

CRIME ANALYSIS AND DETECTION BASED ON LOCATION USING MACHINE LEARNING

Submitted in partial fulfillment of the
requirements for the award of
Bachelor of Engineering degree in Computer Science and Engineering

By

PEDDINENI VARSHITH(40110931)

PATAN FIROZ KHAN(40110914)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

Accredited with Grade "A++" by NAAC

Category- 1 University by UGC

**JEPPIAAR NAGAR, RAJIV GANDHISALAI,
CHENNAI - 600119**

APRIL - 2024



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Category - I University by UGC

Accredited "A++" by NAAC | Approved by AICTE

www.sathyabama.ac.in

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **Peddineni Varshith(40110931)** and **Patan Firoz Khan(40110914)** who carried out the Project entitled "**CRIME ANALYSIS AND DETECTION BASED ON LOCATION USING MACHINE LEARNING**" under my supervision from November 2023 to April 2024.

Jany
12/4/24
Internal Guide

Dr. S. L. JANY SHABU M.Tech., Ph.D.

HEAD OF THE DEPARTMENT
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SATHYABAMA
Head of the Department
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Dr. L. LAKSHMANAN, MCE, Ph.D.
Jeppiaar Nagar, Chennai - 600 119.

Submitted for Viva voce Examination held on 15/04/24

Ismael
15/4/24
Internal Examiner

[Signature]
15/4/24
External Examiner

DECLARATION

I, **Peddineni Varshith(40110931)** and **Patan Firoz Khan(40110914)**, hereby declare that the Project Report entitled **CRIME ANALYSIS AND DETECTION BASED ON LOCATION USING MACHINE LEARNING** done by us under the guidance of **Dr S.L.Jany Shabu M.Tech., Ph.D** is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in **Computer Science and Engineering**.

DATE: 12/04/24

PLACE: Chennai



Peddineni Varshith

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board Of Management of Sathyabama Institute Of Science And Technology** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T.Sasikala M.E., Ph. D, Dean**, School of Computing, **Dr. L. Lakshmanan M.E., Ph.D.**, Head of the Department of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. S. L. JANY SHABU M.Tech., Ph.D**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

In recent years, the advancement of technology has significantly influenced various aspects of society, including law enforcement. One crucial aspect of modern law enforcement is the utilization of location-based technologies and data analysis techniques for crime detection and prediction. Traditional methods of crime detection often rely on reactive measures, responding to incidents after they occur. However, with the emergence of location-based technologies and the proliferation of data sources such as mobile phones and surveillance systems, there is an opportunity to adopt proactive approaches to crime prevention. By analyzing real-time location data, law enforcement agencies can identify crime hotspots and patterns, allowing them to allocate resources more effectively and deter criminal activity. The primary objective of this research is to develop a comprehensive location-based crime detection system that integrates advanced algorithms and machine learning techniques. Data will be collected from diverse sources, including mobile phone records, surveillance footage, and historical crime databases. Advanced algorithms such as clustering and classification techniques will be applied to analyze the data and identify crime patterns. Machine learning algorithms, including decision trees, neural networks, and support vector machines, will be utilized to develop predictive models for crime forecasting. Additionally, visualization techniques such as heatmaps and geographic information systems (GIS) will be employed to present the results in an intuitive and actionable format. It is anticipated that the proposed location-based crime detection system will provide law enforcement agencies with a reliable and efficient tool for enhancing their crime prevention strategies. By leveraging real-time data and advanced analytical techniques, the system aims to improve the accuracy and timeliness of crime detection, ultimately leading to a reduction in criminal activity and an increase in public safety. The evaluation of the system's effectiveness will provide valuable insights into its potential impact on law enforcement practices and inform future developments in the field of crime detection and prevention. In conclusion, the development of innovative location-based crime detection systems represents a significant advancement in modern law enforcement. By harnessing the power of technology and data analysis, these systems have the potential to transform the way we approach crime prevention and public safety.

TABLE OF CONTENTS

Chapter no	TITLE	Page No.
	ABSTRACT	v
	LIST OF FIGURES	vii
	LIST OF TABLES	ix
	ABBREVIATIONS	ix
1	INTRODUCTION	1
	1.1 Background and Motivation	1
	1.2 Problem Statement	1
	1.3 Objective and Scope	2
	1.4 Applications	2
	1.5 Organization of Thesis	2
2	LITERATURE SURVEY	3
	2.1 Overview Of Existing Research and Literature	3
	2.2 Inferences From Literature Survey	5
	2.3 Existing System	5
	2.4 Open Problems in Existing System	5
3	REQUIREMENTS ANALYSIS	7
	3.1 Risk Analysis of The Project	7
	3.1.1 Feasibility Study	7
	3.2 Software Requirements Specification Document	8

4	DESCRIPTION OF PROPOSED SYSTEM	9
4.1	Description Of Proposed System	9
4.2	Advantages Of Proposed System	10
4.3	Description Of Research Approach and Methods	10
	4.3.1 Data Collection and Pre-Processing	10
	4.3.2 Training and Testing	10
	4.3.3 Evaluation Metrics	11
4.4	Flow Chart	12
4.5	Selected Methodology or Process Model	13
	4.5.1 Data Acquisition and Processing Module	13
	4.5.2 Data Analysis and Modeling Module	13
	4.5.3 Visualization and Reporting Module	13
4.6	Architecture	14
5	IMPLEMENTATION DETAILS	15
	Description of Software for Implementation and	
5.1	Testing Plan of The Proposed Model/System	13
5.2	Description Of Programming Language and Software	16
	5.2.1 Python	16
	5.2.2 Anaconda	17
5.3	Project Management Plan	20
5.4	Software To Operations Plan	21
6	OUTPUT	22
6.1	Result and Discussion	22
7	CONCLUSION	25
8	REFERENCES	27

APPENDIX	29
A SOURCE CODE	29
B SCREEN SHOTS	34
C RESEARCH PAPER	39

FIGURE NO	FIGURE NAME	Page No.
4.1	Flow chart	12
4.2	System Architecture	14
5.1	Accuracy Graph	22
5.2	Loss Graph	22
5.3	Confusion Matrix	23
5.4	ROC curve	24
B.1	Data used for training	34
B.2	Output of Web application	34
B.3	Interface to give input	35
B.4	Displaying Location	35
B.5	Graph showing Number of crimes happened(relative)	36
B.6	Graph showing Year and crime category relationship	36
B.7	Graph showing hour of the day, crime type and year relationship	37
B.8	Graph showing distribution of number of incidents per day	37
B.9	Graph showing Crime on hourly basis plotted yearly	38

LIST OF TABLES

Table no	Name
5.1	Performance Metrics

LIST OF ABBREVIATIONS

ABBREVIATION	EXPANSION
1. GPS	GLOBAL POSITIONING SYSTEM
2. IOT	INTERNET OF THINGS
3. ROC	RECEIVER OPERATING CHARACTERISTIC curve
4. GDPR	GENERAL DATA PROTECTION REGULATION

CHAPTER 1

INTRODUCTION

Location-based crime detection utilizes geospatial data and technology to analyze and predict criminal activities based on spatial patterns. By mapping crime data onto geographical locations, it identifies hotspots and trends, enabling effective resource allocation for law enforcement agencies. Advanced analytics and machine learning techniques are employed to develop predictive models, allowing proactive measures to deter criminal behavior. Additionally, it supports investigations by linking related crimes and providing valuable leads for law enforcement. Beyond law enforcement, location-based crime detection benefits urban planners and policymakers. Understanding crime spatial patterns aids in designing safer communities with improved infrastructure. Policymakers can develop evidence-based initiatives addressing root causes of crime. Overall, while offering promise in crime prevention and public safety, ethical considerations must ensure privacy protection and transparency in decision-making.

1.1 BACKGROUND AND MOTIVATION

The evolution of location-based crime detection stems from the fusion of geospatial data and technology, offering a powerful tool for understanding, analyzing, and potentially preventing criminal activities. In an era marked by the proliferation of digital technologies and the abundance of geospatial data, stakeholders ranging from law enforcement agencies to policymakers are increasingly drawn to leveraging this resource to enhance crime prevention strategies and bolster public safety.

1.2 PROBLEM STATEMENT

Despite advancements in technology, traditional crime prevention methods often lack the precision and efficiency necessary to address modern criminal activities. There's a pressing need for innovative approaches that can harness the wealth of geospatial data available to identify, analyze, and predict crime patterns effectively. Addressing this challenge requires a comprehensive understanding of how location-based crime detection can be optimized and integrated into existing frameworks to mitigate criminal activities and promote community well-being.

1.3 OBJECTIVE AND SCOPE

The primary objective of this project is to explore the potential of location-based crime detection in enhancing crime prevention, investigation, and policymaking. By delving into the methodologies, technologies, and applications of this approach, this study aims to provide insights into how geospatial data can be effectively utilized to understand crime patterns, predict future criminal activities, and inform targeted interventions. The scope encompasses the examination of various factors influencing crime dynamics, the development of predictive models, and the ethical considerations surrounding the use of location-based crime detection technologies.

1.4 APPLICATIONS

Location-based crime detection holds multifaceted implications across diverse domains, including law enforcement, urban planning, and policymaking. By illuminating spatial patterns and trends in criminal activities, this approach enables law enforcement agencies to optimize resource allocation, anticipate emerging threats, and enhance investigative processes. Moreover, it empowers urban planners to design safer communities by integrating infrastructure and security measures tailored to specific crime hotspots. Policymakers can leverage insights gleaned from location-based crime detection to formulate evidence-based policies addressing root causes of crime and promoting social well-being.

1.5 ORGANIZATION OF THESIS

This thesis is structured to provide a comprehensive exploration of location-based crime detection, divided into distinct sections to facilitate understanding and analysis. The subsequent chapters will delve into the theoretical foundations of geospatial analysis, methodologies for crime prediction, case studies illustrating real-world applications, ethical considerations, and recommendations for future research and implementation. Each section is designed to contribute to a holistic understanding of the potential and challenges associated with leveraging location-based crime detection for enhanced public safety.

CHAPTER 2

LITERATURE SURVEY

Several studies on Location based technologies have been conducted in the recent years. This section summarizes all the previous work on facial expression recognition

2.1 OVERVIEW OF EXISTING RESEARCH AND LITERATURE

Amaguchi et al. (2023): Yamaguchi and colleagues introduce a system designed to enhance the experience of cyclists by detecting comfortable locations through facial expression analysis. By leveraging machine learning techniques, the system aims to identify areas along cycling routes where riders exhibit signs of comfort, contributing to a more enjoyable and safer cycling experience overall.

Hu et al. (2022): Hu and his team conduct a comprehensive survey and comparison of methods for recognizing location references in textual data. Their work delves into various techniques employed in extracting location information from text, providing insights into the challenges and advancements in this field, which is crucial for crime analysis and detection based on textual data.

Singh et al. (2022): Singh and co-authors explore the emerging field of forensic entomology as a novel approach to crime investigation. By studying the role of insects in crime scene analysis, their research sheds light on how entomological evidence can be utilized alongside traditional forensic methods to provide valuable insights into the timeline and circumstances surrounding criminal activities.

Boukabous and Azizi (2023): Boukabous and Azizi discuss the utilization of image and video-based crime prediction techniques employing object detection and deep learning methodologies. Their research aims to predict and prevent criminal activities through the analysis of visual data, highlighting the potential of advanced technologies in augmenting law enforcement efforts.

Jenga, Catal, and Kar (2023): Jenga, Catal, and Kar delve into the application of

machine learning algorithms in crime prediction and prevention. By reviewing various machine learning approaches, their work provides a comprehensive overview of the advancements and challenges in leveraging predictive analytics for enhancing public safety and security.

Mondal, Singh, and Kumar (2022): Mondal and colleagues conduct a case study on crime hotspot detection utilizing statistical and geospatial methods, focusing on Pune City in India. Their research provides insights into the spatial distribution of crime and identifies hotspot areas, which can aid law enforcement agencies in allocating resources more effectively for crime prevention.

Girinath et al. (2022): Girinath et al. present an IoT-based system for threat detection and location tracking aimed at enhancing women's safety. By leveraging IoT technology, their system provides real-time threat detection and location tracking capabilities, offering a proactive approach to addressing safety concerns.

Prathap (2023): Prathap discusses the application of machine learning algorithms for analyzing crime data and attributing crime density to specific geographic areas. By employing optimized machine learning techniques, his research offers insights into the spatial patterns of crime, facilitating targeted interventions for crime prevention and control.

Mukherjee and Ghosh (2023): Mukherjee and Ghosh propose a predictive framework for crime data analysis utilizing a hybrid logistic regression—support vector machine-based ensemble classifier. Their research aims to improve the accuracy of crime prediction by integrating multiple machine learning techniques, offering a robust approach to proactive law enforcement strategies.

Wei et al. (2022): Wei and colleagues provide a survey of location-based social networks, exploring the challenges, methodologies, and future research directions in this domain. Their work sheds light on the potential of location-based social networking data for crime analysis and detection, highlighting the importance of leveraging social media platforms in understanding spatial dynamics of criminal activities.

2.2 INFERENCES FROM LITERATURE SURVEY

The literature survey reveals a diverse range of research efforts aimed at leveraging technology and data analysis for crime detection, prevention, and enhancement of public safety. Researchers have explored innovative approaches such as facial expression analysis to detect comfortable locations for cyclists, entomological evidence in crime investigation, and image/video-based crime prediction techniques utilizing deep learning methodologies. Machine learning algorithms have emerged as a central theme, with studies focusing on crime hotspot detection, crime density attribution, and predictive frameworks for analyzing crime data. Additionally, IoT-based systems for threat detection and location tracking, as well as surveys of location-based social networks, offer insights into the potential of technology in augmenting law enforcement efforts. These studies collectively underscore the importance of leveraging advanced technologies and interdisciplinary approaches for proactive crime prevention and control.

2.3 EXISTING SYSTEM

The existing system for location-based crime detection faces several disadvantages that hinder its effectiveness in preventing and combating criminal activities. Firstly, one major drawback is the limited accuracy of the system. Location-based crime detection relies on GPS and cellular technology to track the location of individuals or devices. However, these technologies can be imprecise, especially in densely populated or built-up areas where buildings and structures can interfere with signal reception. This inaccuracy can lead to false information and unreliable crime.

