Volume 9, Issue 7, July – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/24jul1852

Sentiment Analysis of Incoming Voice Calls

Mahesh Kumar Chaudhary [1]; Mahima Sahu[2]; Manu Priya K[3]; Pujashree V[4]; Suguna A[5]

[5]Professor

[1, 2,3,4,5] Computer Science and Engineering, Sri Sairam College of Engineering, Anekal, Bengaluru-562106

Publication Date: 2025/06/27

Abstract: This project aims to meet the increasing need for real-time sentiment analysis within voice call interactions, acknowledging the rising significance of voice-based engagements in today's telecommunications realm. The proposed framework utilizes advanced natural language processing (NLP) techniques and machine learning models to promptly evaluate emotional nuances, integrating voice signal processing, feature extraction, and sentiment classification to ensure adaptability across diverse linguistic and cultural contexts. This initiative not only introduces a robust framework for real-time sentiment analysis but also tackles challenges specific to voice-based communication. Its wide-ranging applications span across industries such as customer service, market research, and social monitoring, offering valuable insights for organizations to comprehend and effectively respond to sentiments expressed within the dynamic landscape of real-time voice communication.

Keywords: *Sentiment Analysis, NLP, NLTK, Voice to Text.*

How to Cite: Mahesh Kumar Chaudhary ; Mahima Sahu; Manu Priya K; Pujashree V; Suguna A (2024) Sentiment Analysis of Incoming Voice Calls. *International Journal of Innovative Science and Research Technology, 9(7),* 3575-3584. https://doi.org/10.38124/ijisrt/24jul1852

I. INTRODUCTION

The main objective of this project is to create a real-time sentiment analysis system specifically designed for incoming voice calls. By leveraging advanced natural language processing (NLP) techniques and machine learning models, this system aims to bridge the gap left by traditional sentiment analysis tools, which primarily focus on written text and often overlook the subtleties of emotional expression in spoken language.

Given the increasing prevalence of voice-based interactions in customer service and virtual assistants, there's a pressing need to decode emotional cues in real-time. This system will utilize NLP to capture these nuances and machine learning algorithms to ensure adaptability across diverse contexts.

A key distinguishing factor of this project is its real-time capability, which allows for instant analysis of voice calls. This is particularly crucial in scenarios like customer support, where quick comprehension of sentiment can significantly enhance service quality. The envisioned system promises organizations transformative insights, enabling improved customer engagement, enhanced service quality, and informed decision-making based on real-time emotional intelligence.The call can immediately be recorded and uploaded in our system.

By uncovering the emotional dynamics embedded within voice calls, businesses cantailor their responses more effectively, identify areas for improvement, and establish deeper connections with their audiences.

II. LITERATURE REVIEW

According to Yang, L., Li, Y., Wang, J., & Sherratt, R. S. [1] at, Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. Their study delved the new SLCABG model for detection of emotions. Machine Learning-Based Sentiment Analysis of Incoming Calls on Help Desk.

According to Kokane, C. D.,[2], this paper discusses how to convert the user's audio input to text and analyze it. The Effect of Different Occupational Background Noises on Voice Recognition Accuracy.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/24jul1852

According to Li, S., Yerebakan [3], this paper discusses the regular Convolutional Neural Network (CNN)-based voice recognition algorithm and Auto Speech Recognition (ASR)-based model with a denoising module. Application of probabilistic neural network for speech emotion.

According to Deshmukh S [4], this paper introduces SER system using Probabilistic Neural Network Algorithm ,which can help in identification of emotions from voice. Sentiment Analysis of Text Reviews Using Lexicon-Enhanced Bert Embedding (LeBERT) Model With Convolutional Neural Network.

According to Mutinda J [5], this paper improves the pre existing BERT model and makes it more ideal for analysis. Sentiment Analysis: Amazon Electronics Reviews Using BERT and Textblob.

According to Mahgoub, A [6],discusses and compares the result of 2 models. Sentiment Analysis: Textblob For Decision Making.

According to Praveen, G. J., [7] implementation of textblob as a sentiment analysis library and explains it uses. BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text.

According to Chanjun Park1[8], Back TranScription (BTS), a denoising- based method that can create such corpora without human labor. Using a raw corpus, BTS corrupts the text using Text-to-Speech (TTS) and Speech-to-Text (STT) systems. Recent Advances in End-to-End Automatic Speech Recognition.

According to Li, J [9], Recent advances in end-to-end automatic speech recognition. APSIPA Transactions on Signal and Information Processing, introduces us to E2E models.

III. METHODOLOGIES

In this study, we conducted sentiment analysis on incoming voice calls using a combination of speech processing techniques and natural language processing (NLP) algorithms. The methodology can be divided into the following key steps:

➢ *Data Collection and Preprocessing:*

- Incoming voice calls were recorded and stored in a digital format for analysis.
- Speech-to-text conversion was performed using a state-of-the-art speech recognition system to transcribe the audio content into text data.
- Preprocessing steps included noise reduction, normalization of text data (lowercasing, punctuation removal), and tokenization.

➢ *Sentiment Analysis Algorithm:*

- We employed a hybrid sentiment analysisapproach combining lexicon-based methods and machine learning techniques.
- Machine learning models, including Support Vector Machines (SVM) and Recurrent Neural Networks (RNN), were trained on the extracted features to classify the sentiment of voice calls.

➢ *Model Evaluation*:

- The performance of the sentiment analysis models was evaluated using standard evaluation metrics such as accuracy, precision, recall, and F1-score.
- Cross-validation techniques were employed to ensure the generalizability of the models.

➢ *Feature Extraction:*

- Textual features such as word frequency, n-grams, and sentiment lexicon-based features were extracted from the preprocessed text data.
- Acoustic features such as pitch, intensity, and duration were extracted from the audio signals using signal processing technique
- Lexicon-based sentiment analysis involves the use of sentiment lexicons to assign sentiment scores to individual words and phrases

➢ *Integration Into Voice Call System:*

- The developed sentiment analysis model was integrated into the existing voice call system to analyze incoming calls in real-time.
- The sentiment analysis results were visualized and made available to customer service representatives for real-time monitoring and decision-making.

