IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

Forecasting Robbery Tendencies and Crime Prediction with AI Insights

Koppula Vineetha[1] and Swathi Terli[2]
Department of MCA, Sir C. R. Reddy College (Autonomous), Eluru[1]
M.Tech (Ph.D), Assistant Professor, Department of MCA, Sir C. R. Reddy College (Autonomous), Eluru[2]

Abstract: *Crime prediction has emerged as one of the most promising applications of Artificial Intelligence (AI) and Machine Learning (ML) in modern law enforcement and public safety. The rapid growth of urban populations and technological advancements have increased the complexity of criminal activities, making traditional crime analysis methods insufficient for proactive policing. This research presents an intelligent crime prediction system, Crime Sense AI, which forecasts robbery tendencies and other crime categories using Machine Learning techniques.*

The proposed system employs the Random Forest Classification algorithm to predict crime types based on multiple influencing variables such as city, locality, area type, weather conditions, day of the week, and time of occurrence. The application generates crime risk scores, visualizes crime hotspots through interactive maps, and provides model insights using statistical evaluation techniques. Furthermore, the system incorporates an Emergency SOS module that allows users to share their live location with emergency contacts during unsafe situations.

A synthetic dataset calibrated using National Crime Records Bureau (NCRB) statistics was developed to overcome the unavailability of public incident-level crime data in India. Experimental evaluation demonstrates that the Random Forest model effectively captures crime patterns and achieves satisfactory prediction accuracy. The study highlights the role of Artificial Intelligence in supporting law enforcement agencies, improving public awareness, and enabling smart-city crime prevention strategies.

Keywords: Artificial Intelligence, Machine Learning, Crime Prediction, Random Forest, Robbery Forecasting, Flask, Crime Analytics

I. INTRODUCTION

Crime is one of the most critical challenges faced by modern societies. Urbanization, population growth, unemployment, and economic inequalities contribute significantly to the increase in crimes such as robbery, burglary, theft, and assault. Traditional policing methods largely depend on historical crime investigations, eyewitness reports, and manual analysis of crime records. While these methods remain important, they provide limited support for predicting future criminal activities.

Artificial Intelligence has transformed several sectors, including healthcare, education, agriculture, transportation, and finance. Recently, AI has gained significant importance in crime analytics due to its ability to identify complex relationships hidden within large datasets. Machine Learning algorithms can discover patterns from historical crime records and estimate the probability of crimes occurring under similar conditions.

The present research develops an AI-enabled web application named **CrimeSense AI**, capable of forecasting robbery tendencies based on spatial, temporal, and environmental factors. The system integrates machine learning prediction with crime hotspot visualization, statistical dashboards, and emergency safety mechanisms. Unlike traditional crime reporting systems, the proposed application focuses on preventive decision-making by informing users about potential crime risks before an incident occurs.

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

The application has been implemented using Python, Flask, SQLite, and Scikit-learn. Random Forest Classification is employed because of its robustness, high prediction accuracy, and capability to handle multiple input variables efficiently.

II. REVIEW OF LITERATURE

Crime prediction has been extensively studied using statistical analysis and machine learning approaches. Early research relied on regression analysis and time-series forecasting, whereas modern studies increasingly employ Artificial Intelligence techniques.

Breiman (2001) introduced the Random Forest algorithm, demonstrating that ensemble learning significantly improves classification accuracy while minimizing overfitting. Due to its stability and interpretability, Random Forest has become one of the most widely adopted algorithms in predictive analytics.

Pedregosa et al. (2011) developed the Scikit-learn framework, which provides efficient implementations of machine learning algorithms and evaluation techniques. The framework has become an industry standard for supervised learning applications.

National Crime Records Bureau (NCRB) reports provide comprehensive statistical information regarding crime patterns across Indian states and metropolitan cities. However, these reports mainly present aggregated statistics rather than incident-level predictive datasets.

