Annual Methodological Archive Research Review
Online ISSN 3007-3197
Print ISSN 3007-3189
com/index.php/Journal/about
Volun
http://amresearchreview.com/index.php/Journal/about

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 6 (2025)

Enhancing Parkinson's Disease Classification: Evaluating SVM, Decision Tree and Ensemble Learning with Advanced Preprocessing Strategies

[1]Kamran Ali, [2]Zainab Fatima, [3]Arooba Sultan Tiwana, [4]Zarka Saeed , [5]Anam Safdar Awan

Article Details

Kamran Ali
Kamranali3037414518@gmail.com
Zainab Fatima
Superior university Lahore, Sargodha.
zainabranjha96@gmail.com
Arooba Sultan Tiwana
University of Sargodha.
aroobasultantiwana@gmail.com
Zarka Saeed
Minhaj University Lahore.
zarkazarka244@gmail.com
Anam Safdar Awan
Superior University Lahore, Sargodha.
anamawan416@gmail.com

ABSTRACT

Parkinson's disease (PD), the second most common neurodegenerative disorder, affects over millions of people worldwide and encompasses a wide variety of motor and non-motor symptoms which immensely impact one's quality of life. This research aims to improve further diagnostic precision for PD using advanced machine learning (ML) algorithms which are essential in identifying and differentiating the condition from other similar neurodegenerative diseases during its preclinical phase. This research utilizes a comprehensive data set acquired from the Telemonitoring Database for Parkinson's disease which contains clinical, genetic, and neuroimaging information from patients, employing a quantitative research design. The dataset consists of 5,875 patient records which included demographic information, assessment of motor and non-motor symptoms, and vocal impairment features vital for PD diagnosis. The stages of PD were classified using five ML models: Support Vector Machine (SVM), Random Forest, Decision Trees, Gradient Boosting, and Neural Networks which were all rigorously trained and tested to ensure precise classification. The models were also measured on accuracy, precision, recall, F1 score, and their cross-validated performance for generalizable reliability. Out of all models tested, Decision Trees came out on top with an impressive but potentially overfitting bias accuracy of 99.32%. Random Forest and Gradient Boosting also performed well with over 96% accuracy demonstrating their effectiveness on complex high dimensional data. Both the SVM and the Neural Networks were less accurate than other methods, but their use in initial screenings and dealing with nonlinear data relationships showed greater potential. The results of this study demonstrate ML models can transform PD diagnostic processes with early and precise detection that drastically improves patient care and management, optimizing treatment strategies and outcomes. This study supports the use of these models in clinical practice as they could provide accurate diagnostics, help track the course of the disease, and enable targeted adjustments to therapy. Further development of these models, broadening the diversity of the datasets, and investigating their practical use to ensure clinical relevance will improve outcomes for patients while addressing their optimal care needs.

INTRODUCTION

Parkinson's Disease (PD) presents both a multidimensional and complex experience due to its myriad symptoms with physical and non-physical manifestations which deteriorate the patient's quality of life Furthermore, it greatly burdens caregivers along with the healthcare industry with considerable challenges in providing adequate care. The condition is neurodegenerative in nature and primarily impacts the motor functions due to the loss of dopamine neurons in the substantia nigra region of the brain which controls movement (Karabayir et al., 2020). PD symptoms include tremors, rigidity, bradykinesia, and postural instability. These symptoms arise as a result of the disease, progressively worsening with time, which means patients increasingly struggle with performing basic daily tasks while increasingly losing autonomy. The non-motor symptoms of Parkinson's disease further complicate diagnosis as they tend to surface first. Gradation occurs with time and olfactory issues along with sleep disturbances, constipation, low mood, and cognitive decline which all occur prior to the emergence of motor symptoms (Kamran et al., 2021). Sleep disturbances in PD include three distinct conditions; restless legs syndrome, sleep apnea, and REM sleep behavior disorder.

Gastrointestinal problems causing constipation emanate from sluggish bowels and disrupt a patient's comfort, as well as their nutritional health. A lack of neurochemicals in the brain caused by PD will both lead to depression and anxiety, common mood disorders in PD patients. Managers need to realize that Parkinson's dementia comes from a broad range of memory-related cognitive decline, which also includes attention and executive function reasoning skills. The non-motor symptoms are still essential because they affect life quality but pose difficult treatment challenges that standard Parkinson's disease medications have limited success in addressing (Heinzel et al., 2019). The intricate nature of Parkinson's disease alongside its infection psychology necessitates a holistic approach, as both the body and mind are deeply intertwined. The mental and emotional effects of this disease are profound, significantly impacting the patient's quality of life. Patients gradually lose the ability to perform daily activities, which ultimately leads to social withdrawal and isolation. The external symptoms invite societal discrimination, which exacerbates depressive feelings and isolation. The progression of Parkinson's disease places an increasing burden of emotional and financial strain on caregivers. Daily care responsibilities involve managing behavioral and cognitive symptoms alongside personal and occupational responsibility maintenance (Ghaneet al., 2022).

Due to focusing on Parkinson's disease (PD), this study's research portrayals are paramount to

medical diagnostic work because PD poses intricate diagnostic problems. The machine learning diagnostic models developed automated algorithms surpass existing paradigms in reliability and efficiency (Byeon, 2020). The advancement of machine learning has resolved many of the critical diagnostic hurdles in PD, radically transforming its diagnosis. The methods used for diagnosing Parkinson's disease rely heavily on a physician's clinical expertise and observable symptoms, which remain judgmentally personal in nature. Such subjectivity leads to varying accuracy in diagnosis due to inconsistency among different healthcare providers and over divergent diagnostics frameworks, affecting condition-consistency and inter-establishment diagnostic consistency. This research's goal is to apply machine learning models to vast datasets to identify patterns of Parkinson's disease to replace subjective clinical evaluations with objective standardized techniques (Araújo et al., 2022). Machine learning systems can process vast sets of clinical documents, genetic materials, and submitted patient results to uncover elusive evaluative patterns and relationships, patterns of data that evaluators may miss. Integrating traditional motor symptom data with early non-motor symptoms and biomarkers yields comprehensive disease surveillance through ML systems.

