

# Speech Recognition System Report

In this task I designed a simple speech recognition system.

Techniques Used:

1. Linear Predictive Coding(LPC) for feature extraction
2. Curve fitting
3. Gaussian Mixture Model

## Linear Predictive Coding:

Linear predictive coding (LPC) is a method used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model. (Referenced from Wikipedia).

I used inbuilt lpc function from MATLAB to calculate coefficients.

<https://in.mathworks.com/help/signal/ref/lpc.html>

## Curve fitting:

Used curve fitting for scenario 4 to fit pitch of speech and gender (0 for male and 1 for female).

[Fit curve or surface to data - MATLAB fit - MathWorks India](#)

## Gaussian Mixture Model:

Gaussian mixture models are a probabilistic model for representing normally distributed subpopulations within an overall population. Mixture models in general don't require knowing which subpopulation a data point belongs to, allowing the model to learn the subpopulations automatically.

I used gmdistribution.fit function.

[Create Gaussian mixture model - MATLAB - MathWorks India](#)

Use this command to get documentation doc gmdistribution.fit

### 1. Scenario 1 'Yes' or 'No' recognition system:

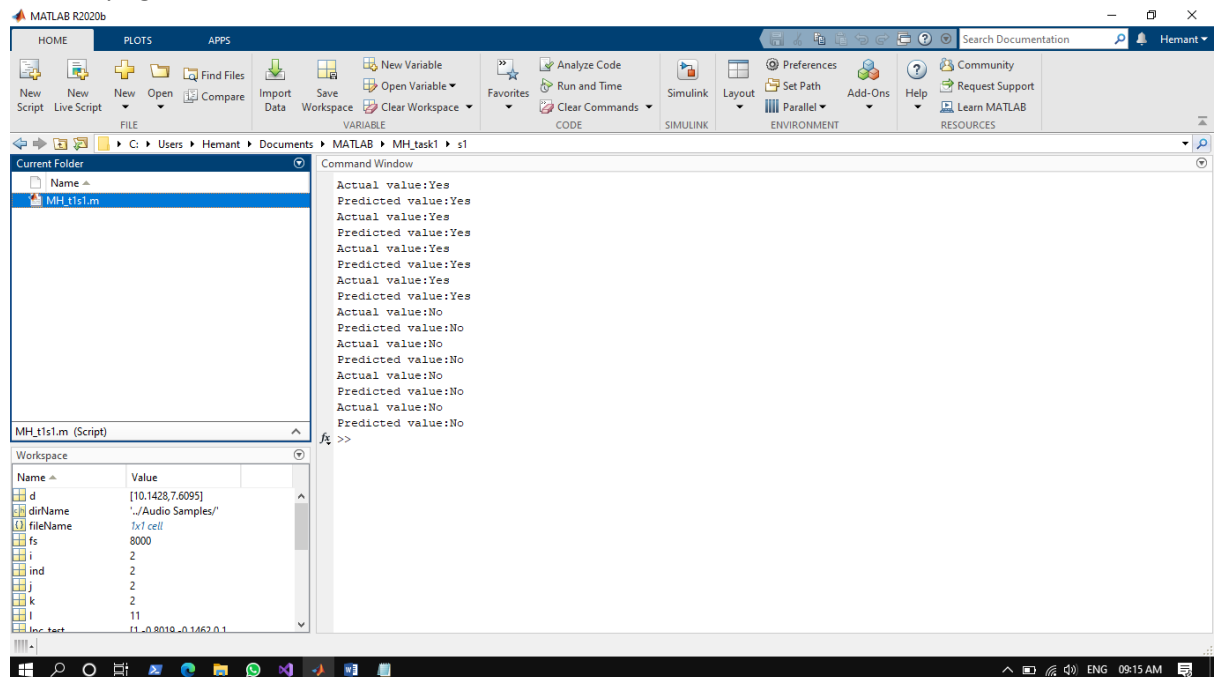
Used 80/20 split for training and testing data. For each unique situation we have 10 audio samples. I used 4<sup>th</sup> and 8<sup>th</sup> sample for testing and others for training.

Steps:

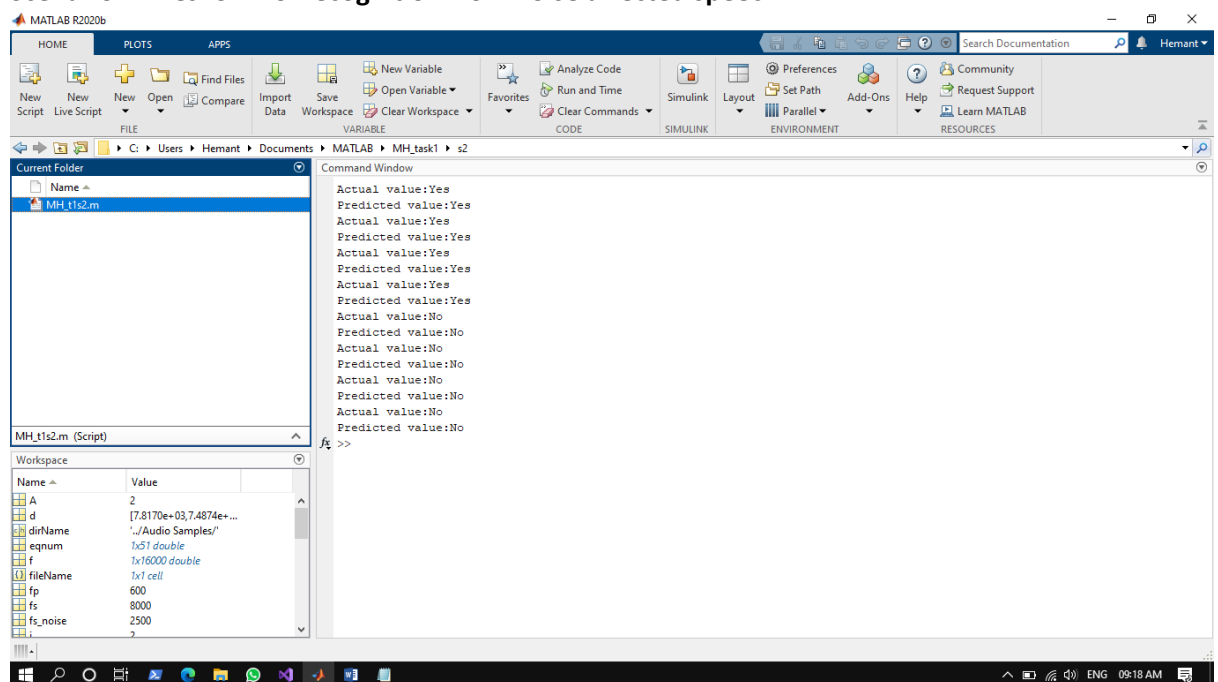
- a. Cleaning
- b. Setup
- c. Reading audio files, doing basic speech Threshold
- d. Create matrix of lpc coefficients
- e. Gaussian modelling

f. Predict

Accuracy I got is 100%.



## 2. Scenario 2 'Yes' or 'No' recognition from noise affected speech



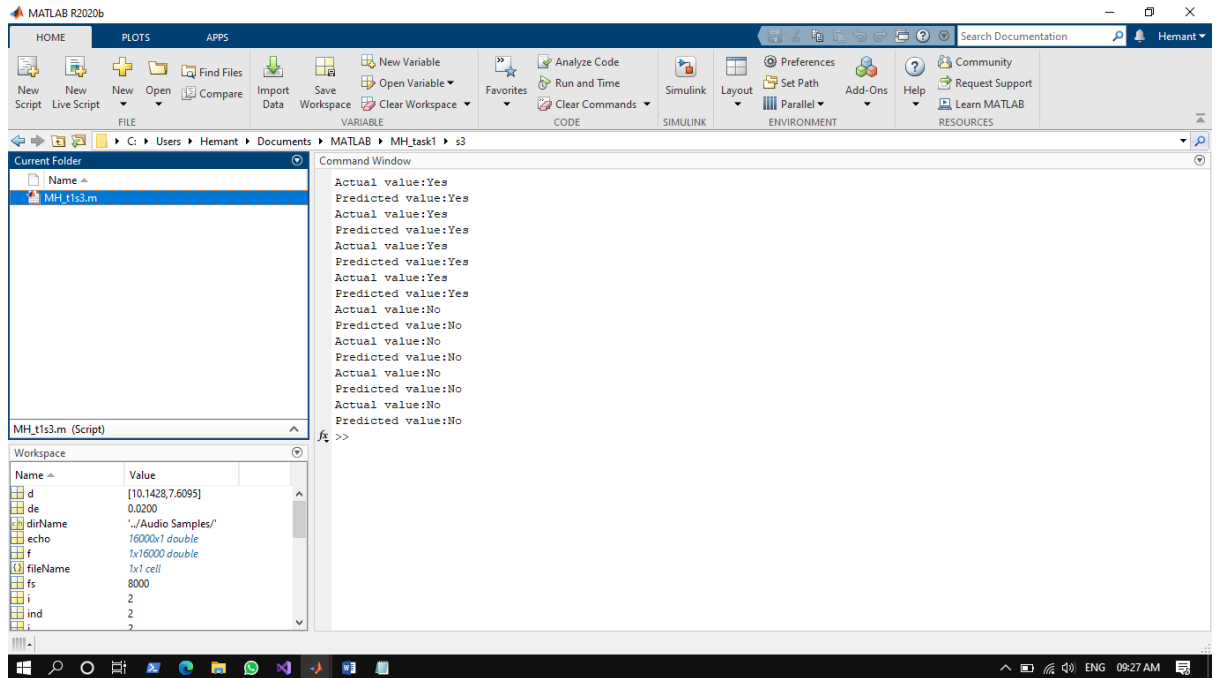
Accuracy is 100%.

