Unveiling Insights Through Question Answering Chatbot using LangChain

[1]Tanishca Soodan, [2]Ritika Kumari, [3]Dr. Poonam Bansal[*]

[1] Student, Department of Artificial Intelligence and Data Sciences, IGDTUW, New Delhi, India

[2] Assistant Professor, Department of Artificial Intelligence and Data Sciences, IGDTUW, New Delhi, India

[3] Professor, Department of Artificial Intelligence and Data Sciences, IGDTUW, New Delhi, India

* Corresponding author

ABSTRACT

In today's world, various chatbots have been created, such as to converse with humans for customer care, ChatGPT to generate text for several purposes, or domain specific tasks. However, the chatbot introduced in this study differs because it provides a convenient way to search and interact with PDF documents. This bot is capable of answering questions about multiple PDF documents. This study aims to create a user centric and intelligent system that enhances information retrieval from PDF documents through natural language queries. This intelligent system is designed to significantly improve user experience by making the search and retrieval process more intuitive and accessible. In this paper, a chatbot is designed as a web application using Streamlit that allows users to chat with multiple PDF documents, extracting information from them and providing responses based on user queries. This study introduces a novel chatbot designed to enhance the process of information retrieval from PDF documents through natural language queries. The paper details the architecture of the chatbot, its implementation, and the potential applications of this technology in various fields, such as research, education, and document management. Overall, this study aims to advance the capabilities of chatbots in handling domain-specific tasks and transforming the way users interact with complex document collections.

Keywords: Chatbot, Portable Document Format files, Conversational AI, LangChain, Information Retrieval

1   Introduction

In the information driven world, Portable Document Format (PDF) files are ubiquitous for storing and sharing valuable content across various domains. However, as the volume of digital data grows exponentially, efficiently extracting pertinent information from these documents becomes increasingly challenging. Traditional manual search and retrieval methods are time consuming and prone to human error, making the process tedious and inefficient. To address this challenge, the Question Answering chatbot allows a person to chat with multiple PDF documents.

The existing chatbots are like ChatGPT that generates textual data based on prompts [2] or as domain-specific information providers [3], or made for customer care [6] however they aren't capable of processing PDF files hence there is a need for an information extracting chatbot that makes the process of retrieving information from documents easier. This limitation creates a gap in functionality, especially for users who need to extract specific information from complex documents. There is a clear need for an intelligent system that can bridge this gap, enabling users to easily retrieve relevant data from PDF files. Such a chatbot would streamline the document search process, providing more efficient and user friendly ways to interact with and extract information from digital documents.

In this study the development and functionality of a specialized Question-Answering (QA) chatbot aimed at simplifying the process of working with PDF documents. Unlike conventional search techniques or general purpose chatbots, this tool enables users to engage in direct, natural language conversations with multiple PDF files, asking questions and receiving precise, contextually relevant answers based on the content of the documents. The key innovation of this chatbot lies in its ability to understand and respond specifically to queries related to the loaded PDFs, making it unnecessary for users to manually scan through documents or rely on external search engines for information. By integrating this capability, the chatbot streamlines information retrieval, offering a more efficient and user friendly experience when working with large or complex PDF collections.

2 Literature Review

Recent advancements in chatbot technology, particularly those leveraging Large Language Models (LLMs), reflect a growing trend toward domain specific applications that enhance user interaction across various fields. One notable application is in healthcare, where LLM powered chatbots have made significant strides in improving the delivery of medical information. For instance, CataractBot offers asynchronous responses to cataract related inquiries by tapping into a curated knowledge base, thereby providing reliable, timely information to patients and healthcare providers [1]. Additionally, the use of advanced models like Generative Pre trained Transformers (GPT 3.5 and GPT 4) has further demonstrated the potential of AI in processing and analyzing research papers related to complex medical topics, such as breast cancer treatment. These models assist in facilitating comprehensive literature surveys, helping researchers and healthcare professionals stay updated on the latest developments in the field [2].

Expanding beyond healthcare, LLMs have proven valuable in enhancing information retrieval in industries like automotive. By applying LLMs to traditional vehicle manuals, researchers have developed systems that significantly improve the way users access technical information. This illustrates the broader applicability of AI technologies beyond healthcare, showcasing their utility in industries that require detailed technical knowledge [4]. Moreover, LLMs are also making strides in emotional support, where their ability to understand and respond to user emotions is being tested. One example is the development of an Empathetic LLM (EmLLM), which integrates physiological data analysis to monitor stress levels, underscoring the increasing focus on emotional intelligence in chatbot systems. This emphasis on emotional intelligence is further highlighted by the design of EBER, a chatbot that adjusts its responses based on the user's mood, particularly benefiting older adults who may require more personalized interactions [5][6].

In addition to healthcare and emotional support, LLM powered chatbots are also addressing complex global issues. Platforms like the Climate Q&A chatbot engage users on topics related to climate change and biodiversity, making scientific concepts more accessible to a wider audience. These platforms serve as an example of how chatbots can bridge the gap between specialized knowledge and public understanding, encouraging informed conversations about critical environmental issues [7]. Similarly, chatbots are being applied to enhance information accessibility in regulatory and educational contexts. Chatbots designed to clarify mining regulations and answer university related inquiries help users navigate complex legal and academic landscapes, demonstrating the versatility of LLM-powered systems in providing specialized, user friendly support [8][9].

Finally, LLM powered chatbots have proven valuable in enhancing consumer rights awareness, with applications such as chatbots that inform air passengers about their rights. This further illustrates the

potential of AI technologies to provide users with easy access to important, domain-specific information, promoting better decision making and user empowerment [10].

Collectively, these studies underscore the expanding potential of LLM-powered chatbots to address a diverse range of user needs across various domains. By providing targeted support in healthcare, emotional well being, environmental advocacy, regulatory clarity, and consumer rights, LLM based systems are transforming how users interact with information and AI. The versatility and growing capabilities of these systems point to a future where LLM-powered chatbots will continue to play a critical role in improving access to specialized knowledge across many industries.

3   Research Methodology

3.1 Architecture

The development of the Question Answering chatbot is a highly intricate process that leverages a diverse array of libraries and tools to create an intelligent system capable of interacting with PDF documents. At the core lies Streamlit, a dynamic framework that serves as the foundation for the interactive web application. Streamlit's intuitive interface and flexibility ensure a seamless user experience, allowing users to easily interact with the chatbot and submit queries without technical barriers. The integration of the Langchain library elevates the chatbot's functionality by providing key features such as the Character Text Splitter, which breaks down large volumes of text into smaller, more manageable chunks, facilitating more efficient processing. Additionally, OpenAI Embeddings allow the chatbot to comprehend and interpret the semantic meaning of the text, enabling it to understand complex queries and respond with accuracy.

To further optimize performance, FAISS is used for efficient vector storage, enabling rapid retrieval of relevant information based on user queries. The Conversation Buffer Memory module ensures the chatbot can maintain context and track conversation history, enabling it to offer more coherent and contextually appropriate responses. The Conversational Retrieval Chain combines these capabilities, synthesizing answers in a manner that feels natural and connected to the ongoing dialogue. Security is also a priority in this system, with the dotenv library ensuring the safe handling of sensitive environment variables, such as API keys. Finally, the PyPDF2 Pdf Reader module plays a crucial role in enabling the chatbot to extract and process text from PDF documents, making it capable of navigating and understanding vast information stored in this format. By combining these advanced technologies, the chatbot offers a powerful tool for interacting with and retrieving information from PDF files, significantly enhancing how users engage with complex documents.

3.2 Implementation

The implementation of the Question Answering chatbot for PDF documents follows a structured approach designed to efficiently respond to user queries by leveraging the content of multiple PDF files. It comprises the following vital steps (shown in Figure 3.2.1).

The process begins with the first crucial step: loading the PDF documents into the application. Using libraries such as PyPDF2's Pdf Reader module, the chatbot extracts the text from the PDFs. Once the text is extracted, it is then broken down into smaller, more manageable chunks, which facilitates the effective processing and analysis of the data. The chunking of text allows the application to handle large volumes of content more efficiently, improving both speed and accuracy in subsequent stages.

Proceedings of 2nd International Conference on Emerging Applications of Artificial Intelligence, Machine Learning and Cybersecurity

247

After the text is chunked, the application employs a language model to generate vector representations, or embeddings, of the text. These embeddings capture not only the context but also the semantic meaning of the text, which allows the chatbot to better understand the nuances of the content within the PDF documents. When a user submits a query, the application compares the question against the extracted text chunks. To do this, it utilizes similarity metrics such as cosine similarity or other neural network-based techniques to evaluate how closely the question aligns with the content of the text chunks.

Once the most relevant text chunks are identified based on their semantic similarity to the user's question, they are passed to the language model for response generation. The model then synthesizes a coherent and contextually appropriate answer using the information from the selected text chunks. This systematic approach ensures that the chatbot provides accurate, context aware responses based on the content of the PDF documents, enabling users to efficiently retrieve the information they need.
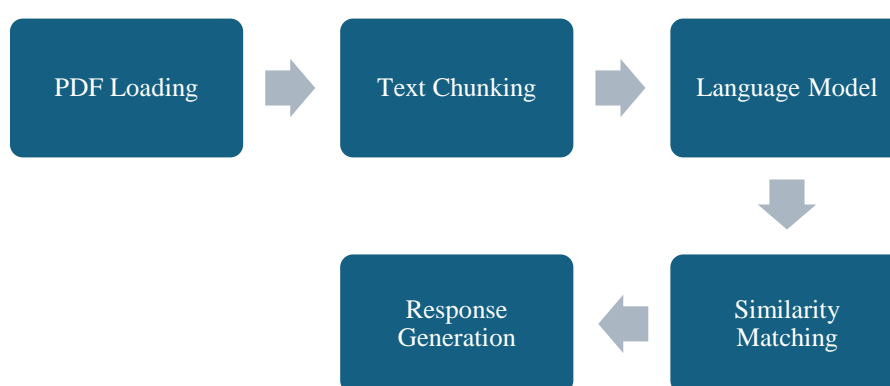
Figure 3.2.1: Steps to show the methodology

This structured method guarantees the reliability of the application in supporting effective information retrieval and interaction with PDF content.

4   Results and Discussion

This explains how users can engage with the Question Answering chatbot in the web application built with Streamlit. Users can upload multiple PDF documents, then ask questions in natural language. The chatbot processes the content of the uploaded PDFs and provides answers based solely on the information within those documents. This interactive setup allows users to easily explore relevant responses, making it a streamlined and efficient tool for retrieving specific information from PDFs without the need for external searches or manual scanning. The chatbot is designed to respond contextually to queries, enhancing the user experience.

The following section presents a series of screenshots illustrating the process and results when a user uploads a textbook and reference materials for an assignment into the web application. After uploading the study materials in PDF format, the user can interact with the Question Answering chatbot by posing questions related to the content of the materials. As shown in the screenshots, once the PDFs are successfully loaded, the chatbot processes the documents and is ready to respond to queries.

When the user asks a specific question about the content, the chatbot retrieves the relevant information directly from the uploaded textbooks and reference materials. The responses provided by the chatbot

are based entirely on the text extracted from these documents, ensuring that the answers are precise and contextually aligned with the study materials. The interface displays the question posed by the user along with the chatbot's corresponding answer, making it easy for users to follow the conversation and gather the necessary information.

The screenshots also demonstrate the seamless interaction between the user and the chatbot. The user interface is intuitive, allowing for easy navigation and query submission, while the chatbot's ability to parse and analyze the text from multiple PDFs ensures that it can effectively provide comprehensive answers. This functionality significantly enhances the efficiency of information retrieval, allowing users to quickly extract key details from textbooks and references without manually searching through large volumes of text. The screenshots visually represent this process, showcasing how the chatbot facilitates a smooth and effective learning experience for students and researchers working with academic or study materials.

Figure 4.1 displays the user interface of the web application, where users can engage with the Question Answering chatbot. This interface allows users to ask questions directly, enabling them to quickly obtain relevant information from the uploaded documents without the need to manually sift through lengthy texts. By streamlining the information retrieval process, the chatbot significantly improves productivity, particularly when working on assignments or research tasks. The intuitive design ensures that users can efficiently access the data they need, ultimately saving time and enhancing the overall efficiency of completing academic or professional work.
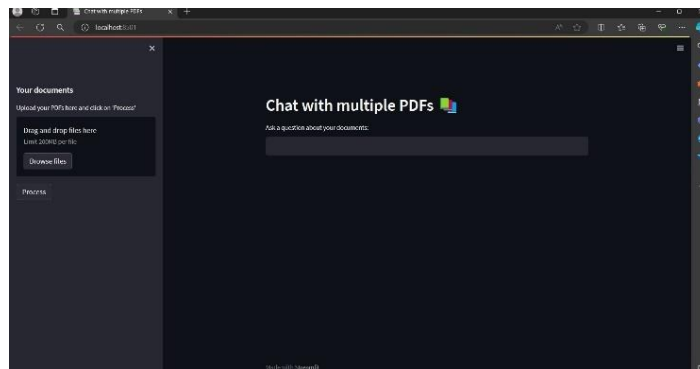

Figure 4.1: UI of Question-Answering chatbot

Figure 4.2 illustrates the process of uploading multiple PDF documents in the web application. This feature allows users to access and utilize various sources of information simultaneously, enhancing the quality and depth of their inquiries. By integrating multiple documents, users can cross reference content and gather insights from a range of materials. This functionality is especially valuable in academic contexts, where students frequently rely on a variety of resources to build a well rounded understanding of complex topics. It enables more efficient research and study, promoting a thorough approach to learning and improving academic productivity.
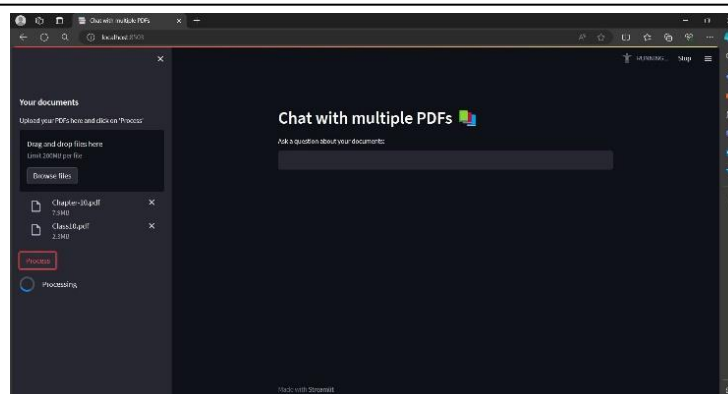
Proceedings of 2nd International Conference on Emerging Applications of Artificial Intelligence, Machine Learning and Cybersecurity

249

Figure 4.2: Processing of PDF documents

Figure 4.3 demonstrates the response of the Question Answering chatbot when a user asks a question. It highlights the chatbot's ability to provide accurate and contextually relevant answers based exclusively on the content of the uploaded documents. This precision ensures that users receive information directly related to their inquiries, enhancing the effectiveness of the chatbot. By delivering specific and reliable answers, the chatbot proves to be an invaluable educational tool, supporting efficient learning and information retrieval. Its responsiveness underscores its role in facilitating targeted and meaningful interactions with the uploaded study materials.
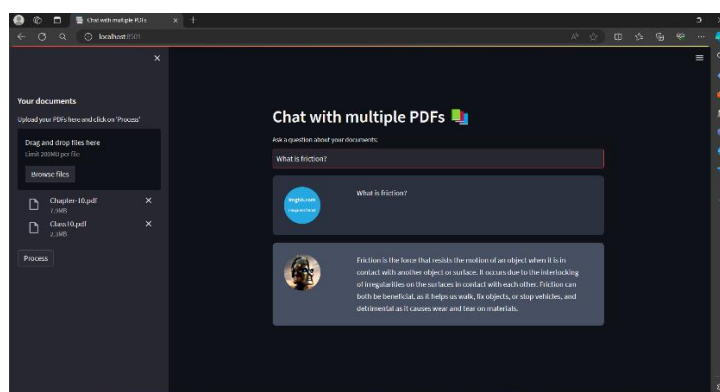

Figure 4.3: Result of the Question-Answering chatbot

The demonstration presented above showcases an innovative method for improving user interaction with academic content. By integrating the Question Answering chatbot into a web application, the system not only boosts user engagement but also simplifies the process of retrieving relevant information. This approach offers a valuable solution for both students and researchers, providing a more efficient way to access and utilize academic resources. Its ability to streamline information retrieval makes it a highly promising tool for enhancing productivity in academic and research settings.

