Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

Detection and Classification of Parkinson disease using various features extraction model and deep Learning Techniques for early stages

Ms. Tejashree Ladhe [1], Dr. Anita S. Walde[2]

[1]M.E. Student, Department of Electronics & Telecommunication, Deogiri Institute of Engineering and Management Studies Aurangabad (M.S), India

[2]Professor, Department of Electronics and Telecommunication Engineering, Deogiri Institute of Engineering and Management Studies Aurangabad (M.S), India

[1]**tejashreeladhe03@gmail.com**

Abstract: Parkinson's disease (PD) is a neurological movement disorder characterized by a slow, progressive worsening of symptoms, such as a tremor in one hand and a generalized sense of stiffness. More than 6 million individuals throughout the globe are afflicted. In the early stages of the illness, when symptoms are hard to identify, there is currently no convincing finding for this condition by non-specialist practitioners. To better understand patients, an RNN-based predictive analytics system is developed. The issue can be resolved with a small margin of error utilizing deep learning methods. If input data sets are to be used for analysis, they should be retrieved from the UCI Machine Learning repository. The goal of this research is to create an auditory feature-based method for detecting Parkinson's disease. Several machine learning methods are used to model the retrieved characteristics. In this study, we apply an RNN-based classification technique to identify PD patients' samples from those of healthy individuals. The RNN network is taught acoustic characteristics and a spectrogram of the speaker's voice. Only auditory characteristics are used in the training of RNN models. Using optimization strategies, Deep Learning (RNN) algorithms, and health care data, this research establishes whether or not people have Parkinson's disease. To test the efficiency and performance of the proposed method, a comparative research is conducted.

Keywords: Machine learning, Deep learning, Parkinson's disease

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

I Introduction

Dopamine deficiency leads to the neurodegenerative disorder known as Parkinson's disease. Common among the elderly, PD alters stride and posture, raising the risk of falls and limiting mobility. Therefore, it has an impact on regular activities and reduces the quality of life for both the sufferer and their loved ones. Parkinson's illness mostly affects motor functions. Symptoms of this movement condition include impaired mobility, slowness, increased muscular tone, and tremors even while at rest. A lack of facial expression, poor motor control, and a fluctuating tone of voice are further indicators. People with PD may suffer from sleep problems and a diminished sense of smell during the REM phase of the sleep cycle. Parkinson's disease is estimated to afflict around 1 percent of the population over 60. In most instances of Parkinson's disease, the underlying etiology is unclear. It has been shown that pathogenic alterations in dopaminergic neurons and the resulting petrochemical dysfunction are the most apparent features of this condition. The bulk of the brainstem's dopamine-producing neurons congregate in a black substance called the substantial Ingra. This region of the brain is intricately linked to other sub cortical regions and plays a role in the generation of typical motor behaviors. The substantial Ingra’s dopaminergic neurons can't produce enough of the neurotransmitter, limiting their range of motion and impacting their ability to move on their own will. As of now, Parkinson's disease cannot be cured.

II Literature Survey

In [1], Early and accurate diagnosis of PD is crucial for halting the illness's evolution and giving patients access to disease-modifying medication. The preemptor period in PD must be watched closely intended for this reason. To determine at an early stage whether or not a person is afflicted by PD based on preemptor traits, a novel deep-learning approach is developed. In this analysis, many indications, including R.E.M. in addition olfactory harm, cerebrospinal fluid data, and dopaminergic imaging markers, were investigated to detect PD in its earliest stages.

According to [2]early PD diagnosis with the use of machine learning methods to telemedicine. Thirty persons with chronic pain (PWP) and 30 people without chronic pain (healthy) have had their MDVP audio data studied while training 4 ML models. Random Forest classifier was shown to be the most effective Machine Learning (ML) methodology for PD identification when compared to other popular ML methods like Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Logistic Regression.

According to [3] the use of machine learning for Parkinson's disease diagnosis is described. Feature selection and classification procedures make up the suggested technique of diagnosis. When it came to selecting features, both the Feature Importance and Recursive Feature Elimination approaches were taken into account. Experiments were conducted using three different methods for classifying Parkinson's patients: classification and regression trees, artificial neural networks, and support vector machines.

According to [4] recent studies have shown that 90% of persons with PD also have difficulties with their voice. A hoarse, strained, or labored voice is the result of a vocal impairment or disorder. Multiple attempts at developing an automated PD diagnostic system based on a speech

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

dataset have been made. In this study, we compare the results of the ANN, KNN, and SVM classifiers on a speech database, analyzing their performance in terms of sensitivity, specificity, total classification accuracy, and geometric mean.

According to [5] recent studies on the early diagnosis of Parkinson's disease (PD) have connected voice problems to symptoms in 90% of PD patients. Thecontributes by using the most recent and comprehensive publicly accessible dataset for PD identification and improving accuracy via the use of less chosen voice cues. Wrappers feature subset selection reduces this public dataset's 754 features to between 8 and 20 that are useful for classification.

According [6] Neurodegenerative disease research is a very important field at the moment since early diagnosis may improve quality of life. There have been encouraging outcomes from recent advances in speech analysis methodology. Several deep learning and ML techniques for PD identification are investigated here. Our primary objective is to demonstrate the voice-signal analysis-based PD diagnosis. Speech processing has enormous promise in the categorization and diagnosis of PD for many years due to the fact that voice measures are non-invasive.

According to [7] various machine learning methods (such as Naive-Bays, K-NN, Decision Tree Classifier, and Gab Classifier) have been implemented. The data set used in this paper was split 70:30, with the larger portion going toward training and the smaller portion going toward testing. From there, hyper parameter tuning was performed to determine which hyper parameters would yield the most accurate models and the best results from the algorithms implemented.

According to [8] we compared the accuracy of several machine learning systems trained on speech characteristics for PD detection. In order to diagnose Parkinson's disease (PD), a range of machine learning algorithms were used, including XGBoost, LightGBM, Random Forest, Bagging, Gadabouts, the Decision Tree, Logistic Regression, Support Vector Machine, K-Nearest Neighbor, & Naive Bays.

According to [9] analysis of vocal problems was used to compare different machine learning (ML) methods for identifying PD. The SVM, KNN, and DT algorithms are all examples of ML techniques that were used in this study. Measures of accuracy, sensitivity, specificity, and area under the curve (AUC) show that the constructed model is more effective than previous methods.

According to [10] identify PD using data including voice recordings of patients who have already been diagnosed, a novel and potent strategy based on multi-level feature selection was developed. Chi-square and L1-Norm SVM (CLS) methods were used for feature selection at the first stage. The characteristics gathered by these algorithms were then blended to make the samples more representative. Finally, the Relief method was used to choose samples that were significantly distinguishable from the combined feature set by assigning weights to the features' relevance. Popular classifiers including KNN, SVM, and DT were utilized for machine learning throughout the classification phase, with KNN achieving the greatest result. The performance of the suggested method was further enhanced by using the Bayesian optimization process to choose the KNN classifier's hyper parameters.

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

III Proposed System Design

Fig.1 shows the process for developing a model to predict Parkinson's disease at an early stage employing ML and DL techniques. It comprises of the stages listed below

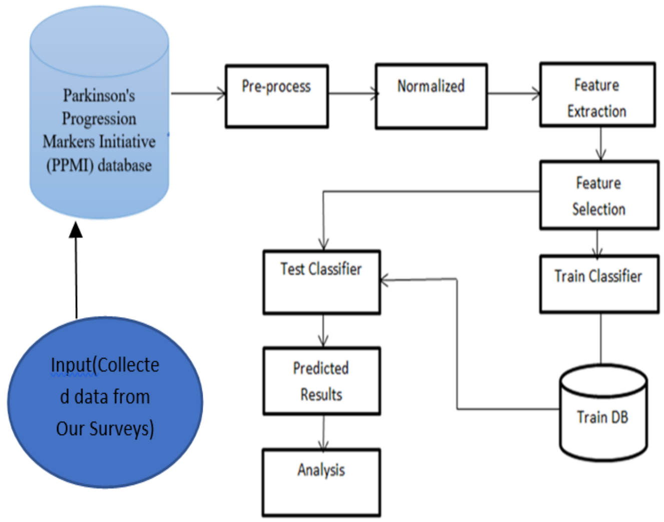


Figure 1: Proposed System Design

Implement Process

Importing datasets: Import data from repository and preprocess it. Normalized data can be use for training and testing using cross validation.

