SENTIMENT ANALYSIS BY AUDIO SPEECH USING TEXTBLOB

Anil Kumar Ahlawat
Professor, Computer Science and Engineering
KIET Group of Institutions, Ghaziabad, U.P., India
anil.ahlawat@kiet.edu

Vibhanshu Verma
KIET Group of Institutions
Ghaziabad, U.P., India
vibhanshu.1923cs1128@kiet.edu

Chetan Shukla
KIET Group of Institutions
Ghaziabad, U.P., India
chetan.1923cs1038@kiet.edu

Pallavi Verma
KIET Group of Institutions
Ghaziabad, U.P., India
pallavi.1923cs1189@kiet.edu

Anand Gupta
KIET Group of Institutions
Ghaziabad, U.P., India
anand.1923cs1140@kiet.edu

Abstract— *Sentiment Analysis, a task in Natural Language Processing, involves identifying whether a text expresses objective or subjective information and the sentiment it conveys - positive, negative, or neutral. It is a powerful tool for gauging people's reactions to a particular post or product.The primary focus of this paper is to examine the utilization of machine learning techniques for sentiment analysis and opinion mining of textual data.*

Integrating machine learning techniques with sentiment analysis can aid in forecasting customer opinions and product reviews related to recently introduced items. In this paper, there's the process involves converting audio and video inputs into text format and subsequently analyzing them using appropriate machine learning methods. The paper conducts an extensive examination of different machine learning approaches utilized for sentiment analysis, assessing their accuracy, benefits, and constraints.

Keywords — Natural Language Processing, objective information, subjective information, machine learning.

1. Introduction

Sentiment analysis has numerous practical applications, including identifying customer sentiments, predicting consumer preferences, and analyzing social media sentiments. As social media continues to grow in importance, the use of sentimental analysis is becoming increasingly vital for businesses to stay competitive. Therefore, further research is needed to develop more innovative and effective techniques for analyzing sentiments and addressing the current challenges, such as identifying indirect opinions, comparative sentences, and sarcastic expressions.

The fast-paced expansion of e-commerce platforms has allowed individuals to incorporate the internet into their everyday routines. People have become accustomed to browsing through user-generated content, such as product reviews and social media posts, which provide insights into people's opinions, emotions, attitudes, and behaviors. In this context, sentiment analysis has emerged as a powerful technique for analyzing and understanding how sentiment is expressed in text.

There are three classification levels in sentiment analysis:

- Document-level Classification
- ·Sentence-level Classification
- Aspect-level Classification

In the document level classification, it aims to classify the opinion as positive or negative overall, is one approach involves considering the complete document as a unified entity. This approach assumes that the document represents a cohesive and consistent piece of writing. The purpose of the sentence-level analysis is to classify the emotions that are being expressed in the various sentences. The first stage in determining whether a sentence is subjective, or objective is at the sentence level. The sentiment analysis algorithm identifies whether a sentence expresses a positive or negative opinion, depending on its subjectivity. The aspect-level analysis aims to classify the sentiment towards specific aspects or items. [1]

It is crucial to acknowledge that audio-based sentiment detection focuses on distinguishing between positive and negative opinions, which is different from speech emotion identification. Researchers in this field commonly use a tandem system architecture, which involves converting speech to text using automated speech recognition (ASR) technology and then applying traditional text-based sentiment detection algorithms to the resulting text. In this way, the speech recognition system's output is analyzed by the text processing system for features (words, phrases, etc.) that convey sentiment. Our approach to sentiment extraction involves using two main systems: automatic speech recognition (ASR) and a text-based sentiment extraction system. We've created a unique way to analyze text and determine the emotions behind it. Our approach involves using part-of-speech tagging to identify key elements of the text, and then using maximum entropy modeling to predict whether the sentiment expressed is positive or negative based on those elements. [2]

Our method utilizes audio to extract emotional information, where we employ a machine learning algorithm to analyze text data derived from the audio. A distinguishing feature of our proposed system is the capability to identify the specific contributions of individual text features to sentiment analysis or estimation.

The proposed system involves capturing audio data from various sources such as speakers, customers, and others, it is then transformed to text. Emotional information is extracted from the text data through sentiment analysis.. After applying the audio sentiment analysis with TextBlob and Speech Recognition, the output of the system is presented as a pie chart that shows the percentage of positive, negative, and neutral sentiment expressed in the text data.

2. Literature Survey

The task of sentiment analysis is vital in natural language processing as it entails detecting the emotional and opinionated aspects conveyed in different mediums like text, speech, and images. It is a complex task that involves various algorithms and techniques to produce accurate reviews of products or feedback on services. Many researchers have conducted studies to improve the accuracy and effectiveness of sentiment analysis using diverse sources, including audio input from conversations, video transcripts, images, and text data from social networking sites or e-commerce platforms. [3]

One such study introduced a generalized model that uses audio input from conversations between two individuals to analyze the speakers' identity and content automatically. The system transcribes the audio to text and performs speaker recognition to identify the emotions and opinions expressed in the conversation. Although the current system has some limitations, the upcoming research aims to address these issues and enhance the system's precision and scalability. [4]

Another study uses data from YouTube to identify sentiment in unscripted, natural speech. The system employs Maximum Entropy (ME) model tuning and feature selection techniques to extract video transcripts and enhance the models' accuracy and domain independence. The study finds that sentiment can be identified in natural audio with high accuracy and produces suitable keywords and tags for use in YouTube videos. [5]

The paper proposes a new methodology for image sentiment analysis that utilizes latent correlations between textual, visual, and sentiment aspects of training images. The approach leverages interrelationships between these different views of the same image data to enhance the accuracy and effectiveness of image sentiment analysis. The study uses a multi-view canonical correlation analysis framework (CCA) to identify the correlations between these three views, resulting in a latent embedding space that maximizes the correlations between the three perspectives. [6]

The recognition of nonverbal cues such as intonation and facial expressions is essential in determining the speaker's true intentions, and is considered a significant factor in determining whether a speaker is expressing their original thoughts or merely repeating someone else's ideas. This aspect of analysis is referred to as multimodality and has broad applications, including lie detection, interview and interrogation analysis, among others. However, more research is needed to enhance the accuracy and effectiveness of Multimodal Sentiment Analysis.[7]

In another study, the authors investigated sentiment analysis of film user reviews, utilizing the Internet Movie Database (IMDb) dataset. They employed Naive Bayes and linear SVM to classify the dataset after data collection and preprocessing to detect opinions, using the Porter stemming algorithm to remove English word suffixes. The model's performance was evaluated using a confusion matrix to determine accuracy, precision, and recall. The study found that drama films were preferred by users based on sentiment orientation, and different words were classified based on polarity.

Lastly, an article provides an extensive comparison of various machine learning techniques, evaluating their accuracy, benefits, and drawbacks. The paper argues that current approaches to sentiment analysis need to be more advanced and effective in dealing with issues such as classifying indirect opinions, comparative sentences, and sarcastic sentences. These studies and their findings emphasize the importance of sentiment analysis in various domains and the need to continuously improve its accuracy and effectiveness.

Sentiment analysis is a crucial area of research in natural language processing that involves identifying emotions and opinions expressed in text data. Recent developments in deep learning techniques, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated promising results in analyzing sentiment by extracting complex features from input data. Word embeddings, like Word2Vec and GloVe, have also played a significant role in improving the accuracy of sentiment analysis by representing words as dense vectors, capturing their semantic relationships.

Sentiment analysis has many practical applications in diverse fields, including marketing, customer service, and politics. In marketing, it helps understand customers' preferences and opinions about products and services, leading to data-driven decision-making and improved customer satisfaction. In customer service, sentiment analysis automatically classifies and prioritizes customer feedback, allowing for efficient responses and improved customer experience. In politics, sentiment analysis can analyze public opinion and detect trends, helping politicians understand their constituents' needs and concerns.

However, sentiment analysis still faces challenges such as language ambiguity, sarcasm, and cultural differences. Additionally, ethical considerations like privacy concerns and potential biases in the analysis process must be taken into account.

In conclusion, sentiment analysis is a vital component of natural language processing, with numerous applications in various fields. Ongoing research and technological advancements can improve its accuracy and effectiveness, addressing current challenges and providing valuable insights into human communication and behavior.

3. Proposed System

In this cutting-edge system proposal, we revolutionize sentiment analysis by incorporating audio input from customers and users in real-time. By saving the audio as a .wav file and extracting the corresponding textual portion, we can then leverage a powerful NLP-based library to determine the sentiment behind each recorded speech. With this breakthrough approach, we aim to provide more accurate and reliable insights into the opinions and feelings of individuals, paving the way for more effective decision-making and improved customer experiences.

