Project Title	**Building a RAG-Powered Web Content Question-Answering System with Citation Support**
Skills take away From This Project	**1. Retrieval Augmented Generation (RAG) implementation** **2. API development using REST principles** **3. Vector database management** **4. Authentication system implementation** **5. UI development (Streamlit/Gradio)** **6. Cloud deployment (AWS/GCP/Heroku)** **7. Git version control**
Domain	**1. Artificial Intelligence & Machine Learning** **2. Natural Language Processing** **3. Information Retrieval**

Problem Statement: Website Retrieval Assignment - Develop a Retrieval Augmented Generation (RAG) system to answer questions based on content from specific websites.

Business Use Cases:

- Question-answering systems for company documentation/knowledge bases
- Automated customer support using website content
- Information retrieval from multiple web sources
- Content-aware chatbots with citation capabilities

Approach:

1. RAG System Implementation:
    - Build a system to extract and process website content
    - Implement vector database storage
    - Create a retrieval mechanism for relevant information
2. API Development:
    - Create endpoints for indexing (api/v1/index) and chat (api/v1/chat)
    - Implement authentication system
    - Deploy on the cloud platform
3. User Interface Development:

    - Build minimalistic UI using Streamlit/Gradio
    - Implement question input interface
    - Display answers with citations

Results: The system should:

- Successfully index and retrieve content from provided websites
- Generate coherent responses with accurate citations
- Handle user queries through API and UI interfaces
- Maintain secure access through authentication

Project Evaluation Metrics:

- Accuracy and relevance of generated answers
- Proper API implementation and deployment
- Correct citation of sources
- Code quality and documentation
- Authentication and UI implementation (bonus)

Technical Tags:

- RAG (Retrieval Augmented Generation)
- API Development
- Vector Databases
- Authentication
- UI Development
- Cloud Deployment
- Natural Language Processing

Data Set: Websites to be processed:

- https://huyenchip.com/2024/07/25/genai-platform.html
- https://lilianweng.github.io/posts/2024-07-07-hallucinatio
- https://jina.ai/news/what-is-colbert-and-late-interaction-and-why-they-matter-in-search/
- https://quoraengineering.quora.com/Building-Embedding-Search-at-Quora

Data Set Explanation: The provided websites contain technical content about AI/ML topics that must be processed, indexed, and made available for question-answering through the RAG system.

Project Deliverables:

1. Private GitHub repository containing:
    - Complete source code
    - Documentation
    - README file with setup instructions and API usage examples
2. Brief report covering:
    - Approach explanation
    - Challenges faced
    - Potential improvements

Project Guidelines:

1. API Development:
    - Follow REST principles
    - Implement proper error handling
    - Include authentication
    - Provide clear documentation
2. Citation System:
    - Accurate source referencing
    - Links to original content
    - Clear citation format
3. UI Development:
    - Clean, minimalistic design
    - Intuitive user interaction
    - Proper display of citations

PROJECT DOUBT CLARIFICATION SESSION (PROJECT AND CLASS DOUBTS)

About Session: The Project Doubt Clarification Session is a helpful resource for resolving questions and concerns about projects and class topics. It provides support in understanding project requirements, addressing code issues, and clarifying class concepts. The session aims to enhance comprehension and provide guidance to overcome challenges effectively.
Note: Book the slot at least before 12:00 Pm on the same day

Timing: Monday-Saturday (4:00PM to 5:00PM)

Booking link :https://forms.gle/XC553oSbMJ2Gcfug9

For DE/BADM project/class topic doubt slot clarification session:

Booking link : https://forms.gle/NtkQ4UV9cBV7Ac3C8

Session timing:

For DE: 04:00 pm to 5:00 pm every saturday
For BADM 05:00 to 07:00 pm every saturday

LIVE EVALUATION SESSION (CAPSTONE AND FINAL PROJECT)

About Session: The Live Evaluation Session for Capstone and Final Projects allows participants to showcase their projects and receive real-time feedback for improvement. It assesses project quality and provides an opportunity for discussion and evaluation.
Note: This form will Open only on Saturday (after 2 PM) and Sunday on Every Week

Timing:

For BADM and DE
Monday-Saturday (11:30AM to 1:00PM)

For DS and AIML
Monday-Saturday (05:30PM to 07:00PM)

Booking link : https://forms.gle/1m2Gsro41fLtZurRA