**Infosys - Milestone 2** **M. Sri Lakshmi (Grp - 1)** **Milestone 1 Link**: [Document(Project)](https://docs.google.com/document/d/1OawuMHj4V0DttoTGeaw3Kwyf6R6ldASt9dRQB5dBfmE/edit?tab=t.0)  
 **Date**: 23/12/2024

**Milestone 2: Real-Time Sentiment, Tone, and Intent Analysis Engine**

**Overview**

This milestone focuses on implementing a real-time system to analyze sentiment, tone, and intent during live conversations, particularly for sales calls. The core of this project integrates speech-to-text transcription, advanced sentiment analysis, tone analysis, and real-time intent detection.

**Work Updates**

**23/12/2024: Introduction to Milestone 2 Components**

* Explored foundational tools and frameworks like OpenSMILE for audio analysis, Whisper for speech-to-text transcription, and the Gemini API for NLP tasks.
* Reviewed methods for detecting pitch, tone, and pace in speech to analyze emotional and linguistic patterns.

**24/12/2024: Detailed Implementation Planning**

* Finalized the pipeline components: Speech-to-Text, Sentiment Analysis, Tone Analysis, Intent Detection, and Integration.
* Set up and tested Whisper for live transcription.
* Researched the Gemini model for NLP-based sentiment and intent analysis.

**26/12/2024: Brainstorming and System Architecture**

* Defined the system architecture for integrating real-time transcription, analysis, and visualization.
* Key decision: Use Gemini for NLP analysis to leverage its advanced language understanding capabilities.
* Planned a user-friendly dashboard for live insights.

**27/12/2024: Tone Analysis with OpenSMILE**

* Began working with OpenSMILE to extract tonal features (e.g., pitch, volume) for emotion detection.
* Analyzed tonal variations in speech for emotion and sentiment evaluation.

**28/12/2024: Sentiment Analysis with Gemini**

* Utilized the Gemini model for analyzing sentiment from live transcribed speech.
* The analysis pipeline outputs sentiment classifications (e.g., Positive, Neutral, Negative) with confidence scores.

**29/12/2024: Tone Analysis Integration**

* Combined audio feature extraction using OpenSMILE with NLP-based sentiment outputs.
* Enhanced tone analysis for emotion detection during speech.

**30/12/2024: Real-Time Intent Detection with Gemini**

* Successfully implemented intent detection using the Gemini model.
* Achieved high accuracy in identifying user intents (e.g., Inquiry, Complaint, Negotiation).

**31/12/2024 - 2/01/2025: System Integration**

* Integrated Speech-to-Text, Sentiment Analysis, Tone Analysis, and Intent Detection into a unified pipeline.
* Ensured real-time performance and synchronization of outputs.
* Conducted initial testing for edge cases and system stability.

**Documentation for Real-Time Sentiment Analysis During a Call**

This document outlines the process for implementing real-time sentiment analysis during a call. The steps include Speech-to-Text, Sentiment Analysis, Tone Analysis, Real-Time Intent Detection, and Integration.

**Overview**

This project aims to process live conversations during sales calls by:

1. Transcribing speech to text.
2. Analyzing the sentiment of the text to gauge the user's mood.
3. Evaluating the tone from audio to detect emotions.
4. Detecting the user's intent in real time.
5. Integrating all components into a seamless real-time pipeline.

**Steps**

**1. Speech-to-Text (Transcribing)**

* **Objective**: Convert audio input into textual data for processing.
* **How it Works**:
  + The system captures audio through a microphone, processes it using speech recognition tools, and generates a textual transcript in real time.
* **Tools and Techniques**:
  + APIs like Google Speech-to-Text or OpenAI Whisper for transcription.
  + Libraries for capturing audio input, such as PyAudio.
* **Output**: A continuously updated text transcript of the call.

**2. Sentiment Analysis**

* **Objective**: Identify the sentiment (positive, negative, neutral) of the text.
* **How it Works**:
  + The system uses pre-trained machine learning models to classify sentiment based on linguistic features of the transcript.
* **Tools and Techniques**:
  + Pre-trained libraries like TextBlob or Hugging Face Transformers.
  + Custom fine-tuned models for domain-specific sentiment detection.
* **Output**:
  + Sentiment labels (e.g., Positive/Negative/Neutral) or scores (e.g., Positive: 0.85, Neutral: 0.10, Negative: 0.05).

**3. Tone Analysis**

* **Objective**: Evaluate the speaker's emotional tone (e.g., anger, happiness, sadness).
* **How it Works**:
  + The system analyzes audio features (pitch, volume, speech rate) to identify emotions in the speaker's tone.
* **Tools and Techniques**:
  + Audio feature extraction libraries like librosa.
  + APIs like IBM Watson Tone Analyzer for advanced tone detection.
* **Output**: Emotional tone classifications with scores (e.g., Happiness: 0.7, Anger: 0.2).

**4. Real-Time Intent Detection**

* **Objective**: Detect the purpose or intent behind the speaker’s statements.
* **How it Works**:
  + Intent detection leverages language models (LLMs) or fine-tuned intent classifiers to understand the context and classify the user's intent.
* **Tools and Techniques**:
  + LLMs like OpenAI GPT or Google Gemini.
  + Pre-trained or custom models fine-tuned on intent datasets.
* **Output**: Intent categories (e.g., Inquiry, Complaint, Negotiation).

**5. Integration**

* **Objective**: Combine all components into a unified real-time pipeline.
* **How it Works**:
  + Speech-to-Text transcribes the audio.
  + The transcript is passed to sentiment and intent analysis models.
  + Tone analysis runs parallelly on the audio stream.
  + All outputs are displayed in real-time on a dashboard or console.
* **Tools and Techniques**:
  + Real-time frameworks like Flask or Streamlit for UI.
  + Multithreading or asynchronous programming for smooth integration.
* **Output**: A seamless real-time system providing live insights into sentiment, tone, and intent.