

## DETECTING PARKINSON'S DISEASE USING MACHINE LEARNING

Raksha Sharma

GITAM University, Hyderabad, India.

### ABSTRACT

Parkinson disease is a neurodegenerative disorder that affects nervous system and the root cause of it is falling rates of dopamine levels in the forebrain. It is a chronic degenerative disease with progressive illness, which means it develops new symptoms over time. This happens with progressive neuronal loss in the substantia nigra of brain. People with PD cannot do their works as a normal human. Though clinical assessments considered ample amount of data that include various features, sometimes it is hard to decide whether a person is suffering from PD or not based on the type of data, feature selection methods help to solve this issue. Various methods are developed, proposed, and analyzed to detect the Parkinson disease, given the required data. This paper is a survey of predicting Parkinson disease using machine learning algorithms, various new technologies applied, and their accuracies achieved.

**Keywords:** PD (Parkinson Disease), dopamine, SVM (Support Vector Machine), KNN (K Nearest Neighbor), ANN (Artificial Neural Network).

### INTRODUCTION

Parkinson's disease is a progressive disorder of the central nervous system affecting movement and inducing tremors and stiffness. It has 5 stages to it and affects more than 1 million individuals every year in India. This is chronic and has no cure yet. It is a neurodegenerative disorder affecting dopamine-producing neurons in the brain. Biomarkers derived from human voice can offer insight into neurological disorders, such as Parkinson's disease (PD), because of their underlying cognitive and neuromuscular function. PD is a progressive neurodegenerative disorder that affects about one million people in the United States, with approximately sixty thousand new clinical diagnoses made each year. Historically, PD has been difficult to quantify and doctors have tended to focus on some symptoms while ignoring others, relying primarily on subjective rating scales. Due to the decrease in motor control that is the hallmark of the disease, voice can be used as a means to detect and diagnose PD. With advancements in technology and the prevalence of audio collecting devices in daily lives, reliable models that can translate this audio data into a diagnostic tool for healthcare professionals would potentially provide diagnoses that are cheaper and more accurate. We provide evidence to validate this concept here using a voice dataset collected from people with and without PD.

This paper explores the effectiveness of using supervised classification algorithms, such as deep neural networks, to accurately diagnose individuals with the disease. Our peak accuracy of 85% provided by the machine learning models exceed the average clinical diagnosis accuracy of non-experts (73.8%) and average accuracy of movement disorder specialists (79.6% without follow-up, 83.9% after follow-up) with pathological post-mortem examination as ground truth.

### IMPLEMENTATION

In this Python machine learning project, using the Python libraries scikit-learn, numpy, pandas, and xgboost, we will build a model using an XGB Classifier. We'll load the data, get the features and labels, scale the features, then split the dataset, build an XGB Classifier, and then calculate the accuracy of our model.

It is a disease which is a disorder in the nervous system. Parkinson's disease affects the movement of the human body. In today's world, around 1 million people are suffering from this disease. This is a disorder which produces neurodegenerative dopamine-producing neurons in the brain. The following system will detect Parkinson's symptoms in the human body.

