**Idea 1**

**Meeting Summary and Analysis tool**

Create application that will examine a meeting recording, convert speech to text using Whisper

(May give user option to save video after transcription which would impact the output options.)

Text will be fed into LLM and summarized (main points)

User will be able to select a main point and review summary.

User will be able to ask questions about the meeting.

Output: User can select to play the video at the starting point of the text they select or have the program convert text to audio and listen.

**Idea 2**

**Chat GPT tool for visual impaired.**

Use case: Visually Impaired people take a class about a specific topic. Create a program that will allow them to study the class content using audio interface.

Class content transcribed and put into LLM

Model creates summary (main points)

Initial Output: Main point count read to user. User prompted to provide a point they want to hear. Program plays audio summary of the point they selected.

Optional: Would be good to allow the user to ask general questions about the content and have answer read as well.

Notes & Random

Nvidia Language models

Look at Riva for transcribing

<https://catalog.ngc.nvidia.com/resources?filters=&orderBy=weightPopularDESC&query=&page=&pageSize=>

**NVIDIA NGC Catalog**

[**Jupyter Notebooks and Model Scripts | NVIDIA NGC**](https://catalog.ngc.nvidia.com/resources?filters=&orderBy=weightPopularDESC&query=&page=&pageSize=)

Step-by-step instructions to build models for video analytics, image classification, language translation, text-to-speech and more.