Audio Topic Analysis

Overview

This project aims to develop an automated pipeline that processes audio files, analyzes the spoken content to understand the main topics discussed, and then synthesizes a spoken summary of the analysis. The pipeline will leverage Language Model (LLM), and Speech-to-Text (STT) technologies to facilitate a thorough understanding and summarization of audio conversations. This system can be useful in various scenarios like summarizing key points from meetings, conference calls, or customer service interactions.

Core Components

Audio Input System
: Create a user-friendly interface for uploading audio files in common formats (e.g., MP3, WAV).

Speech-to-Text (STT) Conversion
: Employ an STT engine to accurately transcribe the audio files into text.

Text Analysis and Summary Generation
: Utilize a Language Model to analyze the text, identify main topics, and generate a summary of the conversation.

Output System
: Design a user-friendly interface to present the text summary, play the audio summary, and allow users to download the audio summary.

API Endpoints
: Develop API endpoints using FastAPI for external systems or users to interact with the pipeline.

Container
: Dockerize the application.

Note: The focus of this project is the backend development, the interface can be simple.