Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

An Enhanced Framework for Intelligent Crime Pattern Recognition

Iftikhar Ahmad
Selinus University of Sciences and Literature, Ragusa - Italy

Usama Arif
Information Technology University, Lahore

Muhammad Khurram Afzaal
Qubit Cloud Nexus

Muhammad Rahman Arif
QUBIT Cloud Nexus

Ahmed Shahzad
The University of Faisalabad

ABSTRACT

Crime pattern recognition is one of the most significant activities of the common of the intelligent and LEO (Law Enforcement Organizations) all over the world. Typically, they collect national and international crime-related records to prevent future attacks and use a limited number of law enforcement resources in an optimum manner. Now a days, crime ratio is increasing at a higher rate. It is tough to predict the occurrence of crime due to many different elements. It is difficult to predict which criminal would commit the crime, but the place and the possibility of the incident of high crime rate in a specific time of day and high crime rate in a month of the year can very well be predicted. And also, the most common occurrence of a type of crime can be predicted. The system developed will help to speed up the method of solving crimes for law enforcement agencies. Data mining is a very good tool that can be used efficiently for studying huge databases and extracting important analytical results. In this work, Crime data for the last 15 years is analyzed using two different data-processing approaches. Two predictive models, Random Forest and boosted decision tree, are implemented, and a crime prediction accuracy obtained 96% when predicting crime.

Keywords: Crime Pattern Recognition, Data Mining, Predictive Models, Random Forest, Boosted Decision Tree

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

INTRODUCTION

Define Crime

Crime rate is rapidly increasing with the passage of time because of multiple reasons. Due to the rapid increase in population and technological evolution, the vastness of crimes is also a major issue, so in this modern world, the detailed examination and analysis of crimes has become a very important and essential activity. There are different organizations that are working and gathering crime data on a national and international level. These organizations include the law enforcement organizations and intelligence agencies that collect data regarding crimes to fend off future attacks. (Akinmuyisitan et al., 2025) During the data collection process about crimes, they face different types of difficulties because they have to collect a large amount of data, and they use different methodologies to analyze this data, which has a vast variation. As a result of this process, the use of manpower is not so effective, and even productivity decreases. These are the major problems that are faced by different intelligence agencies and law enforcement organizations (Akinmuyisitan & Cosmas, 2024).There are different reasons and factors due to which it is difficult to predict the occurrence of crimes. There are different methodologies and techniques available to prevent crimes, but despite these techniques, it is not easy to control crimes to a considerable extent. There are different experts who can give predictions on behalf of their knowledge and experience about the chances of occurrence of different types of crimes and even about the area and place (Akinmuyisitan et al., 2025).But it is hard to forecast which criminal would commit the crime. Crimes of different types are increasing day by day, so computerized techniques and systems are used to trace the crimes. The analysts of computer data are helping the officers of law enforcement to increase and improve the procedure of solving crimes. Data mining paradigm methodology is very helpful today as it is additionally used at a faster rate to solve crimes (Mandalapu et al., 2025).For the purpose of constructing the spatial (facts and figures) data mining technique, a practical clarification for a wide range of audience of law enforcement officers, the important and necessary technology has become reasonable, availabl,e and economical. Now there are different types of software for data analysis, computer hardware, and tools for mapping that are commonly available to police departments at very reasonable rates, which help them to enable the visualization of opaque spatial data. A good quality of data is required in large quantities to see the pattern in offensive incidents over the course of a remarkable socioeconomic cycle has been available. Moreover, many technologists who are very skilled in data mining are now appearing. Given the following diagram, show how can predict the crime (Elluri & Mandalapu,2024).

Figure 1: Crime Prediction

There are different types of terms that are used in police departments and criminal justice. These terms differentiate and contrast in a data mining system. Suspicious refers to the person who is believed to have done or committed the offence and crime. The suspicious may be recognized or unrecognized (Ștefănoaia et al., 2025). The suspicious person cannot be punished by the law until his crime and offense is not proved. However, more restrictions are required regarding narcotics

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

and crimes. Information about the criminals of sex is made publically to warn and alert others in the area, but the recognition and identification of the victim and the sufferer remain in secret. So as a data minor, the analyst has to deal with different situations and issues like he has to deal with all these private versus public data problems and issues so that data mining modeling procedure does not violate or break on these lawful boundaries (Ștefănoaia et al., 2025).

Data mining has a relation with the interdisciplinary area of knowledge discovery in databases. Through the process of analysis and inference, raw data is collected in data mining. This data is used to provide correct forecasting and predictions, and its application is made in real-world situations. Association, classification, clustering, and forecasting of data is a very important thing, and their assumption is bellowed:

Association

Identifying and recognizing the connection among data setting up relationship and association between data that exit together in a stated record.

Classification

In this process data is recognize after discovering and sorting procedure on the bases of similarities. The most common application which is used in data mining classification is decision tree, if then rules and neural networks are those tools which are used for classification analysis.

Clustering

Group of those facts which are unknown previously and even not taken into notice they are finding and visually presenting in clustering process. Survival analysis and neural networks are the common and basic tools which are used in clustering. Data of heterogeneous form is categorized into homogeneous clusters.

Forecasting

It is a process through which data and patterns are discovered and may lead to appropriate prediction (Wu et al., 2018). Growth of economy and the quality of life is badly affected by the crime which is a socio-economic problem. There are different factors which affect the crime rates like poverty, climate, unemployment and education. Crime prediction about these factors is given in previous researches (Rossi & Bianchi, 2025).

Analysis of Crimes

Crimes of different types are analyzed by the analysts who are appointed by the law enforcement agencies. There are different reasons to analyze the crimes some of them depend upon the geographic efforts and community culture etc. But the major reasons of crime analysis are given below.

Crimes analysis is necessary to get the benefits of plentiful information existing in criminal justice system, law enforcement organizations and public domain.

To increase the practice of limited law enforcement resources crimes are analyzed.

Through crimes analysis information about series in an ongoing, timely manner, pattern and about specific and general crime trends is given to law enforcement

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

agencies.
To meet and fulfill the law enforcement requirements of a converting society
To understand the behavior of criminals crimes are analyzed.
For the purpose to detect the crimes and how to prevent from these crimes their analysis is necessary (Abuzari, 2024)

Issues and Challenges on Crime

Data Collection and Integration

Input information is imperative to utilize in the testing system & training system in the crime investigation methods. For the conduction of the crime type, the training procedure is utilized. For validation of the algorithm, the testing method is employed (Li & Han, 2025). News, social media, various detectors, and unlawful activity indexes can provide input data. These streams collected data from the state agencies. Compiled information is broad proportions of data as an impact. Formats may vary of the compiled information, the data may be unstructured. Various databases are used to store data. (Abuzari, 2024) Challenges like preparation, transformation, and integration of data rise with problems of data collection. Researchers are working for finding the solutions to these issues. However, it is a massive challenge to explore and extract overlooked proficiency from enormous sizes of data. Methods like entity extraction or grouping and filtering are beneficial to collect & integrate data. (Abuzari, 2024).

Pattern of Crime

The subjects of crime models are treating with determining and prophesying the mysterious crime. These days, there is a continual rise in the crime index and the patterns of crime are constantly altering. It is the reason that the aspects of crime are problematic to be clarified and indicated (Li & Han, 2025). For detection of violations, researchers focus on finding such a way to develop models that will help in crime prevention and detection. The test is demonstrating the crime assault practices that help misdemeanor recognition even though crime designs are advancing. The prescience and analytical procedures may be efficient to find and manipulate the prototype of crime. The crime paradigm should be capable of augur and recognize illegal conduct. (Situmeang et al., 2024).

Inter- Region Spatial Patterns

A pattern that gives suggestion to the geographical impact in the city among regions is called an inter-region spatial pattern (Khan & Rahman, 2025). Due to the differing distribution of the data it is difficult to build a steady and consistent framework for all administrative divisions (Situmeang et al., 2024). There is a potentiality to utilize transfer culture to assemble exact offense forecasting for all the boroughs that are motivated by the presence of civic transportation between various administrative divisions.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

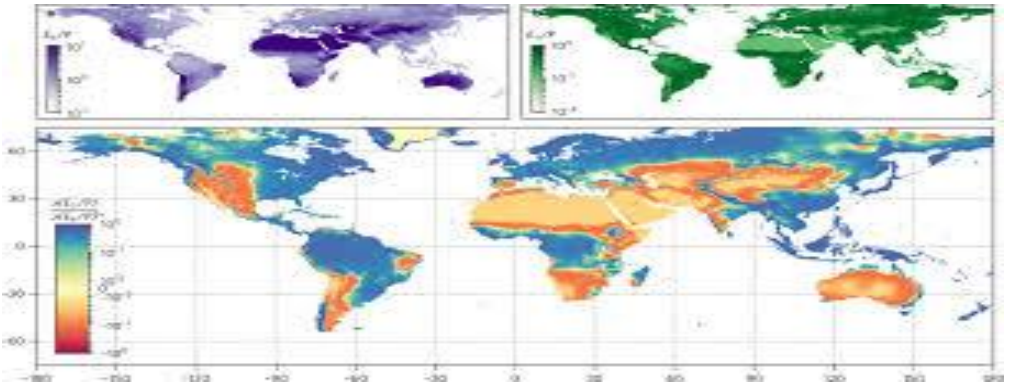


Figure 2: Inter Region Spatial Patterns

Within a borough for dissimilar regions we account inter-region spatial patterns in which almost twain spatially adjacent regions have comparable offense numbers and when the spatial space between two regions increases then the difference of crime tends to expands and then regions from different boroughs do not act upon according to above mentioned inter region spatial pattern. (Mohammadinia & Alizadeh, 2023)

Intra-Region Temporal Pattern

These patterns assist us to perceive and find out that how crime expands over time in a city for a region and intra region sequential designs that offenses modify fluently over closest period and modify over succeeding weeks in a region periodically (Reddy & Rao, 2024). Afterwards we grip transfer learning techniques on the basis of parameters of single borough to teach and instruct the parameters of other administrative division within the city predictions of the crimes based on such parameters that are well trained. (Mohammadinia & Alizadeh, 2023)

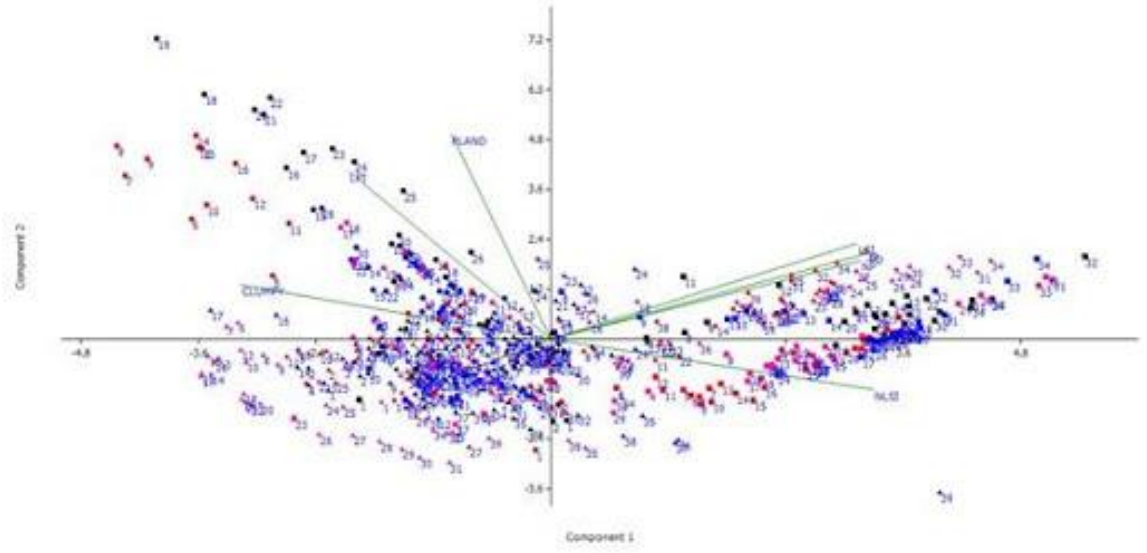


Figure 3: Intra Region Temporal Patterns

We provide such framework which gives predictions of crimes of each and every area

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

in the coming of different boroughs to police departments within the city, afterwards help them to intercede to fend off crimes (Adegoke & Babatunde, 2024).

REVIEW OF LITERATURE

Different research scholars have pointed out the issues concerning crime limitation and have suggested diverse criminal offense foresight algorithmic programs. The precision of the prediction relies upon the properties chosen and the dataset is utilized as a source of perspective (Adegoke & Babatunde, 2024).

Misconduct erudition mining is the exercise of information mining techniques for transgression analysis. Various inquiries have been resolved in this field and here we have explained a fraction of this information, and. Criminal activities can be separated into subcategories subject to multiple models (Hassan & Al-Jubouri, 2023). Eight transgression classes are given. They are unlawful transportation misdemeanors, sex crimes, robbery, deception, torching, drug offenses, cybercrimes and brutal transgressions. They have given definitions for every classification at the nearby law requirement level and public law implementation level. IPTC (international press media transmission chamber) has given an alternate order where violations are separated into the atrocity corporate transgression coordinated transgression, and so on .There are different transgression information mining strategies accessible (Adegoke & Babatunde, 2024). The most normally utilized techniques are element extraction, bunching strategies affiliation rule mining, successive example mining and grouping. There were diverse attempts to inspect various categories of crimes using computerized maneuverings however there is no brought together arrangement representing how to implement those modes to multiple transgression sorts (Adegoke & Babatunde, 2024). they have introduced a structure that incorporates a connection between the transgression information mining strategy also, offense type features. They have proposed a structure that incorporates a connection between the transgression information mining strategies also transgression type qualities (Ahmed & Wu, 2023).

A few existing skeletons use transgression information-digging designs for crime inquiry for illustration a regional offense research program a part of information investigating structure for crime design discriminating facts and an opiates network in the Tucson police office (Ahmed & Wu, 2023) (Mandalapu et al., 2023). An assortment of criminal examination steps is given. Among them, steps for example, hotspot discovery, transgression examination, transgression design perception are critical (Mahmoudi & Kazemi, 2024). In transgression design representation a period arrangement can be drawn between the transgression recurrence and the time and utilizing it fascinating transgression patterns can be distinguished. In expansion to these means has given some different investigation steps, for example transgression clock episode location and closest police headquarters discovery (Ahmed & Wu, 2023).

Jeffrey T. Ward, James V. Beam, Kathleen A. Fox built up a MIMIC model for Exploring contrasts in discretion across sex, race, age, schooling, and language and make an inference that separated from race, testing bunch contrasts in poise with a watched scale score is to a great extent fair. Testing bunch contrasts in components utilizing watched sub-scores are oftentimes one-sided and by and largely unsupported

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

(Ahmed & Wu, 2023).

Bromley and Nelson uncover worldly examples of transgression to anticipate liquor-related transgression in Worcester city. They likewise give a significant understanding of the spatial attributes of liquor-related transgression. The creators inspect the examples of transgression and turmoil at the road level by recognizing hotspots (Ehsanpour, 2025). Ratcliffe proposes three sorts of worldly and spatial hotspots for transgression design recognition (Ehsanpour, 2025).

The creator additionally shows how the spatial and worldly qualities consolidate through his hotspot grid. In any case the creator didn't make a difference in any AI methodology to foresee transgression. Creators dissect four classifications of transgression information that incorporate alcohol law infringement attacks and batteries defacement and clamor protests. Various classes of transgression show distinctive worldly examples. Brower and Carroll explain the transgression development through the city of Madison utilizing GIS planning. The creators examine the connections between high-thickness liquor sources and various areas (Ehsanpour, 2025). distinguish transgression hotspots utilizing Kernel Density Estimation (KDE) to foresee spatial transgression designs. Also, in another examination, Nakaya furthermore, Yano makes transgression hotspots with the assistance of KDE. In any case, they consolidate worldly highlights with transgression hotspots investigation. Nath utilized a semi-administered grouping strategy for identifying transgression designs. In the creators propose an example identification calculation named Arrangement Finder to identify examples of a transgression naturally. Creators study crime percentage surmising issues utilizing Point-Of-Interest information and taxi stream information. Focal point information and taxi stream information are utilized to improve the segment data and the geological vicinity relationship (Ehsanpour, 2025) (Rodriguez et al., 2024). Early investigations in criminology have attempted to show the connection between transgression and different affecting elements for example socioeconomics financial matters, and joblessness. Indeed as of late, these examinations are being kept on showing connections among transgression and different affecting elements. For instance, Kelly thought about the connection between disparity also, a transgression for metropolitan areas in the USA and exhibited that socially distraught individuals carried out most rough violations; the examination presumed that the most hindered individuals from society who lived in regions of high imbalance confronted more prominent weight and motivators to submit the transgression, prompting perpetrating fierce violations (Chen, 2025) (Rodriguez et al., 2024). Hojman considered disparity, joblessness, also, transgression in Latin American urban communities by assessing the variety among urban areas and utilizing a relapse examination that showed the function of obstacles, destitution, and disparity as reasons for transgression (Chen, 2025). Poveda examined financial and vicious transgression in seven urban areas in Columbia. Their investigation demonstrated that urban areas' monetary hardship and high populace thickness are solid factors in manslaughter rates (Chen, 2025).

MATERIAL AND METHOD

In this project we use Vancouver crime dataset, we used year, month, day, and neighborhood as input to predict the type of crime committed. First we converted the

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

type of crime values into numeric format.

This dataset was fed into two different models decision tree and random forest. Both of these were trained separately, after training we used some data for prediction if both model predicted the same then its good, if both predicted differently which mean there is 50% chance of both.

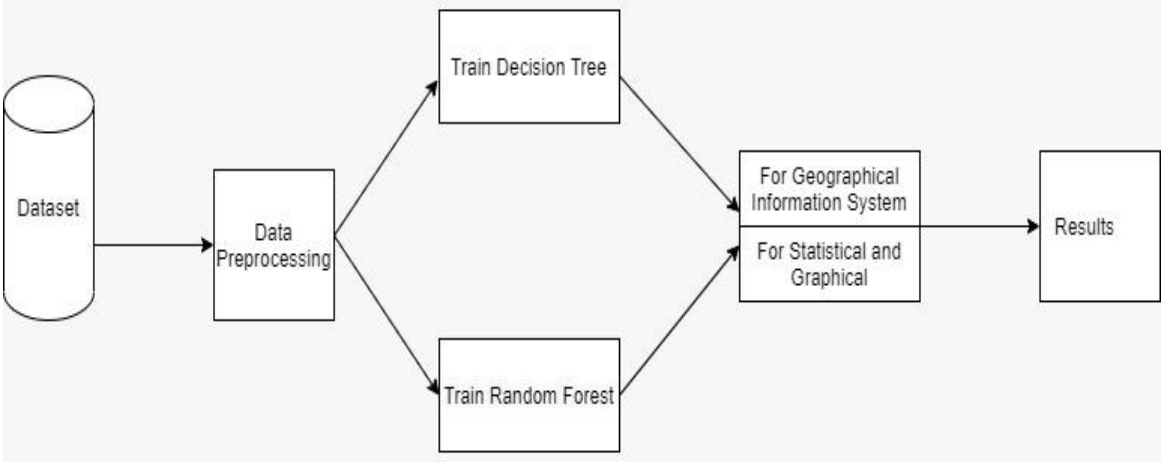


Figure 4: Methodology

Dataset

The initial datasets were collected from the city of Vancouver's open data catalog. For this initiative, there are two types of datasets used: crime and community. Since 2003, the criminality dataset has been gathered by the VPD and is reorganized every Sunday morning. It contains statistics on the type of crime dedicated and the time and place of the crime. In the Geographic Information System, the community dataset includes the limits for the 22 local areas of the city (GIS). The criminal dataset is used to evaluate the data in this project and neighborhood datasets are used to draw maps.

O38912

	A	B	C	D	E	F	G	H	I	J
38910	Break and Enter Commercial	2004	10	17	11	0	X WATER S	Central Bu	492414.4	5458982
38911	Break and Enter Commercial	2004	10	23	2	57	X WATER S	Central Bu	492414.4	5458982
38912	Break and Enter Commercial	2005	11	30	0	1	X WATER S	Central Bu	492414.4	5458982
38913	Break and Enter Commercial	2008	9	3	3	50	X WATER S	Central Bu	492414.4	5458982
38914	Break and Enter Commercial	2012	4	21	18	0	X WATER S	Central Bu	492414.4	5458982
38915	Break and Enter Commercial	2012	7	24	23	2	X WATER S	Central Bu	492414.4	5458982
38916	Break and Enter Commercial	2012	9	30	17	0	X WATER S	Central Bu	492414.4	5458982
38917	Break and Enter Commercial	2016	2	29	0	36	X WATER S	Central Bu	492414.4	5458982
38918	Break and Enter Residential/Other	2004	5	13	9	0		Mount Ple	492807	5457236
38919	Break and Enter Residential/Other	2006	6	21	4	0		Oakridge	491854	5452682
38920	Break and Enter Residential/Other	2006	3	19	5	39		Riley Park	492276	5454816
38921	Break and Enter Residential/Other	2003	8	30	3	30	10XX ALBE	West End	491059.5	5459122
38922	Break and Enter Residential/Other	2003	9	19	15	20	10XX ALBE	West End	491059.5	5459122
38923	Break and Enter Residential/Other	2003	12	1	10	0	10XX ALBE	West End	491059.5	5459122
38924	Break and Enter Residential/Other	2004	6	9	17	0	10XX ALBE	West End	491059.5	5459122
38925	Break and Enter Residential/Other	2004	11	3	10	30	10XX ALBE	West End	491059.5	5459122
38926	Break and Enter Residential/Other	2005	3	18	3	0	10XX ALBE	West End	491059.5	5459122
38927	Break and Enter Residential/Other	2009	5	11	11	11	10XX ALBE	West End	491059.5	5459122
38928	Break and Enter Residential/Other	2013	1	29	18	0	10XX ALBE	West End	491059.5	5459122

Figure 5: Dataset

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Preprocessing

In order to fill empty cells, remove needless columns and increase many functions, the original dataset must be preprocessed. Fig: 3.3 Datasets preprocessed.

	A	B	C	D	E	F	G	H	I	J	K	L
1	TYPE	YEAR	MONTH	DAY	HOUR	MINUTE	HUNDRED_BLOCK	NEIGHBOU	X	Y	Latitude	Longitude
2	Other Theft	2003	5	12	16	15	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
3	Other Theft	2003	5	7	15	20	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
4	Other Theft	2003	4	23	16	40	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
5	Other Theft	2003	4	20	11	15	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
6	Other Theft	2003	4	12	17	45	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
7	Other Theft	2003	3	26	20	45	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
8	Break and Enter Residential/Other	2003	3	10	12	0	63XX WILTSHIRE ST	Kerrisdale	489325.6	5452818	49.22805	-123.147
9	Mischief	2003	6	28	4	13	40XX W 19TH AVE	Dunbar-So	485903.1	5455884	49.25556	-123.194
10	Other Theft	2003	2	16	9	2	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
11	Break and Enter Residential/Other	2003	7	9	18	15	18XX E 3RD AVE	Grandview	495078.2	5457221	49.26773	-123.068
12	Other Theft	2003	1	31	19	45	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
13	Mischief	2003	9	27	1	0	40XX W 21ST AVE	Dunbar-So	485853	5455684	49.25376	-123.194
14	Break and Enter Residential/Other	2003	4	19	18	0	18XX E 3RD AVE	Grandview	495093.7	5457230	49.26781	-123.067
15	Break and Enter Residential/Other	2003	9	24	18	30	18XX E 3RD AVE	Grandview	495103.8	5457221	49.26773	-123.067
16	Break and Enter Residential/Other	2003	11	5	8	12	63XX WINDSOR ST	Sunset	493790.5	5452631	49.22643	-123.085
17	Break and Enter Commercial	2003	9	26	2	30	10XX ALBERNI ST	West End	491067.7	5459114	49.28471	-123.123
18	Break and Enter Residential/Other	2003	10	21	10	0	18XX E 3RD AVE	Grandview	495119.3	5457230	49.26781	-123.067
19	Other Theft	2003	1	25	12	30	9XX TERMINAL AVE	Strathcona	493906.5	5457452	49.2698	-123.084
20	Offence Against a Person	2003	2	12			OFFSET TO PROTECT PRIVACY		0	0	0	0

Figure 6: Preprocess of Dataset

Before applying algorithms, the dataset changed into a usable form. We executed two approaches for data processing and matched the results

Approach 1:

Both variables are changed into binary variables 0 and 1 in the first approach. All the areas and days were transformed into characteristics. Just "1" is assigned the required variable and "0" is assigned to all other variables. Both variables of "0" are essentially dummy. This allows the algorithm to generate extra variables and avoids skewing of the data on one side. Skewed data experiments provide 98.9 percent incorrect, inaccurate precision.

Approach 2:

Categorical variables with unique IDs are transformed in the second method into numerical variables. There are separate IDs for all forms of crime and communities. The crime form ID 10 is given to vehicle robbery for example. The same algorithms with the same parameters and validation procedures are applied for both approaches. The validity of the classifier algorithm shall be assessed with a 5-fold cross-validation. The goal of the algorithm is the category of crime. Cross-validation avoids the issue of over fitting and makes sure that the forecast model performs well on new

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)
Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

invisible results.

Decision Tree

Decision tree analysis is typically a method for prediction modeling and can be used in many fields. Decision trees could be built using an algorithmic method, which could segment the data set according to various conditions. Decision trees are the most efficient algorithms in the supervised algorithm group. It is an ensemble technique that is well than a single decision tree as by averaging the outcome it decreases the over-fitting.

You will use them both for regression and classification activities. The two key entities in a tree are decision nodes, in which the data is separated and the outcomes are left. The following is an example of a binary tree to predict whether an individual is suitable or unsuitable for providing different information such as age eating habits and practice.

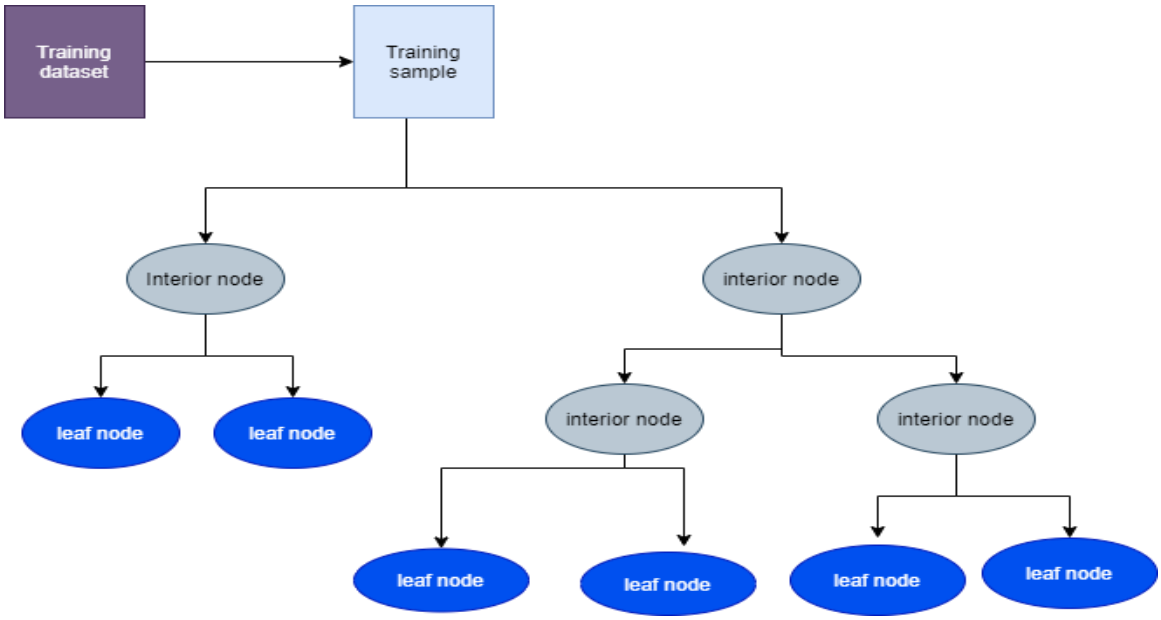


Figure 7: Decision Tree

There are decisions nodes and final conclusions in the above-mentioned decision tree. We have two kinds of decision-making bodies.

Classification Decision Trees: The variable of decision is categorical; the latter decision tree is a case in point of the decision tree classification.

Regression Decision Trees: The decision vector is constant in such decision-making trees.

Classification

Sklearn Classifier- Decision Tree

The decision tree model, as defined in the associated work, is a supervised

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

classification method of learning. The purpose of the classification is to develop a method, through basic decision-making rules extracted from evidence, that forecasts the value of a target variable. A deeper tree implies nuanced guidelines for decision making and that allows a better model fitting. If y is a predictable attribute and X are the predictive attributes, then the syntax of the classifier can be written as:

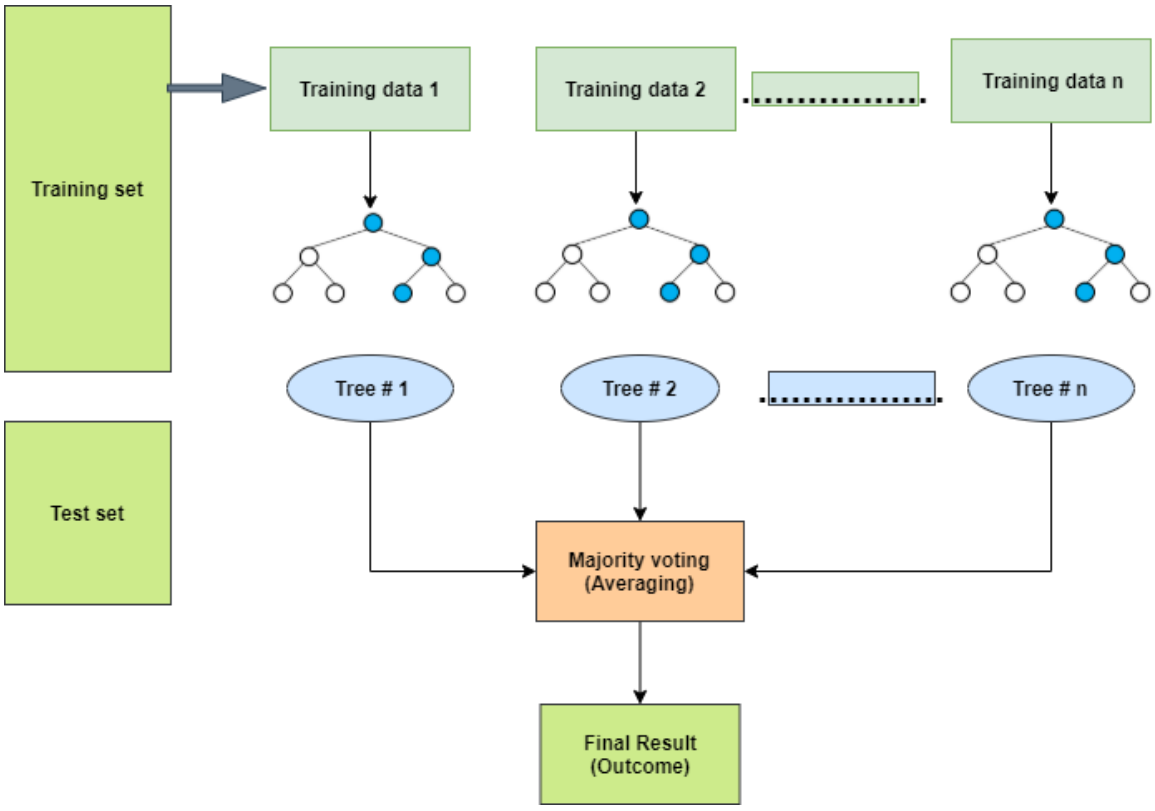
Decision Tree Classifier from sklearn tree import
dt = Decision Tree Classifier (criterion = 'entropy')
#Object Criteria Object of Decision Tree Classifier
dt.fit(X, y)

Entropy offers the knowledge obtained by a rule of decision. Therefore, Entropy was used to separate branches for this predictive model.

Random Forest:

Random forest is an algorithm in supervised learning used both for classification and regression. But it is also used with issues of grouping. As we know, a forest of trees means more robust forest. It is more trees. Likewise, random forest algorithms build decision trees on data samples and then predict each and pick, by vote, the best solution.

1 − First, begin with random sample collection from a specific dataset.
2 − This algorithm will create for each sample a decision-tab. Then any decision tree will deliver the prediction outcome.
3 − Any expected outcome will be voted through this step.
4 − Ultimately take the most voted answer for the forecast as the final outcome.



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Figure 8: Random forest

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Sklearn Classifier- Random Forest

This classifier suits a variety of decision tree classifiers to different sub-samples of the data set monitors the over-fitting of the data and increases predictive accuracy by means of averaging. In the case of singular decision-making trees this classifier is used to ensure strong exactness. If y is a predictable attribute and X are the predictive attributes then the syntax of the classifier can be written as:

Random Forest Classification from sklearn.tree import

dt = Random Forest Classifier () #Object of classifier

dt.fit (X, y) [10]

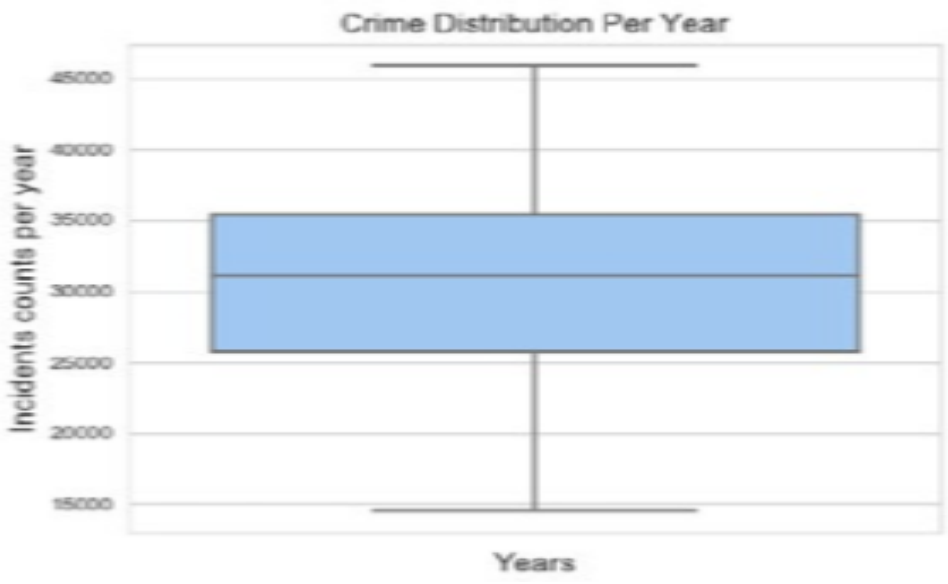
In both methods, we used an improved decision tree algorithm to compare the results. We used the ensemble process and decision-making style tree for both methods to Adaptive Boosting (AdaBoost). AdaBoost is a meta-algorithm merging many poor students to increase the classification. The number of branches was overall 20. Accuracy and time for training for approach 1 respectively was 54.47 percent, while approach 2 was close.

After preprocess the dataset we will show the result in different types. Given the following three types of results.

1. Statistical
2. Trend Base
3. Geographical Information system

Statistical Analysis

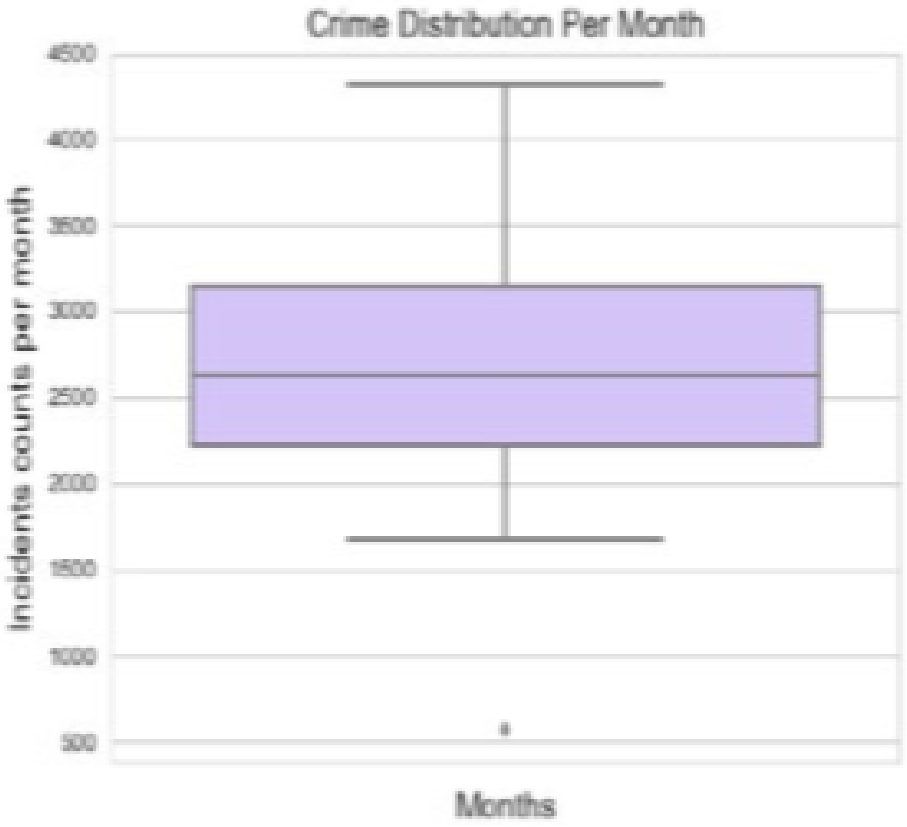
First of all we will perform our dataset results on statistical form as like given the following diagram. The distribution by year, month and day of the crime-based data set out in the following Figure.

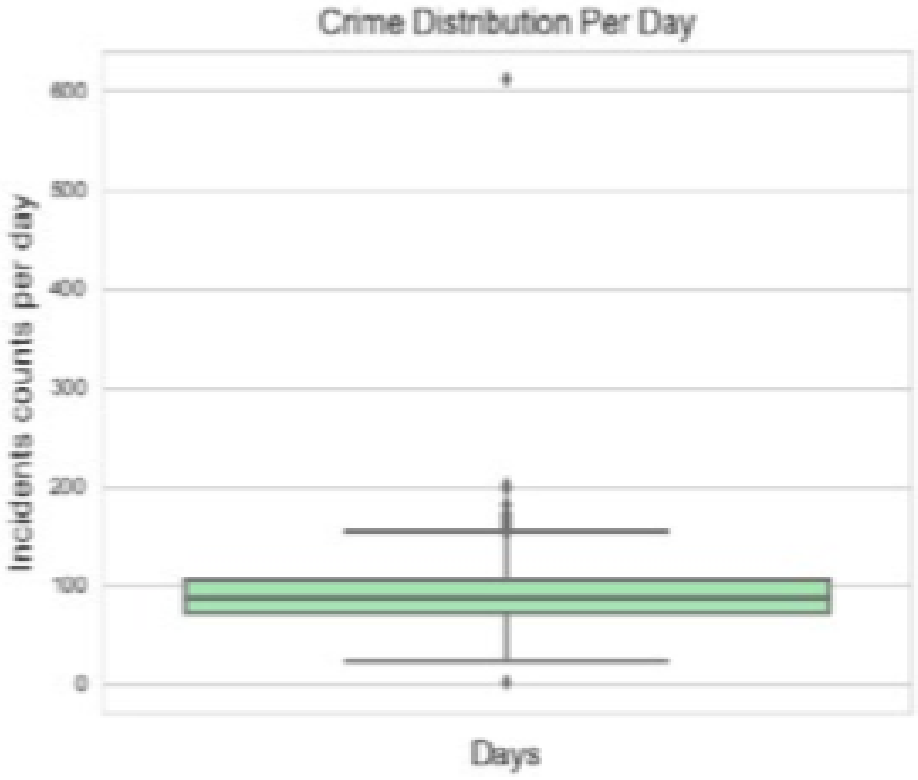


Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about





Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Figure 9: Statistical Result

Trend Analysis

In the second we will show our result as a trend base as like given the following diagram. Trend means a procedure used in technological analysis that efforts to forecast the future offense rate stand on recently viewed trend data is called trend analysis.

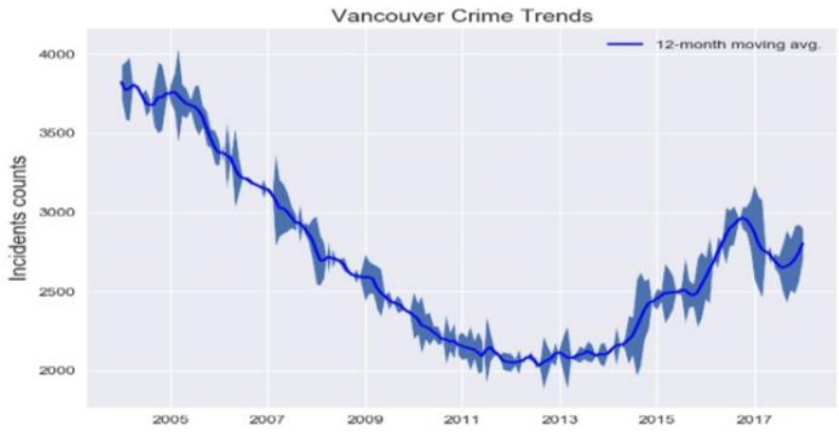


Figure 10: Trend Base Result

Geographical Analysis

At the end we will show the results of crime in geographic information system. A potent computational method for crime mapping is the Geographic Information System (GIS). It displays the positions of the crime series on a single map with separate geographical details, which allows police officers to make organizational and tactical decisions.

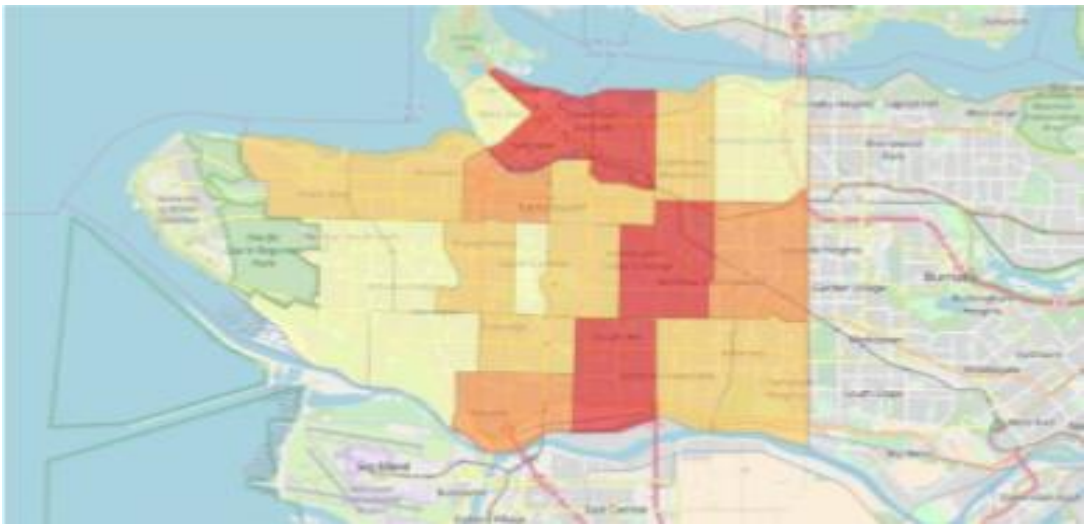


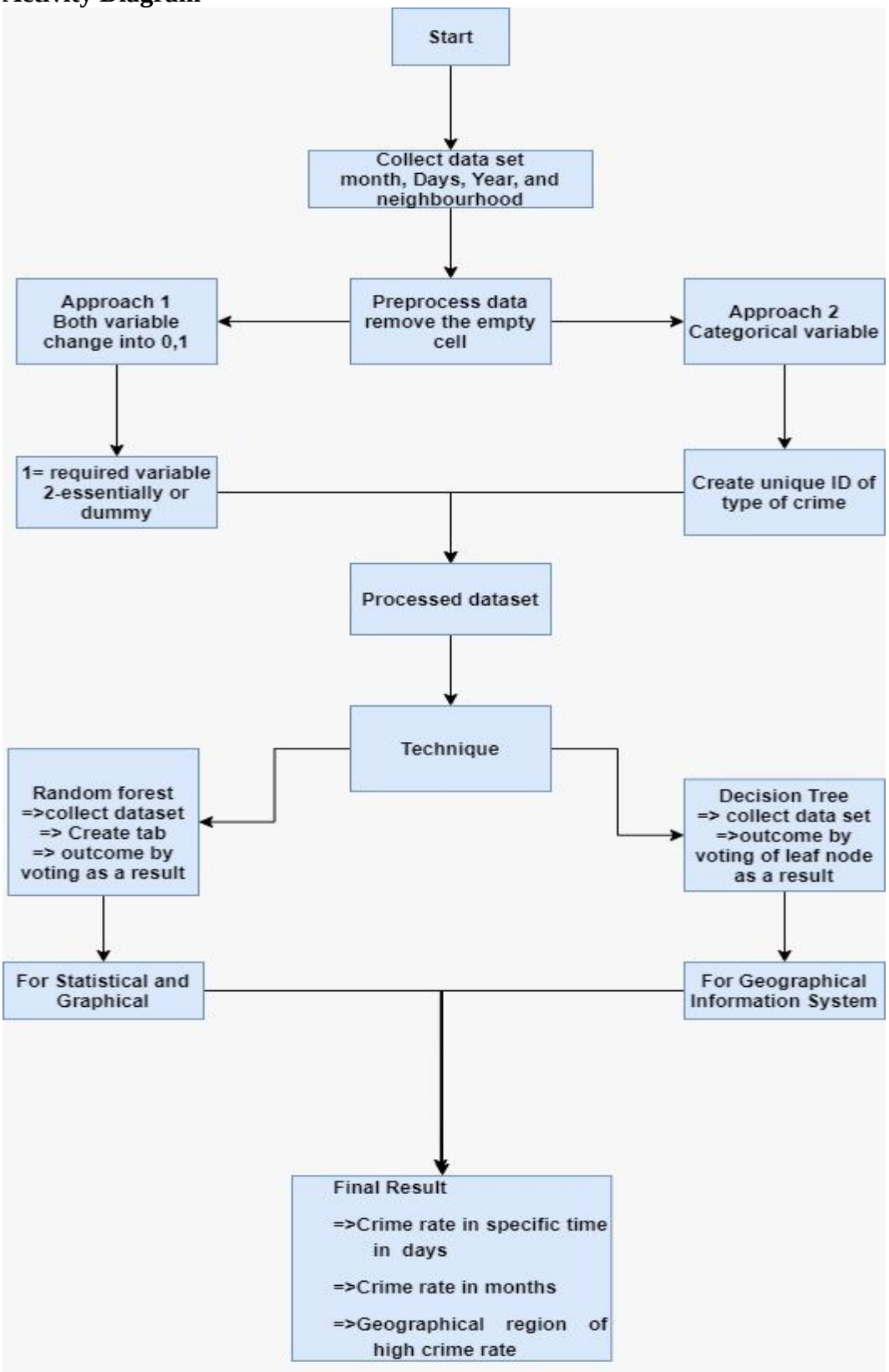
Figure 11: GIS

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Activity Diagram



Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Figure 12: Activity Diagram

RESULTS

All the types of crime variable have altered into numerical variables with unique IDs. Eleven types of crime have different IDs.

Crime id	Crime type
0	Break and Enter Commercial
1	Break and Enter Residential/Other
2	Homicide
3	Mischief
4	Offence Against a Person
5	Other Theft
6	Theft from Vehicle
7	Theft of Bicycle
8	Theft of Vehicle
9	Vehicle Collision or Pedestrian Struck (with Fatality)
10	Vehicle Collision or Pedestrian Struck (with Injury)

Our dataset contain several different types of crimes that happened during 2003 and 2017.

The diagram 4.1 shows the crime rate per month over the year. And we can look the 4546 high crime rate in May 2003.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about



Figure 13: Crime Rate Per Month Over Year

Trend of Crime

Fig.4.2, the trend shows that number of crimes is decreasing per month from 2003 to 2013 but increasing per month from 2003 to 2016, and again fell slightly to about 3000 crime per year in 2017.

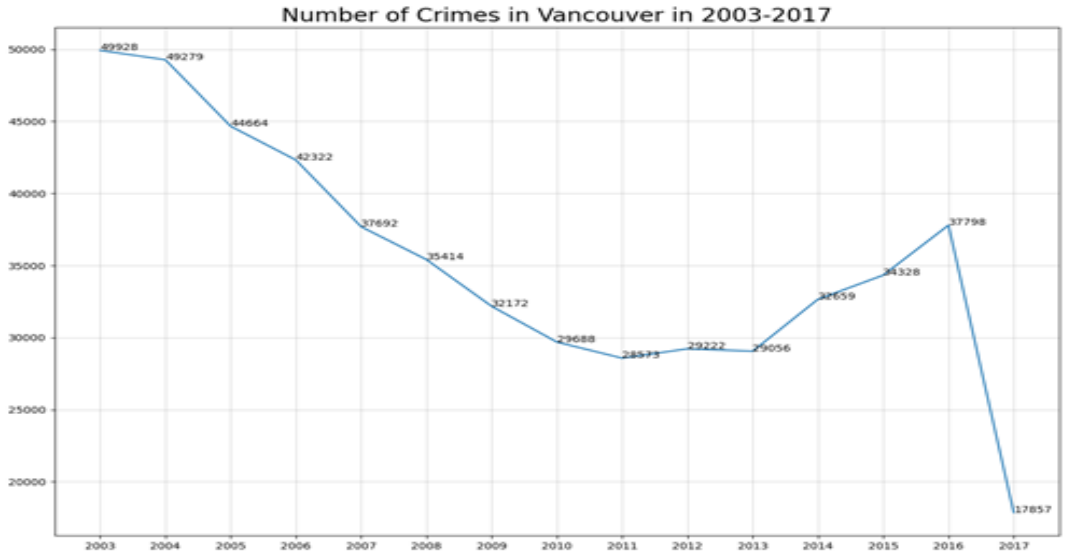


Figure 14: Trend of Crime Per Year

In Fig. 4.3, show the result on the base of per day and per month. We can look in the diagram the crime rate is high in starting day of the month. And when the last month of year the crime on peek point.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

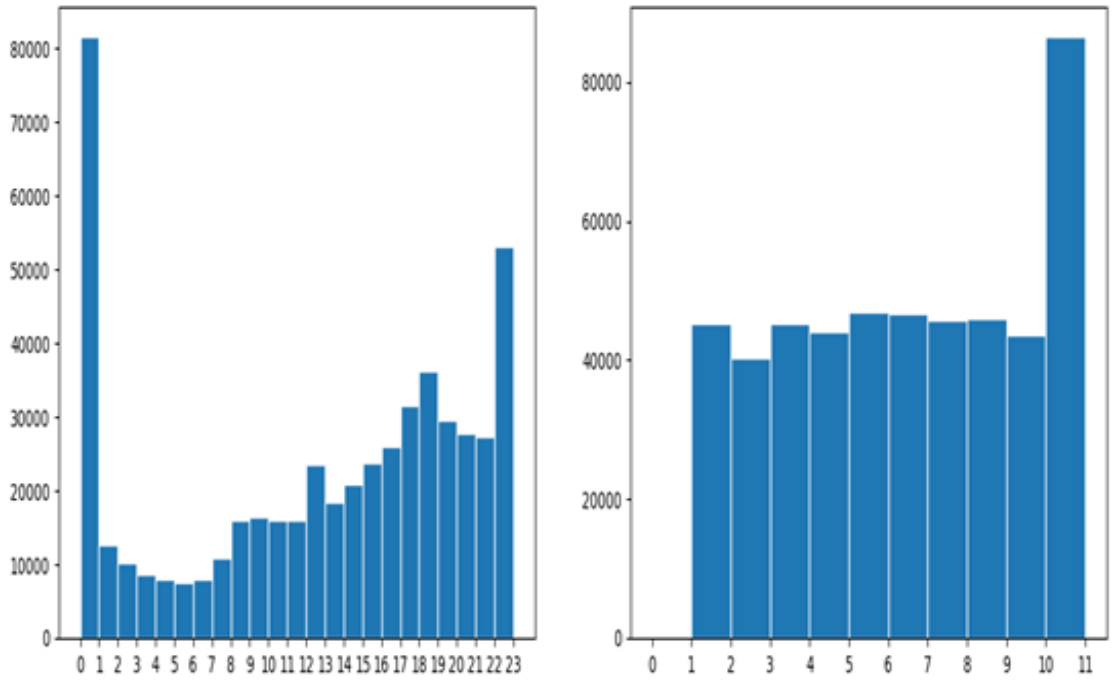


Figure 15: Crime Per Hour and Crime Per Month

Geographical Result

In following map figures shows the result of the crime. The high light areas in which crime have occurred

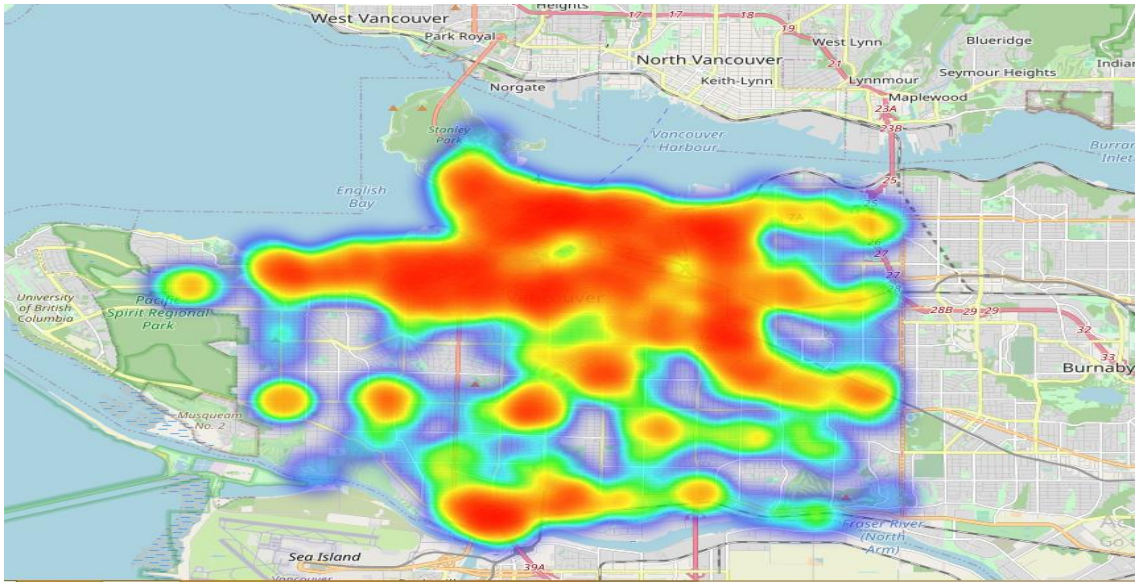


Figure 16: Break and Enter Commercial Crime

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

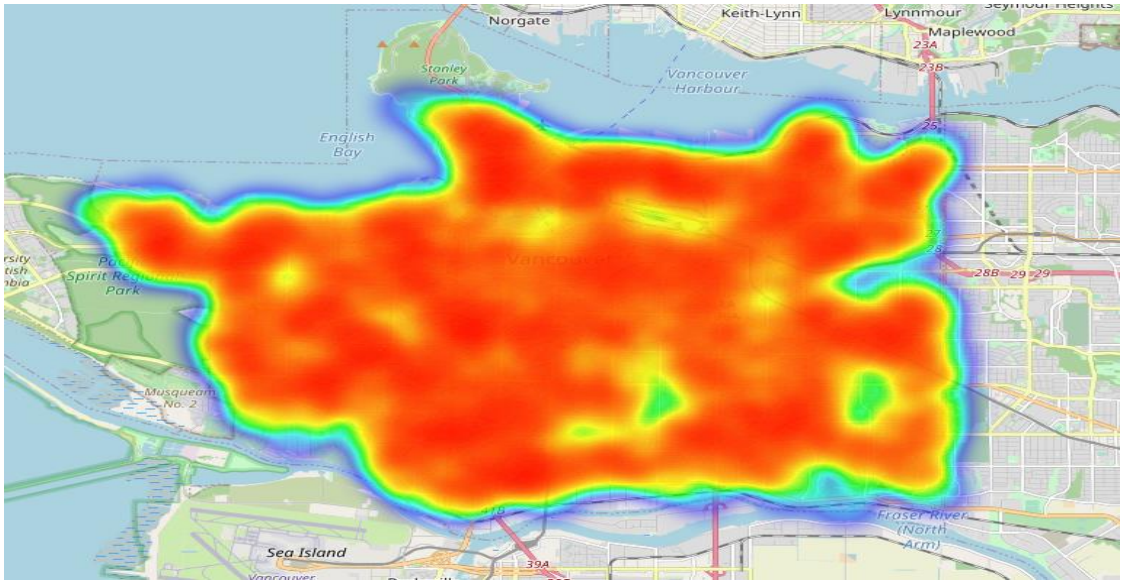


Figure 17: Break and Enter Residential/Other

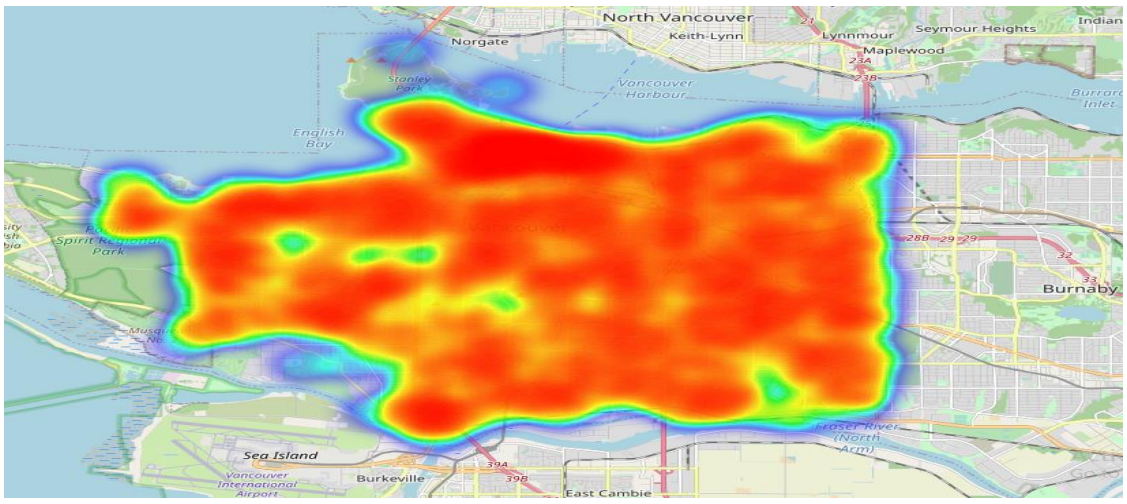


Figure 18: Mischief

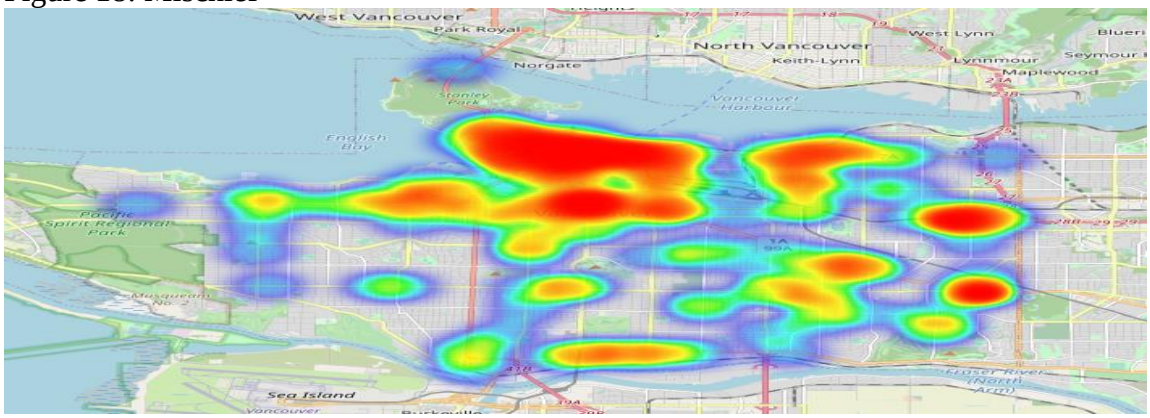


Figure 19: Other Theft

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

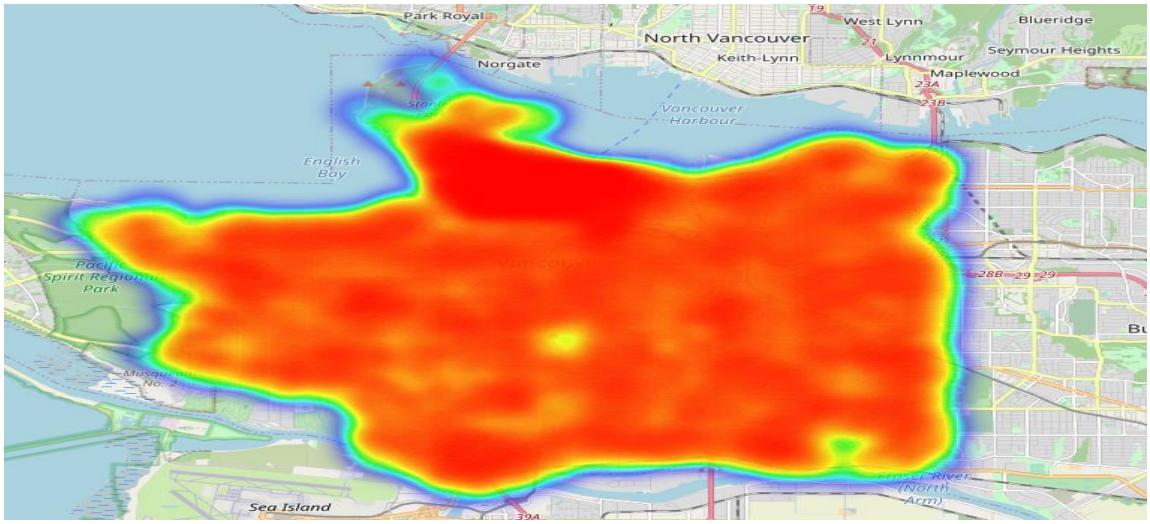


Figure 20: Theft from Vehicle

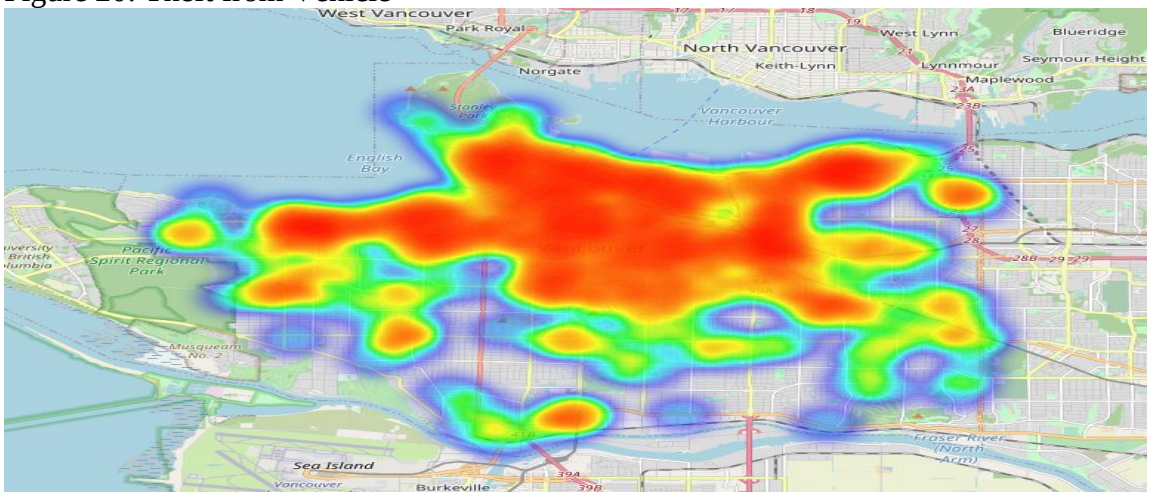


Figure 21: Theft of Bicycle

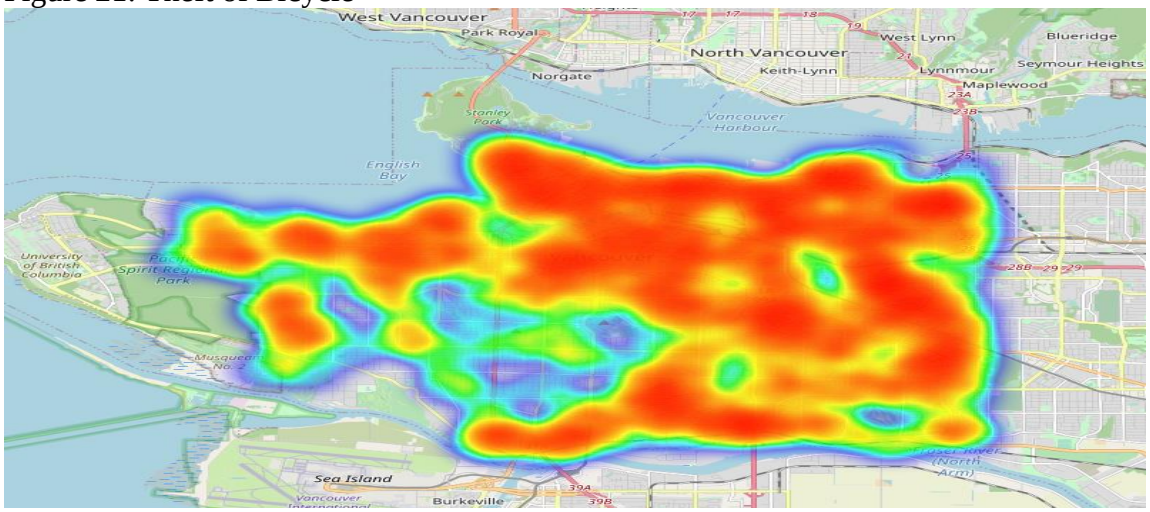


Figure 22: Theft of Vehicle

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

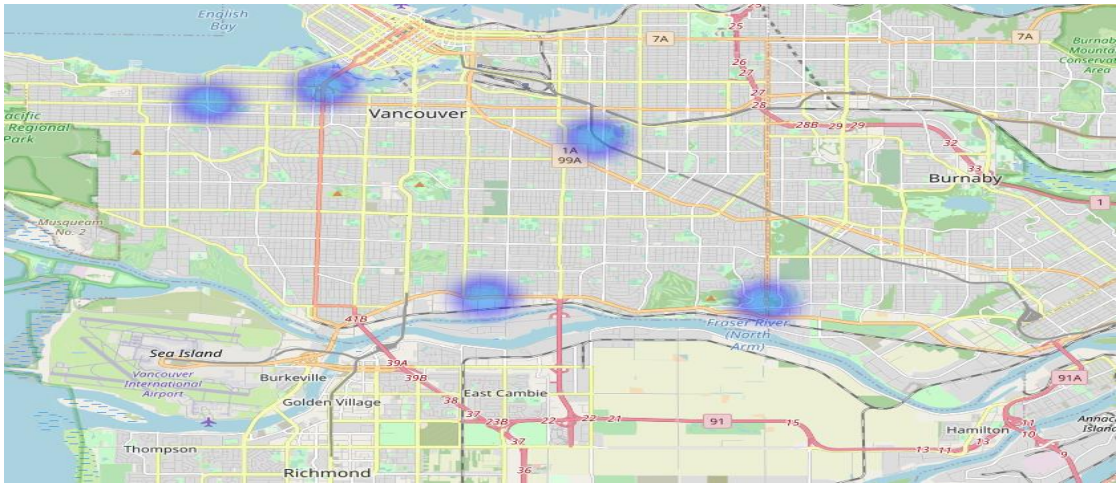


Figure 23: Vehicle Collision or Pedestrian Struck (with Fatality)

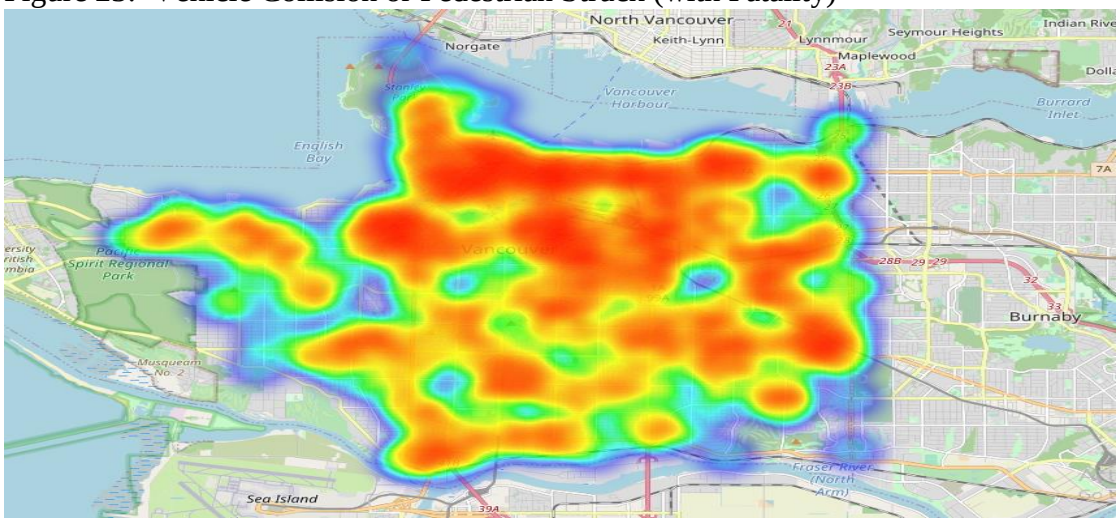


Figure 24: Vehicle Collision or Pedestrian Struck (with Injury)

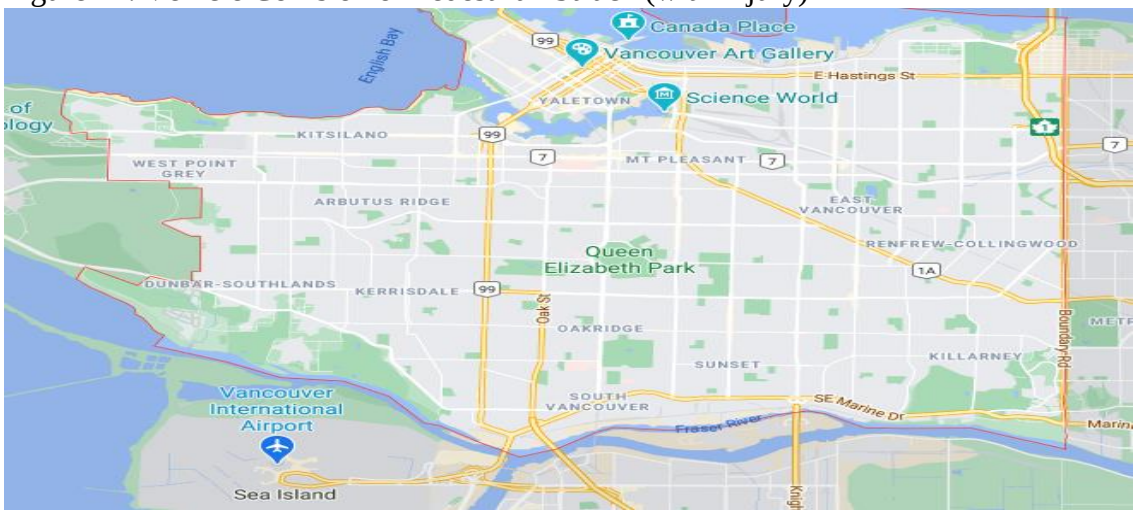


Figure 25: Homocide

The following diagram showing that the crime rate of theft from vehicle is very high as compare to other crime. And after this mischief is high crime rate. Similarly the

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

crime of other theft and theft of vehicle is same crime rate from each other. The crime rate of theft of bicycle is low from break and enter commercial. At the end Homicide crime rate is very low from all types of crime.

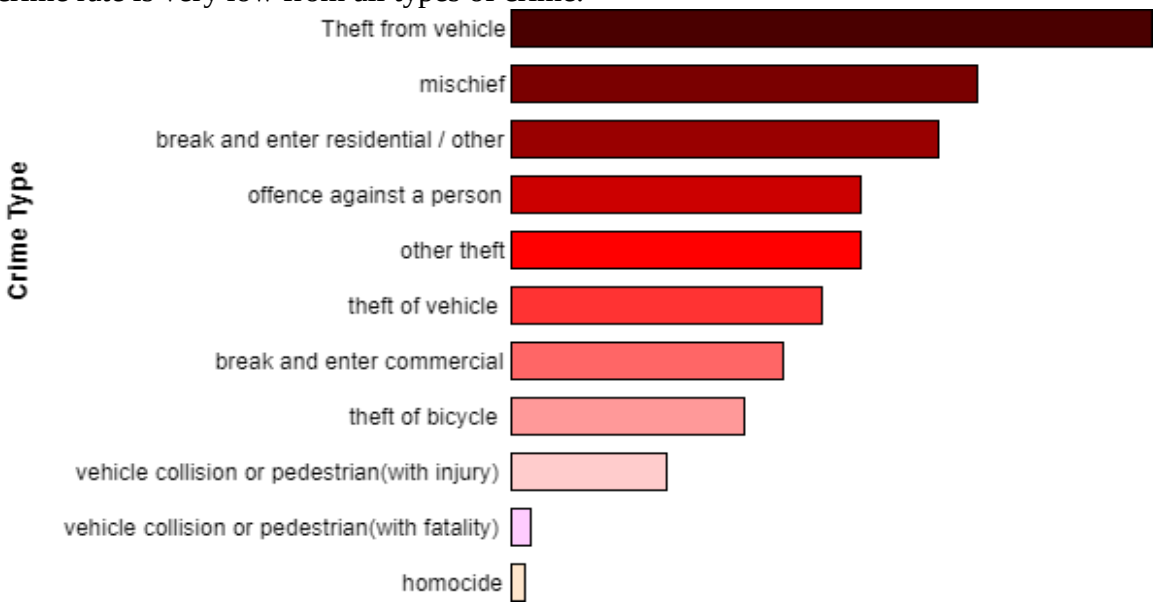


Figure 26: Overall Result of Eleven Types of Crime

Confusion Matrix

Given the following confusion matrix of random forest and decision tree algorithm, looking at the confusion matrix we can say that it is easy to predict crime number 4 and 6, while other crime such as 1, 3 and 5 has probability of being predicted.

Decission Tree

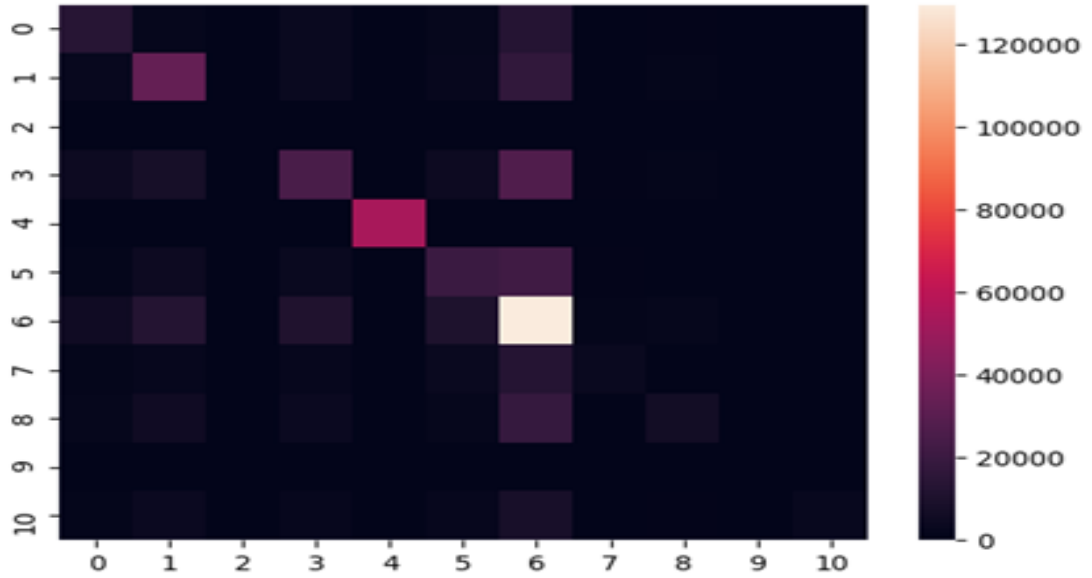


Figure 27: Confusion Matrix-Decission Tree

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Random Forest

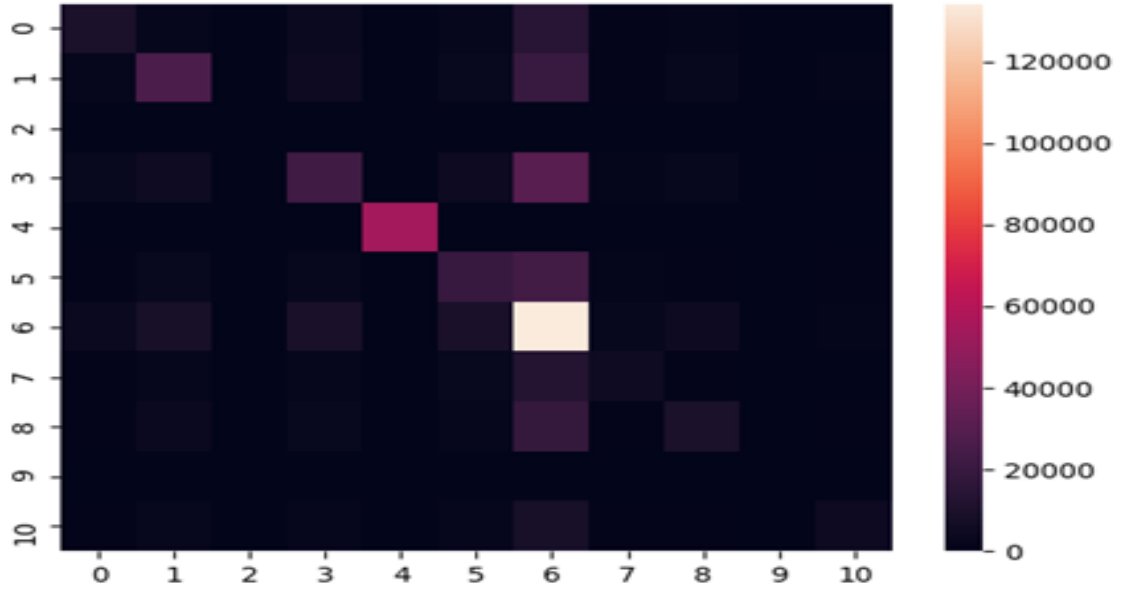


Figure 28: Confusion Matrix-Random Forest

Two types of model are using in ensemble technique random forest and decision tree. Decision tree model got accuracy of 54.47% and Random forest got accuracy of 53.96%. The last one diagram is the result of the ensemble model confusion matrix

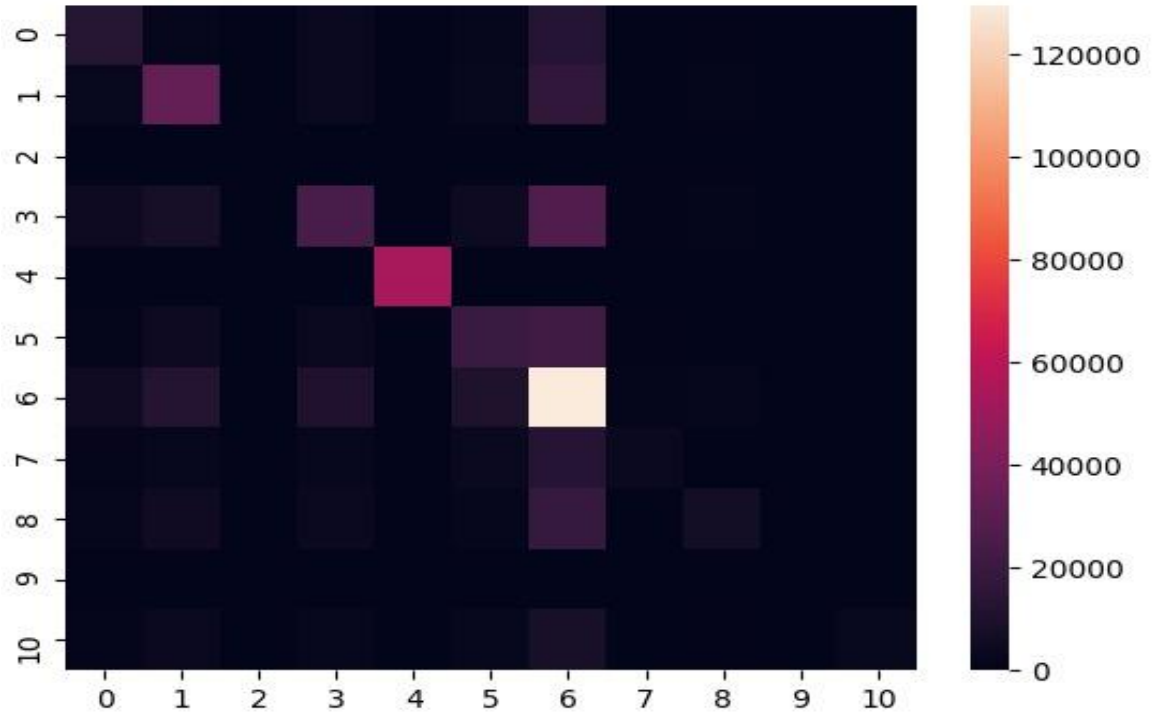


Figure 29: Confusion Matrix-Ensemble Model

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 11 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

SUMMERY

Crime rate is increasing with the passage of time because of multiple reasons. Due to the rapid increase in population and technological evolution the vastness of crimes is also a biggest issue so in this modern world the detailed examination and analysis of crimes has become very important and essential activity. But it is hard to forecast that which criminal would do the crime. Crimes of different types are increasing day by day so computerized techniques and systems are used to trace the crimes. The analysts of computer data are helping the officers of law enforcement to increasing and improving the procedure of solving crimes. Data mining paradigm methodology is very helpful today as it is additionally used at a faster rate to solve the crimes. In developed project we use Vancouver Crime dataset (VPD), we used year base, month base, and day base, and neighborhood as input to forecast the sort of crime committed. First we altered the type of crime values into numeric arrangement.

REFERENCES

T. M. Akinmuyisitan and J. Cosmas, "Enhanced Crime Prediction with Computer Vision-Yolov4 Approach," Global Journal of Computer Science and Technology, vol. 24(D1), pp. 55–65, 2024.

T. M. Akinmuyisitan, J. Cosmas and Y. A. Giwa, "Advancing Real-Time Crime Weapon Detection and High-Risk Person Classification in Pre-Crime Scenes: A Comprehensive Machine Vision Approach Utilizing SSD Detector," Global Journal of Computer Science and Technology, vol. 25, issue 1, pp. 1–12, 2025.

V. Mandalapu, L. Elluri, P. Vyas and N. Roy, "Enhancing Crime Prediction with Machine Vision- Convolutional Neural Network Approach," International Journal of Computer and Information Security, vol. 18, no. 4, pp. 211–225, 2025.

L. Elluri and V. Mandalapu, "Crime Prediction Using Machine Learning and Deep Learning," International Journal of Scientific Research in Science Engineering and Technology, vol. 11(3), pp. 08–15, 2024.

S. Ștefănoaia, M. Oprea and D. Stoica, "AI in crime prediction: current trends and challenges," International Journal of Security and Intelligent Systems, vol. 18, no. 2, pp. 101–118, 2025.

A. Rossi and L. Bianchi, "Contestable AI for criminal intelligence analysis: improving decision-making through semantic modeling and human oversight," Frontiers in Artificial Intelligence, vol. 9, article 1602998, 2025.

H. Abuzari, "Combating Delinquency in the Age of Artificial Intelligence: Prediction as Prevention," Journal of Criminal Law and Criminology, vol. 115, issue 2, pp. 450–475, 2024.

T. Situmeang, L. Napitupulu, T. Butarbutar and R. Manalu, "The Role of Artificial Intelligence in Criminal Justice," Journal of Law and Social Development and Applications, vol. 12, no. 3, pp. 210–225, 2024.

B. Mohammadinia and A. Alizadeh, "AI-Driven Risk Assessment in Predictive Policing Frameworks," International Journal of Computer and Information Security, vol. 15, no. 4, pp. 112–125, 2023.

A. O. Adegoke and B. T. Babatunde, "Enhanced Crime Prediction Model using

Online ISSN 3007-3197
Print ISSN 3007-3189

Bagging and Boosting Algorithms," Due Process International Journal of Applied Science, vol. 17, no. 1, pp. 40–55, 2024.

H. S. Ahmed and F. L. Wu, "An Intelligent Data Fusion Framework for Proactive Crime Forecasting," Knowledge-Based Systems, vol. 278, p. 110901, 2023.

S. Ehsanpour, "Ethical Considerations in Deploying Enhanced Predictive Policing Algorithms," European Journal of Criminology, vol. 22, issue 3, pp. 501–520, 2025.

X. Chen, "Improving the Trial Efficiency of Criminal Cases with the Assistance of Artificial Intelligence," Journal of Artificial Intelligence and Law Enforcement, vol. 5, no. 1, pp. 32–45, 2025.

J. S. Rodriguez, P. P. Alvarez and M. A. Gomez, "Predicting crime using Twitter and kernel density estimation: A Multimodal Deep Learning Approach," International Journal of Innovative Science and Research Technology, vol. 9, issue 1, p. 1021, 2024.

M. Mahmoudi and M. B. Kazemi, "Deep learning models for anomaly detection in surveillance data: A security-oriented approach," Journal of Criminal Policy and Management, vol. 10, no. 4, pp. 1–18, 2024.

V. Mandalapu, L. Elluri, P. Vyas and N. Roy, "Spatio-Temporal Deep Learning for Crime Hotspot Identification," IEEE Transactions on Intelligent Transportation Systems, vol. 24, issue 11, pp. 11001–11015, 2023.

A. S. Hassan and R. R. Al-Jubouri, "Developing an Enhanced Intelligent Framework for Crime Clustering and Visualization," International Journal of Forensic Science and Technology, vol. 8, no. 3, pp. 150–165, 2023.

S. G. K. Reddy and P. S. Rao, "Adaptive Ensemble Learning for Fine-Grained Crime Classification in Smart Cities," Expert Systems with Applications, vol. 245, p. 122905, 2024.

M. Z. A. Khan and I. U. S. Rahman, "Hybrid Deep Neural Networks for Time-Series Crime Forecasting and Pattern Analysis," Neurocomputing, vol. 601, p. 120510, 2025.

W. Li and J. Han, "Optimizing Resource Allocation via Predictive Policing Frameworks using Reinforced Deep Learning," ACM Transactions on Intelligent Systems and Technology, vol. 16, issue 2, p. 18, 2025.