The use of technology, like machine learning, can identify PD symptoms and biomarkers years in advance, which allows diagnosis even earlier than what is possible through conventional medical evaluation (Blauwendraat et al., 2020). The application of machine learning in the PD field has also shown benefits in optimizing their diagnostic processes. Algorithms based on machine learning scan data far quicker than traditional diagnostic techniques, providing insights well ahead of what standard methods would yield. Machine learning's enhanced diagnostic accuracy enables swifter tailored diagnostic care for each patient, optimizing healthcare resources for the organization by assisting in diverse operate with efficient patient flow. The influence of machine learning in early and accurate PD diagnosis shifts greatly the treatment strategies and outcomes for patients. Timely intervention greatly improves care strategies (Chen et al., 2023).

LITERATURE REVIEW

When analyzing the paralysis tremor, also known as shaking palsy, Dr. James Parkinson in 1817 noted the loss of motor function in his patients—he called it Parkinson's disease (PD) ("An Essay on the Shaking Palsy"). PD therapy breakthroughs stemmed from the PD anatomical and clinical framework established by Parkinson's. That basis enabled later innovations in treatment approaches of PD that came about in the latter half of the 20th century (Noor et al., 2020). During the early 1960s, the breakthrough underlying the deficiency of dopamine produced within

the brains of patients diagnosed with Parkinson's, greatly fueled treatment development. This led to the creation of Levodopa, now considered the first-line medication for PD. Because of a better understanding of the multi-faceted nature of the disease, treatment options expanded to include dopamine agonists, MAO-B inhibitors, and even deep brain stimulation (Pahuja & Nagabhushan, 2021). Throughout the progression of Parkinson's, the destruction of the dopaminergic neurons occurs at the substantia nigra pars compacta area which coordinates muscle movement through the biochemistry of the brain. The damaged neurons reduce the amount of dopamine-containing neurotransmitters in a large portion of the brain tissue which is crucial for the cessation of smooth muscle actions (Priyadharshini et al., 2024). In patients suffering from PD, proteinaceous Lewy bodies, which contain the alpha-synuclein protein, are found in their brains and are a central component to the progression of PD, Currently research shows that the spreading misfolding of alpha-syncline plays an important role in the progression of PD.

CLINICAL DIAGNOSIS

Diagnosing Parkinson's Disease (PD) is challenging due to a lack of clear biological markers. To make a clinical diagnosis, physicians evaluate core motor symptoms, which include: tremor, bradykinesia, rigidity, and postural instability. Patient history and neurological examinations are also important. Most clinicians follow The UK Brain Bank Criteria which mandates the presence of bradykinesia and at least one of the following: muscle rigidity, rest tremor, or postural instability. Diagnosis is invalid unless essential tremor, multiple system atrophy (MSA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), or drug-induced Parkinsonism are ruled out. While categorically diagnosing PD, a positive response to Levodopa (dopaminergic medication) reinforces the diagnosis, but is not conclusive as many other Parkinsonian disorders respond partially or fully to treatment.

SUBJECTIVITY AND VARIABILITY IN CLINICAL ASSESSMENT

The evaluation and diagnostic stages of Parkinson's disease (PD) are mostly holistic. This is because a doctor's evaluation is dependent on how well the clinician understands and interprets the signs and symptoms presented by the patient. The early signs of PD are easily overlooked, and they may be confused with the usual signs of aging or other neurological disorders. As a result of this subjectivity, the rate of misdiagnosis is very high. There are studies that show that as high as 15% of individuals diagnosed with PD could be misdiagnosed with other disorders. Also, a lot of PD cases are diagnosed late because the symptoms presented do not conform to textbook characteristics.

Symptoms often begin asymmetrically, with one side more affected than the other. The rest symptoms that include anosmia, sleep, mild cognitive decline, and other non-motor features tend to emerge some time before movement symptoms develop; however, they are not included in most diagnostic criteria.

INCONSISTENCIES IN DIAGNOSIS

Outcomes differ among practitioners for a given diagnosis due to blending of training and interpretation, along with a lack of standardized evaluation methods in routine practice. This inconsistency causes delays or misdiagnoses, exacerbating challenges during the initial phases of many conditions. Given that the symptoms of PD and other neurodegenerative conditions tend to overlap, objective diagnostics are crucial for early accurate detection.

THE NEED FOR OBJECTIVE DIAGNOSIS

Early intervention and PD treatment greatly benefits from an accurate prognosis and timely implementation, but observation on its own cannot be clinically relied on. The development of biomarkers and neuroimaging, as well as ML-based diagnostics, may create new opportunities for accuracy. Distinguishing PD from other similar conditions has the potential to become more reliable with these technologies. By incorporating objective tests into benchmarks and standard protocols, the subjectivity of the testing process may be minimized, thereby improving early detection and overall patient outcomes.

MOTOR AND NON MOTOR SYMPTOMS

The Parkinson's disease (PD) is a type of neurodegenerative disorder that manifests with various symptoms, both motor and non-motor, which are interwoven and intricate challenges that a patient needs to address. PD is usually characterized by motor symptoms; however, patients tend to develop non-motor symptoms significantly prior to any motor deficits, which greatly diminishes their quality of life. All these symptoms considerably hinder acknowledging the existence of PD, evaluating its progression, and measuring the effectiveness of the treatment (Nilashi et al., 2022). Some of the symptoms are fellows:

Tremor – This condition usually results in a characteristic shaking motion, beginning with one hand or leg before progressing to both sides of the body. It differs from essential tremor because PD shaking is static during voluntary movements whereas essential PD shaking is worsened during activity (Pereira et al., 2018). Bradykinesia (Slowness of Movement).

Bradykinesia –- Another feature of the PD syndrome is bradykinesia. It manifests as a distinct drop in physical activity and increase in time taken to complete tasks such as walking, writing, or

speaking. Patients found it hard to initiate movement, develop a stoppage of motion, and reduced facial expressiveness, which is termed as hypomania (Raj et al., 2024).

Postural Instability – Over the course of PD progression, balance and coordination become increasingly difficult, resulting in frequent falls and risk of injury (Lin et al., 2021). Postural instability arises in the later stages of PD and becomes one of the major contributing factors of disability. In the earlier phases of the disorder, the difficulties with movement are primarily unilateral rather than bilateral. Advancement of the disease results in more widespread symptoms that adversely affect the patients' mobility during ambulation and their capacity to live independently. The hallmark features of Parkinson's disease are primarily caused by dysfunction of basal ganglia pathways controlling movement. However, non-motor symptoms tend to precede motor symptoms by years to decades, arising alongside or together, decades to years prior. Systemic manifestations are crucial to the progression of the disease, leading to greater challenges for the patients (Wang et al., 2024).

