CLOUD COMPUTING – CS623

A Cloud Based Smart Meeting Transcription & Analysis Platform

Murali Krishna Kancheti – yr1529
Karthik Veeravalli – df8025
Sunnyhita Guda – dj2361

Problem Statement:

In the modern business environment, transcribing and analyzing meetings manually is a time-consuming and error-prone process, which creates obstacles to effective collaboration.

Solution:

To overcome the difficulties posed by manual meeting transcription and analysis, our solution involves creating a Cloud-Based Smart Meeting Transcription & Analysis Platform. This platform will utilize advanced technologies such as Google Cloud Speech-to-Text API, OpenAI, and FFmpeg to automate real-time transcription and extract valuable insights from meetings. Leveraging the cloud's capabilities, the platform will foster seamless collaboration, leading to substantial reductions in transcription time and errors, while promoting effective knowledge sharing and decision-making during and after meetings.

Technologies used:

Google Cloud: Speech-to-Text API: The Speech-to-Text API by Google Cloud is an advanced service that transforms audio into written text with remarkable accuracy, even in challenging environments.Additionally, it includes features like automatic punctuation, word-level confidence scores, and the option to customize for specific terms or accents.

Google Cloud: Storage API : The Google Cloud Storage API offers developers and enterprises a scalable and versatile data storage solution, enabling seamless storage and retrieval of data, regardless of the size. It supports standard HTTP/HTTPS protocols and provides multiple storage classes, such as regional, multi-regional, nearline, and coldline, ensuring robust security and high availability through fine-grained access controls.

Open AI API: OpenAI's API opens up access to robust machine learning models specifically designed for natural language processing tasks,The API is user-friendly, making it accessible to both seasoned AI experts and novices in the field. With comprehensive documentation and a supportive community, it has gained widespread popularity for seamlessly integrating cutting-edge AI capabilities into various applications.

FFMPEG: FFmpeg, a popular open-source software project, comprises a collection of tools and libraries designed for multimedia data management. It offers users the ability to convert, record, stream, and manipulate video and audio files in diverse formats. With its flexibility and reliability, FFmpeg supports a broad spectrum of codecs, filters, and formats, making it a favored choice for both professionals and hobbyists involved in multimedia-related tasks.

Workflow and Architecture :

The audio processing stage involves utilizing FFmpeg to prepare audio files from meetings. Next, Google Cloud Speech-to-Text API transcribes the processed audio into text, with the transcripts stored in Google Cloud Storage for future access. The transcripts are then analyzed by ChatGPT from OpenAI's API to generate summaries or identify tasks. The insights derived from ChatGPT's analysis are used for follow-up actions. ChatGPT also generates AI-generated responses, which may include concise meeting summaries or clear lists of tasks. These responses are consolidated into a text file, presenting the AI's analysis in the requested format, encompassing summaries, tasks, or other valuable insights.