# **Objectives:**

Processing and extracting features from audio signals to build a binary classifier distinguishing between normal individuals and Parkinson's patients.

# **Methodology:**

## Voice Data collection:

First, one subject will be instructed to sustain the phonation of sustained vowels, such as "aah" or "eee," for 5 seconds.

## Data Preprocessing:

Recorded data will be pre-processed by following steps:

### **Plotting and Checking the Frequency spectrum:**

Firstly, the input voice signals time domain representation and frequency domain representation were checked to find the cut off frequency of our filter for filtering unwanted noise components. A random recorded voice signals time domain and frequency spectrum is shown below:

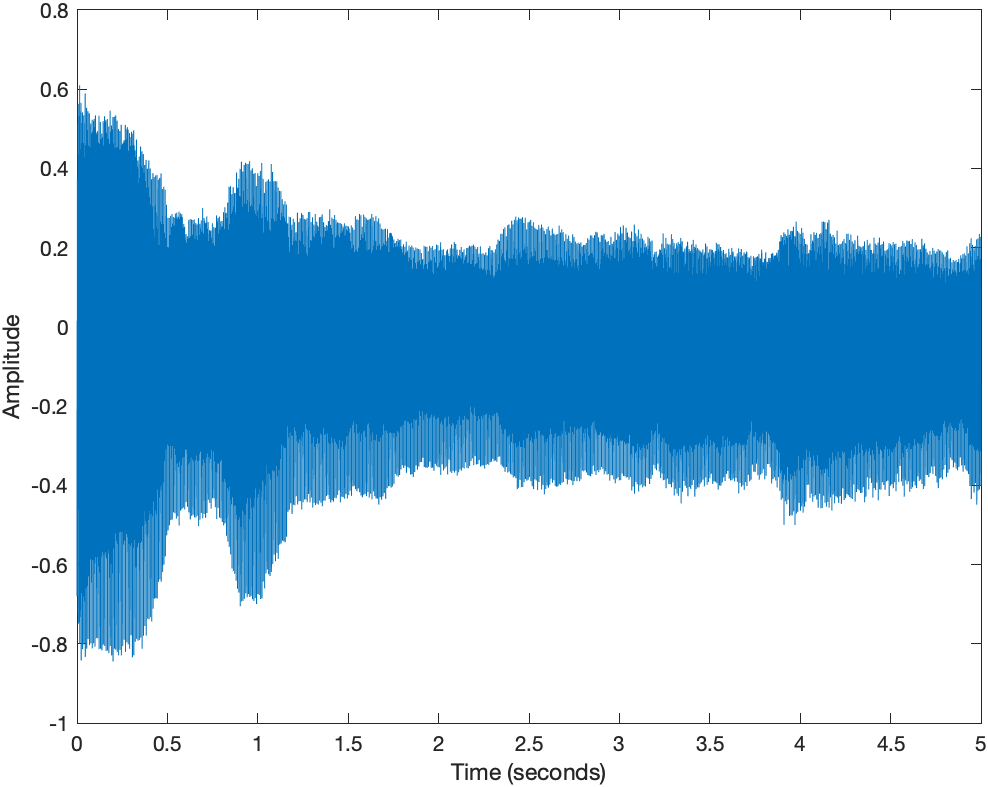


Fig:01

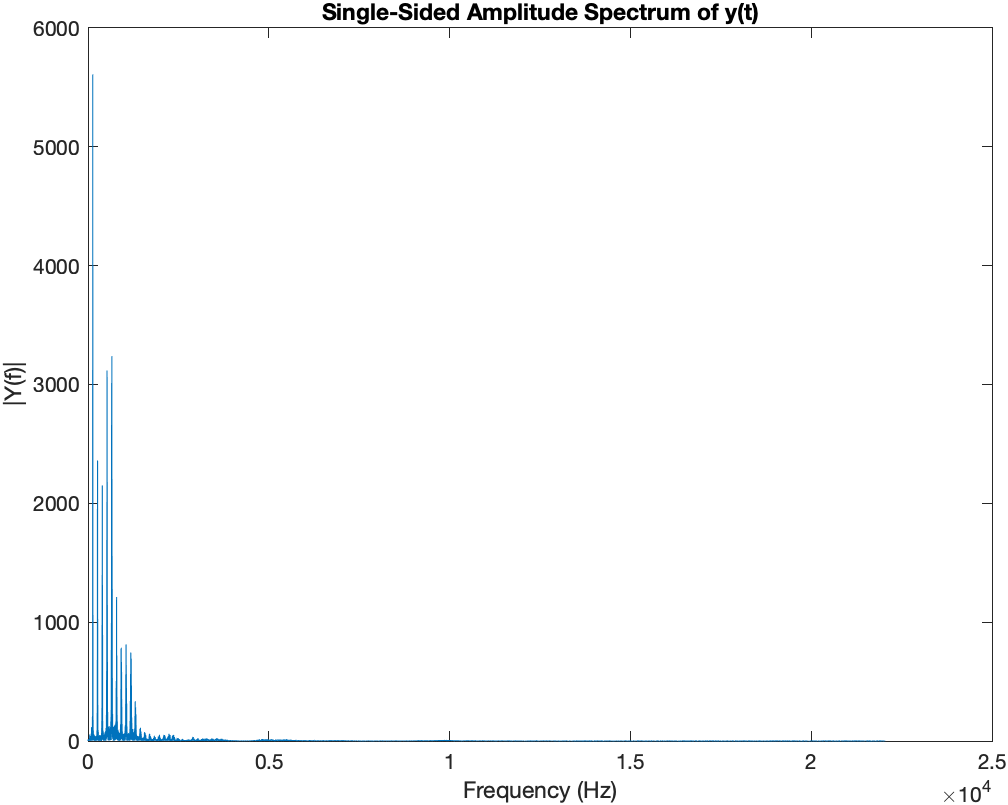


Fig:02

### **Filtering:**

The background noise and other unwanted signal frequency components were removed using a 5th order butterworth lowpass filter. The cut-off frequency for the filter was selected as 10 KHz. The frequency response of the filter is shown below:

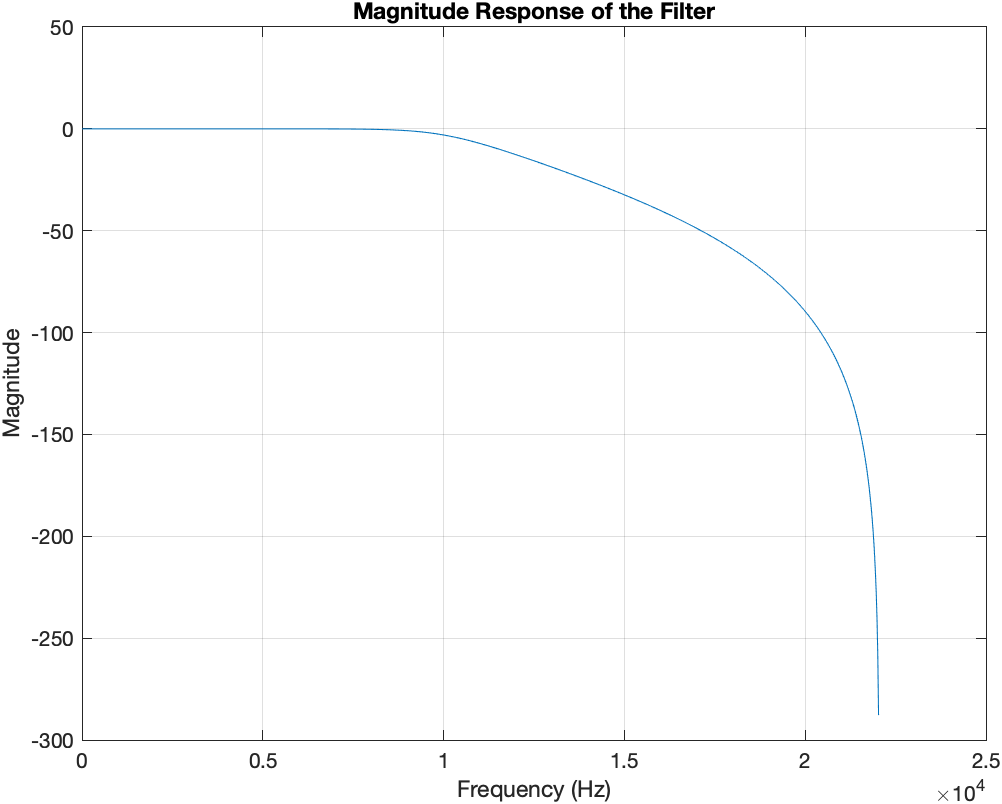


Fig:03

### **Segmentation:**

The signal was segmentized where each segment contained 2.5 seconds duration of the original signal. The last segment was padded with zero if the required sample number was insufficient for the last segment.

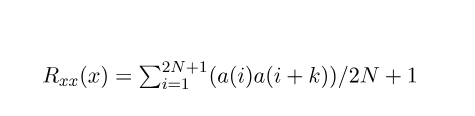
### **Normalization:**

Each segment was normalized using standard deviation to scale the signal in a way that is robust to variations in amplitude and noise. This normalization centers the data around zero and scales the data to have unit variance. It ensures that the signal's spread of values is consistent across different segments, which can be beneficial for feature extraction algorithms.

## Feature Extraction:

Features to be extracted:

1. **Pitch:** Mathematically computing the fundamental frequency (F0) of an audio signal sample involves analyzing the signal to identify its periodicity. One common method for estimating the fundamental frequency is the autocorrelation method. The following mathematical model was used.



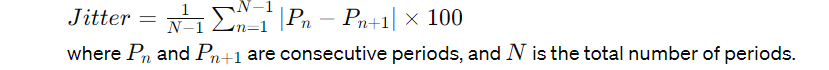
Based on the above expression,

N = Window/Segment size

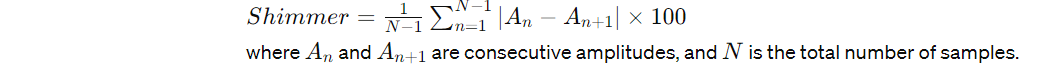
a = signal samples

k = zero lag locations

1. **Jitter :** Jitter refers to a measure of the cycle-to-cycle variation in the fundamental frequency (F0) of a periodic signal, such as the human voice. Jitter is also known as frequency perturbation or pitch variability. It reflects the irregularity or instability in vocal fold vibration, which can occur due to vocal fold pathology or neurological conditions such as Parkinson's disease. Jitter is typically expressed as a percentage (%), representing the relative change in F0 from one vocal cycle to the next. Higher jitter values indicate greater variability in F0, which may be indicative of vocal fold tremor or dysphonia associated with Parkinson's disease. The following mathematical model was implemented in MATLAB:



1. **Shimmer:** Shimmer is a measure of the cycle-to-cycle variation in the amplitude or intensity of the voice signal. It quantifies the irregularity or instability in vocal fold vibration related to changes in vocal fold stiffness, mass, or tension. Shimmer is typically expressed as a percentage (%), representing the relative change in signal amplitude from one vocal cycle to the next. The typical formula for shimmer calculation is the average absolute difference between consecutive amplitudes, expressed as a percentage:



1. **HNR :** The Harmonics-to-Noise Ratio (HNR) is a measure of the ratio of harmonic components (periodic signal) to non-harmonic components (noise) in the voice signal. It provides an indication of the clarity or purity of the voice signal. In the context of Parkinson's disease, changes in vocal fold dynamics or vocal tract resonance may result in alterations in HNR. Lower HNR values may be indicative of increased noise or turbulence in the voice signal, which can occur due to vocal fold irregularities or dysphonia associated with Parkinson's disease.

In order to compute the HNR for each of the segment the following method was implemented:

* Compute the power spectrum of the voice signal using the Fast Fourier Transform (FFT).
* Estimate the harmonic components of the signal by identifying peaks in the power spectrum corresponding to the fundamental frequency and its harmonics.
* Estimate the noise floor by averaging the power spectrum outside the harmonic regions.
* Calculate the HNR as the ratio of the sum of harmonic components to the noise floor.

1. Labeling:

Each segmentized voice signal features was given a label either 0 or 1. 0 was used for normal people and 1 was used for parkinson patients. These labels are going to be fed in the binary classifier which will be implemented using the support vector machine.

# **Project Output:**

## Statistical Analysis:

After collecting voice signals from all our subjects and extracting necessary features, all the feature data were plotted as scatter plots.

