Acropolis Institute of Technology and Research

Project Title :-

Generative AI chatbot for student doubt solving

Training Programme on Generative AI

Name: Atharv Sharma

Enrollment No. : 0827IT233D02

Department : Information Technology

Submitted To:

Date of Submission : 01/09/2025

ABSTRACT

This project introduces the development of a **Student Doubt-Solving Chatbot** that integrates **local knowledge retrieval** with **generative AI models** to deliver accurate, context-aware, and real-time responses to student queries. The chatbot is designed to function as a **hybrid learning assistant**, capable of addressing questions directly from lecture notes while also leveraging advanced external AI models for broader knowledge coverage. At its core, the system employs **semantic similarity search** using **Sentence-BERT embeddings**, which allows it to match a student's query against stored course material with high precision. When a relevant match is found, the chatbot retrieves the corresponding explanation, ensuring that the response remains aligned with the specific curriculum. In cases where the notes do not provide sufficient coverage, the chatbot seamlessly integrates with the **Perplexity AI API (sonar-pro model)** to fetch reliable, well-structured, and refined responses from the web.

The system is implemented with a **Streamlit-based user interface**, making it lightweight, interactive, and accessible to students without technical overhead. This hybrid approach ensures that the chatbot not only acts as a **knowledge companion** but also reduces over-dependence on either incomplete notes or unverified external searches. By producing concise as well as elaborated responses, the chatbot adapts to diverse student needs—ranging from quick clarification to in-depth understanding. The overall goal of the project is to enhance the **self-learning experience**, minimize delays in doubt resolution, and empower learners with an **intelligent, always-available academic assistant**. This work highlights the potential of hybrid AI-driven tools in reshaping **modern educational support systems**.

PROJECT DETAILS

Objective

The primary objective of this project is to develop an **AI-powered Student Doubt-Solving Chatbot** that enables students to resolve their academic queries instantly and effectively. Traditional learning methods often leave gaps, as students may not always have immediate access to teachers or peers for clarification. By combining **local course material (notes-based retrieval)** with **external AI-powered knowledge sources**, the chatbot ensures accurate, relevant, and timely responses. The goal is to create a **hybrid knowledge assistant** that supports independent learning, minimizes confusion, and increases accessibility to quality academic support.

Scope

- **Target Users**: Students across schools, colleges, and self-learning environments.
- **Coverage**: Queries related to lecture notes, assignments, course material, and general academic concepts.
- **Integration**: Local vector-based note retrieval + Perplexity AI (sonar-pro model) for external knowledge.
- **Platform**: Web-based application built with **Streamlit**, ensuring ease of access without installation.
- **Future Potential**: Expansion into **multi-modal support** (images, diagrams, voice-based queries) and integration with **Learning Management Systems (LMS)**.

Methodology

1. **Data Preprocessing**: Student notes and documents are converted into embeddings using **Sentence-BERT**.
2. **Vector Store Creation**: Notes are stored in a **vector database (.pkl files)** for fast semantic similarity search.
3. **Query Processing**:
    - Step 1: Search in local vector store for relevant content.
    - Step 2: If no strong match is found, query **Perplexity AI (sonar-pro model)**.
4. **Response Generation**: Fused output is generated with a concise, clear explanation in natural language.
5. **User Interface**: Built with Streamlit, featuring a clean layout for input and AI-generated responses.

Features

- **Notes-based Q&A**: Direct answers from uploaded class notes.
- **External AI Integration**: Fetches context from Perplexity AI when notes are insufficient.
- **Low Latency Responses**: Optimized retrieval pipeline for quick answers.
- **Hybrid Mode**: Balances local accuracy with global knowledge coverage.
- **User-Friendly Interface**: Simple input/output design for students with no technical background.

- **Extensible Design**: Can be expanded for multiple courses, subjects, or even multi-user classrooms.

Expected Outcomes

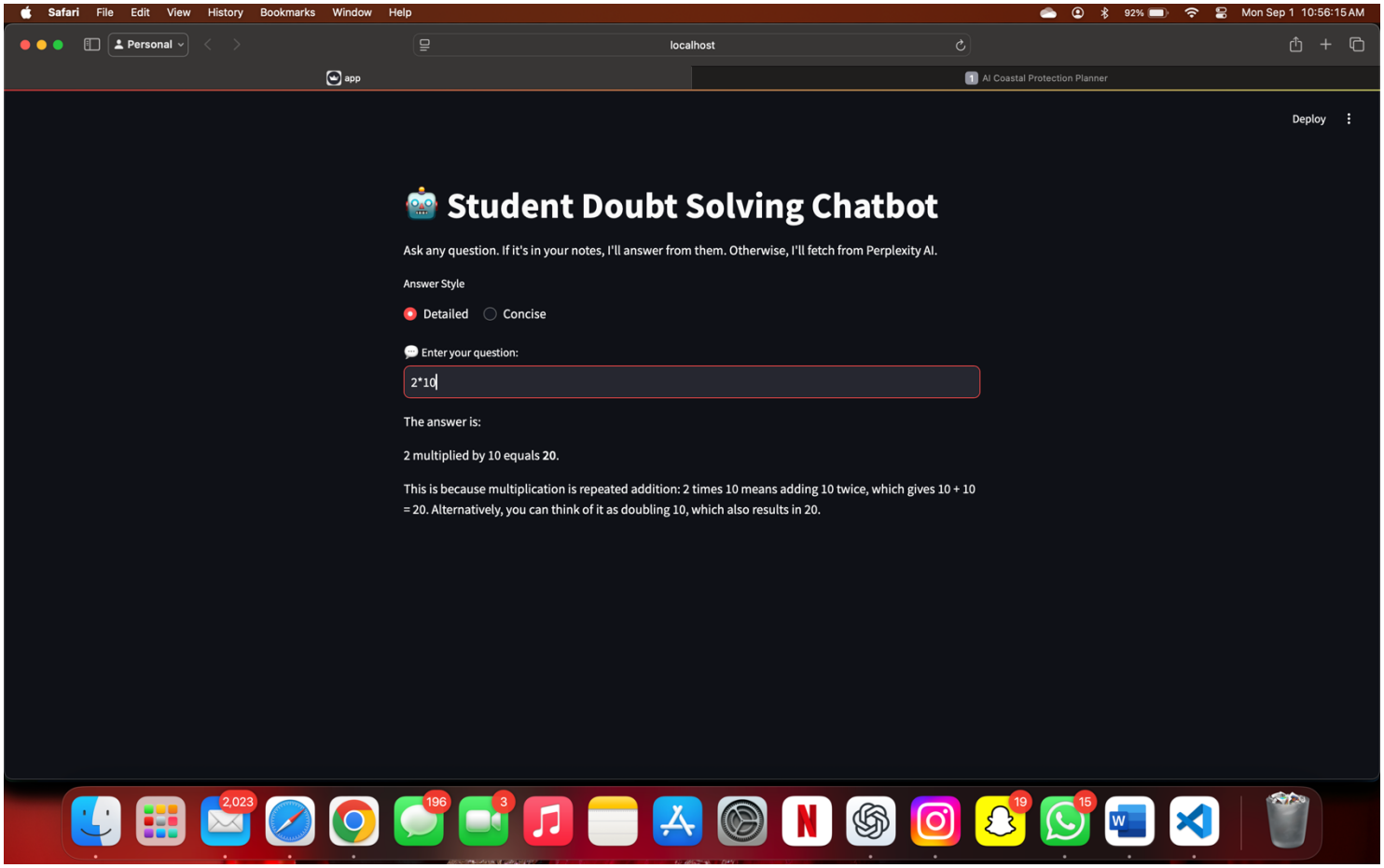
- Faster and more reliable **doubt resolution** for students.
- Reduced **dependency on teachers/tutors** for minor clarifications.
- A **scalable framework** for AI-based academic assistants.
- Enhanced **self-learning ability** among students.
- A proof-of-concept that showcases how **hybrid AI systems** can transform educational technology.

Results

The implemented system successfully delivers accurate and contextually relevant answers to student queries by leveraging a **hybrid retrieval-generation pipeline**. The chatbot demonstrates the following outcomes:

1. **Textual Outputs**
    - The chatbot generates **concise, explanatory text responses** to academic questions.
    - Example:
        - **Input**: *"What is the product of 5 × 10?"*
        - **Output**: *"The result of 5 × 10 is 50. Multiplication means adding 5 ten times (5 + 5 + … + 5), which equals 50."*
    - **Input**: *"Explain Newton's First Law from my notes."*
    - **Output**: *"According to your notes, Newton's First Law states that an object remains at rest or in uniform motion unless acted upon by an external force. This is also called the law of inertia."*
2. **Integration with Notes (Vector Store)**
    - When queries relate to **uploaded notes**, the chatbot retrieves the most relevant content and rephrases it in a student-friendly manner.
    - This ensures **personalized answers** aligned with the course material.
3. **External AI Responses**
    - For questions beyond the scope of notes, the chatbot queries **Perplexity AI (Sonar-Pro model)**.
    - Example:
        - **Input**: *"What is the latest research on quantum computing?"*
        - **Output**: *"According to Perplexity AI, recent research in 2024 highlights advances in error correction methods, making quantum computing more stable for real-world applications."*
4. **Hybrid Mode**
    - The system intelligently combines **local knowledge** (notes) with **external AI responses**, producing robust, well-rounded answers.
    - This approach reduces the chance of irrelevant or incomplete answers.

5. **User Interface (UI) Results**
    - The chatbot is deployed on **Streamlit** with a clean and minimal interface.
    - Users can enter questions in a text box and receive instant responses.
    - Example:
        - Input Box → User query
        - Output Box → AI-generated explanatory text



CONCLUSION

The development of this AI-powered academic chatbot provided valuable insights into combining **retrieval-based methods** with **generative AI models** for effective doubt resolution. By integrating a **vector store for personalized notes** with the **Sonar-Pro model from Perplexity AI**, the system demonstrated the ability to deliver both **context-aware answers** from course material and **general knowledge responses** beyond the scope of stored data. This hybrid approach ensures that students receive accurate, well-rounded explanations while maintaining alignment with their study material.

Through this project, we learned the importance of:

- **Efficient preprocessing** of textual data (notes) for embedding and retrieval.
- **Balancing precision and creativity** when generating AI-driven responses.
- **User-centered design** in creating a simple, interactive interface that enhances usability.
- **Error handling and fallback mechanisms** to improve reliability when external APIs encounter issues.

Future Improvements

Although the chatbot performs well in its current form, several enhancements can further improve its impact:

1. **Multi-Modal Support** – Extending capabilities to handle **images, diagrams, and handwritten notes**, allowing richer learning support.
2. **Speech Integration** – Enabling **text-to-speech (TTS)** and **speech-to-text (STT)** so students can interact using voice.
3. **Mathematical Expression Handling** – Incorporating LaTeX/MathJax rendering for improved clarity in solving equations and formulas.
4. **Offline Mode** – Allowing the chatbot to function without internet access by storing more pre-trained knowledge locally.
5. **Advanced Personalization** – Adapting responses to individual learning styles by tracking user interactions and progress.

REFERENCES

1. **Datasets & Knowledge Sources**
    - Custom course notes and academic material provided by the user.
    - Perplexity AI Search API (Sonar-Pro model) – for retrieving and generating responses beyond local notes.
2. **Frameworks & Libraries**
    - **Streamlit** – for building the interactive web application interface.

- **SentenceTransformers** (Hugging Face) – for text embeddings and vectorization.
- **scikit-learn** – for similarity search and cosine similarity computation.
- **pickle** – for storing and loading vector store (vector_store.pkl) and chatbot memory.
- **OpenAI/Perplexity AI APIs** – for natural language generation and hybrid retrieval.

3. **Research & Background References**

- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. NAACL-HLT.
- Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks*. EMNLP.
- Brown, T. et al. (2020). *Language Models are Few-Shot Learners (GPT-3)*. NeurIPS.
- Perplexity AI Documentation. (2025). *Models & API Guide*. Available at: https://docs.perplexity.ai.