PaperInsight

(Track-1 cactus ChatWithAnyScientificDocument)

By Group- Gaussian Guys
Members:
Soham Ravindran
Gaurav Boob
Kaushal Kulkarni
Varad Unhale

Institute:
Pune Institute Of Computer Technology

Problem Statement:

- Research and scientific endeavours heavily rely on the analysis and dissemination of information contained within various document formats.
- While PDF remains a prevalent format for sharing research findings, scientists, academics, and professionals often utilize a diverse array of formats such as DOC, DOCX, TEX, and PPT for creating and presenting their work.
- However, existing solutions primarily cater to parsing PDF documents, leaving a gap in efficiently handling these alternative formats.

Introduction

- Documents are fundamental to research and professional communication, but existing solutions primarily focus on parsing PDF files. This leaves a gap in handling diverse formats like DOCX, TEX, and PPT, hindering efficient information extraction.
- Our project aims to develop a robust framework to parse multiple formats beyond PDF, integrated with language models for intuitive user interaction and query answering.
- This will streamline document analysis and knowledge dissemination, fostering a more efficient ecosystem for research and collaboration.

Our Approach

- Document Parsing and Preprocessing:
  - For text files (.txt), straightforward reading can be done using Python's built-in file handling.
  - For PowerPoint (.pptx) and Pdf (.pdf) files, we have used Convert-api.
  - Utilised libraries such as docx for parsing data from a docx file.
  - for latex (.tex) files we are sending the code to the model.
  - Preprocess the extracted text data to remove any irrelevant information, such as metadata , formatting artefacts or stop words.
- Language Model Integration:
  - Integrated Gemini 1.0 Pro, our chosen language model, for natural language understanding tasks and processing user queries.

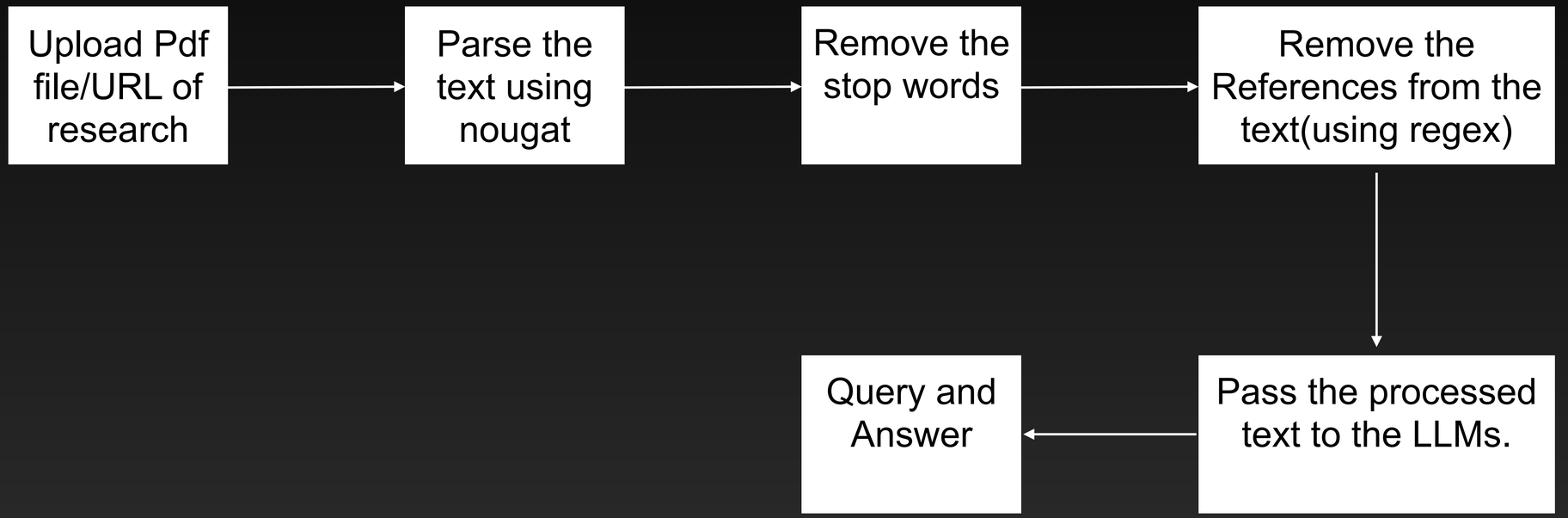
Our Approach

- User Interface Development:
  - Developed a web-based interface using Streamlit and Django frameworks to provide a user-friendly experience.
  - Used tiktoken for authentication and authorization functionalities if required.
  - Implemented interactive features allowing users to upload documents and input queries.
- Document Analysis and Query Answering:
  - Used Gemini 1.0 Pro for document analysis, extracting key insights, generating summaries and answering users query.

