**Q: Executive Summary (Group)**: A brief overview of what has been accomplished this week.

A:

**Data**

1. Collected sustainability reports for target companies in the past 5 years (together)

2. Came up with sample prompts and expected answers for Q&A model (together)

**Research**

Researched on Reinforcement Learning with Human Feedback (Michelle & Jasper)

**Code**

1. Compared different vector databases, retrievers, and LLM model performances for the Q&A model (Yi)

2. Built summarization model, compared different LLM models, different chain types (refine and map-reduce), and did prompt engineerings to improve model performances (Alisa & Yanni)

3. Tested different prompts for Q&A model and prompt engineering(Yi & Michelle)

4. Multiple documents loading and table reader(Jasper)

**Q: Challenges and Roadblocks (Group)**: Enumerate any obstacles faced and how they were managed or are planned to be managed.

A:

**1.**

**Problem**

Summarization output is not ideal (inaccurate data, too general output, not well-formatted, etc.)

**Solution**

We will do research on why Open-Source LLMs do not perform well and how to solve the issue. We will also try out bigger sizes of the T5 model (XL, XXL). Finally, we will try out more prompts to see if we can get better answers.

**2.**

**Problem**

The retriever did not retrieve the correct chunk, which resulted in the failure to obtain correct results even though the model was good enough.

**Solution**

Adjust chunk sizes and set overlap to obtain more coherent data. Also research on better methods to load data, such as separating table and text.

**Q: Next Steps (Group)**: Outline the plan for the upcoming week.

**A:**

**Data Collection**

1. Collect 2-3 years quarterly reports
2. Label tables in reports manually (100-5000)
3. Get financial reports as the dataset for fine-tuning

**Research**

1. Possible reasons for bad results of summarization model
2. Possible reasons for bad results on complicated questions

**Visualization**

* Flow graph of summarization model

**Optimization**

1. Use the flan-t5-xl or flan-t5-xxl instead of flan-t5-l on flan-t5 model
2. Add metadata into the input to improve the retrieval part (year, page, etc.)
3. Further prompt engineerings for model improvement
4. Find way to better distinguish table and images in documents
5. Fine-tune the model by feeding 100,000 financial report

**Q: Executive Summary (Individual)**: A brief overview (one or two sentences) of what has been accomplished this week. Group coordinators collect and report the progress made by individual team members.

A:

**Yi Lu (yl5118)**: Collected sustainability reports for target companies in the past 5 years; Came up with sample prompts and expected answers for Q&A model; Compared different vector databases, retrievers, and LLM model performances for the Q&A model; Tested different prompts for Q&A model and prompt engineering

**Yanni Chen (yc4179)**: Collected sustainability reports for target companies in the past 5 years ; Came up with sample prompts and expected answers for Q&A model; Built summarization model, compared different LLM models, different chain types (refine and map-reduce), and did prompt engineerings to improve model performances

**Xiaolin Sima (xs2483)**: Collected sustainability reports for target companies in the past 5 years ; Came up with sample prompts and expected answers for Q&A model; Built summarization model, compared different LLM models, different chain types (refine and map-reduce), and did prompt engineerings to improve model performances

**Junyuan Huang (jh4608)**: Collected sustainability reports for target companies in the past 5 years ; Came up with sample prompts and expected answers for Q&A model; Researched on Reinforcement Learning with Human Feedback; Finished multiple documents loading and table reader

**Michelle Sun (ms6514)**: Collected sustainability reports for target companies in the past 5 years ; Came up with sample prompts and expected answers for Q&A model;Researched on Reinforcement Learning with Human Feedback; Tested different prompts for Q&A model and prompt engineering