**CSCE- 5310 Methods of Empirical Analysis**

**PROJECT PROPOSAL**

**Emotion Classification Using Audios**

**Sai Karthik Poojari (11518167)**

**Shashi Sekhar Ranganathan (11533071)**

**Leela Varadatta Sai Addanki (11617487)**

**University of North Texas, Denton TX-76201**

**Idea Description:**

Speech Emotion Recognition, abbreviated as SER, is the act of attempting to recognize human emotion and affective states from speech. This is capitalizing on the fact that voice often reflects underlying emotion through tone and pitch. This is also the phenomenon that animals like dogs and horses employ to be able to understand human emotion.

The best example of it can be seen at call centers. If you ever noticed, call centers employees never talk in the same manner, their way of pitching/talking to the customers changes with customers. Now, this does happen with common people too, but how is this relevant to call centers? Here is your answer, the employees recognize customers’ emotions from speech, so they can improve their service and convert more people. In this way, they are using speech emotion recognition.

**Goals and Objectives:**

1. To apply DL techniques to predict the stock market
2. To learn to engineer to extract features from audio data
3. Apply various techniques to increase the accuracy of the models
4. To split the dataset for training the model and testing of the model

**Motivation:**

Trying to predict the emotion of the humans using their voices because voices reflect the various aspects of the human beings. And this will help in understanding humans in a more accurate way and may automate our daily tasks which consume most of our time. It will help us understand the behaviors of customers, patients and in many other fields in real life.

**Literature Survey:**

1. Emotion Recognition is an important area of work to improve the interaction between human and machine. Complexity of emotion makes the acquisition task more difficult. Quondam works are proposed to capture emotion through unimodal mechanisms such as only facial expressions or only vocal input. More recently, the inception of the idea of multimodal emotion recognition has increased the accuracy rate of the detection of the machine. Moreover, deep learning techniques with neural networks extended the success ratio of machines in respect of emotion recognition. Recent works with deep learning techniques have been performed with different kinds of input of human behavior such as audio-visual inputs, facial expressions, body gestures, EEG signals and related brainwaves. Still many aspects in this area to work on to improve and make a robust system will detect and classify emotions more accurately. In this paper, we tried to explore the relevant significant works, their techniques, and the effectiveness of the methods and the scope of the improvement of the results.
2. With increasing accessibility and availability to online data, it is very motivating and interesting to study huge data for sentiment and emotion analysis. Emotion Analysis is an extension of sentiment analysis. It is the process of analyzing text and classifying text into different emotion classes.In current scenario, emotion detection has become a trend because of its use in various domains like marketing, pervasive computing, recommendation systems, political science, etc. A lot of research work done so far deals with issues like context-dependency, word sense disambiguation and co-reference resolution and to resolve these issues and improve the design and implementation of systems is strictly needed. Treating emotion detection as a single-label classification problem is not a good idea since a particular affective word can be mapped to multiple classes.

**Features:**

1. Predicting Errors
2. Improving Accuracy
3. Classifying Emotions
4. Detecting Emotions from Voices

**Expected Outcome:**

To train the DL Neural Network that can classify no human emotions from their voices with the help of Modern AI techniques. Learn how to extract features from audio data and how to handle the audio data.

**References:**

1. Ahmed Ali Abdalla Esmin, Roberto L. de Oliveira Jr., and Stan Matwin. 2012. Hierarchical classification approach to emotion recognition in twitter. In 11th International Conference on Machine Learning and Applications, ICMLA, Boca Raton, FL, USA, December 12-15, 2012. Volume 2, pages 381–385.
2. Saif M. Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. Semeval-2018 Task 1: Affect in tweets. In Proceedings of International Workshop on Semantic Evaluation (SemEval-2018), New Orleans, LA, USA.

Github: <https://github.com/Shashisekharranganathan/Emotion-Classification-using-Audio>