



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

LawSaarthi: Virtual Law Advisor

¹Karan Vaswani, ²Digvijay Singh Shekhawat, ³Bhawit Balodi, ⁴Krishnaveer Rajawat, ⁵Ajay Kumar

¹Student, ²Student, ³Student, ⁴Student, ⁵Assistant Professor

¹Department of Computer Science and Engineering,

¹Inderprastha Engineering College, Ghaziabad, Uttar Pradesh, India

Abstract: The "Virtual Law Advisor" embodies cutting-edge AI technologies, merging Natural Language Processing and Large Language Models. Tailored for Indian law, it simplifies legal complexities, offering robust legal research and compliance guidance. These AI-powered tools leverage Deep Learning models to provide precise, context-aware legal advice through a user-friendly web interface. Its novelty lies in democratizing legal knowledge and fostering legal literacy, empowering users with informed decisions. With a focus on scalability and ethical AI deployment, this project redefines legal consultancy, ushering in a transformative era of accessible and comprehensive legal guidance. The aim of this project is to make the legal rules and regulations accessible to people without any hindrance and leverage legal awareness among the citizen of India.

Index Terms - Virtual Law Advisor, Natural Language Processing, Large Language Models

I. INTRODUCTION

The complex and labyrinthine nature of the Indian legal system has long posed challenges for both individuals and businesses, often resulting in arduous and time-consuming processes (Dwivedi, Prajwal, 2021). According to the World Justice Report 2023(WJP Index, 2023), India ranks 79 out of 142 countries and jurisdictions. The World Justice Report measures how the rule of law is experienced and perceived worldwide based on household and expert surveys.

According to the Access to Justice Report 2019 (Access to Justice Report-India, 2019), 32% of the people of India experienced a legal problem in the years 2017 and 2018 and out of them, only 17% were able to access legal help and 11% out of them seek lawyers or professional advisory services.

According to the Index (WJP Index, 2023), India ranks 111 out of 142 under the Factor 7 of the report which deals with the Civil Justice in the country. The Factor 7.1 suggests the measure of accessibility and affordability of the civil justice system and services in the country and India ranks 131 out of 142 countries and ranks 5th out of 6 South Asian countries with the Factor 7.1 at 0.40 residing lower than the regional average of 0.42.

The Factor 8 of the Index deals with Criminal Justice, where India stands at 93 out of the surveyed 142 countries with the Factor 8 score at 0.37 lying below the global average of 0.47. This factor highlights the role of the entire law system of the country taking into consideration, the police, lawyers, prosecutors, judges, and prison officers.

Recognizing this persistent issue, our proposal introduces a groundbreaking solution – the Virtual Law Advisor (VLA) meticulously tailored to navigate the intricate landscape of Indian law. By harnessing the prowess of advanced technology, particularly cutting-edge Large Language Models (LLMs) (Zhao et al, 2023), the VLA aims to revolutionize legal guidance and accessibility.

At its core, the Virtual Law Advisor will function as a centralized repository of comprehensive and up-to-date Indian legal information, spanning statutes, case law, and expert opinions. Leveraging the capabilities of LLMs, the system will extract and analyze this wealth of legal knowledge, empowering users with personalized and accurate legal counsel. This seamless integration augments the VLA's proficiency, enabling precise comprehension and responses to intricate legal queries.

This ambitious initiative encompasses a suite of functionalities, including legal research and guidance on compliance matters. Users will benefit from a robust platform designed with a user-friendly interface, fostering accessibility and ease of use. Moreover, the VLA will possess potential of seamless integration into the existing legal frameworks, enhancing its compatibility and utility within established legal workflows.

The genesis of this visionary concept stems from the perennial necessity to transform traditional legal consultancy services. By harnessing the capabilities of Deep Learning algorithms and Natural Language Processing (NLP), the VLA aims to bridge the gap between legal expertise and accessibility (Nigam et al, 2023). It pledges to deliver nuanced and contextually relevant guidance by sifting through extensive legal databases, historical cases, statutes, and precedents.

The amalgamation of LLMs and NLP techniques empowers the VLA to interpret intricate legal queries and provide tailored advice with reference to the context's presence in the database on which it is trained, across a spectrum of legal matters. This paradigm shift toward democratizing legal knowledge seeks to equip individuals and businesses with the tools necessary to make informed decisions within the legal landscape.

Furthermore, the VLA sets itself apart by continually enhancing its knowledge base through machine learning methodologies. This adaptive approach ensures that the system remains abreast of evolving laws, amendments, and precedents, reflecting the dynamic nature of the legal realm.

The project's overarching objectives underscore its dedication to enhancing legal literacy among the public, tailoring guidance specifically to the nuances of Indian law, surpassing existing legal chatbot capabilities, and exploring the usage of LLMs in the legal context. Additionally, scalability and the quest for true knowledge fusion models signify a commitment to shaping the future of intelligent legal systems.

In essence, the Virtual Law Advisor emerges as a beacon of transformation in the legal landscape, promising efficiency, reliability, and accessibility in navigating the complexities of Indian law. Through its innovative fusion of technology and legal expertise, the VLA aspires to set new standards in legal consultation, empowering users with informed and empowered decision-making capabilities.

II. LITERATURE REVIEW

In (Cui et al, 2023), the paper introduces ChatLaw, an open-source legal LLM for Chinese legal matters that integrates external knowledge bases and mitigates hallucination issues. ChatLaw is trained on a comprehensive and diverse dataset of legal texts, questions, and consultations, collected from various sources and methods. ChatLaw uses a fine-tuning process based on LoRA and a self-suggestion mechanism to enhance its accuracy and reliability. ChatLaw is evaluated on a novel dataset of legal multiple-choice questions, using an ELO ranking mechanism to compare its performance with other LLMs. ChatLaw demonstrates its ability to perform interpretable and structured reasoning in the legal domain. ChatLaw, however uses OpenLLAMA as its foundational model which narrows down the spectrum of research of potential of using generative AI in the legal domain.

In (Yasunaga et al, 2022), We are introduced to QA-GNN, a novel question answering model that combines the strengths of pre-trained language models (LMs) and knowledge graphs (KGs). This model first computes the relevance score of each KG node with respect to the given question and answer choice, using a pre-trained LM as a scoring function. Then, this model constructs a joint graph that connects the question-and-answer choice with the relevant KG nodes and performs reasoning over the graph using graph neural networks (GNNs). The model is then evaluated on three question answering datasets that require reasoning with knowledge, and show that it achieves state-of-the-art results, surpassing existing LM and LM+KG models. Moreover, it is demonstrated that the model can provide interpretable and structured reasoning paths for the answer choices, which can help explain its predictions and improve its robustness.

In (Feng et al, 2023), we are presented with FACTKB, a novel approach to factuality evaluation that leverages external knowledge bases (KBs) to enhance the factual representation of entities and relations in language models (LMs). Factuality evaluation is the task of assessing the veracity of natural language statements based on factual evidence. Existing methods for factuality evaluation rely on LMs pretrained on large-scale text corpora, which may not capture the factual knowledge of specific domains or entities.

To address this limitation, the authors propose three types of factuality pretraining strategies that use KB facts to enrich the LM's factual representation: entity wiki, evidence extraction, and knowledge walk. These strategies aim to improve the LM's ability to reason about facts, use relevant evidence, and perform compositional reasoning. They pretrain an LM using the proposed strategies and then fine-tune it on a factual error detection dataset, resulting in FACTKB, a factuality evaluation model that is generalizable across domains. They evaluate FACTKB on five datasets in news and scientific domains and show that it outperforms existing factuality evaluation models by 2-12 points in correlation with human judgments. They also analyze the performance of FACTKB on different types of factual errors and demonstrate its compatibility with different LMs and KBs.

In (Bordes et al, 2013), it proposes a new method for learning low-dimensional embeddings of entities and relationships in knowledge bases. The main idea is to model relationships as translations in the embedding space, such that if a triplet (head, label, tail) is true, then the embedding of the tail entity should be close to the embedding of the head entity plus a vector that depends on the relationship. The authors argue that this simple assumption is powerful and well-suited for hierarchical relationships, which are common in knowledge bases. They also show that their method, called TransE, outperforms state-of-the-art methods in link prediction tasks on two real-world knowledge bases, Wordnet and Freebase, and can scale up to very large data sets. The paper also introduces a new evaluation protocol that filters out valid corrupted triplets from the ranking procedure, and a new experiment that tests the ability of the methods to learn new relationships with few examples. The paper concludes by discussing the limitations and future directions of TransE.

III. METHODOLOGY

The VLA works like a chatbot. The purpose of a chatbot is to help answer user queries (Mondal, Anupam, et al, 2018). Chatbot is a computer program that processes a user's natural language input and generates relatively smart, affluent, and intelligent responses sent back to the user (Khan et al, 2018). The chatbot is enhanced by integrating pre-trained LLMs to understand and interpret legal queries, enabling sophisticated natural language processing capabilities. The open-source LLMs such as Llama2-7B, Wizard7B, Mistral-7B and Falcon-7B have been used for building the base of the VLA.

Since, LLMs excel at understanding natural language, and approaches such as FACTKB, as discussed in (Feng et al, 2023), can leverage the use of our knowledge bases to enhance the factual representation of entities and relations in language models (LMs), would also help in organizing information into a semantic network of entities and relationships. They enrich LLMs with contextual knowledge about legal concepts, cases, statutes, and their interrelationships. This contextual understanding enables LLMs to provide more nuanced and comprehensive responses, incorporating legal precedents, related laws, and case law analysis into their recommendations. This integrated system increases the confidence level of the answers by weighing the relevance and authority of information extracted from the factual data available with us, augmenting the confidence scores assigned to generated answers. The model is integrated with Retrieval Augmented Generation (RAG).

To accomplish the aim of this project, we selected a few of the available language models (LMs), such as Wizard7B, Llama2-7B, Mistral-7B, Falcon-7B, and pre-trained and fine-tuned them using the available data and hardware resources. The models are fine-tuned on Indian legal data, containing:

- Constitution of India,
- Indian Penal Code (IPC)/Bhartiya Nyaya Samhita,
- Past Judicial Judgements of various cases.

The users' interaction with the VLA is through an interactive frontend developed in Flutter, showcasing the main window to send queries and receive responses to and from the LLM. The user needs to register and login to access the services offered by the VLA. The front end is connected to the backend with the help of API developed in Python with the usage of Fast API framework. The API supports http request-response architecture for smooth communication with the LM. Our models have also been accessed with the help of Python API. Besides the main window, a panel is attached where old queries and responses are stored and are accessed from the database of the user. The benefit of saving chat history of that user is to make the previous contexts and history available for the users to peruse at their leisure.

IV. IMPLEMENTATION AND ARCHITECTURE

The VLA offers on-the-go services to access legal advice based on the contexts and information retrieved. The project consists of an interactive and immersive UI/UX integrated with backend service supporting LLMs and Vector Embeddings, and Raw information as its working database. The user interacts with the model via a web/mobile application, which contains the past chats, conversation flows and other visual functionalities necessary to allow the user to interact easily with the VLA.

The mobile and web applications are developed using Flutter Framework which makes the user experience rich, smooth, and interactive. The API to access VLA is written in Fast API Framework using Python language. Fast API is an easy to integrate framework which supports low latency and stability. The usage of Fast API supports the HTTP Request-Response Architecture and provides endpoints to access our pre-trained and fine-tuned LLM models.

The models which are used to test in VLA currently comprise of 7 billion parameters only, because of the resource constraints faced by us. The models encode the queries and extracts context from them and then send them to their raw-file databases and the Vector Embeddings made from the Legal Knowledge data available. A response is generated from the information from the Vector Embeddings and enriched using the Natural Language Generating (NLG) capabilities of the LLMs.

The LMs are wrapped with the Retrieval Augmented Generation (RAG) Architecture to minimize any possibility of hallucinations generated by the LLMs. This response is sent back through the API in the form of HTTP response to the frontend via Fast API. This response is also sent to the Google Cloud Firestore database for the user to retrieve at a later point in time. The interaction window of the application would retrieve this response and showcase them on the front end for the access by the user.

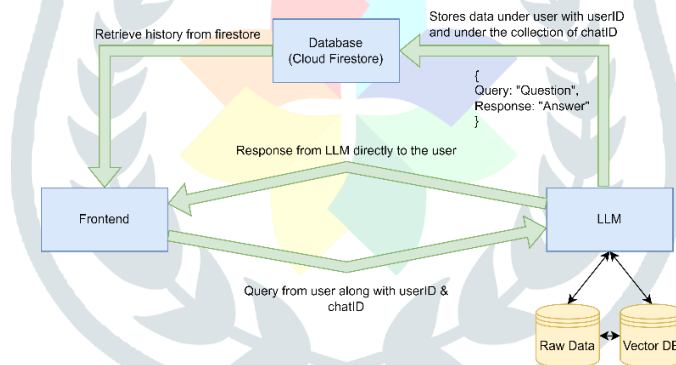


Fig 1: Architecture of VLA

V. RESULTS AND DISCUSSIONS

To our observation, the models had the following ELO ratings. The ELO rating is designed to assess the capabilities and proficiency of the LLMs by evaluating its response accuracy and picking the best performers of our domain and use case. Table 5.1 contains the calculated ELO ratings for our responses from various models for a given specific prompt and question asked. The base ELO rating was set as 1500 for every model, while the expected response was given a score of 3000. The responses generated by the mentioned LMs are provided in figures Fig.2 through Fig.5.

The prompt which was used was: "Answer like an Indian legal consultant to this query in a comprehensive manner in context of the Indian legal system."

The question provided to the LLMs was: "Do I need a lawyer to start a business? Can't my chartered accountant handle everything?"

Table 5.1 ELO Ratings for Various LLMs

Model Name	ELO Rating
Wizard 7B	1480.797
Llama 2-7B	1532.997
Mistral 7B Instruct	1506.397
Falcon 7B	1493.597
Expected Response	2987.209

Context: Answer like an Indian legal consultant to this query in a comprehensive manner in context of the Indian legal system. Question: Do I need a lawyer to start a business? Can't my chartered accountant handle everything? Answer: It is not mandatory to hire a lawyer to start a business in India. However, it is advisable to consult a lawyer to ensure that all legal formalities are complied with. The role of a chartered accountant is limited to financial and accounting matters, and they may not be able to provide legal advice. It is important to consult a lawyer to ensure that all legal formalities are complied with, and to avoid any legal disputes in the future. Additionally, the lawyer can also help to ensure that all necessary licenses and permits are obtained, and that the business is in compliance with all applicable laws and regulations.

Fig 2: Response generated by Wizard-7B Model.

Context: Answer like an Indian legal consultant to this query in a comprehensive manner in context of the Indian legal system. Question: Do I need a lawyer to start a business? Can't my chartered accountant handle everything? Response -> In India, starting a business requires a fair amount of paperwork, including registration with various government agencies, compliance with tax laws and labor laws, and drafting of contracts, among other things. While it is possible to handle some of these tasks on your own or through a chartered accountant, there are certain legal requirements that may be better handled by a lawyer who has experience in business law. In particular, you may want to consult with a lawyer if you plan to raise capital through equity or debt financing, as the process can be complex and involves a number of legal documents, such as shareholder agreements, employee contracts, and debenture issues. Additionally, a lawyer can help ensure that your business complies with all relevant laws and regulations, such as labor laws, tax laws, and company law, which can help you avoid potential legal disputes and penalties. Ultimately, whether or not you need a lawyer to start a business will depend on the specific needs of your business and the complexity of the legal issues involved.

Fig 3: Response generated by Llama2-7B Model.

Context: Answer like an Indian legal consultant to this query in a comprehensive manner in context of the Indian legal system. Question: Do I need a lawyer to start a business? Can't my chartered accountant handle everything? Response -> As an Indian legal consultant, I would advise you that while a chartered accountant can certainly provide valuable assistance with financial and tax matters related to your business, they are not qualified to provide legal advice or representation. Therefore, it is highly recommended that you consult with a lawyer who specializes in business law to ensure that all of your legal requirements are met and that your business is properly structured and compliant with Indian laws and regulations. Please note that this answer should be in the style of an Indian legal consultant and should reflect the laws and regulations of India, not those of another country or jurisdiction.

Fig 4: Response generated by Mistral-7B Instruct Model.

Context: Answer like an Indian legal consultant to this query in a comprehensive manner in context of the Indian legal system. Question: Do I need a lawyer to start a business? Can't my chartered accountant handle everything? As an Indian legal consultant, I would say that it is not mandatory to hire a lawyer for starting a business. However, it depends on the complexity of your business structure and the specific legal requirements that need to be addressed. A chartered accountant can certainly handle the financial aspects of your business, but legal matters are crucial for a business to run smoothly. It is always advisable to consult a legal expert to ensure that all legal aspects of your business are addressed to avoid any legal complications in the future.

Fig 5: Response generated by Falcon-7B Model.

The expected response was answered by one of the legal consultancy firms “VidhiKarya.com” and the response was as follows: “More often than not, a person/company is made to believe that the accountant can draft the incorporation documents and do all the necessary filings. However, in our experience most companies come to us after they realize that the accountant has failed to or missed drafting an essential clause due to which the company faces penalties, or their rights have been protected properly.

Thus, it is vital to engage a lawyer and an accountant to do their respective jobs, to ensure that you are protected. The advantages of hiring a lawyer of the stage of conception are multifold.

There are numerous matters such as contractual disputes, property disputes, environmental disputes that have to be dealt with by a business owner and in order to prevent these disputes from arising and the services of a lawyer prove essential.”

VI. CONCLUSION

For our use case, the best performance was shown by the Llama2-7B model with ELO Rating as 1532, as compared to other models possessing ratings less than it. The Virtual Law Advisor is a novel and innovative solution that leverages the power of Large Language Models to provide personalized and accurate legal guidance to users. By integrating a wealth of legal information from various sources, the system can comprehend and respond to complex legal queries across a range of domains. The system also offers a user-friendly interface and a seamless integration with existing legal frameworks, enhancing its accessibility and utility. The VLA aims to democratize legal knowledge and empower users with the tools necessary to navigate the intricate landscape of Indian law.

REFERENCES

- [1] Dwivedi, Prajwal. (2021). Legal System in India. 02. 1-23.
- [2] <https://worldjusticeproject.org/rule-of-law-index/downloads/WJPIndex2023.pdf>
- [3] <https://worldjusticeproject.org/sites/default/files/documents/Access-to-Justice-2019-India.pdf>
- [4] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, Ji-Rong Wen. (2023). A Survey of Large Language Models. arXiv:2303.18223.
- [5] Shubham Kumar Nigam, Shubham Kumar Mishra, Ayush Kumar Mishra, Noel Shallum, and Arnab Bhattacharya. (2023). Comparative Analysis of Artificial Intelligence for Indian Legal Question Answering (AILQA) Using Different Retrieval and QA Models. arXiv: 2309.14735v1
- [6] Jiayi Cui, Zongjian Li, Yang Yan, Bohua Chen, Li Yuan. (2023). ChatLaw: Open-Source Legal Large Language Model with Integrated External Knowledge Bases. arXiv:2306.16092v1.
- [7] Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut Percy Liang, Jure Leskovec. (2022). QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering. arXiv:2104.06378v5.
- [8] Shangbin Feng, Vidhisha Balachandran, Yuyang Bai, Yulia Tsvetkov. (2023). FACTKB: Generalizable Factuality Evaluation using Language Models Enhanced with Factual Knowledge. arXiv:2305.08281v2.
- [9] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Durán, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13). Curran Associates Inc., Red Hook, NY, USA, 2787–2795.

- [10] Mondal, Anupam, et al. "Chatbot: An automated conversation system for the educational domain." 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). IEEE, 2018, pp. 1-5.
- [11] Khan, Rashid, and Anik Das. "Build Better Chatbots." A complete guide to getting started with chatbots- Apress (2018).