Sleep Disorders – Comprised of insomnia, excessive daytime sleepiness, restless legs syndrome, and RBD, these disorders are often present in PD patients. Sleep disturbances negatively impact the patients' health, which in turn exacerbates motor symptoms (Su et al., 2020).

Autonomic Dysfunction – A person with Parkinson's disease faces complications involving the autonomic nervous system, which causes symptoms such as hypotension, over-urination, abnormal sweating, and digestive issues. In addition to these complications, patients may struggle with fatigue, lightheadedness, and gastrointestinal distress which complicates managing the disease (Sigcha et al., 2023).

CHALLENGES IN THE DIAGNOSIS OF PARKINSON'S DISEASE

Diagnosing One of the most significant clinical challenges in diagnosing and treating Parkinson's disease within the field of neurology is the lack of objective, definitive biomarkers along with overlapping symptoms with other conditions, and devoid of clinical assessments. PD's primary symptoms, including tremors and other motor features, become increasingly more difficult to rely in early stage detection, contributing to a misdiagnosed diagnosis (Yang et al., 2021). PD poses one of the most challenging conditions to identify in clinical practice. Detecting Parkinson's disease (PD) early is important for optimizing treatment strategies. PD presents with a gradual tempo. Its earliest signs include olfactory loss, constipation, sleep problems, and other low-grade symptoms which are vague and insidious. For many patients, non-motor signs present up to a decade prior to the motor symptoms including tremors, rigidity, and bradykinesia.

During this time, patients are often overlooked as they are construed as normal aging processes or unrelated to any underlying health issues, resulting in prolonged delays in diagnosis. Once PD is diagnosed, significant and irreversible neuronal damage has often been sustained. The overlap with the early signs of multiple system atrophy (MSA), progressive supra nuclear palsy (PSP), essential tremor, and other neurodegenerative disorders further muddles precise differentiation adding to the diagnostic difficulty. Current diagnostic paradigms, such as the UK Brain Bank Criteria, rely heavily on the presence of motor sign and necessitate subjective clinical evaluation.

Trusting clinicians to make diagnosis based solely on observation causes significant differences between doctors. Each clinician's unique personal experience and training can color symptom interpretation and diagnostic results. While the Unified Parkinson's Disease Rating Scale (UPDRS) and Hoehn and Yahr staging systems are common, both are inconsistently applied, especially during the early phases where symptoms are mild or vague. The stages of PD that are critical for treatment decisions may be poorly aligned even while assessing the same patient. Also, in the absence of objective criteria, so many cases go misdiagnosed or without diagnosis, and patients often initiate therapy only after advancing to a stage where approach is much less effective. Although non-motor symptoms are impactful, they are not formally integrated into chronological frameworks, which assists with further reducing the chances of identifying a patient earlier. To overcome these challenges, unilateral approaches focused on neuro biomarkers such as genetic and biochemical markers are PD-relevant to detect early biochemical shifts associated with the condition. New prospects for precise and timely diagnosis also stem from brain imaging and developing technologies like artificial intelligence (AI). Fluently distinguishing PD from other similar disorders, subtle telltale signs in large clinical, imaging and genetic datasets using machine learning (ML) algorithms trained on them far exceeds what human clinicians could achieve.

The applications of these AI tools can streamline the processes of determining a diagnosis, lessening bias, and enhancing precision in clinical evaluations. Limitations of Neuroimaging in PD Diagnosis.

In the case of diagnosing Parkinson's Disease. (It's essential to have specialized neuroimaging procedures like Positron Emission Tomography (PET) scans and Dopamine Transporter (DaT) scans, as these differentiate between PD and other types of Parkinsonism by identifying certain brain functions and structures that are abnormal. PET scans evaluate the metabolic function of the basal ganglia, and DaT scans measure the dopamine transporter

concentration, as both are diminished in PD.

CONCEPTUAL FRAMEWORK

This investigation's research framework aims to develop an advanced ML model for multi-class PD diagnosis. The PD framework integrates the theoretical facets of PD with the practical clinical ML model applications for diagnosing PD, addressing various PD problems simultaneously (Priyadharshini et al., 2024). All the components of the research framework are consolidated into one diagram, which aligns with the research outline. Data Input.

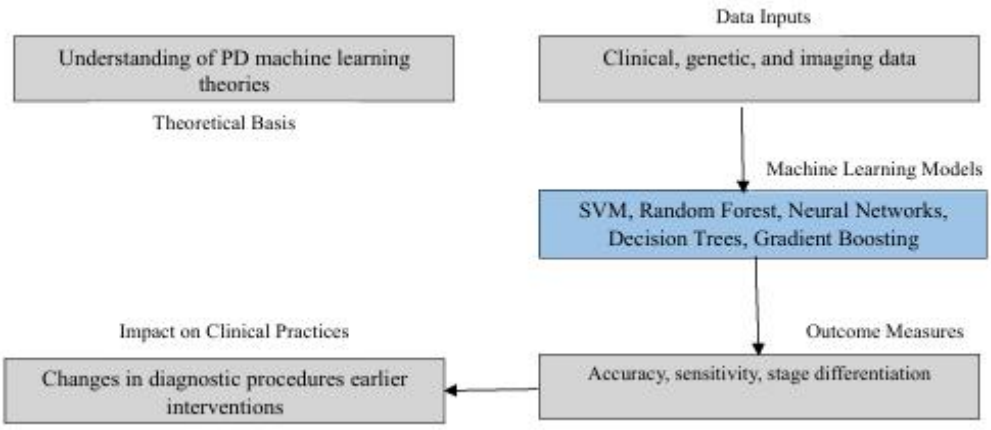

FIGURE 0- 1 CONCEPTUAL FRAMEWORK

METHODOLOGY

DATA COLLECTION

The source for this study's Parkinson's disease datasets is primarily from the Kaggle platform which is known for housing medical and healthcare datasets. To build machine learning models for diagnosing the disease, patients' clinical records, along with their genetic profiles and neuroimaging data, are sufficient. Showing the opportunity for accessible data-based studies has the potential for wider impact through increasing public trust in research results and making them available for scrutiny before publication. For the proposed system, these datasets are necessary in sufficient volume to achieve target goals and expose the model to adequate training and validation. The study requires three primary information sources. Clinical information system consists of monitoring patients' motor and non-motor symptoms, disease progression, and treatment response. This dataset includes data on the physical features and development patterns of Parkinson's disease, diagnostic criteria and longitudinal data, and also contains psychiatric evaluation data for mental impairment and somatic disorders which allows to perform adequate medical holistic assessments. The study aims to obtain genetic data associated with PD

by examining mutations and other genetic factors known to influence the disease's pathology. Such genetic information is critical in establishing the hereditary components of the disease, especially when differentiating various stages of PD and Glibre disorder and other neurodegenerative diseases relevant disease stages.