Several researchers have combined Geographic Information Systems (GIS) with crime prediction to identify hotspot regions requiring additional police surveillance. Recent studies have also explored the integration of weather, population density, traffic conditions, and socio-economic indicators to improve crime forecasting.

Despite these advancements, existing systems generally focus on prediction alone and seldom integrate user-oriented safety mechanisms such as Emergency SOS alerts or interactive crime dashboards. The present study attempts to bridge these limitations through a comprehensive AI-enabled platform.

III. RESEARCH METHODOLOGY

The proposed research follows a quantitative machine learning methodology involving data collection, preprocessing, model training, prediction, and evaluation.

Since incident-level crime data with geographical coordinates are not publicly available in India, a synthetic dataset containing approximately 6,000 records was generated. The dataset includes:

The generated dataset was calibrated using official NCRB city-level crime statistics to preserve realistic crime distributions.

Data preprocessing involved cleaning missing values, encoding categorical variables, extracting temporal features, and splitting the dataset into training and testing subsets.

The Random Forest Classifier was selected due to its ensemble-learning capability, resistance to overfitting, and superior performance with mixed categorical and numerical variables.

S.No	City	Locality	Area Type	Weather	Date	Time	Day	Crime Type	Risk Score	Risk Level	Latitude	Longitude
1	Eluru	Ashok Nagar	Residential	Cloudy	2026-01-12	21:30	Monday	Robbery	86	High	16.7105	81.0952
2	Eluru	Powerpet	Commercial	Clear	2026-	22:15	Sunday	Theft	78	High	16.7132	81.1036

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

					01-18							
3	Eluru	Fire Station Centre	Commercial	Rainy	2026-02-02	20:45	Tuesday	Robbery	82	High	16.7109	81.1015
4	Eluru	Ramachandra Rao Pet	Residential	Cloudy	2026-02-14	19:20	Saturday	Burglary	65	Medium	16.7165	81.0912
5	Eluru	Sanivarapupeta	Commercial	Clear	2026-03-06	23:10	Friday	Robbery	91	High	16.7097	81.1008
6	Eluru	Tangellamudi	Residential	Foggy	2026-03-21	05:50	Saturday	Theft	58	Medium	16.7213	81.0784
7	Eluru	Satrampadu	Residential	Cloudy	2026-04-03	18:15	Friday	Assault	61	Medium	16.7302	81.0864
8	Eluru	Chataparru	Industrial	Sunny	2026-04-18	16:45	Saturday	Theft	55	Medium	16.6897	81.0705
9	Eluru	Ponangi	Residential	Rainy	2026-05-07	22:05	Thursday	Burglary	72	High	16.6968	81.1121
10	Eluru	Vatluru	Residential	Cloudy	2026-05-26	21:50	Tuesday	Robbery	88	High	16.7015	81.1183
11	Eluru	Mini Bypass Road	Commercial	Clear	2026-06-09	23:40	Tuesday	Robbery	90	High	16.7138	81.1048
12	Eluru	Railway Station Area	Commercial	Rainy	2026-06-24	20:20	Wednesday	Theft	75	High	16.7108	81.1055

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

13	Eluru	Duggirala	Residential	Sunny	2026-07-04	17:10	Saturday	Assault	49	Medium	16.7056	81.0847
14	Eluru	Narasimharao Pet	Residential	Cloudy	2026-07-19	21:00	Sunday	Burglary	69	High	16.7140	81.0938
15	Eluru	Gavaravaram	Residential	Rainy	2026-08-01	22:30	Saturday	Robbery	92	High	16.7262	81.0816

IV. SYSTEM DESIGN

The proposed CrimeSense AI system follows a three-tier architecture consisting of the Presentation Layer, Application Layer, and Data Layer. The front-end is developed using HTML5, CSS3, JavaScript, Chart.js, and Leaflet.js for visualization. The backend is implemented using Flask, which manages user authentication, prediction requests, dashboard generation, and Emergency SOS services. SQLite serves as the database for storing user accounts, prediction history, emergency contacts, and SOS alerts.

The trained Random Forest model predicts crime types and calculates a confidence-based crime risk score.

V. MACHINE LEARNING MODEL

The Random Forest algorithm constructs multiple decision trees using random subsets of training data.

Each tree independently predicts a crime category.

The final prediction is determined through majority voting.

Advantages include:

High classification accuracy

Reduced overfitting

Robust performance

Automatic feature importance estimation

Efficient handling of categorical variables

Feature importance analysis revealed that Area Type, Time of Day, and Locality contribute most significantly to crime prediction.

VI. RESULTS AND DISCUSSION

The developed system successfully predicts crime categories using historical crime patterns.

Performance evaluation employed:

Accuracy

Precision

Recall

F1-score

Confusion Matrix

The trained Random Forest model achieved approximately **60% prediction accuracy**, which is substantially better than the random baseline for a five-class classification problem.

Interactive dashboards display:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

Crime distribution
City-wise crime statistics
Monthly crime trends
High-risk localities
Crime hotspot maps
The Emergency SOS module enhances the practical utility of the system by allowing users to instantly notify emergency contacts with live location information.
Overall, the experimental results demonstrate that Artificial Intelligence can effectively support crime prediction and public safety initiatives.

VII. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed system offers several advantages over conventional crime information systems.
Predicts future crime risks rather than merely reporting historical incidents.
Provides personalized crime risk assessments.
Integrates Machine Learning with GIS-based hotspot visualization.
Offers statistical dashboards for administrators.
Supports Emergency SOS alerts.
Easy deployment using Flask and SQLite.
Highly scalable for future cloud deployment.
User-friendly interface.

VIII. LIMITATIONS

Although the developed system performs effectively, several limitations exist.
Uses a synthetic dataset instead of real police incident records.
Crime prediction depends on historical patterns and cannot anticipate completely unexpected events.
Social and psychological variables were not included.
Prediction accuracy may improve with larger real-world datasets.
Weather and event information are not updated dynamically.

IX. FUTURE SCOPE

The proposed work can be enhanced through several future developments.
Integration with real-time police crime databases.
Incorporation of Deep Learning algorithms such as LSTM and Neural Networks.
Real-time weather API integration.
Mobile application development.
Voice-enabled Emergency SOS.
CCTV-based crime detection.
Cloud deployment with distributed databases.
Integration with Smart City surveillance infrastructure.

X. CONCLUSION

This research successfully developed an Artificial Intelligence-enabled crime prediction system capable of forecasting robbery tendencies and other crime categories using Machine Learning techniques. The Random Forest algorithm demonstrated reliable predictive performance while maintaining interpretability through feature importance analysis.
The integration of predictive analytics, hotspot visualization, statistical reporting, and Emergency SOS functionality distinguishes the proposed system from conventional crime reporting applications. Although the current implementation

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, July 2026
Impact Factor: 8.2

uses a synthetic dataset calibrated with NCRB statistics, the framework is flexible enough to incorporate real-world police datasets in future deployments.

The study confirms that Artificial Intelligence can play a significant role in improving public safety, supporting law enforcement agencies, and contributing to the development of smart and secure cities.

REFERENCES

[1]. Breiman, L. (2001). *Random Forests*. Machine Learning, 45(1), 5–32.

[2]. National Crime Records Bureau (2022). *Crime in India 2022*. Ministry of Home Affairs, Government of India.

[3]. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, 2825–2830.

[4]. Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.

[5]. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. (2017). *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Morgan Kaufmann.

[6]. Géron, A. (2022). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3rd ed.). O'Reilly Media.

[7]. Flask Documentation. https://flask.palletsprojects.com/

[8]. Scikit-learn Documentation. https://scikit-learn.org/

[9]. Python Software Foundation. https://docs.python.org/

[10]. Twilio Documentation. https://www.twilio.com/docs/sms