Artificial Intelligence-Based Parkinson's Disease Identification: A Comparative Analysis of Recurrent Neural Networks and Machine Learning Models

Swati Chauhan[1], Sachin Kumar[2*]

[1] Computer Science and Engineering, The NorthCap University, Gurugram, India

[2]IT Department, TATA Consultancy Service, Gurugram, India

swati20csd010@ncuindia.edu[1], kumar.sachin3@tcs.com[2]

*Corresponding author: Swati Chauhan (swati20csd010@ncuindia.edu)

ABSTRACT

The neurodegenerative disorder weakens the brain's neurological, physiological, and behavioural systems; accurate identification is challenging in the early stages due to mild variances. Parkinson's disease is a major health concern. General signs of this condition include sluggish movements called 'bradykinesia'. Parkinson's disease, a neurological disease, causes tremors, rigidity, imbalance, and loss of motor control. Most symptoms progress over time. This study suggests using supervised algorithms like Logistic regression, SVM, Decision tree, K-NN, Random Forest, Bagging, XGBoost Classifier, ANN /NN and LSTM (part of RNN) to generate the models. to diagnose subjective diseases. The proposed method involves feature selection using filter and wrapper methods, and classification processes. Data for this research were acquired from the UCI Machine Learning Repository. The dataset comprises biological voice recordings from 31 subjects, including 23 patients affected by Parkinson's disease and 8 healthy individuals. Based on the results of the experiment, it can be stated that XGBoost has the highest accuracy with 97%, followed by KNN with an accuracy 95%, while RNN's model LSTM achieved 97%, along with 50 epochs. Compared with prior research, the current study's findings indicate that the proposed approach yields equivalent or even superior outcomes.

Keywords: Parkinson's Disease, Machine Learning models, RNN, Data Science, Voice Measurements.

1. Introduction

Insomnia, changes in mood and behaviour, impaired motor skills, and physical pain are only some of the symptoms of Parkinson's disease (PD), a degenerative brain ailment. The condition primarily results from the degeneration of nerve cells that produce dopamine, a neurotransmitter essential for regulating and coordinating body movements [1]. In this disease, a person first notices tremor on one hand, which may indicate Parkinson's disease in its initial stages. In India, the medical genomics firm MedGenome, in collaboration with the Parkinson's Research Alliance of India (PRAI), conducted the first study of its kind in the Indian population and discovered frequent and unusual genetic variants linked to Young Onset Parkinson's disease (YOPD). India has a slightly lower Parkinson's disease onset age than Westerners. This study states that although the disorder is most often diagnosed around the age of 60 among populations in Western regions, in India, it is roughly 10 years earlier, before 40, and 1 in 20 people are diagnosed. Around 18 per 1,000 over-65s are impacted [2]. Since 1990, there has been an average yearly increase of 3.8% in the mortality rate per 100,000 individuals in India due to Parkinson's disease, a rise of 87.9% [3]. A simple reminder of the four main signs of Parkinson's disease is the abbreviation "TRAP". The letter T symbolises Tremor at rest, R indicates rigidity, A indicates Akinesia, which refers to slow movement, and P indicates posture, which refers to a stooped posture. Typically, these symptoms affect only one side of the body [4]. While the underlying causes of Parkinson's disease are still a mystery. A familial predisposition has been associated with a higher likelihood of disease. Exposure to pesticides, solvents, or air pollution may increase the risk [5]. Nonetheless, the diagnosis cannot be made without first doing a neurological examination. Physical, occupational, and speech therapy in conjunction with medication can successfully manage the symptoms of Parkinson's disease,

Proceedings of ICICCS-2026 (Poster Papers), DMPedia Lecture Notes in Computer Science & Engineering.
ISBN: 978-81-993813-1-5

while there is no cure at this time. Medications may provide temporary relief, but it's not uncommon for symptoms to return or even worsen first thing in the morning—a phenomenon called "morning off" [6]. It has been seen that men are more likely to experience this disease than women. Because there aren't enough high-quality interventions for older people, more studies are needed to develop early diagnostic tools [7]. The death of cells that produce, release, or respond to dopamine in the striatum reduces the number of dopamine transporters, which may cause the condition. These neurons release dopamine, a neurotransmitter critical for regulating motor function and coordinated movement. The degeneration of these cells leads to diminished dopamine availability, a characteristic feature of Parkinson's disease. The cholinergic pathway in the brain delivers dopamine to the caudate-putamen ganglia [8]. A reduction in dopamine-producing cells in the body impairs muscle coordination, hindering movement in Parkinson's disease [9]. Cure from Parkinson's disease relies heavily on dietary changes. Though it affects many parts of health, nutrition cannot stop its progression. Constipation, delayed stomach emptying, and swallowing difficulties are all symptoms of Parkinson's disease, which can also cause a loss of smell and taste as well as a slowdown in the digestive tract's peristaltic motions. Dietary intervention alleviates discomfort and diminishes hospitalisation requirements in individuals with Parkinson's disease, who have an elevated risk of nutrition [10]. Different people have different ideas about what constitutes an early onset of PD. When looking for the early beginning of Parkinson's disease (EOPD), some research has used a cutoff age of 40, while other studies have used a cutoff age of 50 as per Figure1,2 and 3.

Five to ten per cent of PD cases are early-onset, according to most research. Juvenile onset, which occurs before age 21, is a subset of early onset. Similarly, the reported age for late-onset PD (LOPD) varies, but it often includes cases when the onset occurs after the age of 60. Cases of PD confirmed by pathology have been utilised in very few EOPD research studies. Similarly, the reported age for late-onset PD (LOPD) varies, but it often includes cases when the onset occurs after the age of 60. Cases of PD confirmed by pathology have been utilised in very few EOPD studies [11]. According to WHODAS-2, the productive cycle of life mainly explains data, except for cognition deficits, where EOPD persons showed worse impairments than LOPD ones. In employment-related activities, EOPD patients had more impairments than LOPD patients as per Table 1 [12]. In Figure 2, Availability includes parameters such as: MD Specialist, Investigations - Imaging & Genetics, Medical treatment, Surgical treatment, Social support, Rehabilitation and physiotherapy, and PD support groups. Affordability includes the lack of health insurance; some people cannot afford costly medicines, etc.

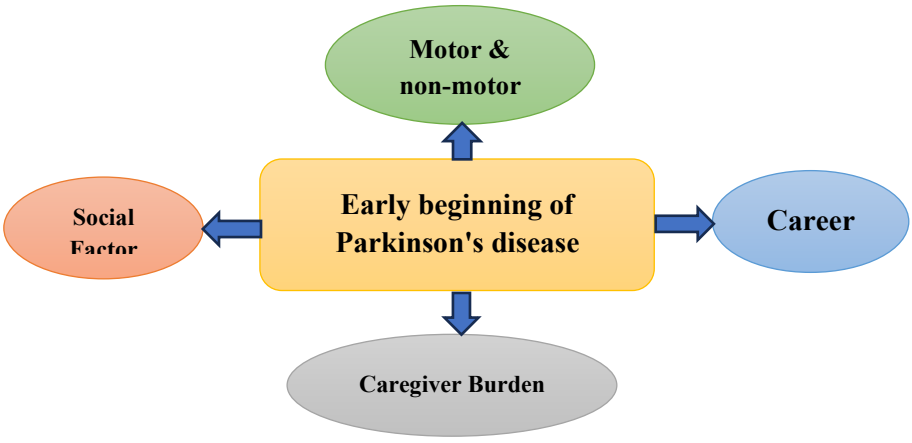


Figure 1: Early-stage Parkinson's diagnosis and management challenges.

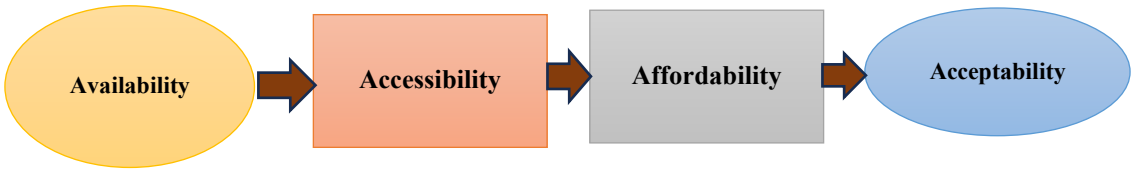


Figure 2: Critical elements involved in managing early-onset Parkinson's disease.

Table 1: There are mainly two types of symptoms, motor and non-motor [11].

Signs	Description	Experience
Motor (movement)	Motor impairments may lead to difficulties with coordination, movement, and balance.	Tremors, stiffness, slow movement, falls, dizziness, freezing, dystonia, muscle cramps, gait disruption, impaired handwriting, grip force, speech impairments, etc.
Non-Motor (non-movement)	A broad spectrum of symptoms falls under the category of non-motor manifestations.	Non-motor symptoms like depression, loss of smell, constipation, and RBD often emerge long before Parkinson's disease is diagnosed and can significantly affect daily functioning.

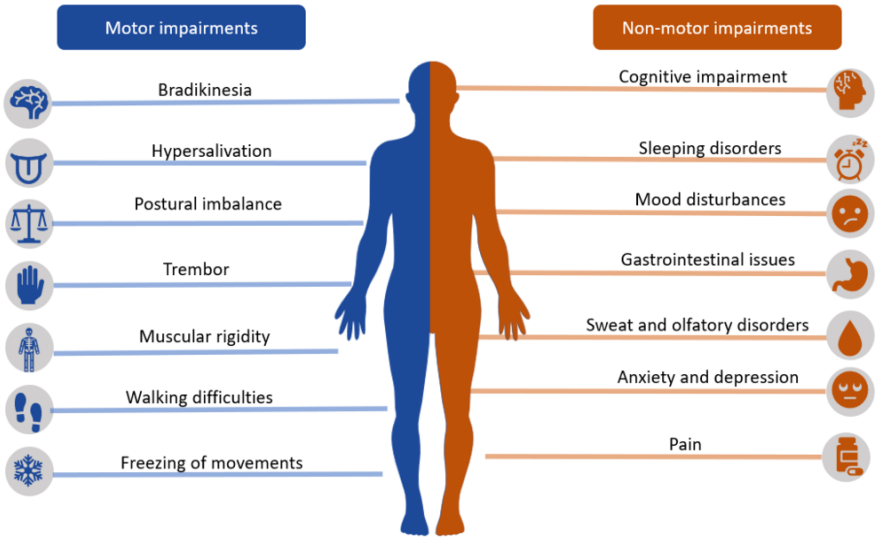


Figure 3: Showing the general issue with motor and non-motor impairments [12].

