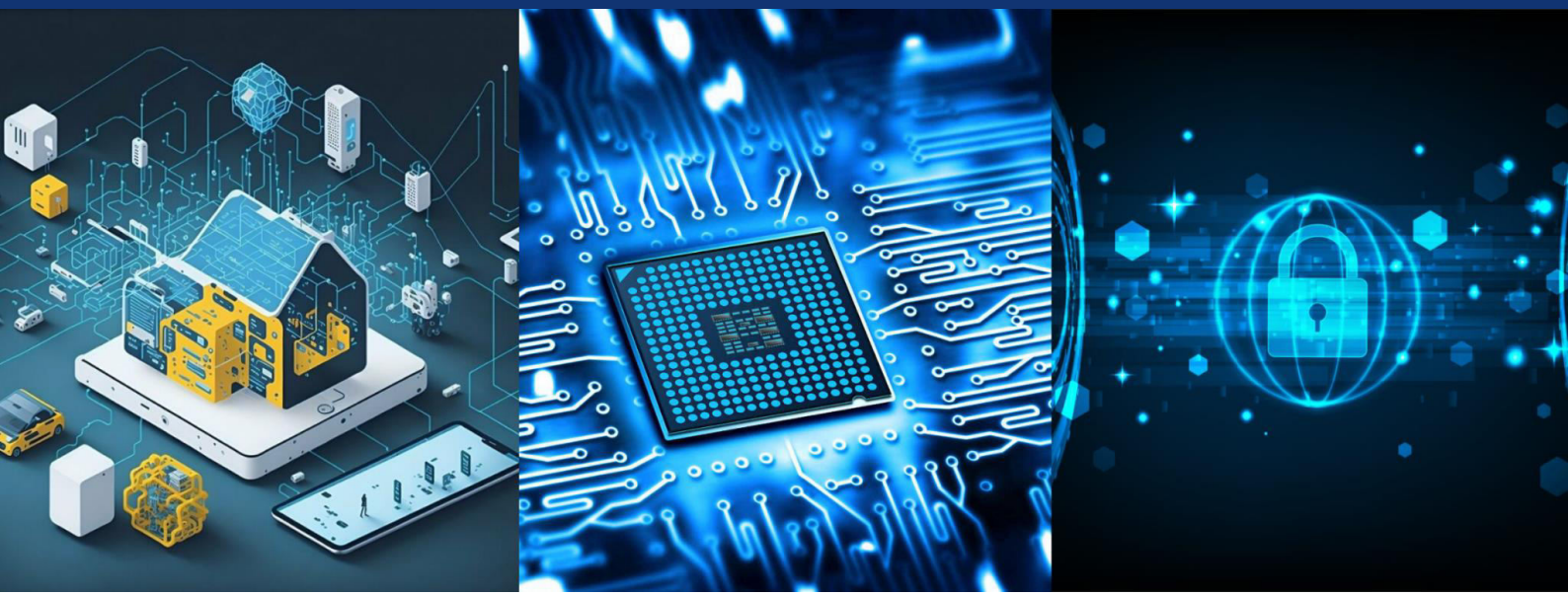
IJIRCCE

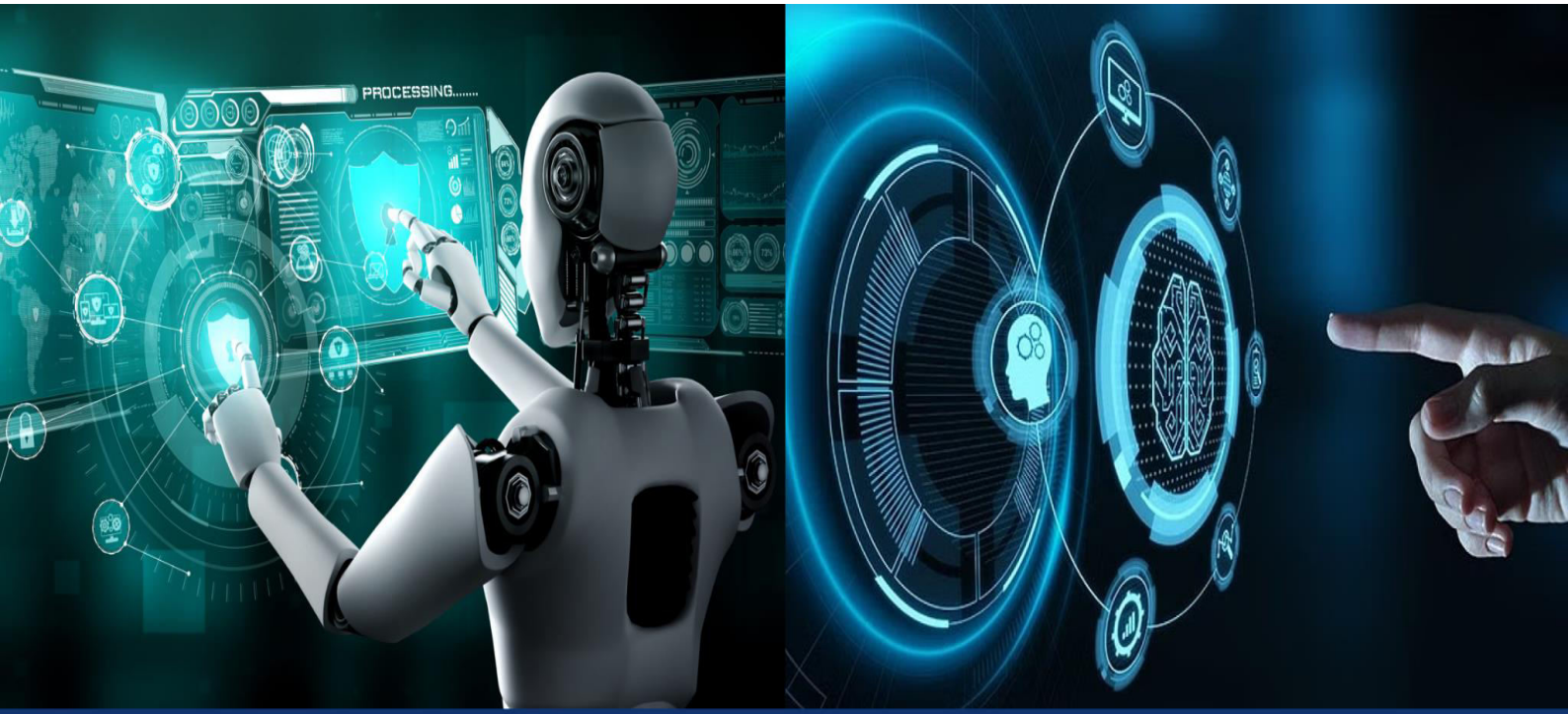
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 6, June 2026

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406021

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Crime Prediction using Machine Learning and Deep Learning

Bollu Poojitha, P. Rajasekhar

PG Student, Dept. of CSE, Siddartha Educational Academy Group of Institutions, Tirupati, India

Assistant Professor, Dept. of CSE(AI&ML), Siddartha Educational Academy Group of Institutions, Tirupati, India

ABSTRACT: Crime has become a major concern for governments and law enforcement agencies worldwide. The increasing availability of crime-related data and advancements in Artificial Intelligence (AI) have enabled the development of predictive systems capable of forecasting criminal activities. Crime prediction aims to identify potential crime hotspots, estimate crime occurrence probabilities, and support proactive policing strategies. This research presents a comprehensive study of crime prediction using Machine Learning (ML) and Deep Learning (DL) techniques. Various algorithms including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks are analyzed for crime forecasting. Historical crime datasets are preprocessed and transformed into meaningful features for model training. Experimental results demonstrate that deep learning models outperform conventional machine learning approaches in capturing complex temporal and spatial crime patterns. The proposed framework assists law enforcement agencies in resource allocation, crime prevention, and decision-making. The study concludes that integrating ML and DL techniques can significantly improve prediction accuracy and contribute to safer and smarter cities.

