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

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

**Week 1 – Project Proposal**

**Project Details**

This project will develop a **medical chatbot** using **retrieval-augmented generation (RAG)** and compare it to an LLM-only baseline. The focus is on ensuring accuracy, faithfulness to sources, and privacy when answering user medical questions.

**Project Timeline**

* Week 1 – Project definition and proposal submission (DONE)
* Week 2 – Proposal revision and evaluation design
* Week 3 – Collect medical documents and build knowledge base
* Week 4 – Implement retrieval pipeline and integrate with LLM
* 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

**Planned Work for the Week**

Define the scope of the project, identify tools (LangChain, HuggingFace, Pinecone, Flask, OpenAI), and submit the proposal.

**Progress for the Week**

* Drafted and submitted project proposal.
* Outlined objectives, methodology, expected outcomes, and timeline.
* Identified evaluation metrics (accuracy, faithfulness, privacy).

**Plan for Next Week**

Revise the proposal with stronger literature grounding and begin defining the evaluation design more clearly.