Pre-Processing: The data has taken from different sources, sometime it contains some invalid objects, and this phase can eliminate such values as well as records and gives the normalized dataset. Here we have current selected matrix of data and training set for classification, system generate the co-relation weight for selected features matrix and according to current it will classify with respective training label.

Build classifies: The RNN has used for module training as well as testing for achieve good accuracy.

Analysis: With the help of confusion matrix we carried out the system performance with different experiment analysis. The accuracy as well as time complexity will be the major factors has used for system performance analysis

Once complete execution ash done; we visualize the accuracy graphs of proposed system and show the effectiveness of proposed system**.**

IV Algorithm Design

Process for Training

Input: Train dataset [], activation functions AF [].

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

Output: Trained module in array list for entire spited dataset

Step 1: Start both algorithmsTrain_dataset [], AF [], epoch_size

Step 2: Extracted Features ← Extract Features (Train dataset [])

Step 3: Selected _Features [] ← optimized (Extracted_FeatureSetx)

Step 4: Train array list←Build Classification (Selecte_Features [])

Step 5: Return Array list from Training

Process for Testing

Input: User-defined threshold The, Test dataset [] as a testing case set or separate patient record, Train_arraylist as a repository of training-level background knowledge

Output: Classifier's recommended optimal instance is represented by the output Predicted_ class_label, Similarity weight...

Step 1: Use the following formula to decipher all test logs:

$$test\_Feature(m) = \sum_{t1=1}^{n} (.feature\_Set[A[i] \ldots\ldots A[n] \leftarrow \text{Test\_DB})$$

Step 2: Use the formula below to get the desired attribute features from the whole test record test Feature (t1).

Extracted_ Feature Setx [t……n] = $\sum_{x=1}^{n}(\text{t}) \leftarrow test\_Feature$ (t1)

Extracted_FeatureSetx [t] contains the feature vector of respective domain

Step 3: Use the below function to get every training instance out of the modules that have been trained.

$$train\_Feature(m) = \sum_{w=1}^{n} (.feature\_Set[A[i] \ldots\ldots A[n] \leftarrow \text{Train.pkl})$$

Step 4: Take test Feature (w) and use the following equation to extract each feature as a hot vector (or input neuron).

Extracted_Feature Set_Y [t……n] = $\sum_{x=1}^{n}(\text{t}) \leftarrow test\_Feature$ (w)

Extracted_FeatureSet_Y[t] includes a class-wide feature vector in its storage.

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

Step 5: At present, use all of the training elements to assess each test case.

$$calc\_weight = calcSim\ (\mathrm{Feature\_Set\_x} \mid\mid \sum_{i=1}^{n} \mathrm{Feature\_Set\_y[y]})$$

Step 6 : Returncalc weight

V Mathematical Model

We'll call everything in the system S= {I, P, D, and O}

I = Data entry for Parkinson's disease.

P = Process:

D = Dataset

Step1: The user is going to input the query.

Step2: Following the entry of the query, the operations that follow will be carried out.

Step3: The first processing of data.

Step4: The process of extracting features and selecting features.

Step5: Datasets for both training and testing.

Step6: Classification.

Step7: Classifier with final output that has been tuned and its performance indication.

O= Output (Predicted class label)

VI Results

The procedure for the implementation was finished off in an open-source Java environment. The hardware has a Java 3-tier analytics platform, a distributed Intel i5 CPU operating at 3.0 GHz, and 4 gigabytes of random access memory (RAM). Using the Parkinson's disease dataset, it was possible to assess whether or not Parkinson's disease should be classified as a disease. We have carried out experiment study on the use of deep learning in order to validate the results.

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

Table I. Comparison Table

Reno	Classifiers	Accuracy
1	SVM Classifiers	97%
2	Gadabouts Classifiers	95%
3	RF Classifiers	90%
4	RNN Classifiers	98%

Correlating quantitative evaluation and clinical ratings makes sense and may help with various difficulties in decision-making. Several machine learning models were used on the created training library to identify trends in frequent, dubious, and potentially harmful behaviors. Different cross-validation models were used with the behavior categorization training-database to evaluate and test machine learning approaches. The uniformity across all implementations can be traced back to the fold classification approach shown in Figure 3 below, which was applied to all parameters.

