RAG Task

Assignment: Build a RAG Model

Objective: Develop an application powered by a **single LLM-based agent** that follows RAG flow

1. Attempt to answer user queries using a RAG-based Knowledge Base

All decision-making and logic must reside within one unified LLM agent.

Scope of work:

1. Unified Agent Design (LLM-based)
    a. Implement a single-agent framework that:
        i. First performs knowledge retrieval via RAG.

2. RAG Layer:
    a. Build a Knowledge Base using a vector DB (e.g., **ChromaDB**, **Weaviate**, or **FAISS**). or any open-source DB available
    b. Optimize for:

        - Chunking strategy
        - Embedding model selection
        - Retrieval accuracy & performance
        - Context injection format for LLM

3. UI & Deployment
    a. Simple frontend to input questions and display the final model response (any open source is fine)

Deliverables

- **Source Code** with modular structure
- **README** including:
    - Architecture diagram
    - Setup instructions
    - Optimization explanations
    - Sample queries and expected outputs

Evaluation Criteria:

1. Full functional working prototype
2. Architecture Diagram
3. Optimisation logics will be given more weightage

Document Store:

1. Please try to use google drive for the Knowledge store