iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 14 **Issue:** IV **Month of publication:** April 2026

**DOI:** https://doi.org/10.22214/ijraset.2026.80136

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 14 Issue IV Apr 2026- Available at www.ijraset.com*

# Automated Crime Reporting and Charge Sheet Generation System Using Reinforcement

Vikash Mallick[1], Tanmayi Jaiswal[2], Sujal Dhurve[3], Arpita Bankar[4], Rakesh Moharle[5], Sushama Telrandhe[6]
[1, 2, 3, 4]*Dept of CSE, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India*
[5]*Assistant Professor, Dept of CSE, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India*
[6]*Associate Professor, Dept of CSE, Gurunanak Institute of Engineering and Technology, Nagpur, Maharashtra, India*

***Abstract:*** ***Preparing crime records, First Information Reports (FIRs), and charge sheets by hand continues to be a serious obstacle in policing due to human mistakes, procedural slowdowns, and inconsistent record-keeping. To overcome these shortcomings, this study presents an AI-powered framework that handles FIR creation, crime categorization, IPC-based charge sheet drafting, and punishment estimation automatically. The framework combines machine learning, natural language processing, and legal data analytics to identify offense categories, link them to suitable IPC sections, and produce structured legal paperwork. Logistic Regression paired with text vectorization methods allows dependable crime classification, while predictive analytics assists in estimating sentences using past case records. This method boosts operational productivity, lowers dependence on manual effort, and guarantees consistency and transparency across law enforcement processes. Future developments may involve blockchain-based data protection and live integration with national crime databases.***
***Keywords: Automated Crime Reporting, FIR Generation, Charge-Sheet Automation, Crime Classification, Logistic Regression, Natural Language Processing, Indian Penal Code (IPC), Judicial Automation, Sentencing Prediction, Predictive Analytics, Machine Learning in Law Enforcement, Legal Document Processing, Crime Data Analysis, AI-Assisted Policing, Digital Forensics Integration.***

## I. INTRODUCTION

Even with major steps taken toward digital governance, the core elements of India's criminal justice framework still depend greatly on manual processes such as writing FIRs, preparing charge sheets, recording evidence, and applying IPC-based classifications. This reliance on manual effort raises the chances of human error, process delays, and uneven interpretation of legal standards [4]. As police stations handle thousands of incident reports each day, these inefficiencies lead to overwhelmed courts, growing case backlogs, and slow delivery of justice.

AI-powered automation holds the potential to transform these shortcomings. NLP methods accurately pull meaningful information from written complaint text, making it possible to automatically handle narrative descriptions [1], [12]. ML classification algorithms, particularly Logistic Regression, Naïve Bayes, and XGBoost, have been widely applied to categorize legal papers and crime types with strong accuracy [2], [5]. Similarly, predictive systems have delivered promising By combining these capabilities, the proposed system supports automatic FIR creation, offense categorization, IPC mapping, and sentence prediction. This kind of automation matches global directions where legal AI tools are increasingly being used for document screening, case-law lookup, fraud identification, and judicial decision assistance [12], [15]. The system therefore improves efficiency, transparency, and fairness across law enforcement procedures.

## II. LITERATURE SURVEY

AI and NLP have become essential instruments in modern legal informatics because of their capacity to interpret, sort, and examine complex written content. Several significant studies highlight how they are applied in criminal justice:

[1] NLP for Legal Documentation

Research by Vattikuti (2024) shows that transformer-based language models such as BERT and GPT greatly enhance clause identification, anomaly spotting, and document review precision in legal settings [1]. These tools lower the need for manual reading and improve uniformity across documents.

[2] Automated Crime Classification

Ku and Leroy (2013) created a decision-support tool using NLP and Naïve Bayes classifiers to sort crime reports. Their system reached an accuracy rate of 94.82%, outperforming human analysts [2]. This proves the dependability of ML in automated crime categorization.

[3] Charge Prediction and Legal Reasoning

Ye et al. (2018) built a text-to-text model that produces judicial reasoning from case descriptions. Their label-guided Seq2Seq structure improved explainability and showed strong usefulness for predicting charges [3].

[4] Ethical and Human Rights Concerns

Završnik (2020) stressed that AI applied in criminal justice must account for fairness, openness, and the risk of bias to prevent deepening existing inequalities [4]. These concerns highlight the need for AI systems that are explainable and accountable.

[5] Neural Networks for Crime Forecasting

Walczak (2021) and Shah et al. (2021) demonstrated how neural networks and computer vision systems can forecast crime hotspots, detect unusual patterns, and classify incidents using data from multiple sources [10], [11].

[6] Deep Learning and Crime Prediction

Safat et al. (2021) applied deep learning models to large datasets from Chicago and Los Angeles, showing strong results in crime forecasting while also pointing out difficulties such as biased datasets and unbalanced data [9].

Across all reviewed literature, it is clear that combining NLP and ML with legal domain expertise can greatly improve the automation of crime reporting and judicial documentation processes.

## III. SYSTEM ARCHITECTURE

The architecture uses a multi-module design where each component contributes toward automation and decision support.

*1)* Complaint Processing Interface: Users submit written descriptions of incidents. NLP techniques including tokenization, Named Entity Recognition, and dependency parsing are used to pull out key details. Similar methods have been applied in legal information retrieval and document categorization.

*2)* NLP-Based FIR Generator: NLP models convert free-form written input into properly formatted FIR fields covering crime nature, incident timeline, location, and individuals involved. Earlier research confirms the strength of NLP in extracting legal entities and producing structured documents.

*3)* Crime Classification Module: A Count Vectorizer turns text into numerical data, and Logistic Regression then classifies the offense as either cognizable or non-cognizable. Logistic models are commonly used in law-related text classification because they are easy to interpret.

*4)* Evidence Repository: Digital evidence including images, videos, and documents is stored in a central database. Such storage systems allow for efficient retrieval and cross-case linking, as supported by forensic AI research.

*5)* Charge-Sheet Generation Engine: Based on the identified offense type, the system connects the incident to the relevant IPC sections using pre-set legal rules combined with ML-based inference. Comparable systems blending rule-based and ML approaches have been put forward in legal analytics research.

*6)* Sentencing Prediction Module: Using historical case records, the system estimates the severity of sentencing with ML models. Predictive tools have been used in various legal settings, including predicting European Court of Human Rights judgments, US Supreme Court rulings, and general legal outcome modeling.

## IV. FIR GENERATION

The FIR is produced automatically using a pre-built structured template. NLP supports the extraction of the following details:

*1)* Crime description along with the surrounding context

*2)* Legal classification as cognizable or non-cognizable

*3)* Location and time-related information

*4)* Identification of suspects and witnesses through Named

Entity Recognition

*1)* Summary of evidence drawn from uploaded documents

*2)* Assigned **officer details and current case status**

## V. OUTPUT AND INTERPRETATION

The system delivers the following outputs:

*1)* Crime category label (cognizable or non-cognizable)

*2)* Formally structured FIR produced through NLPprocessing

*3)* IPC-mapped charges applicable to cognizable offenses

*4)* Suggested sentencing ranges based on ML-generatedpredictions

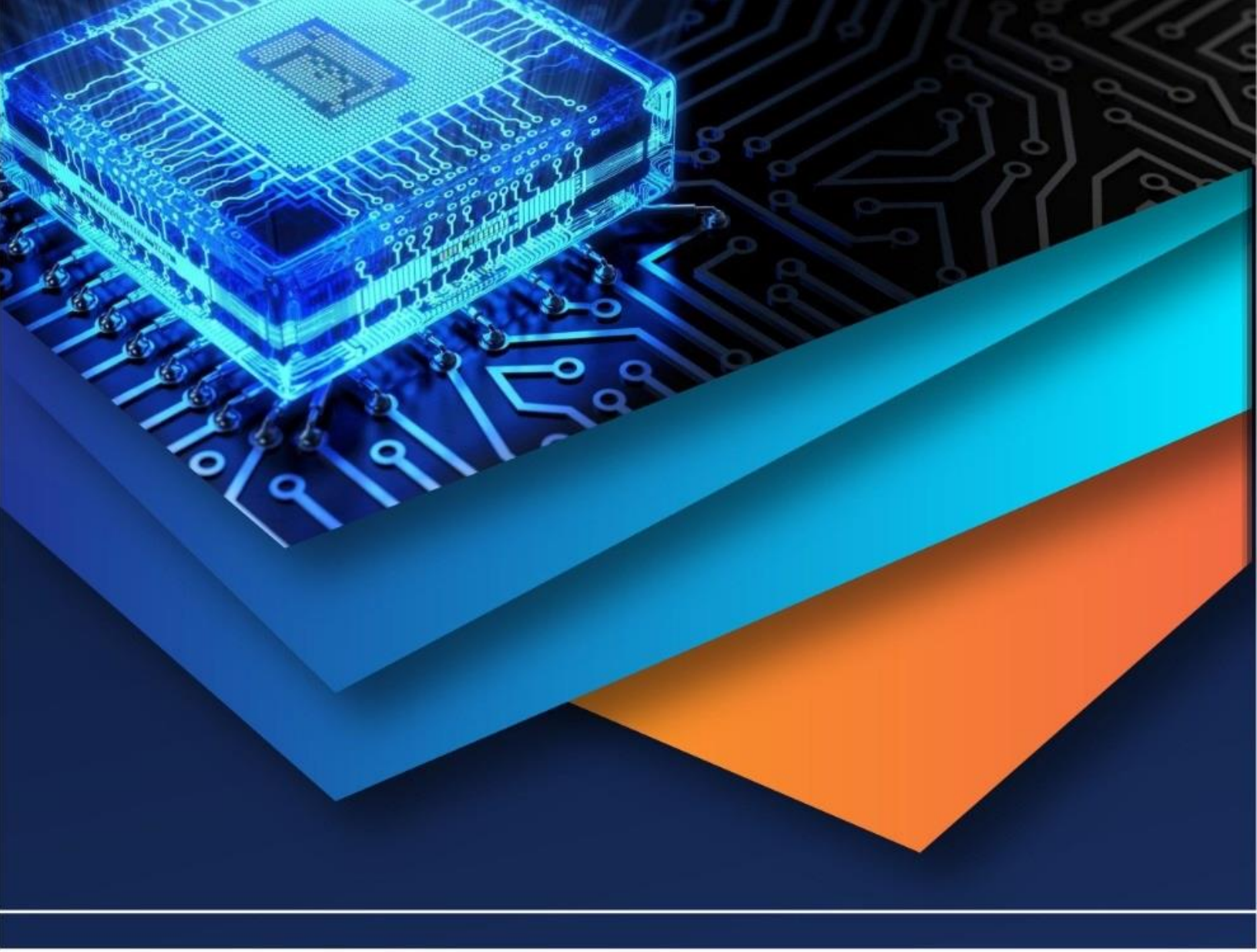
The reliability of ML-based classification is measured using standard metrics including accuracy, precision, recall, and F1-score, in line with evaluation methods used in crime forecasting studies [9], [10].

## VI. CONCLUSION

This research establishes that combining NLP and ML can effectively automate the most critical parts of criminal justice workflows. Tasks such as FIR drafting, offense classification, IPC mapping, and sentencing prediction — which were previously handled manually and consumed significant time — can now be reliably managed by machines. The proposed system draws from well-established approaches including Logistic Regression, Seq2Seq reasoning models, crime prediction frameworks, and forensic AI tools [1]–[14], [16]. Future enhancements may include integration of transformer-based language models, blockchain-supported audit records, and real-time connectivity with national crime databases for complete end-to-end automation. These steps represent a major advancement toward a criminal justice system that is transparent, efficient, and driven by modern technology.

## REFERENCES

[1] Vattikuti, Manoj Chowdary. "Natural Language Processing for Automated Legal Document Analysis and Contract Review." International Journal of Sustainable Development in Field of IT, vol. 16, no. 16, 2024.
[2] Ku, Chih-Hao, and Gondy Leroy. "Automated Crime Report Analysis and Classification for E-Government and Decision Support." Proceedings of the 14th Annual International Conference on Digital Government Research, 2013.
[3] Ye, Hai, et al. "Interpretable Charge Predictions for Criminal Cases: Learning to Generate Court Views from Fact Descriptions." arXiv preprint arXiv:1802.08504, 2018.
[4] Završnik, Ales. "Criminal Justice, Artificial Intelligence Systems,and Human Rights." ERA Forum, vol. 20, no. 4, 2020.
[5] Jenga, Karabo, Cagatay Catal, and Gorkem Kar. "MachineLearning in Crime Prediction." Journal of Ambient Intelligence and Humanized Computing, vol. 14, no. 3, 2023, pp. 2887–2913.
[6] Jeong, Doowon. "Artificial Intelligence Security Threat, Crime,and Forensics: Taxonomy and Open Issues." IEEE Access, vol. 8, 2020, pp. 184560–184574.
[7] King, Thomas C., et al. "Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions." Science and Engineering Ethics, vol. 26, 2020, pp. 89–120.
[8] Sushina, Tatyana, and Andrew Sobenin. "Artificial Intelligence inthe Criminal Justice System: Leading Trends and Possibilities." 6th International Conference on Social, Economic, and Academic Leadership (ICSEAL-6-2019), Atlantis Press, 2020.
[9] Safat, Wajiha, Sohail Asghar, and Saira Andleeb Gillani. "Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques." IEEE Access, vol. 9, 2021, pp. 70080–70094.
[10] Shah, Neil, Nandish Bhagat, and Manan Shah. "CrimeForecasting: A Machine Learning and Computer Vision Approach to Crime Prediction and Prevention." Visual Computing for Industry, Biomedicine, and Art, vol. 4, no. 1, 2021.
[11] Walczak, Steven. "Predicting Crime and Other Uses of NeuralNetworks in Police Decision Making." Frontiers in Psychology, vol. 12, 2021, 587943.
[12] Surden, Harry. "Artificial Intelligence and Law: An Overview."Georgia State University Law Review, vol. 35, 2018, p. 1305.
[13] Yang, Wenmian, Weijia Jia, Xiaojie Zhou, and Yutao Luo. "Legal Judgment Prediction via Multi-Perspective Bi-Feedback Network." arXiv preprint arXiv:1905.03969, 2019.
[14] Medvedeva, Masha, Michel Vols, and Martijn Wieling. "Using Machine Learning to Predict Decisions of the European Court of Human Rights." Artificial Intelligence and Law, vol. 28, no. 2, 2020, pp. 237–266.
[15] Ashley, Kevin D. Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age. Cambridge University Press, 2017.
[16] Zhong, Haoxi, Zhipeng Guo, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu, and Maosong Sun. "Legal Judgment Prediction via Topological Learning." Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pp. 3540–3549, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY