

Research Paper

A Data Science Approach to Analyzing and Understanding Crime Patterns for Improved Law Enforcement Strategies

¹Battu Denny Paul

¹M.Tech, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
dennypaul989@gmail.com

²Mr. Ch. Venkata reddy

²Assistant Professor, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
venkat95dnr@gmail.com

Abstract— Crime analysis is essential for understanding regional crime trends and supporting effective law enforcement strategies. Conventional approaches to crime analysis often depend on manual examination and statistical reporting, which may limit the identification of complex relationships across geographical regions and crime categories. This study presents a data science approach for analyzing and understanding crime patterns across different states and union territories in India using historical crime data. The proposed approach integrates data preprocessing, exploratory analysis, K-Means clustering, and supervised machine learning techniques for crime pattern analysis and prediction. K-Means clustering is used to group regions according to similarities in crime trends and identify areas with high, medium, and low crime incidence. Random Forest, Logistic Regression, and Support Vector Machine classifiers are applied to analyze historical crime patterns and predict crime categories. Random Forest further provides insights into the relative importance of crime-related features, while classification models support predictive analysis. The integration of clustering and classification provides both descriptive and predictive perspectives for improved law enforcement planning, resource allocation, and targeted crime prevention.

Keywords— Crime Pattern Analysis, Data Science, Machine Learning, K-Means Clustering, Random Forest, Logistic Regression, Support Vector Machine, Crime Prediction, Spatial Analysis, Law Enforcement.

I. INTRODUCTION

Crime is an important social and administrative challenge that affects public safety, economic development, and the quality of life of communities. Understanding crime patterns is essential for supporting effective law enforcement, improving public safety, and developing evidence-based prevention strategies. Historical crime records contain valuable information about crime categories, geographical regions, frequencies, and changes across different years. These patterns can vary according to geographical location, population characteristics, socioeconomic conditions, and temporal factors. Recent advances in data science and machine learning have created opportunities for intelligent crime analysis, crime classification, pattern identification, and predictive law enforcement strategies [1]–[3].

Conventional crime analysis generally relies on manual examination of records, statistical summaries, and traditional reporting methods. Although these approaches provide useful information about recorded criminal activities, they have limited capability to efficiently identify complex relationships across large historical datasets. Machine-learning techniques can learn relationships from historical crime observations and generate predictions for different crime categories and regional patterns [3], [5]. Classification algorithms such as Random

Forest, Logistic Regression, and Support Vector Machines can further identify relationships between crime-related features and target crime categories [5], [6].

Crime analysis is important because crime patterns can change across geographical regions and over time. Clustering and data-mining approaches have therefore been explored to identify groups of regions with similar crime characteristics and to discover recurring crime patterns [6], [8]. Spatial and temporal crime analysis can additionally support the identification of crime hotspots and changing regional trends [8]–[10]. However, differences in datasets, missing observations, variations in crime reporting, feature selection, geographical characteristics, and model performance remain important challenges in developing reliable crime-analysis systems [1], [2], [11].

Prediction alone cannot provide complete support for effective law enforcement. A practical data-driven crime-analysis system should integrate data preprocessing, exploratory analysis, regional clustering, crime classification, pattern identification, and decision support. Such an integrated approach can provide both descriptive and predictive information for understanding crime behavior and supporting targeted prevention strategies.

This study therefore aims to:

1. Analyze historical crime data across Indian states and union territories.
2. Examine temporal and geographical variations in different crime categories.
3. Apply K-Means clustering to identify regions with similar crime patterns.
4. Compare Random Forest, Logistic Regression, and Support Vector Machine classification approaches.
5. Identify important crime-related features influencing classification results.
6. Develop an integrated framework combining clustering and classification for crime-pattern analysis.
7. Identify evaluation requirements, research challenges, and future directions for data-driven law enforcement strategies.

The remainder of this paper discusses existing crime-analysis and prediction approaches, the proposed comprehensive crime-pattern analysis framework, evaluation requirements, research challenges, future research directions, and concluding observations.

II. EXISTING METHODS FOR CRIME PATTERN ANALYSIS AND PREDICTION

Existing crime pattern analysis approaches can be broadly classified into traditional statistical analysis, geographical and hotspot analysis, clustering approaches, machine-learning classification, predictive crime analysis, and integrated data-driven law enforcement frameworks.

