

ENHANCING PUBLIC SAFETY: CRIME CATEGORY PREDICTION THROUGH FUSION-BASED ENSEMBLE LEARNING WITH XG-BOOST

Khadija Tahira^{*1}, Sidra Mushtaq², Najam Us Sahar³

^{*1,3rd} Minhaj University Lahore

^{2nd} NUML University

^{*1}khadija.csir@mul.edu.pk, ²sidra.mushtaq@numl.edu.pk, ³najamussahar.it@mul.edu.pk

DOI: <https://doi.org/10.5281/zenodo.22870559>

Keywords

Crime Analytics, Public Safety, Crime categorization, Machine Learning models, Pattern Recognition, Fusion Model.

Article History

Received: 14 June 2026

Accepted: 05 September 2026

Published: 21 September 2026

Copyright @Author

Corresponding Author: *

Khadija Tahira

Abstract

The rate of crime has been rising on daily basis all around the world, leading to an increase in security concerns, Closed-Circuit Television (CCTV) cameras have been deployed with the intention of reducing crime and enhancing public protection. The correctness of criminal detection is greatly increased by the use of CCTV cameras. However, the amount of data produced every day by modern monitoring technologies makes it challenging for experts in computer vision to evaluate, since it requires a lot of time and human involvement. The development of a system that can recognize the patterns of human by monitor humans in real-time is necessary. Decision trees are an effective supervised learning strategy for categorizing and comprehending crime trends because they divide data into essential qualities such as time, location, and victim information, resulting in a visible, rule-based framework. This interpretability makes them useful for detecting influential aspects in crime, but its simplicity restricts usefulness in practical criminology and public safety. While individual classifiers such as Decision Trees, Random Forests, or Support Vector Machines constitute the basis of crime prediction research, they frequently fail to capture the intricate spatiotemporal and categorical correlations in crime datasets. To address these limitations, this study provides a hybrid fusion modeling framework that integrates numerous classifiers using approaches like soft voting to improve forecast accuracy and robustness. The use of advanced algorithms like XG Boost, as well as feature engineering strategies and ways for regulating class imbalance, assures better performance in detecting and predicting criminal patterns. This approach not only takes advantage of the characteristics of several classifiers, but it also offers a scalable and dependable solution for real-world crime analysis.

I. Introduction:

In recent era, the primary goals of government are to protect citizens from harm and violence. An online reporting and surveillance platform enhances communication between law enforcement and the community, providing a fast

and user-friendly method for reporting crimes. These crime reports are crucial for detectives, leading to numerous arrests, case resolutions, and property recovery.[3]

Surveillance is essential for a variety of reasons, including enhancing security and safety by preventing crime in areas at risk, enabling quick incident response, and giving essential proof for investigations. It also helps to monitor behavior in a variety of situations, ensuring compliance with regulations. However, surveillance poses significant challenges. Individuals may feel uncomfortable about being constantly tracked, raising privacy concerns. Additionally, keeping track of large volumes of footage can be challenging and costly, creating concerns about privacy and access. Traditional systems frequently create fake warnings, resulting in unnecessary resource allocation, and fixed cameras may leave blind patches, decreasing coverage. The total cost of setting up and maintaining surveillance systems can be important and technological flaws risk data integrity. Addressing these problems is crucial to enhancing surveillance effectiveness while balancing concerns about privacy and security.

For public safety and law enforcement organizations, crime prediction is an essential responsibility since it allows for the proactive use of resources and well-informed decision-making to lower crime rates and improve community well-being. Machine learning (ML) approaches have become increasingly effective tools for modeling and predicting criminal activity as a result of the growing availability of historical crime data. Because of their relative interpretability and ease of use, individual machine learning classifiers including Decision Trees, Random Forests, Support Vector Machines, and Logistic Regression have been the mainstay of existing crime prediction studies. When faced with the inherent complexity of real-world crime data, these models frequently perform poorly, despite their demonstrated usefulness in identifying certain trends within crime datasets. In particular, crime statistics usually show:

Spatial-temporal dependencies, in which the probability of an event is significantly impacted by both time-related and geographic elements.

Categorical heterogeneity, which results from a

variety of crime kinds and related characteristics that don't readily fit into straightforward parametric connections.

Unbalanced distributions, which result in skewed forecasts and poor performance on uncommon but important classes, because some crime categories like stealing are far more common than others like homicide.

Decision Tree is a widely used supervised machine learning algorithm that models' decisions and their possible consequences in a tree-like structure. It works by recursively splitting the dataset based on the most significant features, creating branches that represent decision rules and leaf nodes that indicate outcomes. This structure makes it easy to visualize and interpret, allowing researchers to clearly understand how different input variables such as time of crime, location, or victim demographics contribute to predicting the crime category.[5]

Transparency and interpretability are two of a decision tree's main benefits. A decision tree offers simple "if-then" principles, which are particularly helpful in delicate fields like public policy and crime investigation, in contrast to intricate black-box models. Along with having a high prediction capacity, it also identifies the most important aspects, allowing for data-driven choices to be made in community safety, urban planning, and law enforcement. With an emphasis on crime matching the process of connecting comparable episodes to find possible suspects.[4]

To address these challenges, this research provides a hybrid fusion-based architecture that combines linear and tree-based classifiers, including sophisticated approaches like XG Boost, to successfully capture the complex spatiotemporal and categorical correlations found in crime data, use of fusion techniques like soft voting, which use the complimentary capabilities of different models to increase overall prediction accuracy and robustness. To achieve reliable evaluation and avoid overfitting, as illustrated in Fig. 1, experiments on a real-world crime dataset indicate that the proposed framework consistently outperforms individual classifiers and traditional ensemble approaches across a wide range of

assessment metrics. In instance, using the F1-score gives a more balanced performance measure for dealing with class imbalance, while accuracy verifies predictive potential. Ultimately, by stressing interpretability and investigating practical applications, this study demonstrates the potential of the suggested fusion technique as a decision-support tool for law enforcement agencies and lawmakers.

II. Literature Review:

Traditional techniques for machine learning struggle to incorporate human domain expertise, which is crucial in disciplines such as crime investigation. In [5] author describes a human-centered technique that uses text mining to evaluate unstructured crime reports and establish correlations between criminal components. By integrating an associative questioning model with partition clustering, the approach makes it easier to uncover crime patterns and suspect networks. It improves decision-making by increasing human contact with machine learning through visualization and expert input, as demonstrated using an anonymous burglary dataset.

Today's surveillance systems produce massive amounts of video data, making it challenging to manually search for unforeseen operations. In [2] author focuses on deep learning-based system for intelligent detection of anomalies that efficiently extracts spatiotemporal information from video frames using a pertained Convolutional Neural Network (CNN). The traits are then identified as abnormal or normal using a Bi-directional Long Short-Term Memory (Bi-LSTM) model. This technique improves detection capacities while lowering the need for human intervention, hence increasing public safety. [2]

Violence is a major societal concern, and effective prevention depends on precise detection, which has traditionally necessitated large human resources. In [3] author proposes an automated violence detection system that requires minimal human intervention while providing simple, immediate, and widespread monitoring. The technology not only detects but also anticipates aggression by evaluating verbal and nonverbal

cues. Furthermore, it uses external data sources to improve the accuracy of its forecasts. [3]

The rising crime rate has raised security concerns, forcing law enforcement to concentrate on proactive crime detection and prevention. While CCTV systems are commonly utilized in public as well as private locations, human monitoring might result in discrepancies due to the vast number of images. In [7] author uses a Deep Learning-Based Real-Time Crime Detection Technique. This approach uses the YOLO (You Only Look Once) object detection framework to analyze live video streams and delivers instant notifications to right next cybercrime units when illegal behavior is discovered. This technology is efficient, to processing photos at 45 frames per second.

Cybercrimes include illicit behaviors involving computers or communication tools, such as child pornography, online stalking, identity theft, credit card fraud, and cyberterrorism. These offenses frequently result in violations of privacy, security concerns, money losses, and harm to properties both public and private. In [9] author investigates methods for recognizing and combating cybercrime, focusing on the many types of crime and the threats they bring to privacy and security. It examines how cybercriminals operate and assesses current detection and prevention measurements, noting their advantages and disadvantages. Finally, it makes proposals for developing more effective cybercrime detection systems.

In [16] author proposes a decision support system for linking repeated crimes, emphasizing the need for quantitative tools among crime researchers in China. The method employs pairwise classification based on resemblance to assess whether instances belong to the same series. A case study on a robbery dataset indicates that it works, and the system has been used successfully in a public security bureau for over a year with positive results. By automating illegal linking and allowing for updates and expert input, the method improves inquiry productivity and organizes resources for law enforcement. [16]

In [8] author introduces YOLO (You Only Look Once), an innovative object detection algorithm

that approaches detection as a regression problem. YOLO uses a single neural network to predict boundaries and class probabilities from entire photos in one assessment, allowing for end-to-end optimization. The default model analyzes images at 45 frames per second, while Fast YOLO handles 155 frames per second, beating existing real-time detectors in mean Average Precision (MAP). Although it has more localization errors, YOLO produces less false positives and excels at generalizing across domains, outperforming established approaches such as DPM and R-CNN. In [12] The author focuses on need for technology that can detect and anticipate crimes in response to rising crime rates. It looks into deep learning methods for crime detection, emphasizing the importance of effective data pre-processing to manage missing values and the tendency of violent crime to cluster in specific areas. Identifying crime hotspots is critical to crime reduction. The research highlights the effectiveness of deep learning algorithms, which are frequently utilized in a variety of disciplines, in improving forecast accuracy. It focuses on the function of genetic algorithms in crime detection and urges for more resources to support these intelligent systems. [12]

In [9] The author studies human activity recognition and its applications in areas such as surveillance and human-computer interaction. It generates a dataset of two-person interactions using low-cost depth sensors such as the Microsoft Kinect, which includes synchronous video, depth, and motion capture. The study examines various characteristics for indexing and retrieving motion capture data and finds that geometric relational features based on joint distances are the most successful. It also explores Multiple Instance Learning (MIL) approaches, concluding that MIL classifiers outperform Support Vector Machines (SVMs) in categorizing lengthy sequences containing interactions. [9]

In [21] The author paper presents a statistical method for criminal linkage analysis that detects and organizes offenses associated with the same culprit, prioritizing suspects for further investigation. It uses Bayes factors to assess evidence strength and hierarchical clustering with agglomeration to compare crime series. The algorithm performs well with real-world data, attaining an 82% true connection rate with a 5% false positive rate and recognizing 74% to 89% of additional crimes in a set of 50 occurrence.

Author	Model / Technique	Challenges Addressed	Achievements	Accuracy / Performance	Limitations
[1]	Human-centered text mining with associative questioning and partition clustering	Incorporating human expertise; unstructured crime reports	Links crime patterns and suspect networks through visualization and expert input	Not specified	May rely heavily on manual validation and expert interpretation
[2]	CNN + Bi-directional LSTM for anomaly detection	Manual review of massive surveillance video data	Efficient spatiotemporal anomaly detection; reduces human workload	Not specified	May require significant computational resources for video processing
[3]	Automated violence detection using multimodal cues	High resource demands for traditional violence detection	Predicts and detects aggression with minimal human intervention	88%	Potential privacy concerns and reliance on data quality

[4]	Deep Learning Real-Time Crime Detection with YOLO	Human monitoring inconsistencies in CCTV	Real-time detection and instant alerts to cybercrime units	Processes 45 frames per second	Localization errors; potential false positives
[5]	Methods for cybercrime recognition and prevention	Diverse cyber threats compromising privacy and security	Reviews threats, detection, prevention, and proposes improvements	92%	Existing methods have gaps; evolving cyber threats
[6]	Decision Support System with pairwise classification and expert input	Linking repeat crimes efficiently	Automated linking improved inquiry productivity; real-world deployment	Positive operational results in China	May depend on data similarity thresholds and manual updates
[7]	YOLO Object Detection (Fast YOLO variant)	Need for fast, accurate object detection	Achieves high speed (45–155 FPS) and better generalization	Fast YOLO: 155 FPS; mAP superior to prior detectors	Higher localization error compared to some methods
[8]	Deep learning and genetic algorithms for crime prediction	Managing missing data; identifying crime hotspots	Improved prediction accuracy and hotspot detection	Not specified	Data preprocessing complexity; resource demands
[9]	MIL classifiers and geometric relational features for human activity recognition	Accurate recognition of human interactions	MIL classifiers outperform SVMs in complex sequences	MIL better accuracy over SVMs (percentages not specified)	Focused on two-person interactions only
[10]	Statistical linkage with Bayes factors and hierarchical clustering	Organizing offenses and prioritizing suspects	High true linkage rate with low false positives	82% true connection, 5% false positives, 74–89% extra crimes linked	May need large historical datasets for effectiveness
11	Proposed Model(XG Boost)	Temporal; Forecasting	Crime prediction & categorization enhancement	97%	~

Table 1: Limitations of previous Work

III. Research Methodology

Framework of this research paper is to integrate six heterogeneous base learners to capture diverse patterns and relationships within the crime dataset in Fig:1, By averaging several decision tree outputs,

Random Forest an ensemble of decision trees trained by bootstrap aggregation (bagging), lowers variance and enhances generalization. Random Forests are especially good at handling noisy data and simulating non-linear feature interactions.

$$\{y\} = \text{mode} \{h_{1(x)}, h_{2(x)}, \dots, h_{K(x)}\} \quad (1)$$

$$\{y\} = \frac{1}{K} \sum_{k=1}^K h_{k(x)} \quad (2)$$

XG Boost (Extreme Gradient Boosting) is an optimized ensemble learning technique that uses gradient boosting decision trees. It creates numerous weak learners in a sequential manner, with each new tree correcting the errors of the prior ones by focusing on misclassified instances.

The model uses techniques including regularization, parallelization, and tree pruning to avoid overfitting and increase generalization. These qualities make XG Boost extremely efficient and accurate, which explains why it outperforms most crime categories in the findings.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (3)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (4)$$

$$\{y\}_i = \sum_{k=1}^K f_{k(x_i)}, \text{quad } f_k \in \text{Trees} \quad (5)$$

A **decision tree classifier** divides the feature space recursively into areas with uniform labels using a single tree-based model. Decision trees are

computationally efficient and may capture distinct decision rules, although they are more prone to overfitting than ensemble approaches.

$$H(S) = - \sum_{i=1}^n p(i) \log p(i) \quad (6)$$

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v) \quad (7)$$

The MLP Classifier (Multilayer Perceptron) is a neural network comprising input, hidden, and output layers that employs nonlinear activation functions to learn intricate patterns. It is trained via backpropagation and gradient descent, with

weights adjusted to reduce mistakes. While good in capturing nonlinear relationships, its efficacy is strongly reliant on parameter optimization. MLP outperformed XG Boost and Random Forest in some categories, but it was less consistent overall.

$$z = \sum_{i=1}^n w_i x_i + b \quad (8)$$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (9)$$

$$a = \sigma(Wx + b) \quad (10)$$

$$L = - \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1-y_i) \log(1-\hat{y}_i)] \quad (11)$$

This study uses a systematic strategy to create and assess a crime prediction model that makes use of multiple classifiers on past data trends, the method forecasts crime occurrences by combining data collection, preprocessing, model training,

and performance evaluation. The purpose of this research is to develop an interpretable and useful predictive framework that can help law enforcement agencies detect high-risk locations and possible criminal activity by using XG Boost.

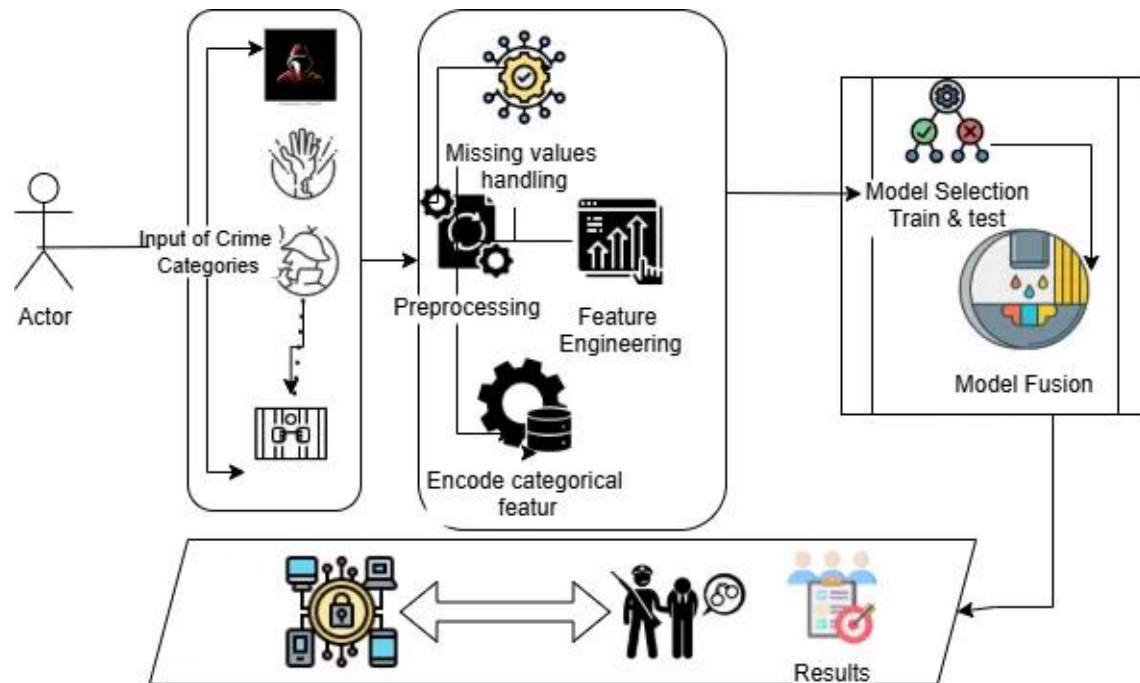


Fig.1(Proposed Model for crime prediction)

Fill numerical missing values with median

The median of each column is used to impute missing numerical values in variables like Weapon Used Cd and Premis Cd. In order to reduce the impact of outliers and maintain the data's underlying distribution, median imputation was chosen. Missing entries were substituted with zero where a missing value suggested the absence of a feature (for example, no weapon used).