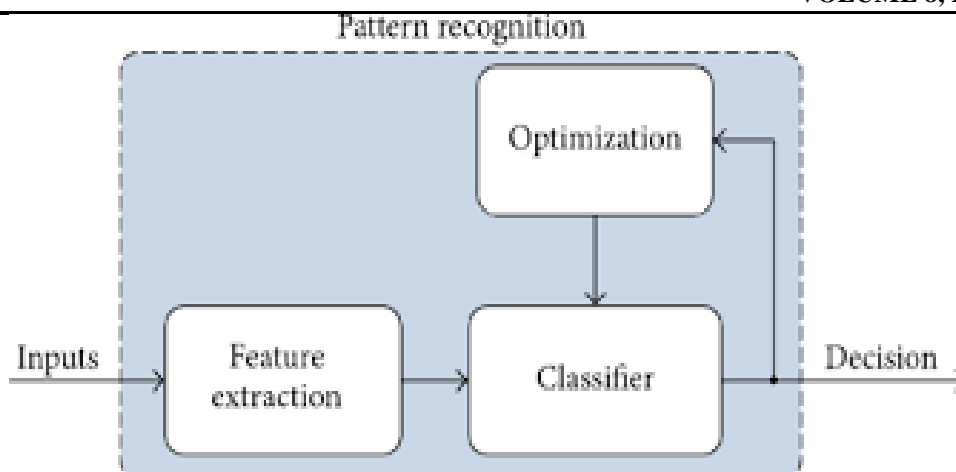


Figure 1: Parkinson's Disease Detection Flowchart

The project will be made by a new machine learning algorithm called the XG Boost. XG Boost is a new Machine Learning algorithm designed with speed and performance in mind. XG Boost stands for Extreme Gradient Boosting and is based on decision trees. In this project, we will import the XGB Classifier from the XG boost library; this is an implementation of the scikit-learn API for XG Boost classification. To build a model to accurately detect the presence of Parkinson's disease in an individual.

## METHODOLOGY

**XG Boost-** is an implementation of gradient boosted decision trees designed for speed and performance that is dominative competitive machine learning. XGBoost provides a wrapper class to allow models to be treated like classifiers or regressors in the scikit-learn framework. This means we can use the full scikit-learn library with XGBoost models.

**XGB Classifier-**The XGBoost model for classification is called XGB Classifier. Initialize an XGBClassifier and train the model. This classifies using eXtreme Gradient Boosting- using gradient boosting algorithms for modern data science problems. It falls under the category of Ensemble Learning in ML, where we train and predict using many models to produce one superior output. We can create and fit it to our training dataset. Models are fit using the scikit-learn API and the model.fit() function. To detect the presence of Parkinson's Disease in individuals using various factors. We used an XGBClassifier for this and made use of the sklearn library to prepare the dataset.

## RESULTS

In this Python machine learning project, we learned to detect the presence of Parkinson's Disease in individuals using various factors. We used an XGBClassifier for this and made use of the sklearn library to prepare the dataset. This gives us an accuracy of 94.87%, which is great considering the number of lines of code in this python project.

```
y_pred=model.predict(x_test)
print(accuracy_score(y_test, y_pred)*100)

94.87179487179486
```

[ ]:

Figure 2: Output Screenshot

## CONCLUSION

The aim of the study is to analyze which algorithm provide the high accuracy of prediction for the Parkinson's disease dataset, here the classification accuracy was studied and compared, with good performance and fast implementation XGBoost achieved a high accuracy with 94%. This system provides the comparison between machine learning classifiers of LR and XGBoost in PD disease diagnosis with high dimensional data.

## REFERENCES

- 1) Machine-learning-based FoG prediction using estimated gait characteristics with a wearable accelerometer. *Sensors*. 18:E3287. doi: 10.3390/s18103287
- 2) PubMed Abstract | CrossRef Full Text | Google Scholar
- 3) Artyomov, E., and Yadid-Pecht, E. O. (2005). Modified high-order neural network for invariant pattern recognition. *Pattern Recogn. Lett.* 26, 843–851. doi: 10.1016/j.patrec.2004.09.029
- 4) CrossRef Full Text | Google Scholar
- 5) Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.
- 6) Google Scholar
- 7) Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* 2, 303–314. doi: 10.1007/BF02551274
- 8) CrossRef Full Text | Google Scholar
- 9) Djuric-Jovicic, M., Petrovic, I., Jecmenica-Lukic, M., Radovanovic, S., Dragasevic-Miskovic, N., Belic, M., et al. (2016). Finger tapping analysis in patients with Parkinson's disease and atypical parkinsonism. *J. Clin. Neurosci.* 30, 49–55. doi: 10.1016/j.jocn.2015.10.053
- 10) PubMed Abstract | CrossRef Full Text | Google Scholar.