LOKMANYA TILAK COLLEGE OF ENGINEERING

Sector No. 4., Vikas Nagar, Koparkhairane, Navi Mumbai -400 709.

Department of Computer Engineering

A. Y. 2020-21

Mid- Semester Progress Report for Major Project-I

Title of the Project- Speech Emotion Recognition System

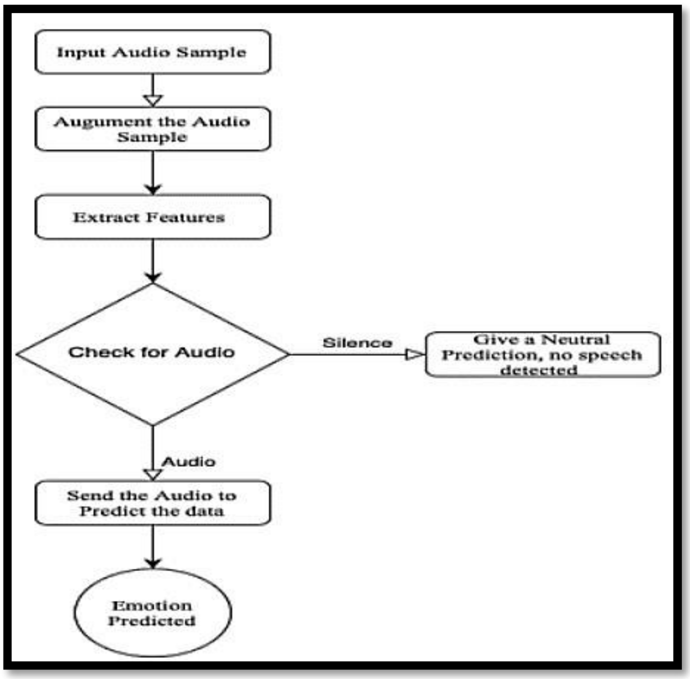
Name of the Guide- Prof. Sonal Bankar

Abstract – Speech emotion recognition is one of the latest challenges in speech processing. Emotion recognition from the speaker's speech is very difficult because of the following reasons: In differentiating between various emotions which particular speech features are more useful is not clear. Because of the existence of the different sentences, speakers, speaking styles, speaking rates accosting variability was introduced, because of which speech features get directly affected. In this Python Project, we will use the libraries librosa, sound file, and sklearn (among others) to build a model using a MLP Classifier MLP Classifier stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network.

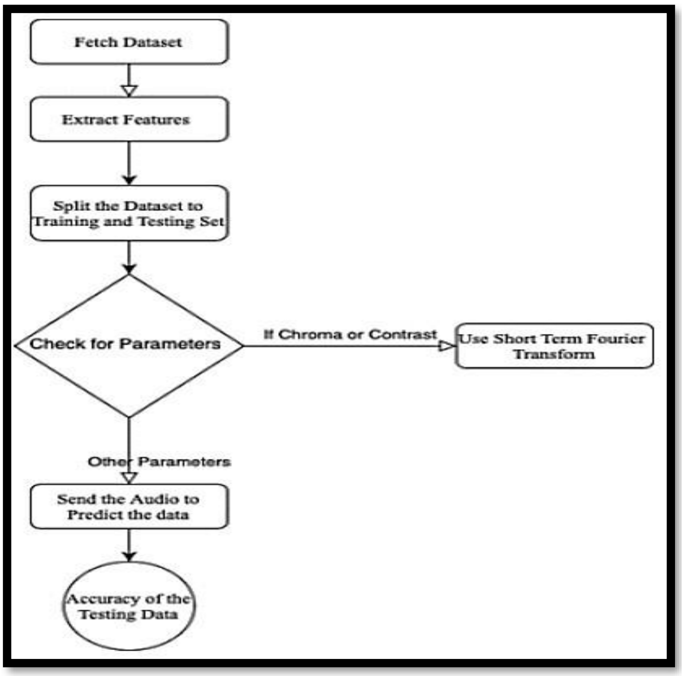
Requirement Analysis- Jupyter Notebook for implementation and RAVDESS Dataset (has voice records of 24 different people)

Analysis & Design-

Training Process Workflow


Input Audio Sample
Augument the Audio Sample
Extract Features
Check for Audio
Silence
Give a Neutral Prediction, no speech detected
Audio
Send the Audio to Predict the data
Emotion Predicted

Testing Process Workflow


Fetch Dataset
Extract Features
Split the Dataset to Training and Testing Set
Check for Parameters
If Chroma or Contrast
Use Short Term Fourier Transform
Other Parameters
Send the Audio to Predict the data
Accuracy of the Testing Data

Algorithm:

1. Make the necessary imports
2. Define a function extract_feature to extract the mfcc, chroma, and mel features from a sound file. This function takes 4 parameters- the file name and three Boolean parameters for the three features:

   i. mfcc: Mel Frequency Cepstral Coefficient, represents the short-term power spectrum of a sound
   ii. chroma: Pertains to the 12 different pitch classes iii. mel: Mel Spectrogram Frequency

3. Define a dictionary to hold numbers and the emotions available in the RAVDESS dataset, and a list to hold those we want- calm, happy, fearful, disgust.
4. Now, load the data with a function load_ data()
5. split the dataset into training and testing sets
6. Observe the shape of the training and testing data. And get the number of features extracted.
7. Now, initialize an MLPClassifier. This is a Multi-layer Perceptron Classifier; it optimizes the log-loss function using LBFGS or stochastic gradient descent.
8. Fit/train the model.
9. To calculate the accuracy of our model, call up the accuracy_score() function imported from sklearn

Future Plan –

- Splitting the data for training (75%) and testing (25%)
- Initializing the MLP classifier
- Training and testing data
- Calculate the accuracy of the model

Progress Timeline-

Duration	Planned Task	Remarks (To be filled by the Project Guide) (Complete/Pending)
Date of Synopsis submission to Presentation I (Week of 20th Aug)	Synopsis Submission Date – 23 July 2020 Presentation Date – 20 August 2020	Completed
31 Aug- 04 Sept	Deciding, Studying And Downloading Data Set	Completed
07 Sept-12 Sept	Studying Of Algorithm	Completed

14 Sept-19 Sept	Installation of Jupyter Notebook,Imported libraries such as Librosa, Soundfile, os, glob, pickle and numpy
	from sklearn.model_selection import train_test_split
	from sklearn.neural_network import MLPClassifier
	from sklearn.metrics import accuracy_score
		Completed

21 Sept - 26 Sept	Loading Dataset ,Deciding Features and Emotions	Completed
28 Sept - 03 Oct	Feature Extraction	Pending
	Overall Remark by the Guide (Good/ Satisfactory/ Unsatisfactory)	

Signature of the Guide