A-Comprehensive-Study-on-Crime-Rates-and-Trends-in-Los Angeles

KARTHEEK
kabr4568@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

AJAY TATA
ajta4172@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

SAI RAM REDDY APPIDI
saap9704@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

Abstract

The research work studies and projects crime incidence based on several aspects such as, but not limited to, type of crime, geographical location, time of occurrence, and antecedent records of crime. The research is done for crime-related data and analysis using machine learning-based models: Logistic Regression, Random Forest, XGBoost, Support Vector Machines (SVMs), and Multi Layer Perceptrons (MLP). All models take historical crime data about Los Angeles into their proper analysis for attempting to derive details, patterns, and correlations. The gathered information will now be used to predict future crime trends, points at the greatest risk, and possibilities for improving public safety.

Data was prepared through comprehensive preprocessing techniques; missing value treatments, categorical encoding, standardization, outlier removal, and so on, which maximized the performance of machine learning models. The best was achieved by MLP, producing excellent metrics-record accuracy of 99.68

The research accentuates the significance of data-driven decisions for solving societal problems, such as preventing crime and allocating resources, behind an evidence-based approach to law enforcement. It shows the effectiveness of analytical power across disciplines in finding meaningful gains from huge crime databases while paving the way toward much more innovative developments, such as using real-time data streams or utilizing interpretability methods more advanced than these, on the path to further improving prediction abilities.

Keywords

Crime Analysis, Machine Learning Models, Crime Prediction, Los Angeles Crime Data, Multi-Layer Perceptron (MLP), Temporal Crime Trends, Geospatial Crime Analysis, Public Safety, Data-Driven Decision Making, Predictive Modeling, Law Enforcement Optimization, Urban Development, Resource Allocation, Crime Hotspots, Exploratory Data Analysis (EDA), Crime Severity, Data Preprocessing, Outlier Detection, Hyperparameter Tuning, Real-Time Crime Analysis

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/XXXXXXX.XXXXXXX

ACM Reference Format:
KARTHEEK, **AJAY TATA**, and **SAI RAM REDDY APPIDI**. 2018. A-Comprehensive-Study-on-Crime-Rates-and-Trends-in-Los Angeles. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Identifying patterns is increasingly becoming one of the most important avenues to understanding changing crime scenarios, as this provides information for urban development and public safety. Trend analysis is a data-oriented approach scrutinizing historical crime data to bring out patterns and relationships in emerging threats or very rarely flagged threats. Such capabilities are critical to law enforcement and policymakers to effectively fight the ever-complex and changing nature of crime, particularly in an urban setting. Predictive features and also trend forecasting are key to making resourcing allocations, prevention, and community safety more effective. Data was prepared through comprehensive preprocessing techniques; missing value treatments, categorical encoding, standardization, outlier removal, and so on, which maximized the performance of machine learning models. The best was achieved by MLP, producing excellent metrics-record accuracy of 99.68

The research accentuates the significance of data-driven decisions for solving societal problems, such as preventing crime and allocating resources, behind an evidence-based approach to law enforcement. It shows the effectiveness of analytical power across disciplines in finding meaningful gains from huge crime databases while paving the way toward much more innovative developments, such as using real-time data streams or utilizing interpretability methods more advanced than these, on the path to further improving prediction abilities.

The objectives of this study include:

(1) Identifying crime trends based on the features like date time area and many other factors
(2) Predicting the different type of crimes in different areas using the data of the area.
(3) Analyze the data and identify the patterns to detect the high risk areas and also to predict the crime hotspots.
(4) Providing actionable insights for strategic deployment of the law enforcement resources
(5) showing the value of machine learning for societal problems, like crime prevention and enhancing urban safety .

This report will be used to make better decisions for policing departments, policymakers, researchers and planners. This work shows that crime trend analysis can be applied to public safety, resources allocation, and crime reduction by identifying meaningful patterns and their impacts. The results can be used to guide targeted interventions and evidence-based practices for safer communities. It is structured as follows: the datasets and preprocessing were described, the machine learning models were developed and tested, the results were presented and the impact on public safety and police protection, and it concludes with the main findings and directions for future research to expand the scope of this work to real-time data integration and predictive modeling.