1.1 Brain alterations in PD

Degenerative basal ganglia substantia nigra neurons are characteristic of Parkinson's. Brain impulses to the basal ganglia cause muscles to contract. Basal ganglia neurons release

neurotransmitters that stimulate nerve cells. Degenerating basal ganglia nerve cells produce less dopamine and connect less. This hinders the basal ganglia's ability to direct muscle activity, leading to tremors, bradykinesia, hypokinesia, and difficulties with posture, walking, and coordination. Basal ganglia nerve cells are brain-deep. They assist with [13]:

- Start and refine the voluntary muscular actions you want to perform.
- Restrict uncontrollable motions
- Arrange for shifts in body positioning

The basal ganglia in the deep brain, which include the Caudate nucleus (C-shaped with a slender tail), Globus pallidus, subthalamus, Black Substantia, which start, smooth, and repress involuntary muscular movements and coordinate postural changes, are located near the putamen as per Figure 4 [13].

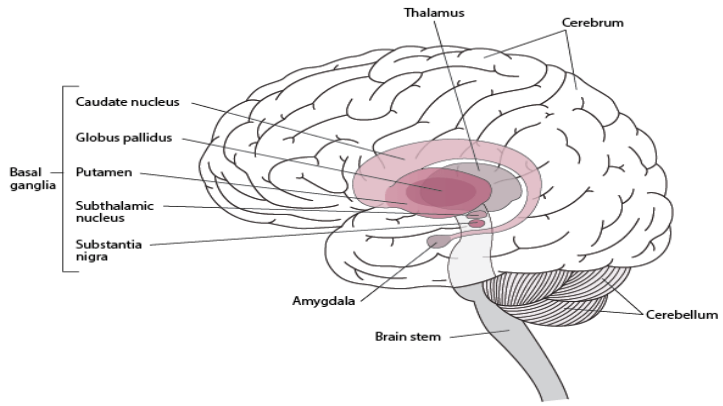


Figure 4: Showing the declining substantia nigra neurons, an important basal ganglia area.

1.2 Cause of PD [14]

Although the underlying cause of Parkinson's disease is not fully understood, evidence indicates that inherited and environmental factors contribute to its onset:

(i) Genetic causes: Parkinson's is rarely inherited. Only a few cases of inherited Parkinson's disease have been recorded recently. Many families have multiple cases of Parkinson's disease, suggesting a faulty gene has been handed down. However, this is highly unlikely. Scientists are investigating whether particular genes contribute to the hereditary causes of Parkinson's as they study how these genes affect risk, onset, and course. This research may lead to advances in medicine.

(ii) Root reasons in the environment: Some studies suggest that harmful chemicals, viruses, bacteria, and heavy metals may kill dopamine-producing neurons, causing Parkinson's disease. A huge population relates Parkinson's disease to herbicides and chemical exposure.

(iii) Combined causes: According to current scientific consensus, both genetic susceptibility and environmental factors contribute to the multifactorial origin of Parkinson's disease. From one individual to the next, the relative importance of each component differs.

Some kinds of Parkinson's disease occur in families and are linked to genetic differences. Parkinson's is assumed to be genetic, yet most instances do not run in families. Increasing evidence suggests that environmental factors, especially chemical exposure, interact with genetic propensity to develop PD. This research utilises advanced computational approaches to create a non-invasive voice analysis framework for Parkinson's disease detection and monitoring

1.3 Major Medical Treatments of PD

Treatments for Parkinson's disease can alleviate motor and non-motor symptoms to a considerable extent, but there are no disease-modifying medications available at this time. We are describing some of the latest medical treatments here.

1.3.1 Levodopa

Medications based on levodopa constitute the backbone of present-day PD treatment for the purpose of replenishing the dopamine in the damaged striatum. Standard treatment begins with a low levodopa dose and gradually increases it based on response and adverse effects. Levodopa, the metabolic precursor to dopamine, is an effective Parkinson's disease treatment. Since levodopa is inert, decarboxylation to dopamine generates its therapeutic and detrimental effects [15]. Most individuals need 150–1000 mg daily, divided into several doses. We'll learn later that dosing increases the risk of major side effects [16]. The clinical reaction to levodopa early in an illness can last several hours [16]. However, the drug's effects wane after shorter durations and require more frequent administration as the condition worsens. A portion of levodopa is metabolised into dopamine by DOPA decarboxylase outside the central nervous system, leading to various unwanted side effects; this is called peripheral conversion. The severity of these adverse effects can be reduced when levodopa is used with peripheral DOPA decarboxylase blockers [16]. Dyskinesias, hyperkinetic, twisted motions that the patient cannot control, often occur at the highest dose of the drug but can also occur as the drug wears off or during off-periods [17]. Balancing motor symptom control and adverse effects is difficult, but reducing levodopa dose may be used to address bothersome dyskinesias as they occur [18]. Future improvements in clinical efficacy and side-effect profiles may be possible if researchers continue to focus on developing additional long-acting oral formulations and other drug-delivery methods [19].

1.3.2 Dopamine agonists

They have a similar effect to dopamine in the brain. Some examples of these medications are ropinirole (Requip), pramipexole (Mirapex), and rotigotine (Neupro), where it wasn't until 1978 that dopamine receptor agonists became commercially available to treat Parkinson's disease [20]. Common dopamine agonists include ropinirole, pramipexole, and rotigotine. While these therapies can improve motor complications such as dyskinesia and fluctuations more effectively than levodopa, they may also increase the risk of adverse effects, including gastrointestinal symptoms, insomnia, oedema, dizziness, hallucinations, and fatigue [21,22].

1.3.3 Agents That Inhibit Monoamine Oxidase B

MAO-B inhibitors such as selegiline, rasagiline, and safinamide enhance dopamine availability by limiting its breakdown, contributing to improved symptom control in Parkinson’s disease. The advantage is usually minimal, and some people don't feel better after using them [23]. Most people use MAO-B inhibitors without problems, but stomach pain is common. Joint pain, depression, weariness, dry mouth, insomnia, lightheadedness, nausea, vomiting, hallucinations, disorientation, nightmares, and flu-like symptoms are further side effects [24].

1.3.4 Targeting Catechol-O-Methyltransferase (COMT)

Entacapone (Comtan), Tolcapone (Tasmar) and opicapone (Ongentys) can extend and improve levodopa's efficacy. Patients with motor abnormalities and "wearing off" at the end of levodopa are the main candidates for this medication. No benefit comes from using these medications alone [25]. Dyskinesia, hallucinations, confusion, nausea, diarrhoea, orange urine, and low blood pressure after standing up are the most common COMT inhibitor adverse effects. Tolcapone should be monitored for liver enzymes in the first year, but not entacapone, because

they may rise. Regular assessment of liver function is recommended during tolcapone therapy due to its potential hepatic effects. Less frequently reported adverse reactions include drowsiness, nausea, reduced appetite, vomiting, dizziness, orange discolouration of urine, hallucinations, headaches, respiratory difficulties, confusion, dry mouth, and abdominal discomfort.

1.3.5 Anticholinergics

Parkinson's disease patients under 70 without significant gait impairment can use anticholinergics for tremors. Dopamine agonists or levodopa can be used in combination with anticholinergics. Use trihexyphenidyl, benztropine, orphenadrine, procyclidine, and biperiden for Parkinson's. Many believe these drugs work similarly [26]. Common adverse reactions include vision problems, constipation, tiredness, incontinence, vertigo, impaired deglutition, uncontrolled body movements, and disorientation. Anticholinergic medications may be administered to reduce excessive drooling by decreasing saliva production [26].

1.3.6 Amantadine

Amantadine, an antiviral, may treat moderate tremors, akinesia, and stiffness in Parkinson's patients. Quick-release (Symmetrel) and extended-release (Gocovri) versions are available. Amantadine with levodopa may diminish dyskinesia in advanced Parkinson's patients. Extended-release Gocovri increases the risk of "wearing off." Two or three oral doses of immediate-release amantadine per day are acceptable; one extended-release dosage is advised [27].It starts low and is titrated up, like other anti-Parkinson medications. Syrup and tablets are available. Although generally well-tolerated, amantadine may induce adverse effects. These include hallucinations, poor focus, disorientation, swelling legs, impaired vision, nausea, vomiting, appetite loss, sweating, anxiety, and nightmares [21].

2. Literature Review

There is no proven standard for Parkinson's imaging or lab tests. However, DaTscan was approved by the FDA in 2011 [28]. This method details the brain's dopamine system. A small radioactive tracer is administered during a DaTscan to facilitate SPECT imaging, comparable to an MRI. The drug binds to brain dopaminergic receptors. A deficit of dopaminergic neurons in Parkinson's disease results from the loss of axons that produce dopamine. DaTscans can confirm or rule out Parkinson's mimics but not prove Parkinson's [28]. Effective machine learning methodologies are increasingly used for prompt diagnosis of Parkinson's disease as per Figure5.

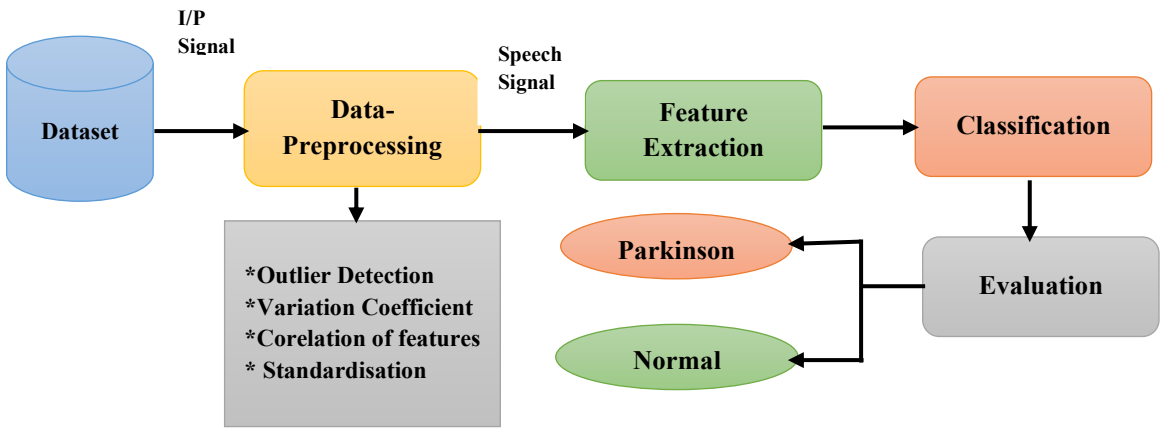


Figure 5: Procedure to detect Parkinson's disease using machine learning models.

