MACHININE LEARNING-BASED PREDICATION TECHNIQUE FOR ANALYSING COMMUNITIES AND CRIME

Project submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

BY

T.POOJITHA (18C91A05A1)
S.DEEPAK KUMAR (18C91A0595)
T.HARSHWARDHAN SINGH (18C91A0598)

Under the Esteemed guidance of

Mr.V.VEERESH M.tech(Ph.D)

Associate professor in CSE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING-14

HOLY MARY INSTITUTE OF TECHNOLOGY & SCIENCE-16

(COLLEGE OF ENGINEERING) 11

(Approved by AICTE New Delhi, Permanently Affiliated to JNTU Hyderabad, Accredited by NAAC with 'A' Grade)

Bogaram (V), Keesara (M), Medchal District -501 301. - 11

2021 - 2022

HOLY MARY INSTITUTE OF TECHNOLOGY & SCIENCE

(COLLEGE OF ENGINEERING)

(Approved by AICTE New Delhi, Permanently Affiliated to JNTU Hyderabad, Accredited by NAAC with 'A' Grade)

Bogaram (V), Keesara (M), Medchal Dist-501301.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the mini project entitled "MACHININE LEARNING-BASED PREDICTION TECHNIQUE FOR ANALYZING COMMUNITIES AND CRIME" is being submitted by T.POOJITHA (18C91A05A1), S.DEEPAK KUMAR (18C91A0595),T.HARSHAWARDHAN (18C91A0598), in Partial fulfillment of the academic requirements for the award of the degree of Bachelor ofTechnology in "COMPUTER SCIENCE AND ENGINEERING" HOLY MARY INSTITUTE OF TECHNOLOGY & SCIENCE, JNTU Hyderabad during the year 2021-2022.

INTERNAL GUIDE

Mr. V.VEERESH M.Tech(Ph.D)
Associate professor in CSE
Dept. of Computer Science & Engineering.

HEAD OF THE DEPARTMENT

DR .B.NARSIMHA M.Tech,Ph.D.
Professor & HoD
Dept. of Computer Science & Engineering

EXTERNAL EXAMINER-14

ACKNOWLEDGEMENT -17

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, who's constant guidance and encouragement crowns all effort with success.

I take this opportunity to express my profound gratitude and deep regards to My Guide **Mr. V.VEERESH, Assistant Professor**, Dept. of Computer Science & Engineering, Holy Mary Institute of Technology & Science for his / her exemplary guidance, monitoring and constant encouragement throughout the project work.

My special thanks to **Dr. B. Narsimha, Head of the Department**, Dept. of Computer Science & Engineering, Holy Mary Institute of Technology & Science who has given an immense support throughout the course of the project.

I also thank to **Dr. P. Bhaskara Reddy,** the **honorable Director** of my college Holy Mary Institute of Technology & Science for providing me the opportunity to carry out this work.

At the outset, I express my deep sense of gratitude to the beloved **Chairman A. Siddartha Reddy** of **Holy Mary Institute of Technology & Science**, for giving me the opportunity to complete my course of work

I am obliged to **staff members** of Holy Mary Institute of Technology & Science for the valuable information provided by them in their respective fields. I am grateful for their cooperation during the period of my assignment.

Last but not the least I thank **ALMIGHTY** and My **Parents**, and **Friends** for their constant encouragement without which this assignment would not be possible.

T.POOJITHA **(18C91A05A1)**

S.DEEPAK KUMAR **(18C91A0595)**

T.HARSHWARDHAN SINGH **(18C91A0598)**

DECLARATION

This is to certify that the work reported in the present project titled **"MACHININE LEARNING-BASED PREDICTION TECHNIQUE FOR ANALYZING COMMUNITIES AND CRIME"** is a record of work done by me in the Department of Computer Science & Engineering, Holy Mary Institute of Technology and Science.

No part of the thesis is copied from books/journals/internet and wherever the portion is taken, the same has been duly referred in the text the reported are based on the project work done entirely by me not copied from any other source.

T.POOJITHA **(18C91A05A1)**

S.DEEPAK KUMAR **(18C91A0595)**

T.HARSHWARDHAN SINGH **(18C91A0598)**

CONTENTS

LIST OF FIGURES

Figure Names page no.

ABSTRACT

We live in a society where crime is both a dominant factor and an alarming phenomenon. Every day, crimes are committed in huge numbers. These crimes disturb the lives of many citizens, making them restless. Machine learning have become a vital part of crime detection and prevention. The purpose of this paper is to evaluate machine learning techniques and their performances that can be used for analyzing the collected data about the past crimes. We identified some of the appropriate machine learning techniques like Random Forest, KNN to analyze the collected data from different source. Some attributes of the dataset will contain the crimes that are committed/happened in our communities.

