PARKINSON'S DISEASE DETECTION BASED ON HYBRID MACHINE LEARNING ALGORITHMS

D. Karishma [1], K. Aashritha[2], K. Harathi Reddy[3], B. Faseeha Siddiqa[4], and R. Raja[5]

[1]Assistant Professor in Dept. of CSE in RGMCET , Nandyal

[2] Department Of Computer Science Engineering, Student in RGMCET, Nandyal,

[3] Department Of Computer Science Engineering, Student in RGMCET, Nandyal,

[4] Department Of Computer Science Engineering, Student in RGMCET, Nandyal,

[5] Department Of Computer Science Engineering, Student in RGMCET, Nandyal,

ABSTRACT

Parkinson's disease is a neurodegenerative disorder affecting millions worldwide. Early detection can help in better management and treatment of the disease. Machine learning can help in accurate and efficient diagnosis of Parkinson's disease. Our project aims to develop a machine learning model for the detection of Parkinson's disease. The model uses data on voiceand motor function to predict the presence of Parkinson's disease. The model has a high accuracy rate and can be used for early detection of Parkinson's disease. Our project has the potential to contribute to improved patient outcomes and quality of life.

Directly detecting Parkinson's complaint(PD) at an early stage is clearly necessary for decelerating down its progress and furnishing cases the possibility of penetrating to complaint- modifying remedy.

Beforehand discovery can help in better operation and treatment of the disease. Machine literacy can help in accurate and effective opinion of Parkinson's disease. Our design aims to develop a machine literacy model for the discovery of Parkinson's disease. The model uses data on voice and motor function to prognosticate the presence of Parkinson's disease. The model has a high delicacy rate and can be used for early discovery of Parkinson's disease. Our design has the implicit to contribute to bettered patient issues and quality of life. still, Naïve Bayes, Decision Tree Classifier, If a person has Parkinson disease or not is prognosticated by comparing different machine learning algorithms similar as Random Forest. In this proposed work, the delicacy score of the algorithms is compared and the stylish algorithm is chosen that gives the loftiest delicacy rate in prognosticating the Parkinson complaint.

If a person has Parkinson disease or not is predicted by comparing different machine learning algorithms such as Random Forest, Naïve Bayes, Decision Tree Classifier, K-NearestNeighbours (KNN) and Extreme Gradient Boosting (XG Boost). In this proposed work, the accuracy score of the algorithms is compared and the best algorithm is chosen that gives the highest accuracy rate in predicting the Parkinson disease.

INDEX TERMS : Parkinson's disease, Machine learning, Classification algorithm, Random Forest, XG Boost algorithm, K-Nearest Neighbours.

I. INTRODUCTION

Parkinson's disease is a neurological disorder that leads to tremors, stiffness, and difficulty with movement and coordination. Symptoms typically occur gradually and worsen over time. As the disease progresses, individuals with Parkinson's may experience difficulties with walking and speaking. They may also suffer from cognitive and emotional changes, sleep problems, depression, memory problems, and fatigue. Parkinson's disease is caused by the death or dysfunction of certain neurons in the brain that produce dopamine, a chemical messenger. The lack of dopamine causes abnormal brain activity, leading to the characteristic motor and non-motor symptoms of Parkinson's disease. Therefore, it is crucial to be able to diagnose and manage this disease effectively.

Among the signs and side effects of Parkinson's defilement are:

- Shiver: A quake, or shaking, if all else fails, begins in one of your furthest points, most all things considered in your grasp or fingers. A pill-moving shiver happens when you rub your thumb and index finger to a great extent.

• Moved back progress: Parkinson's defilement moves back your movability after some time, making routine undertakings more extraordinary and repetitive. Precisely when you walk, your means could end up being more confined. Moving away from a seat can challenge. Precisely when you attempt to walk, your feet could drag.

• Unfaltering Muscles: Muscle heartiness can strike whenever and in any piece of your body. Solid muscles can be unusual and confine your degree of improvement.

• Position and concordance issues: because of Parkinson's defilement, your position could become stooped, and you could encounter balance issues.

• Loss of body Balance: It's conceivable that you'll acquire a couple of more enthusiastically encounters doing careless activities like flickering, grinning, or swinging your arms as you walk.