2.1 Algorithm and techniques with related work

There is extensive research by scientists showing a rapid rise in the number of articles discussing PD diagnosis utilising DL and machine learning methods [29]. The study evaluated several classification approaches, including decision trees, regression trees, and artificial neural networks for recognising Parkinson's disease [30]. This study recommends employing supervised algorithms (SVM, K-NN, naïve Bayes, and ANN) for arbitrary-purpose disease diagnosis, where the recommended feature selection method combines a filter, wrapper, and classifier. In this research paper, a few clinical test findings are needed to diagnose PD; therefore, this technique may save time and money [31]. In this article, eight typical machine learning models were used to gradually incorporate PD-associated characteristics into risk estimations, balancing cost and accessibility for full assessment. Shapley Additive Explanations evaluated each factor's impact. Demographics, hospital entrance exams, clinical assessment, and polygenic risk score models beat invasive biomarkers. Penalised logistic regression and XGBoost outperformed eight PD risk models [32]. This article used machine learning to predict PD and assist doctors in prescribing tailored medications. This research study uses rough set theory to pick acceptable characteristics. Dimensionality reduction was achieved through Principal Component Analysis. At the same time, categorising performance was evaluated using deep neural networks, random forests, and support vector machine techniques [33]. Multiple machine learning techniques, including Logistic Regression and SVMs, as well as deep learning architectures such as AlexNet, GoogleNet, VGG-16, ResNet-50, and Inception-V3, were deployed for the identification of Parkinson's disease. Among these methods, AlexNet demonstrated superior performance [34]. Neural networks are the most widely used deep learning and machine learning approaches, according to this survey. This is due in large part to neural networks outperforming more standard machine learning approaches in terms of feature extraction accuracy and area under the receiver operating characteristic (ROC) curve (AUC). The findings indicate that handwriting and EEG data have been extensively used in the hunt for distinctive PD traits—the bulk of published studies aimed to test and categorise individuals as controls or PD patients. The categorical difficulty of Parkinson's disease staging, as well as the more practical and challenging problem of identifying PD symptoms and early biomarkers, are both of interest to physicians and scientists but have received very little attention in the literature [35]. This study addresses multicollinearity and simplifies the input feature space using principal component analysis (PCA) on the highlighted dataset. To evaluate PD prediction accuracy, the proposed DNN model is trained with a smaller input feature space and an adjusted parameter norm penalty (L2) to predict Motor and total UPDRS scores. Then execute a battery of tests and compare the results to previous approaches on the same dataset to evaluate the model. The fitness parameters MAE, RMSE, and R2 evaluate the model's prediction accuracy. The total UPDRS readings are 1.334, 2.221, and 0.956. The method improves the prediction of Motor and Total-UPDRS scores. The UPDRS score-predicting method is viable and useful in PD, according to this study [36]. Several artificial intelligence methods are used for diagnosing Parkinson's disease. These methods included Boosting, Innovative Support Vector Machines, Logistic Regression, K-Nearest Neighbours, Classification Trees, Feed-Forward Neural Networks and Random Forest. Furthermore, cognitive biomarkers like olfactory impairment and REM Sleep Behaviour Disorder (RBD) were analysed using deep learning techniques to differentiate between unaffected controls and those with Parkinson's disease. Deep learning automates the acquisition of linear and nonlinear representations from Parkinson's disease data, reducing the laborious process of feature extraction. However, multi-modal verification enhances illness-specific properties [37]. We trained industry-standard classifiers using voice and spiral imaging datasets. Last, ensemble-based sickness classification is applied. The final diagnosis depends on detecting vocal or spiral ensemble PD. AI models

such as SVMs, decision trees, and Naïve Bayes use speech and spiral test data. Accuracy, sensitivity, and specificity measure classifier model performance. Ensemble and SVM classify illnesses well on spiral imaging datasets. PD suspects benefit from early diagnosis and treatment [38]. There are various research papers with gaps that could be addressed by implementing effective, efficient techniques as per Table 2.

Table 2: Summary of CNN-Based Frameworks for Lung Nodule Detection and Delineation

Journal and dataset	**Method**	**Outcomes**	**Gaps**
IEEE Access, 2023 [39] NewHandPD [40]	The genetic algorithm K-Nearest Neighbour enhances feature extraction and promotes transfer learning.	Accuracy >95%, Precision = 98%, AUC area under curve = 0.90	Accuracy could be improved from 95% to reduce false positive results
Journal of Health care Engineering, 2022 [41] PD Dataset [42]	Support vector regression ensembles with PD data clustering can predict motor- and total-UPDRS. Comparing neuro fuzzy, multiple linear regression, and support vector regression prediction learning results.	MAE=0.5665, RMSE= 0.4186, R^2 = 0.9018, Computation Time (ms)= 363422	Need to focus on reducing computation time,
Computational Intelligence and Neuroscience & 2022 [43] Private Dataset	The Wavelet-cleft fuzzy algorithm aggregated features, and the signal-error-drop standardised approach pre-processed them. DBIN classifies, and the Firming Bacteria Foraging Algorithm finds aberrant traits.	Accuracy =99.8%	Need a larger dataset with additional samples from additional severity classes, patient-voice data, and features such as movement and handwriting.
Computing and Communication: Innovations & 2023, PD data from Istanbul University Cerrahpasa Medical Faculty Neurology Department [44].	Two common validation approaches are Stratified K-Fold Cross-Validation and early stopping regularisation.	Train Accuracy= 98.7%, Test Accuracy =97%.	Need advanced Various deep learning architectures, such as RNNs, CNNs, and related neural network frameworks.

Journal of Disability Research & 2024 [45] https://archive.ics.uci.edu/dataset/174/parkinsons [53]	ML Algo- SVM, KNN, Random Forest, Logistic Regression, AdaBoost	Accuracy: SVM=95, KNN=92, RF=95, AB=93, LR=86,	Only used the Speech dataset
Scientific Reports & 2024[46] Parkinson's data set, and are stored in the UCI repository for machine learning—the Parkinson's disease database [47,48].	L_1 SVM & DNN	Accuracy =100, specificity=100, sensitivity =100%. Dataset2, Accuracy= 96.87, sensitivity = 95.83, Specificity = 100.0.	PSP, MSA, CBS, and DLB were not distinguished in this study. Health facilities could employ model importance for real-time diagnosis.
Neural Computing and Applications & 2024 [49] EEG (Electroencephalography) dataset (version 1.0.4) [50]	Decomposing EEG data with the variable mode and empirical mode, creating subbands. Sub-band decomposed signal features were significant in chi-squared testing. Classification used SVM, FNN, KNN, and DT ML models, each with important features.	Accuracy =73.23/74.84, Sensitivity = 65.33/70, and Specificity = 80.63/79.38.	Less Dataset
Journal of Medical Internet Research 2021 [51] & Data from 726 US and overseas cohorts (262 PD and 464 controls).	Speech analysis added jitter, shimmer, and deep embeddings. Next, machine learning and Shapley adaptive explanations were used to assess how each attribute affects model output.	area under the curve of 0.753, SVM= 0.735, Random Forest= 0.745, Light GBM=0.72, XGBoost=0.74	Unbalanced Dataset, fewer cross-validation techniques
EU-wide iPrognosis app data from seven nations. The iPrognosis app received 29,048 II-A voice calls from 835	For 498 multilingual people (392/106 self-reported HC/PD patients), multiple- and single-instance language-aware learning classifiers	Best model 95% CI: English, German, Greek, and Portuguese cohorts had	It should be used to monitor daily activity during the day and night hours.

users, including HC and PD patients.	predicted speech and demographic data.	AUCs of 0.69, 0.63, 0.68, and 0.83, respectively.	

3. Methodology

Several machine learning models have been used to identify Parkinson's disease, achieving high accuracy and reducing false-positive detection. This study employed a diverse set of machine learning and deep learning models, including Logistic Regression, Support Vector Machine, K-Nearest Neighbours, Random Forest, Decision Tree, Gradient Boosting, XGBoost, Bagging, and Artificial Neural Networks.

Logistic Regression & SVM: Logit models are used in classification and predictive analytics. Since the result is a probability, the dependent variable can only be 0 or 1. Logistic regression transforms odds, i.e., the ratio of successes to failures. The logistic regression model relates the logit-transformed probability to the predictor variable. x. Parameter estimation is commonly performed using maximum likelihood estimation, while the logistic function represents the natural logarithm of the odds of the outcome [53]. Both types of problems are solved well using SVM machine learning. It helps regression and classification. SVM classification divides the dataset into two categories using the optimal hyperplane. Finally, the model can predict new noisy data issues as per Figure 6 [54,55].

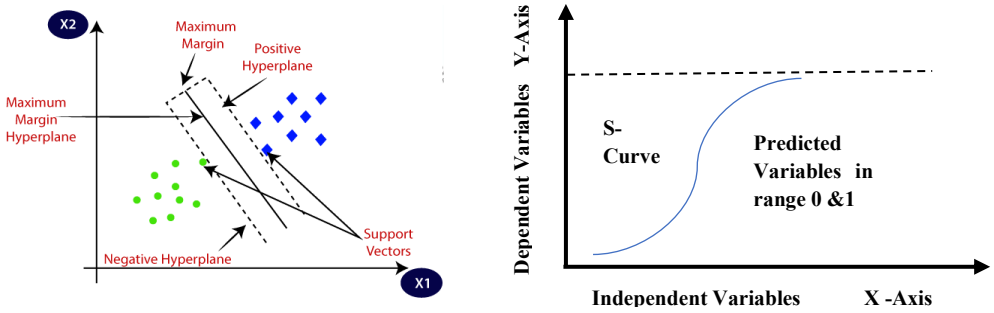
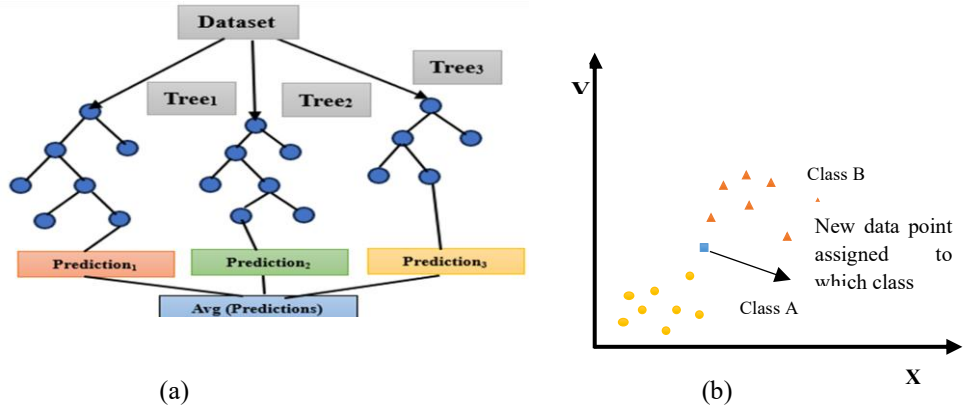


Figure 6: General structure of SVM and Logistic Regression models.

KNN and Random Forest:

KNN is a supervised machine learning algorithm that assumes that related parameters are close to each other to solve regression and classification problems. To detect Parkinson's disease, some modified KNN algorithms are also used to achieve high detection accuracy, such as the weight-based KNN algorithm and Fuzzy KNN [55]. This algorithm uses KNN: K is set to an initial value, and the geometric distance between new and training data points is calculated. The group is sorted in ascending order by Euclidean distance, with K-value labels. The standard distance calculation determines K, which classifies the new data point as its nearest neighbour. **Random Forest (**RF) employs the CART algorithm to grow many trees and average their predictions, an extension of bagging. Each tree's high variance should be reduced to improve diagnostic performance. Random and clustering methods, such as bootstrap, improve diagnostic predictions by strengthening weak methods as shown in Figure7 [56].



(a) (b)

Figure 7: General structure of KNN and Random Forest in figures (a) and (b).