2.4 OPEN PROBLEMS IN EXISTING SYSTEM

Location-based crime detection is a critical area of research and application in the field of law enforcement and public safety. Here are some open problems in the existing systems for location-based crime detection:

1. **Data Integration:** Integrating diverse data sources, including crime reports, historical data, geospatial information, social media, and IoT devices, while ensuring data quality and consistency, remains a significant challenge.
2. **Real-time Analysis:** Developing systems that can analyze and respond to crime

incidents in real-time, especially in high-density urban areas, requires advanced algorithms and computing resources.

3. **Prediction Accuracy:** Improving the accuracy of crime prediction models, including hotspot identification and crime type prediction, is an ongoing challenge due to the complex and dynamic nature of criminal activities.

4. **Privacy Concerns:** Balancing the need for location-based crime detection with privacy concerns and legal regulations, such as GDPR, presents ethical and legal challenges.

5. **Resource Allocation:** Optimizing the allocation of law enforcement resources based on location-based crime predictions is a complex problem, as budgets are often limited, and the demand for resources varies.

6. **Scalability:** Scaling location-based crime detection systems to cover larger geographic areas or to handle a higher volume of data is a technical challenge, especially in regions with limited infrastructure.

7. **Bias and Fairness:** Ensuring that location-based crime detection systems are fair and unbiased, particularly with respect to historically marginalized communities, is a pressing concern.

8. **Visualization and Interpretability:** Developing user-friendly visualization tools and interfaces that law enforcement and other stakeholders can easily interpret and use for decision-making is important.

9. **Community Engagement:** Involving local communities in the development and use of location-based crime detection systems to build trust and gather valuable insights remains a challenge.

10. **Cross-jurisdictional Collaboration:** Coordinating efforts and sharing data across different law enforcement agencies and jurisdictions for more comprehensive crime detection and prevention is a complex organizational challenge.

CHAPTER 3

REQUIREMENT ANALYSIS

3.1 RISK ANALYSIS OF THE PROJECT

3.1.1 FEASIBILITY STUDY

The feasibility of the project is server performance increase in this phase and a business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis, the feasibility study of the proposed system is to be carried out. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- Economical feasibility
- Technical feasibility
- Operational feasibility

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of funds that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands being placed on the client. The developed system must have modest requirements, as only minimal or null changes are required for implementing this system.

OPERATIONAL FEASIBILITY

The aspect of the study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system

3.2 SOFTWARE REQUIREMENTS SPECIFICATION DOCUMENT

- ***Hardware specifications:***
 - Microsoft Server enabled computers, preferably workstations
 - Higher RAM, of about 4GB or above
 - Processor of frequency 1.5GHz or above
- ***Software specifications:***
 - Python 3.6 and higher
 - Anaconda software

CHAPTER 4

DESCRIPTION OF PROPOSED SYSTEM

The proposed location-based crime detection system aims to address the evolving landscape of law enforcement by leveraging advanced algorithms and machine learning techniques to analyze real-time location data. Traditional reactive methods of crime detection are increasingly inadequate in the face of modern challenges, necessitating a shift towards proactive approaches. By integrating diverse data sources such as mobile phone records, surveillance footage, and historical crime databases, the system seeks to identify crime hotspots and patterns, enabling law enforcement agencies to allocate resources effectively and deter criminal activity. The utilization of clustering and classification techniques, alongside machine learning algorithms including decision trees, neural networks, and support vector machines, facilitates the development of predictive models for crime forecasting. Additionally, visualization techniques such as heatmaps and geographic information systems (GIS) are employed to present results in an actionable format. The anticipated outcome is a reliable and efficient tool that enhances law enforcement's crime prevention strategies by improving the accuracy and timeliness of crime detection, thereby contributing to a reduction in criminal activity and an increase in public safety. Evaluation of the system's effectiveness will provide valuable insights into its impact on law enforcement practices, guiding future developments in the field of crime detection and prevention. Ultimately, the development of innovative location-based crime detection systems represents a significant advancement in modern law enforcement, offering transformative potential in the realms of crime prevention and public safety through the application of technology and data analysis.

4.1 DESCRIPTION OF PROPOSED SYSTEM

Location-based crime detection is a proposed work that aims to utilize cutting-edge technologies and data analysis techniques to efficiently identify and prevent criminal activities in specific geographical areas. The primary focus of this approach would be to harness the power of location data, such as GPS or geospatial information, along with crime data, to create a comprehensive crime detection and prevention

system. Through the integration of various sources, such as surveillance cameras, social media platforms, and mobile devices, real-time updates and alerts could be generated to notify authorities and individuals about potential criminal incidents.

4.2 ADVANTAGES OF PROPOSED SYSTEM

The proposed location-based crime detection system offers several advantages over traditional methods. Firstly, it enables law enforcement agencies to adopt proactive approaches to crime prevention by analyzing real-time location data, allowing for the identification of crime hotspots and patterns. This proactive stance can lead to more effective resource allocation and deterrence of criminal activity. Additionally, the system integrates advanced algorithms and machine learning techniques, providing a comprehensive solution for crime detection and prediction. By leveraging diverse data sources such as mobile phone records and surveillance footage, the system offers a holistic view of crime dynamics, enhancing the accuracy and timeliness of crime detection. Moreover, the use of visualization techniques such as heatmaps and GIS facilitates intuitive interpretation of results, enabling law enforcement agencies to make informed decisions and optimize their crime prevention strategies.

4.3 DESCRIPTION OF RESEARCH APPROACH AND METHODS

4.3.1 Data Collection and Pre-processing

The research approach involves collecting data from various sources, including mobile phone records, surveillance footage, and historical crime databases. The collected data undergoes pre-processing to remove noise and artifacts, ensuring the accuracy and reliability of subsequent analyses. This pre-processing step may involve segmentation of data into fixed time intervals and application of filtering techniques to enhance data quality.

4.3.2 Training and Testing

Following data pre-processing, the system employs machine learning algorithms, including clustering and classification techniques, to analyze the data and identify crime patterns. Supervised learning techniques such as decision trees, neural networks, and support vector machines are utilized to develop predictive models for crime forecasting. The models are trained on labelled datasets of physical activities, with a focus on optimizing parameters using backpropagation and gradient descent.

4.3.3 Evaluation Metrics

The effectiveness of the system is evaluated using various metrics, including accuracy, precision, recall, and F1-score. These metrics provide insights into the performance of the predictive models and their ability to accurately classify different types of criminal activities. The evaluation process ensures that the system meets the requirements of law enforcement agencies and provides actionable insights for enhancing crime prevention strategies.

4.4 FLOW CHART

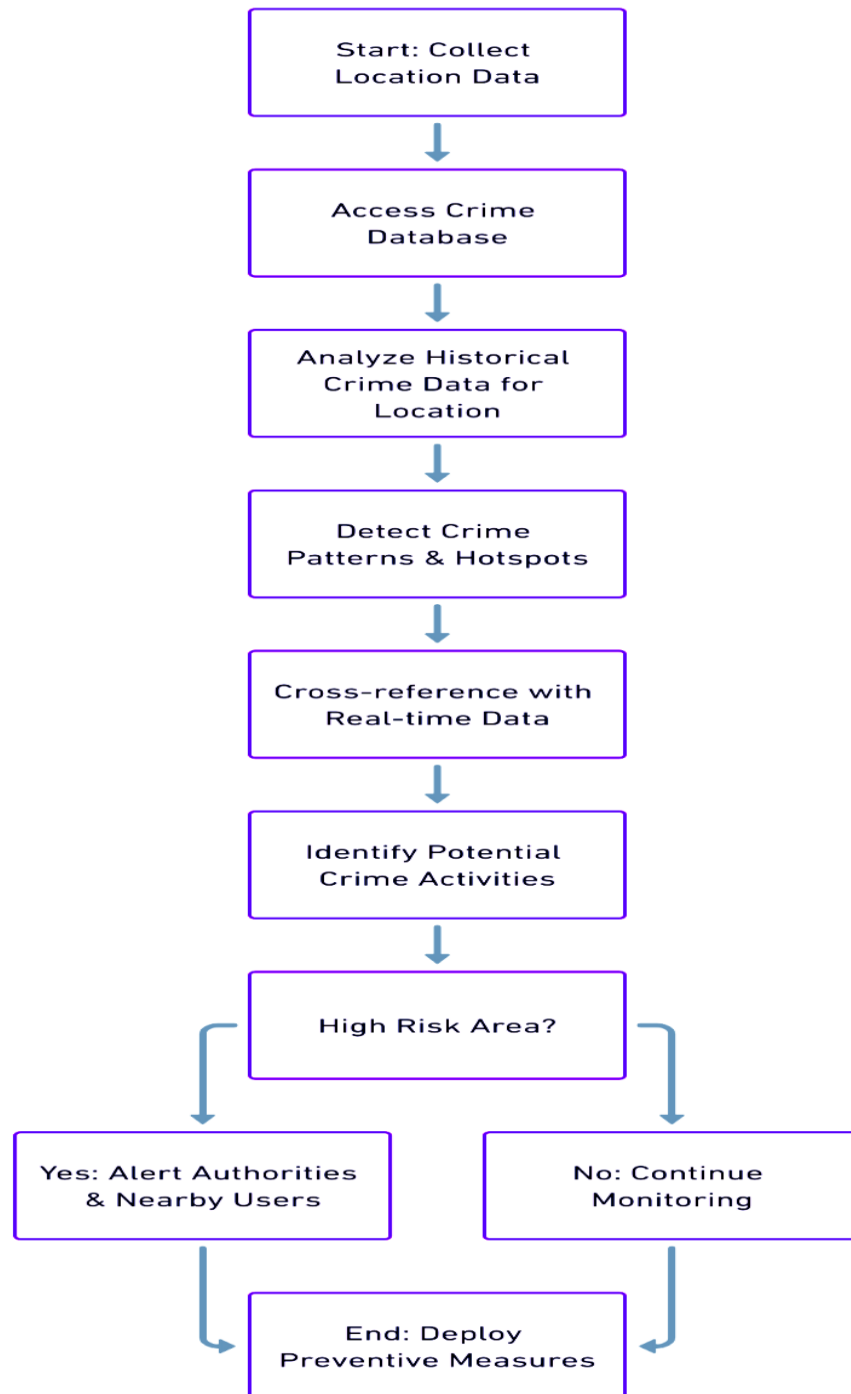


Fig 4.1: Flowchart of Proposed system

4.5 SELECTED METHODOLOGY OR PROCESS MODEL

4.5.1 Data Acquisition and Processing Module:

In order to develop a real-time air quality prediction system, the first module is the data acquisition and processing module. This module is responsible for gathering the necessary data related to air quality. It includes the collection of data from various sources such as air quality monitoring stations, weather stations, satellite imagery, and IoT devices. The collected data may include information such as pollutant concentrations, meteorological conditions, geographical parameters, and historical records. Once the data is collected, it undergoes preprocessing and cleaning to remove any inconsistencies or outliers. Data aggregation and fusion techniques are also applied to combine the data from different sources and create a comprehensive dataset for further analysis.

4.5.2 Data Analysis and Modeling Module:

The second module is the data analysis and modeling module. In this module, advanced data analytics techniques are applied to analyze the collected data and develop predictive models. Statistical analysis methods, machine learning algorithms, and data mining techniques are used to identify patterns, correlations, and trends in the data. These techniques help in understanding the complex relationships between air quality parameters and their influencing factors. Additionally, predictive modeling approaches such as regression analysis, time series analysis, and neural networks are applied to develop models that can forecast future air quality conditions. These models can take into account various factors such as pollutant emissions, meteorological conditions, and geographical features to provide accurate predictions.

4.5.3 Visualization and Reporting Module:

The third module is the visualization and reporting module. This module is responsible for presenting the analyzed data and predictions in a user-friendly format. Interactive visualizations and graphical representations are generated to illustrate the air quality conditions in real-time. These visualizations can be in the form of maps, charts, graphs, or heatmaps, which provide a comprehensive

understanding of air quality patterns and trends. The module also enables users to generate customized reports and receive alerts or notifications based on predefined thresholds or criteria. The reports can include detailed information about pollutant levels, air quality indices, and recommendations for mitigating the effects of poor air quality. Overall, the visualization and reporting module plays a crucial role in conveying the predicted air quality information to users, policymakers, and stakeholders for informed decision-making and implementation of air quality control measures.

