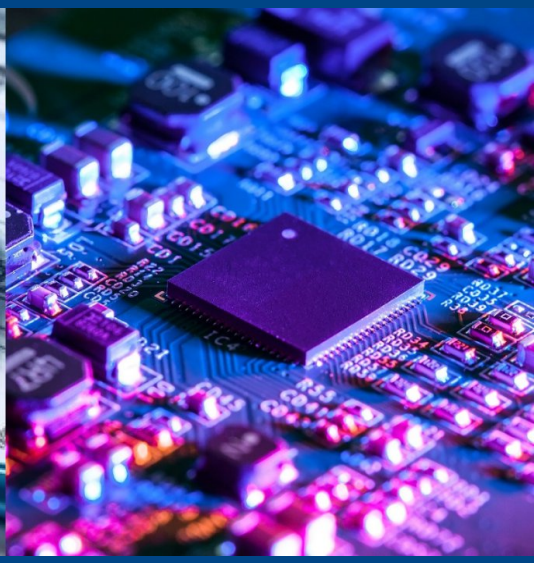
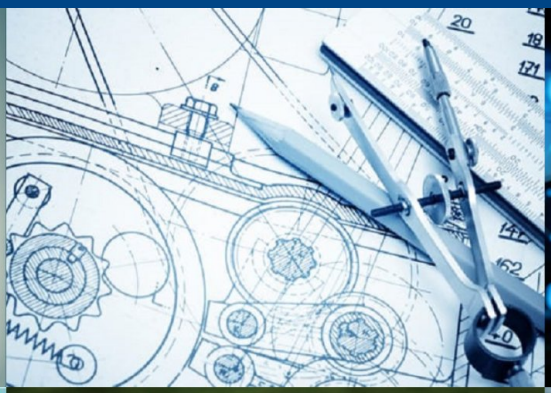
ISSN: 2582-7219

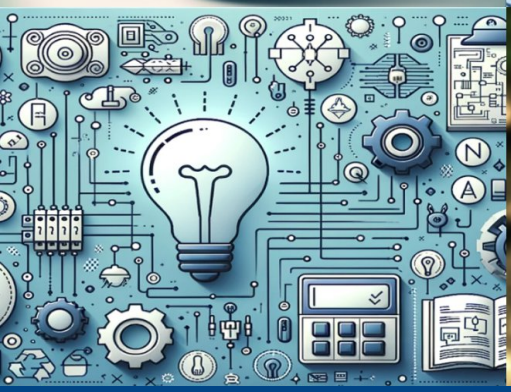




International Journal of Multidisciplinary Research in Science, Engineering and Technology

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



Impact Factor: 8.206 **Volume 9, Issue 4, April 2026**

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

Crime Pattern Detection and Hotspot Analysis System Using Machine Learning System

Lavanya .s [1], Musfira Zaka [2], and Ms.Sabaria S [3], H Vignesh[4], S Adhil Harshath[5], and Dr. R Shanthi[6]

Third Year B. Tech Student, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[1]

Third Year B. Tech Student, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[2]

Assistant Professor, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[3]

Third Year B.Tech Student, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[4]

Third Year B.Tech Student, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[5]

Assistant Professor, Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India[6]

ABSTRACT: Crime analysis and hotspot identification play a critical role in enhancing public safety and assisting law-enforcement agencies in proactive crime prevention. Tradition crime monitoring methods rely primarily on manual reports and static statistical summaries, which Patterns and hotspot localization's are frequently not visible in the raw data of large datasets. This paper proposes a machine learning-driven framework for crime hotspot detection and spatial crime analysis using crime data collected from various districts of Tamil Nadu. The proposed system utilizes crime attributes including crime type, geographic location to analyze crime distribution patterns.

Multiple data analysis Spatial analysis and machine learning techniques are employed to pre-process crime data, perform cluster analysis, produce visualization maps to help locate hot spots and spatial models that describe the environmental causes of crime and detect high-risk regions. Spatial visualization methods such as geographic heat maps are employed to represent crime density across different locations, enabling intuitive identification of crime hotspots.This work relies heavily on a web-based crime data mining system that uses historical crime incident data and geographical information system data to build a crime forecasting model. Visualization dashboard is developed to present crime density maps and analytical insights for decision support. The proposed approach enables efficient crime monitoring, identification, and supports data-driven policing strategies for better law enforcement resource allocation

I. INTRODUCTION

Crime, necessitating advanced analytical methods, is among the most significant issues globally. Law enforcement addresses crime due to its focus on the way prevention appears in complex environments. The issue involves manual observation and basic statistics, and the hidden patterns of data reflect the analyst's idea of the connection between urbanization and population.

