**PART – A**

Develop a **Retrieval-Augmented Generation (RAG)** model to answer financial questions based on company financial statements (last two years).

**CAI Group 87 Members:**

|  |  |
| --- | --- |
| SATHISH KUMAR .S | 2023aa05698 |
| SATHISH N M | 2023aa05812 |
| ROBERTJOHNSON S | 2023aa05813 |
| SAURABH SHARMA | 2023aa05626 |

**Financial Chatbot - Advanced RAG**

**🚀 Setup Instructions**

**1️⃣ Create and Activate Virtual Environment**

To set up the project, create a virtual environment and activate it:

python -m venv env\_financial\_report\_Chatbot

env\_financial\_report\_Chatbot\Scripts\activate

**2️⃣ Install Dependencies**

Install the required dependencies using:

pip install -r requirements.txt

**3️⃣ Run the Application**

To start the chatbot, use one of the following commands:

**Run the Basic RAG Version:**

streamlit run app.py

**Run the Advanced RAG Version:**

streamlit run C:\Robert\GitWorkspace\CAI\_Financial\_Chatbot\_AdavancedRAG\AdvanceRAG.py

**🌍 Project Links**

* 📂 **GitHub Repository**: [CAI Financial Chatbot](https://github.com/robertjohnson88/CAI_Financial_Chatbot)
* 💰 **Live RAG Chatbot**: [Basic RAG](https://cai-financial-chatbot.streamlit.app/)
* 📈 **Advanced RAG Chatbot**: [Advanced RAG](https://cai-financial-advancedrag.streamlit.app/)

**📝 Query Examples**

**✅ High-Confidence Query (Relevant Financial Question)**

**Example:** "GOOG company revenue in year 2020"

**⚠️ Low-Confidence Query (Future Financial Data)**

**Example:** "GOOG company revenue in year 2045"

**❌ Irrelevant Query (General Knowledge - Not Supported)**

**Example:** "What is the capital of France?"

🔹 *Note: The chatbot does not handle general knowledge queries unrelated to financial data.*

|  |  |
| --- | --- |
| 1. Data Collection & Preprocessing | Download the last two years of financials (Use any one group member's company earning statements, if nothing is available use any company's data freely available). Clean and structure the data for retrieval. |
| 2. Basic RAG Implementation | Implement a simple RAG model: - Convert financial documents into text chunks. - Embed using a pre-trained model - Store and retrieve using a basic vector database |
| 3. Advanced RAG Implementation | Improve retrieval by: - Using BM25 for keyword-based search alongside embeddings. - Testing different chunk sizes & retrieval methods for better accuracy. - Implementing re-ranking. |
| 4. UI Development (e.g., Streamlit) | Build an interactive UI: - Accept user queries. - Display answer & confidence score. - Ensure clear formatting & responsiveness. |
| 5. Guard Rail Implementation | Implement one guardrail: - Input-Side: Validate and filter user queries to prevent irrelevant/harmful inputs. - Output-Side: Filter responses to remove hallucinated or misleading answers. |
| 6. Testing & Validation | Ask 3 test questions: - A relevant financial question (high-confidence). - A relevant financial question (low-confidence). - An irrelevant question (e.g., "What is the capital of France?") to check system robustness. |

1. Data Collection & Preprocessing

A screen shot of a computer

AI-generated content may be incorrect.

Basic RAG Implementation

A computer screen shot of a program

AI-generated content may be incorrect.

A screen shot of a computer

AI-generated content may be incorrect.

A screen shot of a computer program

AI-generated content may be incorrect.

Advanced RAG Implementation:

A screenshot of a computer program

AI-generated content may be incorrect.

UI Development using Streamlit

A screenshot of a computer

AI-generated content may be incorrect.

A relevant financial question (high-confidence)

(e.g., GOOG company revenue in year 2020)

A screenshot of a computer

AI-generated content may be incorrect.

A relevant financial question (low-confidence).

(e.g., GOOG company revenue in year 2045)

A screenshot of a computer

AI-generated content may be incorrect.

An irrelevant question

(e.g., "What is the capital of France?")

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.