# Design and Implementation of Human Emotion Recognition System from Voice Samples Using MFCC and SVM

Jahnavi Vardhan. Ambati Dept of Mtech.Integrated Sofware Engineering

VIT chennai University Email:ambatijahnavivardhan @gmail.com Sanjana Reddy.Jada Dept of Mtech.Integrated Sofware Engineering

VIT chennai University Email:sanjanareddyjadaa@gmai 1.com

#### Abstract—

Emotion Recognition through voice samples is a recent research topic in the Human-Computer Interaction (HCI) field. The need for it has arisen for a more easy communication interface between humans and computers because computers have become the main part of our lives. A huge amount of work currently going on to improve the interaction and understanding between humans and computers.

To achieve this goal, a computer would have to be able to differentiate its current situation and respond differently depending on that particular observation. This process involves understanding a user's emotional state and to make the human-computer interaction more natural, the main objective is that the computer should be able to recognize the emotional states of humans in the same as a human does. The capacity of the emotion recognition system depends on the number of type of features extracted and the classifier used which for the detection of emotions. The proposed system aims at the identification of basic emotional states such as anger, joy, neutral, and sadness from human voice samples.

While classifying different emotions, features like Mel Frequency Cepstral Coefficient and Energy are used. Standard Emotional Database that is can be an English Database is used which gives the exact detection of emotions than recorded samples of emotions. Besides compared to human facial expressions, speech has proven to be one of the most promising modalities for automatic human emotion recognition. Speech is a spontaneous medium of perceiving the depth of emotions. The information which is related to different states of a human being. This methodology describes and compares the performances of Learning Multiclass

Support Vector Machine (SVM) and their combination of emotion recognition. Multilevel SVM classifier is used for the identification of four discrete emotional states namely happy, sad, anxious, and neutral.

Keywords— ANN- Artificial Neural Network, DCT- Discrete Cosine Transform, DFT- Discrete Fourier Transform, FFT- Fast Fourier Transform, GUI- Graphical User Interface, HCI- Human Computer Interaction, HMM- Hidden Markov Model, LPCC-Linear Prediction Cepstrum Coefficient, MATLAB- Matrix Laboratory, MF- Mel Filtering, MFCC-Mel-frequency cepstral coefficients, SER- Speech Emotion Recognition, **SVM- Support Vector Machine** 

#### I. INTRODUCTION

Speech is one of the basic and natural ways of communication among humans begins. Emotion can make speech more expressive and also effective. Speech Emotion detection can be an easy task for humans beings, but a difficult one for computers. So, there is a need for such emotion recognition systems that can also make Human and Computer Interaction (HCI) easier. To make human-computer interaction easier, the computer should be able to recognize emotional states in the same as a human being does. The efficiency of the voice emotion recognition system depends on the type of features extracted and classifiers used for the detection of different types of emotions. By giving the machines the ability to recognize different types of emotions and make them able to behave in accordance with the emotional conditions of the user, human-computer interfaces will become more efficient.

There are two segments: The First part is the training part in which the system is trained to identify further proceedings. In the first part, samples of every voice category are taken and their features are fetched after the successful partition of the voice file and further on saved into the database. The other part is the testing part in which a sample of voie is taken and all the required properties are fetched and matched with the saved database values. The previous technologies which use different classifiers for emotion recognition are reviewed.

The classifiers are used to different emotional states such as angry, disgust, fear, happy, neutral, sad, and surprise.

