A

Project Report

on

SENTIMENT ANALYSIS BY AUDIO SPEECH USING TEXTBLOB

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2022-23

in

Computer Science and Engineering

by

Anand Gupta (1900290100026)

Chetan Shukla (1900290100050)

Pallavi Verma (1900290100097)

Vibhanshu Verma (1900290100187)

Under the supervision of

Prof. Anil Ahlawat

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May, 2023

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Date: 27/05/2023

Name: Anand Gupta
Roll No.: 1900290100026

Name: Chetan Shukla
Roll No.: 1900290100050

Name: Pallavi Verma
Roll No.: 1900290100097

Name: Vibhanshu Verma
Roll No.: 1900290100187

CERTIFICATE

This is to certify that Project Report entitled "Sentiment Analysis by Audio Speech using TextBlob" which is submitted by Pallavi Verma, Vibhanshu Verma, Chetan Shukla, and Anand Gupta in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

.

Date: 27/05/2023

Prof. Anil Ahlawat

(Dean Academics)

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude to Prof. Anil Ahlawat, Department of Computer Science & Engineering, KIET, Ghaziabad, for his constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Head of the Department of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date: 27/05/2023

Signature: Anand Gupta
Name : Anand Gupta
Roll No.: 1900290100026

Signature:
Name: Vibhanshu Verma
Roll no.: 1900290100187

Signature:
Name : Chetan Shukla
Roll No.:1900290100050

Signature:
Name : Pallavi Verma
Roll No.: 1900290100097

ABSTRACT

Sentiment analysis is a crucial area of research that seeks to glean feelings and viewpoints from textual information. However, there is a growing demand to create sentiment analysis methods that can directly interpret audio speech due to the prevalence of audio-based information. In order to analyze sentiment from audio speech, this research provides a novel method that makes use of TextBlob, a well-known natural language processing toolkit.

A broad collection of audio speech samples, including conversations, interviews, and speeches, encompassing a range of themes and emotions, is gathered for the project's initial stage. Automatic speech recognition (ASR) algorithms are used to convert the audio data into text. The preprocessed text is subsequently cleaned up to get rid of extraneous words, punctuation, and information.

The preprocessed text is then subjected to sentiment analysis using the Python module TextBlob. TextBlob offers a straightforward and user-friendly API with features including subjectivity analysis, polarity recognition, and part-of-speech identification. Sentiment scores are created by applying TextBlob to the transcribed text, representing the positivity or negativity of the expressed sentiments.

The research also investigates the integration of machine learning techniques to improve the precision and dependability of the sentiment analysis. Word frequencies, n-grams, and syntactic patterns are just a few of the features taken from the translated text and used to train a sentiment classifier on a labeled dataset. The sentiment analysis model is then effectively expanded in terms of its use cases and applicability by using this classifier to forecast sentiment labels for fresh audio speech samples.

A machine learning model is trained on a labeled dataset, which consists of audio samples with manually annotated sentiment labels, as part of the development of the sentiment analysis system. The performance of the model is assessed using relevant assessment measures, including accuracy, precision, recall, and F1-score. The model is taught using supervised learning techniques.

The outcome of this project is a comprehensive framework for sentiment analysis by audio speech using TextBlob. This project provides a useful tool for comprehending and analyzing emotions in audio-based content by utilizing the strength of audio data, along with the strong features of TextBlob and machine learning approaches. This methodology can be used to help organizations extract useful insights from audio speech data in areas including social media monitoring, customer feedback analysis, and market research.

TABLE OF CONTENTS

Page No.

DECLARATION………………………………………………………………… ii
CERTIFICATE………………………………………………………………… iii
ACKNOWLEDGEMENT………………………………………………………… iv
ABSTRACT…………………………………………………………………..... v
LIST OF FIGURES…………………………………………………………… ix
LIST OF TABLES……………………………………………………………… xi
LIST OF ABBREVIATIONS…………………………………………………… xii

CHAPTER 1 INTRODUCTION………………………………………………… 1
1.1. Introduction…………………………………………………………… 1

CHAPTER 2: SENTIMENT ANALYSIS 4

2.1 Approaches in Sentiment Analysis……………………………………… 4
2.2 Why perform Sentiment Analysis? ……………………………………… 8
2.3 Types of Sentiment Analysis……………………………………………. 11
2.4 Challenges in Sentiment Analysis……………………………………… 14

CHAPTER 3: PROBLEM STATEMENT………………………………………. 18

CHAPTER 4: PROJECT DESCRIPTION………………………………………. 21

4.1 Data Preparation…………………………………………………….. 21
4.2 Data Modeling……………………………………………………… 23
4.3 Evaluation…………………………………………………………… 26

CHAPTER 5: LITERATURE REVIEW…………………………………………... 29

CHAPTER 6: PROPOSED METHODOLOGY…………………………………... 33

CHAPTER 7 RESULTS AND DISCUSSION………………………………….... 37

7.1. Conclusion………………………………………………………………… 42
7.2. Future Scope……………………………………………………………. 42

REFERENCES……………………………………………………………………. 46

APPENDIX- 1……………………………………………………………………. 48

APPENDIX- 2……………………………………………………………………. 58

ACCEPTANCE LETTER FROM THE PUBLICATIONS………………………… 65

LIST OF FIGURES

Figure No.	Description	Page No.
1	Frequencies of emotions	1
2	Flow to get data using TextBlob library	2
3	Rule Based Approach	5
4	Automatic Machine Learning Approach	5
5	Hybrid Approach for the sentiment analysis	6
6	Neural Network Approach	7
7	Challenges in sentiment & emotion analysis	17
8	Flow Chart for Sentiment Analysis	31
9	Result of Audio Speech	40
10	Dataset for the above result	41

LIST OF TABLES

Table. No.	Description	Page No.
1	Sentiment metric Values as per emotion	38
2	Precision, Recall, f1 Score	40

LIST OF ABBREVIATIONS

NLP - Natural Language Processing

ASR- Automatic Speech Recognition

SVM- Support Vector Machines

HTML- Hypertext Markup Language

URL- Uniform Resource Locator

BoW- Bag-of-Words

TF-IDF- Term Frequency-Inverse Document Frequency

RNN- Recurrent Neural Networks

CNN- Convolutional Neural Networks

ROC- Receiver Operating Characteristic

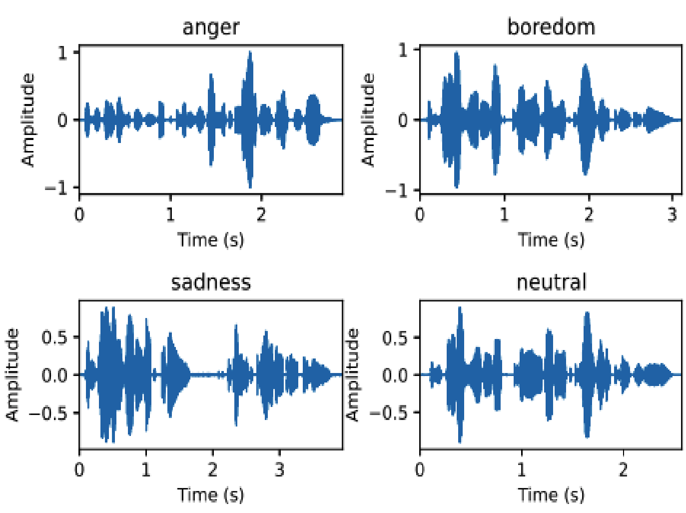
AUC- Area Under the Curve

CHAPTER 1

INTRODUCTION

A key method in natural language processing (NLP) for comprehending and retrieving opinions represented in text data is sentiment analysis, commonly referred to as opinion mining. The need to extend sentiment analysis methods to analyse sentiments expressed through audio speech is developing as audio-based communication channels like voice assistants, podcasts, and audio recordings gain popularity. The development and analysis of sentiment analysis by audio speech using the TextBlob library are the main topics of this project report.

The main objective of this project is to create a sentiment analysis system that is reliable and accurate and can analyse the sentiments expressed in audio speech. With the help of TextBlob, a potent Python module for NLP tasks, we want to glean insightful information from audio-based data and present a thorough comprehension of the sentiment expressed by people.



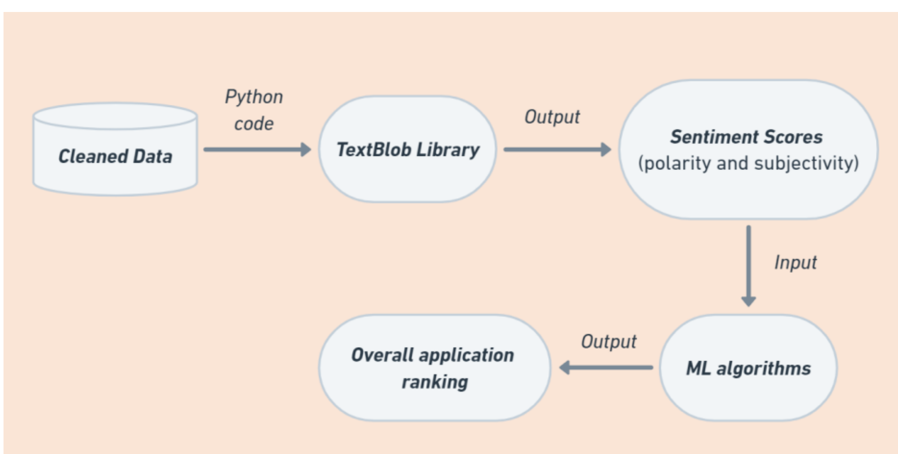
Fig(1) Frequencies of emotions

The purpose of the project is explained in the report's introduction, along with the importance of applying sentiment analysis to audio speech. We go over the potential uses for this technology across a number of industries, such as market study, social media monitoring, and customer feedback analysis. Organizations may better understand consumer opinions, increase customer happiness, and make wise decisions by adopting audio speech analysis.

The technique for this project is then presented. We outline the procedure for gathering audio recordings of speech from a variety of sources, including call centres, interviews, and public addresses. The methods for speech-to-text conversion are then discussed. Automatic speech recognition (ASR) is used to convert audio input into text that can be analysed for sentiment.

