# RECOGNITION OF SPEECH EMOTION USING NAÏVE BAYES CLASSIFIER

**Dr. Ch. Raja**
Associate Professor, Department of ECE
Mahatma Gandhi Institute Of Technology, Hyderabad-500075

## ABSTRACT:

Speech emotion recognition is one of the major challenges in speech processing. Besides facial expressions or gestures, speech has proven as one of the most promising modalities for automatic emotion recognition. To identify the emotions from the speech signal, many systems have been developed. Recognizing basic emotion through speech is the process of recognizing the intellectual state. Emotion identification through speech is an area which increasingly attracts attention within the engineers in the field of pattern recognition. Emotions play an extremely important role in human life. It is an important medium of expressing a human's viewpoint or feelings and his or her mental state to others. Humans have the natural ability to recognize emotions through speech information. Emotional computing has gained enormous research interest in the development of Human Computer Interaction over the past ten years. With the increasing power of emotion recognition, a logical computer system can provide a more friendly and effective way to communicate with users in areas such as video surveillance, interactive entertainment, intelligent automobile system and medical diagnosis. Automatic detection of emotions has been evaluated using standard Mel-Frequency Cepstral Coefficients (MFCCs), and pitch related features extracted from a speech segment[1]. These obtained features are then classified using Naïve Bayes classifier. Recognition accuracy for these features is considered as it mimics the human ear perception. So, emotion recognition using these features are illustrated.

**KEYWORDS:** MFCC, Naïve Bayes classifier

## CLASSIFIER USED:

**NAÏVE BAYES:** Naïve Bayes classifier is based on the so-called Bayesian theorem with the naïve assumption of independence between every pair of features. This classifier is inspite of the apparently over-simplified assumptions has worked quite well in many real-world situations. It is very fast and has a good performance, better in some cases than more sophisticated methods. The main advantages of this classifier are the conditional independence assumption, which helps to obtain a quick classification, the probabilistic hypothesis (results obtained as probabilities of belonging of each class), high accuracy [2].

## PRINCIPLE OF WORKING:

In this paper we have collected the speech signals data from University of Toronto and University of California. Main reason to collect this data base instead of recording is signal is that, signal which we have collected is with low noise and then normalization is done. The pre-emphasized speech signal is then blocked into frames of N sample points with adjacent frames being separated by M. In our work, the frame length N = 256(10ms). Here we used hamming window to reduce the side lobes of the speech signal. Hamming window technique is used because it is better among all the windowing techniques. Short-time Fourier transform converts each frame from time domain signals into frequency domain and obtain frequency response of each frame. It involves extracting important information associated with the given speech and removing all the remaining useless information. Features such as energy, pitch, power and MFCC are extracted. MFCCs enable a signal representation that is closer to human perception. A Mel is a unit of measure of perceived pitch or frequency of a tone MFCC's can be calculated by applying a Mel-scale filter bank to the Fourier transform of a windowed signal. In the process of training Each collected signal in the dataset has labelled with the different emotional state [3]. A classifier (Naïve Bayes) is made by trained dataset and algorithm is form on the basis of the dataset. Trained dataset is perfectly classified under six different headings, such as anger, happy, sad, surprise, disgust and fear., this is basically called a machine learning. In testing block when an input is given to test the speech signal the MFCC values of the signal is compared with the trained signals MFCC values. If the MFCC values matches then the output box is displayed which contains the emotion. The following figure 1 shows the working principle of speech emotion recognition using Naïve Bayes classifier

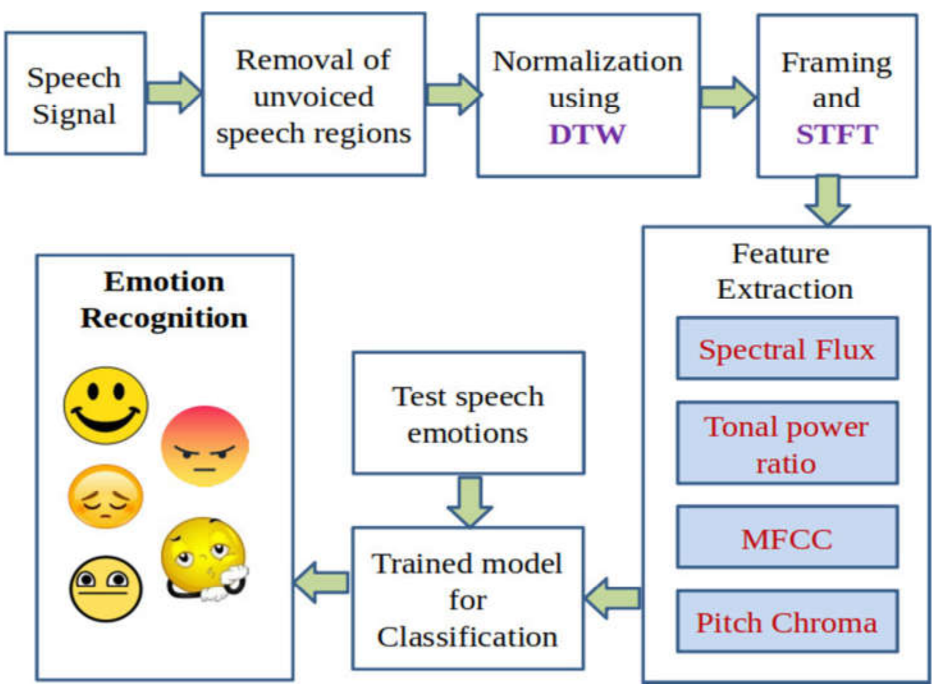


Figure 1: Speech emotion recognition using Naïve Bayes classifier

## RESULTS:

To get required results first the classifier must be trained. The database collected from University of Toronto and University of California is used for this purpose. During the training the classifier itself will classify the emotions into six different types based on name given by us. Then we will calculate the required features namely pitch, formants and MFCC. These features are given as input to the classifier for training purpose. Then the trained classifier is used in the output for identifying the unknown speech emotion [4-5].

