**COMP4703 Natural Language Processing**

**Project Assignment 2 Report**

**Semester 2, 2024**  
**Due Date:** Friday, 11 August 2024, 3:00 PM

**1. Introduction**

* **Objective:**  
  Provide a brief overview of the project, focusing on the Retrieval Augmented Generation (RAG) system and the Multi-Hop Retrieval problem. Describe the importance of evaluating different retrieval and generation strategies in solving factoid question answering tasks that require information from multiple documents.
* **Background:**  
  Discuss the basics of Retrieval Augmented Generation Systems (RAG) and Multi-Hop Retrieval. Explain the stages involved: retrieval of top-k documents and the use of a Large Language Model (LLM) for answer generation.

**2. Methodology**

* **Stage 1: Retrieval**  
  Describe the methods used for the retrieval stage, including any algorithms or techniques implemented. Explain how different algorithms were compared and any reranking strategies applied.
* **Stage 2: Generation**  
  Detail the Large Language Models (LLMs) used for generating answers from the retrieved documents. Include any modifications or configurations applied to the LLMs.
* **Evaluation Metrics:**  
  List and describe the metrics used to evaluate the performance of each system, including precision, recall, F1 score, or any additional metrics you introduced.

**3. Experiments**

* **Experiment Setup:**  
  Outline the experimental setup, including the configuration of the retrieval algorithms and LLMs. Detail the data used, such as the corpus and queries.
* **Results:**  
  Present the results of the experiments. Include tables, figures, or diagrams that summarize the performance of different models and configurations.
* **Comparison:**  
  Provide a detailed comparison of the various algorithms and models used. Discuss their strengths, weaknesses, and overall effectiveness in solving the Multi-Hop Retrieval problem.

**4. Conclusions**

* **Summary of Findings:**  
  Summarize the key findings from your experiments. Highlight which methods performed best and why.
* **Insights:**  
  Provide any insights gained from the experimentation process. Discuss any unexpected results or challenges encountered.
* **Future Work:**  
  Suggest potential improvements or future research directions based on your findings.

**5. References**

* Include a list of references for any papers, articles, or tools used in your research.

**Additional Sections (if applicable)**

* **Code and Scripts:**  
  Provide a brief description of the code and scripts used for the experiments. Mention any additional scripts or tools that were particularly important.
* **Intermediate Results:**  
  Include any important intermediate results that support the findings of your experiments.
* **Appendices:**  
  Add any supplementary material, such as detailed tables, additional figures, or extended results, if necessary.

**Submission Checklist**

* **MyREADME.md** - Instructions to run the code and environment changes.
* **report.pdf** - The experimental report.
* **requirements.txt** - Generated requirements file.
* **ranker[A-Z].py/ipynb** - Retrieval baselines.
* **reranker[A-Z].py/ipynb** - Reranking baselines.
* **rag[A-Z].py/ipynb** - LLM text generation baselines.
* **output/ranker[A-Z].json** - Results from retrieval baselines.
* **output/reranker[A-Z].json** - Results from reranking baselines.
* **output/rag[A-Z].json** - Results from LLM text generation baselines.
* **evalXX.[py/json]** - Additional evaluation metrics code.