Project Title: Sanskrit Document Retrieval-Augmented Generation (RAG) System

1. Objective

The objective of this assignment is to design and implement a Retrieval-Augmented Generation (RAG) system capable of processing and answering queries based on Sanskrit documents. The system must operate fully on CPU-based inference.

2. Description

You are required to develop an end-to-end RAG pipeline that performs the following tasks:

1. Ingest Sanskrit documents (in `.txt` or `.pdf` format) on some domain.
2. Preprocess and index these documents for efficient retrieval.
3. Implement a query interface that accepts user input in Sanskrit or transliterated text.
4. Retrieve relevant context chunks from the indexed corpus.
5. Generate coherent responses using a CPU-based Large Language Model (LLM) integrated into the pipeline.

The RAG architecture should follow standard practices, ensuring modularity between the retriever and generator components.

3. Technical Requirements

- **Language of Source Documents:** Sanskrit (attached in mail)
- **Model Inference:** CPU only (no GPU usage permitted)
- **Core Components:**
    - Document Loader and Preprocessor
    - Retriever (Vector or Keyword based)

    - Generator (LLM-based text generator)

- **Frameworks (allowed):** Any open-source framework or library for retrieval and generation

- **Deployment Environment:** Local or lightweight containerized setup

4. Expected Deliverables

1. **Technical Report** detailing:
    - System architecture and flow
    - Details of the Sanskrit documents used
    - Preprocessing pipeline for Sanskrit documents
    - Retrieval and generation mechanisms
    - Performance observations (latency, accuracy/relevant metric, resource usage)
2. **Codebase** containing:
    - Complete, runnable implementation
    - Clear instructions in a `README.md` for setup and execution
    - Configuration files (if applicable)
3. **Demonstration Video (Optional)** showing end-to-end query–response flow.

5. Evaluation Criteria

Criteria	Description
System Architecture	Clarity, modularity, and alignment with RAG principles
Functionality	End-to-end working retrieval and generation on Sanskrit text
CPU Optimization	Efficient inference without GPU support
Code Quality	Clean, documented, and reproducible code

Report Quality	Depth of technical explanation and clarity

6. Submission Format

- Folder Name: `RAG_Sanskrit_<InternName>`
- Contents:
    - `/code/` – implementation scripts
    - `/data/` – sample Sanskrit documents used
    - `/report/` – final PDF report
    - `README.md` – setup and usage instructions

Submit your completed assignment in the **Github** repository