This internship task is designed to provide you with an opportunity to work with the latest AI technologies throughout your entire internship duration of 3 months. If you are excited to work with the API that powers ChatGPT and build useful applications, completing this task correctly will increase your chances of securing the internship. We give high priority to the task submission and weigh it more heavily than the interview in our selection process.

Task Requirements:

Step Number 1: In this step, you will be required to write code to extract the entire transcript from a YouTube video using OpenAI's Whisper API. You can sign up for a free account and receive $20 in trial credits, which should be more than enough for this task. The YouTube video that you will be transcribing is available at https://www.youtube.com/watch?v=QxU-JrfA824.

Note that this task may take some time to complete, and you can test your code on smaller YouTube videos before submitting the final version.

Step Number 2: In this step, you will store the transcription in MongoDB and use the text file along with Langchain, Pinecone, and OpenAI to create a question-answering system using a combination of the three. We have included a couple of tutorials to help you get started:

https://www.youtube.com/watch?v=h0DHDp1FbmQ
https://towardsdatascience.com/4-ways-of-question-answering-in-langchain-188c6707cc5a

Step Number 3: In this step, you will create a simple frontend for end-users to interact with the application in a question-answer format.

NOTE:
1. Every thing has to be made in MERN Stack and not python as given in youtube examples
2. You can search for npm packages for each service and use chatgpt to rewrite the python given in youtube tutorials to node code
    a. Pincone package - https://docs.pinecone.io/docs/node-client
    b. Open AI - https://www.npmjs.com/package/openai
    c. Langchain - https://langchainers.hashnode.dev/getting-started-with-langchainjs
3. Above tutorials will only guide you how to make the search, but you have to write the code in node js.

Task Deliverables:

● A GitHub repository containing the source code of the question-answering system.
● A deployed version of the question-answering system on a cloud platform of your choice.
● A detailed documentation of the question-answering system, including installation instructions, usage instructions, and any additional information that you think would be useful.

Evaluation Criteria:

- Technical Skills: The quality of code, the use of best practices, and the ability to build a functional question-answering system using Langchain, Pinecone, and OpenAI.
- AI Technologies: The ability to use OpenAI's Whisper API, Langchain, Pinecone, and OpenAI to build a question-answering system.
- Deployment: The ability to deploy the question-answering system to a cloud platform of your choice.
- Documentation: The quality and completeness of the documentation provided.

You are free to make assumptions, use tutorials on the internet, ChatGPT to write code, GitHub Copilot, or anything else available to you. Look forward to your submissions :)