Decision Tree, Gradient Boosting:A decision tree is a flowchart with nodes for attribute checks, branches for consequences, and leaves for class labels. The root is the highest node in the tree. Classification and regression use decision trees. DT is a popular machine-learning method due to its robust predictive performance, low computational cost, interpretability, robustness to noise, and tolerance for missing data. Handles redundant predicted attribute values [57].

Gradient Boosting enhances predictive accuracy by sequentially integrating weak learners and refining model weights based on previous prediction errors, thereby reducing errors and improving accuracy. Continuous target columns employ a Gradient Boosting Regressor, and classification issues use a Classifier. "Loss function" is the only difference. Gradient descent adds weak learners to reduce the loss function. Log-likelihood for classification and MSE for regression are possible loss functions, as they are used. We classify binary data via gradient boosting. We follow the Gradient Boosting Regressor's stages, but use a different loss function. When the target column was continuous, we used mean squared error. Now we use log-likelihood [58].

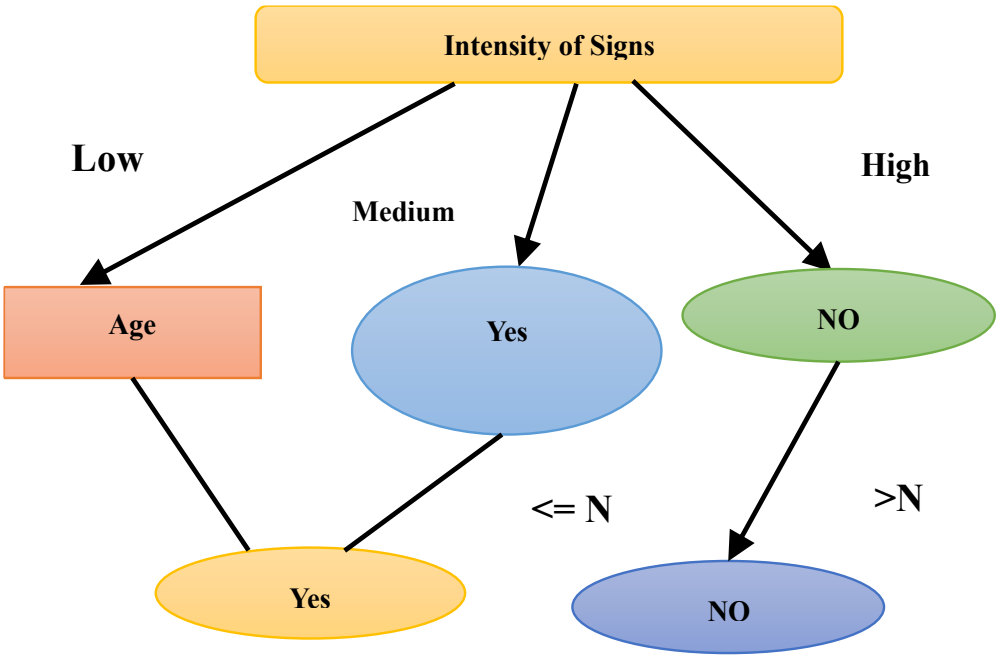


Figure8(a)

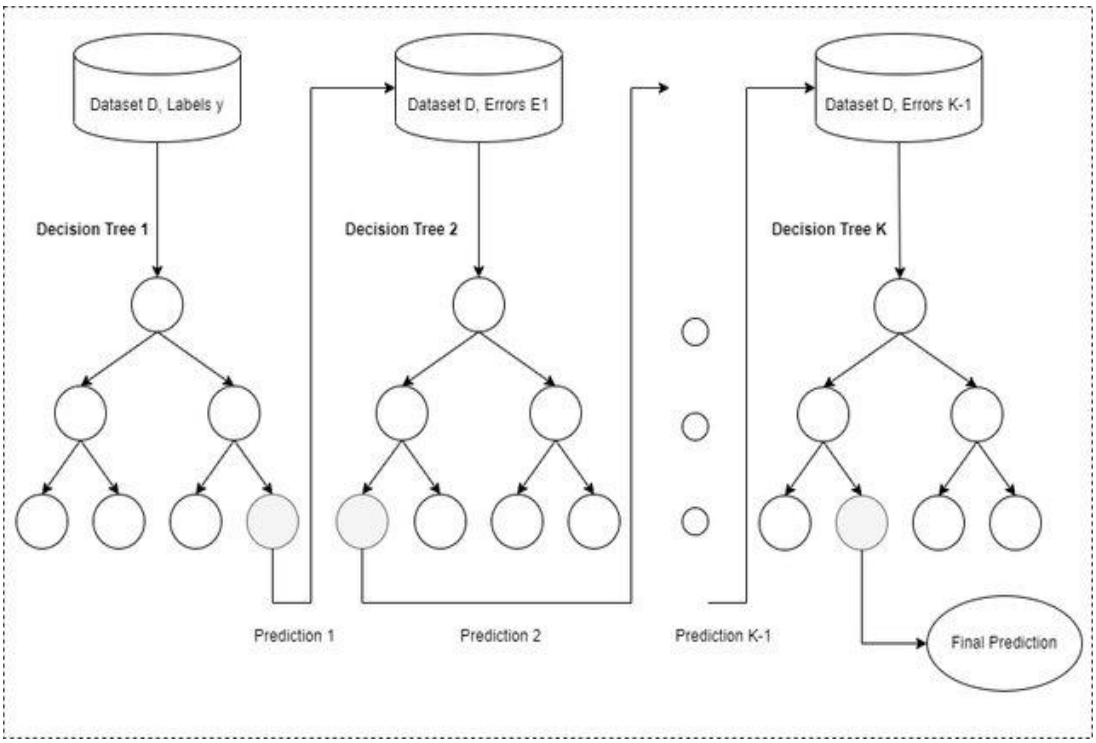
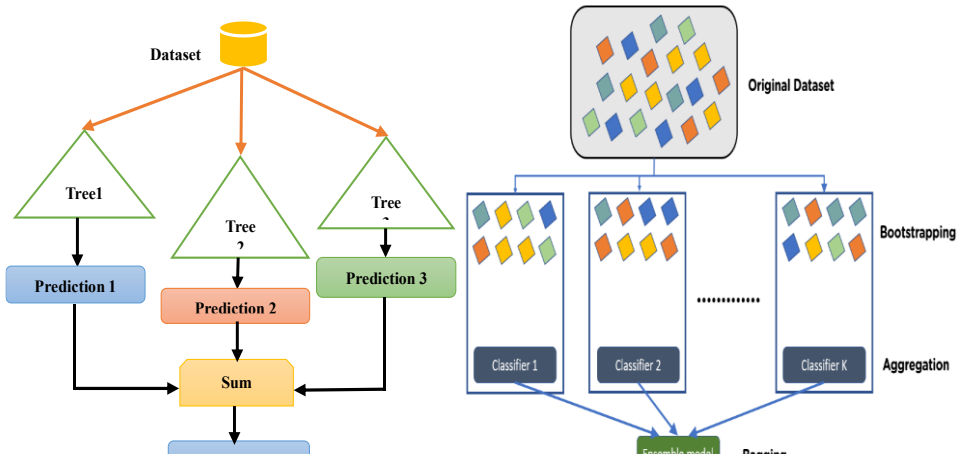


Figure8(b)

Figure 8: Figures (a) and (b) showing the general structure of the decision tree and gradient boosting.

XGBoost and Bagging: It uses gradient boosting for classification, regression, and ranking. Gradient boosting, in which each tree seeks to rectify the flaws of the previous one, is an effective strategy for generating decision tree ensembles. XGBoost has won many Kaggle competitions for its fast model training and high performance. It relies on boosting, which adds weak models to an ensemble and trains them to correct previous errors. To prevent overfitting, XGBoost has built-in regularisation algorithms, such as L1 and L2. Accelerating XGBoost training is achieved through distributed computing, which uses multiple processors on a single machine. By pruning decision trees, XGBoost improves generalisation and reduces model complexity. Cross-validation helps update hyperparameters and reduce overfitting in XGBoost. XGBoost stops training when validation set performance stops improving to save time and avoid overfitting [59].

Bagging, also known as bootstrap aggregation, is a basic yet effective ensemble method for unstable classification. Bagging ensembles use all algorithms. All algorithms construct unique data models, and the ensemble output mixes predictors' outputs. Changing algorithms or data may change models. Each method can be run on a different processor to speed up processing, since they run in parallel. Multicore CPUs are common on desktop PCs, making boosting less preferable. The ensemble's processing time is comparable to that of a single algorithm using this architecture. Just the decision function, which integrates all algorithm outputs, consumes time. One way to make learning algorithms more consistent is to bag them. Because missing values in high-dimensional data can increase variance, which in turn increases the likelihood of overfitting and hinders proper generalisation to new data sets, this is especially useful in such cases as per Figure 9[60].



(a) (b)

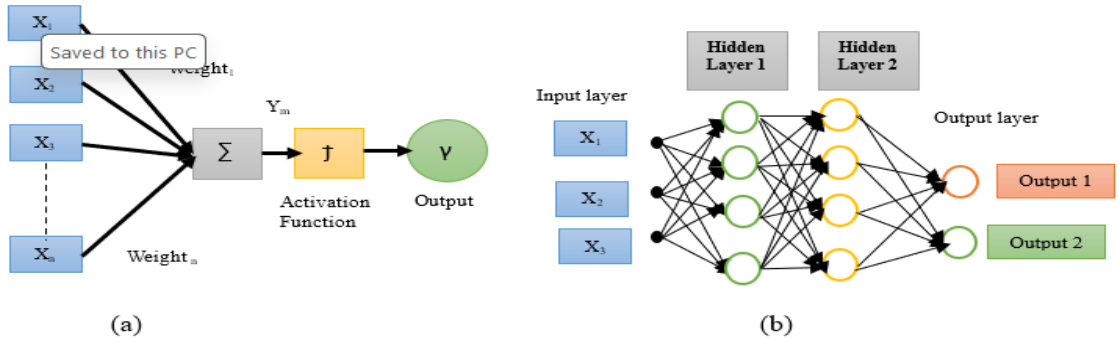
Figure 9: Figures (a) and (b) showing the general structure of the XGBoost model and the Bagging technique [48].

Artificial Neural Network and Neural Network:

In medical data processing, modelling, and interpretation, artificial neural networks are powerful. Medical applications of artificial neural networks primarily involve classification tasks. Computer scientists created artificial neural networks to imitate the human brain in parallel. ANNs have neurons, or nodes, arranged in one or more layers. In ANNs, information flows through weighted connections between neurons, facilitating the network's acquisition [60].

Neural Network: Neural networks use "neural units," computational building pieces, to solve problems like the brain does, using large clusters of neurons linked by axons. Each brain unit connects to many others. Without explicit programming, these systems learn and train themselves [61]. Machine learning models called neural networks imitate the human brain by simulating neuron interactions to notice events, evaluate alternatives, and form conclusions. Each node initiates and, in subsequent layers, connects to other nodes using its weight and threshold. This condition must be met to transfer data to the next network tier.

