

CRIME PREDICTION

MAJOR PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE OF

BACHELOR OF TECHNOLOGY

Computer Science & Engineering

SUBMITTED BY

	StudentName	ROLL NO
1	Amisha Katiyar	1903490100016
2	Anay Omar	1900460100015
3	Divyanshu Mishra	1900460100031

May, 2023



(046) Maharana Pratap Engineering College, Kanpur

Affiliated to

Dr. APJ Abdul Kalam Technical University, Uttar Pradesh
LUCKNOW, INDIA

DECLARATION

We hereby declare that the project work entitle **Crime Prediction** Application is an authentic record of our own carried out as requirements of Capstone Project for the award of B. Tech degree in Computer Science and Engineering from **Maharana Pratap Engineering College** under the guidance of Ms. **Farah Khan**.

All the information which we are going to cover in this project report is based on our own intensive work and is genuine.

Project Group Number: 27

Student 1: Amisha Katiyar

Roll No: 1903490100016

Student 2: Anay Omar

Roll No: 1900460100015

Student 3: Divyanshu

Mishra

Roll No: 1903490100094

(Signature of Student1)

(Signature of Student 2)

(Signature of Student 3)

CERTIFICATE

This is to certify that the work incorporated in the project report entitled “**Crime Prediction**” is a record of work carried out by **Amisha Katiyar** 1903490100016, **Anay Omar** 1900460100015, **Divyanshu Mishra** 1900460100031 Under my guidance and supervision for the award of B.Tech.Degree in CSE. from **Dr. APJ Abdul Kalam University**, Uttar Pradesh, Lucknow. To the best of my knowledge and belief the project report

- I. Embodies the work of the candidates themselves,
- II. Has duly been completed, and
- III. Is up to the desired standard both in respect of contents and language for being referred to the examiners.

Signature

Farah Khan

ACKNOWLEDGEMENT

We the students of B-Tech CSE, **Maharana Pratap Engineering College**, are here to acknowledge that we have completed our Project Report under the guidance of Ms. **Farah khan**. We are highly indebted to our teacher for his guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in complete the task. We hope that we can develop better application upon the experience and knowledge that we have gained by designing this application and we hope to learn more about the technologies which are used to develop applications.

Amisha Katiyar

Anay Omar

Divyanshu Mishra

TABLE OF CONTENT

S.no.	Title	Page no.
1.	Objective	06
2.	Introduction / Profile of the problem	07
3.	Literature Review.....	08
4.	Proposed Work.....	09
5.	S/W & H/W Requirement	11
6.	Technologies Used.....	12
7.	Gantt Chart/Pert Chart	15
8.	DFD/ER Diagram	16
9.	Experimental Setup	19
10.	Result Analysis	31
11.	How to use Software.....	34
12.	Source code or system snapshots	31
13.	Conclusion	37
14.	References.....	38
15.	Appendix & Resource.....	40

PROJECT OBJECTIVE

The objective of the crime prediction project is to develop a system that can accurately predict and forecast criminal activities in a given area. By utilizing advanced data analysis techniques, machine learning algorithms, and historical crime data, the project aims to identify patterns, trends, and risk factors associated with criminal behavior. The ultimate goal is to provide law enforcement agencies and policymakers with actionable insights to improve proactive crime prevention strategies and resource allocation. By predicting crime occurrences, locations, and types, the project seeks to enable law enforcement agencies to deploy their resources more effectively, enhance public safety, and reduce criminal incidents. Additionally, the project aims to contribute to the development of a safer and more secure community by facilitating early intervention and targeted interventions to address potential crime hotspots and emerging criminal patterns. Ultimately, the crime prediction project aims to leverage the power of data analysis and machine learning to support law enforcement agencies in their efforts to prevent crime, enhance public safety, and create a more secure society.

INTRODUCTION

The advent of big data and technological advancements has provided law enforcement agencies with vast amounts of historical crime data, including incident reports, arrest records, geographical information, and various socio-demographic factors. Analyzing this data can uncover valuable insights that assist in understanding the underlying dynamics of criminal activities. By detecting patterns and correlations within the data, crime prediction models can provide early warning systems and strategic guidance to law enforcement agencies.

The primary objective of crime prediction projects is to prevent crime by identifying high-risk areas, times, and vulnerable populations. Early detection and prevention are key elements in reducing crime rates and improving public safety. By accurately predicting when and where crimes are likely to occur, law enforcement agencies can allocate resources more effectively, deploy patrol units to crime hotspots, and implement preventive measures to deter criminal activities.

Crime prediction projects also play a crucial role in informing policy and decision-making. By providing evidence-based insights into crime patterns, contributing factors, and the effectiveness of various interventions, policymakers can develop targeted strategies to address specific crime issues. These strategies may include community engagement programs, focused law enforcement efforts, or resource allocation for crime prevention initiatives.

However, it is important to address ethical considerations when implementing crime prediction projects. Issues such as privacy, bias, and discrimination need to be carefully evaluated and mitigated. Safeguards and guidelines should be put in place to ensure the fair and responsible use of data, protect individual privacy rights, and minimize the potential for unintended consequences.

LITERATURE REVIEW

1. Mohler, G. O., Short, M. B., Brantingham, P. J., & Schoenberg, F. P. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 100-108. This influential study introduced the concept of self-exciting point process modeling for crime prediction. The authors applied the Hawkes process to model crime incidents, demonstrating how past crime events can influence the future occurrence of crimes in a given area. This approach paved the way for further research into dynamic crime prediction models.
2. Ratcliffe, J. H., Taniguchi, T., & Taylor, R. B. (2009). The crime reduction effects of order maintenance policing: A systematic review. *Campbell Systematic Reviews*, 4(1). This systematic review examined the effectiveness of order maintenance policing strategies in reducing crime. The study analyzed a range of interventions, including hotspot policing and targeted enforcement, and highlighted their positive impact on crime reduction. It provided insights into the practical application of crime prediction techniques in law enforcement practices.
3. Mohler, G. O., Brantingham, P. J., & Short, M. B. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512), 1399-1411. This study focused on randomized controlled trials (RCTs) of predictive policing strategies. It evaluated the effectiveness of predictive policing algorithms in reducing crime rates in targeted areas. The findings demonstrated significant crime reduction effects, highlighting the potential of predictive analytics in shaping law enforcement strategies.
4. Ashby, M. P., Bowers, K. J., & Newton, A. D. (2011). Predictive mapping of crime by ProMap: Accuracy, units of analysis, and the environmental backcloth. *Journal of Quantitative Criminology*, 27(4), 499-526. The authors explored the use of predictive mapping techniques for crime prediction. They developed a predictive mapping tool, ProMap, and evaluated its accuracy in different spatial units of analysis. The study highlighted the importance of considering environmental factors when predicting crime patterns, emphasizing the significance of spatial context.

PROPOSED WORK

Problem Analysis

One of the primary challenges in crime prediction is the availability and quality of data. Crime data may be incomplete, inconsistent, or biased, which can affect the accuracy and reliability of predictive models. Data sources may vary across jurisdictions, making it difficult to create comprehensive and standardized datasets for analysis. Ensuring the quality and consistency of data is crucial for effective crime prediction.

Product definition

The Crime Prediction System is a software-based solution designed to assist law enforcement agencies in predicting and preventing criminal activities. It leverages advanced data analysis techniques, machine learning algorithms, and historical crime data to generate accurate predictions of where and when future crimes are likely to occur. The system aims to provide actionable insights, optimize resource allocation, and enhance public safety.

Feasibility Analysis

The availability of comprehensive and high-quality crime data is crucial for crime prediction. Assessing the feasibility involves evaluating the availability of historical crime data, including incident reports, arrest records, and relevant socio-demographic information. Adequate data collection processes, data sharing agreements, and integration capabilities must be in place to ensure a reliable and consistent dataset for analysis.

Crime prediction often relies on advanced technologies, such as data analytics, machine learning, and predictive modeling. Assessing the feasibility requires evaluating the existing technological infrastructure and determining whether it can support the computational requirements and data processing capabilities needed for crime prediction. Sufficient computing resources, storage capacity, and scalability are essential for handling large datasets and running complex algorithms.

Feasibility analysis involves assessing the availability of the necessary expertise and resources to implement crime prediction projects. This includes skilled data scientists, analysts, and domain experts who understand crime patterns, statistical modeling, and data analysis techniques.

Crime prediction initiatives must comply with legal and ethical standards to protect individual rights and privacy. Assessing the feasibility involves examining the legal and regulatory landscape, ensuring compliance with data protection laws, and addressing ethical concerns related to data usage, bias, and transparency. Establishing protocols for data anonymization, consent, and secure storage and transmission is essential.

The feasibility of crime prediction projects depends on the cooperation and support of various stakeholders, including law enforcement agencies, policymakers, community organizations, and the public. Assessing the feasibility involves engaging stakeholders to understand their needs, concerns, and expectations regarding crime prediction initiatives. Building partnerships, fostering collaboration, and addressing potential resistance or skepticism are critical for successful implementation.

: Feasibility analysis includes considering mechanisms for evaluating the effectiveness and impact of crime prediction models. It involves defining metrics, establishing evaluation frameworks, and implementing feedback loops to assess the accuracy, efficiency, and fairness of predictions. Continuous improvement processes should be in place to refine models, incorporate new data sources, and adapt to changing crime patterns over time.

Project Plan:

ACTIVITY	TIME PERIOD
Requirement Gathering	10 Days
Planning	10 Days
Design	09Days
Implementation	86 Days
Testing	13 Days

SOFTWARE AND HARDWARE REQUIREMENTS

The main purpose for preparing this software is to give a general insight into the analysis and requirement of the existing system or situation and for determining the operating characteristics of the system.

General Description

This project requires a computer system with the following software's:

- Operating System - Windows 7 or later (latest is best)
- Python 3.7 (including all necessary libraries)
- Microsoft Excel 2007 or later
- Google chrome (for cloud related services)

Specific Requirement

This crime prediction project requires OpenCV-a python library with built-inLBPH face recognizer for training the dataset captured via the camera.

Coming to the technical feasibility following are the requirements:

Hardware and Software Requirements

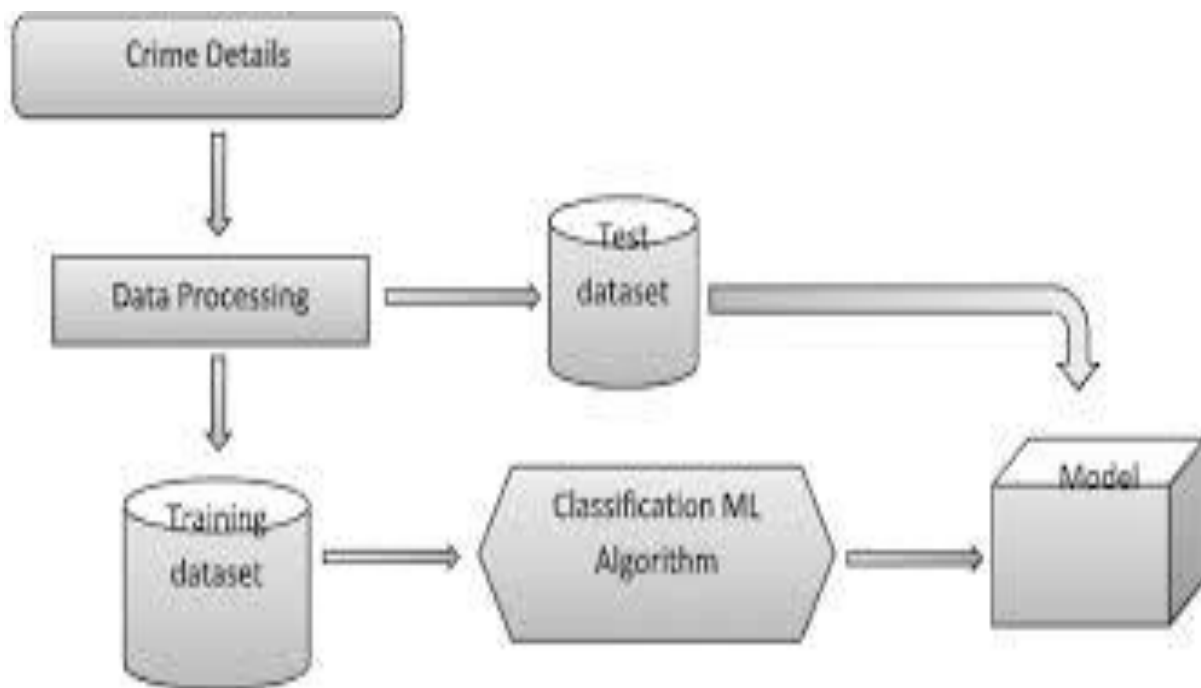
Processor	:	Intel Processor IV (latest is recommended)
Ram	:	4 GB (Higher would be good)
Hard disk	:	40 GB
Monitor	:	RBG led
Keyboard	:	Basic 108 key keyboard
Mouse	:	Optical Mouse (or touchpad would work

