# **Learning Log: AI-Powered PDF Summarizer with Streamlit & Hugging Face**

## **Project Summary**

This project aimed to build a web-based PDF summarization tool that allows users to upload a PDF and receive an AI-generated summary using Hugging Face transformers. The frontend was created using Streamlit with a focus on interactivity and UI polish.

## **Technologies Used**

* **Streamlit** – Web app framework for Python
* **Hugging Face Transformers** – Text summarization (facebook/bart-large-cnn)
* **PDFMiner** – For extracting text from PDFs
* **Python** – Core logic
* **UUID** – Session-based uniqueness for file resets
* **HTML/CSS** – Custom styling of buttons, title, and layout

## **Topics Learned**

| **Concept** | **Description** |
| --- | --- |
| streamlit.file\_uploader | Handling file uploads in Streamlit |
| streamlit.download\_button | Letting users download generated summaries |
| Hugging Face pipeline() | Using pretrained models for summarization |
| pdfminer.high\_level.extract\_text() | Extracting text content from PDF files |
| Session state management | Using st.session\_state to persist data across reruns |
| UI/UX enhancements | Customizing button styles, hover effects, download icons |
| App reset mechanics | Resetting session + file upload using dynamic keys |
| Chunking large text | Splitting PDF text into smaller parts for model limits |
| st.rerun() | Reloading the app to start fresh |

## **Issues Encountered & Solutions**

| **Issue** | **Solution** |
| --- | --- |
| set\_page\_config() not first | Moved st.set\_page\_config() to the first Streamlit command |
| File upload not clearing on reset | Added key=str(uuid.uuid4()) to st.file\_uploader and regenerated key on reset |
| st.experimental\_rerun() not found | Replaced with st.rerun() for Streamlit v1.31+ |
| “Download Summary” button not centered | Used st.columns([1, 2, 1]) and placed button inside center column |
| Session reset not clearing file upload | Manually deleted session keys and changed upload key |
| Summary not persisting | Stored summary in st.session\_state to persist across reruns |
| summary not defined error | Ensured summary is generated and accessed within the same conditional scope |

## 

## **Final App Features**

* Upload PDF
* Extract text from PDF using PDFMiner
* Summarize using BART model (Hugging Face)
* Download as .txt
* Reset with “Summarize Another PDF”
* Styled UI: buttons, title, hover, spacing
* File state reset across reruns

## **Commands:**

* pip install streamlit transformers torch pdfminer.six
* Create [app.py](http://app.py) file
* streamlit run app.py