Capstone Project

_____

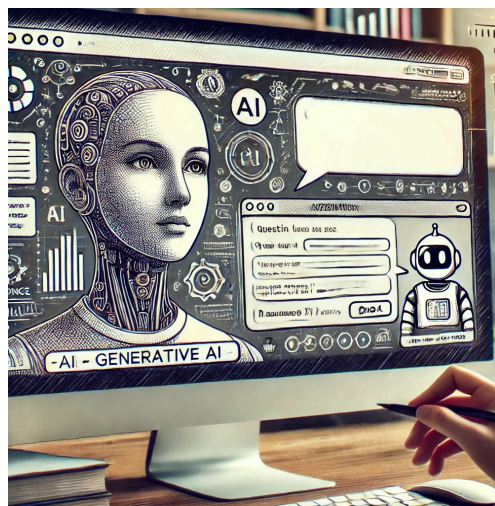**Project Title**: Research Paper Answer Bot

Project Type: 2 Weeks Capstone Project

Organization: Analytics Vidya

Learner Name: <ENTER HERE>

Project Description / Business case:

The scenario is that you are working for a company like [ArXiv](#) which has a repository of the latest and best papers around AI and Data Science. They want to build an intelligent chatbot which can help answer questions which people ask on popular topics based on research papers in Generative AI. Considering the limited time-frame of the project, we have limited the dataset to a few documents.

The broad idea of this project is to build a RAG system on top of some of the famous seminal research papers around Generative AI and LLMs, this system should be able to index the papers into a vector database, use good retrieval strategies to retrieve relevant contextual documents based on your input queries and generate proper responses. The focus of this project is not just to build a simple RAG system but explore various approaches and methodologies in each component when building the RAG system and then finally choose the approach and methodology which works best for you.

You are not compelled to use the sample data we provide for this project, if you have your own documents or data, you are more than welcome to use it (but validate your data and idea with us first), the project steps mentioned below would remain unchanged.

Project Goals:

The major project goals or objectives would include the following. You need to do all the key goals mentioned in the compulsory goals section and **at least 1 stretch goal** (you are welcome to do more).

Compulsory Goals:
- Get your own dataset or download the PDFs from here
- Load the files and index them in a vector database
- Experiment with different embedding models (open-source from huggingface and commercial ones like OpenAI)
- Experiment with various retrieval strategies (simple cosine to hybrid search and rerankers)
- Connect your vector database to an LLM and build a RAG pipeline
- Test the RAG pipeline on sample queries
- Try to also show the source of the generated response (which context documents were used to generate the response - top 3 will do)

Stretch Goals (at least 1):
- Advanced Option 1: Enhance this system into a multi-user conversational RAG system. Check out this example and try to adapt it
- Advanced Option 2: Build a streamlit or chainlit app on top of your RAG system
- Advanced Option 3: Try to enhance your system with web-search using Agentic Corrective RAG patterns (covered in the module on Agents in the LangChain course) or you can use our article as a reference

Milestones:

- Load and index RAG documents

- Experiment with Embeddings and build a Vector Database

- Experiment with various Retrieval strategies and finalize your retriever

- Build basic RAG system with answer sources

- Build enhanced RAG system or application based on stretch goals

Evaluation:
You are recommended to meet with a mentor at least once a week to review on your progress and then do a final project walkthrough and demo of your project once you are ready and then submit it.