Speech Emotion Recognition

Speech Emotion Recognition is of identifying emotion based on human speech.

It is used in emotion-aware entertainment and education by creating adaptive and personalized content based on the user's mood and preferences. Assists in forensic and security investigations by identifying emotional states of suspects or witnesses. Supports in mental health and well being detecting signs of stress, depression, or anxiety in speech and providing interventions. Improves human-machine interaction.

At first we train our CNN(Convolutional Neural Network) model with Ravdess dataset and save the model, also store the weights of the model. Now load the model with weights in to gui for detecting emotion for his/her speech.

Dataset link: RAVDESS Emotional speech audio (kaggle.com)

Here we can learn recording audio and storing in a file. Passing the audio file to our model to detect emotion based on pitch, shift. We also remove noise in the audio for better detection.

Here we extract features for our audio and pass them to our model, it returns different values which are not understandable, those are transformed in to useful information by using LabelEncoder, which results in emotion of audio.