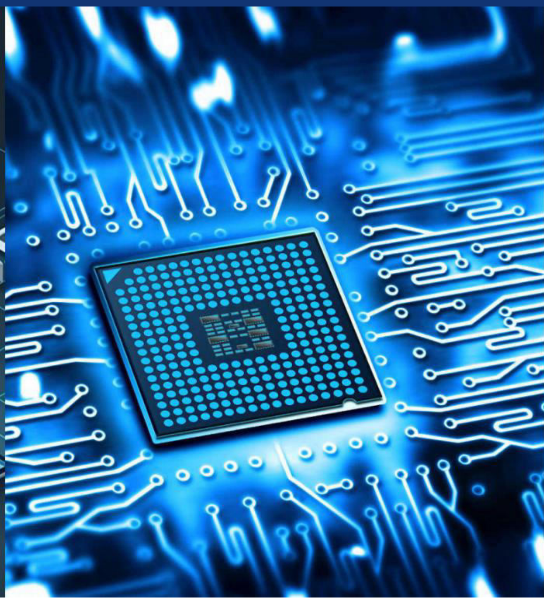
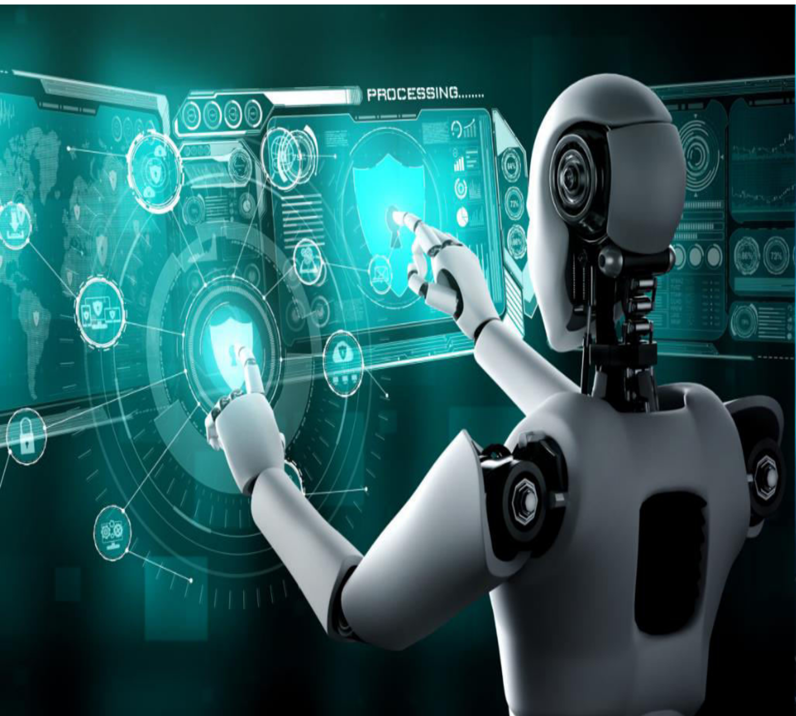
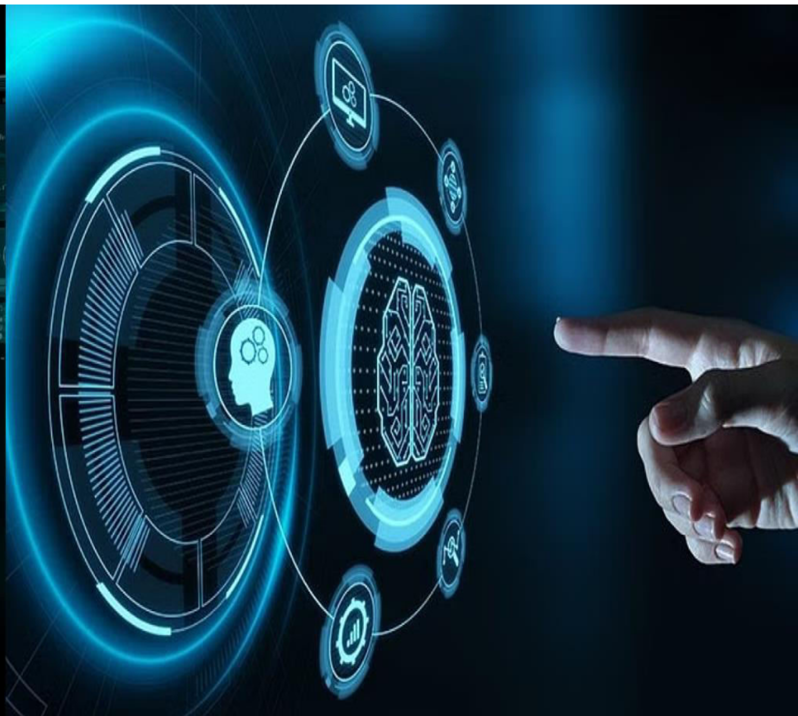
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 5, May 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sentiment Analysis of Incoming Calls on Helpdesk

Siddhartha G[1], Ankita H S[2], Keerthana[3], Shaik Mohammed Adil[4], Mr.Yamanappa[5]

UG Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, Karnataka, India

UG Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, Karnataka, India

UG Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, Karnataka, India

UG Student, Department of Computer Science and Engineering, Presidency University, Bengaluru, Karnataka, India

