International Journal of Latest Trends in Engineering and Technology
Vol.(17)Issue(3), pp.041-044
DOI: http://dx.doi.org/10.21172/1.173.06
e-ISSN:2278-621X

STRESS DETECTION FROM SPEECH SIGNAL USING MFCC, SVM AND MACHINE LEARNING TECHNIQUES

Dr. Bageshree Pathak[1], Chinmayi Dhole[1], Harshada Hajare[1], Mrunal Zambare[1]

Abstract- Due to overtime and frequent rotating shifts, police officers experience stress. This results in poor mental and physical health. Central Police Research (CPR), in collaboration with Cummins College of Engineering for Women, Pune and SNDT Arts & Commerce College will be developing a Stress Detection System from Speech. Feature extraction techniques like Mel Frequency Cepstral Coefficients (MFCCs) and classifiers like Support Vector Machine (SVM) are implemented in this project for detecting stress in speech.

Keywords: Machine Learning, Feature Extraction, MFCC, SVM

I.INTRODUCTION

Due to general work sector problems such as long working hours, heavy workload, working under dangerous conditions especially for police officers dealing with crime and being unhappy with poor work management which leads to extra duties and sometimes physical and mental harassment at work are faced by the people.
In this project, we have developed a system that can recognize whether an individual is in stress or non-stress, given audio with various techniques like speech signal processing, machine learning, human psychology, and many more.
Finally, we have implemented the SVM classification technique and MFCC feature extraction technique separately for speech. The pattern recognition system offers speech stress detection as an application in which patterns of derived speech feature such as MFCC. MFCC is a representation of short term power spectrum of a sound and SVM is supervised learning algorithm that separates output in two classes as stressed and non-stressed.

II.LITERATURE SURVEY

[1]researched on Stress Detection from Speech Using Spectral Slope Measurement. In that convolution of the impulse response of vocal track with glottal source excitation signal is used. SWIPE algorithm for feature selection is implemented. The final output results into stressed and neutral PDF curves.
[2]states that Detection of Depression in Speech with Multi Classifier System, in that speech Signal processed with following feature extraction technique: Linguistic and Acoustic, Energy, Intensity, Loudness, Jitter, HNNNR, MFCC and classified with the help of Support Vector Machine.
[3]researched on Detection of Adolescent Depression from Speech using Optimized Spectral Roll-Off Parameters. Data is classified using ORI-DB data set, feature extraction technique as Spectral Category and SVM classifier.
[4]has researched on Detection through Emotional Speech Analysis and Study of design of a classification system of stress levels using emotional speech. Linear discrimination combined with bootstrapping technique to implement classifiers. Error rates are lower than 33%.

[1] *Dept. of Electronics and Telecommunications, MKSSS's Cummins College of Engineering for Women Pune Maharashtra-India*

[5]researched on Automatic Stress Detection from Speech. The database contains 400 isolated spoken Malayalam words. They have used Discrete Wavelet Transform feature extraction and Artificial Neural Network (ANN) for training and testing purpose. The system has 85% accuracy.

[6]describes about stress detection with speech signal based on deep learning algorithm. In that, the algorithm extracts MFCC features using pre-processed speech signal and predicts the output using binary decision making criteria using Long-Short Term Memory and feed forward networks those results into 66.4% of accuracy.

III.DATABASE GENERATION

Database is generated in two ways for this project:

3.1. Database generated at CPR:

CPR department in collaboration with Cummins College of Engineering and SNDT Arts & Commerce College, have given a problem statement for detection of stress in speech. We visited to training center of police to collect speech samples. Around 50 voice samples we have collected from officers and for that we have provided Google form to collect their voice record and some questionnaires. And based on their answers, Psychology department of SNDT College validated the stressed and non-stressed speeches. This validation is done for better understanding of the database and to verify result after getting an output as stressed or non-stressed.

3.2. Database by Recording Samples:

We have collected samples of stressed and non-stressed from YouTube videos. By considering the recent incidents such as, nationwide pandemic situation of COVID-19, Acid attack survivors speech, other judicial victims of various cases etc., from such situational videos stressed samples are collected and for non-stressed we have collected samples of family members, relatives during when they are in relax and happy i.e., non stressed phase.

IV.METHODOLOGY

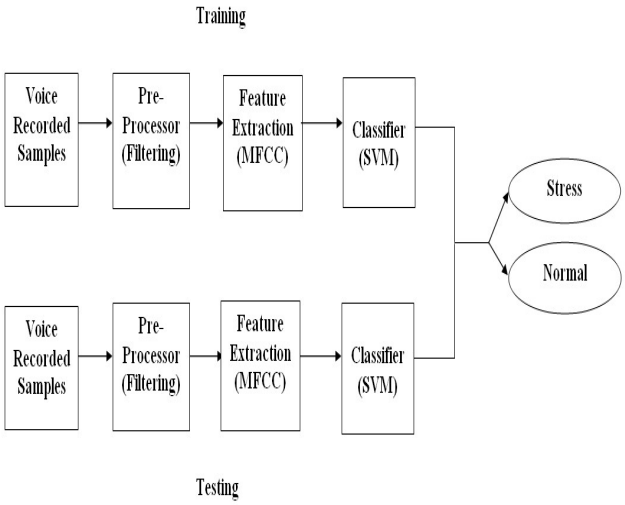

Figure-1Block Diagram

Figure-1 shows the block diagram of the whole system with training and testing phase.

4.1. Signal Pre-processing:

In signal processing pre processing of speech is a more important stage in the development of an automatic speech recognition system.

Here in this step noise removal technique is used which afterward results in the plot of the original signal. For that Librosa library is imported from python.

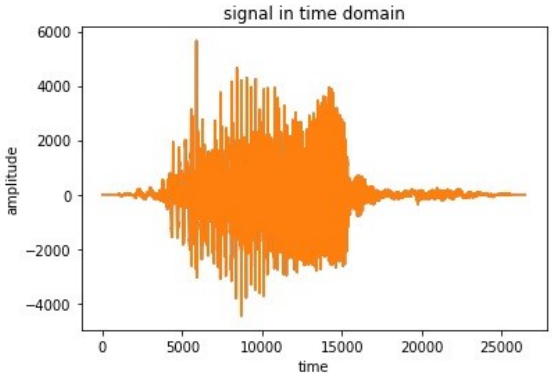

Figure-2 Original Signal

Figure-2 shows the original speech signal where on X-axis shows the time in milliseconds and on Y-axis there is amplitude.

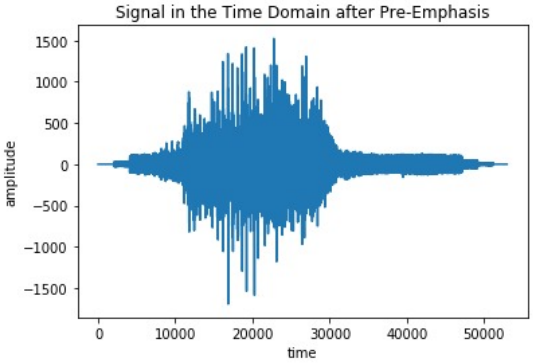

Figure-3 Pre processed Signal

Figure-3 shows the pre processed signal of figure-2 with the default sampling rate of 22.05 kHz. This flattens the audio channel to mono.

4.2. Feature Extraction:

In feature extraction initially the raw data is reduced to a more feasible data for processing.
So for that, we have used Mel-frequency Cepstral Coefficients (MFCC) feature extraction technique.
Representation of Power Spectrum of a sound is called as MFC (Mel Frequency Cepstrum).

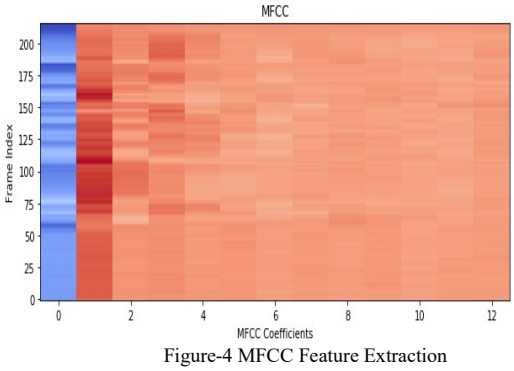

Figure-4 MFCC Feature Extraction

Figure-4 shows the plot of MFCC features as on X-axis there are MFCC coefficients and on Y-axis there is frame index which shows number of frames. We have taken 13 MFCC coefficients per frame for our dataset.