II. LITERATURE REVIEW

undertaken to predict the crime using machine learning techniques. Previous studies employed K Nearest Neighbors (KNN) and Decision Trees, yielding fundamental accuracy with interpretability. Recent research is directed towards more complicated models, for example Random Forest model, Support Vector Machines (SVM) and Deep Learning

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

approaches such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM). Another approachConvolutional Neural Networks (CNN) are also used for crime prediction by extracting spatial features from crime data. In addition, Long Short-Term Memory (LSTM) networks are applied to capture temporal patterns in crime occurrences. These advanced machine learning and deep learning techniques improve prediction accuracy and help law enforcement agencies in decision-making and resource allocation.

Machine learning techniques are widely used in crime prediction to analyze historical crime data and identify patterns. Methods such as K-Nearest Neighbors (KNN), Decision Trees, and Random Forest are commonly applied for classification and prediction tasks. Clustering methods are also used to group similar crime incidents based on location and time. These approaches help in understanding crime trends and support effective law enforcement strategies.

III. DATASET DESCRIPTION

The dataset includes several important attributes that affect crime analysis and prediction. These features were chosen to reflect both the time and place of criminal activities.
The main attributes considered in this study are:

- Crime ID
- Crime type(e.g.Theft,Assault,Cyber Crime,Robbery)
- Location (City/Region)
- Date and Time of Occurrence
- Crime Category (Domestic / Non-Domestic)
- Latitude and Longitude (for geographic mapping)
- Number of Crimes per Location
- Crime Density Information

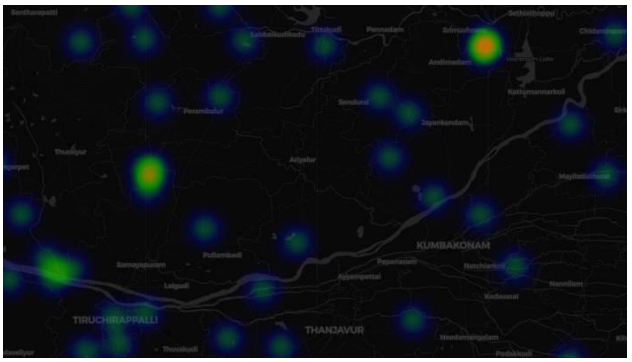


Figure 1 illustrates the geographic visualization of crime distribution, highlighting different regions where crime intensity varies. The map provides a clear understanding of high-density and low-density areas using heatmap techniques.

IV. METHODOLOGY

A. *Data Collection*
The data is collected from various government, non government websites, available datasets, and other websites. Also some of the datasets are collected from other resources such as online news using web scraping, etc.
B. *Data Preprocessing*
The data is cleaned and pre-processed to remove redundancy and fill the gaps in the data for achieving a smooth and complete data set. This dataset results in a smooth and accurate prediction. The data is arranged as required.
C. *Data Analysis*
The data is analyzed for required information which will become an input to the predicting algorithm later. Data analyses helps to know the data and take required measures for the machine learning model to perform accurately.
D. *Data Prediction*
The data is then feed to the prophet tool which predicts the crime rate of certain crimes in a specific area. This tool works on the date time column i.e., the time series, to produce its output.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

E. *Data Visualization*

This website provides various forms in which the data can be visualized such as heat map, pie chart, bar graph, etc. It helps in understanding large datasets.

FRAMEWORK

A. Flask

Flask is a web framework that allows developers to build lightweight web applications quickly and easily with Flask Libraries. It was developed by Armin Ronacher, leader of the International Group of Python Enthusiasts(POCCO). It is based on the WSGI toolkit and Jinja2 templating engine

LIBRARIES

A. Pandas

Pandas is a Python library designed for data manipulation and analysis, providing specialized data structures and functions tailored for working with numerical tables and time series data. The library is built upon another library, NumPy

B. Matplotlib

Matplotlib, a plotting library compatible with both Python and its numerical mathematics extension NumPy, stands as a potent tool for individuals engaged in Python and NumPy based endeavors, offering extensive capabilities for creating visualizations. And for making statistical interference, it becomes very necessary to visualize our data and Matplotlib is the tool that can be very helpful for this purpose.

