

2024 AI challenge - Taxation Assistant

Product

Problem

LLMs have transformed natural language processing, but despite their advantages, LLMs struggle with specialized fields with specific information, where precise expertise is crucial. Despite the prevalence of many chatbots in the medical [1], [2], [3], [4] and educational domains, there has not been much work done in the taxation domain. Taxation is highly intricate, with regulations often spanning hundreds or thousands of pages with sophisticated and cryptic tax regulations. Consequently, individuals face significant obstacles in obtaining reliable taxation information and guidance relying on human expertise in navigating tax functions. The problem is particularly prevalent in countries like Pakistan where tax laws can be obscure and poorly articulated.

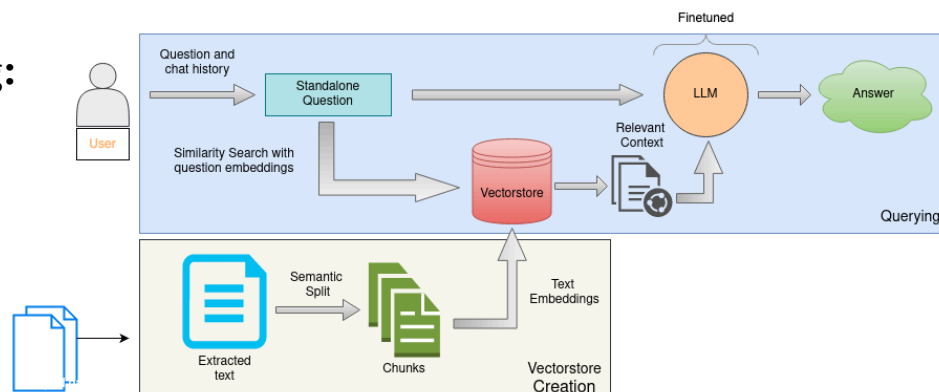
Solution

We introduce Taxation Assistant, a specialized Retrieval Augmented Generation (RAG) system powered by the OpenAI GPT-2 LLM, designed specifically for income taxation. Complemented by a meticulously curated dataset tailored to the intricacies of income taxation, Taxation Assistant leverages the RAG pipeline to mitigate model hallucinations, enhancing the reliability of generated responses. Through a blend of qualitative and quantitative evaluation methodologies, we rigorously assess the accuracy and usability of Taxation Assistant, establishing its efficacy as an income tax advisory tool.

Target Audience

A survey carried out from March to May 2023, which included more than 1,800 legal and tax professionals from the U.S., U.K., and Canada, revealed that 82% of legal professionals and 73% of tax professionals consider ChatGPT useful for legal or tax-related tasks (Reuters 2023). Therefore, our target audience comprises professional tax consultants who struggle to efficiently locate relevant information.

Engineering:



Overall flow of our pipeline:

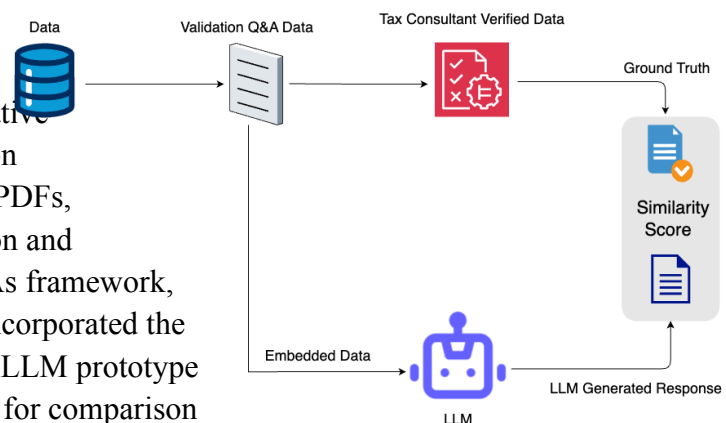
We developed Taxation Assistant, an intelligent chatbot designed to provide easy access to income tax information in Pakistan. Initially, we gathered a dataset from the latest publicly available Income Tax Manual on the Federal Board of Revenue website, encompassing diverse tax categories such as individual and corporate taxes, deductions, exemptions, and accounting guidelines. We also supplemented this data with information from Taxationpk.com, including procedural data and best practices. For processing, text from PDF documents was extracted and semantically split using the Semantic-Split library, forming vector embeddings that were stored in the Faiss vector store. These embeddings facilitated efficient retrieval of information. Our approach uses the Retrieval-Augmented Generation (RAG) framework, which combines standalone questions, embedding generation, similarity searches, and response generation through OpenAI's GPT-2 model, ensuring adaptability to various user queries and evolving tax topics.

Technology:

We used Langchain FAISS vector databases for the retrieval and hugging face for the open source LLM.

Results & Analysis:

We conducted a two-part evaluation of our Taxation Assistant system: quantitative and qualitative. The quantitative evaluation included data extraction from tax-related PDFs, constructing a test set with manual curation and synthetic data generation using the RAGAs framework, and verification by tax consultants. We incorporated the verified question-answer dataset into our LLM prototype and calculated the BERT similarity score for comparison with ground truths. Additionally, we assessed the system's performance using the RAGAs framework, focusing on context relevance, faithfulness, and answer relevancy, with high scores indicating strong system performance. Our RAG system's excellent performance, reflected in its RAGAS score which is 0.90, outperforms other sensitive information retrieval applications.



Going Forward

Our immediate next goal is to compare the RAG system presented to fine-tuned models on the same dataset. We also plan to convert the chatbot into a copilot for Tax and legal consultants. Additionally, we plan to explore a combined approach that leverages both the RAG pipeline and a fine-tuned model to provide relevant, coherent, and accurate responses to tax-related queries.

Citation

1. [Arxiv.2309.04646.](#)
2. [2303.14070](#)
3. [2304.06975](#)
4. [Arxiv.2304.01097.](#)

All diagrams have been taken from our own research paper which is published at FLAIRS-37 Conference.