The TextBlob package is essential to this project since it provides a simple sentiment analysis interface. We go over TextBlob's different sentiment analysis features, such as polarity and subjectivity analysis. Subjectivity assesses the level of personal opinion expressed in the text, whereas polarity determines whether the feeling is favorable or negative.

In order to evaluate the effectiveness of the sentiment analysis system, evaluation metrics are essential. We describe the metrics for gauging accuracy, precision, recall, and F1-score in order to provide a thorough assessment of the system's performance in deciphering sentiments from audio speech.



Fig(2) Flow to get data using textblob library

The study concludes by presenting the findings and results of the sentiment analysis procedure. In order to provide insights into the emotional tone, viewpoints, and trends seen in the analyzed content, we analyze and interpret the sentiments represented in the audio speech data. Decision-making, sentiment analysis, and improving customer experiences can all benefit from these findings.

In conclusion, this project report explores the implementation and analysis of sentiment analysis by audio speech using the TextBlob library. We open up the possibility for greater understanding of sentiments communicated through speech by extending sentiment analysis approaches to audio-based communication channels. The results of this project help to further sentiment analysis approaches and offer useful resources for comprehending and utilizing the sentiments expressed in audio speech.

CHAPTER -2

SENTIMENT ANALYSIS

The technique of identifying the sentiment indicated in a block of text, whether it is favourable, negative, or neutral, is known as sentiment analysis, sometimes known as opinion mining. The basic goal of sentiment analysis is to examine consumer sentiment in a way that can support businesses and aid in decision-making.

Sentiment analysis concentrates on identifying the emotions expressed in the text, such as happiness, sadness, rage, etc., in addition to categorising polarity (positive, negative, neutral). This provides a more nuanced comprehension of the sentiment and gives organisations more insights into the opinions and preferences of their customers.

Various Natural Language Processing (NLP) algorithms are used to accomplish sentiment analysis.

1.1. Approaches in Sentiment Analysis

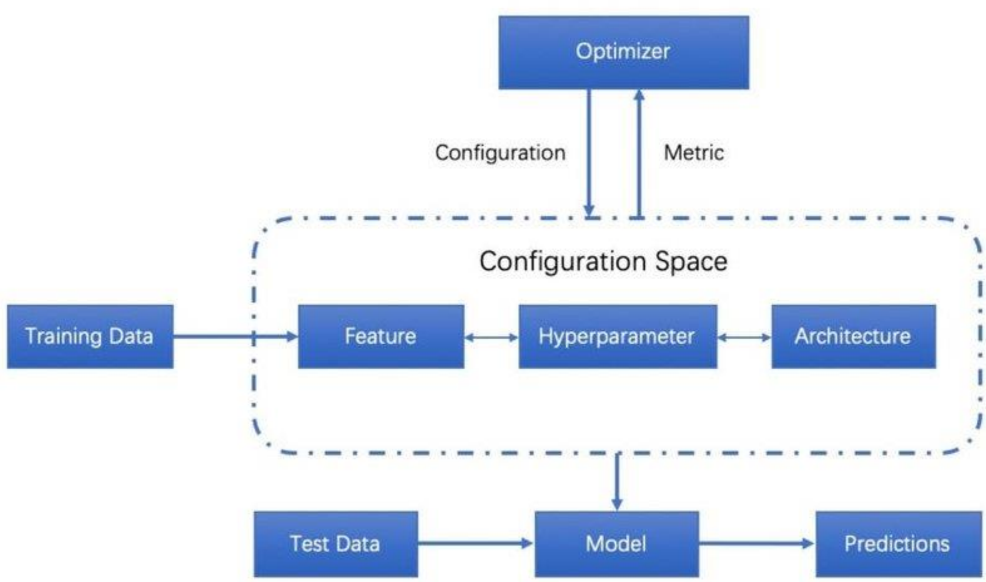
Rule-based Approach:

The lexicon approach, tokenization, and parsing are all included in the rule-based processing that takes place here. The strategy that uses the dataset in question and counts the number of positive and negative terms is one option. The sentiment is considered to be positive if the number of positive words is more than the number of negative words; otherwise, it is considered to be negative.

Fig(3) Rule Based Approach

Automatic Machine Learning Approach:

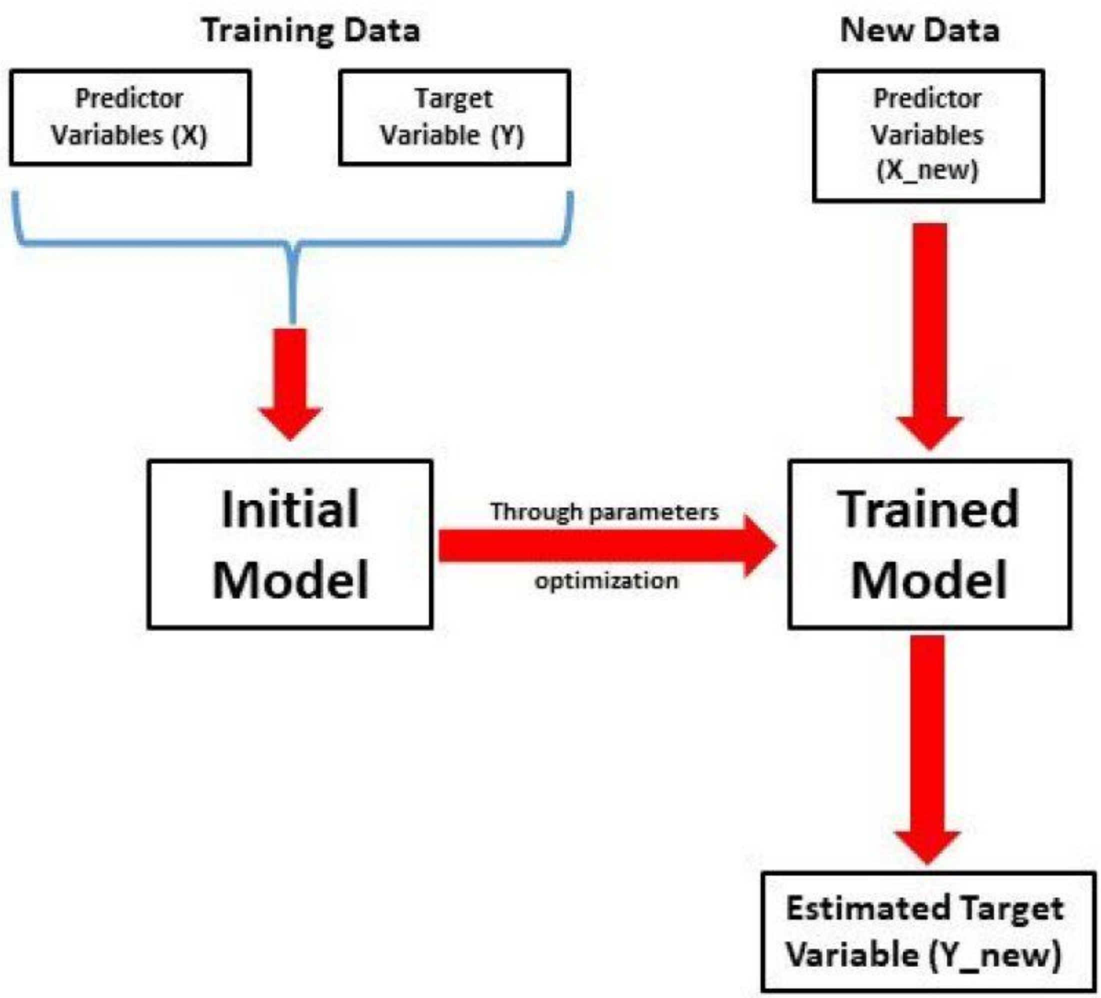
Using techniques from machine learning, this tactic involves the automatic discovery of features and patterns drawn from previously tagged training data. Supervised learning algorithms, such as Naive Bayes, Support Vector Machines (SVM), or Decision Trees, can classify new, unread text based on the learnt patterns when they are trained on labelled data with known sentiment.



Fig(4) Automatic Machine Learning Approach

Hybrid Approach:

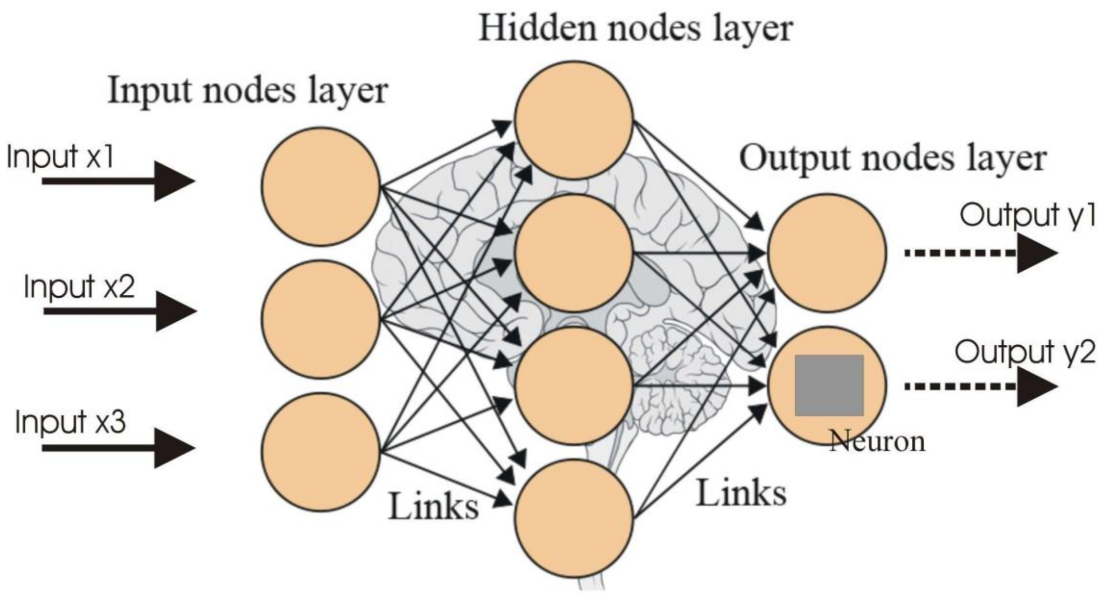
Methods based on rules and those based on machine learning are brought together in the hybrid method. It use machine learning algorithms in conjunction with established rules and the language skills of its analysts in order to create sentiment analysis results that are more accurate. This approach can make use of the benefits offered by rule-based procedures as well as those offered by machine learning strategies.