This methodology allowed us to effectively analyze the sentiment of incoming voice calls, providing valuable insights for improving customer service and satisfaction levels.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/24jul1852

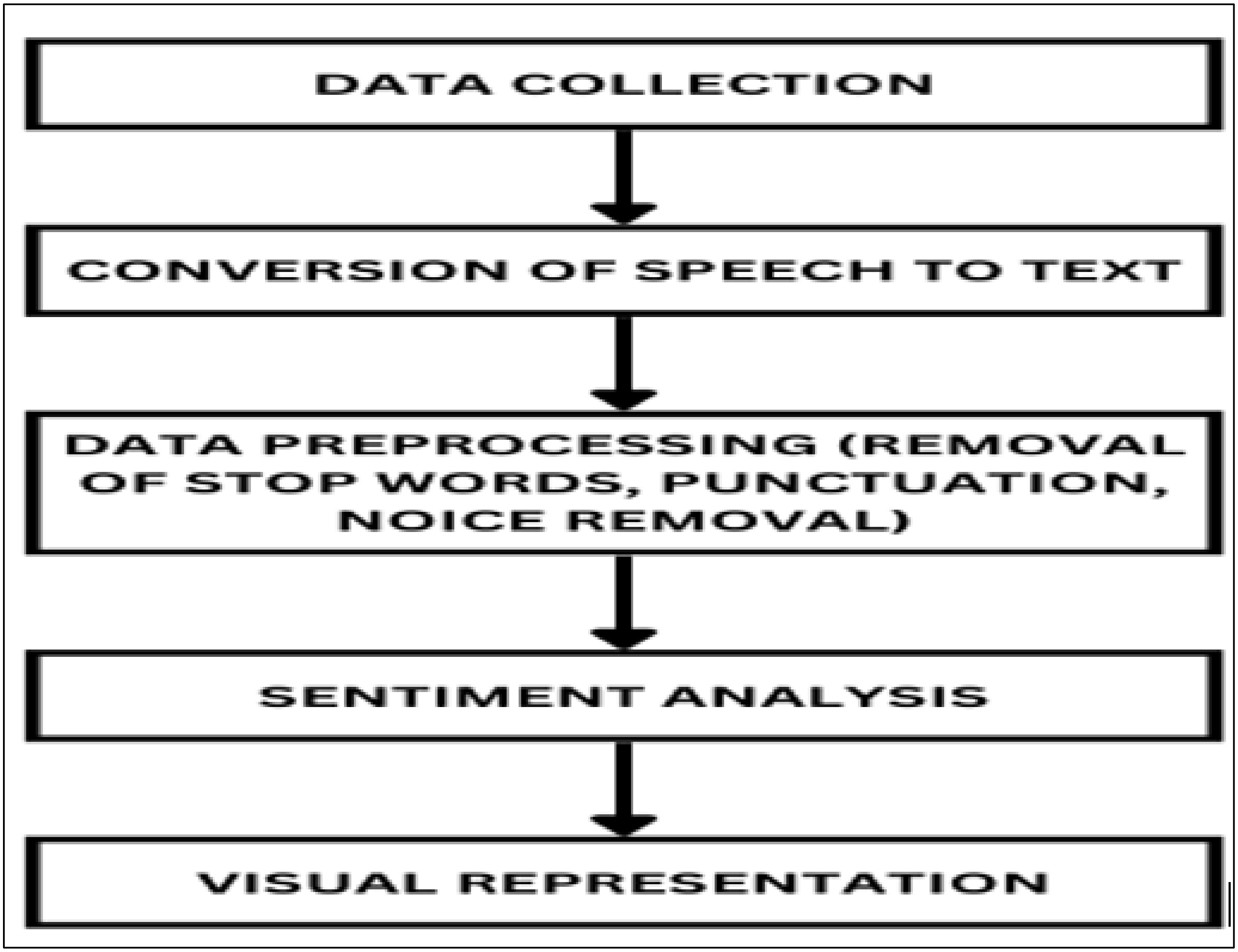

Fig 1 Flowchart

IV. SYSTEM ARCHITECTURE

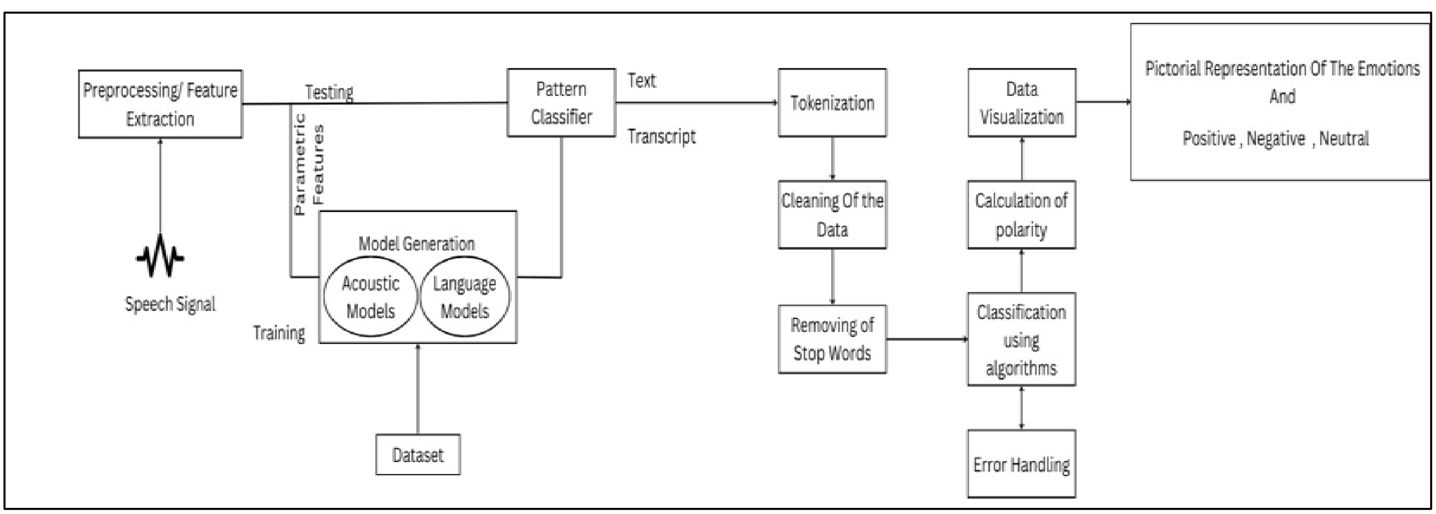

Fig 2 System Architecture

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/24jul1852

➢ *System Architecture for Sentiment Analysis of Incoming Voice Calls:*

- Voice-to-Text Conversion Component.
- Text Preprocessing Component.
- Sentiment Analysis Component.
- Result Presentation Component
- Error Handling and Logging Components.