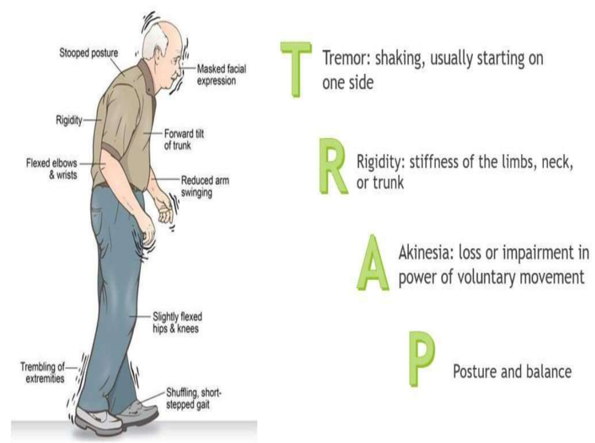
• Shuddering of Voice: You have the choice of talking cautiously, expedient, slurring, or stopping going before talking. Rather than the standard verbalizations, your discussion might be more repetitive.

• Bother recorded as a printed rendition: It could become testing to make, and your shaping could show up minimal as such.

• Resting Disorders: Sleep issues are run of the mill in individuals with Parkinson's difficulty, and can solidify stimulating endlessly all through the range of the evening, getting rolling early, or nodding off during the day.

Expedient eye improvement rest direct knot, which consolidates proceeding with dreams, is in this way an open door. Arrangements could have the decision to assist you with your rest issues.

Individuals with different degrees of personality disorder may experience a change in their development. Because of the variety of the disease. Patients with Parkinson. Severe symptoms such as tremors may be observed during resting. Dif- Certain tremors can be severe such as those felt in the hands or limbs. Discipline and apprehension. Generally, two types. The symptoms of PD can be differentiated by their movement-related characteristics. Movement is separate from the motor, and not related to it. In fact, The impact is greater for patients with non-motor symptoms than those who are motorized. Whose main symptoms are motor. Non-motor symptoms may. Include disorders of depression, sleep disorder, and loss of consciousness. Of smell, and cognitive impairment. It has been reported. As per the CDC's guidelines. The 14th most prevalent cause of PD complications is unclear. Death in the United States. To this day, PD is the reason behind its development in some individuals. Rests principally unknown. Particularly, the economic burden. This manuscript is under the review of an associate editor. It was given the green light by Fatih Emre Boran for publication. Due to the direct and indirect costs of PD treatment, Estimates suggest that social security payments and lost income amount to approximately 900,000 pounds. The United States' yearly revenue is roughly $52 billion. Alone. To be honest, the quantity of individuals suffering from PD has increased. Exceeded 10 million worldwide. Remember that the. The early identification of PD leads to prompt treatment and quick recovery. Alleviate symptoms significantly as reported in [2]. Therefore, The early detection of PD is undoubtedly a crucial factor. To slowing down its progression and could give patients the. Access to therapy that can modify the disease when it becomes more severe. Available. There is currently no method of diagnosing Parkinson's disease. Ease (PD) [2].



The most obvious signs and consequences of Parkinson's disease occur when nerve cells in the basal ganglia, a region of the brain that controls movement, become weakened and die. Normally, these nerve cells, or neurons, produce a crucial brain substance known as dopamine. When the neurons fail or become depressed, they produce less dopamine, which causes the movement problems associated with the disease. Scientists still know little about what causes the neurons to fail.

People with Parkinson's disease also lose the nerve endings that produce norepinephrine, the highly specialized messenger of the autonomic nervous system, which controls various parts of the body, such as heart rate and blood pressure. The lack of norepinephrine could help explain some of the non-movement features of Parkinson's, such as depression, irregular heartbeat, reduced movement of food through the gastrointestinal tract, and sudden drop in blood pressure when a person stands up from a sitting or lying position.

Many brain associations of people with Parkinson's disease contain Lewy bodies, abnormal masses of the protein alpha-synuclein. Scientists are trying to more effectively understand the normal and abnormal roles of alpha-synuclein and how inherited changes affect Parkinson's and Lewy body dementia.

Some cases of Parkinson's disease appear to be hereditary, and a few cases can be traced to specific genetic changes. While genetics is known to play a role in Parkinson's, often the disease does not seem to run in families. Many researchers now believe that Parkinson's results from a combination of genetic and environmental factors, such as exposure to toxins.

