# AI POWERED VOICE MAIL PRIORITIZATION SYSTEM

A PROJECT REPORT

Submitted by


**FATHIMA NADWA RAMEEZ (*MES21AD019*)**
**MUHAMMED SHAHIQ A R (*MES21AD046*)**
**MOHAMMED SINAN P (*MES21AD040*)**
**MUHAMMED AJAS (*MES21AD041*)**


to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirement for the award of the Degree

of

Bachelor of Technology

in

*Artificial Intelligence and Data Science*


**Department of Artificial Intelligence and Data Science**
**[B.Tech. Programme ]**
MES College of Engineering Kuttippuram
Thrikkanapuram South P.O, Malappuram Dt, Kerala, India 679582


2024 - 2025

# DECLARATION

We hereby declare that the project report "AI POWERED VOICEMAIL PRIORITIZATION SYSTEM", submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done under the supervision of Mr. Abin C Jose, Assistant Professor, Dept. of Artificial Intelligence and Data Science. This submission represents our ideas in our own words and where ideas or words of others have been included, We have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place: Kuttippuram
Date: 25/03/2025

FATHIMA NADWA RAMEEZ

MOHAMMED SINAN P

MUHAMMED AJAS A P

MUHAMMED SHAHIQ A R

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

MES COLLEGE OF ENGINEERING, KUTTIPPURAM

# CERTIFICATE

This is to certify that the report entitled **"AI POWERED VOICEMAIL PRIORITIZATION SYSTEM"** submitted by **FATHIMA NADWA RAMEEZ, MOHAMMED SINAN P, MUHAMMED AJAS A P,** and **MUHAMMED SHAHIQ A R**, to the APJ Abdul Kalam University in partial requirements for the award of the Degree of Bachelor of Technology in Artificial Intelligence and Data Science is a bonafide record of the project work carried out under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

| Project Guide: | Project Coordinator: | Head of Department: |
|---|---|---|
| Mr. Abin C Jose | Mr. Sherikh K K | Dr. Govindaraj E |
| Assistant Professor | Assistant Professor | Associate Professor |
| Dept of ADS | Dept of ADS | Dept of ADS |

# ACKNOWLEDGEMENT

At the outset, we would like to thank the Almighty for all his blesssings that led us here.

We are grateful to ***Dr. Rahumathunza I***, *Principal, MES College of Engineering, Kuttippuram*, for providing the right ambiance to complete this project. We would also like to extend our sincere gratitude to ***Dr. Govindaraj E*** *, Head of Department, Artificial Intelligence and Data Science, MES College of Engineering, Kuttippuram.*

We are deeply indebted to the project coordinator, ***Mr. Sherikh K K****, Assistant Professor, Department of Artificial Intelligence and Data Science* for his continued support.

It is with great pleasure that we express our deep sense of gratitude to our project guide, ***Mr. Abin C Jose****, Assistant Professor, Department of Artificial Intelligence and Data Science*, for his guidance, supervision, encouragement and valuable advice in each and every phase.

We would also like to express our sincere thanks and gratitude to all staff members of the department, our friends and family members for their cooperation, positive criticism, consistent support and consideration during the preparation of this work.

**FATHIMA NADWA RAMEEZ**

**MOHAMMED SINAN P**

**MUHAMMED AJAS A P**

**MUHAMMED SHAHIQ A R**

# Abstract

AI-powered voicemail prioritization system is a system that enhances voicemail management by promptly addressing high-priority messages. The system uses real-time analysis to assess urgency through tone, language cues, and keywords like ”urgent” or ”emergency,” leveraging NLP tools such as spaCy for keyword detection and sentiment analysis. It includes a voice-to-text transcription feature for converting audio to text, to streamline user experience. Automated forwarding directs urgent messages to the right team members, while encryption safeguards sensitive data, making it suitable for industries requiring confidentiality, such as healthcare or finance. This system ultimately aims to boost voicemail efficiency, reduce response times, and improve productivity in critical communication environments.

# CONTENTS

# LIST OF FIGURES

# ABBREVIATIONS

| | |
|---|---|
| AI | Artificial Intelligence |
| BERT | Bidirectional Encoder Representations from Transformers |
| NER | Name Entity Recognition |
| NLP | Natural Language Processing |
| API | Application Programming Interface |
| PII | Personally Identifiable Information |

# CHAPTER 1

# INTRODUCTION

In Today's fast-paced, information-driven world, effective communication is essential, especially in environments where timely responses are critical. Traditional voicemail systems often lack the ability to prioritize messages based on urgency, resulting in potential delays that can affect efficiency, customer satisfaction, and even critical outcomes. For organizations that handle large volumes of voicemails daily, these limitations can lead to missed opportunities and slower response times.

This project introduces an AI-powered voicemail prioritization system aimed at transforming voicemail management. Using natural language processing (NLP), the system evaluates and ranks voicemails based on urgency by analyzing keywords and emotional tone. This enables immediate attention to important messages, reducing the chances of delayed responses where quick action is essential.

To improve usability, the system transcribes voicemails to text, allowing users to read messages instead of listening to lengthy audio recordings. Summaries highlight key information, making it easier to understand message content at a glance. High-priority voicemails are automatically directed to the appropriate team members, minimizing manual sorting. Additionally, the system uses encryption to protect sensitive information, meeting the security requirements of industries that handle confidential data.

This system holds significant potential across diverse high-stakes industries. In healthcare and emergency services, it can ensure urgent patient or dispatch calls receive immediate attention, directly impacting outcomes. Customer service departments can use it to address critical issues swiftly, boosting satisfaction. Financial and legal sectors benefit from prioritized client messages with added data security, while real estate managers can respond faster to tenant and client needs. Public sector agencies, handling vast message volumes, can also expedite responses to urgent citizen concerns. By transcribing messages, summarizing key points, and securely directing high-priority messages to relevant team members, the system enhances efficiency, timely responses, and trust across various professional environments.

By implementing this system, organizations can enhance communication efficiency, ensure timely responses, and reduce risks associated with delayed action. This report covers the design, features, and benefits of the system, focusing on its potential to improve productivity and responsiveness in critical communication settings.

# CHAPTER 2

# LITERATURE SURVEY

Research into voicemail urgency detection has explored various approaches to classify voicemail messages based on their content. The paper ”Voicemail Urgency Detection Using Context Dependent and Independent NLP Techniques” investigates methods to prioritize voicemail messages. The authors discuss the importance of factors such as sender position, time of sending, and textual content. The research focuses on using machine learning models for text vectorization(BERT and TF-IDF) and classification(SVM) to determine the urgency of messages, comparing different combinations to identify the most effective approach for business communication [1].

The transcription and summarization of voicemail messages is another critical area of research. In ”Transcription and Summarization of Voicemail Speech,” the authors propose a hybrid connectionist approach(HMM and MLP) for voicemail transcription, achieving competitive performance with context-independent models. The system integrates statistical and prior knowledge sources for term weighting to deliver concise message summaries via a GSM Short Message Service (SMS) gateway. The study emphasizes the efficiency of combining statistical models with real-world knowledge for summarization tasks [2].

A study on ”Urgent Voicemail Detection Focused on Long-term Temporal Variation” presents a novel method for detecting urgent speech based on speech rhythm. This research highlights the inadequacy of using short-term features alone and introduces long-term temporal features such as envelope modulation spectrum(EMS) and temporal statistics of Mel-frequency cepstrum coefficients(MFCC). By combining long-term and short-term features using neural networks, the proposed method significantly reduces error rates in detecting urgency compared to conventional techniques [3].

In the paper ”Privacy-Preserving Personal Identifiable Information (PII) Label Detection Using Machine Learning”, the focus is on safeguarding privacy in voicemail transcription. The research demonstrates that competitive performance in transcription tasks can be achieved using context-independent systems (TF-IDF), while also integrating privacy-preserving techniques to detect sensitive information in voicemail content [4].

Research comparing Named Entity Recognition (NER) software such as StanfordNLP, NLTK, OpenNLP, SpaCy, and Gate provides essential insights into entity detection in text. The paper ”A Replicable Comparison Study of NER Software” evaluates various NER tools in the context of voicemail transcription, highlighting discrepancies in performance across tools. The study finds that StanfordNLP often delivers the best results for entity recognition, which is a crucial component in understanding voicemail content [5].

The paper ”Automated News Summarization Using Transformers” explores transformer-based models for text summarization, focusing on handling large volumes of information by generating concise summaries. The authors compare models such as BART, T5, and PEGASUS, which are fine-tuned for abstractive summarization tasks, using the BBC news dataset for evaluation. The study emphasizes the effectiveness of these models in summarizing structured text, highlighting T5's superior performance based on ROUGE scores. Although computationally intensive, transformer models show promise in creating coherent and contextually accurate summaries, making them relevant for voicemail summarization tasks. This research provides valuable insights into using transformers for generating concise summaries, which can be adapted to improve voicemail processing and prioritization systems [6].

From the research reviewed, we understand that voicemail prioritization involves a combination of various NLP techniques that focus on both context-independent and dependent factors. The urgency of voicemails can be determined through speech content, temporal speech features, and entity recognition. Studies emphasize that combining short-term features with long-term temporal variations improves accuracy in detecting urgency. Furthermore, privacy-preserving techniques ensure sensitive information is handled securely during transcription. Named Entity Recognition plays a vital role in extracting important information from voicemails, contributing to more accurate prioritization. These insights guide us in developing a robust voicemail urgency detection system that balances performance and privacy concerns.

# CHAPTER 3

# SYSTEM ANALYSIS

## 3.1 Existing System

Existing voicemail systems primarily focus on basic functions such as voicemail-to-text transcription, timestamp-based prioritization, and keyword filtering. These systems, like those offered by Google Voice or Apple Voicemail, allow users to read messages but do not prioritize based on urgency or emotional tone, making them less effective for high-stakes scenarios.Some systems in customer service settings include basic sentiment analysis to gauge customer dissatisfaction, but this is often limited to identifying negative keywords and does not fully capture urgency or emotional context. Additionally, existing systems rarely implement secure handling of voicemails; encryption and the protection of Personally Identifiable Information (PII) are usually minimal, leaving sensitive data at risk.

## 3.2 Proposed System

The proposed system aims to develop an AI-powered voicemail prioritization tool that intelligently ranks voicemails based on urgency and emotional tone, helping organizations respond swiftly to critical messages. This user-friendly application will allow users to manage voicemails more efficiently by capturing calls, transcribing content, identifying high-priority messages, and securely directing them to appropriate team members. The workflow includes voice analysis, urgency detection, transcription, and secure data handling to create an efficient, privacy-focused voicemail management experience.

The proposed system consists of the following modules:

1. **Voice Capture and Analysis**
2. **Voice-to-Text Transcription and PII Masking**
3. **Urgency and Sentiment-based Ranking**
4. **Automated Forwarding and Departmental Sharing**
5. **User Dashboard and Analytics**

## 3.3 Objectives of Proposed System

- Develop a sophisticated voicemail system that ranks messages based on urgency and emotional tone using machine learning and natural language processing techniques, allowing organizations to identify and respond promptly to high-priority messages.
- Implement voice-to-text transcription enabling users to read and quickly review voicemail content without the need to listen to lengthy recordings, saving time and enhancing productivity.
- Integrate robust data protection measures, including automatic masking of Personally Identifiable Information (PII) and end-to-end encryption, to ensure secure storage and handling of sensitive information, meeting industry compliance standards.

- Automate the sorting and forwarding of high-priority voicemails to relevant departments or team members, reducing manual sorting and ensuring that time-sensitive messages reach the appropriate personnel efficiently.
- Design an intuitive user dashboard equipped with analytics to monitor voicemail patterns, assess response times, and identify peak communication times, allowing organizations to optimize their voicemail management practices.

### 3.3.1 Advantages of Proposed System

- Prioritizes voicemails based on urgency and emotional tone, ensuring prompt responses to critical messages, improving efficiency and customer satisfaction.
- Provides transcription, allowing users to quickly scan messages, reducing time spent on voicemail management.
- Enhances data security with PII masking and encryption, protecting sensitive information and meeting regulatory standards in fields like healthcare and finance.
- Automates sorting and forwarding, reducing manual workload and ensuring timely delivery to relevant departments.
- Offers analytics through a user dashboard, providing insights for optimizing response times and resource allocation.
- Streamlines user experience with a simple interface, improving satisfaction and organizational communication efficiency.

## 3.4 User Requirement Specification

The user requirements specification outlines the needs and expectations of the users for the AI-powered voicemail prioritization system. This includes identifying the specific tasks the system is intended to perform, the types of input it will handle, and the expected output. This document formally describes the requirements of the proposed system, detailing functional and non-functional requirements and including descriptions of the user interface and operational expectations.

**Functional Requirements**

1. The system should capture voicemail audio and perform voice analysis to detect urgency and emotional tone for prioritization.

2. The system should convert voicemail audio to text with automatic masking of Personally Identifiable Information (PII) for privacy.

3. The system should rank voicemails by urgency, enabling timely responses to high-priority messages.

4. The system should automatically forward high-priority messages to relevant team members or departments, streamlining response workflows.

5. The system should offer a user-friendly dashboard for managing voicemail messages, displaying priority levels, and providing analytics.

**Non-Functional Requirements**

1. The system should process and analyze voicemail data efficiently, even when handling high volumes of incoming messages.

2. The system should be scalable to handle an increasing number of voicemails over time without degrading performance.

3. The system should incorporate robust security measures to protect sensitive data and prevent unauthorized access.

4. The system should have an intuitive and accessible user interface, suitable for users with varying technical experience.

5. The system should ensure high availability and reliability, with minimal downtime or delays in voicemail processing.

## 3.5 Feasibility Study

The feasibility study for the AI-powered voicemail ranking system examines its practicality, user impact, and resource efficiency. The analysis includes:

- Technical Feasibility
- Economic Feasibility
- Operational Feasibility

### 3.5.1 Technical Feasibility

The technical feasibility assesses the capability to develop and implement the voicemail ranking system effectively. Key considerations include:

- **Availability of Resources**: The hardware and software required for this system, such as high-performance processors, voice processing APIs, and machine learning libraries, are readily accessible. The infrastructure needed for voice transcription, sentiment analysis, and PII data-hiding can be supported by cloud computing or on-premise systems.

- **Team Expertise**: The team possesses essential knowledge in AI and machine learning, including experience with natural language processing (NLP), sentiment analysis, voice recognition, and secure data handling. Familiarity with classification models and API integration is critical to achieving the system's goals.

- **Project Complexity**: The system will be developed in phases, starting with simple voice prioritization and transcription, then adding sentiment and urgency analysis. This staged approach reduces technical risks and ensures manageable milestones for incremental improvements.

- **Compatibility**: The system will be designed for seamless integration with existing voicemail and communication platforms, supporting compatibility across a variety of industries, including healthcare, legal, and customer service.

### 3.5.2 Economic Feasibility

The economic feasibility evaluates the financial viability of the project by analyzing:

- **Development Costs**: Estimated costs cover software licenses, hardware for processing, and development team salaries. Expenses related to transcription APIs, data storage, and security protocols will also be considered, especially if high-compliance industries (e.g., healthcare, legal) are targeted.

- **Potential Benefits**: By prioritizing and automating voicemail handling, the system is expected to save organizations time and resources, improving workflow efficiency. Increased responsiveness to high-priority calls can lead to better client satisfaction, potentially enhancing customer loyalty and reputation.

- **Return on Investment (ROI)**: A cost-benefit analysis will consider the potential savings in labor and response times against development costs. Monetization through subscription plans, licensing for high-security industries, or partnerships with customer service solutions will be explored.

### 3.5.3 Operational Feasibility

The operational feasibility assesses the system's effectiveness in real-world settings by examining:

- **Resource Availability**: All necessary tools, software, and datasets are accessible, including datasets for model training on sentiment and urgency detection. Access to scalable computing resources is ensured for smooth operation and growth as the system is adopted.

- **Compatibility with Existing Processes**: The system is designed to enhance, not disrupt, current voicemail workflows. Integration feedback from potential users will guide adjustments, ensuring that the system complements standard practices without causing interruptions.

- **Impact on Stakeholders**: The system's benefits for end users, such as support staff, healthcare providers, and legal teams, are significant. By allowing prioritized responses to urgent voicemails, it can streamline operations and support informed, timely decision-making, leading to a positive reception among users. Stakeholder engagement throughout development will ensure alignment with user expectations.

## 3.6 System Specification

A system requirements specification outlines the necessary conditions for successful implementation. Non-functional requirements impose constraints that affect design and implementation, such as performance standards. The system requirements for the project include:

### Hardware Specification

For optimal performance, the following minimum hardware specifications are recommended for both servers hosting the system and client machines accessing it:

- **Operating System**: Windows 10 or Linux to provide a stable environment for application execution.
- **CPU**: Multi-core processor (Intel i5 or better) to ensure efficient processing capabilities.
- **RAM**: Minimum of 4GB to handle memory requirements for processing voicemails and machine learning tasks.
- **Storage**: At least 256 GB SSD to accommodate software installations, datasets, and generated outputs.

### Software Specification

To ensure the proposed system operates effectively, the following software must be installed and running on the server/client:

- **Platform**: Google Colab, Jupyter Notebook, or Visual Studio Code to provide a conducive development environment.
- **Programming Language**: Python 3.8 as the primary programming language for implementing machine learning and NLP tasks.
- **Frameworks and Libraries**:
  - PyTorch for machine learning and natural language processing tasks.
  - Flask to facilitate web-based interaction and serve as the backend framework for the application.

# CHAPTER 4

# PROPOSED METHODOLGY AND FRAMEWORK

## 4.1 Methodology

This chapter describes the Methodology of the AI-powered voicemail prioritization system. Each major system component is detailed to illustrate the data flow and processing stages that enable voicemail prioritization, transcription, redaction, and routing.

For the ease of implementation, we implemented the project in the healthcare sector, where efficient voicemail management is crucial for handling patient inquiries, emergency cases, appointment scheduling, and follow-ups. This sector was chosen due to the for the ease of implementation, data collection, and explanation. Healthcare organizations frequently receive voicemails related to appointments, prescriptions, emergencies, and patient inquiries, making it easier to categorize and prioritize messages.

### 4.1.1 System Architecture

Figure 4.1 illustrates the architecture and workflow of the voicemail prioritization system. The primary components are:

1. **Voice Capture and Storage**: Voicemails are initially captured and stored for analysis, utilizing Twilio for automated recording and retrieval.

2. **Speech-to-Text Conversion**: An automated speech recognition system con-

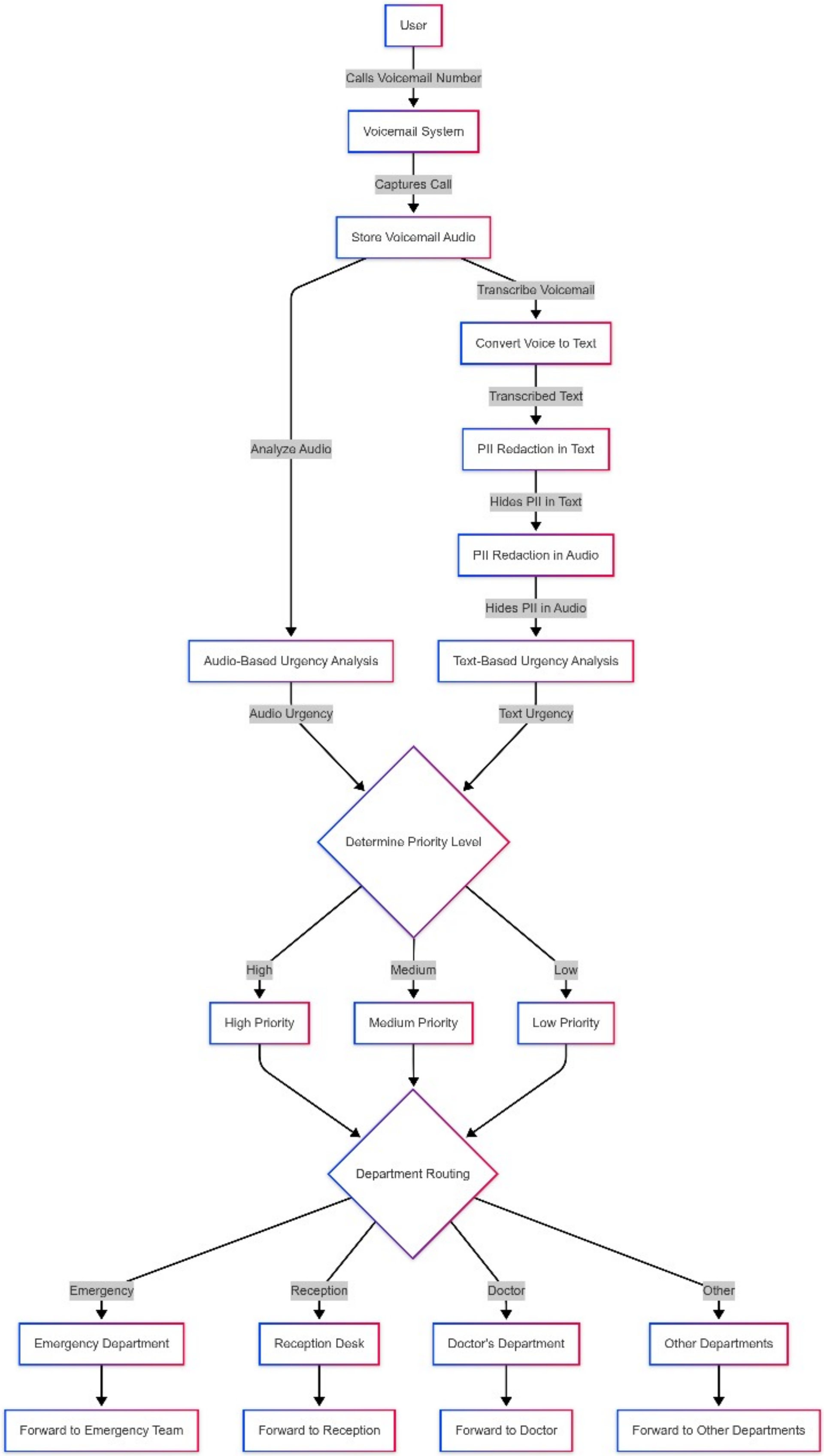


Figure 4.1: System Architecture of AI-Powered Voicemail Prioritization System

verts the voicemail audio to text, enabling further textual analysis.

3. **PII Information Redaction**: Natural language processing (NLP) techniques like spaCy identify and redact Personally Identifiable Information (like names, addresses, phone numbers) within the text, ensuring compliance with privacy requirements. Redacted sections are substituted with placeholders, with authorized users having options to view the original text.

4. **Prioritization Analysis**: Voicemails are evaluated for urgency based on both audio and textual content.

   - **Tone Analysis**: OpenSMILE or librosa libraries analyze audio tone, focusing on pitch, speed, and loudness to detect urgency indicators.

   - **Text Analysis for Urgency**: Using SpaCy, the transcribed text is scanned for urgency-related keywords (e.g., ”urgent,” ”ASAP,” ”emergency”). The presence of specific terms increases the urgency score.

   - **Combined Urgency Score**: The tone and text scores are combined via a weighted approach. A Classification model is then trained on labeled urgency data to classify voicemails into predefined urgency levels.

5. **Departmental Routing**: Based on the urgency and topic, the system routes voicemails to the appropriate department or team for a timely response.

6. **User Interface**: The user interface presents voicemails, summaries, and urgency indicators, tailored to different roles within the organization (e.g., administrators and department representatives).

## 4.2 Data Collection and Preprocessing

### 4.2.1 Voice Capture and Transcription

Incoming voicemails are captured automatically and saved for processing. An automated transcription system then converts the audio into text format, which allows for deeper text-based analysis and facilitates urgency scoring.

### 4.2.2 Data Preprocessing and Redaction

After transcription, a natural language processing system scans the text for sensitive data, such as personal identifiers, and redacts these elements. Authorized users can access the unredacted text if necessary, while default views display placeholders to protect privacy as shown in Figure 4.2.

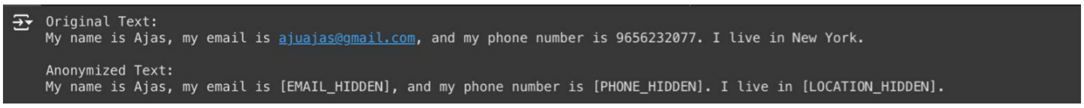


Figure 4.2: Example of PII Data Redaction

## 4.3 Feature Extraction and Urgency Classification

### 4.3.1 Audio and Textual Feature Analysis

The system assesses urgency by analyzing both audio and text features:

- **Audio Features**: Acoustic characteristics like pitch, speed, and volume are measured, indicating emotional tone and urgency. See Figure 4.3 for Audio analysed Dataset.

Figure 4.3: Example of Audio Analysed Dataset

- **Textual Cues**: Urgency-related words and phrases are detected within the transcription, contributing to the voicemail's prioritization score.

| Voicemail ID | Transcription | Emotion Label | Urgency Level | Department/Category |
|---|---|---|---|---|
| 1 | "I need help! I'm having severe chest pain right now!" | Distressed | High | Emergency Department |
| 2 | "Can you please call me back regarding my prescription?" | Neutral | Medium | Pharmacy |
| 3 | "I'd like to schedule an appointment for next week." | Calm | Low | Scheduling |
| 4 | "I'm feeling dizzy and need advice on what to do." | Anxious | High | Medical Assistance |
| 5 | "My child has a fever, and I'm worried. What should I do?" | Distressed | High | Pediatric Care |
| 6 | "Can I get my lab results over the phone?" | Neutral | Medium | Lab Services |

Figure 4.4: Example of Dataset

### 4.3.2 Training and Classification

A Random Forest model is trained on both audio and text features, using historical urgency data as shown in Figure 4.4. Once trained, this model automatically categorizes new voicemails into urgency levels based on combined feature scores, directing higher-priority voicemails for immediate attention.

## 4.4 Departmental Routing

Voicemails are categorized by urgency and topic, then routed to the relevant departments. High-priority messages are flagged for urgent response, while lower-priority messages are scheduled for routine review. This routing ensures that voicemails reach the appropriate team efficiently.

## 4.5 User Interface

The user interface is designed to present voicemail summaries, urgency levels, and additional details. Different user roles have specific access privileges; for instance, administrators can manage the entire queue, while department-specific users can view messages relevant to their teams. The interface is intended to streamline access and response to voicemails, improving response times and organizational efficiency.

This design establishes a solid foundation for developing an AI-powered voicemail prioritization system that efficiently categorizes and routes messages by urgency. With advanced transcription, privacy safeguards, and urgency analysis, it ensures timely and secure voicemail handling tailored to organizational needs.

# CHAPTER 5

# Coding

## 5.1 Front End

### 5.1.1 HTML

The front-end interface of the AI-powered voicemail ranker is built using **HTML** to structure the web pages. It provides the framework for displaying voicemail transcriptions, urgency rankings, and department classifications. The pages include elements such as tables for voicemail records, buttons for forwarding/deleting voicemails, and a dashboard for visualizing priority levels.

### 5.1.2 CSS

**CSS** is used for styling the web pages to ensure a clean and professional look. It defines the layout of voicemail tables, urgency indicators (color-coded priority levels), and department-based views. The design follows a responsive approach, ensuring accessibility across different devices.

### 5.1.3 JavaScript

JavaScript is used for **client-side interactivity**, including sorting voicemails by urgency, filtering messages by department, and dynamically updating the dashboard without requiring a page reload. AJAX is implemented to send requests to the back end for voicemail forwarding and deletion without refreshing the page.

## 5.2 Back End

### 5.2.1 Python (Flask)

The back end is developed using **Flask**, which handles API routing, voicemail processing, and integration with the **Twilio API**. Flask manages:

- **Downloading voicemails** from Twilio
- **Processing voicemail transcriptions**

- **Running AI-based classification models**
- **Handling database operations (CSV storage)**
- **Serving the web dashboard**

### 5.2.2 Speech-to-Text Transcription

The **transcription model** converts voicemail audio into text using **a fine-tuned BERT-based model**. It processes raw voicemail audio, applies noise reduction, and generates text transcriptions, which are later analyzed for urgency and department classification.

### 5.2.3 Urgency Detection (Text & Audio-Based)

- **Text-Based Urgency Detection**: NLP techniques such as **TF-IDF, sentiment analysis, and keyword extraction** determine urgency levels based on transcribed text.
- **Audio-Based Urgency Detection**: Audio features such as **MFCC, pitch (F0), jitter, and loudness** are extracted using **OpenSMILE** and analyzed using a **Random Forest classifier** to detect urgency from tone.
- **Final Priority Score**: A hybrid model combines text and audio urgency to assign a **High, Medium, or Low** priority.

### 5.2.4 Department Classification

A **fine-tuned BERT model** classifies voicemails into one of four departments: **Emergency, Pharmacy, Doctors, or Reception**. The model encodes voicemail text, extracts context, and predicts the most relevant department. The front-end interface allows users to filter voicemails by department and by urgency levels (High, Medium, Low priority) to quickly identify critical messages.

### 5.2.5 PII Redaction (Text & Voice)

- **Text-Based PII Redaction**: Sensitive information such as phone numbers, emails, and addresses are identified and replaced with placeholders (e.g., `[EMAIL_HIDDEN]`, [PHONE_HIDDEN]). This is achieved using **Regular Expressions (regex)** to detect specific patterns such as numeric sequences for phone numbers and standard email formats. Additionally, **Named Entity Recognition (NER)** models are used to recognize context-based personally identifiable information (PII). The redacted text is then securely stored to ensure privacy while being displayed in the web interface.

- **Voice-Based PII Redaction**: Voice-Based PII Redaction: Audio PII is redacted by detecting sensitive words in the transcription and replacing them with a beep sound in the corresponding audio segment. The system estimates word duration based on the total audio length, identifies PII positions, and overlays a beep while muting the original speech in those sections. The anonymized audio is saved separately for secure playback.

### 5.2.6 Database & Storage (CSV)

Processed voicemail details, including **caller info, transcription, urgency score, priority, and assigned department**, are stored in `voicemails.csv`. The Flask app reads and updates this file dynamically for real-time retrieval.

# CHAPTER 6

# INTERFACE

The user interface provides an intuitive dashboard for managing voicemails, ensuring ease of navigation and accessibility. It is designed to be user-friendly, responsive, and efficient, allowing users to seamlessly interact with the voicemail system.

## 6.1 Landing Page (/index)

- A Login page for the User to log into given Department of the Voicemail System as shown in Figure 6.2.
- Displays all received voicemails, including **caller details, transcription, urgency score, and assigned department** as shown in Figure 6.3.
- Users can **sort and filter** voicemails based on priority (High, Medium, Low) and department.
- Action buttons allow users to **forward or delete voicemails** directly from the dashboard shown in Figure 6.4.

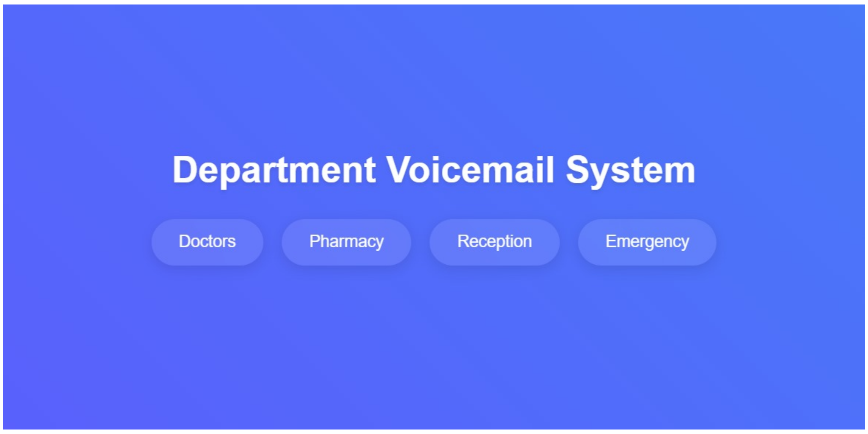


Figure 6.1: Main Page of the Voicemail System

## 6.2 Department-Specific Dashboards

Each department has a dedicated dashboard that displays voicemails assigned to it:

- **/pharmacy-dashboard** – Shows voicemails relevant to the Pharmacy department.
- **/doctors-dashboard** – Lists voicemails assigned to doctors.
- **/emergency-dashboard** – Displays high-priority emergency voicemails.
- **/reception-dashboard** – Contains general inquiries, with **PII redacted** for privacy.

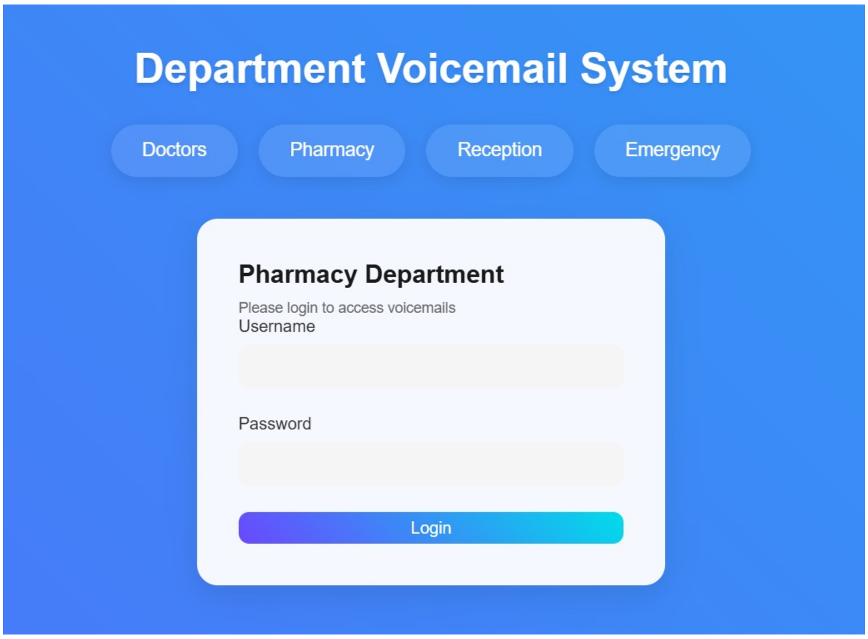


Figure 6.2: Login Page of the Specific Department

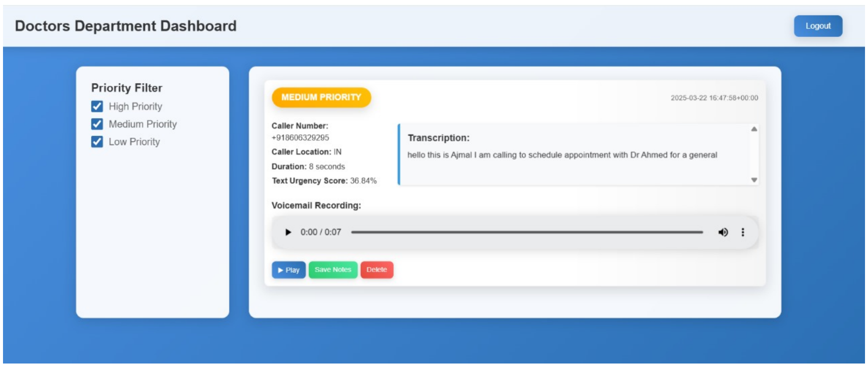


Figure 6.3: Department-Specific Dashboards

## 6.3 Voicemail Management Actions

- **View Details** – Users can click on a voicemail to see its full transcription, priority level, and classification.
- **Forward to Department** – Voicemails can be reassigned to the correct department for processing.
- **Delete Voicemail** – Users can permanently remove unnecessary voicemails.
- **Search & Filter** – Users can search for specific voicemails based on **urgency level and department**.

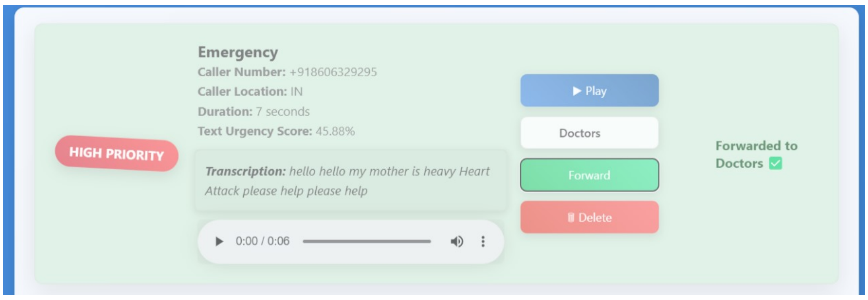


Figure 6.4: Voicemail Management Actions

## 6.4 Security Features

- **PII Redaction** ensures that **phone numbers, emails, and addresses** are hidden in text-based transcriptions.
- **Audio Beep Redaction** replaces PII in voicemail recordings with a beep sound for privacy compliance.

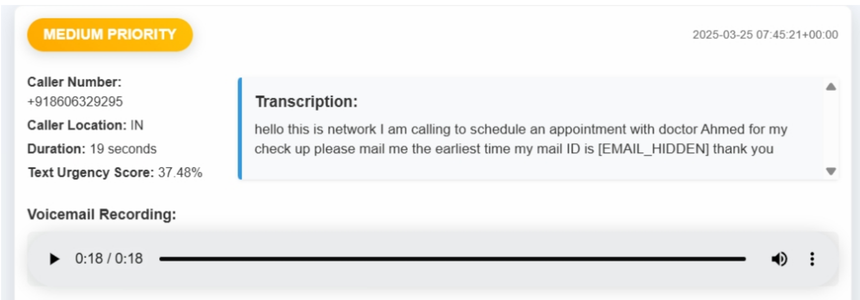


Figure 6.5: PII Redaction in Text and Audio

The system provides a structured and efficient interface for users to access and manage voicemails with minimal effort, ensuring quick response times and streamlined communication.

# CHAPTER 7

# TESTING AND RESULTS

Test and Result Analysis The testing phase of the voicemail system project was conducted to evaluate its performance across three critical functionalities: text-based urgency classification, audio-based urgency classification, and department sharing accuracy. Each test was designed to assess the system's ability to correctly classify and route voicemail messages under varying conditions. This section analyzes the outcomes of each test, identifies areas of success and concern, and offers recommendations for further optimization.

## 7.1 Test Case 1: Text-Based Urgency Classification

The first test evaluated the voicemail system's ability to classify the urgency of text-based messages into three categories: High, Medium, and Low. The classification report Figure 7.1 shows an overall **accuracy of 73%**, with a macro average F1-score of 0.71 and a weighted average F1-score of 0.72. The detailed metrics indicate varying performance across categories: the High urgency class achieved a precision of 0.65, a recall of 1.00, and an F1-score of 0.79 with a support of 11 samples; the Medium class had a precision of 1.00, a recall of 0.44, and an F1-score of 0.62 with a support of 9 samples; and the Low class recorded a precision of 0.78, a recall of 0.70, and an F1-score of 0.74 with a support of 10 samples. The perfect recall for High urgency messages (1.00) suggests the system successfully identified all critical messages, though the lower precision (0.65) indicates some false positives. Conversely, the Medium class's low recall (0.44) highlights a significant number of missed classifications, potentially due to overlapping features with other classes. The confusion matrix Figure 7.2 further illustrates this, showing that out of 11 High urgency messages, 9 were correctly classified, but 1 was misclassified as Low and 1 as Medium. For Low urgency, 19 out of 21 messages were correctly identified, but 1 was misclassified as High and 1 as Medium. The Medium class struggled the most, with only 5 out of 9 messages correctly classified, 3 misclassified as Low, and 1 as High. This suggests that the system's text-based classification model may require better feature engineering or additional training data to improve differentiation, particularly for Medium urgency messages.

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| High | 0.65 | 1.00 | 0.79 | 11 |
| Medium | 1.00 | 0.44 | 0.62 | 9 |
| Low | 0.78 | 0.70 | 0.74 | 10 |
| accuracy | | | 0.73 | 30 |
| macro avg | 0.81 | 0.71 | 0.71 | 30 |
| weighted avg | 0.80 | 0.73 | 0.72 | 30 |

Figure 7.1: Classification Report for Text-Based Urgency Testing

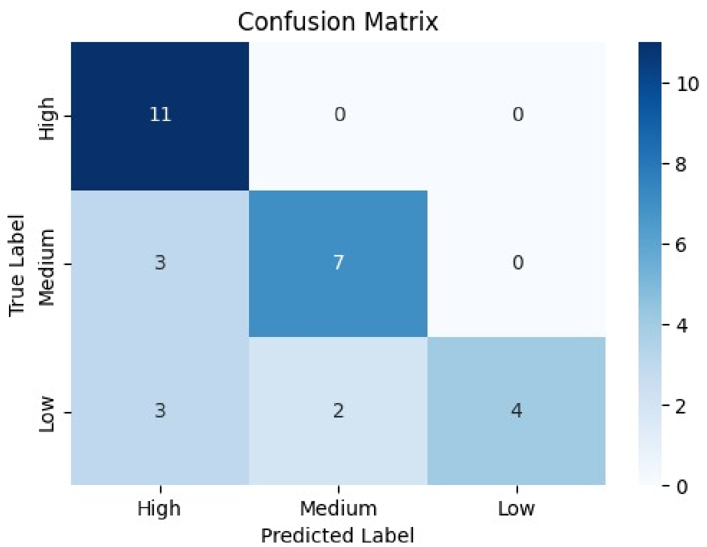


Figure 7.2: Confusion Matrix for Text-Based Urgency Testing

## 7.2 Test Case 2: Audio-Based Urgency Classification

The second test assessed the system's performance in classifying the urgency of audio-based voicemail messages, also categorized as High, Medium, and Low. The classification report Figure 7.3 reveals an improved overall **accuracy of 82%**, with a macro average F1-score of 0.80 and a weighted average F1-score of 0.82. The High urgency class achieved a precision of 0.82, a recall of 0.90, and an F1-score of 0.86 with a support of 10 samples; the Low class had a precision of 0.83, a recall of 0.90, and an F1-score of 0.86 with a support of 21 samples; and the Medium class recorded a precision of 0.83, a recall of 0.56, and an F1-score of 0.67 with a support of 9 samples. The confusion matrix Figure 7.4 shows that all 10 High urgency messages were correctly classified, with no misclassifications. For Low urgency, 19 out of 21 messages were correctly identified, with 2 misclassified as Medium. The Medium class again underperformed, with only 5 out of 9 messages correctly classified, 3 misclassified as Low, and 1 as High. The higher accuracy in audio-based testing compared to text-based testing (82% vs. 73%) suggests that audio features, such as tone or speech patterns, may provide more distinguishable cues for urgency classification.

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| High | 0.82 | 0.90 | 0.86 | 10 |
| Low | 0.83 | 0.90 | 0.86 | 21 |
| Medium | 0.83 | 0.56 | 0.67 | 9 |
| accuracy | | | 0.82 | 40 |
| macro avg | 0.83 | 0.79 | 0.80 | 40 |
| weighted avg | 0.83 | 0.82 | 0.82 | 40 |

Figure 7.3: Classification Report for Audio-Based Urgency Testing

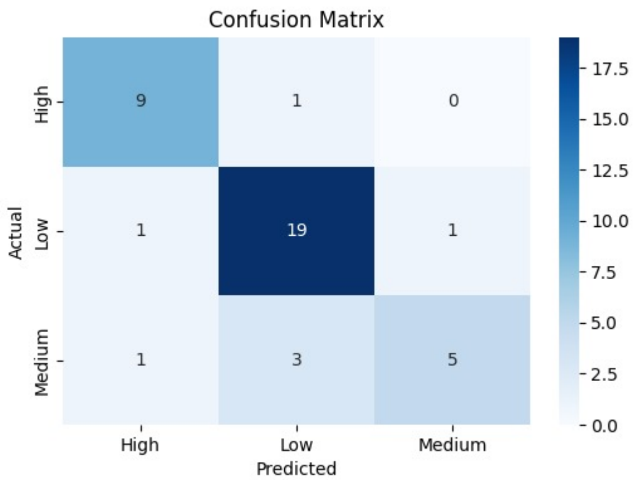


Figure 7.4: Confusion Matrix for Audio-Based Urgency Testing

To understand the contributing factors behind this performance, a Random Forest Feature Importance analysis was conducted on the audio features used in the training data shown in Figure 7.5. The analysis highlights the relative importance of various audio features in the classification model. The most influential feature is **MeanUnvoicedSegmentLength**, with an importance score of approximately 0.25, followed closely by **MeanVoicedSegmentLengthSec** at around 0.22. These features, which measure the average duration of unvoiced and voiced segments in the audio, likely capture critical patterns in speech rhythm and pauses that correlate with urgency levels. Features such as **mfcc2_sma3_amean and mfcc1_sma3_amean (Mel-frequency cepstral coefficients)** also contribute significantly, with importance scores of around 0.15 and 0.12, respectively, indicating that spectral characteristics of the audio play a role in distinguishing urgency. Lower-ranking features include **HNRdBACF_sma3nz_amean (Harmonics-to-noise ratio)**, **jitterLocal_sma3nz_amean (Frequency instability)**, **F0semitoneFrom27.5Hz_sma3nz_amean (Pitch-Related Feature)**, and **loudness_sma3_amean (Loudness  Voice Quality Feature)**, each with importance scores below 0.10. The dominance of segment length features suggests that temporal aspects of speech are more predictive of urgency than spectral or noise-related features in this model. However, the relatively low importance of features like loudness (often associated with urgency in human perception) indicates potential areas for improvement, such as incorporating more dynamic loudness variations or exploring additional features like pitch contour. Enhancing the model by focusing on these underrepresented features could further improve classification accuracy, particularly for the Medium urgency class, which continues to show lower recall.

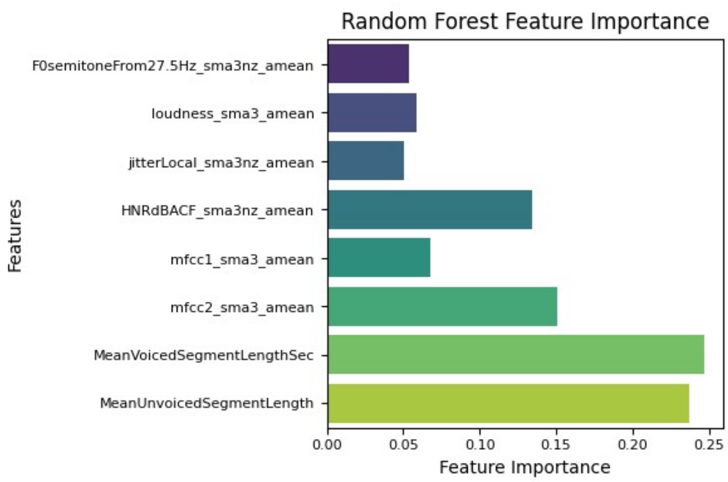


Figure 7.5: Audio Features for Audio-Based Urgency

## 7.3 Test Case 3: Department Routing

The third test evaluated the voicemail system's ability to route messages to the correct department—Doctors, Emergency, Pharmacy, or Reception—based on content analysis. The classification report Figure 7.6 indicates an overall **accuracy of 75%**, with a macro average F1-score of 0.73 and a weighted average F1-score of 0.73. The Doctors department achieved a precision of 0.77, a recall of 0.91, and an F1-score of 0.83 with a support of 11 samples; the Emergency department had a precision of 1.00, a recall of 0.85, and an F1-score of 0.92 with a support of 13 samples; the Pharmacy department recorded a precision of 0.61, a recall of 1.00, and an F1-score of 0.76 with a support of 11 samples; and the Reception department had a precision of 0.67, a recall of 0.31, and an F1-score of 0.42 with a support of 13 samples. The confusion matrix Figure 7.7 provides further insight: out of 128 Doctors messages, 126 were correctly classified, with minimal misclassifications (1 each to Emergency, Pharmacy, and Reception). The Emergency department had 59 out of 64 messages correctly classified, with 2 misclassified as Pharmacy and 3 as Doctors. Pharmacy performed well, with 111 out of 112 messages correctly classified, though 4 Doctors messages were misclassified as Pharmacy. Reception struggled significantly, with only 206 out of 221 messages correctly classified, and 10 misclassified as Doctors, 3 as Pharmacy, and 1 as Emergency. The high recall for Pharmacy (1.00) and Doctors (0.91) indicates strong performance in these categories, but the low recall for Reception (0.31) suggests the system struggles to identify messages intended for this department, possibly due to ambiguous content or overlapping features with other departments. Improving the model's ability to distinguish Reception messages, perhaps through additional training data or more specific keyword mapping, is recommended.

| Classification Report: | precision | recall | f1-score | support |
|---|---|---|---|---|
| Doctors | 0.77 | 0.91 | 0.83 | 11 |
| Emergency | 1.00 | 0.85 | 0.92 | 13 |
| Pharmacy | 0.61 | 1.00 | 0.76 | 11 |
| Reception | 0.67 | 0.31 | 0.42 | 13 |
| accuracy | | | 0.75 | 48 |
| macro avg | 0.76 | 0.77 | 0.73 | 48 |
| weighted avg | 0.77 | 0.75 | 0.73 | 48 |

Figure 7.6: Classification Report for Department Routing Testing

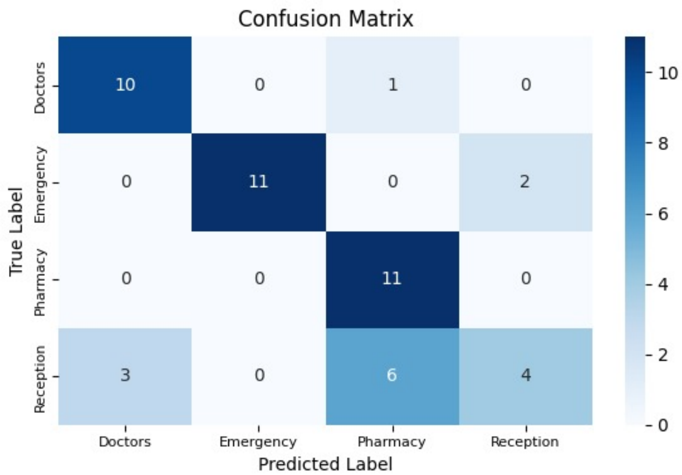


Figure 7.7: Confusion Matrix for Department Routing Testing

## 7.4 Overall Analysis and Recommendations

- Audio-based urgency classification achieved the highest accuracy (82%), followed by department sharing (75%) and text-based urgency classification (73%).
- The system excels in identifying High urgency messages and routing messages to the Doctors and Pharmacy departments, as evidenced by high recall and F1-scores.
- The feature importance analysis for audio-based testing in Figure 7.5 highlights the significance of temporal features like MeanUnvoicedSegmentLength and MeanVoicedSegmentLengthSec, contributing to higher accuracy.

- Consistent challenges emerge with the Medium urgency class in both text and audio tests, where low recall suggests difficulty in differentiation.
- The Reception department's low recall in the department sharing test indicates a need for improved classification logic.

### 7.4.1 Recommendations

- Retrain the urgency classification model with additional Medium urgency samples to improve feature differentiation.
- Incorporate more advanced audio feature extraction techniques, such as dynamic loudness variations or pitch contour, to enhance audio-based classification.
- Refine the department sharing model with more Reception-specific training data and clearer keyword or context-based rules.

These improvements could help the system achieve more consistent and reliable performance across all categories.

# CHAPTER 8

# CONCLUSION AND FUTURE WORK

The AI-powered voicemail prioritization system represents a transformative approach to managing and streamlining voicemail communication within organizations. By integrating advanced tools such as speech-to-text conversion, natural language processing, urgency classification, sentiment analysis, and PII redaction, the system facilitates rapid and efficient responses to high-priority messages. This automated workflow reduces manual effort, enhances message organization, and ensures that voicemails are routed promptly to the appropriate departments. Additionally, the incorporation of privacy-preserving mechanisms such as PII redaction ensures compliance with data security standards while maintaining the confidentiality of sensitive information. In doing so, the system reduces manual workload, organizes messages efficiently, and ensures that voicemails are routed promptly to the appropriate personnel, significantly enhancing organizational communication and operational responsiveness.

Through extensive testing and evaluation, the system has demonstrated promising accuracy in urgency classification, with audio-based classification achieving the highest performance. The feature importance analysis highlighted key attributes influencing the model's decision-making, such as temporal features in speech data. However, challenges remain in classifying medium-urgency messages and improving recall for specific department routing, particularly for the Reception category. Addressing these issues will be crucial for enhancing the reliability and robustness of the system in real-world applications.

## 8.1 Future Work

As the project progresses, several key areas of development and enhancement will be pursued:

- **Multilingual and Dialect Support:** Expand models to handle multiple languages and dialects, ensuring accessibility and effectiveness in diverse hospital environments.

- **Voicemail Summarization:** Extend the summarization feature to voicemail messages, providing concise overviews of lengthy messages for quicker decision-making.

- **Enhanced Sentiment and Emotion Analysis:** Improve the system's sentiment and emotion detection capabilities to better understand caller urgency and tone.

- **Automated AI Responses:** Develop AI-powered automated replies for common inquiries, reducing manual workload and improving response time.

- **Adaptability Across Industries:** Modify the system for different sectors such as healthcare, customer support, banking, and legal services, enabling broader application of voicemail prioritization technology.

- **Model Optimization:** Improve the accuracy of urgency classification by incorporating additional training data, especially for medium-urgency messages, and experimenting with ensemble learning techniques for better prediction stability.

- **Enhanced Audio Processing:** Implement more advanced audio feature extraction techniques, such as pitch contour analysis, dynamic loudness variations, and prosody-based sentiment detection, to refine the accuracy of voice-based urgency classification.

- **Integration of Contextual Understanding:** Enhance the NLP capabilities to better understand the context of messages, improving routing efficiency and minimizing misclassification.

- **Real-Time Processing and Deployment:** Optimize the system for real-time voicemail processing, ensuring minimal latency in classification and forwarding. Deployment on cloud-based or edge-computing platforms will be explored for scalability.

- **User Feedback and Adaptability:** Conduct user testing and collect feedback to refine the system's usability and effectiveness. Adaptive learning techniques could be integrated to allow the system to improve over time based on real-world interactions.

By focusing on these areas, the AI-powered voicemail prioritization system will continue to evolve, delivering more efficient, accurate, and secure voicemail management solutions. Future iterations will emphasize greater accuracy, enhanced privacy measures, and seamless integration into existing communication frameworks, ensuring that organizations benefit from a reliable and intelligent voicemail processing system.

# REFERENCES


[1] Trabelsi, A., Soussilane, S., Helbert, E. (2023). Voicemail Urgency Detection Using Context Dependent and Independent NLP Techniques (pp. 450–456). https://doi.org/10.5220/0011685800003393.

[2] Koumpis, K., Renals, S., Dept. of Computer Science, University of Sheffield. (2001). TRANSCRIPTION AND SUMMARIZATION OF VOICEMAIL SPEECH.

[3] Kamiyama, H., Ando, A., Masumura, R., Kobashikawa, S., Aono, Y., NTT Corporation, NTT Media Intelligence Laboratories. (2021). Urgent Voicemail Detection Focused on Long-term Temporal Variation.

[4] Jaikumar, J., Mohana, N., Suresh, P. (2023). Privacy-Preserving Personal Identifiable Information (PII) Label Detection Using Machine Learning. https://doi.org/10.1109/icccnt56998.2023.10307924.

[5] A replicable comparison study of NER software: StanfordNLP, NLTK, OpenNLP, SPACY, GATE. (2019). [Conference-proceeding].

[6] Gupta, A., Chugh, D., Anjum, and Katarya, R. "Automated News Summarization Using Transformers." Proceedings of the 23rd International Conference on Machine Learning, Delhi Technological University, New Delhi, India, 2021.

[7] Dr Python. (2023, October 16). How To Encrypt Text Files In Python - MINI PROJECT2023 [Video]. YouTube.https://www.youtube.com/watchv=1JtfxOZYkeM.

[8] AK Python. (2021, April 11). Voice based sentiment analysis using python . [Video]. YouTube.https://www.youtube.com/watch?v=euOw1DGqEpo.

[9] Beesly AI — Revolutionize Your Voicemail with AI-Powered Assistant. (n.d.). https://www.beesly.ai/

[10] Audeering. (n.d.). opensmile/config/egemaps/v01a at master · audeering/opensmile. GitHub. https://github.com/audeering/opensmile/tree/master/config/egemaps/v01a

[11] Build software better, together. (n.d.). GitHub. https://github.com/topics/speech-to-text