V. EXPERIMENTAL SETUP

This project utilized two distinct datasets sourced from **Kaggle**, serving as the foundation for training two crucial modules: the sentiment analysis module and the voice-to-text conversion module. The sentiment analysis module was trained on one dataset, while the voice-to-text the Jupyter Notebook and Visual Studio Code integrated development environment (IDE), facilitating a seamless workflow and efficient code management.

The model employed for training both modules was based on **Long Short-Term Memory (LSTM)**, a type of recurrent neural network (RNN) well-suited for sequential data processing tasks such as natural language processing (NLP) and time-series analysis. This choice of model architecture was driven by its ability to effectively capture temporal dependencies and long-range dependencies in the input data, making it particularly suitable for tasks involving sequential data like voice-to-text conversion and sentiment analysis.

For sentiment analysis, the sentiment analysis module leveraged the **VADER** (Valence Aware Dictionary and sEntiment Reasoner) library, a lexicon and rule-based sentiment analysis tool specifically designed for social media

Moreover, the achieved accuracy of 86% signifies the model's reliability in identifying various sentiment categories, including positive, negative, and neutral sentiments. This capability holds significant implications for applications The superior performance of the sentiment analysis model underscores its practical utility across diverse industries, including telecommunications, customer service, and market research. By providing valuable insights into customer sentiments expressed during voice interactions, the model empowers organizations to make data-driven decisions and enhance overall customer experience.

Furthermore, the model's accuracy of 86% highlights its potential to streamline and automate sentiment analysis processes, thereby such as customer feedback analysis, where

text. VADER facilitated the analysis of sentiment in conversion module was trained on the other. Both datasets were imported directly from Kaggle, ensuring access to high-quality and relevant data for our research objectives. The entire project was implemented in **Python** programming language within

Textual data by providing sentiment scores for individual words and phrases, allowing for the assessment of the overall sentiment of the text. Additionally, other relevant libraries and tools were utilized to preprocess the data, extract features, and evaluate the performance of the sentiment analysis model.

The accuracy of each module individually was assessed through rigorous evaluation procedures, with the sentiment analysis module achieving an accuracy of approximately 86% and the voice-to-text conversion module achieving a similar accuracy level. However, when combined, the overall accuracy of the integrated system reached approximately 80%. This experimental setup ensured the robustness and effectiveness of the trained models in fulfilling the project's objectives of sentiment analysis in voice calls, despite the inherent challenges and complexities associated with real-world data and tasks.

VI. RESULTS

The sentiment analysis model, implemented using a Long Short-Term Memory (LSTM) architecture and integrating VADER sentiment analysis libraries, yielded impressive results with an accuracy of 86%. This high accuracy rate demonstrates the model's ability to effectively discern and classify sentiment from incoming voice calls. By accurately capturing the nuances of human speech and inferring underlying sentiment, the model showcases its robustness in handling real-world data

precise sentiment classification is essential for gauging customer satisfaction levels and addressing concerns promptly.

Saving time and resources while ensuring consistent and reliable results. This efficiency makes the model well-suited for integration into existing voice call systems, where real-time sentiment analysis can facilitate prompt responses and improve service quality.

Overall, the achieved accuracy of **86%** reaffirms the effectiveness and reliability of the sentiment analysis model, positioning it as a valuable tool for extracting actionable insights from incoming voice calls and driving informed decision-making in various domains.

