IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, November 2022
Impact Factor: 6.252

Speech Emotion Recognition using Machine Learning

Jordan S. Choudhari[1], Rajdatt N. Kokate[2], Pratham Murkute[3], Anjali Ghopane[4], Noorin Khot[5]
Assistant Professor, Department of Computer Engineering[1]
Students, Department of Computer Engineering[2,3,4,5]
Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract: *In day to day life Language is an essential requirement for humans to communicate, with speech serving as the primary medium. The human voice having so many attributes such as pitch, vocal, accent and etc. Emotion recognition is technology that extracts emotion features from computer speech signals, compares them, and analyses the feature parameters and the obtained emotion changes. Like someone is clam and next moment they will be excited then also the emotion is changes. Speech emotion recognition is a trending research topic these days, with its main motive to improve the human machine interaction. over an acoustic channel This is what we're talking about here, Speech recognition entails digitizing and converting sound waves into basic language units called phonemes, creating words from phonemes, and contextually assessing the words to ensure correct spelling for words that sound similar. Speech Recognition refers to a computer's capacity to detect a caller's responses and move them along the cell's flow. You can use your voice to provide input to an application using speech recognition.*

Keywords: Speech Signal, Communication, interaction, MFCC(Mel Frequency CepstralCoeff), CNN, DSCNN, Berlin, Training & Testing , Emotion Recognition

I. INTRODUCTION

Speech is the most crucial , widespread and proficient form of communication method for people to commune with each other. Speech emotion recognition can play an effective role in the daily human-machine interaction and increases the accuracy, speed and intimacy of interaction. Human are comfortable with speech hence person would also like to communicate with machine via speech, rather than via primitive interface such as keyboard and mouse pointer.[1]
The main challenges in this project is detection of emotion in speech.
A robust device has the capacity to improve or decrease the dependability of standard keyboard and mouse input. There are numerous tools for emotion detection. Basically, it involves speaking to a computer and having it successfully understand what you're saying by allowing the software to respond as it should be or by converting the input speech to another form of verbal communication that is similarly perceptible by all other software that could manner it well and provide the user with the desired result.[3]

II. LITERATURE SURVEY

HAN Zhiyan, WANG Jian College of Engineering, Bohai University, Jinzhou 121013 in their paper[1] titled "Speech Emotion Recognition Based on Deep Learning and Kernel Nonlinear PSVM" In this paper Deep Neural Network (DNN) and Extreme Learning Machine (ELM) and Proximal Support Vector Machine (PSVM) is used. Feature extracted in two ways integration of DBN automatic features and traditional features. DBN error rate is less than traditional features that's why is able to extract some meaningful features and the whole extraction process will be very close to human emotion recognition. For better recognition of emotion they combine facial expression with speech signals.
Ashwin V. Gatty, G. S. Shivakumar, Kiran Shetty in their paper [2] titled "Speech Emotion Recognition using Machine Learning" they focus on understanding the word on sound Frequency & also Lusing the inbuilt marco model we can identify similar words & also can correct the grammar. The modelling of speech units and grammar using the hidden Markov model and neural networks is emphasized. You can use your voice to provide input to an application using speech recognition . In future we can use it for multiple Support multiple language.[2]

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, November 2022
Impact Factor: 6.252

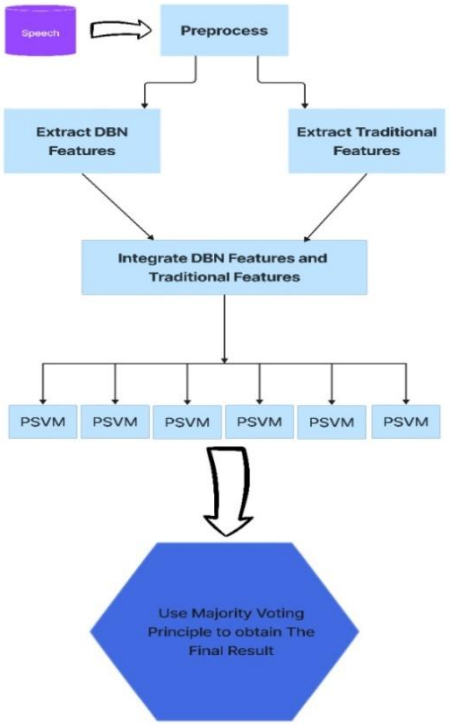

Figure 1: (Courtesy) Speech Emotion Using PSVM[1]

Chaitanya Singlaa, Sukhdev Singhb, Monika Pathakc in their paper [3] Titled "Automatic Audio Based Emotion Recognition System" they looking on the application of this would be vast in areas to converse emotion conversion from audio but challenges will be here like noise, disruption & Various gaps in communication that could degrade the accuracy or change the meaning of statement. Making a machine grasp the emotion of the responses is a difficult undertaking.[3]

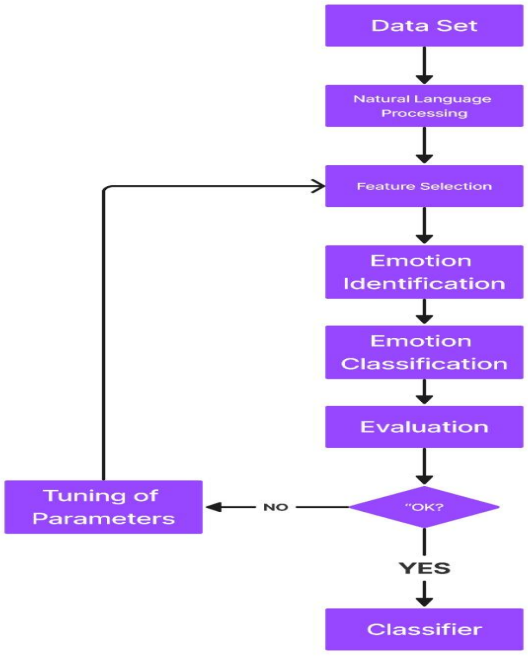

Figure 2: (Courtesy) Automatic Audio Based Emotion Recognition System[3]

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, November 2022
Impact Factor: 6.252

Ajay Gupta, Siddhesh Morye, Mukul Sitap, Supriya Chaudhary in their paper [4] the Speech of in which the Speech Could be slow or fast so data base model should be trained accordingly to do this paper we have CNN & SVM is discussed. It will keep two emotion models trained & test to predict the correct emotion. The proposed approach is based upon the Mel Frequency Cepstral coefficients (MFCC) uses Crema-D database of emotional speech. Data Augmentation is performed on input data audio file, such as Noise, High Speed, Low Speed etc. are added. As we are training data model in future the result will be accurate.

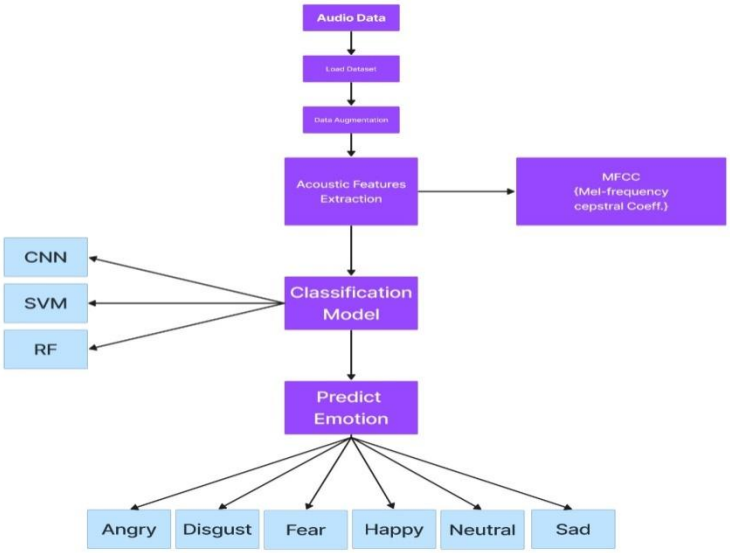

Figure 3: (Courtesy) Speech Based Emotion Using ML[4]

III. COMPARISON BETWEEN LITERATURE PAPERS

Features	[1]	[2]	[3]	[4]
Purpose	Better recognition of emotion	System is extended to integrate with the robot.	Decoding emotion from regional language	Predict more human emotion
Algorithm	Deep Neural Network (DNN) and Extreme Learning Machine (ELM) and Proximal Support Vector Machine (PSVM)	Support vector Machine (SVM)	Convolutional Neural Network (CNN)	CNN SVM RF
Focus on	Detect facial Expression	Speech speed	understanding the word on sound Frequency	Emotion conversion from audio

IV. PROPOSED SYSTEM

As we know that during all speech recognition accuracy is main factor which is matter for the user when use the system. We the help of CNN we have great advantage of classify many type of emotion at same time. We have to develop our system in such way that time complexity should be minimize. Implementation with any hardware is possible.

V. CONCLUSION

From this all paper we studied that accuracy is main factor which affect in speech emotion recognition system which is previously made by other people. It's all about using different types of algorithm to increase efficiency of system.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 3, November 2022
Impact Factor: 6.252

REFERENCES

[1]. Saikat Basu, Jaybrata Chakraborty, Arnab Bag and Md. Aftabuddin." A Review on Emotion Recognition using Speech" (IEEE 2017)

[2]. Ashwin V. Gatty, G. S. Shivakumar, Kiran Shetty. "Speech Emotion Recognition using Machine Learning"(IJRESM 2021)

[3]. Chaitanya Singlaa, Sukhdev Singhb, Monika Pathakc "Automatic Audio Based Emotion Recognition System: Scope and Challenges" (2021 IEEE)

[4]. Ajay Gupta, Siddhesh Morye, Mukul Sitap, Supriya Chaudhary. "Speech based Emotion Recognition using Machine Learning" (IRJET 2021)