## A PROJECT REPORT

on

**SENTIMENT ANALYSIS OF INCOMING CALLS ON HELPDESK**

Submitted in partial fulfillment of the requirements for the award of the degree of

## BACHELOR OF TECHNOLOGY

in

### CSE (Artificial Intelligence & Machine Learning)

Submitted by

**P. Venkata Naga Sai Hasini (21UP1A6638)**

**K. Sai Varshitha (21UP1A6625)**

**K. Harini (21UP1A6624)**

**K. Aswitha (22UP5A6603)**

Under the Guidance of

# Dr.M. THEJOVATHI., M.Tech., Ph.D

Assistant Professor, CSE(AI&ML)



**DEPARTMENT OF CSE (ARTIFICIALINTELLIGENCE& MACHINELEARNING)**

### VIGNAN’S INSTITUTE OF MANAGEMENT AND TECHNOLOGY

### FOR WOMEN

#### (An Autonomous Institution) Approved by AICTE and Affiliated to JNTUH

Accredited to NBA, NAAC, A+ (CSE&ECE)

Kondapur (Village), Ghatkesar (Mandal), Medchal (Dist.), Telangana, Pincode-501301

[www.vmtw.edu.in](http://www.vmtw.edu.in/)

2024-2025



## DEPARTMENT OF CSE(AI&ML)

**CERTIFICATE**

This is to certify that project work entitled “**SENTIMENT ANALYSIS OF INCOMING CALLS ON HELPDESK”** submitted by **P. Venkata Naga Sai Hasini** **(21UP1A6638),** **K. Sai** **Varshitha (21UP1A6625), K. Harini** **(21UP1A6624), K. Aswitha (22UP5A6603)** in the partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **CSE(AI&ML)** **VIGNAN’S INSTITUTE OF MANAGEMENT AND TECHNOLOGY FOR WOMEN** is a record of bonafide work carried by them under my guidance and supervision. The results embodied in this project report have not been submitted to any other University or institute for the award of any degree.

**Signature of Signature of Signature of**

**Project Guide Project Coordinator HOD(AI&ML)**



## DEPARTMENT OF CSE(AI&ML)

**DECLARATION**

We here by declare that the work reported in the present project entitled “**SENTIMENT ANALYSIS OF INCOMING CALLS ON HELPDESK**” is a record of bonafied work duly completed by us in the Department of CSE (AI&ML) from Vignan’s Institute of Management and Technology for Women, affiliated to JNTU, Hyderabad. The reports are based on the project work done entirely by us and not copied from any other source. All such materials that have been obtained from other sources have been duly acknowledged.

The result embodied in this project report have not been submitted to any other University or Institute for the award of any degree to the best of our knowledge and belief.

**SI No. Roll Number Student Name Signature of the Student**

1 21UP1A6638 P. Venkata Naga Sai Hasini

2 21UP1A6625 K. Sai varshitha

3 21UP1A6624 K. Harini

4 22UP5A6603 K. Aswitha

## ACKOWLEDGEMENT

We would like to express our sincere gratitude and indebtedness to our project guide,

**Dr. M. Thejovathi., M.Tech., Ph.D,** Assistant Professor for her valuable suggestions and interest throughout the course of this project.

We would like to extend our gratitude to **Mr. S. Sandeep Babu., M.Tech**, Project Coordinator, Department of CSE- Artificial Intelligence and Machine Learning from Vignan's Institute of Management and Technology For Women, for his valuable suggestions and timely help while the project.

We are also thankful to **Dr. D. Shanthi., Ph.D**, Head of the Department of CSE- Artificial Intelligence and Machine Learning from Vignan's Institute of Management and Technology for Women, Hyderabad for providing excellent infrastructure and a nice atmosphere for completing this project successfully as a part of our Bachelor of Technology Degree (CSM).

We convey our heartfelt thanks to the lab staff for allowing us to use the required equipment whenever needed.

Finally, we would like to take this opportunity to thank our family for their support through the work. We sincerely acknowledge and thank all those who gave directly or indirectly their support in the completion of this work.

**P. Venkata Naga Sai Hasini (21UP1A6638)**

**K. Sai varshitha (21UP1A6625)**

**K. Harini (21UP1A6624)**

**K. Aswitha (22UP5A6603)**

|  |
| --- |
| **TABLE OF CONTENTS**    **CONTENTS PAGE.NO** |
|  |
| **CHAPTER 1: INTRODUCTION** 1   * 1. Problem Statement 2   2. Objective 2   3. Motivation 3   4. Existing System 4   5. Proposed System 7   6. Scope and Purpose 9 |
| **CHAPTER 2: LITERATURE REVIEW** 10 |
| **CHAPTER 3: SYSTEM REQUIREMENTS AND SPECIFICATIONS**   * 1. Software Requirements 16   2. Hardware Requirements 16 |
| **CHAPTER 4: SYSTEM DESIGN**   * 1. System Architecture or Block Diagram 17   2. Data Flow Diagram 18   3. UML Diagrams 20      1. Use Case Diagram 21      2. Class Diagram 22      3. Sequence Diagram 25      4. Activity Diagram 26      5. Deployment Diagram 28 |
| **CHAPTER 5: CONCLUSION** 31   * 1. Future Enhancement 31 |
| **CHAPTER 6:**  6.1 References 33 |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No** | **Figure Name** | **Page No** |
| 1.4.1 | Google Cloud Natural Language API | 4 |
| 1.4.2 | Affectiva | 5 |
| 1.4.3 | IBM Watson | 6 |
| 4.1 | Architecture | 17 |
| 4.2 | Dataflow Diagram | rrr1 19 |
| 4.3.1 | Use case Diagram | 212 21 |
| 4.3.2 | Class Diagram | 232 23 |
| 4.3.3 | Sequence Diagram | 252 25 |
| 4.3.4 | Activity Diagram | 277 27  27 |
| 4.3.5 | Deployment Diagram | 292 29 |

**LIST OF ALGORITHMS**

|  |  |  |
| --- | --- | --- |
| **Algorithm No.** | **Algorithm Name** | **Page No** |
| 1.2 | Naïve Bayes Classifier | 9-10 |

## 

## ABSTRACT

Sentiment analysis of incoming calls on helpdesk systems aims to enhance customer experience by automating the identification of customer emotions expressed during interactions. This project leverages Natural Language Processing (NLP) techniques and machine learning algorithms to analyze call transcriptions and classify sentiments as positive, negative, or neutral. By extracting valuable insights from customer conversations, the system enables helpdesk teams to respond proactively, improving issue resolution and customer satisfaction. The project demonstrates the potential of AI-driven sentiment analysis to streamline customer service operations, providing real-time feedback and actionable intelligence to helpdesk operators. Additionally, this approach reduces manual effort and fosters data-driven decision-making to enhance overall operational efficiency.

# CHAPTER-1

**1.INTRODUCTION**

In today's customer-centric business environment, helpdesk services play pivotal role in maintaining client satisfaction. A critical factor in assessing the quality of helpdesk interactions is understanding customer emotions during their communication. Sentiment analysis has emerged as a powerful tool for automatically detecting the emotional tone in textual and spoken data. Applied to helpdesk operations, sentiment analysis enables businesses to evaluate customer sentiments, identify dissatisfaction early, and improve overall service performance.

