Project 6: Conversational Chatbot (NLP)

1. Description

A simple virtual assistant answers natural language questions (e.g., “What career paths does program X offer?”) using official descriptions, returning concise, digestible answers with cited sources.

2. Objectives

- Build a retrieval-based (RAG) or retrieval+generation QA system.
- Provide a chat interface with source citations.
- Maintain conversational context (basic multi-turn).

3. Deliverables

- Backend retrieval + answer synthesis pipeline.
- Chat UI with history and source links.
- Evaluation suite (accuracy, human rating).

4. Weekly plan (5 weeks suggested)

- Week 1: Ingest and preprocess official documents; build chunking strategy.
- Week 2: Implement retrieval (BM25 + embeddings + FAISS).
- Week 3: Answer generation/synthesis (templates or light seq2seq model).
- Week 4: Chat UI, multi-turn handling, evaluation.
- Week 5 (optional): Improve persona, add multimodal document upload.

5. Main tasks

- Clean and chunk the corpus (PDF → text).
- Build retrieval index and baseline retrieval.
- Implement answer generation (concise synthesis + citation).
- Add conversation context handling and logging.
- Evaluate on a test set and via user judgments.

6. Recommended stack

Python, HuggingFace / sentence-transformers, FAISS, BM25 (Whoosh/Elastic), FastAPI, React, Docker.

7. Data & resources

- Official program descriptions, FAQs, institutional PDFs.

8. Evaluation criteria

- Exactness of factual answers, user satisfaction scores.
- Usage logs and time-to-answer.
- Ability to cite sources.

9. Optional extensions

- Allow students to upload documents and answer from them (on-the-fly ingestion).
- Integrate with RAG + larger LLM for improved fluency (with caution on hallucinations).