Project Proposal
Project Title: Speech Emotion Recognition

Team Members: Jai Raj Choudhary CWID: A20517449
                 Shruti Shankar Shete CWID: A20518508

Problem Statement:
Speech can be considered as a major means by which humans interact with machines. Machines can become more user friendly if they are able to respond according to the user's current emotional status. However, this being an arduous job we focus on few of the major challenges as follows : To begin with, there are a number of speech features and it's even harder to find the appropriate speech feature which is powerful enough to distinguish between emotions. Moreover, it's highly subjective how an individual expresses their emotion, it depends on various factors. We in this project focus primarily on monolingual emotion classification. Our sole aim with this project is to predict the emotion captured throughout the speech using Machine Learning techniques.

Previous work:
Various machine algorithms have been used to recognize emotions from speech such as Hidden Markov models (HMM) which deal with variations in speech, Gaussian Mixture Model (GMM) utilized for representing acoustic features. Deep learning techniques such as Convolutional Neural Network(CNN) and Recurrent Neural Network(RNN) were used to understand prominent features from raw audio data using spectrograms such as Mel-frequency cepstral coefficients (MFCCs) and low-level descriptors (LLDs).
Recently, hybrid methods i.e. combining different algorithms together for increasing efficiency have been tried. Hybrid modalities like DNN-HMM, GMM-HMM are able to achieve better results.
As speech is composed of words and sounds, researchers have also tried a combination of language information and acoustic features for emotion recognition. This method includes identifying key phrases and assessing the speech signals.
In this proposed project we are going to use MLPClassifier to create a model. Also we will use different machine learning and deep learning algorithms for emotion prediction and compare their accuracies .

Preliminary Plan:

Collect suitable dataset and preprocessing	02/10 - 06/10
Working on features	07/10 - 15/10
Building models	16/10 - 28/10
Testing the model	29/10 - 02/11

References:

1. Kalani Wataraka Gamage, Vidhyasaharan Sethu, and Eliathamby Ambikairajah, "Salience based lexical features for emotion recognition," in Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, 2017, pp. 5830–5834.
2. Speech emotion recognition using hidden Markov models Tin Lay Nwe ,, Say Wei Foo , Liyanage C. De Silva
3. MULTIMODAL SPEECH EMOTION RECOGNITION USING AUDIO AND TEXT Seunghyun Yoon, Seokhyun Byun, and Kyomin Jung
4. GMM SUPERVECTOR BASED SVM WITH SPECTRAL FEATURES FOR SPEECH EMOTION RECOGNITION Hao Hu, Ming-XingXu, and Wei Wu.
5. Relative Speech Emotion Recognition Based Artificial Neural Network Liqin Fu Xia Mao Lijiang Chen.
6. Emotion Recognition of Affective Speech Based on Multiple Classifiers Using Acoustic-Prosodic Information and Semantic Labels Chung-Hsien Wu, Wei-Bin Liang.
7. HIDDEN MARKOV MODEL-BASED SPEECH EMOTION RECOGNITION Bjorn Schuller, Gerhard Rigoll, and Manfred Lang.
8. Techniques and Applications of Emotion Recognition in Speech S. Lugović*, I. Dunđer** and M. Horvat.
9. Speech Emotion Recognition with Deep Learning Pavol Harár1 , Radim Burget1 and Malay Kishore Dutta.
10. Speech emotion recognition based on convolution neural network combined with random forest. Li Zheng, Qiao Li, Hua Ban, Shuhua Liu.
11. Multimodal Approach of Speech Emotion Recognition using Multi-Level Multi-Head Fusion Attention-Based Recurrent Neural Network. Ngoc-Huynh Ho, Hyung-Jeong Yang, Soo-Hyung Kim, Gueesang.