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

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

**Week 3 – Knowledge Base Construction**

**Planned Work for the Week**

Collect documents and build embeddings for retrieval.

**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
* 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**

* Gathered reference-quality data from *The Gale Encyclopedia of Medicine*.
* Converted documents into text chunks for indexing.
* Implemented HuggingFace sentence-transformer embeddings (all-MiniLM-L6-v2).
* Created a Pinecone index (medibot) and uploaded vectors.

**Roadblocks/Issues**

Indexing large chunks took more time than expected due to embedding size and API latency.

**Plan for Next Week**

Implement retrieval pipeline and connect it with the LLM.

**Resources for the Week**

* Guha, M. (2007). *The Gale Encyclopedia of Medicine (3rd edition).* Detroit: Thomson Gale.
* Hugging Face. (n.d.). Sentence Transformers: all-MiniLM-L6-v2. https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2
* Pinecone Documentation. https://docs.pinecone.io