In today's competitive business environment, customer satisfaction plays a pivotal role in determining an organization's success. Helpdesks serve as a primary touchpoint for customers seeking assistance, and the quality of interactions directly impacts their perception of the company. However, analyzing and understanding the emotions conveyed during helpdesk calls is a challenging task, especially when relying solely on manual efforts. Sentiment analysis, a branch of Natural Language Processing (NLP), offers a powerful solution by enabling the automated identification and classification of customer emotions—whether positive, negative, or neutral. This project aims to leverage sentiment analysis techniques to analyze incoming calls on helpdesk systems. By transcribing spoken conversations into text and applying machine learning algorithms, the project seeks to uncover valuable insights into customer sentiments.

These insights can help organizations improve service quality, identify recurring issues, and enhance overall customer experience. The project’s focus is not only on detecting sentiment but also on presenting the results in actionable formats for decision-makers. Through the integration of cutting-edge technologies, this system aspires to transform traditional customer service practices, making them more data-driven and efficient. Helpdesk operators face difficulty understanding customer emotions based on call transcripts, leading to delayed or inadequate responses.

Sentiment analysis leverages various AI techniques, but the most common are natural language processing (NLP) models. These models analyze text to identify emotions and opinions, often using techniques like machine learning and deep learning. Today, companies have large volumes of text data like emails, customer support chat transcripts, social media comments, and reviews. As customers express their reviews and thoughts about the brand more openly than ever before, sentiment analysis has become a powerful tool to monitor and understand online conversations. Further, you can use this analysis to tailor your products and services to meet your customer’s needs and make your brand successful.

**1.1 PROBLEM STATEMENT**

Existing systems fail to analyse large volumes of calls in real-time.

Speech to text transcriptions of call recordings followed by text analysis of NLP for detecting emotions.

**1.2 OBJECTIVE**

The objective of sentiment analysis on incoming helpdesk calls is to enhance customer experience by identifying and categorizing caller emotions (positive, negative, or neutral).

Insights gained can improve service quality, customer satisfaction, and operational strategies.

The objective of sentiment analysis of incoming calls and help desk is to understand customer emotions and intent, which can help improve customer experience and brand reputation:

* **Personalize interactions**

Agents can use sentiment analysis to tailor their interactions to the customer’s emotional state and preferences. For example, if a customer is frustrated about slow internet speeds, an agent can offer a discount on an upgraded internet package or provide troubleshooting tips.

* **Improve customer experience**

Sentiment analysis can help identify issues and pain points, and address them proactively.

* **Enhance customer engagement**

Sentiment analysis can help identify opportunities for positive interactions with customers.

* **Monitor brand reputation**

Sentiment analysis can help track what customers are saying about a brand, product, or service.

* **Identify root causes**

Sentiment analysis can help identify the underlying causes of customer dissatisfaction.

* **Provide agent and call center performance metrics**

Sentiment analysis can help identify top-performing agents and implement coaching programs for low-performing agents.

* **Proactive issue resolution**

Since call center sentiment analysis lets you identify the negative sentiments in real time, you can intervene early in the customer journey before customers escalate .it saves customers and makes them happy, building customer loyalty.

Sentiment analysis uses natural language processing and machine learning to identify customer tone, intent, and sentiment. It can be used to analyse customer interactions across various channels, including voice, chat, online reviews, and feedback.

**1.3 MOTIVATION**

Call center sentiment analysis can benefit businesses of all sizes by improving customer satisfaction, identifying areas for improvement, and enhancing customer interactions. It offers valuable insights that can be applied to enhance the customer experience, regardless of the business's scale.Analyzing the sentiment of incoming calls at a helpdesk is essential for improving customer service and operational effectiveness. This process helps detect the emotional tone in customer conversations, allowing the helpdesk to respond to issues with greater empathy and efficiency. By examining sentiments, organizations can prioritize urgent matters, enhance response times, and customize solutions to fit customer requirements. Furthermore, insights derived from sentiment analysis enable companies to assess customer satisfaction and recognize trends or recurring problems. This data-driven methodology promotes ongoing enhancements in service performance. Additionally, comprehending the emotions of callers assists in training helpdesk personnel, ensuring they are well-prepared to manage various scenarios. Adopting sentiment analysis not only enhances customer trust and loyalty but also offers critical insights for strategic planning, ultimately providing a competitive advantage in the industry.

This technology enables helpdesks to prioritize calls based on urgency and emotional intensity, ensuring that highly dissatisfied or frustrated customers receive immediate attention. It also assists in monitoring agent performance by evaluating how effectively they address customer concerns and de-escalate negative emotions. Furthermore, sentiment analysis helps in identifying recurring issues or patterns in customer complaints, allowing organizations to address root causes and improve service quality proactively.

From a strategic perspective, leveraging sentiment analysis fosters a data-driven approach to enhancing customer experience, reducing churn, and building long-term loyalty. It empowers helpdesk managers to deploy resources efficiently and train agents on emotional intelligence and conflict resolution. Ultimately, sentiment analysis transforms a reactive support system into a proactive, empathetic, and customer-focused ecosystem.

**1.4 EXISTING SYSTEM**

1. **Google Cloud Natural Language API** - It detects the sentiment of a block of text (positive, negative, or neutral) and provides a sentiment score and magnitude. It has a limit on the size of each text document you can analyse.

**Sentiment Analysis**:Detects the overall sentiment in a block of text.

**Score**: Ranges from -1.0 (negative sentiment) to +1.0 (positive sentiment).

**Magnitude**: Measures the strength of the sentiment, regardless of positivity or negativity. Higher values indicate stronger emotions.

**Entity Sentiment Analysis**: Identifies specific entities (e.g., products, people, locations) in the text and analyses the sentiment associated with each entity.

**Syntax Analysis**: Analyses the structure of sentences, identifying parts of speech, sentence tokens, and word dependencies.

**Entity Recognition**: Extracts entities such as names, dates, numbers, and more, along with associated metadata.

**Multi-Language Support**: Supports sentiment analysis in multiple languages.



Fig 1.4.1: Google Cloud Natural Language API

1. **Affectiva -**The Emotion Speech API analyses a pre-recorded audio segment, such as an MP3 file, to identify emotion events and gender.

**Facial Expression Analysis**: Uses computer vision to analyse subtle facial movements and micro-expressions to determine emotional states such as joy, sadness, anger, or surprise. Captures data in real-time via video feeds or pre-recorded media.

**Multimodal Sentiment Analysis**: Combines facial expressions, voice tone, and other physiological signals (like skin conductance) for a holistic understanding of emotions. Supports context-aware sentiment analysis.

**Technology Stack**: Employs supervised learning techniques with large datasets of annotated emotional expressions to train models. Uses neural networks to enhance the accuracy of emotion

Fig 1.4.2: Affectiva

1. **IBM Watson Natural Language Understanding -** uses deep learning to extract meaning and metadata from unstructured text data.

**Targeted Sentiment**: Analyses sentiment toward specific entities, keywords, or concepts within the text.

**Sentence-Level Analysis**: Determines sentiment on a per-sentence basis, offering fine-grained understanding.

**Emotion Detection**: Identifies emotions such as joy, anger, sadness, and fear, often complementing sentiment insights.

Supports a wide range of languages for sentiment detection, including English, Spanish, French, German, and more.

**Custom Models**: Allows customization for domain-specific sentiment needs by training Watson NLU on your unique data.

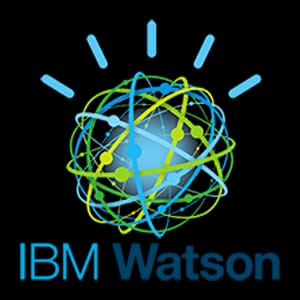


Fig 1.4.3: IBM Watson

**DRAWBACK OF EXISTING SYSTEMS**

1. **Sensitivity to Background Noise:** Current systems often have difficulty in detecting emotions during incoming calls because of background noise. This can hide important sound features that are essential for understanding customer feelings, which can result in misclassifying how customers truly feel.
2. **Inaccurate Transcriptions:** Many systems use automatic speech recognition (ASR) to turn spoken words into text. When the transcription is inaccurate, it can lead to incorrect labels for customer emotions.

**3. Limited Sentiment Scope:** concentrate mainly on detecting negative emotions, often overlooking positive and neutral sentiments.

**1.5 PROPOSED SYSTEM**

### ****1.User-Friendly Visualization and Reporting Dashboard****

### ****Sentiment Trends Over Time****: Graphs and charts that track changes in sentiment across different time frames, helping identify patterns in customer satisfaction.

### ****Real-Time Feedback Monitoring****: Displays live sentiment scores during ongoing calls, providing agents with immediate insights to adjust their approach.

### ****Interaction Highlights****: Summarizes key moments in customer interactions that significantly impacted sentiment, such as escalations or resolutions.

### ****Sentiment Scoring****: Assigns each call a sentiment score ranging from -1.0 (highly negative) to 1.0 (highly positive), with 0 representing neutrality. These scores are color-coded for quick identification of areas needing improvement.

### ****Filter and Drill-Down****: Provides filters to segment data by customer demographics, agent performance, or specific product/service categories. This enables targeted analysis and strategy development.

### ****Custom Reports****: Allows users to generate and export reports tailored to business needs, making it easier to share insights across teams.

### ****2. Emotional Intensity Detection****

### ****Emotion Classification****: Identifies the type of sentiment (e.g., joy, anger, sadness, fear) expressed during interactions.

### ****Intensity Scoring****: Assigns a scale (e.g., 0–100%) to each emotion, distinguishing between mild and extreme expressions. For example:

### ****Frustration Levels****: Differentiate between slight dissatisfaction and extreme anger, helping prioritize escalation cases.

### ****Positive Feedback Levels****: Distinguish between moderate satisfaction and enthusiastic advocacy, enabling tailored responses to happy customers.

### ****3.Sentiment Correction Based on Customer Feedback****

### ****Customer Surveys****: At the end of each call, request feedback through a brief survey asking customers how they felt about the interaction (e.g., satisfied, neutral, dissatisfied).

### ****Feedback Comparison****: Compare survey results with the sentiment scores generated by the AI model to identify discrepancies.

### ****Feedback Loop Integration****: Integrate these adjustments into the system periodically to keep the model updated and aligned with real-world customer expectations.

### ****4.Enhanced Noise Filtering Algorithms****

### ****Spectral Subtraction****: Reduces background noise by estimating the noise spectrum and subtracting it from the audio signal. This helps isolate human speech for clearer analysis.

### ****Deep Learning-Based Noise Reduction****: Leverage advanced neural networks to distinguish between speech and background noise, providing superior noise cancellation even in challenging environments.

### ****Dynamic Noise Profiling****: Adaptively adjusts to different noise conditions (e.g., office chatter, traffic noise) for consistent performance.

### ****Impact on Sentiment Analysis****: A cleaner audio signal improves the accuracy of sentiment detection, especially for subtle emotional cues like tone and pitch variations.

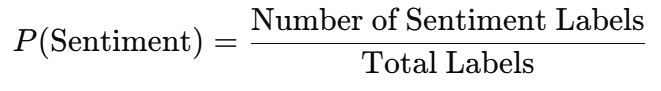
### ALGORITHMS USED:

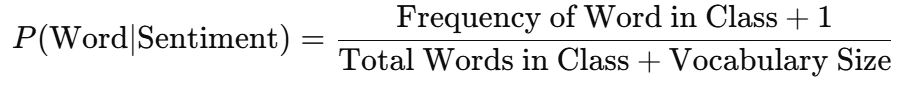
### Naive Bayes Algorithm:

### Assume Features Independence:

* Treat each word in the text as an independent feature.

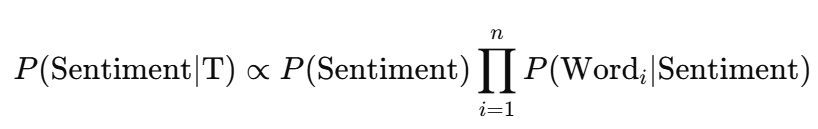
**2. Calculate Probabilities:**

* Compute the prior probability of each sentiment class:
* Compute the likelihood of each word given the sentiment class:



* Use Laplace smoothing to handle zero probabilities

**3.** **Classify Sentiment:**

* For a given call transcript 
* Assign the sentiment label with the highest probability.

### 1.6 SCOPE AND PURPOSE

### Sentiment analysis on phone calls is a powerful tool that uncovers the opinions and emotions of customers during phone calls. Using machine learning, natural language processing, and computational linguistics, sentiment analysis creates a score for each call—positive, neutral, or negative. Sentiment analysis systems help companies better understand their customers, deliver stronger customer experiences and improve their brand reputation. Discover the power of integrating a data lake house strategy into your data architecture, including enhancements to scale AI and cost optimization opportunities. Sentiment analysis, also referred to as opinion mining, is an approach to natural language processing (NLP) that identifies the emotional tone behind a body of text. This is a popular way for organizations to determine and categorize opinions about a product, service or idea. Impact Analysis is a technique designed to unearth the unexpected negative effects of a change on an organization. It provides a structured approach for looking at a proposed change, so that you can identify as many of the negative impacts or consequences of the change as possible.

### Voice call sentiment analysis is transforming how contact centers operate by providing deeper insights into the emotions and interactions of customers. Sentiment analysis can give management information that they can use to boost revenue, enhance customer satisfaction.

### 

### 

### 

### CHAPTER-2

**LITERATURE SURVEY**

### [1] A Deep Learning System for Sentiment Analysis of Service Calls

### [A Deep Learning System for Sentiment Analysis of Service calls](https://aclanthology.org/2020.ecnlp-1.4.pdf)

### Key Features and description:

* **Real-World Data Application**: The study focuses on sentiment analysis within real-world multi-party conversations, specifically service calls, highlighting the practical application of the proposed system.
* **Multimodal Feature Extraction**: The system extracts both acoustic and linguistic features from the service calls, enabling a comprehensive analysis of sentiment from multiple modalities.
* **Novel Aggregated Method**: A new framework for voice sentiment recognition is introduced, which aggregates the extracted features to enhance the accuracy of sentiment detection.
* **Interaction Pattern Analysis**: The research investigates individual sentiment patterns of each party during communication, as well as the interaction sentiment patterns between all parties involved, providing insights into the dynamics of sentiment in conversations.

### Methodology

### The study begins with data collection, gathering call transcripts and preprocessing them through cleaning, tokenization, and handling domain-specific jargon. Sentiment labels (positive, neutral, negative) are assigned via manual or semi-supervised annotation. Word embeddings like Word2Vec, GloVe, or BERT are used for feature extraction. A deep learning model, potentially based on CNNs, LSTMs, or transformers, is trained to analyze sentiment effectively. Model performance is evaluated using metrics like accuracy, F1-score, and AUC, with strategies to address challenges such as class imbalance and noisy data. Finally, the system is integrated for real-time sentiment prediction during live service calls, incorporating feedback for continuous improvement.

**Limitations**:

* **Data Quality Challenges**: The variability and noise inherent in real-world service call data can affect the performance of the sentiment analysis system.
* **Dependency on Accurate Transcriptions**: The effectiveness of linguistic feature extraction relies heavily on the accuracy of transcriptions, which can be compromised by factors such as background noise or strong accents.
* **Generalization to Other Domains**: The system is tailored to service call data, and its applicability to other types of conversations or domains may require further adaptation and validation.

**Publication Details:**

* **Conference**: Proceedings of the 3rd Workshop on e-Commerce and NLP (ECNLP 3)
* **Date**: July 10, 2020
* **Location**: Online (originally scheduled for Seattle, WA, USA)
* **Publisher**: Association for Computational Linguistics
* **Pages**: 24–34
* **DOI**: 10.18653/v1/2020.ecnlp-1.4

**[2] How Help Desk Agents Help Clients**

[How Help Desk Agents Help Clients](https://ieeexplore.ieee.org/document/4464071)

**Description and Key Features:** This study examines telephone helpdesk interactions, utilizing conversational analysis to understand how agents assist clients. The research focuses on the following aspects:

* **Conversational Analysis Framework:** Applying conversational analysis to dissect the structure and dynamics of helpdesk calls, identifying patterns in communication between agents and clients.
* **Problem-Solving Strategies:** Investigating the methods agents employ to diagnose and resolve client issues, including questioning techniques, information gathering, and providing instructions.
* **Client-Agent Interaction Dynamics:** Exploring how agents manage interactions to maintain clarity, ensure understanding, and effectively guide clients through problem resolution.

**Methodology:**

Steehouder employs a qualitative research approach, analysing recorded help desk interactions to identify effective communication strategies. He uses discourse analysis techniques to examine the structure and content of these interactions, focusing on how agents manage client inquiries and provide assistance. The study involves coding and categorizing different communication behaviours, such as questioning techniques, information provision, and empathy expressions. By analysing these elements, Steehouder aims to uncover patterns that contribute to successful client-agent interactions. The findings are intended to inform training programs and improve help desk services by highlighting the communication practices that lead to effective client support.

**Limitations:**

* **Scope of Data:** The study focuses on telephone interactions, which may not encompass the full range of communication channels used in modern helpdesk environments, such as email or chat.
* **Cultural and Contextual Factors:** The findings may be influenced by specific cultural or organizational contexts, potentially limiting the generalizability of the results to other settings.
* **Temporal Relevance:** Given the rapid evolution of communication technologies and practices, the relevance of the findings to current helpdesk operations may be affected over time.

**Publication Details:**

* **Conference:** IEEE International Professional Communication Conference (IPCC)
* **Date:** 2007
* **Publisher:** IEEE
* **DOI:** 10.1109/IPCC.2007.4464071

**[3] Machine Learning-Based Sentiment Analysis of Incoming Calls on Helpdesk**

[Machine Learning-Based Sentiment Analysis of Incoming Calls on Helpdesk](https://www.researchgate.net/publication/376359963_Machine_Learning-Based_Sentiment_Analysis_of_Incoming_Calls_on_Helpdesk)

**Description and Key Features:**

This research proposes an adaptive methodology for analysing incoming calls using sentiment analysis, aiming to identify spam or unwanted calls based on the conversation content. The key features of the study include:

* **Speech-to-Text Conversion:** The system processes user conversations in audio format, converting speech to text for further analysis.
* **Keyword Identification:** After pre-processing the transcribed text, the system detects keywords that are indicative of the call's nature.
* **Word Embedding with Word2Vec:** The Word2Vec technique is employed to represent words in a vector space, generating 150-200dimensional word vectors for effective analysis.
* **Sense Mapping with Lexical Resources:** Freely available lexical resources such as WordNet, SemCor, and OMSTI are utilized for sense mapping and keyword identification, enhancing the system's understanding of context.
* **Sentiment Analysis for Decision Making:** Based on the sentiment analysis of input calls, the system determines whether to accept or reject calls, effectively identifying spam numbers and sharing this information with other users.

**Methodology:**

The proposed system processes incoming call audio by converting it into text using speech-to-text technology. The textual data undergoes preprocessing steps, including noise reduction, tokenization, and normalization, to prepare it for analysis. Key phrases and terms are then extracted using the Word2Vec embedding technique, which transforms words into vector representations within a 150-200dimensional space. These vectors facilitate the identification of semantic relationships and the disambiguation of word senses through resources like WordNet. Subsequently, a supervised machine learning model analyses the sentiment of the conversation to determine the likelihood of the call being spam. The system's adaptive nature allows it to update its detection criteria based on new data, thereby improving its performance over time. This methodology has demonstrated superior results in identifying and managing unsolicited calls, contributing to more efficient helpdesk operations.

**Limitations:**

* **Dependence on Speech-to-Text Accuracy:** The effectiveness of the system relies heavily on the accuracy of speech-to-text conversion, which can be affected by factors such as background noise, accents, or speech clarity.
* **Scalability and Real-Time Processing:** The study does not extensively address the system's scalability or its ability to process calls in real-time, which are critical factors for practical deployment in helpdesk environments.
* **Evaluation Metrics and Comparative Analysis:** The paper lacks detailed evaluation metrics and comparative analysis with existing systems, making it challenging to assess the relative performance and effectiveness of the proposed methodology.

**Publication Details:**

* **Journal:** International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)
* **Volume:** 11
* **Issue:** 9
* **Pages:** 21–27
* **DOI:** 10.17762/ijritcc.v11i9.8113

**[4] Call Center Customer Sentiment Analysis Using ML and NLP**

[Call Center Customer Sentiment Analysis Using ML and NLP](https://www.researchgate.net/publication/377141570_Call_Center_Customer_Sentiment_Analysis_Using_ML_and_NLP)

**Description and Key Features**: This study explores the integration of sentiment analysis into call center operations to enhance the emotional intelligence of callbot interactions. By leveraging machine learning (ML) and natural language processing (NLP) techniques, the authors aim to improve customer satisfaction and operational efficiency. The key features of the paper include:

* **Sentiment Analysis Integration**: Incorporating sentiment analysis into call center workflows to assess and respond to customer emotions during interactions.
* **Machine Learning Models**: Utilizing ML algorithms to analyze customer sentiments, enabling the system to learn and adapt to various emotional cues.
* **Natural Language Processing Techniques**: Applying NLP methods to process and understand the textual data from customer interactions, facilitating accurate sentiment detection.
* **Enhanced Callbot Interactions**: Improving the responsiveness and empathy of automated callbots by equipping them with the ability to recognize and appropriately react to customer sentiments.

**Methodology:**

The study employs a comprehensive approach to sentiment analysis in call centers by integrating ML and NLP techniques. Initially, customer interactions are transcribed using speech-to-text technologies. The textual data then undergoes pre-processing steps, including tokenization, normalization, and the removal of stop words, to prepare it for analysis. Subsequently, feature extraction methods, such as word embeddings, are applied to capture semantic meanings within the text. The processed data is then fed into machine learning models—such as support vector machines (SVM), random forests, or deep learning architectures like recurrent neural networks (RNNs)—which have been trained to classify sentiments. The models are evaluated using metrics like accuracy, precision, recall, and F1-score to ensure their effectiveness. This methodology enables the real-time analysis of customer sentiments, providing actionable insights to improve service quality and customer satisfaction

**Limitations**:

* **Data Quality and Diversity**: The effectiveness of the sentiment analysis system depends on the quality and diversity of the training data. Limited or biased datasets may affect the model's performance across different customer demographics and scenarios.
* **Real-Time Processing Challenges**: Implementing real-time sentiment analysis requires significant computational resources and efficient algorithms to ensure prompt responses during live customer interactions.
* **Contextual Understanding Limitations**: NLP models may struggle with understanding nuanced or context-dependent sentiments, potentially leading to misinterpretation of customer emotions.

**Publication Details:**

**Conference**: 2022 IEEE International Conference on Smart Information Systems and Technologies (SIST)

**Date**: 2022

**Publisher**: IEEE

**DOI**: 10.1109/SIST54437.2022.9945865

**CHAPTER-3**

**SYSTEM REQURIMENTS AND SPECIFICATIONS**

**3.1 SOFTWARE REQUIREMENTS**

* Programming Language: Python 3.x
* Natural Language Processing (NLP) Libraries: NLTK, spaCy, genism
* Machine Learning Libraries: scikit-learn, TensorFlow, PyTorch
* Deep learning Libraries: Keras
* Visualization Libraries: Matplotlib, seaborn
* Audio Processing Libraries: PyAudio, Librosa
* Database Management System: MySQL, PostgreSQL
* Web Framework: Flask, Django
* Operating System: 64-bit Linux (e.g., Ubuntu) or Windows 10
* Text Editor/IDE: PyCharm, Visual Studio Code, Sublime Text

**3.2 HARDWARE REQUIREMENTS**

* Server: Dedicated server with at least 16 GB RAM, 1 TB storage, and a quad-core processor
* Storage: Separate storage device for storing audio files, transcripts, and analysis results
* Audio Equipment: Good quality audio interface and headphones for listening to calls
* CPU: Intel Core i7 or AMD Ryzen 7 processor
* Memory: At least 16 GB RAM
* Storage: At least 1 TB storage
* Network: High-speed internet connection for data transfer and communication

**Additional Requirements**

* Call Recording Software: Software for recording incoming calls, such as CallRail or Twilio
* Transcription Software: Software for transcribing recorded calls, such as (link unavailable) or Trint
* Sentiment Analysis Tools: Tools for sentiment analysis, such as Meaning Cloud or IBM Watson Natural Language Understanding

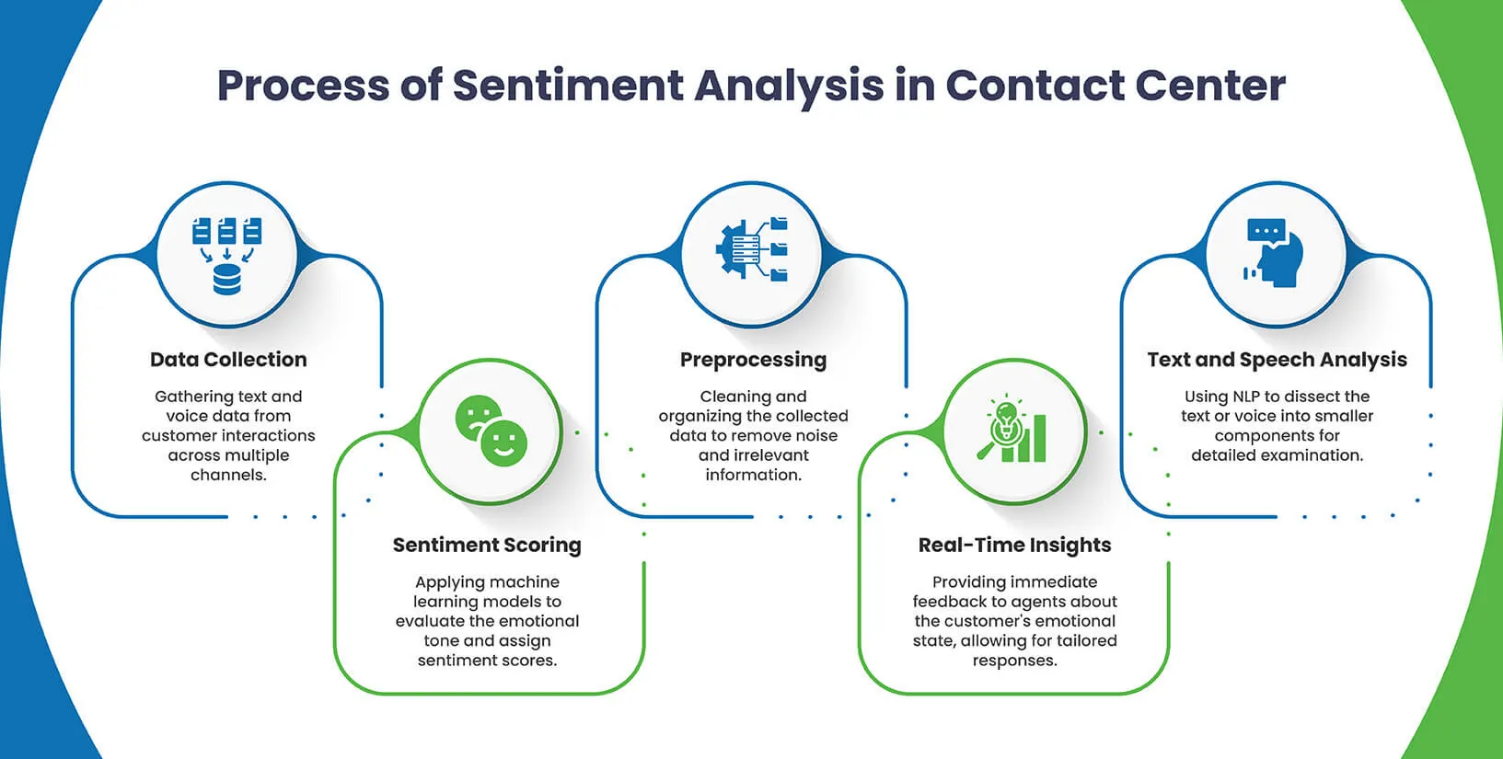
#### 

#### 

#### CHAPTER-4

#### SYSTEM DESIGN

**4.1 ARCHITECTURE**



#### Fig 4.1: Architecture

An architecture diagram is a visual representation of a system or application's components, layout, and structure. It helps people involved in the project understand how the system works, and can be thought of as a blueprint.

**Data Collection:** The process of gathering and measuring information to answer questions, test hypotheses, and evaluate outcomes. It's a key part of research, analysis, and decision-making in many fields, including business, social sciences, and healthcare.

**Pre-processing:** Data pre-processing is the process of preparing raw data to make it easier to analyse. This process can include:

* **Data cleaning**: Removing or fixing incorrect, corrupted, duplicate, or incomplete data.
* **Data reduction**: Reducing the size of the original data while preserving its integrity.
* **Data transformation**: Converting the data into a usable format for analysis. This can include normalizing, scaling, and weighting variables.
* **Data integration**: Combining data from multiple sources into a single dataset.

**Text And Speech Analysis:** Speech and text analytics is a set of features that uses natural language processing (NLP). It automatically analyses content immediately after the completion of an interaction. This analysis gives businesses insight into customer-agent conversations.

* Transcribe voice interactions
* Identify topics and categories
* Analyse sentiment
* Spot trends
* Highlight areas of improvement, concern, or recognition

**Sentiment Scoring:** A sentiment score is a rating of the emotional tone of a text, indicating whether it's mostly positive, negative, or neutral.