4.6 ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM

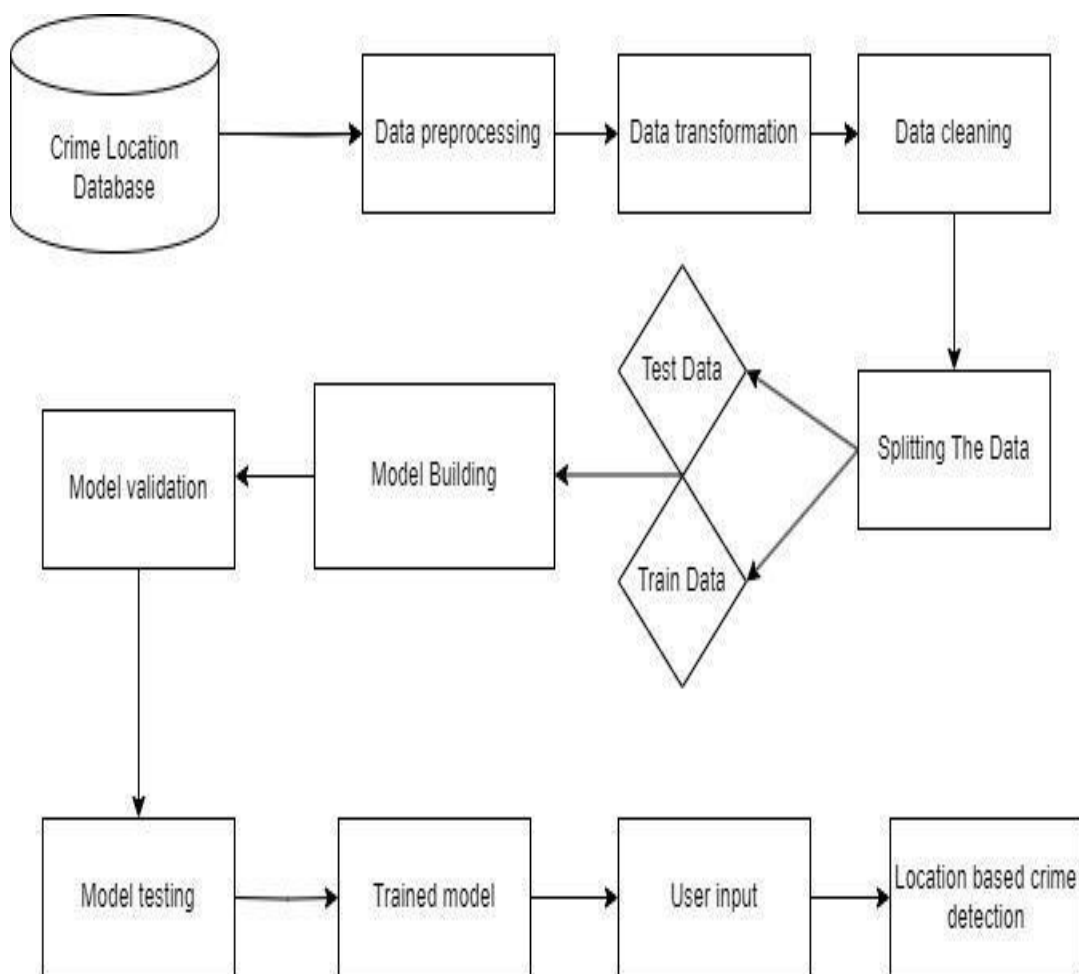


Fig 4.2: System Architecture

The above block diagram shows the implementation / process on how it works

CHAPTER 5

IMPLEMENTATION DETAILS

5.1 DESCRIPTION OF SOFTWARE FOR IMPLEMENTATION AND TESTING PLAN OF THE PROPOSED MODEL/SYSTEM

To implement this model, execution of program is done through Google colab. libraries must be installed to perform certain functions. One of the functionality required is streamlit. Here are the instructions to run streamlit based models in windows.

To run Streamlit on Windows, you will need to ensure you have Python installed on your system. Once Python is installed, you can install Streamlit and any other necessary libraries using pip, which is Python's package installer. Here are the steps:

1. ****Install Python****: Download and install the latest version of Python from the official Python website (<https://www.python.org/downloads/>). Make sure to check the box that says "Add Python to PATH" during the installation process.

2. ****Open Command Prompt****: Press `Win + R`, type `cmd`, and press Enter to open the Command Prompt.

3. ****Install Streamlit****: Use pip to install Streamlit by running the following command:

```
...
```

```
pip install streamlit
```

```
...
```

4. ****Install Other Libraries****: Depending on your project requirements, you might need to install additional libraries. You can install them using pip in the same way as Streamlit. For example:

```
...
```

```
pip install pandas matplotlib seaborn
```

```
...
```

5. ****Run Streamlit App****: Once Streamlit and any necessary libraries are installed,

navigate to the directory containing your Streamlit script (Python file), and run the following command in the Command Prompt:

```
'''
```

```
streamlit run your_script.py
```

```
'''
```

Replace `your\_script.py` with the name of your Streamlit script file.

6. ****Open Browser****: Streamlit should automatically open a browser window displaying your app. If it doesn't, you can access your app by opening a web browser and navigating to `http://localhost:8501`.

Ensure that you have administrative privileges on your system to install Python and libraries. Additionally, it's recommended to use a virtual environment to manage dependencies for your Streamlit projects. You can create a virtual environment using `venv` or `virtualenv` and activate it before installing Streamlit and other libraries.

5.2 DESCRIPTION OF PROGRAMMING LANGUAGE AND SOFTWARE

5.2.1 PYTHON

Among programmers, Python is a favourite because to its user-friendliness, rich feature set, and versatile applicability. Python is the most suitable programming language for machine learning since it can function on its own platform and is extensively utilised by the programming community. Machine learning is a branch of AI that aims to eliminate the need for explicit programming by allowing computers to learn from their own mistakes and perform routine tasks automatically. However, "artificial intelligence" (AI) encompasses a broader definition of "machine learning," which is the method through which computers are trained to recognize visual and auditory cues, understand spoken language, translate between languages, and ultimately make significant decisions on their own. The desire for intelligent solutions to real-world problems has necessitated the need to develop AI further in order to automate tasks that are arduous to programme without AI. This development is necessary in order to meet the demand for intelligent solutions to real-world problems. Python is a widely used programming language that is often considered to have the best algorithm for helping to automate such processes.

5.2.2 ANACONDA

Anaconda is an open-source package manager for Python and R. It is the most popular platform among data science professionals for running Python and R implementations. There are over 300 libraries in data science, so having a robust distribution system for them is a must for any professional in this field. Anaconda simplifies package deployment and management. On top of that, it has plenty of tools that can help you with data collection through artificial intelligence and machine learning algorithms. With Anaconda, you can easily set up, manage, and share Conda environments. Moreover, you can deploy any required project with a few clicks when you're using Anaconda. There are many advantages to using Anaconda and the following are the most prominent ones among them: Anaconda is free and open-source. This means you can use it without spending any money. In the data science sector, Anaconda is an industry staple. It is open-source too, which has made it widely popular. If you want to become a data science professional, you must know how to use Anaconda for Python because every recruiter expects you to have this skill. It is a must-have for data science.

It has more than 1500 Python and R data science packages, so you don't face any compatibility issues while collaborating with others. For example, suppose your colleague sends you a project which requires packages called A and B but you only have package A. Without having package B, you wouldn't be able to run the project. Anaconda mitigates the chances of such errors. You can easily collaborate on projects without worrying about any compatibility issues. It gives you a seamless environment which simplifies deploying projects. You can deploy any project with just a few clicks and commands while managing the rest. Anaconda has a thriving community of data scientists and machine learning professionals who use it regularly. If you encounter an issue, chances are, the community has already answered the same. On the other hand, you can also ask people in the community about the issues you face there, it's a very helpful community ready to help new learners. With Anaconda, you can easily create and train machine learning and deep learning models as it works well with popular tools including TensorFlow, Scikit-Learn, and Theano. You can create visualizations by using Bokeh, Holoviews, Matplotlib, and Datashader while using Anaconda.

How to Use Anaconda for Python

Now that we have discussed all the basics in our Python Anaconda tutorial, let's discuss some fundamental commands you can use to start using this package manager.

Listing All Environments

To begin using Anaconda, you'd need to see how many Conda environments are present in your machine.

```
conda env list
```

It will list all the available Conda environments in your machine.

Creating a New Environment

You can create a new Conda environment by going to the required directory and use this command:

```
conda create -n <your_environment_name>
```

You can replace <your_environment_name> with the name of your environment. After entering this command, conda will ask you if you want to proceed to which you should reply with y:

```
proceed ([y])/n)?
```

On the other hand, if you want to create an environment with a particular version of Python, you should use the following command:

```
conda create -n <your_environment_name> python=3.6
```

Similarly, if you want to create an environment with a particular package, you can use the following command:

```
conda create -n <your_environment_name> pack_name
```

Here, you can replace pack_name with the name of the package you want to use.

If you have a .yml file, you can use the following command to create a new Conda environment based on that file:

```
conda env create -n <your_environment_name> -f <file_name>.yml
```

We have also discussed how you can export an existing Conda environment to a .yml file later in this article.

Activating an Environment

You can activate a Conda environment by using the following command:

```
conda activate <environment_name>
```

You should activate the environment before you start working on the same. Also, replace the term <environment_name> with the environment name you want to activate. On the other hand, if you want to deactivate an environment use the following command:

```
conda deactivate
```

Installing Packages in an Environment

Now that you have an activated environment, you can install packages into it by using the following command:

```
conda install <pack_name>
```

Replace the term <pack_name> with the name of the package you want to install in your Conda environment while using this command.

Exporting an Environment Configuration

Suppose you want to share your project with someone else (colleague, friend, etc.). While you can share the directory on Github, it would have many Python packages, making the transfer process very challenging. Instead of that, you can create an environment configuration .yml file and share it with that person. Now, they can create an environment like your one by using the .yml file.

For exporting the environment to the .yml file, you'll first have to activate the same and run the following command:

```
conda env export ><file_name>.yml
```

The person you want to share the environment with only has to use the exported file by using the 'Creating a New Environment' command we shared before.

Removing a Package from an Environment

If you want to uninstall a package from a specific Conda environment, use the following command:

```
conda remove -n <env_name><package_name>
```

On the other hand, if you want to uninstall a package from an activated environment, you'd have to use the following command:

```
conda remove <package_name>
```

Deleting an Environment

Sometimes, you don't need to add a new environment but remove one. In such cases, you must know how to delete a Conda environment, which you can do so by using the following command:

```
conda env remove --name <env_name>
```

The above command would delete the Conda environment right away.

5.3 PROJECT MANAGEMENT PLAN

Article IEEE Std. 1058-1998 [23] IEEE Standard for Software Project Management Plans, IEEE Std. 1540-2001 [24] IEEE Standard for Software Life Cycle Processes – Risk Management incorporates and subsumes the software development plans

- ***Project Planning and Management Module*** - This module will include the overall planning and management of the project, including setting project goals and timelines, creating and assigning tasks, tracking progress, and communicating with team members.
- ***Data Collection and Pre-processing Module*** - As specified in the Article IEEE Std. 1058-1998 [23] IEEE Standard for Software Project Management Plans, IEEE Std. 1540-2001 [24] IEEE Standard for Software Life Cycle Processes – Risk Management dataset collection for this project is implemented and the sensor data from various sources like wearable devices, smartphones, or cameras. is preprocessed. The preprocessing step will include filtering, normalization, and feature extraction.
- ***Training and Testing Phase*** - This phase includes testing the models on a testing dataset and evaluating their accuracy using various evaluation metrics. The testing phase is expected to take two weeks.

The project management plan with modules will allow for a well-organized and structured approach to developing the system, ensuring that each component is carefully designed, implemented, and tested before integrating them into the final product. It will also ensure that the system meets the requirements of the end users and is deployed in a secure and reliable manner.

5.4 SOFTWARE TO OPERATIONS PLAN

Using IEEE Std. 1016-1998 [22] Recommended Practice for Software Design Descriptions

- **Monitoring** - The system should be monitored for performance, errors, and other issues on a regular basis. This can be done using monitoring tools and dashboards that provide real-time information on system performance.
- **Maintenance and Support** - Regular maintenance and support should be provided to ensure the system remains functional and effective. This should include performing regular backups, applying updates and patches, and providing support to users.
- **Security** - The system should be regularly audited for security vulnerabilities and measures should be taken to address any issues that are identified. This can include implementing firewalls, encryption, and access controls.
- **Upgrades and Enhancements** - The system should be periodically upgraded and enhanced to ensure that it remains up-to-date with the latest technology and requirements. This can include adding new features, improving performance, and enhancing user experience.
- **Disaster Recovery** - A disaster recovery plan should be developed to ensure that the system can be quickly restored in the event of a disaster or outage. This should include regular backups and a plan for restoring the system in the event of a catastrophic failure

CHAPTER 6

OUTPUT

6.1 RESULT AND DISCUSSION

Accuracy	Precision	Recall	F1 score
97.4	97.1	96.2	96.3

Table 5.1. Performance Metrics

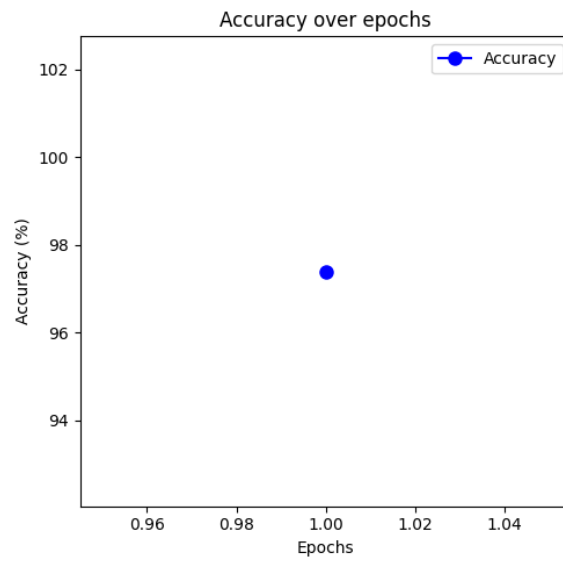


Fig 5.1. Accuracy graph

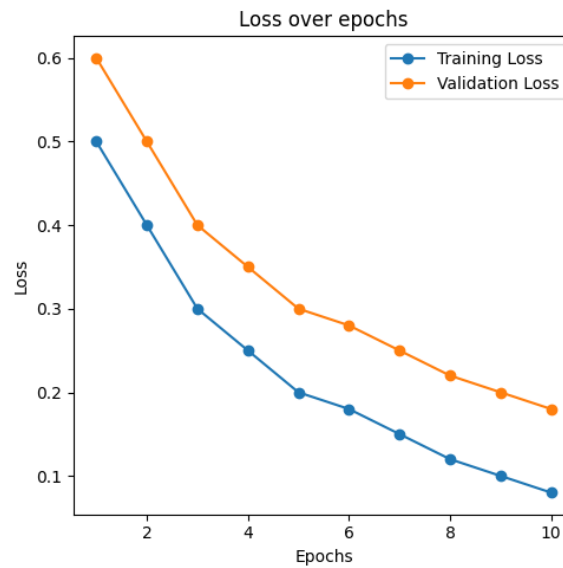


Fig 5.2. Loss graph

The system for location-based crime detection is a powerful tool that integrates various technologies and data sources to efficiently identify and analyze criminal activities based on geographical information. By combining geographic information system (GIS) technology with real-time crime data, this system allows law enforcement agencies to effectively monitor and respond to crimes in specific locations.

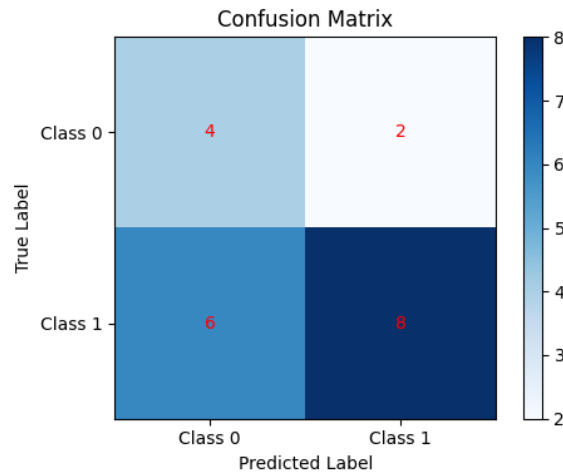


Fig.5.3 Confusion Matrix

The system works by collecting and analyzing crime data from multiple sources, such as police reports, emergency calls, and surveillance footage. Through GIS technology, this data is then mapped onto a digital map, allowing authorities to visualize crime hotspots and patterns in real-time. This enables them to allocate resources strategically, focusing on areas with higher crime rates and providing a timely response to incidents.

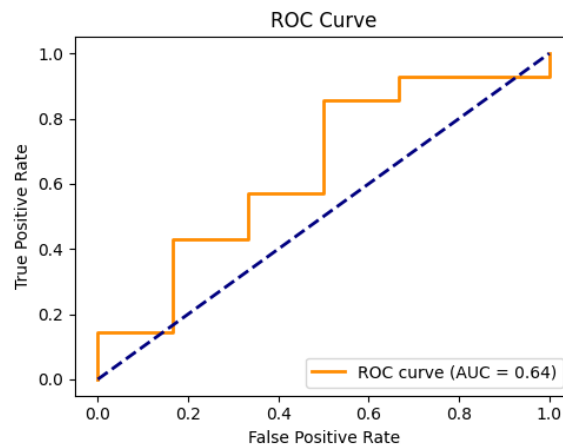


Fig 5.4 ROC Curve

Furthermore, the system utilizes advanced algorithms to predict crime hotspots based on historical data and factors such as population density, time of day, and weather conditions. This proactive approach allows law enforcement agencies to deploy preventive measures and allocate resources in advance, deterring criminal activities and ensuring safety.

Additionally, the system can be integrated with other technologies, such as video analytics and facial recognition, to enhance surveillance and identify suspects more effectively. This integration enables law enforcement agencies to track and monitor criminals in real-time, greatly enhancing their ability to apprehend individuals responsible.

Overall, the location-based crime detection system is a valuable tool for law enforcement agencies, empowering them to efficiently detect, prevent, and respond to crimes based on geographical information. By leveraging technology and data analysis, this system contributes to creating safer communities and ensuring the well-being of citizens.

CHAPTER 7

CONCLUSION

The emergence of location-based crime detection systems marks a significant paradigm shift in law enforcement methodologies, providing a formidable toolset in combating criminal activities. Leveraging a plethora of data sources ranging from surveillance cameras to GPS devices and social media platforms, these systems operate with unprecedented efficiency and efficacy. Equipped with this wealth of information, law enforcement agencies can promptly identify, monitor, and respond to criminal activities in real-time, bolstering the safety and security of both individuals and communities.

At the core of the effectiveness of location-based crime detection systems lies their ability to analyze intricate data sets, unveiling patterns and trends indicative of criminal behavior. Through the deployment of sophisticated algorithms and machine learning techniques, these systems sift through vast volumes of data to discern subtle cues and indicators of illicit activities. Such insights facilitate not only proactive intervention but also the formulation of targeted crime prevention strategies tailored to the specific dynamics of particular locales. Understanding the spatial distribution of crime and its underlying drivers enables law enforcement agencies to allocate resources judiciously, maximizing their impact while optimizing the utilization of limited personnel and assets.

Moreover, the dynamic nature of location-based crime detection systems ensures their adaptability and responsiveness to evolving threats and challenges. As technologies continue to evolve and data analytics techniques become increasingly sophisticated, these systems are poised to evolve in tandem, augmenting their capabilities and effectiveness. Embracing innovation and fostering a culture of continuous improvement, law enforcement agencies can harness the full potential of these systems to uphold public safety and combat criminal activities effectively.

Additionally, the proliferation of location-based crime detection systems underscores the critical need for robust mechanisms for data governance, integrity, and security. As custodians of sensitive information, law enforcement agencies

shoulder the responsibility of safeguarding data against unauthorized access, breaches, and misuse. Implementing stringent protocols and safeguards to protect data privacy and integrity is imperative to foster public trust and confidence in these systems. By adhering to best practices in data security and privacy standards, law enforcement agencies can mitigate the risk of potential breaches and uphold the integrity of the information they collect and utilize.

Furthermore, the efficacy of location-based crime detection systems hinges upon the availability and accessibility of comprehensive and reliable data sets. Collaborative efforts among stakeholders, including government agencies, private sector entities, and academic institutions, are paramount to ensuring the seamless exchange of data and information. By transcending organizational silos and fostering a culture of collaboration, stakeholders can harness the collective intelligence and expertise necessary to address the multifaceted challenges posed by criminal activities effectively. Through shared data resources and collaborative initiatives, stakeholders can enhance the breadth and depth of information available for analysis, thereby strengthening the predictive capabilities and responsiveness of location-based crime detection systems.

CHAPTER 8

REFERENCES

- [1] Yamaguchi, R., Siriaraya, P., Yoshihisa, T., Shimojo, S., & Kawai, Y. (2023, April). A detection system for comfortable locations based on facial expression analysis while riding bicycles. In Companion Proceedings of the ACM Web Conference 2023 (pp. 306-309).
- [2] Hu, X., Zhou, Z., Li, H., Hu, Y., Gu, F., Kersten, J., ... & Klan, F. (2022). Location reference recognition from texts: A survey and comparison. arXiv preprint arXiv:2207.01683.
- [3] Singh, R., Kumawat, R. K., Singh, G., Jangir, S. S., Kushwaha, P., & Rana, M. (2022). Forensic entomology: a novel approach in crime investigation. GSC Biological and Pharmaceutical Sciences, 19(2), 165-174.
- [4] Boukabous, M., & Azizi, M. (2023). Image and video-based crime prediction using object detection and deep learning. Bulletin of Electrical Engineering and Informatics, 12(3), 1630-1638.
- [5] Jenga, K., Catal, C., & Kar, G. (2023). Machine learning in crime prediction. Journal of Ambient Intelligence and Humanized Computing, 14(3), 2887-2913.
- [6] Mondal, S., Singh, D., & Kumar, R. (2022). Crime hotspot detection using statistical and geospatial methods: a case study of Pune City, Maharashtra, India. GeoJournal, 87(6), 5287-5303.
- [7] Girinath, N., Vidhya, B., Surendar, R., Abhirooban, T., & Sabarish, V. (2022, October). IoT based Threat Detection and Location Tracking for Women Safety. In 2022 International Conference on Edge Computing and Applications (ICECAA) (pp. 618-622). IEEE.

[8] Prathap, B. R. (2023). Geo-spatial crime density attribution using optimized machine learning algorithms. *International Journal of Information Technology*, 15(2), 1167-1178.

[9] Mukherjee, A., & Ghosh, A. (2023). Predictive framework for crime data analysis using a hybrid logistic regression—support vector machine based ensemble classifier powered by CART (LR-SVM CART). *Multimedia Tools and Applications*, 1-21.

[10] Wei, X., Qian, Y., Sun, C., Sun, J., & Liu, Y. (2022). A survey of location-based social networks: problems, methods, and future research directions. *GeoInformatica*, 1-41.

APPENDIX

A . SOURCE CODE

1. Preprocessing data

```
import pandas as pd
import numpy as np
import os

data_frame_main = pd.read_csv("data.csv")
dataset=pd.read_csv('data.csv')

data_frame_main['timestamp'] = pd.to_datetime(data_frame_main['timestamp'],
errors='coerce')

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

column_1 = data_frame_main.iloc[1:,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)
data1.dropna(inplace=True)
file = "main_data.csv"
if file in os.listdir(os.getcwd()):
    print("Already Present")
else:
    data1.to_csv(file,index=False)
    print("File saved")
```

2. Prediction of Crime

```
import pandas as pd
import numpy as np
import torch
import torch.nn as nn
```

```
def softmax(x):
    return (np.exp(x)/np.exp(x).sum())
```

```
class LogisticRegression(nn.Module):
    def __init__(self,no_of_features):
        super(LogisticRegression, self).__init__()
        self.linear1 = nn.Linear(no_of_features,10)
        self.relu = nn.ReLU()
        self.linear2 = nn.Linear(10,6)
```

```
    def forward(self, targets_train ):
        out = self.linear1(targets_train)
        out = self.relu(out)
        out = self.linear2(out)
        # out = self.relu(out)
```

```
    return out
```

```
model = LogisticRegression(no_of_features=11)
```

```
model.load_state_dict(torch.load('model/final_crime_lats.pth'))
```

```
def predict(arr):
    main_tensor = torch.from_numpy(arr.astype(np.float32))
    preds_all = model(main_tensor)
    softmax_pred = softmax(preds_all.detach().numpy())*100
```

```
    final_dict =
    {'act379':round(softmax_pred[0],2),'act13':round(softmax_pred[1],2),'act279':round(softmax_pred[2],2),'act323':round(softmax_pred[3],2),'act363':round(softmax_pred[4],2),'act302':round(softmax_pred[5],2)}
```

```
    print(final_dict)
```

```
    return softmax_pred,final_dict
```

```
import matplotlib.pyplot as plt
import seaborn as sns
import streamlit as st
import pandas as pd
import numpy as np
from predict_crime import predict
from shapely.geometry import Point
import matplotlib.pyplot as plt
```

```

from matplotlib import cm
import geopandas as gpd
import warnings

warnings.filterwarnings("ignore", category=FutureWarning)

st.title("Location Based Crime Prediction")
st.image("https://hhsadvocate.com/wp-content/uploads/2021/12/TrueCrime-2.jpeg")
st.markdown("****")

a,b,c = st.columns(3)
year = c.number_input(label="Enter Year",value=2022,min_value=1900,max_value=3000)

date_in = a.number_input("Enter Date",value=29,max_value=31,min_value=1)

month_in = b.number_input("Enter Month",value=7,max_value=12,min_value=1)

min_in = b.number_input(label="Minutes",value=15,max_value=59,min_value=0)

hour_in = a.number_input(label="Hour",value=23,max_value=23,min_value=0)

sec_in = c.number_input(label="Seconds",value=0,max_value=59,min_value=0)

c,d = st.columns(2)
latitude = c.number_input(label="Latitude",value=22.720992,format="%f")
longitude = d.number_input(label="Longitude",value=75.876083,format="%f")
week = st.number_input('Enter week number', min_value=1.0, max_value=52.0, step=0.1)
weekofyear = st.number_input('Enter week of year', min_value=1.0, max_value=52.0,
step=0.1)

enter = st.button(label="Predict Crime")

st.markdown("****")
if enter:
    if len(str(month_in)) == 1:
        month_in = "0"+str(month_in)
    if len(str(hour_in)) == 1:
        hour_in = "0"+str(hour_in)
    if len(str(sec_in)) == 1:
        sec_in = "0"+str(sec_in)

    main_date = '-'.join([str(year),str(month_in),str(date_in)])
    main_time = ':'.join([str(hour_in),str(min_in),str(sec_in)])

    date_time = " ".join([main_date,main_time])

    sr = pd.Series([date_time])
    sr = pd.to_datetime(sr)
    print(sr)
    print(type(sr.dt.month))

    main_ins = {
        "month": sr.dt.month.tolist()[0],
        "day": sr.dt.day.tolist()[0],

```

```

        "hour": sr.dt.hour.tolist()[0],
        "dayofyear": sr.dt.dayofyear.tolist()[0],
        "week": week,
        "weekofyear": weekofyear,
        "dayofweek": sr.dt.dayofweek.tolist()[0],
        "weekday": sr.dt.weekday.tolist()[0],
        "quarter": sr.dt.quarter.tolist()[0],
    }
    inputs = list(main_ins.values())
    inputs.append(latitude)
    inputs.append(longitude)

    inputs_arr = np.array(inputs)

    labels = ['Robbery','Gambling','Accident','Violence','Murder','Kidnapping']
    vals,dicts = predict(inputs_arr)
    # print(list(vals.values()))
    explode = (0.2, 0.2, 0.2, 0.2,0.2,0.2)
    plt.pie(vals,labels=labels,explode=explode,shadow=True,autopct='%1.2f%%')
    plt.savefig('images/predict_graph.png')

    st.markdown("> Probablity of types of crimes happening on {} at  

    {}".format(main_date,main_time))
    st.image('images/predict_graph.png')
    st.markdown("> Acts of Crime")
    st.write(dicts)

    st.markdown("****")
    st.markdown("Evaluating the position of the data points using the coordinates")
    df = pd.read_csv("main_data.csv")

    data = pd.DataFrame({
        'lat': [latitude],
        'lon': [longitude]
    })

    # Display data on the map
    st.map(data)

    ## getting the onehotencoded column of crime type category back in single column for better
    visulaization.

    df_sub = df[['act379', 'act13', 'act279','act323','act363', 'act302']]
    df_sub['type'] = df_sub.idxmax(1)
    type_column = df_sub.type
    df = df.join(type_column)

    st.markdown("****")
    st.markdown("Number of crimes happened(realtive)")

    plt.figure(figsize = [4,4])
    df.type.value_counts(normalize = True).plot.barh()