The features extracted from the speech are the energy, pitch, Zero crossing rate, linear prediction cepstrum coefficient (LPCC), Mel-frequency cepstrum coefficient (MFCC) for automatic recognition of the speaker's emotional states. The classification is based on the extraction features. Multilevel SVM (Support Vector Machine) classifier is used for identification of seven discrete emotional states

### **Mel Frequency Cepstral Coefficients**

Mel frequency Cepstral coefficients (MFCC) algorithm is one of technique which takes voice sample as inputs. After the processing, it can calculate coefficients that are unique to a particular sample. In the project, the software called MATLAB 7.14 Version R2012 is used to perform MFCC. The procedure for the implementation of MFCC makes it the preferred technique for voice recognition. MFCC is the most rapidly used feature in speech recognition. It was introduced by David and Mermelstein. The purpose of the MFCC processor is to mimic the behavior of human understanding

Mel-frequency cepstral coefficients are very important coefficients that collectively make up an MFC. They are derived from a type of cepstral representation of all the audio clips. The difference between the cepstrum and the Mel-frequency cepstrum is that in the MFC, that the frequency of bands is equally spaced with the Mel scale, which almost with the human auditory system's response that is more closely than the linearly-spaced frequency bands used in the normal cepstrum. This frequency of warping can allow for better

representation of voice.

## **Support Vector Machine**

Support Vector machines (SVM) can be defined as the systems which can use hypothesis space of the linear functions in the high dimensional feature space, which is trained with a learning algorithm from the optimization theory that implements learning bias derived from statistical learning theory. It can also be used for many applications, such as handwriting analysis, face analysis, and so on, especially for pattern classification and regression-based applications.

The Support Vector Machine can be used as a classifier for emotion recognition. SVM is a computer algorithm that can be used in pattern recognition for data classification and regression. The classifier can be used for classifying or separating features from other features. Support Vector Machine performs the classification by constructing an N-dimensional hyperplane which optimally separates the data into many categories. SVM is a simple and efficient algorithm it has a very good classification performance compared to the other classifiers

### II. LITERATURE REVIEW

Recently, automatic speech recognition has reached very high levels of performance. This current state of performance is largely due to improvements in the algorithms and techniques that are used in this field. As a result, the accuracy level of Automatic Speech Recognition systems is improved especially when combining of various algorithms and techniques. Performance of speech recognition systems is calculated in terms of accuracy and speed. The speed here in emotion recognition systems have various processes. Some of the important steps are features extraction, subset selection and classification. There are some different features and classifiers used in various papers have been analyzed here.

# **MFCC (Mel- Frequency Cepstral Coefficients)**

Mel- Frequency Cepstral Coefficients (MFCCs) are one of the most recently used feature extraction techniques used in speech recognition based on frequency domain using the Mel scale which is based on the human ear scale.MFCC is used as an audio feature extraction technique which extracts all the detailed parameters from the speech.

In the experiments, the audio pitch is linearly scaled in the frequency range 0-1000 Hz. Above 1000 Hz, the scale is logarithmic. MFCCs are extracted by performing and transforming the speech signal which is complex and difficult between glottal pulse and the vocal tract impulse. These both glottal pulse and vocal tract impulse response into a sum of two components known as the cepstrum that can be separated by band pass linear filters.

MFCC features, first performed as standard Fourier analysis, and then converted the power-spectrum to a Mel-frequency spectrum. Therefore, MFCC issecured by taking the logarithm of that spectrum and by computing its inverse Fourier transform.

## **SVM (Support Vector Machine)**

A SVM is a binary nonlinear classifier. It is capable of guessing whether an input vector x belongs to a class 1 or to a class 2. SVM is used as a classier method that performs classification tasks by constructing hyper planes in a multidimensional space which separates cases of different class labels. For this project, the SVM is used as a classifier to classify different emotional states in the Speech. It can be used for many applications including image recognition, character recognition, text categorization, speech recognition, biometric applications, fault detection, diagnostic applications etc.

The main goal in this project is to find the optimal decision function, where it can be easily seen that there is an infinite number of optimized solutions for this problem, in that case we can separate the training samples without any errors. Traditional classifiers such as Support vector machine (SVM) have been used for almost all proposed speech emotion recognition systems.

# **Hidden Markov Model-Based Speech Emotion Recognition**

Hidden Markov model is one of the technique frequently used in Speech Emotion Recognition. There are two methods that are used. The first method that is introduced here is global statistics framework of an utterance which is classified by Gaussian mixture models. The model is used for deriving features from the raw pitch and energy contour of the speech signal. A second method that used here is increased temporal complexity by applying continuous hidden Markov models considering several states using low-level features instead of global statistics

# Speaker Emotion Recognition Based on SVM/HMMs

For speech emotion recognition we mainly use Support Vector Machine and Hidden Markov Model. SVM is used to divide the emotions into a groups and HMM is differentiate emotions in those groups. For more detailed estimation, we combine all four HMM classifiers into a system. The recognition result of the system will be compared with the isolated HMMs using Mandarin data.

After the experiment, the results shows that comparing with the method based on only HMMs, the proposed system is more effective and the average of the recognition rate reaches 76.1% when speaker is independent.

#### III. DESIGN AND IMPLEMENTATION

## **Proposed Work**

The speech emotion recognition system aims to automatically identify the emotion of human beings from a speaker's voice. It is related to the speech signal, extracting the features which information emotions from the speaker's speech, and using the appropriate method to recognize the emotion. The existing systems of speech emotion recognition based on the Gaussian Mixture Model (GMM) and Hidden Markov Model(HMM)

This project shows better performance compared to existing. Noise levels are taken so that the emotion can be identified even though if the voice a signal is highly noised.

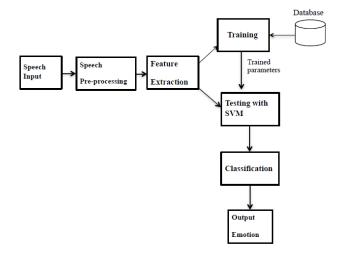
SVM classifiers have advantages over GMM and HMM which includes the global optimality of the training algorithm.

The existence of excellent data-dependent generalization bounds.

# Algorithm

The design and implementation of the human emotion recognition from voice samples using the MFCC and SVM techniques is with respect to the proposed system in detail.

#### 3.1 Architectural Diagram of Proposed System



Basic Block Diagram of the System

The basic structure of speech emotion system can be showed as above. The proposed methodology of human emotion recognition system has majorly 6 components: input speech signals, preprocessing Technique, feature extraction Technique and selection Technique, training and testing Techniques, classification Technique and finally emotions recognition.

### **Speech**

The major means of communication happens with speech for Human. It is a Complete complex signal which consists of information about messages, speakers, languages, emotional state.

#### **Emotions**

Emotions are said as change in physical, psychological feelings which influence action thinking of human. It is connected with temper, personality, State of mood, motivations and energy.

The speech is obtained as the input to System with a input technology microphone, after the speech signal is pre-processed the silent side of signal the noise is detached and feature extraction is did the speech signal contain different types of parameter for the property of speech that is defined. Speech feature basically do not understand easily as of the change in behaviour and temper adjustment makes this tasks highly tedious. MFCC, Energy feature is used. The speech signal is recorded with sample rate of 16000 Hz with microphone.

Train and test is done to compared the all signal train is the process of familiar the system with the emotion character of a speaker, while test is recognise task. Classifier SVM is used to obtain Speaker Emotion.

#### **Emotional Speech Input:**

The input system is voice sample for the person, The voices samples are stored in a database in the computer. The system can identify all six emotion from input speech.

### 3.2.1 Label Grouping

To decrease the general error, the emotion are made to many small group instead of 14 distinct label, as given above. The running hundred of random label, the best group are recorded with decrease in error reduced as expected, in the order of 25-30% test error. These label generally order the least error at across 20 to 15%. While checking each group, feature selecting was also given to reduce test error other than train error, make the result optimistic.

### 3.2.2 Types of Emotions:

- HAPPY
- SAD
- ANGER
- FEAR
- ANXIETY
- NEUTRAL

#### 3.3 Pre-Processing

Pre processing is majorly done to clear the silent side of the speech that do not consist any information and the removal of the noise from the sample. The reason of pre-processing is to increase the large frequency of a signal to get low frequency spectrum of signal and frequency.

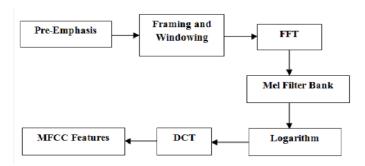
window function make speech frames for window functions are Hamming window and Rectangular window.

## 3.4 Feature Extraction

All emotional state could be analysedby certain speech feature which could be quality features. Some qualityfeatures can be obtained directly include Pitch, Intensities, Energy are most widely used for features in the emotion recognition domain. It is impossible to acknowledge some emotion state using only thesefeature, It is very tough when it comes to emotion statewith equal level of stimulations.

3.4.1 MFCC (Mel-frequency cepstral coefficients) Mel-frequency cepstral coefficients -MFCCs is a parametric inclusion of speech signal used in emotion recognition system,. It is successful for all purposes as well, as the speaker identification and emotion recognition. MFCC is an audio feature extraction technique which include parameter from speech same to one that are used by humans for hearing speech.

Thescale was develop by experiment with the human ear interpretations of the pitch in 1940. The majoruse of the experiment is to demonstrate the humans auditory systems on a linear scale



#### 3.4.3 Framing and Windowing

Framing the speech signal is generally divide to short blocks, frames the spectrum analysis is done in this frame. These are due to reason the human speech signal is slow vary with time can be treatas quasi stationary process. The emphasized speech signalsisafter blocked to frame of N sample point with regard frame being separate by M lower than N. The intialframe is compose of N sample point. The second frame begun the M sample point after intial frames and overlap by NM sample point . The process continue till each are accommodate in more frames. Frame length  $N=25616 \mathrm{ms.overlap~8ms~N-M=128}$  overlap between both adjacent frame to makecommon between frame.

The Hamming window decreasethe frequencies of resolutions of the spectrumanalysis and reduce the side level window transfer .

### 3.4.4 FFT, Logarithm and Discrete Cosine Transform

The Fast Fourier TransformFFT converts every frame of N sample of data sequencesto frequencies of the spectrum. FFT - Fast Fourier Transform is very fast algorithm can be DFT -Discrete Fourier Transform can be said as set of N sample. Logarithms operation simulate loudness. The Mel Frequency Cepstral Coefficients-MFCCfor power of filtered bank bylogarithms . The operation can map the logarithm amplitude of spectrum by the Mel scale.

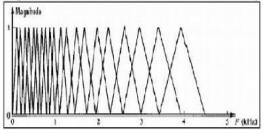
#### 3.4.5 Zero Cross Rate -ZCR

A zero cross that occur the success sample hadmany algebra sign. In other word we say ZCR of a audio signal is the stableof many time signal cross zero amplitude line of passage from a positive to negative . It is useful for detect voice and unvoicepart signal well for the intialpoint & end-point detection. The audio signal is divideto temper segment by rectangle function of zero crossing rate of three possible value: +1, 0, -1 depend on the sample is positive, zero or negative.

#### 3.4.6 Mel Filter

Mel filter is a group of triangles band width pass filter that simulate the character of the human ear is attached to the spectra of speech . This process is called Mel filter . In This step map of every frequency spectrum to Mel scale is done. Psychonomics study have show the human thinking of frequencies content of sound of speech signal does not follow a linear scale. Thus for each tone has actual frequencies f, measure in Hz, a subject pitch is measure of the scale the 'mel' scale.

The mel frequency scale has linear frequencies space less than 1000 Hz and thelogarithmicspace greater than 1000 Hz. Therefore they use the formula to compute the mel for a the frequency f in Hz. The Mel filter bank has consist of overlap triangle filter with frequency has determine by centre frequencies of two filter. The Mel filter are graphical



#### 3.4.7 Train and Test

Emotion recognition system has two phase training dataset and testingdatset. Train set is process of familiar system the emotion character of the speaker. Test is the actual recognition task. Take the real time databases so we should train. Thetrained database 70% of data base require train. We use many databases to combine different feature to build different training model and analyse their recognition accuracy.

. It show its sound with database accuracy. The train we will be take few voice sample of each and every categories by the classification.

#### 3.4.8 CLASSIFIERS

In this project speech emotion recognition system, it consists of two stages:

- 1. The appropriate features which are available in speech data are extracted using front-end processing unit, and
- 2. A classifier (SVM) that decides the underlying emotion of the speech.

## 3.4.9Support Vector Machine (SVM)

SVMs is one of the effective classical pattern recognition problems, a logical progression was to apply it to classification of phonetic segments in speech and it is also have proven. It is also the Determining vowel data. This SVM classifier, is mostly used to benchmark nonlinear classification algorithms, and not for immediate interest in continuous speech recognition because of a lack of variation in the phonetic context. For constructing a multi-class classifier for K-class problem through the construction of basic binary SVMs there are various

ways and approaches. We have chosen One-versus-One Classification based on voting scheme which is most successful.

In this case, to discriminate two training speech signals by considering every possible pairs the individual classifiers are designed (in our experiment K=10). So we constructed 45 One-vs-One classifiers. There are some advantages for choosing this approach as: for differentiating each speech signal from others speeches, we require different kernel functions which reduce generalization error rate and training time by reducing the number of support vectors on small amount subsets of total data. A voting scheme was initially assigned with fixed weights to cast one vote on favor of the class chooses on running of all One-vs-One classifier for the test input data. This voting scheme works well as One-vs one classifiers are forced to choose between only two classes.

These votes all are distributed uniformly so that we can classify the correct classes of the speech signals. The One-vs-One classifier takes more time and space for training data but it differentiates the speech signals very well which will be helpful for improving the efficiency of a system in a faster manner.

## 3.4.10 SVM CLASSIFIER

In this project Support Vector Machine will classify different emotional states such as Anger, sadness, fear, happy, boredom. SVM gives very good performance in classification comparing to the others and it is also simple and efficient. SVM is a most famous learning method for classification, regression and other tasks. Kernel function is the mainly behind the support vector machine (SVM) classifier, which leads to get optimum classification in this new feature space. Ittransforms the original feature set to a high dimensional feature space.

There are some functions in Kernel like linear, polynomial, radial basis function (RBF) can be used in SVM model for some extent. SVM classifier is used in some of the main applications like pattern recognition and classification problems, and because of that it is used in the speech emotion recognition system. SVM is one of the best classifier. Because of SVM emotions can be classified to huge margin.

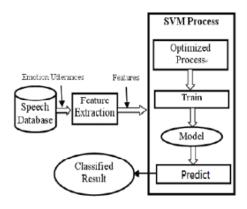


Figure 3.3 : SVM Architecture

The process of the SVM system is as follows: Step1: Extracting the emotions in the speech from utterances.

Step2: To improve the accuracy rate of the classification in SVM, the main task is optimized process.

Step3: After optimizing, the system trains the optimized model for classification.

Step4: Finally, system gives a classification result (class label or recognition rate) about test samples

#### 3.5 FLOW CHART

Nutraceuticals are characterized as a food or food fixings that forestall and treat diseases. They have antioxidative job that can decrease the degree of ROS and free revolutionaries. There is a connection between ROS creation and oxidative pressure that assume a part on redox motioning from the organelle to the cytosol to core. Oxidative pressure is unevenness between the development of free revolutionaries and cancer prevention agent guard instrument. The oxidative pressure diminish in the body with utilization of foods grown from the ground including high measures of against oxidative nutraceuticals and thus, frequency of malignant growth and cardiovascular sicknesses decline.

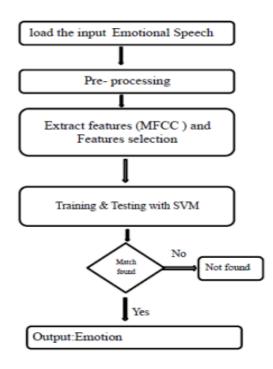


Figure 3.4 : Flow Chart of the system.

The process of emotion recognition from voice sample as follows:

- 1. For input, emotional speech is given.
- Now pre-processing technique is done for the input and all the noise will be removed in this step.
- 3. Features like pitch, intensity and energy is extracted after pre-processing process. And this process is done using MFCC (Mel Frequency Cepstral Coefficients).
- 4. In this step after feature extraction Training and Testing is done using SVM (Support Vector Machine).
- 5. During Training and Testing the data, if the match is not found, it displays the result "not found" else it displays the output emotion.

## 3.6 APPLICATION

MFCC and SVM process can be used in some of the application like:

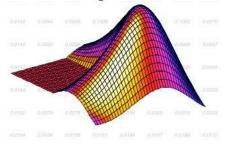
- 1. Human-Computer Interaction
- 2. Lie Detection
- 3. In medicine counselling, it can be used to detect Client's emotional state
- 4. Intelligent Tutoring System
- In call-center system, it is used to recognize voice and detect anger in speech to prioritize angry.

## 3.7 EMOTIONAL AND SPEECH PARAMETERS

In the process of human emotion recognize system, a person's speech can be varied by various changes in the automatic nervous system and some effective technologies. For example, if speech produces fear, anger, or joy with loud and fast, a higher and wider range in pitch, in speech and some emotions such as sadness or tiredness generates slow and low pitched speech. Some emotions have been found to be more easily identified, such as anger or approve.

The processing technologies recognizes the user's emotional state using computational analysis of speech features in this project. Pitch variables can be analyzed by Vocal parameters and prosodic features and through pattern recognition techniques speech rate can be analyzed.

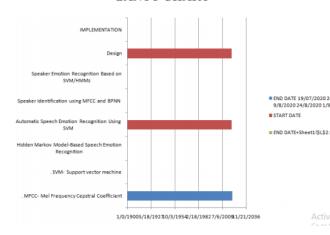
## **3.8 Software Requirement** *3.9.1 MATLAB 2012*



## IV. RESULTS AND DISCUSSIONS

The performance of speech emotion recognition depends upon many factors such as the quality of speech that is obtained without noise, features extracted from speech signals, and the classification method used. We elaborate on our own speech wave sample for the feature extraction in the condition. Another part of the wave sample which is in another format we can convert in .wave format. We use both databases, combine different features to build different training models, and analyze their recognition accuracy. First off, we want to analyze and feature extract a small collection of audio samples storing their feature data as training data in .WAV format.

#### **GANTT CHART**



	Anger	Happiness	Sadness	Fear	Surprised
Rate	Slightly faster	Faster or slower	Slightly slower	Much faster	Very much faster
Pitch Average	Very much higher	Much higher	Slightly lower	Very much higher	Very much lower
Pitch Range	Much wider	Much wider	Slightly narrower	Much wider	Slightly wider
Intensity	Higher	Higher	Lower	Normal	Lower
Voice Quality	Breathy, chest	Breathy, blaring tone	Resonant	Irregular voicing	Grumble chest tone
Pitch Changes	Abrupt on stressed	Smooth, upward inflections	Downward inflections	Normal	Wide, downward terminal inflections
Articulation	Tense	Normal	Slurring	Precise	Normal

#### V. CONCLUSION

The recent work is done in the field of Speech Emotion Recognition with the most used methods of feature extraction and many other classifier performances can be reviewed. The probability of success of speech emotion recognition is dependent on appropriate feature extraction and also exact classifier selection from the sample speech with emotion.

MFCC technique can be used as the feature extraction. It is a well-known technique which can be implacable in speech recognition to describe the characteristics of every individual signal. The SVM's algorithm overall performance can also be tested because the training is done for the learning module section to initiate the database selection.

#### Future Work

There are a number of ways that this project could be extended. Perhaps one of the most common tools in the emotion detection system is the neural networks and support vector machine for automatic, recognition. Also, the integration of other modalities such as video based or manual interaction will be investigated further.