Parkinson's disease detection using machine learning algorithms consists of data collection from various sources, feature extraction to identify informative characteristics, pre-processing to ensure data quality, training data split for model training and evaluation, machine learning algorithms including SVM, Random Forests, Neural Networks, or Gradient Boosting, model evaluation using metrics like accuracy and cross-validation, and model selection based on performance. This comprehensive framework enables accurate and reliable prediction of Parkinson's disease, aiding in early detection and effective management of the condition.

Parkinson's disease detection using machine learning algorithms involves patient registration, data collection from various sources, and feature extraction to identify informative characteristics. The system captures relevant demographic information and medical history during patient registration. Data collection includes motor symptom measurements, speech recordings, and other relevant features from sources like clinical databases, wearable devices, or mobile applications. Feature extraction techniques are then applied to extract the most informative characteristics from the collected data.

II. DATA & METHODOLOGY

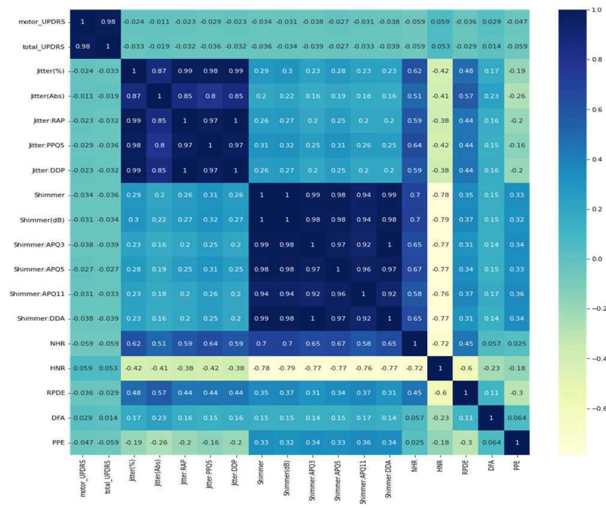
A. Data analysis

Exploratory Data Analysis

Produce a visual representation of the dataset, which should include information about the number of columns and the metropolises used in the exploratory data analysis. Before removing any columns, it's important to examine and estimate the data. use the test and train datasets to induce AQI values for different metropolises. Identify and remove any columns with a high correlation from the dataset.

Correlation Analysis

The dependent variable is removed from the dataset using correlation analysis, which also identifies the independent and dependent variables. When two variables move in the same direction, it indicates a positiverelationship.However, the other will also increase, If one variable increases. For illustration, running on a routine for a longer duration will burn further calories. When two variables have a negative relationship, it means they move in contrary directions. For case, if an increase in one measure leads to a drop in the other, accelerating a auto will reduce the time it takes to reach yourdestination.However, they're inapplicable to each other, If there's no correlation between two variables.



Data Pre-processing

Data pre-processing, a fashion used in data mining, enables the running and application of deficient data. Data drawing involves tasks similar as filtering out noisy data, removing outliers, attributing missing values, and resolving data conflicts. Missing data can make prognostications delicate due to the performing query, and it's important to handle them duly to avoid prejudiced prognostications. One approach is to use a boxplot to identify outliers and remove them from the dataset. They include Ignore tuples This methodology is applied when a sizable piece of the data in a tuple is missing. Use the data's mean worth also again, if essential, redundant principally indistinguishable characteristics, to fill in any openings.

SL no.	Feature group	Feature name
1	ID	Subjects's identifier
2	Recording	Number of the recording
3	Gender	0 = Man; 1 = Woman
4	Pitch local perturbation measure	Relative jitter (Jitter_rel), absolute jitter (Jitter_abs), relative average perturbation (Jitter_RAP) pitch perturbation quotient (Jitter_PPQ)
5	Amplitude perturbation measures	Local shimmer (Shim_loc), shimmer in dB (Shim_dB), 3-point amplitude perturbation quotient (Shim_APQ3), 5-point amplitude perturbation quotient (Shim_APQ5)
6	Harmonic-to-noise ratio measures	HNR05, HNR15, HNR25, HNR35, HNR38
7	Mel frequency cepstral coefficient-based spectral measures	MFCC0, MFCC1,..., MFCC12) and their derivatives (Delta0, Delta1,..., Delta12)
8	Recurrence period density entropy	RPDE
9	Detrended fluctuation analysis	DFA
10	Pitch period entropy	PPE
11	Glottal-to-noise excitation ratio	GNE
12	Status	0 = Healthy; 1 = PD

B. Building Model

Once the dataset is stoked, a model is erected exercising the arbitrary timber and XG Boost algorithms to compare their performance in terms of delicacy and other hyperparameters. The dataset is resolve into training and testing data, and the separate evaluations are applied to them.

