Justice Lens: An AI-Powered Legal Assistant for Simplifying Indian Law

Archana V S*, Dolus K Shyju*, Rose Saniya P X*, Sreeraj S P* and Assistant Prof Amulya P M†Assistant Prof Divya Jose

∗Department of Computer Science and Engineering
†ICCS College of Engineering And Management Thrissur, Kerala, India

DOI: 10.64823/ijter.2605015

***Abstract*:** Obtaining legal information in India is a major barrier to entry for the common individual due to the statutory language being dense and the exorbitant prices associated with hiring legal experts. Although conversational LLMs are a great starting point in helping users find answers, they often tend to hallucinate cases and fail to cite credible sources. In this work, we introduce Justice Lens, a legal expert powered by AI. The framework is based on the Retrieval-Augmented Generation paradigm, which allows it to provide accurate and reliable legal advice. Moving away from keyword-driven search to semantic search through the use of high dimensional vector embeddings, the architecture enables instant and intent-based queries. Our RAG model makes use of a Pinecone vector store with Approx- imate Nearest Neighbors (O(log V) retrieval time) along with the Google Gemini LLM to convert complex Indian laws into simple English. We demonstrate through our experiments that our framework is successful in avoiding any AI hallucination by ensuring a ”duty to verify” constraint in every output.

***IndexTerms*:** Artificial Intelligence, Legal Technology (LawTech), Semantic Search, Vector Embeddings, Transformer Models, Indian Judicial System, Explainable AI (XAI).

I. Introduction

The India’s legal system is based on a strong consti- tutional framework, the system continues to be out of reach for many people. The use of complex legal language, an extensive collection of case laws, and expensive consultation costs present a daunting challenge for ordinary citizens seeking information about their rights. In reality, “access to justice” becomes a challenge for disadvantaged communities who are unable to afford to navigate through this labyrinthine process.

There is a critical need for technological innovation to connect this complex judicial system with its citizens. Although the digitization of court records may be consid- ered a positive move towards this aim, there still exists a problem of how to turn this data into practical knowledge. There have been attempts to resolve this problem through developments in Legal AI technology. However, these devel- opments appear to exist on two opposite ends of the spectrum. The former are useful in practice due to their ability to perform certain operations with relative ease; however, they still rely on outdated TF-IDF algorithm for search processes. This means that while such technologies are capable of producing results, these results can be wrong if a user does not ask questions using precise legal terminology. On the other hand, more sophisticated models may provide accurate forecasts, but they will not be able to explain them.

We propose a legal assistant chatbot named Justice Lens, using artificial intelligence to merge conversation intelligence with explainability in its output. Unlike conventional search engines, Justice Lens employs a highly sophisticated semantic search engine, driven by the latest deep learning vector embed- ding technique. This means that unstructured text documents are transformed into high-dimensional vectors for better un-

derstanding of the context in which the users intend to use certain keywords. Thus, Justice Lens ensures that everyone – no matter how knowledgeable they are about legal terms – can express themselves in natural language.

To make the tool even more valuable for legal professionals and common peoples, Justice Lens includes advanced modules of Case Outcome Prediction and Explainable AI (XAI). The prediction module employs a specially designed time-decay algorithm that gives preference to recent judicial decisions, assuming that they better capture the contemporary legal standards than old cases of history. Mathematically, this can be represented as:

$$W\ (t)\ =\ e^{-\lambda(tcurrent-tprecedent)}$$

whereλ is the decay factor. However, Justice Lens does not want to turn into an opaque machine. It has a dedicated layer of explainability that relies on knowledge graphs to give logical reasons behind its answers. In this way, the tool mentions specific precedents from history and elaborates on legal considerations, thus giving a solid basis for decision- making.

The scope of justice lens is designing, development, and implementation of Justice Lens as a responsive web applica- tion which will be uniquely catered to the needs of the Indian justice system. The idea behind this platform is that of using it as a means of first contact when it comes to legal advice, providing users with immediate access to their legal rights using a user-friendly interface. While being highly useful for legal research purposes and general knowledge dissemination, Justice Lens is not intended to be used as a substitute for professional legal advice.

II. Literature Survey and Gap Analysis

The intersection of Artificial Intelligence and Law has resulted in the development of various novel paradigms; how- ever, there are still numerous issues to be tackled.

