

Implementing a Retrieval-Augmented Generation (RAG) System

Objective:

Implement a simple RAG system using either LangChain, LlamaIndex or any other framework of your choice. Experiment with different embedding models, chunking strategies, pdf readers, etc and document your findings. PDF for RAG: [Link to PDF](#)

Requirements:

1. RAG Implementation
 - Choose either LlamaIndex or LangChain as your framework
 - Use the provided PDF for your knowledge base
 - Implement basic RAG functionality
2. Experiments
 - Document your experimental setup, process, and results
3. Documentation
 - Provide clear explanations for each experiment
 - Include qualitative analysis of results
 - Mention the shortcomings of your current implementation in real scenario
4. Optional
 - Add relevant citations to support your findings and methodologies
 - Discuss the impact of different models and strategies on RAG performance

Deliverables:

1. Source code for your RAG implementation
2. A detailed report documenting your experiments and findings

Evaluation Criteria:

- Working implementation of RAG system
- Depth and variety of experiments conducted (more weightage to unique ideas in experimentation)
- Clarity and thoroughness of documentation

Resources:

- [LlamaIndex documentation](#) / [LangChain documentation](#)
- <https://blog.lancedb.com/chunking-techniques-with-langchain-and-llamaindex>