**4.2** **DATAFLOW DIAGRAM**

A data flow diagram (DFD) is a visual representation of how data moves through a system or process. DFDs are used to simplify complex systems and make it easier to understand how data is collected, processed, and used. They are a popular tool in software engineering, business management, and process improvement. In a DFD, the system is represented through various symbols. A process is depicted as a circle or rectangle, which indicates any action or transformation occurring on the data. Data sources and destinations, such as external entities or data stores, are represented by squares or rectangles. Arrows in a DFD show the movement of data between processes, external entities, or data stores. These arrows indicate data flows and specify the direction of data transfer.

The main objective of a DFD is to provide a clear and understandable representation of how data is transferred and processed. This visual representation aids system designers and analysts in identifying potential issues, inefficiencies, or gaps in the flow of data, allowing for improvements and optimization. Additionally, DFDs can serve as a helpful tool for communicating system structure with both technical and non-technical stakeholders. Creating a DFD requires a thorough understanding of the system's data sources, data flow, processes, and destinations. It also requires collaboration among developers, analysts, and users to ensure the diagram accurately reflects how the system functions and where improvements can be made. Data Flow Diagrams are essential tools for system analysis and design, offering insights into both high-level and detailed perspectives of data movement within a system.

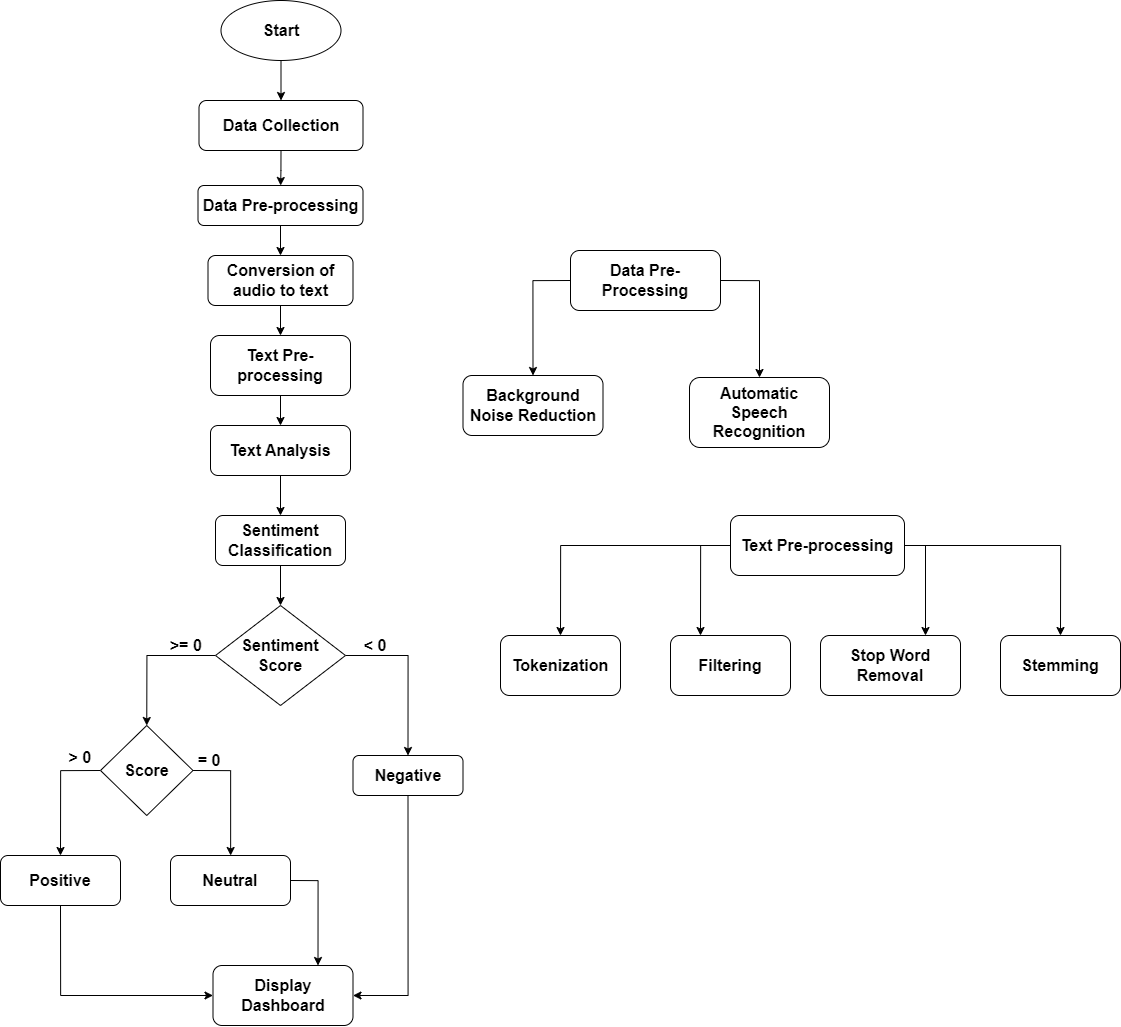


Fig 4.2: Data flow Diagram

**4.3 UML DIAGRAMS**

A UML diagram is a way to visualize systems and software using Unified Modelling Language (UML). Software engineers create UML diagrams to understand the designs, code architecture, and proposed implementation of complex software systems. UML diagrams are also used to model workflows and business processes. Coding can be a complicated process with many interrelated elements. There are often thousands of lines of programming language that can be difficult to understand at first glance. A UML diagram simplifies this information into a visual reference that’s easier to digest. It uses a standardized method for writing a system model and capturing conceptual issues.

The two main categories of UML diagrams are structural diagrams and behavioural diagrams. Structural diagrams represent the static aspects of the system, such as its classes, objects, and their relationships. Common structural diagrams include Class Diagrams, Component Diagrams, and Deployment Diagrams. Class Diagrams, for example, show the structure of the system by displaying classes, their attributes, methods, and the relationships between them (such as inheritance, association, and dependency).

Behavioural diagrams, on the other hand, represent the dynamic aspects of the system, such as its interactions, activities, and states over time. Key examples of behavioural diagrams are Use Case Diagrams, Sequence Diagrams, and Activity Diagrams. Use Case Diagrams describe the functionality of the system from the user's perspective, showing how users (actors) interact with the system to achieve specific goals. Sequence Diagrams depict how objects interact with each other in a sequence, highlighting the flow of messages. Activity Diagrams represent the flow of control within a process or system, often used for modelling workflows and business processes.

UML diagrams provide a common language that both developers and stakeholders can understand. They help in visualizing the system architecture, making it easier to communicate design decisions and understand complex systems. By using UML, development teams can identify design flaws early, reduce ambiguities in requirements, and ensure that the system meets the desired objectives.

**4.3.1 USECASE DIAGRAM**

Use-case diagrams describe the high-level functions and scope of a system. These diagrams also identify the interactions between the system and its actors. The use cases and actors in use-case diagrams describe what the system does and how the actors use it, but not how the system. operates internally

#### 

#### Fig 4.3.1: Use case Diagram

1.**Helpdesk Agent**:

* Interacts with the system to submit input, likely audio or text data.

2. **Sentiment Analysis System**:

* The central system responsible for processing the input and analysing sentiment. It consists of several steps:

a. **Submit Input**:

* The entry point where the helpdesk agent provides input (e.g., audio recording or text data).

b. **Collect Audio Data**:

* If the input is in the form of audio, it is collected at this stage.

c. **Noise Reduction**:

* The system applies noise reduction techniques to clean the audio data, ensuring accurate processing.

d. **Speech-to-Text Conversion**:

* Converts the cleaned audio into text using a speech recognition model. This is essential for text-based sentiment analysis.

e. **Sentiment Analysis**:

* The converted text is analysed by the system to determine the sentiment (positive, negative, or neutral) using a trained model.

f. **Generate Sentiment**:

* The processed sentiment data is generated and formatted for presentation.

g. **Display the Dashboard**:

* The analysed sentiment is displayed on a dashboard for the helpdesk agent to review.

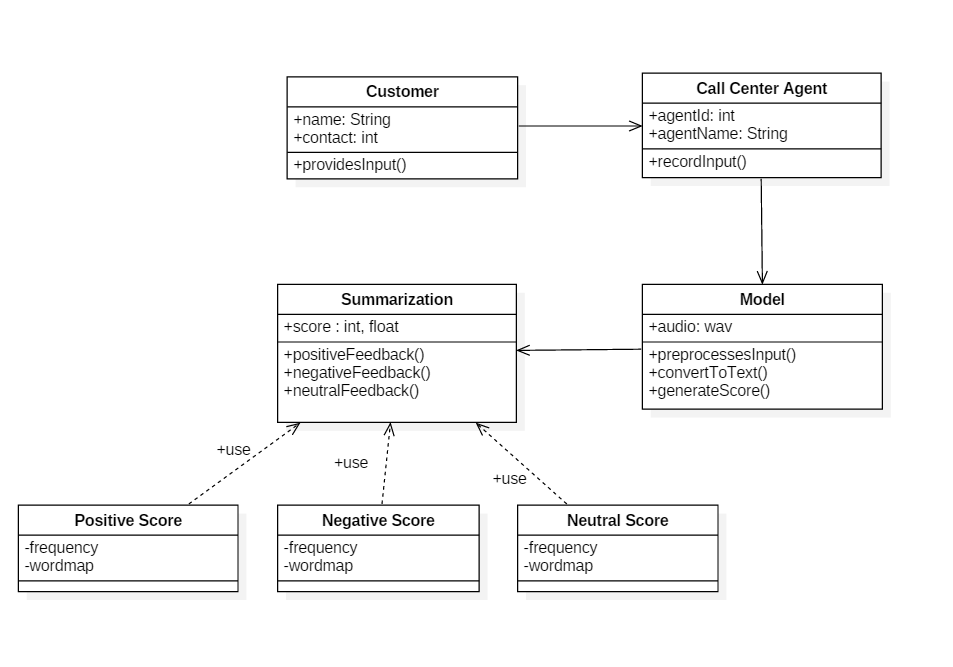
3. **Model**:

* Represents the trained machine learning model used for:
  + Speech-to-text conversion.
  + Noise reduction.
  + Sentiment analysis.

**4.3.2 CLASS DIAGRAM**

The class diagram depicts a static view of an application. It represents the types of objects residing in the system and the relationships between them. A class consists of its objects, and also it may inherit from other classes. A class diagram is used to visualize, describe, document various different aspects of the system, and also construct executable software code. It shows the attributes, classes, functions, and relationships to give an overview of the software system. It constitutes class names, attributes, and functions in a separate compartment that helps in software development. Since it is a collection of classes, interfaces, associations, collaborations, and constraints, it is termed as a structural diagram.

The class diagram is the main building block of [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) modelling. It is used for general [conceptual modelling](https://en.wikipedia.org/wiki/Conceptual_model) of the structure of the application, and for detailed modelling, translating the models into [programming code](https://en.wikipedia.org/wiki/Programming_code). Class diagrams can also be used for [data modelling](https://en.wikipedia.org/wiki/Data_modeling). The classes in a class diagram represent both the main elements, interactions in the application, and the classes to be programmed.

. Fig 4.3.2: Class Diagram

#### 1. Customer Class

#### Attributes:

#### name: A string representing the customer's name.

#### contact: An integer for the customer's contact information.

#### Method:

#### Provides Input (): A method indicating that the customer provides input, possibly as feedback.

#### 2. Call Center Agent Class

#### Attributes:

#### Agent Id: An integer representing the unique ID of the agent.

#### Agent Name: A string representing the agent's name.

#### Method:

#### Record Input (): A method for recording the feedback or input provided by the customer.

#### 3. Model Class

#### Attributes:

#### audio: An attribute to store the audio file of the customer's input (in .wav format).

#### Methods:

#### Preprocesses Input (): A method to preprocess the input audio data.

#### Convert To Text (): Converts audio input into text.

#### Generate Score (): Generates a score (positive, negative, or neutral) based on the feedback.

#### 4. Summarization Class

#### Attributes:

#### score: Can be an integer or a float, representing the summarized score of the feedback.

#### Methods:

#### Positive Feedback (): Identifies and processes positive feedback.

#### Negative Feedback (): Identifies and processes negative feedback.

#### Neutral Feedback (): Identifies and processes neutral feedback.

#### 5. Positive Score Class

#### Attributes:

#### frequency: Tracks how often positive words appear.

#### Word map: Maps specific words to their significance in positive feedback.

#### Relationship:

#### Linked to the Summarization class with a "use" relationship, as positive scores are part of the summarization process.

#### 6. Negative Score Class

#### Attributes:

#### frequency: Tracks how often negative words appear.

#### Word map: Maps specific words to their significance in negative feedback.

#### Relationship:

#### Linked to the Summarization class similarly to the Positive Score class.

#### 7. Neutral Score Class

#### Attributes:

#### frequency: Tracks how often neutral words appear.

#### Word map: Maps specific words to their significance in neutral feedback.

#### Relationship:

#### Also linked to the Summarization class in the same way.

#### 4.3.3 SEQUENCE DIAGRAM

#### A sequence diagram in Unified Modeling Language (UML) is a diagram that shows how objects interact and operate with each other in a sequence. Sequence diagrams are used by software developers and business professionals to document processes or understand requirements for new systems.

#### 

#### Fig 4.3.3: Sequence Diagram

#### Customer → Call Center Agent: Initiates the Call

#### The process starts when the customer initiates a call to the call center agent. This action is represented as a message from the Customer to the Call Center Agent.

#### Call Center Agent → Customer: Respond to the Call

#### The call center agent responds to the customer's call. This step ensures a two-way interaction is established between the customer and the agent.

#### Call Center Agent → Model: Records the Call

#### The call center agent records the conversation using a system (likely integrated with the Model). This recorded audio serves as the input for further processing.

#### Model: Processing the Input

#### Once the call is recorded, the Model processes the input. This includes steps like preprocessing the audio (e.g., noise removal, segmentation) and converting it to text for analysis.

#### Model: Sentiment Analysed

#### The Model analyses the processed input to identify the sentiment (e.g., positive, negative, or neutral). This is likely based on predefined algorithms or machine learning techniques.

#### Model → Call Center Agent: Visualization

#### Finally, the analyzed sentiment or feedback data is visualized (e.g., as a report or dashboard) and shared with the Call Center Agent for further action, insights, or decision-making.

#### 4.3.4 ACTIVITY DIAGRAM

#### An activity diagram is a type of Unified Modelling Language (UML) flowchart that shows the flow from one activity to another in a system or process. It’s used to describe the different dynamic aspects of a system and is referred to as a ‘behaviour diagram’ because it describes what should happen in the modelled system.

Activity diagrams are an essential part of the [Unified Modelling Language (UML)](https://www.geeksforgeeks.org/unified-modeling-language-uml-introduction/) that help visualize workflows, processes, or activities within a system. They depict how different actions are connected and how a system moves from one state to another. By offering a clear picture of both simple and complex workflows, activity diagrams make it easier for developers and stakeholders to understand how various elements interact in a system.

