Emotion Recognition using Speech Signals

S. Harsha Vardhan[1], M. P. Rahul[2], P. Kavyasri[3], A. Sraavani[4],

B. Tech Students, Department of Information Technology[1,2,3,4]

Prasad V Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

Abstract: *Communication is the key to express one's thoughts and ideas. Among all the forms of communication, the most preferred form of communication is speech. In the era of Intelligent systems, the computer human interaction plays a major role in the functionality. This python project uses the libraries present to analyse the audio files. This project gives an overview of the deep learning techniques which are based on feature extraction and model creation which recognizes the emotion of the speaker. This project was made more efficient by the usage of data, efficient methods for feature extraction and classification. The existing models are capable of analysing only the three emotions namely Happy, Angry and Neutral, using this project we can identify Eight different emotions namely Sad, Happy, Calm, Angry, Fearful, Disgust, Surprised and Neutral. The emotion detection through speech improves the functionality of the Intelligent Systems.*

Keywords: Sound File, NumPy, Librosa, Sklearn, MFCC, MEL, CHROME, MLP Classifier, Confusion Matrix.

REFERENCES

[1]. https://colab.research.google.com/drive/1pXYDKsBVenXQhiFLrII4tv4Wp46LqMl4#scroll To=5I6CDcGoCzG6

[2]. https://www.kaggle.com/datasets/uwrfkaggler/ravdess-emotional-speech-audio

[3]. https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.htm l

[4]. https://www.analyticsvidhya.com/blog/2021/06/confusion-matrix-for-multi-class-classificati on/#:~:text=Confusion%20Matrix%20is%20used%20to,number%20of%20classes%20or%20 outputs.

[5]. https://www.analyticsvidhya.com/blog/2021/06/mfcc-technique-for-speech-recognition/

[6]. https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53

[7]. https://www.ee.columbia.edu/~dpwe/resources/matlab/chroma-ansyn/

[8]. https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance

[9]. -measures/

[10]. https://www.sharpsightlabs.com/blog/scikit-train_test_split/

[11]. https://pysoundfile.readthedocs.io/en/latest/

[12]. https://www.javatpoint.com/librosa-library-in-python

[13]. https://www.w3schools.com/python/numpy/numpy_intro.asp

[14]. https://www.sciencedirect.com/science/article/pii/S1746809418302337

[15]. H. Cao, R. Verma, and A. Nenkova, "Speaker-sensitive emotion recognition via ranking: Studies on acted and spontaneous speech," Comput. Speech Lang., vol. 28, no. 1, pp. 186–202, Jan. 2015.

[16]. L. Chen, X. Mao, Y. Xue, and L. L. Cheng, "Speech emotion recognition: Features and classification models," Digit. Signal Process., vol. 22, no. 6, pp. 1154–1160, Dec. 2012.

[17]. T. L. Nwe, S. W. Foo, and L. C. De Silva, "Speech emotion recognition using hidden Markov models," Speech Commun., vol. 41, no. 4, pp. 603–623, Nov. 2003.