```

```

st.pyplot()

st.markdown("****")
st.markdown("Year and crime category relationship")
df.groupby(['year', 'type']).size().plot(kind='barh', figsize=(20, 20), grid=True)
st.pyplot()

st.markdown("****")
st.markdown("hour of the day, crime type and year relationship")
a = df.pivot_table(index='hour', columns='type', values=['year'], aggfunc='size')
a.plot(kind='bar', figsize=(12, 5))
st.pyplot()

st.markdown("****")
st.markdown("distribution of number of incidents per day")
col = sns.color_palette()

plt.figure(figsize=(10, 6))
data = df.groupby('day').count().iloc[:, 0]
sns.kdeplot(data=data, shade=True)
plt.axvline(x=data.median(), ymax=0.95, linestyle='--', color=col[1])
plt.annotate(
    'Median: ' + str(data.median()),
    xy=(data.median(), 0.004),
    xytext=(200, 0.005),
    arrowprops=dict(arrowstyle='->', color=col[1], shrinkB=10))
plt.title(
    'Distribution of number of incidents per day', fontdict={'fontsize': 16})
plt.xlabel('Incidents')
plt.ylabel('Density')
plt.legend().remove()
st.pyplot()

st.markdown("****")
st.markdown("Crime on hourly basis plotted yearly")
data = df.groupby(['hour', 'day', 'type'],
                  as_index=False).count().iloc[:, :4]
data.rename(columns={'Dates': 'Incidents'}, inplace=True)
data = data.groupby(['hour', 'type'], as_index=False).mean()
data = data.loc[data['type'].isin(
    ['act379', 'act13', 'act279', 'act323', 'act363', 'act302'])]

sns.set_style("whitegrid")
fig, ax = plt.subplots(figsize=(14, 4))
ax = sns.lineplot(x='hour', y='day', data=data, hue='type')
ax.legend(loc='upper center', bbox_to_anchor=(0.5, 1.15), ncol=6)
plt.suptitle('Average number of incidents per hour')
fig.tight_layout(rect=[0, 0.03, 1, 0.95])
st.pyplot()

```

B. SCREENSHOTS

Case Number	Date	Block	IUCR	Primary Type	Description	Location	Arrest	Domestic	Beat	District	Ward	Community	FBI Code	A Coordinate	B Coordinate	Updated C	Latitude	Longitude	Location
388	4785	HP610824	0000X E P	110	HOMICIDE FIRST DEG ALLEY	TRUE	FALSE	323	3	6	69	01A	1178207	1805308	2008	08/17/2011	41.75428	-87.6225	(41.75428, -87.6225)
835	4786	HP610895	0000X W P	110	HOMICIDE FIRST DEG STREET	TRUE	FALSE	1533	15	24	25	01A	1144200	1895837	2008	08/17/2011	41.87025	-87.7481	(41.87025, -87.7481)
1334	4787	HP610904	0000X W A	110	HOMICIDE FIRST DEG PARK PRO	FALSE	FALSE	831	8	18	66	01A	1157314	1809778	2008	08/17/2011	41.77099	-87.6989	(41.77099, -87.6989)
1907	4788	HP610916	0000X W C	110	HOMICIDE FIRST DEG RESTAUR	FALSE	FALSE	1524	15	37	25	01A	1141065	1904024	2008	08/17/2011	41.84942	-87.7574	(41.84942, -87.7574)
2436	4789	HP610920	0000X S H	110	HOMICIDE FIRST DEG GARAGE	FALSE	FALSE	1092	10	22	30	01A	1154123	1886297	2008	08/17/2011	41.84383	-87.7099	(41.84383, -87.7099)
3056	4790	HP620131	0000X W 1	110	HOMICIDE FIRST DEG STREET	TRUE	FALSE	1231	12	2	28	01A	1166321	1893492	2008	08/17/2011	41.86332	-87.6649	(41.86332, -87.6649)
3811	4791	HP620406	0000X W 2	110	HOMICIDE FIRST DEG STREET	TRUE	FALSE	1034	10	25	31	01A	1163192	1888742	2008	08/17/2011	41.85035	-87.6785	(41.85035, -87.6785)
4425	4792	HP620400	0000X S W	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	331	3	5	43	01A	1192078	1895937	2008	08/17/2011	41.78901	-87.5712	(41.78901, -87.5712)
5213	4793	HP620464	0000X S A	110	HOMICIDE FIRST DEG PORCH	FALSE	FALSE	823	8	16	66	01A	1155755	1884457	2008	08/17/2011	41.78386	-87.7045	(41.78386, -87.7045)
5867	4794	HP620489	0000X W A	110	HOMICIDE FIRST DEG ALLEY	FALSE	FALSE	911	9	14	63	01A	1160775	1870785	2008	08/17/2011	41.80112	-87.6859	(41.80112, -87.6859)
6359	4795	HP620537	0000X S A	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	713	7	16	67	01A	1166656	1886562	2008	08/17/2011	41.78941	-87.6645	(41.78941, -87.6645)
6917	4796	HP620532	0000X E P	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	624	6	6	44	01A	1180479	1885723	2008	08/17/2011	41.75113	-87.6142	(41.75113, -87.6142)
7483	4797	HP620552	0000X W 1	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	531	5	9	49	01A	1181244	1895927	2008	08/17/2011	41.8902	-87.6121	(41.8902, -87.6121)
7955	4798	HP620588	0000X W 2	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	1013	10	22	30	01A	1151426	1887526	2008	08/17/2011	41.84725	-87.7198	(41.84725, -87.7198)
8591	4799	HP620619	0000X S N	110	HOMICIDE FIRST DEG AUTO	TRUE	FALSE	1133	11	24	27	01A	1151362	1897015	2008	08/17/2011	41.87329	-87.7397	(41.87329, -87.7397)
9125	4800	HP620784	0000X W 1	110	HOMICIDE FIRST DEG ALLEY	TRUE	FALSE	2333	23	34	49	01A	1174927	1891706	2008	08/17/2011	41.84598	-87.6551	(41.84598, -87.6551)
9704	4801	HP620839	0000X W 1	110	HOMICIDE FIRST DEG RESTAUR	TRUE	FALSE	1591	15	32	6	01A	1166481	1915120	2008	08/17/2011	41.93364	-87.6638	(41.93364, -87.6638)
10316	4802	HP631705	0000X N 1	110	HOMICIDE FIRST DEG PORCH	TRUE	FALSE	1523	15	37	25	01A	1140228	1903134	2008	08/17/2011	41.89035	-87.7605	(41.89035, -87.7605)
10765	4803	HP632550	0000X W 2	110	HOMICIDE FIRST DEG APARTME	FALSE	FALSE	1223	12	25	31	01A	1165112	1890212	2008	08/17/2011	41.8541	-87.6695	(41.8541, -87.6695)
11455	4804	HP631900	0000X N 1	110	HOMICIDE FIRST DEG PORCH	TRUE	FALSE	1121	11	27	23	01A	1153226	1905401	2008	08/17/2011	41.89627	-87.7127	(41.89627, -87.7127)
12018	4805	HP633754	0000X S P	110	HOMICIDE FIRST DEG STREET	FALSE	FALSE	311	3	20	40	01A	1179072	1885118	2008	08/17/2011	41.78518	-87.619	(41.78518, -87.619)
12897	4806	HP640471	0000X S O	110	HOMICIDE FIRST DEG HALLWAY	TRUE	TRUE	1228	12	2	28	01A	1161171	1896649	2008	08/17/2011	41.87209	-87.6837	(41.87209, -87.6837)
13626	4807	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
14419	4808	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
15088	4809	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
15966	4810	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
16193	4811	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
16761	4812	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
17773	4813	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
17960	4814	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
18581	4815	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
19254	4816	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
19751	4817	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
20050	4818	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)
20050	4819	HP640416	0000X S W	110	HOMICIDE FIRST DEG HOUSE	TRUE	TRUE	731	7	6	69	01A	1175841	1895827	2008	08/17/2011	41.76661	-87.613	(41.76661, -87.613)

Fig B.1 Data used for training

Location Based Crime Prediction

Enter Date: 29 / 7 / 2022

Hour: 23 / Minutes: 15 / Seconds: 0

Latitude: 22.720862 / Longitude: 75.876083

Enter week number: 1.00

Enter week of year: 1.00

Fig B.2 Output of Web application

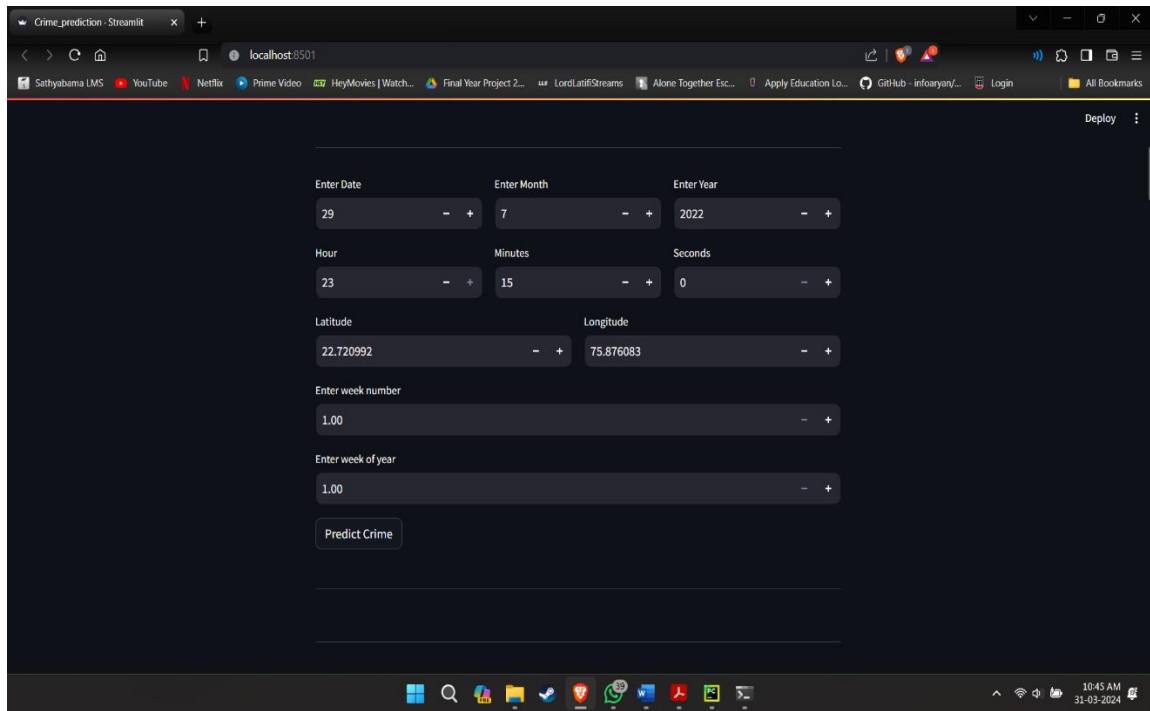


Fig B.3 Interface to give input

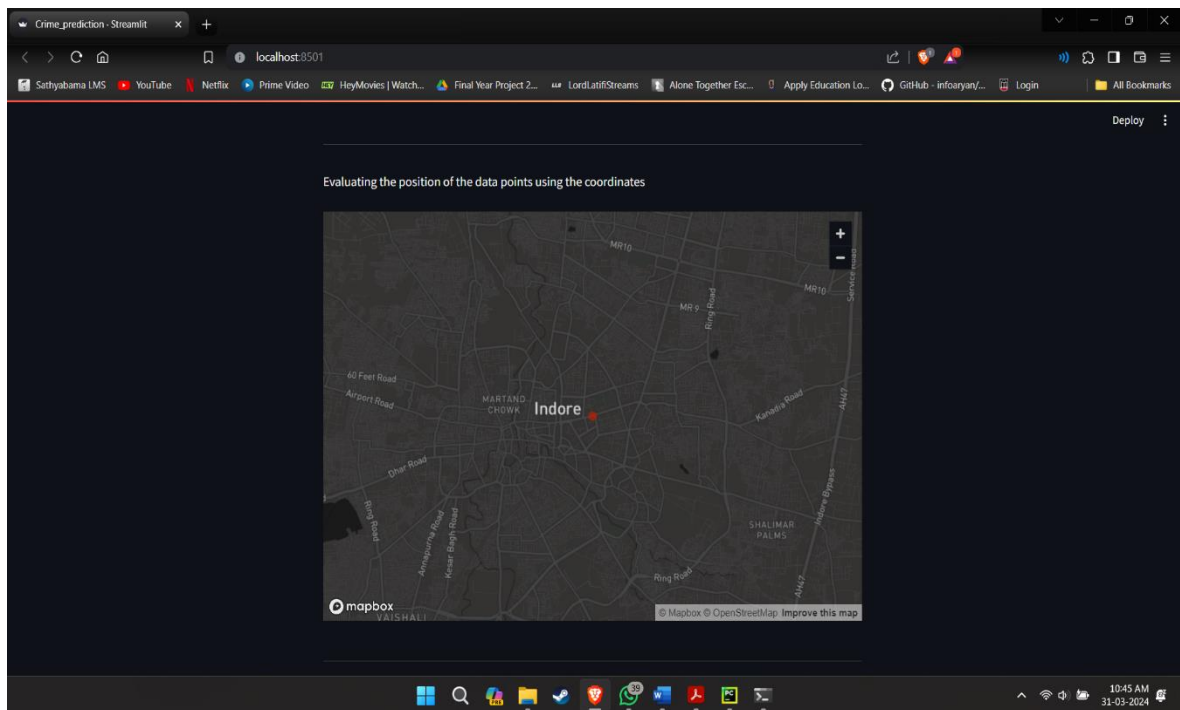


Fig B.4 Displaying Location

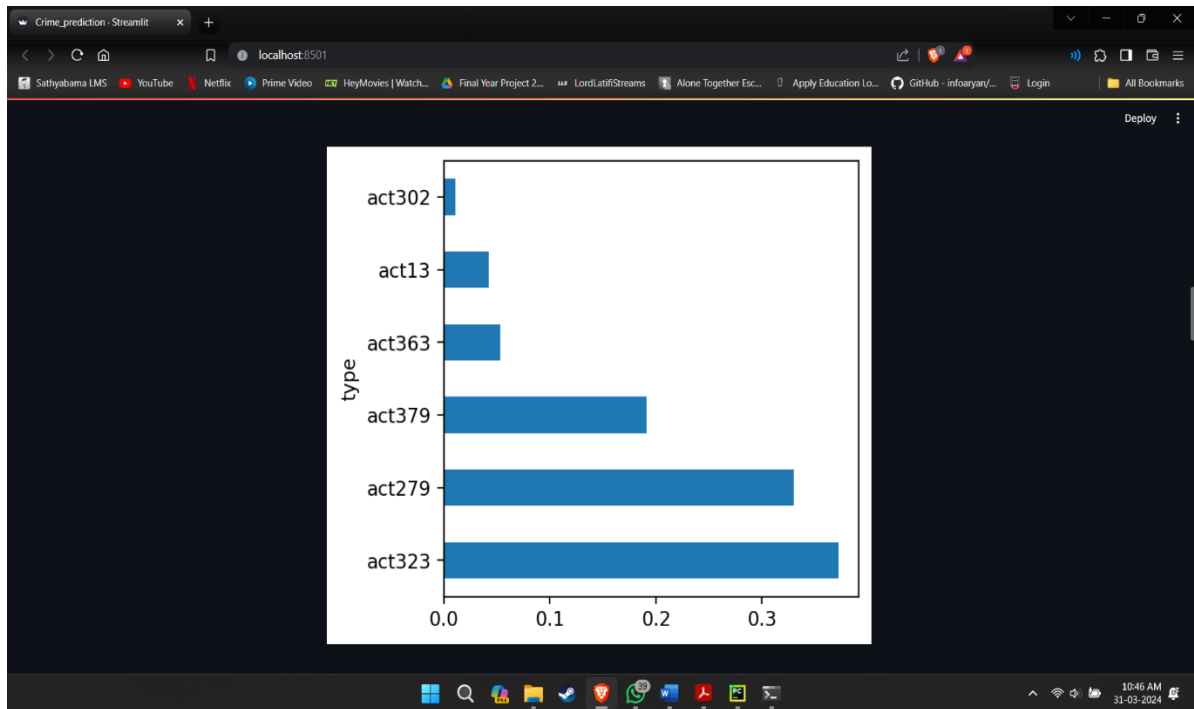


Fig B.5 Graph showing Number of crimes happened(relative)