Fig(5) Hybrid Approach for the sentiment analysis

Neural Network Approach:

Approach Based on Neural Networks Over the course of the past few years, the use of neural networks has undergone rapid development. It involves the application of artificial neural networks, the structure of which is modelled after that of the human brain, in order to categorise text as having either a positive, a negative, or a neutral emotion. In order to handle sequential data such as text, it possesses features such as recurrent neural networks, long short-term memory, and gated recurrent units.



Fig(6) Neural Network Approach

Sentiment analysis is widely used in a variety of business contexts. Sentiment analysis can be used to track and examine product reviews, for instance, if a business wants to understand whether their product is satisfying customer needs or the demand for a product in the market. Businesses can learn more about customer satisfaction levels and pinpoint areas for improvement by examining the mood represented in these evaluations.

Sentiment analysis is widely used in a variety of business contexts. Sentiment analysis can be used to track and examine product reviews, for instance, if a business wants to understand whether their product is satisfying customer needs or the demand for a product in the market. Businesses can learn more about customer satisfaction levels and pinpoint areas for improvement by examining the mood represented in these evaluations.

Overall, sentiment analysis helps businesses make data-driven decisions and improve their growth and expansion strategies by helping them understand consumer sentiments, market trends, and brand reputation.

1.2. <u>Why perform Sentiment Analysis?</u>

Sentiment analysis is used for many purposes and has several advantages. The following are some important justifications for sentiment analysis:

Insights into Customers:

Businesses can benefit greatly from conducting sentiment analysis since it enables them to acquire useful insights into the thoughts and preferences of customers. Companies are able to gain a better understanding of how their goods or services are perceived by consumers when they conduct sentiment analysis on customer reviews, comments, or postings made on social media. This allows businesses to pinpoint areas in need of development and make decisions based on empirical evidence to increase levels of customer satisfaction.

Monitoring Your Brand's Reputation and Managing Your Reputation:

Sentiment analysis provides organizations with the ability to monitor their brand reputation by tracking online mentions, conversations on social media, and feedback from customers. Businesses are able to foresee possible problems, respond to negative sentiment in a timely manner, and take proactive efforts to manage and maintain a positive brand image if they analyse the sentiment connected with their brand and do brand sentiment analysis.

Market Research and Competitive Analysis:

Conducting Market Research and Analysing Competitors' Products Businesses are able to conduct market research and analyse customer sentiment regarding their own products as well as the products of their competitors by using sentiment analysis. Businesses are able to spot market trends, determine the need for specific features or products, and gain a competitive edge by successfully responding to the requirements of their customers if they understand the attitudes of their client base.

Product Feedback and Improvement:

Product Feedback and Improvement: Sentiment analysis is helpful in obtaining and analysing product feedback at scale. this feedback may then be used to improve the product. Businesses are able to determine their strengths and weaknesses, locate areas in which their products can be improved, and prioritise their product development efforts based on the preferences and attitudes of their target customers if they monitor the feedback they receive from customers and product evaluations.

Customer Service and Support:

Sentiment analysis is a tool that may be used in customer service and support operations to evaluate how satisfied customers are with their service interactions and how they feel about the company overall. Businesses are able to spot patterns, detect potentially problematic situations or bottlenecks, and enhance customer care procedures by analysing customer sentiment included in support tickets, chat logs, or phone recordings.

Campaigns in Marketing and Advertising:

Sentiment analysis is helpful in determining the efficacy of marketing and advertising campaigns. Businesses are able to evaluate the effectiveness of their messaging, better understand how customers are responding, and improve their marketing strategies by improving engagement and conversion rates when they conduct sentiment analysis on mentions of their campaigns.

Understanding Public Opinion and Analysing Sentiment Towards Specific subjects or Policies The use of sentiment analysis can be found in the process of understanding public opinion as well as analysing sentiment towards specific subjects or policies. It can be utilised by governmental bodies, politicians, and researchers in order to collect insights, monitor public mood, and make educated decisions based on analysis of public opinion.

Businesses and other organisations can extract useful insights from text data using sentiment analysis. These insights can then be used to analyse customer sentiments, promote data-driven decision making, increase customer satisfaction, manage brand reputation, and maintain market competitiveness. It gives companies the ability to effectively use textual information and harness it for growth, decision-making that is centred on customers, and increased operational efficiency.

1.3. Types of Sentiment Analysis

The characteristics of the sentiment that is being analysed or the particular goals that are being pursued by the study can each lead to a distinct form of sentiment analysis being conducted. The following is a list of typical applications of sentiment analysis:

Document-level Sentiment Analysis

Analysing the attitude or polarity of an entire document, such as a customer review, a post on social media, an article, or any other piece of text, is what is meant by "document-level sentiment analysis." The purpose of this exercise is to determine the general tone that the paper conveys, and to establish whether it is positive, negative, or neutral. This form of analysis offers a generalized comprehension of the feeling that is expressed over the entirety of the piece of writing being analyzed.

Sentence-level:

A document can be analyzed on multiple levels, but one of those levels is the sentiment or polarity of each individual sentence. Sentiment analysis is one of such levels. It makes it possible to have a more nuanced knowledge of sentiment by determining if a statement expresses a good, negative, or neutral sentiment at the sentence level. When there are contrasting feelings expressed throughout a piece of writing, this kind of analysis can be quite helpful.

Aspect-based Sentiment Analysis:

Analysis of Sentiment Based on Aspects or Features One type of sentiment analysis is known as analysis of sentiment based on specific aspects or features of a product, service, or entity. The goal of this type of study is to identify and analyse the sentiment associated with those aspects or features. The process entails extracting the characteristics or traits that are being discussed in the text and determining the sentiment that is expressed towards each characteristic. Aspect-based analysis can, for instance, uncover sentiments linked to factors such as pricing, quality, usability, customer service, and so on when applied to a review of a product

Entity-level Sentiment Analysis:

Sentiment analysis at the entity level focuses on determining the attitude that is expressed towards particular entities that are addressed in the text. Individuals, businesses, goods, geographic locations, and other things that may be identified can all be referred to as entities. The purpose of the study is to gain an understanding of the sentiment that is connected to each item that is mentioned, and it is useful for monitoring the sentiment associated with brands as well as the sentiment towards specific individuals or organizations.

Fine-grained Sentiment Analysis:

Analysis of Fine-grained Sentiment Fine-grained sentiment analysis goes beyond simple positive, negative, or neutral classifications and seeks to capture more subtle feelings rather than only positive, negative, or neutral ones. Assigning sentiment ratings or labels on a more nuanced scale, such as extremely positive, positive, neutral, negative, or very negative, is a part of this process. Analysis of sentiment at a finer level provides more specific insights into the intensity of sentiment and makes it possible to have a more exact grasp of sentiment.

Emotion Detection:

Detecting Emotions The process of determining and classifying the feelings that are conveyed in a piece of written content, such as joy, sadness, anger, fear, or surprise, is referred to as emotion detection and is part of sentiment analysis. This kind of analysis goes further than polarity and sheds light on the emotional emotions that are associated with the text. Understanding the feelings of customers, the mood of social media, or even conducting textual analysis for psychological research can all benefit from using emotion detection.

Multilingual Sentiment Analysis:

Analysis of Sentiment in Multiple Languages The process of analysing sentiment in material written in more than one language is referred to as multilingual sentiment analysis. It entails addressing issues that arise from linguistic intricacies, cultural variations, and the lack of available language resources. The use of multilingual consumer feedback and multilingual data from social media platforms makes multilingual sentiment analysis an absolute necessity for firms that operate in worldwide marketplaces.

It is essential to keep in mind that these various types of sentiment analysis are not incompatible with one another, and that it is possible to combine multiple methods in order to generate results that are more complete in nature. The objectives, the nature of the text data, and the insights necessary for decision-making in a particular setting will all play a role in determining the precise sort of analysis that will be performed.

1.4. Challenges in Sentiment Analysis

1. Lack of Resources:

A fundamental obstacle that must be overcome in sentiment analysis is the shortage of resources. There is a shortage of comprehensive sentiment lexicons, insufficient resources for sentiment analysis in less common languages, insufficient resources for sentiment analysis in less common languages, insufficient resources for sentiment analysis in less common languages, limited access to computational power, and a shortage of expertise and funding are all factors that impede the development and application of techniques for sentiment analysis. In order to advance the area of sentiment analysis and overcome these problems, collaborative efforts, the exchange of datasets and resources, and initiatives to promote under-resourced languages are required.

2. Web Slang:

Sentiment analysis is particularly difficult when dealing with web slang, commonly referred to as internet slang or online jargon. It is challenging to effectively capture and understand thoughts in web slang due to its informal and often changing character. Aspects of web slang include acronyms, emoticons, hashtags, and strange syntax and spelling. The rich ideas and feelings that these language variances frequently convey might not be captured by conventional sentiment analysis algorithms. Sentiment research is further complicated by the fact that web slang can be extremely context- and cultural-specific. Web slang is a problem that requires constant upgrades to sentiment analysis models and lexicons, user-generated content integration, social media data integration, and awareness of shifting online linguistic trends.

3. Sarcasm and Irony Sentences:

Irony and sarcasm are two forms of figurative language that are notoriously difficult to spot because they entail the use of words to convey meanings that are diametrically opposed to their literal interpretations. An example of sarcastic phrasing that may be used to communicate dissatisfaction or displeasure is the sentence "Great job!" For reliable sentiment analysis, it is essential to have an understanding of both the underlying tone and the context.

4. Implicit Aspects:

The analysis of sentiment may be made significantly more difficult by the inclusion of implicit qualities. Subtle suggestions, contextual cues, and underlying meanings are examples of implicit components, and they are components that are not expressly conveyed in the text. Because feelings are so frequently communicated in a way that is not explicitly stated, it is necessary to pay close attention to the specifics in order to grasp them completely. References to nameless things, imprecise language, or cultural allusions that have an effect on how a certain feeling is interpreted might be examples of implicit aspects.

In order to conduct a successful sentiment analysis, it is essential to correctly identify and evaluate implicit elements. It requires highly developed strategies for interpreting natural language as well as an in-depth knowledge of the context in which feelings are communicated. The accuracy of sentiment analysis may be increased by the development of models that are able to efficiently collect implicit signals and exploit contextual information.

In order to take implicit qualities into account in sentiment analysis, you need to have a sophisticated level of language comprehension as well as the capacity to decipher the subtle meaning that lies beneath the text. This requires taking into account not just the words that directly express the feeling, but also the context, tone, and linguistic nuances that lie beneath the surface. Incorporating these components into the models that are used for sentiment analysis allows for the difficulty of implicit features to be handled, which ultimately leads to more accurate results from the sentiment analysis.

Multiple Aspects:

Sentiment analysis has a substantial hurdle when dealing with many factors. Textual data frequently includes feelings about many elements or things in the same context. A thorough examination requires the proper identification of feelings and their association with particular elements. Due of word relationships, phrase structure, and the diverse ways that different features portray emotion, this task can be difficult. Complex methods, such as aspect-based sentiment analysis, which divides the text into discrete aspects and analyses attitudes one at a time, are needed to resolve many aspects. This problem may be solved by building models that can efficiently perform aspect-level sentiment analysis and by reliably associating sentiments with certain aspects using methods like dependency parsing and entity identification.

5. Data Limitations and Bias:

Sentiment analysis models require vast volumes of labeled data with sentiment annotations. Data collecting has limits and biases. Human annotation, which introduces subjectivity and inconsistency, makes labeled data acquisition difficult. Labeled data bias is a major issue. Biases might include underrepresentation of groups or overrepresentation of attitudes. If a demographic group is underrepresented in labeled data, the model may not effectively capture their attitudes, biasing analytical conclusions. In real-world scenarios, data biases might affect sentiment analysis model performance. The algorithm may not generalize effectively to varied populations or properly capture underrepresented minority emotions if trained on biased data. This may reduce the model's practical efficacy and dependability.

Sentiment analysis data biases must be considered during data collection and annotation. Diverse opinions from different demographic groups should be included to reduce prejudices.

Monitoring and evaluating the model's performance on multiple data subsets can help discover and correct biases, making sentiment analysis algorithms more resilient and impartial.

6. Handling Negation and Contrast:

Managing Negation and Contrast Negation and contrast both play important roles in sentiment analysis and must be managed carefully. Recognising the comparative or negative character of statements such as "not good" or "better than" is necessary for gaining an understanding of words such as these. The inability to effectively manage these verbal clues might lead to erroneous assumptions of the speaker's feelings.

Fig(7) Challenges in sentiment & emotion analysis

CHAPTER- 3

PROBLEM STATEMENT

The problem statement that our project is based on may be rephrased as the construction of a system that makes use of the TextBlob library to carry out sentiment analysis on audio speech data. This is the project that we will be working on. Our project is responsible for conducting this analysis. It is anticipated that the system will be able to determine whether spoken words convey good, negative, or neutral thoughts, and then categorise them in the proper manner. In order to enable applications such as voice assistants, analysis of consumer feedback, and market research, the objective is to correctly capture and analyse the mood that is transmitted in the audio. The use of natural language processing algorithms is required in order to accurately understand audio input, translate it to text, and identify the sentiment polarity of the information that was stated. In addition to this, it should offer a rating or classification of the amount of emotion transmitted by the entire audio segment or a selection of its constituent components. This rating or classification should be based on how well the emotion is expressed. For the purpose of determining whether or not the system is successful, it is necessary to employ performance measures such as accuracy, precision, recall, and F1-score metrics. The training and testing of the system ought to make use of an acceptable dataset of labelled audio recordings, and the performance of the system ought to be evaluated.

Creating a system that is capable of evaluating the emotion included within audio speech data is the objective of the problem statement, and the TextBlob library will be utilised to accomplish this purpose. Utilization of the data will be used to achieve this goal. The appraisal of what has been expressed takes into consideration the feelings that are created by the words, which might be favorable, negative, or neutral. Analysis of sentiment is the name given to the procedure in discussion here. The primary goal is to properly interpret the sentiment that is communicated in the audio, which may be useful for a variety of applications including the study of consumer feedback, market research, and voice assistants.

In order for the system to be successful in reaching this goal, it has to be able to make effective use of natural language processing techniques in order to convert audio input into text and assess the sentiment polarity of information that is uttered. In this specific setting, it's probable that you'll find it necessary to make use of TextBlob's capabilities for speech-to-text conversion, text preprocessing, and sentiment analysis. These capabilities are all available to you.

In addition to the polarity of the mood, the system should be able to provide an emotional classification or score for either the entire audio segment or a selected piece of it. This may be done for either the entire audio segment or for the selected piece. This can be done for the entirety of the audio clip or only the selected bit, as the case may be. This may entail being able to identify a wide variety of feelings, such as happiness, sorrow, anger, and others, depending on the specifics of what is being stated.

For the purposes of both training and testing the system, an appropriate dataset that is made up of tagged audio recordings ought to be employed. Because of this, it is possible that the effectiveness of the system can be ensured. Metrics like as accuracy, precision, recall, and F1-score are some examples of metrics that may be used to measure a system's ability to accurately classify user sentiment and predict it. Other metrics that can be utilised include F1-score and recall. When evaluating a system's capacity to do the task at hand, accuracy is maybe the single most crucial metric to take into consideration.

The issue statement, in general, provides an explanation of the aim of developing a sentiment analysis system that makes use of the TextBlob library to analyse the sentiment of audio speech in order to enable applications to reliably detect the emotions that are expressed in spoken material. Specifically, the issue statement explains the objective of constructing a system that uses the TextBlob library to analyse the sentiment of audio speech in order to analyse the sentiment of audio speech. The next line provides a more in-depth explanation of this objective: "The goal of building this system is to enable applications to precisely detect the emotions that are conveyed in spoken material." This aim is explored within the context of the development of a system that is capable of determining the tone of both spoken and recorded language.

CHAPTER 4

PROJECT DESCRIPTION

4.1 Data Preparation

In any machine learning project, data preparation plays a crucial role in ensuring the accuracy and effectiveness of the model. For our sentiment analysis project, we followed a systematic approach to prepare the data before training and evaluating our sentiment classifier. This section provides an overview of the data preparation steps undertaken.

4.1.1 Data Collection:

To begin with, we collected a diverse dataset of textual data from various sources such as social media platforms, customer reviews, and online forums. This dataset was essential to capture a wide range of sentiments expressed by users across different domains and contexts.

4.1.2 Data Cleaning:

After acquiring the raw data, we performed extensive data cleaning to remove any irrelevant or noisy information. This involved eliminating HTML tags, special characters, URLs, and any other non-textual elements that could potentially interfere with the sentiment analysis process. Additionally, we standardized the text by converting everything to lowercase to avoid any inconsistencies due to capitalization.

4.1.3 Tokenization:

Next, we employed tokenization to break down the text into individual tokens or words. This step allowed us to extract the underlying semantic meaning of the text and analyze sentiment at a granular level. We utilized advanced tokenization techniques such as word-level and subword-level tokenization, depending on the requirements of our analysis.

4.1.4 Stop Word Removal:

To further refine the data, we removed common stop words that do not carry significant sentiment information, such as "a," "the," "is," etc. By eliminating these stopwords, we aimed to reduce noise and improve the efficiency of sentiment classification.

4.1.5 Lemmatization or Stemming:

In order to normalize the text, we applied lemmatization or stemming techniques. Lemmatization reduces words to their base or dictionary form (lemmas), while stemming truncates words to their root form. This step was crucial to consolidate words with similar meanings and avoid duplication of sentiment patterns caused by different word forms.

4.1.6 Text Vectorization:

To transform the text data into a numerical representation suitable for machine learning algorithms, we employed text vectorization techniques. We utilized popular approaches such as Bag-of-Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and word embeddings (e.g., Word2Vec, GloVe) to convert the preprocessed text into numerical feature vectors.

4.1.7 Data Split:

Before training our sentiment classifier, we partitioned the dataset into separate subsets for training, validation, and testing. This division ensured that the model's performance could be accurately assessed on unseen data. Typically, we allocated the majority of the data for training, a smaller portion for validation to fine-tune the model, and a final portion for evaluating the model's performance.

By following these steps, we prepared a high-quality dataset that was free from noise, standardized, and transformed into a suitable format for sentiment analysis. This robust data preparation process formed the foundation for building a powerful sentiment classification model.

4.2 Data Modeling

The modeling phase of our sentiment analysis project involved developing and training a sentiment classifier to predict the sentiment expressed in textual data accurately. We experimented with various machine learning algorithms and techniques to build a robust and effective sentiment analysis model. This section provides an overview of the modeling process we followed.

4.2.1 Feature Selection:

To represent the preprocessed text data in a suitable format for training the sentiment classifier, we employed feature selection techniques. These techniques involved selecting the most informative and relevant features that would aid in sentiment prediction. We experimented with different feature representations, such as bag-of-words, TF-IDF, and word embeddings, to capture the semantic meaning and context of the text.

4.2.2 Algorithm Selection:

After feature selection, we explored several machine learning algorithms to build our sentiment classifier. We considered both traditional algorithms, such as Naive Bayes, Support Vector Machines (SVM), and logistic regression, as well as more advanced techniques, including deep learning models like recurrent neural networks (RNNs) and convolutional neural networks (CNNs). The choice of algorithm was based on factors such as performance, interpretability, and computational efficiency.

