CSC354 Course-Code

Dr Saeed ur Rehman

Comsats University Islamabad (Wah Campus)

**Student Name: Student ID:**

Ahmad Hasham FA19-BCS-028

M. Abubaker Abid FA19-BCS-170

**Contents**

[**Introduction to Parkinson Disease** 2](#_Toc99363302)

[**Data set, description, and links** 3](#_Toc99363303)

[**Expected algorithm that will be used** 4](#_Toc99363304)

# **Introduction to Parkinson Disease**

**What is Parkinson disease?**

Parkinson Disease is a brain neurological disorder. It leads to shaking of the body, hands and provides stiffness to the body. No proper cure or treatment is available yet at the advanced stage. Treatment is possible only when done at the early or onset of the disease. These will not only reduce the cost of the disease but will also possibly save a life. Most methods available can detect Parkinson in an advanced stage; which means loss of approx 60% dopamine in basal ganglia and is responsible for controlling the movement of the body with a small amount of dopamine. More than 145,000 people have been found alone suffering in the U.K and in India, almost one million population suffers from this disease and it’s spreading fast in the entire world.

**Parkinson’s Disease Symptoms:**

A person diagnosed with Parkinson’s disease have other symptoms that include:

1. Depression

2. Anxiety

3. Sleeping, and memory-related issues

4. Loss of sense of smell along with balance problems.

**Parkinson’s Disease Causes:**

What causes Parkinson’s disease is still unclear, but researchers have research that several factors are responsible for triggering the disease. It includes:

1. Genes- Certain mutation genes have been found by research that are very rare. The gene variants often increase the risk of Parkinson’s disease but have a lesser effect on each genetic marker.

2. Environment- Due to certain harmful toxins or chemical substances found in the environment can trigger the disease but have a lesser effect

Although it develops at age of 65 15% can be found at young age people less than 50.

**Parkinson’s Disease Stages:**

The Parkinson’s disease is progressive and is marked by five different stages.

**Stage 1:** Mild symptoms that do not typically interfere with daily life, including tremors and movement issues on only *one* side of the body.

**Stage 2:** Symptoms continue to become worse with both tremors and rigidity now affecting *both* sides of the body. Daily tasks become challenging.

**Stage 3:** Loss of balance and movements with falls becoming frequent and common. The patient is still capable of (typically) living independently.

**Stage 4:** Symptoms become severe and constraining. The patient is unable to live alone and requires help to perform daily activities.

**Stage 5:** Likely impossible to walk or stand. The patient is most likely wheelchair bound and may even experience hallucinations.

While Parkinson’s cannot be cured, **early detection along with proper medication can *significantly* improve symptoms and quality of life,** making it an important topic as computer vision and machine learning practitioners to explore.

# **Data set, description, and links**

The following dataset was created by Max Little of the University of Oxford, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals.

<https://archive.ics.uci.edu/ml/machine-learning-databases/parkinsons/>

**Data Set Information:**

This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Each column in the table is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column). The main aim of the data is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD. The data is in ASCII CSV format. The rows of the CSV file contain an instance corresponding to one voice recording. There are around six recordings per patient, the name of the patient is identified in the first column.

**Attribute Information:**

Matrix column entries (attributes):  
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.

This following dataset is collected from UCI Machine Learning Repository through the following link:

[https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification#](https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification)

**Dataset:**

<https://www.kaggle.com/code/parhamzm/parkinson-s-disease-pd-classification/data>

The data used in this study were gathered from 188 patients with PD (107 men and 81 women) with ages ranging from 33 to 87 (65.1Â±10.9) at the Department of Neurology in CerrahpaÅŸa Faculty of Medicine, Istanbul University. The control group consists of 64 healthy individuals (23 men and 41 women) with ages varying between 41 and 82 (61.1Â±8.9). During the data collection process, the microphone is set to 44.1 KHz and following the physicianâ€™s examination, the sustained phonation of the vowel /a/ was collected from each subject with three repetitions.

**Dataset:**

<https://www.kaggle.com/datasets/kmader/parkinsons-drawings>

Images of healthy and patients with Parkinsons drawing spirals and waves. The images are further divided into training and testing groups for comparing (or reproducing) the results of the original publication.

# **Expected algorithm that will be used**

We will use here XGBoost, KNN Algorithm, Support Vector Machines (SVMs), Random Forest Algorithm and utilize the data-set available on UCL Parkinson Data-set.

**XGBoost** is a new Machine Learning algorithm designed with speed and performance in mind. XGBoost stands for eXtreme Gradient Boosting and is based on decision trees. In this project, we will import the XGBClassifier from the xgboost library

**Support vector machine algorithm** for the analysis of classification and regression. It is a supervised machine algorithm used. Image classification and hand-written recognition are where the support vector machine comes to hand used. It sorts the data in one out of two categories and displays the output with the margin between the two as far as possible.

**K-Nearest Neighbors** is one of the most powerful utilized algorithms of machine learning that is widely used both for regression as well as classification tasks. To predict and examine the class in which data points fall, it examines the label of chosen data points surrounded by the target point.

**Random forests** are an ensemble version of many choice bushes, wherein each tree will specialize its focus on a specific feature while maintaining a top-level view of all capabilities.