LSTM: A deep learning sequential neural network that preserves information is the Long Short-Term Memory Network. This version addresses vanishing gradients, a typical problem encountered by recurrent neural networks. Hochreiter and Schmidhuber developed LSTM to address problems with RNNs and ML techniques. The Keras module in Python implements LSTM models as per Figure 10 and 11[62].



(a) (b)

Figure 10: The overarching framework of the ANN and NN models in (a) and (b).

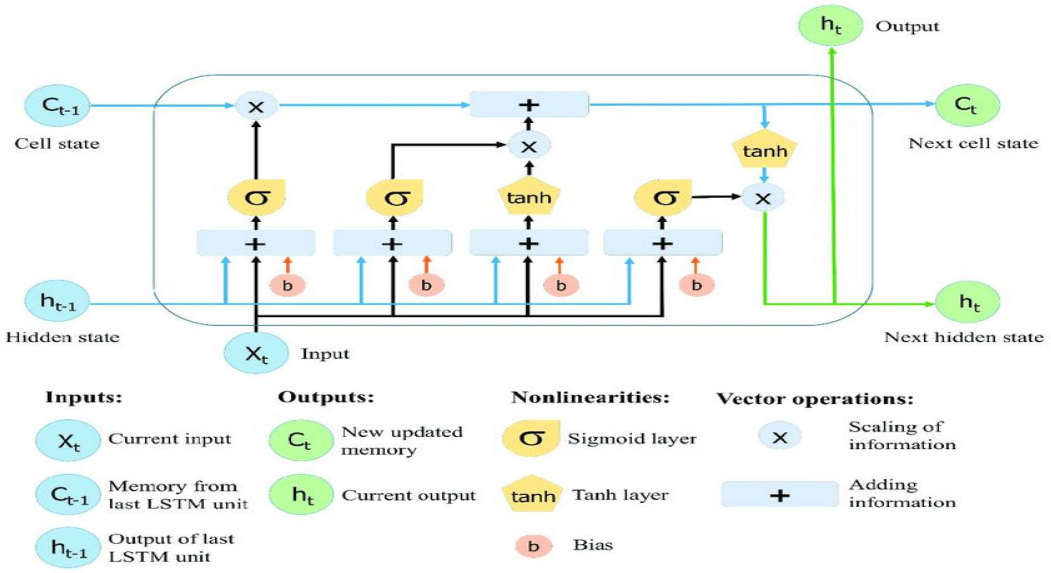


Figure 11: Fundamental framework of long-term short memory.

4. Evaluation Metrics

The Parkinson's dataset was utilised to assess the accuracy, recall, precision, and F1 scores of classification methods. These metrics evaluate classification models best. Equations show the statistical measures for successful classification and misclassification (FP and FN). Where terms TP-True Positive, TN-True Negative, FP-False Positive, FN-False Negative are used as per Table 3.

Table 3: Diverse evaluation metrics to assess [42].

Metrics	Formula
Accuracy	$\text{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)}$ (1)
Precision	$\text{Precision} = \frac{TP}{(TP+FP)}$ (2)
Recall	$TPR/Sensitivity/Recall = \frac{TP}{(TP+FN)}$ (3)
F1-Score	$\text{F1 Score} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision}+\text{Recall}}$ (4)

4.1 Public Dataset:

A total of 31 individuals, including 23 with a Parkinson's disease diagnosis, contributed 195 voice recordings to the collection. Whether the status variable reads "normal"(0) or "abnormal" (1) indicates the presence or absence of several voice-related characteristics in each record. Participants ranged in age from 46 to 85 years, with disease durations spanning 0–28 years. On average, each subject contributed 6 phonation recordings, each lasting 1-36 seconds [63]. Another dataset that includes spiral drawings to identify Parkinson's disease in patients at an early stage shows that patients with Parkinson's disease can be identified by having them draw a spiral and then monitoring their performance. There are two parts to the image dataset: the

training set and the testing set. Each contains the following elements: Training and testing form a spiral. Training and testing in a single wave as per Table 4 [64].

Table 4: Description of the used dataset.

Factures	Description [40]
mdvp_fo_hz	This is the human voice's base hertz frequency. With a tiny standard deviation, the data seem broad. No obvious outliers exist.
mdvp_fhi_hz	This feature provides the highest voice frequency. New data has a greater standard deviation and a more dispersed distribution. Voice recordings may differ more than thought.
mdvp_flo(hz)	This trait matches the lowest speaking frequency. Similar to the last feature, the data has a wide range and a high standard deviation.
mdvp_Jitter(%)	This feature measures the percentage change in fundamental frequency. Despite rare examples of increased volatility, the low standard deviation and mean suggest consistent jitter levels.
mdvp_Jitter(abs)	This feature tracks the % change in fundamental frequency. High volatility is rare, but the low standard deviation and mean indicate continuous jitter.
mdvp_rap and mdvp_ppq	These properties impact numerous jitter variables. As with prior jitter measures, they have low means and standard deviations.
Jitter:DDP	Jitter is linked to this trait. It matches earlier jitter measurements in standard deviation and means.
mdvp_Shimmer and mdvp_Shimmer (dB)	Jitter is linked to this trait. It matches earlier jitter measurements in both the mean and the standard deviation.
Shimmer_dda	This method also measures shimmer. Like other shimmer-related properties, the data distribution appears moderately spread out.
Nhr & hnr	The voice noise-to-tone ratio is NHR. Records have a stable noise-to-tonal balance due to low mean and standard deviation—the ratio of harmonics to noise. A larger mean and a lower standard deviation ensure a constant harmonic-to-noise ratio.
status	This trait may indicate health (1 for Parkinson's, 0 for healthy). The mean suggests a large fraction of the dataset has Parkinson's.
rpde and dfa	They appear to be numerical measures with moderate variability, as indicated by their standard deviations.

spread1 and spread2	Both features exhibit moderate variability, though spread2 has a higher mean than spread1.
d2& ppe	This lowers the dimension. Data appears moderately varied. This is related to nonlinear dynamical complexity. Moderate data variability.

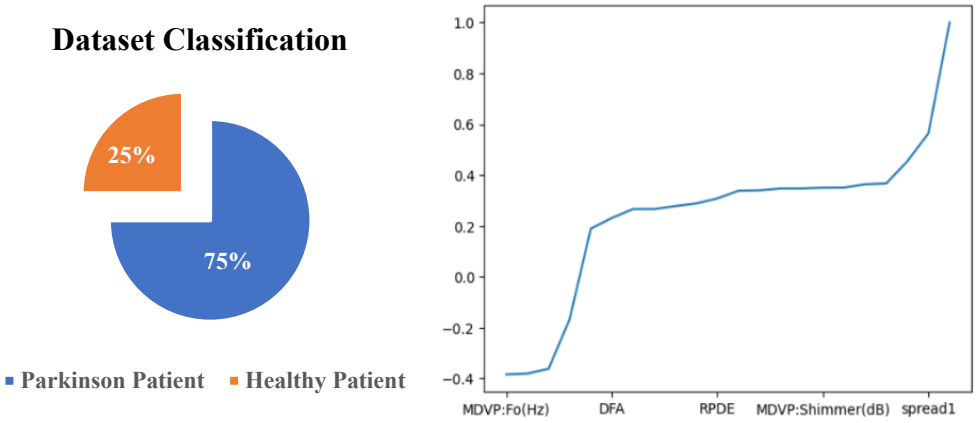


Figure 12: Factor Description of the dataset with co-linearity of features.

This dataset's units and feature values are adjustable because the independent variables are measured in Hz, dB, %, and absolute units. All other attributes are evenly distributed above the median. After identifying essential statistics (central values, spread, tails, correlations, etc.) for each feature using univariate and bivariate analysis. The dataset's Average voice fundamental frequency (mdvp_fo_hz) shows slight rightward skewness, with a mean of 154.2286 and a skewness of 0.5917. Maximum datapoints are 110-130 Hz. Most Parkinson's patients have a 90–190 Hz vocal fundamental frequency (mdvp_fo_hz). Despite this, some healthy patients' fundamental vocal frequencies are 110-130 or 170-180 Hz as per Figure 13 and 14[64].

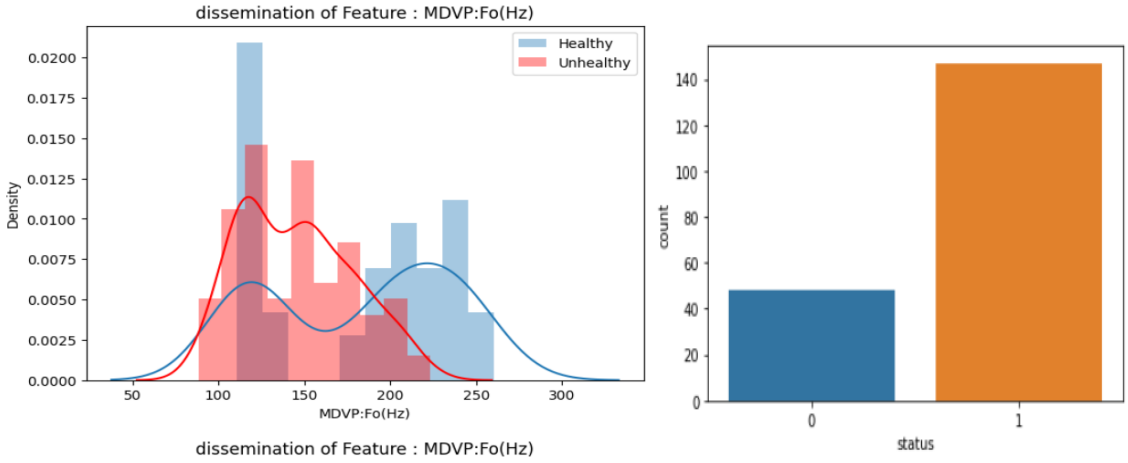


Figure 13: Correlation of features of the dataset and the pictorial representation of status values of the dataset.

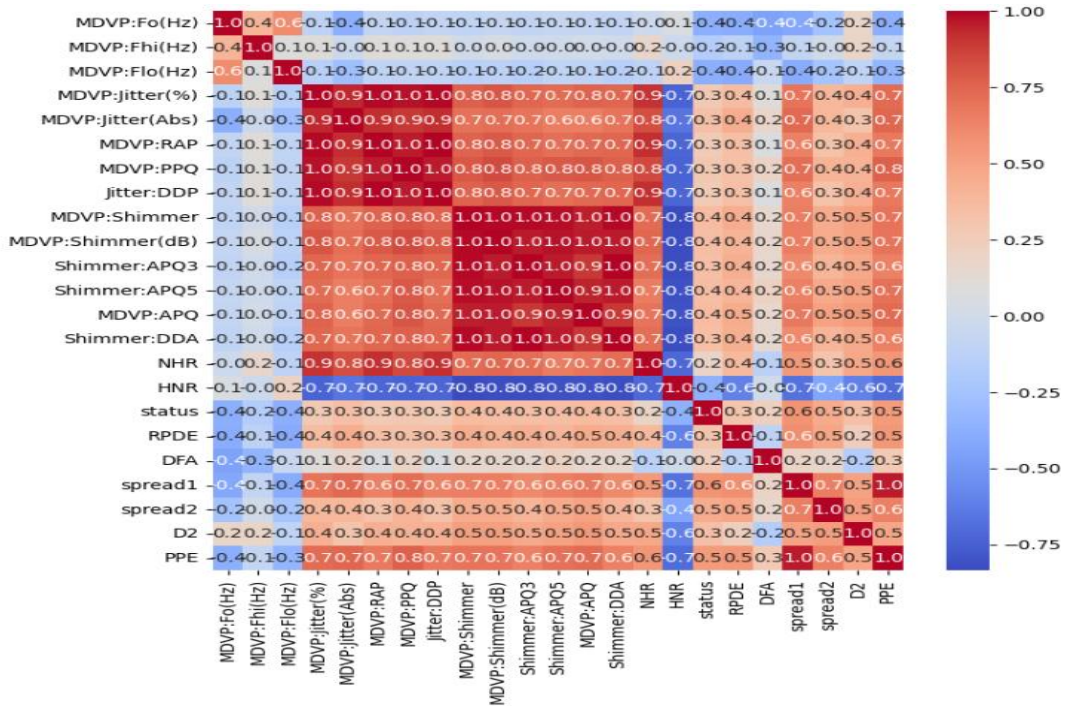


Figure 14: Heatmap of the features' relation in the dataset.

5. Implementation

Ten separate machine learning models, ensemble learning methods and deep learning models are used to evaluate the recognition accuracy. The following machine learning models are included: Random Forest, Logistic Regression, Decision Tree Support, AdaBoost (Bagging and Boosting), Vector Machines, ANN, KNN, Ensemble learning methods, Long short-term memory, and so on.

6. Environment Setting

The methodology incorporates both machine learning and deep learning models, as well as max pooling and batch normalisation, for efficient feature learning over 10–20 epochs. The suggested technique is trained and tested on this hardware: CPU: Intel i9-9900 K @ 3.60, 15.9 GB for Kaggle; 12.0 CUDA, GPU: NVIDIA-SMI. The network frameworks are written in Windows using PyTorch 2.0.0 and a Tesla T4 GPU (Colab).

6.1. Proposed Algorithm:

Algorithm: Input ◄— Dataset

Step 1: Analysis of the Dataset:

- Determine distributions, patterns, and relationships to reveal data trends, events, and relationships.
- Investigate the target's bivariate relationships in great detail.

Step 2: Preprocessing Steps:

- Eliminate extraneous elements and remove any unnecessary components
- Find and handle outliers and missing values.
- Data with categories should be encoded.

Step 3: Apply Models:

- Use Logistic Regression, SVM, decision trees, Random Forest, KNN, Adaboost, and XGBoost classification models.
- Make sure that people with Parkinson's disease are identified thoroughly and comprehensively by prioritising achieving a high recall for class 1.
- Apply network architecture, LSTM

Step 4: Analyse and Contrast the Models' Efficiency

- To assess the models' accuracy, efficacy, recall, F1-score and precision.

6.2 Outcomes

Categorical cross-entropy guided training, validation, and testing, while iterative parameter optimisation improved the accuracy and robustness of machine learning and deep learning models for Parkinson's disease detection as per Table 5

.

Table 5: Model's performance to detect Parkinson's disease in the dataset.

S.No.	Models	Accuracy %	Recall Score	Precision	F1-Score Score
1	Support Vector Machine	87	100	88.8	94.1
2	Logistic Regression	90	100	89.1	94.2
3	K Nearest Neighbour	95	100	94.1	96.9
4	Random Forest	95	96.8	93.9	95.3
5	Decision Tree	94	94	93.7	94
6	AdaBoostClassifier	91	96.5	92.5	96
7	XGBoost	97	97	97	95
8	Bagging	90	91	97	94
9	Neural Network/Artificial Neural Network	99.14 (Epochs-600),97.5 epochs-300			
10	Long-Short-Term Memory	97% Accuracy with a loss of 0.13 at epoch 50.			

The ANN model achieved 98% accuracy at 300 epochs, while the LSTM reached 97% accuracy at 50 epochs. In this experiment, the ANN model achieved the highest accuracy at 300 epochs, whereas ML with XGBoost achieved 97% for more refer to the below given figure, Figure15, 16, 17, 18, 19, 20 and 21.

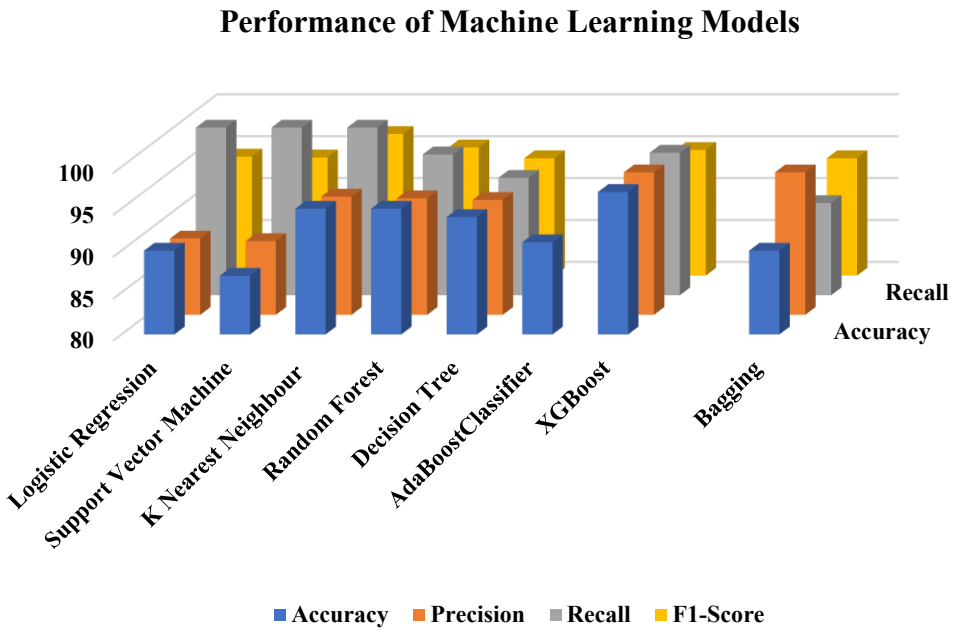


Figure 15: Assessment of machine learning model performance using key classification metrics.

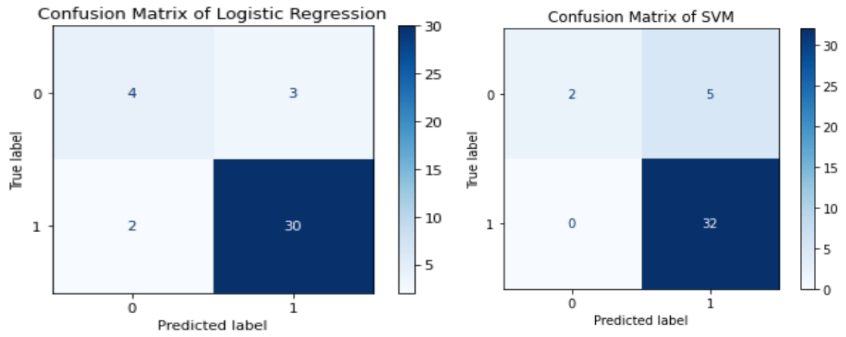


Figure 16: Confusion matrix of logistic regression and SVM.

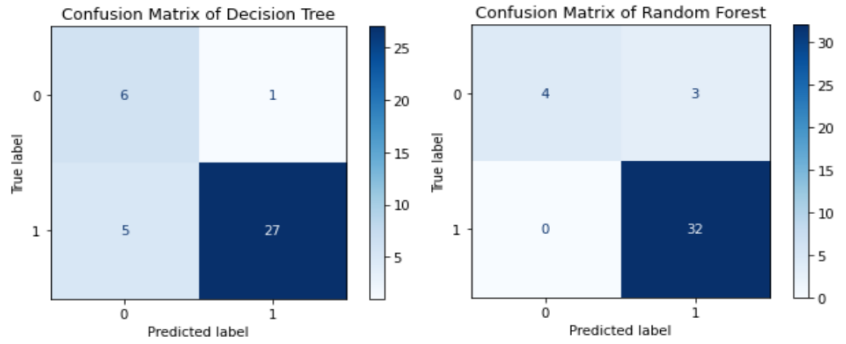


Figure 17: Confusion matrix of Decision Tree and Random Forest.

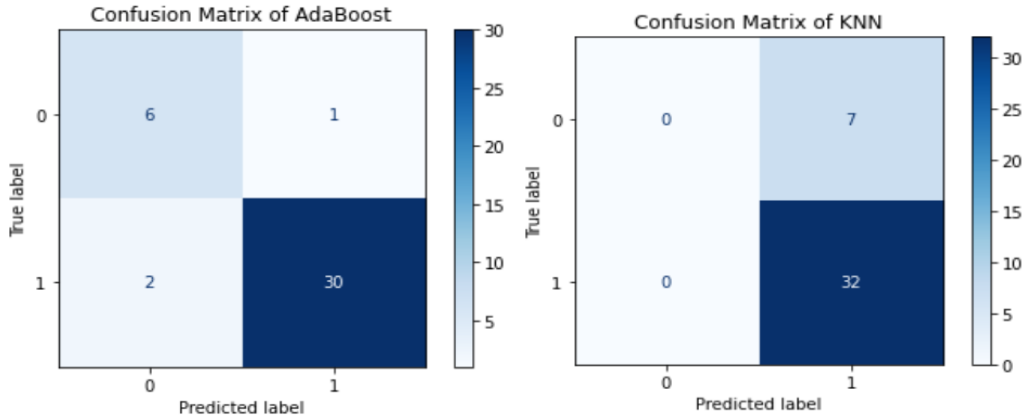


Figure 18: Confusion matrix of Adaboost and KNN Model.

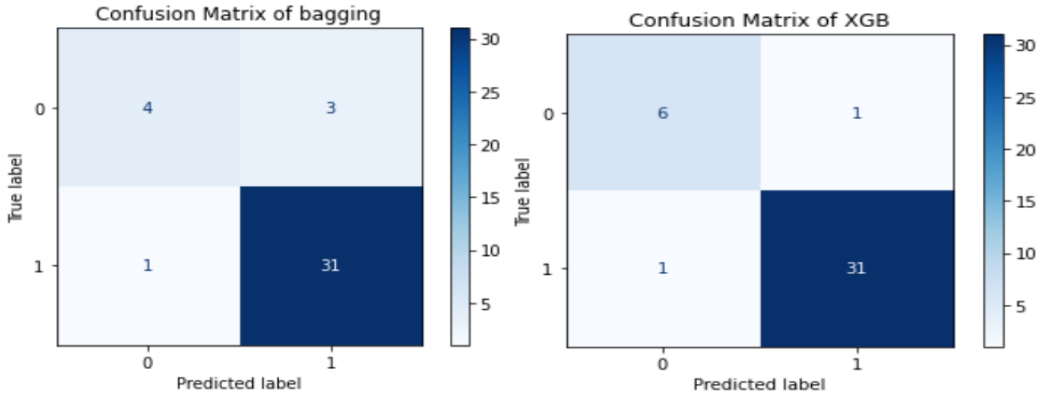


Figure 19: Confusion matrix of bagging and XGB Model.

