GPT-Powered Legal Document Summarization and Intelligent Question Answering System for Efficient Legal Research, Content Analysis, and Decision Support

[1]Mr. N. Lakshmi Narayana,[2]Siva Naga Sai Deepthi Vulli,[3]Putta Tirumala Subba Reddy, [4]Yeragalla Chakravarthi

[1]Assistant Professor, Dept of Computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

[2,3,4]B. Tech Student, Dept of Computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

ABSTRACT

GPT-Powered Legal Document Summarization and Intelligent Question Answering System for Efficient Legal Research, Content Analysis, and Decision Support presents a GPT-powered Legal AI Assistant built with Streamlit that helps users understand legal documents more easily. It solves the problem of reading long and complex legal files by allowing users to upload documents, get clear summaries, and ask questions about the document. This system uses the AI models, document reading tools, and smart search techniques to find relevant information and generate accurate answers. As a result, it saves time, improves understanding, and supports better legal research and decision-making for users.

KEY WORDS: ***Legal AI, Document Summarization, Question Answering, Legal Research, Streamlit, Large Language Models, Retrieval-Augmented Generation, FAISS, Natural Language Processing***

INTRODUCTION

Legal professional and researchers often work with large amount of documents like contracts, case files and legal reports. Reviewing these documents manually is time consuming. With rapid growth of legal data, there is a strong need for intelligent tools that can assist in legal research. This project introduces a GPT-powered legal AI assistant which is built using Streamlit and allows users to upload legal documents, generate summaries, and ask questions related to the document. The goal of this project is to make legal research easier for better decision-making. Finally, its very useful in real time applications.

LITERATURE SURVEY

Chalkidis et al. (2020) introduced LEGAL-BERT, a transformer model pretrained on legal texts to improve legal NLP tasks, but

it is limited to offline analysis and lack interactive features. Lewis et al. (2020) proposed the Retrieval-Augmented Generation(RAG), which combines document retrieval with text generation, but not specified for legal documents. This proposed work differs by applying RAG to legal domain with document processing, summarization, and real-time question answering through an interactive Streamlit based application.

RELATED WORK

Several research studies have explored the use of Artificial Intelligence and Natural Language Processing to support legal document analysis. Earlier works are more focused on applying transformer based models. For example, LEGAL-BERT was trained on legal documents to better understanding of legal language but does not allow user interaction. The Retrieval-Augmented Generation (RAG) which improves answer accuracy by fetching relevant text before generating responses. These systems are usually general purpose and not designed specifically for legal documents.

EXISTING SYSTEM

The existing system Legal Document Summarizer focuses only on generating brief summaries of legal documents. While this helps the user to get a quick idea of the legal document, but the system does not allow the user to ask questions about the document content or interact system in a conversational manner which limits flexibility and ease of use. The system lacks an interactive interface, making it difficult for users to ask questions, explore specific details or clarify doubts. Due to these limitations this system is not suitable for in-depth legal research or decision support.

PROPOSED SYSTEM

The proposed system is a GPT-powered legal AI Assistant built using streamlit that enables users to upload documents in PDF or DOCX format. The system automatically generates concise and customizable summaries of the document using large language models with prompt-based abstractive summarization, accessed through the OpenRouter API. And supports question answering using a retrieval-augmented generation approach, where relevant document sections are retrieved using semantic search to produce accurate, context-aware answers.

SYSTEM ARCHITECTURE

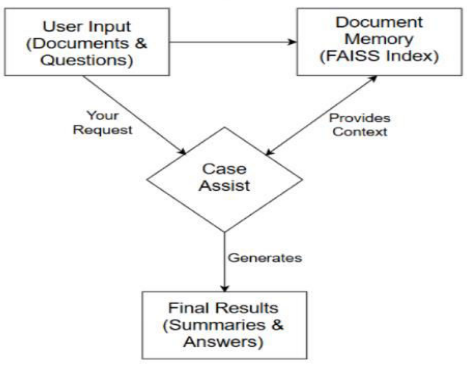

Fig 1: System Architecture

METHODOLOGY DESCRIPTION

The proposed system uses a retrieval-augmented generation approach to analyze legal documents efficiently. Uploaded PDF or DOCX files are processed to extract text, which is split into smaller chunks and converted into semantic embeddings stored in a FAISS vector database. For summarization, a large language model generates concise summaries using prompt-based instructions. For question answering, relevant document chunks are retrieved using semantic search and combined with the user's query to produce accurate, context-aware responses. This approach enables effective legal research, content analysis, and decision support through an interactive interface.

RESULTS AND DISCUSSION

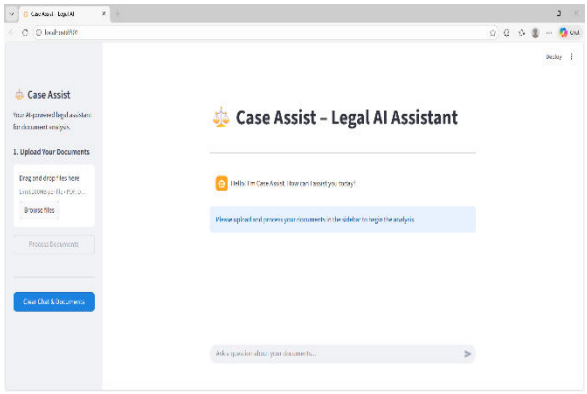

Fig 2: Home Page

An AI-powered legal assistant that helps you upload, summarize, and analyze legal documents with ease. Ask questions and get accurate, context-aware answers to support efficient legal research and decision-making.

