Nepali PDF Chatbot - Working Flow

Introduction

The Nepali PDF Chatbot is a Streamlit-based application that enables users to upload a Nepali PDF, extract and process its content, and ask questions to retrieve relevant information in English or Nepali. This document outlines the working flow of the chatbot.

1. Starting the Application

- The application is launched using the following command:
  - Stremlit run chatbot.py
- After launching, the console displays a local URL and a network URL where the chatbot can be accessed.

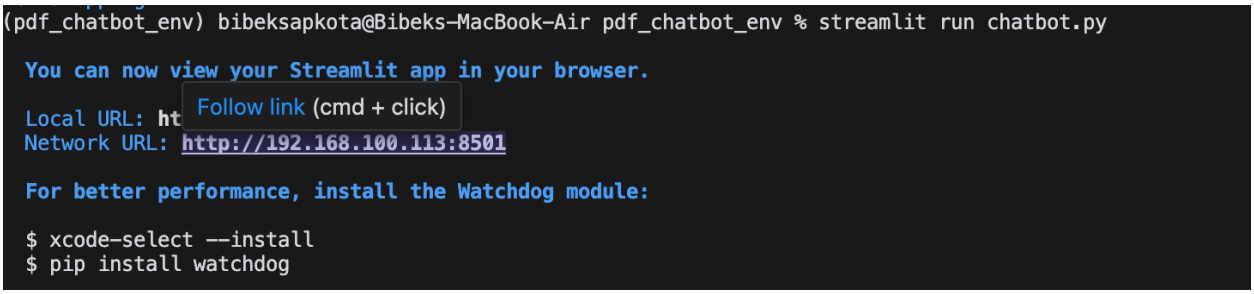

Fig1- showing Streamlit running in the terminal with a URL

2. Uploading a Nepali PDF

- Users can drag and drop or browse to upload a PDF document.
- The application supports PDF files up to 200MB.
- Users can also select their preferred response language: English or Nepali.

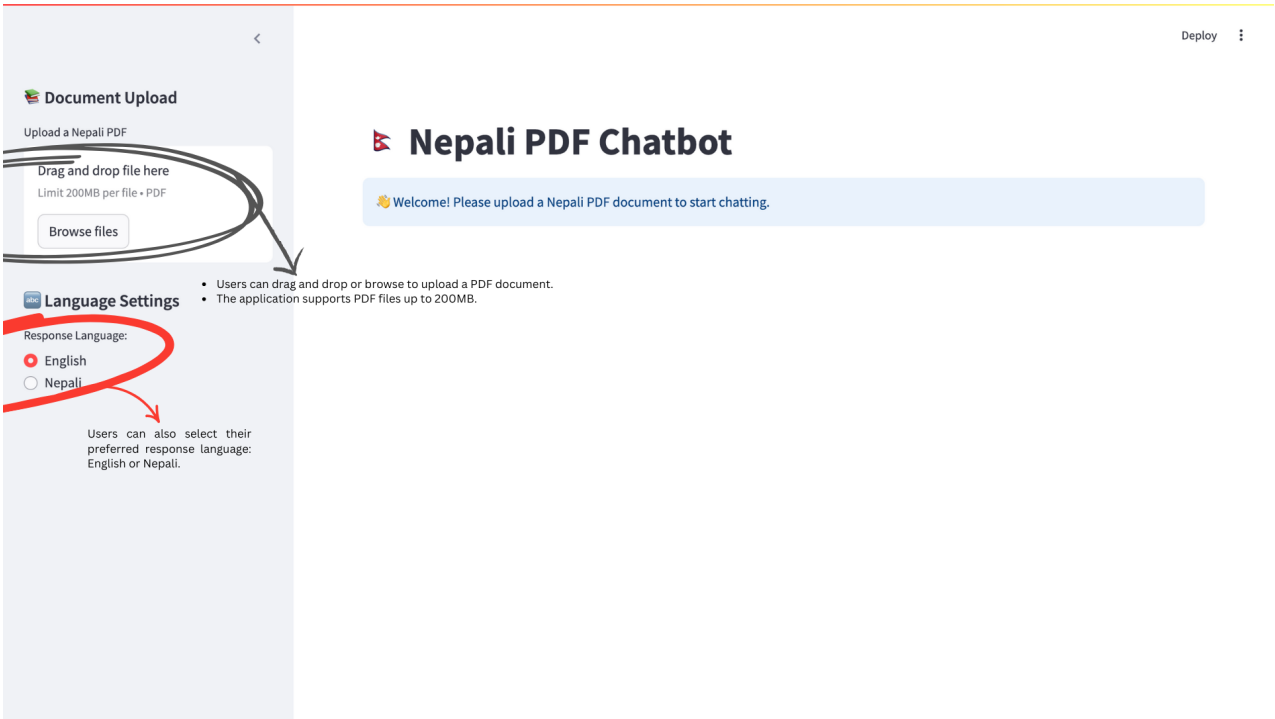

Fig2- showing the file upload interface.

