JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Parkinson's Diseases Detection Using ML

Mr. Shivam Dubey[1], Mr. Amrut Gaikwad[2], Mr. Roshan Gupta[3], Mr. Pawan Rana[4],

Prof. Shrikrishna. d. patil[5]

[1]*Student, Department of electronics and telecommunication, MCT's Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*
[2]*Student, Department of electronics and telecommunication, MCT's Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*
[3]*Student, Department of electronics and telecommunication, MCT's Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*
[4]*Student, Department of electronics and telecommunication, MCT's Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*
[5]*Associate professor, Department of electronics and telecommunication, MCT's Rajiv Gandhi Institute of Technology, Mumbai, Maharashtra, India*

Keywords

Parkinson's disease, nervous system illnesses, mobility, speech problems, machine learning, KNN algorithm, early detection, local binary pattern, dataset, accuracy, MDVPs, Jitter, Shimmer, NHR, DFA, RPDE, prediction, website, CSS, HTML, user interface.

Abstract

One of the most destructive and moderate sensory system sicknesses that influence versatility is Parkinson's. It is one of the most dangerous neurological diseases, second only to Alzheimer, because it shortens life expectancy and affects posture, both of which have no known cure.

Discourse issues impact around 90% of the people who have this ailment. In real-world applications, a variety of machine learning methods are used to generate information. The early discovery of ailments utilizing AI calculations assists more seasoned individuals with living lengthier. The characteristics are taken from the audio of people with Parkinson's disease who have a local binary pattern.

While thinking about "Parkinson's," discourse attributes are the significant accentuation. The author of this study uses a variety of machine learning techniques, such as KNN, to predict Parkinson's disease using data from a dataset and user input. There are 195 people in the Parkinson's dataset, of which 48 had the disease and 147 were in good health. There were three records for every patient. The accuracy on the validation set was higher than 87 percent, which is comparable to other methods that are currently in use.

A number of parameters, including MDVPs, Jitter, Shimmer, NHR, DFA, and RPDE, are included in the dataset. Based on these factors, the author predicts the algorithm with greater accuracy. When treating patients when they are still in the early stages, prediction is essential. The author has developed a website on which the results can be viewed. CSS and HTML have been used to create websites. It furnishes us with exact information, precise data, and results. AI can be utilized to finish this method. The result may be either 0 or 1 after the user interface has received the parameter value.

1 Introduction

Parkinson's disease, a neurological condition that can cause stiffness, shivering, and trembling, as well as difficulty walking and maintaining balance, is the second most dangerous. Cell death in the nervous system was primarily the cause. Parkinson's disease can cause side effects that aren't caused by engines. Gradualness of development, unbending nature, issues with equilibrium, and quakes are the engine side effects. If this disease continues, patients may have difficulty walking and talking. Non-motor symptoms include anxiety, difficulty breathing, depression, a loss of smell, and changes in speech. If the individual demonstrates any of the aforementioned symptoms, the information is recorded in the records.

Neurodegenerative disorders were caused by neuron loss and progressive tearing in various parts of the nervous system. Neurons are the functional units of the brain. They are coterminous instead of steady.

1.1 Factors in the terrain

- Contact with fungicides and heavy essence like lead and aluminium.
- Quality of the Air Contamination brings about respiratory ails
- Water quality The presence of biological and abiotic contaminants in water results in water impurity.
- Unfortunate lifestyle It prompts strength and a dormant lifestyle.
- Mental pressure It raises the position of a hormone called stress, which weakens the functions of neurons.

1.2 growing element:

There are numerous causes of Parkinson's disease, one of which is growing. In India, the creator guarantees that out of 1,065, 070, 6072 individuals have Parkinson's protest. converse Explanation Factors PD discovery systems use a variety of instruments to focus on determining the inflexibility of symptoms.

1.3 Speech articulation factors:

The oral issue is perhaps of the most widely recognized side effect, and most extreme cases observer oral renunciations in the beginning phases of the protest. The Parkinson's disease causes some speech-language pathology, such as changes in voice, articulation, and swallowing. There are many ways that a person with Parkinson's disease (PD) can be affected.

- The voice becomes softer and further slurred.
- blackening of speech.
- The person's speech becomes sluggish as a result of difficulty chancing the right words.

Signs and symptoms of Parkinson's complaint.

A. Motor symptoms:

This is a symptom that can be treated on one's own. Bradykinesia, temblors, severity, freezing, or any voluntary muscle movement are all suggestions of this condition.

B. Non-Motor symptoms:

Examples of nonmotor symptoms include signs that are unrelated to mood and affect disorders like apathy, cognitive dysfunction, and complex behavioural disorders. Specialists distinguish two distinct subtypes of Parkinson's disease based on primary and secondary symptoms.

2 Problem statements and Objectives

2.1 Problem statements

- The basic point is to expect the assumption efficiency that would be important for the patients who are encountering Parkinson and the level of the contamination will be diminished.
- Subsequently, early identification of PD is basic to serve patients. Finding the prediction model—also known as a machine learning method—that can distinguish Parkinson's patients from healthy individuals is the primary objective of this study.
- The experimental study makes use of the voice dataset of Parkinson's patients that was downloaded from Kaggle.

The project that involves creating a website to diagnose Parkinson's disease by entering patient data is a useful tool for early detection. A machine learning algorithm is used on the website to look at patient data and figure out how likely Parkinson's disease is.

2.2 Objectives

The following objectives are included in the project:

- The primary goal of this project is to determine whether a person has Parkinson's disease.
- The goal mainly focuses on using the most effective algorithms to find people with Parkinson's disease early.
- By recognizing the Parkinson sickness at beginning phase, we can cut the cost on clinical finding.

The creation of a Parkinson's disease diagnosis website has the potential to improve patient outcomes and raise awareness of the condition. This makes the project look promising. With further advancements in machine learning algorithms and data analysis techniques, online diagnostic tools for Parkinson's disease and other neurological disorders can be refined and improved.

3 Research Methodology:

3.1 System design:

- For this project, we must devise a method for identifying the affected person by utilizing speech data and its various parameters.
- At the point when you partition your information into a preparation set and a testing set, you consider this a "train-test split." The testing set is used to test your model, while the preparation set is used to prepare the model. Consequently, you can train your models on the training set prior to assessing their accuracy on the unobserved testing set.
- The primary goal of splitting the dataset into a validation set is to stop our model from overfitting—that is, the model gets very good at classifying the samples in the training set but can't accurately generalize to data it hasn't seen before.

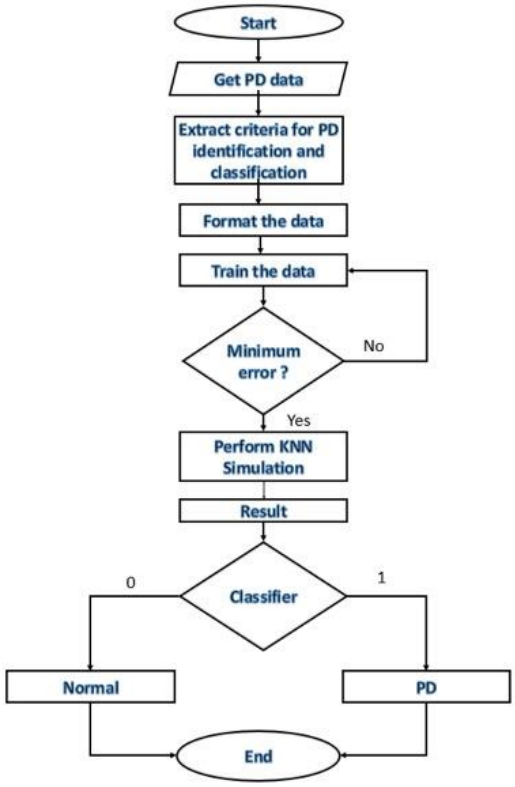


Fig 3.1.1: Block Diagram

3.2 Dataset Collection:

The primary goal of this step is to locate and collect all data-related issues. We might want to recognize the different information sources during this step since information are habitually assembled from various sources, like records and data sets, like Discourse DATASET PRE-Handling Information Preparing Information APPLY AI Calculations KNN NAVE BAYES Strategic REGRESION TEST Information Result 18. The result's adequacy still up in the air by the amount and nature of the gathered information. The more information that is available, the more accurate the prediction will be. We've assembled our data from the Kaggle site.

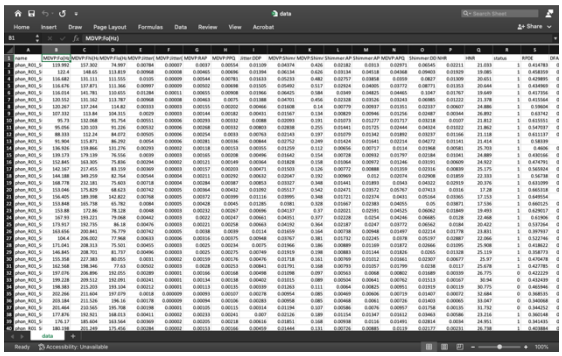


Fig-3.2.1 Speech data set

The figure above depicts the speech dataset gathered from the Kaggle website. This acquired dataset contains approximately 756 patient data and 755 distinct voice features per row. However, in order to locate the prediction, we selected ten primary characteristics in this paper.

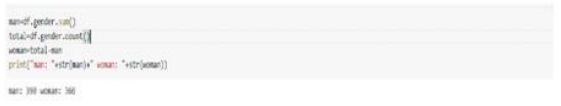
The highlights are recorded beneath:

- I.d.
- gender
- PPE (Pitch Period Entropy)
- DFA (Detrended Fluctuation Analysis)
- RPDE (Recurrent Period Density Entropy) 19
- numPulses
- numPeriodPulses
- meanPeriodPulses
- stdDevPeriodPulses
- locPctJitter
- locAbsJitter
- rapJitter
- locShimmer, etc.

For the dataset, we selected a CSV (Comma Separated Value) file. Read the data from the CSV file into the Google colab, also known as a Python notebook, is the next step. In our project, Python scratch pad is used for model correlation, highlights determination, and information pre-handling.

Fig 3.2.2 Read information from CSV record

we have advised the most ideal way to examine data from CSV records using the inbuilt python works that are fundamental for the 20 pandas library.
Parkinson's disease mostly affects men rather than women when comparing genders. We chose this dataset because it contains more men.



In Fig-3.2.3 We have demonstrated that there are more men than women

Testing Data A variety of data points are used to test the Parkinson's disease prediction model following the model's training on the pre-processed dataset. By giving the model a test dataset, the rightness and precision of the model are checked during this testing step. To decide the best model to utilize, each preparing technique should be confirmed. We involved this model to anticipate values for the test dataset subsequent to fitting it to the preparation information. These predicted values from the testing data are used in model comparison and precise calculation.

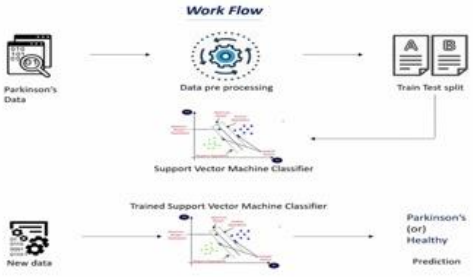


Fig 3.2.4: Flow of Process

4 Requirements:

A. Software:
   - Google Colab
   - VS Code

B. Operating System: Windows 10
C. Tools: Web Browser
D. Languages: HTML, CSS, Python(flask)
E. Python Libraries: numpy, pandas, matplotlib, seaborn, sklearn, pickle, flask, imblearn.

5 Experimental Analysis and Performance:

The most effective method or algorithm available to medical science for the early diagnosis of Parkinson's disease.
We have utilized a few measures to assess the prescient outcomes. Accuracy, Precision, Recall or Sensitivity, Specificity, F1-score, LR-, LR+, odd, and youden score are these measurements.

1. Accuracy:

It is the extent of the absolute number of forecasts that were right. It can be obtained by dividing the total number of Samples by the sum of true positive and true negative instances.

2. Precision:

The ratio of true positive to predicted yes instances is called precision. The ratio of correct positive results to the total positive results predicted by the system is another name for it.

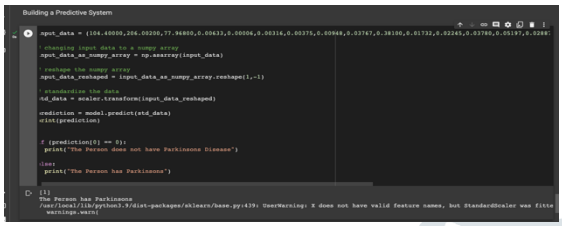
6 Result:

Using machine learning technology, we can predict the presence of Parkinson's disease in a patient's body through this procedure. The process is made simpler for our user by this approach. The field of neurodegenerative disease research is currently of great importance, and early diagnosis can improve a patient's quality of life.
Promising outcomes have been achieved by the most recent advancements in speech analysis methodologies. For the purpose of detecting PD, various deep learning and machine learning algorithms have been investigated in this paper. By analyzing the voice signals, our primary objective is to demonstrate the PD diagnosis. Since the voice estimations are harmless, in this manner discourse handling has been generally utilized in numerous different applications and has fantastic possible in the characterization and finding of PD for a long time. This paper is planned to investigate and determine the presentation of numerous grouping models. Patients no longer need to make in-person visits to clinics because these models can effectively monitor and diagnose PD remotely. When various voice datasets were used to test the various classifiers, it was discovered that the random forest outperformed the

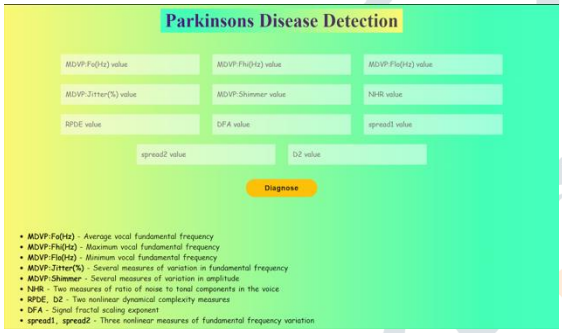
other classifiers in machine learning algorithms with an accuracy of more than 99 percent. The neural network classifier, one of the deep learning methods, also produced results with a reasonable accuracy of 99.49 percent. As a result, artificial intelligence techniques have been widely utilized for the early diagnosis and detection of Parkinson's disease.

We test the remaining test data with three different algorithms to demonstrate the project's outcomes. After that, our trained model is prepared to determine whether or not the disease is present. The test exactness is finished in the Google colab which is our python journal.

Using Google Colab:



Output using Flask:



7 Acknowledgment:

Everyone who contributed to the successful completion of this project on the application of machine learning to the prediction of Parkinson's disease deserves my sincere gratitude.
To begin, I would like to express my appreciation to my supervisor for providing me with guidance, support, and useful advice throughout the project. This study's design and success have greatly benefited from your expertise and knowledge.
Also, I might want to offer my thanks to the members who liberally contributed their information to this review. This task could never have been conceivable without their help.
I should convey my appreciation to the makers of the datasets used in this survey for making them unreservedly available. The scientific community has been able to learn more about Parkinson's disease as a result of your efforts, which have made it easier for research.
To wrap things up, I might want to offer my thanks to my friends and family for their help and consolation all through this undertaking. I've been inspired and motivated by your words of encouragement and belief in my abilities.

8 Reference

[1] K. Taha, et al., "Classification of speech intelligibility in Parkinson's disease," Biocybernetics and Biomedical Engineering, vol. 34, pp. 35-45, 15, July 2015.

[2] Pahuja, Gunjan, and T. N. Nagabhushan. "A comparative study of existing machine learning approaches for Parkinson's disease detection." IETE Journal of Research 67, no. 1 (2021): 4-14.

[3] Sriram, T.V., Rao, M.V., Narayana, G.S., Kaladhar, D.S.V.G.K. and Vital, T.P.R., 2013. Intelligent Parkinson disease prediction using machine learning algorithms. Int. J. Eng. Innov. Technol, 3, pp.212-215.

[4] Mamun, M., Mahmud, M.I., Hossain, M.I., Islam, A.M., Ahammed, M.S. and Uddin, M.M., 2022, October. Vocal Feature Guided Detection of Parkinson's Disease Using Machine Learning Algorithms. In 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0566-0572). IEEE.

[5] Safi, K., Aly, W.H.F., AlAkkoumi, M., Kanj, H., Ghedira, M. and Hutin, E., 2022. EMD-Based Method for Supervised Classification of Parkinson's Disease Patients Using Balance Control Data. Bioengineering, 9(7), p.283.Quan, C., Ren, K. and Luo, Z., 2021. A deep learning based method for Parkinson's disease detection using dynamic features of speech. IEEE Access, 9, pp.10239-10252.

[6] Tanveer, M., Rashid, A.H., Kumar, R. and Balasubramanian, R., 2022. Parkinson's disease diagnosis using neural networks: Survey and comprehensive evaluation. Information Processing & Management, 59(3), p.102909.

[7] Rana, A., Dumka, A., Singh, R., Panda, M.K., Priyadarshi, N. and Twala, B., 2022. Imperative Role of Machine Learning Algorithm for Detection of Parkinson's Disease: Review, Challenges and Recommendations. Diagnostics, 12(8), p.2003.

[8] Amato, F., Saggio, G., Cesarini, V., Olmo, G. and Costantini, G., 2023. Machine Learning-and Statistical-based Voice Analysis of Parkinson's Disease Patients: A Survey. Expert Systems with Applications, p.119651.

[9] Quan, C., Ren, K., Luo, Z., Chen, Z. and Ling, Y., 2022. End-to-end deep learning approach for Parkinson's disease detection from speech signals. Biocybernetics and Biomedical Engineering, 42(2), pp.556-574.

[10] Wang, W., Lee, J., Harrou, F. and Sun, Y., 2020. Early detection of Parkinson's disease using deep learning and machine learning. IEEE Access, 8, pp.147635-147646.

[11] Almeida, J.S., Rebouças Filho, P.P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R. and de Albuquerque, V.H.C., 2019. Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. Pattern Recognition Letters, 125, pp.55-62.

[12] Mathur, R., Pathak, V. and Bandil, D., 2019. Parkinson disease prediction using machine learning algorithm. In Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018 (pp. 357-363). Springer Singapore.

[13] Grover, S., Bhartia, S., Yadav, A. and Seeja, K.R., 2018. Predicting severity of Parkinson's disease using deep learning. Procedia computer science, 132, pp.1788-1794.

[14] Hazan, H., Hilu, D., Manevitz, L., Ramig, L.O. and Sapir, S., 2012, November. Early diagnosis of Parkinson's disease via machine learning on speech data. In 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel (pp. 1-4). IEEE.