

Parkinson's Disease Prediction

Prathamesh Patil¹, Vaibhav Patil², Rohit Morankar³, Karan Pawar⁴, Yogeshwari Hardas⁵

^{1,2,3,4,5}Department of computer engineering Atma Malik Institute of Technology and research, Shahapur, Maharashtra, India

Corresponding Author Email: vaibhavpatil1862001@gmail.com

Abstract Voice-based biomarkers can help diagnose symptoms of dementia such as Parkinson's disease, PD is a modern neurodegenerative disease affecting about 7 million people worldwide (usually adults), with about 150 thousand new scientific diagnoses performed each year. Historically, PD has been difficult to find and documents tend to focus on a few symptoms and even ignore some, depending on the scores of independent points. Due to the decline in motor manipulation which is a sign of illness, the term can be used as a means of detecting and diagnosing PD. Common sense has meant that physicians often focus on the symptoms of PD while ignoring the other. By using independent measurement scales, the term can be used to diagnose and diagnose the disease. This paper presents evidence to support the concept of supervised classification, which can be used to diagnose individuals with diseases such as diabetes and pulmonary fibrosis. Through Linear Regression, Logistic Regression, Decision Trees, Support Vector Machine, Random Forest, XGBoost, Neural Network and Adaboost we were able to achieve a peak accuracy of 100% for diagnosing pathological conditions. The project also uses various Evaluation Methods and Metrics such as Confusion Matrix, Classification Report, F1 - Score, Accuracy, Precision, Recall.

I. INTRODUCTION

Parkinson's disease is a neurodegenerative disorder that results in unintended or inevitable movements in our body. It is estimated to affect numerous people across the globe and has negatively affected the quality of life of those who are affected. If not treated on time and not given the effective treatment, it will start to hamper the health on worst level. So, a proper treatment in proper time, is a must for Parkinson's disease.

A program model can be taught to detect the disease with the help of large dataset and numerous training sets. So, a proper treatment in proper time, is a must for Parkinson's disease.

II. LITERATURE SURVEY

Parkinson's disease (PD) is a neurodegenerative disorder affecting millions globally. Early detection is crucial for better disease management. Machine learning and data mining techniques are showing promise in predicting PD. Here's a breakdown of key areas to explore in literature survey:

1. L. Naranjo, C. J. Pérez, J. Martín and Y. Campos-Roca, "A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications," *Computer Methods and Programs in Biomedicine*, vol. 142, pp.147-156, 2017.
2. G. Ahmad, S. Alanazi, M. Alruwaili, F. Ahmad, M. A. Khan *et al.*, "Intelligent ammunition detection and classification system using convolutional neural network," *Computers, Materials & Continua*, vol. 67, no. 2, pp. 2585-2600, 2021.
3. F. Alhaidari, S. H. Almotiri, M. A. A. Ghamdi, M. A. Khan, A. Rehman *et al.*, "Intelligent software-defined network for cognitive routing optimization using deep extreme learning machine approach," *Computers, Materials and Continua*, vol. 67, no. 1, pp. 1269-1285, 2021.
4. S. Y. Siddiqui, I. Naseer, M. A. Khan, M. F. Mushtaq, R. A. Naqvi *et al.*, "Intelligent breast cancer prediction empowered with fusion and deep learning," *Computers, Materials and Continua*, vol. 67, no. 1, pp. 1033-1049, 2021

5. Wroge, Timothy J., Yasin Özkanca, Cenk Demerol, Dong Si, David C. Atkins, and Reza Hosseini Ghomi. "Parkinson's disease diagnosis using machine learning and voice." In *2018 IEEE signal processing in medicine and biology symposium (SPMB)*, pp. 1-7. IEEE, 2018.
6. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018 American Diabetes Association Diabetes Care 2018; 41(Supplement 1): S13–S27.
7. Gullapalli, Ajay Sankar, and Vinay Kumar Mittal. "Early detection of Parkinson's disease through speech features and machine learning: A review." *ICT with Intelligent Applications: Proceedings of ICTIS 2021, Volume 1* (2022): 203-212.

III. METHODOLOGY

1. DATA ACQUISITION

Specify the source of your data. Public datasets of Parkinson's patients are available online. Describe the data characteristics:

Number of participants (patients and healthy controls)

Type of data collected (clinical information, genetic data, voice recordings, etc.)

