

¹Ayan J. Shaikh²Dr. Zainab Mirza³Fazal J. Parkar⁴Mohd Maaz A. Shaikh⁵Afeef Ahmed⁶Abdul Qayyum A. R. Khan

Chat Kanoon: A Novel Approach to Legal Assistance in India



Abstract: - This paper presents ChatKanoon, an innovative multilingual AI chatbot tailored for the Indian legal system. Utilising advanced language models like GPT-4 and Llama2 70B, ChatKanoon employs instructional techniques rather than traditional model fine-tuning, to provide contextually relevant legal advice. This approach addresses critical challenges in India's legal sector, including limited access to legal information, high service costs, and the scarcity of specialised guidance. ChatKanoon's development involved using a diverse array of legal documents and case laws, which enabled the chatbot to deliver precise and reliable legal information. The paper describes the application of instructional techniques in guiding language models, the process of dataset utilisation, and the development of an intuitive chatbot interface. ChatKanoon is positioned as a tool to democratise legal information, making it more accessible and affordable, thereby enhancing the efficiency of legal procedures in India. The paper concludes by discussing the current limitations and future potential of ChatKanoon, along with the broader impact of AI-driven legal assistance tools in developing countries. ChatKanoon exemplifies the transformative role of AI in the legal domain, with the potential to revolutionise legal aid in India and beyond.

Keywords: ChatKanoon, Generative AI, Natural Language Processing (NLP), GPT-4, LLAMA 2, Large Language Models (LLMs), Indian Law, Ethical AI, Societal Impact of AI, Indian Legal System

1 INTRODUCTION

The Indian legal system, despite its robust framework, faces significant challenges, chief among them being limited access to legal information and assistance. The high demand for legal services, juxtaposed against the complexities and costs of legal processes, presents a substantial barrier, particularly for a significant portion of the population and legal professionals. The integration of artificial intelligence (AI) in various domains offers a unique opportunity to mitigate these challenges. This paper presents ChatKanoon, an innovative AI chatbot tailored to provide multilingual legal assistance within the Indian legal framework, utilising the GPT-4 and Llama2 70B APIs.

1.1 Background of the Indian Legal System and Its Challenges

The Indian legal system, with its vastness and complexity, faces numerous challenges. Limited access to legal information and assistance is a significant issue, compounded by intricate procedures and a diverse linguistic landscape. This results in prolonged legal processes and an imbalance in the justice delivery system, with high legal service costs further alienating economically weaker sections of the population.

1.2 The Need for Accessible Legal Information and Assistance

The high demand for legal services in India highlights a stark gap in legal necessities and their fulfilment. This gap affects not only the general public but also legal professionals who contend with the evolving and intricate nature of Indian laws. There is a clear need for innovative solutions to make legal assistance more accessible, affordable, and efficient.

1.3 Introduction to ChatKanoon and Its Objectives

ChatKanoon is a pioneering multilingual AI chatbot designed for the Indian legal framework. It utilises advanced language models like GPT-4 and Llama2 70B, employing instructional tuning techniques instead of traditional model fine-tuning. This method allows ChatKanoon to provide accurate, reliable, and contextually relevant legal

¹ Ayan J. Shaikh

² Dr. Zainab Mirza

advice. The primary objective of ChatKanoon is to simplify legal processes for legal professionals and the public, democratising access to legal information, reducing legal advice costs, and enhancing the efficiency of legal procedures in India.

1.4 Overview of the Paper's Structure

This paper outlines the development journey of ChatKanoon, focusing on the use of instructional techniques with GPT-4 and Llama2 70B APIs, dataset utilisation, and chatbot interface design. It will explore ChatKanoon's features, its impact on the Indian legal system, and its potential to transform legal aid in India and other developing countries. The paper also discusses ChatKanoon's current limitations, future enhancements, and the wider implications of AI in legal assistance.

2 RELATED WORK

Taking a look at existing AI-driven legal assistance tools, as well as the unique challenges posed by the Indian legal context and the role of language models in legal AI, goes on to help set the context for where ChatKanoon fits in. This gap will be ascertained through a critical analysis of the current landscape of legal AI tools in general, and more specifically their applicability to the intricate and diverse Indian Legal system. This section lays the groundwork for examining how ChatKanoon addresses these specific needs and challenges.

2.1 Review of Existing AI-Driven Legal Assistance Tools Globally

The global legal sector has witnessed a growing integration of artificial intelligence, which promises a significant transformation in legal services delivery. This technological disruption challenges traditional business models in the legal sector, as seen in the UK legal services sector. These transformations are spurred by developments in automation and AI, alongside innovative new entrants in the legal tech space (Brooks, Gherhes, & Vorley, 2020).

2.2 Specific Challenges and Needs in the Indian Legal Context

India's legal system presents unique challenges that existing AI-driven legal assistance tools have yet to fully address. These challenges include the vast and intricate nature of Indian laws, the need for multilingual support, and the adaptation of AI tools to cater to state-specific legal systems.

2.3 The Role of Language Models in Legal AI

The role of advanced language models in legal AI is crucial, especially in the context of adapting these models to the unique requirements of different legal systems. This involves not only understanding complex legal language but also providing contextual information relevant to specific legal frameworks.

2.4 Highlighting the Gap that ChatKanoon Aims to Fill

ChatKanoon aims to address the gaps left by existing AI-driven legal assistance tools in the Indian context. Its focus on multilingual support and understanding the specific intricacies of Indian law positions it as a unique solution in the legal AI landscape.

3 USER INTERACTION FLOW AND EXAMPLE SCENARIOS

3.1 Scenario:

User is Seeking Legal Advice.

1. User opens ChatKanoon and selects his/her language.
2. User writes his/her legal query into the chat window, i.e. "The question is "If a person is found guilty of cyberbullying, what specific legal sections or statutes would apply to him and what potential penalties or consequences could he face based on these laws presently?".

3. ChatKanoon query processing is done by its language model which returns a detailed answer explaining the cyberlaw basis of Indian law.

4. Therefore, ChatKanoon addresses user’s follow-up questions that are preceded with clarifications and pragmatically gives information and counseling.

3.2 Scenario:

User Conducts a Legal Concept Research

1. The user opens ChatKanoon and picks up his/her desired language.
2. The user is interested in knowing more about certain legal concepts, for example, " What is the distinction between civil and criminal law? "
3. ChatKanoon simply discusses the distinction between civil and criminal law by showing examples.
4. Human talks to ChatKanoon in a dialogue, asking further questions to get a deeper understanding of the legal principle. In both cases ChatKanoon proves to be a suitable candidate for providing accurate legal information and advice in a format convenient for the audience who seeks guidance within the framework of Indian law.

4 DEVELOPMENT OF CHATKANOON

ChatKanoon's development revolves around the integration of the Llama2 70B and GPT-4 APIs, with a specific emphasis on enhancing cost-effectiveness within the context of Indian legal scenarios. The decision to leverage these APIs is grounded in their robust natural language processing capabilities, a crucial factor for addressing the intricacies of legal language and ensuring accurate contextual understanding. This section delves into the practical aspects of incorporating these APIs and the preprocessing steps are taken to optimise the performance of ChatKanoon on legal datasets.

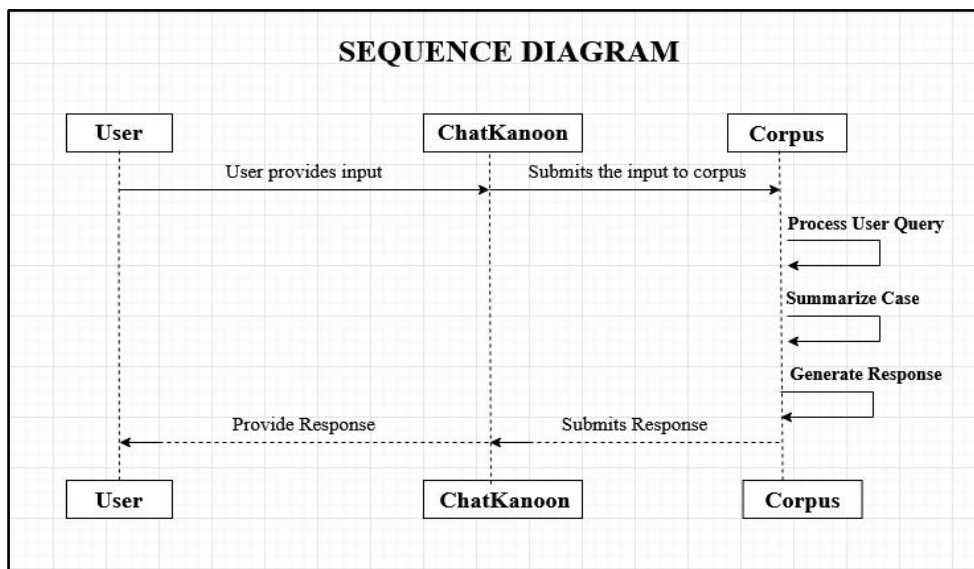


Figure 1: Sequence Diagram of ChatKanoon

4.1 Architecture and Technology Stack

- The Next.js Framework

The Next.js framework, which is a static web pages production environment based on React, is used by ChatKanoon to improve the loading speeds by providing server-side rendering. The reason Next.js is picked is

because of its scalability, user-friendly nature, and ability to deliver a smooth and interactive user platform. The framework's ability to assist in the integration of the API routes makes the process of backend integration easier for applications like the integration language models like Llama2 70B and GPT-4. It adopted a component-based architecture, with the use of Nextjs, which was important in simplifying the reusability and maintenance. Node.js, as a bridge between the front end and the language models, is used for server-side logic and API regulation. Vercel is used to host and deploy the Next.Js application due to its native support for the framework of Next.Js projects.

- **Folder structure**

The folder structure displayed in our application ChatKanoon, implemented by the Next.js framework, shows a respectable approach to structuring the application's assets and codebase. This construction is crucial to keep a sustainable, maintaining, and efficient application. Below is a description concerning folder structure based on the public and source (src) folders with the branch beneath them as well as their major components.

- **Public Folder:**

In Next.js there are projects such as public folders which act as standard directories for files to be accessed similarly by a browser using static files that can be accessed directly. Here, such resources that come out of the react component tree, include images, icons, and any other kinds of assets.

- **Images:**

This directory only has image files that are used in dynamic several views of the ChatKanoon application on the whole.

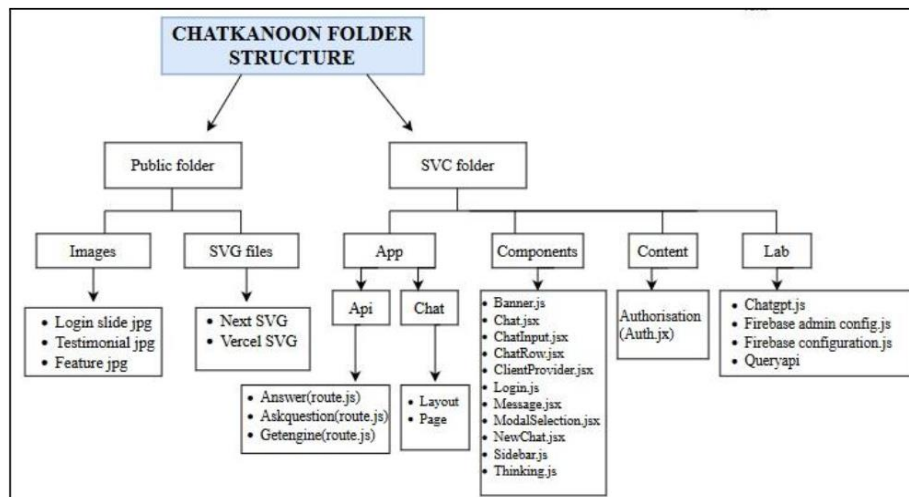


Figure 2: Folder Structure of ChatKanoon

- **loginside.jpg:**

A graphic that was most probably incorporated in the login and authentication pages to maintain eye-catching effects.

- **testimonial.jpg:**

One of the oldest standing features, utilised for displaying testimonials from users or endorsements, that help build credibility and provide security to users.

- **features.jpg:**

Features of ChatKanoon; some of the components that may appear in the section dealing with feature description for example details of the user completed the task.

- **next.svg and vercel.svg:**

These are SVG files, likely logos, that represent the technologies employed in the project – Next.js and Vercel.

- **Source (src) Folder:**

The source or src directory is where the directions of React and JavaScript are associated with the greater part of the application.

- **App Folder:**

Comprises functional units, about the core functionality of the application.

- **API:**

These API files like answer, ask a question, and get engine .route.js files may be hosted in this directory and have been critical in processing user queries and interactions with the chatbot.

- **Chat:**

Includes page and layout files, of which are essential for page arrangement as well as for sustaining chat interface pages.

- **Components Folder:**

Components used throughout the application are available under the Reusable React Components directory. Banner.js, Chat.jsx, ChatInput.jsx, ChatRow.jsx, and so on are code respective for various UI elements such as chat interfaces, input boxes, and message displays.

- **ClientProvider.jsx:**

As an example, it is likely that Likely also supplies the kind of context or state management of data that pertains to the client. Login.js, Message.jsx, ModalSelection.jsx, NewChat.jsx, Sidebar.js, Thinking.js: They are responsible for such aspects of inter-user communication as logging In, messages display, and managing the UI of a chat.

- **Context Folder:**

Homes the application's context files, such as Auth.js, which manages authorization and Authentication Context to allow each component to share User Data through the component tree without prop drill.

- **Lib Folder:**

Includes utility libraries or settings that make available the backward side and the front of the application.

Contains JS files for chatting with ChatGPT and interacting with Firebase (which provides outsourced backend services and database interactions), plus API querying capacities.

This structured approach not only provides an orderly separation of concern awareness, to segregate UI components, business logic, and static assets but also allows for the improvement of the development workflow that makes it easier for the developer to navigate and maintain the code. By following this organisation, ChatKanoon guarantees a modular, scalable architecture of the application, which can be developed along with the growing project.

5 FUNCTIONALITY AND FEATURES

ChatKanoon is tailored for this purpose. The major language models used in ChatKanoon include GPT-4 Turbo and Llama2 70B which cover several functions to enhance the legal process, give users access to legal documents, and follow the rule standard. This paragraph is about the reliability of ChatKanoon and giving legal advice to the app users.

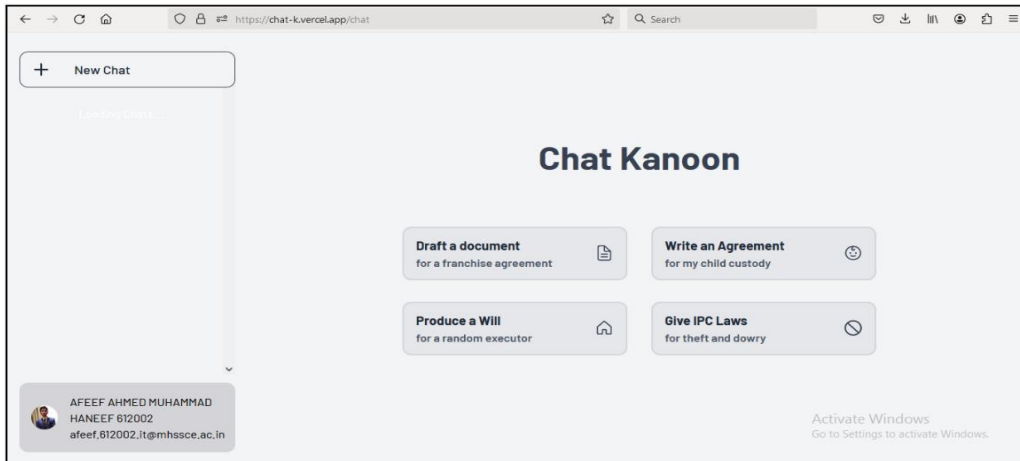


Figure 3: UI of ChatKanoon's homepage

5.1 Multilingual support:

One of the main advantages of ChatKanoon is that it provides services in different Indian languages including Hindi, Gujarati, Marathi, Tamil, Telugu Malayalam, Punjabi, Urdu, and Malayalam. Nevertheless, this makes it possible for people to obtain legal advice and detailed information of their choice in their language.

5.2 Legal advice:

ChatKanoon provides advice that follows the rules of good language standards called GPT-4 Turbo and Llama2 70B with recommended style, and offers legal information. The chatbot has learned to deal with legal issues by analysis of the text and providing the correct responses as per Indian laws.

5.3 Tailored for the Indian legal context:

Only for the Indian legal scene, ChatKanoon, which is uniquely made as to the extent and intricacy of Indian laws. ChatKanoon is also able to understand and obtain sensitive information on Indian laws and it also assists users with important information and state-wise facts.

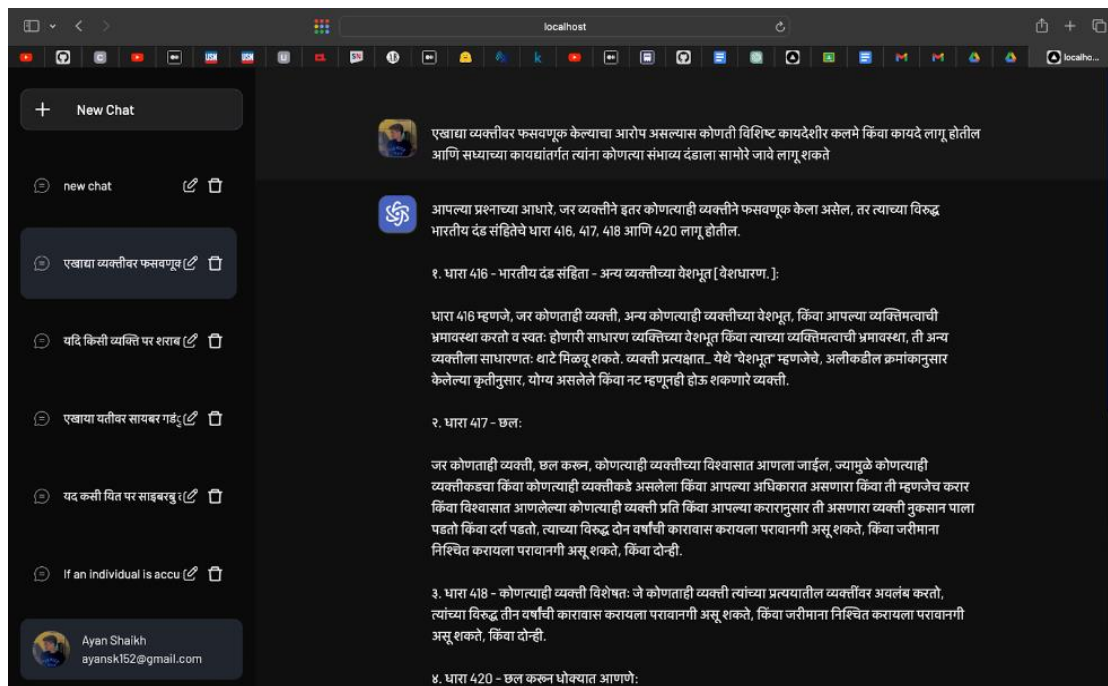


Figure 4.1: Prompt given in Marathi language

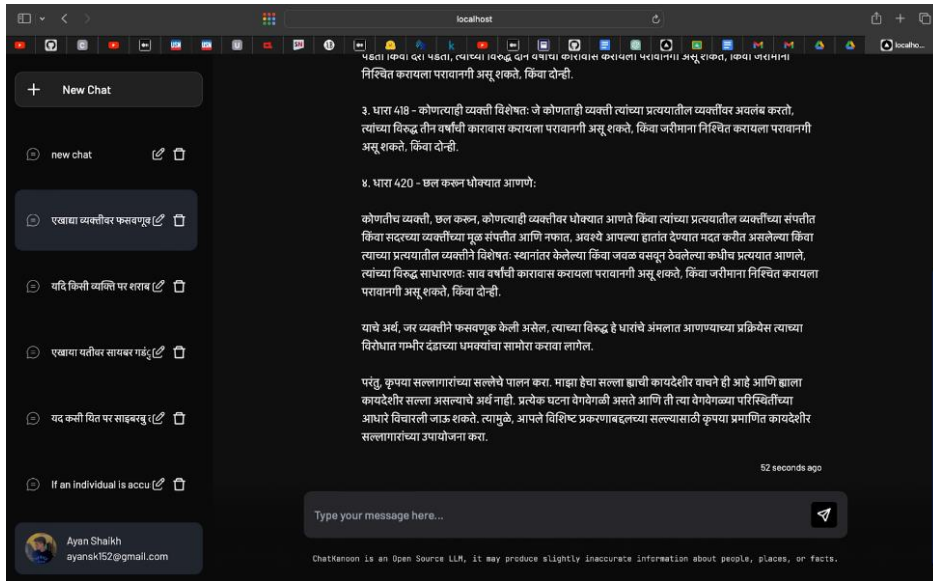


Figure 4.2: Prompt given in Marathi language

5.4 Legal advice:

ChatKanoon provides advice that follows the rules of good language standards called GPT-4 Turbo and Llama2 70B with recommended style, and offers legal information. The chatbot has learned to deal with legal issues by analysis of the text and providing the correct responses as per Indian laws.

5.5 Simplify the legal process:

The smart work enables ChatKanoon to give the correct answers to the legal questions on time and thus simplifies the law process. Thus, In the future it will help in reducing the accumulation of cases in the Indian courts and hence will lead to an overall improvement in the efficiency of the judiciary in the future.

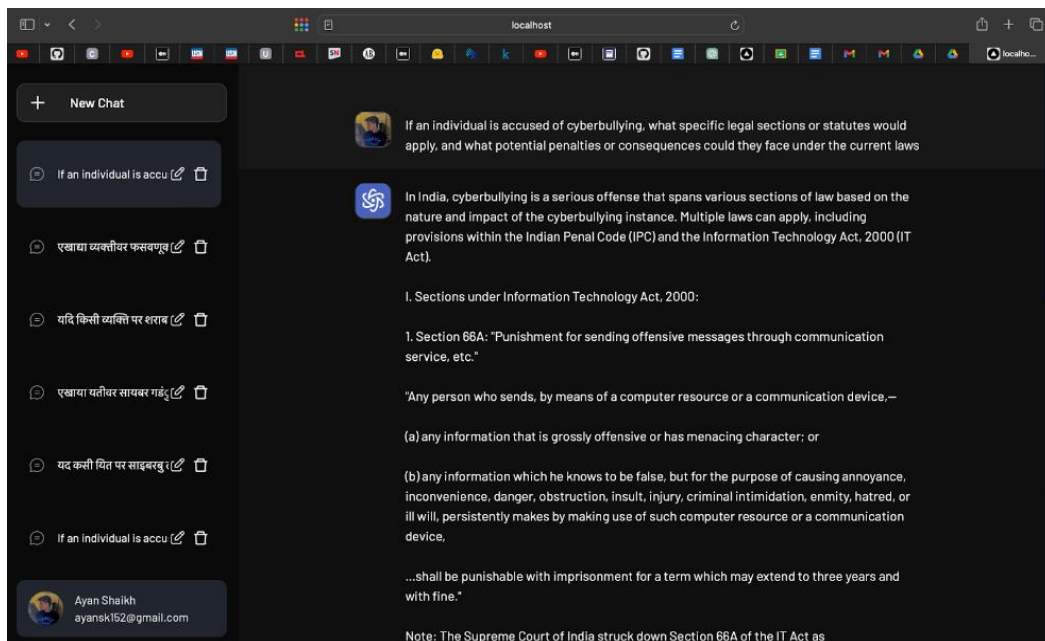


Figure 5.1: Prompt given in English language

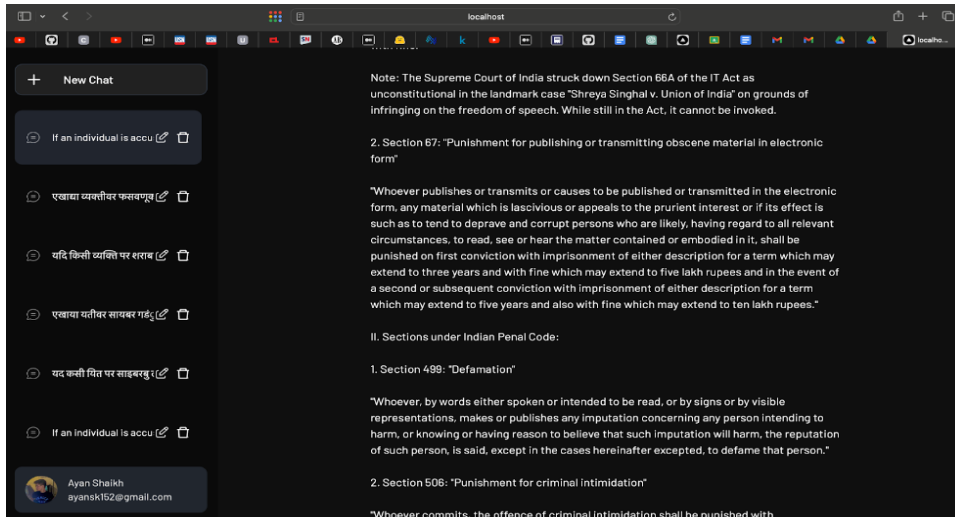


Figure 5.2: Prompt given in English language

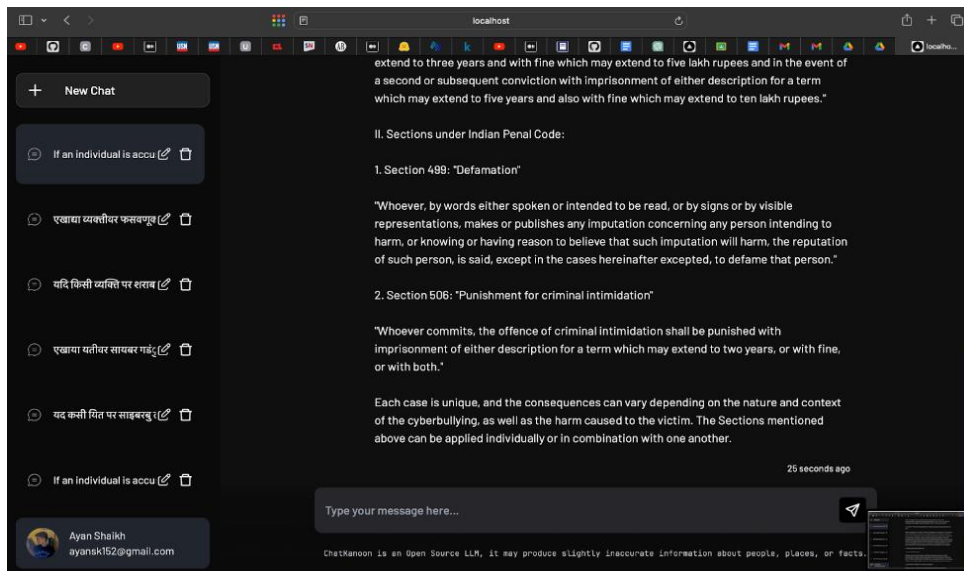


Figure 5.3: Prompt given in English language

6 IMPACT OF CHATKANOON ON THE INDIAN LEGAL SYSTEM

6.1 Enhancing Accessibility and Affordability

- ChatKanoon offers cost-effective legal advice using AI-supported interfaces, aiding low-income earners.
- It aligns to provide affordable legal services to the general population.

6.2 Increasing Efficiency and Reducing Case Backlogs

- ChatKanoon streamlines legal processes by providing timely and informed answers to legal queries.
- It reduces backlog in Indian courts, enhancing the overall effectiveness of the justice system.

6.3 Promoting Legal Literacy and Awareness

- ChatKanoon educates the public by offering accessible explanations of legal concepts and rights.
- Also empowers individuals with knowledge, enabling them to make informed decisions regarding legal matters.

6.4 Customization to Indian Legal Context

- ChatKanoon understands the history and context behind Indian laws, making it relevant to users.
- And provides state-specific legal information, catering to the diverse legal environments across the country.

6.5 Bridging the Digital Divide:

- ChatKanoon reaches rural areas, overcoming the gap in quality legal services between urban and rural populations.
- Promotes digital literacy by encouraging interaction with technology for legal information, contributing to technological skill-building.

7 LIMITATIONS AND CHALLENGES

7.1 Computational Requirements:

Our model requires high processing power and resources, as our chatbot utilises the models GPT-4 and Llama2 70B which demand significantly high computational resources due to their large size and complexity. So we lack the necessary computational resources.

7.2 User Interaction and Control:

The model may produce unexpected results, particularly if there are no specific constraints or directions. It is still a challenge to ensure control and user-friendliness of interfaces; to create more predictable and user-controlled results, it is necessary to improve the model's responsiveness to user inputs and the clarity of instructions.

7.3 User Data Privacy and Security Concerns:

Concerns regarding the protection of user data is one of the obstacles in our chatbot model. Since our model involves the exchange of information between individuals and the system, there are concerns about how sensitive data is being safeguarded. To combat these potential risks, it is crucial to have strong encryption protocols, secure data storage methods, and strict privacy policies.

7.4 Lack of Granularity in Instruction Fine-Tuning:

When fine-tuning a language model, precise instructions are vital to customise the model's responses effectively. However, finding the right level of detail in the instructions can be challenging. The instructions should be specific enough to capture small details, but not too specific to avoid over-specification. Striking the right balance is crucial. A new method called instruction tuning can help with this process.

7.5 Limited Regional Language Support:

Our chatbot model faces limitations in supporting all regional languages, impacting its ability to comprehensively address linguistic diversity in India. Even if it supports major languages like Hindi, Marathi, Urdu, Tamil, Telugu, Malayalam, and Punjabi, its inability to support some regional languages creates obstacles to achieving complete linguistic diversity and accessibility. This highlights the need for further efforts to enhance linguistic diversity and ensure accessibility for users across various linguistic backgrounds in the Indian context.

8 FUTURE ENHANCEMENTS

8.1 Advancing Ethical and Legal Competence:

Our ChatKanoon model stresses the importance of human oversight in legal contexts, urging users to verify advice with legal professionals. Future updates will focus on integrating larger datasets to cover more legal frameworks and ethical issues. We'll enhance the model with advanced NLP algorithms and regularly update its knowledge base with new legal developments. These enhancements aim to improve the model's ability to provide accurate and contextually relevant legal judgments, minimising the need for extensive verification.

8.2 Document Processing:

To improve our chatbot model, we plan to expand its ability to handle various document formats, including PDFs, Word documents, and spreadsheets. This development aims to enhance the model's versatility, making it easier for users to input and extract information from diverse sources. By accommodating various formats, the model will adapt to the evolving demands of users from different industries and contexts, promoting a more inclusive and user-friendly experience.

8.3 Enhancing user experience and usability:

Our future focus is on enhancing user experience and model usability, prioritising smooth, intuitive interactions via advanced natural language processing for concise, user-friendly responses. We will introduce interactive features, intuitive interfaces, and accessible prompts, leveraging user feedback to refine our model continuously. Emphasising user-centric design, we aim to create a chatbot that delivers accurate information and a fulfilling experience tailored to user needs.

8.4 Multilingual Fine-Tuning:

The upcoming improvements to the ChatKanoon project will focus on customising the AI chatbot to different languages to better suit specific legal systems and needs. This involves teaching language models like GPT-4 and Llama2 70B the nuances of various languages used in the Indian legal system to provide accurate and contextually relevant legal advice in multiple languages, thereby serving a broader user base and meeting diverse linguistic requirements within the Indian legal domain.

8.5 Enhancing Legal Knowledge Base for Improved Proficiency:

We'll add vast, recent legal datasets to ChatKanoon's knowledge base to update and grow it to improve it and overcome training restrictions. This seeks to increase its competence in certain legal areas and enhance comprehension of the intricacies of Indian law. By expanding its expertise, ChatKanoon will be able to provide more accurate guidance, improving its dependability and efficiency as a tool for legal help.

8.6 Facilitating User Connections with Legal Professionals:

Future enhancements to ChatKanoon will introduce a feature suggesting legal professionals to users based on their specific queries, facilitating a seamless connection between AI-driven assistance and expert human advice. This addition allows users seeking personalised or detailed legal guidance to directly contact qualified professionals whose details, expertise, and credentials will be integrated into the ChatKanoon website, enhancing user access to comprehensive legal assistance and tailored advice.

9 BROADER IMPLICATIONS AND CONCLUSION

9.1 The role of AI-driven legal assistance tools in developing countries:

AI-based legal assistance services have the potential to democratise legal access in developing countries, mitigating high costs and sparse legal infrastructure. While enhancing diversity in legal environments, cultural and contextual variations must be considered alongside ongoing collaboration with legal experts. AI integration requires both advanced AI expertise and deep legal knowledge, with tools necessitating regular updates to keep up with evolving AI and legal landscapes. This innovation promises efficient and accessible legal research, particularly in underdeveloped legal infrastructures.

9.2 The transformative power of AI in the legal domain:

Integrating advanced AI language models like GPT-4 and Llama2 into the Indian legal system brings major benefits. These AI tools help legal professionals analyse vast data, understand complex laws, and provide relevant legal information, revolutionising the legal field. However, as AI advances the legal landscape, ongoing ethical discussions and human intervention are crucial to ensure the efficiency and justice of Indian legal procedures.

9.3 The promise of ChatKanoon in revolutionising legal aid:

ChatKanoon has the potential to revolutionise the delivery of legal aid in areas where access to legal resources is limited. By leveraging specialised data sets and proactive language models, ChatKanoon aims to bridge the gap between the general public and legal knowledge. The chatbot's ability to deliver reliable and relevant legal information in an approachable manner can benefit anyone seeking advice on a range of legal matters. The rapid pace of both AI and legal developments means that tools will need to be updated regularly. Implementing AI in legal services is indeed technically challenging, but the rewards in terms of efficiency, accessibility, and speed are substantial.

9.4 A recap of key findings and contributions:

The integration of ChatKanoon, an innovative multilingual AI chatbot tailored for the Indian legal system, has the potential to significantly impact the legal landscape in India. By leveraging advanced language models like GPT-4 and Llama2 70B, ChatKanoon aims to democratise legal information, making it more accessible and affordable, and enhancing the efficiency of legal procedures in the country. The chatbot's ability to deliver reliable and relevant legal information in an approachable manner can benefit anyone seeking advice on a range of legal matters.

REFERENCES

- [1] “LexNLP: Natural Language Processing and Information Extraction for Legal and Regulatory Texts.” *Edward Elgar Publishing eBooks*, 2021, <https://doi.org/10.4337/9781788972826.00017>.
- [2] Kalyan, Katikapalli Subramanyam. “A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4.” *Social Science Research Network*, Jan. 2023, <https://doi.org/10.2139/ssrn.4593895>.
- [3] Jiao, Wenxiang, et al. “Is ChatGPT a Good Translator? Yes With GPT-4 as the Engine.” *arXiv (Cornell University)*, Jan. 2023, <https://doi.org/10.48550/arxiv.2301.08745>.
- [4] Mehrotra, Anay, et al. “Tree of Attacks: Jailbreaking Black-Box LLMs Automatically.” *arXiv (Cornell University)*, Dec. 2023, <https://doi.org/10.48550/arxiv.2312.02119>.
- [5] Chen, Weixin, and Bo Li. “GRATH: Gradual Self-Truthifying for Large Language Models.” *arXiv (Cornell University)*, Jan. 2024, <https://doi.org/10.48550/arxiv.2401.12292>.
- [6] Zhang, Xuanming, et al. “ProLex: A Benchmark for Language Proficiency-oriented Lexical Substitution.” *arXiv (Cornell University)*, Jan. 2024, <https://doi.org/10.48550/arxiv.2401.11356>.
- [7] Liu, Alisa, et al. “Tuning Language Models by Proxy.” *arXiv (Cornell University)*, Jan. 2024, <https://doi.org/10.48550/arxiv.2401.08565>.
- [8] Amato, Flora, et al. “An Intelligent Conversational Agent for the Legal Domain.” *Information*, vol. 14, no. 6, May 2023, p. 307. <https://doi.org/10.3390/info14060307>.
- [9] “LAW-U: Legal Guidance Through Artificial Intelligence Chatbot for Sexual Violence Victims and Survivors.” *IEEE Journals & Magazine | IEEE Xplore*, 2021, ieeexplore.ieee.org/document/9539229.
- [10] Hasal, Martin, et al. “Chatbots: Security, Privacy, Data Protection, and Social Aspects.” *Concurrency and Computation: Practice and Experience*, vol. 33, no. 19, June 2021, <https://doi.org/10.1002/cpe.6426>.
- [11] Devaraj, Prathab, et al. “Development of a Legal Document AI-Chatbot.” *arXiv (Cornell University)*, Nov. 2023, <https://doi.org/10.48550/arxiv.2311.12719>.
- [12] “Generative Artificial Intelligence: Trends and Prospects.” *IEEE Journals & Magazine | IEEE Xplore*, 1 Oct. 2022, ieeexplore.ieee.org/document/9903869.
- [13] Ahuja, Kabir, et al. “MEGA: Multilingual Evaluation of Generative AI.” *arXiv.org*, 22 Mar. 2023, arxiv.org/abs/2303.12528.

- [14] Brooks, Chay, et al. “Artificial Intelligence in the Legal Sector: Pressures and Challenges of Transformation.” *Cambridge Journal of Regions, Economy and Society*, vol. 13, no. 1, Jan. 2020, pp. 135–52. <https://doi.org/10.1093/cjres/rsz026>.
- [15] Muñoz-Soro, José Félix, et al. “A Neural Network to Identify Requests, Decisions, and Arguments in Court Rulings on Custody.” *Artificial Intelligence and Law*, Jan. 2024, <https://doi.org/10.1007/s10506-023-09380-9>.
- [16] Chalkidis, Ilias. “ChatGPT May Pass the Bar Exam Soon, but Has a Long Way to Go for the LexGLUE Benchmark.” arXiv.org, 9 Mar. 2023, arxiv.org/abs/2304.12202.