Name: Shamaparvin Shaikh

Annual Sales Strategy Meet Audio Analysis Tool

1. Overview

The Annual Sales Strategy Meet Audio Analysis Tool is a Python-based application designed to analyze audio recordings of sales strategy meetings. The tool transcribes the audio, generates a summary, and provides participation metrics for each speaker. The user interface (UI) is built using the `tkinter` library, and the audio processing is handled by the 'speech_recognition' and 'pydub' libraries. The summary is generated using Hugging Face's `transformers` pipeline.

2. Key Features

1. Audio Input: Users can upload audio files in formats such as `.wav` and `.mp3`.

2. Transcription: The tool transcribes the audio into text using Google's Speech Recognition API.

3. Summary Generation: A concise summary of the transcribed text is generated using Hugging Face's summarization pipeline.

4. Participation Metrics:

- Identifies speakers (e.g., Speaker 1, Speaker 2).

- Calculates speaking duration and the number of contributions.

- Displays key contributions in a tabular format.

5. User Interface:

- Clean and organized layout.

- Displays full transcription, summary, and participation metrics.

3. Implementation Details

3.1 Code Structure

The code is structured into a single class, `AudioAnalyzerGUI`, which handles the UI and audio processing logic. Key components include:

File Input: Users can browse and select an audio file.

Audio Processing:

- Converts the audio to WAV format for compatibility.

- Transcribes the audio using `speech_recognition`.

- Summarizes the transcription using Hugging Face's `transformers`.

Metrics Calculation:

- Splits the transcription into parts and assigns them to speakers (currently assumes 2 speakers).

- Calculates speaking duration and contributions.

UI Display:

- Displays the transcription, summary, and metrics in a tabular format.

3.2 Libraries Used

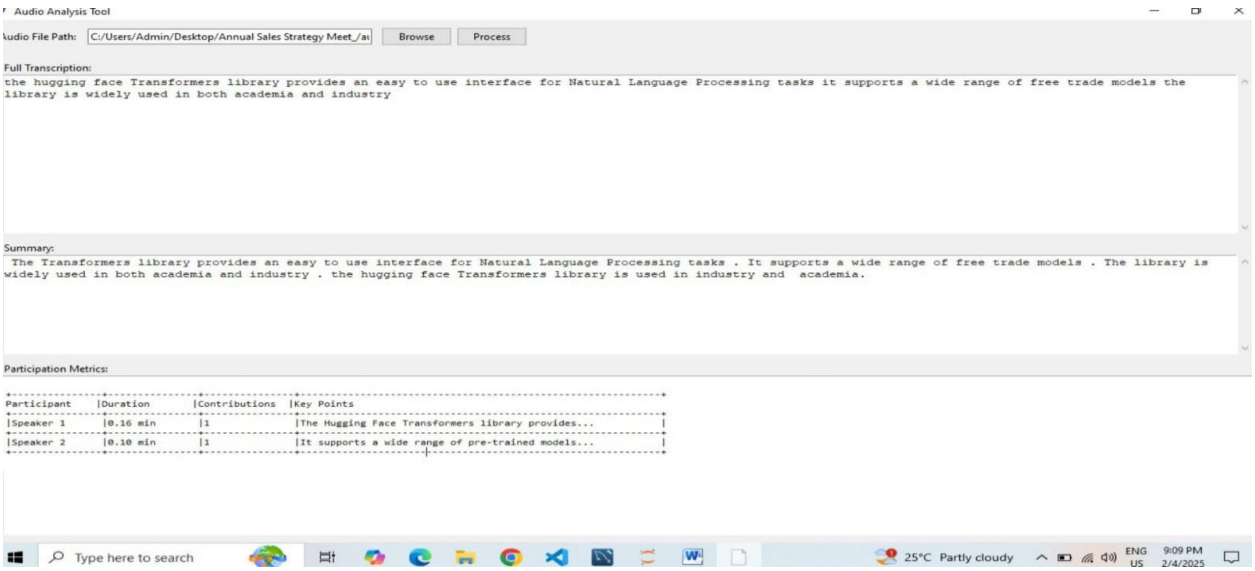
- tkinter: For building the graphical user interface.

- speech_recognition: For transcribing audio to text.

- pydub: For handling audio file conversions.

- transformers: For generating summaries using pre-trained models.

4. Sample Output



5. Evaluation Criteria

1. Accuracy of Transcription:

   - The tool uses Google's Speech Recognition API, which provides high accuracy for clear audio.

   - Challenges may arise with low-quality audio or overlapping speech.

2. Completeness and Clarity of Summary:

   - The Hugging Face summarization pipeline ensures a concise and coherent summary.

   - The summary length is adjustable (currently set to 150 words max).

3. Correct Identification of Participants:

   - The tool currently assumes 2 speakers and splits the transcription evenly.

   - For real-world use, speaker diarization (e.g., using `pyAudioAnalysis`) should be implemented.

4. Presentation and Organization:

   - The UI is clean and intuitive, with clear sections for transcription, summary, and metrics.

   - The metrics table is well-formatted and easy to read.

5. Quality and Usability of UI Design:

- The UI is responsive and user-friendly.

- The use of scrollable text boxes ensures readability for long transcriptions.

6. Limitations and Future Improvements

1. Speaker Diarization:

- The tool currently assumes a fixed number of speakers. Implementing speaker diarization would improve accuracy.

2. Handling Overlapping Speech:

- The tool may struggle with overlapping speech, which is common in meetings.

3. Customizable Summary Length:

- Allow users to adjust the summary length based on their needs.

4. Export Options:

- Add functionality to export the transcription, summary, and metrics as a PDF or CSV file.

5. Support for Multiple Languages:

- Extend the tool to support non-English audio.

7. Conclusion

The Annual Sales Strategy Meet Audio Analysis Tool is a robust solution for analyzing meeting recordings. It provides accurate transcriptions, concise summaries, and detailed participation metrics. With further enhancements, such as speaker diarization and multilingual support, the tool can become an indispensable resource for sales teams and other professionals.

8. Final Submission Requirements

1. Full Transcription: Displayed in the UI.

2. Refined Summary: Generated using Hugging Face's summarization pipeline.

3. Participation Metrics Table: Includes speaker details, duration, contributions, and key points.

4. UI Design: Clean and organized interface showcasing all data.

9. Code Execution

To run the tool:

1. Install the required libraries:

   pip install tkinter pydub speechrecognition transformers

   pip install SpeechRecognition

   pip install transformers

   pip install SpeechRecognition pydub transformers

2. Save the provided code in a Python file (e.g., `audio_analyzer.py`).

3. Run the script:

   python audio_analyzer.py

4. Use the UI to upload an audio file and view the results.

This report provides a comprehensive overview of the tool's functionality, implementation, and potential improvements. Let me know if you need further assistance!