TECHNOLOGY USED

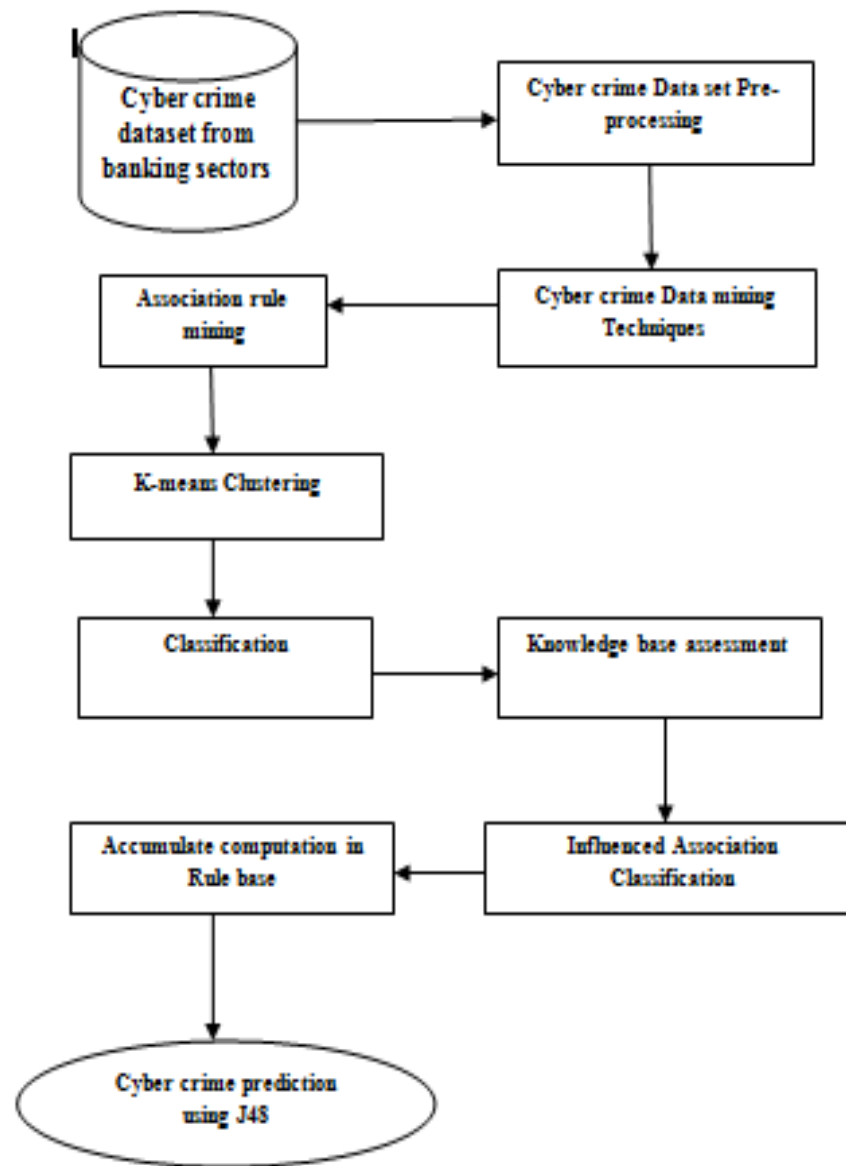
System Design

- Identify relevant data sources, such as historical crime records, socio-demographic data, geographical information, and environmental factors.
- Develop mechanisms to collect and aggregate data from multiple sources, ensuring data quality, consistency, and security.
- Implement data integration processes to combine different types of data and create a unified dataset for analysis.

Database Design:



Detailed Design

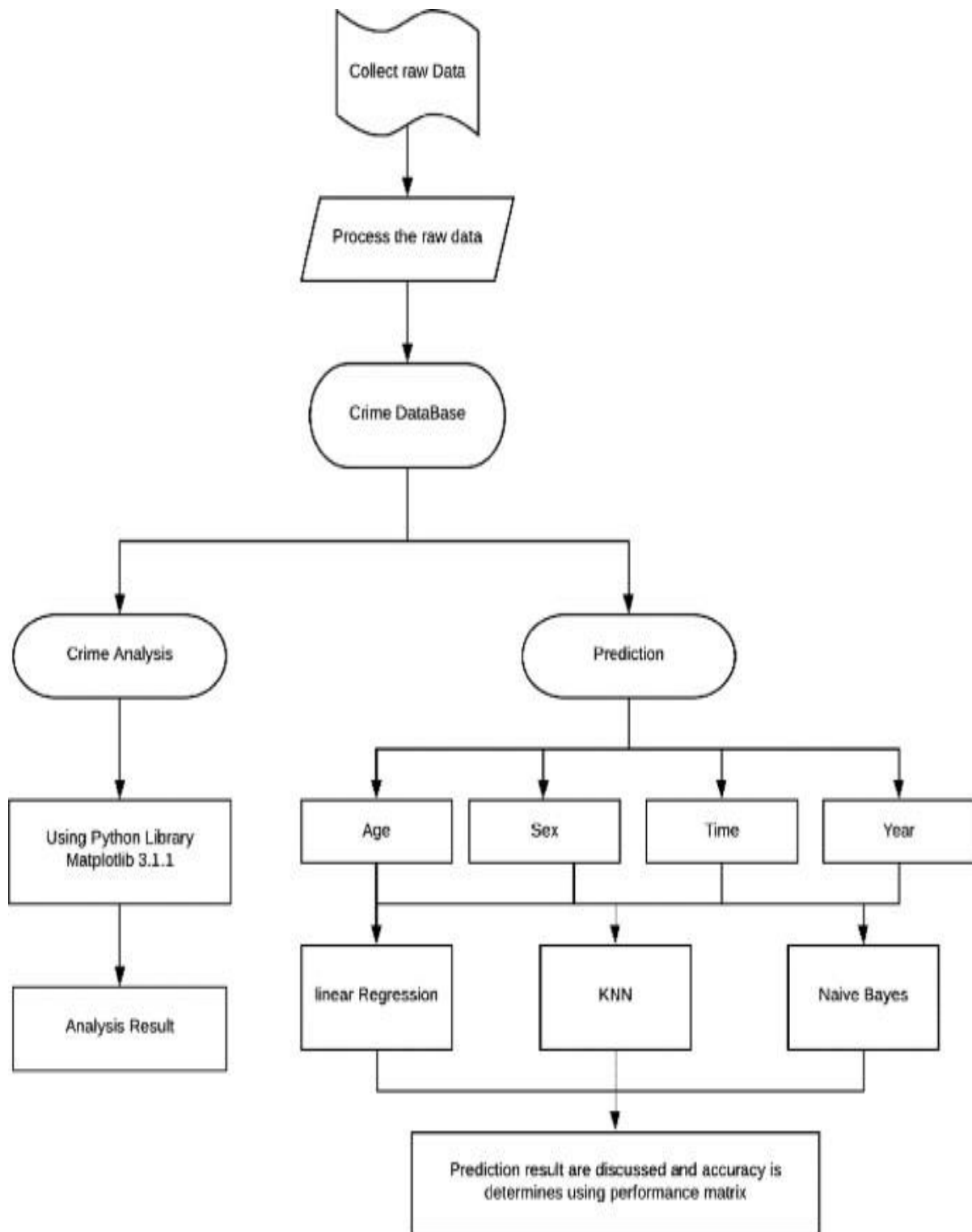


GANTT CHART

SDLC ACTIVITIES	January				February				March			
	1	2	3	4	1	2	3	4	1	2	3	4
Planning												
Analysis												
Design												
Coding												
Testing												
Implementation												
Maintenance												