Extreme Gradient Boost Algorithm

It's similarly a decision tree- grounded gathering AI calculation which uses the inclination supporting arrangement. The figure issues which consolidate the unshaped data will in this way play out the colossal number of different plans. Right when the data is in the position of little to medium also the decision tree calculation choice is stupendous. This usages inclination plunge evaluation inciting the dwindling in messes up and in moderate models. This calculation isn't unequivocally unclear from others in its momentous ways. Can be watched out for in wide collecting of purposes like break conviction, gathering, organizing. It has unknown authentically straightforward nature in Windows, Linux, and OS X. It stays apprehensive of in all the essential programming vernaculars.

Naive Bayes Classifier

The Naive Bayes algorithm is a probabilistic machine learning algorithm that's generally used for bracket tasks. It's grounded on Bayes' theorem, which provides a way to calculate the probability of a particular event grounded on previous knowledge of conditions that might be related to the event. In the environment of the Naive Bayes algorithm, this means that it can calculate the probability of a particular class given some features or attributes of the data. The algorithm assumes that the features are conditionally independent of each other, which is why it's called" naive". This supposition simplifies the computation of chances, making the algorithm computationally effective and easy to apply. To train the Naive Bayes algorithm, it uses a set of labelled data, where each data point is associated with a particular class. The algorithm estimates the tentative probability of each point given each class grounded on the training data. It also uses Bayes' theorem to calculate the probability of each class given the features of a new data point. The class with the loftiest probability is also assigned to the new data point. Naive Bayes is frequently used in textbook bracket tasks, similar as spam discovery or sentiment analysis, where the features are generally words or expressions. It has also been used in other fields, similar as healthcare and finance, for vaticination tasks. Despite its simplicity and hypotheticals, Naive Bayes has been shown to perform well in numerous real- world scripts.

Decision Tree Classifier

Decision tree classifier is a popular machine learning algorithm that works by recursively unyoking the dataset into subsets grounded on the most significant features until a stopping criterion is reached. Each split is made grounded on the point that results in the maximum information gain or Gini contamination reduction. This creates a tree- suchlike structure, where the internal bumps represent the decision rules grounded on features, and the splint bumps represent the class markers or the outgrowth of the bracket. Decision trees are easy to understand, interpret and fantasize, and can handle both categorical and nonstop data. still, they tend to overfit the training data, leading to poor conception performance on new data. To overcome the overfitting problem, several ways have been proposed, including pruning, setting the minimal number of samples needed at a splint knot, and limiting the depth of the tree. also, ensemble styles similar as Random Forest and Gradient Boosting can be used to combine multiple decision trees to achieve better performance. Decision trees have been applied to a wide range of operations, including medical opinion, credit threat assessment, and object recognition. Overall, decision tree classifier is a important and flexible algorithm that can be used for both double and multiclass bracket tasks, with the capability to handle both numerical and categorical data.

Random Forest Classifier

The arbitrary timber is a supervised literacy algorithm that combines multiple decision trees to ameliorate delicacy and reduce overfitting. It uses averaging and meta- assessment ways to fit a number of decision trees for both bracket and retrogression tasks. The model adds randomness to the decision- making process by combining the prognostications of numerous trees, with each tree furnishing a different split point, performing in increased delicacy and stability of prognostications. Random timber has analogous hyperparameters to a decision tree, including a classifier. still, it introduces fresh parameters similar as the number of trees and the number of features considered for each split, allowing for better running of missing values and maintaining high delicacy indeed for large datasets. The strike is that it can be computationally precious and may not perform well with noisy or inapplicable features.

Nearest Neighbours

K Nearest Neighbour(KNN) is an essential AI algorithm that utilizes colorful data values to prognosticate affair values. It's one of the simplest bracket algorithms and workshop by organizing the data of interest grounded on the aggregation of its neighbours. KNN collects new data grounded on its similarity position with the preliminarily stored data of interest. For illustration, if we've a dataset of tomatoes and bananas, KNN will store analogous attributes similar as shape and colour. When a new item is introduced, it'll compare its similarity with the stored data grounded on colour and shape. The K value in KNN determines how numerous nearest neighbours are used to organize new data of interest.

