

Hackathon Project

Of

AI STUDY ASSISTANT

Group Members:

Samiya Zahid	47446	samiyazahid004@gmail.com
Reshail Herman	47895	271062380@formanite.fccollege.edu.pk
Ahmad Ali Arshad	47839	ahmadaliarshad23@gmail.com
Muhammad Nazeem	48094	nazeemwani1954@gmail.com

Leader Name:	Abdur Rahman
Unique id:	47779
Gmail:	abdur.bscs2401@uoch.edu.pk

1 Introduction

The rapid growth of AI models has transformed how students learn, revise, and interact with study material. However, many advanced AI tools require heavy computational resources or paid APIs, making them inaccessible to beginners. This project solves this issue by building a **zero-cost, ultra-fast AI Study Assistant** using **Groq's free LLM API, Google Colab, LangChain, and Chroma vector database**.

The assistant can read PDFs, summarize chapters, generate MCQs, and answer questions directly from uploaded notes. This makes it extremely useful for students preparing for exams, teachers creating assessments, and developers learning how to work with Retrieval-Augmented Generation (RAG).

1.1 Document Purpose

The purpose of this document is to provide a detailed, professional, and structured explanation of the **AI Study Assistant** project. It outlines the motivation behind the system, the technologies used, the workflow, and the functionality of the solution. This document serves as a complete reference for instructors, evaluators, and team members to understand the design, development, and capabilities of the AI Study Assistant.

It clearly describes how the system uses **Groq API, LangChain, ChromaDB, Python, and Google Colab** to create an efficient learning support tool. The document also includes diagrams, workflow explanations, and a breakdown of every component to ensure clarity. Overall, the document aims to communicate the technical foundation, working process, and educational value of the project in a professional academic format.

1.2 Project Scope

The AI Study Assistant is a cloud-based intelligent learning system that helps students understand their study material quickly and easily. Instead of manually reading long PDF notes, the system allows users to upload their documents and get automatic summaries, MCQs, and answers to their questions. It uses modern AI technologies such as Groq LLM, LangChain, and Chroma Vector Database to retrieve information accurately and provide fast responses.

This platform addresses common academic challenges such as time-consuming revision, difficulty understanding complex topics, and lack of personalized study help. It gives students a smarter way to study, teachers a tool to prepare assessments, and learners at all levels a way to get instant explanations from their own notes.

Main Benefits:

Efficiency: Students can understand long chapters in minutes through AI-generated summaries.

Accuracy: Answers are generated directly from the uploaded PDF, ensuring content-based correctness.

Automation: The system automatically processes notes, creates MCQs, and retrieves relevant text without manual effort.

Accessibility: Works online on any device using Google Colab—no installation or paid tools

required.

Speed: Groq's ultra-fast AI models provide quick, real-time responses.

Convenience: Users can ask unlimited questions from their own study material.

The main goal of this system is to make studying easier, faster, and more effective by using advanced AI technology to support learners in understanding, revising, and practicing academic content.

1.3 Project Objectives

The main goal of this project is to create a complete end-to-end AI system that:

- Reads and understands PDF documents.
- Converts text into vector embeddings for accurate searching.
- Retrieves relevant document sections based on user questions.
- Uses Groq LLM to generate **fast, accurate answers**.
- Creates summaries and MCQs instantly.

All of this is designed to work smoothly inside **Google Colab**, with no paid services required.

1.4 Tools and Technologies Used

- **Google Colab:** A free cloud-based environment used to run all Python code without installations.
- **Groq API:** Groq provides extremely fast inference for models like Llama 3. The free API key is enough for high-speed responses.
- **Python:** Used for coding the entire workflow.
- **LangChain:** A powerful framework that simplifies creating LLM pipelines, retrieval systems, and RAG applications.
- **PyPDF:** Allows extraction of text from PDF study notes.
- **ChromaDB:** A vector store used to save and search document embeddings. It helps the AI retrieve exactly the right text segments.
- **HuggingFace Embeddings:** Used to convert PDF text into numerical vectors so the AI can understand relationships between different chunks of content.

2 System Workflow

The project follows a clear and efficient pipeline:

Step 1 — PDF Loading

The system uses **PyPDFLoader** to read the uploaded PDF and convert it into text documents.

Step 2 — Text Splitting

Large PDF content is broken into smaller chunks (800 characters with 150 overlap). This improves accuracy when the user asks questions.

Step 3 — Vector Database Creation

Each text chunk is converted into embeddings using the model **sentence-transformers/all-MiniLM-L6-v2**.

These embeddings are stored in a **ChromaDB** vector store.

Step 4 — LLM Setup (Groq)

The project uses the **Llama 3 8B** model through Groq's API. This model is fast, efficient, and completely free.

Step 5 — Retrieval QA Chain

LangChain's **RetrievalQA** chain is used to:

1. Search the vector database,
2. Identify the most relevant chunk,
3. Provide the answer using Groq LLM.

Step 6 — Summarization

A custom summarizer was built that takes any text and converts it into simplified bullet points.

Step 7 — MCQ Generation

The assistant can generate **10 MCQs with answers** from any text chunk.

2.1 Key Features of AI Assistant

- **PDF-Based Question Answering**

Students can upload lecture slides or ebooks and ask any question directly from the content.

- **Smart Summaries**

Long chapters are reduced into short, understandable bullet points.

- **Automatic MCQ Generation**

The model creates exam-style multiple-choice questions for revision.

- **Fast and Free**

Thanks to Groq, responses are much faster than most paid APIs.

- **Works in Any Device**

Since everything runs in Colab, there is no installation required.

2.2 Use cases

- **Students preparing for exams using class notes:**

The AI Study Assistant helps students quickly revise lengthy chapters by generating concise summaries, clear explanations, and topic-focused MCQs. Instead of reading entire PDFs, learners can ask direct questions and get instant, content-based answers, making exam preparation more efficient.

- **Teachers generating quizzes and summaries:**

Educators can use the system to automatically create high-quality MCQs, chapter summaries, and question banks derived directly from their teaching material. This reduces manual workload and supports the creation of engaging classroom activities and assessments.

- **Developers learning about vector databases and RAG:**

The project serves as a practical example for developers interested in learning how Retrieval-Augmented Generation works. It demonstrates how to combine embeddings, vector search, and LLMs to build intelligent applications, making it a great learning resource for AI and NLP enthusiasts.

- **Researchers analyzing PDFs efficiently:**

Researchers who work with large documents, reports, or academic papers can use the assistant to extract important information, highlight relevant sections, and ask questions directly from the text. This speeds up literature review and data extraction.

- **Hackathon participants showcasing AI-powered tools:**

Because the entire system is fast, free, and easily deployable on Google Colab, it is ideal for hackathons. Participants can demonstrate a fully functional AI tool capable of text analysis, summarization, and interactive Q&A—all powered by modern open-source technologies.

3 Conclusion

This project successfully demonstrates how powerful and practical modern AI systems can be when built using free and accessible tools. By integrating Groq LLM, LangChain, and ChromaDB, the AI Study Assistant showcases a complete end-to-end solution capable of delivering fast, accurate, and context-aware academic support. The system effectively performs Retrieval-Augmented Generation (RAG), allowing users to receive precise answers directly from their uploaded study materials rather than generic AI responses.

Beyond simple question answering, the assistant offers a full suite of learning tools, including PDF text extraction, intelligent summarization, and automatic MCQ generation. These features work together to create an adaptive study environment that supports students in understanding complex topics, revising efficiently, and testing their knowledge.