

Dr. Mahalingam College of Engineering & Technology  
Department of Artificial Intelligence & Data Science  
19ADPN6601 – Innovative and Creative Project  
AI Legal Advisor

Lakshmitha A  
727622BAD018

Shambuvaruni M  
727622BAD022

Deekshiba PV  
727622BAD046

**Abstract:**

In many industries, including the legal field, the development of artificial intelligence has created new opportunities. The AI Legal Advisor is a web-based application that automates the petition processing process to modernize and expedite the redressal of legal grievances. Conventional systems are frequently sluggish and prone to mistakes, which causes delays and incorrect classification. This technology reduces administrative effort by classifying petitions into civil, criminal, or invalid categories using an improved BERT model. It also detects duplicates using semantic approaches like TF-IDF and cosine similarity. Additionally, it has a lawyer-matching module that matches suitable attorneys according to the district and type of case. The system, which was developed with Python, Flask, and SQLite, has a role-based interface for administrators, attorneys, and users. It also has dashboards specifically designed for tracking cases in real time. Its modular and scalable design enhances efficiency, transparency, and accessibility in legal grievance management.

**1. Introduction:**

Artificial intelligence (AI) has revolutionized data handling in many industries, and there is a lot of promise for its use in public grievance resolution. Citizen petitions are essential feedback mechanisms in democracies, but manual processing frequently results in delays, inefficiencies, and a lack of transparency.

To overcome these obstacles, the AI Legal Advisor project uses natural language processing and deep learning to automate petition classification, prioritize cases, monitor progress, and assign suitable attorneys based on location and context. The system, which was created with BERT, Flask, and SQLite, combines clever back-end processing, real-time updates, and an intuitive user interface. The concept, architecture, and implementation of the system are examined in this study, emphasizing its potential as a scalable solution for updating the processing of administrative and judicial grievances.

**2. Key Technologies:**

**2.1 Natural Language Processing:**

A fundamental component of artificial intelligence, Natural Language Processing (NLP) aims to provide robots the ability to comprehend, interpret, and react to human language in a meaningful way. NLP is essential in the legal field for automating the examination of unstructured text, including court decisions, legal notices, and petitions. Tokenization, named entity recognition, part-of-speech tagging, and syntactic parsing are some of the methods that NLP uses to help extract useful information from vast amounts of legal text.

Because of these features, NLP is especially good at comprehending the structure, context, and intent of legal documents - which are frequently verbose and complex. Including NLP in legal tech applications improves decision-making effectiveness

and streamlines the user experience for petitioners and attorneys alike.

## **2.2 Term Frequency–Inverse Document Frequency:**

A popular statistical technique in text mining and information retrieval, TF-IDF measures the significance of words in a document in relation to a broader corpus. It computes two metrics: inverse document frequency, which measures how uncommon or unique a word is across all documents, and term frequency, which measures how frequently a word appears in a document. These two measures can be used to find keywords and related subjects since they highlight terms that are both common in a particular document and unique throughout the corpus.

TF-IDF can be used in legal applications to find the main concerns in petitions or to find similarities between documents for tasks like case grouping or duplicate identification. It serves as a foundational approach for more advanced semantic analysis techniques in text-based legal systems.

## **2.3 Bidirectional Encoder Representations from Transformers:**

By providing deep bidirectional context awareness, Google's transformer-based language model BERT has made tremendous strides in the field of natural language processing. BERT reads simultaneously in both directions, which allows it to understand a word's entire context based on the surrounding text, in contrast to standard models that read text either left-to-right or right-to-left. Because of this, BERT is particularly effective at classifying tasks, answering questions, and identifying entities. BERT is used in legal systems to categorize petition content into administrative, criminal, and civil categories, increasing the precision and

effectiveness of processing legal documents. It can generalize well across a variety of domains thanks to its pre-training on large datasets, and it can perform even better on domain-specific tasks by fine-tuning on legal datasets.

## **2.4 Longformer Encoder-Decoder Tokenizer:**

The problem of processing lengthy documents that beyond the token limits of conventional transformers such as BERT is addressed by the LED Tokenizer, which is a component of the Longformer Encoder-Decoder paradigm. Thousands of words are frequently found in legal filings, with important information strewn throughout. Using sparse attention techniques, the LED Tokenizer effectively parts and processes these lengthy texts, enabling the model to expand to longer inputs without seeing a notable rise in computing cost. This feature is critical for legal applications because correct analysis and classification depend on maintaining a document's whole context. The LED Tokenizer improves the system's capacity to comprehend complex meanings and long-range dependencies in legal language by keeping a larger portion of the original information.

## **2.5 Legal BERT:**

Pre-trained on a vast collection of legal texts, such as statutes, case law, contracts, and court rulings, Legal BERT is a domain-specific version of the original BERT model. Legal BERT is better equipped to comprehend the structure, terminology, and semantics specific to legal language thanks to this specialist training. Legal BERT is specifically designed to capture legal jargon and context-specific meanings, which makes it extremely effective for tasks like legal document classification, citation prediction, and legal question answering.

Generic language models sometimes struggle with these nuances.

Legal BERT greatly increases the accuracy of legal petition classification and strengthens the system's capacity to make context-aware recommendations when incorporated into programs such as the AI Legal Advisor. It is a significant step toward using AI in the legal field more accurately and consistently.

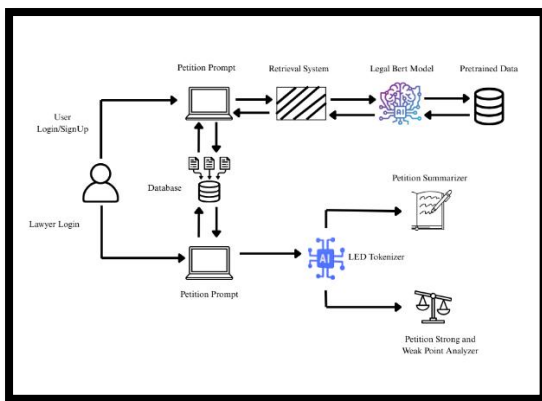


Figure 1. Proposed Block Diagram

### 3. Modules

Two main user interfaces—one for petitioning individuals and another for legal experts (lawyers)—are the foundation of the AI Legal Advisor platform. With the help of a centralized backend, each module is expertly crafted to satisfy the distinct requirements and duties of its user group. Routing, authentication, and model inference are handled by the Flask framework, and permanent storage for user information, petitions, and case updates is handled by SQLite. From petition filing to resolution, the system's coordinated functioning, scalability, and maintainability are guaranteed by its modular design.

#### 3.1 User Module

The user module is designed specifically for citizens who want to submit online legal petitions. Users are granted protected access to a dashboard where they may

submit grievances, monitor the status of petitions, and examine lawyer assignments after completing a registration process that requires a name, email address, and password. A refined BERT model is used to analyse submitted petitions and categorize them into three predetermined groups based on contextual textual understanding: civil, criminal, and invalid. To prioritize issues and expedite administrative processes, duplicate or recurring petitions are also identified using the TF-IDF and cosine similarity algorithms.

Users can track a petition's progress through the following phases after it is submitted: "Filed," "Reviewed," "Assigned," and "Resolved." The technology shows the designated attorney's name and provides real-time updates through flash notifications. To promote openness and participation, users can also start a conversation with the designated attorney. For non-technical users, the legal grievance process is easy to understand and engaging thanks to the user interface's emphasis on accessibility and simplicity.

#### 3.2 Lawyer Module

Registered attorneys who have undergone secure admin validation are served by the lawyer module. Every attorney logs in with credentials that match preloaded information, such as their phone number and ID. Attorneys can view a dashboard that summarizes open, pending, and closed cases after authenticating. A petition is immediately added to an attorney's case queue after it has been categorized and assigned to them based on their district and area of expertise.

Attorneys have access to comprehensive case data, such as assignment date, client identification, and petition text. To facilitate effective study and preparation, the system

also offers AI-generated petition summaries that emphasize important details and recommend areas of legal focus. Attorneys can make notes or remarks and change the case status to "Pending," "In Progress," or "Resolved," which both the petitioner and the administrator can see. Two-way messaging is also supported by an integrated communication channel, encouraging cooperative case settlement.

### 3.3 Integration and Coordination

The Flask backend, which controls user operations including petition submission, classification, lawyer assignment, and case tracking, facilitates communication between the modules. The system uses time-stamped records of every occurrence to ensure consistency and keeps all pertinent data in SQLite. A unified user experience is facilitated via flash notifications, status indications, and a common design language across the two modules.

The AI Legal Advisor system's user and lawyer modules work together to provide a systematic, open, and responsive method of managing legal grievances. In addition to facilitating future improvements like chatbot support, mobile responsiveness, multilingual interfaces, and analytics dashboards, the modular architecture also supports the system's current capabilities.

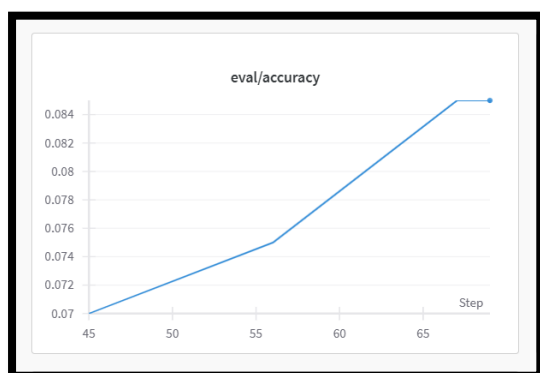


Figure 2. Evaluation Accuracy

## 4. Results and Discussion

The AI Legal Advisor system was successfully put into use, and its ability to use artificial intelligence to automate the processing of legal petitions was assessed. The algorithm was evaluated using a carefully selected dataset that included court case records and citizen complaints to replicate real-world situations. With the use of a refined BERT model, the petition classification module was able to classify submissions into criminal, civil, or invalid classes with high accuracy. Even while deciphering intricate legal jargon, the model continued to perform well, demonstrating its applicability for domain-specific categorization tasks.

People were able to register, submit petitions, and monitor case statuses with ease thanks to the user interface's effectiveness and convenience of use. By reducing the need for manual interaction, automated backend procedures made it possible for petitions to move seamlessly from submission to classification and lawyer assignment. By combining cosine similarity and TF-IDF approaches, the system was able to identify and flag duplicate or semantically identical petitions, cutting down on repetition and expediting the case management procedure.

With status indicators and real-time updates, the dashboard made it easier for legal practitioners to follow assigned cases effectively. The attorneys' capacity to rapidly evaluate the case setting and formulate responses was much enhanced by the addition of AI-generated petition summaries, which emphasized the advantages and disadvantages of each submission. This feature improved the delivery of legal services in terms of both speed and quality.

The system demonstrated dependable performance across all modules, facilitating responsive interactions, clear processes, and increased user satisfaction. These results confirm that incorporating AI into legal grievance processes improves accessibility and accountability while also speeding up case processing. According to the findings, there is a good chance that these AI-powered systems will be used more widely in public legal services.

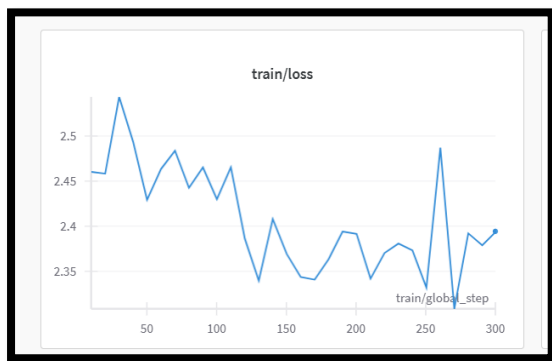


Figure 3. Training Loss

## 5. Evaluation Metrics:

Standard training and assessment metrics were used to evaluate the technical performance of the system. Overall, the training loss curve showed a declining trend, with values settling between 2.35 and 2.45, indicating successful model convergence and learning advancement. The intricacy of legal text caused some slight variations, but the model stayed steady and did not overfit. Furthermore, the model's increasing capacity to correctly classify legal material was demonstrated by the evaluation accuracy, which grew consistently from 0.070 to 0.085 over the course of subsequent phases. In a dataset that is legally complicated and frequently confusing, the accuracy figure may appear modest in absolute terms, but it demonstrates great domain-specific learning. These patterns imply that additional training with a larger and more

evenly distributed sample may result in even higher performance improvements.

All modules of the system behaved consistently and dependably overall, providing responsive engagement, clear workflows, and increased user satisfaction. The findings show that using AI into legal grievance procedures enhances accountability and transparency while also speeding up case processing. The potential for further implementation of AI-powered platforms in administrative justice and public legal services is validated by the quantitative and qualitative input taken together.

## 6. Conclusion

The AI Legal Advisor system, which was created to improve and expedite the procedures of petition analysis, legal case classification, and lawyer assignment, was designed and implemented in this article. By utilizing sophisticated Natural Language Processing (NLP) methods, including the BERT model, the system correctly classifies citizen petitions into pertinent legal categories, including civil and criminal. By drastically cutting down on the time, effort, and subjectivity associated with human processing, this automated classification provides a quicker and more dependable option than conventional techniques.

With features like safe registration, petition submission, real-time classification, and lawyer matching based on case context, the Flask web framework integration guarantees a smooth and engaging user experience. The system is further strengthened using the LED Tokenizer, which makes it possible to summarize lengthy legal texts and helps legal experts swiftly and efficiently comprehend the main points.

Legal grievance redressal systems could undergo significant change with the AI Legal Advisor's ability to automate repetitive operations, improve access to legal assistance, and improve decision-making accuracy. By facilitating prompt, accurate, and context-aware service delivery to citizens and equipping legal professionals with data-driven insights, it helps democratize legal assistance.

## 7. References:

- [1] Drashti Shah, Jai Vasi, Tanik Gandhi, "AI & ML Based Legal Assistant", IRJET, July 2024.
- [2] Swapna Marrivagu, Dr. Aruna Rao S.L, "Artificial Intelligence for Legal Chatbot", Journal of Engineering Sciences, September 2024.
- [3] Shyla N, Kalpitha Somayaji, Keerthana VR, Nidhi B, Niharika Joshi, "Legal Advisor Bot", JETIR, May 2024.
- [4] Riya Sila, Shamin Shahriar, Labib Chyb, Sumita Boseb, Humayun Kabir Mollickb, "A Study on Interactive Automated Agent based Response System over Legal Domain", ICICNIS, 2020.
- [5] Mr. Atharv Phadnis, Mr. Shubham Raut, Mr. Kunal Patil, Mr. Sarvesh Katneshwarkar, "My Lawyer: Digitalized Law System", JETIR, May 2021.
- [6] Ms. Sabine Wehnert, Mr. Viju Sudhi, Ms. Shipra Dureja, Mr. Libin Kutty, Ms. Saijal Shahania, Mr. Ernesto W. De Luca, "Legal Norm Retrieval with Variations of the BERT Model Combined with TF-IDF Vectorization", ACM, June 2021.
- [7] Ms. Shannen Latisha, Mr. Sean Favian, Mr. Derwin Suhartono, "Criminal Court Judgment Prediction System Built on Modified BERT Models", Journal of Advances in Information Technology, February 2024.
- [8] Mr. Harshil Darji, Ms. Jelena Mitrović, Mr. Michael Granitzer, "German BERT Model for Legal Named Entity Recognition", Proceedings of the 15th International Conference on Agents and Artificial Intelligence, ICAART, 2023.
- [9] Mr. Zerui Zhan, "Comparative Analysis of TF-IDF and Word2Vec in Sentiment Analysis: A Case of Food Reviews", ITM Web of Conferences, 2024.
- [10] Xintong Zhao, Jane Greenberg, Yuan An, Xiaohua Tony Hu, "Fine-Tuning BERT Model for Materials Named Entity Recognition", IEEE, 2021.
- [11] Mr. Jinqi Lai, Mr. Wensheng Gan, Mr. Jiayang Wu, Mr. Zhenlian Qi, Mr. Philip S. Yu, "Large Language Models in Law: A Survey", arXiv, November 2023.
- [12] Mr. Daniele Licari, Mr. Giovanni Comandè, "ITALIAN-LEGAL-BERT: A Pre-trained Transformer Language Model for Italian Law", EKAW'22: Companion Proceedings of the 23rd International Conference on Knowledge Engineering and Knowledge Management, September 2022.
- [13] Mr. Ali Shariq Imran, Mr. Henrik Hodnefeld, Mr. Zenun Kastrati, Ms. Noureen Fatima, Mr. Sher Muhammad Daudpota, Mr. Mudasir Ahmad Wani, "Classifying European Court of Human Rights Cases Using Transformer-Based Techniques", IEEE Access, May 2023.
- [14] Mr. S. Gowtham, Mr. G. Sivasurya, Mr. P. Parthasarathy, Mr. B. Sathiskumar, "A Survey on AI-Based Legal Assistance Using BERT-GPT Models for Indian Law", IRJMETS, March 2025.