C. Evaluation Metrics

Various metrics are utilized to evaluate the quality of a machine learning or artificial intelligence model, including accuracy, confusion matrix, and f1-score. These metrics are compiled into a classification report.

Confusion Matrix:

The evaluation of a classification model's performance on a test dataset with known actual values is a common practice, which can be misleading if there are unequal numbers of observations in each class or multiple classes in the dataset. It provides a summary of the model's performance on the problem and there are specific terms that need to be defined such asrepresents the Confusion Matrix which consists of Predicted values as X axis and Actual Values as Y axis

True Positive (TP)	Diagnosed PD vs Actual PD
False Positive (FP)	Diagnosed PD vs Healthy
True Negative (TN)	Diagnosed Healthy vs Actual Healthy
False Negative (FN)	Diagnosed Healthy vs Actual PD

Recall:

The recall score measures the accuracy of the model in correctly identifying True Positives. Therefore, it indicates the number of patients with coronary disease who were accurately diagnosed with the condition. Mathematically, it can be expressed as:

$$Recall = \frac{TP}{TP + FN}$$

F1-Score:

The F1 score is a measure of the model's performance that combines precision and recall using a specific formula. It indicates the balance between the precision and the recall of the model and is calculated based on the confusion matrix. Generally, it is the harmonic mean of the precision and recall of the model. Higher F1 score means better performance of the model.

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

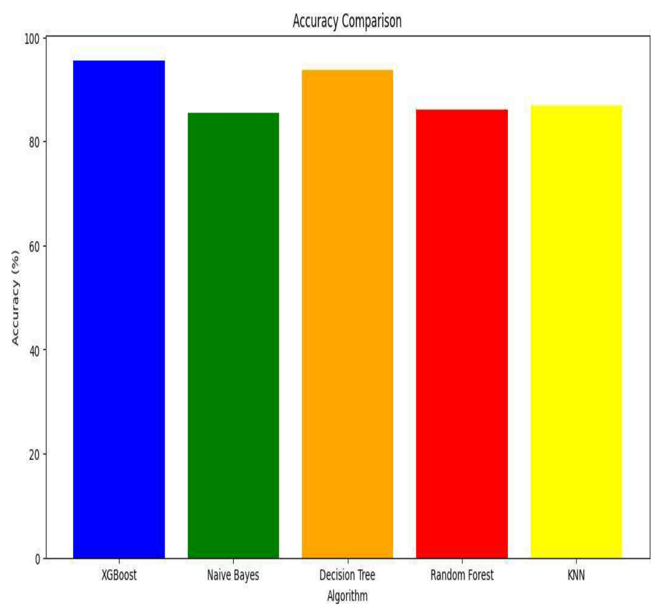
$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

III. Experimental Results:

We use the following to evaluate the effectiveness of machine methods for discriminating between Parkinson patients. Criterion: Accuracy =TP+TN/TP+FP+TN+FN , Sensitivity =TP/TP+FN , Specificity =TN/TN+FP , Precision =TP/TP+FP. F1 = 2. Precision. Sensitivity/Precision+Sensitivity =2TP/2TP+FP+FN . In terms of true positives, false negative and TN, respectively, with TP = true + FP; and false negativity = false. What does each represent? In addition to the five metrics mentioned earlier, we can also use the area under the receiver operating characteristic curve (AUC) The accuracy of predictions is determined by their proportion. Greater precision implies superior overall accuracy. This system was proposed to be a more accurate classification of features so that the feature dimensions were reduced, and therefore voice-loss was detected in people with PD. This section will examine the experimental outcome of the newly developed classification method. Initially, Parkinson disease was predicted using four classifiers: Gradient Boosting (a higher-order neuroprotective index), Extreme Gradant boosting with an extra tree classification system, and Bagging. Following the three feature selection method, we select some additional features. Finally, Following this, we apply the classifier from above that is significant to the other features.



represents Accuracy Plot for XGBoost, Naïve Bayes, Decision tree,Random forest and KNN algorithms.