HISTORY

The formalized processes of analyzing crime and establishing specialized units to perform these tasks are relatively new to law enforcement. If you are just beginning to develop a crime analysis program, you may be asked to acquaint your agency's personnel with the history of this most interesting field. The following material is provided to assist you with such a presentation. Analyzing crime to identify suspects and/or crime patterns can be traced to the feudal period of England. Populations of geographic areas were small, people knew each other well, and villagers often spent an entire lifetime in the same village or town. The local constable knew who the troublemakers were, where they lived, and had a good working knowledge of how they committed their crimes. By the late 1800s, crime in Europe had risen to the point that it seemed only Sherlock Holmes, the fictional detective created by Sir Arthur ConanDoyle, was capable of ferreting out the criminals. Holmes' ability to scientifically analyze bits and pieces of information made him a legendary hero; the real police officer was viewed as a buffoon. The reason, of course, was that use of scientific methods to study and analyze crime was not as advanced in reality as it was in the imaginative mind of Mr. Holmes' creator. As the years rolled on, new scientific discoveries were made and many came to be of significant benefit to law enforcement. Physical examination of a wide variety of materials enabled police officials to place suspects at scenes of crimes. At the same time, the developing field of psychology began to explain the nature of human personality and the establishment of habit patterns.These and other outgrowths of scientific inquiry allowed law enforcement to accurately determine the nature of criminal activity, predict its future occurrence, and identify its perpetrators.

PURPOSE

After finding and understanding various distinct methods used by the police for surveillance purposes, we determined the importance of each method.Each surveillance method can perform well on its own and produce satisfactory results, although for only one specific characteristic, that is, if we use a Sting Ray, it can help. Us only when the suspect is using a phone, which should be switched on. Thus, it is only useful when the information regarding the stake out location is correct. Based on this information, we can see how the ever-evolving technology has yet again produced a smart way to con duct surveillance. The introduction of deep learning ML and computer vision techniques has provided us with a new perspective on ways to conduct surveillance. This is an intelligent approach to surveillance because it tries to mimic a human approach, but it does so 24 h a day, 365 days a year, and once it has been taught how to do things it does them in the same manner repeatedly.Although we have discussed the aspects that ML and computer vision can achieve, but what are these aspects essentially? This brings us to the main point of our paper discussion, Le, our proposed idea, which is to combine the point aspects of Sting Ray, body cams, facial recognition, number plate recognition, and stake outs. New features include core analytics, neural networks, heuristic engines, recursion processors, Bayes networks, data acquisition, cryptographic algorithms, document processors, computational linguistics, voice print identification, natural language processing, gait analysis, biometric recognition, pattern mining, intel in interpretation, threat detection, threat classification. The new features are completely computer dependent and hence require human interaction for development; however, once developed, it functions without human inter action and frees humans for other tasks. Let us understand the use of each function.

1.INTRODUCTION

At present, the criminal cases that are pending in India are rapidly increasing with the number of crimes committed are increasing. To solve a case based upon a particular data there should be a thorough investigation and analysis that is to be done internally . With the amount of crime data that is present in India currently the analysis and decision making of these criminal cases is too difficult for the officials. Identifying this a major problem this paper concentrates on creating a solution for the decision making of crime that is committed.

Machine Learning is the branch of science where computers decide without human intervention. In recent times Machine Learning is being used in various domains one of the examples of such cases is automated or self-driving cars. By Machine Learning algorithms there is a way where we can predict certain results based upon our inputs given and provide a solution to solving crime cases in India. The two common types of prediction techniques are classification and regression. This crime data prediction is a domain where classification is applied. Classification is a supervised prediction technique and it has been used in various domains like forecasting stock, medicinal area, etc.. The main aim of this paper is to consider some algorithms which can be used to predict and analyse the crime data and improve the accuracy of those models by data processing in order to obtain better results. The purpose is to train the required model to predict the data using the training data set by validation of the test data set . The models which are being used here are Logistic Regression, Decision Tree classification, Random Forest classification.

PROBLEM STATEMENT:

The main problem is that day to day the population is going to be increased and by that the crimes are also going to be Increased in different areas by this the crime rate can‘t be accurately predicted by the officials. The officials as they focus on many issues may not predict the crimes to be happened in the future. The officials/police officers although they tries to reduce the crime rate they may not reduce in full-fledged manner. The crime rate prediction in future may be difficult for them. There has been countless of work done related to crimes.Large datasets have been reviewed, and information such as location and the type of crimes have been extracted to help people follow law enforcements.Existing methods have used these databases to identify crime hotspots based on locations. There are several maps applications that show the exact crime location along with the crime type for any given city .Even though crime locations have been identified, there is no information available that includes the crime occurrence date and time along with techniques that can accurately predict what crimes will occur in the future.

OBJECTIVES:

The main objective of the project is to predict the crime rate and analyze the crime rate to be happened in future. Based on this Information the officials can take charge and try to reduce the crime rate . The concept of Multi Linear Regression is used for predicting the graph between the Types of Crimes (Independent Variable) and the Year (Dependent Variable).The system will look at how to convert crime information into a regression problem, so that it will help detectives in solving crimes faster.Crime analysis based on available information to extract crime patterns. Using various multi linear regression techniques, frequency of occurring crime can be predict.Data mining can be used to

examine many large datasets involving a large set of variables beyond what a single analyst, or even an analytical team or task force, can consider correct, whereas machine learning uses neural networks, predictive model and automated algorithms to make the decisions. Like any other problem solving method, the task of data mining begins with a problem definition. The identification of the data mining problem enables the determination of the data mining process and the modeling technique. Machine learning is a subfield of data science that deals with algorithms able to learn from data and make accurate predictions . Data mining gives law enforcement agencies the opportunity to learn about crime trends, how and why crimes are committed. Using data mining methods and machine learning improves crime analysis and help reduce and prevent crime based on territorial distribution of existing data and Crime recognition.

MOTIVATION:

The objective is to develop an application which provides the details of a prediction of crimes. The purpose is to design an application which automates the processes involved and allows users to perform various operations.

SCOPE OF PROJECT:

This paper presented the techniques and methods that can be used to predict crime and help law agencies. The scope of using different methods for crime prediction. and prevention can change the scenario of law enforcement agencies. Using a combination of ML and computer vision can substantially impact the overall functionality of law enforcement agencies. In the near future, by combining ML and computer vision, along with security equipment such as surveillance cameras and spotting scopes, a machine can learn the pattern of previous crimes, understand what crime actually is, and predict future crimes accurately without human intervention. A possible automation would be to create a system that can predict and anticipate the zones of crime hotspots in a city. Law enforcement agencies can be warned and prevent crime from occurring by implementing more surveillance within the prediction zone.

This complete automation can overcome the drawbacks of the current system, and law enforcement agencies can teck depend more on these techniques in the near future. De signing a machine to anticipate and identify patterns of such crimes will be the starting point of our future study. Although the current systems have a large impact on crime prevention, this could be the next big approach and bring about a revolutionary change in the crime rate, prediction, detection, and prevention, i.e., a "universal police officer".

CHAPTER-2

LITERATURE SURVEY

EXSITING SYSTEM:

Latest technical developments in sophisticated tools of data analytics and visualization are helping the society in numerous ways to investigate the info of social relevance. One among such socially relevant activities is crime details of various demographic places. The analysis of the crime data will help higher cognitive process agencies to require precautionary steps to regulate the rate over demographic places. Advancements within the field of knowledge technology, publicly available information and services, somehow help criminals to attain their misdeeds and involve them in much serious crimes than earlier. As a result, rate is increasing with a really high rate in developed and under-developed nations. supported the previous year crime details in Indian states, It present statistical models through Weighted Moving Average, Functional Coefficient Regression and Arithmetic-Geometric Progression based prediction of the crime in coming years. Difference between actual records and our predicted values for both years gives the accuracy of the proposed approaches between the range 85% and 90%. In future, this work are often modified by using Machine Learning (ML) models for forecasting crime, because the data points will sufficiently increase to use ML models. This will also increase the accuracy of the predictions. Further, statistical modeling's methods may also be clubbed with ML models so calculate weighted accuracy for a part, this will make the answer more robust. Machine learning is a computer system's method of learning by way of examples. There are many machine learning algorithms available to users that can be implemented on data sets. However, there are two major types of learning algorithms: supervised learning and unsupervised learning algorithms. Supervised learning algorithms work by inferring information or "the right answer" from labeled training data. The algorithms are given a particular attribute or set of attributes to predict. Data preprocessing process includes methods to remove any null values or infinite values which may affect the accuracy of the system. The main steps include Formatting, cleaning and sampling. Cleaning process is used for removal or fixing of some missing data there may be data that are incomplete.

Crimes Prediction ways:

 To utilize the resources identify the hotspots of crimes and allocate vigilante resources such as policeman, police cars, weapons etc.

PROPOSED SYSTEM:

The proposed system is made on the basis of the research work that is done by going through various such documentations. Nearly all of the crimes are predicting based on the location and the types of crimes that are occurring in those areas.On surveying previous works, Linear Regression, Decision Tree and Random Forest tend to give good accuracy so these models are used in this paper to predict crimes. The dataset used in this paper is from data.world.com. The data set contains different types of crimes that being committed in India according to the state and year respectively . This paper

takes types of crimes as input and gives the area in which crimes are committed as output. The data pre-processing involves data cleaning, feature selection, dropping null values, data scaling by normalizing and standardizing. After data pre-processing the data is free of null values which m ay alter the accuracy of the model significantly and feature selection is used to select only the required features that won't affect the accuracy of model.
After data pre-processing the models chosen i.e., Logistic Regression, Decision Tree and Random Forest are trained by splitting the data into as train and test data. As the output required is a categorical value classification models are used here. Python language is used for the data prediction.