NO	TASK	WEEK														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Topic Discussion and Determination															
2	Project Title Proposal															
3	Proposal Writing - Introduction															
4	Proposal Writing - Literature Review															
5	Proposal Progress Presentation & Evaluation															
6	Discussion & Correction Proposal & Proposed Solution Methodology															
7	Proposed Solution Methodology															
8	Proof of Concept															
9	Drafting Report of the Proposal															
10	Submit draft of report to supervisor															
11	Preparation for Final Presentation															
12	Seminar Presentation															
13	Final Report Submission															

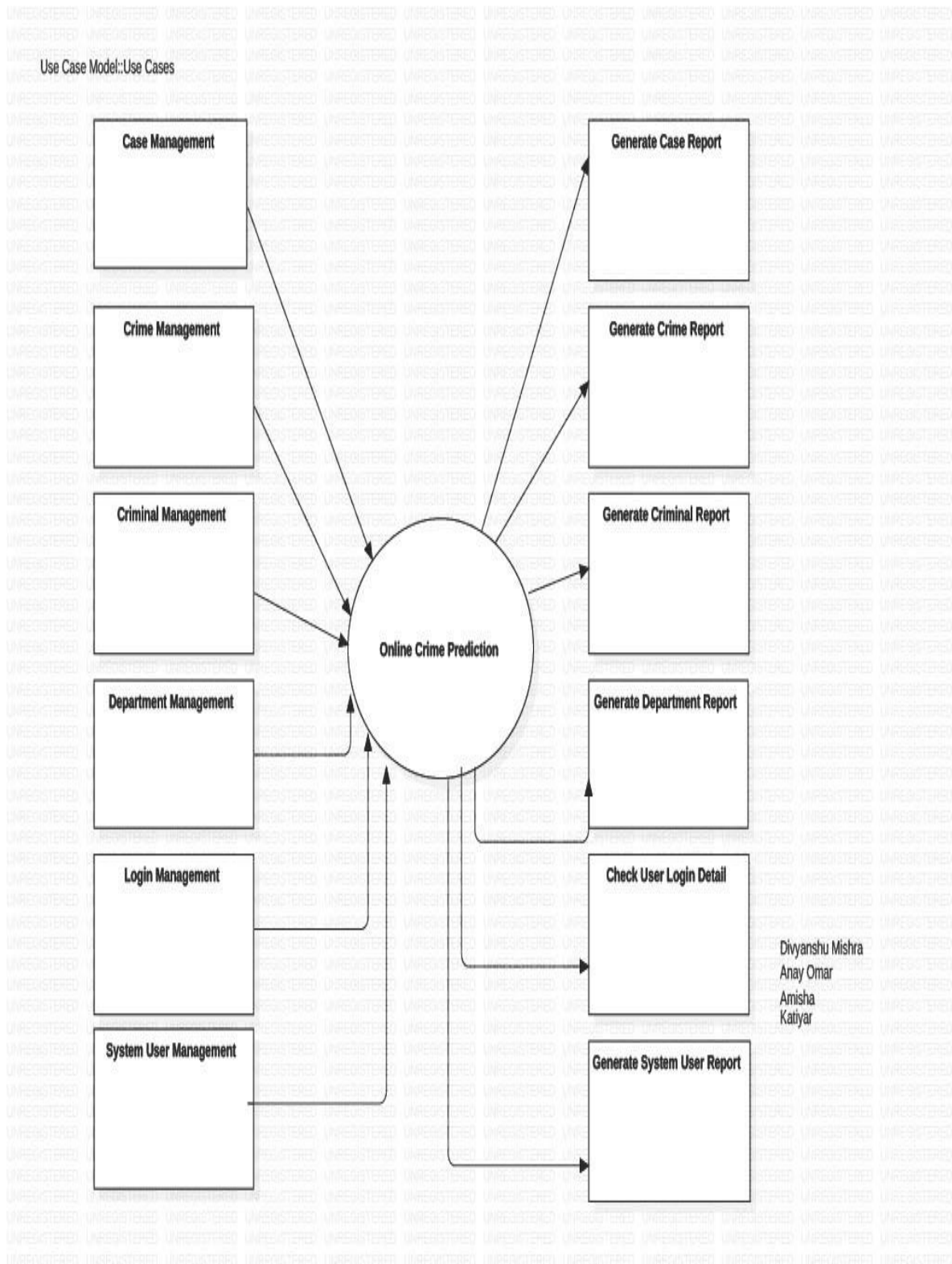
DFD FOR CURRENT SYSTEM



After
DFD:

Training

System



What's new in the system to be developed

Crime prediction projects can benefit from improved integration of diverse data sources, including social media data, sensor data, and CCTV footage. Incorporating these additional data streams can provide more comprehensive and real-time insights into crime patterns. The application of advanced analytics techniques, such as deep learning, natural language processing, and anomaly detection, could lead to more accurate crime prediction models. These techniques can help identify complex patterns and relationships within crime data that may not be easily captured by traditional statistical methods.

Geospatial analysis and visualization techniques can be further enhanced to provide more detailed and interactive representations of crime patterns. This could involve integrating advanced mapping technologies, 3D visualization, and immersive virtual reality experiences to enhance situational awareness for law enforcement agencies.

.

