**Milestone-01 Documentation**

**Overview**

This milestone focuses on creating a voice-interactive Real-Time AI Sales Assistant by integrating speech-to-text and text-to-speech functionalities with Google Generative AI. The assistant listens to user inputs, processes queries using AI, and responds through both text and voice, providing a seamless and engaging user experience.

**Modules and Functions Used**

**1. audio\_record**

* **Purpose**: Records user speech and converts it to text using Speech-to-Text functionality.
* **Source**: Imported from assignment\_02.

**2. text\_to\_speech**

* **Purpose**: Converts text responses to speech for user interaction.
* **Source**: Imported from assignment\_02.

**3. text\_response**

* **Purpose**: Processes user input using Google Generative AI and returns an appropriate response.
* **Source**: Imported from assignment\_03.

**Functional Workflow**

**1. Welcome Message**

* The program starts by greeting the user and explaining how to exit the session.
* Example Message:
* 🛒 \*\*Welcome to the Real-Time AI Sales Assistant!\*\* 🛒
* Say 'exit' to end the chat.

**2. Continuous Interaction Loop**

* The assistant:
  + Listens for user input using the audio\_record function.
  + Transcribes the recorded audio into text.
  + Checks if the user says "exit" to end the session.
  + If not exiting, passes the query to text\_response for AI processing.
  + Outputs the AI-generated response as both text and speech using text\_to\_speech.

**Key Features**

1. **Voice Interaction**: Enables hands-free communication by accepting and responding through voice.
2. **AI-Driven Responses**: Generates professional and empathetic replies using Google Generative AI.
3. **Real-Time Processing**: Provides immediate interaction and feedback.

**Functional Steps**

1. **Start the Program**: Displays a welcome message.
2. **Capture User Input**: Records speech and converts it into text.
3. **Check Exit Condition**: Ends the session if the user says "exit".
4. **Process AI Query**: Sends input to text\_response and retrieves AI-generated output.
5. **Output Response**: Displays and vocalizes the AI response.

**Example Interaction**

**Input**

* User says: "Can you explain the product features?"

**Output**

* AI Response:
* "Our product offers cutting-edge automation to simplify your tasks, boost productivity, and ensure a smooth user experience. How else can I assist?"
* The response is also vocalized for the user.

**Advantages**

1. **Hands-Free Convenience**: Allows users to interact using only their voice.
2. **Engaging Experience**: Combines text and audio for enhanced communication.
3. **Professional Responses**: Tailored and empathetic AI replies enhance user satisfaction.

**Limitations**

1. **Speech Recognition Dependency**: May struggle with noisy environments or varied accents.
2. **API Reliance**: Requires an active API key for functionality.
3. **Limited Error Handling**: Future improvements needed for handling unexpected inputs.

**Conclusion**

Milestone-01 marks the successful integration of speech processing and generative AI, paving the way for a user-friendly Real-Time Sales Assistant. This milestone demonstrates the capability to deliver engaging, professional, and dynamic voice-based interactions. Future enhancements may include multilingual support and improved speech recognition accuracy to cater to a broader audience.