Professor, Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

ABSTRACT: The proposed project is a full-fledged system for sentiment detection with respect to audio that accesses the latest advancements of NLP and deep learning technologies for extracting as well as identifying the emotional tone of spoken language, packed into a Kivy-based graphical user interface (GUI) for real-time or file-based audio input. The audio is transcribed using Google Speech Recognition, and the emotion realized through the audio transcription is detected from the fine-tuned BERT (Bidirectional Encoder Representations from Transformers) model. The pipeline begins with normalizing and cleaning the raw input text before moving on to aggressive preprocessing steps like tokenization, stop-word removal, and lemmatization. The pipeline builds on a labelled dataset to train and fine-tune a BERT-based transformer classifier to categorize transcriptions into positive, negative, or neutral sentiment using techniques such as label encoding and training optimizations like dynamic adjustments of learning rate. The model, producer, and label encoder will be serialized for seamless incorporation into the GUI. The application allows interactions in various ways, such as text input, voice recording in real-time, or audio file upload, after which the transcription will be done automatically, and it will conduct the analysis of sentiments before visually displaying it. Then, every interaction by any user can be saved locally in CSV format along with timestamp details for later reference. This makes it an important feature for customer service, call centre support, and mental health monitoring. With modularity, the design will allow any new input into the system, such as multilingual support, fine sentiment scoring, or real-time conversation analysis. It also speaks volumes about the sophisticated integration of transformer-based deep learning, speech processing, and user-friendly interface design for interpreting human sentiment through voice.

KEYWORDS: Sentiment Analysis, NLP, multilingual speech recognition, BERT, customer sentiment.

I. INTRODUCTION

Sentiment analysis is also known as opinion mining and is well established here in the NLP domain[1]. The sensing and cataloguing of human sentiments, opinions, and attitudes in terms of texts or spoken words are what this branch really deals with. Sentiment analysis actually helps customer experience management to be used effectively in social media monitoring, brand reputation management, and customer service. In help desk scenarios, sentiment analysis means call differentiation based on personality characteristics; hence it helps organizations in prioritizing cases that require immediate action along with increasing the quality-of-service responses and increasing overall customer satisfaction. Sentiment analysis can mainly be categorized into three classes: lexicon analysis-based approach, which is the approach based on lists of pre-compiled dictionaries specifically for discerning sentiment; secondly its machine learning-based applications where models learn some classes from already labeled datasets and this will classify the sentiments; and finally, hybrid approaches which use both for better accuracy[2].The general classification of sentiments is based on positive to negative sentiment; however, the other complex models deal with complex emotions such as frustration, anger, or jubilation. Many challenges are there in sentiment analysis, for example, but not limited to the repetition of irony, ambiguity, and context-oriented emotional expressions, especially in utterance. However, as deep learning and artificial intelligence reach new heights, strides have been made in recent research in sentiment classification. It makes sentiment analysis a perfect position in business for better customer experience enhancement and also better decision-making.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sentiment analysis accounts for a wide range of applications of Natural Language Processing (NLP) in the interpretation and classification of human emotions distilled from text or speech[1]. The code-decodes human language into coherent parts such as the word, the phrases, and case-syntax by machines. Textual data analysis utilizes various techniques such as tokenization, stemming, and lemmatization for data preparation while others employing advanced techniques such as named entity recognition and dependency parsing enhance contextual understanding. Sentiment analysis is carried out with the basic form of three categories being 1) lexicon-based approaches, employing pre-made dictionaries of sentiment in which emotions polarity can be inferred from these dictionaries; 2) machine learning-based approaches, which train some models from labelled datasets and evaluate them to determine the sentiment of any post using that model; and 3) deep learned methods such as Long Short-Term Memory (LSTM) and transformer based models (like BERT), which allow the computations of very complex patterns of languages[3]. The overall pipeline of sentiment analysis contains preprocessing of data, extraction of features through methods like TF-IDF or word embeddings, and measuring of machine learning model for sentiment classification. These allow companies to curl information straight from customer feedback and open better chances for improving the quality of service and satisfaction of customers.

Bidirectional Encoder Representations from Transformers (BERT) has quickly been recognized as one of the top natural language processing (NLP) models developed by Google and has changed how we analyze sentiment by increasing contextual adoption of language. Previous models had capabilities to only analyze sentiment in a uni-directional manner (to read text, they either read left or right); with BERT, researchers can read text in both directions and have a more meaningful record of sentiment in terms of contextual meaning. BERT was pre-trained on very large corpora of text from Wiki and BooksCorpus[20]. Fine-tuning is a possibility for some tasks, like sentiment classification. BERT can represent and be fine-tuned to a certain extent, even with very little domain (i.e., domain specific) training data. Further, BERT can be utilized in applications like Flask or Kivy, where it automatically performs sentiment analysis of either user inputs as text, or of transcribed audio, in near-real-time[15]. BERT provides state-of-the-art accuracy and context localization, BERT is made up of a transformer architecture that allows for an easily trainable language based model that is not for one specific language. This means developers could create data sets that are not supervised and support adding more users easily[19]. As such, BERT continues to create opportunities for levels of accuracies in sentiment analysis systems and has raised the level of intelligent, user-centered, applications.

To perform an sentiment analysis, one would require machine learning models and frameworks for creating interactive applications and deploying models. Flask is commonly used and considered a lightweight web-based Python framework to create sentiment analysis applications[15]. A web-based, lightweight framework allows the user/developer to create APIs that will accept user input,either through text of voice, perform any type of algorithm needed and produce results for consumption in real-time . Flask is the way forward, it can serve sentiment models and provide a web interface, mobile app or chatbot application. On the alternative, Kivy is also a Python framework, but allows the development of cross platform applications, therefore we can use Kivy to produce a mobile sentiment analysis toolkit[18]. Moreover, using Kivy the developer/designer can create graphical user interfaces to collect user feedback, run sentiment analysis and display the results on mobile devices. Uber, Lyft and both Go are also supporting frameworks, FastAPI that can provide high performance as a swap for Flask for constructing APIs for sentiment analysis, or Django which is more sophisticated complete with actual data base managment capabilities. The above frameworks are often combined with machine learning libraries such as scikit-learn or Natural language processing (NLP) tools like NLTK so that practitioners can leverage sentiment analysis capabilities by producing simpler end-user friendly systems and systems to drive real-time customer feedback and support[17].

A helpdesk serves as the primary channel for clients to access assistance, information, or any other services that a client may require[5]. The nature and urgency of calls received by a helpdesk tend to differ, mostly, from customer to customer. Such calls typically include aspects of technical support where clients seek troubleshooting help with hardware or software, service requests such as account creation or changes in passwords, or billing questions such as invoice clarification or payment disputes[6]. In many cases, unsatisfied customers would make complaints and escalation calls that require maximum attention. Some calls may just pertain to inquiries on specific products or services. There are also calls that request feedback or conduct surveys whereby customers pass forth their views for improvement of the services. Responding to and addressing these different types of calls results in customer satisfaction and enhanced service productivity.

This analysis seeks to interpret the emotions in customer conversations and their influence on helpdesk calls. The helpdesk voice channel will initially be accessed; calls will be converted to text via Speech to-Text Software[7]. Next, the generated text will be analyzed using Natural Language Processing techniques that churn it to a status of positive or negative or neutral. Emotion Machine Learning models like Vadar, TextBlob, or any other deep-learning-based emotion analyzers will scan the tone, important words and phrases, and contextualize it to derive the emotions of the customer[17].This processed information is relayed to the dashboard in graphical format for helpdesk agents to highlight critical cases, prioritize escalation cases, and review retention strategies. The implementation of sentiment analysis on customer support calls would enhance customer support, productivity of agents, and satisfaction of customers.

II. RELATED WORK

The evolution of technology and increased use of deep learning in NLP has greatly changed how companies offer their customer services with the implementation of advanced sentiment analysis tools. I focus on the above problem in my research with regard to belief representation and reasoning models tailored for deep learning based on human emotions and language processing. The model attempts to uncover patterns related to accurate sentiment classification by analyzing call center data from past customer service interactions[5]. Communication is considered sequential in nature, thus requiring a Recurrent Neural Network (RNN) or Long-Short Term Memory (LSTM) approach to model spoken language, taking into account its contextual hierarchy. With appropriate tuning of parameters and system optimization for these deep learning frameworks, sentiment detection was significantly improved even in the most challenging scenarios of multi-turn conversations[1]. It is noted that deep learning techniques outperform traditional methods in recognizing subtle emotional nuances, such as sarcasm or sudden changes in sentiment, during interactive exchanges. This certainly highlights the unparalleled performance of deep learning in improving the adaptability of customer care systems in regard to real-time precise sentiment analysis.

The increase in automated sentiment analysis calls for enhanced machine learning solutions to improve service efficiency in call centers[15]. This research applies supervised learning techniques with Support Vector Machines (SVM), Random Forest, and Gradient Boosting for sentiment classification across conversations. There is dual focus on feature extraction with TF-IDF and word embeddings as well as on Mel Frequency Cepstral Coefficients (MFCCs) and pitch contours for tone variation analysis. Experimental results indicate improved performance with ensemble methods in classification accuracy. It is also true that other conventional models suffered from lacking empathy towards contextual subtleties like sarcasm, or emotion detection, thus needing advanced deep learning models like Recurrent neural Network (RNN) and BERT transforming Attention models capable of contextual comprehension and understanding of order sensitivity dynamics[2]. This study certainly addresses the gaps in automated sentiment analysis in call centers and shifts the focus towards more precise metrics for modular systems designed for automated systems.

Emotional Intelligence, which is becoming one of the key factors in enhancing customer experience in support interactions, emphasizes the importance of sentiment detection[6]. This paper proposes a solution that combines verbal and non-verbal elements for better sentiment analysis accuracy in telephone conversations. The system ensures simultaneous extraction of temporal and contextual dependencies on a signal level by applying deep learning techniques known as CNNs for acoustical feature extraction of pitch, tone, voice, and modulation, and Bi-directional Long Short-Term Memory (BiLSTM) networks for text analysis[9].The study presents a hybrid approach for sentiment classification by employing feature and decision-level fusion for even more sophisticated sentiment analysis. It is this fusion approach that provides for better predictive performance in detecting complex emotions like sarcasm, frustration, and empathy, often difficult for more traditional models to detect. The attention mechanism enables a stricter control of the model context, thereby improving its sensitivity to changes in the conversation flow[10]. With these features, the model is expected to greatly improve sentiment detection accuracy and advance the development of intelligent, emotionally responsive customer support systems.

Real-time sentiment analysis of incoming voice calls sets many dilemmas for dynamic environments, with different consuming noise and quality of the audio for every call. This paper proposes a holistic framework that integrates speech- to-text conversion, text sentiment analysis, and audio sentiment analysis to address most of these problems. The most advanced ASR systems convert speech into text which can be used for analyzing sentiment with the use of pretrained models like BERT and RoBERTa that demonstrate high and superior capabilities in the comprehension of contextual meanings[8]. It also acts simultaneously whereby CNNs analyze acoustic features which include pitch, tone,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

and speech patterns in ways that the audio signals can then be classified as having emotional cues. Text and audio-based analyses combined give better and consistent classifying ingong customer emotions. Thus, a dual-modal system enhances robustness, especially in a noisy environment, because text alone would not sensibly classify emotions in such a case. Therefore, this framework improves real-time sentiment detection efficiency, thus resulting in better quality interaction with customers and efficient management of emotional conversations.

A complex technique is necessary in identifying consumer emotions in interactions concerning services rather than just the textual modality. Thus, this paper enhances the sentiment detection through customer calls by incorporating acoustic and lexical features into the system. Voice modulation, pitch, speech rate, and tone are analyzed in conjunction with the text, making it possible to appreciate emotions quite well. Again, for account feature extraction, viewing or calculating Mel-Frequency Cepstral Coefficients helps the model grasp distinctions in speech that can refer to anything from frustration to urgency[13]. While it does so, word embeddings for lexical analysis allow the capture of words bearing sentiments and contextual meanings. This multimodal study should be reserved for particular authentication purposes relative to recognizing concealed sentiments like sarcasm, hesitation, and annoyance that present insolvable issues in textual sentiment analysis[14]. Such a research shows that multimodal sentiment analysis will always refer to the use of combinations of verbal and nonverbal signs in increasing accuracy in identifying customer emotions in service contexts.

With the complexities of interaction, sentiment analysis escalated from just a simple transcription of voice Call Centers into sophisticated audio-visual processing techniques. The old paradigm often classified issue arbitrary based on algorithms that analyzed the textual data from machine learning-based transcripts; however, modern sentiments analysis has shifted drastically to require the depth and immersion into different approaches and modalities. Related to all sorts of customer-directed subjective modeling, they can be categorized as textual, auditory, or hybrid. Text usually refers to models that analyze the transcription of the conversation to detect emotions by extracting features using natural language processing and word embedding. In contrast, audio-based models go even deeper than that and extract features like intonation, pitch, and speech speed to make evocation stretch miles beyond the limits of the functionalities provided by written words[4]. The combination of the best qualities from both these systems is what's called hybrid models, where both would join to leverage the best attributes of each model to provide a more accurate interpretation of sentiment. Also, the application of end-to-end neural networks, which is part of deep learning technique advances, has led to very promising improvements in sentiment detection as understanding the subtlety of contexts turns possible with this method[11]. Further advances, such as transfer learning and pre-trained language models, have increased performance levels and ensured that sentiment analysis could not be easily disrupted by background noise, differences through accents, and those ambiguities related to context. These progressions underscore the necessity of a multimodal approach to have a more accurate and clearer picture of customer emotions.

III. METHODOLOGY

1) SYSTEM ARCHITECTURE

The system architecture for the multilingual sentiment analysis tool facilitates interaction with the different components in a way that allows for efficient audio processing and sentiment analysis. It was developed in order to give the user a seamless and responsive experience, with one application managing disparate input applications and audio processing functions.

The UI was completed utilizing Kivy, a Python framework that supports cross-language development. Making it user-friendly and accessible was the key goal they constructed several critical functionalities to allow end users convenient and easy access to the program. The user can select their preferred language from a drop-down menu so that the transcriptions take into account variances based on languages. This system will allow the user to live record audio to analyze on the fly, but also allows the user to upload recorded audio files to be processed. The system will then show the user the sentiment analysis results organized in documents containing transcriptions, detected sentiments, and emotions. The user will also have navigation controls to start the analysis, delete previous entries, and export the data into CSV format for a readout for final evaluation.

The Voice Recognition module is utilizing the speech_recognition library to achieve speech to text conversion. Specifically, it supports real-time transcription through microphone input, which involves capturing live audio and processing the audio data respectively. Users can also import audio files in different formats, and the module operates on the use of a range of different speech recognition techniques to convert the audio into text. Also, the language adaptability at the user level allows users to practice in multiple languages, and all transcription and feedback for mistakes will be done as accurately as possible.

For Emotion Evaluation functionality, the system relies on two libraries: TextBlob and VADER. TextBlob provides sentiment polarity detection giving a -3 to +3 score based on positive, neutral, or negative classifications. VADER extends this analytics to capture emotional nuances that could be very subtle; it is especially useful at for analyzes conversation speech and social media text. Together, the two libraries give a user a profound and extensive view about sentiments and emotional tone of the transcription contents.

The Data Management module utilizes SQLite for efficient storage and retrieval of sentiment analysis data. It guarantees that all analytics include original text, translated text (if available) and sentiment scores as well as detected emotions, and corresponding timestamps are securely stored. The system also supports history management, which means users can reference sentiment analysis history easily. The CSV export feature will also allow the users to download their sentiment analysis results to be further analysed or to present.

In order to enable Integration and Processing for all system components, the workflow is intended to run seamlessly integrated. The flow begins with users uploading an audio file or recording in real-time. Once the audio is ingested it is then processed through speech recognition, turning that audio into text. Upon transcribing the audio to text, the text then goes through sentiment and emotion detection to generate insights. The results are then saved to be stored in the database while also displayed on the UI, ensuring that users can easily see their data retrieved from a sentiment analysis process to act on.

The integration and processing steps in the workflow are intended to function tightly, without interruption, across all systems. The workflow starts with the users uploading a specific audio file or recording real-time audio. The audio file is then run through a speech recognition process that outputs text. The text is then sentiment and emotion analyzed, which produces valuable meaning. The resulting information is stored in the database and shown on the UI, which allows users to view and interact with their sentiment analysis data.This integrated environment produces a fluid and robust user experience that provides real-time and batch processing of sentiment analysis, while ensuring accurate results that are readily consumable.

2) DATA COLLECTION

Data is gathered through two modalities of devices in order to give users additional choices. The first modality is Live Audio Recording, the application allows user to record audio from the user's microphone in real time. There are several notable features included in this modality. Language Selection offers a list of languages for users to select from and allows the speech recognition library to draw from phonetics and vocabulary based on the selected language. The Real-Time Processing allows the application to capture and segment the audio while transcribing simultaneously giving the user short feedback loops to help with learning. Lastly, the Ambient Noise Adjustment allows the user to fine-tune the gain of the audio before recording the audio to improve the audio quality which turns into better transcription. The second modality of devices is Pre-Recorded Audio Files, the user either can record live audio or they can upload an audio file to be analyzed later. There are some notable features found in the pre-recorded audio file modality as well even though they are fewer compared to the other modality[2]. Supported Formats provides the user the ability to work with a variety of audio formats(WAV) the whole time during playback and listening. The format conversion converts the users uploaded mp3 audio file to a wav format to ensure the audio file used for the audio file can be processed by the speech recognition algorithms and interprets audio's best quality specification to allow for processing.In conclusion, Persistent Execution offers the ability for serial processing for your audio during upload, like with live recordings, does sentiment analysis on the transcribed text, and analyzes the emotional state as well. The application can analyze audio recordings live and audio recordings you have already done, offering a simple and practical approach to sentiment analysis in a plethora of environments to a vastly different user base.

3) SPEECH RECOGNITION

Speech Recognition is a core part of the application that translates spoken language into transcribed text. The entire process is fairly involved, but it consists of several important steps. Once a user starts the speech recognition, the user is prompted to select a language among a handful of options, such as English, Spanish, Hindi, and others. This is important to ensure the transcriptions are as accurate as possible as the algorithms run speech recognition are specific to selected languages. The application updates the parameters for recognition based on the selected language, because it has to account for phonetics, accents and dialects. Before the application listens for any audio input, it attempts to lessen background sounds through noise-cancellation techniques as to have cleaner audio input, which adds to the best chance at accurate transcription.

As the user speaks, the application only listens for inputs through the microphone in small chunks, and processes the audio input simultaneous to recording the audio. This allows the end user to receive real-time feedback with minimal lag and that makes the application an interactive experience.The official texts are produced by Google's speech recognition service, which is known for its high accuracy and multilingual support. Google speech recognition uses machine learning algorithms to recognize spoken words and then convert those spoken works into written text. The speech recognition engine allows the machine to understand contextual meaning, which is essential to help distinguish between homophones and phrases with multiple interpretations important for recognizing conversational speech.

4) SENTIMENT ANALYSIS

User sentiment and emotions assessed through meaningful insights in transcribed text create an important process in the application. Text Analysis involves the first step in the sentiment analysis process where the text is brought under examination utilizing powerful libraries such as TextBlob to process textual data. It assigns a sentiment polarity value between -1, with positive meaning a negative sentiment, and +1 meaning a positive polarity. This is helpful in further analyzing how to better anchor an overall discussion of the text sentiment. Furthermore, for greater detail in emotional subtleties, the application makes use of VADER Sentiment with the emotion squirrelled away in a social media context or with everyday language.

VADER takes some of these features into consideration: punctuation to enhance wording with strength due to its use; capitalization allows for the importance of caps (LOVE vs love); negations for words like in the example "not happy." The next step deals with the analysis of emotion detection: this is based on keyword-based emotion detection where events are triggered by taking key stones that relate to different emotional categories in the emotion lexicon. With an emotion lexicon like POSE, the transcribed text is scanned for those keywords making the flag concerning the emotion:. Putting this, the system can detect multiple emotions from the same piece of text. Following up would be structuring the outputs into organized sets that present the transcribed audio text by the original language, the sentiment analysis scores from TextBlob and VADER, and a specific area of detected emotions keyword analysis from the audios in their original language from which the text description is derived[15].

The results of these processes then go on to be kept within the SQLite database for lightweight yet efficient data handling and easy retrieval so that users can view the results of past analyses through an interactive interface built for fast review and comparison, while also including the option to export to CSV for external analysis or reporting from the application. The entire procedure also includes a feedback loop, which in essence is the solicitation from the users for opinions regarding how accurate sentiment analysis has been once the results had been made available out to them. This undoubtedly is a very important aspect which will keep improving the keyword lexicon and sentiment estimation algorithms. If these are to be achieved, then the sentiment analysis module will be responsible for transformation from transcribed text to useful insights assisting the user to get a better feeling over What is said in speech or language.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

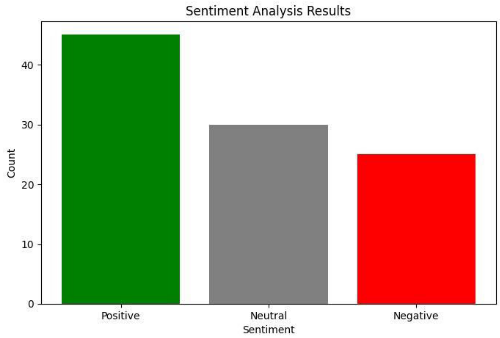


Fig:3.4.1 Sentiment Analysis

5) DATABASE MANAGEMENT

The Database Management System focuses on maintaining suitable storage, retrieval, and data maintenance of results from sentiment analyses. SQLite, which acts as a primary database management system, facilitates better data management. A few important steps need to be followed in the management of databases for smooth running of data processing and ensuring security.

