Project Part 1

This sub-project is worth 50% of the overall project grade.

Task: Interview/Debate Audio Analysis

This task is to leverage some of the existing models to perform an analysis of interviews/debates. This tool could be used to identify media bias/impartiality

The objective of this task is to develop a notebook that will accept an audio file (mp3 or wav) of an interview/debate.

Speaker Diarisation Analysis:

Using pre-built Speaker Diarisation models, the first task is to identify who spoke when – i.e. the model should output the times that each speaker started and stopped talking. This should enable calculating how many seconds/minutes each speaker spoke for.

Speech to Text Analysis:

Once the separate speakers have been identified, the next task is to use a Speech to Text model to create a transcript of the audio with the speakers identified in the transcripts:

E.g:

[Speaker 1] Hi, how are you today?

[Speaker 2] I'm good and you?

[Speaker 1] Good thanks, how about….

This should enable an analysis of how many words each person spoke and might enable a frequency analysis of particular words (i.e. how many times a particular word was used by a particular individual).

Large Language Model Analysis:

Once you have the transcript with speakers identified and annotated, you can input this into a LLM to query it on sentiment etc.

E.g. ask it can it identify what the names of the speakers are/ask it whether speakers associate with more right-wing/left-wing values etc….

Testing:

I will give you an audio file for research purposes (from the Harris V Trump 2024 US Presidential Debate) – Note: this file is not to be hosted on a public GitHub repo.

However, you should test on another (potentially more complex) file (with more speakers/multiple speakers of the same sex) and evaluate/annotate the performance of the components.

Development/Documentation:

All documentation and development should be performed within a Jupyter Notebook. Regular commits should be made to a **private** GitHub repository. You must add me (brianmcgatu) as a collaborator. The audio you tested on should be also hosted on your private repo.

Your final committed notebook should be complete with all code cell outputs populated (i.e. it doesn't require a viewer (me) to replicate the environment and re-execute the notebook to view the results)

Your repository should also include a README – detailing how to recreate your environment for the notebook to execute etc.

Any submission that does not have a full and incremental git history with informative commit messages over the course of the project timeline will be accorded a proportionate mark.

Research:

All research must be captured through the notebook and collated in a bibliography at the end of the notebook. An academic referencing style must be used.

If doing research/comparisons between different models, include that analysis in a separate research notebook and refer to the research in the main notebook. This is to keep the main notebook analysis concise.

Rubric:

Note: Your final mark will not solely be based on your final results but also on your methodology/approach.

Consistency: 20%

Research: 25%

Development: 30%

Documentation: 25%

As usual, you are bound by ATU Student Code of Conduct (Student Code_Final_August_2022.pdf).