2 EXPLORATORY DATA ANALYSIS

In the exploratory data analysis we have provide the detailed description of the project like data set used in our project and preprocessing steps that are taken for the data to gain the readbility of the data.The data set which we are used in our project is losangles data set until March 2024.in the data cleaning process there are many characteristics and challaneges like missing values ,ientifying the outliers

2.1 DATA SET

The dataset for this crime analysis project was obtained from the Los Angeles Police Department (LAPD) website [data source] . The data reflects incidents of crime in the City of Los Angeles dating back to 2020. The LAPD transcribes this data from original crime reports, which are initially written on paper and then manually entered into the system. This data is released publicly by the LAPD to allow for transparency and analysis of crime patterns within the city.

It's important to note that starting from March 7th, 2024, the LAPD will be transitioning to a new crime reporting system that complies with the FBI's National Incident-Based Reporting System (NIBRS).

2.2 DATA CLEANING

The data cleaning process involve the handeling the missing values in the data and removing outliers in the data .

Handling Missing Values missing values in a data set will have a dramatic effect on the results of ML models and statistics which can result in biased estimates and prediction errors. It can be because of various causes including the errors in data collection, system crash or incomplete entries. Missing value processing was a key preprocessing operation for this project to make sure the data is in good shape and the model is stable.

```
# Check for missing values in the dataset
missing_values = df.isnull().sum()

# Display columns with missing values
print(missing_values[missing_values > 0])

# Unwanted Columns
unwanted_columns = ['DR_NO', 'Mocodes', 'Cross Street', 'Crm Cd 1', 'Crm Cd 2', 'Crm Cd 3', 'Crm Cd 4']

# Drop unwanted columns
crime_data_cleaned = df.drop(columns=unwanted_columns)

# Verify the cleaning process
crime_data_cleaned.info()
```

Removing outliers in the data Outliers include all those correlations that are pretty much deviant to the rest of observations and can affect even a great analysis regarding the conclusion. Outliers in crime data may mean data entry error, rare occurrence, or exceptional condition. Bringing down the outliers is a must to make sure that the machine learning architectures are performing perfectly and accurately.

In this project, outlier detection and removal by the interquartile range (IQR) technique were applied. IQRs can be calculated as: IQR=Q3-Q1 Where: Q1 is the median of lower half data in the data set. Q3 is the median of the upper half of dataset points below Q1-1.5×IQR or above Q3+1.5×IQR were termed and eliminated as outliers. This ensures that the dataset remains as representative

```
if 'Weapon Used Cd' in df.columns:
   df['Weapon Used Cd'] = df['Weapon Used Cd'].fillna('Unknown')


if 'Weapon Desc' in df.columns:
   df['Weapon Desc'] = df['Weapon Desc'].fillna('Unknown')

# Convert Age to integer
if 'Age' in df.columns:
   # Fill missing age values with median age (you can adjust the strategy)
   df['Age'] = df['Age'].fillna(df['Age'].median()).astype(int)
numeric_columns = df.select_dtypes(include=['float64', 'int64']).columns


for col in numeric_columns:
   Q1 = df[col].quantile(0.25)
   Q3 = df[col].quantile(0.75)
   IQR = Q3 - Q1
   lower_bound = Q1 - 1.5 * IQR
   upper_bound = Q3 + 1.5 * IQR
   # Capping outliers
   df[col] = df[col].clip(lower=lower_bound, upper=upper_bound)
```

as possible of typical patterns while eliminating extreme values that may skew the results of trends in crime analysis and possible predictions through machine learning.

2.3 DATA Visualizations

As the data cleaning and processing is done on the crime dataset was appropriately set for analysis, ensuring that it contained no missing values, that every variable had been appropriately scaled, and that no outliers were present. Missing values were dealt with via the usage of SimpleImputer, which replaced numerical and categorical variables with the mean and most frequent values, respectively.

With these thorough preprocessing steps, all guarantees-tautologically noted here-about the reliability and quality of the data together spell a ready opportunity for producing powerful visualizations and revealing crime trends and patterns across locations, timeframes, and types of crime data. These findings become the bases to provide accurate modeling and prediction-enabled treatment.

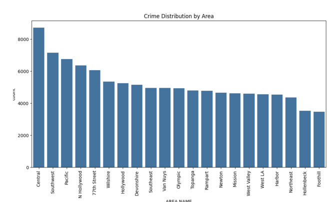

Figure 1: area plot

Distribution of Crimes-The pattern shown in Figure 1 reveals that the Central area is more crime-prone, reporting more than 8,000 incidents, making it a central hotspot for criminal activities. Other regions are reported as high crime numbers, including Southwest, Pacific, and North Hollywood, with decreases set at just below 6,000. Others, such as Foothill, Hollenbeck, and Harbor, consistently report lower crime rates of fewer than 4,000. Such a method of presentation illustrates how certain areas are crime-concentrated while others maintain a subdued level of crime.

The graph provides a lucid illustration of crime occurrence rates in the countries. The leading crime among all incidents is motor vehicle theft at slightly above 17000 cases. Vehicle burglary and minor theft cases follow, thereby indicating that most crimes are property offenses. Vandalism and trespassing come in the middle, which would refer to crimes mostly associated with personal or community security. Interestingly, battery and identity theft score

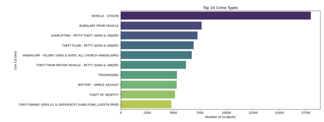

Figure 2: Top 10 Crime Types: Frequency of Incidents Highlighting Vehicle Theft as the Leading Offense

lower on the scale but remain significant, reflecting a wide range of diverse crimes. Understanding these trends may help in basic resource allocation for crime prevention and policy making.

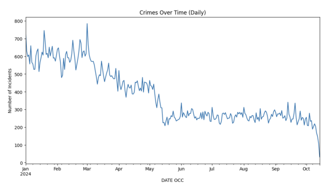

Figure 3: Daily Crime Trends in 2024

The graph illustrates the tendency of daily crime across the year 2024, which indicates the number of occurrences captured within each day. During the early part of the year, daily crimes are recorded at high numbers approaching between 700-800, trending downwards after a peak to reach a significant decline beginning in April with the averages further decreasing to ranges around 200-400 at mid-year. The trend is also observed to have stabilized at low rates after, with minor inconsistencies; the end of the year sees sharp drops, probably seasonal or data-related.

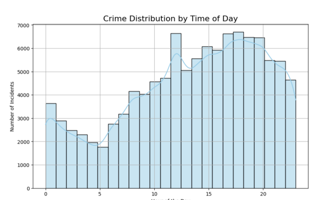

Figure 4: Crime Activity by Hour: Peaks During Afternoon and Early Evening

This displays a graphical representation of the crime by hour of the day that indicates when the crime incidents are likely to occur. The x-axis shows hours of the day (from 0 to 23), while the y-axis shows the count of reported incidents.Meeting a relatively high value in the overnight period (from midnight to 1 a.m.), the incident rates fall for the remaining hours between, mostly during early morning hours (from 2 a.m. to 6 a.m.) and suggest reduced activity at night. Around 7 a.m., crime rates rise sharply, peaking in the afternoon between 3 p.m. and 6 p.m. These hours coincide with an increased human presence in office locations and high traffic during work and commuting hours. After 8 p.m., numbers taper off gradually but remain at a fairly high and level amount well into the evening.

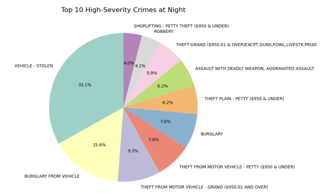

Figure 5: Top 10 Nighttime High-Severity Crimes: Dominance of Vehicle-Related Incidents

This pie chart shows how the ten most severe high-quality crimes are associated with the night when comparing the types. Highest proportions which amount to 33.1

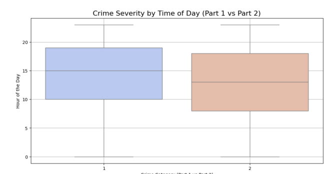

Figure 6: Boxplot comparison of the time of day for Part 1 (severe) and Part 2 (less severe) crimes, showing the distribution of their occurrences during the day

This box plot delineates crime occurrence for different time hours of the day under two classifications: Part 1 crimes, which includes severe crimes, and Part 2 crimes, which are less severe. The x-axis defines the different crimes, while the y-axis defines hours of the day (0-24). Looking at the Part 1 crimes, which are the blue box, there is a relatively wider range of occurrence time having median in the afternoon (15:00). However, in the case of Part 2 crimes (orange) it looks slightly similar distribution with a bit of skewness towards the later hours. This illustration is supposed to present time variations relating to severity as far as crimes are concerned, thereby ensuring the appropriate application of law enforcement resources during prime hours.

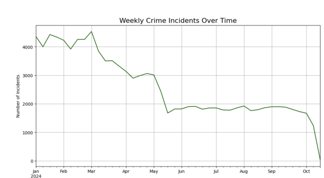

Figure 7: Weekly Crime Incidents in 2024

figure 7 This line graph shows the trend of crime over the weeks in the year 2024-with notable decreasing time-to-time crime occurrences as the year advances. In January and February, crime incidents remain at about 4,000 per week. However, from April onward, there is a sharp decrease, falling below 2000 by May. The numbers continue to worsen after their stabilization during the summer months, with crime coming almost to zero by October.

The violin plot figure 8 depicts crime occurrences across different statuses and their distribution over the hours of the day. The

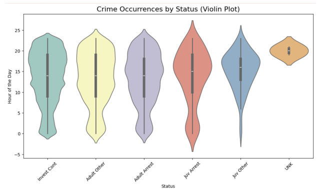

Figure 8: Crime Occurrences by Status

statuses include "Invest Cont," "Adult Other," "Adult Arrest," "Juv Arrest," "Juv Other," and "UNK." The plot shows the density of crimes at specific hours, with boxplots embedded to highlight median values, interquartile ranges, and variability. For example, "Adult Arrest" and "Juv Arrest" have denser distributions around particular time windows, indicating peak activity. This visualization effectively reveals temporal crime patterns based on status classifications.

Model Implementation and Performance

To predict crime patterns effectively, we implemented a suite of machine learning models, each tailored to address the complexities of the dataset. Below, we provide a detailed account of the models, their configurations, and performance metrics.

1. Logistic Regression

Logistic Regression was implemented as a baseline model due to its simplicity and interpretability. Categorical features were encoded numerically, and the model was trained with regularization techniques to handle potential overfitting. Despite its efficiency, the model achieved a moderate accuracy of **91.62%**, indicating its limitations in capturing complex relationships within the data.

2. Random Forest

The Random Forest algorithm, an ensemble learning technique, was employed for its ability to handle nonlinearity and high-dimensional data. Using 100 trees with optimized maximum depth and minimum samples per split, the model demonstrated robust performance, achieving an accuracy of **96%**.

3. Gradient Boosting

Gradient Boosting was utilized to iteratively minimize prediction errors by building sequential models. Hyperparameters such as learning rate and the number of estimators were fine-tuned to prevent overfitting while maintaining efficiency. The model achieved an accuracy of **95.77%**, reflecting its capability to adapt to the dataset's complexity.

4. XGBoost

XGBoost, a scalable and efficient gradient boosting framework, was implemented with advanced regularization techniques. Learning rate, max depth, and subsample ratio were optimized to balance accuracy and computation time. This model achieved an accuracy of **95.62%**, making it a strong contender for crime prediction tasks.

5. Support Vector Machine (SVM)

The SVM algorithm, configured with a radial basis function (RBF) kernel, was employed to classify the data by maximizing the margin between classes. Cost and gamma parameters were tuned to enhance generalization. The SVM model achieved an accuracy of **96.26%**, indicating its effectiveness for linear and mildly non-linear data.

6. K-Nearest Neighbor (KNN)

The KNN algorithm classified crime types by analyzing the k-nearest data points in feature space. The optimal value of k was determined experimentally to reduce bias-variance trade-off. The model achieved an accuracy of **93%**, demonstrating good performance for neighborhood-based classification.

7. Naive Bayes

Naive Bayes, based on Bayes' theorem with an assumption of feature independence, was implemented for its simplicity. Gaussian Naive Bayes was selected to handle numerical features. Despite its computational efficiency, the model achieved a relatively lower accuracy of **90%**, likely due to the independence assumption being violated in the data.

8. Linear Discriminant Analysis (LDA)

LDA was employed to find a linear combination of features that best separates the classes. The model was trained on scaled features to ensure proper class separation and achieved an accuracy of **93.05%**, confirming its ability to handle linearly separable data.

9. Decision Tree

The Decision Tree model was designed to partition data into subsets using feature-based splits. To avoid overfitting, the tree depth and minimum samples per split were restricted. The model achieved an accuracy of **93.11%**, offering interpretability and reasonable predictive power.

10. Multi-Layer Perceptron (MLP)

The MLP neural network was implemented with multiple hidden layers, ReLU activation functions, and an adaptive learning rate. Hyperparameter tuning was extensively performed, including optimization of hidden layer sizes and learning rate schedules. The MLP achieved the highest accuracy of **99.68%**, with precision, recall, and F1-scores exceeding **99%**, making it the most reliable model for this study.

11. AdaBoost

The AdaBoost algorithm was applied to strengthen weak classifiers iteratively. Hyperparameters such as learning rate and the number of estimators were adjusted to maximize performance. This model achieved an accuracy of **92%**, demonstrating competitive results in crime classification tasks.

Discussion of Results

Among all the models implemented, the **Multi-Layer Perceptron (MLP)** emerged as the most effective, capturing intricate relationships within the dataset and achieving outstanding metrics across all evaluation criteria. Ensemble methods such as **Random Forest** and **XGBoost** also performed well, albeit with slightly lower accuracies. Simpler models like **Naive Bayes** and **Logistic Regression** provided baseline performance but struggled with the dataset's complexity.

This comprehensive exploration of models highlights the importance of aligning model capabilities with dataset characteristics, further validated by extensive hyperparameter tuning that enhanced model performance across the board.

=== Tuned Model Performance Summary ===

	Model	Train Accuracy	Test Accuracy	Precision	Recall	F1 Score	ROC AUC
0	Logistic Regression	0.915989	0.916207	0.909384	0.828005	0.866788	0.964491
1	Random Forest	0.999379	0.960884	0.960192	0.919308	0.939305	0.993198
2	Gradient Boosting	0.977684	0.957713	0.944721	0.925729	0.935129	0.989768
3	XGBoost	0.964745	0.956150	0.948174	0.916934	0.932292	0.991459
4	Support Vector Machine	0.964676	0.962585	0.954604	0.930616	0.942457	0.990069
5	K-Nearest Neighbors	0.959482	0.937994	0.920573	0.888315	0.904156	0.976225
6	Naive Bayes	0.907405	0.904808	0.911176	0.787659	0.844927	0.949204
7	Linear Discriminant Analysis	0.930410	0.930548	0.949570	0.833310	0.887650	0.953956
8	Decision Tree	0.945693	0.931145	0.898327	0.891805	0.895054	0.964779
9	Multi-Layer Perceptron	0.999885	0.998621	0.998184	0.997627	0.997905	0.999936
10	AdaBoost	0.922780	0.920803	0.902486	0.851459	0.876230	0.974284

Figure 9: Metrics of each model.

Tuned Model Performance Summary:

Comparison of machine learning models based on train accuracy, test accuracy, precision, recall, F1 score, and ROC AUC. The Multi-Layer Perceptron demonstrates the highest performance with a test accuracy of 99.86%, followed by ensemble methods like Random Forest and Gradient Boosting, indicating their effectiveness in crime prediction tasks. Metrics highlight the trade-offs between precision and recall across various models.

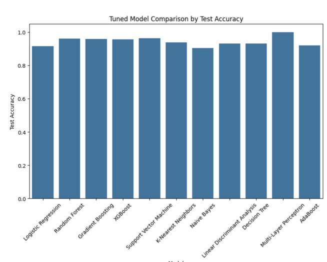

Figure 10: Accuracies of each model.

Conclusion and Result

Comparison of machine learning models based on train accuracy, test accuracy, precision, recall, F1 score, and ROC AUC. The Multi-Layer Perceptron demonstrates the highest performance with a test accuracy of 99.86%, followed by ensemble methods like Random Forest and Gradient Boosting, indicating their effectiveness in crime prediction tasks. Metrics highlight the trade-offs between precision and recall across various models.

Key Insights and Discoveries:

Crime Patterns and Hotspot Areas: Central LA and 77th Street repeatedly showed up as the most crime-ridden locations, especially for vehicle-related offenses like theft and burglary. Visual tools such as heatmaps and bar charts pinpointed these areas as key hotspots, assisting law enforcement in prioritizing their interventions.

Temporal Patterns: Crime rates exhibited significant temporal trends, peaking between 12 PM and 8 PM, while showing fewer incidents in the early morning hours. This understanding is useful for enhancing police patrol timing and allocating resources during peak crime periods.

Crime Severity and Demographic Factors: Crimes related to vehicles made up most of the occurrences, followed by assaults and burglaries. Victim demographics indicated a focus within the 25–35 age range, emphasizing particular groups that need specialized safety initiatives.

Model Effectiveness and Crime Forecasting: The Multi-Layer Perceptron (MLP) surpassed other models, attaining an impressive test accuracy of 99.68%, along with outstanding metrics in precision, recall, and ROC AUC. Ensemble techniques such as Random Forest and Gradient Boosting have also shown strong performance, reinforcing the effectiveness of machine learning for predicting crime.

Significance and Findings:

These results offer a strong foundation for:

Proactive Crime Prevention: Anticipating areas and times with high crime rates allows law enforcement to implement preventive strategies, effectively lowering crime levels.

Data-Driven Decision Making: Understanding victim demographics, crime hotspots, and time-based trends assists policymakers in creating focused safety initiatives and improving resource distribution.

Urban Planning and Policy Development: Illustrating crime trends assists urban planners and policymakers in tackling socio-economic and environmental issues that lead to criminal behaviors.

Future Improvements:

Although the project met its goals, there are numerous opportunities for enhancement and further investigation:

Data Granularity: Including more detailed datasets, like specific offender demographics or socio-economic indicators, may reveal more profound correlations and insights.

Real-Time Crime Assessment: Establishing instant data streaming and analysis functionalities would allow for swift action regarding rising crime patterns.

Improved Modeling Approaches: Investigating sophisticated deep learning frameworks or ensemble methods might enhance prediction precision and reliability.

Wider Focus: Incorporating comparative studies with different cities or areas may yield significant benchmarks and reveal distinctive crime trends.

Potential Use Cases:

The findings derived from this project have a variety of uses, such as:

Law Enforcement Enhancement: Positioning patrol units and resources in high-crime zones during critical times informed by predictive analytics.
Community Involvement: Creating safety initiatives focused on the requirements of at-risk groups.
Urban Development: Guiding zoning and infrastructure strategies to address socio-environmental factors that lead to crime. **Policy Development:** Creating data-informed policies to improve public safety, tackle disparities, and distribute resources more efficiently.

3 References:

1. S. V. S, J. Retna Raj, A. A, S. Srinivasulu, Gowri and Jabez, "Crime Analysis Framework for Predicting Criminal Behavioral Patterns with Machine Learning," 2023 IEEE Renewable Energy and Sustainable E-Mobility Conference (RESEM), Bhopal, India, 2023, pp. 1-5, doi: 10.1109/RESEM57584.2023.10236417. keywords: Training;Support vector machines;Renewable energy sources;Machine learning algorithms;Urban areas;Data visualization;Prototypes;Machine learning;Crime prediction;Data analysis;Security and privacy;Crime analysis;Crime pattern,

2. M. V. R. S. Harsha, M. H. S. Akhil, R. D. A, A. M. J and Y. F. A, "Unravelling State-wise Crime Patterns with Ensemble Machine Learning and Data Visualization for Public Safety," 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), Namakkal, India, 2024, pp. 1-5, doi: 10.1109/AIMLA59606.2024.10531626. keywords: Visualization;Data visualization;Machine learning;Predictive models;Public security;Data models;Ensemble learning;ensemble machine learning;data visualization;crime patterns;public safety;detection,

3. K. Vinothkumar, K. S. Ranjith, R. R. Vikram, N. Mekala, R. Reshma and S. P. Sasirekha, "Crime Hotspot Identification using SVM in Machine Learning," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 366-369, doi: 10.1109/ICSCDS56580.2023.10104689. keywords: Support vector machines;Training;Machine learning algorithms;Data visualization;Machine learning;Pressing;Prediction algorithms;Crime hotspot;Machine Learning;SVM;Crime Data;Crime Prediction,

4. P. Tamilarasi and R. U. Rani, "Diagnosis of Crime Rate against Women using k-fold Cross Validation through Machine Learning," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, pp. 1034-1038, doi: 10.1109/ICCMC48092.2020.ICCMC-000193. keywords: Support vector machines;Machine learning algorithms;Data analysis;Linear regression;Government;Decision making;Machine learning;Crime Rate;Feature scaling;k-fold cross validation;Machine Learning Algorithms,

5. D. M S and S. Shankaraiah, "Crime Analysis and Prediction using Machine Learning Algorithms," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-7, doi: 10.1109/MysuruCon55714.2022.9971801. keywords: Machine learning algorithms;Linear regression;Clustering algorithms;Predictive models;Prediction algorithms;Market research;Classification algorithms;Crime Analysis and Prediction;Naive Bayes;Linear Regression;K-means Clustering,

6. R. S. Kumar, N. P. Saravanan, K. N. Devi, P. Jayanthi, N. Krishnamoorthy and S. Karthi, "Empirical Analysis on Crime Prediction using Machine Learning," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICCCI56745.2023.10128304. keywords: Support vector machines;Radio frequency;Technological innovation;Machine learning algorithms;Measurement uncertainty;Predictive models;Prediction algorithms;Crime;Machine Learning;Support vector;Naïve Bayes;Logistic Regression,

7. P. Sarzaeim, Q. H. Mahmoud and A. Azim, "A Framework for LLM-Assisted Smart Policing System," in IEEE Access, vol. 12, pp. 74915-74929, 2024, doi: 10.1109/ACCESS.2024.3404862. keywords: Predictive models;Task analysis;Law enforcement;Data models;Artificial intelligence;Recurrent neural networks;Adaptation models;Few-shot learning;Large language models;Crime prediction;fine-tuning;few-shot prompting;large language models;LLM;zero-shot prompting

8. D. R, K. M, T. C, A. L, M. M. D. J. T. Hansika and S. S. Chandrasiri, "The Future of Crime Prevention: Police Case Analysis Using Machine Learning," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023, pp. 454-459, doi: 10.1109/ICAC60630.2023.10417168. keywords: Text analysis;Law enforcement;Machine learning;Predictive models;Public security;Faces;Accidents;Police case analysis;Accident analysis;Crime analysis;Case document classification;Crimes against women;Future prediction;Machine learning models;Forecast;Pattern,