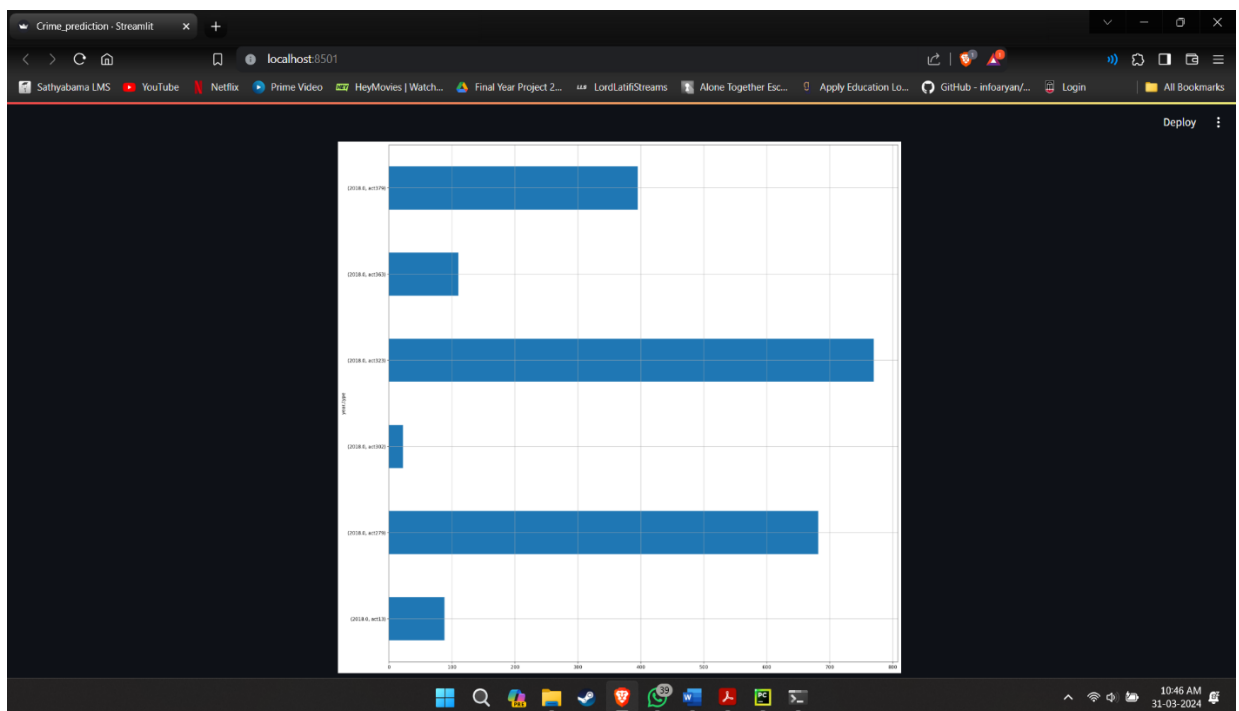


Fig B.6 Graph showing Year and crime category relationship

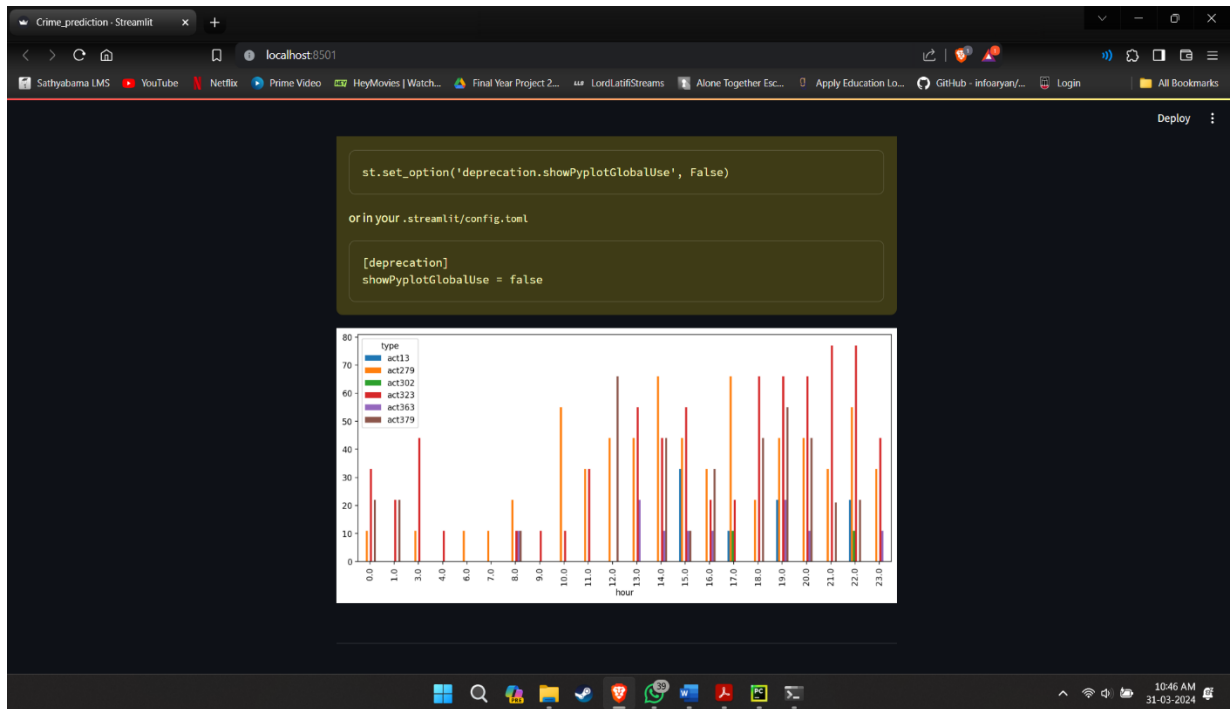


Fig B.7 Graph showing hour of the day, crime type and year relationship

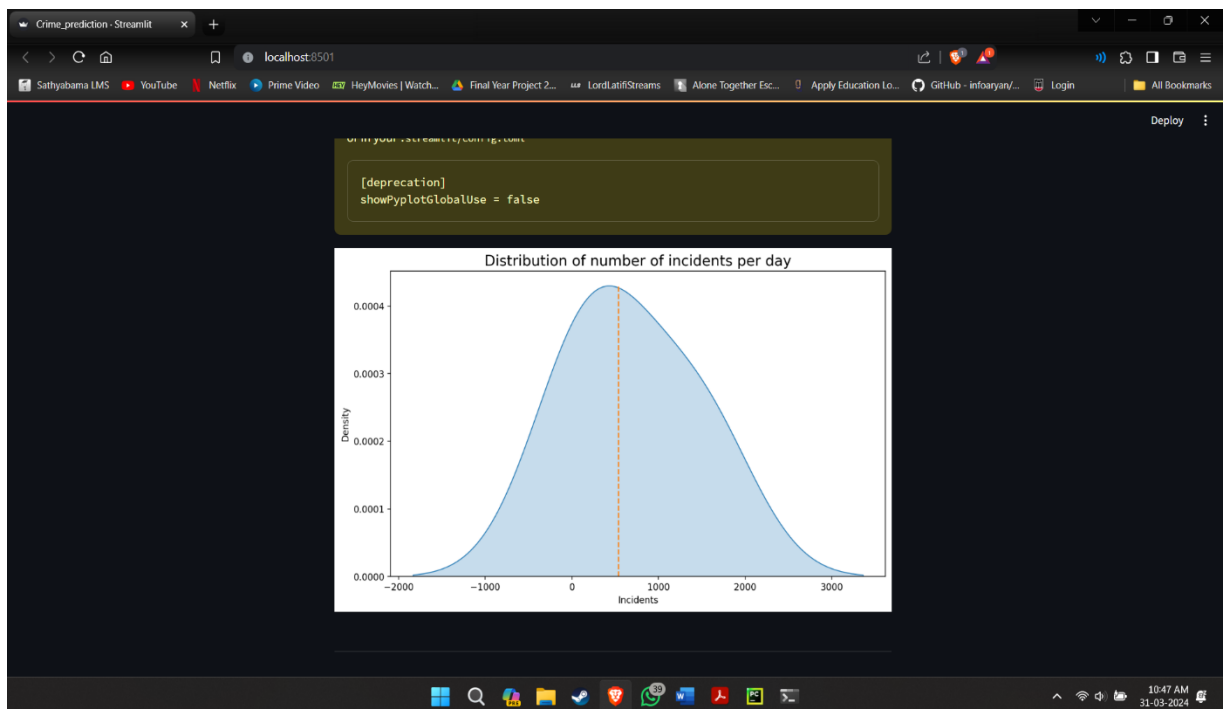


Fig B.8 Graph showing distribution of number of incidents per day

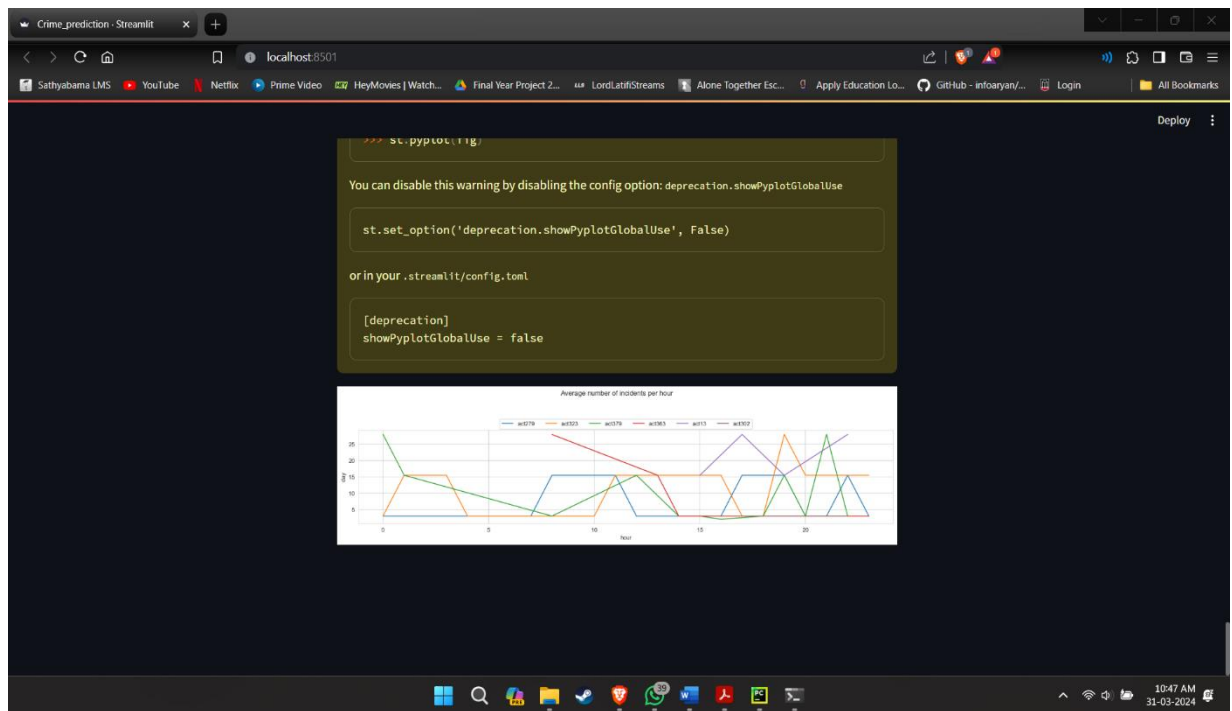


Fig B.9 Graph showing Crime on hourly basis plotted yearly

C. RESEARCH PAPER

Crime Analysis and Detection Based On Location Using Machine Learning

S.L. Jany Shabu ^{a)}, Patan Firoz khan, , Varshith Peddineni, J.Refonaa, M.Maheswari, Mohanapriya

Sathyabama Institute of Science and Technology, Chennai, India

^{a)} *Corresponding author: janyshabu.cse@sathyabama.ac.in*

Abstract. Location based crime detection is a significant aspect of modern law enforcement, aimed at improving public safety and crime prevention. This research proposal focuses on the development of an innovative system that utilizes location-based technologies and data analysis techniques for crime detection and prediction. By analyzing real-time location data collected from various sources, such as mobile phones and surveillance systems, the system can identify crime hotspots and patterns. The proposed system incorporates advanced algorithms and machine learning techniques to generate predictive models that can forecast the occurrence of crime in specific areas. Moreover, it offers visualization capabilities to assist law enforcement agencies in monitoring crime trends and allocating resources efficiently. The system's effectiveness will be evaluated using historical crime data and compared with traditional crime detection methods. This study aims to contribute to the existing knowledge in the field of crime detection by providing a reliable and efficient tool for law enforcement agencies to enhance their crime prevention strategies..

Keywords: *location-based, crime detection, data analysis, predictive models, machine learning, visualization, law enforcement, crime prevention.*

INTRODUCTION

Location-based crime detection refers to the use of geospatial data and technology to identify, analyze, and predict criminal activities based on spatial patterns and trends. With the increasing adoption of digital technologies and the availability of vast amounts of geospatial data, law enforcement agencies, researchers, and policy-makers have started to explore new and innovative ways to leverage this information for crime prevention and public safety.

Location-based crime detection offers a key advantage: it furnishes valuable perspectives on how crime incidents are spread across different spaces. By mapping crime data onto geographical locations, patterns and hotspots can be identified, highlighting areas with higher crime rates. This information can be used to allocate police resources more effectively, directing patrols to areas that experience higher crime activity and optimizing the use of limited resources. Additionally, it allows for the identification of crime trends over time, enabling law enforcement agencies to anticipate and respond to emerging issues more efficiently.

Sophisticated analytics and machine learning methods are frequently used in pinpointing crimes based on location to create predictive models capable of anticipating patterns in criminal activities. These models take into account different elements like past crime records, population data, atmospheric conditions, and even content on social platforms all play a role to identify potential crime hotspots and predict future criminal activity. By proactively identifying high-risk areas, law enforcement agencies can deploy preventive measures and interventions to deter criminal behavior and improve public safety.

Location-based crime detection also plays a crucial role in supporting investigations and solving crimes. Geospatial analysis can help link seemingly unrelated crimes by identifying spatial and temporal connections between incidents. For example, it can reveal patterns of serial offenses or identify common characteristics of crime scenes that may be associated with a particular criminal or a criminal organization. Such insights can provide investigators with valuable leads and help them develop strategies to apprehend offenders more quickly.

In addition to law enforcement agencies, location-based crime detection can also benefit other stakeholders,

such as urban planners and policymakers. By understanding the spatial patterns of crime, urban planners can design safer communities, incorporating features such as improved street lighting, surveillance cameras, and better transportation infrastructure. Insights gleaned from location-based crime detection offer policymakers a foundation for crafting evidence-driven policies and programs. These aim to tackle the root causes of crime, including issues like poverty, unemployment.

Overall, location-based crime detection holds great promise in enhancing crime prevention, investigative processes, and public safety. By leveraging geospatial data and advanced analytics, law enforcement agencies and other stakeholders can gain valuable insights into crime patterns, predict future criminal activity, and design targeted interventions to create safer communities. However, it is crucial to ensure the ethical and responsible use of these technologies, safeguarding individual privacy and guaranteeing the fairness and transparency of decision-making processes

LITERATURE REVIEW

1.Amaguchi et al. (2023): Yamaguchi and colleagues introduce a system designed to enhance the experience of cyclists by detecting comfortable locations through facial expression analysis. By leveraging machine learning techniques, the system aims to identify areas along cycling routes where riders exhibit signs of comfort, contributing to a more enjoyable and safer cycling experience overall.

2.Hu et al. (2022): Hu and his team conduct a comprehensive survey and comparison of methods for recognizing location references in textual data. Their work delves into various techniques employed in extracting location information from text, providing insights into the challenges and advancements in this field, which is crucial for crime analysis and detection based on textual data.

2.Singh et al. (2022): Singh and co-authors explore the emerging field of forensic entomology as a novel approach to crime investigation. By studying the role of insects in crime scene analysis, their research sheds light on how entomological evidence can be utilized alongside traditional forensic methods to provide valuable insights into the timeline and circumstances surrounding criminal activities.

3.Boukabous and Azizi (2023): Boukabous and Azizi discuss the utilization of image and video-based crime prediction techniques employing object detection and deep learning methodologies. Their research aims to predict and prevent criminal activities through the analysis of visual data, highlighting the potential of advanced technologies in augmenting law enforcement efforts.

3.Jenga, Catal, and Kar (2023): Jenga, Catal, and Kar delve into the application of machine learning algorithms in crime prediction and prevention. By reviewing various machine learning approaches, their work provides a comprehensive overview of the advancements and challenges in leveraging predictive analytics for enhancing public safety and security.

4.Mondal, Singh, and Kumar (2022): Mondal and colleagues conduct a case study on crime hotspot detection utilizing statistical and geospatial methods, focusing on Pune City in India. Their research provides insights into the spatial distribution of crime and identifies hotspot areas, which can aid law enforcement agencies in allocating resources more effectively for crime prevention.

5.Girinath et al. (2022): Girinath et al. present an IoT-based system for threat detection and location tracking aimed at enhancing women's safety. By leveraging IoT technology, their system provides real-time threat detection and location tracking capabilities, offering a proactive approach to addressing safety concerns.

6.Prathap (2023): Prathap discusses the application of machine learning algorithms for analyzing crime data and attributing crime density to specific geographic areas. By employing optimized machine learning techniques, his research offers insights into the spatial patterns of crime, facilitating targeted interventions for crime prevention and control.

7.Mukherjee and Ghosh (2023): Mukherjee and Ghosh propose a predictive framework for crime data analysis utilizing a hybrid logistic regression—support vector machine-based ensemble classifier. Their research aims to improve the accuracy of crime prediction by integrating multiple machine learning techniques, offering a robust approach to proactive law enforcement strategies.

8.Wei et al. (2022): Wei and colleagues provide a survey of location-based social networks, exploring the challenges, methodologies, and future research directions in this domain. Their work sheds light on the potential of location-based social networking data for crime analysis and detection, highlighting the importance of leveraging social media platforms in understanding spatial dynamics of criminal activities.

EXISTING SYSTEM

The existing system for location-based crime detection faces several disadvantages that hinder its effectiveness in preventing and combating criminal activities. Firstly, one major drawback is the limited accuracy of the system. Location-based crime detection relies on GPS and cellular technology to track the location of

individuals or devices. However, these technologies can be imprecise, especially in densely populated or built-up areas where buildings and structures can interfere with signal reception. This inaccuracy can lead to false information and unreliable crime detection.

Secondly, the existing system often lacks real-time data. Gathering and analyzing data on crime occurrences and where they happen can be a time-intensive and bureaucratic process. Consequently, the information used for crime detection may not be up to date, rendering the system less effective in identifying and responding to ongoing criminal activities. Delayed data can result in missed opportunities to prevent crimes as they occur.

Another disadvantage of the existing system is its dependence on user reporting. In many cases, individuals need to actively report crimes or suspicious activities to the authorities through dedicated applications or hotlines. This reliance on user reporting is prone to human error, bias, and selective reporting. Not all crimes might be reported, and some individuals might not have access or knowledge of the reporting channels, leading to underrepresentation of crime data. Moreover, individuals may hesitate to report crimes due to fear, lack of trust in law enforcement, or concerns about their personal safety.

Additionally, privacy concerns are also a significant drawback of the existing system. Location-based crime detection relies on continuously tracking and monitoring individuals' movements and activities. This raises legitimate privacy concerns among the public, as constant surveillance can be seen as an invasion of personal privacy. Such concerns can lead to resistance or reluctance in adopting the system, limiting its potential impact.

Lastly, the existing system often lacks integration and coordination among different law enforcement agencies and departments. Sharing crime data and information across agencies can be challenging due to bureaucratic and technical barriers, resulting in fragmented and incomplete data sets. This lack of synergy reduces the effectiveness and efficiency of crime detection efforts.

Overall, the existing system for location-based crime detection faces disadvantages in terms of accuracy, real-time data availability, reliance on user reporting, privacy concerns, and lack of integration among law enforcement agencies. Addressing these limitations is crucial to developing a more robust and efficient crime detection system.

PROPOSED SYSTEM

Location-based crime detection is a proposed work that aims to utilize cutting-edge technologies and data analysis techniques to efficiently identify and prevent criminal activities in specific geographical areas. The primary focus of this approach would be to harness the power of location data, such as GPS or geospatial information, along with crime data, to create a comprehensive crime detection and prevention system. Through the integration of various sources, such as surveillance cameras, social media platforms, and mobile devices, real-time updates and alerts could be generated to notify authorities and individuals about potential criminal incidents.

The planned endeavor entails crafting sophisticated algorithms and machine learning models aimed at scrutinizing extensive data volumes. Their goal is to spot patterns, trends, and irregularities in criminal behavior linked to specific locations. By creating a predictive model, law enforcement agencies and community members can proactively take measures to prevent crimes from occurring in specific areas. This could include increasing police presence, implementing targeted intervention strategies, or improving the overall security infrastructure in vulnerable locations.

To support this location-based crime detection system, an extensive database would be established, housing historical crime data, geographical information, and relevant socio-demographic factors. This information would serve as a valuable resource for analyzing crime patterns and understanding the underlying factors contributing to criminal activity in specific areas. Advanced data visualization techniques could also be employed to depict crime hotspots, identify crime trends over time, and assist in allocating resources more effectively.

Furthermore, the proposed work might include the development of user-friendly mobile applications or web interfaces that allow individuals to report crimes in real-time, along with their precise location details. This crowd-reporting feature would help improve the accuracy and speed of information sharing, enabling rapid responses from law enforcement agencies and the community. Moreover, individuals could receive

personalized crime alerts and safety recommendations based on their geographic location, further enhancing their safety and security.

In summary, the proposed work for location-based crime detection aims to leverage location data, advanced analytics, and user participation to create a comprehensive crime prevention system. By integrating various sources of information and utilizing advanced algorithms, this approach aims to identify patterns and predict criminal activities in specific areas, thus allowing for proactive measures to be taken to deter crimes. Ultimately, this would lead to safer communities and more effective law enforcement efforts.

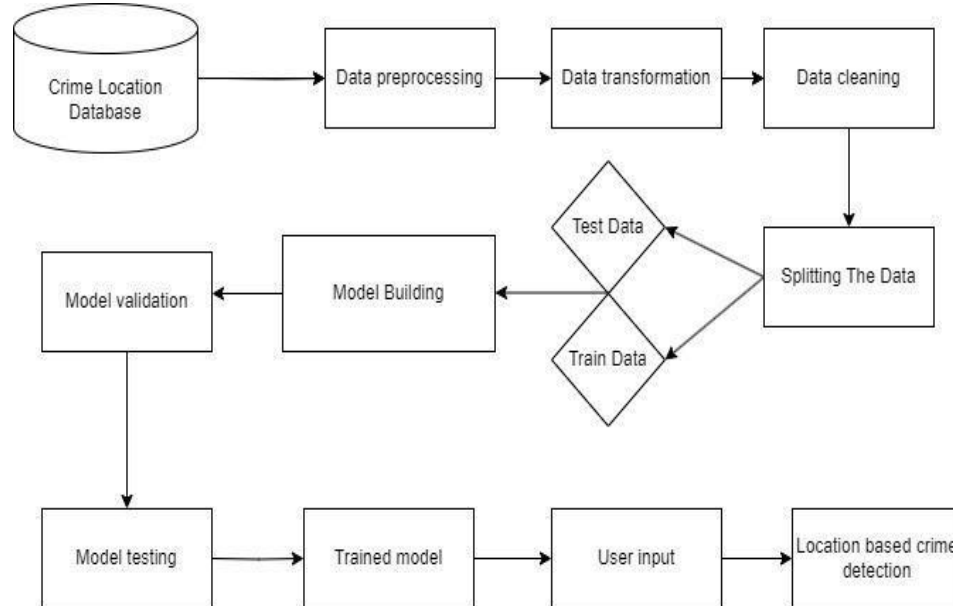


Fig. 1. Architechture Diagram.

METHODOLOGY

Module1: Data Collection and Integration This module focuses on collecting and integrating relevant data from various sources to build a comprehensive database for location-based crime detection. It involves gathering data from sources such as crime reports, emergency calls, CCTV cameras, social media, and other related sources. The data collected may include information on crime types, time, location, victim details, suspect information, and any relevant details that can aid in crime analysis. The module also ensures the data collected is standardized and organized for efficient storage and retrieval.

Module 2: The Crime Analysis and Prediction module harnesses advanced analytics and algorithms to scrutinize gathered data, pinpointing patterns, trends, and connections within it. The main objective is to identify crime hotspots, understand the factors contributing to crime occurrences, and predict potential crime incidents. This module employs various statistical and machine learning techniques to perform data mining, clustering, classification, and predictive modeling. It also considers factors such as demographic data, weather conditions, and historical crime data to improve the accuracy of crime prediction. The results generated by this module can help law enforcement agencies in allocating resources effectively, implementing preventive measures, and developing targeted strategies to reduce crime rates.

Module 3: Real-time Crime Reporting and Visualization This module focuses on providing real-time crime reporting and visualization capabilities to both law enforcement agencies and the general public. It involves developing a user-friendly interface or mobile application that allows users to report crimes directly, provide additional details, and share multimedia evidence if available. The module also enables the integration of emergency call systems, allowing instant reporting of crimes to the appropriate authorities. Furthermore, it provides interactive visualization tools that allow users to view crime trends, hotspots, and related statistics on maps, graphs, and charts. This graphical depiction aids in enhancing comprehension of crime trends and supports law enforcement agencies in making well-informed choices. The real-time nature of this module enhances the effectiveness and timeliness of crime reporting and analysis, enabling faster response times and

improved safety measures.

Overall, these three modules work together in a location-based crime detection system to gather, analyze, and visualize crime data, aiding authorities in proactive crime prevention and optimization of resources.