This approach fulfills the two functions of predicting possible susceptibility and forecasting disease progression to facilitate prompt intervention. The studies will collect neuroimaging data through MRIs, PET scans, and Dopamine Transporter (DaT) scans. These imaging modalities can detect changes in the brain associated with Parkinson's disease.

DATA PROCESSING

Preprocessed data is a vital element necessary for effective models in machine learning, especially in the case of diagnosing Parkinson's disease since it ensures that the datasets are accurate and well-organized. Remove all irrelevant errors including inconsistencies and missing gaps, as well as "outlier" values, using statistical imputation methods such as the mean, median, or mode. This increases accuracy and enhances performance significantly. Further, if any value is removed or changed, retaining the meaningful relevance of the model is crucial for establishing reliability. Normalization is another method in which certain features can be scaled, like the aforementioned clinical scores, genetic markers or any other pertinent variable that needs monitoring while being measured in different scales and units to be used effectively to prevent large values from dominating the variables. Lastly, conducting feature selection aids in retaining attributes that hold significance in recognizing Parkinson's disease. Through expert clinical guidance, correlation analysis and scoring feature importance helps in eliminating excess redundancy to effectively lower the number of informal predictors.

Choosing important features enhances a model's efficiency and reduces the risk of overfitting, which strengthens the model. For imbalanced datasets, such as uneven representations of PD stages, augmentation techniques like SMOTE are used to create synthetic imbalance and balance the dataset. This ensures that all minority classes are present, thus allowing the algorithm to learn without bias toward majority classes.

MACHINE LEARNING ALGORITHMS

This particular research project tests different machine learning algorithms to assess their effectiveness in classifying the stages of Parkinson's Disease. The chosen algorithms showed better results in handling complex datasets and were successful in past classification tasks. The assessment includes a range of machine learning algorithms which comprise of:

Support Vector Machine (SVM): SVM is very useful in performing classification operations on multi-dimensional inputs; SVM does well on data where several features are present. Its mathematical model solves by finding a splitting hyperplane which optimally separates the given points into their respective regions for some maximum distance between them, especially if the classes are separable.

Random Forest: Predictive accuracy increases in an algorithm that makes combined prediction based on more than one decision tree's prediction. The method lends itself to large-scale datasets with multiple high-dimensional features as commonly encountered in medical databases. The feature importances of the diagnostic characteristics of Random Forest could be used by healthcare professionals to identify which characteristics are the most relevant to diagnose PD.

Gradient Boosting: The analytical approach followed by this model generates models sequentially to address errors unearthed by preceding models. The advantage of the algorithm is its capability to handle complicated data associations simultaneously on classification problems with balanced data distribution.

Neural Networks: When it comes to their power of deep learning, neural networks show killing capacity of finding complex patterns in large unordered data sets. The neural networks show the ability to recognize complex patterns in time series data sequences and processed images and improve the diagnosis of the stages of PD.

Decision Trees: A decision system organizes information by layering decisions with their expected consequences in a tree structure. Decision Trees are excellent when interpretation is required because of their straightforwardness. When determining the level of Parkinson's disease symptoms and staging the progression of the disease, clear and ample data partitions can be obtained using Decision Trees.

To evaluate the performance of the algorithms, they need to be trained utilizing the available dataset for determining the stages of Parkinson's disease. Effectiveness of the algorithms will be evaluated using multiple assessment metrics including accuracy, sensitivity, specificity, and F1 score. A comprehensive machine learning approach to studying the progression of Parkinson's disease will be guided by a robust framework of diagnostic testing.

MODEL EVALUATION AND VALIDATION

Distinct performance benchmarks will evaluate the implementation and reliability of the machine

learning models developed for this research study. Several metrics are critical for assessing how well the models perform stage classification of PD and how well they can detect various neurological disorders. Evaluation of this study incorporates an evaluative process that utilizes these techniques:

Accuracy: The Accuracy metric captures the simplest frame of evaluation by defining correct model outputs in relation to total possible outputs for each classification level. Accuracy assesses the ratio of right estimates made to the total estimates made which comprises both positive and negative estimates whether true or false. Relying on accuracy alone can be dangerous especially with datasets that are unbalanced as the dominant class tends to bias the results.

Sensitivity (True Positive Rate): With sensitivity, we assess how well the model uncovers patients with Parkinson's disease among all patients that received a diagnosis. The accuracy of a medical model hinges upon its sensitivity as this is the extent to which the model is able to capture all positive cases. Important and pivotal in medical diagnosis is the level of sensitivity due to the clinician's preposterous mistakes of overlooking real cases (false negatives) which may lead to worsening treatment or worsening of the patient's condition.

Specificity (True Negative Rate): In specificity, it assesses how accurately the model identifies and diagnoses healthy subjects as people who do not have Parkinson's disease (true negatives). The inappropriate designation of non-Parkinson's disease patients as healthy individuals requires specification to avoid the risk of exposing them to unnecessary interventions.

F1-score: The F1-score derives from a model's precision as well as its recall (sensitivity), using harmonic mean to synthesize both metrics into one. The F1-score is adept at performing on tasks with data imbalance because it takes into account the model's ability to recall positive cases (recall) and how well it avoids making erroneous optimistic predictions (precision). The F1-score analyzes model performance better than accuracy, especially in scenarios with imbalanced datasets.

Cross Validation: Cross-validation techniques will ensure proper model generalization and defend against overfitting based on training data.The method trains models using unique training-testing pairs and splits datasets into groups of subsets. In the K-fold cross-validation method, the data is divided into K pieces and for each piece, the model is trained and evaluated, resulting in K evaluations with different testing subsets.