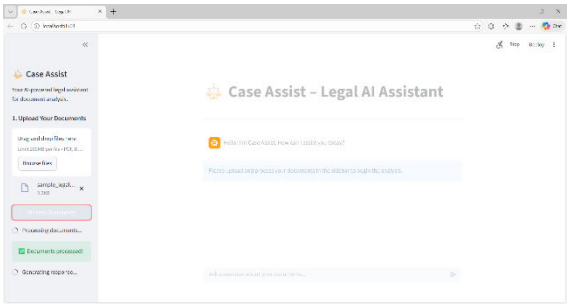

Fig 3: Processing Page

This page shows the system actively handling the uploaded legal document. It indicates that the document has been successfully uploaded.

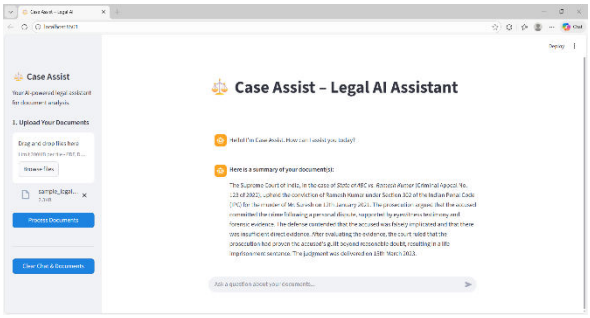

Fig 4: Summary Page

The system automatically displays a clear, concise summary highlighting the key facts, parties involved, legal arguments, and final judgment of the case.

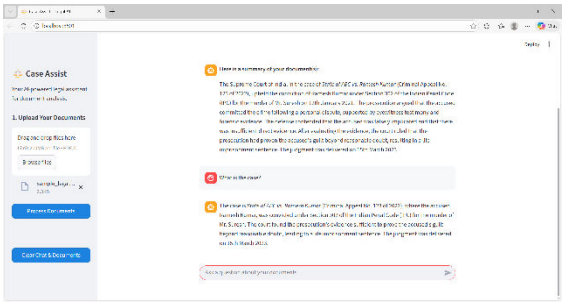

Fig 5: Question -answering Page

The system retrieves the relevant information from the processed document and provides a clear, accurate answer based strictly on the document content after a question is asked.

CONCLUSION

This project successfully builds a Legal AI Assistant that helps users understand legal documents easily. It allows users to upload documents, get clear summaries, and ask questions based on the document content. By using AI and smart search techniques, the system saves time, reduces manual effort, and makes legal document analysis faster and more user-friendly.

FUTURE SCOPE

The system can be extended to support multilingual translation of documents, summaries, and answers. Important clauses and keywords can be highlighted for faster understanding. Voice-based input and output can be added for easier interaction.

REFERENCES

1. Kesavulu, O. S. C., & Harini, P. (2013). Enhanced packet delivery techniques using crypto-logic riddle on jamming attacks for wireless communication medium. *Int. J. Latest Trends Eng. Technol*, *2*(4), 469-478.

2.Chalkidis, I., Kampas, D., Aletras, N., & Androutsopoulos, I. "LEXGLUE: A Benchmark Dataset for Legal Language Understanding," Proceedings of ACL, 2022.

3.Zhong, H., Guo, Z., Tu, C., Xiao, C., & Liu, Z. "Legal Judgment Prediction via Topological Learning," Proceedings of EMNLP, 2018.

4.Bommarito, M. J., & Katz, D. M. "A Mathematical Approach to the Study of the United States Code," *Physica A*, vol. 389, no. 19, pp. 4195–4200, 2010.

5.Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proceedings of NAACL, 2019.

6.Lewis, P., Perez, E., et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Proceedings of NeurIPS, 2020.

7.Guu, K., Lee, K., Tung, Z., Pasupat, P., & Chang, M.-W. "REALM: Retrieval-Augmented Language Model Pre-Training," Proceedings of ICML, 2020.

8.Kornilova, A., & Eidelman, V. "BillSum: A Corpus for Automatic Summarization of US Legislation," Proceedings of ACL, 2019.

9.Zhang, J., Zhao, Y., Saleh, M., & Liu, P. "PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization," Proceedings of ICML, 2020.

10.Radev, D. R., et al. "Introduction to the Special Issue on Summarization," *Computational Linguistics*, vol. 28, no. 4, pp. 399–408, 2002.

11.Rabelo, J., et al. "COLIEE 2021: Legal Information Retrieval and Entailment," Proceedings of ICAIL, 2021.

12.Tran, V., et al. “Automatic Legal Question Answering Using Pretrained Transformers,” arXiv preprint arXiv:2009.11402, 2020.

13.Liu, Y., & Lapata, M. “Text Summarization with Pretrained Encoders,” Proceedings of EMNLP-IJCNLP, 2019.

14.OpenAI. “GPT-4 Technical Report,” arXiv preprint arXiv:2303.08774, 2023.

15.Touvron, H., et al. “LLaMA: Open and Efficient Foundation Language Models,” arXiv preprint arXiv:2302.13971, 2023.

16.Manning, C. D., Raghavan, P., & Schütze, H. Introduction to Information Retrieval. Cambridge University Press, 2008.

17.Johnson, J., Douze, M., & Jégou, H. “Billion-scale Similarity Search with FAISS,” Proceedings of IEEE Big Data, 2019.

18.Zheng, L., et al. “Judicial Reading Comprehension: A New Dataset for Legal AI,” Proceedings of EMNLP, 2020.