Parkinson's Disease Prediction: Early Diagnosis through Machine Learning

- Dev Karnik

Problem Statement:

Parkinson's Disease (PD) is a neurodegenerative disorder that affects millions of people worldwide, with symptoms ranging from tremors and muscle rigidity to impaired balance and difficulty in walking. Early detection of PD is crucial for effective treatment and improved quality of life for patients. However, diagnosing PD in its early stages remains challenging, often leading to delayed intervention and missed opportunities for timely treatment.

This project aims to develop a machine-learning model that can accurately predict the likelihood of an individual developing Parkinson's Disease based on specific features and biomarkers. By harnessing advanced data analytics techniques and machine learning algorithms, we aim to create a reliable predictive tool that can aid healthcare professionals in identifying individuals at risk of PD at an early stage.

Dataset Information:

Data Set Characteristics: Multivariate

Number of Instances: 197

Number of Attributes: 23

Associated Tasks: Classification

Missing Values: N/A

Attribute Information:

Name - *ASCII subject name and recording number*

MDVP:Fo(Hz) - *Average vocal fundamental frequency*

MDVP:Fhi(Hz) - *Maximum vocal fundamental frequency*

MDVP:Flo(Hz) - *Minimum vocal fundamental frequency*

MDVP:Jitter(%),MDVP:Jitter(Abs),MDVP:RAP,MDVP:PPQ,Jitter:DDP

- Several measures of variation in fundamental frequency

MDVP:Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA - *Several measures of variation in amplitude*

NHR, HNR - *Two measures of ratio of noise to tonal components in the voice status –*

 Health status of the subject (one) - Parkinson's, (zero) - healthy

RPDE, D2 - *Two nonlinear dynamical complexity measures*

DFA - *Signal fractal scaling exponent*

spread1, spread2, PPE - *Three nonlinear measures of fundamental frequency variation*