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

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

**Week 6 – Debugging and Refinement**

**Planned Work for the Week**

Debug evaluation and refine how RAG processes inputs.

**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
* Week 8 – Final evaluation, analysis, and presentation

**Progress for the Week**

* Adjusted similarity threshold and filtering for relevant documents.
* Fixed errors where str was passed instead of Document objects (caused page\_content issues).
* Improved evaluation by separating “no answer due to no context” from incorrect answers.
* Re-ran tests and confirmed stable CSV output of metrics.

**Roadblocks/Issues**

RAG sometimes underperforms LLM in raw “accuracy” because it refuses to answer without context. However, this behavior is consistent with **safer medical practice.**

**Plan for Next Week**

Develop chatbot interface with Flask + HTML/CSS.

**Resources for the Week**

* Guha, M. (2007). *The Gale Encyclopedia of Medicine (3rd edition).* Detroit: Thomson Gale.