3. Processing the PDF

- Once a PDF is uploaded, the system processes the document.
- The application extracts text from the PDF pages and converts them into embeddings for efficient searching.
- The success message “PDF processed successfully” appears once the process is complete.

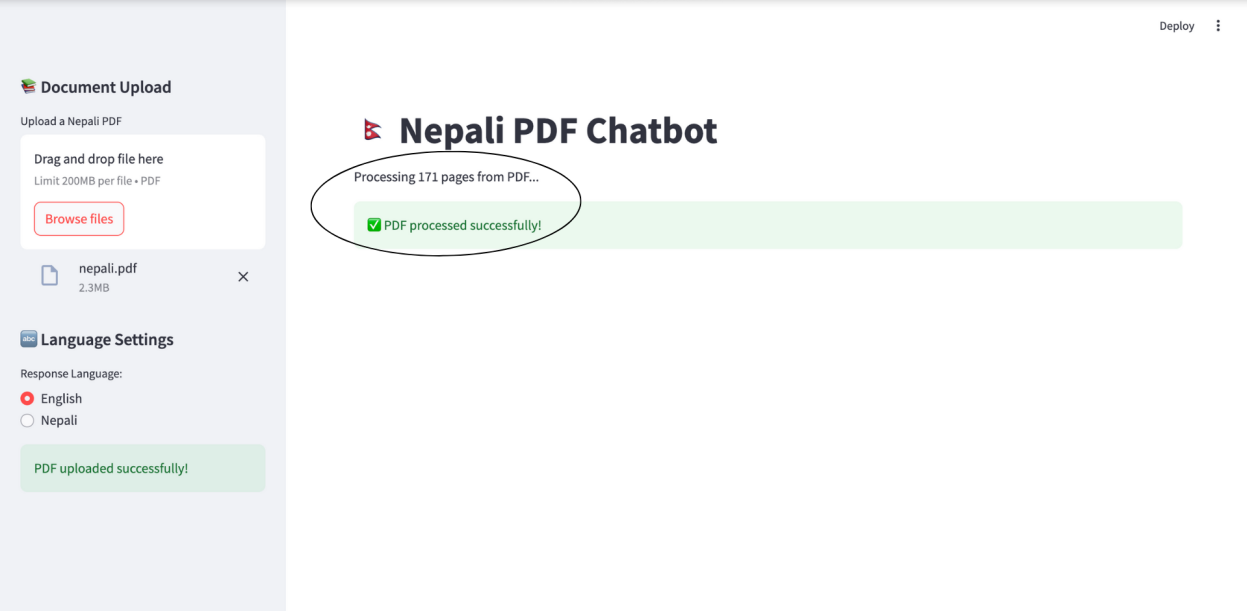

Fig3- showing the document processing screen.

4. Asking Questions

- Users can type their queries in the chat input field.
- The chatbot retrieves relevant answers based on the extracted text and displays the response in the selected language (English or Nepali).