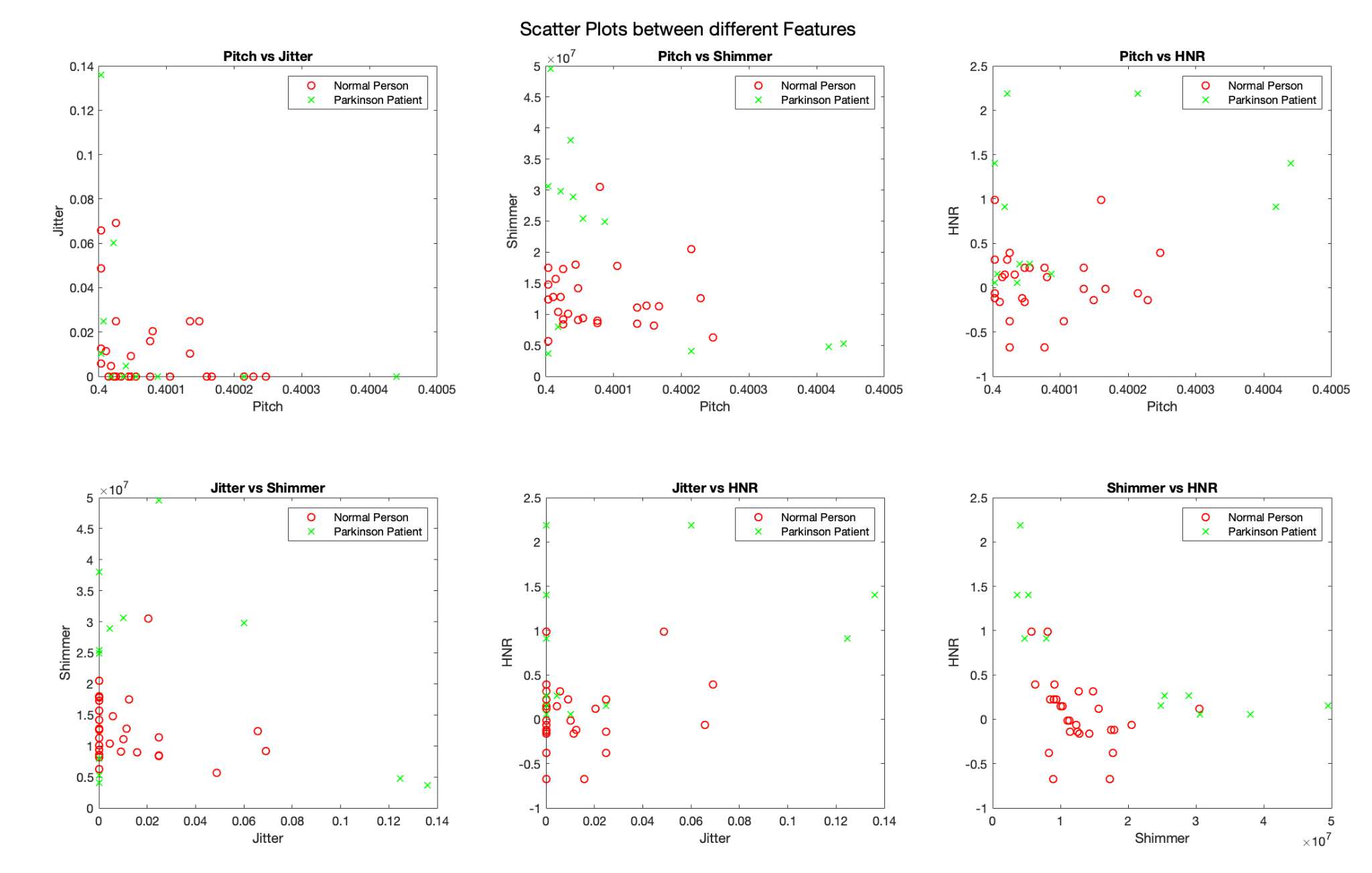


Fig:04

The scatter plots were plotted against any two of the features for both classes of our subjects. Some of the scatter plots do not have any significant correlation or other properties to distinguish between two of our classes.