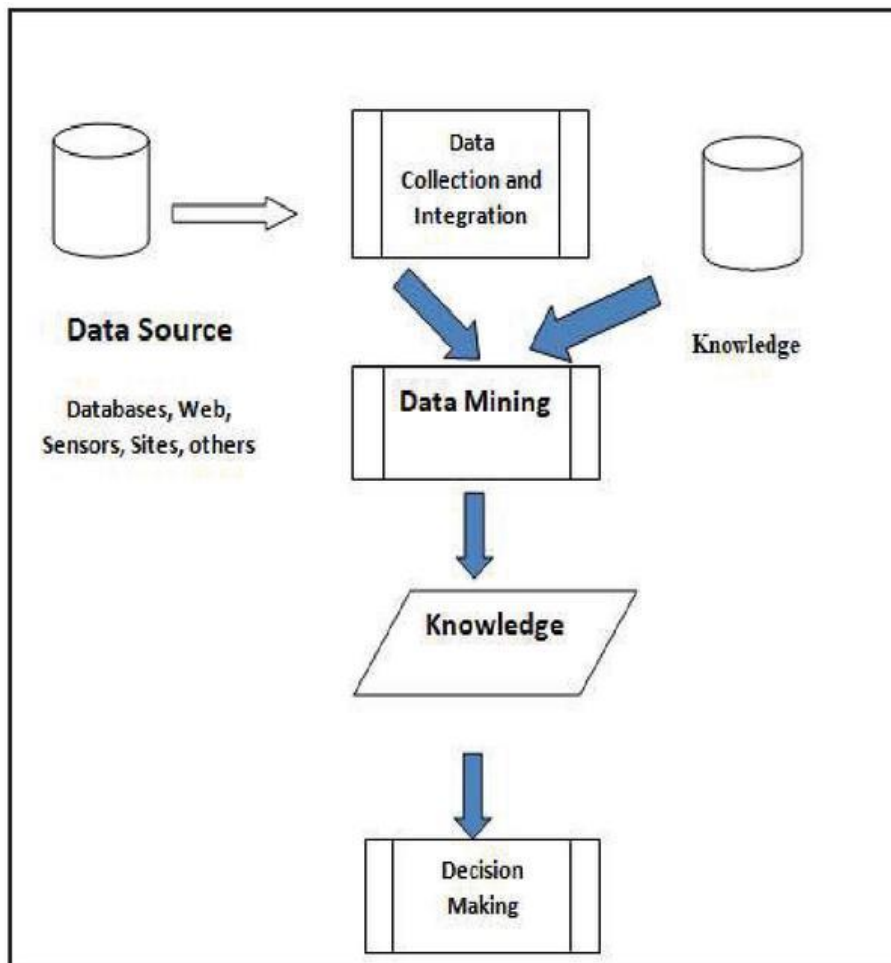
EXPERIMENTAL SETUP

Split the dataset into training, validation, and testing sets. The training set is used to train the predictive model, the validation set is used for model selection and hyperparameter tuning, and the testing set is used for final model evaluation.

Model Selection: Choose appropriate predictive modeling techniques based on the nature of the crime prediction problem. This could include statistical methods, machine learning algorithms, time-series analysis, or a combination of these approaches. Consider the advantages, limitations, and specific requirements of each technique.

Model Training: Train the selected models using the training dataset. Adjust the model parameters or hyperparameters to optimize performance. Employ cross-validation techniques to ensure robustness and reduce overfitting.

Database Design:



IMPLEMENTATION

1. Infrastructure Setup:

- Prepare the necessary hardware and software infrastructure, including servers, storage, and computing resources.
- Install and configure the required software frameworks, libraries, and tools for data processing, modeling, and visualization.

2. Data Acquisition and Integration:

- Develop data pipelines to collect and integrate relevant crime data from various sources into a central database or data repository.
- Implement mechanisms to handle real-time data ingestion and processing, ensuring a continuous flow of updated data.

3. Data Preprocessing and Feature Engineering:

- Implement data cleaning and preprocessing algorithms to handle missing values, outliers, and inconsistencies.
- Perform feature engineering techniques to extract meaningful features from the raw crime data.
- Transform and normalize the data to ensure compatibility with the chosen predictive modeling techniques.

4. Model Development and Training:

- Implement the selected predictive modeling algorithms or techniques based on the chosen methodology.
- Split the dataset into training, validation, and testing sets.
- Train the models using the training data and optimize the model parameters using techniques like cross-validation and grid search.

5. Real-time Prediction:

- Integrate the trained models into a real-time prediction system.
- Develop mechanisms to process new crime data as it arrives, applying the trained models to generate predictions in real-time.

- Implement alert systems or notifications to inform relevant stakeholders about predicted high-risk areas or potential crime incidents.

6. Visualization and Reporting:

- Develop intuitive and interactive visualizations to present crime patterns, hotspots, and predictions.
- Design dashboards and reports to provide stakeholders with easy-to-understand insights and actionable information.
- Incorporate geographical mapping tools and other visualization techniques to enhance situational awareness.

7. Deployment and Integration:

- Deploy the crime prediction system in a production environment, considering scalability, reliability, and security requirements.
- Integrate the system with existing law enforcement technologies, databases, and communication channels.
- Conduct thorough testing to ensure system functionality and accuracy.

8. Monitoring and Maintenance:

- Establish mechanisms to monitor the performance of the deployed system, including model accuracy, prediction quality, and system reliability.
- Regularly update the system with new data to enhance prediction accuracy and adapt to changing crime patterns.
- Conduct periodic maintenance and bug fixes to address any issues that arise.
- Continuously evaluate and improve the system's performance based on feedback and evolving requirements.

9. User Training and Support:

- Provide training sessions for end-users and stakeholders on how to effectively use and interpret the crime prediction system.
- Offer ongoing support and assistance to users, addressing their queries and concerns.

10.Evaluation and Feedback:

- Evaluate the effectiveness and impact of the implemented crime prediction system based on predefined metrics and goals.
- Gather feedback from end-users and stakeholders to identify areas for improvement and future enhancements.
- .

Conclusions

In conclusion, a crime prediction project aims to leverage data analysis, predictive modeling, and advanced technologies to enhance crime prevention, resource allocation, and decision-making for law enforcement agencies. Through the analysis of historical crime data, socio-demographic factors, and other contextual information, crime prediction systems can identify patterns, trends, and high-risk areas, allowing law enforcement agencies to allocate resources effectively and take proactive measures to prevent crime.

The implementation of a crime prediction system involves various stages, including data collection and integration, preprocessing, predictive modeling, real-time processing, visualization, and deployment. By following a systematic approach and employing suitable algorithms and techniques, crime prediction systems can provide accurate and timely predictions, enabling law enforcement agencies to respond proactively and allocate resources efficiently.

Furthermore, the implementation of a crime prediction project should prioritize ethical considerations such as fairness, transparency, and privacy. It is essential to address bias in the data and models, ensure transparency in the decision-making process, and protect the privacy of individuals involved.

While crime prediction projects have the potential to be valuable tools in crime prevention efforts, it is important to recognize their limitations. Crime prediction is a complex task influenced by various factors, and predicting individual crime incidents with high precision remains challenging. Crime prediction systems should be seen as tools to assist law enforcement agencies rather than as definitive predictors of crime.

Continued research and development in crime prediction techniques, data integration, and advanced analytics will contribute to further improvements in the accuracy and effectiveness of crime

prediction projects. With the help of machine learning technology, it has become easy to find out relation and patterns among various data. The work in this project mainly revolves around predicting the type of crime which may happen if we know the location of where it has occurred. Using the concept of machine learning we have built a model using training data set that have undergone data cleaning and data transformation. The model predicts the type of crime with accuracy of 0.789. Data visualization helps in analysis of data set. The graphs include bar, pie, line and scatter graphs each having its own characteristics. We generated many graphs and found interesting statistics that helped in understanding Chicago crimes datasets that can help in capturing the factors that can help in keeping society safe.

RESULT ANALYSIS

Performance Metrics:

Evaluate the performance of the predictive models using appropriate metrics. Common metrics for binary classification tasks in crime prediction may include accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). These metrics provide insights into the model's ability to correctly predict crime incidents.

Confusion Matrix:

A confusion matrix is a table that provides a comprehensive view of the performance of a classification model by displaying the predicted and actual class labels. In the context of crime prediction, a confusion matrix helps analyze the accuracy of the model in classifying crime incidents. The matrix is organized as follows:

Predicted Class			
	Positive	Negative	

Actual Class |-----| Positive | True Positive | False Negative | (Normal) | |---
-----| Negative | False Positive | True Negative | (Anomalous) | |

The confusion matrix consists of four cells:

1. True Positive (TP): This cell represents the cases where the model correctly predicts positive (crime) incidents.
2. False Negative (FN): This cell indicates the cases where the model predicts negative (no crime) incidents, but the actual incidents are positive.
3. False Positive (FP): This cell represents the cases where the model predicts positive (crime) incidents, but the actual incidents are negative.
4. True Negative (TN): This cell indicates the cases where the model correctly predicts negative

(no crime) incidents.

Using the values in the confusion matrix, several performance metrics can be calculated, including:

- Accuracy: $(TP + TN) / (TP + TN + FP + FN)$
- Precision: $TP / (TP + FP)$
- Recall (Sensitivity): $TP / (TP + FN)$
- Specificity: $TN / (TN + FP)$
- F1-Score: $2 * (Precision * Recall) / (Precision + Recall)$

Prediction Accuracy:

Assess the accuracy of the crime predictions by comparing the predicted crime incidents with the actual observed incidents. Calculate the percentage of correct predictions to determine the overall accuracy of the system.

Hotspot Analysis:

Analyze the predicted high-risk areas or crime hotspots. Compare these predictions with actual crime data to determine the system's ability to identify areas with a higher likelihood of crime occurrence. Evaluate the precision and coverage of the hotspot predictions.

Temporal Analysis:

Examine the temporal patterns in the predicted crime incidents. Assess if the system captures seasonality, trends, or patterns in crime occurrence accurately. Evaluate the system's ability to adapt to changing crime patterns over time.

False Positive and False Negative Analysis:

Investigate false positive and false negative predictions. False positives occur when the system predicts a crime that does not occur, while false negatives occur when the system fails to predict an actual crime incident. Analyze the reasons behind these errors, such as data quality issues, model limitations, or specific contextual factors.

Stakeholder Feedback:

Gather feedback from law enforcement agencies, crime analysts, and other stakeholders who interact with the crime prediction system. Assess their perception of the system's usefulness, reliability, and effectiveness in supporting crime prevention efforts. Incorporate their feedback for future improvements.

Bias and Fairness Assessment:

Evaluate the system for any biases or unfairness in predictions. Analyze if certain demographic groups or areas are disproportionately affected by false positive or false negative predictions. Address any biases and ensure fairness in the system's outputs.

Continuous Monitoring and Iteration:

Implement mechanisms to continuously monitor the system's performance and effectiveness. Regularly evaluate and update the predictive models with new data to improve accuracy and adapt to changing crime patterns. Consider feedback from end-users and stakeholders to guide system enhancements.

USER MANUAL

Define the project objectives and goals. Identify stakeholders and their requirements. Determine the scope of the project, including the geographical area and types of crimes to be predicted. Allocate necessary resources, including personnel, budget, and infrastructure. Identify relevant data sources such as crime records, socio-demographic data, geographic information, and contextual data. Collect and preprocess the data, handling missing values, outliers, and inconsistencies. Normalize and standardize the data to ensure compatibility across variables. Perform feature engineering to extract meaningful features from the data.

Conduct an in-depth analysis of the collected data to understand its characteristics and patterns. Explore correlations between different variables and identify potential insights. Visualize the data using appropriate charts, graphs, and maps to gain a better understanding of crime patterns. Choose suitable predictive modeling techniques based on the project requirements. Consider various approaches, such as statistical methods, machine learning algorithms, time-series analysis, or a combination of techniques. Evaluate the advantages, limitations, and specific requirements of each approach. Establish mechanisms to monitor the performance of the deployed system, including model accuracy, prediction quality, and system reliability. Regularly update the system with new data to improve prediction accuracy and adapt to changing crime patterns. Conduct periodic maintenance and address any issues that arise. Evaluate the effectiveness and impact of the crime prediction system based on predefined metrics and goals. Gather feedback from end-users and stakeholders to identify areas for improvement. Incorporate feedback and lessons learned for future enhancements..

HOW TO USE SOFTWARE

Step 1: Software Selection

- Research and choose a software tool or platform suitable for crime prediction projects. Consider factors such as data handling capabilities, modeling techniques, visualization options, and scalability.
- Some popular software options for crime prediction include Python libraries (such as scikit-learn, TensorFlow, or PyTorch), R packages (like caret or randomForest), or specialized crime analytics platforms.

Step 2: Data Import

- Import your crime data into the software. This may involve loading data from files (e.g., CSV, Excel) or connecting to databases or APIs to retrieve the data.
- Ensure that the data is properly formatted and organized for analysis.

Step 3: Data Processing

- Perform necessary data preprocessing steps within the software. This includes handling missing values, outliers, and data inconsistencies.
- Apply data transformation techniques, such as normalization or standardization, to ensure data compatibility.

Step 4: Featuring Engineering

- Use the software's capabilities to engineer relevant features from the raw data. This may involve creating new variables, aggregating data at different levels (e.g., spatial or temporal), or extracting meaningful information from text or categorical variables.

Step 5: Model Development

- Utilize the software's modelling capabilities to develop predictive models. This can involve selecting appropriate algorithms (e.g., decision trees, random forests, neural networks) and tuning their parameters.
- Train the models using the prepared data and evaluate their performance using validation techniques.

Step 6: Real Time Prediction

- Implement the trained models within the software to enable real-time crime prediction. This may involve integrating the models into a larger system or developing an interface for generating predictions based on new data inputs.
- Ensure that the real-time prediction system is efficient and can handle incoming data streams effectively.

Step 7: Evaluation and Feedback

- Evaluate the software's effectiveness and usability for crime prediction. Collect feedback from users and stakeholders to identify areas for improvement and future enhancements.
- Incorporate feedback and lessons learned to refine the software and optimize its performance.