5   Conclusions

The Question Answering chatbot represents a significant shift in how users access and interact with information, offering a more efficient and user-friendly alternative to traditional document searching methods. Built with LangChain, the chatbot simplifies the process of retrieving information from PDF documents. Users can ask questions, and the chatbot analyzes the content of the PDFs to provide accurate, contextually relevant answers. A key feature of the chatbot is its ability to handle multiple PDF documents at once. Users can upload several files, and the chatbot processes the content of each document, offering responses based on the combined information. This multi document functionality greatly enhances the chatbot's versatility, enabling users to draw insights from a variety of sources without switching between

multiple documents. The user interface is designed to be intuitive and accessible, catering to users of all skill levels. The straightforward design ensures that even those with minimal technical expertise can navigate and interact with the chatbot seamlessly. Additionally, the chatbot's conversational interface promotes a more engaging and interactive experience. Users can ask questions in a natural, conversational manner, making the process of retrieving information feel more like a dialogue than a search. Overall, the development of the Question Answering chatbot is a breakthrough in information retrieval. By enabling seamless interactions with multiple documents and offering a user friendly interface, it significantly streamlines the process of accessing and understanding PDF content. This chatbot has the potential to revolutionize how individuals, especially students and researchers, engage with and extract information from documents, paving the way for more efficient and effective workflows.

Declarations

Competing Interests

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Tanishca Soodan, Ritika Kumari, Poonam Bansal (2025). Unveiling Insights Through Question Answering Chatbot using LangChain. *AIJR Proceedings*, 245-251. https://doi.org/10.21467/proceedings.178.26

References

[1] P. Ramjee *et al.*, "CataractBot: An LLM-Powered Expert-in-the-Loop Chatbot for Cataract Patients," *arXiv*, 2024. [Online]. Available: https://arxiv.org/abs/2402.04620v1. [Accessed: Feb. 7, 2024].

[2] A. D. Silva *et al.*, "AI Insights: A Case Study on Utilizing ChatGPT Intelligence for Research Paper Analysis," *arXiv*, 2024. [Online]. Available: https://arxiv.org/abs/2403.03293v1. [Accessed: Mar. 5, 2024].

[3] C. Kang *et al.*, "Domain-Specific Improvement on Psychotherapy Chatbot Using Assistant," *arXiv*, 2024. [Online]. Available: https://arxiv.org/pdf/2404.16160v1. [Accessed: Apr. 24, 2024].

[4] T. Medeiros *et al.*, "Analysis of language-model-powered chatbots for query resolution in PDF-based automotive manuals," *Vehicles*, vol. 2023, no. 5, pp. 1384-1399.

[5] P. Dongre *et al.*, "Integrating physiological data with large language models for empathetic human-AI interaction," *arXiv*, 2024. [Online]. Available: https://arxiv.org/pdf/2404.15351v1. [Accessed: Apr. 14, 2024].

[6] S. G. Mendez *et al.*, "Entertainment chatbot for the digital inclusion of elderly people without abstraction capabilities," *IEEE Access*, vol. 9, pp. 75878-75891, May 2021.

[7] N. D. L. Calzada *et al.*, "ClimateQ&A: Bridging the gap between Climate Scientists and the General Public," *arXiv*. [Online]. Available: https://doi.org/10.48550/arXiv.2403.14709. [Accessed: Mar. 18, 2024].

[8] V. Khandelwal *et al.*, "Chatbot to Respond to Text Queries Pertaining to Various Acts, Rules, and Regulations Applicable to Mining Industries," *Journal of Emerging Technologies and Innovative Research*, vol. 10, Nov. 2023.

[9] Dinh H. *et al.*, "EduChat: An AI-Based Chatbot for University-Related Information Using a Hybrid Approach," Applied Sciences, vol. 13, no. 22, pp. 12446, Nov. 2023.

[10] M. Taranukhin *et al.*, "Empowering Air Travelers: A Chatbot for Canadian Air Passenger Rights," *arXiv*, 2024. [Online]. Available: https://arxiv.org/pdf/2403.12678v1. [Accessed: Mar. 19, 2024].

Proceedings of 2nd International Conference on Emerging Applications of Artificial Intelligence, Machine Learning and Cybersecurity

251