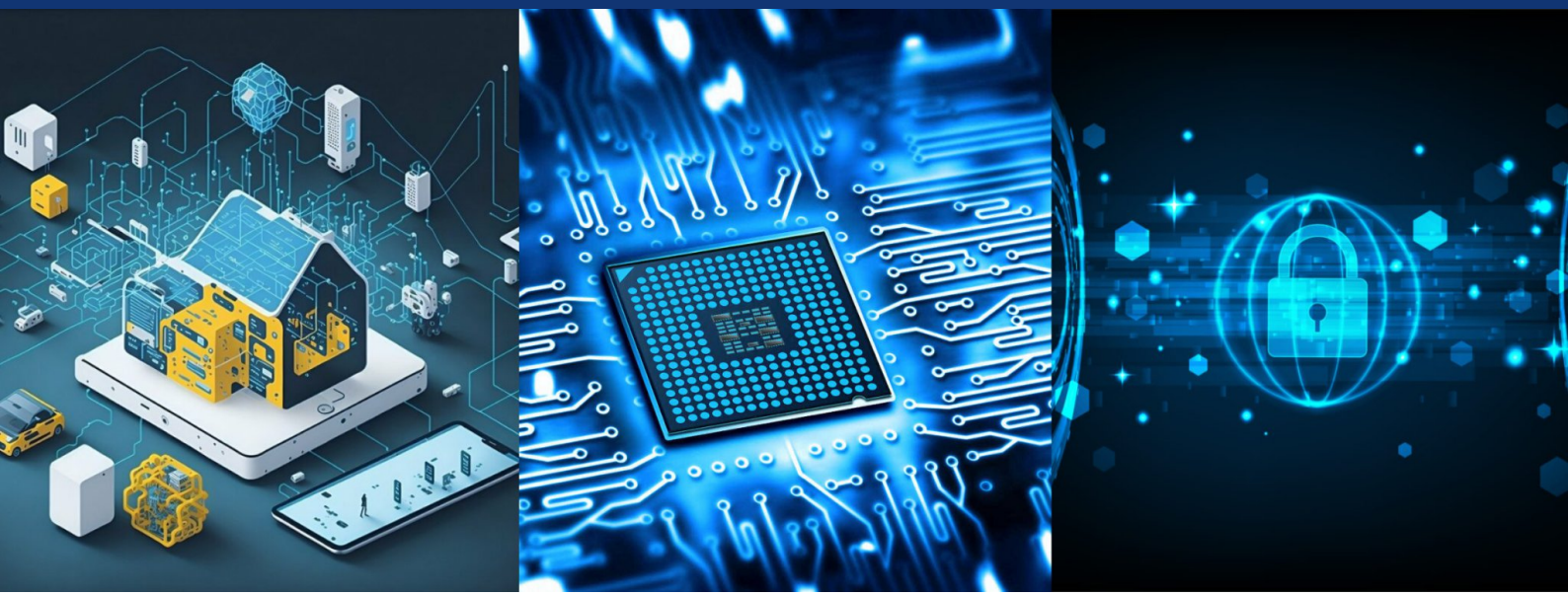
IJIRCCE

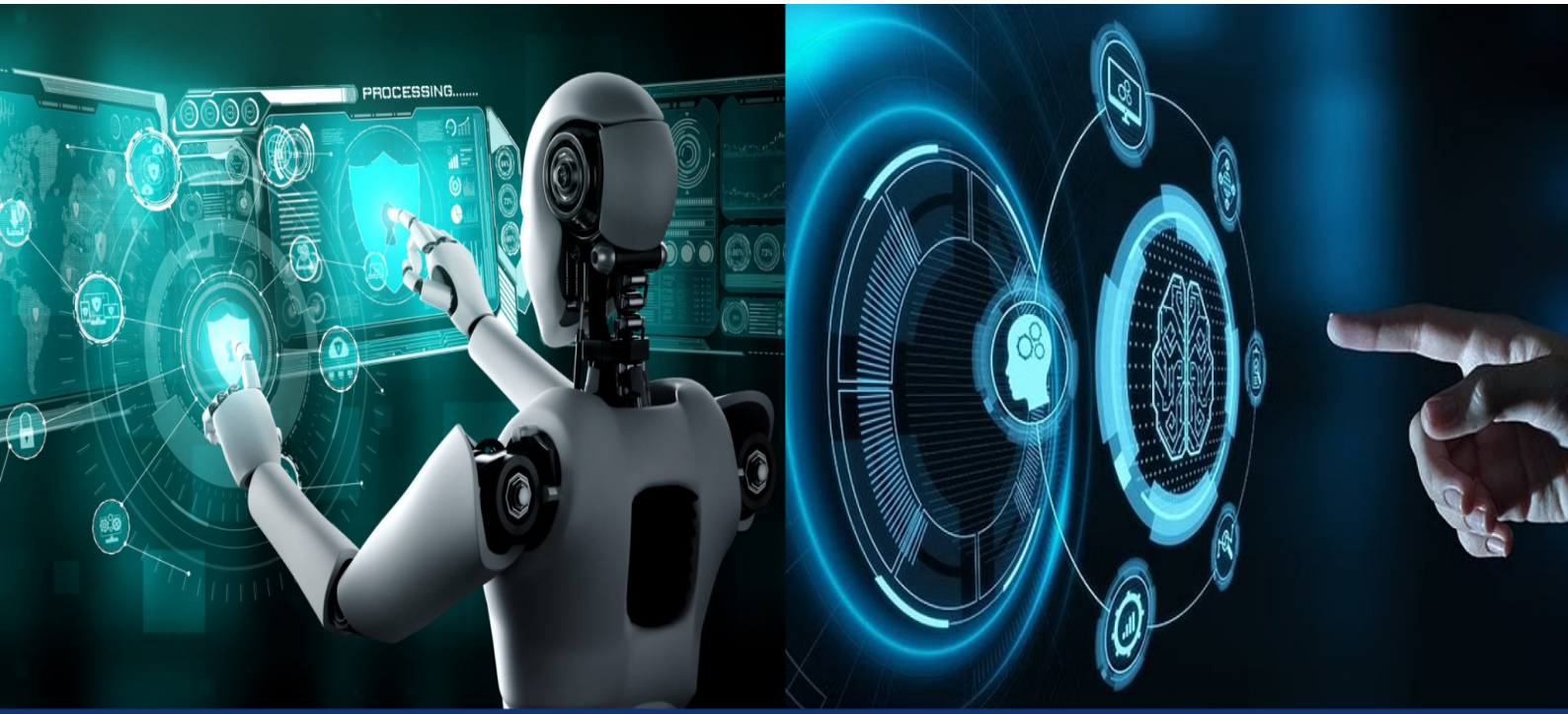
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 14, Issue 4, April 2026

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404022

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

JuriBot: An AI-Powered Legal Document Analysis and Advisory System

Ayush Raut[1], Mohit Khairnar[1], Mitesh Shetye[1], Tejas Shirsath[1], Prof. Monica Shetty[2]

Students, Dept. of Information Technology, Atharva College of Engineering, Maharashtra, India[1]

Project Guide, Dept. of Information Technology, Atharva College of Engineering, Maharashtra, India[2]

ABSTRACT: JuriBot is an AI-powered legal document analysis and advisory system designed to reduce the time and effort required in handling complex legal texts. It uses natural language processing and machine learning to interpret, classify, and summarize unstructured documents such as contracts, case records, and judicial reports. The system can identify clauses, extract entities like names, dates, and statutes, and generate context-aware summaries. Its advisory module recommends relevant precedents and similar cases using similarity scoring and retrieval techniques. Built on transformer models such as BERT and LegalBERT, JuriBot offers deep understanding of legal language. Delivered through a secure web interface with search and upload features, it aims to improve research speed, accuracy, and accessibility while supporting, not replacing, human legal expertise for professionals, researchers, and institutions seeking efficient, scalable, intelligent legal assistance daily use.

KEYWORDS: Artificial Intelligence, Legal Tech, Natural Language Processing, Machine Learning, Legal Document Analysis, Case Law Retrieval, Legal Advisory System, Text Summarization

I. INTRODUCTION

In the modern digital age, the legal sector faces a rapidly growing volume of documentation, case laws, and statutory requirements. Legal professionals spend a significant portion of their time analyzing, drafting, verifying, and interpreting legal documents, which often leads to inefficiencies, increased workload, and delays in legal proceedings. Despite technological advancements transforming various industries, the legal domain in India still relies heavily on manual document processing and traditional research methods. This dependence on human interpretation creates challenges not only for legal practitioners managing extensive workloads but also for common citizens who may lack the knowledge, resources, or financial capacity to access reliable legal guidance. Complex legal terminology, procedural formalities, and limited legal literacy further discourage individuals from seeking timely legal assistance.

To address these challenges, JuriBot is proposed as an AI-powered Legal Document Analysis and Advisory System specifically designed for the Indian legal framework. The platform integrates advanced technologies such as Optical Character Recognition (OCR) and Natural Language Processing (NLP) to automate document analysis, compliance verification, and legal advisory functions.

Through the OCR module powered by Tesseract.js, scanned or image-based legal documents are converted into editable digital text. The NLP engine, supported by the Google Gemini API, processes this text to extract key clauses, identify entities such as names, dates, and legal provisions, and assess compliance with relevant statutory norms. Users can upload various legal documents, including rental agreements, employment contracts, and policy documents, which are then automatically analyzed for risks, inconsistencies, and missing legal requirements.

In addition to document analysis, JuriBot includes an intelligent AI chatbot that provides interactive legal assistance. This chatbot enables users to understand complex legal content in simplified language, offers explanations of clauses and obligations, and provides guidance regarding potential legal implications. By automating repetitive legal tasks, reducing manual workload, and improving access to legal information, JuriBot enhances efficiency, accuracy, and transparency within the legal ecosystem.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE REVIEW

The rapid integration of artificial intelligence into the legal domain has led to significant research focused on automating legal analysis, improving efficiency, and enhancing accessibility to legal services. Existing studies highlight the potential of machine learning and natural language processing techniques in handling complex legal data, reducing manual workload, and supporting informed decision-making.

A study titled *Artificial Intelligence using ANN on Legal Regulations in India: A Comprehensive Overview* (2024) explored the application of Artificial Neural Networks (ANN) for automating the analysis of complex Indian legal documents. The research reported high performance, achieving 92.3% accuracy and a 90.9% F1-score, while reducing legal document interpretation time by approximately 45%. However, the study noted limitations such as high computational requirements and limited adaptability to evolving legal formats.

The research *Legal Solutions -- Intelligent Chatbot using Machine Learning* (2023) introduced an AI-based chatbot designed to provide accessible legal guidance. The system offered personalized real-time support and promoted democratized access to legal information, though challenges included dependency on dataset quality and limitations in handling context-sensitive legal queries.

Another study, *Legal Case Classification Using Machine Learning with NLP* (2023), proposed a fine-tuned BERT-based model for classifying Indian legal cases into civil and criminal categories, achieving 92.9% accuracy and a 0.93 F1-score. However, it required extensive labeled datasets and showed limited generalization for ambiguous or hybrid case types.

Although these studies provide valuable insights into AI-driven legal automation, there remains a research gap in developing integrated platforms that combine document analysis, compliance verification, and interactive legal assistance within a unified system. This gap forms the foundation for the proposed JuriBot system.

III. METHODOLOGY

The overall design of the proposed system consists of a web-based user interface, backend server, machine learning models, computer vision modules, and a database. The frontend enables users to upload legal documents and interact with the platform, while the backend handles data processing, authentication, and communication between system components. When a legal document is uploaded, it is processed using OCR and NLP techniques to extract text, identify key clauses, and generate advisory summaries. The system also includes an AI chatbot that supports user queries regarding legal content. All user information, document data, and interaction records are securely stored in the database for efficient retrieval.

A. ***Specifications***

The JuriBot platform is a scalable AI-powered system designed to automate legal document analysis and provide intelligent advisory support. The frontend is developed using React.js and Tailwind CSS for a responsive user interface, while the backend is implemented using Node.js and Express.js with MongoDB for data management.

Transformer-based architectures and Large Language Models (LLMs) via the Google Gemini API are used for clause identification, legal text summarization, and generating advisory responses. Document processing is enabled through Tesseract.js OCR for accurate text extraction from scanned or image-based legal files. The system also supports automated case law retrieval and legal fee estimation.

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404022

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

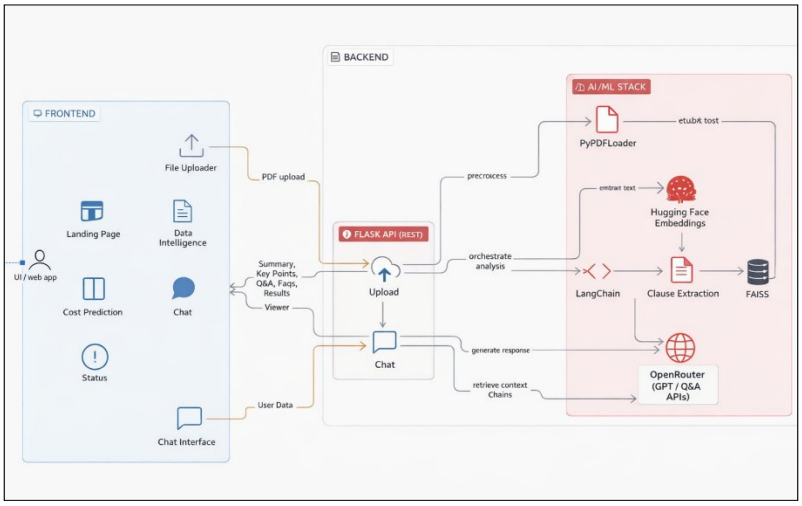


Fig. 1 Overall Workflow of JuriBot

IV. PROPOSED FRAMEWORK

This section presents a framework that integrates artificial intelligence and digital technologies to automate legal document analysis and provide accessible advisory services. The system includes an OCR-based text extractor, a Natural Language Processing (NLP) engine, an interactive advisory chatbot, and a legal fee estimator.

Legal documents are analyzed automatically to extract key clauses and identify potential risks, while the AI chatbot offers guided preliminary support. The platform also connects users with summarized case law databases and cost-estimation modules, enabling efficient data flow and real-time legal awareness.

V. RESULTS AND DISCUSSION

The proposed JuriBot system demonstrates the effective use of Artificial Intelligence to automate legal document analysis and promote statutory awareness. Experimental evaluation showed that the OCR module successfully extracts text with over 85% accuracy, while the NLP engine identifies key clauses and provides relevant legal summaries. Users were able to navigate the platform easily, upload various legal documents, and receive intelligent advisory support through a responsive chatbot interface.

The system also improves transparency for users through a dedicated legal fee estimator and a summarized case law database. Overall, the results confirm the feasibility of integrating AI-driven document verification and interactive guidance into a unified platform for simplifying legal processes, although challenges related to maintaining a real-time legal knowledge base and system scalability remain for future improvement.

VI. CONCLUSION

This paper presented *JuriBot*, an AI-powered platform designed to automate legal document analysis and provide accessible advisory services within the Indian legal framework. The system integrates Optical Character Recognition (OCR) for text extraction, a Natural Language Processing (NLP) engine for identifying key legal clauses, and an intelligent advisory module to offer preliminary guidance in simplified language. Additionally, the platform features a case law search engine and a legal fee estimator, bridging the information asymmetry often faced by the average citizen. The results demonstrate that the proposed framework achieves high text extraction accuracy and provides near-instantaneous responses, enhancing the efficiency of legal research and document verification. Overall, the study highlights the potential of Artificial Intelligence to democratize legal information, empower users with statutory awareness, and foster a more equitable and transparent judicial ecosystem.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Department of Information Technology, Atharva College of Engineering, for providing the necessary facilities and support to carry out this project. The authors also thank the project guide and faculty members for their valuable guidance and continuous encouragement throughout the development of this work.

REFERENCES

[1] M. M. Rahman, N. G. Md, and M. M. Rahman, "Natural Language Processing in Legal Document Analysis Software: A Systematic Review of Current Approaches, Challenges, and Opportunities," International Journal of Innovative Research and Scientific Studies, Vol. 8, No. 3, pp. 5026–5042, 2025.
[2] S. Ajay Mukund et al., "Optimizing Legal Text Summarization Through Dynamic Retrieval-Augmented Generation and Domain-Specific Adaptation," IEEE Conference Publication | Recent Trends in Legal AI, May 2025.
[3] J. K. Verma et al., "Context-Aware Legal Summarization Using Reinforcement Learning," in 2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), IEEE, Ghaziabad, India, 2025.
[4] A. Bouhouche et al., "Predicting the Duration of Judicial Cases Using Hybrid Systems Based on Language Models," International Journal of Advanced Computer Science and Applications, Vol. 16, No. 1, 2025.
[5] Peddarapu et al., "DocSum: A Universal PDF Summarizer Using ASP.NET Core for Professional Legal Insights," Journal of AI and Law, February 2025.
[6] S. Gao and W. Liao, "Transformer-Based Models for Automatic Legal Case Summarization: A Comparative Study of T5 and BART," Artificial Intelligence and Law, Vol. 29, No. 4, pp. 567–588, 2024.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details