Activity diagrams show the steps involved in how a system works, helping us understand the flow of control. They display the order in which activities happen and whether they occur one after the other (sequential) or at the same time (concurrent). These diagrams help explain what triggers certain actions or events in a system.

#### Fig 4.3.4: Use case Diagram

#### Start

#### The process begins (indicated by the black circle at the top).

#### User Initiates the Call

#### A customer initiates a call, which starts the data collection process.

#### Collect Audio Data

#### The system records the audio data of the conversation between the customer and the call center agent.

#### Reduce Noise from Audio

#### The audio data undergoes noise reduction to improve clarity and quality, making it suitable for analysis.

#### Convert Audio to Text

#### The cleaned audio data is converted into text using speech-to-text technology. This step allows for textual analysis.

#### Pre-Process the Text

#### The text is pre-processed, which may involve:

#### Tokenization (breaking text into words or phrases).

#### Removing stop words (common words like "is", "the", etc.).

#### Lowercasing and normalization.

#### Stemming or lemmatization (reducing words to their base forms).

#### Analyze the Sentiment

#### The pre-processed text is analyzed to determine its sentiment (positive, negative, or neutral). This is done using sentiment analysis techniques or machine learning models.

#### Decision: Sentiment Score Generated?

#### A decision point is reached where the sentiment score is evaluated:

#### If the score is greater than or equal to zero, it indicates a positive sentiment.

#### If the score is less than zero, it indicates a negative sentiment.

#### Positive Sentiment

#### If the sentiment is positive, the corresponding result is displayed.

#### Negative Sentiment

#### If the sentiment is negative, the corresponding result is displayed.

#### End

#### The process concludes (indicated by the black circle at the bottom).

#### 

#### 4.3.5 DEPLOYMENT DIAGRAM

#### Deployment diagrams model the physical architecture of a system. Deployment diagrams show the relationships between the software and hardware components in the system and the physical distribution of the processing.

The deployment diagram is mostly employed by network engineers, system administrators, etc. with the purpose of representing the deployment of software on the hardware system. It envisions the interaction of the software with the hardware to accomplish the execution. The selected hardware must be of good quality so that the software can work more efficiently at a faster rate by producing accurate results in no time.

#### 

#### Fig 4.3.5: Deployment diagram

#### 1. Customer Client

#### Represents the user or an application interacting with the system.

#### Communicates via HTTP to send requests to the Web Server.

#### 2. Web Server

#### Acts as the entry point for the customer’s requests.

#### Handles HTTP requests and forwards them via a REST API to the API Gateway.

#### 3. API Gateway

#### Serves as a central hub to manage API calls.

#### Routes incoming requests to various downstream services, ensuring secure and efficient communication.

#### 4.Performs the following tasks:

#### Sends sentiment data to the Analytics Engine.

#### Stores data in the Sentiment Database.

#### 5. Sentiment Database

#### Stores data related to user interactions or sentiments (e.g., customer feedback, reviews).

#### Provides updates or raw data for other components like the Analytics Engine.

#### 6.Analytics Engine

#### Processes the data (e.g., performs sentiment analysis) received from the API Gateway and the database.

#### Outputs processed information (e.g., reports or insights).

#### 7. Report Server

#### Receives processed report data from the Analytics Engine.

#### Generates and serves reports to stakeholders.

#### Data Flows

#### HTTP: Customer Client → Web Server

#### REST API: Web Server → API Gateway

#### Data Storage: API Gateway → Sentiment Database

#### Sentiment Data: API Gateway → Analytics Engine

#### Data Updates: Sentiment Database → Analytics Engine

#### Report Data: Analytics Engine → Report Server

#### CHAPTER-5

#### CONCLUSION

#### The Sentiment Analysis of Incoming Calls on Helpdesk project successfully addresses the challenges of real-time emotion detection in customer interactions by integrating advanced machine learning models and natural language processing techniques. The system provides real-time sentiment scoring ranging from -1 (very negative) to +1 (very positive), ensuring immediate insights during calls. With enhanced noise reduction using convolutional neural networks, transcription accuracy is significantly improved, and emotional intensity analysis allows differentiation between mild and extreme emotions. A user-friendly dashboard visualizes sentiment trends over time, highlights customer interactions, and provides actionable insights for helpdesk operators. Continuous learning is achieved through feedback from post-call surveys, enabling the system to adapt to evolving customer behavior and language patterns. This project represents a substantial advancement in customer service by allowing organizations to understand and address customer emotions in real time, thereby improving satisfaction and operational efficiency. Future developments could include multilingual support, voice tone analysis, and predictive analytics to forecast recurring issues and sentiment trends, paving the way for smarter, data-driven customer service strategies.

#### 5.1 FUTURE ENHANCEMENT

#### Future enhancements for the Sentiment Analysis of Incoming Calls on Helpdesk project could significantly improve its effectiveness and usability. One potential enhancement is advanced emotion detection. The system can be upgraded to classify specific emotions such as anger, frustration, happiness, or confusion rather than just positive, negative, and neutral sentiments. Additionally, incorporating intensity scoring for emotions would allow for a more nuanced understanding, distinguishing between mild and extreme emotional states.

#### Another enhancement could focus on providing real-time feedback for helpdesk agents. The system can be designed to offer contextual suggestions during calls, recommending phrases or actions to de-escalate tension or show empathy. Furthermore, dynamic call handling strategies could adapt in real time based on sentiment trends, such as escalating calls to senior staff if the detected sentiment worsens.

#### Improved visualization and reporting tools are also crucial for future development. Enhancements like sentiment heatmaps to visualize emotion variations during calls and dashboards that monitor agent-specific sentiment trends can help identify training needs. Comparative analysis across teams, regions, or timeframes would offer deeper insights into performance and customer satisfaction trends.

#### Integration with CRM and ticketing systems could add another layer of functionality. By analyzing customer interaction history, the system could provide more accurate sentiment predictions and prioritize tickets based on sentiment scores, ensuring that urgent cases receive prompt attention.

#### Finally, combining speech and text analysis could enhance sentiment detection accuracy. A hybrid approach that incorporates paralinguistic cues such as tone, pitch, and pace, alongside text transcription analysis, would provide a holistic view of the customer’s emotional state. These enhancements could transform the project into a robust tool for improving customer satisfaction and agent performance.

#### 

#### 

#### 

#### CHAPTER-6

#### REFERENCES

#### [1] A Deep Learning System for Sentiment Analysis of Service Calls, Yanan Jia, Businessolver / Bellevue, WA Professional Communication Conference, 2007. IPCC 2007. IEEE International

#### [2] Machine Learning-Based Sentiment Analysis of Incoming Calls on Helpdesk, C.D. Kokane et al., October 2023, International Journal on Recent and Innovation Trends in Computing and Communication 11(9):21-27

#### [3] Sentiment Analysis of Call Centre Audio Conversations using Text Mining: This study explores the application of text mining techniques on transcribed audio recordings to detect speakers' emotions, aiming to predict customer satisfaction levels.

#### [4] Call Center Customer Sentiment Analysis Using ML and NLP, Imad Aattouri, Hicham Mouncif, Mohamed Rida, 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA)

#### [5] Sentiment Analysis of Call Centre Audio Conversations using Text Classification, Souraya Ezzat, Neamat El Gayar, Cairo University, Moustafa M. Ghanem

#### [6] Sentiment Analysis and Emotion Recognition from Speech Using Universal Speech Representations, by Bagus Tris Atmaja and Akira Sasou