Clinical Depression Detection Using Speech Feature With Machine Learning Approach

Ms. Anjum Shaikh, Ms. Firdos Shaikh, Mr. Suhaib Ramzan, Prof. M. M. Patil

Department of Electronics & Telecommunication Engineering, Sinhgad Academy of Engineering, Pune, Maharashtra, India

ABSTRACT

Depression is a general mental health disorder that presents state of low mood, negative thoughts, mental disturbance, typically with lack of energy , difficulty in maintaining concentration, guilty, irritable, restless and cognitive difficulties such as lose interest in different new things. Clinical depression is a major risk factor in suicides and is associated with high mortality rates, therefore making it one of the leading causes of death worldwide every year. The landmark World Health organisation(WHO) Global Burden of Disease (GBD) quantified depression as the second highest leading cause of disability world-wide[1]. It is observed that, there is increase in tendency of clinical depression in adolescents (i.e. age between 13–20 years) has been linked to a range of serious problem, basically an increase in the number of suicide attempts and deaths. This is making public health concern. In this project we are detecting whether the person is in depression or not using tensor flow software. There various biomarkers of depression like facial expressions, speech, pupil, T-body shape, MRI, EEG, etc. Here we are processing on speech feature extracted from database by SVM technique. Again among features of speech like TEO, MFCC, pitch, etc. Here we are extracting MFCC feature of speech from database.

Keywords: Machine Learning, Supervised Learning, Neural Networks, Topic Detection, Natural Language Processing

I. INTRODUCTION

Speech is a natural form of communication for human beings and has been recognized as one of the potential sources in providing cues for depression. Clinical depression has been related to severe affect disturbances of a person's feelings and moods, and can be conceptualized in the failure to regulate emotions [1]. For example, in one of the letters that was written from a parent of a depressed adolescent, it was commented that *"I have always believed that I could pretty accurately identify the episodes and severity of depression just from the sound of her voice.... As a mother, I think I am familiar with every nuance of the sounds my children utter. I can hear the depression"*. This reveals supporting evidence that when someone is depressed or could be depressed but pretends to be well; listeners attend to the tone of the voice rather than the linguistic content to make social judgments of the person's behavioural traits. By paying particular attention to ones tone and word emphasis, the voice of a person during social interactions may expose important information that can express a person's emotional state, moods, attitude and personality. It makes a large difference where the emphasis or stress is placed on certain words. For example, consider the same sentence spoken in various ways by altering the pitch and tonal patterns on how the words are stressed in the following sentence "I would appreciate if you pick up around the house more." By paying close attention to the emphasis placed on the word "appreciate", the tone of the voice could indicate a person being hostile, sarcastic, bitter, pleased, angry, painful, etc. (i.e., a loud voice tone on the word "appreciate" may indicate a person being angry or sarcastic).According to psychologists, the ways these emotions are expressed through a person's voice are one of the communication channels (other channels are face, body and gestures) that clinicians look out for in the

search of possible tell-tale signs in depressive symptoms. Therefore, this study delves into the realm of clinical depression in people to investigate the MFCC property of speech as potential indicators of depression.

II. EASE OF USE

In this project we are detecting whether the person is in depression or not using tensor flow software. There various biomarkers of depression like facial expressions, speech, pupil, T-body shape, MRI, EEG, etc. Here we are processing on speech feature extracted from database available with us by SVM technique. Again among features of speech like TEO, MFCC, pitch, etc. Here we are extracting MFCC feature of speech from database.

III. LITERATURE SURVEY

Kuryati Kipli, Abbas Z. Kouzani researched on depression detection using MRI in a paper named, "Evaluation of Feature Selection Algorithms for Detection of Depression from Brain sMRI Scans".[In this paper[1], the brain structural MRI (sMRI) volumetric features are investigated to determine features that contribute towards more accurate depression detection, whether a person is in depressed or non-depressed forms. It gives accuracy upto 80% but it is costly system[1].

D.P.X. Kan and P.F. Lee researched on depression detection using EEG in a paper named[2], "Decrease Alpha Waves In Depression: An Electroencephalogram (EEG) Study". They stated that, In the present study, electroencephalography (EEG) will be used to monitor the brain waves of subjects. EEG is a device to measure the electrical activity of the brain.. They got the accuracy about 70%.but usually people don't prefer to go through EEG for detection of depression[2].

Naomi JaneScott and Robin Stewar tSamuel Kramer researched on depression detection using facial expressions in a paper named[3] "Facial cues to depressive symptoms and their associated personality attributions". This experiment had two step, the first which contains the collection of photographs and a measure of depressive symptoms to create our face stimuli. In the second stage, these stimuli were used to investigate the ability of observers to discriminate between faces representing high and low depression symptom occurrence and to measure observer impressions from the high and low images of important social traits. By using this technique they

got accuracy 58%[8]. But many time facial expressions change with mood, so there is a possibility of detecting the false depression[3].

M. Dietrich et al. [4] "The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice," The essential target of this review was to explore the frequency of stress, anxiety, and depression for patients with normal pathologies influencing voice. The pathologies focused on were MTD and PVFMD because of their putative connection to mental conditions. The glottal insufficiency due to its lack of obvious relation. Before proceeding with a discussion of the outcomes, two preventative statements are in order. (1)The severity of an individual's voice problem was not included in the information examination. Although the inclusion of vocal handicap information would have given extra significant knowledge, we have chosen to defer the investigation of quantitative issues to future studies that will build upon this exploratory and preliminary study. Thus, any conclusions must be tempered with these perceptions and caught up with further, forthcoming reviews[4]

L. S. A. Low et al. [5] "Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents," In this paper we analyzed the impacts of joining different acoustic low-level descriptors (LLDs) with two baseline features i.e. MFCC and TEO-CB-Auto-Env for the recognition of depression in young people. In the underlying examination, TEO-CB Auto - Env beat the MFCC gauge include in the order. Combining the baseline feature of TEOCB-Auto-Env together with acoustic LLD features resulted in a classification increment over the first standard element (TEO-CB-Auto-Env). For male subjects, the general increment in order precision was 77.82%. For the female subjects, utilizing TEO-CB-Auto-Env include alone gave the most elevated segregation of 74.74%[5].

IV. METHODOLOGY

A. Introduction

Open Source IDE for software developers. That will provide all the basic tools needed to develop professional desktop, enterprise and mobile applications with the languages C, C++, and other related language such as, JavaScript and PHP and Ruby. NetBeans IDE 8.0.2 is easy to install and use and it run on many platforms like Linux, Windows.

B. Block Diagram:

F. Result

Fig 2.Coding Screen with Command Prompt Window
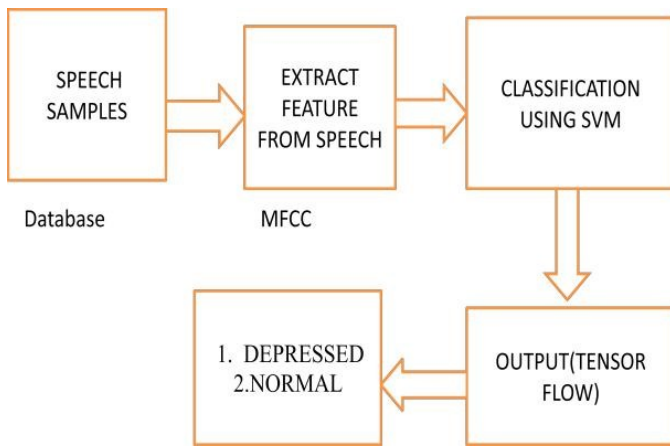
C. Database:

We are using Berlin Dataset Of Emotional Speech. The Recordings Took place in the anechoic chamber of the Technical University Berlin, Department of Technical Acoustics.

D. Steps to calculate MFCC:

1) First take the speech sample, then apply pre-emphasis followed by normalization and hanning windoing used for framing.
2) Frame the signal into short frames.
3) For each frame calculate the periodigram and estimate the power spectrum.
4) Apply the mel filterbank to the power spectra, sum the energy in each filter.
5) Take the logarithm of all filterbank energies.
6) Take the DCT of the log filterbank energies.
7) Keep DCT coefficients 2-13, discard the rest

E. Coding Part:

Machine learning is carried out in two half i.e Training and execution. The database containing audio files is trained first and then used in execution of program.
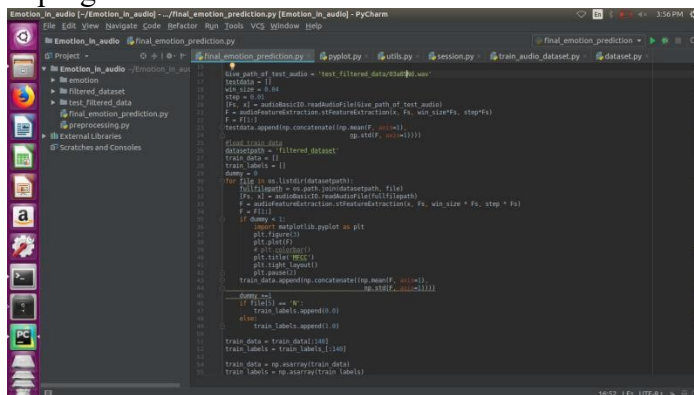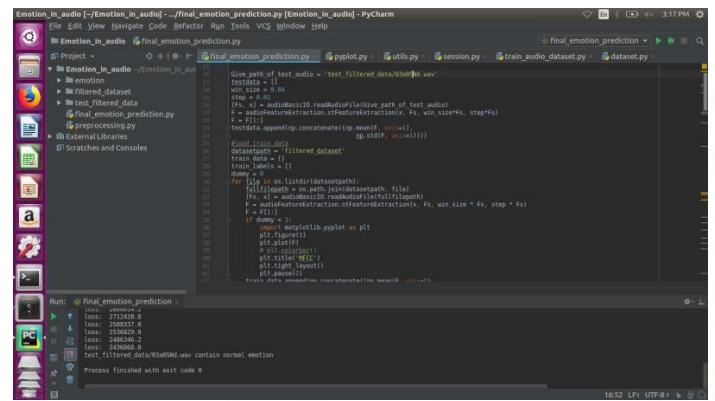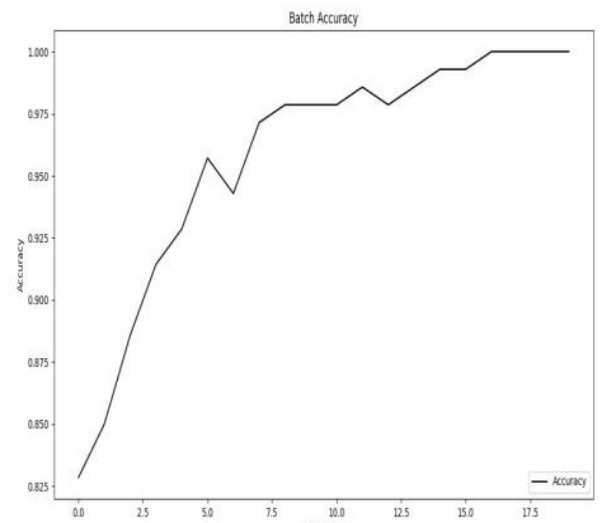