4.3. Classification:

We have implemented the system using SVM classifier**.** It uses kernel for transformation of data that helps to find an optimal boundary between possible outputs**,** here we have used gamma kernel because it gives better output as compared to linear kernel.

[1] *Dept. of Electronics and Telecommunications, MKSSS's Cummins College of Engineering for Women Pune Maharashtra-India*

V. RESULTS

Figure-5 shows the loss model of SVM classifier that indicates how bad the model's prediction is. As we have used the SVM classifier it gives 83.4% accuracy. On the X-axis there is an epoch of signal and on Y-axis there is a loss of a model. We have used 100 recorded samples to train the model. We got 83.4% accuracy by using the SVM classifier.

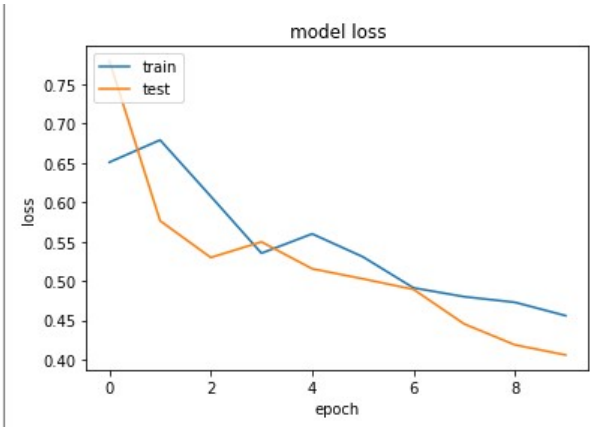

Figure-5 Loss Model

Confusion Matrix from SVM classifier is obtained which describes the performance of classifier using validation of data and it is given in the table-1 given below:

Total samples=100	Stress	Non-stress
Stress	52	7
Non-stress	6	35

Table-1

VI. CONCLUSION

In this project, we have developed an algorithm to detect whether a person is under stress or not.
For this, we used Python 3.7 software & Spyder IDE compiler to implement python code.
We generated database by two means firstly from CPR department's trainee officers and secondly by using media coverage and youtube videos for stressed & non-stressed speech signal that are verified by the psychology department of SNDT College.
We have implemented SVM classifier to train system and final result testing accuracy obtained by SVM classifier is 83.4%. To increase the efficiency of system more training data is required.

REFERENCES

[1] Sharma, N., Gedeon, T.: Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Comput. Methods Programs Biomed. **108**(3), 1287–1301 (2012)
[2] Murray, I.R., Baber, C., South, A.: Towards a definition and working model of stress and its effects on speech. Speech Commun. **20**(1), 3–12 (1996)
[3] Hailiang Long, Xia Wu, Bin Hu**:** Detecting Depression in Speech: A Multi-classifier System with Ensemble Pruning on Kappa-Error Diagram (2017)
[4] Olympia Simantiraki, Giorgos Giannakakis, Anastasia Pampouchidou, Manolis Tsiknakis: Stress Detection from Speech Using Spectral Slope Measurements (2017)
[5] Stress Detection Through Emotional Speech Analysis: INMA MOHINO University of Alcala Signal Theory and Communications 28805 Alcala de Henares ´ SPAIN, ROBERTO GIL-PITA, LORENA ALVAREZ P ´ EREZ (2012)
[6] Sharifa Alghowinem "A Comparative Study of Different Classifiers for Detecting Depression in Speech: Multi Classifier system", IEEE International Conference on Acoustics, Speech and Signal Processing, (2013)
[7] Melisa N Stolar and Margaret Lech and Shannon J Stolar & Nicholas B Allen "Detection of Adolescent depression from speech using optimized spectral Roll-Off parameters", Biomedical Journal (2018)
[8] in Zuo, Tian Li, and Pascale Fung "A Multilingual Database of Natural Stress Emotion", Proceedings of the 8th International Conference on Language Resources and Evaluation(LREC'12), (2012)