Accuracy metrics: shown in the confusion matrix will provide in-depth insights into the model's data classification performance. It displays the accuracy metrics relating to PD patient

identification alongside true and false values, exhibiting dual-classification of PD and healthy subjects plus erroneous assignment of PD and healthy patients. Through the confusion matrix, researchers grasp the workings of the model including its detection capabilities and the strength, weaknesses, and patterns of various misclassifications between the recognized classes.

RESULTS

The data for this research was collected from the Tele monitoring Database of Parkinson's Disease, which contains the records of various metrics of patients suffering from Parkinson's disease. This vast database consists of 5,875 records which capture the health of each patient by measuring motor and non-motor symptoms associated with evolving Parkinson's disease. In addition, the dataset includes some demographic information such as the age of participants ranging from 36 to 85 years, with an average age of 64.8 years, and the proportion of 31.78% females to 68.22% males. The study sample indicates that female patients constitute 31.78% while male patients constitute 68.22% of the total sample. The dataset also includes the Motor UPDRS ratings, total UPDRS scores as well as the voice metrics Jitter and Shimmer, used to gauge the level of symptoms associated with Parkinson's disease. These voice traits are essential in identifying vocal problems that happen as a result of 37 Parkinson's diseases. Machine learning requires a stable automated system which this project aimed for, and in preparation, there were several steps required to refine the dataset to stable conditions. For this analysis, the researchers looked into incomplete information sets and applied appropriate filling techniques to maintain the structure of the data freeze. The research group scaled all numerical features to prevent bias, as explained by Shojaie et al, and eliminated bias through feature normalization. The feature selection processes achieved their goals using some combination of correlation evaluations and feature importance metrics to remove irrelevant or redundant characteristics which reduced computational effort while optimizing model performance. An essential model and evaluation bias-free accuracy check requires a split data preparation technique where a 70:30 ratio is standard for dividing training test datasets. The processing steps designed an infrastructure sufficient for deploying algorithms and achieving effective results in the diagnosis of Parkinson's disease.

SUPPORT VECTOR MACHINE

Utilizing features from the dataset, SVM was able to reach a discriminating success of 88.32% in distinguishing patients with Parkinson's disease from those without. SVM output is so complex that I will analyze it step by step:

TABLE 1 SVM CLASSIFICATION REPORT

SVM Accuracy		0.8831537152580828	
SVM Classification Report			
Precision	**Recall**	**F1-score**	**Support**
0.86	0.91	0.89	876
0.91	0.86	0.88	887
Accuracy	-	-	0.88	1763
Macro avg	0.88	0.88	0.88	1763
Weighted avg	0.88	0.88	0.88	1763

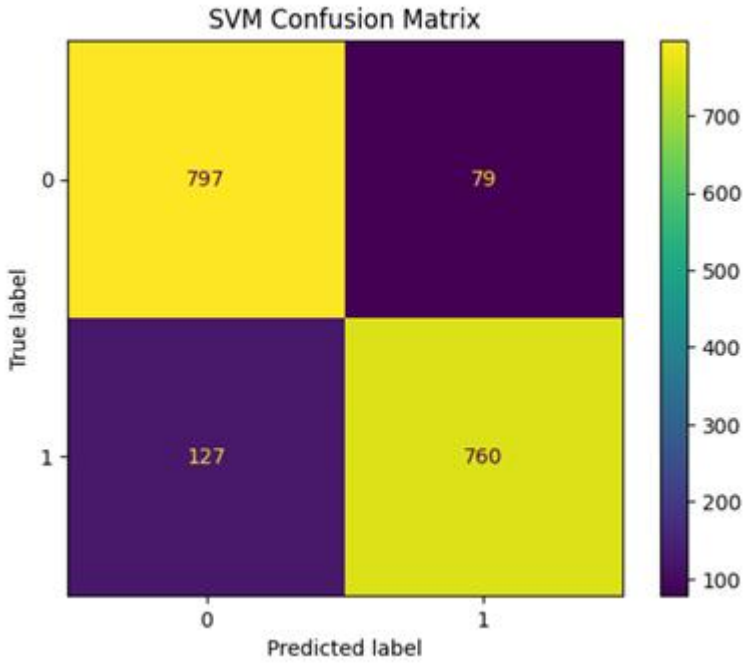

FIGURE 2 CONFUSION MATRIX ANALYSIS

RANDOM FOREST

According to the results of a study, the effect of Random Forest was 96.31% of accurate model predictions. The identification system demonstrates its high classification accuracy between PD patients and non-PD subjects from the input data characteristics. The quality of the model performance indicators are revealed in the evaluation provided as follows:

TABLE 2 RANDOM FOREST CLASSIFICATION REPORT

Random Forest Accuracy		0.9631310266591038		
Random Forest Classification Report				
Precision	**Recall**	**F1-score**	**Support**	
0.97	0.96	0.96	876	
0.96	0.97	0.96	887	
Accuracy	-	-	0.96	1763
Macro avg	0.96	0.96	0.96	1763
Weighted avg	0.96	0.93	0.96	1763

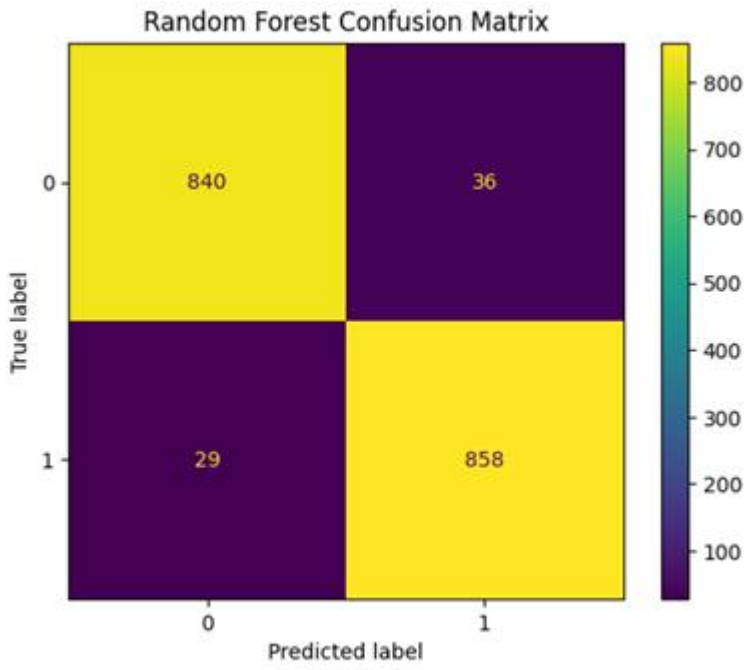

FIGURE 3 CONFUSION MATRIX ANALYSIS

GRADIENT BOOSTING

Gradient Boosting is a good diagnostic tool for Diagnosis of Parkinson's Disease due to 97.73% Accuracy. The subsequent section of this subsection shows model evaluation full inspection:

TABLE 3 GRADIENT BOOSTING CLASSIFICATION REPORT

Gradient Boosting Accuracy			97.73%	
Gradient Boosting Classification Report				
Precision	**Recall**	**F1-score**	**Support**	
0.97	0.98	0.98	876	
0.98	0.97	0.98	887	
Accuracy	-	-	0.98	1763
Macro avg	0.98	0.98	0.98	1763
Weighted avg	0.98	0.98	0.98	1763

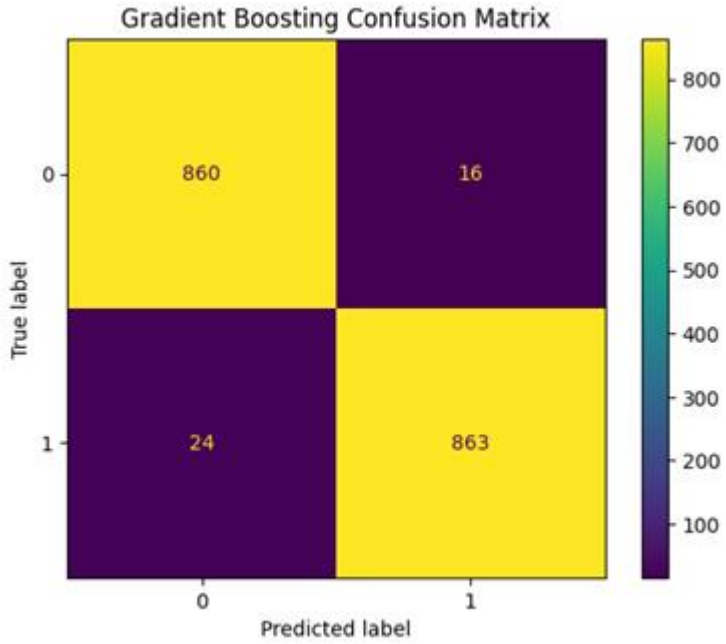

FIGURE 4 CONFUSION MATRIX ANALYSIS

NEURAL NETWORK

The examined models, the Decision Tree model proved highly accurate by reaching 99.32% in detecting Parkinson's or non-Parkinson's conditions. We will focus on elaborating results obtained from model evaluation in this segment:

TABLE 4 NEURAL NETWORK CLASSIFICATION REPORT

Neural Network Accuracy		96.65%		
Neural Network Classification Report				
Precision	**Recall**	**F1-score**	**Support**	
0.96	0.97	0.97	876	
0.97	0.96	0.97	887	
Accuracy	-	-	0.97	1763
Macro avg	0.97	0.97	0.97	1763
Weighted avg	0.97	0.97	0.97	1763

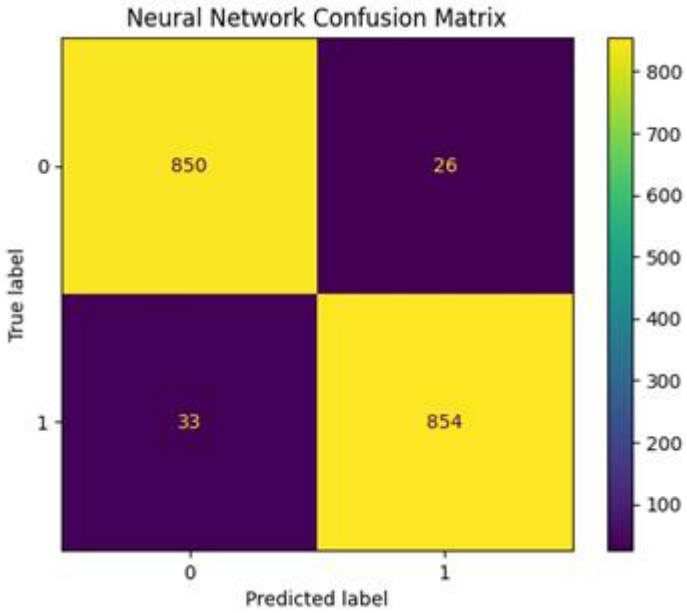

FIGURE 5 CONFUSION MATRIX ANALYSIS

DECISION TREES

The examined models, the Decision Tree model proved highly accurate by reaching 99.32% in detecting Parkinson's or non-Parkinson's conditions. We will focus on elaborating results obtained from model evaluation in this segment:

TABLE 5 DECISION TREE CLASSIFICATION REPORT

Decision Tree Accuracy			0.9931934203062961
Decision Tree Classification Report			
Precision	**Recall**	**F1-score**	**Support**
0.99	1.00	0.99	876
1.00	0.99	0.99	887
Accuracy	-	-	0.99	1763
Macro avg	0.99	0.99	0.99	1763
Weighted avg	0.99	0.99	0.99	1763

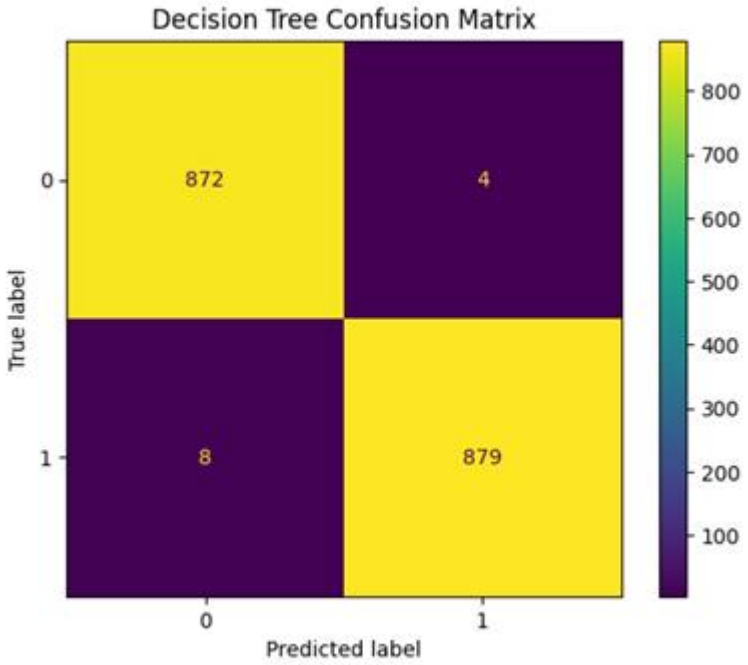

FIGURE 6 CONFUSION MATRIX ANALYSIS

COMPARATIVE ANALYSIS

Different algorithms display varying levels of accuracy based on the characteristics of the data and the complexity of the task at hand. For instance, while the Support Vector Machine (SVM) is popular for classification tasks, it had the lowest classification performance in this comparison analysis due to its noise sensitivity and struggles with handling large datasets featuring complex