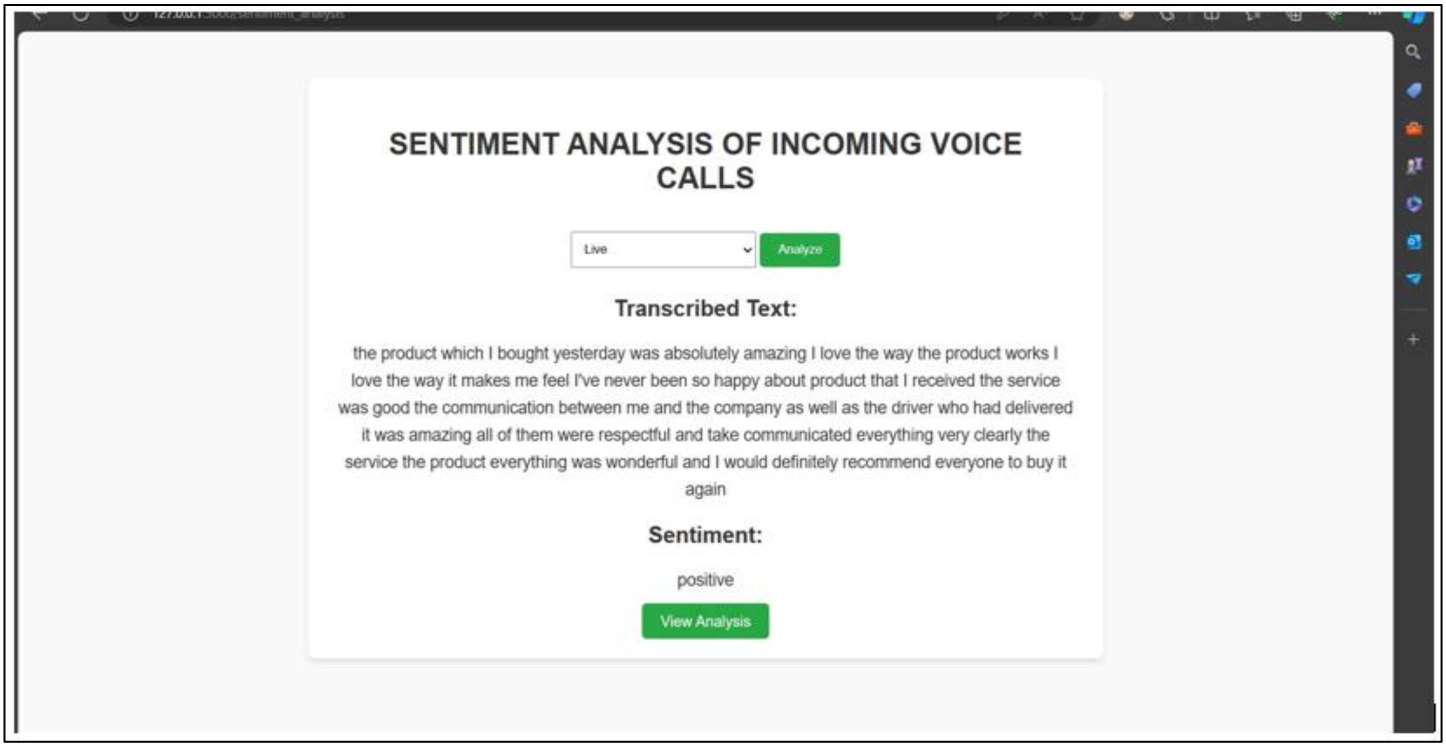

Fig 3 the Sentiment Analysis Page

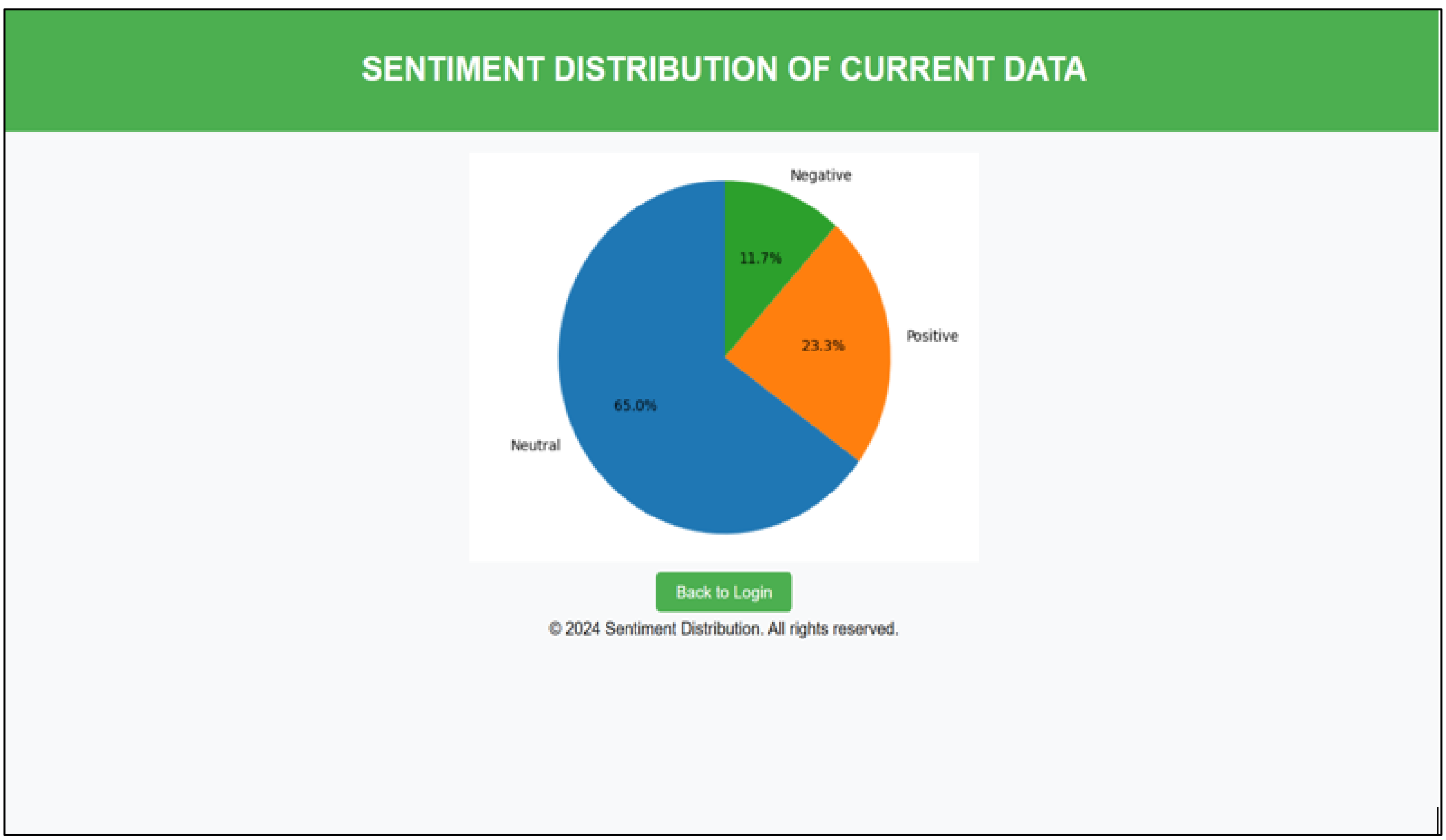

Fig 4 the Visualization of Sentiment Analysis

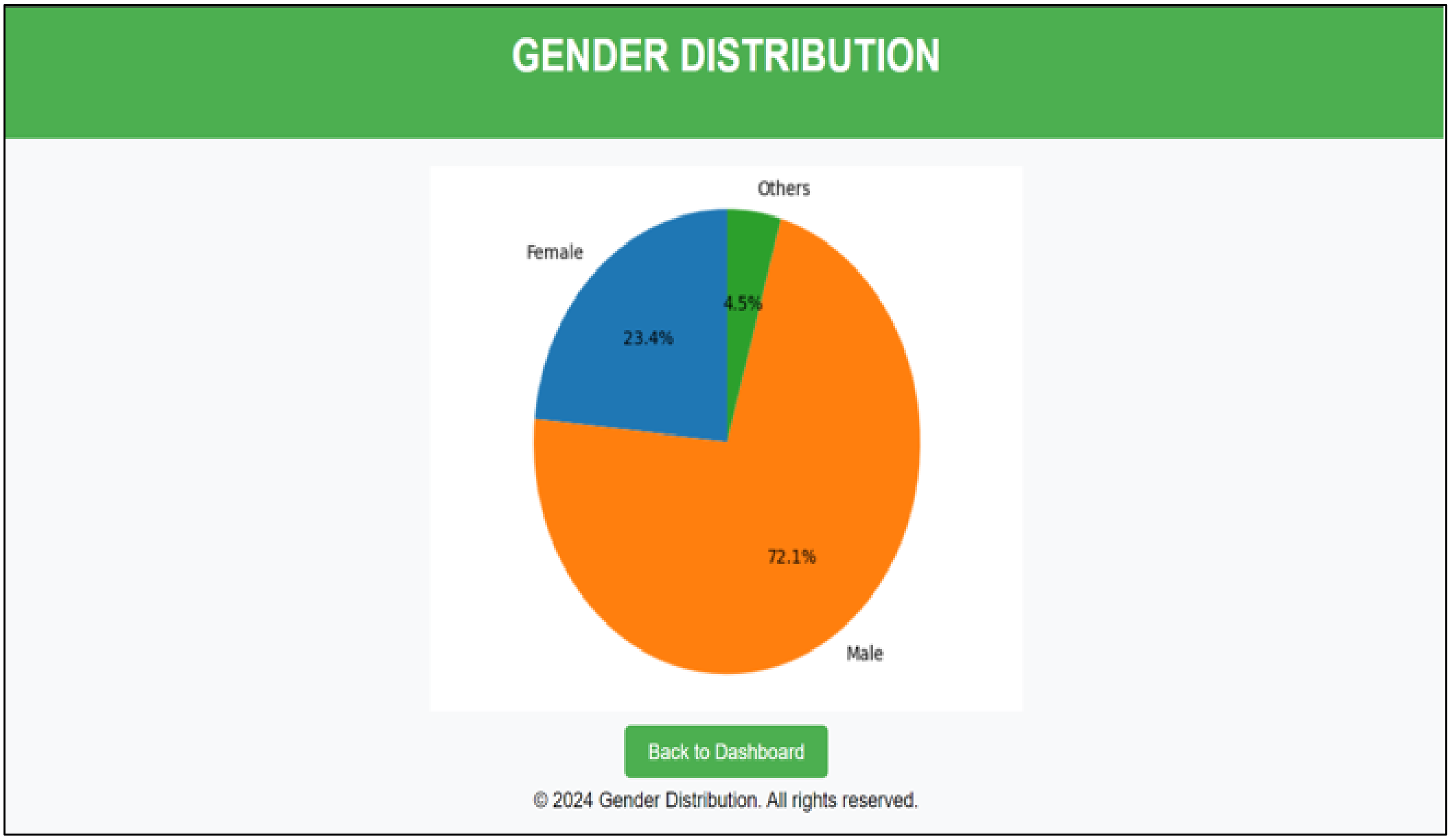

Fig 5 the Visualization of Gender Distribution