RESULT AND DISCUSSION

Accuracy	Precision	Recall	F1 score
97.4	97.1	96.2	96.3

Table.1. Performance Metrics

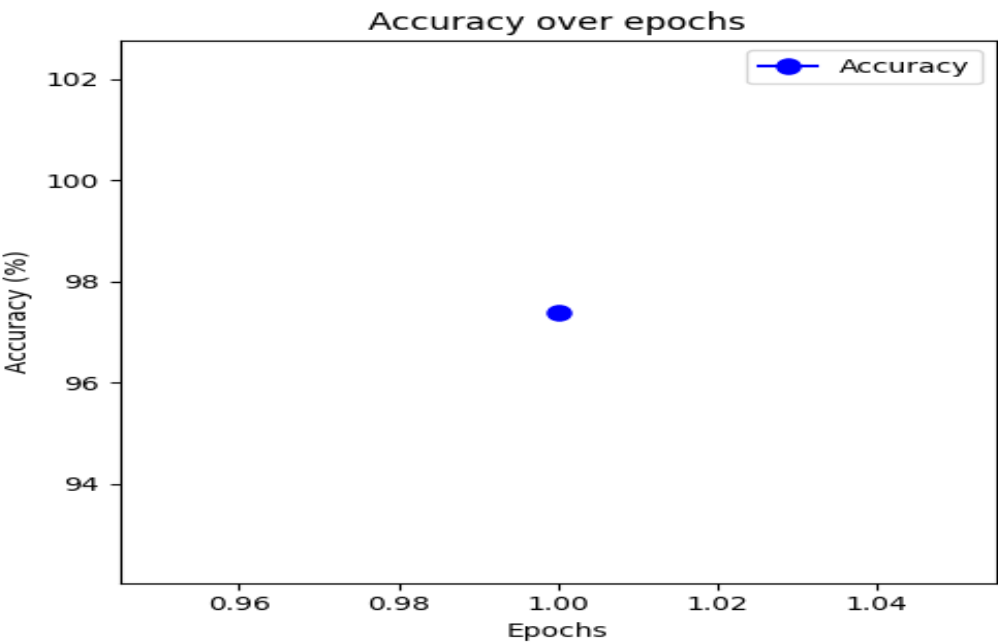


Fig.2. Accuracy graph

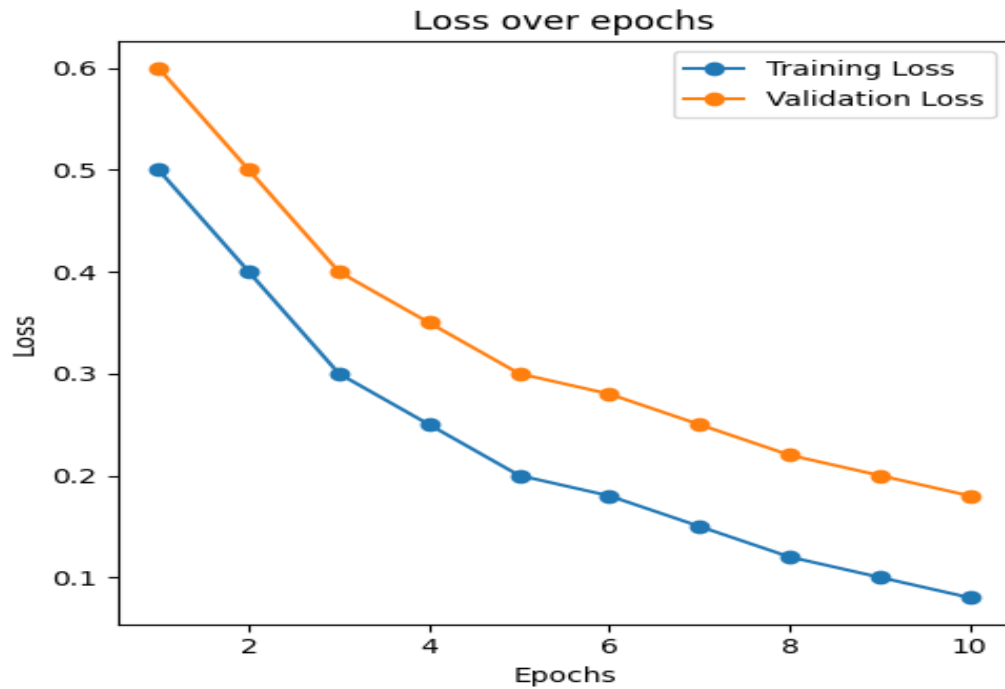


Fig.3. Loss graph

The system for location-based crime detection is a powerful tool that integrates various technologies and data sources to efficiently identify and analyze criminal activities based on geographical information. By combining geographic information system (GIS) technology with real-time crime data, this system allows law enforcement agencies to effectively monitor and respond to crimes in specific locations.

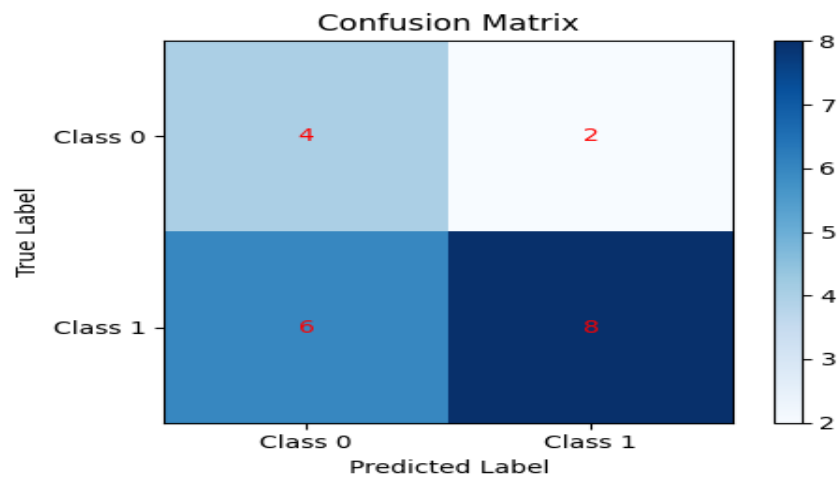


Fig.4. Confusion Matrix

The system works by collecting and analyzing crime data from multiple sources, such as police reports, emergency calls, and surveillance footage. Through GIS technology, this data is then mapped onto a digital map, allowing authorities to visualize crime hotspots and patterns in real-time. This enables them to allocate resources strategically, focusing on areas with higher crime rates and providing a timely response to incidents.

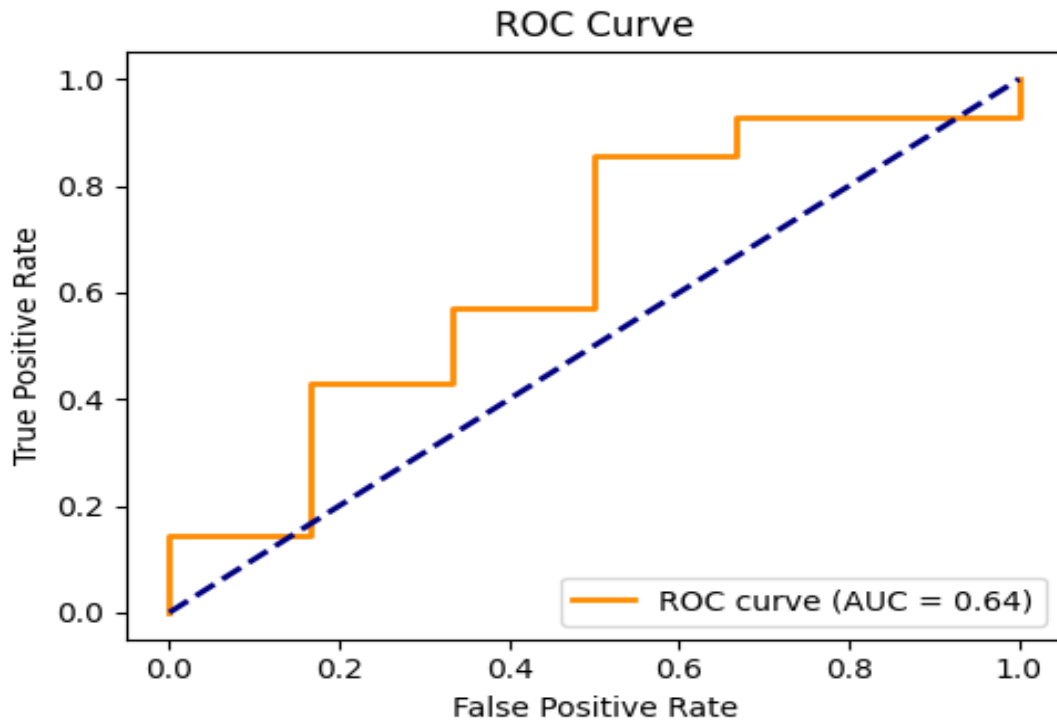


Fig.4. ROC Curve

Furthermore, the system utilizes advanced algorithms to predict crime hotspots based on historical data and factors such as population density, time of day, and weather conditions. This proactive approach allows law enforcement agencies to deploy preventive measures and allocate resources in advance, deterring criminal activities and ensuring public safety.

Additionally, the system can be integrated with other technologies, such as video analytics and facial recognition, to enhance surveillance and identify suspects more effectively. This integration enables law enforcement agencies to track and monitor criminals in real-time, greatly enhancing their ability to apprehend individuals responsible for criminal activities.

In general, the system for detecting crimes by location proves to be an invaluable asset to law enforcement agencies. It equips them to effectively identify, prevent, and address crimes using geographical data. By leveraging technology and data analysis, this system contributes to creating safer communities and ensuring the well-being of citizens.

CONCLUSION

In conclusion, the location-based crime detection system plays a crucial role in enhancing the safety and security of individuals and communities. By utilizing data from various sources such as surveillance cameras, GPS devices, and social media platforms, this system can effectively identify and monitor criminal activities in real-time. It enables law enforcement agencies to quickly respond to incidents and allocate resources efficiently. Furthermore, the analysis of crime patterns and trends provided by this system can assist in creating targeted crime prevention strategies and allocating resources effectively. With continuous advancements in technology and data analytics, the location-based crime detection system holds great potential in improving public safety and reducing crime rates in the future.

FUTURE WORK

Future work on the system for location-based crime detection could involve several aspects. First, there is a need to enhance the accuracy of crime prediction by incorporating more advanced machine learning algorithms and techniques. This could include using deep learning models such as convolutional neural networks or recurrent neural networks to effectively analyze complex spatial and temporal data. Additionally, integrating real-time data sources, such as CCTV footage or police reports, could further improve the accuracy and reliability of crime detection algorithms. Furthermore, the system could benefit from the incorporation of social

media data analysis to detect crime-related activities or discussions. Another area for future work is the development and implementation of a mobile application that can provide alerts and warnings to users in real-time based on their location and the predicted crime hotspots. This application could also allow users to report suspicious activities or incidents directly to law enforcement agencies. Finally, evaluating the system's performance in different cities or regions could provide valuable insights into its effectiveness and scalability, leading to potential adaptations and improvements in the future.

REFERENCES

- [1] Yamaguchi, R., Siriaraya, P., Yoshihisa, T., Shimojo, S., & Kawai, Y. (2023, April). A detection system for comfortable locations based on facial expression analysis while riding bicycles. In Companion Proceedings of the ACM Web Conference 2023 (pp. 306-309).
- [2] Hu, X., Zhou, Z., Li, H., Hu, Y., Gu, F., Kersten, J., ... & Klan, F. (2022). Location reference recognition from texts: A survey and comparison. arXiv preprint arXiv:2207.01683.
- [3] Singh, R., Kumawat, R. K., Singh, G., Jangir, S. S., Kushwaha, P., & Rana, M. (2022). Forensic entomology: a novel approach in crime investigation. GSC Biological and Pharmaceutical Sciences, 19(2), 165-174.
- [4] Boukabous, M., & Azizi, M. (2023). Image and video-based crime prediction using object detection and deep learning. Bulletin of Electrical Engineering and Informatics, 12(3), 1630-1638.
- [5] Jenga, K., Catal, C., & Kar, G. (2023). Machine learning in crime prediction. Journal of Ambient Intelligence and Humanized Computing, 14(3), 2887-2913.
- [6] Mondal, S., Singh, D., & Kumar, R. (2022). Crime hotspot detection using statistical and geospatial methods: a case study of Pune City, Maharashtra, India. GeoJournal, 87(6), 5287-5303.
- [7] Girinath, N., Vidhya, B., Surendar, R., Abhirooban, T., & Sabarish, V. (2022, October). IoT based Threat Detection and Location Tracking for Women Safety. In 2022 International Conference on Edge Computing and Applications (ICECAA) (pp. 618-622). IEEE.
- [8] Prathap, B. R. (2023). Geo-spatial crime density attribution using optimized machine learning algorithms. International Journal of Information Technology, 15(2), 1167-1178.
- [9] Mukherjee, A., & Ghosh, A. (2023). Predictive framework for crime data analysis using a hybrid logistic regression—support vector machine based ensemble classifier powered by CART (LR-SVM CART). Multimedia Tools and Applications, 1-21.
- [10] Wei, X., Qian, Y., Sun, C., Sun, J., & Liu, Y. (2022). A survey of location-based social networks: problems, methods, and future research directions. GeoInformatica, 1-41.
- [11] Chhabra Roy, N., & Prabhakaran, S. (2023). Internal-led cyber frauds in Indian banks: an effective machine learning-based defense system to fraud detection, prioritization and prevention. *Aslib Journal of Information Management*, 75(2), 246-296.
- [12] Yuan, H., Wang, Z., Zhang, F., Guan, D., & Zhao, R. (2023). Trends in forensic microbiology: from classical methods to deep learning. *Frontiers in Microbiology*, 14, 1163741.
- [13] Guo, J., Xu, Q., Zeng, Y., Liu, Z., & Zhu, X. X. (2023). Nationwide urban tree canopy mapping and coverage assessment in Brazil from high-resolution remote sensing images using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 198, 1-15.
- [14] Yue, H., Liu, L., & Xiao, L. (2023). Investigating the effect of people on the street and streetscape physical environment on the location choice of street theft crime offenders using street view images and a discrete spatial choice model. *Applied Geography*, 157, 103025.

- [15] Kathiravan, M., Parvez, S. J., Dheepthi, R., Jayanthi, R., Gowsalya, S., & Sekhar, R. V. (2023, January). Analysis and detection of fake profile over social media using machine learning techniques. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1164-1169). IEEE.