KEYWORDS: Crime Prediction, Machine Learning, Deep Learning, Artificial Intelligence, LSTM, Random Forest, Predictive Policing, Smart Cities.

I. INTRODUCTION

Crime prediction has emerged as a significant research area due to the increasing need for effective law enforcement and public safety. Traditional crime analysis methods rely heavily on manual investigation and historical trends, which often fail to provide timely insights for preventive action. With the rapid growth of data analytics and artificial intelligence, predictive models can now analyze large volumes of crime-related information and identify hidden patterns that may indicate future criminal activities.

Machine Learning and Deep Learning techniques have demonstrated remarkable success in various domains, including healthcare, finance, transportation, and security. Their ability to learn from historical data and generate accurate predictions makes them suitable for crime analysis. Crime prediction systems utilize information such as crime type, location, date, time, weather conditions, and socioeconomic indicators to estimate future crime occurrences. Predictive policing has gained attention as a strategic approach to reduce crime rates by enabling law enforcement agencies to deploy resources efficiently. However, developing accurate and reliable crime prediction models remains challenging due to the dynamic nature of criminal behavior and the complexity of influencing factors. This research investigates the application of machine learning and deep learning algorithms for crime prediction and evaluates their effectiveness in identifying crime hotspots and forecasting criminal incidents.

II. LITERATURE REVIEW

Several researchers have explored machine learning techniques for crime prediction. Logistic Regression and Decision Trees have been widely used for crime classification because of their simplicity and interpretability. However, these models often struggle with nonlinear relationships and complex data distributions.

Random Forest algorithms have shown improved performance by combining multiple decision trees to reduce over fitting and enhance generalization. Support Vector Machines have also demonstrated effectiveness in crime classification tasks due to their ability to handle high-dimensional datasets.

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406021

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Recent advancements in deep learning have enabled more sophisticated crime prediction models. Artificial Neural Networks can learn complex patterns from large datasets and provide better prediction accuracy than traditional methods. Convolutional Neural Networks (CNNs) have been utilized for spatial crime hotspot detection by analyzing geographical crime distributions.

Long Short-Term Memory (LSTM) networks have become increasingly popular for crime forecasting because they effectively capture temporal dependencies in sequential crime data. Hybrid CNN-LSTM architectures further improve performance by combining spatial and temporal feature extraction.

Despite these advancements, challenges such as data imbalance, privacy concerns, and model interpretability remain significant issues in predictive policing research.

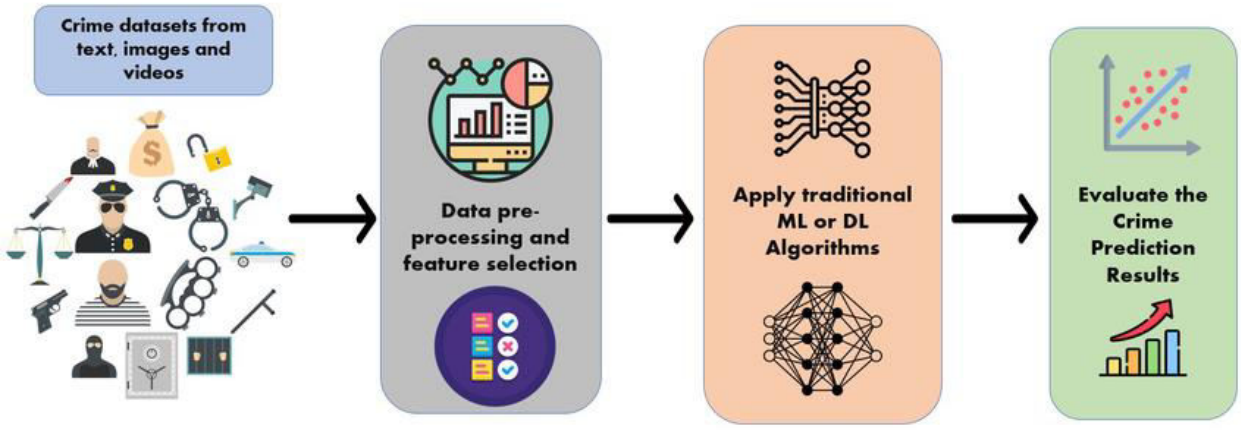
III. PROBLEM STATEMENT

Crime incidents often occur due to various interconnected social, economic, and environmental factors. Traditional crime prevention methods are reactive rather than proactive. Law enforcement agencies require intelligent systems capable of predicting potential crime occurrences before they happen.

The objective of this research is to develop a crime prediction framework using machine learning and deep learning techniques that can:

- Analyze historical crime data.
- Predict future crime occurrences.
- Identify crime-prone regions.
- Support strategic police deployment.
- Improve public safety through early intervention.

IV. PROPOSED METHODOLOGY



Data Collection

The crime dataset contains the following attributes:

- Crime ID
- Crime Type
- Date
- Time
- Location
- Latitude
- Longitude
- Population Density

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Weather Conditions
- Socioeconomic Indicators

Publicly available crime datasets from police departments and government agencies can be used for experimentation.

Data Preprocessing

Data preprocessing is essential for improving model performance.

Missing Value Handling

Missing values are replaced using mean, median, or mode techniques depending on the attribute type.

Data Cleaning

Duplicate records and inconsistent entries are removed.

Feature Encoding

Categorical variables such as crime type and location are transformed using one-hot encoding.

Normalization

Numerical attributes are scaled using Min-Max normalization to ensure consistent feature ranges.

Feature Engineering

Important features include:

- Day of Week
- Month
- Season
- Holiday Indicator
- Population Density
- Weather Parameters
- Previous Crime Frequency

Feature selection techniques help identify the most relevant variables influencing crime occurrences.

Machine Learning Models

Logistic Regression

Logistic Regression is used for binary crime classification. It estimates the probability of crime occurrence using a logistic function.

Advantages:

- Simple implementation
- Easy interpretation
- Computationally efficient

Limitations:

Limited capability for nonlinear relationships

Decision Tree

Decision Trees classify crime data based on hierarchical decision rules.

Advantages:

Easy visualization

Handles categorical and numerical data

Limitations:

Overfitting risk

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Random Forest
Random Forest combines multiple decision trees to improve prediction accuracy.

Advantages:
- High accuracy
- Reduced overfitting
- Robust performance

Support Vector Machine
SVM identifies optimal hyperplanes for classification.

Advantages:
- Effective in high-dimensional spaces
- Strong classification performance

Limitations:
Computationally expensive for large datasets

Deep Learning Models
Artificial Neural Network (ANN)
ANN consists of multiple interconnected neurons organized into input, hidden, and output layers.

Advantages:
- Learns complex nonlinear patterns
- High predictive capability

Convolutional Neural Network (CNN)
CNN extracts spatial features from geographical crime maps.

Applications:
- Crime hotspot detection
- Spatial crime analysis

Long Short-Term Memory (LSTM)
LSTM networks are designed to process sequential data and capture long-term dependencies.

Applications:
Crime trend forecasting
Temporal pattern analysis

Advantages:
- Handles time-series data effectively
- Reduces vanishing gradient problems

Hybrid CNN-LSTM Model
The CNN-LSTM model combines spatial and temporal feature extraction.

