A MINI PROJECT REPORT

# CRIME RATE PREDICTION AND ANALYSIS USING MACHINE LEARNING

***Submitted by***


**S.ABHINAYA - 21TK1A6626**


***In partial fulfillment for the award of the degree of***

**BACHELORS OF TECHNOLOGY**

**IN**

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

***Under the esteemed guidance of***


**Dr. G. Ranjith**

**Asst. Professor, Department of AI&ML**



**SVS GROUP OF INSTITUTIONS-SCHOOL OF ENGINEERING**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING**

**Bheemaram, Warangal 506015, TELANGANA**


**2024-2025**

SVS GROUP OF INSTITUTIONS

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

# CERTIFICATE

This is to certify that **S.ABHINAYA -21TK1A6628** of the B tech has satisfactorily completed the dissertation work entitled **“Crime Rate Prediction And Analysis Using Machine Learning”** in the partial fulfillment of the requirements of the B. Tech degree during this academic year 2024-2025.

| **INTERNAL** | **HOD(AIML)** | **PRINCIPAL** |
| --- | --- | --- |
| Dr. G. RANJITH | Dr. G. RANJITH | Dr. R RAMESH |
| Asst. Professor | Asst. Professor | Dept of (AI&ML) |
| **Internal Guide** | **Head of the Department** | SVSGOI |
| Dept of (AI&ML) | Dept of (A&ML) | BHEEMARAM |

**DEPARTMENT OF COMPUTER SCIENCE&ENGINEERING**
**SVS INSTITUTE OF TECHNOLOGY**
**BHEEMARAM ,HANAMKONDA- 506015**

ISO 9001:2008 CERTIFIED

## DECLARATION BY THE CANDIDATE

I bearing the **S.ABHINAYA H.T No 21TK1A6628**,to hereby certify that the Project Report entitled **"CRIME RATE PREDICTION"**, under the guidance of **Dr. G.RANJITH.** Department of CSE(AI&ML), SVS Institute of Technology & Science, Hanamkonda, submitted in partial fulfillment of the requirements for the Award of the Degree of Master of Technology in Artificial Intelligence & Machine Learning.

This is a record of Bonafide work carried out by me and the results embodied in this Project Report has not been produced / copied from any source. The results embodied in this Project Report have not been submitted to any other University or Institute for the Award of any other Degree or Diploma.

PLACE: Hanamkonda

DATE: 30-01-25

**S.ABHINAYA - 21TK1A6628**

**DEPARTMENT OF COMPUTER SCIENCE&ENGINEERING**
**SVS INSTITUTE OF TECHNOLOGY**
**BHEEMARAM ,HANAMKONDA- 506015**

ISO 9001:2008 CERTIFIED

## DECLARATION BY THE CANDIDATE

This is to certify that the Project Report entitled **"CRIME RATE PREDICTION"** being submitted by ,**S.ABHINAYA** H.T No **21TK1A6628**,in partial fulfillment of the requirements for the Award of the Degree of the Master of Technology in Artificial Intelligence & Machine Learning.is a record of Bonafide work carried out by him under my guidance.

The results of investigation enclosed in the report have been verified and found satisfactory. The results embodied in this thesis have not been submitted to any other University for the award of any degree or diploma.

Place:Hanamnonda

Date:30-01-25

**Dr.G.RANJITH**
Asst Professor &HOD

## ACKNOWLEDGMENT

The completion of this project work gives me an opportunity to convey my gratitude to all those who helped me to complete the project successfully.

First, I gratefully acknowledge my deep sense of gratitude to Almighty for spiritual Guidance blessings shown to complete the Project. I thank my parents for unconditional support to improve myself throughout my life.

My sincere thanks to Dr. E. THIRUMAL RAO, Chairman, SVS Group of Institutions, for providing this opportunity to carry out the Project in the institution.

I owe my respectable thanks to Dr .R RAMESH, Principal of SVS Group of Institutions, for providing all necessary facilities and encouraging words for completion of this project.

My sincere thanks to Mr. RATHOD, Vice-Principal , SVS Group of Institutions, for encouraging and supporting for completion of this project.

I gratefully acknowledge Dr .G.RANJITH, Associate Professor, Head of the Department, Computer Science and Engineering, for his encouragement and advice during the this project work,

My sincere thanks to MR. MD AMJAD PASHA Assistant Professor, Guide SVS Group of Institutions, for continuous support for doing this project. I would like to express my thanks to all the faculty members of Department of Computer Science and Engineering, and Non-teaching staff of SVS Group of Institutions who have rendered valuable help in making this project successful.

**S.ABHINAYA - 21TK1A6628**

TABLE OF CONTENT

# ABSTRACT

Day by day crime data rate is increasing because the modern technologies and hi-tech methods are helps the criminals to achieving the illegal activities According to Crime Record Bureau crimes like burglary, arson have been increased while crimes like murder, sex, abuse, gang rap have been increased Crime data will be collected from various blogs, news and websites. The huge data is used as a record for creating a crime report database. The knowledge which is acquired from the data mining techniques will help in reducing crimes as it helps in finding the culprits faster and also the areas that are most affected by crime day by day crime data rate is increasing because the modern technologies and hi-tech methods are helps the criminals to achieving the illegal activities .According to Crime Record Bureau crimes like burglary, arson have been increased while crimes like murder, sex, abuse, gang rap have been increased .

# CHAPTER 1

# INTRODUCTION 1

## .1Purpose:

The purpose of crime rate prediction and analysis using machine learning is to leverage advanced Computational techniques to identify patterns, trends, and potential future occurrences of criminal activity. This approach enables law enforcement agencies to allocate resources more efficiently, enhance public safety, and develop proactive strategies to prevent crime. By analysing historical data and realtime inputs, machine learning models can provide valuable insights into high-risk areas, optimal patrol routes, and potential crime hotspots, thereby improving the effectiveness of crime prevention and intervention efforts.

## 1.2Background of project:

crime rate prediction and analysis using machine learning stems from the growing need for innovative solutions to tackle the persistent issue of crime in urban environments. Traditional crime prevention methods, which often rely on reactive measures and historical crime data, are insufficient in addressing the dynamic and complex nature of criminal activities. With the advent of big data and machine learning technologies, there is an opportunity to harness vast amounts of data from various sources, such as police reports, social media, weather conditions, and demographic information, to build predictive models that can provide more accurate and timely insights into crime trends. Machine learning algorithms excel at detecting patterns and making predictions based on large datasets, which are often beyond the capability of manual analysis. By implementing these technologies, the project aims to transform raw data into actionable intelligence that can predict future crime hotspots, identify potential perpetrators, and optimize resource allocation for law enforcement agencies. This proactive approach not only enhances the efficiency of crime prevention strategies but also contributes to the overall safety and security of communities by enabling pre-emptive measures to thwart criminal activities before they occur

## 1.3 Scope of project:

Crime rate prediction and analysis using machine learning holds significant potential for improving public safety and resource allocation. By leveraging historical crime data and advanced algorithms, machine learning models can identify patterns and trends, enabling law enforcement agencies to predict future crime hotspots and times with greater accuracy. This proactive approach allows for optimized deployment of police resources, timely interventions, and better-informed policy decisions. Moreover, integrating socio-economic factors and real-time data can enhance the precision and effectiveness of these predictive models.

## 1.4 Project Features:

1. **Spatial Features:** Use geographic information to create features like distance to police stations, crime hot spots, and neighbourhood characteristics

2. **Model Selection:** Implement and compare various machine learning algorithms.

3. **Data Collection:** Gather historical crime data, socioeconomic data, weather data, and other relevant information.

4. **Prediction and Visualization:** Generate predictions, create risk maps and dashboards, and provide actionable insight

# CHAPTER 2

# SYSTEMREQUIREMENTS

## 2.1 Hardware Requirements:

- HDD 500 GB
- RAM 4 GB
- Processor intel i5
- Keyboard
- Mouse

## 2.2 Software requirements:

**Frontend**:

- JUPYTER

**Backend**:

- Python

**Operating System:**

- Windows 10

## 2.3 Existing System:

1. Data Mining Clustering
2. Naïve Baye
3. Spatial Analysis
4. Logistic Regression

### 2.3.1Drawbacks of existing system:

1. Data Quality and Bias: The accuracy of predictions heavily depends on the quality and completeness of historical crime data. Inconsistent, incomplete, or biased data can lead to flawed predictions and reinforce existing biases in policing.

### 2.Privacy Concerns:

These systems often require access to vast amounts of personal and sensitive data, raising concerns about privacy and the potential for misuse of information.

### 3.Over-policing:

Predictive models might lead to over-policing in certain areas, particularly those with higher recorded crime rates, potentially exacerbating community tensions and unfairly targeting specific demographics.

### 4.Ethical Concerns:

There are significant ethical concerns regarding the potential for machine learning systems to perpetuate systemic biases and inequalities present in the historical data used to train them.

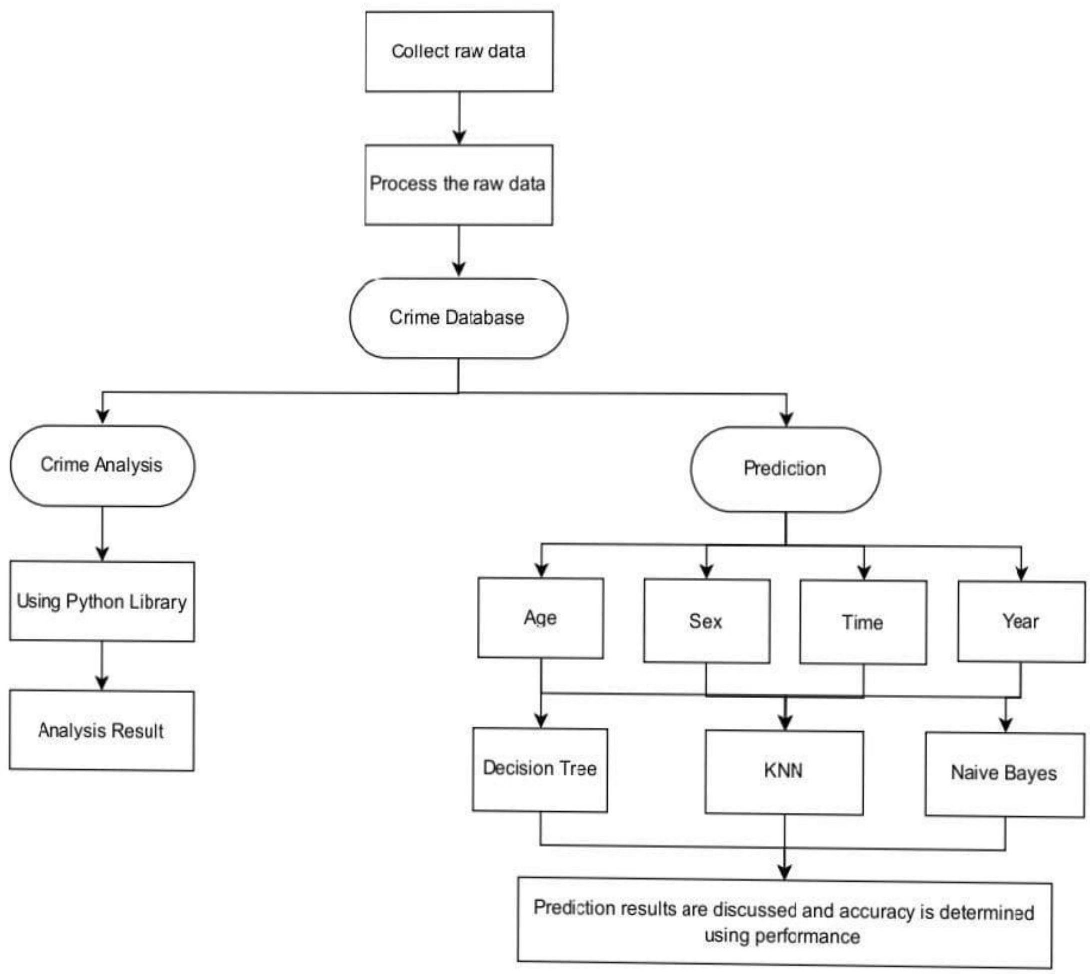
## 2.4 Proposed System:

It consists of crime information like location description, type of crime, date, time, latitude, longitude. Before training the model data preprocessing will be done following this feature Neighbour selection and scaling will be done so that the accuracy obtained will be high. The K-Nearest (KNN) classification and various other algorithms (Decision Tree and Random Forest) will be tested for crime and propose one with better query-based use for training. Visualization of the dataset will be done in terms of graphical representation of many cases, for example at which time the crime rates are high or at which month the criminal activities are high . The sole purpose of this project is to give a just idea of how machine learning can be used by law enforcement agencies to detect, predict and involve crimes at a much faster rate and thus reduce the crime rate. This can be used in other states or countries depending upon the availability of the data set. We will be using the technique of machine learning and data science for crime prediction of crime data sets. The crime data is extracted from the official portal of police

# CHAPTER 3

# SYSTEM DESIGN

## 3. 1 System Architecture



### 3.2 UML Diagrams

#### 3.2.1 Use case diagram

Use Case during requirement elicitation and analysis to represent the functionality of the system. Use case describes a function by the system that yields a visible result for an actor. The identification of actors and use cases result in the definitions of the boundary of the system i.e., differentiating the tasks

