

# Detailed Explanation

This project implements an end-to-end AI voice assistance pipeline, which processes audio input to generate meaningful text responses and converts those responses back into speech. The pipeline is broken down into the following steps:

## 1. Voice-to-Text Conversion:

- **Library:** whisper (OpenAI)
- **Model:** base
- **Description:** The Whisper model is employed to transcribe the audio input into text. This model is selected for its robustness in handling various accents and background noise, making it ideal for voice-driven applications.
- **Hyperparameters:**
  - **Language:** English (language="en")
  - **VAD Threshold:** Set to 0.5 for better voice activity detection, reducing background noise impact.

## 2. Text Input into Language Model (LLM):

- **Library:** transformers (Hugging Face)
- **Model:** GPT-2
- **Description:** The transcribed text is processed by a pre-trained language model, GPT-2, to generate a coherent and contextually relevant response. GPT-2 is chosen for its balance between performance and computational efficiency.
- **Hyperparameters:**
  - **Max Length:** 50 tokens, to limit the response length and maintain relevance.
  - **Num Return Sequences:** 1, to ensure only one response is generated.

## 3. Restricting Output Length:

- **Custom Function:** restrict\_output\_length
- **Description:** This function limits the LLM's output to a maximum of two sentences, ensuring concise responses. It improves user experience by avoiding overly verbose answers.

## 4. Text-to-Speech Conversion:

- **Library:** gTTS (Google Text-to-Speech)
- **Description:** Converts the generated text response into speech, saved as an audio file. gTTS is selected for its simplicity and integration with Python, though alternatives like edge\_tts can be used for more natural voice synthesis.
- **Output:** The audio file is saved in MP3 format for easy playback on various devices.

## Diagram

The following flowchart visually represents the data flow through the pipeline:

