**STUDYMATE-AI ASSISTANT**

**Problem Statement:**

Description :

StudyMate is an AI-powered academic assistant that enables students to interact with their

study materials—such as textbooks, lecture notes, and research papers—in a

conversational, question-answering format. Instead of passively reading large PDF

documents or relying on manual searches for specific information, users can simply upload

one or more PDFs and ask natural-language questions. StudyMate responds with direct,

well-contextualized answers, referenced from the source content.

Expected Solutions:

1) Conversational Q&A from Academic PDFs – Enables students to ask natural-language questions and receive contextual answers grounded in their uploaded study materials.

2) Accurate Text Extraction and Preprocessing – Efficiently extracts and chunks content from multiple PDFs using PyMuPDF for high-quality downstream processing.

3) Semantic Search Using FAISS and Embeddings – Retrieves the most relevant text chunks using SentenceTransformers and FAISS for precise question matching.

4) LLM-Based Answer Generation – Uses IBM Watsonx’s Mixtral-8x7B-Instruct model to generate informative, grounded answers from retrieved content.

5) User-Friendly Local Interface – Provides an intuitive Streamlit-based frontend for seamless document upload, question input, and result visualization.

**SOLUTION:**

Proposed Solution / Plan of Action:

Our solution, StudyMate, is an AI-powered academic assistant that transforms static PDFs (textbooks, notes, and research papers) into interactive, queryable knowledge sources. Instead of students manually searching or skimming large documents, StudyMate allows them to upload multiple PDFs and ask questions in natural language. The system uses a Retrieval-Augmented Generation (RAG) approach, where semantic search retrieves the most relevant text chunks and an IBM Watsonx large language model generates clear, contextual answers grounded in the uploaded material.

The implementation plan is structured into four stages:

1. PDF Processing: Extract and clean text using PyMuPDF, then split into overlapping chunks for context continuity.

2. Semantic Retrieval: Encode chunks with SentenceTransformers and index them with FAISS for fast similarity search.

3. Answer Generation: Construct prompts from the retrieved context and query, then use IBM Watsonx LLM to generate concise answers.

4. User Interface: Deliver an intuitive Streamlit-based interface where users can upload PDFs, ask questions, view references, and download Q&A logs for revision.

How it Solves the Challenge:

StudyMate directly addresses the challenge of information overload and inefficiency in academic study. It enables:

Faster Learning: Students get instant, accurate answers without skimming through lengthy materials.

Higher Accuracy: Every answer is linked to source paragraphs for verification, reducing misinformation.

Better Exam Preparation: Q&A history can be saved as a transcript for revision.

Cross-Document Insights: Supports multi-PDF uploads, making it useful for research comparisons.

This solution meets the requirements by combining transparency, usability, and AI-driven intelligence into a single, student-friendly platform. It bridges the gap between static study materials and personalized, interactive learning.