

IIR – AIML Engineer- Project Report Analyzer- Technical Assessment

Objective

Create a question-answering system that enables users to query multiple project reports and receive accurate, sourced responses.

Core Requirements

Document Management

- Allow users to upload multiple project reports (PDF format)
- Support simultaneous querying across all uploaded documents
- Maintain session persistence of uploaded documents

Question Answering System

- Implement RAG (Retrieval Augmented Generation) to provide accurate answers from the documents
- For each answer, identify and cite the specific sections of the document(s) used
- Handle queries that may require information from multiple documents
- Provide coherent responses that combine information when needed

Response Format

Each answer should include:

- The response to the query
- Source attribution indicating which document(s) and section(s) were referenced
- Direct quotes or references from the source material where relevant

Technical Requirements

- Backend must be implemented in Python
- Use an LLM of your choice (e.g., OpenAI, Anthropic, open-source models)
- Implement appropriate vector storage for document embeddings
- Include proper error handling for cases such as:

- * Irrelevant questions
- * Questions without sufficient information in the documents
- * Invalid document formats

User Interface

- Create a simple interface that allows:
 - * Document upload
 - * Question input
 - * Clear answer display with source attribution
- Can use Streamlit or any other presentation method of your choice
- Focus is on functionality rather than design sophistication

Evaluation Criteria

1. RAG Implementation

- Accuracy of responses (*does not need to be perfect... just try your best*)
- Appropriateness of source attribution
- Handling of multi-document queries

2. Code Quality

- Clean, well-structured code (*does not need to be perfect... just try your best*)
- Proper error handling (*does not need to be perfect... just try your best*)
- Clear documentation

3. Functionality

- System reliability
- Accuracy of source attribution

Deliverables

1. Working application with interface

2. Source code with documentation
3. Brief explanation of:
 - Choice of technologies
 - RAG implementation approach
 - Any challenges faced and how they were solved

Time Allocation

- 10 days (feel free to send it in earlier). Remember the output does not have to be perfect. We're generally more concerned with the thought-process and approach and we understand that the time is limited.

Example Queries to Support

- "What could have been some main challenges faced in this Project?" *(if one project report is uploaded)*
- "Compare the timelines of these projects." *(if multiple project reports are uploaded)*
- "What was the budget allocation and location of this Project?" *(if one project report is uploaded)*
- "What were the budget allocations and locations of these Projects?" *(if multiple project reports are uploaded)*