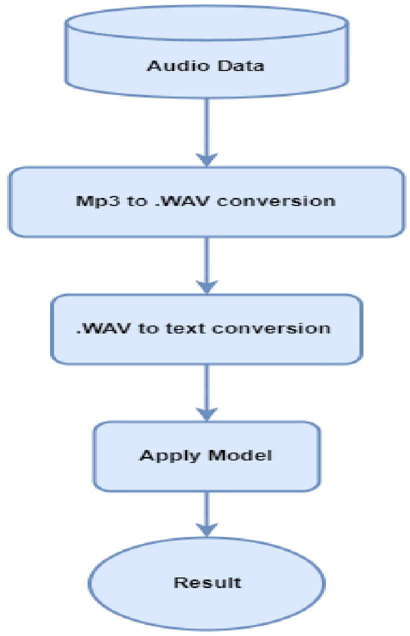

Fig-1: Proposed system

4. Working

In this model, used the NLP(Natural Language Processing) and AIML(Artificial Intelligence and Machine Learning) based library named as :

- Automatic Speech Recognition
- TextBlob

The above Libraries were described below.

4.1. Automatic Speech Recognition

Automatic Speech Recognition (ASR) is a technology that uses Machine Learning and Artificial Intelligence (AI) to convert human speech into written text. It has experienced tremendous growth in recent years and is now found in many popular applications, including social media platforms, podcast players, and video conferencing software, where it is used for real-time transcription and closed captioning. With ASR technology becoming increasingly accurate, it is anticipated that many more applications will begin incorporating it to improve the accessibility of audio and video data. Additionally, affordable and reliable Speech-to-Text APIs like AssemblyAI are making it easier for developers to access and integrate ASR technology into their products.

4.2. TextBlob

This Python library referred to is widely recognized for its efficient handling of textual data, particularly in the realm of natural language processing (NLP), where it is commonly used for tasks like sentiment analysis, part-of-speech tagging, and language translation. It provides a simple and easy-to-use API for performing various NLP operations on text data.

TextBlob uses the Natural Language Toolkit (NLTK) library as its base and extends its capabilities to provide an even more user-friendly and intuitive interface. It also incorporates pattern.en, a Python library for web mining and natural language processing, and the WordNet lexical database.

TextBlob's sentiment analysis module is particularly noteworthy, as it provides a pre-trained model for analyzing the sentiment of textual data, classifying it into positive, negative, or neutral. This feature makes it a popular choice for sentiment analysis tasks in social media monitoring, customer feedback analysis, and other similar applications.

Overall, TextBlob is a useful and versatile library for NLP tasks in Python, particularly for those who are new to the field and want an easy-to-use interface for performing text analysis tasks.The sentiment of the text is either positive or neutral or negative which was considered from the Table 1 mentioned below.

Sentiment Metric	Value
Positive	1
Neutral	0
Negative	-1

Table:-1 Sentiment Metric

	Precision	Recall	f1 score
Negative	0.67	0.48	0.56
Positive	0.25	0.59	0.35
Neutral	0.26	0.15	0.19

Fig-2 : TextBlob Sentiment classification

5. Result and Discussion

The accuracy of audio sentiment analysis using TextBlob, a popular Python library for natural language processing, is influenced by various factors. The quality of the audio data, the sentiment labeling process employed, and the linguistic complexity and variability of the audio data are all essential considerations. TextBlob's sentiment analysis function determines the polarity score of a text fragment by detecting certain keywords and phrases. However, when it comes to audio data, the effectiveness of TextBlob may be limited due to the absence of visual cues and the nuances of spoken language. Therefore, while TextBlob can be a useful tool for sentiment analysis of textual data, it may not be the most effective option for audio data.

To overcome the limitations of TextBlob and other existing sentiment analysis models, researchers may explore alternative approaches such as deep learning and multimodal analysis. Multimodal sentiment analysis involves incorporating data from multiple modalities, such as audio and visual information, to gain a more comprehensive understanding of human emotions and intentions. Additionally, deep learning models, such as convolutional neural networks and recurrent neural networks, have shown promise in enhancing the accuracy and precision of sentiment analysis, particularly in complex and dynamic data such as audio.

Furthermore, ethical considerations must be taken into account when using audio sentiment analysis. Privacy concerns and potential biases in the analysis process are among the challenges that need to be addressed. Collaboration between experts in natural language processing, machine learning, psychology, and other related fields is essential to address these challenges and advance the field of audio sentiment analysis.

In conclusion, while TextBlob can be useful for sentiment analysis of textual data, its effectiveness may be limited when applied to audio data. To improve the accuracy and precision of audio sentiment analysis, researchers may explore alternative approaches such as deep learning and multimodal analysis. Ethical considerations must also be taken into account to ensure the responsible and ethical use of audio sentiment analysis. Collaboration and interdisciplinary research will play a vital role in advancing the field of audio sentiment analysis and unlocking its potential applications.

