Project Overview: GenAI Video Transcription and Chat

Name: MD ROKIBUL HASAN

ID- 19900

Introduction

This project involves creating a chatbot that can answer questions from a video. It showcases a blend of technologies including Docker, OpenAI, Whisper, Embeddings, Chat completions, Pinecone, and Retrieval-Augmented Generation.

Prerequisites

- OpenAI API Key [https://platform.openai.com/api-keys]
- Pinecone API Key [https://app.pinecone.io]
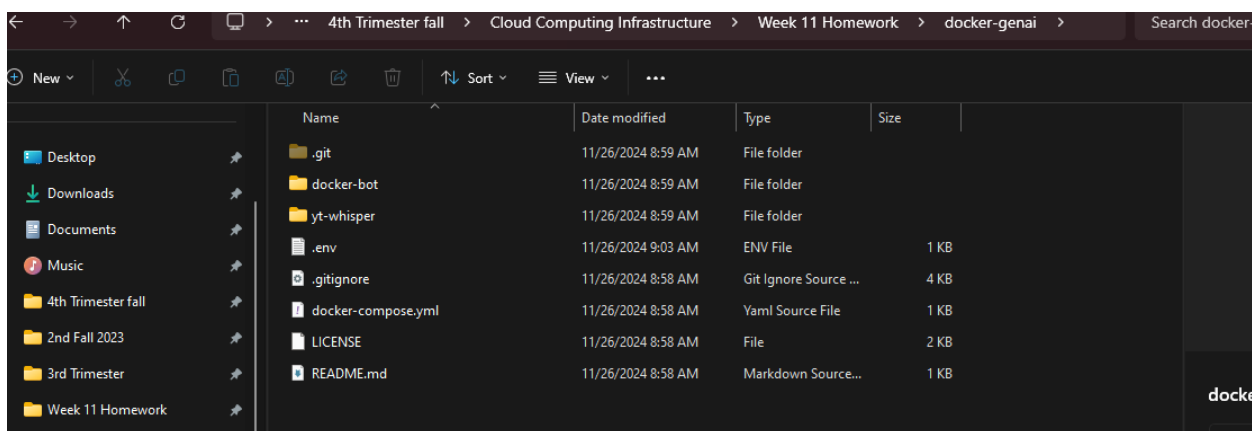- Latest version of Docker Desktop
- Git client

Application Purpose

The chatbot is designed to transcribe video content and answer questions related to the video, Providing relevant timestamps from the source.

Steps to Build and Run the Application

1. Clone Repository: Clone the project's repository using Git.

git clone https://github.com/Davidnet/docker-genai.git

2. Set API Keys

In the project directory, create a ".env" file and fill in your OpenAI and Pinecone API keys.

```
#--------------------------------------------------------------------------------
# OpenAI
#--------------------------------------------------------------------------------
OPENAI_TOKEN= YOUR KEYS HERE
#--------------------------------------------------------------------------------
# Pinecone
#--------------------------------------------------------------------------------
PINECONE_TOKEN=YOUR KEYS HERE
```
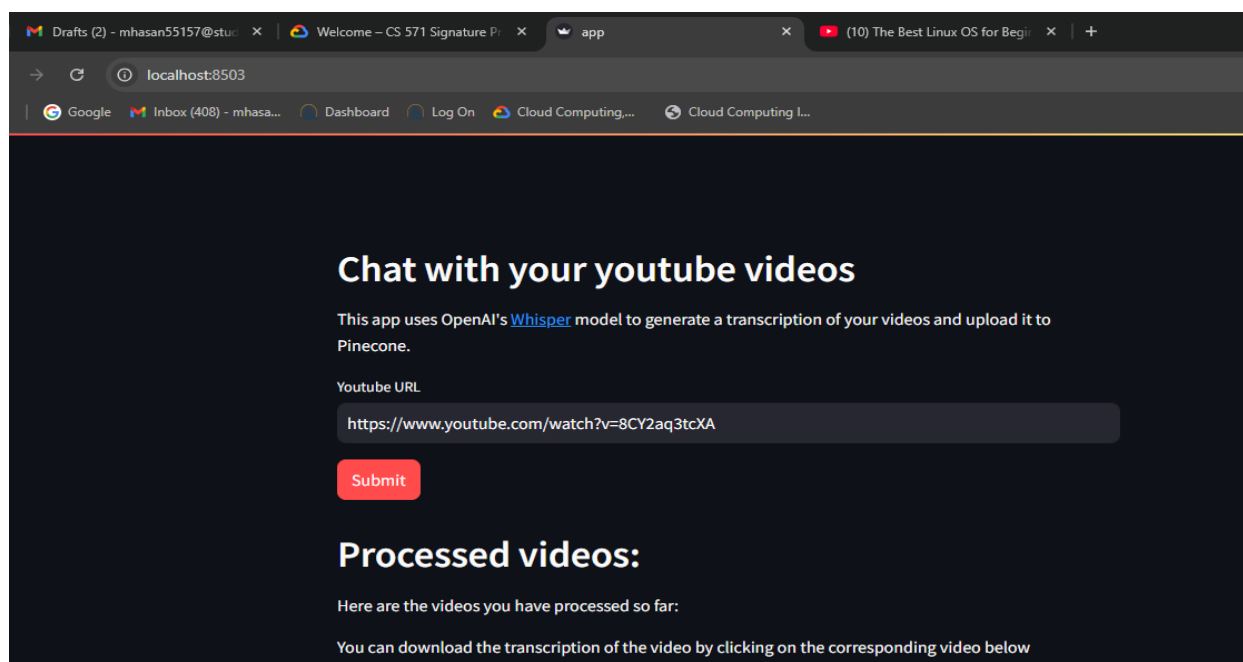
3. Build and Run:

Inside the project directory, execute the below command:

docker compose up --build

Using the Application

When the application is running, you'll see the logs of 2 services in the terminal. You'll see the services are exposed on ports 8503 and 8504. The two services are complimentary to each other.

1. yt-whisper Service: Running on port 8503, this service processes YouTube videos, transcribes them using Whisper, and stores the data in Pinecone.

The yt-whisper service first downloads the audio track from a given video. This audio is then processed using the Whisper technology to transcribe the spoken content into a WebVTT file format, which is made available for download. Then, the service generates embeddings using the text-embedding-3-small model. These embeddings are then uploaded to the Pinecone database. Once a video has been processed, the web application displays a list of videos that have been indexed within Pinecone, and it offers a feature to download the transcription of the video.

2. dockerbot Service: Available on port 8504, this service answers questions about the transcribed videos by querying the Pinecone database and generating responses using GPT.