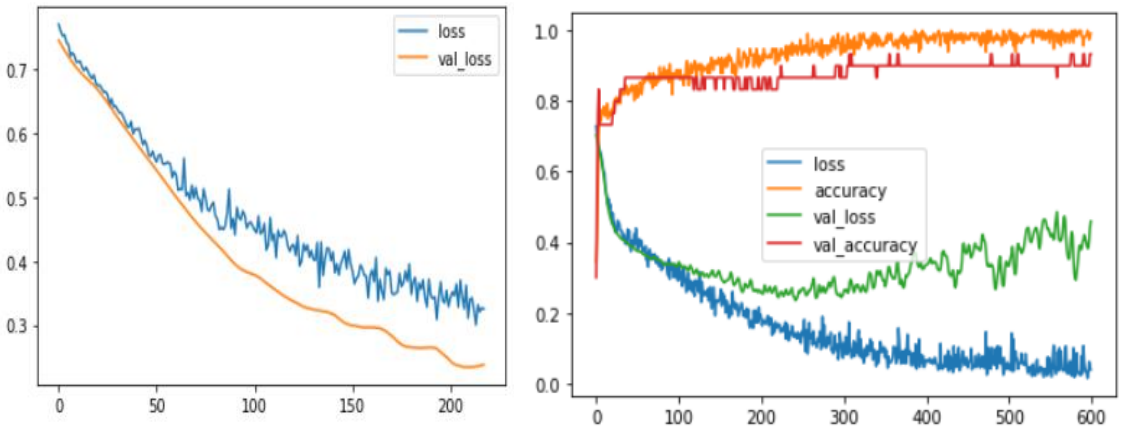


Figure 20: Performance of Neural Network at 600 epochs.

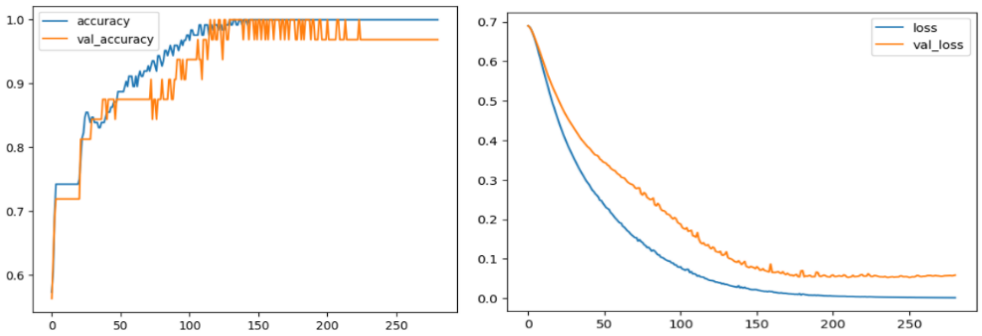


Figure 21: Performance of the artificial neural network at 300 epochs with accuracy and loss changes during the procedure.

7. Conclusion

This work introduces a data-driven framework that integrates machine learning and recurrent neural network models for Parkinson's disease identification from continuous speech recordings collected from patients. Using a real-world dataset, we trained and evaluated a model that could identify PD using only spoken language. The experimental framework was developed to evaluate and compare the performance of the proposed models. To accurately detect PD, the machine learning, data science, deep learning, and RNN communities are using this opportunity to help Parkinson's patients receive diagnoses and therapy faster. Successful models included Logistic Regression (90%), SVM (87%), Decision Tree (94%), Random Forest (95%), KNN (95%), Bagging (90%), and the XGBoost model, which achieved 97% accuracy. The ANN model achieved 98% accuracy at 300 epochs, and the LSTM model achieved 97% accuracy at 50 epochs. In this experiment, the ANN model achieved the highest accuracy at 300 epochs, whereas ML XGBoost achieved 97% accuracy, which could be improved further in the future.

8. Limitations and Future Work

The study is limited by class imbalance and reliance on a single voice-based dataset, which may affect model generalizability. Moreover, voice features alone may not fully represent the multifactorial nature of Parkinson's disease. Although the proposed models achieved high accuracy, further validation on larger, diverse, and multicentre datasets is necessary. Future work should incorporate multimodal biomarkers and external validation to improve robustness and clinical applicability. To further improve model performance, a variety of feature selection strategies can be integrated into models in the future, including keystroke features, to identify tremors and forecast the progression of Parkinson's disease. In addition, we intend to incorporate Natural Language Processing (NLP) and image-processing algorithms to identify a person's posture into our model.

References

[1] Thomas M, Lenka A, Kumar Pal P. "Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions". Mov Disord Clin Pract. 2017 Nov 1;4(6):806-818. doi: 10.1002/mdc3.12552. PMID: 30363367; PMCID: PMC6174397.

[2] Białek, K, Potulska-Chromik, A, Jakubowski, J, Nojszewska, M, Kostera-Pruszczyk, A. "Analysis of Handwriting for Recognition of Parkinson's Disease: Current State and New Study". *Electronics* 2024, *13*, 3962. https://doi.org/10.3390/electronics13193962.

[3] Prasad S, Rakesh K, Kamble N, Holla VV, Mailankody P, Lenka A, Naduthota RM, Stezin A, Mahale R, Yadav R, Pal PK, "Early onset of Parkinson's disease in India: Complicating the conundrum. Parkinsonism Relat Disord" 2022 Dec;105:111-113. doi: 10.1016/j.parkreldis.2022.11.017.

[4] Mehanna R, et al. "Age Cutoff for Early-Onset Parkinson's Disease: Recommendations from the International Parkinson and Movement Disorder Society Task Force on Early Onset Parkinson's Disease" Movement Disorders Clinical Practice. 2022 Oct;9(7):869-78.

[5]https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/diagnosis treatment/drc-20376062.

[6] Fox, Susan H., et al. Al., "Parkinson's Disease: Advanced Disease, Motor Complications, and Management", in David Burn (ed.), Oxford Textbook of Movement Disorders, Oxford Textbooks in Clinical Neurology,2013,

https://doi.org/10.1093/med/9780199609536.003.0010, accessed 5 Apr. 2024.

[7] Müller-Nedebock, A.C., Dekker, M.C.J., Farrer, M.J. et al. "Different pieces of the same puzzle: a multifaceted perspective on the complex biological basis of Parkinson's disease", npj Parkinsons Dis. 9, 110 (2023). https://doi.org/10.1038/s41531-023-00535-8.

[8] Simuni, T., and Rajesh, P., "Parkinson's Disease", New York, NY: Oxford University Press USA.Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. J. R. Stat. Soc. Ser. B (Methodol.). 36, 111–147,2009.

[9] Emamzadeh, Fatemeh N, and Andrei Surguchov, "Parkinson's Disease: Biomarkers, Treatment, and Risk Factors." Frontiers in neuroscience vol. 12 612. 30 Aug. 2018, doi:10.3389/fnins.2018.00612

[10] Knight, Emily et al. "The Role of Diet and Dietary Patterns in Parkinson's Disease." Nutrients vol. 14,21 4472. 25 Oct. 2022, doi:10.3390/nu14214472

[11] Ferguson LW, Rajput AH, Rajput A. "Early-onset vs. Late-onset Parkinson's disease: A Clinical-pathological Study", Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques. 2016;43(1):113-119. doi:10.1017/cjn. 2015.244

[12] Nascimento, Isaíra Almeida Pereira da Silva et al. "Comparison of disability level between Early and Late Onset Parkinson's Disease using WHODAS 2." Frontiers in Neurology, Vol. 14, 1281537. 15 Nov. 2023, doi:10.3389/fneur.2023.1281537

[13] Noyes, Eric et al. "Effect of Levodopa on Heat Hypersensitivity and Complex Motor Parkinsonism." Clinical neuropharmacology vol. 47,1 (2024): 29-32. doi:10.1097/WNF.0000000000000580.

[14] DeMaagd, George, and Ashok Philip. "Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis." P & T: a peer-reviewed journal for formulary management vol. 40,8 (2015): 504-32.

[15] Chauhan, S., Malik, N. & Vig, R., "UNet with ResNextify and IB modules for low-dose CT image denoising", International Journal of Information Technology. Volume. 16(7), 4677-4692, 2024. https://doi.org/10.1007/s41870-024-01898-8

[16] Nord M. Linköping, "Levodopa pharmacokinetics-from stomach to brain: A study on patients with Parkinson's disease", Linköping University Electronic Press; 2017. pp. 1–72.

[17] Fahn S, “The spectrum of levodopa-induced dyskinesias”, Ann Neurol. 2000 Apr;47(4 Suppl 1): S2-9; discussion S9-11. PMID: 10762127.

[18] Jankovic J, “Levodopa strengths and weaknesses”, Neurology. 2002 Feb 26;58(4 Suppl 1):S19-32. doi: 10.1212/wnl. 58. suppl_1.s19. PMID: 11909982.

[19] Murakami H, Shiraishi T, Umehara T, Omoto S, Iguchi Y, “Recent Advances in Drug Therapy for Parkinson's Disease”, Intern Med. 2023 Jan 1;62(1):33-42. doi: 10.2169/internalmedicine.8940-21. Epub 2022 Feb 1. PMID: 35110492; PMCID: PMC9876715.

[20] Deleu D, Northway MG, Hanssens Y, “Clinical pharmacokinetic and pharmacodynamic properties of drugs used in the treatment of Parkinson's disease”, Clin Pharmacokinet. 2002;41(4):261-309. doi: 10.2165/00003088-200241040-00003. PMID: 11978145.

[21] Davies N, Armstrong M, Pigott JS, Nimmons D, Read J, Gardner B, Maydon B, Schrag A, Walters K, “Co-Design of a Facilitated Self-Management Toolkit for People With Parkinson's Disease”, Health Expect. 2024 Dec;27(6):e70104. doi: 10.1111/hex. 70104. PMID: 39545646; PMCID: PMC11565443.

[22] Ceravolo R, Rossi C, Del Prete E, Bonuccelli U, “A review of adverse events linked to dopamine agonists in the treatment of Parkinson's disease”, Expert Opin Drug Saf. 2016;15(2):181-98. doi: 10.1517/14740338.2016.1130128. Epub 2016 Jan 20. PMID: 26646536.

[23] Bianchi MLE, Riboldazzi G, Mauri M, Versino M, “Efficacy of safinamide on non-motor symptoms in a cohort of patients affected by idiopathic Parkinson's disease”, Neurol Sci. 2019 Feb;40(2):275-279. doi: 10.1007/s10072-018-3628-3.