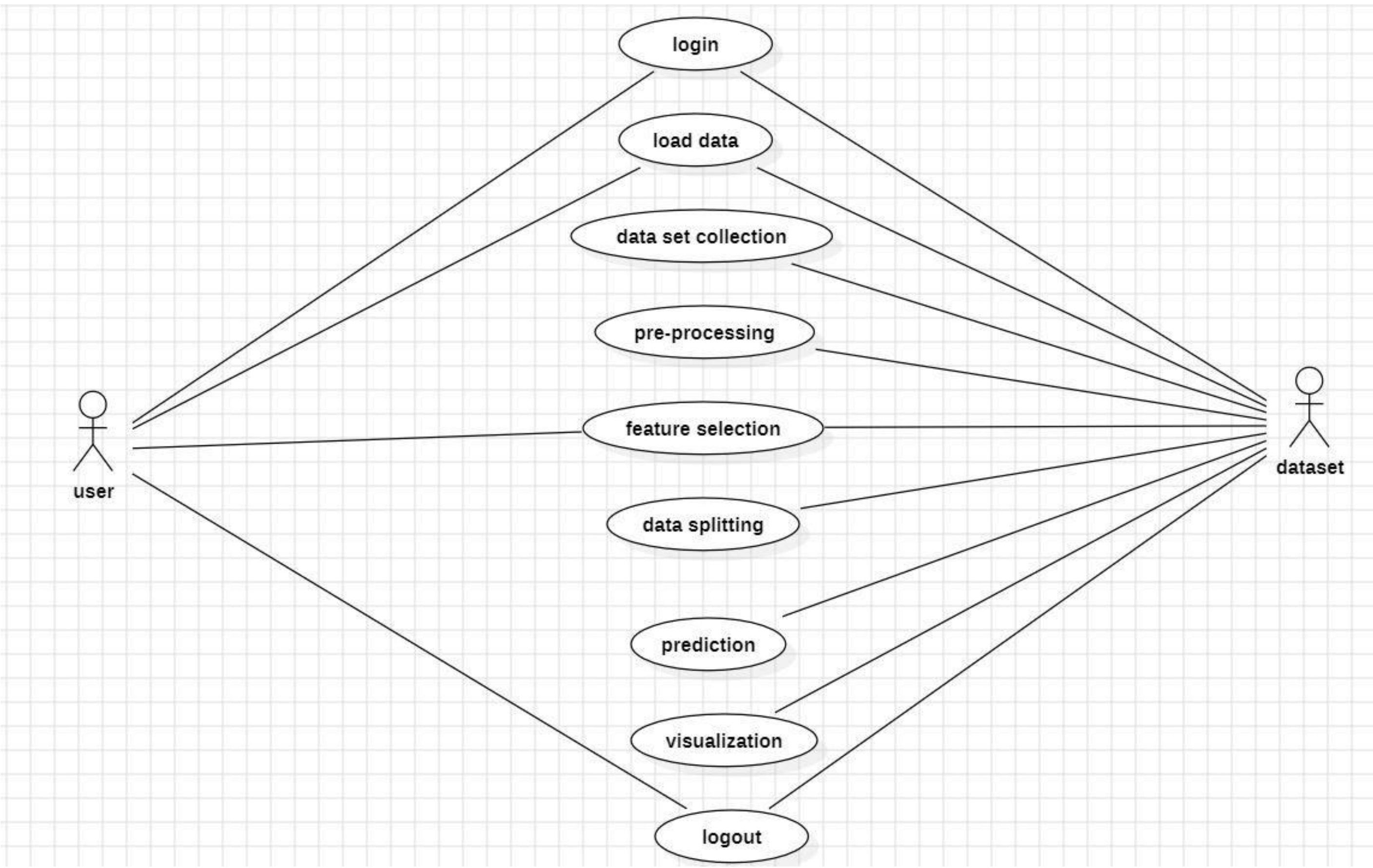



**Fig 3.2.1 Use Case Diagram of CRPAML**

### 3.2.2 Class Diagram

Class diagrams model class structure and contents using design elements such as classes, packages and objects. Classes are composed of three things: name, attributes, and operations. Class diagram also display relationships such as containment, inheritance, association etc. The association relationship is most common relationship in a class diagram. The association shows the relationship between instances o

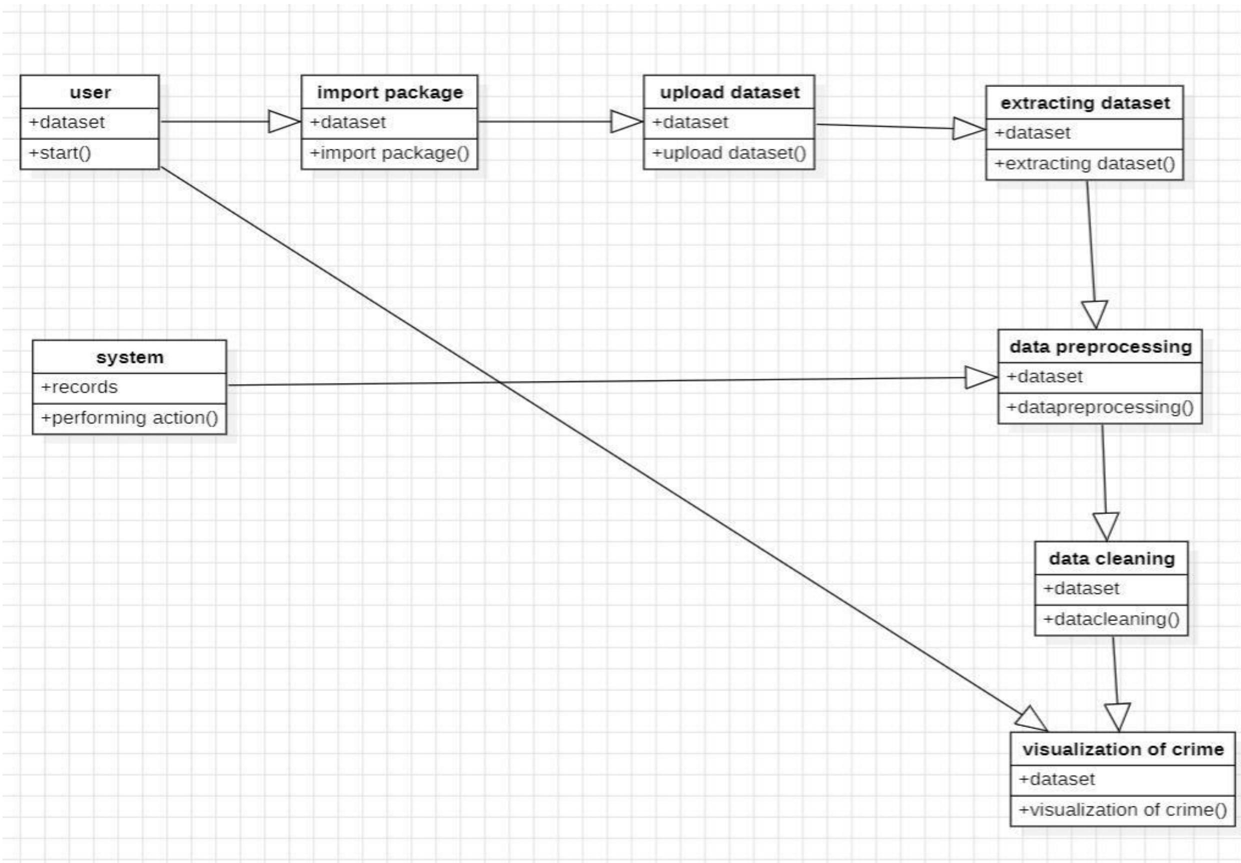



**Fig 3.2.2 Class Diagram of CRPAML**

### 3.2.3 Sequence Diagram

Sequence diagram displays the time sequence of the objects participating in the interaction. This consists of the vertical dimension (time) and horizontal dimension (different objects). Objects: An object can be thought of as an entity that exists at a specified time and has a definite value, as well as a holder of identity. A sequence diagram depicts item interactions in chronological order. It illustrates the scenario's objects and classes, as well as the sequence of messages sent between them in order to carry out the scenario's functionality. In the Logical View of the system under development, sequence diagrams are often related with use case realizations. Event diagrams and event scenarios are other names for sequence diagrams. A sequence diagram depicts multiple processes or things that exist simultaneously as parallel vertical lines (lifelines), and the messages passed between them as horizontal arrows, in the order in which they occur. This enables for the graphical specification of simple runtime scenarios

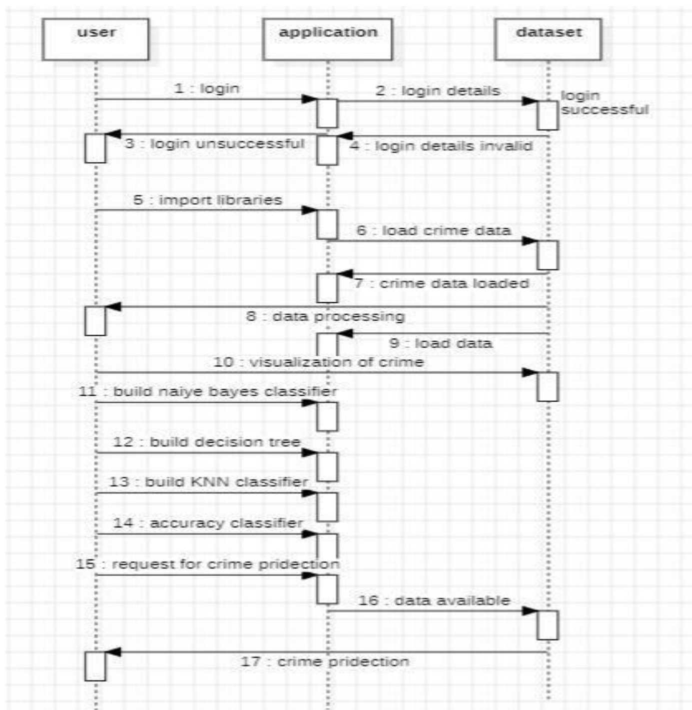



**Fig 3.2.2 Sequence Diagram of CRP**

### 3.2.4 Activity Diagram

The process flows in the system are captured in the activity diagram. Similar to a state diagram, an activity diagram also consists of activities, actions, transitions, initial and final states ,and guard conditions

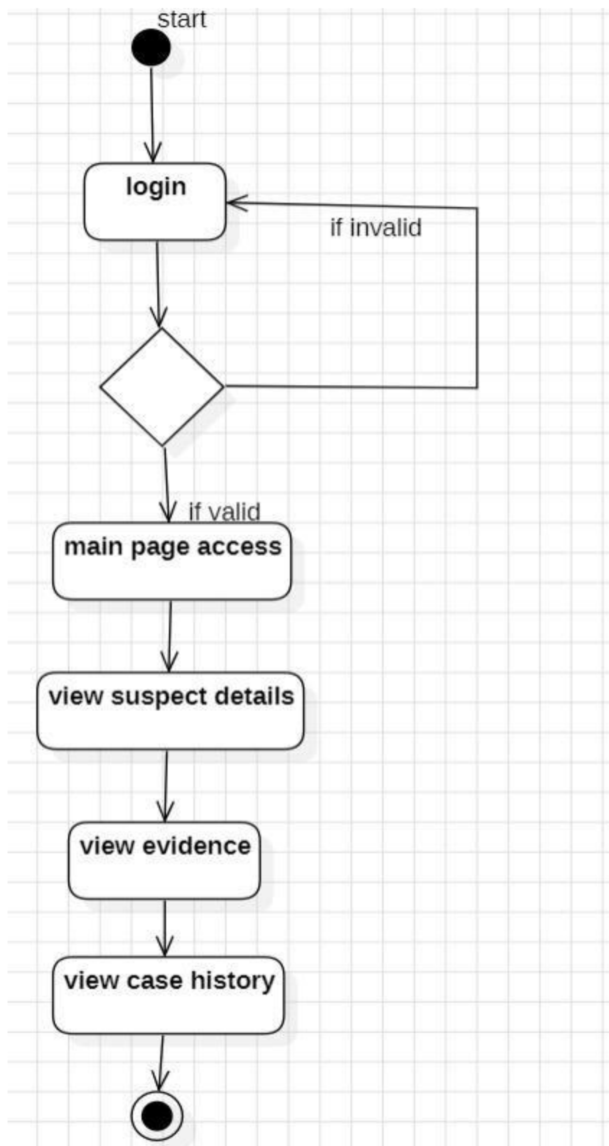



**Fig 3.2.3 Activity Diagram of CRPAML**

# CHAPTER 4

# IMPLEMENTATION

## 4.1: Code

```
{
 "cells": [
  {
   "cell_type": "code",
   "execution_count": 1,
   "metadata": {},
   "outputs": [],
   "source": [
    "import numpy as np\n",
    "import timeit\n",
    "import pandas as pd\n",
    "from sklearn.preprocessing import StandardScaler\n",
    "from sklearn.cluster import KMeans\n",
    "import matplotlib.pyplot as plt\n",
    "import matplotlib\n",
    "from sklearn.decomposition import PCA\n",
    "from sklearn.metrics import silhouette_samples, silhouette_score\n",
    "import matplotlib.cm as cm\n",
    "from sklearn.preprocessing import StandardScaler\n",
    "from sklearn.model_selection import train_test_split\n",
    "from sklearn.preprocessing import OneHotEncoder\n",
    "from sklearn.ensemble import RandomForestClassifier\n",
    "from sklearn.ensemble import GradientBoostingClassifier\n",
    "from sklearn.metrics import confusion_matrix, classification_report, accuracy_score"
   ]
  },
  {
   "cell_type": "code",
   "execution_count": 2,
   "metadata": {},
   "outputs": [
    {
     "data": {
      "text/html": [
       "<div>\n",
       "<style scoped>\n",
       "    .dataframe tbody tr th:only-of-type {\n",
       "        vertical-align: middle;\n",
```

```
"    }\n",
"\n",
"    .dataframe tbody tr th {\n",
"        vertical-align: top;\n",
"    }\n",
"\n",
"    .dataframe thead th {\n",
"        text-align: right;\n",
"    }\n",
"</style>\n",
"<table border=\"1\" class=\"dataframe\">\n",
"  <thead>\n",
"    <tr style=\"text-align: right;\">\n",
"      <th></th>\n",
"      <th>X</th>\n",
"      <th>Y</th>\n",
"      <th>Index_</th>\n",
"      <th>event_unique_id</th>\n",
"      <th>occurrencedate</th>\n",
"      <th>reporteddate</th>\n",
"      <th>premisetype</th>\n",
"      <th>ucr_code</th>\n",
"      <th>ucr_ext</th>\n",
"      <th>offence</th>\n",
"      <th>...</th>\n",
"      <th>occurrencedayofweek</th>\n",
"      <th>occurrencehour</th>\n",
"      <th>MCI</th>\n",
"      <th>Division</th>\n",
"      <th>Hood_ID</th>\n",
"      <th>Neighbourhood</th>\n",
"      <th>Long</th>\n",
"      <th>Lat</th>\n",
"      <th>ObjectId</th>\n",
"      <th>Total</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0</th>\n",
"      <td>-79.405228</td>\n",
"      <td>43.656982</td>\n",
"      <td>7801</td>\n",
"      <td>GO-20152165447</td>\n",
"      <td>2015-12-18T03:58:00.000Z</td>\n",
"      <td>2015-12-18T03:59:00.000Z</td>\n",
"      <td>Commercial</td>\n",
```

```
"     <td>1430</td>\n",
"     <td>100</td>\n",
"text/plain": [
"         X          Y  Index   event_unique_id           occurrencedate  \\\n",
"0 -79.405228  43.656982   7801  GO-20152165447  2015-12-18T03:58:00.000Z   \n",
"1 -79.307907  43.778732   7802  GO-20151417245  2015-08-15T21:45:00.000Z   \n",
"2 -79.225029  43.765942   7803  GO-20151421107  2015-08-16T16:00:00.000Z   \n",
"3 -79.140823  43.778648   7804  GO-20152167714  2015-11-26T13:00:00.000Z   \n",
"4 -79.288361  43.691235   7805  GO-20152169954  2015-12-18T19:50:00.000Z   \n",
"\n",
"               reporteddate premisetype  ucr_code  ucr_ext  offence  ...  \\\n",
"0  2015-12-18T03:59:00.000Z  Commercial      1430      100  Assault  ...   \n",
"1  2015-08-17T22:11:00.000Z  Commercial      1430      100  Assault  ...   \n",
"2  2015-08-18T14:40:00.000Z   Apartment      2120      200      B&E  ...   \n",
"3  2015-12-18T13:38:00.000Z       Other      2120      200      B&E  ...   \n",
"4  2015-12-18T19:55:00.000Z  Commercial      1430      100  Assault  ...   \n",
"\n",
"  occurrencedayofweek occurrencehour              MCI  Division Hood_ID  \\\n",
"0          Friday                  3          Assault       D14      79   \n",
"1          Saturday               21          Assault       D42     118   \n",
"2          Sunday                 16  Break and Enter       D43     137   \n",
"3          Thursday               13  Break and Enter       D43     133   \n",
"4          Friday                 19          Assault       D55      61   \n",
"\n",
"                   Neighbourhood       Long        Lat  ObjectId  Total  \n",
"0               University (79) -79.405228  43.656982      7001      1  \n",
"1  Tam O'Shanter-Sullivan (118) -79.307907  43.778732      7002      1  \n",
"2                   Woburn (137) -79.225029  43.765942      7003      1  \n",
"3  Centennial Scarborough (133) -79.140823  43.778648      7004      1  \n",
"4             Taylor-Massey (61) -79.288361  43.691235      7005      1  \n",
"\n",
"[5 rows x 30 columns]"
]
},
"execution_count": 2,
"metadata": {},
"output_type": "execute_result"
}
],
"source": [
"df=pd.read_csv(\"C:/Users/KIIT/Downloads/MCI_2014_to_2019.csv\")\n",
"df['Total'] = 1\n",
"df.head()"
]
},
{
"cell_type": "code",
```

```
  "execution_count": 3,
  "metadata": {},
n",
     "    <tr>\n",
     "      <th>206423</th>\n",
     "      <td>-79.564438</td>\n",
     "      <td>43.670914</td>\n",
     "      <td>205889</td>\n",
     "      <td>GO-2019333586</td>\n",
     "      <td>2019-02-21T21:30:00.000Z</td>\n",
     "      <td>2019-02-22T06:00:00.000Z</td>\n",
     "      <td>House</td>\n",
     "      <td>2135</td>\n",
     "      <td>210</td>\n",
     "      <td>Theft Of Motor Vehicle</td>\n",
     "      <td>...</td>\n",
     "      <td>Thursday</td>\n",
     "      <td>21</td>\n",
     "      <td>Auto Theft</td>\n",
     "      <td>D22</td>\n",
     "      <td>11</td>\n",
     "      <td>Eringate-Centennial-West Deane (11)</td>\n",
     "      <td>-79.564438</td>\n",
     "      <td>43.670914</td>\n",
     "      <td>205989</td>\n",
     "      <td>1</td>\n",
     "    </tr>\n",
     "    <tr>\n",
     "      <th>206424</th>\n",
     "      <td>-79.365318</td>\n",
     "      <td>43.745293</td>\n",
     "      <td>205890</td>\n",
     "      <td>GO-2019335865</td>\n",
     "      <td>2019-02-21T20:00:00.000Z</td>\n",
     "      <td>2019-02-22T13:04:00.000Z</td>\n",
     "      <td>House</td>\n",
     "      <td>2135</td>\n",
     "      <td>210</td>\n",
     "      <td>Theft Of Motor Vehicle</td>\n",
     "      <td>...</td>\n",
     "      <td>Thursday</td>\n",
     "      <td>20</td>\n",
     "      <td>Auto Theft</td>\n",
     "      <td>D33</td>\n",
     "      <td>42</td>\n",
     "      <td>Banbury-Don Mills (42)</td>\n",
     "      <td>-79.365318</td>\n",
```

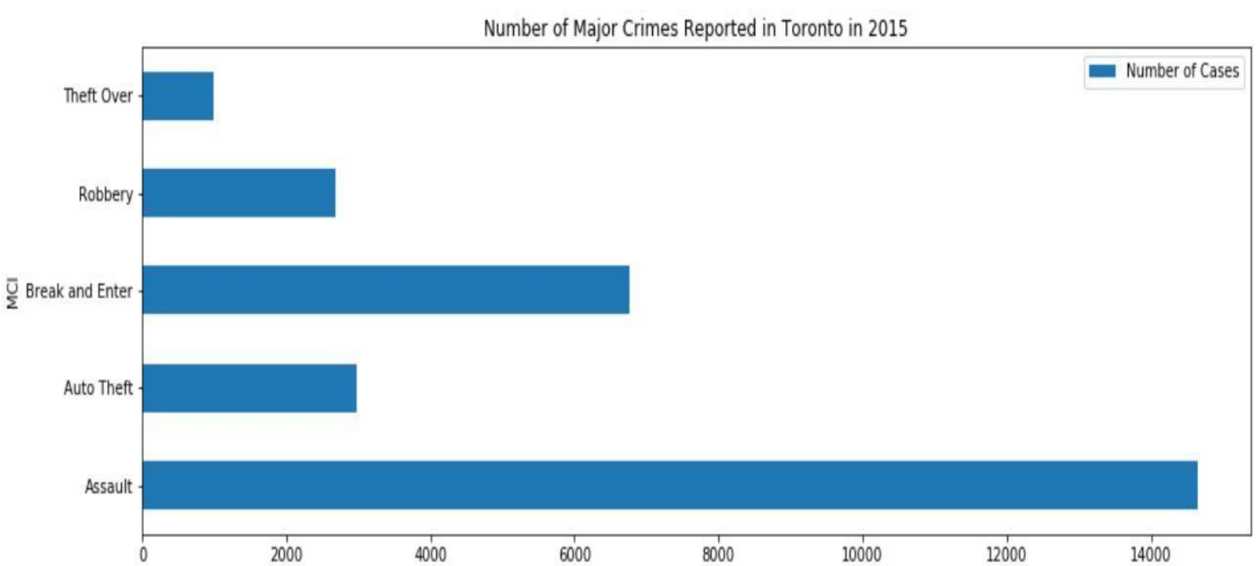
```
"    <td>43.745293</td>\n",
"    <td>205990</td>\n",
"    <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"    <th>206425</th>\n",
"    <td>-79.464493</td>\n",
"    <td>43.668125</td>\n",
"    <td>205891</td>\n",
"    <td>GO-2019332516</td>\n",
"    <td>2019-02-21T20:30:00.000Z</td>\n",
"    <td>2019-02-21T22:42:00.000Z</td>\n",
"    <td>Outside</td>\n",
"    <td>2135</td>\n",
"    <td>210</td>\n",
"    <td>Theft Of Motor Vehicle</td>\n",
"    <td>...</td>\n",
"    <td>Thursday</td>\n",
"    <td>20</td>\n",
"    <td>Auto Theft</td>\n",
"    <td>D11</td>\n",
"    <td>90</td>\n",
"    <td>Junction Area (90)</td>\n",
"    <td>-79.464493</td>\n",
"    <td>43.668125</td>\n",
"    <td>205991</td>\n",
"    <td>1</td>\n",
"   </tr>\n",
"   <tr>\n",
"    <th>206426</th>\n",
"    <td>-79.614281</td>\n",
"    <td>43.745495</td>\n",
"    <td>205892</td>\n",
"    <td>GO-2019399759</td>\n",
"    <td>2019-02-21T12:44:00.000Z</td>\n",
"    <td>2019-03-04T10:19:00.000Z</td>\n",
"    <td>Commercial</td>\n",
"    <td>2135</td>\n",
"    <td>210</td>\n",
"    <td>Theft Of Motor Vehicle</td>\n",
"    <td>...</td>\n",
"    <td>Thursday</td>\n",
"    <td>12</td>\n",
"    <td>Auto Theft</td>\n",
"    <td>D23</td>\n",
"    <td>1</td>\n",
"    <td>West Humber-Clairville (1)</td>\n",
```