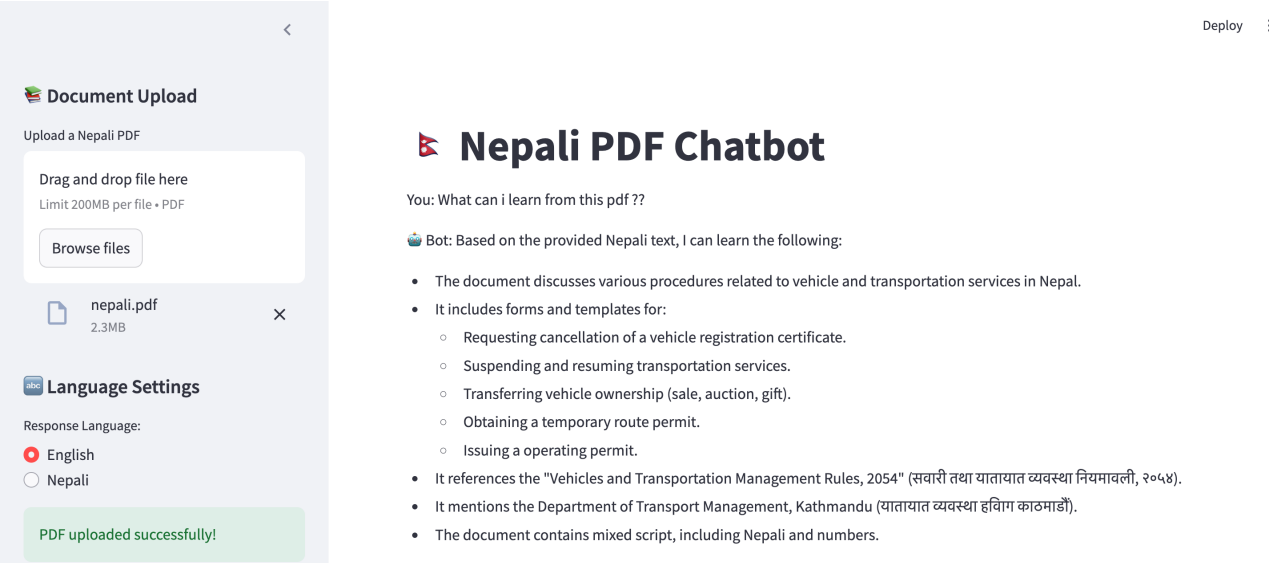

Fig4- showing a chatbot response in English.

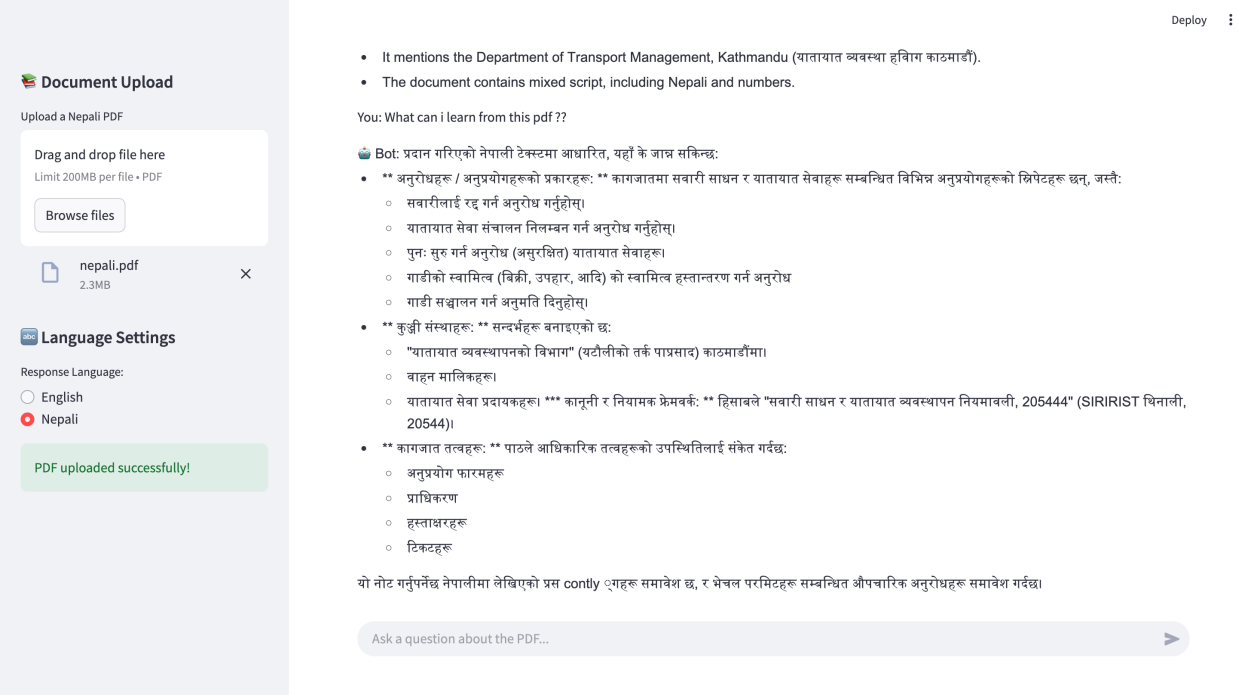

Fig5- showing a chatbot response in Nepali.

Conclusion

The Nepali PDF Chatbot provides an efficient way to process Nepali PDFs, extract information, and interact in both English and Nepali. The application uses FAISS for vector storage, Google Translator for multilingual support, and Gemini AI for intelligent responses.