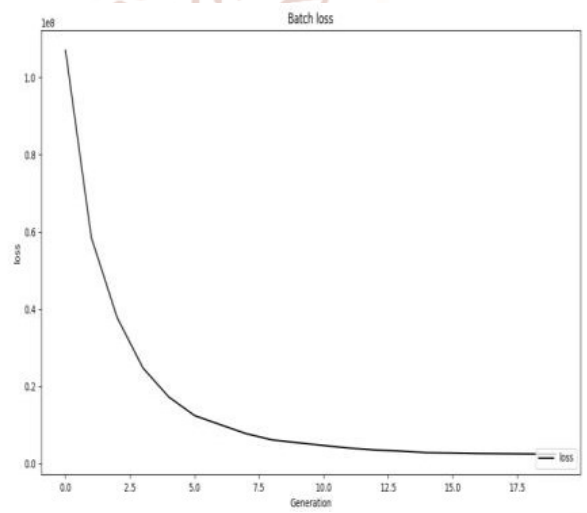Fig.1. Coding Screen

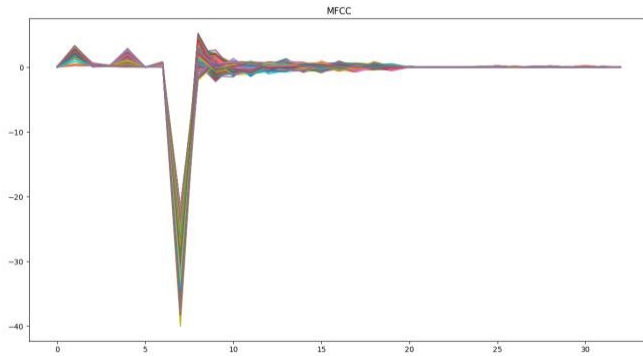Fig.3. Batch Accuracy

Fig.4. Batch Loss

Fig.5. MFCC output

G Algorithm

Take samples of depressed speech and normal speech. Extract the desired feature like MFCC. Classify using SVM. Process the speech samples using software TensorFlow. Calculate the various values of speech sample. If the calculated values of samples are similar to the actual values then say depression is detected.

V. CONCLUSION

By using this project we detect depression at early stage which helps to take appropriate treatment on primary stage. Speech is known to contain important information regarding a person's psychological state. It is one of the best media through which one can identify whether the patient is in controlled state or depressed state. Thus, a speech-based depression detection system could serve as a screening tool to assist mental health professionals in identifying clinically depressed persons. Using speech recorded database, this project reports an investigation of speech features categories such as the TEO and

MFCC. Good results are getting with help of SVM. Tensor flow software is used for speech processing.

VI. REFERENCES

1) Kuryati Kipli and Abbas Z. Kouzani, Member IEEE , "An Algorithm for Determination of Rank and Degree of Contribution of sMRI Volumetric Features in Depression Detection" *35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July, 2013*

2) D.P.X. Kan and P.F., "Decrease Alpha Waves In Depression: An Electroencephalogram (EEG) Study " *2015 International Conference on BioSignal Analysis, Processing and Systems (ICBAPS)*

3) Naomi Jane Scott n, Robin Stewart Samuel Kramer, Alex Lee Jones, Robert Ward , "Facial cues to depressive symptoms and their associated personality attributions" *Bangor University, School of Psychology, Adeilad Brigantia, Gwynedd, LL57 2AS, UK a*

4) M. DIETRICH, K. V. ABBOTT, J. G. SCHMIDT, AND C. A. ROSEN, "The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice," J. Voice, vol. 22, no. 4, pp. 472–488, Jul. 2008.

5) L. S. A. LOW, N. C. MADDAGE, M. LECH, L. B. SHEEBER, AND N. B. ALLEN, "Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents," in Proc. IEEE Int. Conf. Acoustic, 2010, pp. 5154–5157, Speech, Signal Process.