

The background is a dark blue-grey color. It features several thin, gold-colored lines that form abstract, angular shapes. These lines radiate from the central text box, extending towards the corners of the frame. Some lines form sharp, pointed shapes, while others are more horizontal or vertical, creating a dynamic, geometric pattern.

wave

EMOTIONS FROM SPEECH

CSD402 | Internet and Web Systems

Problem Statement

Emotions are fundamental for humans, impacting perception and everyday activities such as communication, learning, and decision-making.

They are expressed through speech, facial expressions, gestures, and other non-verbal clues.

Speech forms the underlying method of communication and expression for humans

To be able to interpret and detect human emotions through speech analysis and further analyse the use cases for the technology across sectors.

SOLUTION

An application that detects and displays human emotions so detected from an input audio stream.

The characterisation of emotions is done based on various elements of the speech - tone, pitch etc.




WAVE



Alexithymia Aid

A disorder which causes one to not be able to detect emotions, the application case serve as an aid to those suffering from the disorder



Analysis

- A web app that allows the user to record and detect emotion based on their speech. The algorithm analyses the recording using an MLP classifier. This is a Multi-layer Perceptron that optimizes the log function using stochastic gradient descent. MLP Classifier has an internal neural network for the purpose of classification
- The ensure the accuracy of the algorithm, we train and test it using the Ravedess Dataset, which has 7356 files rated by 247 individuals 10 times on emotional validity, intensity, and genuineness. The entire dataset is 24.8GB from 24 actors.

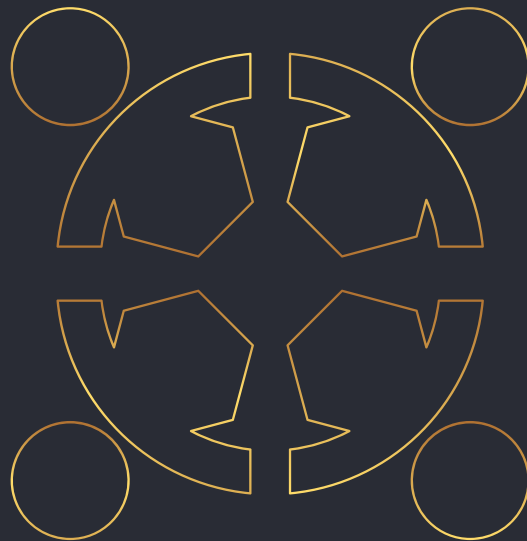
Tech Stack

Python

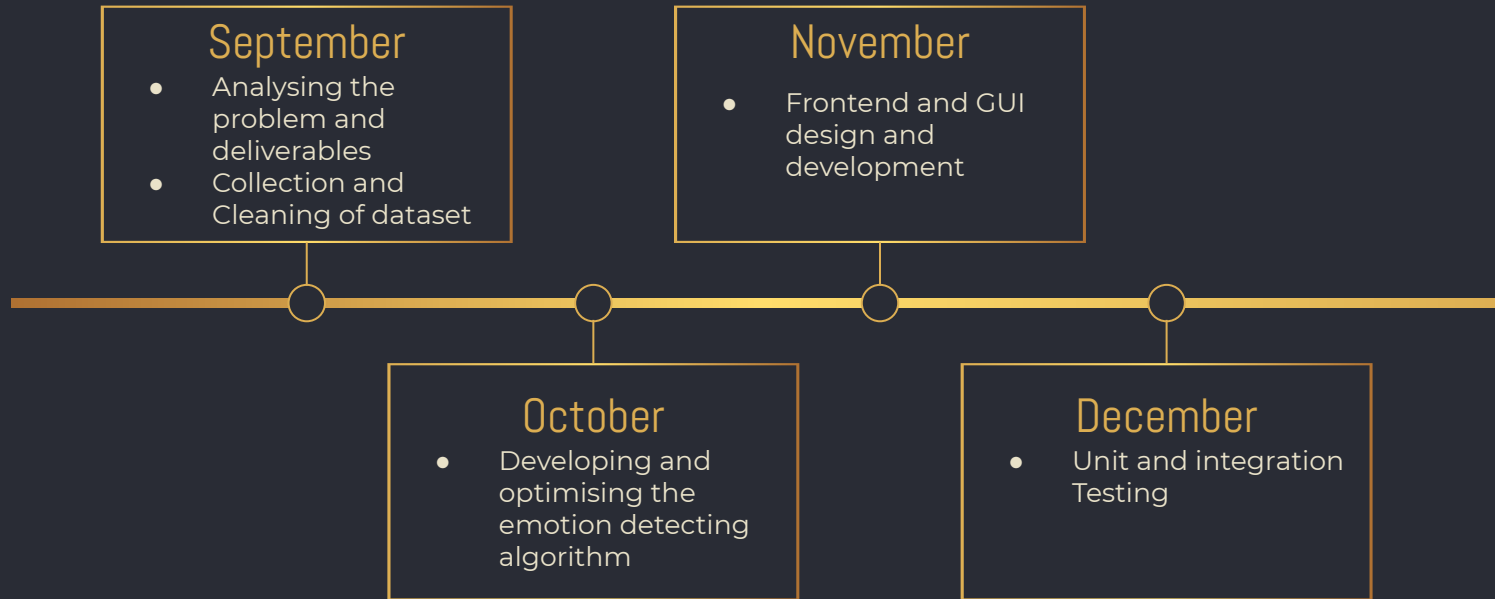
MERN Stack

Heroku

RAVDESS



TIMELINE



Current Progress

File Explorer view showing the directory structure of the dataset:

Path: This PC > LENOVO (D:) > 4th year > IWS > project > speech-emotion-recognition-ravdess-data

Name	Date modified	Type	Size
Actor_01	27-09-2021 23:57	File folder	
Actor_02	27-09-2021 23:57	File folder	
Actor_03	27-09-2021 23:57	File folder	
Actor_04	27-09-2021 23:57	File folder	
Actor_05	27-09-2021 23:57	File folder	
Actor_06	27-09-2021 23:58	File folder	
Actor_07	27-09-2021 23:58	File folder	
Actor_08	27-09-2021 23:58	File folder	
Actor_09	27-09-2021 23:58	File folder	
Actor_10	27-09-2021 23:58	File folder	
Actor_11	27-09-2021 23:58	File folder	
Actor_12	27-09-2021 23:58	File folder	
Actor_13	27-09-2021 23:58	File folder	
Actor_14	27-09-2021 23:58	File folder	
Actor_15	27-09-2021 23:58	File folder	
Actor_16	27-09-2021 23:58	File folder	
Actor_17	27-09-2021 23:58	File folder	
Actor_18	27-09-2021 23:58	File folder	
Actor_19	27-09-2021 23:58	File folder	
Actor_20	27-09-2021 23:58	File folder	
Actor_21	27-09-2021 23:58	File folder	
Actor_22	27-09-2021 23:58	File folder	
Actor_23	27-09-2021 23:58	File folder	
Actor_24	27-09-2021 23:58	File folder	

Collected and
Analysed Ravdess
Dataset

File Explorer view showing the contents of the Actor_01 folder:

Path: This PC > LENOVO (D:) > 4th year > IWS > project > speech-emotion-recognition-ravdess-data > Actor_01

Name	#	Title	Contributing artists	Album
03-01-01-01-01-01-01.wav				
03-01-01-01-01-02-01.wav				
03-01-01-01-02-01-01.wav				
03-01-01-01-02-02-01.wav				
03-01-02-01-01-01-01.wav				
03-01-02-01-02-01-01.wav				
03-01-02-01-02-01-01.wav				
03-01-02-01-02-01-01.wav				
03-01-02-01-02-01-01.wav				
03-01-02-01-02-01-01.wav				
03-01-02-02-01-01-01.wav				
03-01-02-02-02-01-01.wav				
03-01-03-01-01-01-01.wav				
03-01-03-01-02-01-01.wav				
03-01-03-01-02-01-01.wav				
03-01-03-01-02-01-01.wav				
03-01-03-02-01-02-01.wav				
03-01-03-02-02-01-01.wav				
03-01-03-02-02-01-01.wav				
03-01-04-01-01-01-01.wav				
03-01-04-01-01-02-01.wav				
03-01-04-01-02-01-01.wav				
03-01-04-01-02-02-01.wav				
03-01-04-02-01-01-01.wav				
03-01-04-02-01-02-01.wav				
03-01-04-02-01-01-01.wav				
03-01-04-02-02-01-01.wav				
03-01-05-01-01-01-01.wav				

Extracted Features from the dataset

```

1 import librosa
2 import soundfile
3 import os, glob, pickle
4 import numpy as np
5
6 def extract_feature(file_name, mfcc, chroma, mel):
7     with soundfile.SoundFile(file_name) as sound_file:
8         X = sound_file.read(dtype='float32')
9         sample_rate=sound_file.samplerate
10        if chroma:
11            stft=np.abs(librosa.stft(X))
12            result=np.array([])
13            if mfcc:
14                mfccs=np.mean(librosa.feature.mfcc(y=X, sr=sample_rate, n_mfcc=40).T, axis=0)
15                result=np.hstack((result, mfccs))
16            if chroma:
17                chroma=np.mean(librosa.feature.chroma_stft(S=stft, sr=sample_rate).T, axis=0)
18                result=np.hstack((result, chroma))
19            if mel:
20                mel=np.mean(librosa.feature.melspectrogram(X, sr=sample_rate).T, axis=0)
21                result=np.hstack((result, mel))
22        return result
23
24 #@function describing emotions in the dataset
25 emotions={}
26
27 '01': 'neutral',
28 '02': 'calm',
29 '03': 'happy',
30 '04': 'sad',
31 '05': 'angry',
32 '06': 'fearful',
33 '07': 'disgust',
34 '08': 'surprised'
35 }
36
37 #sample emotions to observe
38 observed_emotions=['calm', 'happy', 'fearful', 'disgust']
39
40 x,y=[],[]
41 for file in glob.glob(D:\4th year\DS\project\speech-emotion-recognition-rawdata\data\Actor\1\*.wav"):
42     file_name=os.path.basename(file)
43     #print(file_name.split(".")[-1])
44     emotion=emotions[file_name.split(".")[-1]]
45     if emotion not in observed_emotions:
46         continue
47     feature=extract_feature(file, mfcc=True, chroma=True, mel=True)
48     x.append(feature)
49     y.append(emotion)
50
51 print('Extracted Emotions from files: ')
52 print(y)

```

Roles

Sparsh Ailawadi

Backend Developer

Sajal Ganjewala

Full Stack Developer

Savita G

Frontend And
Application Design

Future Scope

- The algorithm will be optimised to provide real-time speech to emotion detection.
- Incorporating more types of emotions by collecting data from recordings.