C. Seaborn

Seaborn, built upon matplotlib, serves as a Python data visualization library offering a sophisticated interface for crafting visually appealing and informative statistical graphics. Seaborn is a library for making statistical graphics in Python. Expanding upon matplotlib and tightly integrating with pandas data structures, Seaborn's plotting functions are tailored to operate seamlessly on dataframes and arrays encompassing entire datasets. Internally, they execute essential semantic mapping and statistical aggregation, culminating in the creation of insightful plots. Its dataset oriented, declarative us focus on what the different elements of our plots mean, rather than on the details of how to draw them

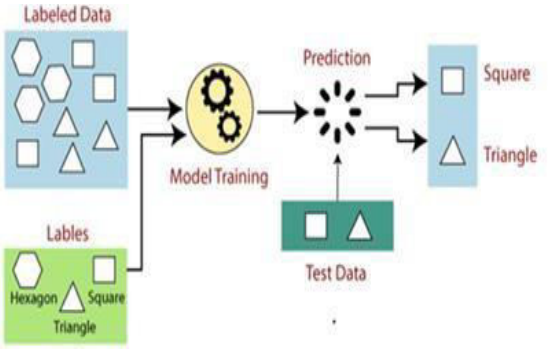
D. Numpy

NumPy enhances Python with support for large, multi-dimensional arrays and matrices, complemented by a vast array of highlevel mathematical functions tailored for manipulating these arrays. NumPy is open-source software and has many contributors

ALGORITHM

A. Supervised machine learning

- This system uses supervised learning algorithm for prediction. Supervised learning falls within the realm of machine learning, where labeled datasets are employed to train algorithms, enabling them to predict outcomes and identify patterns. Labelled data refers to input data that has been pre-assigned with corresponding correct output values. In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly)



© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

ALGORITHM

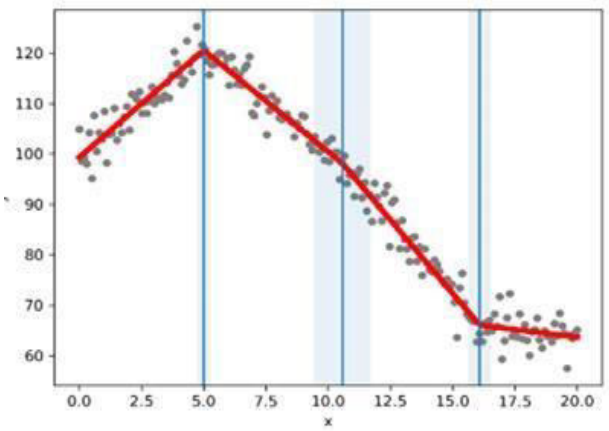
A. Supervised machine learning

- This system uses supervised learning algorithm for prediction. Supervised learning falls within the realm of machine learning, where labeled datasets are employed to train algorithms, enabling them to predict outcomes and identify patterns. Labelled data refers to input data that has been pre-assigned with corresponding correct output values. In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly)

B. LINEAR REGRESSION

The piecewise linear regression model in Prophet captures the overall trend in the data while allowing for flexibility and adaptability to changes over time. This approach differs from traditional linear regression models, which assume a single linear relationship between the predictor variables and the target variable.

By incorporating piecewise linear regression, Prophet can capture complex trends and patterns in the time-series data, making it particularly suitable for forecasting taskswhere the trend may exhibit non-linear behavior or undergo changes over time.



V. PROPOSED SYSTEM

The proposed system uses Machine Learning and Data Visualization to overcome existing limitations

A. Features:

- Data Pre-processing and Cleaning
- Crime prediction using Random Forest Machine Learning algorithm
- Heat map generation using latitude & longitude
- User-friendly visualization

B. ADVANTAGES

- Faster prediction with accuracy
- Identify hotspots where crimes are happening
- Facilitates decision-making for authorities

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

VI. RESULT

The required data is collected and preprocessed as required. Then we have created a heat map for depictions of areas with high crime rate

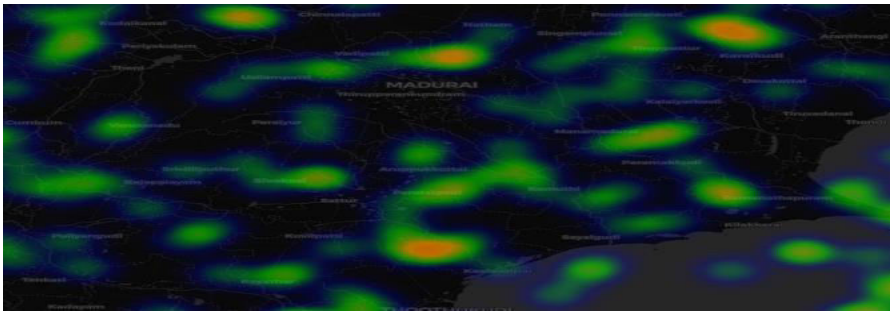


Fig4. Heat map

We have also created crime intensity metrices graph for crime trends over year

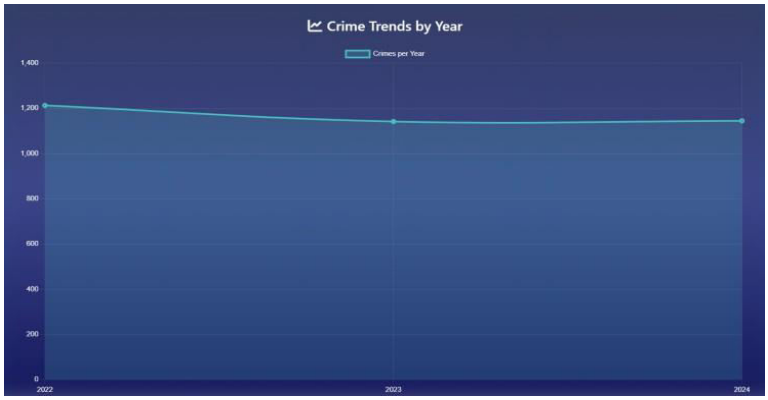


Fig5. Crime Trends

Then we have also created crime distribution for better understanding.

Fig6. Domestic and Non-Domestic

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904370

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

APPLICATION

The proposed crime analysis and prediction system has a wide range of practical applications across multiple domains. By combining data visualization, geospatial analysis, and machine learning, the system supports informed decision-making and enhances public safety.

- Police Departments:

Law enforcement agencies can use the system to identify crime hotspots, monitor trends, and allocate resources more effectively. Predictive insights help in proactive patrolling, faster response planning, and reducing crime rates in high-risk areas.

- Government Agencies:

Government bodies can utilize the system for policy formulation, urban planning, and safety initiatives. It enables data-driven decisions for improving infrastructure, surveillance systems, and law enforcement strategies.

- Crime Analysts:

Analysts can leverage the dashboards and heat-maps to study crime patterns, correlations, and trends over time. The system simplifies complex datasets into meaningful visual insights, aiding in accurate reporting and forecasting.

- Research and Academic Use:

The system serves as a valuable tool for researchers and students to study crime behavior, test machine learning models, and develop advanced analytical solutions.

VII. CONCLUSION

Crime is an unlawful act which disturbs the peace and harmony of the society. This projects aims to successfully predict crime and their locations based on the historical crime data. The project uses machine learning which is an advanced and latest technology for accurate prediction. The web application will display crime rate in various areas. It is extremely useful for both the higher investigating authorities and officers designated to handle low level crime for tracking and stopping the crime. The predictions will help to ensure increased security and thus could help in lowering the crime rate. Overall, the project demonstrates the potential of data analysis and mapping technologies to improve public safety and inform decision-making. Proactive measures can be taken to prevent crime and improve public safety by using data to identify crime hotspots and trends. Although there's more work needed to enhance the precision and breadth of the project, it marks a significant stride towards employing data-driven strategies to tackle intricate social challenges.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

REFERENCES

[1] Varun Mandalapu , Lavanya Elluri, Piyush Vyas, and Nirmala Roy, "Crime Prediction Using Machine Learning and Deep Learning: A

[2] Systematic Review and Future Directions" , IEEE, Volume 11.

[3] Raza, D. M. & Victor, D. B. ''Data mining and region prediction based on crime using random forest'', 980–987 (IEEE, 2021). .

[4] Prakash Maurya, Tahir Shaikh, Imran Ahmed, Amaan Firdosi, Prof. Kiran Deshmukh, "Crime Analysis and Prediction Using Machine Learning", IJRASET, Volume 11, Issue 4.

[5] Tzu-Wei Hung1 ,Chun-Ping Yen, "Predictive policing and algorithmic fairness", Synthese (2023)

[6] Ruaa Mohammed Saeed, Husam Ali Abdulmohsin, "A study on predicting crime rates through machine learning and data mining using text", Journal of Intelligent Systems, Volume 32, Issue 1

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com