AI Video Summarization Project using Mixtral, Whisper, and AWS
Project Overview

Overview
In today's fast-paced digital landscape, there is a significant demand for efficient educational tools. Traditional learning methods are rapidly evolving, with an increasing prevalence of video-based content. However, the abundance of educational videos available online presents a challenge for learners trying to extract key concepts effectively. This highlights the need for a comprehensive tool that can summarize videos and generate quizzes to help learners understand and retain essential information.

The emergence of large language models (LLMs) like Mixtral and Whisper is transforming the landscape of educational technology platforms. These powerful models use advanced natural language processing techniques to quickly understand and process complex educational content.

This project aims to address the growing need for effective educational tools by developing a robust video summarization and quiz generation application powered by state-of-the-art LLMs. By leveraging technologies like Mixtral and Whisper, the platform will extract key insights from educational videos, condense them into concise summaries, and generate quizzes to test comprehension. The project involves implementing a user interface and running the application on AWS cloud infrastructure, including integrating feedback mechanisms to continuously improve the learning experience.

Prerequisite Project: Kindly ensure the completion of the LLM Project for building and fine-tuning a large language model before proceeding with this project.

Note:
Please note that the use of the OpenAI API may require the utilization of allocated free credits or additional purchases (depending on the account status) for implementing the project. Kindly take note of the free credit limit provided for your usage.

Utilizing AWS services for this project may result in charges; it is essential to thoroughly review the AWS documentation to understand the pricing structure and potential costs associated with different resources and usage patterns.

Aim

The aim of the project is to develop a video summarization and quiz generation tool that leverages Large Language Models and AWS. By integrating advanced LLMs like Mixtral and Whisper, the application will extract key concepts from educational videos, generate concise summaries, and create quizzes to assess comprehension, thereby enhancing the efficiency and effectiveness of learning experiences.

Data Description

The dataset comprises tutorial videos covering diverse topics such as linear regression and other ML concepts.

Tech Stack

- ➔ Language: Python 3.10.4
- ➔ Libraries: Flask, FFmpeg, Torch, Transformers, Sqlalchemy
- ➔ Model: Mixtral, Whisper
- ➔ Cloud Platform: Amazon Web Services

Approach

- Initial Setup:
    - Create AWS EC2 instance with GPU for efficient functioning
    - Create environment and Install necessary dependencies and libraries.

- Model Selection:
    - Select Whisper for audio transcription
    - Choose Mixtral for video summarization and quiz generation

- Transcription Functionality:
    - Extract audio from uploaded videos using FFmpeg
    - Transcribe audio to text using the Whisper model

- Quiz Generation:
    - Use Mixtral to summarize and generate quizzes based on transcribed text
    - Implement different complexity levels for quizzes

- Frontend and API Development:
    - Design a user-friendly interface using HTML and JavaScript

    - Implement file upload functionality for videos

- Feedback Mechanism:
    - Develop a feedback form for user input
    - Store feedback data in a database for analysis

Modular code overview:

Once you unzip the code_folder.zip file, you can find the following:

```
├─ app
│   ├─ database.py
│   ├─ data_models
│   │   ├─ feedback.py
│   │   └─ __init__.py
│   ├─ mixtral_prompts.py
│   ├─ mixtral_wrapper.py
│   ├─ routes.py
│   ├─ templates
│   │   └─ index.html
│   ├─ utils.py
│   ├─ whisper_wrapper.py
│   ├─ __init__.py
├─ finetune.py
├─ readme.md
├─ requirements.txt
└─ run.py
```

Here is a brief information on the files:

- The code folder contains all the Python scripts and sample files used in the project.

- The README.MD and the solution methodology file provides essential information about the project, usage, and setup instructions.
  Kindly follow all the instructions for running the code from Readme.md file

- The requirements.txt file has all the required libraries with respective versions. Kindly install the file by using the command **pip install -r requirements.txt**

Project Takeaways

1. Understand the importance of video summarization and quiz generation
2. Gain insights into the decision-making process behind selecting language models and tools
3. Understand the strengths and limitations of popular language models like GPT-3, BERT, and Mixtral
4. Understand the impact of GPU-equipped instances on reducing latency
5. Learn how to set up AWS EC2 instances for running large language models effectively
6. Learn to integrate large language models like Mixtral and Whisper
7. Gain insights into the process of extracting audio from videos using FFmpeg
8. Understand the steps involved in transcribing audio to text using the Whisper model
9. Learn to generate quizzes dynamically based on various prompting strategies
10. Explore the implementation of API and front-end development using Flask, HTML, and JavaScript.
11. Gain experience in storing user feedback data in a database for feedback mechanism