We also did principle component analysis (PCA) to check in how many directions our feature space is varying. But after running the PCA analysis, we saw that almost 99.99 variations of the data were captured within one dimension. This signifies that our prepared 4 dimensional dataset is almost one dimensional. The scatter plot using 2 PCA features is shown below:

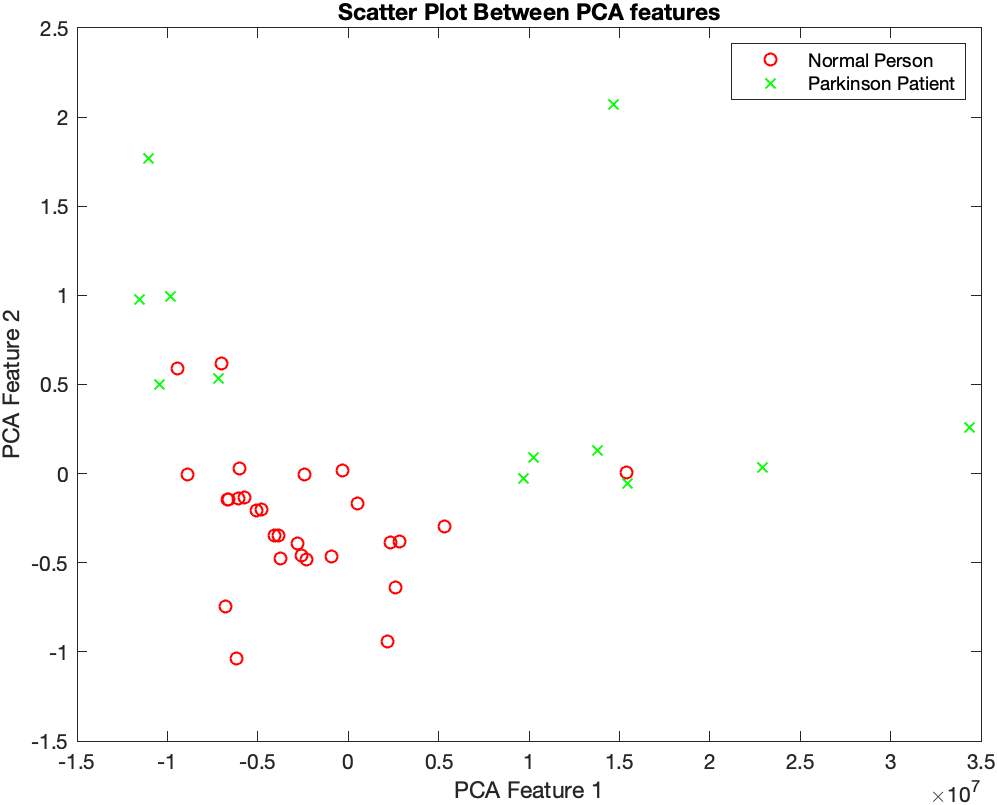


Fig:05

## SVM Binary Classifier:

Though we didn’t have enough voice samples for Parkinson patients, we tried making a binary classifier for distinguishing between the voices of normal people and Parkinson's patients. We used the support vector machine feature of MATLAB and purposefully overfitted our model on the training set to get 100% accuracy on our train set. As we don’t have any more patient samples to test on, we didn’t try to check the accuracy on any kind of test sets.

## Implementation in GUI:

A simple GUI was implemented on app designer, the function of which is to provide a visual representation of the recorded audio signal as well as detecting whether the subject is a patient of Parkinson’s or not.