A. Traditional Crime Data Analysis

Traditional crime analysis depends on historical records, statistical summaries, police reports, and manual examination of crime incidents. It can provide useful information about previously recorded criminal activities and regional crime frequencies, but analyzing large and complex datasets manually becomes difficult. Limited automation and difficulties in identifying relationships among multiple crime categories and geographical regions can restrict effective crime-pattern discovery [1], [3].

B. Geographical and Hotspot Analysis

Geographical crime analysis examines the distribution of criminal activities across different locations and can help identify areas with relatively high concentrations of crime. Hotspot-based approaches provide useful information for understanding spatial crime patterns and supporting police resource allocation. However, geographical analysis may become less effective when temporal variations, multiple crime categories, and large historical datasets need to be analyzed simultaneously [8], [9].

C. Clustering-Based Crime Analysis

Clustering approaches group regions or crime observations according to similarities in their characteristics. K-Means clustering can be applied to historical crime data to identify regions with similar crime trends and classify them into relatively high, medium, and low crime-incidence groups. The main advantage of clustering is its ability to discover patterns without requiring predefined class labels. However, clustering performance depends on feature selection, data quality, and the characteristics of the available dataset [2], [6].

D. Conventional Machine Learning

Machine-learning approaches learn relationships between crime-related features from historical observations. Algorithms such as Random Forest, Logistic Regression, and Support Vector Machines can be applied for crime classification and prediction. These approaches provide automated analysis and can identify patterns within historical crime records. Random Forest can additionally provide information about important features influencing prediction results. However, model performance depends on dataset quality, feature representation, class distribution, and regional variations [1], [5].

E. Predictive Crime Analysis

Predictive crime-analysis approaches use historical crime information to estimate future crime patterns or likely crime categories. Machine-learning models can identify recurring relationships between crime characteristics, geographical regions, and historical trends, supporting proactive law enforcement strategies [3], [7]. Such approaches can provide useful predictive information, although differences in crime reporting, limited historical observations, changing crime behavior, and prediction uncertainty remain important challenges [4], [13].

F. Integrated and Intelligent Crime Analysis

Integrated crime-analysis frameworks combine multiple analytical approaches to provide broader understanding of crime patterns. Spatial analysis, machine learning, clustering, and predictive techniques can be combined to identify crime

hotspots, analyze regional similarities, and support crime prediction [8]–[10]. Although these approaches improve analytical capabilities, practical challenges include data integration, preprocessing requirements, model interpretation, geographical variability, and reliable validation.

TABLE I: COMPARISON OF EXISTING CRIME PATTERN ANALYSIS METHODS

Method	Main Technique	Strength	Major Limitation
Traditional analysis	Statistical reports	Simple interpretation	Time-consuming
Geographical analysis	Spatial analysis	Identifies crime locations	Limited predictive capability
Hotspot analysis	Crime concentration analysis	Identifies high-crime areas	Sensitive to spatial data
Clustering	K-Means	Groups similar regions	Depends on feature selection
Machine learning	RF, LR, SVM	Automated prediction	Data dependent
Predictive analysis	ML-based prediction	Supports proactive planning	Prediction uncertainty
Integrated analysis	Clustering + classification	Descriptive + predictive analysis	Requires comprehensive validation
Proposed framework	Integrated data science	Regional clustering + crime classification + feature analysis	Requires comprehensive validation

The literature therefore indicates a transition from traditional crime reporting toward automated, predictive, and data-driven crime analysis. However, many existing approaches focus primarily on individual prediction, geographical analysis, or hotspot identification tasks. A comprehensive architecture should integrate crime-data preprocessing, exploratory analysis, regional clustering, classification, feature analysis, and predictive decision support into a unified crime-pattern analysis pipeline.

III. PROPOSED DATA SCIENCE-BASED CRIME PATTERN ANALYSIS AND PREDICTION SYSTEM

The proposed system considers crime analysis as an integrated sequence of interconnected processing stages:



The architecture consists of five major functional layers: data acquisition, preprocessing, exploratory and clustering analysis, classification and prediction, and decision support. The first layer collects historical crime information containing crime categories, states and union territories, yearly observations, and related crime attributes. The preprocessing layer handles missing values, removes inconsistencies, prepares relevant features, and converts the historical records into suitable inputs for subsequent analysis.

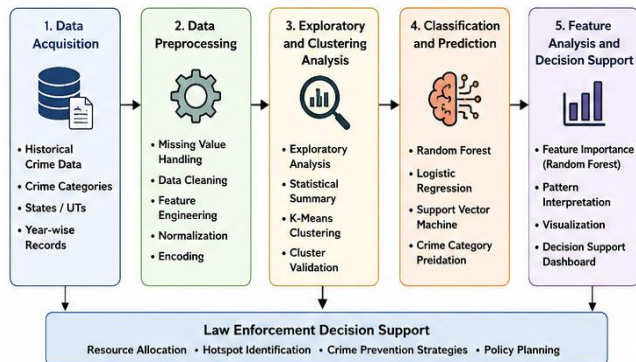


Fig. 1. Overall Architecture of the Proposed Data Science-Based Crime Pattern Analysis and Prediction System

The exploratory and clustering layer analyzes historical crime distributions and applies K-Means clustering to group states and union territories according to similarities in their crime patterns. The resulting clusters provide an analytical representation of regions with relatively high, medium, and low crime incidence. The classification layer applies Random Forest, Logistic Regression, and Support Vector Machine algorithms to analyze historical crime characteristics and predict crime categories.

The feature-analysis component examines the contribution of crime-related attributes to the classification results. Random Forest provides additional information regarding feature importance, supporting the interpretation of factors associated with crime classification. The prediction results can then be used to identify recurring patterns and differences among geographical regions.

Finally, the decision-support layer presents the analyzed patterns, cluster results, classification outcomes, and important crime-related features in an understandable form for policymakers and law enforcement agencies. The architecture is designed as a decision-support framework that assists authorities in understanding crime behavior, allocating resources, identifying priority regions, and developing targeted crime-prevention strategies. This integrated approach enables descriptive analysis, regional pattern discovery, predictive classification, and data-driven law enforcement planning.

IV. CRIME DATA AND CRIME PATTERN REPRESENTATION

A. Crime Data Acquisition and Processing

The proposed system collects historical crime data containing information such as crime categories, states, union territories, yearly crime observations, and related regional attributes. Preprocessing handles missing, duplicate, inconsistent, and abnormal records and prepares the available numerical and categorical information for subsequent analysis.

The structured crime representation considers geographical, temporal, and crime-related attributes. State and union territory

information provides the geographical dimension, while yearly observations represent temporal variations in crime activities. Crime categories provide information about the type of criminal activity recorded within each region.

The collected crime information is analyzed to identify regional variations, recurring trends, and differences among crime categories. Data preprocessing improves the consistency and quality of the historical records and prepares the dataset for clustering and classification. Standardizing numerical attributes can further ensure that variables with different scales receive appropriate consideration during machine-learning analysis.

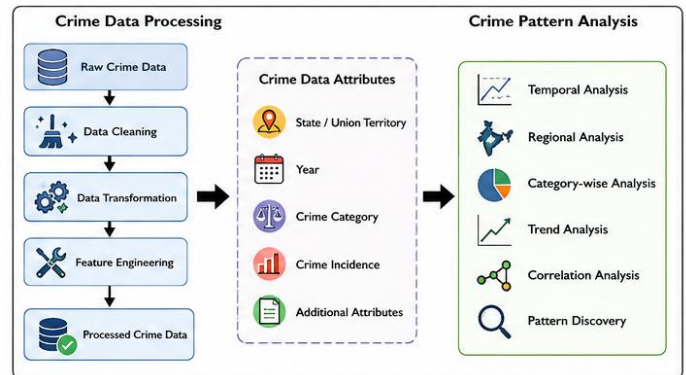


Fig. 2. Crime Data Processing and Crime Pattern Analysis

B. Crime Pattern Representation

The crime pattern of a region is represented using relevant geographical, temporal, and crime-category characteristics. The system extracts:

- State and union territory information
- Year-wise crime observations
- Crime categories
- Crime incidence values
- Regional crime distributions
- Temporal crime variations
- Crime-category frequencies
- Regional crime similarities
- Historical crime trends
- Crime-related predictive features

Representing crime-related information using a unified structure enables the proposed system to analyze relationships between geographical regions, crime categories, and historical variations. The resulting representation can subsequently be provided to the machine-learning models for crime-pattern classification, regional clustering, and crime-category prediction.