Fill categorical missing values with mode

The most frequent value (mode) is used to impute all missing values in categorical columns. The reasoning for this is that mode imputation avoids data loss from row deletion while maintaining the distribution of the majority class. For instance, using the mode to fill in any missing values in the Vict Sex column guarantees consistency throughout the dataset.

Data Type Conversion

Converting data types guarantees that every column is kept in an analysis-ready manner. For instance, numeric columns is saved as strings are cast to numerical types, and date strings are transformed into datetime objects to enable the extraction of elements like as the month and day of the week. Accurate feature engineering and modeling depend on proper conversion.

Feature Engineering

By adding new variables, feature engineering improves the model's capacity for prediction. Features like hour, day of the week, and month can be used to identify trends in crime in temporal data. To identify high-risk areas, spatial features such as latitude and longitude are kept & grouped by region.

Handling Duplicates

Because duplicate records overrepresent some data, they can skew model training. In order to preserve data integrity and prevent skewing the results, duplicate rows are found and eliminated.

Data Transformation

To prepare variables for modeling, data transformation standardizes and encoded. For instance, numerical variables can be scaled or normalized, and categorical variables can be label-encoded or one-hot encoded. This stage guarantees that every feature is in a format that machine learning algorithms can use.

IV. RESULTS AND DISCUSSION

The following hardware setup, which is accessible on Google Colab, was used for the experiments: Processor: 11th Gen Intel(R) Core(TM) i5-1145G7 Graphics, (based on Colab's session allocation and availability).. Disk Storage: The Colab environment provides 32.4 / 107.7 GB of

disk storage. System of operation: Libraries and Frameworks: Pandas, NumPy, and additional machine learning tools as required

The proposed model uses a dataset with 15 categories of crimes, including 14 input variables and 1 output variable, it consists of 290,684 records, with 70% used for training and 30% for testing. Key input features include Time, location, crime type, Victim information, other context, The system proposed in this study uses a ML based approach to Predict the type of crime that will occur in a location & the time, likelihood of occurrence in a given area it also Identify spatial-temporal crime patterns Fig. 1 The first stage involves gathering the data in its raw form. The second stage is preprocessing, where the raw data is cleaned, standardized, and prepared for analysis as shown in the Fig.1 Datasets for this process can be obtained from sources kaggle & having the following results after preprocessing

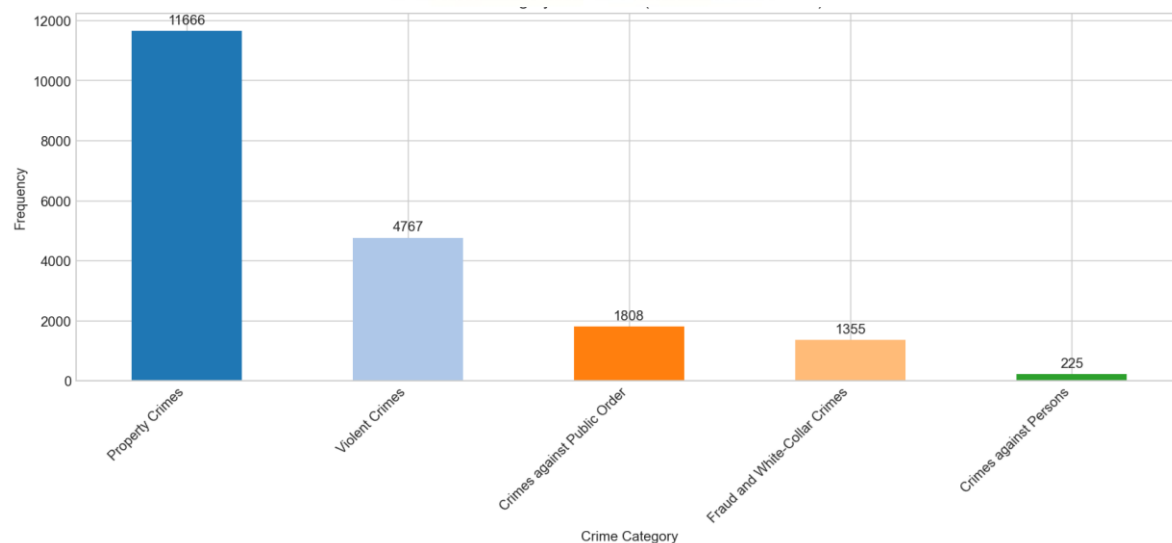


Fig.2

The distribution highlights that property crimes dominate the dataset, making up the majority of reported cases, followed by violent crimes shown in Fig.2. Public order offenses and white-collar crimes appear less frequently, while crimes against

persons are relatively rare. This imbalance indicates that predictive models may lean more heavily toward property and violent crime patterns due to their higher representation.

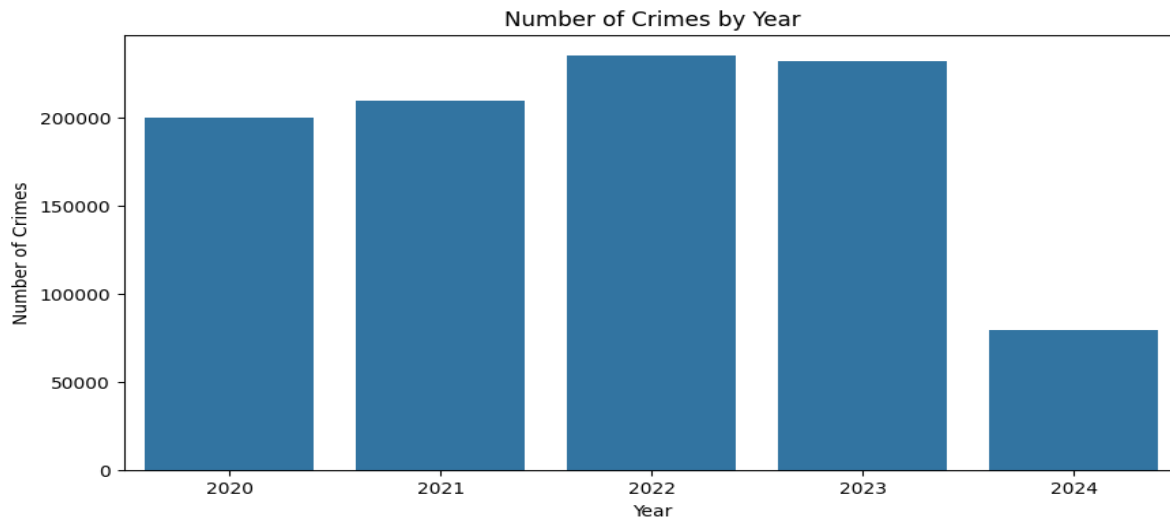


Fig.3

The annual distribution of recorded crimes from 2020 to 2024 is shown in the bar chart in Fig 3. Between 2020 and 2023, crime rates stayed high, reaching a peak of almost 230,000 incidents in 2022. 2023 saw a minor fall, while 2024 saw a

sharp decline as crimes dropped below 100,000. This pattern points to a notable decline in crime rates in recent years. Gaining understanding into the causes of this steep fall may help develop prevention and policy initiatives.

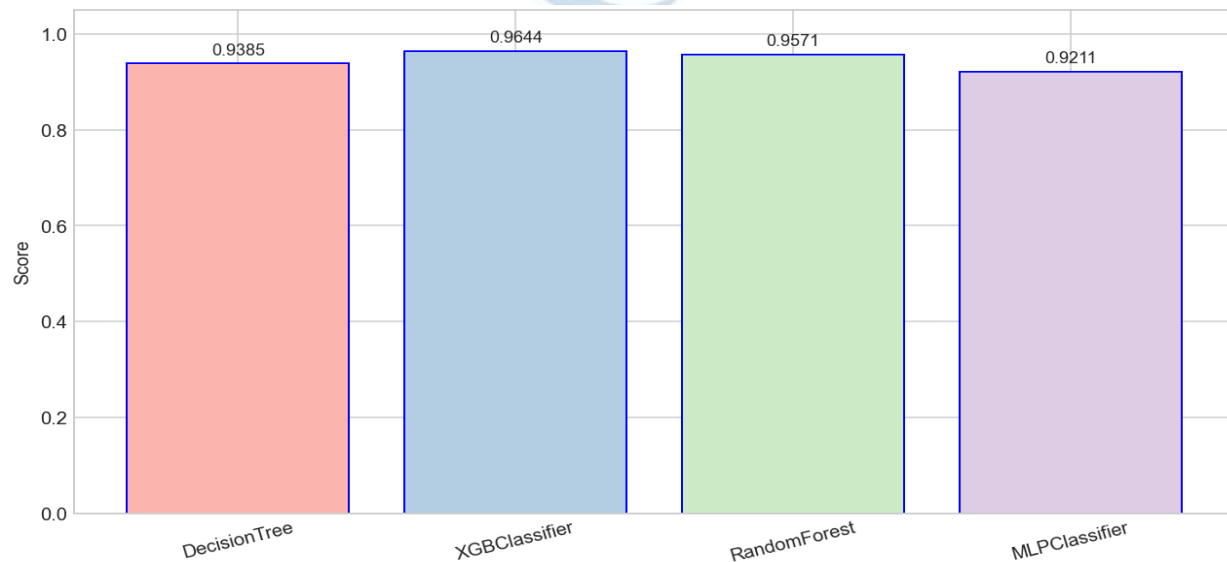


Fig.4

This Fig.4 presents the comparative validation accuracy of four classifiers used in the study. Among them, XG Boost achieved the highest accuracy (96.44%), followed by Random Forest (95.71%), while Decision Tree (93.85%) and MLP

Classifier (92.11%) showed relatively lower performance. The results indicate that ensemble-based methods, particularly XG Boost, demonstrate superior predictive capability for crime data classification.

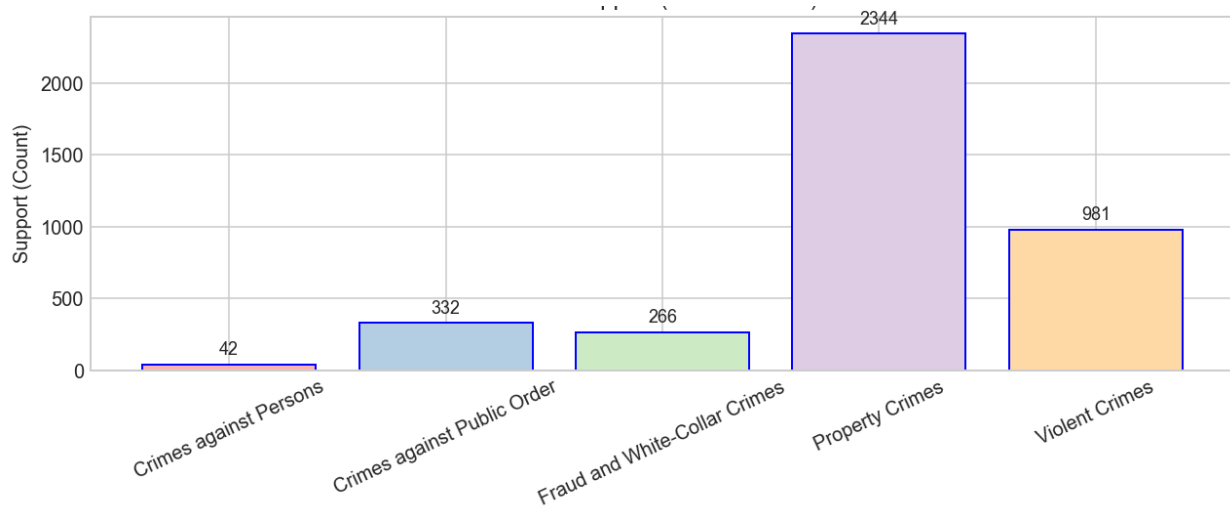


Fig.5

Fig.5 shows the class distribution of the validation dataset across different crime categories. Property crimes dominate the dataset (2,344 cases), followed by violent crimes (981 cases), while fraud and white-collar crimes (266 cases) and crimes against public order (332 cases) are comparatively fewer. The smallest category is crimes against

persons (42 cases), indicating a highly imbalanced dataset. Such imbalance emphasizes the importance of adopting strategies like feature engineering, class rebalancing, and metrics such as F1-score to ensure fair and reliable model performance across all crime types.

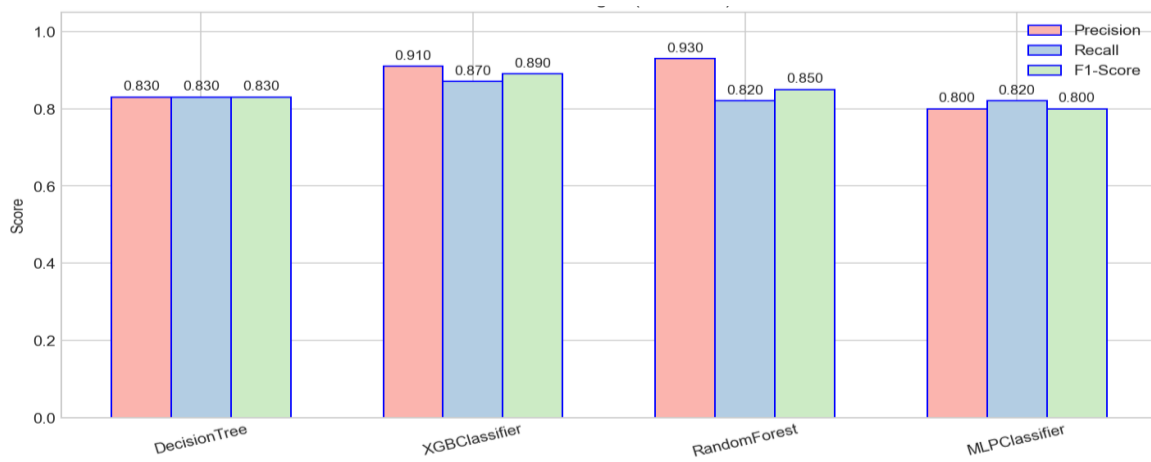


Fig.6

This figure[6] illustrates the macro-averaged Precision, Recall, and F1-scores of different classifiers on the validation set. The results show that XGBoost achieved the best balance, with Precision (0.91), Recall (0.87), and F1-score (0.89),

indicating strong and consistent performance across classes. Random Forest also performed competitively, with the highest Precision (0.93) but slightly lower Recall (0.82), suggesting it is more conservative in predicting positive cases.

Decision Tree produced balanced but moderate scores (0.83 across all metrics), while MLPClassifier had the lowest overall performance (around 0.80). These outcomes highlight the effectiveness of XGBoost in handling imbalanced

data, as it provides the best trade-off between precision and recall, making it more reliable for real-world crime prediction tasks.