2. DATA PREPROCESSING

Explain how you handled missing values and outliers in the data. Describe any data normalization techniques used.

3. FEATURE SELECTION (OPTIONAL)

Not all features in your data may be relevant for prediction.

You can implement feature selection techniques (filter-based or wrapper-based) to identify the most informative features for building the model.

4. MACHINE LEARNING MODEL SELECTION AND TRAINING

Choose a suitable machine learning algorithm for PD prediction. Common choices include:

Support Vector Machines (SVM), Random Forest, XGBoost.

IV. PROBLEM SYSTEM

Parkinson's disease (PD) is a neurodegenerative disorder with a significant impact on quality of life. Early detection is crucial for effective treatment and management.

Ensure data quality through cleaning and pre-processing steps.

V. EXISTING SYSTEM

Several existing systems leverage machine learning techniques for Parkinson's disease prediction. Here's a breakdown of some common approaches:

Machine Learning Algorithms:

- **Ensemble methods:** Random Forest, XGBoost (eXtreme Gradient Boosting) have shown promising results with high accuracy in predicting PD risk.
- **Support Vector Machines (SVMs):** Efficiently classify data points based on learned patterns, achieving good performance in PD prediction.
- **Neural Networks:** Deep learning architectures like Deep Brooke Inception Net have been successful in predicting disease severity based on voice abnormalities.

VI. PROPOSED SYSTEM

Parkinson's disease (PD) is a neurodegenerative disorder with a significant impact on quality of life. Early detection

is crucial for effective treatment and management.

Ensure data quality through cleaning and pre-processing steps.

Feature Engineering:

Extract relevant features from the data that can potentially indicate PD.

Examples include:

Speech characteristics (tremor-induced variations), Movement patterns (rigidity, slowness), Demographic information (age, family history), Genetic markers (if available)

Machine Learning Model:

Train an ML model on the prepared data. Popular choices for PD prediction include:

Support Vector Machines (SVM), Random Forest, X G Boost Deep Neural Networks (for complex feature extraction).

Evaluate different models and select the one with the best performance metrics (accuracy, sensitivity, specificity).

Prediction and Interpretation:

Use the chosen model to predict PD risk for new patients based on their features.

Employ techniques like feature importance analysis to understand which factors contribute most to the prediction.

VII. RESULTS

Disease Name	Algorithm Name	Existing system accuracy
Parkinson's disease	SVM	71%

VIII. CONCLUSION

In this paper, we introduce a unique, machine learning- based Parkinson Disease Prediction using parkinson.csv data file from <https://www.kaggle.com/>.

From this we have taken parkinson.data file after that we change the name to parkinson.csv, after that we have placed in our document folder.

Using kaggle ml and data science community we got easy way for our project. Our program aspect of individual input or also can say that it only take one input at a time. When user

5. REFERENCES

[1] L. V. Kalia and A. E Lang, "Parkinson's disease,." *The Lancet*, vol. 386, no. 9996, pp. 896–912, 2015.

[2] J. P. Iannotti and R. Parker, "The netter collection of medical illustrations-musculoskeletal system,," in *Elsevier Health Sciences*, 2nd edition, vol. 6, 2013.

[3] M. Fjodorova, E.M. Torres and S.B. Dunnett, "Transplantation site influences the phenotypic differentiation of dopamine neurons in ventral mesencephalic grafts in Parkinsonian rats," *Experimental Neurology*, vol. 291, pp. www.ijsssr.com

8-19, 2017.

[4] A. Jamak, A. Savatić and M. Can, "Principal component analysis for authorship attribution," *Business Systems Research*, vol. 3, pp. 49-56, 2012.

[5] M. Can, "Neural networks to diagnose the Parkinson 's disease," *Southeast Europe Journal of Soft Computing*, vol.2, no. 1, 2013.

[6] L.V. Kalia, S. K Kalia and A. E. Lang, "Disease modifying strategies for Parkinson's disease," *Movement Disorders*, vol. 30, pp. 1442-1450, 2015.

[7] N. Singh, V. Pillay and Y.E. Choonara, "Advances in the treatment of Parkinson's disease," *Progress in Neurobiology*, vol. 81, pp. 29-44, 2007.

Retinopathy disease severity scale detailed table", Tech. Rep., 2002.