A. Outcome Prediction and Temporal Dynamics

One of the most crucial issues faced within the realm of legal analytics is known as the ”temporal shift.” Legal precedents evolve over time, and their meaning can vary based on the time when they were established. Frameworks such as PILOT have made breakthroughs within the research domain, utilizing time-decay functions to mitigate the issue, making sure that the past legal precedent does not have too much sway on contemporary predictions. It takes into account the fact that the law in place in 1980 will not carry as much value compared to the groundbreaking verdict issued by the Supreme Court in 2024.

B. Domain-Specific Classifiers

While generic machine learning models (such as SVM or Random Forest) have been tested on legal data, their efficiency is often limited by a lack of domain-specific feature extraction. In early researches of the Indian context they utilized classic classifiers to predict outcomes for murder cases, achieving high accuracy but relying on ”tedious” manual feature selection across small datasets. The limitation of these classifiers lies in their inability to scale; they require human experts to manually label legal factors.Justice Lens advances this by shifting focus to automated classification, whereby the algorithm automatically determines the ”legal factors” from the raw text.

C. Transformers in Legal Information Retrieval

The introduction of Transformer architectures like Legal- BERT and GPT has revolutionized the way in which AI processes judicial language. These models utilize ”attention mechanisms,” to understand long-range dependencies in com- plex legal sentences where the interpretation of one sentence may depend on a word from three paragraphs earlier. Unlike typical language models, these transformer models are trained on particular legal datasets (for instance, ECHR or Supreme Court data), enabling them to understand the distinct subtleties of ”legalese.” [10].

D. *The Shift to Semantic Search*

The most significant paradigm shift in legal technology is the transition from Keyword Matching (TF-IDF) to Semantic Search (Vector Embeddings).This method works around the problem that traditional search engines face in cases where the user uses an alternative term or explains a scenario in ”plain English” instead of using technical legal language. Vectorization: Legal documents get transformed into numeric vectors in terms of semantics. Context Discovery: In a case where a user is looking for information related to ”unfair dismissal,” the software will be able to fetch cases related to ”wrongful termination” because they will share a common mathematical closeness, despite sharing no keywords.

III. System Analysis and Specifications

To ensure the strength of our Justice Lens, we carried out a comprehensive requirement analysis involving both hardware and software limitations.

System analysis for Justice Lens involves a meticulous as- sessment of the functional architecture, operating constraints, and multi-dimensional feasibility that is needed to overcome the legal accessibility challenge. At the fundamental level, the system is developed in such a way as to cater to the entire process of asking a legal question—from entering the query via natural language processing to synthesizing based on the semantics involved. By utilizing a pipeline that leverages the process of feature extraction in conjunction with a vector space model for retrieving information, our system can go beyond simply identifying keywords to comprehend the actual intent of the query. On the other hand, there are also strict requirements set regarding the non-functional qualities of the platform; specifically, sub-second latency and zero-training usability. It is imperative that our solution should be both technologically advanced and accessible to those marginal- ized members of society who are technically incompetent. Moreover, the development of a pipeline that facilitates model training offline ensures that the system remains flexible and resilient to the changing laws in India.

Feasibility Analysis is another way to prove the practicality of this architecture via evaluating different aspects of the project. First, from the technical standpoint, this approach makes use of highly mature open-source technologies like Python, FastAPI, and PyTorch, thus making it possible to implement high-complexity functionalities like Explainable Artificial Intelligence and Knowledge Graphs in the scope of current organizational resources. Secondly, from the opera- tional standpoint, this technology tackles the key issue of trust in the use of AI in legal aid applications as it offers evidence- based reasoning and, therefore, guarantees user acceptance in the population. The fact that it operates as a web application means its universality regardless of a particular environment. Finally, from an economical perspective, Justice Lens becomes a cost-effective way to make a huge impact on society, as it relies on free open-source software and current hardware..

A. *software requirements*

1] Python is an open source and high-level programming language with a very easy syntax, which is why it is widely used in the areas of artificial intelligence and web develop- ment. In the case of Justice Lens, Python is the primary lan- guage that will be used for backend development, along with frameworks such as FastAPI and libraries such as PyTorch and Hugging Face Transformers.

2] JavaScript is the language used for developing front-end web applications that can interact. In our case, JavaScript will be used to develop the client side interface such as the Dashboard as well as the chatting aspect.

3] FastAPI is a high-speed python framework to develop APIs. The backend server is built using FastAPI in order to provide API Gateway services. It will receive user requests from the frontend interface, process the requests, and forward the requests to the AI core.

4] PyTorch is a popular machine learning library that powers various deep learning algorithms, including natural language processing. In the case of Justice Lens, PyTorch forms the basis of training and fine-tuning the Legal-BERT model to develop vector embeddings for the semantic search engine.

5]Vector Database (Pinecone) is a database specifically built for storing and querying high-dimensional vectors. In our case, it plays an essential role in storing vector embeddings of all legal documents. The Semantic Retrieval Engine utilizes this vector database to conduct lightning-fast and precise similarity queries to locate the most appropriate legal document for the end-user.
MongoDB is a NoSQL database that is used for storing any generic data related to an application like user data and case data. As opposed to the vector database that stores the AI- generated embeddings, MongoDB will be used for storing operational data for the application to run effectively.

B. Hardware Requirements

Efficient training and inference require specific hardware capabilities:
1] Processor (CPU): A multiple core processor such as a Core i5 processor is needed to operate the application server and perform data preprocessing. This will be very useful especially when training models and serving multiple clients.
2] GPU (Graphics Processing Unit): A GPU plays an important role in making the calculations related to deep learning faster. A GPU having at least a capacity of NVIDIA GTX 1050 with 4GB VRAM is required for inference purposes, while a GPU of at least 8GB VRAM NVIDIA RTX 3060 will be required for training purposes.
3] RAM (Random Access Memory): 16 GB RAM is required to run the development environment and to handle the large datasets and AI models like Legal-BERT without performance issues during training and inference. Storage (SSD – Solid State Drive):The storage is provided by using 512GB SSD. The speed of accessing data in an SSD is essential to load the required files, which is very important in making the application more responsive.

IV. METHODOLOGY

The Justice Lens functions via a set of processes that are executed sequentially and concurrently to obtain, process, and output legal information contained in judicial documents. It aims to convert raw legal data into useful information by means of intelligent information retrieval, reasoning, and explanation techniques.

A. Knowledge Base preparation

The first phase involves building a comprehensive legal knowledge base, which serves as the foundation for all AI- driven reasoning and prediction tasks.

- Data Collection: A substantial body of legal texts in India, comprising statutes, case laws, and judgements, is sourced from credible open-source databases and legal libraries.
- Preprocessing: To clean and organize the raw data a structured text preprocessing pipeline is implemented . This includes tokenization, removal of noise, normaliza- tion, and chunking of long texts into smaller, context-rich passages for precision retrieval.
- Storage: The cleaned and structured data is indexed for fast search and retrieval, ensuring scalability and efficiency during inference.

B. Engineering the AI Core

JusticeLens uses AI Core as its primary technology which is an advanced machine learning component with a highly accurate algorithm.

- Model Architecture: The system employs a fine-tuned Legal-BERT transformer model specifically adapted for legal language comprehension.
- Contrastive Learning: To enhance contextual accuracy, contrastive learning is used for creating vector embed- dings that capture relationships between legal cases and queries.
- Vector Indexing: All embeddings are stored in a high- speed vector database, allowing instant and semantically rich retrieval during user interactions.

AI core ensures the system understands both the literal and contextual meanings behind user queries,

enabling accurate and relevant responses.

C. *Advanced Analytical Features*

JusticeLens integrates advanced analytical mechanisms to provide decision-support capabilities and interpretability in legal predictions.

- Prediction of Case Outcomes: Our system includes a time decay mechanism, which emphasizes contemporary cases, making sure that any form of legal reasoning adheres to recent judicial trends.
- Explainable AI: Knowledge graphs can assist our project in providing explainable artificial intelligence, whereby there will be logical reasoning for each decision provided by the software.

Together, these features make the system not only intelligent but also explainable, bridging the gap between automation and legal reasoning.

D. *Application Development and Deployment*

The final phase involves transforming the intelligent back- end into a fully functional, user-friendly legal assistant.

- Interface Design: An interface that uses chats is de- veloped using Python, providing users with easy-to-use and engaging interface features. Users can type their queries, and get answers that contain comprehensive and systematic legal information immediately.
- Deployment: The system is deployed on a scalable cloud platform, ensuring high availability, fast performance, and secure data handling for all users.
- Testing and Validation: Thorough testing is performed to guarantee the accuracy of the system. This includes testing through tools such as Postman, conducting tests on the backend, and doing user acceptance testing on the interface and its responses.

E. *Overall Workflow*

JusticeLens follows a logical and modular workflow from data collection and preprocessing, to AI-driven retrieval and explanation, culminating in an interactive user interface. Each layer of the system contributes to its overarching goal: democ- ratizing legal knowledge through artificial intelligence.

V. System Architecture and Implementation

Justice Lens is built on a decoupled, micro-service-oriented architecture.

A. *Component Breakdown*

- **User Interface Layer:** A chat-based interface where users can input queries in natural language form. The layer is responsible for rendering citations and descrip- tions.
- **Application Layer:** A FastAPI Python server responsible for the logic of application processing. It accepts the user query and forwards it to the embedding module before querying the vector database.
- **Data Persistence Layer:** Pinecone vector databases host the vectors, and regular databases persist user session information.

B. *The Semantic Feedback Loop*

To improve accuracy over time, our system implements a feedback loop. When a user finds a retrieved document helpful, the system records this interaction. This data can later be used to fine-tune the embedding model, ”pulling” related legal concepts closer together in the vector space.

VI. Results and Discussion

Initial evaluation of Justice Lens was conducted using a test set of 200 common legal queries. We analyzed the effectiveness of our semantic search engine in comparison to a keyword-based search engine in Elasticsearch.

A. *Accuracy and Relevance*

In terms of the semantic search, there was an increase in the accuracy of finding relevant laws by 35%, in case the keywords were not used in the search query (using "stolen property" instead of "theft," for example). This proves the model's capability to connect legal terminology with common language.

B. *Latency Analysis*

Despite the complexity of vector searching, the use of Pinecone's indexing allowed for an average retrieval latency of under 150ms. This ensures that the chatbot remains responsive, a critical factor for user retention in digital applications.

VII. Conclusion

In this paper, we have managed to develop a highly effective AI-powered framework called Justice Lens, which caters to the essential information gap in the legal domain in India. Lever- aging the shift from the use of keyword searching systems to a neural semantic search framework, the system enables us to bridge the chasm between the vast array of judicial data and practical legal insights. With the use of deep learning-based vector embedding and preprocessing by OCR techniques, our system comprehends all the intricacies of "legalese" in the same way an expert would. Using a combination of online and offline intelligence pipelines, we can deliver legal assistance that is both timely and context-dependent.

The primary value of this study can be attributed to the fact that it manages to bridge the gap in legal aid available to minority populations while at the same time improving the research processes for lawyers. Through the use of explainable artificial intelligence (XAI) and the implementation of a time decay function in the weighting of precedent cases, the model departs from being a black box prediction tool to a more understandable one backed up by data. Nonetheless, limitations associated with the quality and variability of the data used in training as well as the substantial computing power needed in transformer models cannot be overlooked.

In terms of future developments, the modular approach adopted by Justice Lens sets the stage for numerous significant improvements. First, future work on the Predictive Analytics module will aim at increasing the number of patterns observed in courts, as well as building the Automated Document Gener- ation component capable of creating standard legal motions. In addition, including multimodal data in case analysis is likely to be one of the goals of future work. Overall, while Justice Lens is certainly not expected to replace humans, it should act as an essential assistant for legal experts. In conjunction with the state-of-the-art natural language processing algorithms, the system stands to make a substantial contribution towards building a better legal environment.

A. *Future Directions*

Our future work will focus on:

The present-day architecture of Justice Lens serves as an excellent base on which many transformations can be made to take the platform to even greater heights. The key aspect through which this can be achieved is through further developments in the domain of Predictive Analytics. This can be done through using larger and lon- gitudinally structured sets of data pertaining to judgments issued by Supreme Court and High Courts, to analyze trends and predict likely outcomes of various courses of action. In addition to this, the current Answer Synthesis component can be improved into a document generation engine that prepares legal documents and requests based on Indian legal forms. In addition to integrating data from texts, future versions of the platform should incorporate Multimodal Data Integration through speech-to-text and computer vision technologies, thus providing a complete analysis platform that is able

to handle different kinds of case information. Lastly, building a standardized API can enable Justice Lens to work in conjunction with other legal practice management applications, thus making law firm and legal clinic work even easier. Through all of these improvements, as well as a commitment to using ethical AI principles, Justice Lens is sure to improve significantly but still remains an important analytical tool without which lawyers simply cannot do their job.

REFERENCES

[1] M. Dhore et al., "Bettercall: AI based legal assistant", *2024 ICIPCN*, DOI 10.1109/ICIPCN63822.2024.00048.
[2] L. Cao et al., "PILOT: Legal Case Outcome Prediction with Case Law",
[3] *arXiv:2401.15770v2*, 2024.
[4] R. A. Shaikh et al., "Predicting Outcomes of Legal Cases based on Legal Factors using Classifiers", *ICCIDS*, 2019.
[5] Y. Al-Shareef, "CHRExpert: An AI-Driven Court of Human Rights Expert Assistant", *IEEE Access*, 2025.
[6] D. A. Rincon-Riverose et al., "Automation System Based on NLP for Legal Clinic Assistance", *IFAC-PapersOnLine*, 2021.
[7] S. Marrivagu and A. Rao, "Artificial Intelligence for Legal Chatbot",
[8] *Journal of Engineering Sciences*, 2024.
[9] N. Jain and G. Goel, "An Approach to Get Legal Assistance Using Artificial Intelligence," *ICRITO*, 2020.
[10] V. Socatiyanurak et al., "LAW-U: Legal Guidance Through Artificial Intelligence Chatbot for Sexual Violence Victims and Survivors," in IEEE Access, vol. 9, pp. 131440-131461, 2021, doi: 10.1109/AC-CESS.2021.3113172.
[11] Binita Saha, Utsha Saha and Muhammadzubairmalik, QuIM-RAG: Advancing Retrieval-Augmented Generation With Inverted Question Matching for Enhanced QA Performance, IEEE Access (Volume: 12), 10.1109/ACCESS.2024.3513155.
[12] O.Ovadia,M.Brief,M.Mishaeli,andO.Elisha,"Fine-tuning or retrieval? Comparing knowledge injection in LLMs," 2023, arXiv:2312.05934.
[13] P.Lewis,E.Perez,A.Piktus,F.Petroni,V.Karpukhin,N.Goyal,H.Ku¨ttler,M. Lewis, W.-t. Yih, T. Rockta¨schel, S. Riedel, and D. Kiela, "Retrieval augmented generation for knowledge-intensive NLP tasks," in Proc. Adv. Neural Inf. Process. Syst., Jan. 2020, pp. 9459–9474.
[14] Bru¨ninghaus, S., Ashley, K. D. (2006). Progress in textual case- based reasoning: predicting the outcome of legal cases from text. Paper presented at the AAAI.
[15] Mackaay, E., Robillard, P. (1974). Predicting judicial decisions: The nearest neighbour rule and visual representation of case patterns.
[16] Gupta, S. et al. (2020). "A Machine Learning Approach to Generate Le- gal Recommendations from the Indian Constitution." IEEE International Conference on Artificial Intelligence and Law.
[17] Pandey, N. S., Rawat, P., Kathuria, S., Singh, R., Chhabra, G., Pant, G. (2023). Artificial intelligence assistance in the domain of law. In 2023 IEEE International Conference on Contemporary Computing and Communications (InC4) (pp. 1-4). Bangalore, India. https://doi.org/10.1109/InC457730.2023.10263230.
[18] Singh, R., Kumar, B. V., Rawat, K. K., Akram, S. V., Kathuria, S., Pachouri, V. (2023). Intervention of AI in prediction of judgement of cases. In 2023 4th International Conference on Smart Electron- ics and Communication (ICOSEC) (pp. 1268-1272). Trichy, India. https://doi.org/10.1109/ICOSEC58147.2023.10275850.
[19] M., Aldulaimi, S. H. (2022). Human rights and artificial intelligence: Evaluation of legal challenges and potential risk. In 2022 ASU In- ternational Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS) (pp. 361-367). Manama, Bahrain. https://doi.org/10.1109/ICETSIS55481.2022.9888888.
[20] Yupeng Chang, Xu Wang, Jindong Wang, Yuan , Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi: A Survey on Evaluation of Large Language Model. arXiv 2023.
[21] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang,and H.Wang,"Retrieval augmented generation for large language- models: Asurvey," 2023, arXiv:2312.10997.
[22] M.Kulkarni,P.Tangarajan,K.Kim,andA.Trivedi,"Reinforcement learning for optimizing RAG for domain chatbots," 2024, arXiv:2401.06800.