To improve machine-human interface automatic emotion recognition through speech provides some other applications such as speech emotion recognition system used in Aircraft cockpits to provide analysis of the Psychological state of pilots to avoid accidents

Emotion recognition through speech is an area that increasingly attracting attention within the engineers in

\*\*\*DO DATE 19/07/2020 24/8/2020 1/9/2the of pattern recognition and speech signal

\*\*\*START DATE\*\*

processing in recent years. Emotion recognition plays

\*\*\*END DATE-Sheet1/5152-5159\* n important role in identifying the emotional state of the speaker from a voice signal. Emotional speech recognition aims at automatically identifying the emotional human being from his or her voice.

Activate Windows

#### REFERENCES

- [1] Survey on AI-Based Multimodal Methods for Emotion Detection Joanna Kołodziej Horacio González-Vélez 26 March 2019 Part of the <u>Lecture Notes in Computer</u> <u>Science</u>book series (LNCS, volume 11400)
- [2] HUMAN EMOTION RECOGNITION AND DETECTION ON HUMAN HEALTH pooja S Kaje January 2015 Research gate
- [3] Emotion Detection Algorithm Using Frontal Face Image Rohit patla, Abhishek March 2017 Research gate
- [4] Human Emotion Recognition using Convolutional Neural Network in Real Time Arti mishra, Anushree Desmukh May 2018 IEEE
- [5] Facial Expression Emotion Detection for Real-Time Embedded Systems Saeed Turabzadeh,RafiqM. Swash International Journal of Recent Technology and Engineering (2017)
- [6] Human Emotion Recognition: Review of Sensors and Methods Andrius Dzedzickis Signal Processing, Sensors and Artificial Vision (2016)
- [7] Human Emotion Recognition using Convolutional Network in Real Time Abhishek Adivarekar <u>Arti</u> <u>Mishra</u> IEEE,2018
- [8] Emotion Detection and Recognition from Text Using Deep Learning Chew-Yean 2015 November Devblogs
- [9] Facial Emotion Recognition in Real Time, Dan Duncan 2016 January IEEE
- [10] Facial Expression Recognition with Keras, Abhishek Adivarekar 2018 April IEEE
- [11] Shikha Gupta1, Jafreezal Jaafar2, Wan Fatimah, Wan Ahmad and Arpit Bansal, J. Clerk Maxwell, "Feature Extraction Using MFCC": An International Journal (SIPIJ) Vol.4, No.4, August 2013.
- [12] N.Murali Krishna, P.V. Lakshmi, Y.Srinivas, J.Sirisha Devi, "Classification using SVM" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 1, September 2011.
- [13] Cao Et Al "Human Emotion Recognising Using Hidden Markov Model" Beijing ,June 2014.
- [14] Tin Lay New "Emotion Recognition Using SVM" IEEE conference multimedia expo, July 2006, pp. 1653-1656.
- [15] Kshamamayee Dash, Debananda Padhi, Bhoomika Panda, Prof. Sanghamitra Mohanty "Speaker Identification using Mel Frequency Cepstral Coefficient and BPNN" August 2002.
- [16] Akhilesh Watile, Vilas Alagdeve, Swapnil Jain "Emotion Recognition in Speech by MFCC and SVM" IJSERTC vol 6,no.3, ISSN:2278-7798 March 2017.
- [17] G.Zhou, J.Hansen and J.Kaiser "Non Linear Feature based Classification of Speech under Stress" IEEE trans.speech audio process, vol.9, no.3, pp.201-216, March 2001.

- [18] Sujata Pathak, Arun Kulkarni "Recognizing Emotions from Speech", third international Conference,vol.6 pp.107-109, IEEE May 2011.
- [19] Anjali Shelke and Sonali Joshi "Design of Human Emotion Recognition system from speech using particle Swarm Optimization", Nagpur, March 2014.
- [20] Aparna.P and Shankar.N "Human emotion recognition from speech", Amravathi university, ISSN: 2248-9622, July 2014

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.