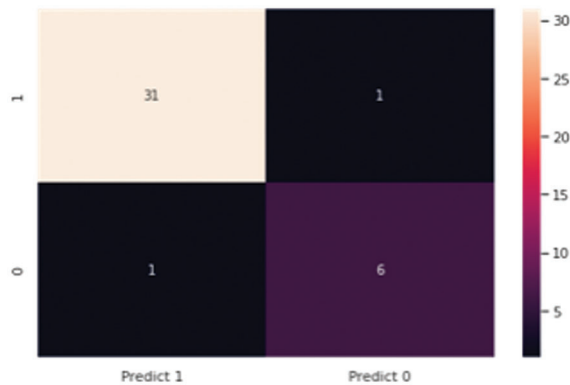
The results of Random forest are shown below (see Fig. 17).

Tab. 5 shows the maximum precision score those get and recall, support in the middle of the predicted classes of the Parkinson disease which produces the training data score is 0.94 and testing score is 1.0, AUC score of 0.91, and F1-score is 0.96. In the end, we generate the total accuracy score of the classifier, as shown in Tab. 5.

Table 4: Logistic regression classifier results

Name of score	Results
Score of the training data	0.8260869565217391
Score of the testing data	0.8717948717948718
AUC score	0.7544642857142856
F1-score	0.923076923076923
Total accuracy score	0.9166666666666666

	precision	recall	f1-score	support
0	0.86	0.86	0.86	7
1	0.97	0.97	0.97	32
accuracy			0.95	39
macro avg	0.91	0.91	0.91	39
weighted avg	0.95	0.95	0.95	39



Score of RandomForest classifier Test Data : 0.9487179487179487
 Score of RandomForest classifier Train Data : 1.0
 Roc AUC score of RandomForest classifier : 0.9129464285714287
 f1 score of RandomForest classifier : 0.96875

Figure 17: Resulting diagram of random forest classifier

Table 5: Random forest classifier results

Name of score	Results
Training data score	1.0
Test data score	0.9487179487179487
AUC score	0.9129464285714287
F1-score	0.96875
Total accuracy score	0.9710144927536232

4.6.5 KNN Classifier

Fig. 18 shows the line chart of the KNN classifier. This classifier reaches the range of 95%.

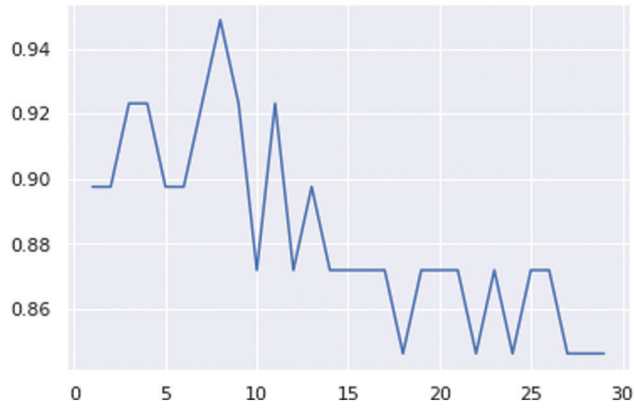
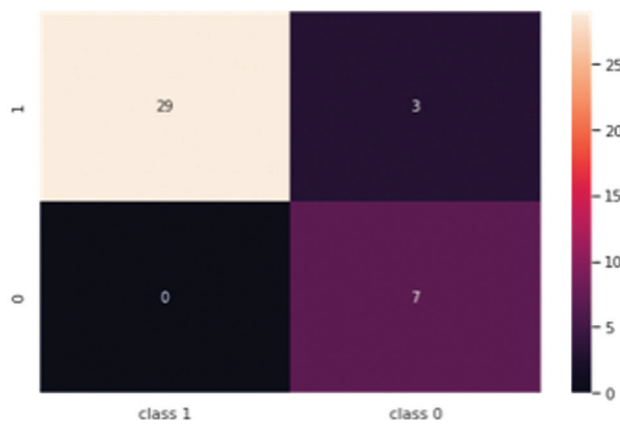


Figure 18: KNN classifier ranges lines chart

Fig. 19 shows the confusion matrix of the KNN classifier, which generates the precision, recall, and shows the best accuracy score of the test data.

	precision	recall	f1-score	support
0	0.70	1.00	0.82	7
1	1.00	0.91	0.95	32
accuracy			0.92	39
macro avg	0.85	0.95	0.89	39
weighted avg	0.95	0.92	0.93	39



Score of Knn Test Data : 0.9230769230769231
 Score of Knn Train Data : 1.0
 Roc AUC score of KNN : 0.953125
 f1 score of KNN : 0.9508196721311475

Figure 19: KNN-classifier of resulting figure and scores

The resulting scores are also arranged in Tab. 6.

Table 6: KNN classifier results

Name of score	Results
Training data score	1.0
Test data score	0.9230769230769231
AUC score	0.953125
F1-score	0.9508196721311475
Total accuracy score	0.9545454545454546

4.6.6 Decision Tree Classifier

In this classifier, we generate the decision tree classifier's training and testing score to reduce the feature of the dataset. The results are shown in [Tab. 7](#).

Table 7: Decision tree classifier results

Name of score	Results
Training data score	0.8173913043478261
Test data score	0.7948717948717948
Total accuracy score	0.9583333333333334

4.7 Comparison of the ML Classifiers

The comparison results between the different algorithms studied in this paper are shown in [Tab. 8](#). Moreover, [Fig. 20](#) clearly shows all classifiers of the result graphically. The results show that random forest outperforms other classifiers with a 97% accuracy score.

Table 8: Comparison results of all algorithms

Classifier	Test data score	Train data score	Total accuracy score
SGD	0.20	0.80	91.666667%
XGB	0.8461538461538461	1.0	95.081967213114
Logistic regression	0.87179487179487	0.82608695	0.916666666666
Random forest	0.9487179487179487	1.0	0.9710144927536232
KNN	0.9230769230769231	1.0	0.9545454545454546
Decision tree	0.7948717948717948	0.8173913043478261	0.9583333333333334

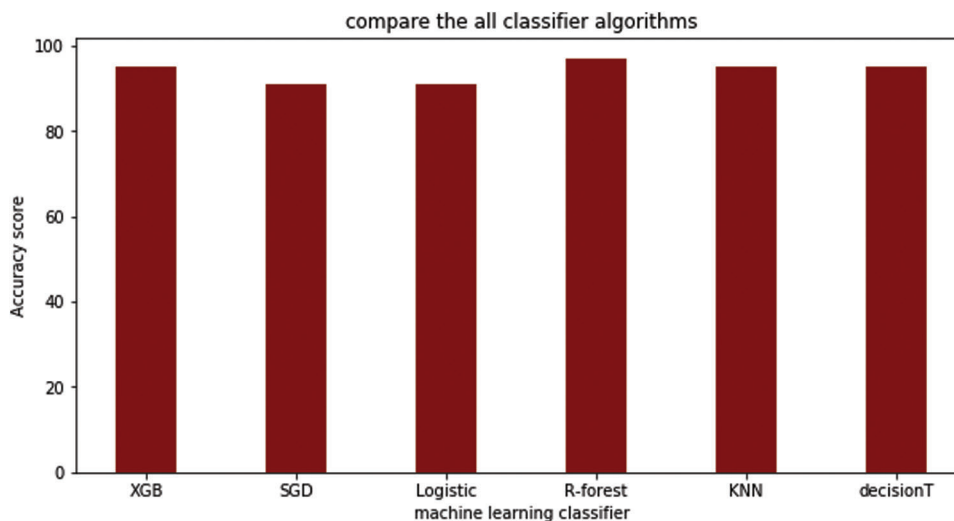


Figure 20: Comparison between all the machine learning classifiers

5 Conclusions and Future Work

Parkinson's disease diagnosis is challenging to manage daily. Thus, an effective screening process will be helpful, especially for cases that do not require a visit to a clinic. Symptoms like vocal characteristics, voice recording, speech, and slow movement are valuable and non-invasive diagnostic tools. This paper used machine learning algorithms to diagnose the disease through the patient's voice patient. This is a practical step to check before meeting with a clinician. A dataset of voices was used as an input to several machine learning models. The results show that the random forest classifier performs with high accuracy. In future work, more datasets of PD patients can be used in order to measure the accuracy of the random forest if the new data was added.

