**Machine Learning-Based Detection of Parkinson's Disease**

A black and white logo

AI-generated content may be incorrect.

**Project Proporsal**

**of Module EC5203**

Team Number: 33

Team Members:

Gunasena P.S. – EG/2022/5052

Karunarathna P.I. – EG/2022/5127

**Group Members**

Member 1:

Name : Gunasena P.S.

Registration Number : EG/2022/5052

Student Email : [eg245052@engug.ruh.ac.lk](mailto:eg245052@engug.ruh.ac.lk)

Member 2:

Name : Karunarathna P.I.

Registration Number : EG/2022/5127

Student Email : [eg245127@engug.ruh.ac.lk](mailto:eg245127@engug.ruh.ac.lk)

**Project Description**

This project aims to develop a machine learning system for the early and non-invasive detection of Parkinson's Disease (PD). PD is a neurodegenerative disorder that often manifests in vocal impairments (known as hypokinetic dysarthria). The system will analyze a set of 22 biomedical voice measurements from individuals. These features capture various aspects of vocal acoustics, such as fundamental frequency, variation in amplitude, and measures of nonlinear dynamics. By training on this data, the model will learn to classify subjects into two categories: healthy or Parkinson's positive. This tool has the potential to assist in early diagnosis and monitoring of the disease.

**Dataset Link**

<https://www.kaggle.com/datasets/jainaru/parkinson-disease-detection>

**Dataset Details**

* **Original Number of Features in the Dataset:** 23 Columns (22 feature variables + 1 target variable)
* **Target Variable:** status
  + 1 - Indicates the presence of Parkinson's Disease.
  + 0 - Indicates a healthy subject.
* **Dataset Variable Description:**
  + MDVP:Fo(Hz): Average vocal fundamental frequency.
  + MDVP:Fhi(Hz): Maximum vocal fundamental frequency.
  + MDVP:Flo(Hz): Minimum vocal fundamental frequency.
  + MDVP:Jitter(%): A measure of period-to-period variation in fundamental frequency (percentage).
  + MDVP:Jitter(Abs): Absolute measure of jitter.
  + MDVP:RAP: Relative amplitude perturbation.
  + MDVP:PPQ: Five-point period perturbation quotient.
  + Jitter:DDP: Average absolute difference of differences between cycles, related to RAP.
  + MDVP:Shimmer: A measure of period-to-period variation in amplitude.
  + MDVP:Shimmer(dB): Shimmer measured in decibels.
  + Shimmer:APQ3: Three-point amplitude perturbation quotient.
  + Shimmer:APQ5: Five-point amplitude perturbation quotient.
  + MDVP:APQ: A different method of measuring amplitude perturbation.
  + Shimmer:DDA: Average absolute difference between consecutive differences between consecutive amplitudes, related to APQ3.
  + NHR: Noise-to-harmonics ratio.
  + HNR: Harmonics-to-noise ratio.
  + RPDE: Recurrence Period Density Entropy (a nonlinear dynamical complexity measure).
  + DFA: Detrended Fluctuation Analysis (a scale-invariant fractal measure of self-similarity).
  + spread1: A nonlinear measure of fundamental frequency variation.
  + spread2: A nonlinear measure of fundamental frequency variation.
  + D2: Correlation dimension (another nonlinear measure).
  + PPE: Pitch Period Entropy.
* **Type of Problem:** Supervised - Classification (Binary Classification)
* **Algorithms Selected:**
  + Algorithm 01: Support Vector Machine (SVM)
  + Algorithm 02: Random Forest Classifier

**Colab Link**

<https://colab.research.google.com/drive/14zwn-Q-xueXGKLF_tBaVu3RkWJQPdUa0?usp=sharing>

**GitHub Repository**

<https://github.com/pamudithasandaru/Parkinson-s-Disease-Detector.git>