Accuracy

Accuracy

Algorithms	accuracy
XG BOOST	96
Random Forest	86
KNN	87
Naive Bayes	86
Decision Tree	94

Prediction:

```
173,5.6431,28.199,34.398,0.00074,4.52E-06,0.00015,0.00038,0.00045,0.0
4229,0.368,0.02345,0.02756,0.0297,0.07035,0.003743,25.448,0.26411,0.7
506,0.82665
parkinson disease detected
```

```
19,19.681,28.695,35.389,0.00823,6.70E-05,0.00392,0.0042,0.01176,0.025
92,0.245,0.01269,0.0155,0.02127,0.03807,0.031999,20.454,0.65404,0.71
5,0.33618
no parkinson disease detected
```

IV. Discussions:

In this review, we present results from published studies that applied machine learning to the diagnosis and differential diagnosis of PD. Since the number of included papers was relatively large, we focused on a high-level summary rather than a detailed description of methodology and direct comparison of outcomes of individual studies. We also provide an overview of sample size, data source and data type, for a more in-depth understanding of methodological differences across studies and their outcomes. Furthermore, we assessed (a) how large the participant pool/dataset was, (b) to what extent new data (i.e., unpublished, raw data acquired from locally recruited human participants) were collected and used, (c) the feasibility of machine learning and the possibility of introducing new biomarkers in the diagnosis of PD. Overall, methodology studies that proposed and tested novel technical approaches (e.g., machine learning and deep learning models, data acquisition devices, and feature extraction algorithms) have repetitively shown that features extracted from data modalities including voice recordings and handwritten patterns could lead to high patient-level diagnostic performance, while facilitating accessible and non-invasive data acquisition. Nevertheless, only a small number of studies further validated these technical approaches in clinical settings using local human participants recruited specifically for these studies, indicating a gap between model development and their clinical applications.

A per-study diagnostic accuracy above chance levels was achieved in all studies that used accuracy in model evaluation (Figure 4A). Apart from studies using CSF data that measured model performance with AUC, classification accuracy associated with 8 other data types ranged between 85.6% (PET) and 94.4% (SPECT), with an average of 89.9 (3.0) %. Therefore, although the small number of studies of some data types may not allow for a generalizable prediction of how well these data types can help us differentiate PD from HC or atypical Parkinsonian disorders, the application of machine learning to a variety of data types led to high accuracy in the diagnosis of PD. In addition, an accuracy significantly above chance levels was achieved in all machine learning models (Supplementary Table 1), while SVM, neural networks and ensemble learning were among the most popular model choices, all yielding great applicability to a variety of data modalities. In the meantime, when compared with other models, they led to the per-study highest classification accuracy in >50% of all cases (50.7, 51.9, and 52.3%, respectively; Supplementary Table 1). Despite the high diagnostic accuracy and performance reported, in a number of studies, data splitting strategies and the use of cross validation were not specified. For data modalities such as 3D MRI scans, when 2D slices are extracted from 3D volumes, multiple slices could be generated for one subject. Having data from the same subject across training, validation and tests sets can lead to a biased data split (Wen et al., 2020), causing data leakage and overestimation of model performance, thus compromising reproducibility of published results.

As previously discussed (Belić et al., 2019), although satisfactory diagnostic outcomes could be achieved, sample size in few studies was extremely small (<15 subjects). The application of some machine learning models, especially neural networks, typically rely on a large dataset. Nevertheless, collecting data from a large pool of participants remains challenging in clinical studies, and data generated are commonly of high dimensionality and small sample size (Vabalas et al., 2019). To address this challenge, one solution is to combine data from a local cohort with public repositories including PPMI, UCI machine learning repository, PhysioNet and many others, depending on the type of data that have been collected from the local cohort. Furthermore, when a great difference in group size is observed (i.e., class imbalance problem), labeling all samples after the majority class may lead to an undesired high accuracy. In this case, evaluating machine learning models with other metrics including precision, recall and F-1 score is recommended (Jeni et al., 2013).

Even though high diagnostic accuracy of PD has been achieved in clinical settings, machine learning approaches have also reached high accuracy as shown in the present study, while models including SVM and neural networks are particularly useful in (a) diagnosis of PD using data modalities that have been overlooked in clinical decision making (e.g., voice), and (b) identification of features of high relevance from these data. For example, the use of machine learning models with feature selection techniques allows for assessing the relative importance of features of a large feature space in order to select the most differentiating ones, which is conventionally challenging using manual approaches. For the discovery of novel markers allowing for non-invasive diagnostic options with relatively high accuracy, e.g., handwritten patterns, a small number of studies have been conducted, mostly using data from published databases. Given that these databases generally included handwritten patterns from a small number of diagnosed PD patients, sometimes under 15, it would be of great importance to validate the use of handwritten patterns in early diagnosis of PD in clinical studies of a larger scale. In the meantime, diagnosing PD using more than one data modality has led to promising results. Accordingly, supplying clinicians with non-motor data and machine learning approaches may support clinical decision making in patients with ambiguous symptom presentations, and/or improve diagnosis at an earlier stage.

V. Conclusion:

Parkinson's disease is a condition that affects the central nervous system (CNS) and is caused by the brain. Unluckily, Parkinson's disease cannot be cured by treating it without early identification. If diagnosed too late, they will die without treatment. Therefore, early detection is crucial. Parkinson's disease early detection is made possible by various AI algorithms. By analysing data on Parkinson's disease, it is determined that XG Boost is the most reliable algorithm for forecasting the disease'onset, with a 96% accuracy rate. Early treatment and the potential to save lives are all factors considered in this. The estimated values when utilizing the XG Boost algorithm to anticipate Parkinson's disease.

VI. Future Work:

This project has the potential to expand its capabilities by incorporating image recognition technology to recognize the handwriting of patients, thereby enabling it to detect Parkinson's patients more comprehensively. By making necessary adjustments to the dataset, this project could also be adapted to identify diseases such as heart disease or kidney disease, as these algorithms have shown promising results in healthcare.

VII. Acknowledgement:

We would like to extend our sincere gratefulness to our guide , Assistant Professor D. Karishma $_{\text{M.Tech}}$, whose unvarying guidance, support, and moxie have been necessary in the successful publication of this specialized Technical paper. Her fidelity to nurturing our exploration chops and her perceptive feedback have been inestimable throughout this trip. I'm also deeply thankful to the entire Computer Science and Engineering Department at RGMCET for furnishing a conducive terrain for academic growth and for the inestimable benefactions of its speakers. Their stimulant and mentorship have played a significant part in shaping our scholarly hobbies.

References

[1] A Novel Framework of Two Successive Feature Selection Levels Using Weight-Based Procedure for Voice-Loss Detection in Parkinson's Disease. IEEE Access PP (99):1-1

DOI:10.1109/ACCESS.2020.2989032 , Kemal Polat, Amira S. Ashour, Majid K Nour, July 2021.

[2] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," Prog. Neurobiol., vol. 81, pp. 29–44, Jan. 2007.

[3] H. Gunduz, "Deep learning-based Parkinson's disease classification using vocal feature sets," IEEE Access, vol. 7, pp. 115540–115551, 2019.

[4] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 884–893, Apr. 2010.

[5] S. Lahmiri and A. Shmuel, "Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine," Biomed. Signal Process. Control, vol. 49, pp. 427–433, Mar. 2019.

[6] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, "Automatic detection of Parkinson's disease based on acoustic analysis of speech," Eng. Appl. Artif. Intell., vol. 77, pp. 148–158, Jan. 2019.

[7] A. Zhao, L. Qi, J. Li, J. Dong, and H. Yu, "A hybrid spatio-temporal model for detection and severity rating of Parkinson's disease from gait data," Neurocomputing, vol. 315, pp. 1– 8, Nov. 2018.

[8] S. Bilgin, "The impact of feature extraction for the classification of amyotrophic lateral sclerosis among neurodegenerative diseases and healthy subjects," Biomed. Signal Process. Control, vol. 31, pp. 288–294, Jan. 2017.

[9] L. Silveira-Moriyama, A. Petrie, D. R. Williams, A. Evans, R. Katzenschlager, E. R. Barbosa, and A. J. Lees, "The use of a color-coded probability scale to interpret smell tests in suspected parkinsonism," Movement Disorders, vol. 24, no. 8, pp. 1144–1153, Jun. 2009.

[10] G. Valenza, S. Orsolini, S. Diciotti, L. Citi, E. P. Scilingo, M. Guerrisi, S. Danti, C. Lucetti, C. Tessa, R. Barbieri, and N. Toschi, "Assessment of spontaneous cardiovascular oscillations in Parkinson's disease," Biomed. Signal Process. Control, vol. 26, pp. 80–89, Apr. 2016.