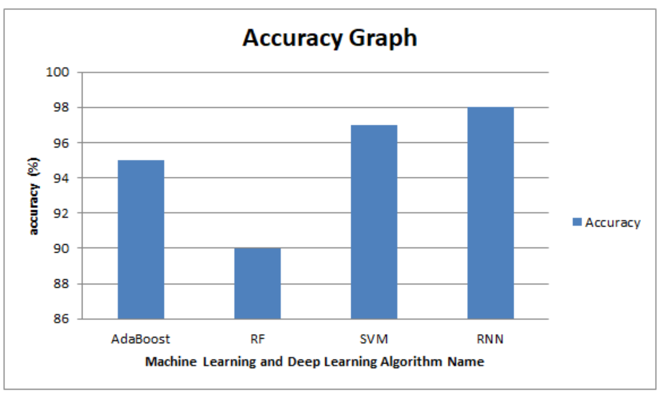


Figure 3 : Analysis ML and DL

Figure 3 displays the system's classification accuracy; utilizing the RF method, the system is around 90% accurate in identifying Parkinson's illness. The SVM algorithm used by the system has an accuracy of about 97% in identifying Parkinson's illness. The system's diagnostic accuracy for Parkinson's disease using the Adaboost algorithm is close to 95%. The proposed method employs the RNN algorithm, which allows for about 98% accurate detection of Parkinson's disease.

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

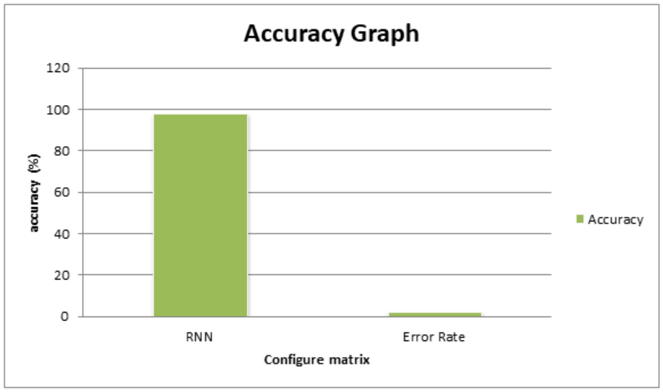


Figure 4: Accuracy of system analysis

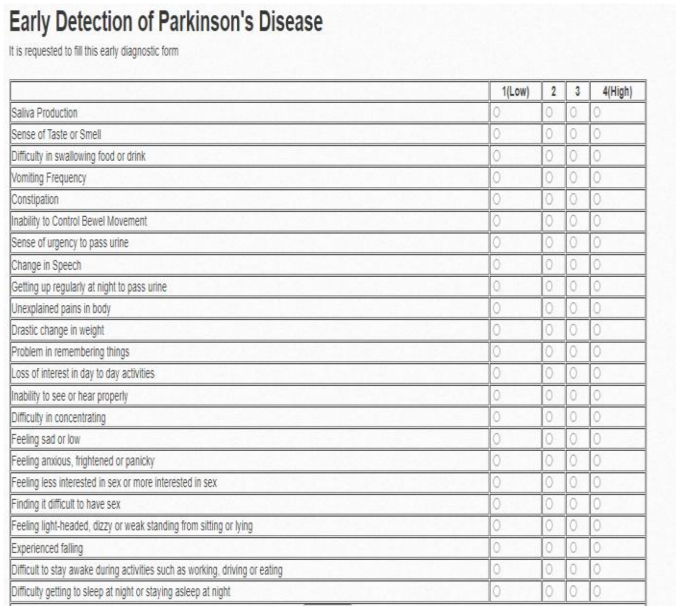
Classification accuracy for the proposed system and comparisons to numerous state-of-the-art systems are shown in Figure 4. Accuracy of deep learning classifiers for predicting Parkinson's illness is shown above. Up to 98% accuracy has been achieved when using the proposed classifier to diagnose PD.

Conclusion

Research into Parkinson's disease is more vital, and earlier diagnosis might improve quality of life for those affected. Recent advancements in technique for speech analysis have led to impressive findings. Our study used a machine learning approach to the problem of Parkinson's disease diagnosis, using a wide variety of machine learning models. The major goal of this project is to demonstrate how speech signals may be used to diagnose Parkinson's disease. Speech processing has long been thought to have enormous promise in the identification of Parkinson's disease since voice measures are non-invasive. The goal of this project is to evaluate and compare the performance of different categorization methods. On a speech dataset, the various classifiers were applied, and various assessment criteria were compared using visualisation and statistical analysis. In machine learning algorithms, it was discovered that the RNN beats all other classifiers. The accuracy of the feature selection technique is 98 % on all feature subsets, which is greater than all state-of-the-art techniques. The following may be suggested based on the findings.