The below figure 2 shows us when the classifier is being trained:

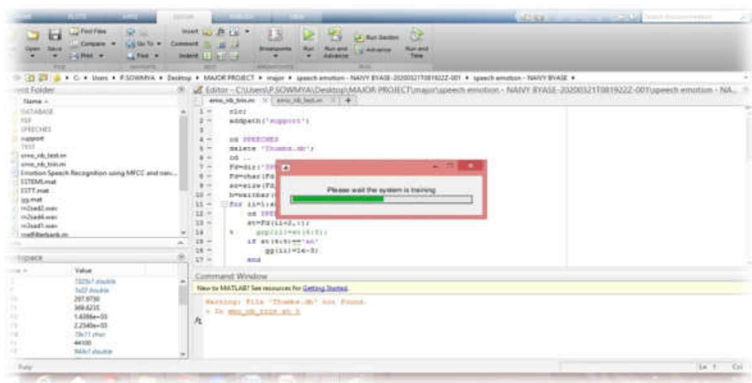

**Figure 2: Classifier being trained**

**Output:**

The following figure 3 shows the output of the system when “Happiness” emotion is being processed.

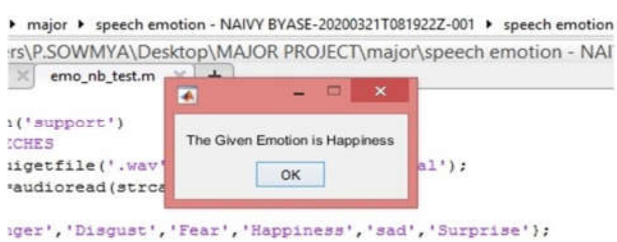


**Figure 3: Result for Happiness emotion is displayed**

**Output:**

The following figure 4 shows the output of the system when “Fear” emotion is being processed.

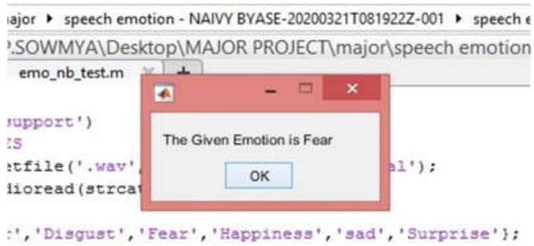


**Figure 4: Result for Fear emotion is displayed**

**Output:**

The following figure 5 shows the output of the system when "Anger" emotion is being processed.

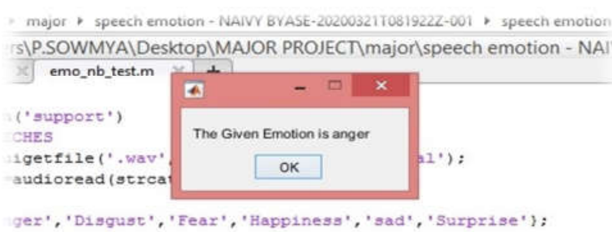


**Figure 5: Result for Anger emotion is displayed**

**Output:**

The following figure 6 shows the output of the system when "Disgust" emotion is being processed.

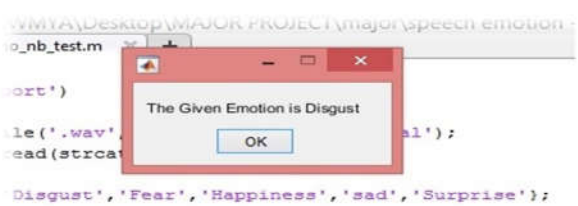


**Figure 6: Result for Disgust emotion is displayed**

## CONCLUSIONS:

As technology evolves interest in human like machines increases. Technological devices are spreading and user satisfaction increases importance. A natural interface which responds according to user needs has become possible with affective computing. The key issue of affective computing is emotions. Any research which is related with detection, recognition or generating an emotion is affective computing. User satisfaction or un-satisfaction could be detected with any emotion recognition system. Besides detection of user satisfaction, such systems could be used to detect anger or frustration. In such cases, user could be restrained like driving a car. In emotion detection tasks, speech or face emotion detections are the most popular ones. Easy access to face or speech data made them very popular. Speech carries a rich set of data. In human to human communication, via speech information is conveyed. Acoustic part of speech carries important info about emotions. MFCC are used for the feature extraction. Algorithm with the Naïve Bayes Classifier, overall performance is tested.

## REFERENCES:


[1] Mehrdad J. Gangeh, AliGhodsi, Mohamed S. Kamel," Multiview Supervised Dictionary Learning in Speech Emotion Recognition," IEEE Transaction on audio, speech, and language processing.

[2] Shikha Gupta1, Jafreezal Jaafar2, Wan Fatimah wan Ahmad3 and Arpit Bansal4 J. Clerk Maxwell, "Feature extraction using mfcc" Signal & Image Processing: An International Journal (SIPIJ) Vol.4, No.4, August 2013.

[3] Sagar K. Bhakre, Arti Bang," Emotion recognition on the basis of audio signal using Naive Bayes classifier," International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 03 November 2016.

[4] Aastha Joshi," Speech Emotion Recognition Using Combined Features of HMM & SVM Algorithm International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Volume 3, Issue 8, August 2013.

[5] Eslam Mansour mohammed1, Mohammed Sharaf Sayed2," LPC and MFCC Performance Evaluation with Artificial Neural Network for Spoken Language Identification," International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 6, No. 3, June,2013