[24] Chauhan, S., Malik, N. & Vig, R., “Pulmonary Nodule Detection Using RBB-Based Optimised Yolov8x-C2fNetwork”, Recent Advances in Electrical & Electronic Engineering; Volume 18, Issue 10, e23520965340330, 2025.DOI: 10.2174/0123520965340330250207072742.

[25] Ahlskog JE, “Slowing Parkinson's disease progression: recent dopamine agonist trials. Neurology”, 2003 Feb 11;60(3):381-9. doi: 10.1212/01.wnl.0000044047.58984.2f. PMID: 12580184.

[26] Brocks DR, “Anticholinergic drugs used in Parkinson's disease: An overlooked class of drugs from a pharmacokinetic perspective”, J Pharm Pharm Sci. 1999 May-Aug;2(2):39-46. PMID: 10952768.

[27] Sawada H, Oeda T, Kuno S, Nomoto M, Yamamoto K, Yamamoto M, Hisanaga K, Kawamura T, “Amantadine Study Group. Amantadine for dyskinesias in Parkinson's disease: a randomized controlled trial”, PLoS One. 2010 Dec 31;5(12):e15298. doi: 10.1371/journal.pone.0015298. PMID: 21217832; PMCID: PMC3013111.

[28] Palermo, Giovanni et al. “Dopamine Transporter Imaging, Current Status of a Potential Biomarker: A Comprehensive Review.” International journal of molecular sciences vol. 22,20 11234. 18 Oct. 2021, doi:10.3390/ijms222011234

[29] Chauhan, S., Malik, N. & Vig, R., AI/ML techniques in servicing LDCT reconstruction: a systematic literature review. Discover Artificial Intelligence,5, 229, 2025.https://doi.org/10.1007/s44163-025-00419-1.

[30] Alshammri, Raya et al. “Machine learning approaches to identify Parkinson's disease using voice signal features.” Frontiers in artificial intelligence vol. 6 1084001. 28 Mar. 2023, doi:10.3389/frai. 2023.1084001

[31] Mei, Jie et al. “Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature.” Frontiers in ageing neuroscience vol. 13 633752. 6 May. 2021, doi:10.3389/fnagi.2021.633752

[32] Zhang, J.; Zhou, W.; Yu, H.; Wang, T.; Wang, X.; Liu, L.; Wen, Y, “Prediction of Parkinson’s Disease Using Machine Learning Methods”, Biomolecules 2023, 13, 1761. https://doi.org/10.3390/biom13121761.

[33] Begum, M.S., Balajee, A., Kulothungan, S. et al. “Parkinson’s disease prediction and drug personalization using machine learning techniques”, Soft Comput 27, 12669–12675 (2023). https://doi.org/10.1007/s00500-023-08744-2

[34] Tran, C., Shen, K., Liu, K. et al. “Deep learning predicts prevalent and incident Parkinson’s disease from UK Biobank fundus imaging”, Sci Rep 14, 3637,2024. https://doi.org/10.1038/s41598-024-54251-1

[35] Shaban, M. “Deep Learning for Parkinson’s Disease Diagnosis: A Short Survey”, Computers 2023, 12, 58. https://doi.org/10.3390/computers12030058.

[36] Shahid, Afzal Hussain, and Maheshwari Prasad Singh. “A deep learning approach for prediction of Parkinson's disease progression.” Biomedical engineering letters vol. 10,2 227-239. 16 Apr. 2020, doi:10.1007/s13534-020-00156-7.

[37] W. Wang, J. Lee, F. Harrou and Y. Sun, “Early Detection of Parkinson’s Disease Using Deep Learning and Machine Learning”, in IEEE Access, vol. 8, pp. 147635-147646, 2020, doi: 10.1109/ACCESS.2020.3016062

[38] Tapan Gandhi and Siddhartha Bhattacharyya and Sourav De et al., “Early detection of Parkinson's disease using data mining techniques from multimodal clinical data”, Advanced Machine Vision Paradigms for Medical Image Analysis, Academic Press, 20121.

[39] S. M. Abdullah et al., “Deep Transfer Learning Based Parkinson’s Disease Detection Using Optimized Feature Selection”, in IEEE Access, vol. 11, pp. 3511-3524, 2023, doi: 10.1109/ACCESS.2023.3233969.

[40] C. R. Pereira, S. A. Weber, C. Hook, G. H. Rosa, and J. P. Papa, ‘‘Deep learning-aided Parkinson’s disease diagnosis from handwritten dynamics,’’ in Proc. 29th Conf. Graph., Patterns Images (SIBGRAPI), Oct. 2016, pp. 340–346.

[41] Chauhan, S., Malik, N. & Vig, R., “Lung Nodule Segmentation Using Machine Learning and Deep Learning Techniques”, Data Analytics and Machine Learning: Navigating the Big Data Landscape. Singapore: Springer Nature Singapore, 289-316, 2024.DOI:10.1007/978-981-97-0448-4_14.

[42] https://archive.ics.uci.edu/ml/ datasets.

[43] R. Sarankumar, D. Vinod, K. Anitha, Gunaselvi Manohar, Karunanithi Senthamilselvi Vijayanand, Bhaskar Pant, Venkatesa Prabhu Sundramurthy, “Severity Prediction over Parkinson’s Disease Prediction by Using the Deep Brooke Inception Net Classifier”, Computational Intelligence and Neuroscience, vol. 2022, Article ID 7223197, 9 pages, 2022. https://doi.org/10.1155/2022/7223197

[44] C, S.., & Hemalatha, S., “A Deep Learning Technique to Clinch the Detection of Parkinson’s Disease using Speech and Voice Attributes”, International Journal on Recent and

Innovation Trends in Computing and Communication, 11(9), 55–61,2023. https://doi.org/10.17762/ijritcc.v11i9.8319

[45] Abdullah H. Al-Nefaie, Theyazn H. H. Aldhyani and Deepika Koundal, “Developing System-based Voice Features for Detecting Parkinson’s Disease Using Machine Learning Algorithms”, JDR. 2024. Vol. 3(1). DOI: 10.57197/JDR-2024-0001

[46] Ali, L., Javeed, A., Noor, A. et al. “Parkinson’s disease detection based on features refinement through L1 regularized SVM and deep neural network”, Sci Rep 14, 1333 (2024). https://doi.org/10.1038/s41598-024-51600-y

[47] Dheeru, D., & Karra Taniskidou, E, “UCI machine learning repository-Parkinsons data set”, http:// archive.ics.uci.edu/ ml (2017).

[48] Dheeru, D., & Karra Taniskidou, E, “UCI multiple voice recordings-Parkinsons data set”, https:// archive.ics.uci.edu/ml/datasets/Parkinson+ Speech+ Dataset+ with++ Multiple+ Types+ of+ Sound+ Recordings (2017).

[49] Latifoğlu, F., Penekli, S., Orhanbulucu, F. et al., “A novel approach for Parkinson’s disease detection using Vold-Kalman order filtering and machine learning algorithms”, Neural Comput & Applic (2024). https://doi.org/10.1007/s00521-024-09569-2.

[50] Rockhill AP, Jackson N, George J, Aron A, Swann NC, “UC San Diego resting state EEG data from patients with Parkinson’s disease”, 2020.https://openneuro.org/datasets/ds002778/versions/1.0.4

[51] Rahman W, Lee S, Islam MS, Antony VN, Ratnu H, Ali MR, Mamun AA, Wagner E, Jensen-Roberts S, Waddell E, Myers T, Pawlik M, Soto J, Coffey M, Sarkar A, Schneider R, Tarolli C, Lizarraga K, Adams J, Little MA, Dorsey ER, Hoque E “Detecting Parkinson Disease Using a Web-Based Speech Task: Observational Study”, J Med Internet Res 2021;23(10):e26305 doi: 10.2196/26305

[52] Laganas, Christos et al. “Parkinson's Disease Detection Based on Running Speech Data from Phone Calls.” IEEE transactions on bio-medical engineering vol. 69,5 (2022): 1573-1584. doi:10.1109/TBME.2021.3116935.

[53] Aditi Govindu and Sushila Palwe. “Early detection of Parkinson's disease using machine learning”, 1877-0509.2023.doi = {https://doi.org/10.1016/j.procs.2023.01.007.

[54]Swati Chauhan, Nidhi Malik et al, “ Lung Nodule Segmentation Using Machine Learning and Deep Learning Techniques”, Book Data Analytics and Machine Learning, March 2024.DOI: 10.1007/978-981-97-0448-4_14

[55] Indu, R., Dimri, S.C. & Malik, P. “A modified kNN algorithm to detect Parkinson’s disease”, Netw Model Anal Health Inform Bioinforma 12, 24 (2023). https://doi.org/10.1007/s13721-023-00420-7.

[56] Ahmed I.A., Senan E.M., Shatnawi H.S.A., Alkhraisha Z.M., Al-Azzam M.M.A, “Hybrid Techniques for the Diagnosis of Acute Lymphoblastic Leukemia Based on Fusion of CNN Features”, Diagnostics. 2023;13:1026. doi: 10.3390/diagnostics13061026.

[57] Charbuty B., Abdulazeez A. “Classification based on decision tree algorithm for machine learning”, J. Appl. Sci. Technol. Trends. 2, 20–28.2021. 10.38094/jastt20165

[58] Sultana, S., Reddy Madhavi, K., Lavanya, G., Swarna Latha, J., Sandhyarani, Prathyusha, B. “Detection of Malicious URLs Using Gradient Boosting Classifier”, In: Bhateja, V., Yang, XS., Ferreira, M.C., Sengar, S.S., Travieso-Gonzalez, C.M. (eds) Evolution in Computational

Intelligence. FICTA 2023. Smart Innovation, Systems and Technologies, vol 370. Springer, Singapore. https://doi.org/10.1007/978-981-99-6702-5_28

[59] E. F. S, E. S. T C and V. D. R S, “Prediction of Parkinson's disease using XGBoost”,8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 1769-1772, doi: 10.1109/ICACCS54159.2022.9785227.

[60] Chauhan, S., Malik, N. & Vig, R., “Experimental Comparisons of Deep Neural Network and Machine Learning Lung Cancer Detection Algorithms for CT Images”, In Innovative Computing and Communications. ICICC 2024, Singapore: Springer Nature Singapore, Lecture Notes in Networks and Systems, vol 1043,2024. https://doi.org/10.1007/978-981-97-4228-8_29.

[61] Shrivastava, G., Peng, S. L., Bansal, H., Sharma, K., & Sharma, M. eds., “New age analytics: Transforming the internet through machine learning, IoT, and trust modelling,” CRC Press, 2020.

[62] B. Nath Saha and A. Senapati, “Long Short Term Memory (LSTM) based Deep Learning for Sentiment Analysis of English and Spanish Data”, 2020 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2020, pp. 442-446, doi: 10.1109/ComPE49325.2020.9200054.

[63] https://archive.ics.uci.edu/dataset/174/parkinsons

[64] Zham Poonam, Kumar Dinesh K et al., “Distinguishing Different Stages of Parkinson’s Disease Using Composite Index of Speed and Pen-Pressure of Sketching a Spiral”, Frontiers in Neurology. 2017.DOI=10.3389/fneur. 2017.00435.