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301



Early Detection of Parkinson's Disease

It is requested to fill this early diagnostic form

	1(Low)	2	3	4(High)
Saliva Production	○	○	○	○
Sense of Taste or Smell	○	○	○	○
Difficulty in swallowing food or drink	○	○	○	○
Vomiting Frequency	○	○	○	○
Constipation	○	○	○	○
Inability to Control Bewel Movement	○	○	○	○
Sense of urgency to pass urine	○	○	○	○
Change in Speech	○	○	○	○
Getting up regularly at night to pass urine	○	○	○	○
Unexplained pains in body	○	○	○	○
Drastic change in weight	○	○	○	○
Problem in remembering things	○	○	○	○
Loss of interest in day to day activities	○	○	○	○
Inability to see or hear properly	○	○	○	○
Difficulty in concentrating	○	○	○	○
Feeling sad or low	○	○	○	○
Feeling anxious, frightened or panicky	○	○	○	○
Feeling less interested in sex or more interested in sex	○	○	○	○
Finding it difficult to have sex	○	○	○	○
Feeling light-headed, dizzy or weak standing from sitting or lying	○	○	○	○
Experienced falling	○	○	○	○
Difficult to stay awake during activities such as working, driving or eating	○	○	○	○
Difficulty getting to sleep at night or staying asleep at night	○	○	○	○

Figure 2 : Data importing and reading

Our model was trained using a large number of classifiers, each of which was then validated and compared to achieve the highest level of accuracy. The evaluated results of each classifier will be sent to the user. When all classifiers have delivered their findings to the user, the user is able to make a decision on whether the Parkinson's disease detection data is "Yes" or "Not" by comparing the result with other results. When displaying each categorization result, graphs and tables will be used so that the information can be understood more easily. The assessment of the performance of suggested algorithms in comparison to the performance of current algorithm detection approaches is shown in Figure 3.

References

[1] Wang, Wu, et al. "Early detection of Parkinson’s disease using deep learning and machine learning." IEEE Access 8 (2020): 147635-147646.

[2] Govindu, Adyta, and Sushi Pale. "Early detection of Parkinson's disease using machine learning." Proscenia Computer Science 218 (2023): 249-261.

[3] Sentry, Zahra Karabiner. "Early diagnosis of Parkinson’s disease using machine learning algorithms." Medical hypotheses 138 (2020): 109603.

[4] Pahoa, Gunman, and T. N. Nagabhushan. "A comparative study of existing machine learning approaches for Parkinson's disease detection." IETE Journal of Research 67.1 (2021): 4-14.

[5] Solana-Lavelle, Gabriel, Juan-Carlos Galan Hernandez, and Roberto Rosas-Romero. "Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using

Received: 2026-05-01
Accepted: 2026-07-20
Published Online: 2026-07-30
DOI: 10.70454/IJMRE.60S102

International Journal of Multidisciplinary Research and Explorer (IJMRE)
E-ISSN: 2833-7298, P-ISSN: 2833-7301

classifiers and a small set of vocal features." Biocybernetics and Biomedical Engineering 40.1 (2020): 505-516.

[6] Nissan, Ire, Waste Ahmad Mir, and Tassel You Sheikh. "Machine learning approaches for detection and diagnosis of Parkinson's disease-a review." 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1. IEEE, 2021.

[7] Banal, Omit, Satyr Jet Raj Pail, and Suresh Sharma. "Early Parkinson disease detection using audio signal processing." Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2022, Volume 1. Singapore: Springer Nature Singapore, 2022. 243-250.

[8] Mammon, Muntasir, et al. "Vocal feature guided detection of Parkinson's disease using machine learning algorithms." 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2022.

[9] Outmode, Same, et al. "A Novel Approach for Parkinson's Disease Detection Based on Voice Classification and Features Selection Techniques." Int. J. Only. Eng. 17.10 (2021): 111.

[10] Demur, Faith, et al. "A simple and effective approach based on a multi-level feature selection for automated Parkinson's disease detection." Journal of Personalized Medicine 12.1 (2022): 55.