iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49497

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

A Review on Detection of Parkinsons Disease Using ML Algorithms

Sakshi M[1], Dr. Sankhya N Nayak[2], Skanda G N[3], Shreyas R Adiga[4], Pavana R[5]
[1, 3, 4, 5]*Research Scholar, Dept. of CS&E, JNNCE, Shimoga, India*
[2]*Associate Prof., Dept. of CS&E , JNNCE, Shimoga, India*

Abstract: ***Parkinson disease prediction is an area of active research in healthcare and machine learning. Even though Parkinson's disease is not well-known worldwide, its negative impacts are detrimental and should be seriously considered. Furthermore, because individuals are so immersed in their busy lives, they frequently disregard the early signs of this condition, which could worsen as it progresses. There are many techniques for Parkinson disease prediction. In this paper we are going to discuss some of the possible technical solutions proposed by researchers.***
Keywords: Parkinson disease

I. INTRODUCTION

The brain and nervous system are both affected by Parkinson's disease, which is a neurodegenerative condition. The loss of dopamine-producing neurons in the basic ganglia is specifically related to it. The illness has negative effects on people, society, and money on a social, professional, and personal level.

Individual symptoms that develop over time and vary from person to person can be divided into two categories. Motor symptoms include stiffness, slowness, also known as bradykinesia, facial expression, fewer swings of the arms, and resting tremor, whereas non-motor symptoms, which affect every system and component of the body, are unseen symptoms. These symptoms of autonomic dysfunction include perspiration, urination, and mood and thought disturbances.

The primary objective of the study is to evaluate the effectiveness of various Supervised Algorithms for enhancing Parkinson Disease detection diagnosis. Parkinson Disease was predicted using K-Nearest Neighbor, Logistic Regression, Decision Tree, Naive Bayes, and XGBoost. The detection of Parkinson's disease is based on the use of different classifiers, such as Accuracy, F1-score, Recall, Precision, R2-score Total UPDRS Motor UPDRS and Confusion matrix.

Amreen Khanum at el. [1] examined the effects of the various Supervised ML Algorithms for upgrading the diagnosis of Parkinson Disease. KNN, LR, DT, NB, and XGBoost were five machine learning techniques used to detect Parkinson's disease. The performance of the classifiers was assessed using precision, accuracy, F1-Score, and recall. Data on Parkinson's disease was obtained for this study from the UCI Machine Learning Repository. 23 speech feature sets are included in the 195 patient records that constituted this dataset. The first step was the extraction of characteristics from datasets related to Parkinson's disease. The study used five supervised learning algorithms to recognise Parkinson's illness. As a result, the performance metrics were evaluated to find the algorithm that outperformed.

Muhtasim Shafi Kader at el.[2] Mushtasim Shafi Kader at the el. [] chose 195 datasets related to Parkinson's disease from the UCI machine learning library in identifying the Parkinson disease. In the specified dataset, there were 24 attributes. After training the data, they were able to identify the machine learning algorithms that were most accurate. Naive Bayes, Adaptive Boosting, Bagging Classifier, Decision Tree Classifier, Random Forest Classifier, XBG Classifier, K Nearest Neighbor Classifier, Support Vector Classifier, and Gradient Boosting Classifier were the nine machine learning algorithms that were utilized to predict the illness. Evaluation metrics analysis and confusion metrics analysis (Precision, Recall, F measure and Accuracy) have been used to calculate the study's outcomes. Algorithm with highest accuracy was found using the above- mentioned metrics analysis.

Mohesh T et al.[3] The input is the Parkinson's disease voice dataset from the UCI device mastering library. Additionally, by combining the spiral drawing inputs of healthy individuals and Parkinson's patients, the gadget delivers accurate findings. It can be inferred that a hybrid approach accurately reads affected individuals' spiral drawings and voice data. This model aims to make this method of expertise a case of Parkinson’s hence, the goal is to apply numerous machines getting to know strategies like SVM, choice Tree, for buying the maximum accurate result.

Ifeoma Oduntan[4] implemented XGBoost (Extreme Gradient Boosted Algorithm) based on its accurate application . Precision and recall metrics were used in evaluating the performance of the classifier . Python 3.8 was used in the implementation of this project because of its flexibility, huge collection of libraries and it is an open-source language. The XGBoost algorithm was applied using different parameter tunings to find the best estimator. GridSearchCV was used, and scoring was set to f1. The best score and parameters were extracted, and this was used in fitting the model. Confusion matrix was used in evaluating the performance of the classifier.

Sonia Singla (2021)[5] Taken from the UCI repository. To determine the optimum algorithm for disease, start identification, algorithms like XGBoost, KNN, SVMs, and Random Forest Algorithm were tested. Confusion matrix and accuracy score were used to evaluate the models.

Anik Pramanik et al. (2020) [6] proposed a model to detect Parkinson's Disease using voice and speech signal data in an efficient and robust manner. The dataset, on which the proposed model was tested, was publicly available as PD Speech data-set from Department of Neurology in Cerrahpa¸sa, Faculty of Medicine, Istanbul University. The dataset contained no less than 750 attributes of 252 persons. Data standardization, Multicollinearity Diagnosis, Principal Component Analysis, and Independent Component Analysis were among the different data-processing methods used. Various algorithms used includes Support Vector Machine, Logistic Regression and Random Forrest.

Anitha R at el. [7] suggested a predictive analytics system that uses K-means clustering and Decision Tree to extract insights from patients. This specific study uses the Parkinson's disease speech dataset from the UCI Machine Learning library as its input. The suggested approach also delivered precise outcomes by combining the spiral drawing inputs of Parkinson's patients and healthy individuals. They suggested a hybrid methodology, which produced technique that detects after evaluating patient speech and spiral drawing data. The drawings were converted into pixels using the Random Forest classification technique, and the extracted values were compared to the training database to generate various characteristics.

Sanghee Moon et al.[8] This retrospective database study includes a total of 1468 people tested at the Parkinson's Disease and Movement Disorder Clinic of the University of Kansas Medical Center. A total of 130 balance and gait features were automatically computed by the Mobility Lab software. Neural networks, Support Vector Machine, kNeareast Neighbors, decision trees, random forests, gradient boosting classifiers, and linear regression were among the classification models. Accuracy, recall, precision, and F1 score were all used to evaluate the classification models.

Timothy J. Wroge et al. [9] Data were gathered through mPower, clinical observational research carried out by Sage Bionetworks with the use of an iPhone app. Before being fed into the feature extraction algorithms, the raw audio is cleaned with VoiceBox's Voice Activation Detection (VAD) method. Scikit-Learn was used to create the decision tree and support vector machine classifiers. Several decision tree classifiers, including additional trees, random forests, gradient boosted decision trees, and normal decision trees, were utilized to categorize the dataset.

Basil K Varghese at el. [10] used the UCI ML repository to access the dataset. The dimension of the data was then reduced using Principal Component Analysis. They used SVM (Support Vector Machine), Decision Trees, Linear Regression, and Neural Networks to predict values from the test dataset. The accuracy was then determined by using the training model to predict values from the test dataset. The dataset used in this study consists of characteristics extracted from voice recordings of 42 individuals who have been diagnosed with early-stage Parkinson's disease. The goal of this work was to use a variety of machine learning techniques to reliably predict the RMSE (Root Mean Square Error) values of motor and total UPDRS scores (referred to as "motor UPDRS" and "total UPDRS").

Srishti Grover at al. [11] Proposed a way for applying deep learning to forecast the severity of Parkinson's illness. In the first stage, voice recordings from PD patients are obtained for analysis. The obtained data is then normalized using min-max normalization. The model after getting data performs training, evaluation, and prediction. A deep neural network containing an input layer, hidden layers, and an output layer is created. And in the end, an evaluation is performed on the resultant DNN classifier.

Tarigoppula V. S Sriram at el. [12] proposed research on Diagnosis of the Parkinson disease through machine learning approaches. Orange v2.0b and weka v3.4.10 was used in the experiment for the statistical analysis, classification, Evaluation, and unsupervised learning methods. A voice dataset for Parkinson's disease was collected from the Center of Machine Learning and Intelligent Systems at UC Irvine. Of the 31 persons in the collection who had biological voice measurements, 23 of them had the disease. The dataset was acquired and utilized for data visualization (parallel coordinates, Sieve graphs, and SOM), classification (Bayes Net, Nave Bayes, Logistic, Simple Logistic, K Star, AD Tree, J48, LMT, and Random Forest), as well as for evaluation and unsupervised learning techniques (Hierarchal clustering).

