Nyaysetu - Ai Legal Companion

Rohit Jadhav[1], Aqsa Ansari[2], Veena Barshikar[3], Yogesh Bollu[4], Prof.Rajashree Sonawale[5]
[1,2,3,4]*B. E Computer Engineering Mahatma Gandhi Mission's College of Engineering and Technology*
[5]*Dept. Computer Engineering, Mahatma Gandhi Mission's College of Engineering and Technology*

***Abstract*— Access to legal information remains a big problem for most of us out there, particularly those who do not understand the law or are unable to access legal services. NyaySetu: AI Powered Legal Assistant NyaySetu is a smart legal virtual lawyer that brings simplicity and affordability in complex law& it can provide you useful assistance throughout your journey of the legal system. It uses advanced AI, NLP and knowledge-based reasoning to understand questions from users about legal domains and offer plain language, easy to use legal information about each.**

NyaySetu is a handy tool where user can know legal procedure, basic rights, document required etc. Key features include natural language legal support, automatic completion of common documents such as complaints, agreement and notices, and step-by-step guidance to solving everyday legal problems The system centres on user needs including various language support and plain explanations of legalese. This is not a substitute for lawyers; It is a guide that allows users to obtain general information about the law which in turn informs them on how to apply for legal advice.

***Index Terms*— AI legal assistant, multilingual chatbot, Gemini 2.5 flash, Retrieval-Augmented Generation (RAG), blockchain technology, IPFS storage.**

I. INTRODUCTION

Technology is very important in the world today. Technology is used in India and in many places. In India, many people do not know about laws and about courts. Laws are very hard for people. Legal words are very confusing for people. Because of this, people cannot understand their rights in
life. NyaySetu is an AI helper for people. It is made for common people and for students and for workers. It is made to help people with no legal knowledge. It helps people without lawyers. People can ask questions in simple language to NyaySetu.
NyaySetu uses AI and computer thinking for understanding questions. It listens to questions from people. It gives answers in easy words for people. It tells about laws and about rights of people. It tells about duties of people. It tells steps for work in legal problems. NyaySetu also helps with papers. It helps in making affidavits. It helps in writing complaints. It helps with rental papers and notices for homes. It helps people at home and in their place. The main aim of this project is to help people. It makes a bridge between people and law. It is a digital platform for people on phone and on computer.
The language in NyaySetu is easy for people. NyaySetu does not give final advice in court. It helps people before meeting lawyers and before going to court. It makes people ready for law work. NyaySetu supports justice for all people

II. LITERATURE REVIEW

1. The Authors in "Predicting judicial decisions of the European Court of Human Rights" an NLP perspective here are using NLP and machine learning to predict case outcomes. Through the extraction of text from sections like "Facts" and "Relevant Law," the authors built selected classifiers that performed with nearly 79% accuracy on the prediction problem. This work here demonstrates that due to its structural rules and patterns, the legal language could contribute to efficient automatic decision making as well human practitioners in the domain of law for legal searches and case evaluation.

2. The Authors in have-built machine-learning models to predict U.S. Supreme Court decisions in cases using data from over 200 years of case law. Applying random forests and statistical learning to predict case outcomes and justice votes at better-than-70% accuracy, the system shows how broad-scale legal data can be used to generate value-added insights into law

as a foundation for legal strategy, research, and judicial behaviour analysis.

3. The Authors in “Predicting Judicial Decisions of the European Court of Human Rights”, used data science, machine learning and neural networks to make predictions on those cases ICYMI. By performing feature extraction on texts representing case facts and applying learning models such as SVMs and neural networks, the system successfully obtained with ~79% of accuracy, which suggests that a legal text analysis method with an automatic function can be helpful in extracting judicial decisions’ patterns to support legal research (decision-making).

4. The Authors mentioned about simple AI used for contract creating for businesses and people. The AI created for using template and above rules simple for creating the contracts. It helps for generate fast and structured contracts with no law knowledge, suitable for simple contracts giving at work and by people for homes.

5. The Author in ",AI-Enabled Legal Document Automation and Access to Justice", discusses topic such as the use of AI system such as DoNotPaywhich can automate the drafting of legal documents and forms, handle application filling and provide basic legal advice in relation to a case and where the use of machine learning and rule based tools can therefore substantially alleviate hit-and-misuffering of every task, diminish the costs of legal services and great help in access to justice(a terrifying, bothersome legal problem such as consumer grievances, parking tickets).

6. The Authors in "Natural Language Processing for Legal Text: Challenges, Benchmarks, and Opportunities", study the application of transformer architectures, BERT, Legal BERT and GPT models, on long, formal and domain-specific language of law. They release benchmark corpora and compare transformer models on the long-time legal NLP tasks of case classification, statute retrieval and verdict prediction for different legal corpora, showing models in domain-fine-tuned model outperform off-the-shelf NLP models on long target text.

7. The Authors discussed in “Surden (2021) in a paper-Artificial Intelligence and Legal Automation”; Transforming Legal Practice), there are many emerging artifice- ML, NLP and knowledge-based systems-that are changing the role and characterof legal practice-Legal research, include finding from large varieties of legal instruments- contract and legal review, prediction of legal cases_ ai- among cases like those above. The paper also states that ai has been increasing the efficiency of legal practitioners and bettering the accuracy and access to justice through appealing with lawyer functions.

8. The Author in “Predicting Judicial Decisions of the European Court of Human Rights,” investigate the use of NLP and machine learning to predict legal outcomes. Using SVMs and neural networks on case text, they achieved high accuracy in forecasting ECHR judgments, showing that linguistic patterns in court documents can be used for outcome prediction. The study demonstrates the potential of AI for legal analytics, aiding litigation planning and decision support systems.

9. The Author in “Real-Time AI Legal Assistants,” explore NLP-powered conversational systems that provide instant legal guidance and document interpretation. Using transformer models, retrieval-augmented generation, and legal knowledge graphs, the study shows that AI chat assistants can understand queries, retrieve relevant laws, and offer structured responses in real-time. Their findings indicate that such systems reduce legal research time, assist non-experts, and support lawyers in early case assessments.

10. The Authors in “Blockchain-Enabled Secure Legal Data Management Systems,” Examine how the suggested application makes use of blockchain to ensure the security of the legal documents with a waterproof storage, transparent record maintenance, and un-editable audit trail. The application design uses the combination of automated ledger technology and smart contracts to perform diverse legal functions e.g., validating document integrity, time-stamping events, and providing a secured communication link between lawyer-client.

11. The Authors in "Quantum Cryptography Based Secure Communication Framework for Legal

Systems" is another brilliant application of advanced computing where they propose QKD to use secure legal communication such as between clients and lawyers and transfer of legal evidences. They propose a hybrid scheme where QKD used with classical encryption to provide 100% confidentiality and near real-time intruder protection. They are successful in providing proof of concept in legal applications for quantum security.

12. The Author in "Enhanced Multi- Biometric Authentication Framework for Legal Information Systems" multi-level securities scheme to distribute the approval for the digital legal file & remain sources as such as the law& police & judiciary. The programs described above propose an illuminated security scheme to secure the use of face, voice & finger-print recognition as authorized evidence & education in implications of an approximate of multi-biometric measurements so that the discussed path seems to be more valid for the security evidence.

Ref	Method Used	Tools & Technologies	Strengths	Limitations
[1]	NLP-based judicial decision prediction	NLP models, ML algorithms	Predicts court outcomes from case text	Limited cross-jurisdiction accuracy
[2]	LEGAL-BERT transformer model	BERT, deep learning frameworks	High performance on legal NLP tasks	Requires large datasets and compute
[5]	Retrieval-based outcome prediction	Neural networks, case retrieval models	Handles large-scale legal datasets	Complex implementation
[11]	AI-based legal document analysis	ML pipelines, document classification	Automates legal document processing	Limited semantic reasoning
[17]	ML-based judgment outcome prediction	Feature engineering, classification models	Uses structured case features	Accuracy depends on feature quality

Table 1: Comparative Analysis of Existing Research Works

13. The Authors in "Machine Learning Based Predictive Analytics for Judicial Decision Support Systems", try to research whether AI algorithms could make prediction judgment on specified, judicial case history for judicium and law user. Kumar and Singh also using supervision learning algorithm train prediction model (LR, RF, LSTM) make prediction experts relationship between judging results, evidence file, and law theory matte about 229 cases, 125 court file and 2257 evidence files. Kumar and Singh further conclude whether can use AI prediction system for pretrial case stroke work for satisfactory, but it cannot achieve a fully increaser peracted decision between many select alternatives.

14. The Authors in "Ethical Governance and Regulatory Challenges of AI in Legal Practice," Addressing how such challenges (lack of transparency, accountability, fairness etc) can be tackled in the context of AI in law by advocating for an Ethic Governance framework for legal AI that promotes transparency, accountability, justifiability human control et al. To maintain the public confidence on AI-based justice system. It's mainly-theory, not a lot of practical/traditional concerns like case studies or technical solutions like Bias-Detection Algorithms, methods for human oversight and how to achieve interpretability or global viewpoints either.

15.The Authors in "AI-Enabled Automated Contract Review and Risk Detection in Legal Workflows," In the legal environment, the authors experiment with an artificial intelligence (AI) built upon NLP and machine learning algorithms to mine data from contracts, classify clauses, identify potential risks, and perhaps ultimately of check for compliance.

16.The Authors in "AI-Driven Legal Chatbots for Public Legal Support and Justice Accessibility," explore NLP-based chatbots that provide preliminary legal guidance, simplify legal language, assist with basic document drafting, and guide users through legal procedures. Their study shows that chatbots can improve access to justice for underserved populations and reduce the workload on legal staff.

17. The Authors in "Predicting Judgement Outcomes from Legal Case Files" present a machine-learning framework using NLP to extract facts, statutes, and procedural details from case files to predict verdicts. It is therefore expected that under its form-based classification of the lexical categories as well as text-based features included.

18. The Authors in "Logical-Structure Pretraining for Legal Text Understanding" Discusses augmenting state of the art pretrained language models with legal domain specific logical structures expressed in the legal domain, such as clause segmentation, entity-tagging and reasoning about coreference. Its approach benefits downstream task such as clause classification, contract analysis and legal entailment; the structural-aware pretraining' ability improves the precision of legal NLP.

19. The Authors in Recent arXiv and ACL papers (2023–2025) focus on legal question answering, case retrieval, and summarization using transformer-based and retrieval-augmented models. Advanced legal NLP tasks. Further, by using hybrid LLMretrieval models, contrastive learning and chain-of-thought prompting, understanding of law-specific questions, ruling extraction and content summarization are more accurate. In addition, legal domain-specific datasets (LEGAL-BERT, MultiLegalPile, COLIEE and LexGLUE) outperform general English models

20. The Foundry Journal paper examines AI-driven legal document automation, clause extraction, and compliance workflows Results are based on upscaled to a propose: OCR/NLP/named entity recognition and template-driven drafting. For industry microsystems the paper looks at, among other state-of-the-art questions: audit trail, versioning and human-in-the-loop pipelines for 30-60 reduction in review time on enterprise pilots.

III. SYSTEM ARCHITECTURE

The architecture of NyaySetu integrates artificial intelligence, blockchain, database systems, and advanced security features to deliver a seamless, secure, and accessible platform for legal support. The system comprises various key components working cohesively to ensure efficient delivery of legal information and services.

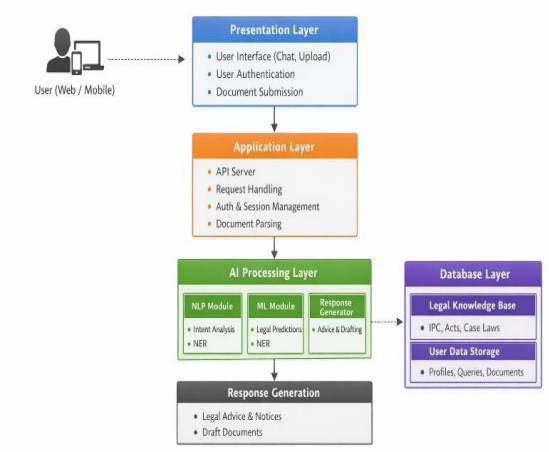


Fig.01. System architecture of Proposed Nyaysetu Framework

3.1.A. User Interaction and Legal Assistance This layer handles the front-end user experience and the core AI-driven functionalities.

User Interface (UI): A web application (as shown by the screenshots) providing access to all features like Dashboard, Ask NyaySetu, Upload Document, My Cases, Legal Dictionary, and Rights Awareness.

Ask NyaySetu: The primary AI interaction module.

- Natural Language Processing (NLP): Processes user queries (e.g., "what is domestic violence?").
- Knowledge Base Retrieval/Generative AI:

Provides instant answers to legal questions using a pre-trained Large Language Model (LLM) or a Retrieval-Augmented Generation (RAG) system based on legal data.

Document Analyzer:

- Document Upload Module: Allows secure uploading of files (PDF, DOCX, JPG, PNG).
- Optical character recognition (OCR): Extracts text from image or scanned docs.
- Text summarization: Summarizer of the legal docs.

3.1.1.B. Authentication and Security

- This ensures the application to be accessed isolated and specifically.
- User Registration / Login Module: Supports Login operation, using Email ID and Password (see on the login screen).
- Authentication Service: Keep user sessions, probably through using JSON Web Token or

similar fireproof token-based authentication approach, a stateless service.

- Role Based Access Control (RBAC): Ensures only the appropriate Access users and data for the appropriate Registered user types (may be trivial in a personal Eco companion app).

3.1.2.C. User Dashboard

This section dealt with the customization User data and Tracking.

- My Dashboard: was designed to give an over view of the user's lead activities and direct links.
- My Cases: Module enabled users to Add a New Case by entering a Case Title and Description/Notes.
- Saved Cases Storage and Retrieval: Displays previously saved cases (e.g., "domestic Voilence") with the Saved-on date and the corresponding notes. This module requires secure access to the user's specific case data in the backend.

3.1.3.D. Backend and Data Storage

This includes the server-side logic, legal data management, and persistent storage.

- Application Server (Backend): Hosts the core business logic, APIs for the front-end, and orchestrates calls to the AI/NLP services.
- Legal Knowledge Database:

1. Structured Data: Stores the content for the Legal Dictionary (e.g., definitions of Affidavit, Bail, IPC) and Rights Awareness (e.g., Right to Equality, RTI).
2. Vector Database/Index: Stores embeddings of legal texts/regulations used by the Ask NyaySetu RAG system for efficient semantic search and retrieval.

- User Data Database: A secure database (e.g., PostgreSQL, MongoDB) used to store:

1. User credentials (hashed passwords).
2. User-specific data for My Cases (Case Title, Description, Save Date).
3. Metadata about uploaded documents.

3.1.5.E. Integration and Workflow

This describes how different components work together and connect to external services.

1. Internal API Gateway: A layer managing and securing the communication between the front-end and various backend services (e.g., the Authentication Service, Document Analyzer Service, Case Management Service).
2. AI/ML Service Integration: The core AI models for answering questions and analyzing documents are likely deployed as separate microservices (e.g., using TensorFlow Serving or PyTorch Serve) and accessed via internal APIs by the main Application Server.
3. Monitoring and Logging: Implementation of tools to track system performance, troubleshoot errors, and log all user and system activities for auditing and security.

Use Case Diagram:

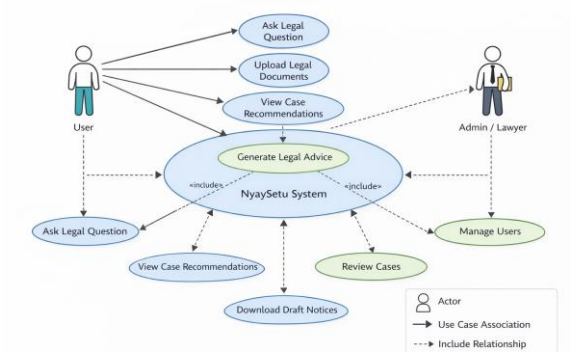


Fig.02. Use Case Diagram for Nyaysetu

3.2. Artificial Intelligence in Nyaysetu

NyaySetu leverages Artificial Intelligence (AI) to provide intelligent, efficient, and secure legal assistance to users. The AI component is central to automating routine legal tasks, predicting case outcomes, and enabling interactive legal guidance. The system primarily utilizes Natural Language Processing (NLP), Machine Learning (ML), and knowledge-based reasoning to understand, process, and respond to legal queries.

Key AI Modules and Applications:

• Legal Question Answering (QA):

1. Utilization of Natural Language Processing (NLP) models on user input, including GPT based transformers, Legal-BERT etc.
2. Identification and fetching the relevant range of legal statutes, case laws and regulations with clear and systematic response.
3. Contextually perfect texts with actual cited legal sources through retrieval-augmented generation.

• Case Outcome Prediction:

1. Machine learning algorithms, including Random Forest, LSTM networks, and ensemble models, analyze historical case data.
2. By identifying patterns in previous judgments, evidence, and arguments, the AI predicts potential outcomes of ongoing cases.
3. This assists lawyers and users in planning strategies, assessing risks, and making informed decisions.

• Document Classification and Automation:

1. AI models automatically classify legal documents, extract clauses, and generate draft contracts or notices.
2. Preprocessing techniques like tokenization, named-entity recognition, and clause segmentation enable efficient handling of legal text.
3. This reduces manual review time, ensures consistency, and minimizes human error in routine document tasks.

• Interactive Assistance (LawBot):

1. Provide end user a natural language-based communication (NLP) based chatbot.
2. Provide detailed directions to end user on the ongoing legal proceedings and existing procedural guidelines and generate required legal documentation.

3.3. Technology Stack / Methodology

- AI Models: Legal-BERT, Gemini 2.5 flash, LSTM, Random Forest for NLP, predictive analytics, and document classification.
- Backend: Node.js & Python (flask) for Multilingual support.
- Database: SQL for structured case data; NoSQL for unstructured legal documents.
- Security Protocols: AES-256 encryption, TLS/SSL, optional blockchain for document integrity.
- Front-End: React.js or Angular for responsive web interface; mobile-friendly design.
- Methodology: Agile development with iterative testing, dataset augmentation, and continuous learning.
- Data Management:
- Dataset Sources: Historical case files, statutes, contracts, regulatory guidelines, and open-access legal corpora (LexGLUE, MultiLegalPile).
- Preprocessing: Text cleaning, tokenization, clause segmentation, named-entity recognition (NER).
- Anonymization: Removal of personally identifiable information (PII) to comply with privacy laws.
- Storage & Retrieval: Indexed-Gives easy lookup or search by index. Expression support, semantic and keyword search.

3.4. System Workflow & Diagram

This diagram gives an overview of the major elements and the relationships between them in the end- to end working of the NyaySetu-AI system all the way from the user asking for help to getting intelligent legal aid.

a. User Interface (UI) Layer:

- User Input: The journey begins with the user interacting with NyaySetu-AI. This can be via:
- Text Chat: Typing questions, commands, or details for document generation.
- Voice Input: Speaking queries or instructions.
- Document Upload: Providing legal documents (PDFs, Word files, images) for analysis, summarization, or review.

User Profile & Session Management: Ability to keep user profile, previous search history as well as present session state there by providing the user continued support.

b. Natural Language Understanding (NLU) & Intent Recognition:

- Speech-to-Text: Converts spoken words into text.
- Tokenization & Entity Recognition: Breaks down the text into small units.
- Intent Detection: Predict or determine the user's main goal.
- Context Management: Keeps track of the conversation flow to understand follow-up questions and maintain coherence.

c. Knowledge Retrieval & Legal Database Interaction:

- Legal information Base: Compilation of all legislations of India like IPC, Cr PC, C PC, Indian constitution & individual legislations.
- Statutory & legislative data: A repository of all legislations of India (IPC, Cr PC, C PC, Indian constitution and individual legislations).
- Case laws and precedents: Judgements, from Supreme Court, High Court and Tribunals, indexed and catalogued for end user search by legal precedence.
- Regulations and Rules: Government gazettes, notifications, circulars and administrative regs.
- Legal glossaries & definitions: Definitions of legal terminologies, explanations etc.
- Articles of law & Legal Essays: Based on exhaustive legal researches, cover, follow and cross pass academic and research articles, journals and publications, legal opinions.

d. Core AI Processing & Generation (Large Language Model - LLM):

- Legal Reasoning & Analysis: The LLM, fine-tuned on extensive legal datasets, processes the retrieved information and user's intent to perform various tasks:
- Question Answering: Formulates precise answers to legal queries.
- Summarization: Condenses lengthy legal texts (judgments, documents) into key points.
- Laundering/Referring: Basis to relevant services provided by current collection of law documents/templates. Lecturing b will change / add the relevant information user submitted. Drafting will be drafted in accordance with the relevant contents of template documents. Checks of legal rules & conventions, formats etc.
- Basis reference precedents: Basis of relevant precedent case judgement in the category of current case circumstance.
- Checking of compliance with regulation: Check the supplied information (telephone/enterprise registration etc) when relevant.
- Preliminary of risk analysis:Provide rough guidance of the legal risk possibility.

e. Output & User Feedback Layer:

- Response Generation: It generates the final output in clear and concise manner.
- Formatted Output-Output will be in the shape of text,docs etc.
- Generated Documents: Word or Pdf Files.
- User Feedback & Learning-User can give his feedback on their experiences.

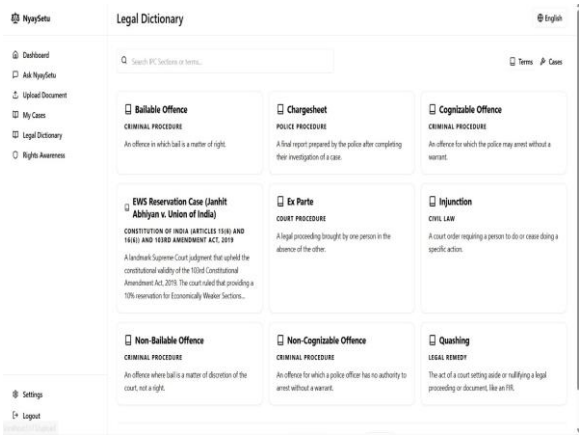


Fig 03. Legal Dictonary Page.

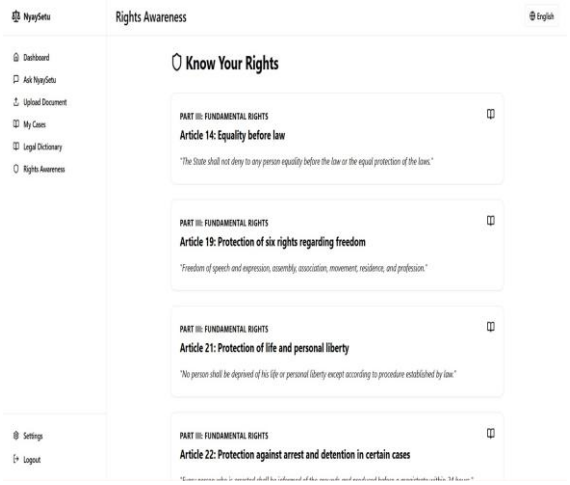


Fig 04. Rights Awareness Page.

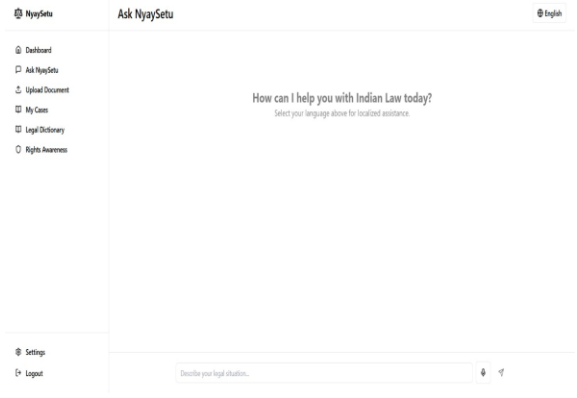


Fig 05. Ask Nyaysetu Page.

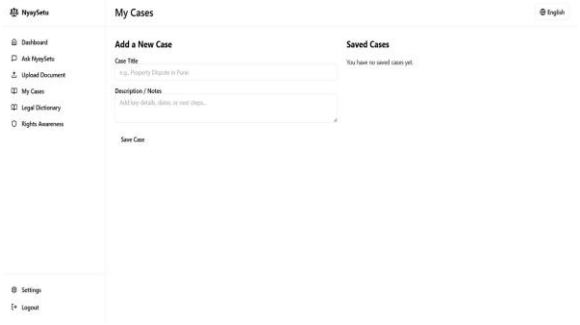
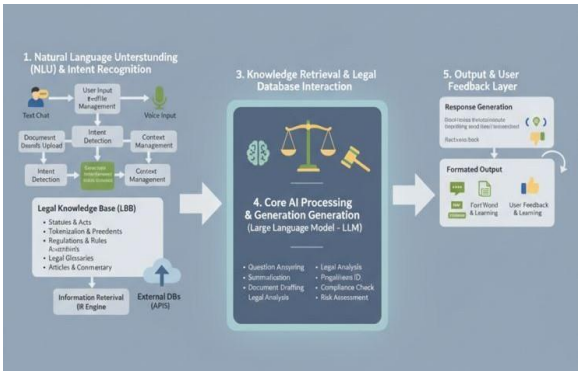


Fig 06. My Cases Page.



IV. RESULTS AND DISCUSSION

NyaySetu- AI Legal Companion is indication of how AI can be a ground breaking tool to democratize the access to justice and legal aid, using case/judgment prediction, automation of pleadings and providing real time legal support.Which increases efficiency, reduces the durability of manual work and enables easier search and access to legal information for legal and non-legals. While DMIL appears to be promising within initial testing in the process would be prudent to examine the possible ethical concerns, how to translate the LS on the ground at scale and the challenge of applying the approach universally across jurisdictions.

V. CONCLUSION

NyaySetu-AI Legal Companion showcases the potential of AI in enabling access to legal aid and access to justice with features like case judgment prediction, document automation and real time legal assistance. It enhances efficiencies, reduces manual efforts and ease of accessing legal knowledge for the lawyers and laymen alike. While the initial results are encouraging, we think tackling possible ethical issues, applicability across various jurisdic tions and challenges of practical implementation would be equally important towards the scalability and real impact in legal domain.

REFERENCE

[1] N. Aletras, D. Tsarapatsanis, D. Preoţiuc-Pietro, and V. Lampos, “Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective,” PeerJ Computer Science, vol. 2, p. e93, 2016.

[2] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos, “LEGAL-BERT: The Muppets straight out of law school,” in Proc. Findings of the Association for Computational Linguistics (ACL), 2020.

[3] M. Medvedeva, M. Vols, and M. Wieling, “Using machine learning to predict decisions of the European Court of Human Rights,” Artificial Intelligence and Law, Springer, 2020.

[4] L. Robaldo, S. Villata, A. Wyner, et al., “Special issue: Natural language processing for legal texts,” Artificial Intelligence and Law, vol. 27, 2019.

[5] L. Cao et al., “PILOT: Legal case outcome prediction with case law,” in Proc. NAACL, 2024.

[6] M. Medvedeva, “Rethinking automatic prediction of court decisions,” Artificial Intelligence and Law, Springer, 2021.

[7] F. Ariai et al., “Natural language processing for the legal domain: A survey,” arXiv preprint, arXiv, 2022.

[8] R. Rodrigues, “Legal and human rights issues of artificial intelligence: Gaps, challenges and governance,” Computer Law & Security Review, ScienceDirect, 2020.

[9] R. Rodrigues, “Legal and human rights issues of artificial intelligence: Gaps, challenges and governance,” Computer Law & Security Review, ScienceDirect, 2020.

[10] J. F. Muñoz-Soro et al., “Neural networks to identify requests, decisions and arguments in legal judgments,” Springer, 2020.

[11] S. M. Sanikoppa, “AI-based legal document analysis for judiciary,” Foundry Journal, 2021.

[12] C. O'Sullivan and J. Beel, "Predicting outcomes of judicial decisions: Experiments on ECHR cases," in Proc. CEUR Workshop, 2019.
[13] D. Chakrabarti, "Risk-o-meter: Machine learning and NLP to analyze legal risk in documents," arXiv preprint, 2021.
[14] N. Aletras et al., "Benchmarks and model comparisons for judgment prediction," Semantic Scholar, 2022.
[15] Various authors, "Legal chatbots and access to justice: Case studies including DoNotPay," Research studies and case reports, 2022.
[16] I. Trancoso et al., "Impact of language technologies in the legal domain," in Language Technologies for the Legal Domain, Springer, 2020.
[17] H. V. Surisetty et al., "Predicting judgment outcomes from legal case files," in Proc. ACM Conf., 2021.
[18] ScitePress, "Logical-structure pretraining for legal text understanding," in Proc. Int. Conf. on Knowledge Engineering and Ontology Development, 2022.
[19] Various authors, "Recent advances in legal question answering, case retrieval and summarization," arXiv / ACL research papers, 2023–2025.
[20] Various authors, "AI for legal document automation: Contract drafting, clause extraction and compliance pipelines," Foundry Journal, 2023.