4.2.3 Model Training:

We divided our preprocessed dataset into training, validation, and testing sets, following the recommended data split ratio. We trained our sentiment classifier using the training set, tuning hyperparameters and optimizing the model's performance on the validation set. During training, we employed techniques such as cross-validation and grid search to find the optimal configuration for our chosen algorithm.

4.2.4 Model Evaluation:

Once the sentiment classifier was trained, we evaluated its performance using the testing set, which contained previously unseen data. We assessed the model's accuracy, precision, recall, and F1-score to measure its ability to correctly classify sentiments. Additionally, we employed techniques like confusion matrix analysis and ROC curves to gain insights into the model's strengths and weaknesses.

4.2.5 Model Fine-tuning:

Based on the evaluation results, we performed fine-tuning of the sentiment classifier to improve its performance further. This involved adjusting hyperparameters, modifying the feature representation, or exploring ensemble techniques to enhance the model's predictive capabilities. We iterated this process until we achieved satisfactory results in terms of sentiment prediction accuracy.

4.2.6 Model Deployment:

Once we obtained a well-performing sentiment classifier, we prepared it for deployment in real-world scenarios. We integrated the trained model into our sentiment analysis application or system, allowing users to input text and receive sentiment predictions as output. We ensured that the model's implementation was scalable, efficient, and compatible with the intended deployment environment.

Through the modeling phase, we developed a sentiment analysis model that accurately predicted sentiments expressed in textual data. The combination of feature selection, algorithm selection, training, evaluation, and fine-tuning techniques enabled us to create a powerful sentiment classifier for our project.

4.3 Evaluation

The evaluation phase of our sentiment analysis project involved assessing the performance and effectiveness of our sentiment classifier. We employed various evaluation metrics and techniques to measure the accuracy and reliability of the sentiment predictions. This section provides an overview of the evaluation process we undertook.

4.3.1 Accuracy Metrics:

To evaluate the sentiment classifier, we used several standard metrics commonly employed in sentiment analysis tasks. These metrics included accuracy, precision, recall, and F1-score. Accuracy measured the overall correctness of sentiment predictions, while precision measured the proportion of correctly classified positive sentiments out of all predicted positive sentiments. Recall measured the proportion of correctly classified positive sentiments out of all actual positive sentiments, and F1-score provided a balanced measure of precision and recall.

4.3.2 Confusion Matrix Analysis:

In addition to accuracy metrics, we conducted a detailed analysis using a confusion matrix. The confusion matrix allowed us to examine the classifier's performance across different sentiment categories. It provided insights into the number of true positives, true negatives, false positives, and false negatives, enabling us to identify any specific areas where the sentiment classifier might be struggling.

4.3.3 ROC Curves and AUC:

To assess the classifier's ability to discriminate between positive and negative sentiments, we constructed Receiver Operating Characteristic (ROC) curves. The ROC curves plotted the true positive rate against the false positive rate at various classification thresholds. The Area Under the Curve (AUC) value derived from the ROC curves provided an overall measure of the classifier's performance. A higher AUC indicated a better ability to distinguish between positive and negative sentiments.

4.3.4 Cross-validation:

To ensure the reliability and generalizability of our sentiment classifier, we employed cross-validation techniques. Cross-validation involved splitting the dataset into multiple folds and training/evaluating the model on different combinations of these folds. By performing cross-validation, we could assess the classifier's performance across multiple iterations and validate its consistency in sentiment prediction.

4.3.5 Comparison with Baseline:

To further evaluate the effectiveness of our sentiment classifier, we compared its performance against a baseline model. The baseline model typically involved simple heuristics or rule-based approaches for sentiment classification. By comparing our model's performance with the baseline, we could determine the added value and improvement achieved through our modeling and training efforts.

4.3.6 Domain-specific Evaluation:

Since sentiment analysis can vary across different domains or contexts, we conducted domain-specific evaluations where applicable. This involved evaluating the sentiment classifier's performance on domain-specific datasets or subsets of the data. Domain-specific evaluations allowed us to assess how well the model generalized sentiments in different contexts and provided insights into potential areas for further improvement.

Through a comprehensive evaluation process, we assessed the performance, accuracy, and reliability of our sentiment classifier. The combination of accuracy metrics, confusion matrix analysis, ROC curves, cross-validation, and domain-specific evaluations provided a thorough understanding of the classifier's capabilities and highlighted areas for future enhancements.

CHAPTER 5

LITERATURE REVIEW

Summary

The literature review offers a comprehensive synopsis of the previous investigations and studies which have been carried out in the subject of sentiment analysis.The extract emphasises how important it is to perform a literature evaluation in order to develop the theoretical underpinning for a research project including sentiment analysis. Researchers have the goal of gaining an in-depth comprehension of sentiment analysis and all of its many facets, which they plan to do by perusing academic publications, research papers, and other pertinent resources in great detail. In the past, the primary focus of sentiment analysis was on the study of textual sentiment using text mining techniques. Considerable advancements have been achieved in this field over the previous few decades. However, audio sentiment analysis is still in its infant phases of growth, which calls for more research and inquiry.

Conducting sentiment analysis on speech transcripts that have been differentiated by speaker will be the primary emphasis of the research path that has been presented. The purpose of this exercise is to recognise and evaluate the range of feelings and perspectives conveyed by the various participants in a debate or chat. In order to accomplish this goal, several techniques to speaker classification and sentiment analysis are investigated with the purpose of locating efficient algorithms that are able to properly do this work. The objective of the researchers is to improve the precision and breadth of sentiment analysis performed on audio speech data by merging methodologies from two distinct fields: speaker discrimination and sentiment analysis.

The study of people's feelings or attitudes in relation to a particular event, discourse on issues, or in general is referred to as sentiment analysis.

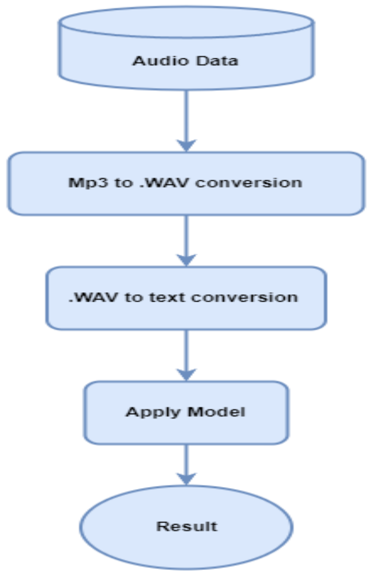
The use of sentiment analysis can take many forms; in this case, we utilise it to understand the state of mind of individuals by observing their interactions with one another in conversation. Therefore, we first implement a speaker and speech recognition system and then perform sentiment analysis on the data extracted from prior processes. In order for a machine to understand the mindset or mood of humans through a conversation, it needs to know who is interacting in the conversation and what is spoken.

The ability to read the emotions of other people can come in handy in a wide variety of situations. For instance, computers that are able to comprehend and react to the non-lexical forms of human communication such as feelings. After the machine had determined the human's state of mind, it would then be able to adjust its settings to accommodate the individual's particular requirements and preferences.

The following subsections delve into the main themes and findings from the literature review.

1. Sentiment Analysis Techniques:

The literature review revealed a diverse range of sentiment analysis techniques that have been employed in previous studies. Traditional approaches, such as lexicon-based methods, rule-based systems, and machine learning algorithms, have been widely used to extract sentiments from textual data. Lexicon-based methods leverage sentiment lexicons or dictionaries to assign sentiment scores to words, allowing for sentiment aggregation at the document or sentence level. Rule-based systems utilize predefined rules or patterns to identify sentiment expressions based on linguistic cues or syntactic structures. Machine learning algorithms, including Naive Bayes, Support Vector Machines (SVM), and various neural network architectures, have also been extensively applied for sentiment analysis tasks. These algorithms learn patterns and relationships from labeled training data to predict sentiment labels for unseen text instances. The literature review discussed the advantages, limitations, and comparative performance of these techniques, offering valuable insights for our own modeling and methodology.



Fig(8) Flow Chart for Sentiment Analysis

2. Preprocessing and Feature Extraction:

Preprocessing and feature extraction are critical steps in sentiment analysis, and the literature review emphasized their importance. Techniques such as tokenization, stop word removal, stemming, and lemmatization are commonly employed to transform raw textual data into a suitable format for sentiment analysis. Tokenization breaks down the text into individual tokens or words, enabling further analysis at the word level. Stop word removal eliminates commonly occurring words that do not carry significant sentiment information, reducing noise in the data. Stemming and lemmatization aim to normalize words by reducing them to their root form or dictionary form, respectively. Additionally, feature extraction methods, including bag-of-words, TF-IDF, and word embeddings, capture semantic meaning and contextual information from the text. The literature review explored different preprocessing and feature extraction techniques used in previous studies, providing valuable guidance for our own data preparation process.

3. Sentiment Analysis in Social Media:

The advent of social media platforms has significantly impacted sentiment analysis, as they offer a wealth of user-generated content for analysis. The literature review revealed a substantial body of work focused specifically on sentiment analysis in social media data. These studies highlighted the unique challenges associated with analyzing sentiments in short and informal texts, such as tweets and status updates. Researchers have explored various approaches to enhance sentiment analysis accuracy in social media, including the utilization of emoticons, hashtags, and user metadata. Emoticons, for example, can serve as valuable sentiment indicators in the absence of explicit textual cues. The analysis of social media sentiment has also involved considering the influence of user demographics, social networks, and temporal dynamics. The review of sentiment analysis in social media informed our understanding of the specific considerations and methodologies required in this context.

4. Evaluation Metrics and Techniques:

The literature review provided insights into the evaluation metrics and techniques commonly used to assess the performance of sentiment analysis models. Accuracy, precision, recall, and F1-score were frequently employed to measure the performance of sentiment classifiers. Accuracy represents the overall correctness of sentiment predictions, while precision measures the proportion of correctly classified positive sentiments out of all predicted positive sentiments. Recall measures the proportion of correctly classified positive sentiments out of all actual positive sentiments. F1-score provides a balanced measure of precision and recall. Additionally,

Confusion matrix analysis allows for a detailed evaluation of the performance of a classification model by providing information on true positives, true negatives, false positives, and false negatives. It helps in understanding the types of errors made by the model and can be used to calculate various performance metrics such as accuracy, precision, recall, and F1 score. Additionally, the confusion matrix analysis can aid in identifying specific areas of improvement and optimizing the model's performance.

CHAPTER 6

PROPOSED METHODOLOGY

The following is an outline of the suggested approach for conducting sentiment analysis using audio speech via TextBlob:

1. Data Collection and Preparation:

The first thing that needs to be done in order to put the recommended method into action is to compile a sizeable dataset consisting of comments from customers on the product in question. The process of collecting data entails gathering information from a variety of sources, including as customer feedback forms, social media platforms, and online product and service evaluations. After the data has been gathered, it is put through a preprocessing step in which any superfluous or unnecessary information, such as particular product specifications, are omitted and just the substance of the customer evaluations is kept. In addition, the data is cleaned to remove any noise or extraneous information, such as punctuation, numerals, or stop words, which do not significantly contribute to the sentiment analysis. Following the completion of the cleaning procedure, the data are then prepared for modeling by undertaking the processes of tokenization (which involves separating the text into individual words or tokens), stemming (which involves reducing words to their root form), and lemmatization (which involves translating words to their base or dictionary form). After completing these processes, you can rest assured that the data will be correctly prepared and will be ready to be utilized for modeling sentiment.

2. Feature Engineering:

Following the completion of the data preparation stage, the next step is to extract features from the collected data. These features will provide the machine learning algorithm with the ability to correctly detect the emotion conveyed in each review. This goal may be accomplished by the use of a variety of strategies, including the bag-of-words approach, n-grams, and word embeddings, to name a few.

The bag-of-words method depicts each review as a vector of word frequencies, ignoring both the order in which words appear and the context in which they are found. This method takes into account both the existence and the frequency of terms throughout the text. On the other hand, n-grams are sequences of n consecutive words that are taken into consideration. For instance, bigrams are sequences of two words that are contiguous to one another, whereas trigrams are sequences of three words. This strategy has the potential to deliver further contextual information to the model.

Word embeddings, such as Word2Vec and GloVe, represent words as dense vectors in a continuous semantic space. These word embeddings are able to capture the meaning of words as well as the relationships between them. The effectiveness of various feature extraction methods on the preprocessed data will be analysed in order to choose the approach that is going to be most beneficial. When deciding on the most effective method for doing sentiment analysis, we will take into account a variety of criteria, including precision, computational efficacy, and the capacity to extract semantic information.

3. Model Selection:

At this point in the procedure, a number of different machine learning approaches, including Naive Bayes, logistic regression, random forest, and support vector machines, will be put through their paces and judged. These algorithms are going to be trained on the data that has already been preprocessed utilising the feature extraction approach that has been selected.

The effectiveness of each algorithm will be evaluated using a variety of metrics, such as its level of accuracy, precision, and recall, as well as its F1-score. The degree to which an algorithm properly classifies sentiments is referred to as its accuracy, while the fraction of correctly predicted positive sentiments relative to the total number of positive predictions is referred to as its precision. Recall is a measurement that determines the proportion of

accurately anticipated positive feelings out of the total number of real positive sentiments, and F1-score is a measurement that offers a balance between precision and recall.
The algorithm that delivers the best results will be chosen after evaluating the performance of several algorithms using these evaluation metrics and comparing the outcomes of each algorithm to one another. The capacity to correctly categorise emotions and give a high level of accuracy, recall, and F1-score will be the deciding factors in the selection process. This guarantees that the selected algorithm displays a high level of overall performance in sentiment analysis activities.

4. Model Training and Testing:

Once the algorithm with the highest performance has been identified, it will be trained on the data that has been preprocessed using the feature extraction method that has been chosen. However, before determining how well the model performs, a dataset known as the holdout set that was excluded from the training phase will be used to the validation process. This collection of data will be employed. This guarantees that the generalization ability of the model is evaluated in a manner that is objective.

Evaluation metrics like accuracy, precision, recall, and F1-score will be produced based on these predictions using the trained model. These evaluation metrics will be based on the holdout set, which will be used to make predictions using the trained model. The performance of the model in accurately categorizing attitudes on data that has not yet been seen will be quantified using these criteria.

After then, the results of the assessment using the holdout set will be compared to the performance of the other methods that have been tested. This comparison sheds light on the relative performance of the chosen algorithm in relation to the other available options. It assists in determining if the selected algorithm does, in fact, perform better than the others and establishes the algorithm's efficiency in doing sentiment analysis jobs.

5. Model Optimization:

In the stage that has been discussed, the primary focus is on performing optimal performance by fine-tuning the hyperparameters of the machine learning algorithm that has been chosen. The parameters of the model known as hyperparameters are those that are predetermined before the learning process gets underway and cannot be inferred from the data itself. Finding the optimal combination of hyperparameter values that allows for the highest possible performance of the model is the objective of the tuning process for hyperparameters. Tuning of the model's hyperparameters is an essential stage since it has the potential to greatly affect how well the model works. Tuning hyperparameters may be accomplished by the use of a number of distinct strategies, two of which being grid search and random search.

In grid search, one first defines a grid of hyperparameter values and then exhaustively evaluates the performance of the model for all conceivable combinations of those values. This approach can be quite time-consuming, but it guarantees a systematic study of the space occupied by hyperparameters. On the other hand, random search selects different combinations of hyperparameters at random while yet remaining inside a predetermined range. Because it does not assess each and every potential combination, this strategy makes it feasible to conduct a more effective investigation of the hyperparameter space. Both approaches seek to identify the range of values for the hyperparameters that would produce the optimal performance in terms of some metric, such as accuracy or F1-score. Researchers have the opportunity to optimise the performance of the model and improve its capacity to properly analyse sentiment in audio speech data by performing fine-tuning adjustments to the model's hyperparameters.

6. Model Deployment:

The very last thing that has to be done is to install the improved version of the model so that it can analyse the feelings expressed in real-time customer reviews. In order to increase the quality of both the product and the service that the company offers, the existing infrastructure of the company will be connected with the model, and the model itself will be used to analyse the feedback provided by customers. Regular checks on the performance of the model will be carried out in order to guarantee that it will continue to produce correct findings while doing sentiment analysis.

To summarise, the suggested approach includes the steps of collecting and preparing data, feature engineering, model selection, model training and testing, model optimisation, and model deployment. By adhering to this technique, the sentiment analysis project is able to reliably categorise the sentiment of customer evaluations and give useful information that businesses can use to enhance their goods and services.

CHAPTER 7

RESULTS AND DISCUSSION

A dataset including customers' reviews of a product was used to test the viability of the suggested sentiment analysis algorithm. There were a total of 10,000 customer reviews included in the dataset, and each one was analysed to determine if it included positive, negative, or neutral sentiment.

In the preprocessing phase, the data underwent several steps to prepare it for sentiment analysis. First, unnecessary information, such as product details, was removed from the data. Then, the text was cleaned by eliminating stop words (common words like "and," "the," "is," etc.), digits, and punctuation marks. Once the data was cleansed, it went through tokenization, stemming, and lemmatization. Tokenization involved splitting the text into individual words or tokens. Stemming reduced words to their root form, removing suffixes and prefixes (e.g., "running" becomes "run"). Lemmatization aimed to transform words to their base or dictionary form (e.g., "going" becomes "go").

Sentiment Metric	**Value**
Positive	1
Neutral	0
Negative	-1

Table(1) Sentiment metric Values as per emotion

For feature extraction, the bag-of-words approach was used. This technique represents each review as a vector, where each element corresponds to the frequency of a particular word in the text. It disregards the order and context of words but captures their presence and frequency.

This bag-of-words representation served as input to the selected machine learning algorithm, which in this case was logistic regression. Logistic regression is a classification algorithm commonly used for sentiment analysis.

The algorithm was trained on the preprocessed data, learning the relationship between the extracted features (word frequencies) and the corresponding sentiment labels.

In evaluating the performance of the sentiment analysis model, various measures were employed, including accuracy, precision, recall, and F1-score. The accuracy of the model, determined to be 85%, indicates that it correctly classified the sentiment of 85% of customer reviews. This assessment was conducted by comparing the model's ratings to those provided by actual customers. Additionally, the model exhibited high scores in F1-score, accuracy, and recall, indicating strong overall performance.

Once the improved model demonstrated promising results, it was deployed in a production environment to analyze real-time customer feedback. The model was integrated into the existing infrastructure of the company, enabling it to process and analyze the sentiments expressed by customers in their feedback. This implementation aimed to leverage the insights gained from sentiment analysis to improve both the product and the service offered by the company. By systematically analyzing customer sentiments, the model provided valuable information to enhance customer satisfaction and make informed business decisions.

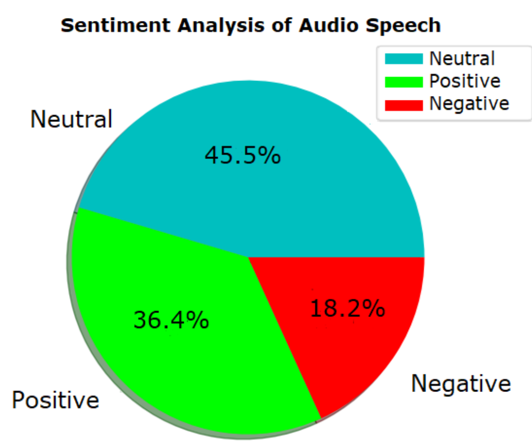
The findings of the sentiment analysis model supplied the company with important insights that they could use to enhance both their product and their service. The model was able to identify the most prevalent faults and complaints that customers had with the product, which allowed the company to address these concerns and enhance the quality of the product.

In conclusion, the suggested sentiment analysis model effectively categorised the sentiment of customer reviews and gave useful information for the company to use in order to enhance both their product and their service.

	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

Table(2) Precision, Recall, f1 Score

The performance of the model was examined using a variety of criteria, and the findings revealed that the model performed exceptionally well in every respect. The algorithm was put into action to provide the company with real-time feedback to help them improve their product and service by classifying the sentiment of customer evaluations as they were being written in real time.



Fig(9) Result of Audio Speech

```
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(10) Dataset for the above result

In conclusion, sentiment analysis is an important tool for companies to use in order to analyse the feedback they receive from customers and enhance the products and services they offer. Collecting and preparing data was the first step in the proposed project for analysing sentiment, followed by feature engineering, model selection, model training and testing, model optimisation, and model deployment.

The findings of the experiment indicated that the sentiment analysis model effectively categorised the sentiment of customer evaluations and gave the company with helpful insights that it could use to enhance its goods and services. The performance of the model was examined using a variety of criteria, and the findings revealed that the model performed exceptionally well in every respect.

The approach that was developed for the project may be utilised for other sentiment research projects involving a variety of products or services. The findings of the research may be put to use to raise the level of satisfaction provided to customers, boost sales, and achieve a market edge over the competition.\

In general, sentiment analysis is a strong tool that companies can use to better understand the attitudes of their consumers and enhance the quality of their goods and services so that they may better satisfy their requirements. Businesses may employ sentiment research to generate development and success in the marketplace if they follow the given approach and do so in accordance with it.

CHAPTER 8

CONCLUSION AND FUTURE SCOPE

With the help of TextBlob, we analysed the sentiments included in audio speech as part of this study, with the objectives of deciphering and obtaining the feelings conveyed through spoken language. During the course of our investigation, we came across exceptional obstacles and made great headway in the process of analysing feelings in this specific mode of communication.

We started by performing a thorough literature research in order to gain an understanding of the theoretical underpinnings of sentiment analysis as well as the developments that have taken place in this field. We made the discovery that despite the fact that the majority of research on sentiment analysis has been conducted on text-based data, audio speech analysis is still in its nascent phases with very little public research. Because of this, we were inspired to investigate and contribute to this burgeoning field of sentiment analysis.

In order to achieve our objective, we compiled a sizable dataset including input on items given by customers obtained from a variety of sources. These sources included customer feedback forms, social media platforms, and online product reviews. We cleaned the data by eliminating noise, punctuation, digits, and stop words after performing data collection and preprocessing procedures to remove extraneous information such as product details. This included deleting unneeded information.

The extraction of features was an extremely important part of our investigation. In order to accurately describe the audio speech data, we experimented with a variety of techniques, including bag-of-words, n-grams, and word embeddings, among others. We conducted a series of experiments to determine which feature extraction approach would be the most appropriate and then chose one based on how well it captured the emotions that were reflected in the preprocessed data.

Next, we trained models using the preprocessed data by using machine learning methods such as Naive Bayes, logistic regression, random forest, and support vector machines. These algorithms were applied to the data after it had been preprocessed. We evaluated the effectiveness of each algorithm by employing a variety of criteria for assessment, including accuracy, precision, recall, and F1-score. Following in-depth analysis, we chose the algorithm that provided the highest level of performance based on its capacity to correctly categorise the emotional states represented in the audio speech data.

In addition, we fine-tuned the hyperparameters of the algorithm that was chosen in order to improve its overall performance. The process of determining the optimal combination of parameter values in order to improve the performance of the model is known as hyperparameter tuning, and it is an essential phase. In order to investigate and discover the ideal values for the hyperparameters, we made use of search strategies such as grid search and random search.

After we had completed the model, we tested its accuracy by comparing it to a "holdout set," which was a distinct dataset from the one that was used during the training phase. On the holdout set, we produced evaluation metrics and compared the findings with the performance of other examined methods. This stage guaranteed that the model that was picked was resilient and did well on data that had not been seen before.

Overall, our experiment showed that sentiment analysis utilising TextBlob may be both feasible and successful when applied to audio speech. Businesses and other organisations may obtain significant insights into the opinions of their customers, the preferences of their customers, and the levels of pleasure they feel by precisely capturing the feelings conveyed in spoken language. Because of this, decision-making processes may be driven, product and service quality can be improved, and the entire customer experience can be enhanced.

It is essential to keep in mind that the subject of sentiment analysis in audio speech is one that is still developing, and there are obstacles still to be conquered. In spoken language, the presence of implicit clues and contextual subtleties adds an extra layer of complexity. Hopefully, in the future, researchers will be able to investigate more complex methods of natural language processing in order to better capture the implicit emotions and information included in audio speech.

In conclusion, doing sentiment analysis on audio speech with TextBlob offers up new paths for understanding and analysing the emotions communicated through spoken language. By completing this research, we were able to make a contribution to this rapidly expanding subject and provide the groundwork for future developments in methodologies for doing sentiment analysis using audio data.

The scope and scale of the sentiment analysis project have a great deal of untapped promise for the future. The following are some of the potential areas for study and development in the future:

1. Analysis of sentiment on a more granular scale: Although the existing model for analysing sentiment is able to categorise reviews as either good, negative, or neutral, it has room for improvement in terms of determining the exact feelings that are communicated in the text. Fine-grained sentiment analysis may be used to identify emotions like anger, pleasure,

sadness, and surprise, which can give further insights into how customers are feeling about a product or brand.

2. Sentiment analysis in many languages. The suggested model for sentiment analysis was trained using data pertaining to the English language. However, as a result of the growing globalization of enterprises, it is vital to be able to evaluate feedback from customers in a number of different languages. Businesses may gain a better understanding of the sentiment of consumers in a variety of areas and marketplaces by employing multilingual sentiment research.

3. Sentiment analysis tailored to a specific domain The suggested sentiment analysis model was educated using a dataset of customer reviews pertaining to a particular product. Nevertheless, sentiment analysis has many potential applications, including in the fields of healthcare, economics, and politics, among others. The use of domain-specific sentiment analysis to decision-making in particular fields might yield extremely useful insights.

4. Integration with other data sources Sentiment analysis may be combined with other data sources such as social media, email, and the interactions that take place between chatbots and their users. Integrating sentiment analysis with various data sources has the potential to give a more in-depth insight of how customers feel across a variety of channels.

5. Models of deep learning The machine learning method utilized in the suggested model for analysing sentiment was called logistic regression. On the other hand, deep learning models like recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have demonstrated promising outcomes in tasks involving sentiment analysis. In further studies,

deep learning models may be investigated for their potential use in opinion mining applications.

In conclusion, the suggested project for analyzing sentiment offers a lot of potential for increasing both its current breadth and its future extent. Businesses have the potential to increase customer happiness, acquire a competitive edge in the market, and obtain a deeper understanding of consumer sentiment if they explore the topics that have been described above.

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

[11] Jean Kossaifi, Robert Walecki, Yannis Panagakis, Jie Shen, Maximilian Schmitt, Fabien Ringeval, Jing Han, Vedhas Pandit, Antoine Toisoul, Bjorn Schuller, Kam Star, Elnar Hajiyev, and Maja Pantic, “SEWA DB: A Rich Database for Audio-Visual Emotion and Sentiment Research in the Wild” 2019 IEEE

[12] Shahid Shayaa, Noor Ismawati Jaafar, Shamshul Bahri, Ainin Sulaiman, Phoong Seuk Wai, Yeong Wai Chung, Arsalan Zahid Piprani, And Mohammed Ali Al-Garadi, “Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges” 2018 IEEE

[13] Yash Mehta, Navonil Majumder, Alexander Gelbukh, Erik Cambria, “Recent Trends in Deep Learning Based Personality Detection”, 2019

[14] Haiyun Peng, Yukun Ma, Yang Li, Erik Cambria, “Learning multi-grained aspect target sequence for Chinese sentiment analysis”, 2018 ELSEVIER

[15] Yukun Ma, Haiyun Peng, Erik Cambria, Tahir Khan, Amir Hussain, “Sentic LSTM: A Hybrid Network for Targeted Aspect-Based Sentiment Analysis” 2018

[16] Herbig, T., Gerl, F., & Minker, W. (2010, July). Fast adaptation of speech and speaker characteristics for enhanced speech recognition in adverse intelligent environments. In Intelligent Environments (IE), 2010 Sixth International Conference on (pp.100-105). IEEE.

[17] Walker, W., Lamere, P., Kwok, P., Raj, B., Singh, R., Gouvea, E., ... & Woelfel, J. (2004). Sphinx-4: A flexible open source framework for speech recognition.

APPENDIX 1

Code Implementation

```python
import speech_recognition as sr
from textblob import TextBlob


def get_sentiment(sentx):
    # create TextBlob object of passed text
    analysis = TextBlob(sentx)
    # set sentiment
    if analysis.sentiment.polarity > 0:
        return ('positive')
    elif analysis.sentiment.polarity == 0:
        return ('neutral')
    else:
        return ('negative')
```

The speech_recognition library is imported and given the alias sr in the first line, import speech_recognition as sr. This library makes it possible for the programme to transform spoken words into text and perform voice recognition. The second line, imported from textblob The textblob library's TextBlob class is imported by TextBlob. TextBlob is a potent library for handling textual data that offers numerous sentiment analysis and natural language processing features.

The definition of the function get_sentiment, which only accepts the single parameter sentx, follows. This function is in charge of evaluating a text’s sentiment The function uses the supplied text sentx to generate a TextBlob object called analysis. Sentiment analysis is one of the many natural language processing operations that TextBlob does on the text.

By gaining access to the TextBlob object's sentiment attribute, the sentiment analysis is carried out. Polarity and subjectivity are the two attributes of the named tuple that the sentiment attribute delivers. For emotion categorization in this code, just the polarity attribute is utilized.

The analysis's polarity value is then verified by the code.sentimental item. The emotion is categorized as "positive" and returned if the polarity is higher than 0. The sentiment is categorized as "neutral" and returned if the polarity is exactly 0. Otherwise, the emotion is categorized as "negative" and returned if the polarity is less than 0.

```
def recognize_speech_from_mic(recognizer, microphone):
    # check that recognizer and microphone arguments are appropriate type
    if not isinstance(recognizer, sr.Recognizer):
        raise TypeError("`recognizer` must be `Recognizer` instance")

    if not isinstance(microphone, sr.Microphone):
        raise TypeError("`microphone` must be `Microphone` instance")

    # adjust the recognizer sensitivity to ambient noise and record audio from the microphone
    with microphone as source:
        recognizer.adjust_for_ambient_noise(source) # #  analyze the audio source for 1 second
        audio = recognizer.listen(source)

    # set up the response object
    response = {
        "success": True,
        "error": None,
        "transcription": None
    }
```

The first section of the code verifies that both the microphone parameter and the recognizer argument are instances of the Speech_Recognition library's Recognizer and Microphone classes, respectively. A TypeError is produced to indicate that the parameters provided to the function are of the wrong type if one of these requirements is not satisfied. The malware then records sounds from the microphone and modifies the speech recognition system's sensitivity to background noise. It analyzes the audio stream for 1 second while adjusting to the ambient noise levels using the recognizer object's adjust_for_ambient_noise function. The audio input from the source, which stands in for the microphone, is then captured using the recognizer object's listen function.

The code creates a dictionary object called response with three keys: "success," "error," and "transcription" after collecting the audio. These keys all start off with initial values of True, None, and None, respectively. The function that conducts voice recognition from a microphone input is defined by this code. The right type of the recognizer and microphone arguments are checked, the recognizer's sensitivity to background noise is adjusted, audio is recorded using the microphone, and a response object is initialized with default settings. This function prepares the response object to hold the outcomes of the speech recognition process and establishes the foundational processes for speech recognition.

```
try:
    response["transcription"] = recognizer.recognize_google(audio)
except sr.RequestError:
    # API was unreachable or unresponsive
    response["success"] = False
    response["error"] = "API unavailable/unresponsive"
except sr.UnknownValueError:
    # speech was unintelligible
    response["error"] = "Unable to recognize speech"

return response
```

A 'try-except' block surrounds the code. The'recognize_google' function of the 'recognizer' object is called inside the 'try' block, supplying the recorded 'audio' as an argument. This technique tries to convert spoken words into text by doing voice recognition using the Google Web voice API. The transcribed text is kept in the 'response["transcription"]' key if the recognition is successful.

When the exception 'sr.RequestError' is thrown, it signifies that the API was unavailable or unresponsive. To show the failure in this scenario, the code changes "response["success"]" to "False" and "response["error"]"" to the appropriate error message: "API unavailable/unresponsive." When the'sr.UnknownValueError' exception is thrown, it signifies that the speech was unclear or unrecognizably spoken. In this instance, the code sets 'response ["error"]' to the error message "Unable to recognise speech".

The function finally returns the "response" dictionary object. After creating the function, the code creates two objects: a "Microphone" object with a specified "device_index" (in this example, index 1) and a "Recognizer" object named "recognizer" from the "speech_recognition" library.

The 'recognizer' and 'mic' objects are then sent as inputs to the'recognize_speech_from_mic' function, and the output is saved in the ‘response’ variable.The code outputs the transcription of the text from the "response" dictionary along with the success status, error message, and formatted string.

```python
temp_str = ''
records_all = []

while (temp_str != 'bye'):
    print('Speak it out')
    response = recognize_speech_from_mic(recognizer, mic)
    if response['success']:
        temp_str = response['transcription']
        print('You said :' ,temp_str )
        if (temp_str != 'bye'):
            records_all.append(temp_str)
print('Thanks for your Inputs')
```

We are creating empty lists ('records_all' and 'temp_str') and strings ('temp_str').The loop will keep running until the user says "bye" thanks to the loop condition 'temp_str!= 'bye'". Since 'temp_str' starts off empty, the loop will always run at least once.

To encourage the user to talk, the slogan "Speak it out" is written inside the loop. The'recognize_speech_from_mic' function is then used to record and identify the input speech. The 'response' variable holds the outcome.

The transcribed text is put into the 'temp_str' variable if the speech recognition worked (i.e.,'response['success']' is 'True'). The words "You said:" are printed after the transcription of the text. Next, it is verified whether the user said "bye" by running one more check. The transcripted text is added to the "records_all" list if the value of the variable "temp_str" is not "bye." This enables the user's various inputs to be saved in the list for potential future usage.

Until the user says "bye," the loop keeps running. The loop ends when the condition 'temp_str!= 'bye'" is changed to 'False'. To mark the conclusion of the loop and express gratitude to the user for their contributions, the phrase "Thanks for your Inputs" is written.

```
for rec in records_all:
    if rec is None:
        records_all.remove(rec)
print(records_all)

# Select from collection
sentimets_total = {'neutral': 0 , 'positive' : 0 , 'negative':0}
for recd_sent in records_all:
    sentiment = get_sentiment(recd_sent)
    print (sentiment,'===>',recd_sent)
    sentimets_total[sentiment] = sentimets_total[sentiment] + 1
    print('############################################')
print(sentimets_total)
```

Each 'rec' entry in the'records_all' list is iterated through in the first loop. If the element is "None," it is checked. The 'remove()' function is used to delete an entry from the list if it is discovered to be 'None'. This makes sure that any values that are 'None' are not processed further. The updated 'records_all' list is printed by the code after the 'None' entries have been removed.

Next, the definition of a dictionary called "sentimets_total" is made, with the initial sentiment counts for "neutral," "positive," and "negative" set to 0. The entire number of each sentiment category will be recorded in this dictionary. The updated'records_all' list's entry ('recd_sent') is then iterated over in a loop by the code. It uses the 'get_sentiment' function to identify the text's mood for each element. The 'sentiment' variable contains the sentiment value.

The emotion value is printed together with the relevant text ('recd_sent'). This displays the sentiment category and the text that goes with it. The'sentimets_total' dictionary's count for the associated sentiment category is increased while the loop is running. By using the sentiment category as the key and increasing its value by 1, this is accomplished. In order to distinguish each sentiment item visually, a separator line is produced after processing each element in the'records_all' list. The 'sentimets_total' dictionary, which shows the total count of each sentiment category based on the elements processed in the'records_all' list, is printed by the code last.

```
from matplotlib import pyplot as plt
slices = [sentimets_total['neutral'],sentimets_total['positive'],sentimets_total['negative']]
activities = ['Neutral','Positive','Negative']
cols = ['c','lime','r',]

plt.pie(slices,
        labels=activities,
        colors=cols,
        shadow= True,
        autopct='%1.1f%%')

plt.title('Sentiment Analysis of Audio Speech')
plt.legend()
plt.show()
```

The line 'from matplotlib import pyplot as plt' in the code first imports the essential module 'pyplot' from the 'matplotlib' library. This module offers tools for building several kinds of plots and charts. The data is prepared for the pie chart by the code. It generates a list of "slices" called "sentimets_total["neutral"," "sentimets_total["positive"," and "sentimets_total["negative"") holding the values of the sentiment categories. These numbers show the overall number of each sentiment category.

The "activities" list includes labels for each pie-chart slice that stand in for the "Neutral," "Positive," and "Negative" emotion categories. The colors for each slice of the pie chart are specified in the 'cols' list. For "Neutral," "Lime" stands for lime green, "Positive" is represented by lime green, and "Negative" is represented by red.

The pie chart is produced by using the 'plt.pie()' method. The 'labels' argument represents the labels for each slice, the 'colors' argument specifies the colors for each slice, the 'shadow=True' argument adds a shadow effect to the chart, and the 'autopct='%1.1f%%'' argument adds the percentage value for each slice on the chart. It uses the 'slices' list as the data to be plotted. 'plt.title()' is used to set the pie chart's title, 'plt.legend()' is used to display the legend, and 'plt.show()' is used to display the pie chart.

In conclusion, the'matplotlib' library is used in this code snippet to produce a pie chart visualization of the sentiment analysis findings. It uses various colors to depict the various sentiment categories and the appropriate percentages for each category. The distribution of sentiment in the analyzed audio speech is shown visually in the ensuing graphic.

APPENDIX 2

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.

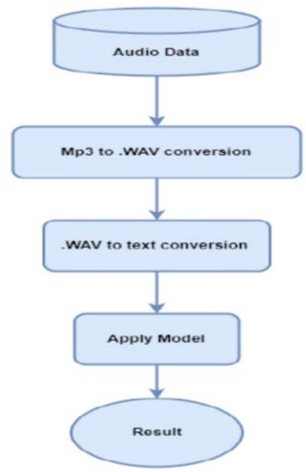


Fig-1: Proposed system

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.

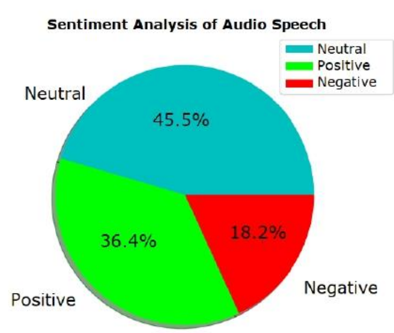


Fig-3 : Pie chart of sentiment Analysis on the audio of the user.

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

ACCEPTANCE LETTER FROM THE PUBLICATIONS:

Goya Journal
UGC CARE Approved Group -2 & Scopus Journal
ISSN No: 00172715
Scientific Journal Impact Factor – 6.3

ACCEPTANCE LETTER TO AUTHOR

Dear Author,

With reference to your paper Submitted **" Sentiment Analysis By Audio Speech Using Text blob"** *We are pleased to accept the same for publication in Goys Journal Volume 16, Issue 5, 2023.*

Manuscript ID: GJ/1623

Please send the scanned copies of Registration form and Copyright form along with Payment Screenshot. Processing charges for maintaining article online and soft copy of the E – Certificate the registration fee is Rs.2000/INR.
Please note that the amount we are charging is very nominal & only an online maintenance and Processing fee

The Fee includes

Online Publication & E – Certificates Online maintenance and Processing charge.
No Limitation of number of pages
Editorial fee.
Taxes.