decision boundaries. On the other hand, Random Forest, an ensemble method, significantly outperformed others because of its robustness against overfitting problems as well as high-dimensional data handling capabilities. The Decision Tree model achieved outstanding results from all models by obtaining the highest accuracy reflecting both strong interpretability and successful pattern learning. Although known as overfitting prone in other contexts, Gradient Boosting also performed impressively illustrating a primary benefit from boosting methods which iteratively learn reducing bias and variance. The Neural Network's results showed that even though trailing slightly behind Gradient Boosting and Decision Tree models, it maintained commendable precision indicating decent ability to capture non-linear relationships albeit needing more data and computational power than preferred. As observed with existing literature supporting tree-based ensembles applied to structured data classifications, Decision Trees followed by Gradient Boosting models provided superior contextual performance within this dataset scope.

The results validate the diagnostic effectiveness of machine learning algorithms that analyze the Parkinson's disease Tele monitoring Database information to detect Parkinson's disease. The clinical applicability of these models became more evident because they brought different advantages and weaknesses to the diagnosis process (Gomati et al., 2024). The accuracy rates of Decision Trees and gradient-boosting techniques reached 99.32% and 97.73%, respectively. The Decision Tree model exhibited nearly flawless performance, yet researchers believe the performance signifies substantial overfitting, as based on (Keserwani et al., 2024). Its ability to explain reasoning processes makes the Decision Tree model suitable for clinical practice because healthcare providers need to understand the basis of diagnoses to build trust and patient compliance. Gradient Boosting provides robust performance against the complex data variations found in clinical settings because it corrects errors in sequential weak learners, which makes the algorithm effective for neurodegenerative condition analysis. The performance of Random Forests and Neural Networks yielded very similar accuracy rates of 96.31% and 96.65%, respectively. The models provide an effective precision-recall balance that minimizes false negative and positive outcomes, which prove essential to medical diagnostic applications. The results demonstrate how ensemble methods paired with deep learning should be used to handle complex biomedical data effectively (Sajal et al., 2020). Support Vector Machine (SVM) maintains adequate performance levels through an 88.32% accuracy rate, which makes it suitable for workflows depending on fast and understandable model types. Before further assessments, the

tool enables a quick method to determine patient condition status.

The report provides accuracy performance charts as a distinct comparison to clearly demonstrate results from different machine learning models. This part elaborate describes performance outcomes for the analyzed models and explore the advantages and disadvantages of the models in focus:

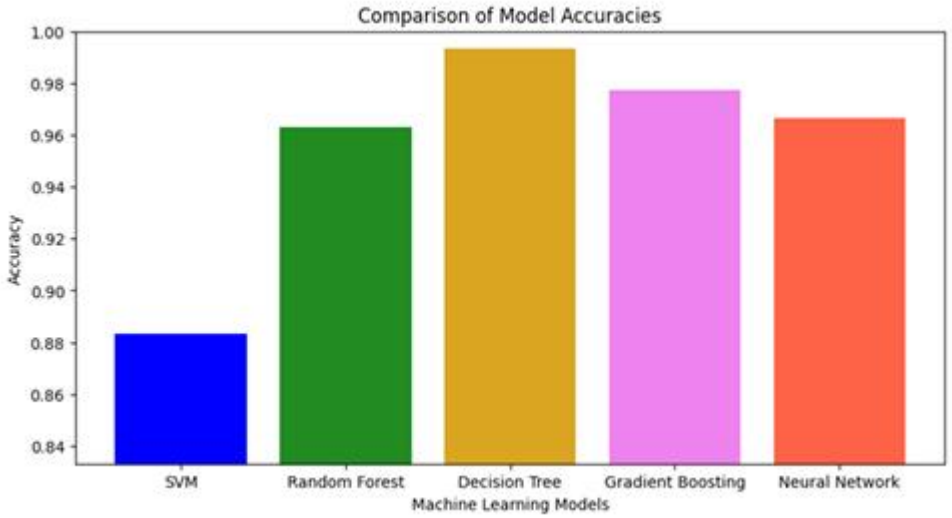

FIGURE *5* COMPARATIVE ANALYSIS

CONCLUSIONS

This study supports the use of machine learning technology as a transformative diagnostic and technology in managing processes for Parkinson's Disease. The research validated diagnostic accuracy through exhaustive model testing with Support Vector Machines, Random Forests, Neural Networks, Decision Trees, and Gradient Boosting methodologies. The constructed models achieve accurate identification of Parkinson's disease while ensuring practical clinical adaptations that streamline patient care with increased precision and multifaceted sophistication. The findings of this study incorporate machine learning techniques into the operational processes of healthcare systems and refine the ongoing literature. These technologies improve medical diagnostic procedures through quicker and more accurate recognition of diseases, especially in patients with Parkinson's disease, so essential in early diagnosis. The research illustrates the potential of machine learning in devising tailored treatment plans, which is an accepted standard of care today. Enhanced machine learning technology improves treatment plans by tailoring them to the individual's needs, thus increasing the efficacy of interventions. Healthcare systems are automating the use of enhanced technologies as part of routine clinical practices. The analysis

highlights the need for sustained investigation to resolve the issues of data quality and expand the scope of model applications as well as their implementation for practical use.

To fully implement these models and revolutionize healthcare, clinicians need to work together with researchers to fully optimize trustworthiness and precision. This study supports the need for advanced machine learning applications in healthcare, especially in the diagnosis and management of Parkinson's Disease. This study demonstrates that machine learning has become a fundamental tool in healthcare due to its powerful impact on diagnostic and treatment processes, especially enhancing life quality for patients suffering from chronic diseases such as Parkinson's. It lays the foundation needed for further research on the use of machine learning in different functional areas of healthcare.

REFERENCES

Abdollahi, S., & Safa, R. (2024). Machine learning and AI for advancing Parkinson's disease diagnosis: exploring promising applications. Big Data and Computing Visions, 4(1), 12-21.

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S., Karthikesalingam, A., King, D., ... & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ digital medicine, 4(1), 65.

Ahmadi Rastegar, D., Ho, N., Halliday, G. M., & Dzamko, N. (2019). Parkinson's progression prediction using machine learning and serum cytokines. NPJ Parkinson's disease, 5(1), 14.

Ahsan, M. M., Luna, S. A., & Siddique, Z. (2022, March). Machine-learning-based disease diagnosis: A comprehensive review. In Healthcare (Vol. 10, No. 3, p. 541). MDPI.

Araújo, D. C., Veloso, A. A., Gomes, K. B., de Souza, L. C., Ziviani, N., Caramelli, P., & Alzheimer's Disease Neuroimaging Initiative. (2022). A novel panel of plasma proteins predicts progression in prodromal Alzheimer's disease. Journal of Alzheimer's Disease, 88(2), 549-561.

Atri, R., Urban, K., Marebwa, B., Simuni, T., Tanner, C., Siderowf, A., ... & Lancashire, L. (2022). Deep learning for daily monitoring of Parkinson's disease outside the clinic using wearable sensors. Sensors, 22(18), 6831.

Badawy, M., Ramadan, N., & Hefny, H. A. (2023). Healthcare predictive analytics using machine learning and deep learning techniques: a survey. Journal of Electrical Systems and Information Technology, 10(1), 40.

Balakrishnan, A., Medikonda, J., & Namboothiri, P. K. (2022). Role of wearable sensors with machine learning approaches in gait analysis for Parkinson's disease assessment: a review.

Engineered Science, 19(13), 5-19.

Bhandari, N., Walambe, R., Kotecha, K., & Kaliya, M. (2023). Integrative gene expression analysis for the diagnosis of Parkinson's disease using machine learning and explainable AI. Computers in Biology and Medicine, 163, 107140.

Blauwendraat, C., Nalls, M. A., & Singleton, A. B. (2020). The genetic architecture of Parkinson's disease. The Lancet Neurology, 19(2), 170-178.

Byeon, H. (2020). Application of machine learning technique to distinguish parkinson's disease dementia and alzheimer's dementia: Predictive power of parkinson's disease-related non-motor symptoms and neuropsychological profile. Journal of personalized medicine, 10(2), 31.

Byeon, H. (2020). Is the random forest algorithm suitable for predicting parkinson's disease with mild cognitive impairment out of parkinson's disease with normal cognition?. International journal of environmental research and public health, 17(7), 2594.

Cardoso, F., Goetz, C. G., Mestre, T. A., Sampaio, C., Adler, C. H., Berg, D., ... & Trenkwalder, C. (2024). A statement of the MDS on biological definition, staging, and classification of Parkinson's disease. Movement Disorders, 39(2), 259-266.

Celik, G., & Başaran, E. (2023). Proposing a new approach based on convolutional neural networks and random forest for the diagnosis of Parkinson's disease from speech signals. Applied Acoustics, 211, 109476.

Chatterjee, K., Kumar, R. P., Bandyopadhyay, A., Swain, S., Mallik, S., Li, A., & Ray, K. (2023). PDD-ET: Parkinson's disease detection using ML ensemble techniques and customized big dataset. Information, 14(9), 502.

Chen, B., Xu, M., Yu, H., He, J., Li, Y., Song, D., & Fan, G. G. (2023). Detection of mild cognitive impairment in Parkinson's disease using gradient boosting decision tree models based on multilevel DTI indices. Journal of Translational Medicine, 21(1), 310.

Chen, M., Sun, Z., Xin, T., Chen, Y., & Su, F. (2023). An interpretable deep learning optimized wearable daily detection system for Parkinson's disease. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, 3937-3946.

Dadu, A., Satone, V., Kaur, R., Hashemi, S. H., Leonard, H., Iwaki, H., ... & Faghri, F. (2022). Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts. npj Parkinson's Disease, 8(1), 172.

Dashtestani, H., Zaragoza, R., Kermanian, R., Condy, E., Anderson, A., Chowdhry, F., ... & Gandjbakhche, A. (2019). The quest for functional biomarkers in the prefrontal cortex 59

using functional near-infrared spectroscopy (fNIRS). In Neurophotonics and Biomedical Spectroscopy (pp. 123-136). Elsevier.

De Oliveira, A. P. S., De Santana, M. A., Andrade, M. K. S., Gomes, J. C., Rodrigues, M. C., & dos Santos, W. P. (2020). Early diagnosis of Parkinson's disease using EEG, machine learning and partial directed coherence. Research on Biomedical Engineering, 36, 311 331.

Dentamaro, V., Impedovo, D., Musti, L., Pirlo, G., & Taurisano, P. (2024). Enhancing early Parkinson's disease detection through multimodal deep learning and explainable AI: insights from the PPMI database. Scientific Reports, 14(1), 20941.

Diaz-Pinto, A., Alle, S., Nath, V., Tang, Y., Ihsani, A., Asad, M., ... & Cardoso, M. J. (2024). Monai label: A framework for ai-assisted interactive labeling of 3d medical images. Medical Image Analysis, 95, 103207.

Garcia Santa Cruz, B., Husch, A., & Hertel, F. (2023). Machine learning models for diagnosis and prognosis of Parkinson's disease using brain imaging: general overview, main challenges, and future directions. Frontiers in Aging Neuroscience, 15, 1216163.

García, A. M., Arias-Vergara, T., C Vasquez-Correa, J., Nöth, E., Schuster, M., Welch, A. E., ... & Orozco-Arroyave, J. R. (2021). Cognitive determinants of dysarthria in Parkinson's disease: an automated machine learning approach. Movement Disorders, 36(12), 2862 2873.

Ghane, M., Ang, M. C., Nilashi, M., & Sorooshian, S. (2022). Enhanced decision tree induction using evolutionary techniques for Parkinson's disease classification. Biocybernetics and Biomedical Engineering, 42(3), 902-920.