Journal of Current Research in Engineering and Science
JCRES
Bi-Annual Online Journal (ISSN: 2581 - 611X)

VOLUME:8, ISSUE:2, AUGUST 2025, PAPER NUMBER:4

PRIMORDIAL PARKINSON'S DISESASES DETECTION USING RANDOM FOREST

Subashini M[1], Badhrinath R[2], Navin D[3], Sai Dhanush K[4]

Dept. of Computer Science Engineering, Sri Sairam Institute of technology, Chennai, India.
Subashini.cse@sairamit.edu.in,Sit20cs040@sairamtap.edu.in,Sit20cs085@sairamtap.edu.in
Sit20cs060@sairamtap.edu.in

Abstract Accurate early diagnosis of Parkinson's disease (PD) is unquestionably crucial to halting the disease's progression and providing patients with access to medication that modifies the condition. To do this, it is important to monitor the premotor stage of Parkinson's disease (PD). Based on premotor traits, a novel machine learning method is presented to rapidly ascertain whether an individual has Parkinson's disease (PD). The highest accuracy, 94%, is averaged here. To achieve this, the premotor stage of Parkinson's disease needs to receive a lot of attention. A machine learning algorithm is presented to quickly determine if an individual has Parkinson's disease (PD) or not based on premotor characteristics.

Keywords: PD-Parkinson's Disease

I. ***INTRODUCTION***

.

1. INTRODUCTION:

Parkinson's disease (PD) is a major degenerative condition of the central nervous system that affects millions of elderly people globally in their quality of life [1]. The disease's variability means that each person's experience with PD symptoms may differ. Parkinson's disease patients may experience symptoms, the most common of which are resting tremors. There are various types of tremors that can happen, such as tremors in the hands, stiffness in the limbs, and problems with walking and balance. Parkinson's disease (PD) symptoms can be divided into two groups: those that are motor-related and those that are not (non-motor). Individuals with non-motor symptoms suffer more than those with motor symptoms as their main complaint.

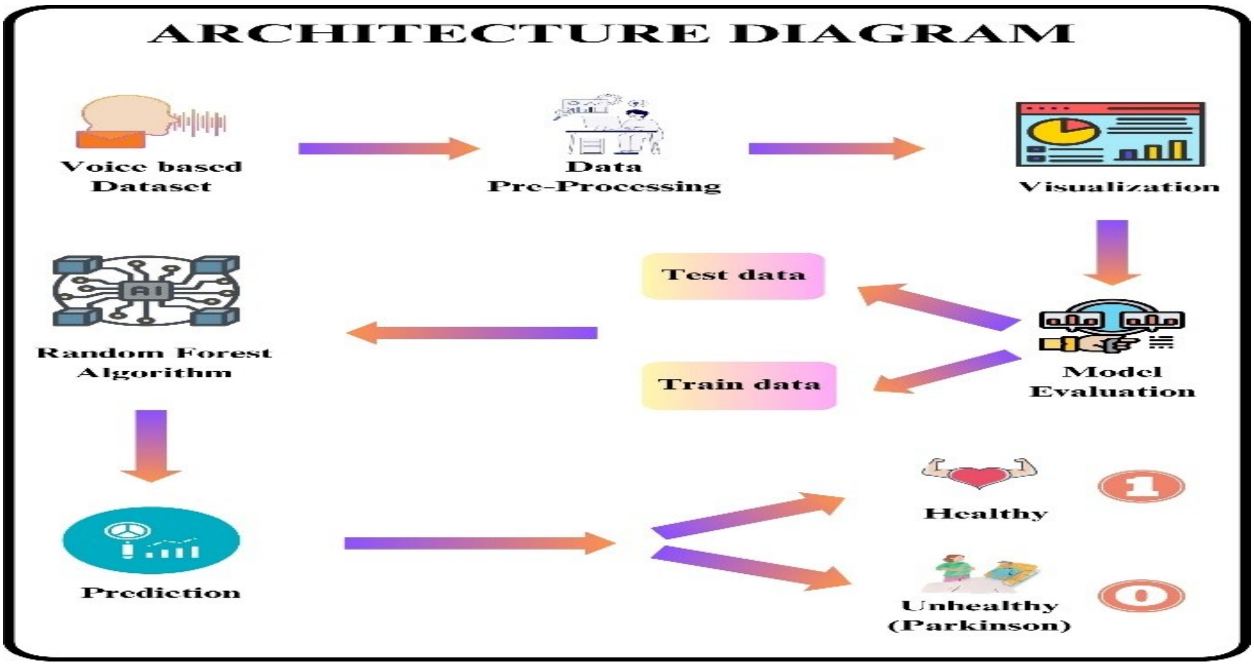
Non-motor symptoms include depression, sleep disorders, behavioural problems, loss of smell, and cognitive impairment. The 14th most common cause of death in the nation, according to the Centres for Disease Control and Prevention (CDC), is complications from Parkinson's disease. As of this writing, the main cause of Parkinson's disease is still unknown.
There are currently more than 10 million PD patients in the globe. Noteworthy, as reported in [2], prompt diagnosis of Parkinson's disease (PD) enables prompt treatment and significantly reduces symptoms. Delaying the progression of Parkinson's disease (PD) and providing patients with access to disease-modifying therapies when they become available can both be achieved by early detection of the disease. One cannot currently diagnose Parkinson's disease (PD) [2].

Nonetheless, certain signs and diagnostic techniques are combined. Researchers have investigated a number of biomarkers to help identify Parkinson's disease early on and slow the disease's progression. Present-day PD treatments all decrease symptoms but do not slow down or halt the progression of the illness. For instance, a variety of methods to assist in the detection of Parkinson's disease (PD) have been built upon speech data [3]– [6], gait patterns [7], force tracking data [8], smell identification data [9], and spontaneous cardiovascular oscillations [10]. A method based on the sawtooth. Speech impairments associated with Parkinson's disease are assessed using the Inspired Pitch Estimator

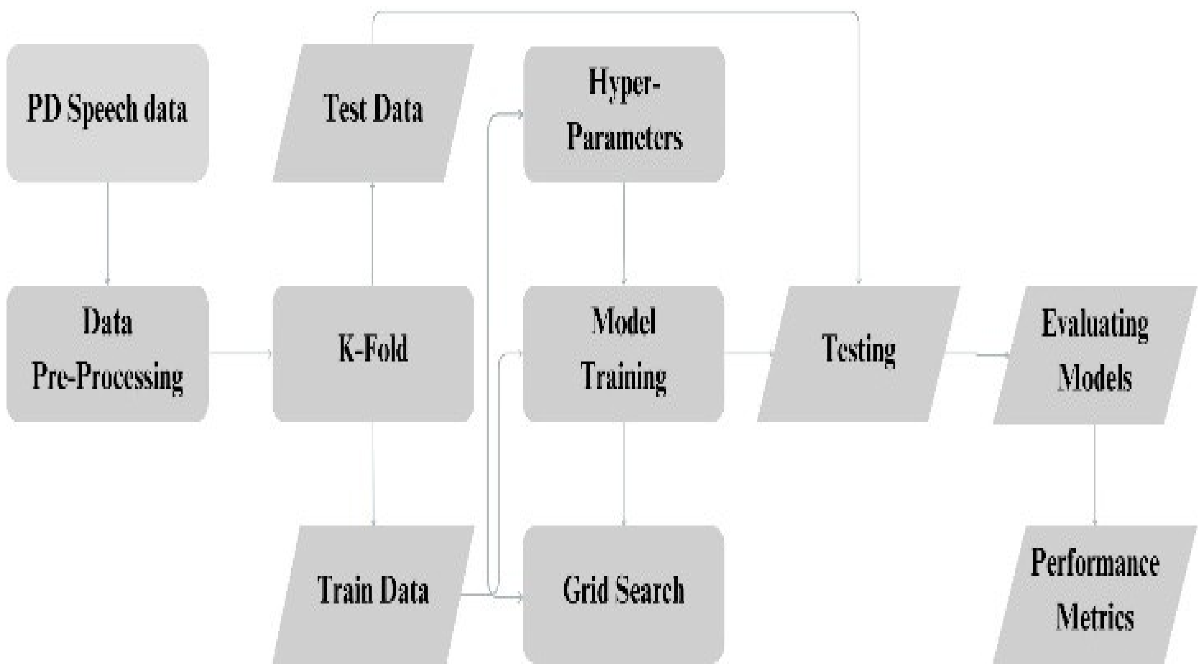
(SWIPE) method. When using the SWIPE method to separate PD and healthy individuals, acceptable results have been obtained. Accurate and timely diagnosis of Parkinson's disease (PD) is crucial because it can provide important information to slow down the disease's progression. To improve the identification of Parkinson's disease (PD), numerous data-driven techniques have been developed over time. Whereas model-based detection techniques also require the prior availability of an analytical model, data-driven strategies only need historical data to be available.

Lately, both academia and industry have expressed interest in studying machine learning (ML) as a possible method for diagnosing Parkinson's disease [20]. This study's consideration of consonants is based on what has been investigated and documented in other publications.

2.ARCHITECTURE:



3.FLOW DIAGRAM:



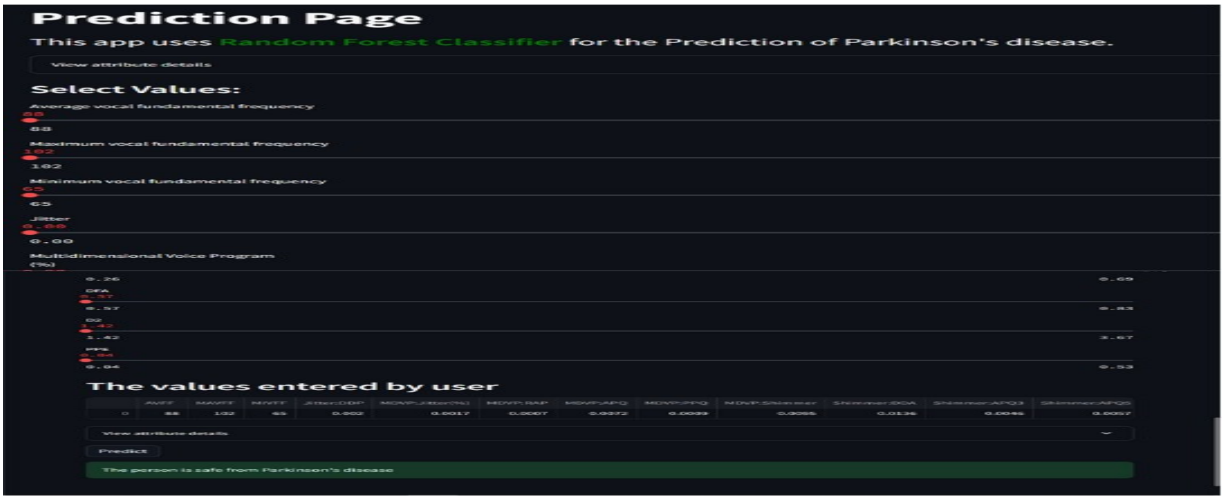
The Parkinson's flow diagram that we have today was created by curators Andrade from the NIATS of the Federal University of Cambridge. Using OpenCV and Python, we will train a model to automatically identify Parkinson's from the voice-based frequency dataset.

4.PICTORIAL REPRESENTATION:

PATIENTS DEMOGRAPHICS AND CLINICAL INFORMATION

Informatics	PWP (MEAN SD)	Control (Mean SD)
Age	71.28	66.91
UPDRS-III (0-132)	25.67	2.64
Duration of PD	4.94	-

Separate classification was done using reduced /a/ and /m/ characteristics to evaluate how well voice features made from sustained /m/ and sustained /a/ could distinguish PWP from controls.



Experimental Analysis

An innovative hybrid model with a class-imbalance data distribution was used to classify the Parkinson's disease dataset. 10-fold cross validation and the holdout technique with a 50/50 train-test split have been used in the decision tree's development and classifier testing. Indicators of performance include area under the curve, precision, recall, accuracy of classification, and Kappa F-measure. The ROC curve has been used to evaluate the suggested method. shows the results obtained from Random Forests. lone classifier in the 50-50 train-test partitioning classification of the PD dataset is a holdout method. Table 2 indicates. The PD dataset was only successfully classified by the decision tree classifier when 10-fold cross-validation was used. It uses a tree to present the results of Parkinson's disease. The combination of employing decision trees and SMOTE classifiers with a holdout partition of 50% train-test ratio for the PD dataset. Results of the combination are shown in the table. of SMOTE and decision tree classifiers in the PD dataset classification process utilizing 10-fold cross-validation.

There are 24 columns and 195 observations

```
In [4]: # We can observe the dataset using the head()function, which returns the first five records from the dataset
parkinsons_data.head()
```

Out[4]:

	name	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP	MDVP:Shimmer	...	Shim
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 × 24 columns

Fig. 1 A sample dataset of primordial Parkinson's disease detection having 194 rows and 24 columns that can handle various voice-based frequency.

Out of the 192 samples that are normally included in the healthy group, 564 samples in the PD dataset have Parkinson disease. Applying the decision tree approach to the PD dataset increased the number of samples from the healthy group from 192 to 767 samples. Within the Parkinson disease category, the number of samples remains unchanged. According to the findings, the hybrid model that was proposed has done a good job of differentiating between Parkinson disease datasets that have problems with class imbalance. There is evidence to show that the pathophysiological alterations linked to Parkinson's disease (PD) may precede the beginning of motor symptoms and may manifest in a variety of nonmotor ways, including sleep difficulties, depression, and cognitive abnormalities. The interest for research focusing on preventative or protective medicines has been fueled by evidence for this preclinical period.

5.SAMPLE OUTPUT:

Three Different Classifier's Accuracy Rate Results:

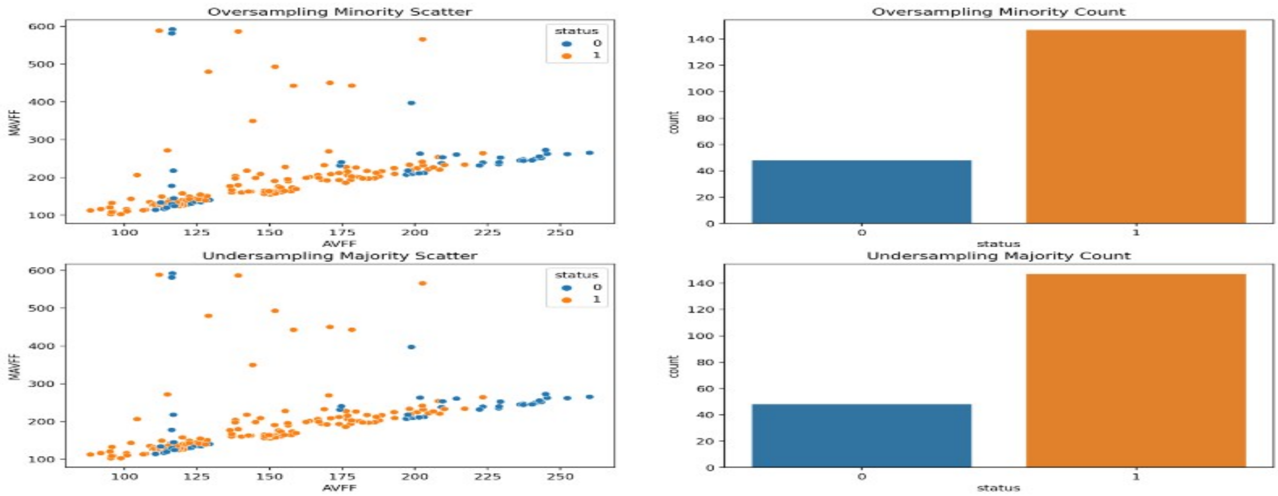
USED CLASSIFIER	CLASSIFICATION ACCURACY RATE
XGBoost Algorithm	82%
Decision tree	86%
Random Forest Algorithm (proposed)	94%

It can be seen from the result proposed model have achieved great accuracy rate compared to previous times algorithms.

The sustained consonant /m/ phonetic features, according to the results, allowed the classification model to distinguish between PWP and healthy controls with a 93% accuracy rate. This categorization accuracy is significantly higher than that of /a/, which was only 70%. The phonation /m/ had an MCC of 0.85, while the phonation /a/ had an MCC of 0.39. The classification model based on /m/ had greater sensitivity and specificity than a model based on /a/...The classification findings indicated that PD may be predicted more accurately using the classification based on phonation properties of /m/

6.CONCLUSIONS:

Early PD diagnosis is essential to better understand the disease's causes, initiate therapeutic approaches, and facilitate the development of effective medications. One example of a premotor characteristic factory loss and a disorder of (RBD) is a deep learning model that was proposed in this work to automatically distinguish people with PD from those who do not. With an accuracy of 96.45%, the suggested deep learning model proved to be an effective detector. The main reason for this is that the deep learning model's beneficial characteristics allow it to learn both linear and nonlinear features from PD data without the need for manually extracted features.



7.ACKNOWLEDGMENT :
This article discusses the development of a risk model that could be used to: (1) identify patients for future studies who may have early-stage Parkinson's disease; and (2) help in the earlier diagnosis of Parkinson's disease (PD) as the cause of symptoms patients present with in primary care.

8.REFERENCES:

[1] "Parkinson's Disease Detection based on Changes of Emotions During Speech," by Justyna Sibasa and Radim Burget, Brno University of Technology 2020 12th International Congress on wide Modern Telecommunications and Control Systems and Workshops, Brno, Czech Republic (ICUMT)

[2] Kim H.-C., S. Chakraborty, and S. Aich (2020). Using artificial neural networks, 3D textural, morphological, and statistical analysis of voxels of interest in 3T MRI scans is performed to identify Parkinson's disease.

[3] A. Abos, H. C. Baggio, B. Segura, A, C. Uribe, D. M. Giraldo, et al (2019). Differentiation between Parkinson's disease and multiple system atrophy using probabilistic tractography-derived structural connectivity.

[4] Arshad M., Ali S., Ali L., Khan S. U., and Anwar M. (2019a). a system with many models for assessing various speech samples with supplementary information regarding Parkinson's illness.