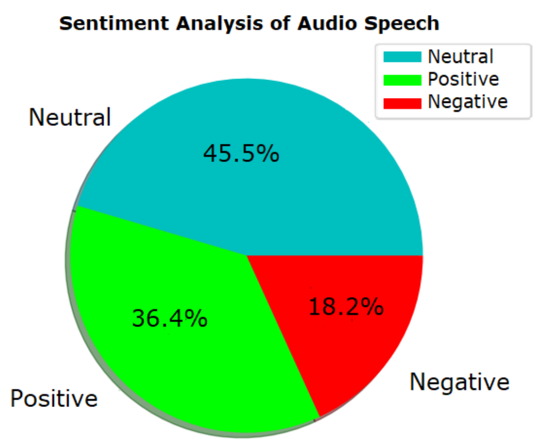

Fig-3 : Pie chart of sentiment Analysis on the audio of the user.

The above pie chart depicts the sentiment in the form of positive, neutral and negative percentages. The cyan color gives the percentage of Neutral. The red color gives the percentage of Negative. The purple color gives the percentage of Positive. By this, the author determines the sentiment of the person in real time.

```
positive ===> Today is a good day
###########################################
neutral ===> well done
###########################################
neutral ===> keep it up
###########################################
positive ===> Always be positive
###########################################
neutral ===> never do that again
###########################################
negative ===> highly disappointing
###########################################
negative ===> i hate you
###########################################
neutral ===> leave me alone
###########################################
positive ===> Have a nice day
###########################################
neutral ===> always be the same
###########################################
positive ===> i love you
###########################################
{'neutral': 5, 'positive': 4, 'negative': 2}
```

Fig-4 : Sentences to be taken as input from the user.

The above sentences were taken as the input from the user considering all the emotions present in the text like the word good present in the first sentence gives a positive impact, and the other values like hate, disappointing gives negative and neutral emotion. The textblob is extracting the words and converting it in numbers to show the outcome of the emotions. The sentences present describe all kind of emotions to show that the project is working properly.

6. Conclusion

Analyzing user reviews of products is crucial for both consumers and manufacturers.

Audio monitoring is now required by social media sites like Facebook, Twitter, and YouTube. This study utilizes audio files from customers and users, which are converted into text data before sentiment analysis is applied. The TextBlob NLP-based Library is used in the system's functionality. Furthermore, this system does not require its own set of training data.

7. Future Work

Audio sentiment analysis is a rapidly evolving field with significant potential for future research and development. To enhance the accuracy of sentiment analysis models, future work may focus on integrating contextual factors, combining information from multiple modalities, adapting models to different languages and cultures, and exploring deep learning techniques. Applying sentiment analysis in real-world scenarios can demonstrate its practical applications, while continued development can lead to better tools for analyzing and interpreting audio data, ultimately improving our understanding of human emotions and ability to communicate effectively.

8. References

[1] Monali Yadav, Shivani Raskar, Vishal Waman and Prof. S.B.Chaudhari," Sentimental Analysis on Audio and Video using Vader Algorithm",2019 IRJET

[2] Dr. Ashwini Rao, Akriti Ahuja, Shyam Kansara and Vrunda Patel, "Sentiment Analysis on User-generated Video, Audio And Text" 2021 ICCCIS

[3] Chhaya Chauhan, SmritiSehgal," SENTIMENT ANALYSIS ON PRODUCT REVIEWS", 2017 IEEE

[4] S. Maghilnan and M. R. Kumar, "Sentiment analysis on speaker specific speech data," 2017 I2C2

[5] Lakshmish Kaushik, Abhijeet Sangwan, John H. L., Hansen ," SENTIMENT .EXTRACTION FROM NATURAL AUDIO STREAMS"2013 IEEE

[6] Marie Katsurai and Shin'ichi Satoh, "Image Sentiment Analysis Using Latent Correlations Among Visual, Textual, And Sentiment Views", 2016 IEEE

[7] Aishwarya Murarka, Kajal Shivarkar, Sneha, Vani Gupta,Prof. Lata Sankpal" Sentiment Analysis of Speech", 2017 IJARCCE

[8] K. Saranya and S. Jayanthy, "Onto-based sentiment classification using machine learning techniques," 2017 ICIIECS

[9] Shweta Rana and Archana Singh," Comparative Analysis of Sentiment Orientation Using SVM and Naïve Bayes Techniques",2016 IEEE

[10] Bhavitha B K, Anisha P Rodrigues and Dr. Niranjan N Chiplunkar," Comparative Study of Machine Learning Techniques in Sentimental Analysis", 2017 IEEE