Steps:

- Cleaning
- Setup
- Created single tone noise signal of 2500 Hz frequency

- d. Reading audio files
- e. Applying noise filter
- f. Create matrix of lpc coefficients
- g. Gaussian modelling
- h. Predict

### 3. Scenario 3 'Yes' or 'No' recognition from echo affected speech

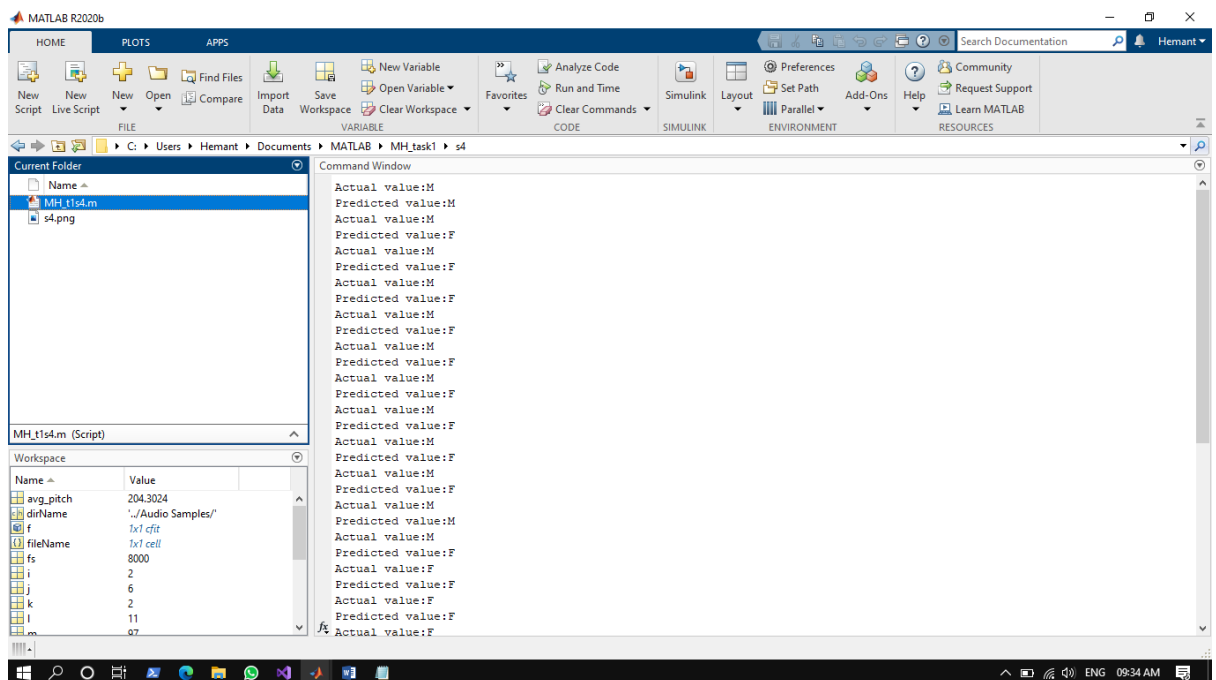
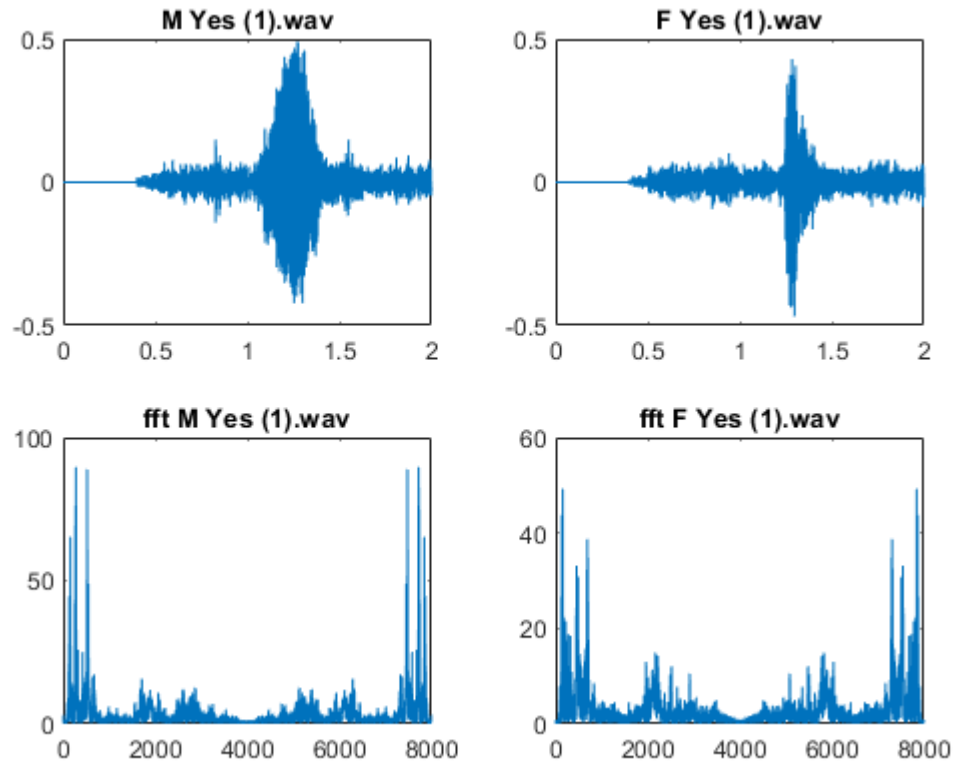