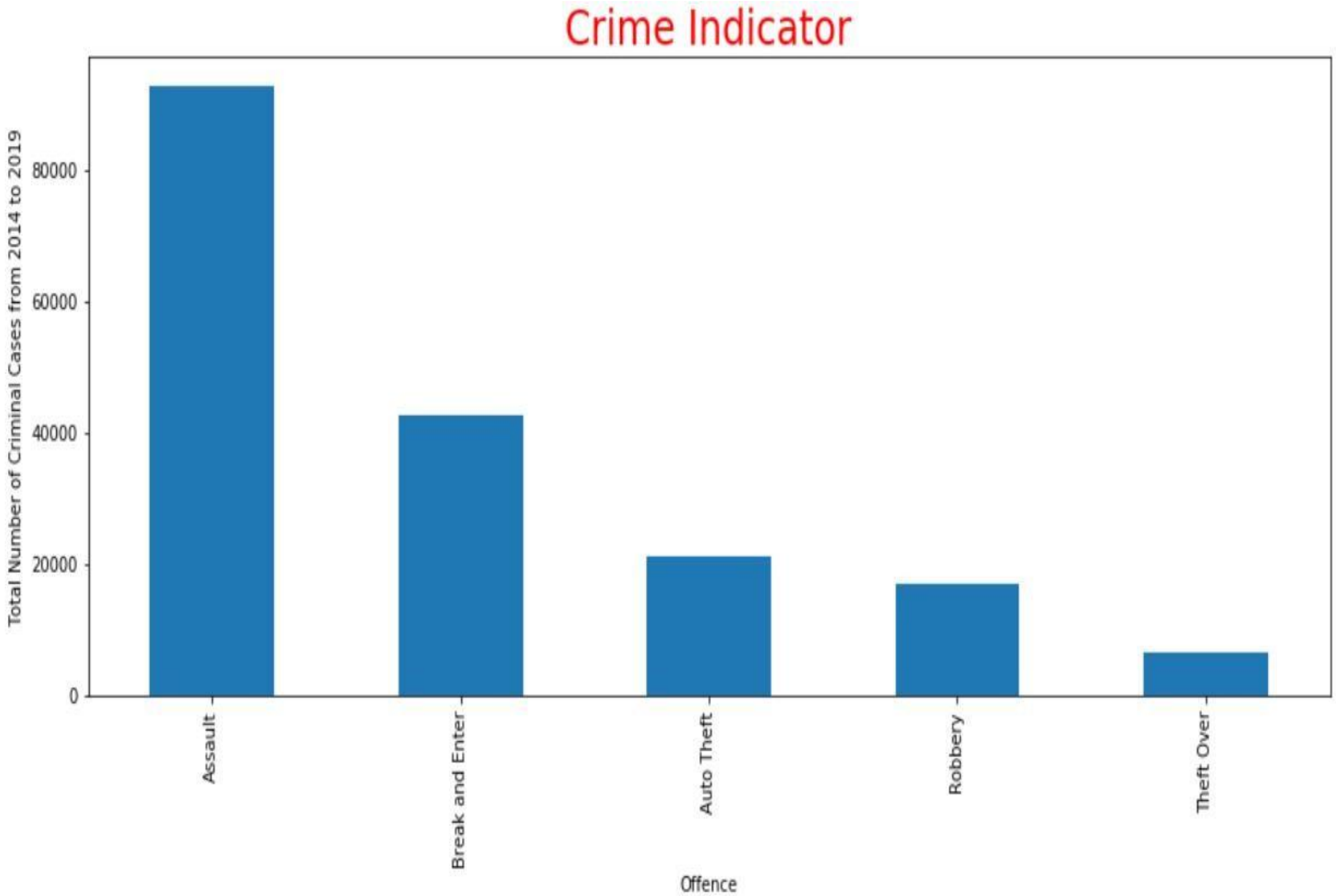
```
"     <td>-79.614281</td>\n",
"     <td>43.745495</td>\n",
"     <td>205992</td>\n",
"     <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"     <th>206427</th>\n",
"     <td>-79.500000</td>\n",
"     <td>43.761066</td>\n",
"     <td>205893</td>\n",
"     <td>GO-2019331091</td>\n",
"     <td>2019-02-21T08:00:00.000Z</td>\n",
"     <td>2019-02-21T17:59:00.000Z</td>\n",
"     <td>Outside</td>\n",
"     <td>2135</td>\n",
"     <td>210</td>\n",
"     <td>Theft Of Motor Vehicle</td>\n",
"     <td>...</td>\n",
"     <td>Thursday</td>\n",
"     <td>8</td>\n",
"     <td>Auto Theft</td>\n",
"     <td>D31</td>\n",
"     <td>27</td>\n",
"     <td>York University Heights (27)</td>\n",
"     <td>-79.500000</td>\n",
"     <td>43.761066</td>\n",
"     <td>205993</td>\n",
"     <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"     <th>206428</th>\n",
"     <td>-79.593422</td>\n",
"     <td>43.738930</td>\n",
"     <td>205894</td>\n",
"     <td>GO-2019333589</td>\n",
"     <td>2019-02-21T23:45:00.000Z</td>\n",
"     <td>2019-02-22T06:00:00.000Z</td>\n",
"     <td>House</td>\n",
"     <td>2135</td>\n",
"     <td>210</td>\n",
"     <td>Theft Of Motor Vehicle</td>\n",
"     <td>...</td>\n",
"     <td>Thursday</td>\n",
"     <td>23</td>\n",
"     <td>Auto Theft</td>\n",
"     <td>D23</td>\n",
"     <td>1</td>\n",
```

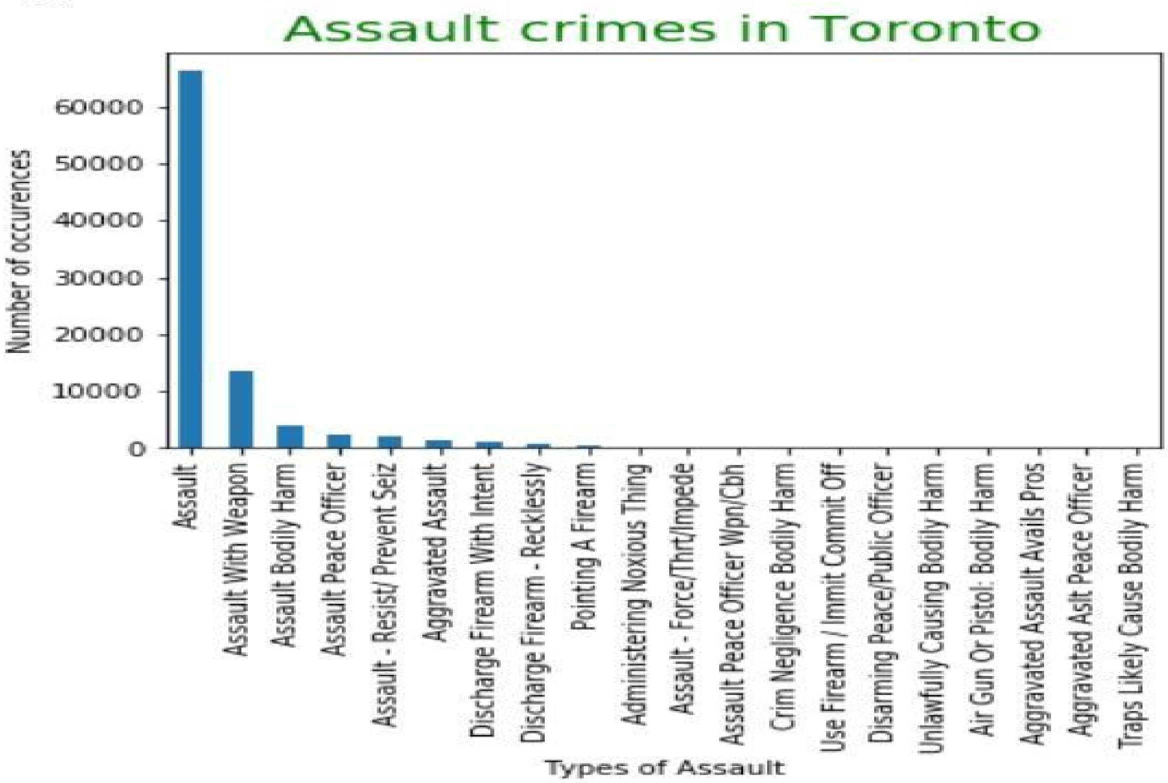
```
"      <td>West Humber-Clairville (1)</td>\n",
"      <td>-79.593422</td>\n",
"      <td>43.738930</td>\n",
"      <td>-79.598869</td>\n",
"      <td>43.743992</td>\n",
"      <td>205998</td>\n",
"      <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>206433</th>\n",
"      <td>-79.348709</td>\n",
"      <td>43.656227</td>\n",
"      <td>205899</td>\n",
"      <td>GO-2019336549</td>\n",
"      <td>2019-02-21T00:01:00.000Z</td>\n",
"      <td>2019-02-22T14:45:00.000Z</td>\n",
"      <td>Commercial</td>\n",
"      <td>2135</td>\n",
"      <td>210</td>\n",
"      <td>Theft Of Motor Vehicle</td>\n",
"      <td>...</td>\n",
"      <td>Thursday</td>\n",
"      <td>0</td>\n",
"      <td>Auto Theft</td>\n",
"      <td>D55</td>\n",
"      <td>70</td>\n",
"      <td>South Riverdale (70)</td>\n",
"      <td>-79.348709</td>\n",
"      <td>43.656227</td>\n",
"      <td>205999</td>\n",
"      <td>1</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>206434</th>\n",
"      <td>-79.217125</td>\n",
"      <td>43.814865</td>\n",
"      <td>205900</td>\n",
"      <td>GO-2019327189</td>\n",
"      <td>2019-02-21T08:00:00.000Z</td>\n",
"      <td>2019-02-21T08:09:00.000Z</td>\n",
"      <td>House</td>\n",
"      <td>2135</td>\n",
"      <td>210</td>\n",
"      <td>Theft Of Motor Vehicle</td>\n",
"      <td>...</td>\n",
"      <td>Thursday</td>\n",
"      <td>8</td>\n",
```

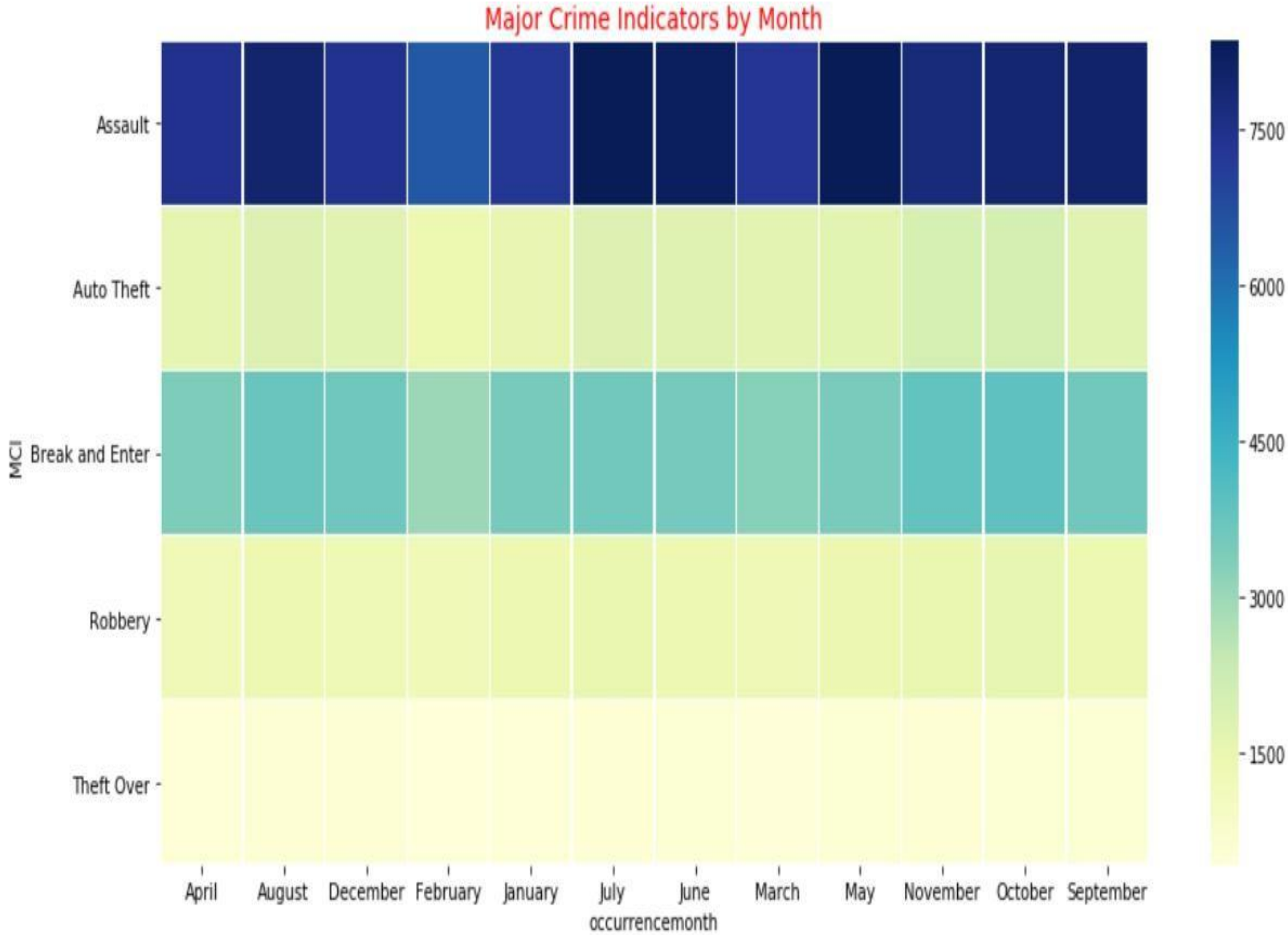
```
"      <td>Auto Theft</td>\n",
"      <td>D42</td>\n",
"      <td>132</td>\n",
"      <td>Malvern (132)</td>\n",
"      <td>-79.217125</td>\n",
"      <td>43.814865</td>\n",
"      <td>206000</td>\n",
"      <td>1</td>\n",
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"<p>206376 rows × 30 columns</p>\n",
"</div>"
],
"text/plain": [
"              X          Y  Index_  event_unique_id  \\\n",
"0     -79.405228  43.656982    7801  GO-20152165447   \n",
"1     -79.307907  43.778732    7802  GO-20151417245   \n",
"2     -79.225029  43.765942    7803  GO-20151421107   \n",
"3     -79.140823  43.778648    7804  GO-20152167714   \n",
"4     -79.288361  43.691235    7805  GO-20152169954   \n",
"5     -79.353310  43.688553    7806  GO-20152123146   \n",
"6     -79.380806  43.649895    7807  GO-20152143239   \n",
"7     -79.380806  43.649895    7808  GO-20152143239   \n",
"8     -79.380806  43.649895    7809  GO-20152143239   \n",
"9     -79.386780  43.663361    7810  GO-20151970205   \n",
"10    -79.611320  43.710689    7811  GO-20151976877   \n",
"11    -79.436020  43.698772    7812  GO-20152060024   \n",
"12    -79.508636  43.720917    7813  GO-20152061875   \n",
"13    -79.416420  43.735794    7814  GO-201521004
```

## 4.2: Output Screens

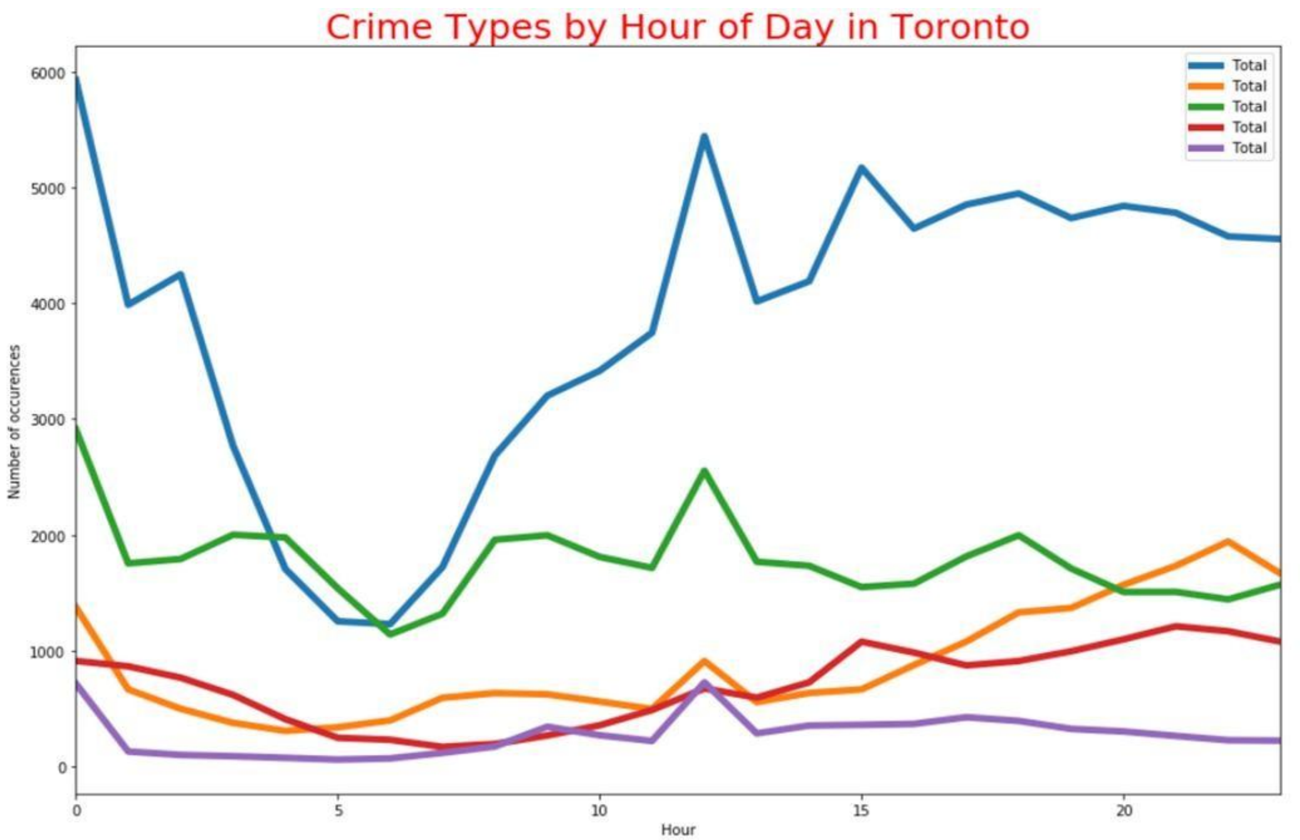



Crime Indicator
Total Number of Criminal Cases from 2014 to 2019
80000
60000
40000
20000
0
Assault
Break and Enter
Auto Theft
Robbery
Theft Over
Offence


Assault crimes in Toronto
Number of occurences
60000
50000
40000
30000
20000
10000
0
Assault
Assault With Weapon
Assault Bodily Harm
Assault Peace Officer
Assault - Resist/ Prevent Seiz
Aggravated Assault
Discharge Firearm With Intent
Discharge Firearm - Recklessly
Pointing A Firearm
Administering Noxious Thing
Assault - Force/Thrt/Impede
Assault Peace Officer Wpn/Cbh
Crim Negligence Bodily Harm
Use Firearm / Immit Commit Off
Disarming Peace/Public Officer
Unlawfully Causing Bodily Harm
Air Gun Or Pistol: Bodily Harm
Aggravated Assault Avails Pros
Aggravated Aslt Peace Officer
Traps Likely Cause Bodily Harm
Types of Assault




| | occurrenceyear | occurrencemonth | occurrenceday | occurrencedayofyear | occurrencedayofweek | occurrencehour | MCI | Division | Hood_ID | premisetype | occurenceday |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 0 | 18.0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1** | 0 | 1 | 15.0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| **2** | 0 | 1 | 16.0 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 |
| **3** | 0 | 2 | 26.0 | 3 | 3 | 3 | 1 | 2 | 3 | 2 | 3 |
| **4** | 0 | 0 | 18.0 | 0 | 0 | 4 | 0 | 3 | 4 | 0 | 0 |




## 4.3: Testing

Software testing is a critical element of software quality assurance and represents the ultimate review of specification, design and code generation.

# CHAPTER 5
# CONCLUSION & FUTURE SCOPE

## 5.1: Conclusion

Machine learning is used for predicting and analysing crime rates. By leveraging vast datasets and sophisticated algorithms, ML models can uncover hidden patterns and correlations within crime data, aiding law enforcement agencies and policymakers in making informed decisions. However, it's essential to acknowledge the limitations and challenges inherent in this approach. Factors such as data biases, privacy concerns, and the dynamic nature of criminal behavior can impact the accuracy and reliability of predictions. Additionally, while ML can provide valuable insights, it should be used in conjunction with human expertise and ethical considerations to ensure fair and responsible use. Despite these challenges, the ongoing advancements in machine learning techniques hold the potential to enhance crime prevention strategies and contribute to safer communication

## 5.2: Future Scope

The future scope for a project focused on crime rate prediction and analysis using machine learning is promising. By leveraging advanced algorithms and vast datasets, such a system could help law enforcement agencies and policymakers better understand crime patterns and trends. Machine learning models could analyse historical crime data to identify patterns, correlations, and predictive factors that contribute to criminal activities. Additionally, the system could incorporate real-time data streams from various sources such as social media, weather patterns, or economic indicators to enhance its predictive capabilities. Overall, the project holds potential to make significant contributions to crime prevention and public safety efforts

# CHAPTER 6

# BIBILOGRAPHY

[1] Mohler, G., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015).

[2] Cholas-Wood, A. (2020).Understanding Risk Assessment Instruments in Criminal Justice Criminology & Public Policy

. [3]: Berk, R. A., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2018).Fair ness in Criminal Justice Risk Assessments.

[4] Wang, T., Rudin, C., Wagner, D., & Sevier, R. (2013).Learning to Detect Patterns of Crime Machine Learning, 90(1)