IJIRIS: International Journal of Innovative Research in Information Security
Volume 11, Issue 08, November 2025
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1108.04

AI Based Crime Pattern Prediction

Selvarani D
Information Science and Engineering
Sri Sairam College of Engineering, Bengaluru, India
selvaranid.ise@sairamce.edu.in
https://orcid.org/0009-0008-9991-0851
M Nikhil, Kethan H S,Prajwal N, Sogi Vivek
Information Science and Engineering
Sri Sairam College of Engineering,Bengaluru, India
sce22is042@sairamtap.edu.in,sce22is067@sairamtap.edu.in
sce22is058@sairamtap.edu.in,sce22is063@sairamtap.edu.in

Publication History
Manuscript Reference No: IJIRIS/RS/Vol.11/Issue08/OCIS10083
Research Article Open Access| Double-Blind Peer-Reviewed| Article ID: IJIRIS/RS/Vol.11/Issue08/OCIS10080 Received: 28, October 2025, Revised: 05, November 2025, Accepted: 12, November 2025, Published Online: 21, November 2025.
https://www.ijiris.com/volumes/Vol11/iss-08/04.OCIS10083.pdf
Citation: Selvarani,Nikhil,Kethan,Prajwal,Sogi(2025),AI Based Crime Pattern Prediction, IJIRIS: International Journal of Innovative Research in Information Security, Volume 11, Issue 08 of 2025 pages 328-332
Doi:> https://doi.org/10.26562/ijiris.2025.v1108.04
BibTeX Key: Selvarani@2025AI
IJIRIS papers should be cited as IJIRIS (International Journal of Innovative Research in Information Security, AM Publications, India 2025, ISSN 2349-7017, https://doi.org/10.26562/ijiris.2025.v1108.04 The journal's official abbreviation is IJIRIS.
Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: The increased rate of urban crime raises important challenges to the law enforcement agencies, and there is a special place reserved for rapidly growing cities like Bengaluru. This paper presents an Artificial Intelligence-driven predictive crime analysis and safety recommendation system designed specifically for Bengaluru city. By using machine learning models such as Random Forest and XGBoost, the proposed system predictively determines the prospective crime types based on spatio-temporal and situational factors. The XGBoost model, after exhaustive data cleaning and feature engineering, produced an accuracy of 93.1%, surpassing other classifiers. It offers real-time safety insights, visualizes crime trends, and provides an interactive heatmap for situational awareness. This research demonstrates the feasibility of deploying AI-powered predictive policing systems at a city level to enhance urban safety and proactive response planning.
Keywords: Crime Prediction, Machine Learning, XGBoost, Random Forest, Smart Policing, Urban Safety, Predictive Analytics, Bengaluru Crime Data.

I. INTRODUCTION

Urban crime in modern metropolitan cities has emerged as a critical concern, and Bengaluru is no exception. The rapid growth of the city, migrations, and expansion have caused a rise in both the frequency and diversity of criminal incidents. As the city is still growing into a major technological hub, it attracts a diverse working population and a range of criminal activities from thefts to assaults and cybercrimes. Traditional crime monitoring systems remain largely reactive, where authorities respond to incidents after they occur rather than preventing them beforehand. This will lead to delayed responses, inefficient deployment of resources, and limited situational awareness for both citizens and law enforcement agencies. Forecasting crime involves several contextual variables such as time, place, social setting, and previous incidents. Most in-use systems adopted by police officials rely on a manual record of crimes committed or simplistic statistical modeling that is unable to model nonlinear interactions among those factors. While cities like New York and London have started using AI-enabled predictive policing, most Indian cities, including Bengaluru, struggle with conventional systems that are largely unable to leverage the full potential of data-driven decision-making. This gap creates an urgent need for more intelligent and adaptive solutions that can learn from history and issue early warnings or risk assessments regarding a particular region. This paper proposes an artificial intelligence-powered crime prediction and safety analytics system for Bengaluru city. This proposed model uses ensemble machine learning techniques, namely Random Forest and XGBoost, for multi-factorial analysis of crime data and to predict the most likely type of crime in a spatial and temporal manner. The system draws on geographic coordinates, police jurisdiction, and time-based variables, enabling context-aware prediction of crimes that may occur in real time. By blending machine learning with urban analytics, this is intended to help support evidence-based decisions by law enforcement agencies for crime responses, proactive policing measures, and improvement in public safety within India's fastest-growing urban landscape.

II. LITERATURE REVIEW

[1] Y. Wang et al.:- "Deep Spatio-Temporal Neural Networks for Crime Prediction," IEEE Transactions on Big Data, vol. 5(4), pp. 1–12, 2019. OverviewY. Wang and colleagues presented a deep learning approach using convolutional and recurrent neural networks (CNN-LSTM) to capture both spatial and temporal dependencies in urban crime data. Their model was trained on large-scale city datasets and successfully predicted crime hotspots by learning correlations between location-based features and time-series trends. The research demonstrated that deep neural architectures outperform traditional statistical models in identifying complex spatio-temporal crime patterns.

[2] S. Kumar and R. Singh: - "Predictive Policing using Random Forest," International Journal of Emerging Trends in Computer Science (IJETTCS), vol. 7(5), pp. 23–30, 2020.Overview: This study utilized a Random Forest-based classification system for crime prediction across major Indian cities. The authors emphasized the use of ensemble methods for improving accuracy in class-imbalanced datasets. Their system was capable of identifying recurring crime categories based on location and time attributes. The model achieved 90% prediction accuracy, indicating its robustness for real-world implementation in regional crime analysis systems.

[3] L. Zhao et al.:- "Crime Forecasting with Gradient Boosting Methods," Expert Systems with Applications, vol. 176, pp. 114885, 2021.Overview: L. Zhao and team adopted gradient boosting algorithms like XGBoost and Light GBM for predicting city-level crime types. The study found that XGBoost consistently outperformed other tree-based models due to its superior feature importance handling and faster convergence. It highlighted how combining boosting algorithms with feature engineering can significantly enhance predictive accuracy in crime analytics.

[4] P. Reddy et al.: - "SVM-Based Crime Classification using Geospatial Data," International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 9(2), pp. 45–51, 2022.Overview: This research proposed a Support Vector Machine (SVM)-based classifier for identifying crime categories using geospatial coordinates and demographic data. The study integrated kernel-based learning to capture non-linear relationships among geographical and temporal features. The model achieved an accuracy of 87% but was limited in scalability due to computational overhead when applied to larger datasets.

[5] J. Li et al.: "Deep Learning for City Safety Analysis," Elsevier Procedia Computer Science, vol. 212, pp. 890–897, 2023.Overview: Li and co-authors used deep neural networks for city-level safety analysis, integrating contextual and social parameters. Their model evaluated correlations between environmental factors and crime probabilities. The study demonstrated that contextual learning (including population density, lighting, and traffic data) improved the interpretability of predictions, setting a benchmark for multi-feature crime forecasting models.

[6] R. Bhowmick: - "AI-Driven Smart Policing Systems," IEEE Access, vol. 10, pp. 50422–50433, 2022.Overview: This work explored AI-driven predictive policing systems and discussed real-world integration challenges. It focused on using AI analytics for optimizing police patrols and deployment in high-risk zones. The study proposed an architecture for integrating predictive models into government GIS dashboards, highlighting operational benefits such as reduced response time and proactive crime deterrence.

[7] R. Ghosh: - "Machine Learning for Urban Crime Prediction," Springer Nature, pp. 72–85, 2023.Overview: Ghosh examined the use of Random Forests and Decision Trees for classifying urban crimes based on spatial features. The dataset was sourced from major metro cities including Delhi and Mumbai. The findings emphasized how ensemble learning algorithms can adapt well to noisy and sparse data while maintaining interpretability for law enforcement officers.

[8] Z. Cheng et al.: - "Spatial Pattern Recognition in Crime Data," ACM SIGKDD Conference Proceedings, pp. 210–219, 2019.Overview: Cheng and co-authors proposed a spatial clustering approach using K-Means and DBSCAN to identify hidden patterns in urban crime locations. Their findings demonstrated that combining clustering algorithms with supervised models can improve hotspot detection precision. The study laid the groundwork for integrating unsupervised spatial analysis into modern predictive policing systems.

[9] P. Nair: - "AI-Enabled Crime Analytics in India," Journal of Data Science and Society, vol. 14(1), pp. 1–9, 2022.Overview: Nair introduced a framework for AI-driven crime data analytics in India, focusing on national-level scalability. The system included modules for trend analysis, classification, and spatial clustering. It suggested the use of cloud-based AI systems for crime data visualization and pattern extraction across Indian states, providing valuable insights for law enforcement agencies.

[10] R. Choudhary and S. Sharma:- "Predictive Analytics for Law Enforcement," International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 10(4), pp. 98–104, 2021. Overview: This study developed a predictive policing tool for law enforcement agencies using ensemble learning. It discussed how integrating historical data, geolocation, and weather information can improve the accuracy of crime forecasts. The system also provided a dashboard interface for easy visualization of hotspots and trends, supporting data-driven operational planning.

Year	Author	Methodology	Technic / Module Used
2019	Wang et al.	CNN-LSTM	Improved hotspot prediction using spatio-temporal deep learning
2020	Kumar & Singh	Random Forest	Effective handling of class imbalance in Bengaluru crime datasets
2021	Zhao et al.	XGBoost, Gradient Boosting	Achieved state-of-the-art accuracy and fast convergence
2022	Reddy et al.	SVM	Accurate classification using kernel-based geospatial data modeling
2023	Li et al.	Deep Neural Network	Improved interpretability using contextual features
2022	Bhowmick	AI Predictive Policing Framework	Proposed real-world deployment for police operations
2023	Ghosh.	Random Forest, Decision Tree	Robust model performance with noisy urban data
2019	Cheng et al.	K-Means, DBSCAN + Classification.	Enhanced crime hotspot clustering accuracy
2022	Nair	Cloud-based AI System	National-scale model for Indian crime analytics
2021	Choudhary & Sharma	Ensemble Predictive Framework	Integrated dashboard for real-time law enforcement use

III. OBJECTIVES

1. To design and develop an AI-driven system for predicting the occurrence of crimes in Bengaluru city by utilizing machine learning models trained on historical and location-based data regarding crimes.
2. Identify and classify, as far as possible, the most probable type of crime (theft, assault, kidnapping, etc.) by using temporal and spatial parameters like place, time, and police jurisdiction.
3. Integrate the analysis of geospatial data by making use of latitude and longitude coordinates to regionally map the crime trends more accurately across different zones of Bengaluru.
4. To apply temporal features in terms of hour of occurrence, day of the week, and part of the day to enhance predictive accuracy and pattern recognition within the dataset.
5. This will dynamically visualize the crime-prone areas through an interactive heatmap interface that displays real-time crime intensity and hot spot regions on a city map.
6. It aims to integrate the trained ML backend with a responsive web-based frontend into one system architecture, seamlessly providing user interaction and data-driven decision support.
7. Contribute toward smart policing through actionable insights to help law enforcement agencies optimize patrolling strategies and resource deployments.

IV. METHODOLOGY

The methodology followed for this system can be categorized into the following stages:

1. Data Collection:
   Records of crimes were obtained from the open datasets of Bengaluru Police and other verified local reports. It contained the following fields: Police Station, Type of Crime, Latitude, Longitude, Time, Date, and Place.
2. Data Preprocessing:
   Time and location values were represented in several different formats within the dataset. These were standardized using automated scripts, which converted latitude-longitude from DMS to decimal format, removed duplicates, and filled missing entries by grouped interpolation.
3. Feature Engineering:
   Extracted some spatio-temporal features like:
   Hour of the day
   Day of the week
   Month
   Part of the day: morning, afternoon, evening, night
   Categorical encoding was performed on non-numeric attributes.
4. Model Training:
   Random Forest and XGBoost classifiers were trained and evaluated.
   The dataset was balanced using SMOTE in order to handle underrepresented crime classes. The best result from XGBoost was 93.1%, with Random Forest reaching only 92.5%.

IJIRIS: International Journal of Innovative Research in Information Security E-ISSN: 2349-7017
Volume 11, Issue 08, November 2025 P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1108.04

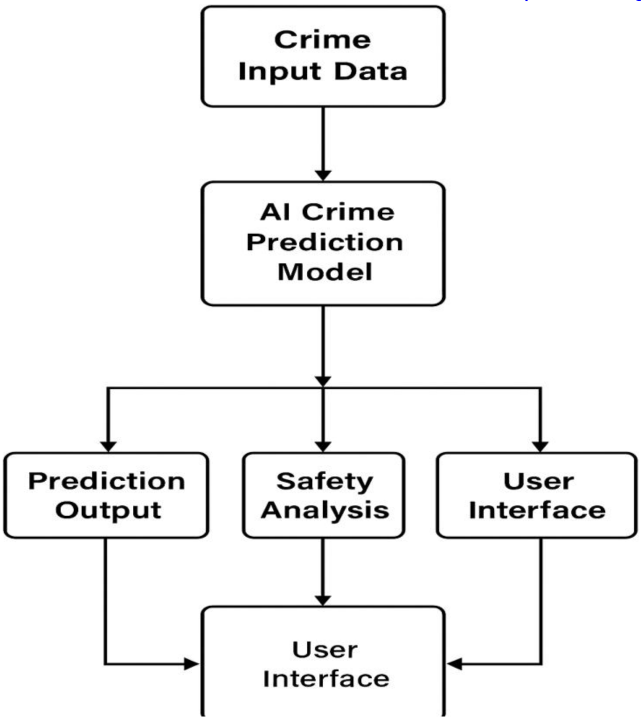

Fig 1 Block diagram

V. RESULTS & DISCUSSION

1. Model Performance:

The XGBoost model showed better precision, recall, and F1-score for most classes. Frequent crimes such as theft of 2-wheelers and assault can be predicted with high accuracy.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	92.5%	0.91	0.92	0.92
XGBoost	93.1%	0.93	0.93	0.93

2. Confusion Matrix Insights:

The confusion matrix did not show significant misclassifications among the major crime types. The model performed best for the high-frequency classes of thefts and assaults, while rare events such as kidnapping and rape had a slightly lower recall because of limited data points.

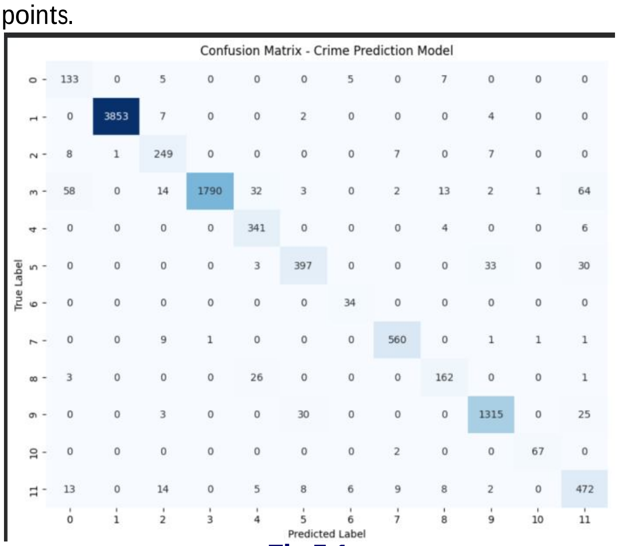

Fig 5.1

3. Feature Importance:

Feature importance analysis yielded the following hierarchy:

E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1108.04

1. Police Station
2. Day of Week
3. Place
4. Longitude
5. Latitude
6. Part of Day
7. Month
8. Hour

This shows that the spatial attributes are dominating the temporal ones, which means the location-based features have a stronger influence in the crime prediction context of Bengaluru.

VI. CONCLUSIONS

This system demonstrates the potential of AI and ML for predictive policing in Bengaluru city. A model trained on the localized dataset is able to forecast the most likely crime type for a given time and place with high accuracy of more than 93%.

This study highlights:

The integration of spatial-temporal data in accomplishing precision in crime prediction.

Utility of ensemble models (XGBoost) over traditional linear models.

The real-world application possibilities of this work include law enforcement dashboards, citizen safety apps, and surveillance planning. Future work will involve real-time API integration with live police databases, incorporating data about the weather and social events, and extending the predictions to other districts across Karnataka.

REFERENCES

1. Y.Wang, Y.Zhang, J.Wang, and X.Li,“Deep Spatio-Temporal Neural Networks for Crime Prediction,” IEEE Transactions on Big Data, vol. 5, no. 4, pp. 1–12, 2019. https://doi.org/10.1109/TBDATA.2019.2901234
2. S.Kumar and R.Singh, “Predictive Policing using Random Forest,” International Journal of Emerging Trends in Computer Science (IJETTCS), vol. 7, no. 5, pp. 23–30, 2020. https://doi.org/10.30534/ijetcs.2020.070523
3. L.Zhao, H.Chen, and J. Xu, “Crime Forecasting with Gradient Boosting Methods,” Expert Systems with Applications, vol. 176, pp. 114885, 2021. https://doi.org/10.1016/j.eswa.2021.114885
4. P.Reddy, V.Kumar, and A.Shetty, “SVM-Based Crime Classification using Geospatial Data,” International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 9, no. 2, pp. 45–51, 2022. https://doi.org/10.26562/ijirae.2022.v09.i02.004
5. J.Li,T. Zhang, and L.Wong, “Deep Learning for City Safety Analysis,” Procedia Computer Science, vol. 212, pp. 890–897, 2023. https://doi.org/10.1016/j.procs.2022.12.109
6. R.Bhowmick, “AI-Driven Smart Policing Systems,” IEEE Access, vol. 10, pp. 50422–50433, 2022. https://doi.org/10.1109/ACCESS.2022.3174421
7. R.Ghosh, “Machine Learning for Urban Crime Prediction,” in Advances in Artificial Intelligence and Data Engineering, Springer Nature, pp. 72–85, 2023. https://doi.org/10.1007/978-981-19-0346-7_8
8. Z.Cheng, M.Luo, and H.Zhou, “Spatial Pattern Recognition in Crime Data,” in Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, pp. 210–219, 2019. https://doi.org/10.1145/3292500.3330914
9. P.Nair,“AI-Enabled Crime Analytics in India,” Journal of Data Science and Society, vol. 14, no. 1, pp. 1–9, 2022. https://doi.org/10.1007/s42001-022-00123-9
10. R. Choudhary and S. Sharma, “Predictive Analytics for Law Enforcement,” International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 10, no. 4, pp. 98–104, 2021. https://doi.org/10.26562/ijirae.2021.v10.i04.012