

Coding Exercise: Building a RAG Pipeline

The PDF document to be used for this exercise is titled: “**BloombergGPT: A Large Language Model for Finance**” found via this link: <https://arxiv.org/pdf/2303.17564>

1. Extract text from the provided PDF document using a library of your choice.
2. Implement a retrieval system to find relevant information from the extracted text. You can use any method or library for this purpose.
3. Use a language model to generate responses based on the retrieved information. You are free to use any LLM and framework.
4. Define and implement metrics to evaluate the quality of the generated responses.
5. Document your code and provide a brief report explaining your approach, the metrics you defined, and any challenges you faced.
6. Submit your code and report within 24-48 hours.

Note: You are allowed to use any framework and language for this exercise. The goal is to understand your way of thinking, code organization, and problem-solving approach.