

Scheme Research Application

~HIMANSHI MAHESHWARI

1. Introduction

The Scheme Research Tool is designed in such a way as to automate the scanning process and retrieval of data from government scheme documents. These are scanned from PDF, and details like benefits of schemes, application processes, eligibility criteria, and documents required get automatically fetched and summated. The answer query facility in the tool allows users to question the data directly for information retrieved faster and easier.

2. Objective

- Automatically fetches and scrapes relevant information from PDF sources.
- Creating text content embeddings so that similarity search can be implemented fast.
- Provides a question answering interface in which it accepts any query about specific schemes and responds accordingly.

3. System Requirements

It is a web application developed by Streamlit, meaning an interactive interface that is very user-friendly and uses OpenAI for language processing and FAISS for good index data performance.

Dependencies:

- Streamlit
- Openai
- faiss-cpu
- requests
- PyPDF2
- sentence-transformers

4. Workflow

- Users will input URLs of scheme PDFs
- This fetches the PDF files and processes them for obtaining content.
- The key information are summarized using Open AI.
- The summarized text is embedded using sentence-transformers and stored in FAISS.

5. Conclusion

This is the Scheme Research WebApp, which efficiently automates scheme information extraction and summarization. Regarding this, access to government scheme data is easier. Therefore, with this question-and-answer capability, usability is improved with readily available quick answers to user queries. The tool demonstrates a strong aptitude for data processing, NLP, and efficient retrieval, making it an asset for government information management.