**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Team Member: 1. Kartik Pandey  2. Aniket Nichat  3. Rohit Thawali  4. Sagar Khekale  Email: 1. [kartikpande12@gmail.com](mailto:kartikpande12@gmail.com)  2. [vrushabhnichat@gmail.com](mailto:vrushabhnichat@gmail.com)  3. [rohitthawali25@gmail.com](mailto:rohitthawali25@gmail.com)  4. sagarkhekale2@gmail.com  Contribution :  Kartik Pandey:   1. Worked on EDA 2. Count of Emotion 3. Compare the MFCC feature for male and female angry audio clips 4. Worked on GRU 5. Deploy Model CNN   Aniket Nichat:   1. Worked on Data Augmentation 2. Noise added in Audio 3. Stretched Audio 4. Shift Audio 5. Worked on LSTM   Rohit Thawali:   1. Worked on Feature Extraction 2. Data Preprocessing 3. Collect Dataset from Kaggle 4. Worked on KNN   Sagar Khekale :   1. Worked on Deploy model 2. Worked on MLP Classifier 3. Work on Confusion Matrix 4. Worked on DecisionTree |
| **Please paste the GitHub Repo link.** |
| https://github.com/sagar234/Speech-Emotion-Recogination |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches, and your conclusions. (200-400 words)** |
| Summary:  Speech Emotion Recognition mean to detect the speech someone who is in front of you and now day we are detect using artificial intelligent and combine with neural network . using various technic like DECISSION TREE ,LSTM,GRU ,MLP,KNN ,CNN . this technic give us a very good accuracy and also a very good model for train our data set , in this project, we divide this project into four-part First EDA, Data Augmentation, Feature Extraction, Model.  First EDA in this Technique we get an insight of data information, using EDA The key feature was  We use MFCC(Mel Frequency Cepstral Coefficients), Mel Spectrogram.  MFCC –  Picture comes about for mfcc show in discourse feeling recognition The MEL Frequency Septral Coefficient (MFCC) was initially recommended to recognize monosyllabic words in reliably talked sentences but not for speaker acknowledgment. MFCC is utilized to distinguish carrier reservations, phone numbers, and voice acknowledgment frameworks for security purposes.  Mel Spectrogram :  A Quick Fourier Change is computed on covering windowed fragments of the flag and we get what is called the spectrogram. This can be fair a spectrogram that portrays sufficiency mapped on a Mel scale.  Now, The Second Part is Data Augmentation .  Data augmentation is the process by which we create new synthetic data samples by adding small disturbance to make our data robust and also use for to reduce the overfeating problem .  To generate syntactic data for audio, we can apply noise , shifting time, changing pitch, and stretch the audio .  Now, Feature Extraction  In Feature Extraction of the main steps in speech processing is the extraction of features and production of parameters from the speech signal which are generally related to the short-term spectrum of a speech signal or to the shapes of the vocal tract. Feature extraction is performed to focus on the information contained in the signal, improve the degree of similarity and dissimilarity between different classes, and reduce the dimension of data and calculations .While extracted features from each pattern play an effective role in classification, this study attempts to improve the efficiency of speech emotion recognition based on the proposed feature extraction method .  Now, Model  In Model after we ran the CNN model, we get model3.h5 file and we store it and we used into the testing purpose for web application  Problem Statement :  Verbal Communication is profitable and looked for after in working environment and classroom situations alike. There's no denying the idea that Indians need verbal communication and consequently lag behind within the work environment or classroom situations. This happens in spite of them having solid specialized competencies. Clear and comprehensive discourse is the imperative spine of solid communication and introduction skill.  Approach :  My approach towards the Project First, I develop the model because this is capitalizing on the fact that voice often reflects underlying emotion through tone and pitch SER is tough because emotions are subjective and annotating audio is challenging then I work on MLP classifier the MLP -classifier is used to classify the emotion from the given wave signal, which chooses learning rate to be adaptive. in the end, I work on the confusion matrix in the confusion matrix we compare the actual label and predicted label so the female clam is going too high in the confusion matrix.  Conclusion:  That's it! We come to the conclusion of our exercise. Starting with stacking the information so distant we have done EDA, Data augmentation , Feature Extraction, and after that model building. we appeared how we will use Machine learning to get the basic feeling from speech sound information and a few bits of knowledge on the human expression of feeling through voice. This framework can be utilized in a assortment of setups like Call Middle for complaints or promoting, in voice-based virtual collaborators or chatbots, in phonetic inquire about, etc.  Drive Link: https://drive.google.com/drive/folders/1ZvA-iRUGot\_9CjeZTBDjQerCPvRD59gN?usp=sharing  . |
|  |