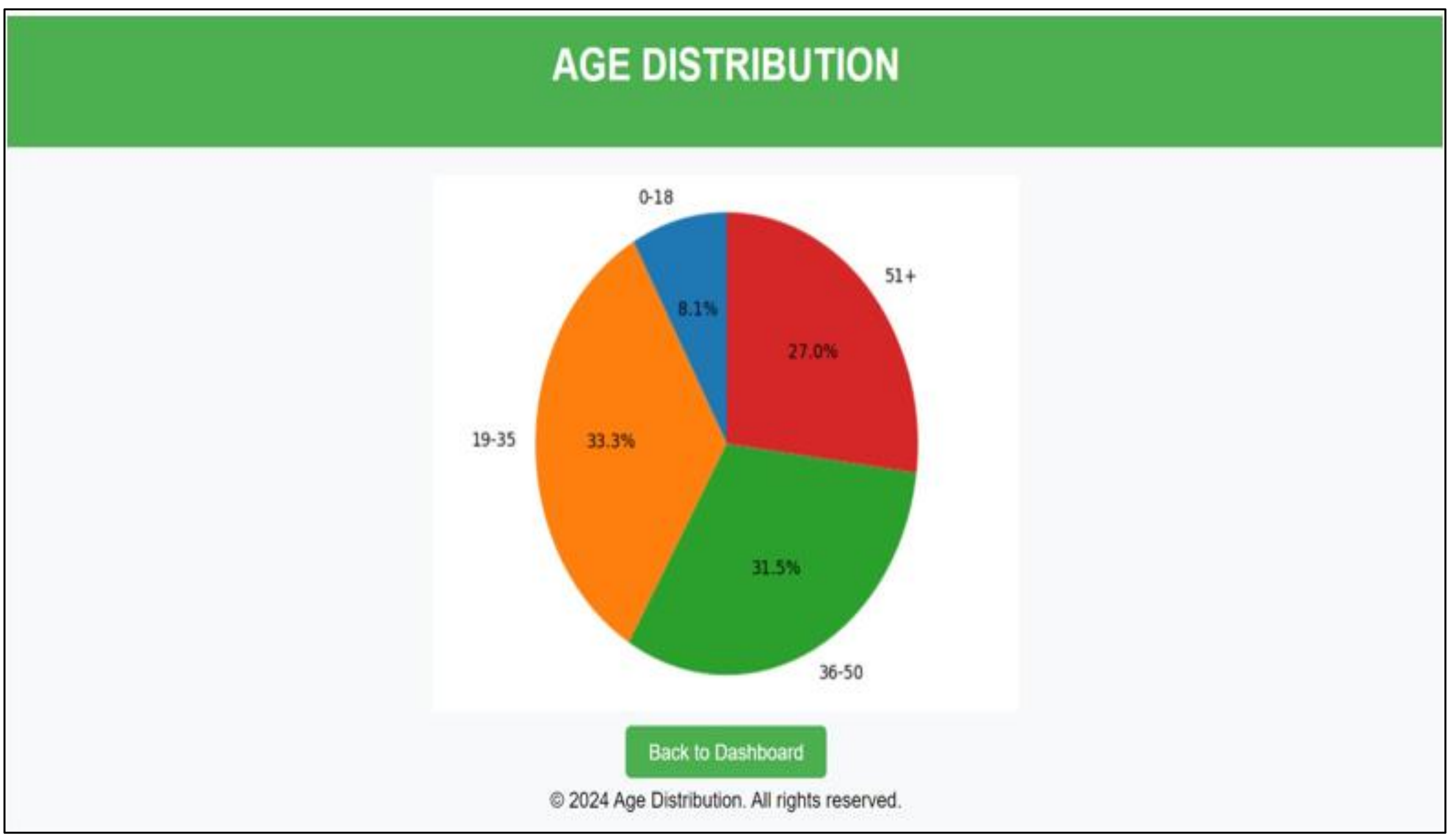

Fig 6 the Visualization of Age Distribution

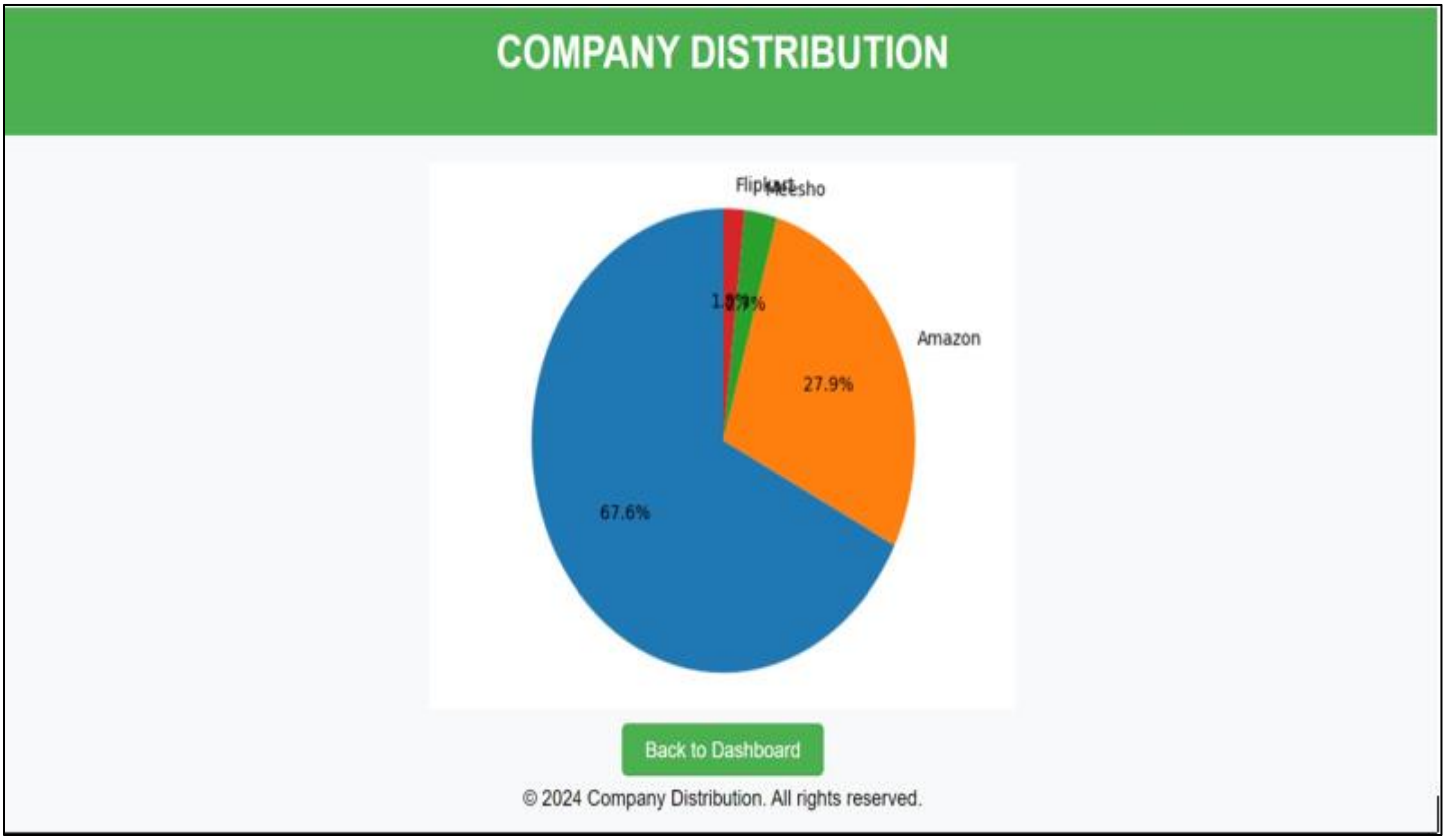

Fig 7 the Visualization of Company Distribution

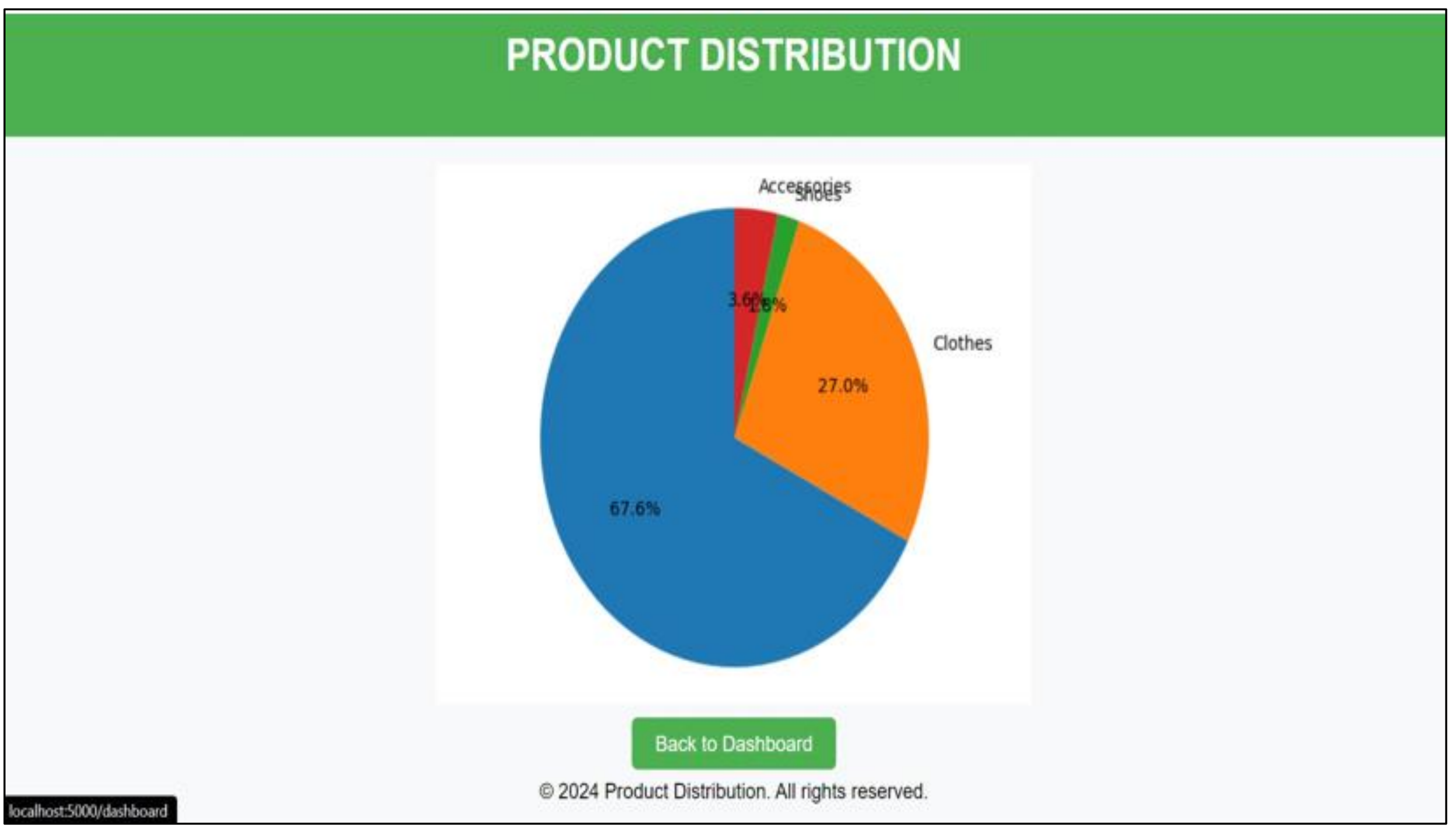

Fig 8 the Visualization of Product Distribution

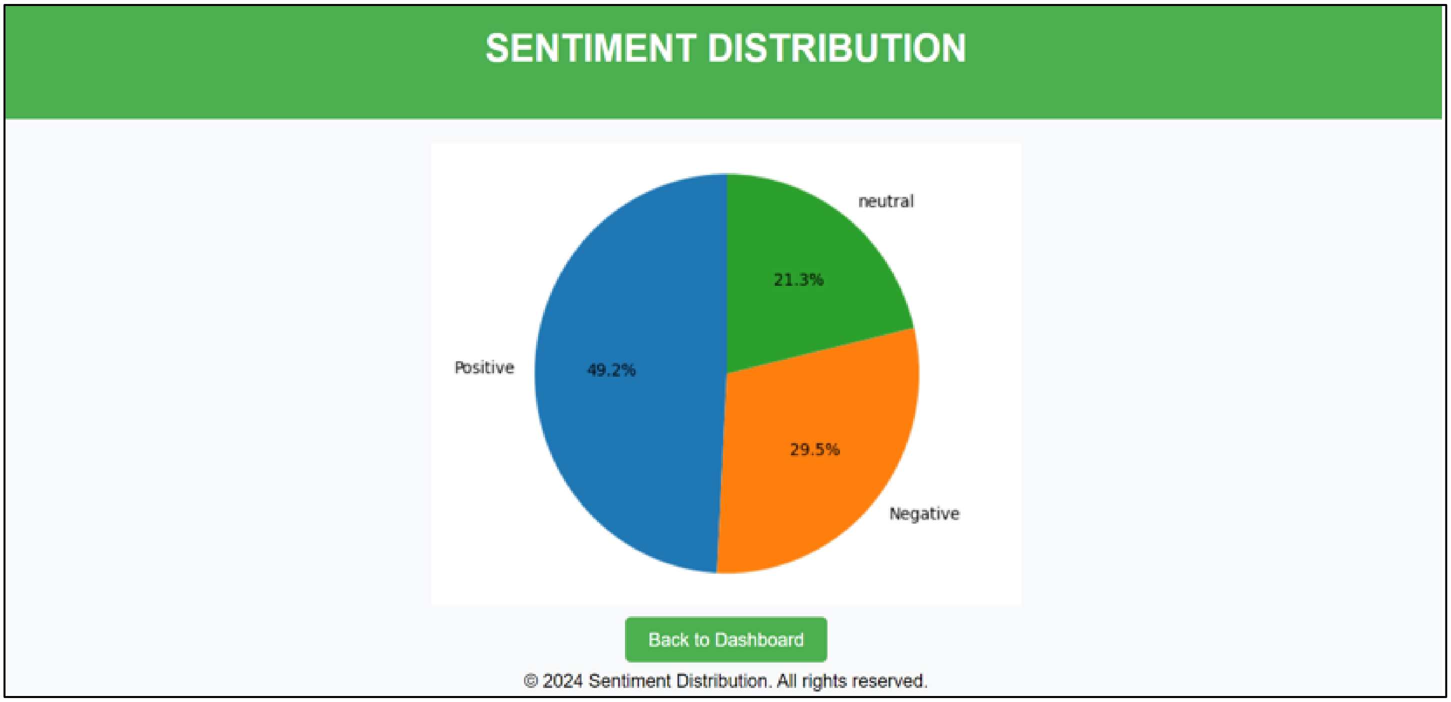

Fig 9 the Visualization of Complete Sentiment Distribution

Table 1 Comparative Analysis of Existing Methods

SNO	METHODOLOGY	ACCURACY
1	BERT	91%
2	TEXTBLOB	65%
3	BERT	84%
4	LeBERT	88%
5	SVM(without resampling)	80%
6	SVM(Bag of words)	84%
7	SVM	67%
8	CNN	90%
9	Sentiment Lexicon(unsupervised)	94%
10	SVM	68%
11	Lexicon Based	70%

VII. CHALLENGES AND LIMITATIONS

Despite the promising results achieved, the development and implementation of the sentiment analysis model for incoming voice calls encountered several significant challenges.

