

Threat Prediction using Speech Analysis

Shreyas Kulkarni¹, Nirantar Kulkarni², Sushil Kukreja³, Kaustubh Kotkar⁴, Shambhavi Kulkarni⁵, Jayanti Kamalasekaran⁶

^{1,2,3,4}B.E. Student, Dept. of Computer Engineering, Sinhgad College of Engineering, Vadgaon, Pune- 411041, Maharashtra, India

Abstract -Speech recognition technology is one of the fast growing Engineering technologies. It has number of applications in different areas and provides potential benefits. Audio surveillance is getting more important and requires large amount of man hours to determine threats from audio. The aim of this system is to identify potential cases of threats, and provide an early warning or alert to such cases. This will be based on voice such as voice chat over telecommunication networks or social media. The intended result will be achieved in the three major steps. This system will utilize latest machine learning algorithms to identify phonetics from the audio and helps convert it to equivalent transcript form. Natural language processing will help in developing sentiment analysis model.

Key Words: Recurrent Neural Network, Sentimental Analysis, Speech Recognition, Natural Language Processing.

1. INTRODUCTION

There has been much advancement in the areas of education, healthcare, etc. However, crime is one of the major issues faced by societies all over the world. The crime investigation team follows a systematic procedure for solving a particular case. At first, there are multiple suspects. Hence, to reach to a conclusion, the police might have to listen to their call recordings. This procedure consumes a lot of time as a single person is required to listen to multiple recordings. This also has a scope for human errors. These errors and time can be minimized using our proposed system. In this system, the user needs to provide an audio as an input and the system displays a warning about threat present in the audio. There are 2 parts in this project. First part converts speech to text and the second part applies sentimental analysis to the text. The total percentage of threat is calculated and if it is above a certain threshold, warning is provided. This system can also be used separately, i.e. only the first part can be used by the user if he/she doesn't want sentimental analysis on the text and the second part can be used be directly providing text to the system to analyze the sentiments in a particular text. Speech to text is converted using signal processing and sentimental analysis is carried out using Naïve Bayes algorithm and neural network. This system is advantageous as it provides the results of the recording directly without a person having to analyze it for hours.

1.1 Literature Survey

1. English Language Speech Recognition using MFCC and HMM:
This model implements and exploits Mel Frequency Cepstral Coefficient and Hidden Markov Model Techniques to evaluate the competence of speech recognition. In this paper they have used data from Google audioSet and converted audio signal into a vector using MFCC. The research is done using Python Programming with the help of Librosa library to generate MFCC.
2. Speech Recognition with Deep Recurrent Neural Network:
In this paper, they have implemented speech recognition technique using Recurrent Neural Network which will also take into consideration context of the sentence. They have taken TIMIT dataset to perform phoneme recognition. This model combines deep, bidirectional Long Short-term Memory RNNs with end-to-end training.
3. Speech Recognition using Deep Learning:
In this paper, audio files are taken as input and Artificial Neural Network is used to predict speech as an output. This research uses Deep Learning for speech recognition using library from Google which has 66.22% accuracy.
4. Sentimental Analysis and Prediction using Neural Networks:
This paper classifies sentences into positive, negative and neutral categories. It uses Artificial Neural Network for sentimental analysis by training and testing data. It also checks the accuracy of ANN for large datasets.

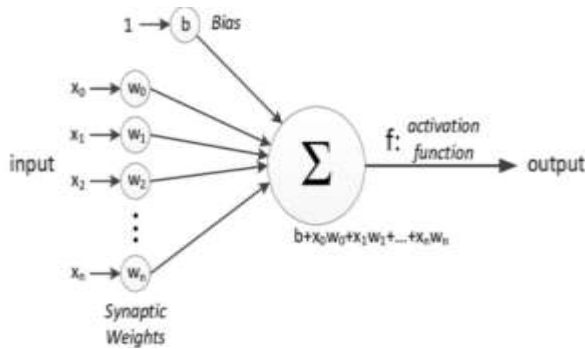
2. Project Overview

Supervision and analysis of audio files manually takes a lot of time. This procedure is very tedious and leaves scope for a lot of human errors. These errors might manipulate the results on a large scale. We have proposed a system that eliminates the human work of supervising the audio files. In sectors like crime investigation, this system might prove to be advantageous as threat in a particular audio file will be analyzed automatically and a warning will be provided as output. This system can also be used separately by the user or actor. The developed model will help the will be of help to various organizations as they can either use it for speech to text conversion or

sentimental analysis or both. There is one user/actor that can use the system by providing an audio file as input.

[A] Artificial Neural Network

Artificial Neural Network is used for processing the information similar to how brain works. It has large number of small processing elements connected to each other, working together to solve a specific problem.



In the above figure, inputs are basically independent variables which are multiplied by their weights and summed up to get the activation function. This function takes decision on whether neuron should be activated or not.

For learning of neural network, it uses the cost function which is the difference between the actual output and expected output. This function is analyzed and proper changes are made to the inputs accordingly. Cost function is reduced to minimum value to get best learning results. We use back propagation on the network continuously till error becomes minimal.

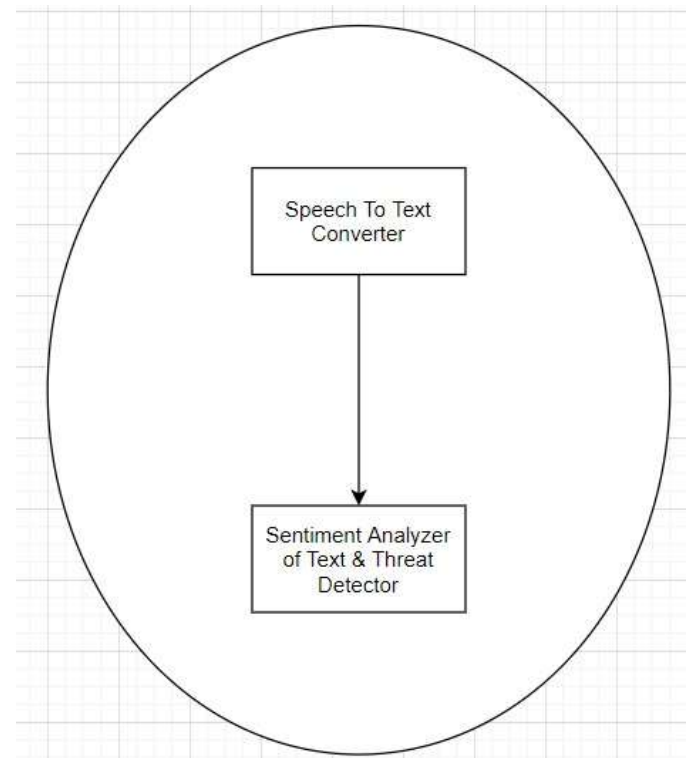
[B] System Architecture Diagram

Large number of audio files are given as input to the system. These audio files are converted to the digital data using signal processing. Then feature extraction and feature selection is performed on digital data. We will get the text output by applying ANN algorithm on those selected features. Text is classified into positive, negative or neutral sentiment using Naïve Bayes classifier. The polarity of the text input is calculated by using ANN on the classified data.

[C] Sentimental Analysis

Sentiment analysis is used to detect threats and give the warning. It is used to analyse the tone of the text and classify it into positive or negative. This process is very useful to detect the threat before so that it can be eliminated or damage can be reduced. Textual data is analysed and the polarity of the text is calculated. All the data is scaled using the data mining methods like standard scalar. Then the data is given to ANN to train the model. This model is supervisedly trained to get the threat warning for the input text.

[D] System Architecture Diagram



3. CONCLUSION

The developed model will help revolutionize audio surveillance as time required to listen and then analyze the audio is highly reduced. Human errors such as skipping some part of audio will also be reduced. Available API's such as Google API may become paid in future and hence our system provides a way to convert speech to text without using Google API.

REFERENCES

- [1] Phoemporn Lakkhanawannakun, Chaluemwut Noyunsan, "Speech Recognition using Deep Learning," Rajamangala University of Technology, Thailand.
- [2] Kanchan Naithani, V.M. Thakkar, Ashish Senwal, "English Language Speech Recognition using MFCC and HMM," GBP Institute of Engineering Technology, Uttarakhand, India.
- [3] Sneha Paliwal, Sunil Khatri, Mayank Sharma, "Sentiment Analysis and Prediction using Neural Networks," Amity University, Noida, Uttar Pradesh, India.