Project Proposal: AI-Powered Student Intelligent Study Companion (MVP)

Submitted To: University Research Board, Dean of Academics
Submitted By: Department of Computer Science
Institution: Institute of Space Technology, Islamabad
Date: 23 APR, 2025

1. Executive Summary

This proposal outlines the development and implementation of an innovative **AI-powered student Success Platform**, beginning with a Minimum Viable Product (MVP) focused on an **Intelligent Study Companion**. This tool is designed to address the evolving academic needs of university students by providing personalized, interactive, and accessible study support. Initially targeting students in **Computer Science, Artificial Intelligence, and Data Science** programs, the platform will align with the **Higher Education Commission (HEC) approved curriculum**.

Leveraging **OpenAI's GPT-4 API (including advanced Vision capabilities)** and **Retrieval-Augmented Generation (RAG)** techniques, the Study Companion will allow students to upload diverse academic materials (notes, PDFs, presentations, images of whiteboards) and interact with them through contextual Q&A, summarization, and automated quiz generation. Accessible via a user-friendly **Streamlit web interface**, the MVP aims to enhance understanding, improve retention, and promote self-directed learning.

Future phases envision expansion to other academic departments and the introduction of faculty-centric modules for AI-assisted lesson planning and assessment generation. This project represents a strategic investment in leveraging cutting-edge AI to enrich the educational experience at IST, positioning the institution as a leader in educational technology adoption. We seek approval and necessary resources to initiate the development of this promising MVP.

2. Introduction & Problem Statement

University students, particularly in rapidly evolving fields like Computer Science, AI, and Data Science, face significant academic challenges. These include managing vast amounts of information from diverse sources (lectures, textbooks, notes, online resources), understanding complex concepts, finding timely answers to specific doubts, and effectively preparing for assessments. Current study methods often lack personalization and immediate feedback, potentially leading to knowledge gaps and decreased engagement.

Furthermore, ensuring consistent understanding aligned with the HEC curriculum standards across various course materials can be difficult for students navigating their studies independently. There is a clear need for an intelligent tool that can centralize learning resources, provide context-aware assistance, adapt to individual learning paces, and make studying more interactive and effective.

This project directly addresses these challenges by proposing an AI-driven companion that acts as a personalized tutor, available 24/7, capable of understanding and processing student-specific course materials.

3. Proposed Solution: AI Study Companion MVP

We propose the development of an AI-Powered Study Companion MVP – a web-based application built using Streamlit for accessibility and ease of use. The core of the solution lies in its ability to:

1. **Ingest and Understand Diverse Academic Content:** Students can upload materials in various formats (PDF, DOCX, PPTX, images of notes/diagrams). The system uses advanced AI (GPT-4 with Vision) to parse, extract, and comprehend this content.
2. **Provide Contextual Interaction:** Through a chat interface, students can ask questions, request summaries, or ask for explanations related to their uploaded documents. The AI maintains context within the conversation and the provided materials.
3. **Facilitate Active Recall and Assessment:** The tool can automatically generate relevant quizzes (e.g., MCQs, short answer prompts) based on the uploaded content, helping students test their understanding and prepare for exams.
4. **Align with HEC Curriculum:** The system will be primed with knowledge of the HEC curriculum outlines for CS, AI, and DS to ensure generated content and explanations are relevant and accurate within the local academic context.

This MVP serves as a foundational step towards a comprehensive AI-driven academic support system for the university.

4. Objectives

The primary objectives of this MVP phase are:

1. **Develop and Deploy a Functional MVP:** Create a working prototype of the AI Study Companion accessible to a pilot group of CS, AI, and DS students.
2. **Implement Core Features:** Successfully integrate functionalities for multi-format content upload, OCR/Vision analysis, RAG-based Q&A, quiz generation, context recall, simplified explanations, and basic progress tracking.

3. **Ensure HEC Curriculum Relevance:** Configure the AI's knowledge base and response generation to align with specified HEC curriculum guidelines for the target courses.
4. **Validate Technical Feasibility:** Demonstrate the effective integration and performance of the chosen technology stack (Streamlit, FastAPI, OpenAI API, LangChain, Vector DB).
5. **Gather User Feedback:** Collect initial feedback from pilot users to assess usability, effectiveness, and identify areas for improvement for future iterations.
6. **Assess Scalability Potential:** Evaluate the architecture's capacity for future expansion to include more users, departments, and features.

5. Scope

In Scope for MVP:

- Target User Group: Students in Computer Science, Artificial Intelligence, and Data Science departments at IST.
- Core Functionality: Features listed under Section 6 (MVP Features).
- Content Focus: HEC-aligned curriculum for the specified departments.
- Platform: Web application accessible via browser (Streamlit frontend).
- Language: English (primary).
- Deployment: Secure cloud environment (e.g., leveraging Supabase or university-approved cloud services).

Out of Scope for MVP:

- Support for departments beyond CS, AI, and DS.
- Faculty-specific modules (Lesson Planner, Faculty Quiz Generator).
- Integration with existing University Systems (LMS, SSO) – planned for future phases.
- Advanced analytics dashboard.
- Multilingual support beyond English.
- Mobile application version.

6. MVP Features (Detailed)

6.1 HEC-Aligned Knowledge Support

- **Mechanism:** The system will be grounded using HEC curriculum documents for CS, AI, and DS. Responses related to course topics will reference these guidelines to ensure relevance and standardization.
- **Benefit:** Provides students with reliable, curriculum-conformant academic assistance.

6.2 Multi-Format Content Upload & Vision Analysis

- **Mechanism:** Utilizes libraries (PyMuPDF, python-docx, python-pptx) for text extraction from documents and OpenAI's GPT-4 Vision API for analyzing images (handwritten notes, diagrams, whiteboard photos). Advanced OCR and visual understanding models extract textual and semantic information.
- **Functionality:** Enables summarization, Q&A, and quiz generation directly from visual or text-based uploaded materials.
- **Benefit:** Allows students to use their preferred study materials, regardless of format.

6.3 Advanced Retrieval-Augmented Generation (RAG)

- **Mechanism:** Uploaded documents are chunked, converted into vector embeddings (using OpenAI Embeddings), and stored in a vector database (FAISS or Supabase Vector Store). When a user asks a question, the system retrieves the most semantically relevant document chunks and feeds them, along with the query, to the GPT-4 model for a contextually accurate, grounded response.
- **Benefit:** Reduces AI "hallucinations" and ensures answers are based directly on the provided course materials.

6.4 Contextual Q&A with Memory

- **Mechanism:** The chatbot maintains conversation history within a session. It leverages this memory, alongside the context from uploaded documents, to understand follow-up questions, resolve ambiguities, and provide coherent, ongoing assistance, mimicking a tutor's conversational flow.
- **Benefit:** Creates a more natural and effective learning interaction.

6.5 Automated Quiz Generator

- **Mechanism:** Students can select uploaded documents or specific topics within them. The AI analyzes the content and generates relevant quizzes (MCQs, True/False, potentially short answer) to test comprehension.
- **Benefit:** Promotes active recall and self-assessment, key elements of effective learning.
-
- **Benefit:** Helps students organize their studies and identify areas requiring further attention (potential for future automated suggestions).

7. Technical Architecture

Layer	Technology / Tool	Purpose
Frontend	Streamlit	Rapid development of an interactive web UI for user interaction

Backend API	FastAPI	Robust, asynchronous API framework for handling requests
AI Core Engine	OpenAI GPT-4 API (incl. Vision, Embeddings)	State-of-the-art LLM for generation, understanding, vision
Orchestration	LangChain	Framework for chaining AI components (parsing, RAG, memory)
Document Parsing	PyMuPDF, python-docx, python-pptx, OCR Libraries	Extracting text and data from various document formats
RAG System	LangChain + Vector Store (FAISS / Supabase Vector)	Document chunking, embedding, retrieval for contextual Q&A
Database & Storage	Supabase (PostgreSQL DB + Object Storage)	Storing user data, processed documents, vector embeddings
Deployment	University Cloud / Supabase Hosting / Other PaaS	Hosting the application infrastructure

Workflow: User uploads content via Streamlit -> FastAPI backend receives request -> LangChain orchestrates parsing & embedding -> Embeddings stored in Supabase Vector Store -> User asks question -> FastAPI receives query -> LangChain retrieves relevant chunks from Vector Store -> LangChain sends query + context to GPT-4 -> GPT-4 generates response -> FastAPI sends response to Streamlit -> User sees answer.

8. Methodology & Implementation Plan

We will adopt an Agile development approach, focusing on iterative development and feedback cycles.

1. **Phase 1: Setup & Core Infrastructure**
    - Set up development environments, cloud resources (Supabase), API keys.
    - Establish backend structure (FastAPI) and frontend base (Streamlit).
    - Integrate basic OpenAI API connectivity.
    - Set up document storage.
2. **Phase 2: Content Processing & RAG Implementation**
    - Implement document parsers for PDF, DOCX, PPTX.
    - Integrate GPT-4 Vision for image analysis and OCR.
    - Set up the vector store (FAISS/Supabase) and embedding pipeline.
    - Develop and refine the RAG system for contextual Q&A.
3. **Phase 3: Feature Development**
    - Build the Quiz Generator module.

- Implement Contextual Memory and ELI5 mode.
- Develop the basic Progress Tracker UI and logic.
- Refine the Streamlit UI/UX based on initial internal testing.

9. Risks & Mitigation Strategies

Risk	Likelihood	Impact	Mitigation Strategy
AI Hallucinations/Inaccuracy	Medium	High	Strong reliance on RAG, prompt engineering, clear sourcing (when possible), user feedback loop for corrections.
OpenAI API Costs & Rate Limits	Medium	Medium	Monitor usage closely, optimize API calls, implement caching where feasible, and explore potential educational grants.
Data Privacy & Security	Low	High	Use secure storage (Supabase), anonymize data where possible, adhere to university data policies, and no PII storage initially.
Low User Adoption (Pilot)	Medium	Medium	Engage pilot users early, focus on UI/UX simplicity, clearly communicate benefits, and actively solicit feedback.
Technical Implementation Challenges	Medium	Medium	Adopt Agile methodology, experienced team, use robust frameworks (LangChain), allocate buffer time in the schedule.
Scope Creep	Medium	Medium	Strictly adhere to the defined MVP scope, manage stakeholder expectations, and document change requests for future phases.

10. Future Scope & Scalability

Following a successful MVP launch and evaluation, the platform has significant potential for expansion:

- **Student-Facing Enhancements:**
    - More sophisticated progress tracking and personalized study plans.
    - Support for collaborative study groups within the platform.
    - Integration with external tools (e.g., calendar, citation managers).
    - Proactive learning suggestions based on performance.
- **Faculty Modules:**
    - **Smart Lesson Preparation:** AI assistant for generating draft lesson plans based on topics, methods, and curriculum requirements (editable by faculty).
    - **AI-Powered Assessment Tools:** Generate diverse quiz types, suggest grading rubrics, analyze student response patterns from uploaded lecture notes/slides.
- **Platform Expansion:**
    - **Cross-Departmental Rollout:** Gradually extend support to Engineering, Business, Medicine, Arts & Humanities, tailoring content alignment for each.
    - **Multilingual Support:** Incorporate models capable of understanding and responding in regional languages.
    - **LMS & University System Integration:** Enable Single Sign-On (SSO), sync progress with Moodle/Google Classroom, provide faculty dashboards via portal integration.
    - **Advanced Analytics:** Develop dashboards for faculty and administration to track usage patterns and learning outcomes (anonymized).

This phased approach ensures manageable development cycles while building towards a comprehensive, university-wide Student Success Platform.

10. Conclusion

The proposed AI-Powered Student Study Companion represents a timely and strategic initiative to enhance the academic experience at IST. By leveraging state of the art AI technology within a focused MVP, we can provide immediate value to students in key technical departments while establishing a foundation for future innovation in teaching and learning across the institution. This project not only addresses current student needs for personalized and accessible academic support but also positions the university at the forefront of educational technology adoption.