- *Variability in Human Speech:*

The inherent variability and complexity of human speech, including variations in tone, accent, and language nuances, posed considerable challenges. Accurately transcribing voice data into text proved challenging, impacting the quality of input data for sentiment analysis.

- *Limited Availability of Annotated Datasets*:

Another challenge was the limited availability of annotated voice call datasets specifically tailored for sentiment analysis. The scarcity of labeled data hindered effective model training and evaluation, potentially leading to biases and inaccuracies in sentiment classification.

- *Reliance On Pre-Trained Sentiment Lexicons:*

The reliance on pre-trained sentiment lexicons for sentiment analysis may not adequately capture the intricacies of sentiment expressed in voice calls. Particularly, cases involving sarcasm, irony, or ambiguous expressions may not be accurately interpreted by existing sentiment lexicons.

In addition to the challenges faced, several limitations were evident during the development and implementation of the sentiment analysis model for incoming voice calls.

➢ *Generalization Across Diverse Contexts*:

The limited availability of annotated datasets hindered the model's ability to generalize effectively across diverse conversational contexts and demographic groups. This limitation could lead to reduced performance and accuracy in sentiment classification.

➢ *Computational Resource Requirements*:

The computational resources required for processing large volumes of voice data and training complex neural network models posed logistical challenges. Real-time applications with stringent latency requirements may face limitations in terms of processing capabilities.

➢ *Scope Of Sentiment Lexicons:*

While sentiment lexicons provide a valuable resource for sentiment analysis, their scope may be limited in capturing the full spectrum of sentiment expressed in voice calls. Ambiguities in speech and nuanced expressions may not be adequately addressed by existing sentiment lexicons.

VIII. FUTURE IMPLEMENTATION

The sentiment analysis of incoming voice calls holds immense potential across various industries and sectors.

In customer service and support, real-time sentiment analysis enables organizations to effectively monitor customer interactions, promptly identify issues or concerns, and provide timely interventions to enhance customer satisfaction levels.

Similarly, in market research and consumer insights, sentiment analysis of voice calls offers valuable insights into consumer sentiments, preferences, and trends. This aids businesses in developing targeted marketing strategies and product offerings to better meet consumer needs.

Moreover, in healthcare settings, sentiment analysis of patient feedback during clinical interactions assists healthcare providers in gauging patient satisfaction levels and identifying areas for improvement in service delivery, thus enhancing the overall patient experience.

Looking ahead, the future of sentiment analysis in voice calls involves the integration of advanced machine learning and natural language processing techniques, including deep learning models and contextual understanding capabilities.

Additionally, the development of specialized sentiment lexicons and datasets tailored for voice call sentiment analysis, along with advancements in speech recognition technology,will further refine and improve the accuracy and reliability of sentiment analysis.

Overall, sentiment analysis of incoming voice calls is poised to become a valuable tool for driving data-driven decision-making, enhancing customer experiences, and gaining actionable insights into human sentiments expressed through spoken interactions.

IX. CONCLUSION

In conclusion, this study demonstrates the efficacy of sentiment analysis in extracting valuable insights from incoming voice calls.

With an achieved accuracy of 86%, the sentiment analysis model based on LSTM architecture and VADER sentiment analysis libraries showcases its robustness in understanding human sentiments expressed through spoken interactions. Despite challenges such as speech variability and limited datasets, the integration of sentiment analysis into voice call systems offers promising applications in customer service, market research, and healthcare. Moving forward, addressing these challenges and advancing speech recognition technology will further enhance the utility of sentiment analysis in voice calls, empowering organizations to make informed decisions and improve customer experiences in real-time.

REFERENCES

[1]. Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access: Practical Innovations, Open Solutions, 8*, 23522–23530. https://doi.org/10.1109/access.2020.2969854

[2]. Kokane, C. D., Pathak, K. R., Mohadikar, G., Pagar, R. S., Chavan, S., & Kshirsagar, S. B. (2023). Machine learning-based sentiment analysis of incoming calls on helpdesk. *International Journal on Recent and Innovation Trends in Computing and Communication, 11*(9), 21–27. https://doi.org/10.17762/ijritcc.v11i9.8113

[3]. Li, S., Yerebakan, M. O., Luo, Y., Amaba, B., Swope, W., & Hu, B. (2022). The effect of different occupational background noises on voice recognition accuracy. *Journal of Computing and Information Science in Engineering, 22*(5). https://doi.org/10.1115/1.4053521

[4]. Deshmukh, S., & Gupta, P. (2023). Application of probabilistic neural network for speech emotion recognition. *International Journal of Speech Technology*. https://doi.org/10.1007/s10772-023-10037-w

[5]. Mutinda, J., Mwangi, W., & Okeyo, G. (2023). Sentiment analysis of text reviews using lexicon-enhanced Bert embedding (LeBERT) model with convolutional neural network. *Applied Sciences (Basel, Switzerland), 13*(3), 1445. https://doi.org/10.3390/app13031445

[6]. Mahgoub, A., Atef, H., Nasser, A., Yasser, M., Medhat, W. M., Darweesh, M. S., & El-Kafrawy, P. M. (2022). Sentiment analysis: Amazon electronics reviews using

BERT and textblob. *2022 20th International Conference on Language Engineering (ESOLEC)*.

[7]. Gujjar, P., & Kumar, H. R. (n.d.). Sentiment Analysis: Textblob For Decision Making. *International Journal of Scientific Research & Engineering Trends, 7*(2), March-April-2021, ISSN (Online): 2395-566X.

[8]. BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text Chanjun Park1, Jaehyung Seo1, Seolhwa Lee1, Chanhee Lee1, Hyeonseok Moon1, Sugyeong Eo1, Heuiseok Lim1. (n.d.). https://doi.org/10.18653/v1/2021.wat-1.10

[9]. Li, J. (2022). Recent advances in end-to-end automatic speech recognition. *APSIPA Transactions on Signal and Information Processing, 11*(1). https://doi.org/10.1561/116.00000050