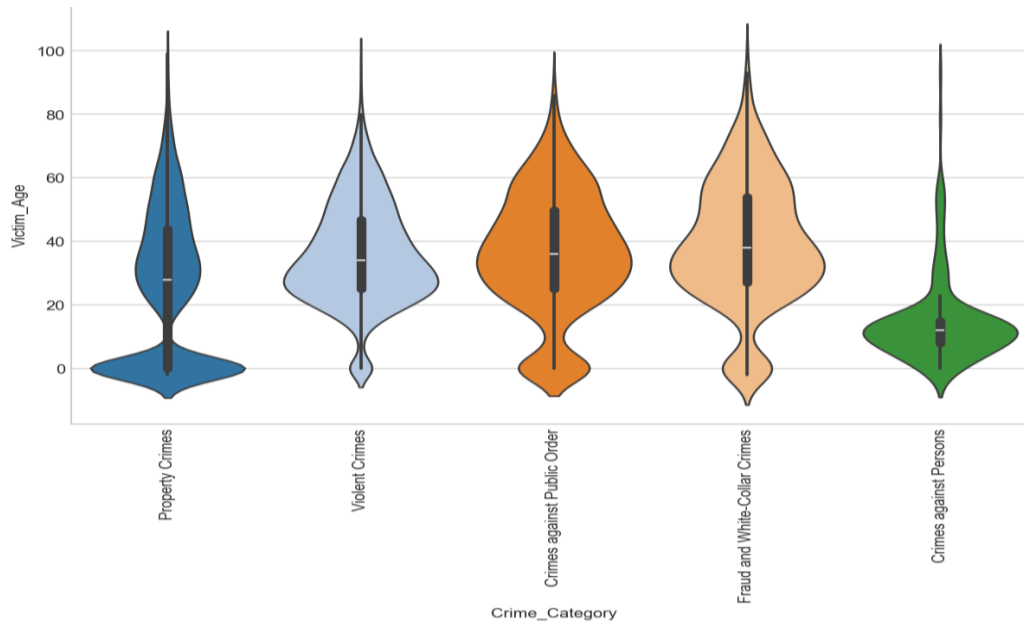


Fig:7(Part 1)

The violin plot illustrates shows in Fig:7(part 1) how victim age varies across crime categories. Property and violent crimes show a wide spread, with most victims concentrated in early adulthood to middle age. Public order and white-collar crimes

affect slightly older populations, while crimes against persons are more frequent among younger victims. This suggests age-related vulnerability patterns that differ by crime type.

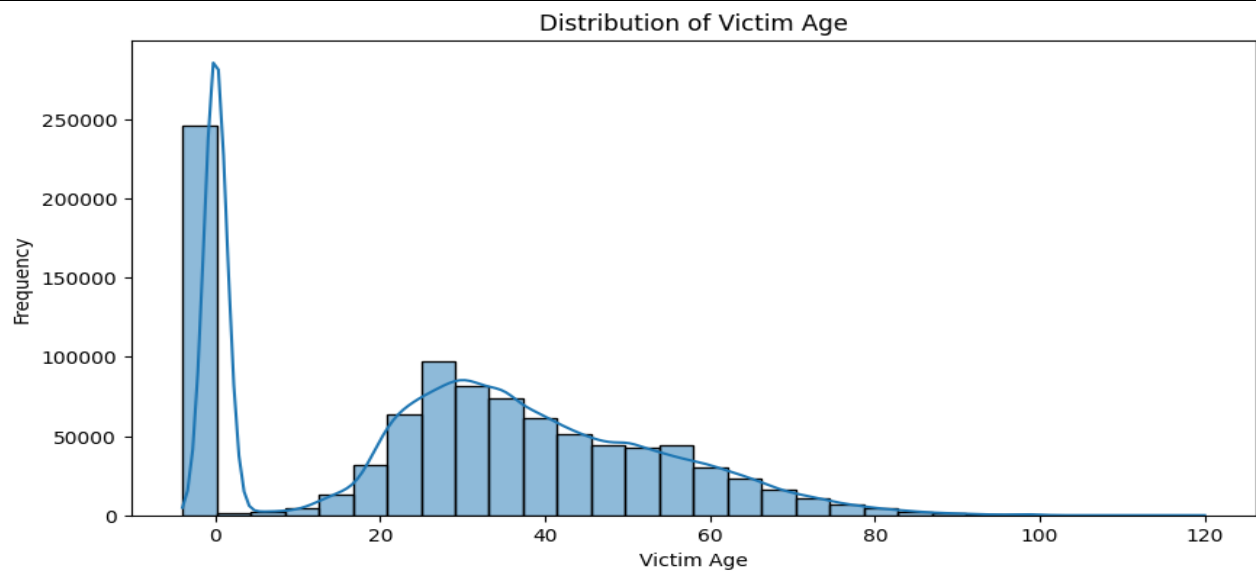


Fig.7(part 2)

The distribution of victim ages in this sample is shown in Fig. 7(Part 2). At age zero, the histogram exhibits a noticeable rise that most likely corresponds to faulty or missing entries that were encoded as zero. Young to middle-aged persons are

most frequently victimized, as seen by the highest frequencies among valid ages, which fall between around 20 and 40 years old. Relatively few casualties are beyond 70, and the distribution progressively tapers off for later age groups.

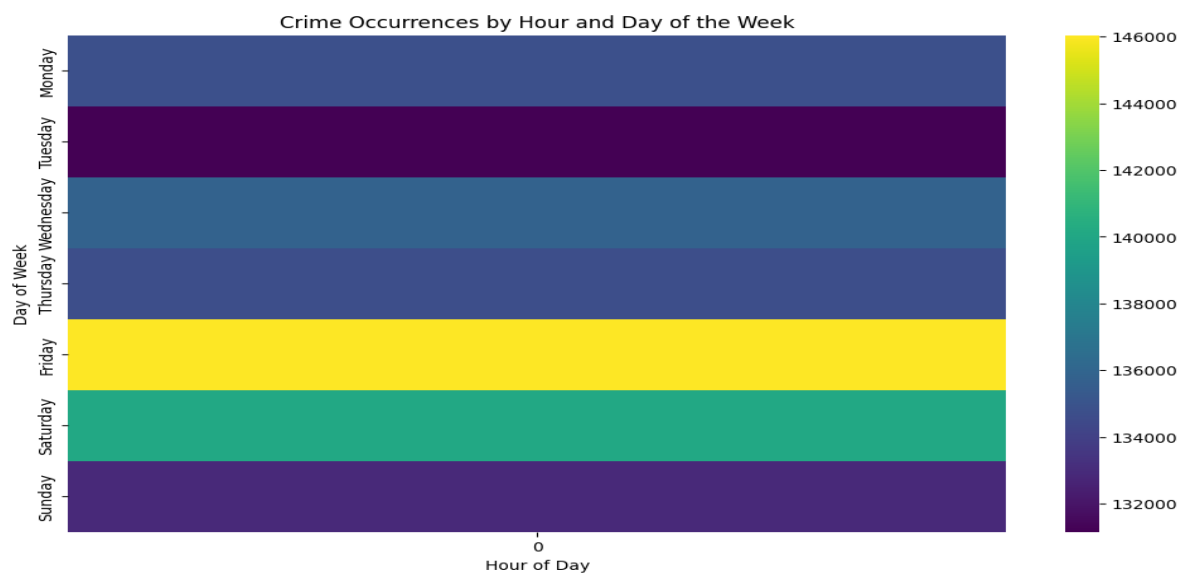


Fig.8

A heatmap showing crime rates by day of the week and hour of the day is shown in Fig. 8. Brighter hues indicate higher frequencies, and the color gradient shows the number of events. When

compared to other days, Friday exhibits the largest concentration of crime reports overall. Tuesday has the lowest crime rates; however, Saturday also sees the highest activity. This pattern highlights

possible temporal trends that could be helpful for predictive modeling, indicating that criminal

activity tends to grow near the conclusion of the week.

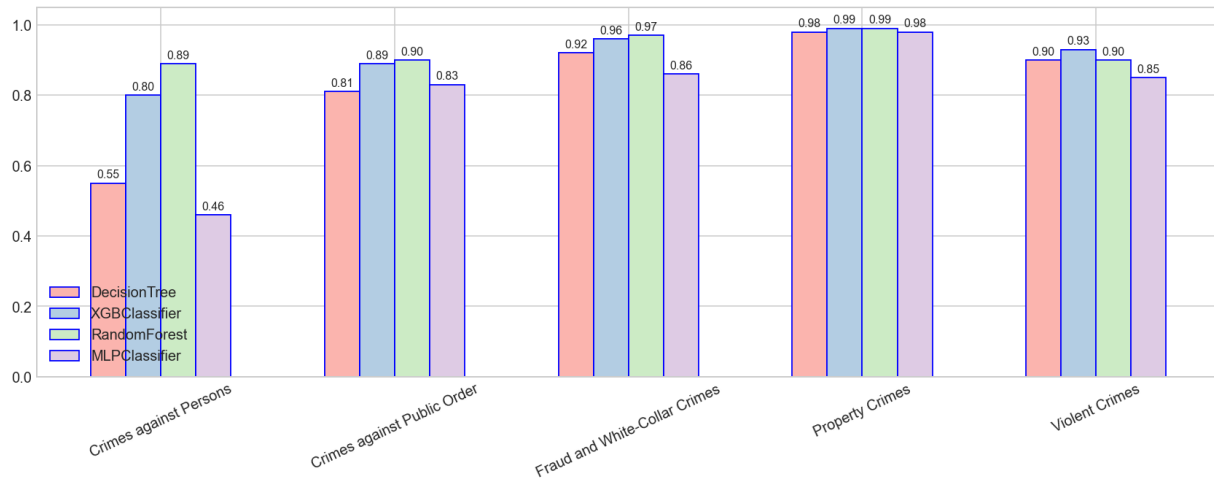


Fig 9:

This bar chart shows the precision scores of four classifiers Decision Tree, XGBClassifier, Random Forest, and MLPClassifier has shown in Fig-9, for different crime categories. Precision measures how many of the predicted positive cases were actually correct. For Crimes against Persons, the Decision Tree and MLPClassifier have lower precision (0.65 and 0.46) compared to XGBClassifier (0.99) and Random Forest (0.89), indicating that the latter

two are more reliable in minimizing false positives. Across other categories—Crimes against Public Order, Fraud/White-Collar Crimes, Property Crimes, and Violent Crimes—all models perform strongly, with precision mostly above 0.85. Particularly, Property Crimes show nearly perfect precision (≈ 0.99) for all classifiers, suggesting that the models can predict this category with very few false alarms.

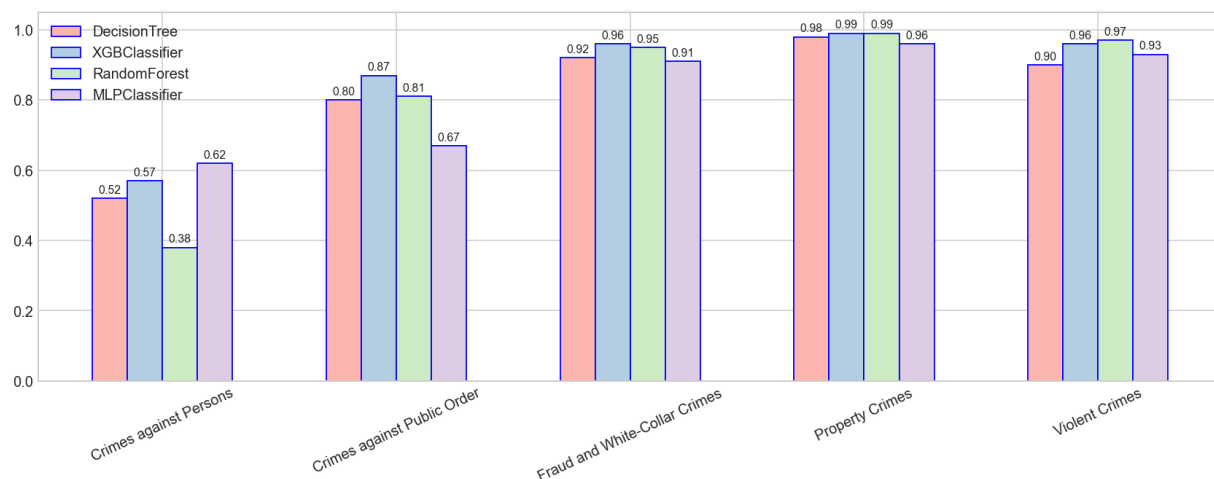


Fig 10:

This chart presents the recall scores of the same classifiers for the above mentioned categories.

Shown in Fig-10, Recall measures how many actual positive cases were correctly identified by the

model minimizing false negatives). For Crimes against Persons, recall is notably lower for Decision Tree (0.52) and MLPClassifier (0.36), but much higher for XGBClassifier (0.92) and Random Forest (0.62). For Crimes against Public Order, Fraud/White-Collar Crimes, Property Crimes, and Violent Crimes, recall values are very

high (mostly 0.87–0.99), indicating that these models successfully captured nearly all true instances of these crime types. This means the models are less likely to miss cases in these categories, making them effective for broader detection.

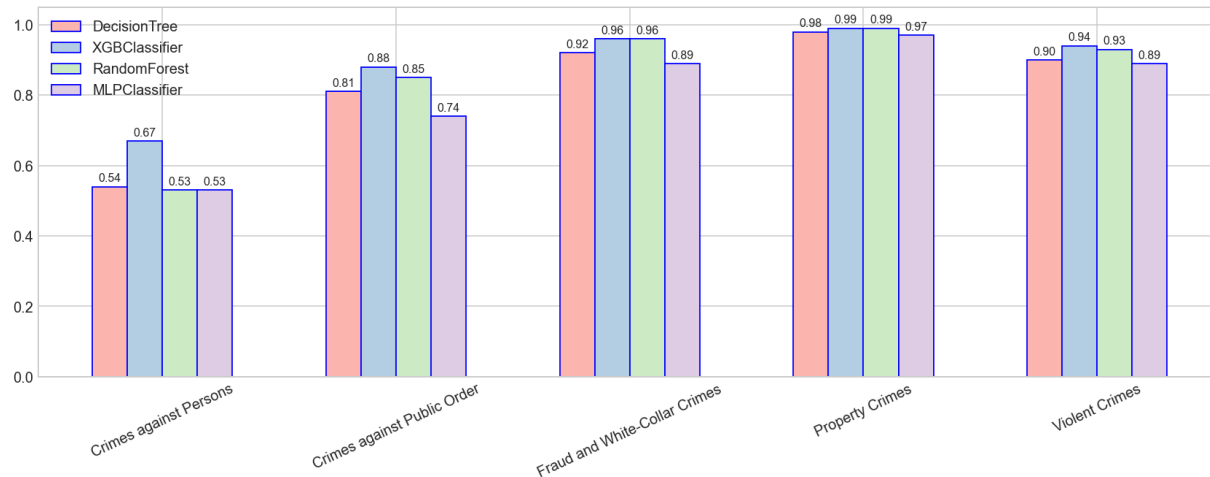


Fig.11

The bar chart supplements the heatmap by providing a more detailed comparison of F1 scores across models and classes as shown in Fig.11. It demonstrates that XGBoost and Random Forest have the highest F1-scores in the majority of categories, particularly in Fraud, White-Collar Crimes, and Property Crimes. The Decision Tree is close behind in these organized categories, but

falls behind in Crimes Against Persons. The MLP Classifier performs well in Violent Crimes and Property Crimes, but falls short in Crimes Against Public Order. Overall, the bar chart reflects the heatmap's trend: while all models perform well in well-defined categories, ensemble techniques regularly give a better mix of precision and recall, resulting in higher F1-scores.

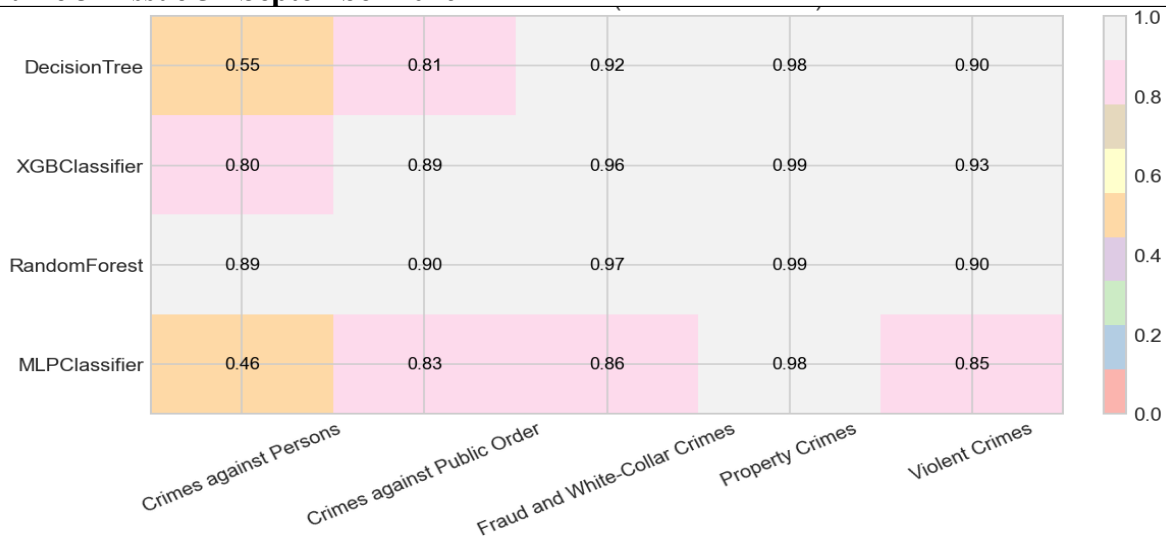


Fig.12

This illustrates the per-class precision of the four models across different crime categories. Random Forest has the highest and most constant precision, with values near 0.99 for property crimes, fraud, and white-collar crimes, and 0.90 for violent crimes. XGBoost also performs well in most classes, particularly Property Crimes (0.99) and Fraud and White-Collar Crimes (0.96), but falls dramatically in Violent Crimes (0.03), indicating a high number of false positives in this

area. The Decision Tree performs moderately, with precision ranging from 0.55 in crimes against persons to 0.98 in property crimes. The MLP Classifier has inconsistent precision, scoring high on Property Crimes (0.98) and Violent Crimes (0.85) but low on Crimes Against Persons (0.46).has shown in Fig.12 These results emphasize that ensemble approaches, notably Random Forest, provide improved precision and reliability across categories.

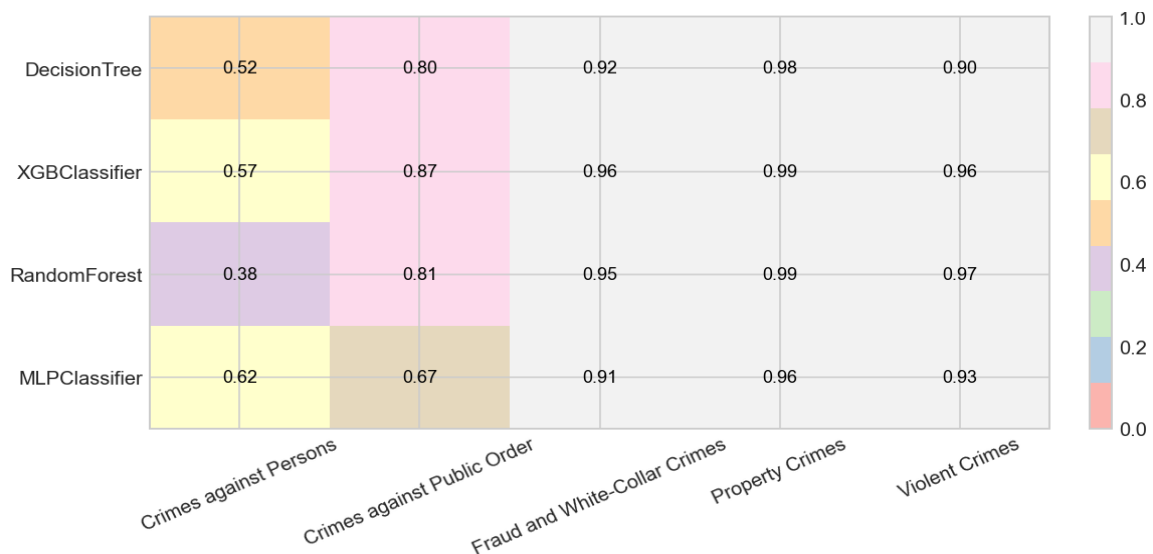


Fig.13

The diagram presents per-class recall, showing how well each model identifies true instances of crimes. Random Forest and XGBoost obtain high recall in most classes, with Random Forest scoring up to 0.99 for Property Crimes and 0.97 for Violent Crimes, but struggling with Crimes Against Persons (0.38). XGBoost maintains balanced recall, with values greater than 0.85 in nearly all categories, indicating constant coverage. The Decision Tree has high recall for Fraud and White-Collar Crimes (0.92) and Property Crimes (0.98),

but slightly lower recall for Crimes Against Persons (0.52). The MLP Classifier outperforms other models in crimes against people (0.62) but has lower recall in crimes against public order (0.67). Overall, shown in Fig.13 the recall findings suggest that ensemble models are effective for the majority of categories, but specific underrepresented classifications, such as Crimes Against Persons, continue to be difficult for all classifiers to predict.

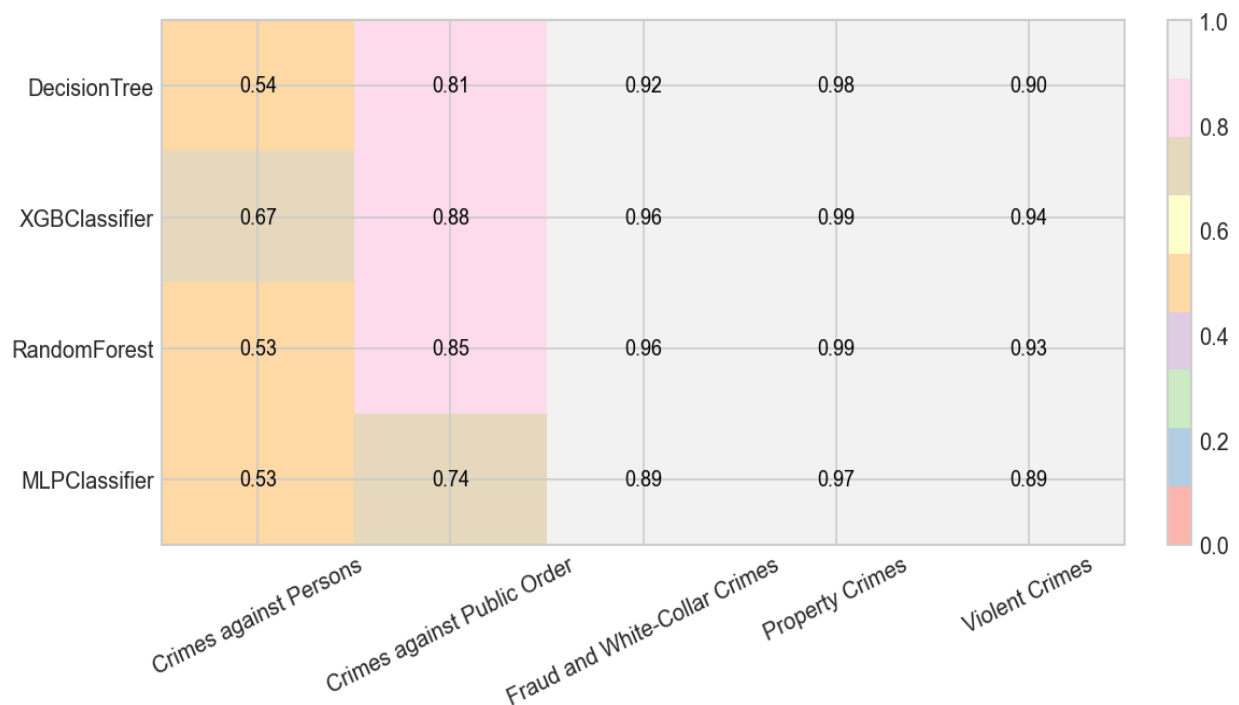


Fig.14

This diagram shows the per-class F1-scores for the four models across various crime categories. Shown in Fig.14 Overall, ensemble approaches Random Forest and XGBoost outperform other models, with near-perfect F1-scores in Property Crimes (0.99) and Fraud and White-Collar Crimes (0.96). XGBoost performs well in most categories, with its lowest F1-score of 0.67 for Crimes Against Persons. The Decision Tree reaches a reasonable balance in categories such as

Fraud and White-Collar Crimes (0.92) and Property Crimes (0.98), but it falls short in Crimes against Persons (0.54). The MLP Classifier has inconsistent performance, performing well in Property Crimes (0.97) but less well in Crimes against Public Order (0.74). These findings highlight how ensemble models deliver more consistent and dependable F1 performance across a wide range of crime categories.

Decision Tree	Actual	Predicted					
			CAP	CAPO	F&WCC	PC	VC
		CAP	22	1	0	0	19
		CAPO	1	267	14	6	44
		F&WCC	0	14	245	4	3
		PC	0	7	1	2307	29
		VC	17	42	6	36	880
XGBoost	Actual						
		CAP	24	2	1	0	15
		CAPO	1	288	8	6	29
		F&WCC	0	10	255	0	1
		PC	0	2	1	2318	23
		VC	5	22	0	15	939
Random Forest	Actual						
		CAP	16	2	2	0	22
		CAPO	0	269	5	11	47
		F&WCC	0	9	254	0	3
		PC	0	1	1	2309	33
		VC	2	19	0	13	947
MLP	Actual						
		CAP	13	5	2	0	22
		CAPO	4	263	19	3	43
		F&WCC	0	11	254	0	1
		PC	0	31	24	2192	97
		VC	6	34	1	8	932

Table 2:

Across all four models, in Table-2 results show varying strengths and weaknesses in crime category prediction. The Decision Tree struggled the most, particularly in distinguishing Crimes against Persons (CAP) (offenses directly targeting individuals) and Violent Crimes (VC) (serious physical harm offenses), with many misclassifications between the two. XGBoost (Extreme Gradient Boosting) consistently outperformed other models, achieving the highest accuracy across categories with minimal errors, especially in Fraud and White-Collar Crimes (F&WCC) (financially motivated, non-violent offenses), Property Crimes (PC) (theft, burglary, and related offenses), and Violent Crimes. Random Forest also delivered strong

results, particularly in Property Crimes and Violent Crimes, though it showed some confusion between Crimes against Persons and Violent Crimes. The MLP (Multilayer Perceptron) Classifier performed reasonably well in Fraud and White-Collar Crimes and Property Crimes but had higher misclassification rates in Crimes against Persons, Crimes against Public Order (CAPO) (offenses disrupting societal norms and rules), and Violent Crimes compared to ensemble methods. Overall, ensemble models, especially XGBoost and Random Forest, proved to be more robust and reliable, while Decision Tree and MLP showed limitations in distinguishing closely related crime categories.

		Accuracy	Precision	Recall	F1-Score
Decision Tree	CAP	0.9385	0.55	0.52	0.54
	CAPO		0.81	0.80	0.81
	F&WCC		0.92	0.92	0.92
	PC		0.98	0.98	0.98
	VC		0.90	0.90	0.90
Random Forest	CAP	0.9571	0.89	0.38	0.53
	CAPO		0.90	0.81	0.85
	F&WCC		0.97	0.95	0.96
	PC		0.99	0.99	0.99
	VC		0.90	0.97	0.93
MLP	CAP	0.9216	0.57	0.31	0.40
	CAPO		0.76	0.79	0.78
	F&WCC		0.85	0.95	0.90
	PC		1.00	0.94	0.96
	VC		0.85	0.95	0.90
XGBoost	CAP	0.9644	0.80	0.57	0.67
	CAPO		0.89	0.87	0.88
	F&WCC		0.96	0.96	0.96
	PC		0.99	0.99	0.99
	VC		0.93	0.96	0.94

Table 3:

The comparative evaluation of classifiers across multiple metrics highlights distinct performance patterns for each crime category as shown in Table-3. For Crimes against Persons (CAP), performance remained the most challenging, with relatively lower recall values across all models, though XGBoost (Accuracy 0.9644, F1 = 0.67) and Random Forest (Accuracy 0.9571, Precision 0.89) showed comparatively better balance than Decision Tree and MLP. In Crimes against Public Order (CAPO), Decision Tree achieved stable performance (Accuracy 0.81, F1 = 0.81), while XGBoost slightly outperformed others with an F1 score of 0.88. For Fraud and White-Collar Crimes (F&WCC), all models demonstrated strong precision and recall, with XGBoost and Random Forest performing nearly identically (F1 = 0.96), reflecting the models' ability to capture financial and non-violent patterns effectively. Property Crimes (PC) showed the highest predictive reliability overall, with Random Forest and XGBoost achieving near-perfect scores (Accuracy and F1 = 0.99), while even MLP reached excellent

results (Accuracy 1.00, F1 = 0.96). Finally, Violent Crimes (VC) were best captured by ensemble models, with XGBoost (Accuracy 0.93, F1 = 0.94) and Random Forest (Accuracy 0.90, F1 = 0.93) outperforming MLP and Decision Tree. Overall, ensemble-based methods, especially XGBoost, consistently delivered superior predictive power, while single models like Decision Tree and MLP showed weaknesses in handling complex categories such as CAP and VC.

V. Conclusion

In this study, a comprehensive crime dataset with over 250,000 records was analyzed to investigate trends in victim demographics, geographic distribution, and temporal trends. A number of preprocessing procedures were used to get the data ready for modeling, such as encoding categorical variables, imputing missing values, and creating extra time-based features to account for possible seasonal impacts. The most commonly reported crimes, according to exploratory research, were car-related offenses and thefts of various kinds.

These crimes were disproportionately committed by people in their early to middle adult years and were typically concentrated near the end of the week. These findings highlight how important it is to correct for class imbalances and choose pertinent predictive characteristics in order to improve model performance.

The study assessed various base classifiers, including Decision Tree, Random Forest, K-Nearest Neighbors, Gaussian Naïve Bayes, Logistic Regression, and XG Boost, each with unique strengths to the prediction framework. Linear and probabilistic models were simple and computationally efficient, whereas tree-based techniques were highly interpretable and capable of capturing non-linear patterns. Ensemble-based approaches, particularly XG Boost, outperformed both accuracy and macro-averaged F1-scores, demonstrating their ability to handle class imbalance and complicated feature interactions. To improve model robustness, considerable feature engineering was performed, ensuring that important features were identified and optimized for classification. Furthermore, the adoption of fusion modeling methodologies such as soft voting increased predicted reliability when compared to a single model, highlighting the need of merging varied classifiers. Overall, the findings show that hybrid ensemble techniques augmented with feature engineering might be useful decision-support tools for predicting crime patterns and guiding law enforcement operations.

VI. REFERENCE

- A. P. Apporva, H. C. Impana, S. L. Siri and M. R. Varshitha, "Automated Criminal Identification By Face Recognition Using Open Computer Vision Classifiers," in Proc. Int. Conf. Computing Methodologies and Communication (ICCMC), 2019.
- P. Chhoriya, "Automated Criminal Identification System using Face Detection and Recognition," Int. Res. J. Eng. Technol., vol. 6, no. 10, Oct. 2019.
- K. Goyal, K. Agarwal and R. Kumar, "Face Detection and Tracking," in Proc. Int. Conf. Electronics, Communication and Aerospace Technology (IFECA), 2017.
- N. I. Hassan, N. M. Tahir and H. Hashim, "People Detection System using YOLOv3 Algorithm," in Proc. Int. Conf. Control System, Computing and Engineering (ICCSCE), Aug. 2020.
- Pratibha, A. Gahalot, Uprant, S. Dhiman and L. Chouhan, "Crime Prediction and Analysis," IEEE, Stanford University, 2017.
- J. Lai and S. Maples, "Developing a Real-Time Gun Detection Classifier."
- J. Azeez and D. J. Aravindhar, "Hybrid Approach to Crime Prediction using Deep Learning," in Proc. Int. Conf. Advances in Computing, Communications and Informatics (ICACCI), 2015.
- J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016.
- T. Ahmad, Y. Ma, M. Yahya, B. Ahmad, S. Nazir and A. ul Haq, "Object Detection through Modified YOLO Neural Network," Hindawi Scientific Programming, vol. 2020, 2020.
- W. Chen, H. Huang, S. Peng, C. Zhou and C. Zhang, "YOLO-face: A Real-Time Face Detector," Springer-Verlag GmbH Germany, part of Springer Nature, 2020.
- C. Kim, L. Lee, T. Han and Y.-M. Kim, "A Hybrid Framework Combining Background Subtraction and Deep Neural Networks for Rapid Person Detection," J. Big Data, vol. 5, no. 1, 2018.
- S. Loganathan, G. Kariyawasam and P. Sumathipala, "Suspicious Activity Detection in Surveillance Footage," in Proc. 2019 Int. Conf. Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, UAE, 2019, pp. 1-4, doi: 10.1109/ICECTA48151.2019.8959600.

- N. Bordoloi, A. K. Talukdar and K. K. Sarma, "Suspicious Activity Detection from Videos using YOLOv3," in Proc. 2020 IEEE 17th India Council Int. Conf. (INDICON), New Delhi, India, 2020, pp. 1-5, doi: 10.1109/INDICON49873.2020.9342230.
- U. M. Kamthe and C. G. Patil, "Suspicious Activity Recognition in Video Surveillance System," in Proc. 2018 Fourth Int. Conf. Computing Communication Control and Automation (ICCUBE), Pune, India, 2018, pp. 1-6, doi: 10.1109/ICCUBE.2018.8697408.
- U. V. Navalgund and K. Priyadharshini, "Crime Intention Detection System Using Deep Learning," in Proc. 2018 Int. Conf. Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 2018, pp. 1-6, doi: 10.1109/ICCSDET.2018.8821168.
- C. V. Amrutha, C. Jyotsna and J. Amudha, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video," in Proc. 2020 2nd Int. Conf. Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 2020, pp. 335-339, doi: 10.1109/ICIMIA48430.2020.9074920.
- K. Ouivirach, S. Gharti and M. N. Dailey, "Incremental Behavior Modeling and Suspicious Activity Detection," Pattern Recognit., vol. 46, no. 3, pp. [pages], 2013.
- N. Nasaruddin, K. Muchtar, A. Afdhal, et al., "Deep Anomaly Detection through Visual Attention in Surveillance Videos," J. Big Data, vol. 7, no. 87, 2020. doi: 10.1186/s40537-020-00365-.
- S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.
- "A Three-Stage Unsupervised Dimension Reduction Method for Text Clustering," J. Comput. Sci., vol. [volume], 2014.
- "An Intelligent Document Clustering Approach to Detect Crime Patterns," Procedia Technol., vol. [volume], 2013.