Funding Statement: This work was supported by Taif University Researchers Supporting Project (TURSP) under number (TURSP-2020/73, Taif University, Taif, Saudi Arabia, and by the Data Science Research Program, Riphah International University, Lahore, Pakistan.

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

References

- [1] Y. Mitra and V. Rustagi, "Classification of subjects with parkinson's disease using gait data analysis," in *Proc. ICACE*, Greater Noida, India, pp. 84–89, 2018.
- [2] H. Ozkan, "A comparison of classification methods for tediagnosis of Parkinson's disease," *Entropy*, vol. 18, no. 4, pp. 115, 2016.
- [3] E. Rastegari, S. Azizian and H. Ali, "Machine learning and similarity network approaches to support automatic classification of Parkinson's diseases using accelerometer-based gait analysis," in *Proc. HICSS*, Honolulu, Hawaii, USA, pp. 4231–4242, 2019.
- [4] D. Weintraub, C. L. Comella and S. Horn, "Parkinson's disease—Part 1: Pathophysiology, symptoms, burden, diagnosis, and assessment," *the American Journal of Management Care*, vol. 14, no. 2, Suppl, pp. S40–S48, 2008.
- [5] C. Salvatore, I. Castiglioni, F. Gallivanone, A. Augimeri, M. Lopez *et al.*, "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy," *Journal of Neuroscience Methods*, vol. 222, no. 1, pp. 230–237, 2014.

- [6] N. Tahir and H. Manap, "Parkinson disease gait classification based on machine learning approach," *Journal of Applied Sciences*, vol. 12, no. 2, pp. 180–185, 2012.
- [7] D. Weintraub, C. L. Comella and S. Horn, "Parkinson's disease—Part 1: Pathophysiology, symptoms, burden, diagnosis, and assessment," *The American Journal of Management Care*, vol. 14, no. 2, Suppl, pp. S40–S48, 2008.
- [8] S. Rewar, "A systematic review on Parkinson's disease (PD)," *Indian Journal of Research in Pharmacy and Biotechnology*, vol. 3, no. 2, pp. 176, 2015.
- [9] S. Arora, V. Venkataraman, A. Zhan, S. Donohue, K. M. Biglan *et al.*, "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study," *Parkinsonism and Related Disorders*, vol. 21, no. 6, pp. 650–653, 2015.
- [10] D. Miljkovic, D. Aleksovski, V. Podpečan, N. Lavrač, B. Malle *et al.*, "Machine learning and data mining methods for managing Parkinson's disease," in *Machine Learning for Health Informatics*, Springer, Cham, Switzerland, pp. 209–222, 2016.
- [11] K. N. R. Challa, V. S. Pagolu, G. Panda and B. Majhi, "An improved approach for prediction of Parkinson's disease using machine learning techniques," in *Proc. SCOPES*, Paralakhemundi, India, pp. 1446–1451, 2016.
- [12] G. S. Lee and S. H. Lin, "Changes of rhythm of vocal fundamental frequency in sensorineural hearing loss and in Parkinson's disease," *Chinese Journal of Physiology*, vol. 52, no. 6, pp. 446–450, 2009.
- [13] B. Lagrue, P. Mignard, F. Viallet and R. Gantcheva, "Voice and Parkinson disease: A study of pitch, tonal range and fundamental frequency variations," in *Proc. ICPHS*, San Francisco, USA, pp. 1811–1814, 1999.
- [14] S. Grover, S. Bhartia, A. Yadav and K. R. Seeja, "Predicting severity of Parkinson's disease using deep learning," *Procedia Computer Science*, vol. 132, no. 1, pp. 1788–1794, 2018.
- [15] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. E. Costello and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *BioMedical Engineering OnLine*, vol. 6, no. 23, pp. 1–19, 2007.
- [16] D. Joshi, A. Khajuria and P. Joshi, "An automatic non-invasive method for Parkinson's disease classification," *Computer Methods and Programs in Biomedicine*, vol. 145, no. 1, pp. 135–145, 2017.