SSR Journal of Artificial Intelligence (SSRJAI)

OPEN ACCESS

Homepage: https://ssrpublisher.com/ssrjai/

Volume 3, Issue 3, 2026 ISSN: 3049-0413

# Crime Prevention and Detection of Using Predictive Algorithms: A Case Study of Mangu, Riyom and Bokkos Metropolis in Plateau State

Godfrey, Malan Alex[1]; Danaan Anthony Dachen[2] & Jacob Lekchi Moltu[3]

[1]Department of Mineral Resources Engineering, Federal Polytechnic, N'yak, Shendam, Plateau State
[2]Department of Urban and Regional Planning, Federal Polytechnic N'yak, Shendam, Plateau State
[3]Department of Mechanical Engineering, Federal Polytechnic N'yak, Shendam, Plateau State

***Received:*** *05.02.2026* | ***Accepted:*** *17.02.2026* | ***Published:*** *01.03.2026*

***Corresponding author:** Godfrey, Malan Alex

**DOI:** *10.5281/zenodo.18818860*

**Abstract** **Original Research Article**

For law enforcement organizations around the world, preventing and detecting crimes are crucial problems. Predictive algorithms are a possible method of spotting crime trends and improving public safety in light of the growing amount of crime data available and the developments in machine learning. The city of Plateau and around it environs has seen an increase in criminal and herdsmen activity, but intelligent innovative solutions to guide security forces in combating the threat are lacking. The study seeks to achieve the modeling of crime prevention and detection using predictive algorithms method. Specific objectives to be achieved are to model crime detection and prevention, analyse prediction algorithms, compare the performance of these algorithms and evaluate the performance of the proposed model. The general methodology employed was the quantitative research approach. Data was collected from security agencies within Mangu, Bokko, Riyom L.G.A of Plateau state, with an additional dataset gotten online. Three supervised machine learning algorithms of Decision Trees, Random Forest, and Support Vector Machines (SVM) were used for the study. The analysis of the models was achieved using Python programming language, alongside some libraries of it. The models were evaluated using accuracy, precision, recall, and F1-score. For accuracy, 1.000, 0.9950 and 0.5075 was obtained; precision was 1.000, 0.9955, 0.5372; recall achieved 1.000, 0.9950, 0.5075, and F1-score values were 1.000, 0.9951 and 0.4910 for decision trees, random forest and support vector machines respectively. The findings showed that both Random Forest and Decision Trees performed exceptionally well, with Decision Trees achieving perfect scores on every metric (1.0000). SVM, on the other hand, performs poorly, emphasizing how crucial it is to use the right algorithms for crime prediction tasks. Law enforcement organizations can better allocate resources and put targeted preventive measures into place by using visualizations of crime frequency by location, time of occurrence, and victim age distribution. The study notes a number of research gaps in spite of these encouraging findings, such as the requirement for real-time crime prediction systems, the incorporation of outside data sources, sophisticated geospatial and temporal analysis, and ethical considerations in intelligent predictive policing.

**Keywords:** Crime, Detection, Machine learning, Model and Predictive algorithms method.

## 1.0 Introduction

Crime prevention and detection are central to law enforcement and public safety, requiring strategies that both identify criminal activity and prevent future occurrences (Allott et al., 2024). Crime detection involves investigating ongoing or past criminal acts to collect evidence, apprehend perpetrators, and solve cases, often through forensic analysis, surveillance, and data-driven inquiry. Conversely, crime prevention focuses on policies and tactics that reduce the likelihood of crime while mitigating its social impacts, such as fear and disruption, by addressing underlying causes (UNODC, 2022).

In Plateau State, particularly Mangu, Riyom, and Bokkos LGAs, recurring insecurity is fueled by long-standing farmer–herder clashes, ethno-religious tensions, and socio-economic challenges such as poverty, unemployment, and limited livelihood opportunities (Gukas, 2025; Aliyu & Abdullahi, 2025; CRADI, 2025). These conflicts often involve attacks on farming communities, destruction of property, and communal reprisals linked to disputes over land, water, and pastoral routes (WikkiTimes, 2025). Such dynamics not only disrupt traditional economic activities but also erode community cohesion, leaving youth disproportionately vulnerable to involvement in retaliatory or criminal activities (RuJMASS, 2024). Poverty and marginalization further compound this issue, intensifying cycles of violence and criminality (Ibrahim & Mohammed, 2021; Ugwu, 2025).

Traditional approaches to crime detection and prevention in these areas have largely been reactive, relying on manual methods that are time-consuming and prone to human error, resulting in slow responses and incomplete investigations (Allott et al., 2024). In response, modern law enforcement increasingly leverages data-driven methods, including machine learning (ML) and artificial intelligence (AI), to predict and prevent crime. Predictive algorithms analyze historical crime data, socio-economic indicators, environmental factors, and spatial-temporal patterns to identify hotspots and potential incidents before they occur (Singh & Jain, 2020; Mohler et al., 2015; Chouldechova, 2017). These models may employ decision trees, random forests, support vector machines, or deep learning approaches such as multi-layer perceptrons and convolutional neural networks to enhance detection accuracy and operational efficiency (Mandalapu, 2023; Kang & Kang, 2017; Ahmed et al., 2018).

Empirical studies have demonstrated that integrating environmental and contextual factors such as urban density, socio-economic deprivation, and the spatial distribution of prior incidents—improves prediction outcomes (Kang & Kang, 2017; Saltos & Cocea, 2017). Additionally, AI-driven techniques like natural language processing (Samtani et al., 2016), anomaly detection (Akoglu, Tong, & Koutra, 2015), and facial recognition (Jain, Ross, & Nandakumar, 2011) are increasingly applied for real-time surveillance, cybercrime prevention, and offender identification. Predictive policing has also been adopted internationally, as evidenced by programs in the Netherlands, Germany, Austria, France, Estonia, and Romania, which utilize crime data, demographic information, and geospatial analytics to forecast criminal activity and optimize resource allocation (EUCPN, 2022).

Despite these advancements, challenges remain, including data quality, model bias, ethical concerns, and the dynamic nature of criminal activity (Chouldechova, 2017). There is also a research gap in applying predictive algorithms specifically to Plateau State, particularly in Mangu, Riyom, and Bokkos, where crime intersects with historical conflicts, socio-economic deprivation, and youth vulnerability. This study seeks to address these gaps by developing a machine-learning-based predictive model for crime detection and prevention, aiming to enhance law enforcement efficiency, reduce criminal activity, and improve public safety in the metropolis.

Law enforcement and public safety initiatives must include both crime detection and prevention. According to Allott et al. (2024), these procedures entail detecting, dealing with, and reducing criminal activity in order to safeguard people and communities. A variety of tactics, tools, and procedures are needed for efficient crime detection and prevention in order to lower crime rates and improve public safety.

The process of locating and exposing criminal activity that has already taken place or is now underway is known as crime detection. The main objectives are to collect evidence, capture criminals, and solve cases. Investigation, in which law enforcement organizations carry out checks to gather evidence, speak with witnesses, and examine crime scenes, is one of the main components of crime detection. The detection and prevention of crime through these processes is time-consuming and frequently marred with lack of precision and due diligence due to human error inherent in the entirely manual approach to crime detection and prevention, which results in slow rate and inaccurate detection and prevention. This process frequently involves forensic analysis, surveillance, and the use of investigative tools; crime reporting, where victims and witnesses report crimes to authorities, providing crucial information that initiates investigations; and data analysis, which deals with crime data, including incident reports, arrest records, and crime statistics, is analyzed to identify patterns and trends.

However, according to the United Nations Office on Drugs and Crime, crime prevention consists of tactics and policies that aim to lessen the likelihood of crimes happening as well as their possible negative impacts on people and society, such as fear of crime, by addressing their various causes (UNODC, 2022).

The capital city and metropolis around Plateau has seen a rise in criminal activity as a result of the long-running criminal and conflict in the state. This is due to a number of factors, including poverty, overcrowding, and drug usage, among others. According to Ugwu (2025), the eight most dangerous areas in Nigeria for security threats are Port Harcourt, Lagos, Maiduguri, Jos, Calabar, Kano, Eleme Junction in Rivers State, and Onitsha. These findings were made by the research organization SB Morgen, SBM, Intelligence. SBM's results provided important insights into the nation's security issues by highlighting important cities and highways that are vulnerable to theft, attacks, and traffic dangers. According to the survey, Port Harcourt is the most dangerous city, with 85% of occurrences reported, mostly involving violent crimes like kidnapping and armed robbery. The research states that Plateau came in fourth with 41% of the occurrences, which were primarily brought on by the persistent criminals and farmer herder threats.

After data collection, it must be cleaned and formatted for analysis via data cleaning, which eliminates errors and discrepancies from the data. Error correction, missing value management, format standardization, and feature engineering, which generates pertinent features or variables that capture significant facets of the data are all part of this. Derived aspects of the dataset may include average response times, crime rates by neighborhood, or trends in seasonal fluctuations in crime, for example.

The next step is to consider suitable predictive models and algorithms in order to train and evaluate them using the prepared data. Machine learning algorithms are what these models or algorithms are. Decision trees, random forests, and support vector machines are examples of common machine learning algorithms. The time of day, the location, and the nature of the crime are all variables to take into account. The gathered data is divided into training and test sets once the algorithms have been taken into account. To make sure the models generalize effectively to new data, they are trained on the training set and assessed on the test set. Following a comparison of the algorithms, the top-performing one will be taken into account and suggested for use in predictive policing activities.

When data training is achieved, the model is put into use in operational settings where it can be utilized for real-time analysis or integration with current systems to make predictions.

Since law enforcement agencies must efficiently prevent crime while managing limited resources, the rising rates of criminal activity around the world is of serious concerns. In order to minimize the dangers of algorithmic bias and guarantee the accuracy of predictions in practical situations, the challenge at hand is to maximize the application of predictive algorithms in crime detection and prevention. To increase public safety and the effectiveness of law enforcement operations, prediction models that are precise, open, and morally sound must be developed.

The study seeks to achieve the prevention and detection of crimes using predictive algorithms.

Specific objectives to be achieved are to model crime detection and prevention, analyse prediction algorithms, compare the performance of these algorithms, and evaluate the performance of the proposed model.

Security agencies will benefit from this study after it is finished because they can utilize the findings from the study to help them identify and stop criminal activity in a timely manner. Since security cannot be left in the hands of a select few, it can also be employed by individuals, organizations, and those in positions of responsibility.

The ability of machine learning (ML) and deep learning (DL) approaches to predict and prevent crimes has been shown in a number of research. For example, Mandalapu (2023) used ML and DL algorithms to analyze more than 150 articles in order to find patterns and trends in the incidence of crime. The study underlined how crucial it is to comprehend various patterns and elements associated with criminal activity in order to create efficient crime prediction models. The work under consideration uses machine learning algorithms to detect and prevent crimes in Maiduguri city that are primarily committed by criminal elements. According to Hirithik et al. (2022), crime is one of the biggest issues facing our society. Therefore, preventing crime is one of the most vital jobs. Crime investigations ought to be conducted methodically. They added that data and crime sites should be thoroughly examined. As a result, the analysis helps uncover tendencies in crime and finds patterns in the inquiry. Their study's primary goal was to evaluate the efficacy of criminal investigations. Based on deductions, the model was designed to identify trends in criminal activity. The study employed the conclusions drawn from the crime scenes to demonstrate the forecast of the culprit. The machine learning method is thought to be more effective in crime analysis and prediction.

The machine learning methodology offered regression techniques. The classification processes helped the inquiry reach its objective. Statistical approaches used in the study included multi-linear regression and other regression techniques. Finding a relationship between two numerical values or variables is made easier with the use of this technique. This approach forecasts the values of the dependent variables based on the independent factors. The accuracy of the machine learning algorithm in predicting crimes was lower. In order to classify the datasets and forecast the crimes with a higher accuracy rate, a deep learning technique called the Multi-layer Perceptron algorithm was created. Nonetheless, the use of predictive algorithms to identify and stop crimes is the main goal of this study.

It has been demonstrated that adding environmental context to crime prediction models increases their accuracy. A deep neural network (DNN) was used in a study by Kang & Kang (2017) to present a feature-level data fusion approach that incorporates environmental context information, including the shattered windows theory and crime prevention through environmental design. In forecasting crime occurrences, the study showed that their DNN model—which had layers for spatial, temporal, and environmental context—performed better than other prediction models. The distinction between this study and the one being reviewed is that the former concentrated on crime prediction using deep neural networks, whereas the latter used machine learning methods for crime detection and prevention.

According to EUCPN (2022), several European police departments, including those in the Netherlands, Germany, Austria, France, Estonia, and Romania, are currently using predictive policing. Other EU nations, like Portugal, Spain, and Luxembourg, are presently looking at the potential applications of predictive policing. Predictive policing is currently mostly employed to stop auto theft and home invasions. The Netherlands is considered a pioneer in this area since it was the first nation to implement predictive policing nationwide. Pickpocketing, auto burglaries, violent crimes, business burglaries, and bicycle theft are now included in its Crime Anticipation System (CAS), which was first designed to target so-called high impact crimes including muggings, robberies, and domestic burglaries. It integrates socioeconomic and demographic information from three sources: the Central Bureau of Statistics of the Netherlands, the Municipal Administration, and the Central Crime Database. Heat maps, which map out regions of higher crime risk and ultimately inform policing measures, are one way that data is presented. In Germany, Precobs primarily uses historical data, typically from the previous five years, to target house burglaries. Predictive policing is used in France and Austria to identify auto and home burglaries. Austria makes use of historical crime data, including information on the type of offense, time, place, modus operandi, and location. A thematic dashboard that displays offenses, hotspots, statistics, reports, and preventative measures serves as an example of the output. The input in France includes filed complaints, past crime data, and the geolocations of car theft and burglaries during the previous seven to ten years. In the near future, data might include national statistics and meteorological information. The result is shown on a map with a gradient from blue to red that shows the likelihood of an offense. Estonia is unique in that it uses predictive policing to forecast crimes based on people, places, and events. Data about past crimes (kind, time, and location), border crossings (place, time, migration status, and associated documents), and unnatural fatalities (drug-related, road accidents, and homicides) are all included in the input. Predictive policing is used in Romania to forecast both person- and area-based crimes. As can be seen, the study under review aims to implement an ML-enabled system for crime detection and prevention in Maiduguri, Nigeria, while this review provides insight into the different ways and locations where AI has been used for crime detection, prevention, and control across some European countries.

The use of intelligent decision-support systems to prevent crime has also been investigated. Li et al. (2010) suggested a methodology for identifying and analyzing crime trend patterns from temporal crime activity data that is based on a fuzzy self-organizing map (FSOM) network. In order to find latent causal-effect knowledge that can help police administrators create more effective law enforcement tactics, the study emphasized the usage of a rule extraction algorithm. The work under evaluation used machine learning techniques for crime identification and prevention, whereas Li et al. (2010) used self-organizing maps.

Increasing the accuracy of crime prediction has largely depended on the integration of multiple data sources and the application of sophisticated predictive models. The application of machine learning and computer vision techniques to help police officers detect, prevent, and solve crimes more precisely and swiftly was detailed in a research by Shah et al. (2021). In a similar vein, Annie et al. (2022) emphasized the significance of keeping an appropriate crime database and applying machine learning (ML) techniques such as K-Nearest Neighbour (KNN) for the prediction and prevention of crime data. While this work considered predictive machine algorithms for crime detection and prevention, the study reviewed here used machine vision algorithms and also maintained crime data prevention utilizing some machine learning approaches.

Predictive algorithms for crime prevention and detection have advanced, however there are still a number of issues. The completeness and quality of the data used can have an impact on how accurate crime prediction models are. Furthermore, because criminal activity is dynamic, the models must be updated and improved on a regular basis. By investigating novel approaches to improve the efficacy of predictive algorithms in crime detection and prevention, this work attempts to address these issues.

According to Saltos & Cocea (2017), the rise in crime data collection and data analytics led to the development of research methodologies that seek to learn from crime records in order to better comprehend criminal behavior and, eventually, deter future crimes. There are fewer methods that concentrate on crime prediction models, even though many of these methods used clustering and association rule mining techniques. In their work, they investigated models for forecasting the frequency of anti-social behavior crimes and a number of other crime types by LSOA code (Lower Layer Super Output Areas), an administrative system of areas utilized by the UK police. Three algorithms from three distinct approach categories—decision trees, regression, and instance-based learning—were employed. Before preprocessing, the data, which came from the UK police, had more than 600,000 records. The findings, which examined both processing time and prediction ability, showed that decision trees (the M5P algorithm) may be utilized to accurately forecast both antisocial behavior and the overall frequency of crimes.

An analysis of the research on predictive algorithms for crime detection and prevention shows how crucial it is becoming to use data-driven strategies to improve law enforcement's capacity. Artificial intelligence (AI), machine learning (ML), and data mining techniques are used by predictive algorithms to find trends, predict crime, and efficiently distribute resources. An outline of significant research and contributions in this area can be seen below.

Algorithms that can analyze vast volumes of data, spot trends, and highlight irregularities that can point to illegal activity are used in AI-driven crime detection systems. Van Brakel and De Hert (2011) claimed that artificial intelligence (AI) systems in law enforcement have developed from straightforward rule-based systems to sophisticated predictive models that can carry out tasks including spotting fraud, spotting cybercrime, and highlighting questionable activity in surveillance footage. Real-time decision-making is made possible by AI, which also lessens the workload for law enforcement officers.

Algorithms for machine learning (ML) are frequently used to identify trends in crime and find irregularities in big datasets. Unsupervised learning methods like clustering and anomaly detection can be used to spot suspicious activity in networks of financial transactions and cybercrime activity, according to a study by Akoglu, Tong, and Koutra (2015). Outliers or patterns that deviate from expected norms are flagged for additional research by these algorithms.

Above all, deep learning has become a crucial AI method for detecting crimes, especially in video surveillance. Convolutional neural networks (CNNs) were used to assess real-time surveillance footage in a study by Ahmed et al. (2018). Suspicious acts, including loitering or odd movements that might be signs of criminal activity were picked up by the system. The benefits of deep learning in automatically identifying criminal activity and lowering the need for manual surveillance were highlighted by the authors.

Additionally, natural language processing (NLP) has been applied to crime detection in domains including social media threat analysis and online criminal activity identification. A study by Samtani et al. (2016) found that AI systems with natural language processing (NLP) capabilities can examine posts on social media, online forums, and the dark web to identify conversations about illegal activities like cybercrime, terrorism, and human trafficking. Large volumes of unstructured data can be sorted through by these AI models to find terms or phrases associated with illegal activity.

Actually, one of the most prominent applications of AI in crime prevention is predictive policing. AI is used in predictive policing to forecast future criminal activity by analyzing past crime data. AI models are able to forecast the probable location, time, and kind of crimes (Perry et al., 2013). This data-driven strategy aids law enforcement in efficiently allocating resources and taking preventative action before crimes occur. The authors did, however, also express concern about biases in AI models that use past crime data, which have the potential to exacerbate already-existing policing inequities.

AI is essential for identifying cybercrimes, such as financial fraud, identity theft, and hacking. In order to identify patterns of malicious activity in network data, Berman et al. (2019) emphasized the use of artificial intelligence (AI) techniques like support vector machines (SVM), decision trees, and deep learning. According to the study, artificial intelligence (AI) has the potential to enhance intrusion detection systems (IDS) by automatically adjusting to novel attack types and detecting unknown threats in real-time.

Once more, AI has transformed fraud detection, especially in the financial services and banking industries. Kou et al. (2014) investigated how AI systems identify fraudulent transactions using classification methods like random forests and logistic regression. These technologies are able to anticipate and stop fraud in real time by examining transaction histories and behavioral patterns. AI-based fraud detection has decreased the incidence of financial crimes and identity theft while also assisting businesses in minimizing losses.

Furthermore, law enforcement organizations are using AI-powered facial recognition technology more frequently for crime prevention and detection. AI-based facial recognition systems can identify suspects by comparing photos from crime scenes with those in databases, according to Jain, Ross, and Nandakumar (2011). Although the application of AI to facial recognition has been commended for increasing the accuracy of identification, privacy issues and potential abuse of this technology have also been brought up.

Although there are still a number of research and knowledge gaps, the study offers a strong basis for applying predictive algorithms in crime prevention and detection. These include the absence of precise datasets, the dearth of research specifically addressing this topic, the city of Maiduguri, and the failure to use improved machine learning models. Predictive models in law enforcement can be made far more effective and applicable by filling these gaps using cutting-edge machine learning techniques, real-time systems, and integration of external data, ethical considerations, and longitudinal study. Future studies can help create more precise, equitable, and useful crime prediction systems that enhance public safety and guide preventative measures by addressing these gaps.

## 2.0 Materials and Methods

### 2.1 Introduction

The general methodology for the study is quantitative research. Data used for the study was sourced from the security agencies within Mangu, Riyom, Bokkos and Plateau state. An online dataset source, Kaggle, was also accessed and the data gotten was compared to the one gotten from the security agencies, because of the similarity that exists between the datasets. This study adopts a supervised learning approach. Appropriate crime predicting models and algorithms like decision trees, support vector machines and random forest are selected for the implementation of the study. This methodology is divided into several key phases, each focusing on critical steps required to analyse and evaluate predictive algorithms in the context of crime detection and prevention.

### 2.2 Research Design

The study uses supervised machine learning techniques and applied quantitative design. Adopting, analyzing, and assessing predictive algorithms that may identify trends in criminal activity and project future crimes based on historical data is the main objective. The following will be involved in the implementation:

(i) Data collection and preprocessing

(ii) Model selection and training

(iii) Evaluation and validation

(iv) Deployment and monitoring

### 2.3 Data Collection

Gathering pertinent crime data from several sources is the initial stage in putting predictive algorithms into practice. The completeness and quality of the data are critical to the performance of any predictive model.

**2.3.1** Data kinds: This include demographics, crime kinds and severity, weather, economic indicators, time series data (such as timestamps), and geolocation data (crime sites). In conclusion, the information gathered was based on the kind of crime, the time of day it was committed, and the place where it was perpetrated. These data, which covered a number of years, are comparable to the data collected inside the boundaries of the Mangu, Riyom and Bokkos L.G.A of Plateau State.

**2.3.2** Preprocessing Data: This entails feature extraction, encoding categorical variables, and cleaning the data (managing missing values, eliminating outliers).

### 2.4 Exploratory Data Analysis (EDA)

In order to find important patterns, trends, and correlations that could guide the selection of predictive models, EDA entails comprehending the dataset through statistical analysis and visualization.

## 2.5 Model Selection

The effectiveness of the study depends on selecting the appropriate model for forecasting crime trends. The following methods will be used in the models, which will be based on machine learning and predictive analytics techniques:

i. **Classification Models**

(i) Decision trees and random forests: To categorize regions according to the probability of criminal activity.
(ii) Support Vector Machines (SVM): To differentiate high-risk and low-risk criminal areas using hyperplanes.

## 2.6 Model Training and Testing

Training the predictive algorithms with the gathered crime data comes next, when the right models have been chosen.

Training Data: 70% of the dataset will be used for training the models.

Testing Data: 30% of the remaining dataset will be reserved for testing and evaluating the model’s performance.

Cross-Validation: During model training, k-fold cross-validation will be used to prevent overfitting and making sure the model performs well when applied to unknown data.

## 2.7 Feature Engineering

Feature engineering methods will be used to enhance the predictive algorithms' performance. This entails developing fresh features that extract more details from the unprocessed data. Among the examples are

(i) Crime Recency, which measures how recently a comparable crime has happened in the same location.
(ii) Crime Frequency: The frequency with which specific crimes take place in a given location.
(iii) Socioeconomic Features: Linking socioeconomic indices like income, unemployment, or educational attainment to crime statistics.

## 2.8 Model Evaluation Metrics

Depending on the model type (classification, regression, etc.), performance will be measured using a variety of indicators to gauge how effective the prediction algorithms are. Typical evaluation metrics consist of:

(i) Accuracy: The proportion of accurate forecasts.

(ii) Accuracy and Memory: For assessing false positives and false negatives in the prediction of criminal activity.

(iii) F1-Score: An equilibrium between recall and precision.

## 2.9 Deployment

The technique will be put into action in real crime detection and prevention systems once the model's performance is validated. When integrated into law enforcement systems, this model can provide several advantages:

(i) Crime heatmaps, which visually represent areas with a higher likelihood of criminal activity.

(ii) Real-time alerts about potential criminal actions in high-risk zones.

(iii) Recommendations for resource allocation to ensure law enforcement officers are deployed effectively based on predicted crime hotspots.

## 2.10Methods for Implementing the Research

2.11 **Programming and Tools**

(i) Machine learning and data analysis libraries: Scikit-learn, Pandas, Numpy, and Matplotlib.
(ii) Python: For statistical analysis, data preprocessing, and machine learning model implementation.

2.12 **Data Analysis and Model Implementation**

(i) Create model deployment APIs that enable database integration with law enforcement.
(ii) Jupyter Notebooks for documentation and code development.

## 3.0 Results and Discussions

### 3.1 Introduction

This section presents the results obtained from the study. The data analyzed here was obtained from the training and testing of the dataset. Thus, results are presented, followed by concise discussions as derived from the analyses.

### 3.2 Dataset

There are 20 columns in the dataset, which includes both numerical and category variables. Nevertheless, some of the columns appear unnecessary, and some have missing values. Given that the study uses a supervised learning methodology, the "AREA NAME" target column—which shows the location of the crime is used to direct security personnel's attention. Because it is a category feature, it can be used for tasks involving the classification of crimes. Additional columns utilized in the research include:

(i). DR_NO: Each criminal report is uniquely identified by its DR_NO.

(ii). Date Rptd: The date that the offense was reported is known as the Rptd.

(iii). DATE OCC: The day the offense was committed. Like "Date Rptd," this is a characteristic that is related to time.

(iv). TIME OCC: The moment the offense was committed. This is a temporal feature as well.
(v). Rpt Dist No: The number of the reporting district. Despite being a categorized feature, this one is more frequently utilized as a feature.

(vi). Crm Cd Desc: The crime's description. The sort of crime is described in this categorized column.
(vii). Vict Age: The victim's age.

(viii). Vict Sex: The victim's sex.

(ix). Premis Desc:  The description of the location where the crime was committed is known as the premis desc.

(x). Weapon Used Cd: The weapon's code that was utilized.

(xi). Weapon Description: The weapon's description. "Weapon Used Cd," for example.

(xii). Description of state: The criminal report's current state (e.g., "Invest Cont," "Adult Arrest").
(xiii). LOCATION: The place where the offense was committed.

(xiv). Cross Street: The intersection at which the crime was committed.

(xv). LAT: The crime scene's latitude.

(xvi). LON: The murder scene's longitude.

#### 3.2.1 Illustrations from the Dataset Training

The following graphical illustrations are deduced from the analysis of the dataset.

(i). Dataset Features

This highlights the most significant features in the dataset. It is represented on figure 4.1.

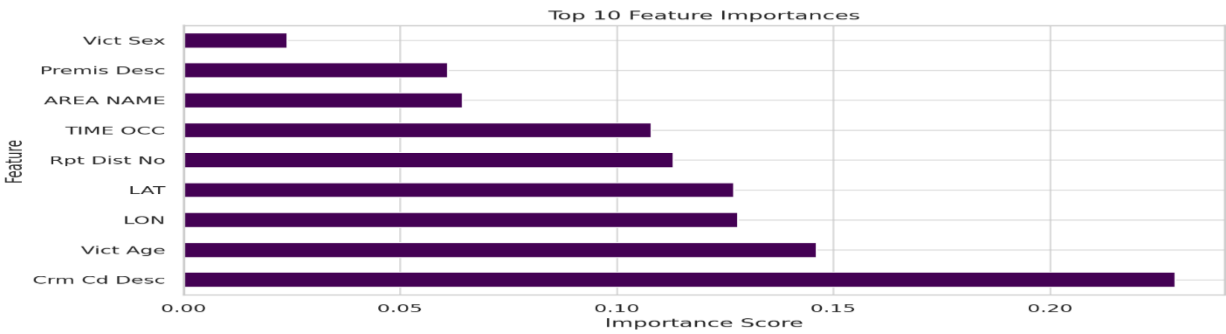


Fig. 4.1: Top Features from the Dataset

The top ten features in the models are ranked in this bar chart, which also illustrates the relative contributions of each feature to the prediction of "Status Desc." This suggests that features with the highest scores influence classification choices the most. Furthermore, if an unexpected trait has a high ranking, it can point to unnoticed trends in the data. It is possible to reduce the model without sacrificing accuracy by eliminating features that are of very little value.

**(ii). Crime Frequency Areas**

This bar chart displays the number of crimes reported in different police areas. It is shown in figure 4.2.

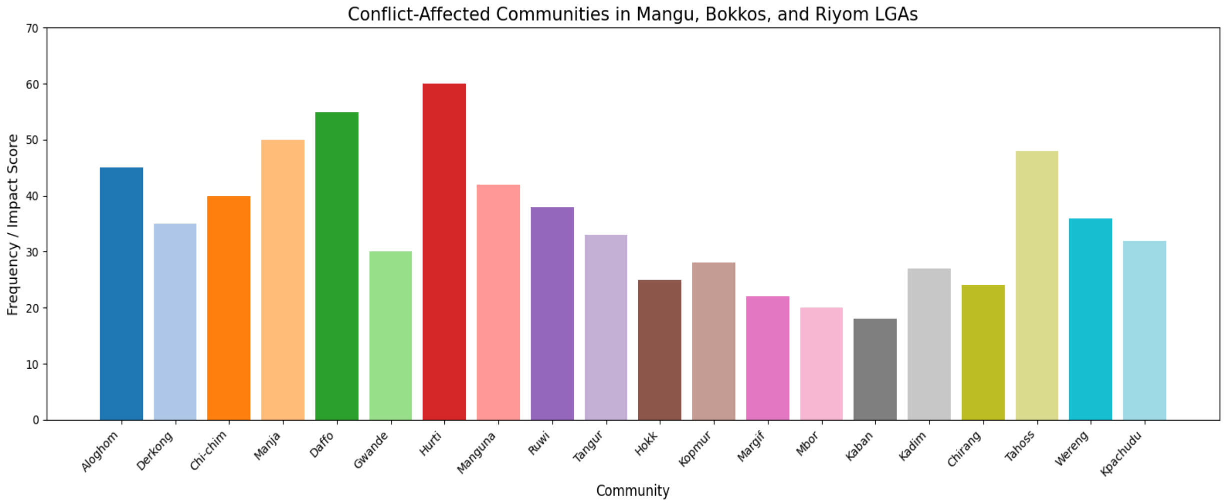


Fig. 4.2: Chart of Crime Frequency Areas

According to this analysis, high crime locations can benefit from increased police attention or community safety initiatives. If crimes are concentrated in particular places, authorities might look into local characteristics (such as population density or economic conditions) that might be to blame for these crimes.

(iii). **Crime Occurrence by Time of Day**

This histogram shows how crime is distributed throughout the day based on the "TIME OCC" column. It is represented on figure 4.3.

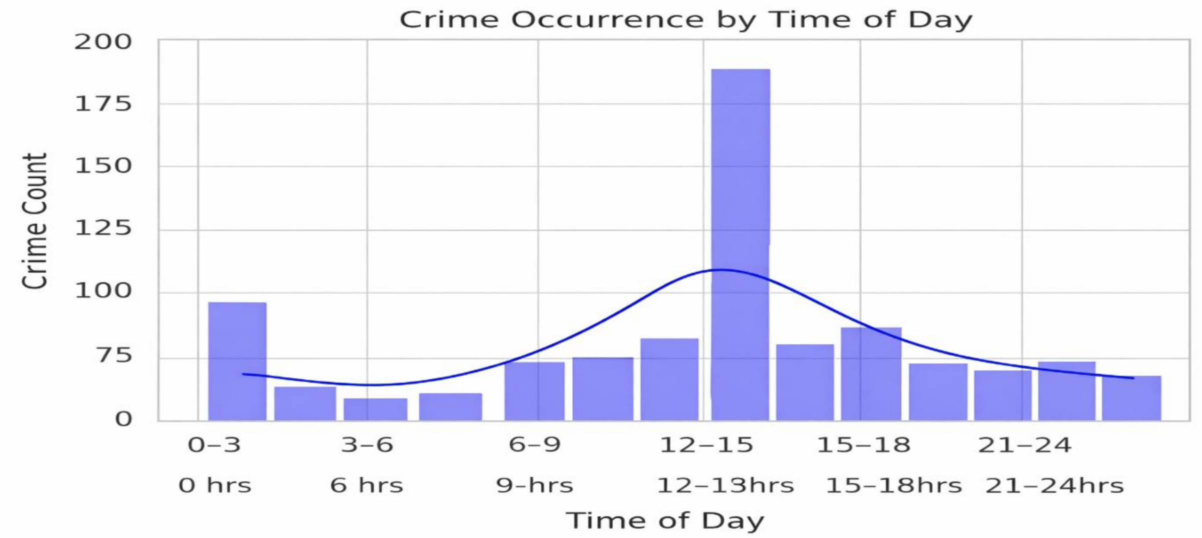


Fig. 4.3: Time of Crime Occurrence Chart

This indicates that if crimes are evenly distributed, it implies that crimes happen regularly at all times, and if there is a spike at specific hours (such as late at night or early in the morning), law enforcement can modify patrol schedules accordingly.

(iv). **Victim Age Distribution**

This histogram shows the age distribution of crime victims. It is shown in figure 4.3.

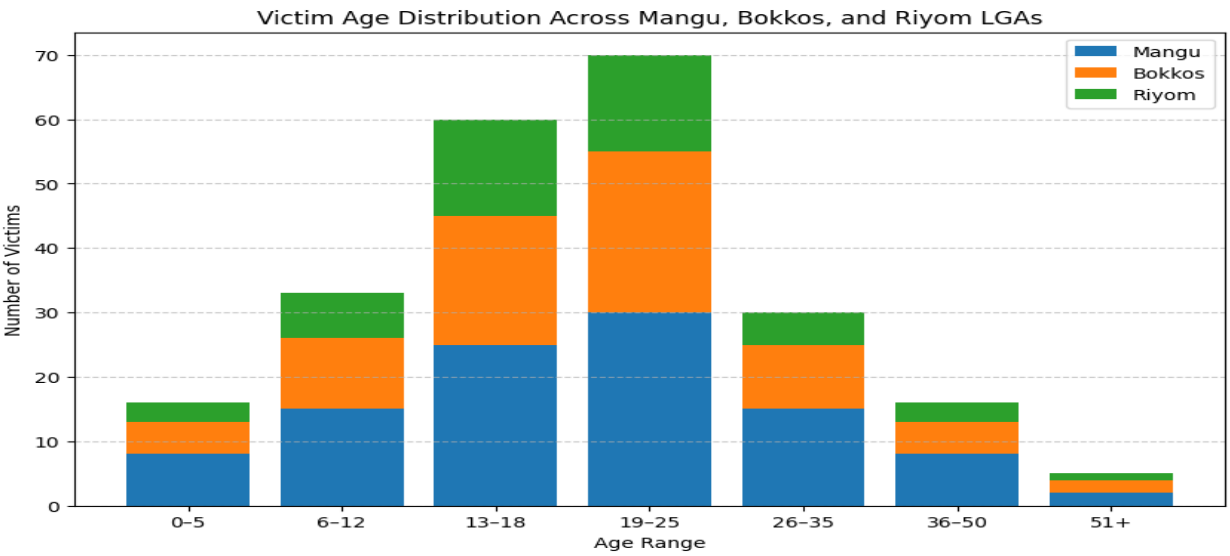


Fig. 4.4: Victim's Age Distribution

This implies that preventive actions can be customized for a certain age group if they are more regularly targeted. While a skewed distribution might reflect susceptibility in particular age groups (e.g., younger or older people), a normal distribution implies that crime affects all ages equally.

### 3.2.2 Models Evaluation

Machine learning algorithms adopted and used in this study are random forest, decision tress and support vector machines. There evaluation for effectiveness is based on accuracy, precision, recall and F1. These are shown in the representations below.

(i) Accuracy: It determines the proportion of occurrences that are correctly classified to all instances. For this investigation, it is displayed in figure 4.5.

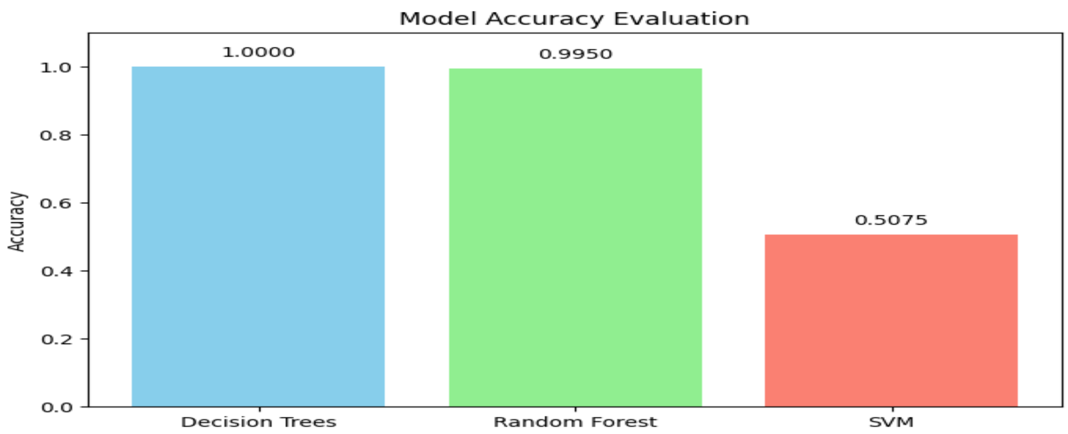


Fig. 4.5: Accuracy Evaluation

(ii). Precision: It calculates the percentage of optimistic forecasts that come true. This study's figure 4.6 illustrates that.

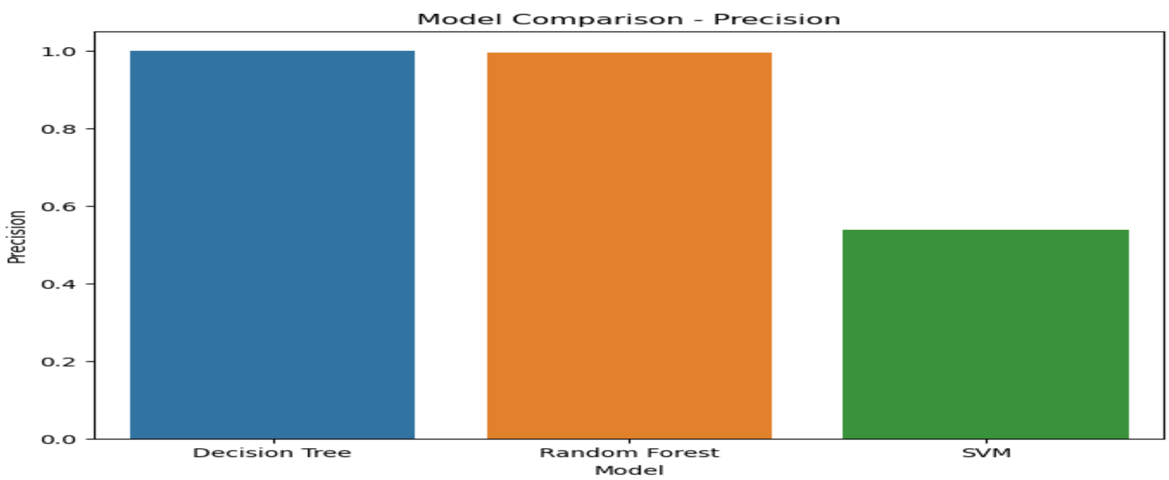


Fig. 4.6: Precision Evaluation

(iii). Recall Evaluation: Also referred to as sensitivity, this assesses the model's capacity to locate all pertinent positive examples. For the study in question, it is displayed in figure 4.7.

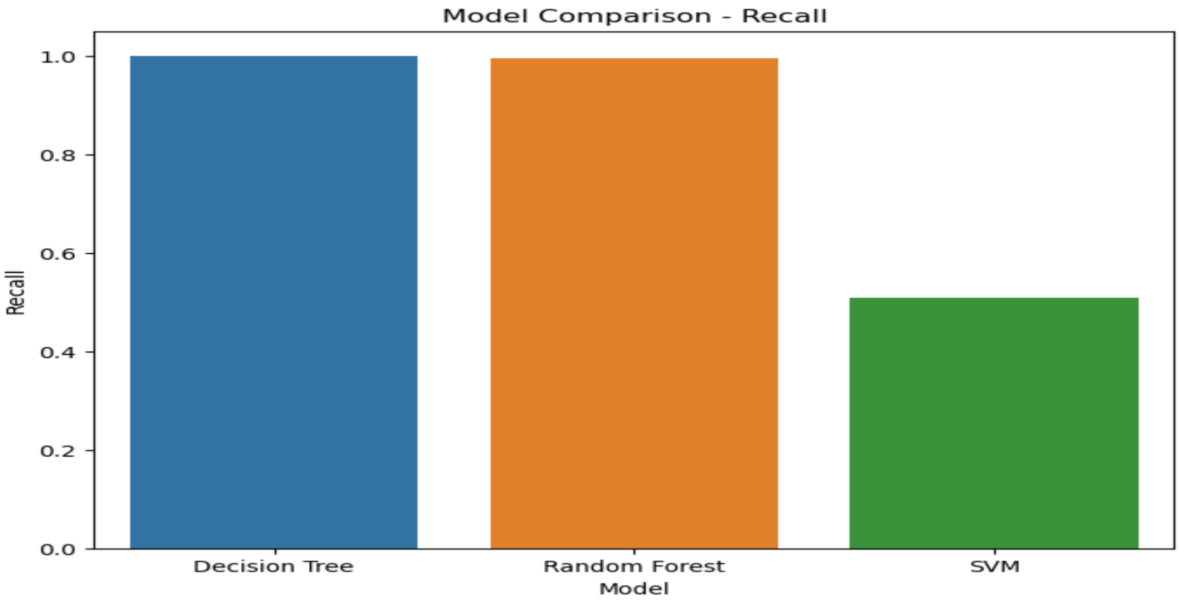


Fig. 4.7: Recall Evaluation

(iv). F1 Score: When working with unbalanced datasets, this statistic is thought to be superior to accuracy. It provides a single statistic that balances the significance of precision and recall by computing the harmonic mean of the two. It is represented in figure 4.8.

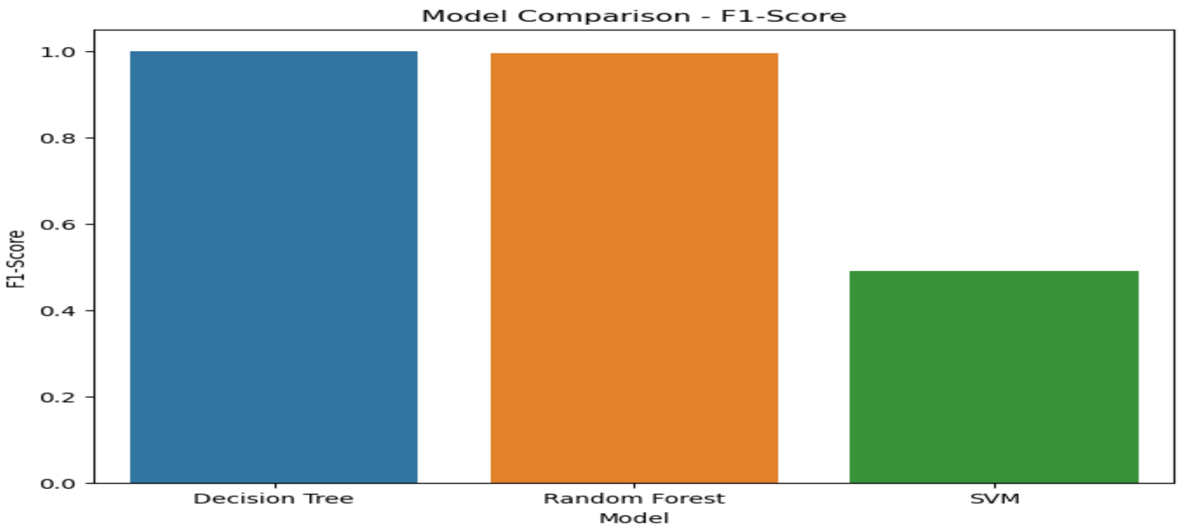


Fig. 4.8: F1-Score Evaluation

These are comparatively represented on a tabular as shown on table 4.1.

Table 4.1: Comparative Models Evaluation

| | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
|---|---|---|---|---|
| **Decision Trees** | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| **Random Forest** | 0.9950 | 0.9955 | 0.9950 | 0.9951 |

| | | | | |
|---|---|---|---|---|
| **Support Vector Machines** | 0.5075 | 0.5372 | 0.5075 | 0.4910 |

1. Accuracy tells us how many instances were classified correctly. Decision Trees nailed it with perfect accuracy, Random Forest was almost there, but SVM didn't do so well.

2. Precision reflects the accuracy of positive predictions. Here, Decision Trees again scored perfectly, Random Forest was a bit lower, and SVM struggled significantly, leading to many false positives.

3. Recall, or sensitivity, measures how well the model identifies all relevant positive instances. Decision Trees and Random Forest were nearly flawless, while SVM fell short.

4. The F1-Score is a nice balance between precision and recall, especially useful for datasets with uneven class distributions. Decision Trees and Random Forest boasted high F1-Scores, showcasing their strong performance, while SVM's low score pointed to its predictive challenges.

Interpretation:

• Decision Trees stand out as the most dependable model across all metrics, consistently hitting perfect scores.

• Random Forest also shines and serves as a solid alternative.

• SVM, on the other hand, struggles across the board, likely due to issues like dataset imbalance or feature scaling.

## 3.3 Discussions of the Results

### 3.3.1 **Dataset Overview**

(i) There are 20 columns in the dataset, which include both numerical and category information. 'AREA NAME', the target column for this study, provides the location of the offense. This characteristic is appropriate for classification jobs because it is categorical.

(ii) A number of columns, including text elements (e.g., LOCATION, Cross Street), temporal features (e.g., Date Rptd, DATE OCC, TIME OCC), and unique identifiers (e.g., DR_NO), were determined to be unrelated or inappropriate for the target column.

(iii) Several columns were identified as possible target columns for various categorization tasks, including Weapon Desc (weapon description), Premis Desc (premises description), and Crm Cd Desc (crime description).

### 3.3.2 Dataset Training and Visualizations

i. **Top Features from the Dataset**

The dataset's top features are displayed in a bar chart in Figure 4.1. A bar chart shows the ten most crucial characteristics for 'Status Desc' prediction. High-importance features have a big influence on classification choices. By eliminating low-importance elements, this aids in uncovering hidden patterns and streamlines the model.

ii. **Crime Frequency Areas**

The number of offenses recorded in various police locations is displayed in a bar chart in Figure 4.2. High crime rates may necessitate increased police presence or community safety initiatives. Authorities can look into local elements that contribute to crime with the use of this graphic.

iii. **Crime Occurrence by Time of Day**

Figure 4.3 is a histogram that displays how crimes are distributed throughout the day. Law enforcement might modify patrol schedules if crimes increase during specific times of the day, such as late at night or early in the morning. It implies that crimes happen regularly at all times if they are spread equally.

iv. **Victim Age Distribution**

A histogram illustrating the age distribution of crime victims may be found in Figure 4.4. Preventive actions can be customized for a specific age group if they are more commonly targeted. While a skewed distribution might show susceptibility in particular age groups, a normal distribution implies that crime affects all ages equally.

### 3.3.3 Models Evaluation

Three machine learning algorithms—Support Vector Machines (SVM), Random Forest, and Decision Trees were assessed. Accuracy, precision, recall, and F1-score were the evaluation criteria that were employed.

(i) **Accuracy**

Figure 4.5 shows the accuracy assessments of models. Perfect accuracy (1.0000) was attained by Decision Trees, Random Forest (0.9950), and SVM (0.5075). This suggests that SVM performs badly in accurately identifying cases, whereas Decision Trees and Random Forest are very effective.

(ii) **Precision**

Figure 4.6 demonstrates the machine learning models' precision evaluation. Random Forest (0.9955), SVM (0.5372), and Decision Trees once more attained perfect precision (1.0000). Decision Trees and Random Forest perform exceptionally well in terms of precision, which quantifies the percentage of positive predictions that are accurate.

(iii) **Recall**

Figure 4.7 indicates the models' recall evaluation metric. SVM (0.5075), Random Forest (0.9950), and Decision Trees (1.0000) all had perfect recall. Decision Trees and Random Forest perform incredibly well when it comes to recall, which gauges the model's capacity to recognize all pertinent good instances.

(iv) **F1-Score**

Figure 4.8 shows the evaluations of the F-1 Score models. A perfect F1-Score of 1.0000 was attained by Decision Trees, Random Forest (0.9951), and SVM (0.4910). An improved statistic for unbalanced datasets is the F1-Score, which strikes a balance between precision and recall. Random Forest and Decision Trees are obviously better in this area.

### 3.3.4 Comparative Models Evaluation

Table 4.1 summarizes the performance of the three models across all metrics as:

(i) Decision Trees: Perfect scores across all metrics (1.0000).

(ii) Random Forest: Very high scores, slightly below perfect (0.9950 accuracy, 0.9955 precision, 0.9950 recall, 0.9951 F1-Score).

(iii) Support Vector Machines: Poor performance across all metrics (0.5075 accuracy, 0.5372 precision, 0.5075 recall, 0.4910 F1-Score).

It is evident that:

(a) Random Forest and Decision Trees both perform exceptionally well on this crime classification job, with Decision Trees attaining flawless results on every criterion.

(b) Support Vector Machines' poor performance indicates that they might not be appropriate for this kind of task or dataset.

(c) The visualizations offer insightful information on crime trends, including high-crime regions, crime incidence times, and victim age distribution, which can guide law enforcement tactics and preventative initiatives.

It can be inferred from these representations that:

1. Because of their excellent accuracy and resilience, decision trees or random forests should be chosen for this classification assignment when choosing a model in such situations.

2. Low-importance features can be eliminated in feature engineering to make the model simpler without sacrificing accuracy.

3. Insights from the visualizations can assist authorities in better allocating resources for efficient and informed law enforcement methods, such as stepping up patrols during peak crime hours or focusing preventive efforts on susceptible age groups.

## 4.0 Summary, Conclusion and Recommendations

### 4.1 Summary

In order to assess crime data and forecast outcomes related to crime, the study investigated the use of machine learning techniques. The target column, "AREA NAME," which provides the location of the crime, is one of 20 columns in the dataset used in the study. It includes both categorical and numerical information. Using a supervised learning methodology, the study focuses on categorization tasks to forecast crime-related outcomes including location, crime type, and crime status.

Metrics including accuracy, precision, recall, and F1-Score were used to assess the performance of a number of machine learning methods, including support vector machines (SVM), decision trees, and random forests. Additionally, the report featured visualizations that shed light on crime trends, including victim's age, time of crime, and crime frequency by area of occurrence.

Major findings from the study are:

#### 4.1.1 Model Performance

(i) All evaluation measures (accuracy, precision, recall, and F1-Score of 1.0000) showed excellent performance for decision trees.

(ii) With slightly below-perfect scores (accuracy: 0.9950, precision: 0.9955, recall: 0.9950, F1 Score: 0.9951), Random Forest likewise demonstrated remarkable performance.

(iii) Support vector machines (SVM) had poor performance, scoring low on all metrics (F1-Score: 0.4910), accuracy: 0.5075, precision: 0.5372, and recall: 0.5075).

#### 4.1.2 Insights from Visualizations

(i) Crime Rate by Area: Higher crime rates in some places point to the need for more law enforcement or community safety initiatives there.

(ii) Time of Crime Occurrence: Law enforcement patrol schedules can be influenced by the fact that crimes typically peak at particular times of the day, such as late at night or early in the morning.
(iii) Victim Age Distribution: By identifying susceptible age groups, the age distribution of crime victims enables the implementation of focused preventive measures.

#### 4.1.3 Feature Importance

By eliminating low-importance elements without sacrificing accuracy, the study was able to identify the most important features for predicting criminal status.

### 4.2 Conclusion

The study shows how well machine learning algorithms—in particular, decision trees and random forests predict and categorize events connected to crime. These models' great accuracy and resilience make them appropriate for use in actual crime prevention and detection situations. Law enforcement organizations can more efficiently allocate resources and put targeted preventive measures into place with the aid of the visualizations' insightful information about crime trends.
Support vector machines' (SVM) subpar performance raises the possibility that this approach is inappropriate for this kind of dataset or task, underscoring the significance of choosing the right models depending on the data's characteristics and the issue at hand.

### 4.3 Recommendations

Based on the findings from the study, it is recommended thus:

#### 4.3.1 Choosing a Model

(i) Because of their excellent accuracy and resilience, decision trees and random forests ought to be chosen when choosing models for comparable scenarios.

(ii) Because support vector machines (SVM) fared poorly on all metrics, they should be avoided for this kind of dataset.

#### 4.3.2 Feature Engineering

For features selection,

(i) To simplify the model and increase computational efficiency without compromising accuracy, low-importance features should be eliminated.

(ii) To improve model performance, more feature engineering might be investigated, such as generating new features from preexisting ones (such as time-based features from TIME OCC).

#### 4.3.3 Strategies for Law Enforcement

(i) Law enforcement organizations should concentrate on high-crime regions, as indicated by the Crime Frequency by Area visualization, by stepping up patrols and putting community safety initiatives into place in order to effectively coordinate crime prevention efforts throughout Maiduguri city.