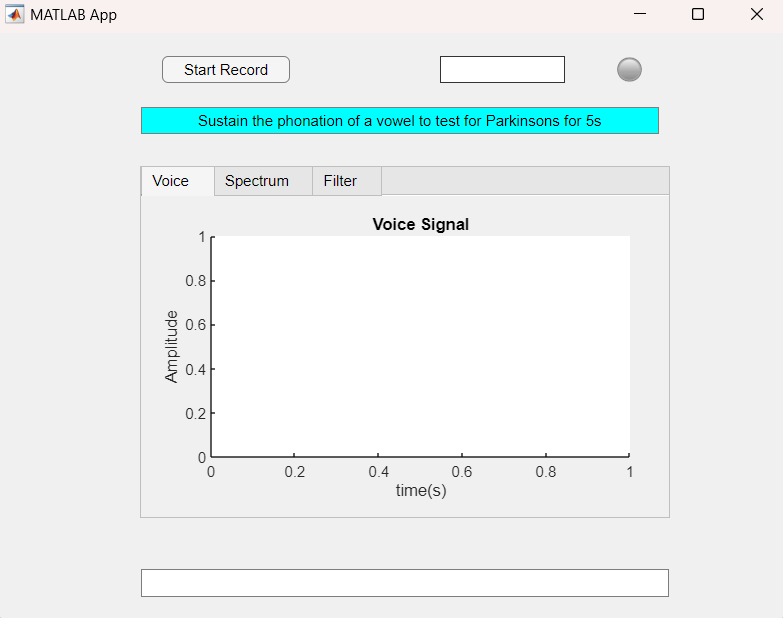
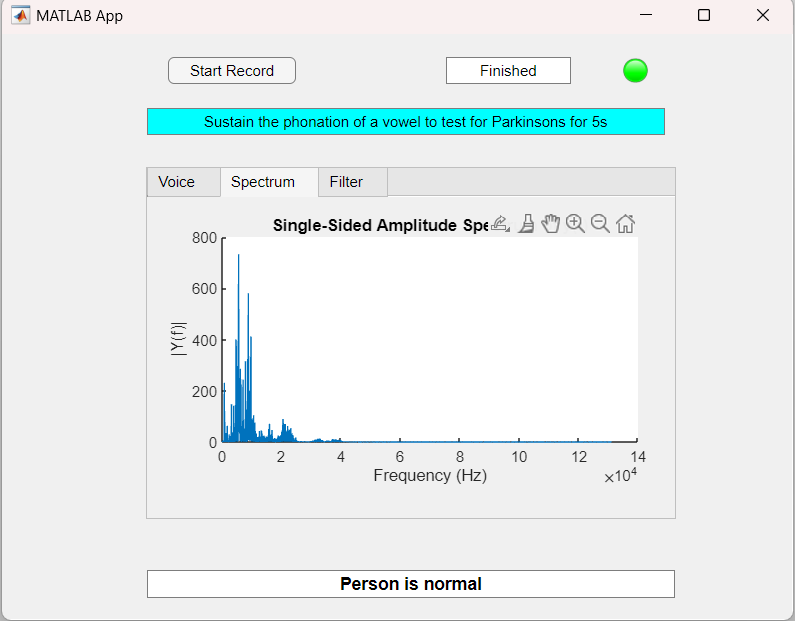
 

Fig:06

# **Scope:**

Our project encompasses the collection of voice data and extraction of features from recorded voice signals. However, the classification accuracy between patients and normal individuals may be limited due to the insufficient number of samples available during training.

# **Limitations:**

Even though our project ran successfully we faced certain limitations. The primary limitation was the lack of patient data. It was really difficult for us to find sufficient data. As a result, our model was overfitted. Another limitation was the lack of proper arrangement to collect the audio signal. We used the built-in microphone of our PCs. a better result could be achieved by recording via sophisticated microphones. Again, we limited our project by extracting only 4 features. If more features were extracted, better results would be obtained.

# **Discussion:**

This project enabled us to explore the domain of signal processing. During the brainstorming, we tried to work with various signals like EEG, ECG, voice signals etc which enabled us to come up with this project. Besides, we studied various journal articles to know the extraction process of various features. In this project we tried to identify Parkinson's disease affected people based on the voice signals. Parkinson’s disease can not be cured, but can be managed if detected early. We tried to develop a user-friendly interface which will enable people to check if they are affected by this disease. The working process of the project is simple. The GUI interface takes the recording of the user, compares with the model we trained prior by extracting features from samples of affected and healthy people, and gives the output if the user is healthy or not by binary classification. We failed to significantly differentiate the patient and healthy patient because of lack of data. Even though we extracted four features distinctly, this project helped us learn various new things in the domain of signal processing which we believe will help us in our future work.

# **Reference:**

[[1] João Paulo Teixeira, Carla Oliveira, Carla Lopes. Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. Procedia Technology, Volume 9, 2013](https://www.sciencedirect.com/science/article/pii/S2212017313002788)