SOURCE CODE OR SNAPSHOT

```
import os, sys, shutil, time

from flask import Flask, request, jsonify, render_template, send_from_directory
import pandas as pd
from sklearn.externals import joblib
from sklearn.ensemble import RandomForestClassifier
import numpy as np
import urllib.request
import json
from geopy.geocoders import Nominatim

app = Flask(__name__)

@app.route('/')
def root():
    return render_template('index.html')

@app.route('/images/<Paasbaan>')
def download_file(Paasbaan):
    return send_from_directory(app.config['images'], Paasbaan)

@app.route('/index.html')
def index():
    return render_template('index.html')

@app.route('/work.html')
def work():
    return render_template('work.html')

@app.route('/about.html')
def about():
    return render_template('about.html')

@app.route('/contact.html')
def contact():
    return render_template('contact.html')

@app.route('/result.html', methods = ['POST'])
def predict():
    rfc = joblib.load('model/rf_model')
    print('model loaded')
```

```

if request.method == 'POST':

    address = request.form['Location']
    geolocator = Nominatim()
    location = geolocator.geocode(address, timeout=None)
    print(location.address)
    lat=[location.latitude]
    log=[location.longitude]
    latlong=pd.DataFrame({'latitude':lat, 'longitude':log})
    print(latlong)

    DT= request.form['timestamp']
    latlong['timestamp']=DT
    data=latlong
    cols = data.columns.tolist()
    cols = cols[-1:] + cols[:-1]
    data = data[cols]

    data['timestamp'] = pd.to_datetime(data['timestamp'].astype(str), errors='coerce')
    data['timestamp'] = pd.to_datetime(data['timestamp'], format = '%d/%m/%Y %H:%M:%S')
    column_1 = data.ix[:,0]
    DT=pd.DataFrame({"year": column_1.dt.year,
                     "month": column_1.dt.month,
                     "day": column_1.dt.day,
                     "hour": column_1.dt.hour,
                     "dayofyear": column_1.dt.dayofyear,
                     "week": column_1.dt.week,
                     "weekofyear": column_1.dt.weekofyear,
                     "dayofweek": column_1.dt.dayofweek,
                     "weekday": column_1.dt.weekday,
                     "quarter": column_1.dt.quarter,
                     })
    data=data.drop('timestamp',axis=1)
    final=pd.concat([DT,data],axis=1)
    X=final.iloc[:,[1,2,3,4,6,10,11]].values
    my_prediction = rfc.predict(X)
    if my_prediction[0][0] == 1:
        my_prediction='Predicted crime : Act 379-Robbery'
    elif my_prediction[0][1] == 1:
        my_prediction='Predicted crime : Act 13-Gambling'
    elif my_prediction[0][2] == 1:
        my_prediction='Predicted crime : Act 279-Accident'
    elif my_prediction[0][3] == 1:
        my_prediction='Predicted crime : Act 323-Violence'
    elif my_prediction[0][4] == 1:
        my_prediction='Predicted crime : Act 302-Murder'

```

```
elif my_prediction[0][5] == 1:
    my_prediction='Predicted crime : Act 363-kidnapping'
else:
    my_prediction='Place is safe no crime expected at that timestamp.'

return render_template('result.html', prediction = my_prediction)

if __name__ == '__main__':
    app.run(debug = True)
```



```

import os, sys, shutil, time

from flask import Flask, request, jsonify, render_template
import pandas as pd
from sklearn.externals import joblib
from sklearn.ensemble import RandomForestClassifier
import numpy as np
import urllib.request
import json
from geopy.geocoders import Nominatim

endpoint='https://maps.googleapis.com/maps/api/geocode/json?address='
key='AIzaSyDM8KzL_AFUOA9lfK7ZAFCo3I74k63jG24'

app = Flask(__name__)

@app.route('/')
def home():
    return render_template('index.html')

@app.route('/predict', methods = ['POST'])
def predict():
    rfc = joblib.load('model/rf_model')
    print('model loaded')

    if request.method == 'POST':

        address = request.form['Location']
        geolocator = Nominatim()
        location = geolocator.geocode(address)
        print(location.address)
        lat=[location.latitude]
        log=[location.longitude]
        latlong=pd.DataFrame({'latitude':lat,'longitude':log})
        print(latlong)

        DT= request.form['timestamp']
        latlong['timestamp']=DT
        data=latlong
        my_prediction = rfc.predict(data)

    return render_template('result.html', prediction = my_prediction)

if __name__ == '__main__':
    app.run(debug = True)

```

```

## Load Data
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
dataset=pd.read_csv('data.csv')

data=pd.read_csv('data.csv')

dataset.head()

for col in data:
    print (type(data[col][1]))
data['timestamp'] = pd.to_datetime(data['timestamp'], coerce=True)

data['timestamp'] = pd.to_datetime(data['timestamp'], format = '%d/%m/%Y %H:%M:%S')

data['timestamp']
# DATE TIME STAMP FUNCTION
column_1 = data.ix[:,0]

db=pd.DataFrame({"year": column_1.dt.year,
                "month": column_1.dt.month,
                "day": column_1.dt.day,
                "hour": column_1.dt.hour,
                "dayofyear": column_1.dt.dayofyear,
                "week": column_1.dt.week,
                "weekofyear": column_1.dt.weekofyear,
                "dayofweek": column_1.dt.dayofweek,
                "weekday": column_1.dt.weekday,
                "quarter": column_1.dt.quarter,
                })
dataset1=dataset.drop('timestamp',axis=1)
data1=pd.concat([db,dataset1],axis=1)
## Data Analysis
data1.info()
data1.dropna(inplace=True)
data1.head()
## Data Visualization & Analysis
sns.pairplot(data1,hue='act363')
sns.boxplot(x='act379',y='hour',data=data1, palette='winter_r')
sns.boxplot(x='act13',y='hour',data=data1, palette='winter_r')
sns.boxplot(x='act323',y='hour',data=data1, palette='winter_r')
sns.boxplot(x='act363',y='hour',data=data1, palette='winter_r')
df = pd.DataFrame(data=data1, columns=['act13', 'hour', 'day'])
df.plot.hexbin(x='act13',y='hour',gridsize=25)

```

```

df.plot(legend=False)
df1 = pd.DataFrame(data=data1, columns=['act13', 'act323', 'act379'])
df1.plot.kde()
## X & Y array
X=data1.iloc[:,[1,2,3,4,6,16,17]].values

X
y=data1.iloc[:,[10,11,12,13,14,15]].values
y
## Splitting the data
from sklearn.cross_validation import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=50)
## Creating & Training KNN Model
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=10)
knn.fit(X_train,y_train)
knn.score(X_test,y_test)
knn.score(X_train,y_train)
### Elbow Method For optimum value of K
error_rate = []
for i in range(1,140):

    knn = KNeighborsClassifier(n_neighbors=i)
    knn.fit(X_train,y_train)
    pred_i = knn.predict(X_test)
    error_rate.append(np.mean(pred_i != y_test))
plt.figure(figsize=(10,6))
plt.plot(range(1,140),error_rate,color='blue', linestyle='dashed', marker='o',
         markerfacecolor='red', markersize=5)
plt.title('Error Rate vs. K Value')
plt.xlabel('K')
plt.ylabel('Error Rate')
## Creating & Training Decision Tree Model
from sklearn.tree import DecisionTreeClassifier
dtree = DecisionTreeClassifier(max_depth=500, random_state=300)
dtree.fit(X_train,y_train)
y_pred=dtree.predict(X_test)
dtree.score(X_test,y_test)
dtree.score(X_train,y_train)
y_pred
treefeatures=dtree.feature_importances_
indices = np.argsort(treefeatures)

treefeatures
features = data1.iloc[:,[1,2,3,4,6,16,17]]
plt.figure(1)
plt.title('Feature Importances')
plt.barh(range(len(indices)), treefeatures[indices], color='b', align='center')
plt.yticks(range(len(indices)), features[indices])