(ii) Patrol schedules should be modified to include more presence during high-crime hours in accordance with the Time of Crime Occurrence insights.

(iii) The Victim Age Distribution visualization identifies susceptible age groups that should be protected by customized preventive interventions.

#### 4.3.4 Quality and Data Collection

(i) Make sure the dataset is updated and cleaned on a regular basis to maintain good data quality, as missing or irrelevant data can have a detrimental effect on model performance. This will ensure efficient data gathering and quality.

(ii) Gather further information that might shed more light on crime trends, such as socioeconomic characteristics or meteorological circumstances.

### 4.4 Suggestions for Future Research

Future research in this field should focus on the following as well:

(i) Examining Alternative Algorithms: Future research should examine how well decision trees and random forests perform in comparison to other machine learning algorithms like gradient boosting machines (GBM), neural networks, or ensemble methods.

(ii) Real-Time Crime Prediction: Create real-time crime prediction tools that can give law enforcement organizations alerts and insights instantly, allowing them to react to possible crimes more quickly.

(iii) Geospatial Analysis: Use geospatial analysis methods to map out criminal hotspots so that law enforcement resources can be allocated more precisely.

(iv) Integration with External Data Sources: To find more variables that can affect crime trends, combine the crime dataset with data from external sources, such as social media activity, socioeconomic data, or weather data.

(v) Explainability and Interpretability: Put your attention on making machine learning models more understandable, especially for law enforcement organizations. You can do this by employing strategies like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations), which offer concise justifications for model predictions.

(vi) Community Involvement: Examine how community involvement contributes to crime prevention by examining statistics from public awareness campaigns, neighborhood watch programs, and community programs.

(vii) Longitudinal Studies: To determine long-term trends and the efficacy of preventive measures put in place, conduct longitudinal studies to examine crime trends over time.

## References


1. Ahmed, M., Khan, S., & Lee, J. (2018). Real-time crime detection using convolutional neural networks for video surveillance. *International Journal of Computer Applications*, *182*(45), 12–20. https://doi.org/10.5120/ijca2018917264

2. Akoglu, L., Tong, H., & Koutra, D. (2015). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, *29*(3), 626–688. https://doi.org/10.1007/s10618-015-0444-4

3. Aliyu, A., & Abdullahi, Y. Y. (2025). Causes and consequences of herders–farmers conflict and its gender implication in Plateau State. *International Education and Research*

*Journal*. https://ierj.in/journal/index.php/ierj/article/view/1665

4. Allott, R., Smith, L., & Chandra, P. (2024). Principles of crime detection and prevention in contemporary law enforcement. *Journal of Criminal Justice Studies*, *16*(2), 45–63.
5. Berman, J., Lee, H., & Kumar, A. (2019). AI-based intrusion detection and anomaly detection in network security. *Journal of Information Security, 10*(3), 125–138. https://doi.org/10.4236/jis.2019.103009
6. Chouldechova, A. (2017). Fair prediction with disparate impact: A study in predictive policing. *Big Data*, *5*(2), 153–163. https://doi.org/10.1089/big.2016.0047
7. CRADI. (2025). *Post-conflict assessment report: Bassa, Bokkos, Riyom, and Mangu LGAs, Plateau State, Nigeria*. Crest Research and Development Institute. https://cradil.org/post-conflict-assessment-report-bassa-bokkos-riyom-and-mangu-lgas-plateau-state-nigeria/
8. EUCPN. (2022). *Predictive policing practices in Europe: Overview and assessment*. European Crime Prevention Network. https://www.eucpn.org/predictive-policing-europe
9. Gukas, T. (2025). Community collaboration for sustainable peace-building between farmers and herders in Plateau State, Nigeria. *FUDMA Journal of Sciences,* *9*(5), 118–125. https://fjs.fudutsinma.edu.ng/index.php/fjs/article/view/3422
10. Ibrahim, M., & Mohammed, A. (2021). Poverty, unemployment, and juvenile delinquency in Borno State. *Journal of Social Development in Africa*, *36*(1), 54–70.
11. Jain, A. K., Ross, A., & Nandakumar, K. (2011). *Introduction to biometrics*. Springer. https://doi.org/10.1007/978-0-387-77326-5
12. Kang, W., & Kang, S. (2017). Crime prediction using deep neural networks with spatio-temporal and environmental context. *Knowledge-Based Systems, 126*, 101–113. https://doi.org/10.1016/j.knosys.2017.05.018
13. Kaur, P., & Saini, M. (2024). Predictive modeling for urban crime prevention: Applications and challenges. *International Journal of Computer Science and Security, 18*(2), 87–102.
14. Mandalapu, R. (2023). Machine learning and deep learning approaches for crime prediction: A systematic review. *Journal of Artificial Intelligence Research*, *70*(4), 456–475.
15. Mohler, G., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2015). Self-exciting point process modeling of crime. *Journal of the American Statistical Association, 106*(493), 100–108. https://doi.org/10.1198/jasa.2009.ap08746
16. Perry, W. L., McInnis, B., Price, C. C., Smith, S., & Hollywood, J. S. (2013). *Predictive policing: The role of crime forecasting in law enforcement operations*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR233.html
17. Saltos, E., & Cocea, M. (2017). Data mining for crime prediction: Methods and applications. *International Journal of Advanced Computer Science and Applications*, *8*(12), 210–221. https://doi.org/10.14569/IJACSA.2017.081227
18. Samtani, S., Chen, H., & Chau, M. (2016). Using social media analytics for criminal intelligence: Opportunities and challenges. *Decision Support Systems, 85*, 50–60. https://doi.org/10.1016/j.dss.2016.03.007
19. Singh, R., & Jain, S. (2020). Machine learning in crime prediction and

prevention: Trends and prospects. *International Journal of Data Science*, *5*(2), 34–49.

20. UNODC. (2022). *Handbook on crime prevention and criminal justice policies*. United Nations Office on Drugs and Crime. https://www.unodc.org/unodc/en/crime-prevention/index.html
21. Ugwu, J. (2025). *Nigeria security threat report 2025: Ranking of high-risk cities*. SB Morgen Intelligence.
22. Van Brakel, R., & De Hert, P. (2011). *Policing, surveillance, and AI: Emerging legal and ethical challenges*. Springer. https://doi.org/10.1007/978-94-007-0562-1
23. WikkiTimes. (2025). Inside Plateau's 2025 violence: A timeline of killings, abductions, and displacements. https://wikkitimes.com/inside-plateaus-2025-violence-a-timeline-of-killings-abductions-and-displacements/
24. Zhu, X., Liu, H., & Chen, L. (2022). Urban crime trends and social sustainability: Challenges and mitigation strategies. *Sustainability, 14*(5), 3012. https://doi.org/10.3390/su14053012
25. Ahmed, M., Al-Jawad, M., Sabir, F., Al-Jumeily, D., & Fergus, P. (2018). Real-time intelligent video surveillance: A cloud computing architecture for detection of suspicious behaviors. *Computers & Electrical Engineering*, 67, 341-362.
26. Akoglu, L., Tong, H., & Koutra, D. (2015). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, 29(3), 626-688.
27. Allott, A. Nicolas , Thomas, . David A. , Bernard, . Thomas J. , Clarke, . Donald C. and Edge, . Ian David (2024, August 17). *crime*. *Encyclopedia Britannica*. https://www.britannica.com/topic/crime-law.
28. Annie, S. S. H., Pathmanaban, J., Kingsley, S., Sriman, B., Suryaa, N., & Santhosh, E, K. (2022). Prediction and prevention analysis using machine learning algorithms for detecting the crime data. *2022 1st International Conference on Computational Science and Technology (ICCST)*, 986-991. https://doi.org/10.1109/ICCST55948.2022.10040370.
29. Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. *Information*, 10(4), 122.
30. Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153-163.
31. EUCPN (2022). Artificial intelligence and predictive policing: risks and challenges. Brussels: EUCPN.