The unified crime representation also supports comparison between different states and union territories. Regions with similar historical crime characteristics can be identified through clustering, while differences between crime categories can be examined through classification and exploratory analysis. This provides a structured basis for understanding how crime patterns vary across different geographical and temporal dimensions.

V. CRIME PATTERN-REGIONAL MATCHING AND CRIME RISK SCORING

A central component of the proposed crime-analysis system is regional crime-pattern matching. Unlike simple examination of individual crime counts, the proposed approach identifies relationships between historical crime characteristics and the corresponding geographical regions. The system considers

crime categories, yearly variations, regional crime incidence, and other available crime-related attributes to determine similarities and differences among states and union territories. K-Means clustering is used to identify groups of regions exhibiting similar crime characteristics. The resulting clusters provide an analytical representation of regions according to their observed crime patterns. Based on the characteristics of the resulting groups, regions can be interpreted as relatively high, medium, or low crime-incidence areas.

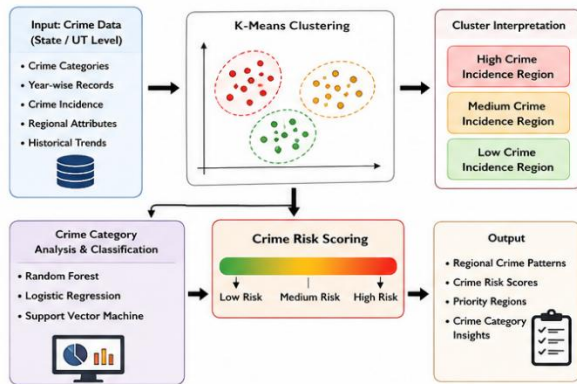


Fig. 3. Regional Crime Pattern Matching and Crime Risk Scoring

Clustering alone cannot determine final crime-analysis decisions because regions with similar overall crime patterns may have substantially higher incidence of specific crime categories. Therefore, the proposed system combines regional clustering with crime-category analysis and supervised classification using Random Forest, Logistic Regression, and Support Vector Machine. These models analyze relationships between crime-related features and crime categories, while Random Forest identifies important features influencing classification results. The combined analysis provides descriptive and predictive insights into regional crime behavior, recurring categories, and historical variations. This information can support law enforcement and policymakers in prioritizing regions, allocating resources, understanding trends, and developing targeted crime-prevention strategies.

VI. EXPLAINABLE CRIME ANALYSIS AND DECISION-AWARE LAW ENFORCEMENT

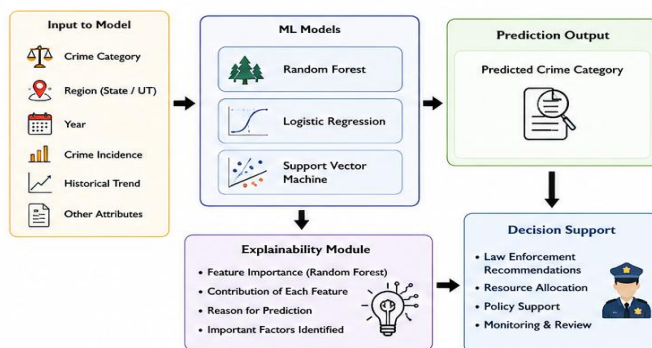


Fig. 4. Explainable Crime Pattern Analysis Framework for Prediction and Decision Support

Machine-learning-based crime analysis should provide more than a predicted crime category or numerical output. Law enforcement agencies and policymakers need to understand why a particular crime pattern or classification result has been generated. Explainable machine-learning techniques can

provide information about the contribution and importance of crime-related features used by the prediction models.

For each prediction, the proposed system can generate an explanation containing:

1. Crime category contribution
2. Regional influence
3. Year-wise crime trend influence
4. Crime incidence contribution
5. Historical pattern influence
6. Important features affecting classification
7. Major factors contributing to the final prediction
8. Suggested law enforcement consideration

For example:

Regional Crime Pattern — High Crime Incidence

- Crime incidence: comparatively high.
- Regional pattern: similar to other high-incidence regions.
- Historical trend: persistent crime activity observed.
- Crime category: major contribution to the classification.
- Important factor: historical crime incidence.
- Cluster assignment: high-incidence regional group.
- Prediction: region classified according to historical crime characteristics.
- Recommendation: prioritize monitoring, resource allocation, and targeted prevention measures.

Such explanations help law enforcement authorities understand whether the generated classification is supported by meaningful historical crime characteristics rather than relying only on the final prediction result.

A. Explainability and Decision Monitoring

Crime classification can be influenced by multiple geographical, temporal, and crime-related attributes. Therefore, the system should examine the contribution of individual features to the prediction results. Feature-importance analysis, particularly from the Random Forest model, can help identify which crime-related attributes have greater influence on the classification outcome.

The proposed framework therefore integrates explainability as a decision-support component. The system should report both: Performance = How accurately does the system classify the crime pattern or category?

Explainability = Can the law enforcement authority understand which crime-related features influenced the prediction?

This dual objective is important because a highly accurate model without understandable reasoning may be difficult to interpret and apply in practical crime-analysis environments.

VII. ADAPTIVE CRIME INTELLIGENCE AND HUMAN-IN-THE-LOOP DECISION SUPPORT

Crime patterns are not static. Crime incidence, crime categories, geographical distributions, and yearly trends can change over time. A crime-classification model trained using historical conditions may therefore become less effective when new crime patterns emerge or regional characteristics change.

Model Update

Adaptive mechanisms can update crime-pattern information, feature distributions, regional clusters, prediction knowledge, and analytical results while maintaining appropriate model version control.

The system can periodically incorporate newly available crime records to improve its understanding of changing crime

conditions. However, updated models should not automatically determine critical law enforcement actions without validation. New model versions should first be evaluated using historical and newly collected crime datasets before being incorporated into the operational analysis framework.

Human-in-the-Loop Workflow

The final crime-analysis decision-support workflow is:

Crime Data Collection → Data Preprocessing → Crime Pattern Analysis → K-Means Clustering → Classification → Prediction → Explanation → Law Enforcement Verification → Final Decision

Law enforcement personnel and policymakers remain responsible for the final decision, particularly when unusual crime patterns are identified or when additional contextual information is required that may not be available in the historical dataset.

This human-in-the-loop design allows authorities to combine data-driven crime predictions with practical law enforcement knowledge and current regional observations. Human feedback and newly available crime information can also support future system improvement while preventing the analytical system from making completely autonomous decisions without appropriate human supervision.

VIII. EVALUATION FRAMEWORK

A comprehensive evaluation framework is adopted to assess the proposed crime-pattern analysis system across classification performance, clustering quality, prediction capability, feature importance, computational efficiency, and robustness. Rather than relying on a single performance measure, the evaluation considers multiple complementary dimensions to determine the practical effectiveness of the proposed framework.

A. Classification Performance

The classification module is evaluated using accuracy, precision, recall, and F1-score. Accuracy measures overall classification correctness, while precision indicates the proportion of correctly predicted instances among the predicted category. Recall measures the ability to identify instances belonging to the appropriate crime category, and F1-score provides a balanced assessment of precision and recall. These metrics provide a comprehensive evaluation of Random Forest, Logistic Regression, and Support Vector Machine models.

B. Clustering Performance

The K-Means clustering component is evaluated according to the quality and separation of the identified regional groups. Cluster analysis examines whether states and union territories with similar historical crime characteristics are appropriately grouped together. The resulting clusters can be analyzed to identify relatively high, medium, and low crime-incidence regions and to determine whether meaningful regional patterns are obtained.

C. Crime Pattern Prediction

The prediction component is evaluated by examining the ability of the machine-learning models to identify crime categories from historical crime characteristics. The performance of Random Forest, Logistic Regression, and Support Vector Machine is compared to determine which approach provides the most reliable classification results. The evaluation also considers differences in performance across crime categories and regional groups.

D. Feature Importance and Explainability

The explainability component evaluates the importance of crime-related features influencing model predictions. Random Forest feature-importance analysis can identify the attributes contributing most strongly to classification results. The analysis helps determine whether regional information, yearly trends, crime incidence, or specific crime categories have greater influence on the predicted outcome.

E. Regional and Temporal Analysis

Regional and temporal evaluation examines variations in crime patterns across Indian states and union territories and across different years. The analysis determines whether the proposed framework can identify recurring trends, changes in crime incidence, and differences between geographical regions. This evaluation provides additional descriptive information beyond classification performance.

F. Efficiency and Robustness

Computational evaluation considers data-processing time, model-training time, prediction time, memory requirements, and scalability. Robustness is evaluated using different subsets of historical crime records, crime categories, regional groups, and temporal periods. This ensures that the proposed system can maintain reliable performance when crime data characteristics change.

TABLE II: EVALUATION DIMENSIONS OF THE PROPOSED CRIME PATTERN ANALYSIS FRAMEWORK

Dimension	Evaluation Metrics / Criteria	Purpose
Classification	Accuracy, Precision, Recall, F1-Score	Evaluate crime-category classification performance
Clustering	Cluster separation, Regional grouping, Cluster interpretation	Evaluate regional crime-pattern grouping
Prediction	Classification performance, Category-wise evaluation	Assess crime-category prediction capability
Explainability	Feature importance, Prediction interpretation	Evaluate decision transparency
Regional Analysis	State-wise and union-territory-wise comparison	Identify geographical crime patterns
Temporal Analysis	Year-wise crime trends and variations	Analyze changes in crime patterns over time
Efficiency	Processing Time, Training Time, Prediction Time, Memory	Assess computational efficiency
Robustness	Dataset variation, Category variation, Regional variation	Evaluate generalization capability

IX. RESEARCH CHALLENGES AND FUTURE ROADMAP

Although data science and machine-learning techniques can enhance crime-pattern analysis and support evidence-based law-enforcement strategies, several challenges must be addressed to develop reliable, explainable, and practically deployable crime-analysis systems.

A. Dataset Availability:

Crime datasets often contain incomplete records, inconsistent classifications, regional reporting differences, and limited historical observations. Reliable analysis requires diverse, representative, and high-quality datasets covering different regions, crime categories, and time periods.

B. Regional Crime Pattern Matching:

Crime characteristics vary considerably across states and union territories. Regions with similar crime incidence may exhibit different crime-category distributions and temporal patterns. Effective regional matching should integrate geographical

information, crime categories, historical incidence, and temporal characteristics.

C. Prediction Reliability:

Crime prediction is influenced by geographical, temporal, and crime-specific factors. Variations in class distributions, regional characteristics, and historical observations can reduce predictive reliability. Robust feature engineering, appropriate model selection, and comprehensive evaluation are therefore essential.

D. Explainability:

AI-generated classifications and predictions should provide understandable explanations. Identifying influential crime categories, regional characteristics, historical trends, and important features can improve transparency and support informed decision-making.

E. Bias and Fairness:

Historical crime records may reflect reporting disparities, regional imbalances, and incomplete observations. Bias-aware evaluation and balanced datasets are necessary to reduce unfair patterns and ensure that model outputs are supported by relevant evidence.

F. Crime Pattern Evolution:

Crime patterns continuously change because of geographical, social, economic, and environmental factors. Future systems should incorporate adaptive learning, updated datasets, evolving feature representations, and periodic model validation.

G. Human-AI Interaction:

AI should support rather than replace expert judgement. Human-centered interfaces should enable law-enforcement personnel to review, interpret, and validate AI-generated results using contextual knowledge.

H. Privacy and Security:

Crime datasets can contain sensitive information. Secure storage, controlled access, anonymization, privacy-preserving processing, and appropriate cybersecurity mechanisms are essential for responsible deployment.

Research Roadmap:

Short Term: Improve preprocessing, regional matching, K-Means clustering, classification, explainability, feature importance, and evaluation.

Medium Term: Develop adaptive clustering, spatio-temporal analysis, robust prediction, diverse datasets, bias-aware evaluation, and cross-regional validation.

Long Term: Establish human-centered crime intelligence integrating explainable prediction, adaptive learning, fairness, privacy preservation, and human-in-the-loop decision support.

X. CONCLUSION

This paper presents a data science framework for analyzing crime patterns across Indian states and union territories by integrating data preprocessing, exploratory analysis, K-Means clustering, supervised machine learning, feature-importance analysis, and decision support. The framework identifies geographical, temporal, and crime-related patterns, grouping regions with similar historical characteristics and classifying crime categories using Random Forest, Logistic Regression, and Support Vector Machine. Feature importance improves interpretability, while human-in-the-loop design ensures AI insights support rather than replace human judgement. Future research should address diverse datasets, spatio-temporal analysis, clustering stability, fairness, privacy, explainability, and reliable decision support for effective crime prevention.

REFERENCES

- [1] V. Mandalapu, L. Elluri, P. Vyas, and N. Roy, "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," *IEEE Access*, vol. 11, pp. 60153–60170, 2023, doi: 10.1109/ACCESS.2023.3286344.
- [2] J. Jenga, C. Catal, and G. Kar, "Machine Learning in Crime Prediction," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 2887–2913, 2023, doi: 10.1007/s12652-023-04530-y.
- [3] J. Yin, "Crime Prediction Methods Based on Machine Learning: A Survey," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 4601–4629, 2023, doi: 10.32604/cmc.2023.034190.
- [4] A. A. Alsubayhin, B. Alzahrani, and M. S. Ramzan, "Review of Crime Prediction Through Machine Learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 5s, pp. 273–281, 2024.
- [5] R. M. Aziz, P. Sharma, and A. Hussain, "Machine Learning Algorithms for Crime Prediction under Indian Penal Code," *Annals of Data Science*, vol. 11, no. 1, pp. 379–410, 2024, doi: 10.1007/s40745-022-00424-6.
- [6] S. R. Bandekar and C. Vijayalakshmi, "Design and Analysis of Machine Learning Algorithms for the Reduction of Crime Rates in India," *Procedia Computer Science*, vol. 172, pp. 122–127, 2020, doi: 10.1016/j.procs.2020.05.018.
- [7] N. Shah, N. Bhagat, and M. Shah, "Crime Forecasting: A Machine Learning and Computer Vision Approach to Crime Prediction and Prevention," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, Art. no. 9, 2021, doi: 10.1186/s42492-021-00075-z.
- [8] R. Ahmad, A. Nawaz, G. Mustafa, T. Ali, M. Tlija, M. A. El-Meligy, and Z. Ahmed, "CHART: Intelligent Crime Hotspot Detection and Real-Time Tracking Using Machine Learning," *Computers, Materials & Continua*, vol. 81, no. 3, pp. 4171–4194, 2024, doi: 10.32604/cmc.2024.056971.
- [9] J. Alghamdi and T. Al-Dala'in, "Towards Spatio-Temporal Crime Events Prediction," *Multimedia Tools and Applications*, vol. 83, pp. 18721–18737, 2024, doi: 10.1007/s11042-023-16188-x.
- [10] "Crime Risk Prediction Incorporating Geographical Spatiotemporal Dependency into Machine Learning Models," *Information Sciences*, vol. 646, Art. no. 119414, 2023, doi: 10.1016/j.ins.2023.119414.
- [11] M. Kaur and M. Saini, "Role of Artificial Intelligence in the Crime Prediction and Pattern Analysis Studies Published over the Last Decade: A Scientometric Analysis," *Artificial Intelligence Review*, vol. 57, Art. no. 202, 2024, doi: 10.1007/s10462-024-10823-1.
- [12] "Artificial Intelligence & Crime Prediction: A Systematic Literature Review," *Social Sciences & Humanities Open*, vol. 6, no. 1, Art. no. 100342, 2022, doi: 10.1016/j.ssaho.2022.100342.
- [13] K. Taha, "Empirical and Experimental Insights into Data Mining Techniques for Crime Prediction: A Comprehensive Survey," *ACM Transactions on Intelligent Systems and*

Technology, vol. 16, no. 2, Art. no. 26, 2025, doi: 10.1145/3699515.

[14] O. Esan and B. Chimbo, "Systematic Literature Review on Crime Prediction Using Machine Learning Techniques," *The Indonesian Journal of Computer Science*, vol. 14, no. 3, 2025, doi: 10.33022/ijcs.v14i3.4881.

[15] A. Sarkar, A. K. Mishra, U. P. Singh, S. Hassan, G. S. Munda, and P. Singh, "Real-Time Crime Prediction in India Using Machine Learning," *International Journal for Research in Applied Science & Engineering Technology*, 2025, doi: 10.22214/ijraset.2025.73053.