The first step is to Initialize the Database that is before the data inputted in the database schema has to be defined. Preparing the tables and columns for the required data types is part of this process. The most general table will have the ID(I.D.- Unique identifier of the entry), Original Text(the text out of the audio input), Translated Text(if any), Sentiment Polarity(sentiment score out of TextBlob and VADER), Detected Emotions(a list of emotions from the text), Language(an original language with respect to the audio), and Timestamp(the time when the analysis was performed). Then each column is assigned by a type that matches it, for example, TEXT, INTEGER, or REAL, according to the efficient storage and retrieval of the data.

Once the database schema is set, the insertion of data starts. The extracted data is inserted into the database after doing sentiment analysis. There is a prepared statement here for making the data secure and in keeping with integrity, so as to avoid SQL injection. Furthermore, batch processing techniques can be applied for better performance when multiple audio inputs are analyzed in one session. An error-handling mechanism is in place to catch and fix a few errors—like duplicates entered into the database or failures to connect to the database, along with making the user aware of the error due to the database.

The system also supports History Retrieval, allowing users to access their past sentiment analysis reports through a Customer Query Interface. This feature enables users to quickly retrieve and review their last few sentiment analysis records, such as the most recent 10 or 20 entries. Users can also view detailed insights into each entry, including the original text, sentiment analysis results, emotion analysis, and timestamps. Additionally, the system includes a search functionality, which allows users to filter records based on date range, keywords, or other criteria, making it easier to find specific sentiment analysis results.

The system is equipped with a data export feature in order to improve the usability. This feature enables users to export sentiment analysis reports in CSV form. This would be beneficial for analysis, reporting, or using with other tools such as Microsoft Excel or Google Sheets for further processing. Users can choose to export all data as is, or filter results by specified date ranges or keywords to create more specific report exports. Before the exportation takes place, there is a check for any data validation to ensure that everything is complete and well formatted before export, thus ensuring a low error margin in the output file.

In a nutshell, the system employs Data Security and Backup measures for the protection of user information and maintenance of data integrity. Only authorized users are permitted access to the database at all times by access control mechanisms, thereby preventing unauthorized modifications. Encryption techniques may also apply to protect sensitive information. The system also provides regular backups in order to prevent data loss. Users are provided with an option of initiating or scheduling automatic backups. All these are meant to ensure that the sentiment analysis records are safe and secure against the system's failure.

Overall, Database Management is an important player in putting sentiment analysis data securely in the store, easily retrievable, and quite well managed. The system also makes a sturdy platform for an effective run of controlling results in sentiment analysis through structured data handling, security provisions, and user-friendly features.

6) USER INTERACTION

The UI of the application is designed with the intent of being intuitive and quite easy to interact with, hence providing a seamless experience for users while using the speech recognition and sentiment analysis functionalities. Key functionalities include the following:

Live Audio Analysis and File Uploads:

- Microphone Access: Users will activate live audio analysis when they click a very conspicuous button activating the microphone. The app will request permission to get access to the user's microphone with a priority of privacy and security over the user.
- Visual feedback: The audience sees the visual feedback while recording in the form of a waveform being displayed or a recording light turned on, which reassures the users that the audio is getting captured in the right way.
- Instant processing: The audio will be processed right after the capturing and thus the audience will get to instantly see the results of sentiment analysis right after the speech recognition is completed.
- Upload Pre-Recorded Files:By clicking an upload button, users can easily upload pre-recorded audio files, which will open a file dialog. Users are informed of the supported formats- like MP3, WAV, etc.- into which files are to be uploaded.
- Processing Feedback: Once a file has been uploaded, then the app sends the feedback about the upload, after which the audio will be processed for analysis.

Language Selection for Speech Recognition: In an easy-to-navigate dropdown menu high on the page, users are allowed to select the language in which they would like to have their speech recognized. Among the listed languages are some most widely used, like English, Spanish, Hindi, et cetera. The application is then enabled to change its recognition settings and adapt different characteristics to provide the higher performance for the selected language in terms of transcription and analysis.

Real-Time Results of Sentiment Analysis: After being processed, users will be able to see the results of the sentiment analysis in real-time. Such information includes the following: the Transcribed Text: verbatim text directly from the audio input being analyzed; A clear indication of the sentiment score ranging from negative to positive (usually color-coded or indicated by emoticons); The listed emotions based on keyword analysis in a very clear format; And then the user can go back and see a history of their analyses in a very neat section of the UI: Users would be able to click on a specific entry and, for example, see detailed info that includes timestamps and sentiment scores, allowing them to keep track of how their results are changing over time.

Database Management Options The database clear function allows users to remove all historical data. Users will receive a confirmation prompt to ensure they do not delete anything by mistake, thus ensuring that they understand fully what they are agreeing to.

7) IMPLEMENTATION TOOLS AND LIBRARIES

1.Kiwi consists of several libraries written in Python which turn PST into an open-source multi-touch interaction environment for project design -cross-platform user interfaces that run on Windows, Mac, Linux, Android, and iOS-creating in minimum time and effort. This library supports multi-touch gestures for enabling very interactive user interactions along with a very strong artwork engine which uses OpenGL ES-based smooth animations and visual effects. This library supports multi-touch gestures for enabling very interactive user interactions along with a very strong artwork engine which uses OpenGL ES-based smooth animations and visual effects. This library supports multi-touch gestures for enabling very interactive user interactions along with a very strong artwork engine which uses OpenGL ES-based smooth for an effect.

It also makes available additional libraries with editions that allow further flexibility in GUI. With it, you can create GUIs that contain the myriad possible components like buttons, scrolls, text input signals, and layouts. You have those various options for developing user interface designs even having the option of adopting multi-touch gesture support for richer user interaction and an excellent artwork engine that takes you through smooth animation and visual effects using OpenGL ES2.

2. Speech: The library serves as a bridge to several addresses for the engine and API that allow spoken word conversion to text. In addition to supporting several libraries, including Google's Web Speech API, CMU Sphinx, and Microsoft Bing, it provides flexible implementations. Real-time audio processing assists live-much-as translation from real-time to pre-recorded audio files on the basis of user ability and need. The library has simple APIs that provide functionality for speech recognition, turning it into a trivial affair for a user with no programming knowledge.

3. TextBlob and VADER, that are essential for emotion detection in texts. The former has a very straightforward process to get the probability of sentiment polarity, which would definitely allow one to garner objective insight on the emotional tone in some text. It fine-tunes sentiment analyses expressed in one's social media status and conversation: more nuanced work, if you will. Thus, together, the two libraries will enhance the understanding of the application regarding the user's sentiment.

8) TESTING AND EVALUATION

In the evaluating and testing stage of the application, it was crucial to demonstrate the applications functionality, performance and satisfaction. An overall approach has been taken to test the various components, for example, speech recognition accuracy, emotion analysis, and user overall experience. Accuracy of speech recognition was of primary interest from the tests of many of the audio input. The researchers tested under quiet and noisy conditions to find out how well the speech recognition component performed. The accuracy tests also included assessments like Word Error Rate. Additionally, documenting user interactions and observing the provided action as an alternative to speech recognition offered the opportunity to amend the system as well.

Another aspect of the evaluation step included reliability of emotion analysis. The reliability of the emotion classifier was measured based on performance with representative types of positive, negative, or neutral emotion-transcribed texts. The classifier output was measured according to proficiency against normative tests that proved the system has the ability to successfully identify emotions of happiness, anger, and sadness correctly. The reliability and accuracy of emotion analysis were evaluated with calculated indices including precisions and recalls. User input was also valuable in assessing the validity of the emotion classification since it would help to understand how the workings of the system under real conditions.Usability and interface functionality were also evaluated to ensure continuity of navigation and interaction. To do this, testing was based on user interaction of the extensive resource using real-life scenarios to complete all evaluation.

IV. CONCLUSION

Through this research paper, the development of the sentiment analysis system is a real advancement in our capability to understand individual's feelings and use technology to interact with those feelings. With the audio input mechanism and transcription process as well as deep learning algorithms such as BERT, the implementation proposed here has the potential to be a powerful way to capture and understand the emotional meaning of spoken communications. In addition to real-time emotion recognition, it also advances human-computer interaction in contexts such as mental health monitoring, customer service, and personal well-being.The ease of use and inclusion of features such as audio recording, live predictions, and the ability to log historical sentiment are critical aspects of the system. These elements encourage users to become self-aware and self-reflective, allowing users to track their emotional changes over time, as well as limited and expanded imaging, and prompt action toward emotional balance. As users identify sentiment patterns and emotional trends, they are more likely to engage in conscious habit-building techniques regarding stress, communication, and overall mental wellness. This research paper builds on the potential of innovative AI models

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305098

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

combined with well-designed user processes to enhance emotional intelligence for the user and the organization. Considering our current awareness of mental health and the increasing need for emotionally intelligent systems, this application is no longer a mere demonstration of technology resource; it is a solution vital in advancing empathy and emotional well-being of our digitized societies.

V. RESULTS

The real-time sentiment analysis apparatus enables the user to input live-recorded speech, or uploaded speech audio samples of any length, which are then transcribed and instantly analyzed with a fine-tuned BERT model. This rapid feedback allows not only for near-real-time prediction of sentiment but also gives 'emotional intelligence' responses for many end-client applications. The sentiment analysis system uses automated transcription through the `speech_recognition` library. It accurately transcribes human speech, converting it into a usable text output while also feeding it into the subsequent step of predicting sentiment, thus providing the user with an easy to read transcript and timed sentiment feedback related to the transcript contents. All predictions, timestamps, and transcripts are automatically saved in CSV format to facilitate users being able to track, backtrack, and observe trends related to emotional behaviors over time. It is largely built to be user-friendly with a Kivy-based GUI, which allows non-technical users to easily upload audio samples, as well as view sentiment and graphics of the model's decisions for their particular audio sample. This thought was also put into the system's flexibility by considering multiple potential-client sectors. While retraining (or modifying) may be required to use this investigation tool and application in areas like healthcare, education, customer service, or personal wellness, it provides single-tool capability across diverse sample types. Moreover, it demonstrates consistent performance across a variety of speech samples. Even if there is different background noise, and tonal characteristics (often related to the speaker) the model appears to maintain the same general predictive sentiments (especially given the lengthy samples analyzed have two coupled results). In relation to the intended use there can also be future flexibility built into the system for more granular recognition of emotions including anger, joy, or sadness.

