**Homework 2**

**Application Development Using Large Language Models**

**Due Wednesday the 19th, June by 8 pm Central time.**

1. The goal of this exercise is to (i) locate a YouTube video or some other platform-based video (with audio) that is not more than 10 minutes long on topics such as sports, news, science, arts, etc. and something non-controversial, (ii) create an English text of the audio, (iii) create an interesting, informative, and interactive set of 5 textual questions and answers (Q&A) from the content which are relevant to the topic or topics in the audio/text. The end goal is to then have either one AI voice (or two different AI voices) render the audio of the Q&A. You will submit the entire code, text generated from the audio, the text of the Q&A, and a single audio file that is rendered from the Q&A to Canvas.

Criteria: The code should be end-to-end automated. (In Colab, it should look like you run all the cells and everything works to get you the final result which is the Q&A audio file). No manual patching of workflow from one stage to another. This requires you to use DL packages, LLM Tools, and traditional Python code. Feel free to use chatbot LLMS to create code, generate ideas, etc. Document diligently where you got your insights from and submit it as well. (Sort of a disclosure document).

1. The second one will be slightly more challenging. Find any online video with many speaker participants, including possibly male and female (example <https://www.youtube.com/watch?v=yX5EJf4R77s> which is a panel discussion). I want you to think about designing a software package that when given this web link, answers the following questions: (1) processes the audio in some fashion (2) identifies the total number of distinct speakers (unique), and 3. Identifies the number of female and male speakers. At this time, I want you to research on a solution – if you want to attempt to solve it, even if it is imperfect, give it a shot, otherwise simply document your ideas for submission. The documentation should have a reasonable of chance of working – not some grand pie in the sky approach. You may not use Vision based approach for this problem for identifying gender or the number of participants.

PS: Some of the underlying research ideas can be open-ended – make reasonable assumptions for a solution. As always, feel free to ask questions – discuss among classmates. I will also turn on a Canvas discussion board.