This is a doc of the initial work I did for my previous ideas for my project.

LLM Prompt Engineering:

- My first plan in this project was to create a pruner for Large Language Models and do different prompt engineering questions for different pruning ratios. This idea aros since I am very interested in prompt engineering and not testing. since this was not shown in the documentation for 'DepGraph: Towards Any Structural Pruning' the issue I ran into when trying to fine tune larger LLMs, I ran into was very long training times during the fine tuning process, around 3 hours per epoch. This was unfeasable due to time constraints and the amount of testings I wanted to complete, so I had to switch to another project. Link to my colab notebook: Notebook

Encoder-Decoder Prompt Engineering:

- After trying to implement LLM Prompt Engineering I switched to trying to run some tests on the code that was provided in the 'DepGraph: Towards Any Structural Pruning' repo, for the Encoder and Decoder architecture, and see what results I was able to achieve at different pruning ratio, but the results I was getting were really uninterpretable and was really not a good project to focus on due to the misinterpretability of the results I was getting. Link to my colab notebook: Notebook

Ended up pursuing my project: Trade-Offs Between Compression and Accuracy for Skin Disease Classification in Edge Devices