Parkinson's Disease Analysis Using Support Vector Machine

# 1. Introduction

Parkinson's disease is a progressive nervous system disorder that affects movement. This project aims to analyze and predict the disease using a dataset containing various voice measurements.

# 2. Dataset Description

The dataset used for this analysis includes the following parameters:  
- 'name'  
- 'mdvp\_fohz'  
- 'mdvp\_fhihz'  
- 'mdvp\_flohz'  
- 'mdvp\_jitter%'  
- 'mdvp\_jitterabs'  
- 'mdvp\_rap'  
- 'mdvp\_ppq'  
- 'jitter\_ddp'  
- 'mdvp\_shimmer'  
- 'mdvp\_shimmerdb'  
- 'shimmer\_apq3'  
- 'shimmer\_apq5'  
- 'mdvp\_apq'  
- 'shimmer\_dda'  
- 'nhr'  
- 'hnr'  
- 'status'  
- 'rpde'  
- 'dfa'  
- 'spread1'  
- 'spread2'  
- 'd2'  
- 'ppe'

# 3. Data Cleaning and Preprocessing

The following steps were performed for data cleaning and preprocessing:  
- Handled missing values  
- Normalized/standardized the data

# 4. Exploratory Data Analysis

The following plots were created during the analysis:  
- Countplot: Analyzes the distribution of categorical variables.  
- Heatmap: Visualizes correlations between variables.  
- Barplot: Displays relationships between categorical and numerical variables.  
- Histplot: Shows the distribution of numerical variables.

# 5. Model Development

The Support Vector Machine (SVM) algorithm was used to classify patients with Parkinson's disease.

# 6. Results

The model's evaluation metrics are as follows:  
Accuracy: 1.00  
  
Confusion Matrix:  
[[ 7 0]  
 [ 0 32]]  
  
Classification Report:  
 precision recall f1-score support  
 0 1.00 1.00 1.00 7  
 1 1.00 1.00 1.00 32  
 accuracy 1.00 39  
 macro avg 1.00 1.00 1.00 39  
weighted avg 1.00 1.00 1.00 39

# 7. Conclusion

The model successfully predicted Parkinson's disease with 100% accuracy, demonstrating its effectiveness in classifying the disease based on voice measurements. Future work may explore additional features or other machine learning models for improved prediction.