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

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

**Week 8 – Final Evaluation and Presentation**

**Planned Work for the Week**

Run final evaluation, analyze results, and prepare presentation.

**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 (DONE)
* Week 6 – Debugging and refinement (DONE)
* Week 7 – Build chatbot interface (DONE)
* Week 8 – Final evaluation, analysis, and presentation (DONE)

**Progress for the Week**

* Completed full evaluation on test questions.
* Compared RAG vs. LLM-only:
  + **RAG:** More faithful and private, but sometimes lower accuracy.
  + **LLM-only:** Higher apparent accuracy but risked hallucinations and unsafe answers.
* Created slides analyzing metrics and explaining trade-offs.
* Prepared final presentation with detailed speaker notes.

**Roadblocks/Issues**

Time constraints in refining privacy evaluation (still simplified to checking if out-of-scope → “don’t know” response).

**Plan for Next Week**

Deliver final presentation and submit project deliverables.

**Resources for the Week**

* Guha, M. (2007). *The Gale Encyclopedia of Medicine (3rd edition).* Detroit: Thomson Gale.
* McGough, B. (2024, July 18). *Building an Accurate, Tested RAG Pipeline for PDF Document Search with LangChain*. *Medium*.