**Medical Chatbot with Retrieval-Augmented Generation (RAG)**

**MSDS692 – Data Science Practicum II**  
Weekly Progress Reports  
**Jonish Bishwakarma**

**Week 4 – Retrieval Pipeline and LLM Integration**

**Planned Work for the Week**

Connect Pinecone retriever to LangChain and test integration with OpenAI model.

**Project Timeline**

* Week 1 – Project definition and proposal submission (DONE)
* Week 2 – Proposal revision and evaluation design (DONE)
* Week 3 – Collect medical documents and build knowledge base (DONE)
* Week 4 – Implement retrieval pipeline and integrate with LLM (DONE)
* Week 5 – Develop evaluation script and run baseline testing
* Week 6 – Debugging and refinement
* Week 7 – Build chatbot interface
* Week 8 – Final evaluation, analysis, and presentation

**Progress for the Week**

* Implemented retriever with top-3 similarity search.
* Integrated retriever with OpenAI GPT model using LangChain’s create\_retrieval\_chain().
* Designed system prompt to enforce rules (use only retrieved context, or answer “I don’t know…”).
* Verified retrieval returned relevant chunks for simple queries.

**Roadblocks/Issues**

Initial errors when passing raw strings instead of Document objects caused chain failures. Fixed by wrapping retrieved docs properly.

**Plan for Next Week**

Build evaluation script to compare RAG and LLM-only answers.

**Resources for the Week**

* Guha, M. (2007). *The Gale Encyclopedia of Medicine (3rd edition).* Detroit: Thomson Gale.
* LangChain Documentation – Retrieval and Chain Setup. <https://python.langchain.com>
* OpenAI API Documentation. <https://platform.openai.com/docs>