PDF Chatbot API Documentation

This API enables users to upload a PDF document and interact with its content using a chatbot powered by Groq's LLM. It leverages FastAPI for serving endpoints, ChromaDB for vector storage, and LangChain + Sentence Transformers for processing.

Base URL:

http://localhost:8000/

Environment Variables:

Create your own .env file with your Groq API key:

For example:
GROQ_API_KEY=your_groq_api_key

Endpoints

POST /UploadingPDF/

Upload and process a PDF file
Description
Extracts text from the uploaded PDF, chunks it, and stores the data as vector embeddings in ChromaDB.

Request

- **Method:** POST
- **Content-Type:** multipart/form-data
- **Body:**
  - file: PDF file (required)

Response

- 200 OK: PDF processed successfully
- 400: Error Reading PDF
- 401: Error Chunking Text/PDF
- 402: Error Creating Embeddings Or Storing in ChromaDB

POST /ChatBot/

Query the PDF content through the chatbot
Description
Sends a question to the LLM, retrieves relevant context from ChromaDB, and returns a contextual, friendly answer.

Request

- Method: POST
- Content-Type: application/x-www-form-urlencoded
- Body:
  - query: Your question (string)

Response
200 OK: Returns the chatbot response (string)
403: Error Retrieving Context
404: Error Generating LLM Response

LLM Behavior

The chatbot uses the llama3-8b-8192 model from Groq's API. Responses are grounded in retrieved context and follow these rules:

- Only use relevant context from the document.
- Avoid hallucinations or fabricated info.
- If the answer isn’t found in the document, respond with:
- "I am sorry, I could not find the answer in the document provided."
- Redirect inappropriate queries.