Name: Rushiraj Suwarnkar  
Roll no.: 381018  
PRN: 22310315  
———————————————————————————————

**Assignment 9: Create a Chatbot Application for a Real-World Scenario**  
———————————————————————————————

**Problem Statement:**  
Develop a chatbot application for a real-world scenario using a Retrieval-Augmented Generation (RAG) approach. The chatbot should answer FAQs for a banking system by retrieving relevant information from documents and generating human-like responses.  
———————————————————————————————

**Objectives:**

* Understand the architecture and working of RAG-based chatbots.
* Implement a chatbot capable of retrieving and generating answers for user queries.
* Apply LangChain and Groq LLM for integrating knowledge retrieval and natural language responses.  
  ———————————————————————————————

**Theory:**  
———————————————————————————————

**Methodology:**  
A RAG-based chatbot combines **retrieval-based** and **generative** models:

1. **Retrieval:** User queries are converted to embeddings, and a vector database searches for the most relevant documents (bank FAQs, policies, etc.).
2. **Generation:** The LLM (Groq) processes the retrieved documents along with the query to generate accurate, context-aware responses.
3. **Integration:** LangChain manages document loading, embeddings, retrieval, and generation in a seamless pipeline.  
   This approach allows the chatbot to provide precise answers based on a knowledge base rather than purely generating from learned patterns.  
   ———————————————————————————————

**Working Principle / Algorithm:**  
———————————————————————————————

**RAG Chatbot Algorithm:**

1. Load FAQ documents or bank-related information into the knowledge base.
2. Generate embeddings for the documents using a vector embedding model.
3. When a user asks a question:
   * Convert the query into embeddings.
   * Retrieve top relevant documents from the vector database.
   * Pass the query and retrieved documents to Groq LLM via LangChain.
4. Generate a natural language response based on the retrieved information.
5. Display the response to the user in the chatbot interface.  
   ———————————————————————————————

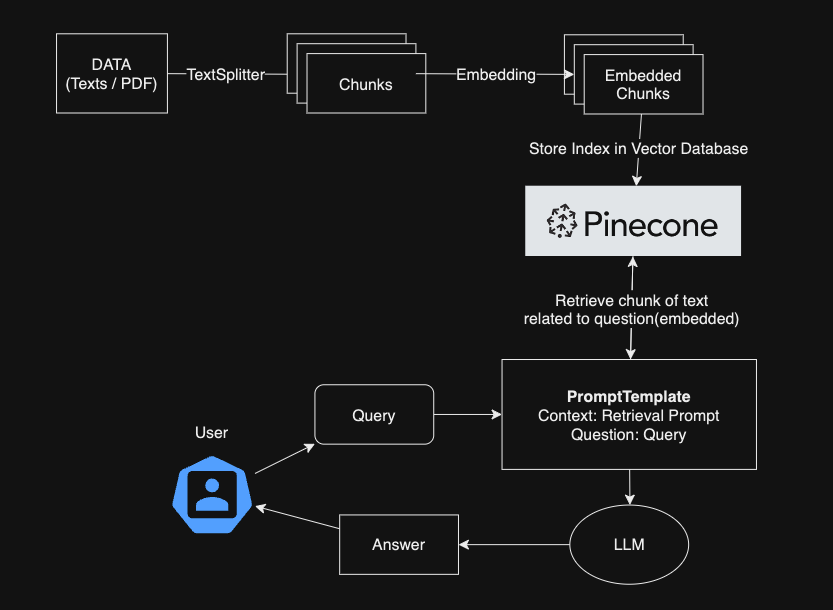
**Advantages:**

* Provides accurate, context-aware answers by leveraging document retrieval.
* Reduces hallucinations by grounding responses in factual knowledge.
* Scalable for large knowledge bases like banks, insurance, or customer support FAQs.  
  ———————————————————————————————

**Disadvantages / Limitations:**

* Requires preprocessing and embedding of knowledge base documents.
* May need frequent updates to reflect new information.
* Latency depends on retrieval and generation pipeline performance.  
  ———————————————————————————————

**Diagram:**

  
———————————————————————————————

**Conclusion:**  
A RAG-based chatbot efficiently integrates retrieval and generation to answer user queries in real-world applications like banking FAQs. Using LangChain and Groq LLM, it provides accurate, context-aware responses while ensuring scalability and reliability for enterprise applications.  
———————————————————————————————