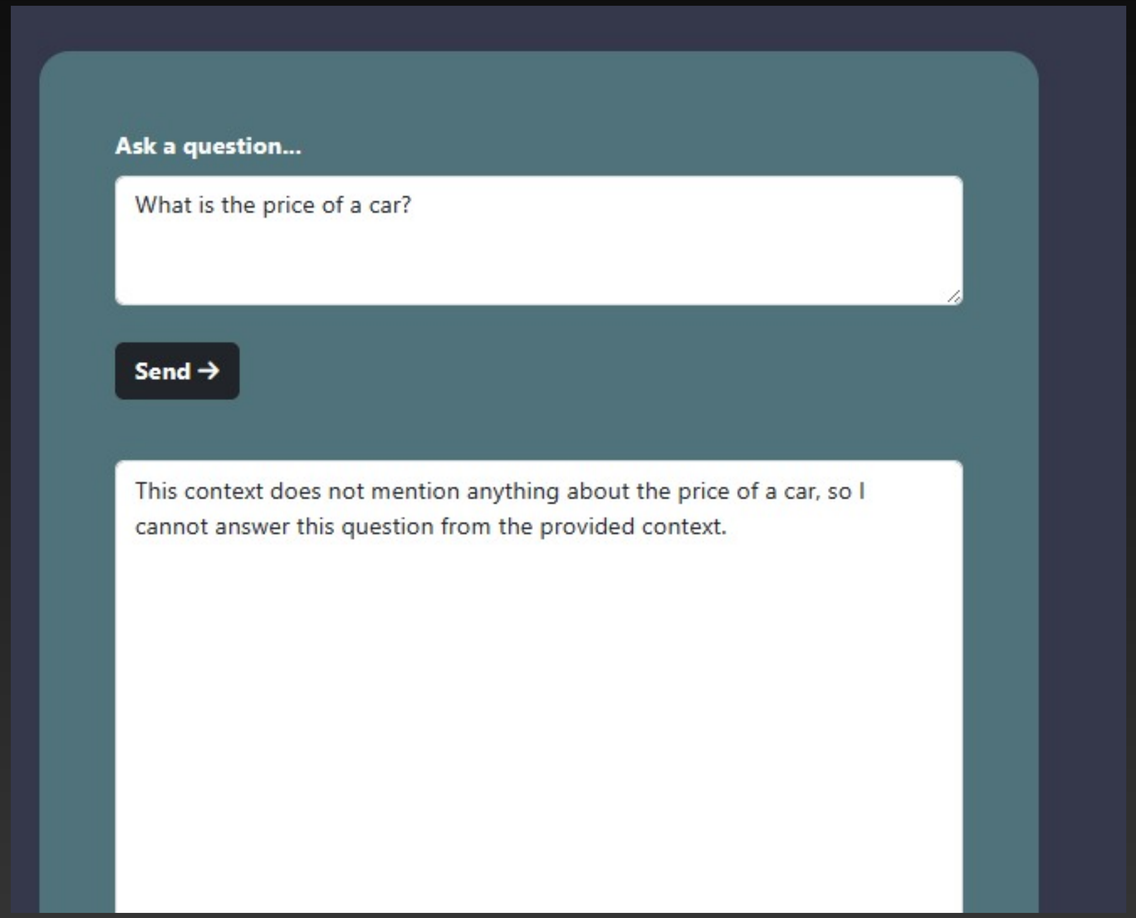
Why should one choose your solution?

- There are several compelling reasons to choose our solution:
  - Comprehensive Format Support: Our solution offers support for a wide range of document formats, including PDF, TXT, PPTX,Tex and DOCX. This ensures versatility and flexibility, allowing users to work with documents in their preferred formats without constraints.
  - Advanced Language Model Integration: We integrated Gemini 1.0 Pro, a powerful language model, into our solution. This enables advanced natural language understanding capabilities, including document analysis, summarization, and query answering. Users can leverage Gemini 1.0 Pro to gain valuable insights from their documents quickly and accurately.
  - User-Friendly Interface: Our solution features a user-friendly web-based interface developed using Streamlit and Django frameworks. This interface makes it easy for users to upload documents, input queries, and interact with the system seamlessly. The intuitive design enhances user experience and productivity.

Nougat Pipeline flow



Output


Ask a question...
What is the price of a car?
Send →
This context does not mention anything about the price of a car, so I cannot answer this question from the provided context.

- Video Demo:

Thank You!