Benefits:
- Improved accuracy
- Better hotspot prediction
- Enhanced temporal forecasting

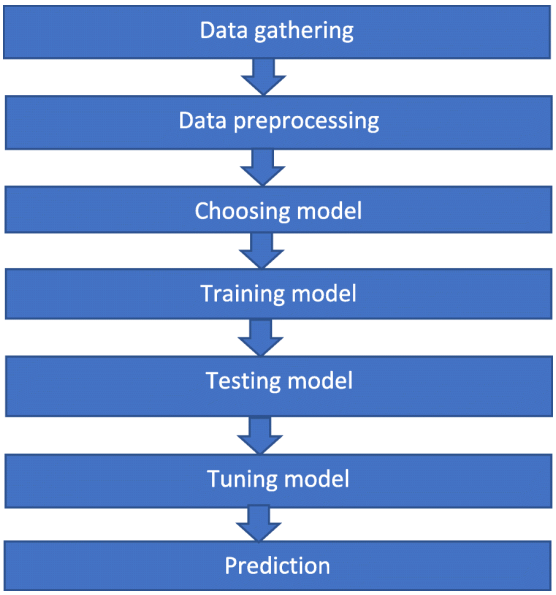
© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406021

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

System Architecture



The proposed system architecture consists of six stages:
Stage 1: Data Acquisition
Crime records are collected from databases and public repositories.
Stage 2: Data Preprocessing
Data cleaning, normalization, and feature extraction are performed.
Stage 3: Feature Engineering
Relevant features are selected for model training.
Stage 4: Model Training
Machine learning and deep learning algorithms are trained using historical crime data.
Stage 5: Prediction
The trained model predicts future crime occurrences and hotspots.
Stage 6: Visualization
Prediction results are displayed through dashboards, maps, and analytical reports.

IV. EXPERIMENTAL RESULTS

The dataset is divided into:
Training Set: 80%
Testing Set: 20%
Performance evaluation metrics include:
Accuracy
Measures the overall correctness of predictions.
Precision
Measures the proportion of correctly predicted crime events.
Recall
Measures the ability to identify actual crime occurrences.
F1-Score
Balances precision and recall.
RMSE
Evaluates prediction error for regression-based forecasting.
Applications
Predictive Policing
Supports police departments in identifying high-risk areas.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Smart City Security
Enhances urban surveillance and crime prevention strategies.
Emergency Response Planning
Optimizes allocation of emergency response teams.
Resource Management
Improves deployment of law enforcement personnel.
Public Safety Monitoring
Provides early warnings regarding potential crime threats.

V. CONCLUSION

Crime prediction using Machine Learning and Deep Learning techniques represents a transformative approach to modern law enforcement. Traditional machine learning algorithms such as Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines provide valuable insights into crime patterns. However, deep learning models, particularly LSTM and CNN-LSTM architectures, demonstrate superior performance in capturing temporal and spatial crime dynamics.

Experimental results indicate that hybrid deep learning models achieve the highest prediction accuracy and offer significant potential for crime hotspot detection and forecasting. The proposed framework enables proactive crime prevention, efficient resource allocation, and improved public safety. Future developments incorporating IoT, GIS, Explainable AI, and real-time analytics are expected to further strengthen intelligent crime prediction systems.

REFERENCES

1. Wang, T., Rudin, C., Wagner, D., and Sevieri, R., "Learning to Detect Patterns of Crime," Machine Learning Journal, 2018.
2. Mohler, G., "Predictive Policing and Crime Forecasting," Annual Review of Statistics and Its Application, 2020.
3. Chainey, S., and Ratcliffe, J., GIS and Crime Mapping, Wiley Publications, 2019.
4. Goodfellow, I., Bengio, Y., and Courville, A., Deep Learning, MIT Press, 2016.
5. Breiman, L., "Random Forests," Machine Learning Journal, 2001.
6. Hochreiter, S., and Schmidhuber, J., "Long Short-Term Memory," Neural Computation, 1997.
7. LeCun, Y., Bengio, Y., and Hinton, G., "Deep Learning," Nature, 2015.
8. Aggarwal, C.C., Neural Networks and Deep Learning, Springer, 2018.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details