II. COMPARISON AMONG MODELS

We compared the work based on metrics, algorithms used and the accuracy on several datasets used by the authors. The work is summarized as shown in Table 1.

Table-1: Comparison Among Models

Reference	Dataset	Tools used	Machine Learning Algorithms Used	Evaluation Metrics	Outcomes of Outperformed Algorithms
Amreen Khanum D et al. [1] 2022	UCI Machine Learning Repository	---	Decision Tree Classifier, Logistic Regression , Naive Bayes, KNN Classifier and XGBoost Classifier	Recall, Precision, F1-Score , R2-Score and accuracy	KNN reached to the highest accuracy: Accuracy=0.966102 F1-Score=0.960000 Recall= 0.923077 Precision=1.00000 R2-Score=0.862471
Mustasim Shafi Kader et al. [2] 2022	195 Parkinson's disease datasets took from the UCI machine learning repository.	Python and the Scikit-learn module.	Naive Bayes, Adaptive Boosting, Bagging Classifier, Decision Tree Classifier, Random Forest classifier, XBG Classifier, K Nearest Neighbor Classifier, Support Vector Classifier and Gradient Boosting Classifier	Confusion metrics analysis and Evaluation metrics analysis using Precision, Recall , F measure and Accuracy	A]Confusion metrics analysis: K Nearest Neighbor with 97% accuracy. Predictive Positive=6,0 Predictive Negative=1,32 B]Evaluation metrics Analysis : K Nearest Neighbor with 97% accuracy Precision=1.00 Recall=0.86 F1-Score=0.92 TP=1.00 TN=0.97
Mohesh T et al. [3] 2022	UCI Parkinson Dataset.	Python 3	Logistic Regression, Support Vector Machine (SVM) , Ada Boost , Gradient Boost, Random Forest, Naive Bayes, Neural Network , XGBoost and Decision Tree	PCA and EDA (TSNE visualization) with accuracy	Decision Tree with an accuracy of 94.7774555%.
Ifeoma Oduntan et al. [4] 2021	Vocal measurement of 195 instances and 24 attributes from 31 people and 23 of them have Parkinson's Disease.	Python 3.8	XGBoost	Accuracy, Precision and Recall	Accuracy=95% Precision=1.00 Recall=0.94
Sonia Singla [5] 2021	UCI Parkinson Dataset	Python 3	XGBoost, KNN, SVMs, and Random Forest Algorithm	Confusion matrix and accuracy score	XGBoost with an accuracy of 94% Matrix: True healthy:[2 5] True Parkinsons:[0 32]
Anik Pramanik et al. [6] 2020	PD Speech data-set from Department of Neurology in Cerrahpa sa, Faculty of Medicine, Istanbul University	----	SVM ,Logistic Regression ,KNN , Ada Boost and Random Forest Algorithm	PCA and ICA for data pre-processing and accuracy	SMV showed the highest overall performance with accuracy 94.1%.
Anitha R et al. [7] 2020	UCI Machine learning library	RStudio and Visual Studio Code	K-Means ,Random Forest and Decision Tree	A]Voice Data Analysis :Accuracy B]Spiral Drawing Analysis : Accuracy and Confusion matrix	A]Voice data Analysis: Accuracy of 88% (Clustering and classification). B]Spiral Drawing Analysis: Accuracy of 83% (Random Forest).
Sanghee Moon et al. [8] 2020	Parkinson's Disease and Movement Disorder Clinic of the University of Kansas Medical Center.	IMU sensors	Neural network(NN), Support vector machines , k-nearest neighbor , Decision tree , Random forest , Gradient boosting and Logistic regression	Accuracy, precision , recall and F1-score.	The F1-score of NN was 0.61,Precision of 0.61 ,Recall of 0.61 and accuracy of 89% ,showed the highest performance among 8 models.

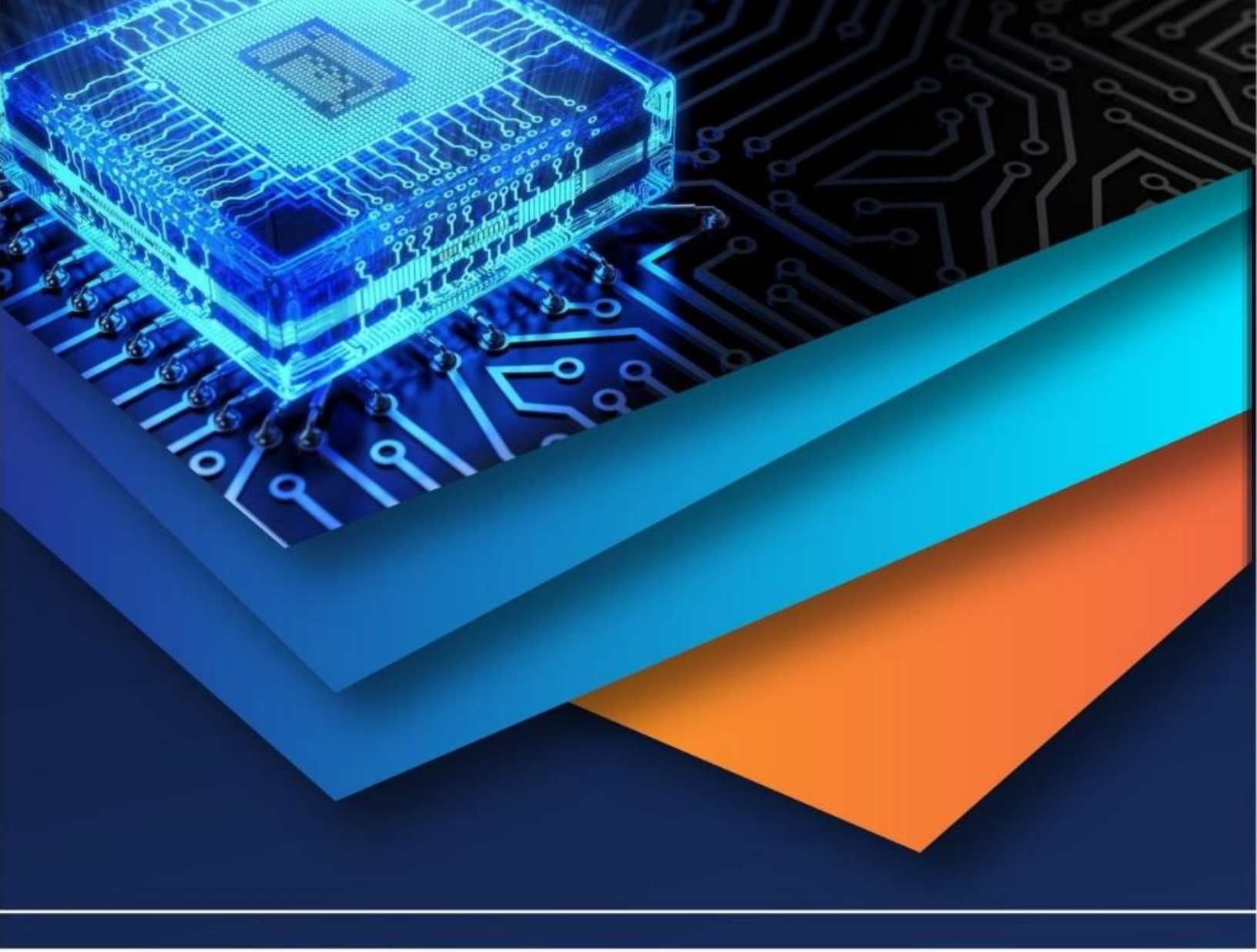
Timothy J. Wroge et al. [9] 2019	The data used for this analysis were collected through mPower, a clinical observational study conducted by Sage Bionetworks using an iPhone app	Scikit-Learn machine learning library as well as the TensorFlow and Keras Deep Learning Libraries	Decision tree ,Support vector machine ,Extra Trees ,Gradient Boosted Decision Tree ,Artificial Neural Network and Random Forest.	Cross validation with accuracy, F-1, recall and precision	Gradient Boosted Decision Tree with 86% and 82% accuracy ,F-1 score of 0.79 and 0.71,Precision of 0.85 and 0.789, recall score of 0.73 and 0.65 for AVEC and GeMaps features.
Basil K Varghese et al. [10] 2019	Parkinson's Telemonitoring dataset from UCI ML repository	----	Support Vector Regression , Decision Tree Regression ,Linear Regression and Resilient Back Propogation	Root Mean Squared Error (RMSE) of Motor and total UPDRS scores ,	Support Vector Regression demonstrated the best results with least RMSE values : 7.49(Total UPDRS) and 6.06(Motor UPDRS)
Srishti Grover et al. [11] 2018	Parkinson's Telemonitoring Voice Data Set from UCI Machine Learning Repository	The Python library, TensorFlow (tf.estimator)	Deep neural networks	Motor-UPDRS and Total-UPDRS accuracy	62.7335% accuracy with total UPDRS score and 81.6667% accuracy with motor UPDRS score
Tarigoppula V.S Sriram et al. [12] 2013	UCI Machine learning repository from Centre for Machine Learning and Intelligent Systems	Orange software v2.0b and weka v3.4.10	Bayes Net , Naïve Bayes , Logistic ,Simple Logistic , KStar , ADTree, J48, LMT and Random Forest	Parallel coordinates , Sieve graph ,ROC visualization and accuracy	Random Forest with an accuracy of 90.26% outperformed other algorithms

III. CONCLUSION

Managing Parkinson disease in day-to-day life is very challenging for an individual. A good screening procedure will be beneficial, especially in circumstances where a physician's treatment is not necessary. We discovered several researchers engaged in the field of Parkinson disease detection during the survey. A decision to select a specific system from the pool of available researchers can be made based on the requirements and resources available.

REFERENCES

[1] Amreen Khanum D, Prof. Kavitha G and Prof. Mamatha H S, "Parkinson's Detection using Machine Learning Algorithms", International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321- 9653 vol 10 Issue X, pp 786-790, Oct 2022.
[2] Muhtasim Shafi and Fizar Ahmed, "Parkinson's Disease Detection Analysis through Machine Learning Approaches" https://www.researchgate.net/publication/359711136, 2022.
[3] Mohesh T,Gowtham K,Vijeesh P and Arun Kumar S, "Parkinson's Disease Prediction Using Machine Learning",IJRASET44075,ISSN:2321-9653,vol 10 Issue VI,2022.
[4] Oduntan Ifeoma, "Prediction of Parkinson's Disease Using Biomedical Voice Measurements Dataset", https://www.researchgate.net/publication/35725620,2021.
[5] Sonia Singla, "Parkinson disease onset detection Using Machine Learning", https://www.analyticsvidhya.com/blog/2021/07/parkinson-disease-onset detection-using-machine-learning/,2022.
[6] Pramanik Anik and Sarker Amlan, "Parkinson's Disease Detection from Voice and Speech Data Using Machine Learning", https://www.researchgate.net/publication/347520593,2020.
[7] Anitha R,Nandhini T, Sathish Raj S and Nikitha V, "Early Detection Of Parkinson's Disease Using Machine Learning",IJARIIE,ISSN(O):2395-4396, Vol 6 Issue 2, pp 505-511,2020.
[8] Sanghee Moon, Hyun-Je Song, Vibhash D. Sharma, Kelly E. Lyons, Rajesh Pahwa, Abiodun E. Akinwuntan and Hannes Devos. "Classification of Parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: a data-driven approach" Journal of Neuro Engineering and Rehabilitation, ISSN: 1743-0003, vol 17, Issue 1,2020.
[9] T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi, "Parkinson's Disease Diagnosis Using Machine Learning and Voice",IEEE Signal Processing in Medicine and Biology Symposium (SPMB), ISSN: 2372- 7241, pp. 1-7, 2018.
[10] Basil K Varghese, Geraldine Bessie Amali D and Uma Devi K S. "Prediction of Parkinson's Disease using Machine Learning Techniques on Speech dataset" Research Journal of Pharmacy and Technology, vol 12 Issue (2), pp 644- 648,2019.
[11] Srishti Grover, Saloni Bhartia, Akshama, Abhilasha Yadav and Seeja K. R, "Predicting Severity of Parkinson's Disease Using Deep Learning", International Conference on Computational Intelligence and Data Science (ICCIDS), vol 132, pp 1788–1794,2018.
[12] Tarigoppula V.S Sriram, M. Venkateswara Rao, G V Satya Narayana, DSVGK Kaladhar and T Pandu Ranga Vital, "Intelligent Parkinson Disease Prediction Using Machine Learning Algorithms", IJEIT, vol 3, pp 212-215,2013.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY