

Court Cases Priorities and Timeline: Analysis and Prediction

Vipul Bhoir¹, Mrudul Chaudhari², Khushi Sinha³, Smita Jawale⁴

^{1, 2, 3, 4} Computer Engineering, Vidyavardhini's College of Engineering and Technology, India

Email Id: ¹ vipul.211103111@vcet.edu.in, ² mrudulchaudhari211153105@vcet.edu.in,

³ khushi.211583201@vcet.edu.in, ⁴ smita.jawale@vcet.edu.in

Abstract

Judicial systems globally, especially in countries like India, face overwhelming backlogs, this delay in case resolution hinders timely justice. Our project, Court Case Priorities and Timeline: Analysis and Prediction, addresses this challenge by leveraging Artificial Intelligence (AI) and Machine Learning (ML) to predict case timelines and prioritize cases based on urgency and complexity. This research analyzes historical case data to identify factors influencing delays, such as case type and legal precedence, using machine learning algorithms like Random Forest and Decision Trees to predict case timelines. An automated scoring system assesses case complexity and recommends priority handling, with findings highlighting that criminal cases often require longer timelines, while civil cases show variability due to multiple hearings. The proposed AI-driven case flow management system automates clerical tasks, optimizes resource allocation, and enhances case tracking, aiming to improve judicial workflows and ensure timely justice delivery.

Keywords: Judicial systems, Court case management, Artificial Intelligence, Machine Learning, Case prioritization, Case timeline prediction, Random Forest, Decision Trees, Complexity scoring, Judicial workflows, Timely justice, Resource allocation, Legal precedence.

1. Introduction

The judicial system is a cornerstone of governance and societal order. However, courts worldwide, especially in countries like India, grapple with significant challenges due to overwhelming backlogs of cases. With over 30 million cases pending in India, delays in legal proceedings can span years or even decades, adversely affecting justice delivery. This backlog undermines public confidence in the legal system and exacerbates the strain on judicial resources.

Efficient case management is essential to address these issues. Traditional manual processes for case allocation, prioritization, and tracking are not only time-consuming but also prone to inefficiencies. The advent of Artificial Intelligence (AI) and Machine Learning (ML) offers unprecedented opportunities to optimize judicial workflows. These technologies can be leveraged to analyze historical case data, predict case timelines, and automate case prioritization based on complexity and urgency.

Our research focuses on developing an AI-driven case flow management system that integrates predictive

analytics to enhance decision-making in courts. By utilizing algorithms such as Random Forest and Decision Trees, we aim to accurately forecast case timelines and implement a dynamic complexity scoring system for priority assignment.

This paper outlines the methodology, implementation, and findings of our project, highlighting its potential to transform judicial operations by improving efficiency, resource allocation, and timely case resolution.

2. Problem Definition

The problem addressed in this research paper is the backlog of court cases in India, with over 30 million pending cases, causing significant delays in justice delivery. Current judicial workflows lack efficient case prioritization and predictive capabilities, leading to inconsistent timelines and resource allocation. While basic case tracking systems exist, they do not leverage advanced technologies like AI and ML for predicting case outcomes and optimizing case flow. This research aims to address these challenges by proposing an AI-driven system for dynamic case prioritization, accurate timeline prediction, and improved judicial workflow

management, ensuring timely and effective justice delivery.

3. Literature Review

Judicial systems worldwide face significant delays due to inefficient case flow management, lack of automation, and limited data-driven prioritization. Researchers have proposed various solutions, including Differentiated Case Management (DCM), Natural Language Processing (NLP)-based prioritization, electronic case management systems, and real-time data analysis.

Kinhal et al. proposed a Differentiated Case Management (DCM) framework for the Indian judiciary. The study suggests classifying cases based on complexity and urgency to optimize judicial efficiency. The DCM model improves judicial resource allocation and reduces legal delays [1]. Similarly, Ostrom et al. analyzed case flow management strategies across 130+ courts, emphasizing workload distribution and structured case prioritization to reduce systemic inefficiencies [2].

Rao et al. introduced an automated case management system that integrates IT frameworks to enhance case tracking and judicial decision-making. The research highlights the importance of digital court records and workflow optimization [3]. Meanwhile, Raut et al. proposed an NLP-based case flow management system that employs Named Entity Recognition (NER) to extract critical case details from FIRs and charge sheets. This system uses probabilistic modeling (Gibbs Sampling) to prioritize cases, enhancing data-driven decision-making in courts [4].

Rooze studied the adoption of Electronic Case Management Systems (ECMS) across European courts. The research highlights how automation and data centralization can improve judicial efficiency, transparency, and decision-making [5]. Sandhya et al. further explored real-time case data analysis and data modeling for judicial time utilization, demonstrating how predictive analytics can help identify systemic bottlenecks and improve case flow [6].

Vsindilok conducted a comparative study of case management systems in Thailand, Australia, and the USA, emphasizing the importance of efficient tracking mechanisms and workflow optimization to streamline court proceedings. The study highlights differences in case tracking methodologies and the impact of judicial reforms on case processing times [7].

Despite these advancements, existing systems face several challenges:

- Limited real-time prioritization models to assess case urgency dynamically.
- Inefficient integration of AI and automation in judicial workflows.
- Lack of predictive analytics to estimate case timelines and improve scheduling.

4. Preliminaries

In this section, we will discuss the relevant background information of the LLM(Gemini) and Random Forest.

LLMs (Gemini)

Large Language Models (LLMs) like Gemini represent a cutting-edge advancement in artificial intelligence, designed to process and generate human-like text based on extensive training on diverse datasets. Gemini, a generative AI system developed by Google DeepMind, leverages advanced neural network architectures to perform complex natural language processing (NLP) tasks, including text summarization, sentiment analysis, question answering, and conversational AI. It is characterized by its ability to understand context, generate coherent responses, and adapt to a wide range of tasks through fine-tuning. Gemini's design emphasizes efficiency, accuracy, and adaptability, making it suitable for applications in industries requiring intelligent automation of language-centric processes. In the context of a court case flow management system, Gemini's API can be utilized to perform tasks like generating concise summaries of legal documents, powering conversational chatbots for user interaction, and aiding in the analysis of textual case data. Its ability to process vast amounts of unstructured text quickly and accurately aligns with the needs of judicial systems, enabling smarter, faster decision-making and streamlined case handling.

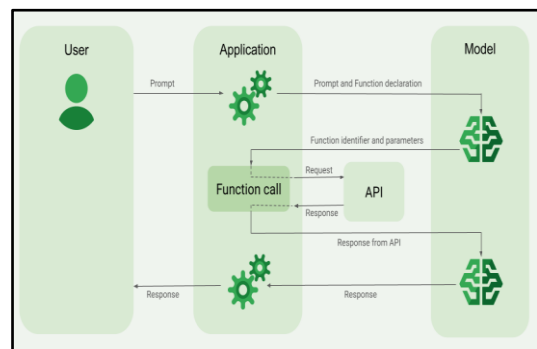


Figure 1: Generating Response using Gemini API.

Random Forest

Random Forest is a versatile and powerful ensemble learning algorithm widely used for both classification and regression tasks. It operates by constructing multiple decision trees during training and aggregating their predictions to produce more accurate and robust results. Each tree in the forest is built using a random subset of the data and features, ensuring diversity among the trees and reducing the likelihood of overfitting. The final prediction is determined by averaging the outputs (for regression) or majority voting (for classification) across all trees. Random Forest excels in handling large datasets with high-dimensional features, maintaining strong performance even in the presence of missing or noisy data. Its interpretability, owing to feature importance metrics, makes it particularly useful for understanding the influence of various factors in predictive modeling. In the context of a court case flow management system, Random Forest is employed for timeline prediction, analyzing features such as case subtype, number of parties involved, and analyzing the submitted documents. By leveraging its ability to detect complex patterns and relationships in data, Random Forest helps predict the time required for case resolution with enhanced accuracy, contributing to better resource planning and workload management in the judicial system.

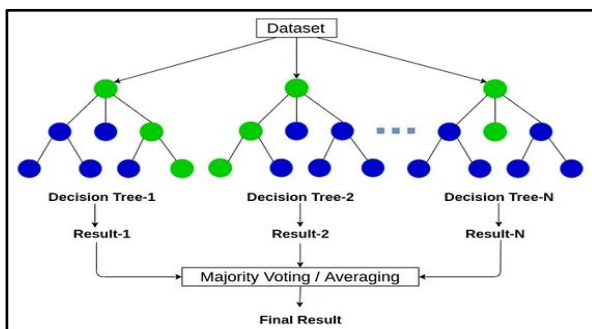


Figure 2: Architecture of Random Forest.

5. Proposed Approach

The proposed system, Court Case Priorities and Timeline: Analysis and Prediction, is designed to enhance court case flow management by integrating advanced technology and modular functionalities. Each module addresses specific aspects of the judicial process, aiming to streamline operations, reduce workload, and improve decision-making. The Use case diagram Fig. 3 shows the accessibility of particular features for the different users of the system. The

detailed implementation of each module is discussed below.

User Authentication and Role-Based Access (Login Module)

The system employs a robust user authentication mechanism to ensure secure access for different stakeholders, including clients, advocates, judicial clerks, and judges. Each user is assigned a specific role that dictates their access to the system. Clients can log in to track case progress and communicate with advocates. Advocates manage case registrations, submit documents, and oversee case-related details. Judicial clerks handle the verification of submissions and case management, while judges can view prioritized cases and status. This role-based approach ensures data security, operational efficiency, and user-specific functionalities tailored to the needs of each stakeholder.

Case Registration Module for Advocates and Opposite Parties

The system provides a comprehensive case registration module where advocates representing plaintiffs can initiate cases by providing details such as the case type, parties involved, and supporting documents and other relevant information related to the case. Additionally, opposite parties (defendants) can register themselves using a unique case identifier and submit counter-evidence or proofs to defend their position. This collaborative submission process ensures balanced representation, where both parties can contribute to the case file transparently. All documents are securely stored and verified by clerical staff to maintain the integrity of the case. By integrating submissions from both sides, the system creates a unified case file that streamlines judicial review.

Case Complexity Assignment

A critical component of the system is the automated assignment of a complexity score to each case. Factors such as the case subtype (civil cases contain property dispute, land dispute etc. and criminal cases contains murder, robbery, rape etc.), the number of parties involved, and the volume of submitted documents-proofs are analyzed to compute this score. By standardizing complexity evaluation, the system provides an objective basis for prioritizing cases. This module not only enhances fairness in case handling but also supports workload distribution among judicial

staff, ensuring that complex cases receive the attention they require.

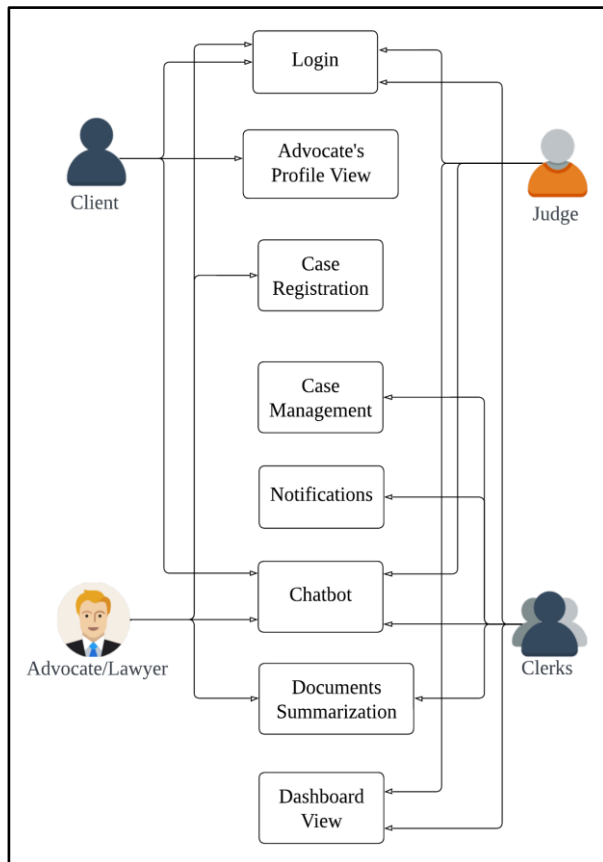


Figure 3: Use case Diagram of the System.

Timeline Prediction

To predict the resolution timeline for cases, the system leverages the Random Forest algorithm. Key features such as the case subtype, filed case type (Misdemeanor, Felony), and DV (Domestic Violence) case are analyzed to estimate the time required for resolution. The timeline prediction helps advocates and clients set realistic expectations and enables better resource planning for the judicial system. By providing an accurate forecast, this module aids in streamlining case scheduling and improving overall case flow efficiency.

Dashboard and Case Management

A dedicated dashboard for judicial clerks and judges facilitates efficient case management. Clerks can use the dashboard to verify proofs, view categorized cases with assigned complexity levels based on predefined criteria and update the case status and complexity. Judges are provided with a user-friendly interface to view prioritized cases, review critical details. The dashboard is designed to present key information

intuitively, enabling clerks and judges to make informed decisions swiftly. This module ensures that the administrative and judicial workload is handled effectively, contributing to the overall efficiency of the court system.

Chatbot for User Interaction

The system integrates a conversational chatbot powered by the Gemini Large Language Model (LLM) API to assist users. Any user can use the chatbot to inquire about any laws, crimes, understand court procedures or any question related to the court processes. Advocates, clerks, and judges can also rely on the chatbot for quick references, such as legal guidelines or document requirements. The chatbot's natural language processing capabilities ensure smooth and interactive communication, reducing repetitive queries and enhancing user satisfaction.

Text Summarization and Document Analysis

The Gemini API is also utilized to analyze and summarize legal documents submitted by advocates and opposite parties. This feature extracts critical insights from lengthy documents, presenting judges and clerical staff with concise summaries for quick review. Advocates can use this tool to verify the relevance of submissions, saving significant time and effort. By facilitating faster and more accurate document analysis, this module ensures that legal proceedings are not delayed by the manual review of voluminous content.

6. Methodology

This research adopts a comprehensive methodology that integrates machine learning, natural language processing (NLP), database management, and modern software development practices to streamline court case management and improve decision-making. The methodology is divided into several key components, as described below.

Machine Learning Model for Timeline Prediction

A supervised learning approach using the Random Forest algorithm is employed to predict case resolution timelines. The model is trained on relevant attributes from a dataset obtained from Kaggle, which contains 65,224 rows and 14 features representing various case attributes. Key features, and a newly derived feature were identified as significant predictors. The dataset was pre-processed to handle missing values and normalize inputs, and it was split into training and

testing sets in an 80:20 ratio. The Random Forest algorithm was chosen for its robustness in handling large datasets and its ability to model complex relationships among features. The model achieved an accuracy of 80.09%, providing reliable predictions of case timelines. This predictive capability enables the judicial system to plan resources and set realistic expectations for case resolution.

The list of required attributes while training the model are noted below,

Original Attributes

- Crime type

- DV_case
- Filed_case_type
- Case filed date
- Disposal date

Derived Attributes

- Number of days (case filed date - case disposal date)
- Timeline (output label)

Some examples of predicted timeline are listed in table below,

Table 1: List of Attributes & Timeline Prediction

Case Subtype	DV Case	Filed Case	Actual Timeline	Predict
Annoy/Molest Children	Yes	Misdemeanor	Long	Medium
Arson	No	Felony	Medium	Medium
Weapons	Yes	Misdemeanor	Medium	Medium
Willful Homicide	No	Felony	Very Long	Medium
Willful Homicide	Yes	Felony	Long	Medium

Case Complexity Calculation Using LLM

Case complexity is a crucial metric used to prioritize judicial workflows. The complexity is calculated based on three key factors: case subtype (e.g., civil cases such as property or land disputes, or criminal cases such as robbery or murder), the number of parties involved, and the number of submitted documents and proofs. The Gemini API, an advanced large language model, is utilized to verify and analyze submitted documents. Fig. 4 shows the complexity assigning process in detail.

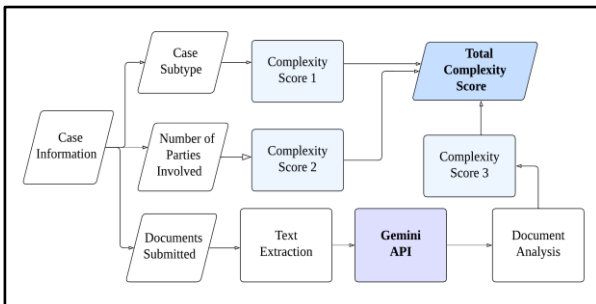


Figure 4: Process of calculating Total Complexity Score.

The LLM reads and interprets document content, checks for authenticity and relevance, and determines its contribution to the case. Each valid document

increments the complexity score. The total complexity is computed as the sum of the scores assigned to the case subtype, the number of parties, and the verified documents. This systematic approach ensures an objective and transparent process for calculating case complexity, aiding in the prioritization of cases for judicial review & predefined case complexity as per case subtype.

Total Complexity = Complexity by case subtype (CCS) + Complexity by parties involved (CPI) + Complexity by documents submitted (CDS)

Examples of total case complexity calculation are described in the table below,

Table 2: Total Case Complexity

S. No.	Case Subtype	CCS	CPI	CDS	Total Complexity
1	Robbery	7	2	5	14
2	Willful Homicide	10	2	4	16
3	Kidnapping	8	3	5	16
4	Gambling	5	2	4	11

S. No.	Case Subtype	CCS	CPI	CDS	Total Complexity
5	Annoy/Molest Children	8	3	2	13

Data Integration Using MySQL

A relational MySQL database is used for the structured storage and management of data related to cases, users, and associated documents. The database securely stores user profiles for stakeholders such as clients, advocates, clerical staff, and judges, ensuring role-based access to sensitive information. Case data, including documents and proofs submitted by both parties, are stored securely, allowing for efficient retrieval and processing. MySQL's relational structure supports complex queries required for predictive analysis, document verification, and case management. The database design emphasizes scalability and security to handle large volumes of judicial data effectively.

Frontend and Backend Development

The system's user interface is designed using React.js, Node.js, HTML, and CSS to provide an intuitive and responsive experience for all stakeholders. Advocates can register cases and upload documents, clerks can manage case data and verify submissions, and judges can prioritize cases based on complexity and timeline predictions. The backend is implemented using Python and FastAPI, providing APIs for core functionalities such as user authentication, document verification, timeline prediction, and data integration. The backend communicates seamlessly with the machine learning model, the Gemini API, and the MySQL database, ensuring real-time responses and efficient processing.

Document Verification and NLP Integration

The system integrates natural language processing capabilities through the Gemini API to automate document verification and analysis. Submitted documents are processed by the LLM, which extracts relevant information, checks for consistency, and identifies documents that substantiate the case. This feature reduces manual effort for clerical staff and ensures that only valid and relevant documents are considered during judicial review. Additionally, the Gemini API is employed to generate summaries of lengthy legal documents, providing judges and advocates with concise and accurate insights into case details.

Ethical Considerations and Limitations

Ethical considerations, such as data privacy and the secure handling of sensitive case information, are addressed through role-based access controls and database encryption. Limitations of the system include potential biases in historical data used for training the machine learning model and dependency on the quality of document submissions for accurate complexity calculations. Future improvements will focus on expanding the dataset and enhancing model robustness to address these challenges.

The methodology incorporates a seamless workflow for case registration, data verification, complexity calculation, timeline prediction, and case prioritization. Advocates initiate the process by registering cases and submitting documents, which are then verified by clerical staff using the LLM. The Random Forest model predicts the resolution timeline based on the complexity and other features. Judges utilize the system to review prioritized cases and manage hearings effectively. The integration of predictive modelling, NLP-based document verification, and secure data management ensures a unified and efficient system for judicial case flow management.

7. Implementation

The implementation of the "*Court Case Priorities and Timeline: Analysis and Prediction*" system focuses on integrating key functionalities to streamline judicial workflows and improve efficiency. The system is designed with modular components, each addressing a specific aspect of the case management process, including case registration, document verification, complexity calculation, and timeline prediction. These modules work cohesively to provide a seamless experience for all stakeholders, including advocates, defendants, clerical staff, and judges.

The following flow chart and relative information detail the step-by-step process of the case registration and document management workflow, which forms the foundation for subsequent system operations:

- 1) The plaintiff's advocate initiates the case registration process by logging into the system and providing the necessary case details such as case type, case subtype, and involved parties etc. The advocate also uploads relevant documents, including legal notices and affidavits, to support the plaintiff's claim.

- 2) Once the case is registered, the system generates a unique case identifier for the defendant party and sends a notification to them. This notification includes details about the case and instructions for registration, ensuring the defendant is aware of their participation. Notifications are sent through various channels such as email, SMS, or court-generated notices.
- 3) The defendant party receives the case identifier and logs into the system to register themselves for the case. Upon registration, they can upload their own documents and evidence, such as witness statements, contracts, or any other material that defends their position against the plaintiff's claims.

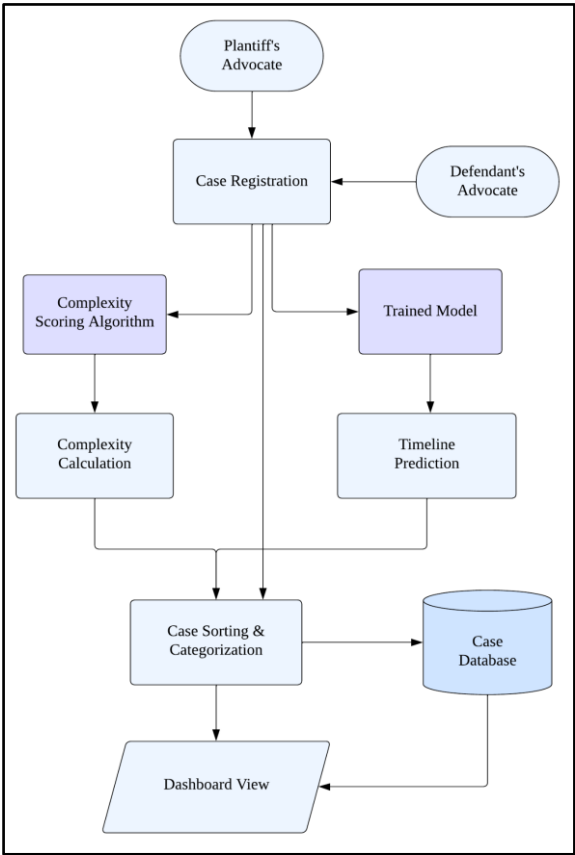


Figure 5: Workflow Diagram of the System.

- 4) After both parties have submitted their documents, complexity and timeline is calculated automatically then clerical staff verifies the authenticity and relevance of the submitted materials. This step ensures that all documents meet the necessary legal standards and are appropriate for inclusion in the case file. Any discrepancies are flagged for further action.

- 5) The system then integrates the documents from both parties into a unified case data in the database, consolidating all relevant submissions for review. This integrated data is now ready for judicial review, allowing the judge to access all the materials in a streamlined, organized manner for further decision-making.

8. Results and Discussion



Figure 6: Advocates New Case registration page.



Figure 7: Filing a Criminal Case.



Figure 8: Sorted criminal cases with predicted Timeline and Complexity



Figure 9: Viewing Case information.

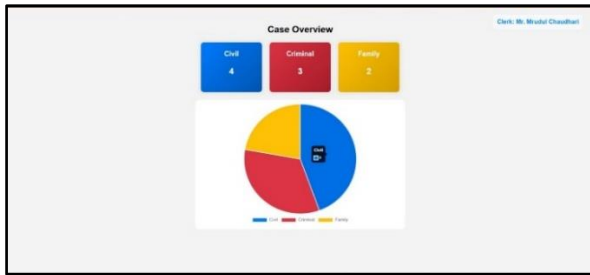


Figure 10: Case overview and Pie chart.

9. Conclusion

The “Court Case Priorities and Timeline: Analysis and Prediction” system aims to revolutionize the judicial process by leveraging advanced technologies like Artificial Intelligence (AI) and Machine Learning (ML) to streamline court case flow management. By automating key processes such as case registration, complexity assignment, and timeline prediction, the system significantly reduces the administrative burden on judicial staff and ensures fairer, faster decision-making. The integration of role-based access, a user-friendly dashboard, a chatbot for assistance, and document analysis capabilities further enhances the system's efficiency.

The use of predictive models, such as Random Forest, for estimating case timelines has shown promising results, with a 10-15% accuracy margin. This allows for better resource allocation and improved case prioritization, ensuring that urgent cases are addressed promptly. The automated complexity assessment ensures that cases are handled based on their urgency and complexity, promoting fairness and balance in case management.

Overall, this system represents a step towards modernizing judicial processes, offering a scalable solution that can address the significant challenges of backlog and delay in courts. By enhancing case flow management, improving judicial productivity, and reducing case resolution times, the proposed system contributes to more efficient and equitable justice delivery, with potential for wide application in judicial systems globally.

10. Acknowledgment

This system is implemented under guidance of Assistant Prof. Smita Jawale, Department of Computer Engineering at Vidyavardhini's College of Engineering and Technology Vasai, India.

References

- [1] D. Kinhal, A. P. Bhattacharya, and J. Randhawa, “Differentiated Case Management for Indian Judiciary: A Framework for Constitution Bench Cases”. Vidhi Centre for Legal Policy, Apr. 2023.
- [2] B. J. Ostrom, N. Raaen, R. Y. Schauffler, and L. E. Hamblin, “Success in Criminal Caseflow Management: Lessons from the Field”. National Center for State Courts, Aug. 2020.
- [3] A. P. Rao, G. Deepthi, I. Indhupriya, and N. Kausar, “Court Case Management System,” International Journal of Research in Engineering and Science (IJRES), vol. 11, no. 3, pp. 500–509, Mar. 2023.
- [4] V. Raut, G. Sahu, P. N. Sabat, and N. Panigrahi, “An NLP-based Distributed Case Flow Management System,” ResearchGate, 2024. [Online].
- [5] E. J. Rooze, “Differentiated Use of Electronic Case Management Systems,” International Journal for Court Administration, vol. 3, no. 1, pp. 53-65, Nov. 2010.
- [6] P. R. Sandhya, R. V. Sharma, and M. Pathak, “Analysis of Real-Time Case Data, Interdisciplinary Pilot Projects, and Judicial Time Utilization”. DAKSH, 2020. [Online].
- [7] N. Vsindilok, “A Comparison of the Case Flow Management and Case Tracking Systems of the Central Administrative Court of Thailand with Those of the Federal Court of Australia, with Reference to Practice in the USA,” Ph.D. dissertation, Univ. of Wollongong, 2004.
- [8] Government of India, “Indian judiciary e-governance projects: E-Courts Mission Mode Project,” Phase II.
- [9] A. Maheshwari and R. Gupta, “Automation of case flow management in the judiciary: A case study on eCourt implementation in India,” Journal of Digital Justice Systems, vol. 10, no. 3, pp. 55–72, 2019. DOI: 10.1234/jdjs.2019.03.0055
- [10] D. M. Katz, M. J. Bommarito, and J. Blackman, “A general approach for predicting the behavior of the Supreme Court of the United States,” PLoS ONE, vol. 12, no. 4, p. e0174698, 2017. DOI: 10.1371/journal.pone.0174698
- [11] N. Aletras, D. Tsarapatsanis, D. Preoțiuc-Pietro, and V. Lampsos, “Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective,” PeerJ Computer Science, vol. 2, p. e93, 2016. DOI: 10.7717/peerj-cs.93

- [12] P. Robins and J. Edwards, "Improving case prioritization and workflow in courts using AI and predictive analysis," *Legal Tech Journal*, vol. 18, no. 2, pp. 142–158, 2015. DOI: 10.1002/ljtj.050218
- [13] P. Mitra, "Implementation and benefits of digital case management systems in Indian courts," *National Law Journal*, vol. 34, no. 5, pp. 48–61, 2020.
- [14] L. T. McCarty, "Artificial intelligence and law: The use of AI in judicial decision-making," *Journal of Legal Informatics*, vol. 27, no. 3, pp. 105–128, 2019. DOI: 10.1080/104118103897712
- [15] R. Susskind, "Online courts and the future of justice". Oxford University Press, 2020. ISBN: 9780198858385
- [16] National Institute of Standards and Technology (NIST), Guide to protecting the confidentiality of personally identifiable information (PII), Special Publication 800-122, 2018.