Accuracy is 100%.

Steps:

- a. Cleaning
- b. Setup
- c. Created echo signal with delay of 20 ms
- d. Reading audio files
- e. Applying echo to signal
- f. Filtering echo (reverse system)
- g. Create matrix of lpc coefficients
- h. Gaussian modelling
- i. Predict

### 4. Scenario 4 Distinguish between male and female speech



Accuracy is 62.5%.

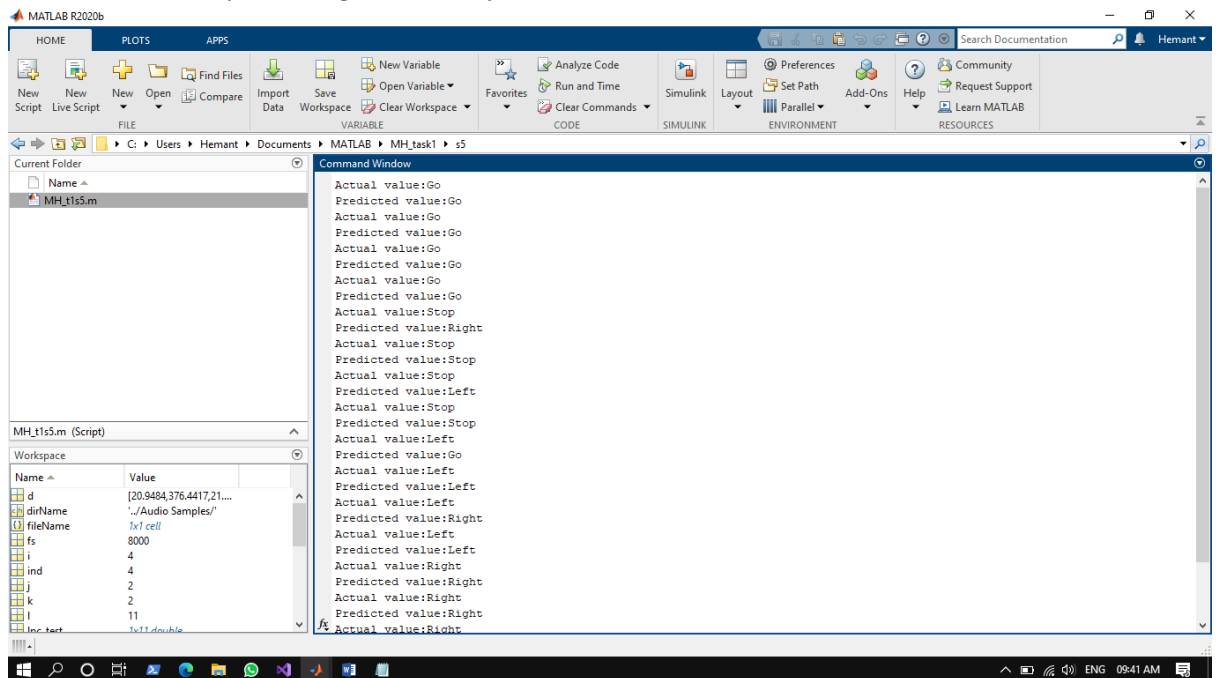
Steps:

- Cleaning
- Setup
- Reading audio files
- Creating matrix for training data (1<sup>st</sup> column is for pitch of speech and 2<sup>nd</sup> column is for male or female label 0 for male 1 for female)

- e. Curve fitting
- f. Predict

## 5. Scenario 4 speech recognition for four words

Words are Go, Stop, Left, Right. Accuracy is 68.7%



Steps:

- a. Cleaning
- b. Setup
- c. Reading audio files, doing basic speech Threshold
- d. Create matrix of lpc coefficients
- e. Gaussian modelling
- f. Predict