ISSN 2348 – 8034
Impact Factor- 3.155

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

MACHINE LEARNING CLASSIFIER FOR EMOTION INTELLIGENCE

Rajni*[1] and Dr. Ajit Singh[2]
*[1]M.Tech (Network Security), BPSMV, India

ABSTRACT

The term "emotional intelligence" does not yet appear in dictionaries. Most of the early research on intelligence focused on problem solving and other things that were easy to measure. In this paper, we have explored speech features to identify is any distinguishable difference in the sentence spoken in different emotions by a group of people. Emotions give instinctive nature to speech. In the proposed work, noticeable dialects of Hindi are considered for the recognizable proof undertaking. The emotions studied in the proposed work are neutral, happy, sad and angry. Prosodic and spectral features extracted from speech are used for discriminating the dialects and emotions. The spectral features of the speech are represented by Mel Frequency Cepstral Coefficients (MFCC) and prosodic features are represented by mode, pitch and energy contours. MFCC as studied by earlier researchers provide 75% efficiency results to get more efficient results is it required to use more features of the sound. This trained dataset is used to study the emotions of the users that can be further used in machine learning.

Keywords- *Emotion Intelligence, Learning Classifier etc.*

I. INTRODUCTION

This emotions insight includes the capacity to screen one's own particular and others' sentiments and feelings, to separate among them, and to utilize this data to guide one's reasoning and activity. Passionate insight addresses the enthusiastic, individual, social, and survival measurements of knowledge. It can be utilized as a part of territories of security, amusement and human machine interface. The feeling acknowledgment normally employments of science picture preparing, discourse handling, motion sign handling and physiological sign handling. **In this paper, algorithm is developed for emotion recognition in the speech processing using the dataset. Machine learning algorithm is developed for better results. This work is to utilize further in robotics and machine learning. So that machines can understand what to do according to the emotions of the human nature.** The emotions can be studied as the study of versification; especially: the systematic study of metrical structure. It is a particular system, theory, or style of versification. These are also called Prosodic features. These features are appeared when we put sounds together in connected speech. It is as important to teach learners prosodic features as successful communication depends as much on **intonation**, **stress** and **rhythm** as on the correct pronunciation of sounds.

II. OBJECTIVES

Speech recognition is the analysis side of the subject of machine speech processing. The synthesis side might be called speech production. These two taken together allow computers to work with spoken language. This study focuses on emotion acknowledgment from a speech. The speech identification, in people, is a huge number of years old. On our planet it could be followed supported a large number of years to the dinosaurs.

- Creating emotional corpus in continuous Hindi speech
- Finding prosodic elements
- Comparative examination of prosodic elements of feelings undertaken
- The first step in any automatic speech recognition system is to find features that are identify the components of the audio signal that are useful for recognizing the semantic substance and tossing the various stuff which conveys data like background noise, emotions and so forth.

The principle point to see about speech is that the sounds produced by a human are separated by the state of the vocal tract including tongue, teeth and so forth. This shape figures out what sound turns out. In the event that we can focus the shape precisely, this ought to issue us a precise representation of the phoneme being delivered. The shape of the vocal tract manifests itself in the envelope of the short time power spectrum, and the job of MFCCs is to accurately represent the envelope.

ISSN 2348 – 8034
Impact Factor- 3.155

III. LITERATURE SURVEY

"Intelligence" did not show up in books before the twentieth century. "Insight" wasn't regular until after 1930.

Generally the hints of spoken languages have been learned at two distinct levels: (1) phonetic segments of spoken words, e.g., vowel and consonant sounds, and (2) acoustic wave designs. A dialect can be isolated into a bit number of vital sounds, called phonemes (English has pretty nearly forty). An acoustic wave is an arrangement of changing vibration designs (by and large in air), in any case we are more habituate to "seeing" acoustic waves as their electrical simple on an oscilloscope (time presentation) or range analyzer (recurrence presentation). Likewise found in stable examination are two-dimensional examples called spectrograms, which show frequency versus time and speak to the sign vitality as the figure force or shading. Daniel Goleman advanced the term 'Passionate Intelligence' in 1995 in the title of his smash hit book, Emotional Intelligence: Why it is more critical than IQ. Goleman characterized enthusiastic insight as 'Understanding one's own emotions, sympathy for the sentiments of others and the regulation of feeling in a manner that upgrades living.' Not everybody concurs with Goleman's model of passionate knowledge, yet there is general assention that passionate insight exists, that it is a variable in individual and expert achievement, and that it can be progressed.

The distribution of Goleman's book Emotional Intelligence in 1995 denoted the start of enthusiastic knowledge as something that was perceived by standard business scholars and authors.

Social-psychological limits help us arrange the social world by illuminating us about the individuals with whom we interface. Out of these limitations our understandings of others' musings and feelings are foremost. Hypothesis of brain (ToM; Wellman, 1990) concerns our gratefulness for individuals' psychological states, for example, convictions and information.Emotion Understanding(EU) alludes to our capacity to recognize obvious passionate responses, to anticipate others' enthusiastic responses, and to admire that individuals have both tangible and private passionate encounters (Denham, 1986; Pons, Harris, & de Rosnay, 2004). Youthful kids' initial comprehension of feelings (e.g., essential full of feeling point of view taking) and comprehension of perception (e.g., false-conviction comprehension) are at first unmistakable limits (Cutting & Dunn, 1999), with diverse relates (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). Then again, ToM and EU are together fundamental for adult social discernment for instance a research of of shrouded feelings (Harris, Donnelly, Guz, & Pitt-Watson, 1986) obliges a joint comprehension of subjective and full of emotions states. Moreover, ToM and EU can be both interestingly and mutually powerful in our thinking and choice making (Pons et al., 2004; Wellman & Banerjee, 1991).

Eric Brill, Radu Florian, John C. Henderson, Lidia Mangu proposed that best in class dialect models for discourse acknowledgment are in light of an exceptionally unrefined semantic model, to be specific molding the likelihood of a word on a little settled number of going before words. Notwithstanding numerous endeavors to join more modern data into the models, the n-gram model remains the best in class, utilized as a part of for all intents and purposes all discourse acknowledgment frameworks. Sameer Maskey, Julia Hirschberg in 2005 introduced aftereffects of an observational investigation of the helpfulness of distinctive sorts of elements in selecting extractive synopses of news telecasts for our Broadcast News Summarization System. Most content based synopsis frameworks depend upon lexical, syntactic, and positional data in figuring out which portions to incorporate in a rundown. They portrayed the element classes we use to foresee sentences to be separated and our system for selecting them, including lexical, auxiliary, prosodic and talk highlights.

In the proposed work, Mel Frequency Cepstral Coefficents (MFCCs) highlight is likewise considered. MFCCs are a component broadly utilized as a part of programmed discourse and speaker acknowledgment. They were presented by Davis and Mermelstein in the 1980's, and have been best in class from that point onward. Before the presentation of MFCCs, Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) and were the fundamental element sort for programmed discourse acknowledgment (ASR).

IV. METHODOLOGY USED

MIRTOOLBOX

MIRtoolbox offers an incorporated arrangement of capacities written in Matlab, devoted to the extraction from sound records of musical components, for example, tonality, rhythm and other measurable investigation on signals.This tool stash proposes a huge arrangement of musical component extractors. MIRtoolbox is a Matlab tool stash devoted to the extraction of musical components from sound documents, including schedules for measurable analysis. The target is to offer a review of computational methodologies in the zone of Music Information Retrieval. The outline is taking into account a secluded structure: the diverse calculations are disintegrated into stages, formalized utilizing a negligible arrangement of rudimentary instruments. These building squares frame the essential vocabulary of the tool kit, which can then be openly explained in new unique ways. These rudimentary components coordinates all the distinctive variations proposed by option approaches - including new methods we have created -, that clients can choose and parametrize. This engineered review of highlight extraction instruments empowers a promotion of the innovation offered by all the option techniques. Furthermore to the essential computational procedures, the toolkit additionally incorporates more elevated amount musical element extraction devices, whose option methods, and their numerous mixes, can be chosen by the client. Olivier Lartillot, Petri Toiviainen, Tuomas Eerola, "A Matlab Toolbox for Music Information Retrieval", in C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker (Eds.), Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and Knowledge Organization, Springer-Verlag in 2008. Olivier Lartillot, Petri Toiviainen, "A Matlab Toolbox for Musical Feature Extraction From Audio",International Conference on Digital Audio Effects, Bordeaux in 2007. MIRtoolbox is taking into account a situated of building hinders that can be parametrized, reused, reordered etc.

V. RESULTS

The pitch values for the neutral and sad emotions are likely to be same. Entropy values for the angry sounds are higher as compared to sad and neutral emotions. The proposed system have found that the emotions, Neutral, joy and anger, are portrayed at a higher frequency than emotions such as sadness

- Anger: Anger can be divided into two types: "anger" and "hot anger". In comparison to neutral speech, anger is produced with a lower pitch, higher intensity, more energy (500 Hz) across the over the vocalization, higher first formant (first solid created) and speedier attack times at voice onset (the start of speech). "Hot anger", in contrast, is produced with a higher, more varied pitch, and even greater energy (2000 Hz).

- Happy: Fear can be divided into two types: "joy" and "anxiety". In comparison to angry speech, happy emotions have a higher pitch, little variation, higher energy, and a slower speech rate with more pauses.

- Sadness: In comparison to neutral speech, sad emotions are produced with a slightly higher pitch, less intensity but more vocal energy (2000 Hz), longer duration with more pauses, and a lower first formant.

MFCCs Results

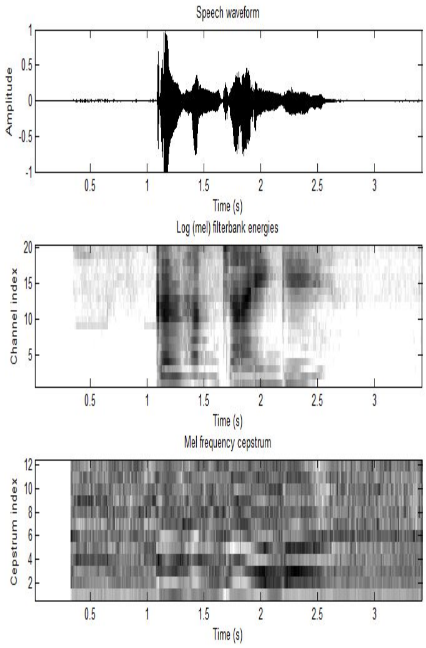

Figure: MFCC for the sound of first user

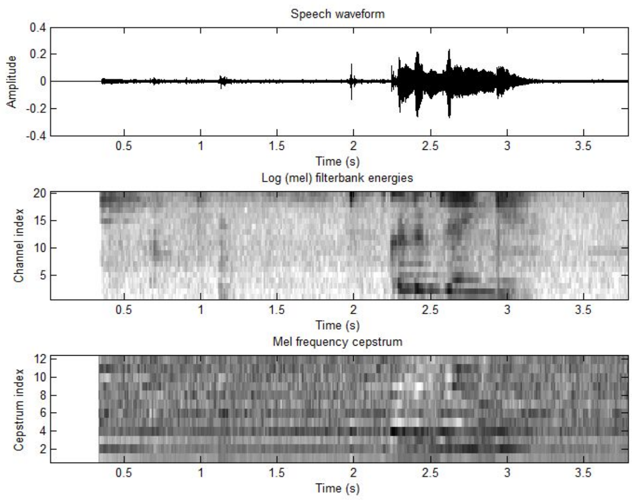

Figure: MFCC for the sound of second user

VI. CONCLUSION

MFCC offers a description of the spectral shape of the sound. The frequency bands are positioned logarithmically (on the Mel scale) which approximates the human auditory system's response more closely than the linearly-spaced frequency bands. This features as studied earlier are also considered in the proposed work are not providing good results. From the features extraction from the sound using MIRToolbox a machine can find the differences between the speeches of the person in contrast to his emotion. This work is proposed to be used by machines to understand what a person wants exactly to be performed. The comparison between basic concepts of emotions (happy, sad, tender, anger, fear) and emotion dimensions (activity, valence, tension).The pitch values for the neutral and sad emotions are likely to be same. Entropy values for the angry sounds are higher as compared to sad and neutral emotions. Way of speaking of all the persons has slight difference. Some persons speak very lightly in sad mood. Some speaks in the similar tone as in neutral form. Trained dataset records the way of talking of the users. This trained dataset is used to study the emotions of the users that can be further used in machine learning.

REFERENCES

1. *K Sreenivasa Rao and Shashidhar G Koolagudi " Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech", SYSTEMICS, CYBERNETICS AND INFORMATICS VOLUME 9 - NUMBER 4 - YEAR 201126 ISSN: 1690-4524 .*
2. *Vayrynen,Eero, "Emotion recognition from speech using prosodic features". University of Oulu Graduate School" , University of Oulu, Faculty of Information Technology and Electrical Engineering, Department of Computer Science and Engineering; Infotech Oulu Acta Univ. Oul. C 487, 2014.*
3. *Somnath Roy, Jawaharlal Nehru University Centre for Linguistics Jnu, New Delhi-110067 Nishant Sinha IMS Engineering College CSE Department Ghaziabad UP, " Duration Modeling in Hindi " International Journal of Computer Applications (0975 – 8887) Volume 97– No.6, July 2014.*
4. *Leena Marya,*, B. Yegnanarayanab, " Extraction and representation of prosodic features for language and speaker recognition", Received 6 April 2006; received in revised form 20 February 2008; accepted 24 April 2008.*
5. *Chung, S., Hirose, K., and Minematsu, N., "Improvement of N-gram language modeling of Japanese using BUNSETSU boundary information and its application to large vocabulary continuous speech recognition," Report for the Spring Meeting, ASJ, 1, 65-66, 2004. (in Japanese).*
6. *Hirose, Y., Ozeki, K., and Takagi, K., "Effectiveness of prosodic features in dependency analysis of read Japanese sentences," Natural Language Processing, 8 (4), 71-89, 2001.*
7. *Inoue, A., Mikami, T., and Yamashita, Y., "Prediction of sentence importance for speech summarization using prosodic features," Proc. Eurospeech, Geneva, 1193-1196,2003.http://www.speech-recognition.de/textbook.html (2003 – 05-18).*
8. *Longbiao Wang, Norihide Kitaoka, Seiichi Nakagawa, "Robust Distant Speaker Recognition Based on Position Dependent Cepstral Mean Normalization", 2005.*
9. *"Dynamic time warping", 16 Dec 2006, http://en.wikipedia.org/wiki/Dynamic_time_warping.*