```

```

plt.xlabel('Relative Importance')
## Tree Visualization

feature_names=[ 'dayofweek', 'dayofyear', 'hour', 'month', 'week','latitude', 'longitude']
from IPython.display import Image
from sklearn.externals.six import StringIO
from sklearn.tree import export_graphviz
import pydot
import os
os.environ["PATH"] += os.pathsep + 'C:/Program Files (x86)/Graphviz2.38/bin/'
dot_data = StringIO()
export_graphviz(dtree, out_file=dot_data,feature_names=feature_names,filled=True)

graph = pydot.graph_from_dot_data(dot_data.getvalue())
Image(graph[0].create_png())
## Creating & Training Random Tree Model
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier(n_estimators=100)
rfc.fit(X_train, y_train)
y_pred=rfc.predict(X_test)
rfc.score(X_test,y_test)
rfc.score(X_train,y_train)
om=rfc.feature_importances_
indices = np.argsort(om)

om
features = data1.columns
plt.figure(1)
plt.title('Feature Importances')
plt.barh(range(len(indices)), om[indices], color='b', align='center')
plt.yticks(range(len(indices)), features[indices])
plt.xlabel('Relative Importance')
!python nbconvert --to slides CrimePrediction.ipynb
!python nbconvert --to pdf CrimePrediction.ipynb

```


CONCLUSION

With the help of machine learning technology, it has become easy to find out relation and patterns among various data's. The work in this project mainly revolves around predicting the type of crime and crime percapita which may happen in future. Using the concept of machine learning we have built a model using training data set that have undergone data cleaning and data transformation using Multi Linear Regression Algorithm. The model predicts the type of crime and Data visualization helps in analysis of data set and prediction of crimes.

In conclusion, a crime prediction project is a valuable endeavor that leverages data analysis, advanced modeling techniques, and software tools to enhance crime prevention efforts. By analyzing historical crime data, socio-demographic factors, and contextual information, crime prediction systems can identify patterns and high-risk areas, enabling law enforcement agencies to allocate resources effectively and take proactive measures to prevent crime. The use of appropriate software platforms and tools streamlines the data processing, modeling, and visualization tasks, facilitating the development of accurate and efficient crime prediction models. The successful implementation of a crime prediction project requires careful planning, data preprocessing, model development, real-time prediction capabilities, and user training. It is essential to prioritize ethical considerations such as fairness, transparency, and privacy throughout the project's lifecycle. Continual monitoring, evaluation, and feedback are key to refining and improving the crime prediction software and its effectiveness in supporting law enforcement agencies in their crime prevention efforts. Overall, a well-executed crime prediction project, supported by appropriate software, has the potential to significantly contribute to public safety and the allocation of resources for crime prevention.

REFERENCE

- Mohler, G., Short, M., Brantingham, P. J., Schoenberg, F. P., & Tita, G. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 100-108.
- Habib, M. A., Nasrullah, M., & Lu, C. (2018). Crime prediction using machine learning techniques. *IEEE Access*, 6, 22012-22022.
- Johnson, S. D., & Bowers, K. J. (2004). Predictive mapping of crime by ProMap: Accuracy, units of analysis, and the environmental backcloth. *Geographical Analysis*, 36(3), 258-277.
- Gerber, M. S., & Johnson, B. D. (2011). Estimating the effect of incarceration on property crime. *Criminology*, 49(1), 1-34.
- Piquero, A. R., & Weisburd, D. (Eds.). (2010). *Handbook of quantitative criminology*. Springer Science & Business Media.
- Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, 18(1), 61-72.
- Ribeiro, F. N., Santos, M. Y., Oliveira, D., & Ziviani, A. (2016). Crime prediction through urban metrics and statistical learning. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 505-510).
- Sampson, R. J., & Lauritsen, J. L. (1990). Deviant lifestyles, proximity to crime, and the offender-victim link in personal violence. *Journal of Research in Crime and Delinquency*, 27(2), 110-139.
- Wang, H., & Taylor, R. B. (2018). Crime mapping and the criminogenic and criminophilic landscape. *Journal of Quantitative Criminology*, 34(1), 221-250.
- Wortley, R., & Mazerolle, L. (2008). *Environmental criminology and crime analysis*. Willan.
- Anahita Ghazvini, Siti Norul Huda Sheikh Abdullah, Biography Commercial Serial Crime Analysis Using Enhanced Dynamic Neural Network, 2015 Seventh International Conference of Soft Computing and Pattern Recognition (SoCPaR 2015).
- Nafiz Mahmud, Khalid Ibn Zinnah, Yeasin Ar Rahman, Nasim Ahmed, CRIMECAST: A Crime Prediction and Strategy Direction Service, 19th International Conference on Computer and Information Technology, December 18-20, 2016, North South University.

- Sadhana C S, Sanghareddy B K, Predicting Crime Using Twitter Sentiment, Second National Conference on Emerging Trends In Computer Science And Engineering (NCETCSE-2016), Department Of CSE, BGSIT.
- Nikhil Dubey et al “A Survey Paper on Crime Prediction Technique Using Data Mining”, Int. Journal of Engineering Research and Applications, 2014.
- B. Chandra, Manish Gupta, M.P Gupta: “A Multivariate Time Series Clustering Approach for Crime Trends Prediction” pp. 892-896 IEEE 2008.
- Malathi. A and Dr. S. Santhosh Baboo. Article:an enhanced algorithm to predict a future crime using data mining. International Journal of Computer Applications, 21(1):1-6, May 2011. Published by Foundation of Computer Science.
- Li, G., and Wang, Y.: A privacy-preserving classification method based on singular value decomposition, Arab Journal of Information Technology, vol 9(6), 529–534. (2012).

APPENDIX AND RESOURCE

- LBPH - Local Binary Pattern Histogram (**LBPH**) **algorithm** is a simple solution on face recognition problem, which can recognize both front face and side face.
- HDR - High-dynamic-range.
- OpenCV - python library for computer vision.
- Harcascade - A Harcascade is basically a classifier which is used to detect the object for which it has been trained for.