ALVA'S INSTITUTE OF ENGINEERING & TECHNOLOGY

(Unit of Alva's Education Foundation ®, Moodbidri)
Affiliated to Visvesvaraya Technological University, Belagavi,
Approved by AICTE, New Delhi, Recognized by Government of Karnataka.
Accredited by NACC with A+ Grade
Shobavana Campus, Mijar, Moodbidri, D.K., Karnataka

2025 – 2026

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

CRIME PATTERN ANALYSIS AND PREDICTION USING MACHINE LEARNING

ABSTRACT

This project focuses on understanding and predicting crime patterns using data-driven technologies. By analyzing large volumes of historical crime data, the system identifies important spatial and temporal trends that influence criminal activity. The work integrates machine learning models—such as Random Forest, KNN, SVM, Decision Trees, and Naive Bayes—to classify high-risk areas, forecast future crime levels, and predict likely crime types. Geospatial mapping techniques further support hotspot detection, offering a clearer view of where crimes tend to cluster. A structured data-processing pipeline ensures clean, consistent, and model-ready datasets, improving the reliability of predictions.

INTRODUCTION

The rise in urbanisation and population growth has made crime analysis an important part of modern public-safety planning. Traditional policing methods largely depend on manual review of crime records and reactive decision-making, which often fail to keep up with fast-changing crime patterns. With the availability of large crime datasets and advancements in artificial intelligence, data-driven approaches now offer a more effective alternative. This project applies machine learning and geospatial analysis to study how crimes vary across locations and time. By examining historical crime data, the system identifies trends, detects hotspots, and predicts future crime behavior. These insights support proactive policing, better resource allocation, and more informed policy decisions.

PROBLEM STATEMENT

Law enforcement agencies struggle to detect, analyse, and predict crime effectively because traditional methods depend on outdated reports, manual review, and limited data integration. As crime patterns change rapidly due to urban growth, socio-economic factors, and population movement, these conventional approaches fail to provide timely or accurate insights. This leads to poor resource allocation, delayed response, and an inability to identify emerging hotspots. Therefore, there is a need for a data-driven system that can analyse large crime datasets, predict future crime occurrences, and support proactive decision-making to improve public safety.

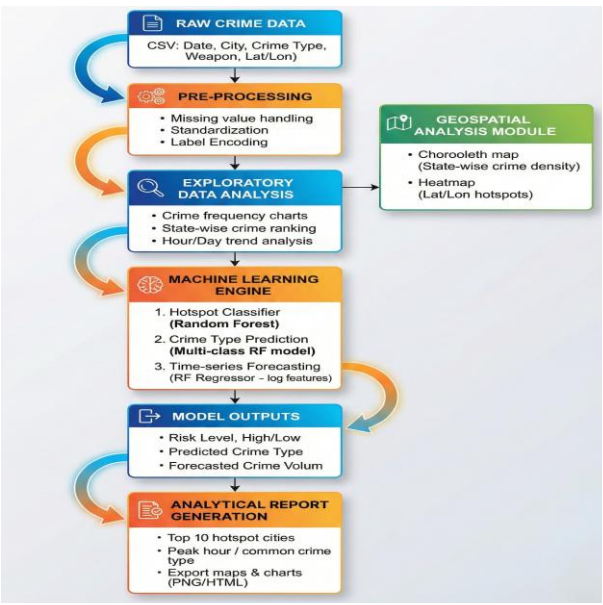
OBJECTIVES

The primary objective of this project is to analyse historical crime data and identify meaningful spatial and temporal patterns that influence criminal activity. The system aims to develop accurate machine learning models capable of predicting crime hotspots, future crime volume, and likely crime types. Another key objective is to apply geospatial techniques for visualizing high-risk locations, enabling better understanding of crime distribution. The project further seeks to automate data processing and reporting so that insights can be easily used by law-enforcement agencies for strategic planning, resource allocation, and proactive crime prevention.

METHODOLOGY

The methodology of this project begins with collecting historical crime data and preparing it through cleaning, handling missing values, and converting timestamps into usable features such as hour, day, and month. Exploratory Data Analysis (EDA) is then performed to understand crime distribution across locations and time periods. Geospatial tools are used to generate hotspot maps and visualize crime density using latitude–longitude information. Machine learning models such as Random Forest, KNN, SVM, Naive Bayes, and Decision Trees are trained to classify high-risk areas, predict crime types, and forecast future crime volume. Each model is evaluated using accuracy, precision, recall, and other performance metrics to ensure reliability. Finally, and summaries to provide actionable insights for law-enforcement use.

FLOW CHART



RESULTS

The project successfully identified key crime hotspots by analyzing spatial and temporal patterns in historical data. The machine learning models especially Random Forest and KNN showed strong predictive accuracy in classifying high-risk areas and forecasting future crime levels. Geospatial heatmaps and choropleth maps clearly highlighted regions with dense criminal activity, enabling better visual interpretation. The automated data-processing pipeline improved dataset quality, which further enhanced model performance. Overall, the system demonstrated that data-driven analysis and ML-based prediction can support proactive policing, improve decision-making, and provide reliable insights for crime prevention.

FUTURE ENHANCEMENTS

Future enhancements of this system can focus on integrating real-time data from CCTV networks, IoT sensors, and emergency response systems to provide more immediate and accurate crime predictions. Advanced AI and deep-learning models can be introduced to improve the adaptability of the system as crime patterns evolve. The platform can also be expanded to include social media sentiment analysis and community feedback to capture public concerns and behavioral signals. Collaboration with urban planners, and law-enforcement experts can help build more comprehensive prevention strategies. Additionally, developing a user-friendly mobile application or dashboard would make the insights easily accessible to police personnel and the general public, supporting faster and more informed decision-making.

CONCLUSION

This project demonstrates that machine learning and geospatial analysis can play a significant role in understanding and predicting crime patterns. By analyzing historical crime data, the system successfully identifies hotspots, classifies high-risk areas, and forecasts future crime occurrences. The models, especially Random Forest and related ensemble techniques, show strong performance in handling complex crime datasets and revealing hidden trends. The integration of automated data processing, visualization tools, and predictive algorithms supports proactive policing and better decision-making. Overall, the project highlights the potential of data-driven crime analysis to improve public safety, assist law-enforcement agencies, and contribute to smarter planning for crime prevention.

TEAM MEMBERS

1. LATHESH KUMAR S R - 4AL23AI400
2. SANKETH PATIL - 4AL22AI043
3. SHIVAMANI M NAYAK - 4AL22AI051
4. TEJASHWINI SHAILENDHRA MURDESHWAR - 4AL22AI060

PROJECT GUIDE

PROF. PRADEEP NAZARETH
ASSOCIATE PROFESSOR
DEPT. OF AIML, AIET