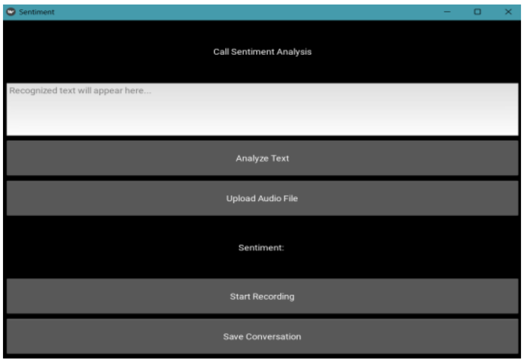


Fig:5.1 Sentiment Analysis Page

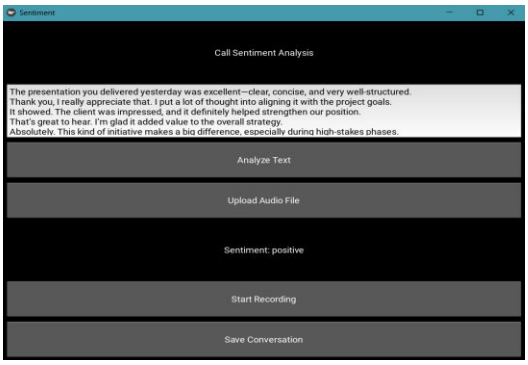


Fig:5.2 Sentiment Analysis - Positive

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

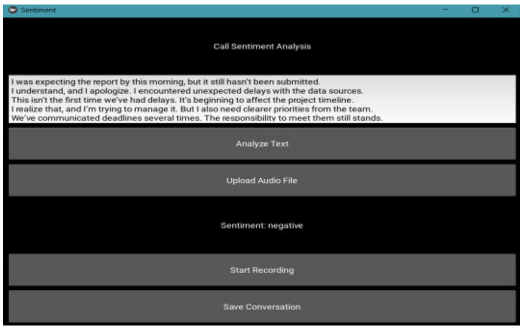


Fig:5.3 Sentiment Analysis – Negative

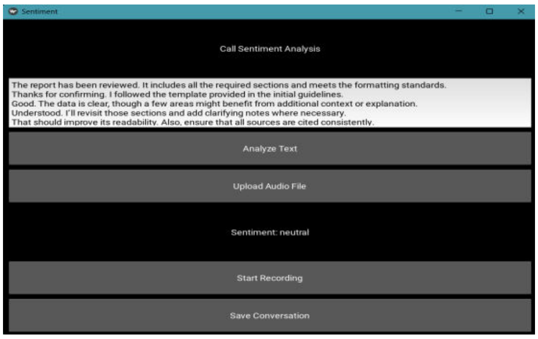


Fig:5.4 Sentiment Analysis – Neutral

REFERENCES

[1] A Deep Learning System for Sentiment Analysis of Service Calls (Yanan Jia, 2020).
[2] A Survey on Sentiment Analysis (R Raja Subramanian, 2021).
[3] Acoustic And Lexical Sentiment Analysis For Customer Service Calls (Bryan Li, 2019).
[4] Automated Speech Recognition System in Advancement of Human - Computer Interaction (Soumya Priyadarsini Panda, 2019).
[5] Call center sentiment analysis systematic review (K J Shrikrishna Narayana, 2023).
[6] Call Redistribution for a Call Center Based on Speech Emotion Recognition (Milana Bojani´c, 2020)
[7] Call-Type Classification And Unsupervised Training For The Call Center Domain (Min Tang, 2003)
[8] Emotion Recognition In Emergency Call Centers: The challenge of real-life emotions (Théo Deschamps-Berger, 2021),
[9] Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis (Saif M. Mohammad, 2022),
[10] Uday Exploring Sentiment Analysis Techniques in Natural Language Processing: A Comprehensive Review (Karthick Prasad Gunasekaran, 2023).
[11] Incorporating End-to-End Speech Recognition Models for Sentiment Analysis (Egor Lakomkin, 2019).
[12] Interactive Voice Response using Sentiment Analysis in Automatic Speech Recognition Systems (Rohit Raj Sehgal, 2018).
[13] Keywords Extraction and Sentiment Analysis using Automatic Speech Recognition (Rachit Shukla, 2020).
[14] Low-Cost Call Type Classification for Contact Center Calls Using Partial Transcripts (Youngja Park, 2009).
[15] Machine Learning-Based Sentiment Analysis of Incoming Calls on Helpdesk (Suhas Chavan, 2023).
[16] Ordinal Learning For Emotion Recognition In Customer Service Calls (Wenjing Han, 2020)
[17] Sentiment Analysis and Emotion Recognition from Speech Using Universal Speech Representations (Bagus Tris Atmaja, 2022)
[18] Sentiment analysis methods, applications, and challenges: A systematic literature review (Yanying Mao, 2024)
[19] Sentiment Analysis of Customer Support Phone Dialogues using Fusion-based Emotion Recognition Techniques (Yitong Wang, 2019)
[20] Sentiment Analysis Using Machine Learning Algorithms (Mohamed Hayouni, 2021)

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING