7th April 2025

AI POWERED VOICE MAIL SYSTEM

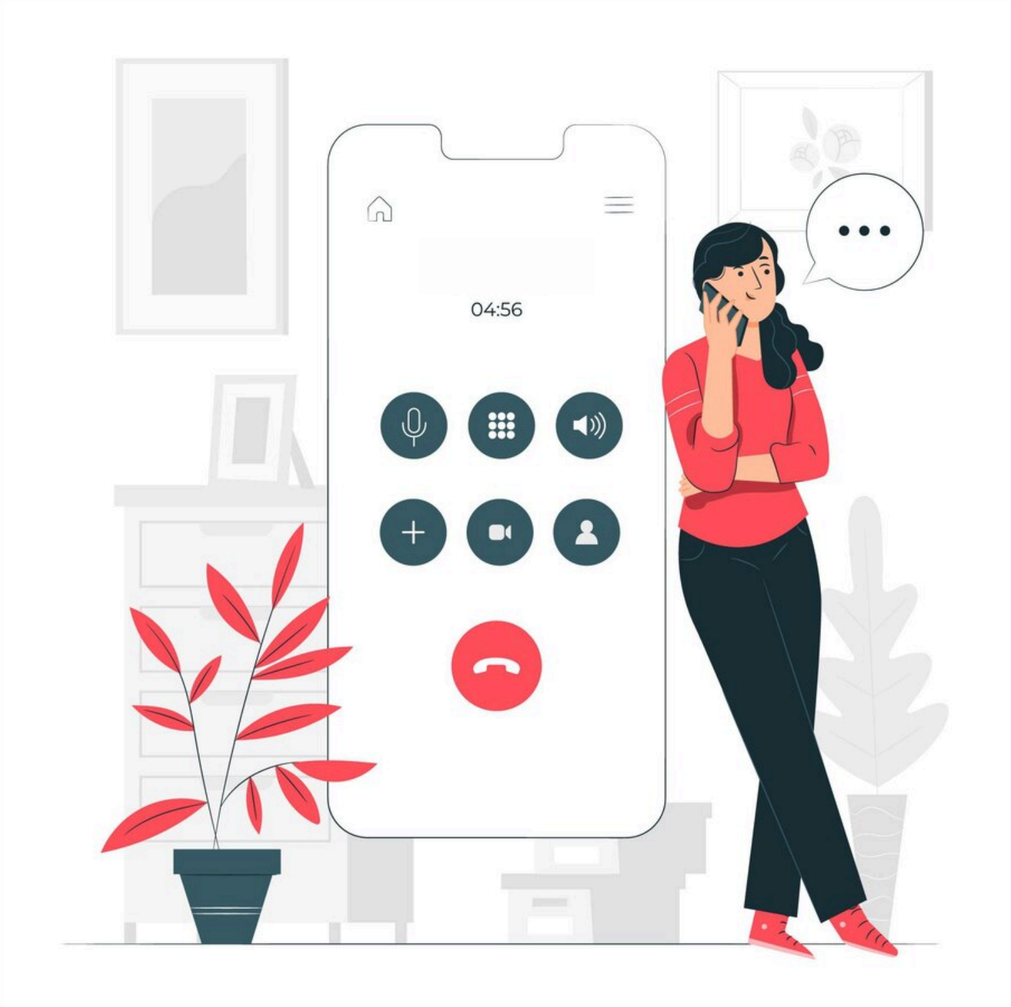
Final Review - Phase 2

GROUP - 2

FATHIMA NADWA RAMEEZ - MES21AD019
MOHAMMED SINAN - MES21AD040
MUHAMMED AJAS - MES21AD041
MUHAMMED SHAHIQ - MES21AD046

Mr. ABIN C JOSE
ASSISTANT PROFESSOR
DEPARTMENT OF ADS
MES COLLEGE OF ENGINEERING

CONTENTS



INTRODUCTION

- In fast-paced industries like healthcare, finance, and customer service, quick and efficient communication is critical.
- Traditional voicemail systems often result in missed or delayed responses, especially for urgent matters.
- This inefficiency can lead to serious consequences, from lost business opportunities to compromised patient care.
- Our project aims to solves this by intelligently prioritizing voicemails based on urgency, tone, and key words, ensuring that critical messages are handled first.

O B J E C T I V E S

- **Voice Analysis & Prioritization:** AI analyzes tone and content to rank messages by urgency.
- **Voice-to-Text Transcription:** Voicemails are converted to text for faster review.
- **Personally Identifiable Information (PII) encryption:** Protects sensitive information, making it secure for industries handling confidential data.
- **Automated Forwarding:** Ensures urgent messages reach the right team or person immediately.

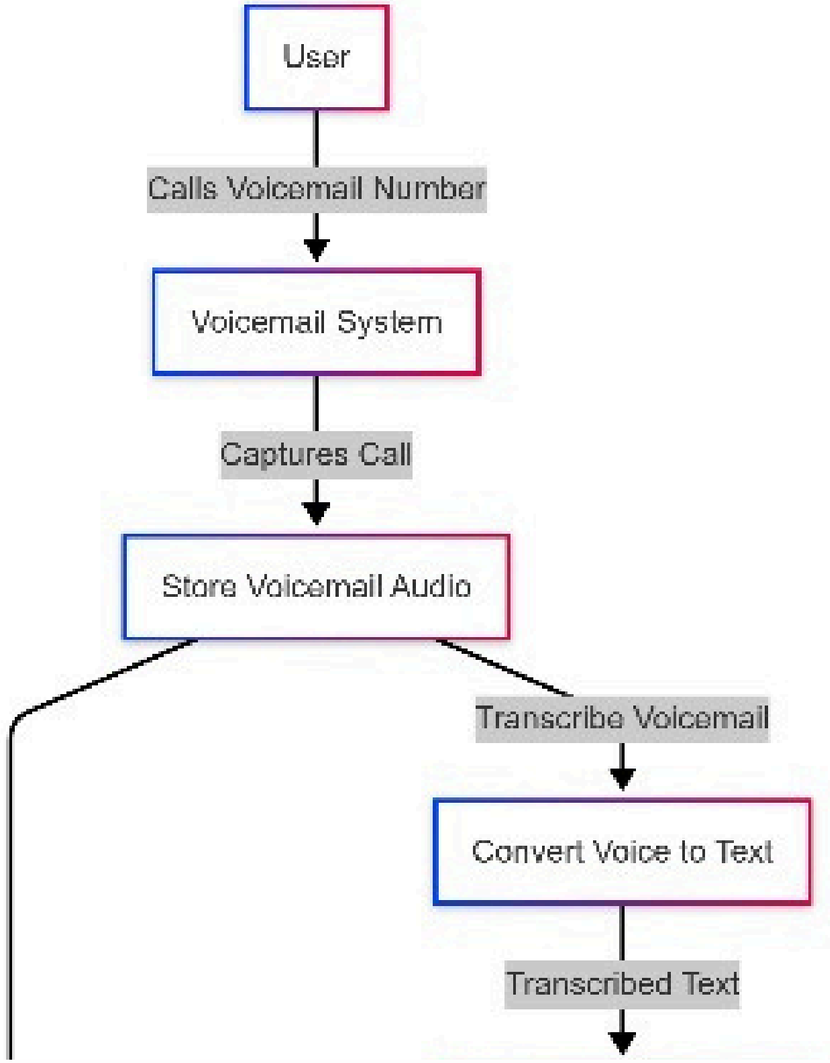
LITERATURE REVIEW

TITLE	MODEL USED	KEYPOINTS
1. Voicemail Urgency Detection Using Context Dependent and Independent NLP Techniques	BERT + SVM, TF-IDF + SVM	Classify voicemail as urgent or non-urgent using NLP techniques TF-IDF struggles with unseen data, poor generalization.
2. Transcription and Summarization of Voicemail Speech	HMM + MLP (Hybrid Approach), NER for key term extraction	Transcribe and summarize voicemail messages using hybrid techniques. High computation, transcription errors, summarization error rate.

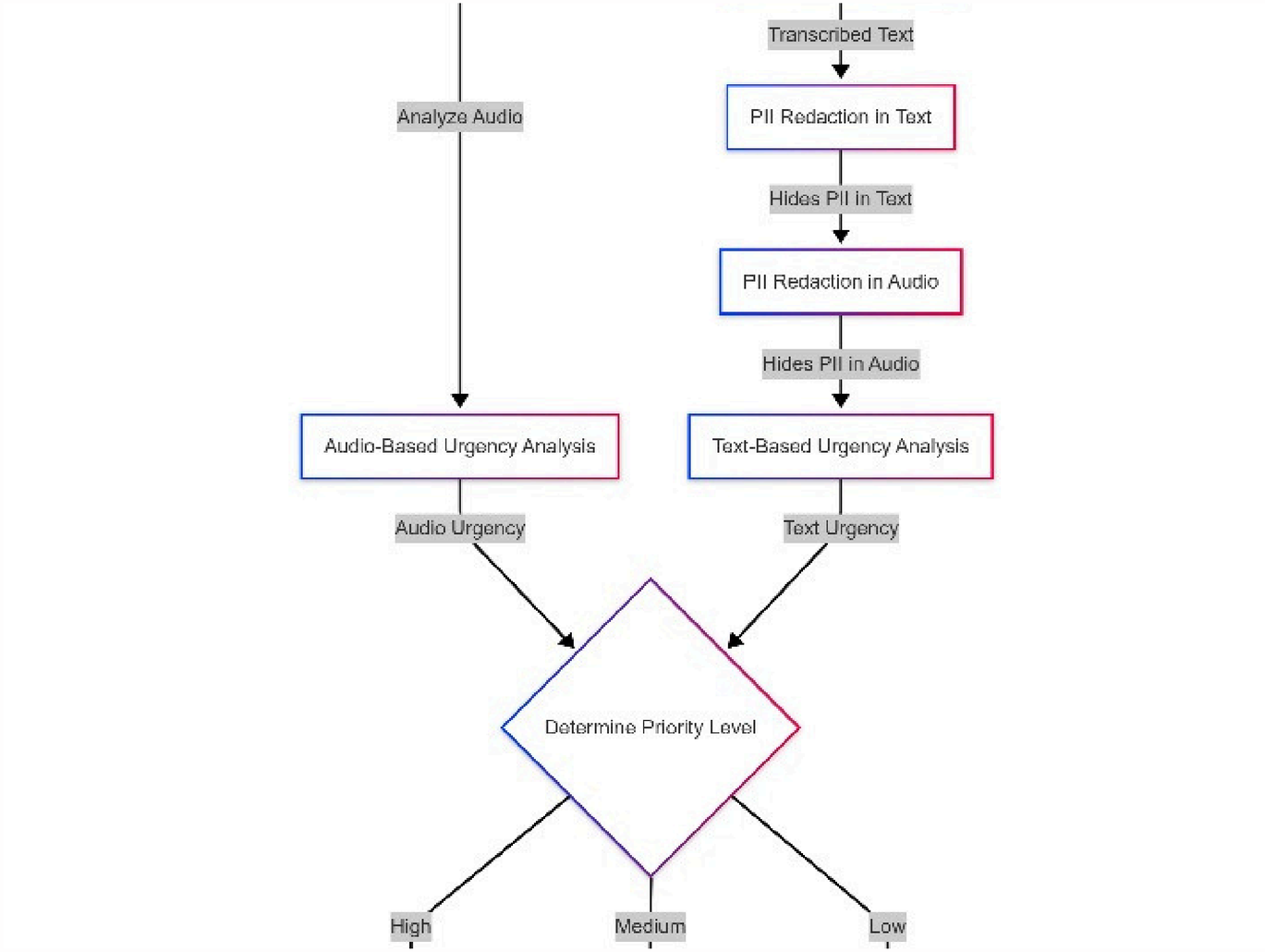
LITERATURE REVIEW

TITLE	MODEL USED	KEYPOINTS
3. Privacy-Preserving PII Label Detection Using Machine Learning	TF-IDF + Random Forest, SVM	Detect and label personal identifiable info (PII) in documents. Class imbalance, data dependency, high resource demand.
4. Urgent Voicemail Detection Focused on Long-term Temporal Variation	MFCC + EMS for feature extraction, RNN with Attention	Use vocal features and rhythm to classify voicemail urgency. High computational cost, requires clean high-quality data.

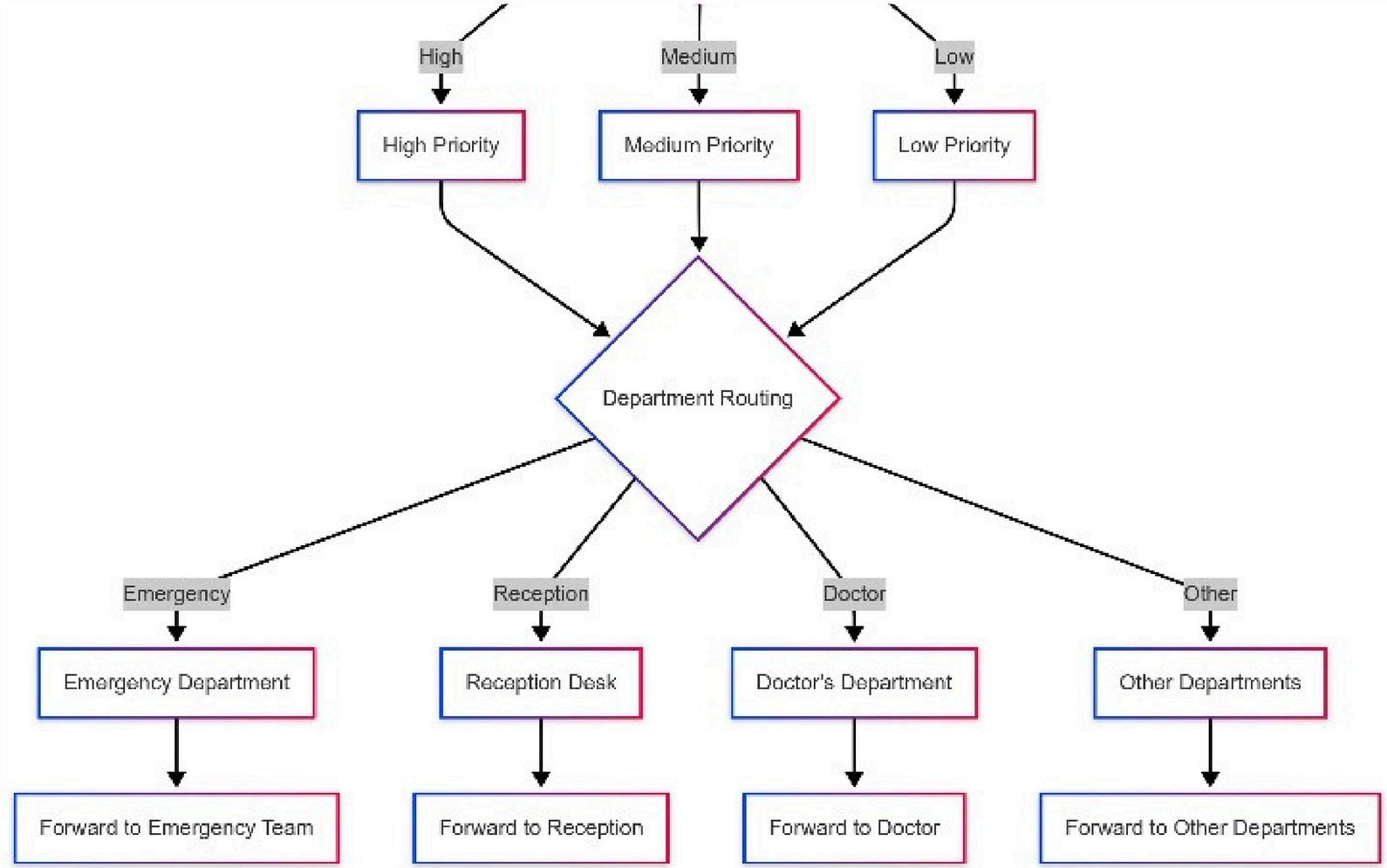
ARCHITECTURE



ARCHITECTURE



ARCHITECTURE



PROPOSED SYSTEM

While this system can be applied in various fields, a hospital setting was chosen for prototyping because

- **Ease of Implementation as a Prototype** – it provides a controlled and structured environment to test voicemail ranking based on urgency.
- **Simpler Data Collection and Labeling** – Creating a dataset from scratch is easier in a hospital since healthcare calls often have well-defined categories (e.g., emergency, phramcy, general inquiries)
- **Ease of Understanding** –The difference between high-priority and low-priority messages is clearer, allowing for better demonstration and evaluation of the system.

PROPOSED SYSTEM

1. **Voice Capture**
   - Tool: Use **Twilio** to capture voicemails.
   - Implementation: Set up a Twilio number to receive voicemails. Use the Twilio API to access these recordings.
2. **Speech-to-Text Conversion**
   - Tool: **SpeechRecognition.**
   - Implementation: This libraries will return the text version of the voicemail.

PROPOSED SYSTEM CONTD…

3. PII Detection and Redaction:

- Tool: Use a library like **SpaCy** for natural language processing.
- After obtaining the text, scan for PII (like Addresses, Phone Numbers) using SpaCy. Replace or redact these with placeholders.

4. Voicemail Prioritization:

- Analyze Tone: Use the **OpenSMILE** library to analyze the captured audio for urgency indicators, such as pitch, speed and loudness.

PROPOSED SYSTEM CONTD…

- Analyze Text for Urgency: Use **SpaCy** to scan the text for keywords **"urgent," "ASAP," "emergency," or "immediately"** adds the urgency score.
- Combine tone and text scores using a weighted sum or average.
- **Random Forest** is used to classify voicemails by training a model on tune .

PROPOSED SYSTEM CONTD…

5. Routing to Departments

- The prioritized voicemails are routed to relevant departments based on the content.

6. Frontend Interface

- Basic interface where users can see the voicemails, their transcription, and urgency levels for different users(Admin, Department, etc).

DATASETS

1. Text Classification and Department forwarding

- Contains transcriptions and their corresponding department labels.
- Used to train the AI model for Text based Prioritization and forwarding them to relevant departments.

Um, hi there... I've been having this really bad headache for like... uh... three days now and it's getting worse. Just wanted to see if I could get an appointment? My number is 555-0123	General	Medium
Hello... *cough* *sniff* I think I might have the flu or something? Been feeling really awful and my fever is... um... pretty high. Should I come in? Call me back at 555-0234	General	medium
Hi, just calling to... uh... schedule my annual check-up. No rush, whenever you have availability. You can reach me at 555-0345	General	low
Hey, I've been having some really bad stomach pains since... um... last night. Not sure if I should be worried? Call me at 555-0456	General	medium
Good morning... er... I need to update my contact information and... um... get a new insurance card issued. My number is 555-0567	General	low
Hi there... I've been feeling really dizzy for the past few days and... uh... just need some advice on what to do. Call me at 555-0678	General	medium
Hello... *clearing throat* just need to... um... get a copy of my vaccination records for work. No rush. 555-0789	General	low
Hi... been having these weird spots appear on my skin and... uh... they're kind of itchy. Should I come in? 555-0890	General	medium
Yes, hello... need to schedule a... um... follow-up appointment for my blood pressure check. Call when you can at 555-0901	General	low

DATASETS CONTD...

2. Generated Audio Data

- Contains voicemail recordings for tone-based urgency analysis.

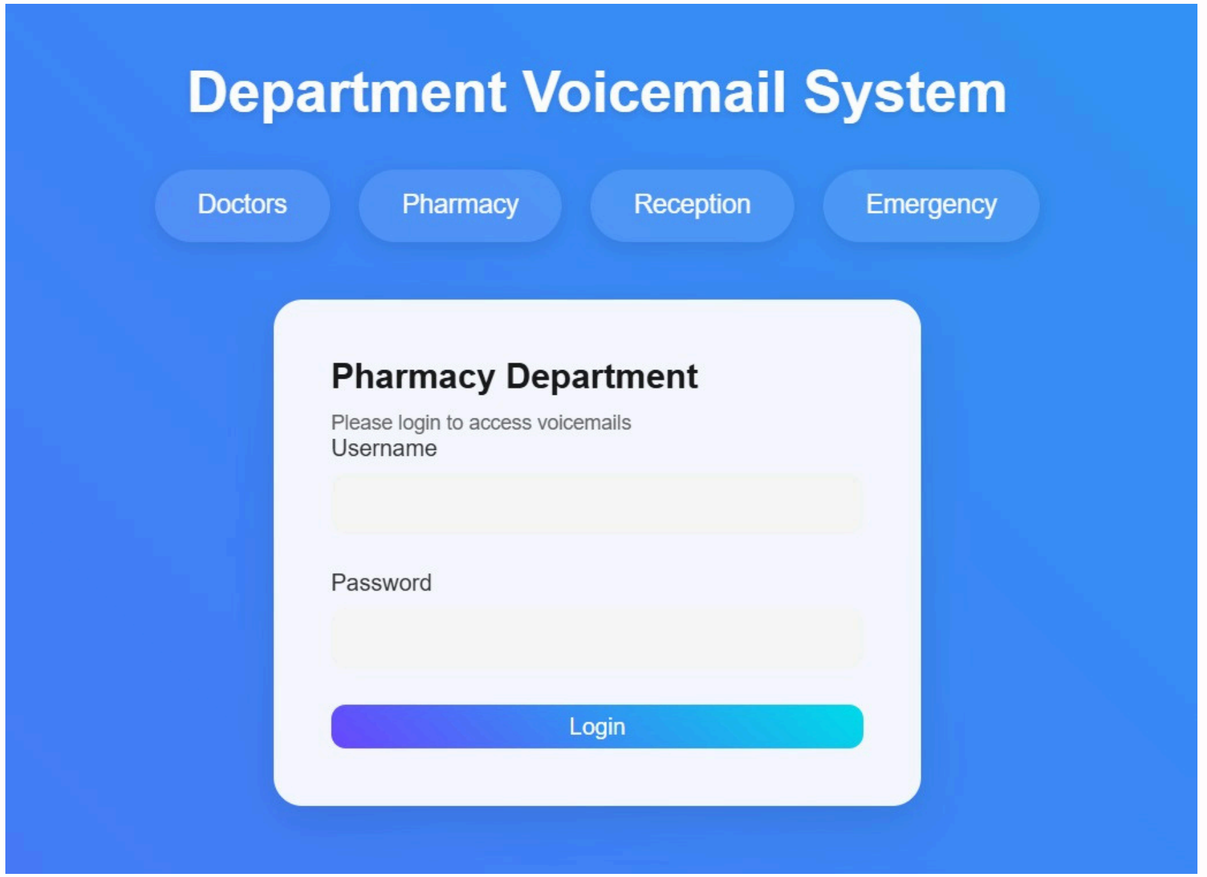
	A	B	C	D	E	F	G	H	I
1	F0semitoneFrom27.5H	loudness_sma3_amean	jitterLocal_sma3nz_ar	HNRdBACF_sma3nz_amean	mfcc1_sma3_amean	mfcc2_sma3_amean	MeanVoicedSegmentLengt	MeanUnvoicedSeg	Urgency
2	39.64327779	0.927096859	0.048211304	8.599468912	27.35936976	-13.91805545	0.503572968	0.233534872	Low
3	25.29886177	0.54672907	0.075266178	3.068153029	24.06722851	-14.76167143	0.260921525	0.05844463	Medium
4	29.66772566	1.01692792	0.053517742	3.405510837	26.89951639	-8.302402133	0.426069334	0.129215364	Low
5	25.22528155	0.418497133	0.050720568	2.660319844	25.93674745	-13.43548623	0.431344624	0.294173347	Low
6	38.69649393	0.663529096	0.038205642	4.565718309	16.70167168	-3.786638487	0.440432961	0.197851957	Low
7	36.74439959	0.593823305	0.080426803	8.970389321	14.43754724	-1.925077369	0.444691823	0.225305646	Low
8	30.84065265	1.939830841	0.082649834	5.491361035	27.06045864	-12.85278018	0.58813314	0.284444456	Low
9	38.29574607	1.607794969	0.05524527	4.130220209	27.33036516	-9.055770808	0.192338465	0.071668581	High
10	39.21954947	0.910055968	0.047839351	3.225319454	18.82098901	7.636671678	0.257821813	0.17743445	High
11	35.50293062	1.353974704	0.085451199	1.102374654	15.83480828	-4.649370663	0.536903567	0.165284917	Low

F R O N T E N D

1. Framework & Technology

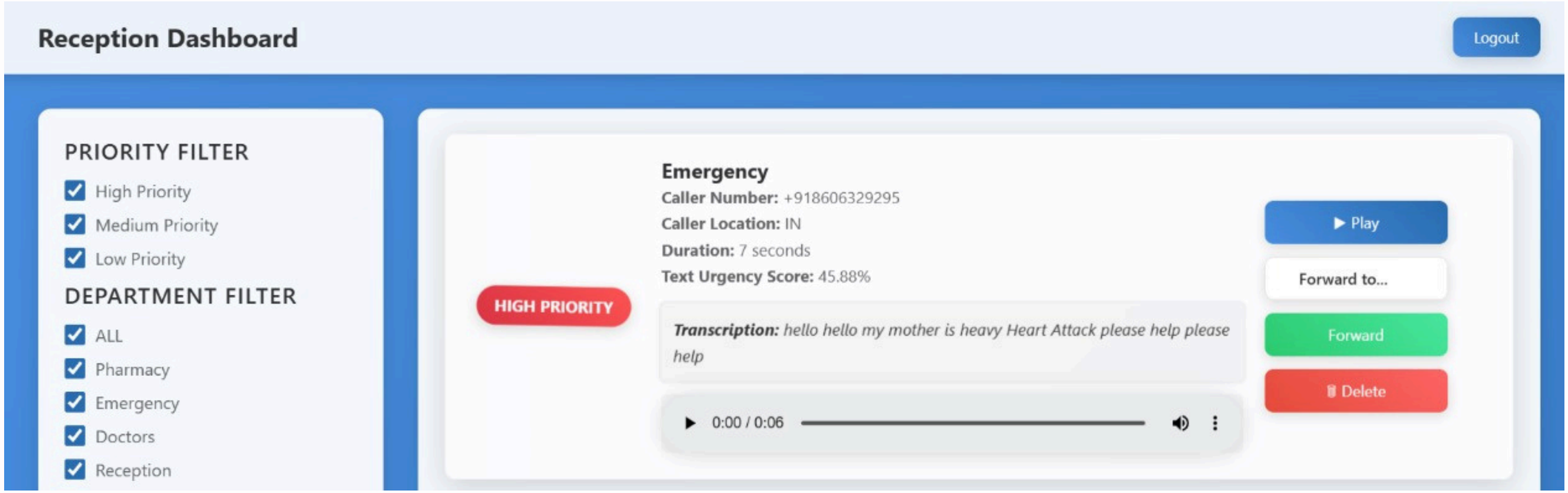
- Flask (Python-based web framework)
- HTML + Jinja2 (For dynamic rendering of voicemail data)
- CSS & JavaScript (For styling and interactivity)

FRONTEND CONTD...



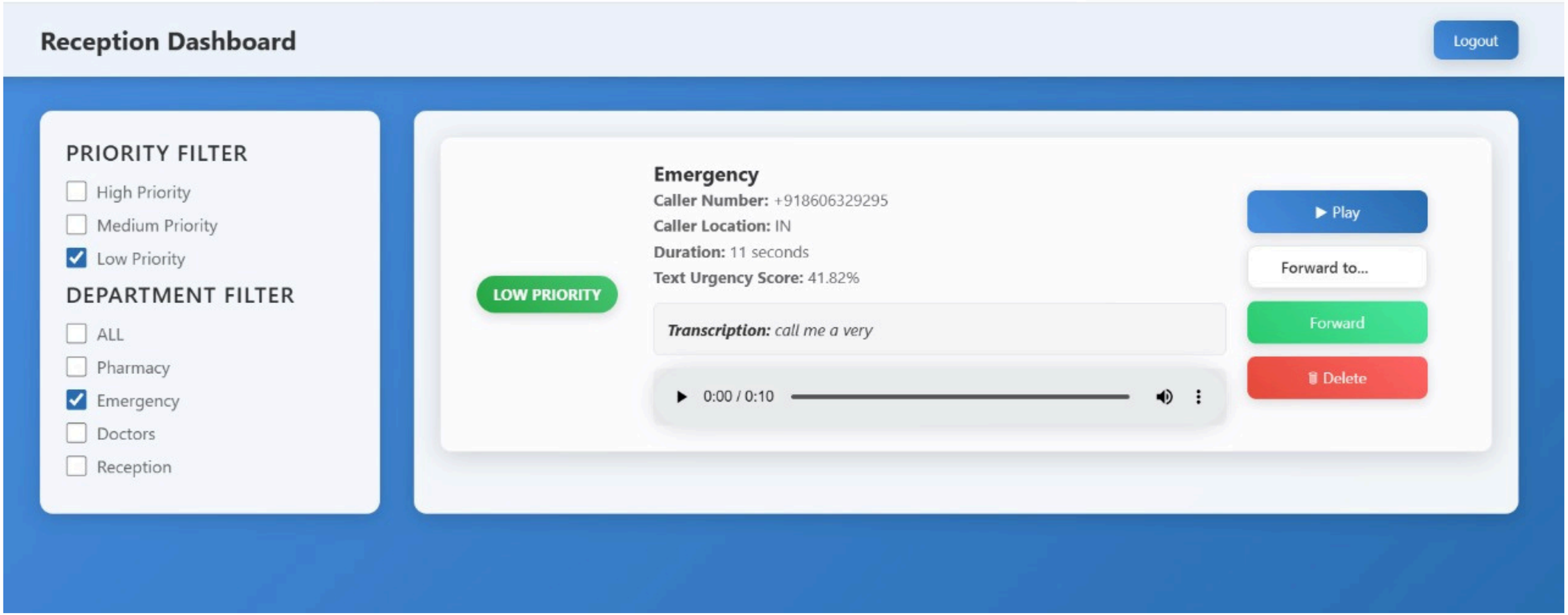
FRONTEND CONTD...

Filter Box



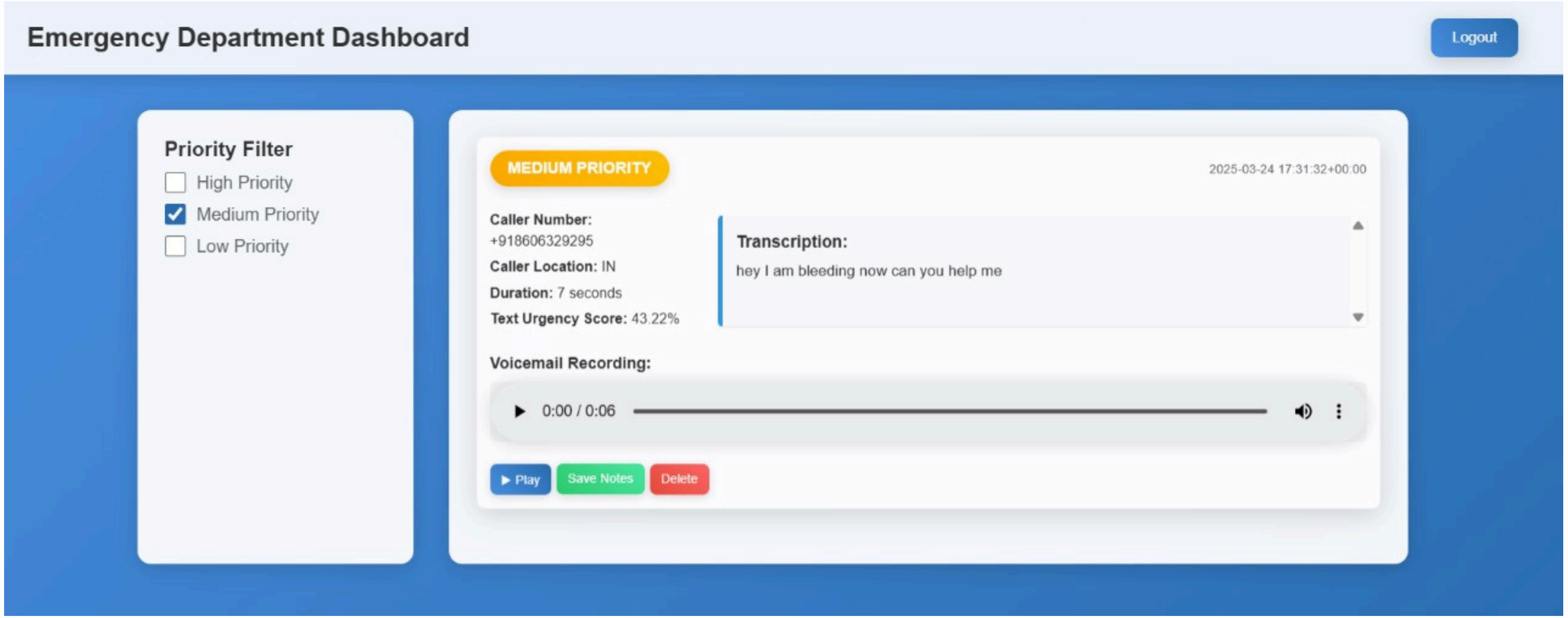
FRONTEND CONTD...

Filtering by Low Priority



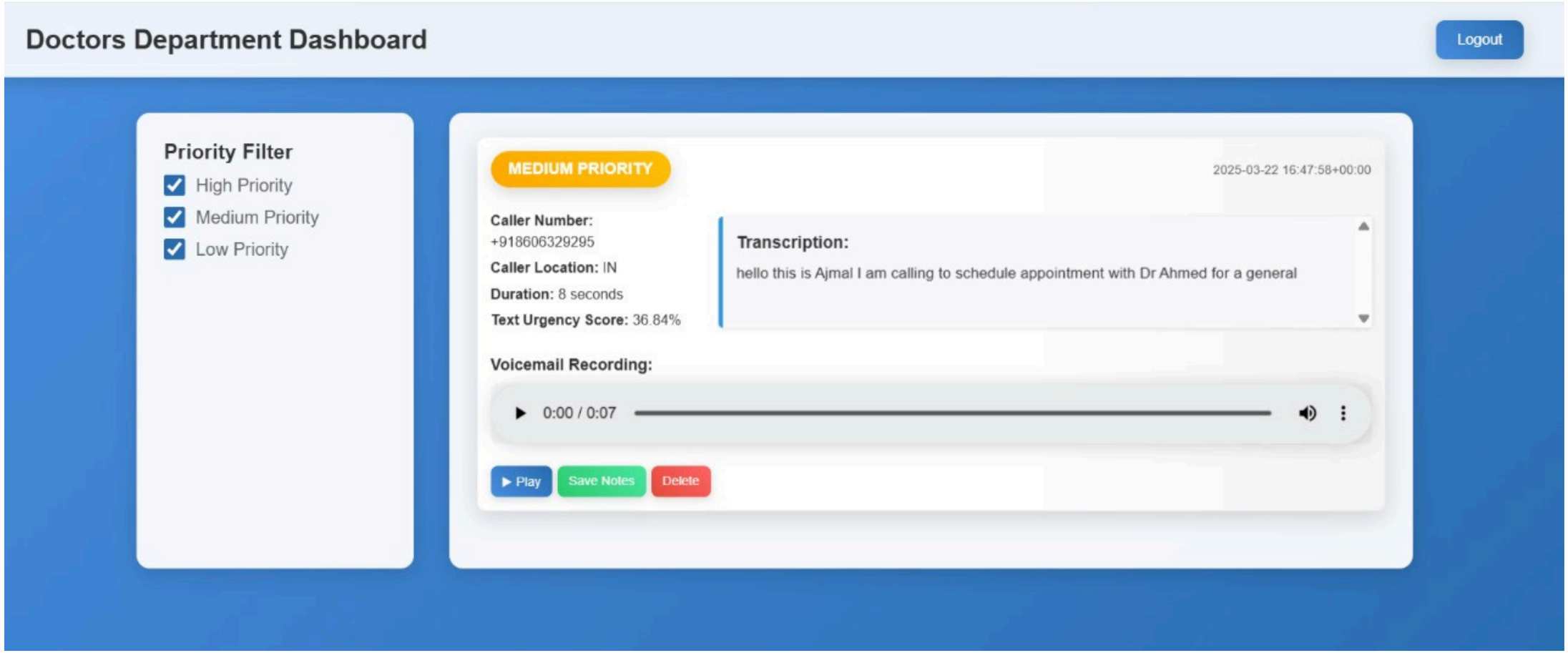
FRONTEND CONTD...

Filtering by Medium Priority



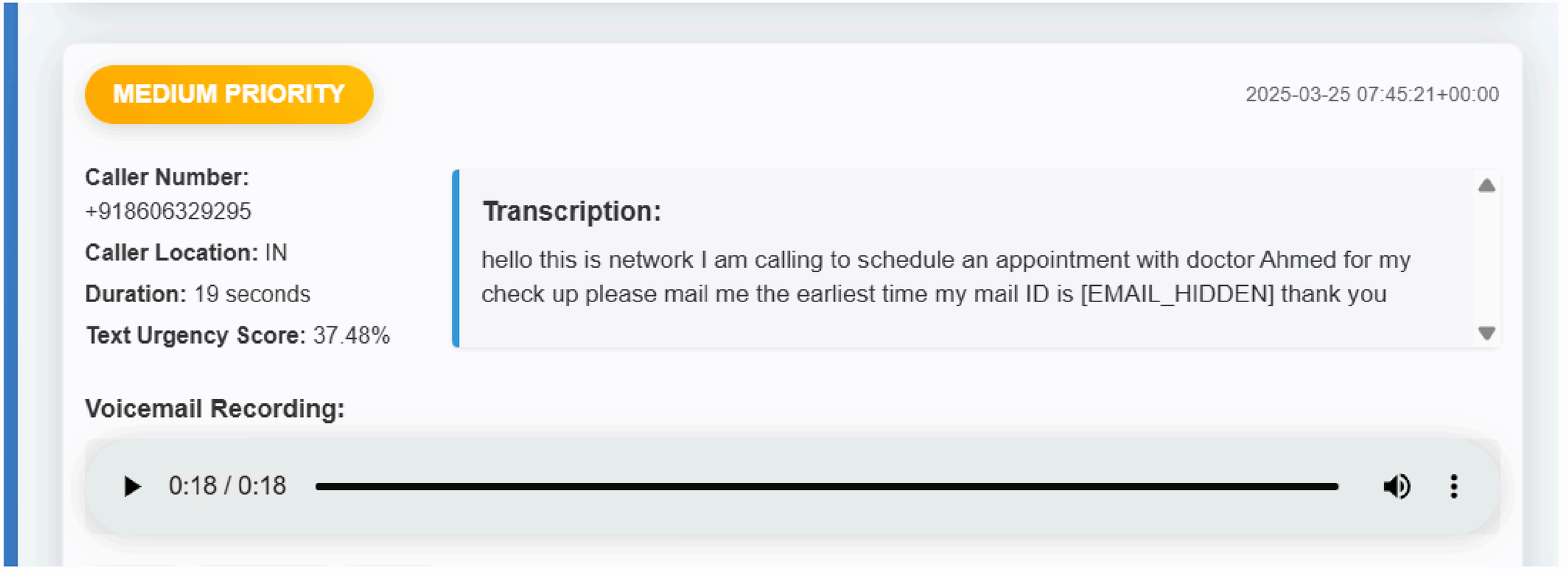
FRONTEND CONTD...

Doctors Dashboard



FRONTEND CONTD...

PII Data Redaction



B A C K E N D

1. Overview of the Backend System

- A web app for healthcare to manage voicemails using Flask and Twilio.
- Grabs voicemail audio from Twilio and saves caller info like phone number and location.
- Transcribes voicemails, rates their urgency, and assigns them to departments.
- Stores all info (caller details, transcription, priority) in a CSV file.
- Shows prioritized voicemails on web pages for different departments, with sorting.

BACKEND CONTD...

2. Prioritization (TEXT)

Converts audio to text and determines its urgency and priority using a trained model.

- Transcribes an audio file into text using **Google's speech recognition**.
- Labels text as **High, Medium, or Low priority** based on the model's prediction.
- Provides a numerical score (0.1 to 1.0) showing how confident the model is in its prediction.
- Higher scores (e.g., 0.9) indicate stronger urgency, while lower scores (e.g., 0.1) suggest less urgency.

BACKEND CONTD...

Text based Prio Model

- **Model: BERT (Bidirectional Encoder Representations from Transformers)**.
- It's fine-tuned for single-label text classification to predict urgency/priority of transcribed voicemail text.
- Trained with a best accuracy of 85%.

B A C K E N D C O N T D . . .

3. Prioritization (AUDIO)

- Extracts **audio features (e.g., pitch, loudness)** from voicemail recordings using OpenSMILE (eGeMAPS v02).
- Saves the extracted features in a CSV file.
- Trains a **Random Forest model** using the extracted features.
- The trained model is saved as random_forest_model.pkl for future predictions.
- Predicts urgency (High, Medium, Low).
- Saves results to CSV and returns urgency level.

BACKEND CONTD...

4. Prioritization(FINAL)

- Takes text urgency (from text analysis) and audio urgency (from audio features) as inputs.
- Converts labels to numbers using a dictionary: **"High" = 3, "Medium" = 2, "Low" = 1.**
- Applies a formula (**0.5 * text_value + 0.5 * audio_value)** where text and audio has same influence **(50%).**
- Calculates the result and rounds it to the **nearest number (1, 2, or 3) to determine the final level.**
- Maps the rounded number back to a label (e.g., 3 → "High") and returns the final urgency.

BACKEND CONTD...

5. Department Classifier

- Uses a pre-trained BERT model fine-tuned to classify text into departments (e.g., Emergency, Pharmacy).
- Converts voicemail text into tokens (max length: 128) using a BERT tokenizer.
- Feeds tokenized text into the model, which outputs scores for each department; the highest score determines the result.
- Takes the predicted number and matches it to a department label.

B A C K E N D C O N T D . . .

6. Security & PII Handling

- Loads spaCy model to detect sensitive info in text.
- Hides emails (e.g., name@domain.com) as [EMAIL_HIDDEN].
- Masks phone numbers (e.g., 8848979016) as [PHONE_HIDDEN].
- Replaces sensitive audio with beep sounds.

RESULT ANALYSIS

- Successfully integrated **Twilio** for voicemail capture.

Call SID and Date	Status	Direction	From	To	Call Type	Duration	STIR Status
CAe0aaa997386b46c2ceaaed630179d2a9 2025-01-08T10:18:11.000Z	Completed	Incoming	+12313106017	+12313106017	Phone	42 Secs	C
CAb863813a6990320f9abb7eece98eb4bd 2025-01-08T10:18:11.000Z	Completed	Outgoing Dial	+12313106017	+12313106017	Phone	42 Secs	C
CA28e2ac85c3b7db6618b7d03a965c81da 2025-01-08T10:17:45.000Z	Completed	Incoming	+12313106017	+12313106017	Phone	14 Secs	C
CA2288ec55e181a54b2cdeec7062a43a25 2025-01-08T10:17:45.000Z	Completed	Outgoing Dial	+12313106017	+12313106017	Phone	14 Secs	C
CA4a3270e2e3a341306943d8ac46155067 2025-01-08T10:16:26.000Z	Completed	Incoming	+12313106017	+12313106017	Phone	11 Secs	C
CA2b170f1f2de0c9289b88fba894648d3e 2025-01-08T10:16:26.000Z	Completed	Outgoing Dial	+12313106017	+12313106017	Phone	11 Secs	C

RESULT ANALYSIS CONTD...

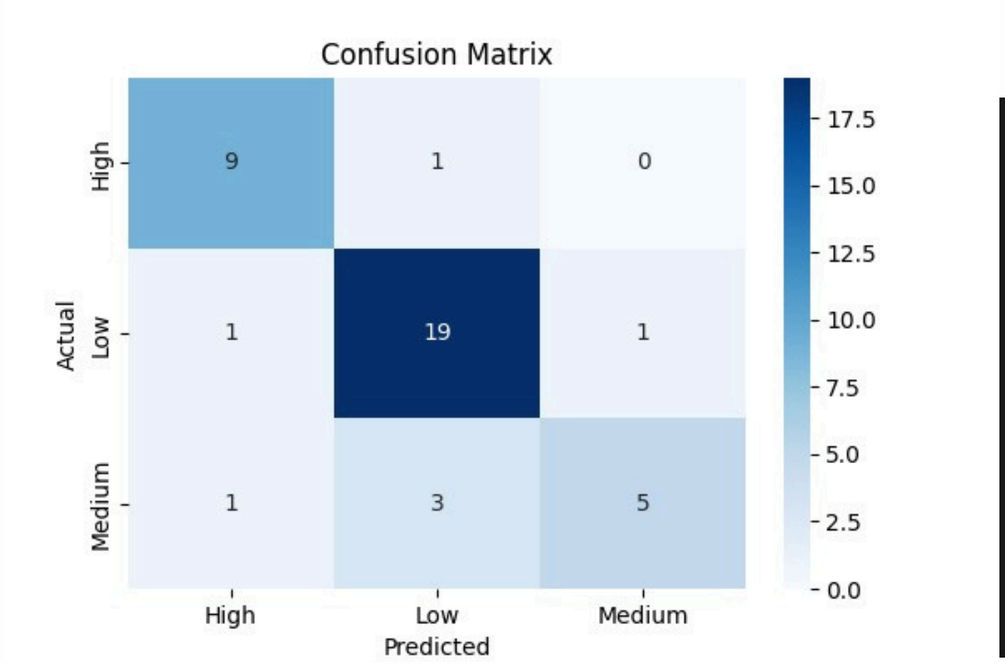
- Stores details of received voicemails for processing, prioritization, and department classification.
- **Caller number, location, duration, transcription, priority, department, etc.**

	A	B	C	D	E	F	G	H	I	J	K	L
1	Recording SID	Caller Number	Caller Location	Date and Time	Duration (Seconds)	Transcription	Priority Label	Text Urgency	Department	Audio Urgency	Final Priority	Audio Path
2	RE e816f9b81af6	9.18606E+11	IN	22-03-2025 22:19	7	hello I am not a	High	0.372774601	Emergency	Medium	High	voicemails/voicemail_RE e816f9b81af6961dcf41a796655b5c78
3	RE06a8c99a5b74	9.18606E+11	IN	22-03-2025 22:18	8	hello my friend	High	0.379073739	Reception	Low	Medium	voicemails/voicemail_RE06a8c99a5b74eb242976b25f919f133
4	RE786e858c0b1b	9.18606E+11	IN	22-03-2025 22:18	9	hello I need a g	High	0.406743497	Reception	High	High	voicemails/voicemail_RE786e858c0b1b7f7d180f5de028f7c24
5	REedfa8013a6fc	9.18606E+11	IN	22-03-2025 22:17	8	hello this is Ajn	Medium	0.368352503	Reception	Medium	Medium	voicemails/voicemail_REedfa8013a6fc115104b3c37e90db38c
6	RE278e875fde72	9.18606E+11	IN	19-03-2025 13:15	8	hello help help	High	0.505383551	Emergency	Low	Medium	voicemails/voicemail_RE278e875fde729227cb1680dcebd086a
7	RE91dbd16cbd7	9.18606E+11	IN	19-03-2025 13:13	13	hello hello site	High	0.445372224	Emergency	High	High	voicemails/voicemail_RE91dbd16cbd7958885e4777095900cb
8	RE24b9ee02feba	9.18606E+11	IN	19-03-2025 13:06	7	hello hello my	High	0.458846807	Emergency	High	High	voicemails/voicemail_RE24b9ee02febabe61e5e46d169e9fbd
9	RE242c8fa26068	9.18606E+11	IN	19-03-2025 12:06	11	call me a very	Low	0.418185174	Emergency	Low	Low	voicemails/voicemail_RE242c8fa26068286913fc55a4cc994c8c.
10												

RESULT ANALYSIS CONTD...

Audio based urgency

- Overall accuracy of **82%** and strong performance on High and Low classes (both at 0.86 F1-score).
- the model struggles to detect Medium cases (0.56 recall).



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

RESULT ANALYSIS CONTD...

- Random Forest feature importance highlights **MeanUnvoicedSegmentLength** and **MeanVoicedSegmentLengthSec** as the most influential features for voicemail classification, with values around 0.25.

Pitch related feature

Loudness and Quality

Frequency Instability

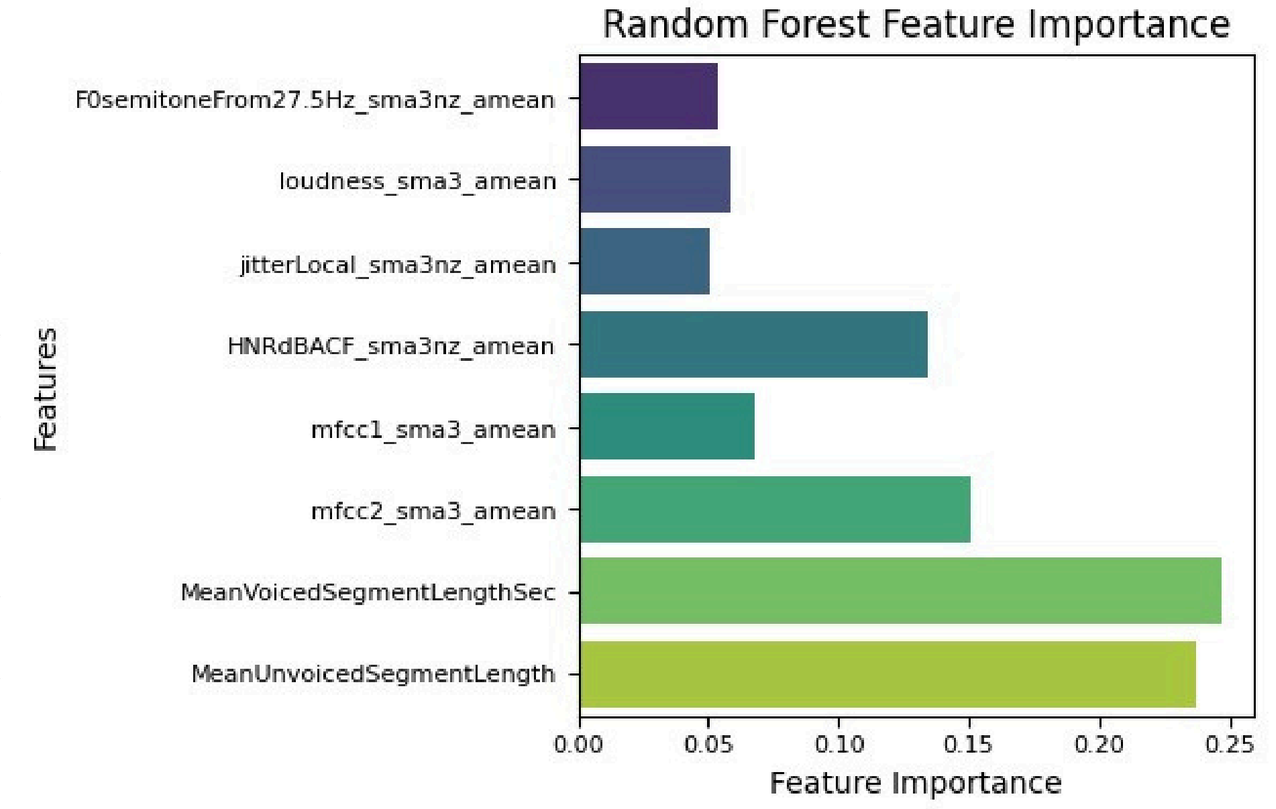
Clarity

Spectral Property

Speech Variability

Duration of Voiced Speech

Duration of Unvoiced Speech



RESULT ANALYSIS CONTD...

Department Classification

- Voicemail system shows mixed performance: Pharmacy and Emergency are accurate (11 correct each), but Reception struggles (3 misclassified as Doctors, 6 as Pharmacy) in a test dataset.

Predictions for new voicemails:

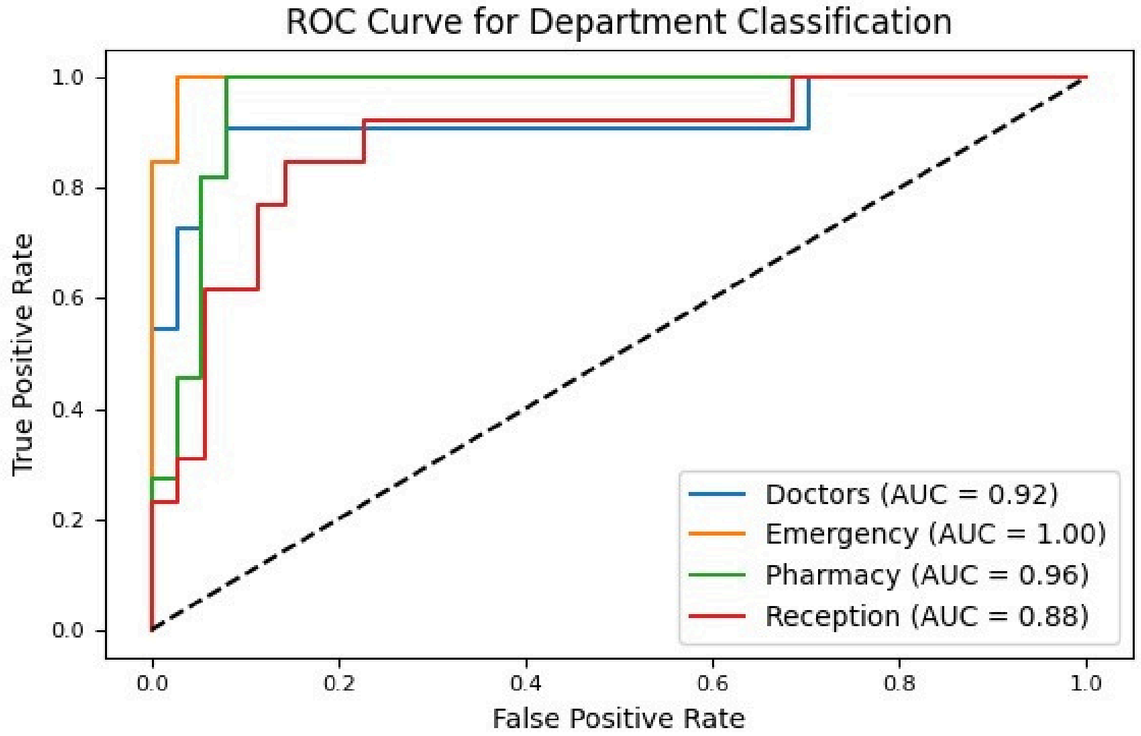
	Text	Predicted_Urgency	Predicted_Department
0	Patient reporting severe chest pain in emergen...	High	Doctors
1	Calling to schedule a routine checkup next week	Medium	Emergency
2	The pharmacy is running low on critical medica...	High	Pharmacy

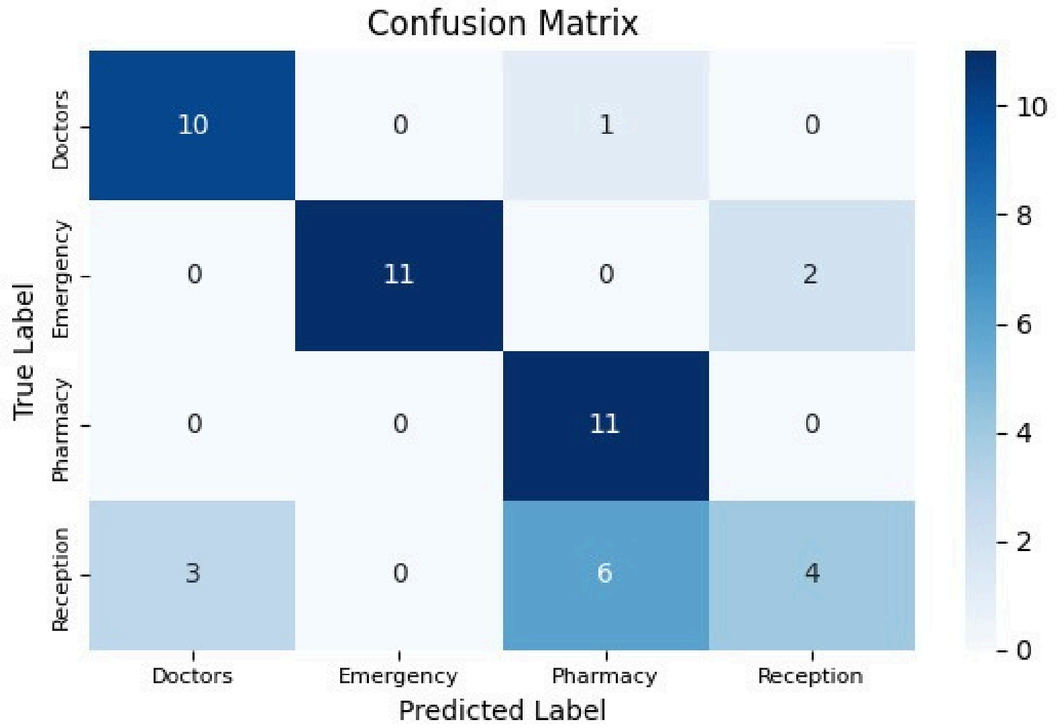
Classification Report:

	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

RESULT ANALYSIS CONTD...

- ROC curve shows strong department classification: Emergency is perfect (AUC 1.00), Pharmacy (0.96) and Doctors (0.92) perform well, while Reception (0.88) has the lowest performance.

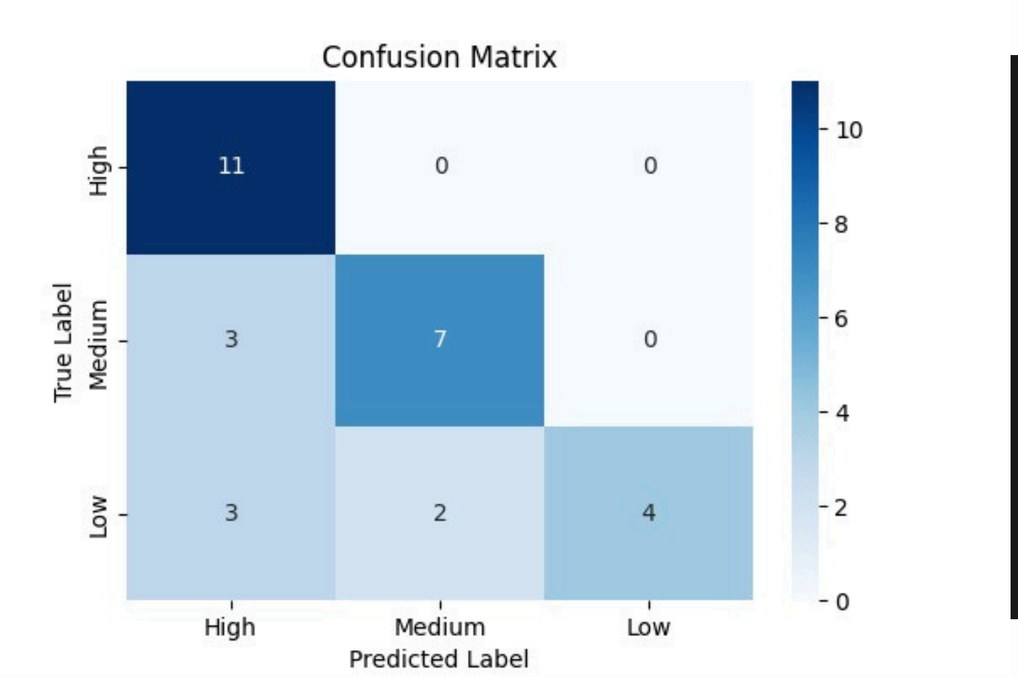




RESULT ANALYSIS CONTD...

Text Urgency

- The rport show that while the model perfectly identifies High cases and performs moderately overall **(73% accuracy),** it struggles to detect many Medium and low cases.

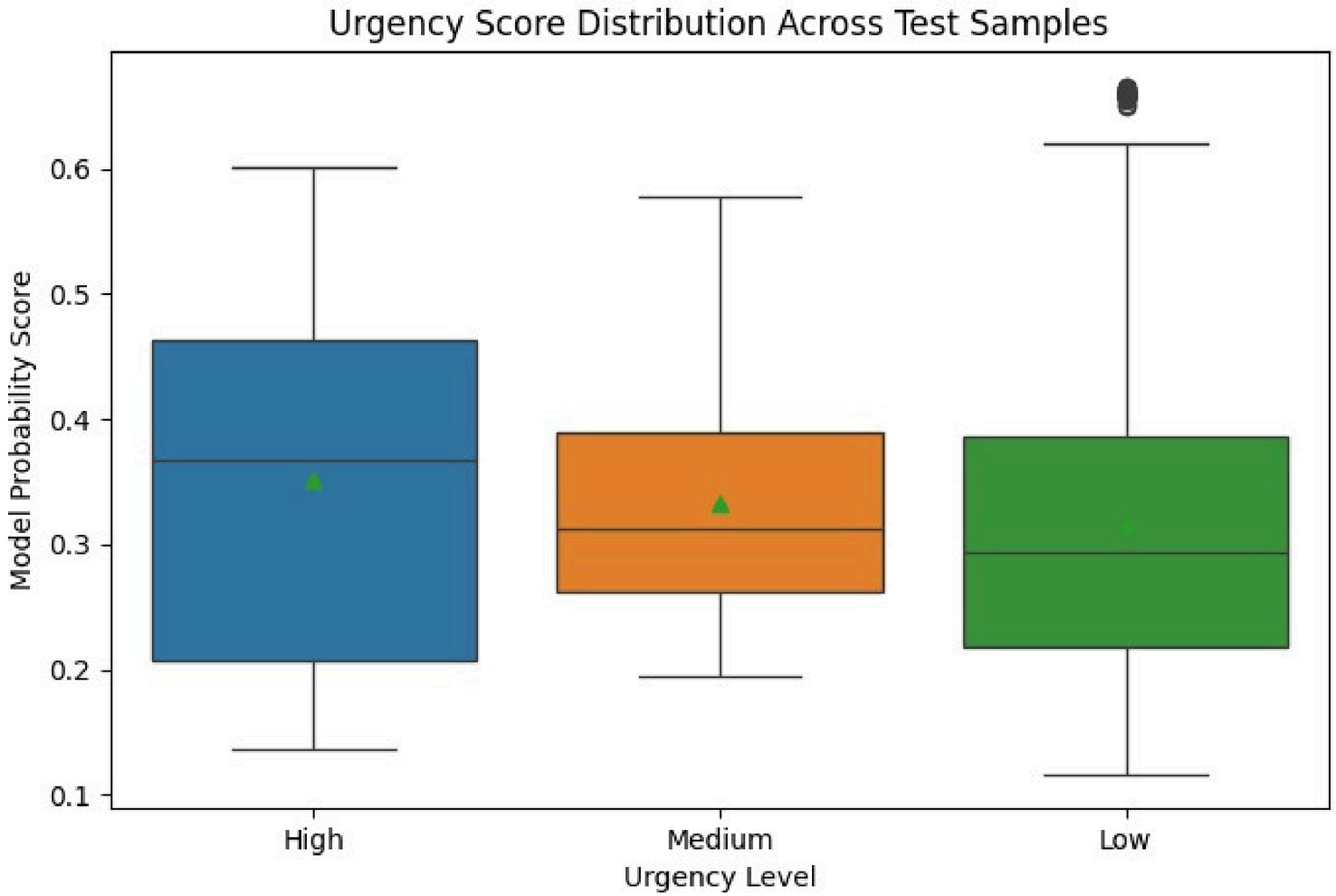



```
Classification Report:
              precision    recall  f1-score   support

        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
```

RESULT ANALYSIS CONTD...

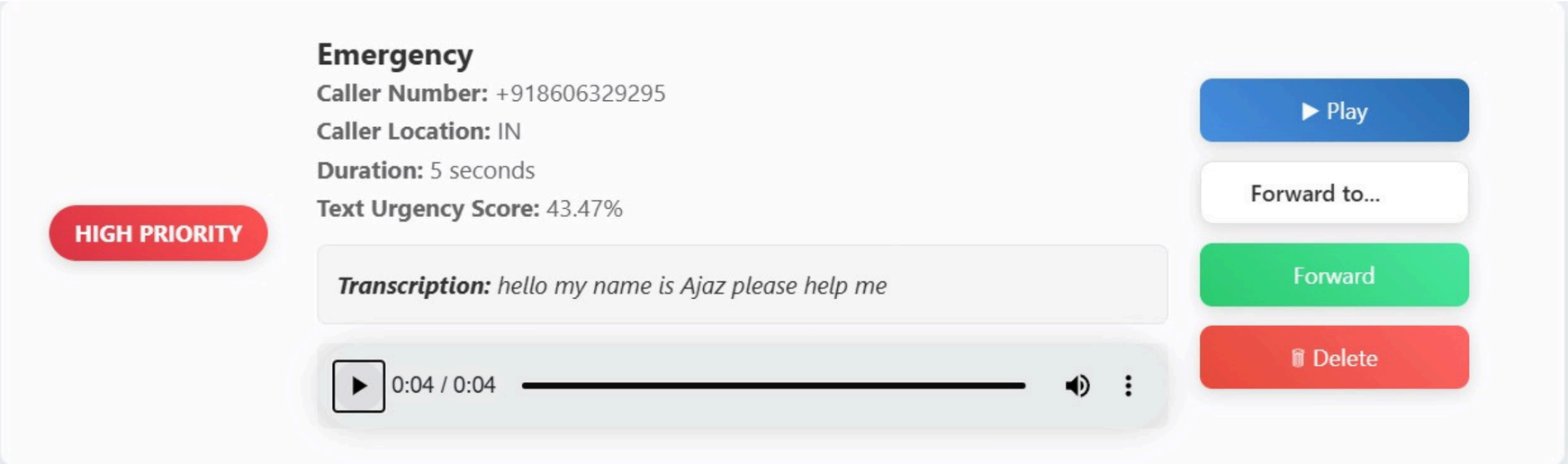


RESULT ANALYSIS CONCLUSION

- The **Audio Urgency** model is **82% accurate** but has some misclassifications. Fine-tuning can improve accuracy.
- The **Department Classification** model is **93.5% accurate** with minor emergency misclassifications.
- The **Text Urgency** Accuracy is **73.0%** with some misclassifications between medium and high urgency.

CHALLENGES

- Model seemed to be biasing towards word “Hello” due to inefficient data.



- Misclassification due to inefficiency in transcription.
- From 83 Audio Features only 8 features were selected for training due computational complexity.

TASK DISTRIBUTION

- Focuses on data generation, preprocessing, and feature extraction for voicemail.
- Designs, trains, and evaluates ML models for text-based prioritization.
- Responsible for creating the user interface for the voicemail ranker.
- Handles API integration, dataset management, and data flow between the interface, tone based prioritization.

FUTURE WORKS

- Expand models to handle multiple languages and dialects for diverse environments.
- Extend Summarization feature to voicemail.
- Improved Sentiment and Emotion Analysis.
- Develop AI-powered automated replies for common inquiries, reducing manual workload.
- Adapt the system for different sectors like customer support, banking, and legal services.

CONCLUSION

- The voicemail prioritization system enhances efficiency in managing voicemails.
- It uses AI and NLP for urgency detection, PII redaction, and transcription.
- The project is technically feasible and cost-effective, utilizing accessible tools.
- This system offers significant benefits for users in various fields, such as customer service, business management, etc improving communication and response times.

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] 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.

[4] 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.

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

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

[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.

Thank you!

Any questions?