[11] V. Illner, P. Sovka, and J. Rusz, "Validation of freely-available pitch detection algorithms across various noise levels in assessing speech captured by smartphone in Parkinson's disease," Biomed. Signal Process. Control, vol. 58, Apr. 2020, Art. no. 101831.

[12] G. Solana-Lavalle, J.-C. Galán-Hernández, and R. Rosas-Romero, "Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features," Biocybern. Biomed. Eng., vol. 40, no. 1, pp. 505–516, Jan. 2020.

[13] I. El Maachi, G.-A. Bilodeau, and W. Bouachir, "Deep 1D-convnet for accurate Parkinson disease detection and severity prediction from gait," Expert Syst. Appl., vol. 143, Apr. 2020, Art. no. 113075.

[14] U. Gupta, H. Bansal, and D. Joshi, "An improved sex-specific and agedependent classification model for Parkinson's diagnosis using handwriting measurement," Comput. Methods Programs Biomed., vol. 189, Jun. 2020, Art. no. 105305.

[15] A. Wagner, N. Fixler, and Y. S. Resheff, "A wavelet-based approach to monitoring Parkinson's disease symptoms," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2017, pp. 5980–5984.

[16] T. Arroyo-Gallego, M. J. Ledesma-Carbayo, A. Sanchez-Ferro, I. Butterworth, C. S. Mendoza, M. Matarazzo, P. Montero, R. Lopez-Blanco, V. Puertas-Martin, R. Trincado, andL. Giancardo, "Detection of motor impairment in Parkinson's disease via mobile touchscreen typing," IEEE Trans. Biomed. Eng., vol. 64, no. 9, pp. 1994–2002, Sep. 2017.

[17] I. D. Dinov, B. Heavner, M. Tang, G. Glusman, K. Chard, M. Darcy, R. Madduri, J. Pa, C. Spino, C. Kesselman, I. Foster, E. W. Deutsch, N. D. Price, J. D. Van Horn, J. Ames, K. Clark, L. Hood, B. M. Hampstead,W. Dauer, and A. W. Toga, "Predictive big data analytics: A study of Parkinson's disease using large, complex, heterogeneous, incongruent, multi-source and incomplete observations," PLoS ONE, vol. 11, no. 8, Aug. 2016, Art. no. e0157077.

[18] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Decision support framework for Parkinson's disease based on novel handwriting markers," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 23, no. 3, pp. 508–516, May 2015.

[19] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease," Artif. Intell. Med., vol. 67, pp. 39–46, Feb. 2016.

[20] F. Harrou, Y. Sun, A. S. Hering, and M. Madakyaru, Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches: Theory and Practical Applications. Amsterdam, The Netherlands: Elsevier, 2020.

[21] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Trans. Biomed. Eng., vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[22] Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning," W. Wang, J. Lee, F. Harrou and Y. Sun IEEE Access, vol. 8, pp. 147635-147646, 2020, Doi: 10.1109/ACCESS.2020.3016062.

[23] Gradient Boosting for Parkinson's Disease Diagnosis from Voice Recordings.Ibrahim Karabayir,Samuel M,Suguna Pappu ,oguz akbilgic Research Gate, January 2020 BMC Med Informatics Decision Making. 2020 Sep 15;20(1):228. date: 10.1186/s12911-020-01250-7. PMID: 32933493; PMCID: PMC7493334.

[24] Feature Selection Based Machine Learning to Improve Prediction of Parkinson Disease.Nazmun Nahar,Ferdous Ara,Md. Arif Istiek Neloy,Anik Biswas,Mohammad Shahadat Hossain &Karl Andersson Research Gate 15 Sep 2021 Journal: Brain Informatics.

[25] Machine Learning Algorithm Using Logistic Regression a for Early Stage Detection of Parkinson's Disease.Shreyas Kar1 , Peter W. Campbell DuPont Manual High School, Louisville, Kentucky, 40208, USA. W University of Louisville, Department of Anatomical Sciences and Neurobiology, Louisville, Kentucky, 40202, Journal of Emerging Investigators 10 OCTOBER 2020.

[26] Diagnosis of Parkinson's Disease Using SVM Classifier,G. Wiselin Jiji,A. Rajesh and P. Johnson Durai Raj.