

Project 14: PDF Summarizer (Upload + Summary)

Description:

This tool allows users to upload a PDF file (e.g., a research paper, report, or document), extracts the text, and provides a concise summary. It's a fantastic use case for knowledge workers, students, or anyone wanting a quick overview of lengthy PDFs.

pdf_summarizer.py

```
import openai
import os
import gradio as gr
import PyPDF2

# Load your OpenAI API key from environment variable
openai.api_key = os.getenv("OPENAI_API_KEY")

# Function to extract text from uploaded PDF
def extract_pdf_text(pdf_file):
    pdf_reader = PyPDF2.PdfReader(pdf_file) # Create a PDF reader object
    full_text = ""

    # Loop through all pages and extract text
    for page in pdf_reader.pages:
        full_text += page.extract_text() or "" # Append each page's text

    return full_text

# Function to summarize the extracted PDF content
def summarize_pdf(file):
    try:
        # Step 1: Extract text from the uploaded PDF file
        extracted_text = extract_pdf_text(file)

        if not extracted_text.strip():
            return "Could not extract readable text from the PDF. Try another"

        # Step 2: Prepare messages for OpenAI to summarize
```

```

messages = [
    {
        "role": "system",
        "content": (
            "You are an AI document assistant. Your job is to read thr
            "sections, or arguments in clear and concise language."
        )
    },
    {
        "role": "user",
        "content": f"Summarize this document:\n\n{extracted_text}"
    }
]

# Step 3: Call OpenAI ChatCompletion API
response = openai.ChatCompletion.create(
    model="gpt-3.5-turbo", # You can switch to GPT-4
    messages=messages
)

# Step 4: Return the summary
return response["choices"][0]["message"]["content"].strip()

except Exception as e:
    # Handle errors
    return f"Error: {str(e)}"

# Gradio interface: Upload PDF and get summary
iface = gr.Interface(
    fn=summarize_pdf, # Core function
    inputs=gr.File(file_types=[".pdf"]), # Upload only PDF files
    outputs="text", # Display summary text
    title="📄 PDF Summarizer", # App title
    description=(
        "Upload a PDF document (e.g., paper, report, article) and get a summar
        "Useful for quickly understanding long documents without reading every
    )
)

# Launch the Gradio web app
iface.launch()

```