

# Legal Solutions – Intelligent Chatbot using Machine Learning

Kalpana R A

Department of Computer Science and Engineering

Sri Sairam Engineering College  
Chennai, India

kalpana.cse@sairam.edu.in

Karunya S

Department of Computer Science and Engineering

Sri Sairam Engineering College  
Chennai, India

karunyasrinivasan2003@gmail.com

Rashmi R

Department of Computer Science and Engineering

Sri Sairam Engineering College  
Chennai, India

rashmi5103r@gmail.com

**Abstract**—In the digital age, navigating legal complexities can be challenging for individuals lacking legal expertise or access to legal counsel. To address this issue, we present an innovative AI chatbot solution. This chatbot empowers users by offering fundamental legal knowledge, personalized instructions, and rapid responses aligned with specific document requirements and procedural nuances. Users also have the unique capability to engage in real-time discussions with qualified attorneys, enabling them to effectively manage intricate legal matters.

Furthermore, our system allows users to customize searches based on their financial circumstances and geographic location, ensuring legal guidance tailored to their individual context. Leveraging advanced machine learning techniques, the chatbot autonomously retrieves and processes information pertinent to users' particular legal concerns, ensuring timely and accurate responses. This comprehensive approach seeks to democratize legal resource access, providing equal opportunities for individuals across society to seek clarity and assistance for their legal needs. This paper outlines the chatbot's architecture, NLP techniques, and machine learning algorithms while highlighting its potential to revolutionize legal support services

**Keywords** – Legal issues, queries, assistance, chatbot, attorney consultation, paraphrase-MiniLML6-v2, natural language processing (NLP), machine learning, artificial intelligence, cosine similarity.

## I. INTRODUCTION

Many people are unaware of the legal methods that can be used to solve a problem. The legal industry will become more competitive in the digital era by modernizing and implementing chatbot technology. The scope and complexity of applications for legal chatbots increase as AI and ML technology develops. Chatbots can analyze many databases and papers for particular legal information using built-in keyword detection capability, providing quicker and more accurate responses.

Additionally, people encounter obstacles that prevent them from getting legal benefits, such as their location, financial situation, etc. There are websites that offer legal advice and inform users on a range of legal matters. These websites also offer information on lawyers and make it easier to get in touch with them. These websites just employ chatbots to gather information about the topic at present shortly before directing the user to an attorney or providing them with contact information for one. However, the chatbot cannot provide answers to the user's queries. This project can help people with such problems as it offers legal remedies. Through this effort, it is hoped to offer equal access and opportunity to everyone in every segment of society, despite existing disparities. This project consists of a chatbot that advises users on legal matters and serves as their own personal attorney. The user needs to first access the chatbot by entering the website. The user will next be required to provide the chatbot a variety of responses to some queries that need to be clarified. The user receives the subsequent solutions by the chatbot. If a user's inquiries are too complex for the chatbot to answer, it offers if they'd like a lawyer consultation and gives them contact information for a lawyer. A lawyer contacts them after a pre-arranged appointment to provide more information. This chatbot can also be incorporated with other software programs or websites run by attorneys who may consider offering legal advice to those who approach their websites via chatbots.

## II. PROPOSED SYSTEM

This project consists of an artificially intelligent chatbot[6] that counsels users on legal matters and serves as their own personal attorney. Fig. 1 shows the proposed architecture of the chatbot. The user must first access the chatbot by going to the website. Then provide the chatbot a variety of information, depending on whether they need to obtain information through the chatbot or speak with a

lawyer. The chatbot is going to request them for some basic details and ask them to choose the category that best describes their problem. In order to provide users precise replies, the chatbot then further categorizes their queries. To discover the necessary phrases in the database, Natural Language Processing Techniques (NLP)[7] like tokenization [7] and lemmatization are used.

#### A. User Module

In a chatbot system, the User Module serves as vital for engaging with people, acquiring data, and giving pertinent answers. A synopsis of the activities and procedures found in the Client Component is provided below:

1) *User Information Collection*: When a user accesses the chatbot, the User Module starts by collecting basic user information, which could include their name, contact details, and any other pertinent data required for the interaction.

2) *Initial Query Handling*: As a user originally makes a query, the module prompts them to specify their initial goal or query type. This includes alternatives like asking attorneys for counsel or clarification on matters of law. The module gives users access to a database of lawyer information if they choose the option to seek legal counsel. The user can then pick the attorney who best meets their requirements.

3) *Problem Identification*: If the user does not want to contact a lawyer immediately, the module offers additional choices to help narrow down the user's query. This entails providing the user with a list of frequent legal issues or categories from which to choose.

4) *Query Processing*: This examines its database for pertinent information and responses after the user selects an issue category about their query. This involves legal articles, FAQs, or predetermined responses. Given its database and understanding, the module analyzes these queries and provides the relevant information in response.

5) *Referral to Lawyers*: The User Module recommends users to contact a lawyer for qualified advice if the user's question is sophisticated or falls outside the chatbot's capabilities. It can offer to set up an appointment with a lawyer or give contact information. Additionally, it has the capability of assisting users in setting up legal consultations with the chosen attorney, simplifying the procedure for users who need legal counsel.

6) *Continuous Interaction*: Throughout the interaction, the component maintains a conversational context, allowing users to ask follow-up questions or seek additional information as needed.

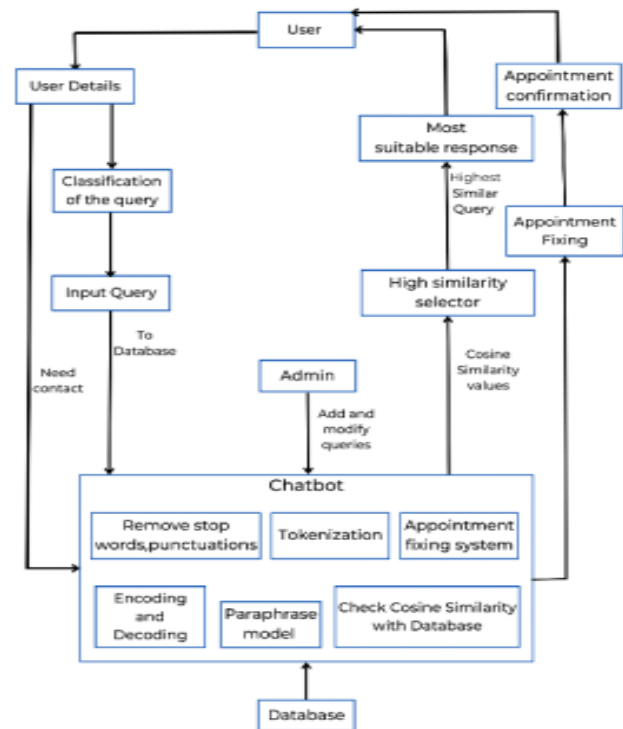


Fig. 1: Chatbot architecture

#### B. Administration Module

The Database Administration Component of the chatbot system serves a critical role in maintaining and updating the database used by the chatbot. Here are the key functions and responsibilities of this component:

1) *Backend Access*: The Database Administration Component is typically accessed from the backend of the chatbot system. This is where database modifications and updates are carried out. It may involve using specialized tools or interfaces designed for administrators.

2) *Database Maintenance*: New data can be entered into the database by the administrator. This includes updating any existing legal content that the chatbot utilizes to answer to user inquiries, such as articles, or frequently asked questions. This is necessary to maintain the accuracy and currency of the material. Legal databases frequently need to be updated frequently to account for changes in laws and regulations.

3) *Query Creation*: This entails specifying the chatbot's responses to particular user inputs or inquiries that might not have been included in the initial database.

4) *Security measures*: It is placed to safeguard the database's integrity. To avoid unauthorized changes or data breaches, only authorized people should have access to this component.

### III. LITERATURE SURVEY

In [1], a standalone application that can also be embedded into websites or mobile applications calculates the stress associated with the given details shortly after being provided with various details. It makes decisions and breaks down the input into various ranges that are then utilized in determining the user's level of stress by applying logistic regression, KNN Classifier, Decision Tree, Boosting, Bagging, Random Forest Classifier, and The Naive Bayes Classifier.

The project in [2] focuses on implementing the WhatsApp app to build a Chatbot platform using Selenium, a Raspberry Pi 4B, as well as a GSM modem through applying Tokenization, Lemmatization, as well as Stemming in the acquired interpersonal interaction. Python employs an algorithm to examine the recently received messages and verify for resembles in the already specified constraints. If matches are identified, a response is delivered; on the contrary, analyzing the incoming information inputs continues to happen.

[3] integrates a chatbot with a college website to provide students the information they require. The chatbot retains input from users and responds each time it receives input from the user. This enables the chatbot, which initially had minimal knowledge, to gain expertise using the accumulated responses. depending on the AIML language. In order to provide the pupils, the necessary information regard- ing the college, it takes advantage of the WordNet Algorithm, SHA-1 Encryption Algorithm, and SHA-256 Encryption Algorithm. By processing the data entered by applying WordNet, the system provides the necessary information.

This paper [4] investigates and assesses how effective chatbot technologies are at assisting users who are learning on e-learning platforms, as well as how these technologies may be able to lessen issues like feelings of alienation and detachment. A chatbot created explicitly to serve as an assistance for online learning is being researched. The retrieval-based model and the QANet model serve as the foundations for this chatbot's NLP core. This hybrid chatbot with two models is intended to work in conjunction with a platform for e-learning. This chatbot's primary response context was created with both course contents and casual conversation in mind, which gives it the appearance of a human companion.

[5] LAW-U is an Artificial Intelligence (AI) chatbot designed to offer legal guidance and advice to victims of sexual assault. This innovative chatbot assists victims by recommending relevant Supreme Court decisions that align with their specific circumstances. It makes use of fictitious dialogue taken from Supreme Court rulings and legal professionals who were trained to create Natural Language Processing (NLP) pipelines for LAW-U. The average similarity score between the Supreme Court cases gathered and the input sets provided by legal experts were

calculated using spaCy. It entails using NLTK's Wordnet corpus to look for synonyms of the frequent terms in the cases that were recovered via the TF IDF algorithm's extraction of common keywords between the cases and the mock-up message inputs.

### IV. IMPLEMENTATION

#### A. User Module

1) *User Interface*: The User Module is highlighted by an intuitive graphical user interface (GUI) crafted using web technologies like HTML, CSS, and JavaScript.

2) *Database Integration*: The database technology of choice, phpMyAdmin, is used. It connects the chatbot to a database on the back end that contains vital data including lawyer profiles, legal queries, and responses. The rapid data retrieval capabilities of this connection will provide users with accurate and recent legal information. In a changing legal context, data accuracy is maintained through routine synchronization.

3) *Decision Tree*: To guide users effectively, a decision tree is implemented. This aids users in selecting the appropriate service—whether it's accessing legal information or connecting with an attorney—based on their specific needs. It begins by analyzing user intent, discerning whether they seek legal insights or wish to engage with an attorney. By offering tailored questions and choices, the decision tree streamlines the user experience, ensuring that users are directed to the most relevant services..

4) *Response Generation*: When users seek information, to comprehend and process user queries accurately, natural language processing (NLP) techniques are employed in order to generate coherent responses. Pre-defined models like paraphrase-MiniLM-L6-v2 are leveraged to enhance response quality.

a) *Natural language processing*: NLP employs computational techniques and algorithms to analyze, model, and create natural language. It aims to enable the comprehension and interpretation of human language by computers. A customer input text stream is tokenized, or broken up into individual tokens like words or letters. Then, lemmatization is carried out which involves reducing a term to its root word to find similarity in meaning. It then builds a vector representation of the actual question after trimming stop words and punctuation using a pandas data frame [7].

b) *Paraphrase-MiniLM-L6-v2*: It is a pre-trained NLP model from sentence transformer which is built using machine learning techniques. It was created

using a large corpus of text data and a paraphrase detection task to encode text into dense embeddings for a range of NLP applications, including semantic search and text categorization. It is based on the MiniLM architecture. The query is cleaned up after pre-processing, which involves changing all the strings to lowercase, and is then encoded as a high-dimensional vector using a sentence transformer model called paraphrase-MiniLM6-v2 [8].

c) *CosineSimilarity()*: It is a function from the PyTorch library, which calculates the cosine similarity[8] between a user query that has been encoded and a set of previously encoded question embeddings. In the cosine similarity function, the dot product of two embeddings is split by the product of their magnitudes. The resultant value ranges from -1 to 1 where 1 indicates that the two embeddings are precisely the same. The function's return result, which is the index of the question with the highest cosine similarity score, is used to extract the related answer from a pandas dataframe that already has columns for questions and responses. If there is a significant degree of resemblance, the appropriate response is displayed; else, a default answer is given.

5) *Appointment Scheduling*: This involves users effortlessly choosing the times and dates for legal consultations. This module works with the backend of the system to verify the accessibility of legal experts and appointments, sending users alerts upon confirmation and maybe adding appointments to their calendars. It simplifies the scheduling procedure, improving user convenience and making sure that appointments with legal experts are successfully scheduled.

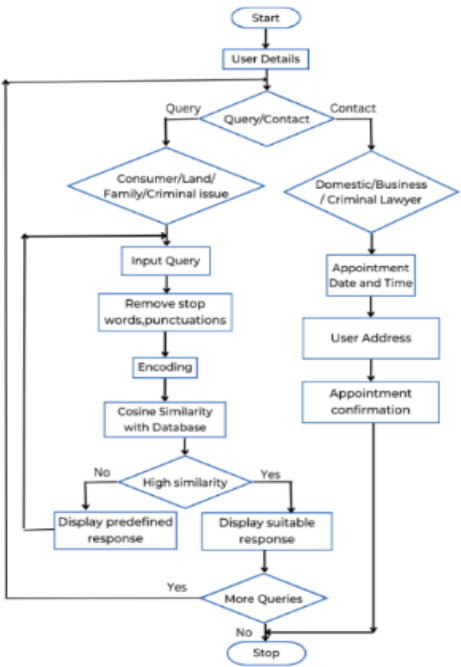


Fig. 2: Flowchart for User Module

### B. Admin Module

1) *Authentication*: Username-password pair mechanism is integrated into the Admin Module, ensuring that only authorized personnel can access and modify the database. This strategy improves security by limiting who can access and change the database to those who are allowed to do so. This protects sensitive data while preserving the system's integrity and secrecy.

2) *Database Management*: The module seamlessly integrates with the phpMyAdmin database system, enabling administrators to manage data by adding, updating, or deleting records. This involves executing SQL commands or NoSQL queries, depending on data management needs. Data validation mechanisms are in place to uphold data integrity, enforcing rules that ensure consistency and prevent erroneous or malicious data entries. This integration ensures efficient data management and safeguards the quality and security of the database.

3) *Logging and Monitoring*: The Admin Module is integrated with robust logging and monitoring tools that track system activities, detect anomalies, and enable swift issue resolution. These tools record and analyze user actions, database changes, and administrative operations, identifying any unusual behavior or security concerns. Real-time alerts notify administrators when issues arise, allowing for immediate response and mitigation, ultimately enhancing system security and reliability.

4) *Security Measures*: The system prioritizes security through stringent measures such as data encryption and access control to protect sensitive data within the database. It employs data encryption to transform information into a secure format, ensuring its confidentiality, and access control restricts system interactions to authorized users with proper permissions, reducing the risk of data breaches. Routine security audits and updates further bolster the system's defenses by identifying and promptly addressing vulnerabilities, reinforcing its resilience against potential threats.

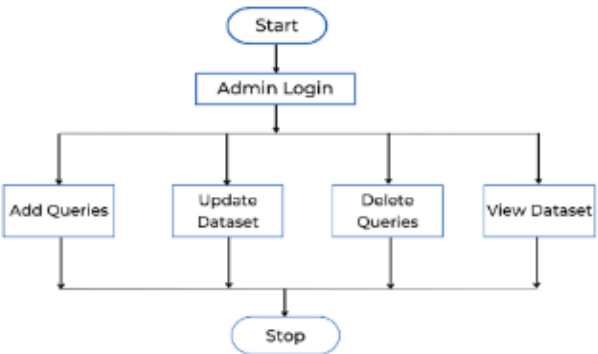


Fig. 3: Flowchart for Admin Module

## V. RESULT

This study introduces a sophisticated AI chatbot tailored to the legal field, aiming to provide users with comprehensive assistance, educate them, and easy access to legal services. Retrieving a response from a chatbot is one based on information. The replies are retrieved by the chatbot, which in this case contains all of the senses that are performed when the user query matches. The chatbot's operation involves several key components and a user-friendly graphical user interface (GUI). When users access the chatbot's GUI, they are prompted to provide essential contact information, encompassing their name, email address, and phone number. This initial registration step serves to personalize the user experience and maintain a channel for future communication. After registration, users are presented with service options. They can choose between engaging in a chat session or seeking assistance from a lawyer. This flexibility ensures that users can opt for the level of assistance that best suits their needs as shown in Fig. 4.

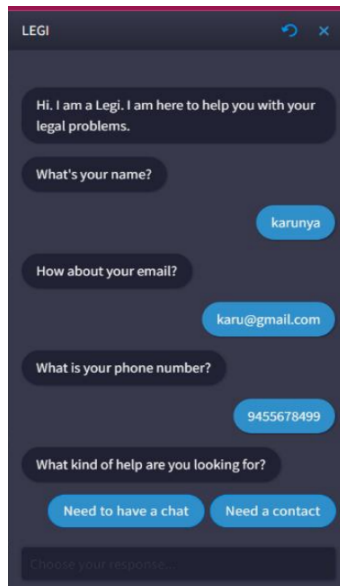


Fig 4: Providing options

For users opting for a chat session, the chatbot initiates a structured process to gather information about their specific legal issue. Through a series of options and questions, the chatbot identifies the nature of the user's problem. This step is crucial in ensuring that the chatbot provides targeted and relevant assistance. When a user enters their legal query into the chatbot's GUI, the chatbot employs a robust query processing mechanism. It matches the user's query with the relevant data in its dataset, which includes a repository of legal information and responses. If a matching query is found within the database, the chatbot promptly presents the corresponding response to the user, offering accurate and tailored information as shown in Fig. 5.

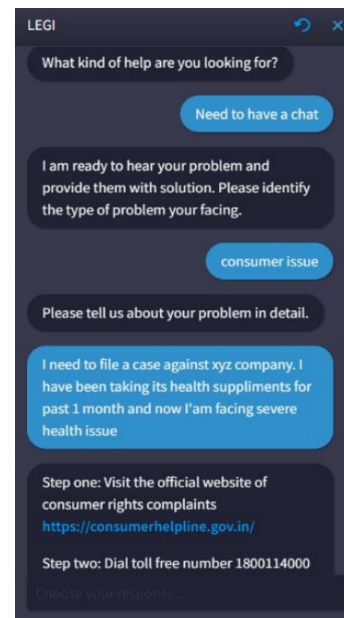


Fig 5: Answering queries

In cases where the chatbot does not find a matching query in its dataset, it doesn't leave the user without guidance. Instead, it provides a predetermined response and extends an option to the user. This option allows users to schedule an appointment with an appropriate lawyer. This recognition of the limitations of the chatbot and the seamless transition to expert legal assistance is a valuable feature for users with complex legal needs. For users seeking legal assistance, the chatbot excels in simplifying the appointment scheduling process. It matches users with registered lawyers available on the website, considering the availability of both parties. This seamless scheduling of appointments enhances user convenience and accessibility to legal services as shown in Fig. 6.

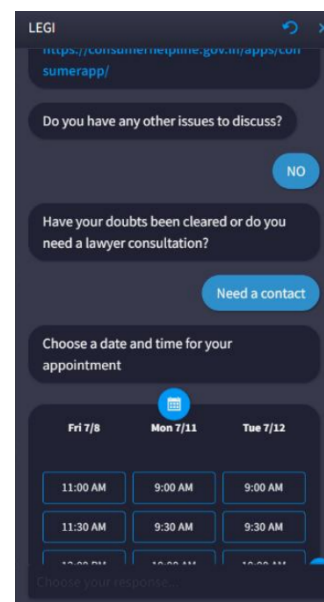


Fig 6: Fixing appointment

By logging into the admin page on the website, the admins can add, update or delete the query, response datasets, and lawyer information into the database. To log in, the admin must provide a username and password. Only with proper authentication is the admin permitted to access the database. Once the administrator is logged into the admin page, the admin can add queries and lawyer information to datasets, modify existing databases, edit predefined statements, delete existing data, and can change the admin login password.

## VI. CONCLUSION

In this project, we created a website integrated with a chatbot that provides legal advice to users and acts as their personal attorney. This system is made up of several blocks, including data creation, Natural Language Processing (NLP)[7] techniques such as tokenization and lemmatization to break down and process sentences, Machine learning models such as paraMiniLM-L6-v2 that compute embedding for the dataset and queries, and Cosine Similarity to retrieve the most appropriate response to the client's query. If the chatbot is unable to resolve the client's issue or if the client requires the services of a lawyer, this system can schedule an appointment for the user with the appropriate lawyer. This website can help people who are unable to obtain legal benefits due to their location, financial situation, or other factors by providing legal answers to their questions. Through online clients, this website also benefits the registered lawyers of the website. The chatbot can be integrated with an attorney's official website, increasing the website's efficiency.

## VIII. FUTURE WORKS

The chatbot can be enhanced with a voice assistant. Voice To-text can be added as an additional option that would be more convenient for users. To improve the chatbot's effectiveness and efficiency in resolving user's queries, as well as to obtain more relatable and accurate responses, the chatbot can be trained with more judicial datasets and the scope of information. The module's efficiency in providing solutions can be increased by exposing it to a variety of languages and topics. The chatbot's latency period can be reduced.

## REFERENCES

- [1] Tushar Sharma, Jitendra Parihar, Saurabh Singh, "Intelligent Chatbot for Prediction and Management of Stress," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence).
- [2] Achmad Ramaditiya, Suci Rahmatia, Aris Munawar, Octarina Nur Samijayani, "Implementation Chatbot Whatsapp using Python Programming for Broadcast and Reply Message Automatically," 2021 International Symposium on Electronics and Smart Devices.
- [3] Hrushikesh Koundinya K., Ajay Krishna Palakurthi, Vaishnavi Putnala, Ashok Kumar K, "Smart College Chatbot using ML and Python," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN).
- [4] Eric Hsiao-Kuang Wu, Chun-Han Lin, Yu-Yen Ou, Chen-Zhong Liu, Wei-Kai Wang, Chi-Yun Chao, "Advantages and Constraints of a Hybrid Model K-12 E-Learning Assistant Chatbot," IEEE Access, Vol. 8, 2020.
- [5] Vorada Socratianurak, Nittayapa Klangpornkun, Adirek Munthuli, Phongphan Phienphanich, Lalin Kovudhikulrungsri, Nantawat Saksakulkunakorn, "LAW-U: Legal Guidance Through Artificial Intelligence Chatbot for Sexual Violence Victims and Survivors," IEEE Access, Vol. 9, 2021.
- [6] Silvia Garcia'ia-Mendez, Francisco De Arriba-Perez, Francisco J. Gonzalez- Castano, JOSé A. Regueiro-Janeiro, "Entertainment Chatbot for the Digital Inclusion of Elderly People Without Abstraction Capabilities," IEEE Access, Vol. 9, 2021.
- [7] Reiss, Frederick, Bryan Cutler, and Zachary Eichenberger, "Natural Language Processing with Pandas DataFrames," 2021.
- [8] Ramnarain-Seetohul, Vidasha, Vandana Bassoo, and Yasmine Rosunally, "Work-in-Progress: Computing Sentence Similarity for Short Texts using Transformer models," 2022 IEEE Global Engineering Education Conference (EDUCON).