APPLICATIONS:

Assume that burglaries of construction sites have been occurring in the southeastern portion of your jurisdiction for a two-month period over the same time frame, armed robberies of liquor stores have been occurring in the northwestern section of your city.
Identification of Evolving or Existent Crime Patterns.In the daily process of reviewing crime reports, the analyst looks for similarities in dates, times and methods by which crimes. are committed. He or she also notes any similarity in suspect descriptions. This procedure identifies a particular suspect who may be responsible for the commission of several offenses. After report examination, the analyst collates and analyzes the data in an attempt to define recurring patterns of criminal activity. If a pattern is noted, relevant information is disseminated to patrol or special tactical action units.
The analyst then continues to monitor the pattern until the pattern ceases, usually because the suspect leaves the area or is arrested.In this example, the analyst notes that, of the construction site burglaries committed, there is nothing to suggest they are being com. mitted by the same person(s). In some instances only lumber was taken; in others, appliances were stolen and lumber was left behind. Some burglaries occurred on weekends; others occurred at night 14 during the week. Witnesses also described variety of suspects.
Some were male white juveniles while others were black adults. At this point, the analyst would say that a crime pattern existed. A crime pattern is the occurrence of similar offenses in a defined. geographic area, either a single reporting district, a beat, or an entire. city. In this case, the analyst has a number of construction site burglaries committed in the same part of town but there is nothing to suggest that the same person or persons are responsible. The armed robberies, however, are another matter. In one of the robberies ,a teenage white suspect used a knife. In two others, the suspects simulated a weapon. One of the suspects was a black male Hispanic adult. The robberies occurred on different and the other a Hispanic adult. days of the week.In four other robberies, however, an adult white male used a chrome plated revolver. This suspect made his victims kneel on the floor. opened the register himself, removed cash and then told his victims to "wait five minutes before calling the cops."He committed all robberies on Wednesdays between nine and eleven o'clock in the evening.Since a number of robberies (same type of crime) were occurring in the northwestern part of town (same geographical area), it would be appropriate to state that a crime pattern was in evidence. However, in the case of the four robberies wherein the chrome plated handgun was used, the analyst would state that he or she was observing a crime series. A crime series is a crime pattern wherein there is reason to believe that the same person or persons are responsible.In the four robberies committed on Wednesdays, the MO in each case was the same and all victims gave similar suspect descriptions.
The same person is probably responsible for committing these crimes.

SUMMARY:

Few crime analysts are able to implement a crime analysis program without encountering some resistance. Few people really like change: police officers are no exceptionWe all tend to feel secure in our comfort zones and don't really appreciate being pulled out of them. Control-oriented individuals may have an especially hard time dealing with change, and this is not hard to understand.Police officers tell people what to do all day long and usually get their way. If they determine someone is going to get a ticket or go to jail. that happens. It doesn't take long for officers to become so accustomed to controlling others that they become resistant to being controlled.
They will accept control from departmental executives and. supervisors (usually), but seldom from anyone else. That this creates 170 problems in their interpersonal relationships is understood by police psychologists throughout the nation.
Officers may see the analyst as yet another person trying to control their actions and shake them out of their comfort zones. You and your program may initially be perceived as tools the department will use to "fix what ain't broke." For the sake of your own self-esteem, realize that this interplay of personal dynamics creates resistance to programs.Expect it, prepare for it, and do not take it personally Be yourself. Learn from the officers, provide them with what they want and need and you will eventually gain the acceptance you desire.

REALTED WORK

In crime prediction is done on Chicago data set in which various machine learning models are used. Comparison of models like KNN, Naïve Bayes, SVM is done this paper. It is seen that prediction varies depending upon the dataset and features that have been selected. The prediction accuracy found in [1] is 78% for KNN, 64% for Gaussian NB, 31% for SVC.

Auto regressive integrated Moving average models was used in to make machine learning algorithms to forecast crime trends in urban areas.

One of the major problems in crimes is detecting and analyzing the pattern of crimes.

Understanding datasets is also an important concept in this case. We surely want to accurately predict so that we don't waste our resources due to false signals.

In paper, Algorithms like KNN and neural networks are developed, tested and crime prediction is done on San Francisco.

It is observed that many machine learning models are implemented on datasets of different cities having unique features, so predictions are different in all cases.

Classification models have been implemented on various other application like prediction of weather, in banking and finances also in security.

Most of the research in crime prediction is finding the location of crimes and doing analysis based on proposed area-specific models using geographical data.

Based on the review and studying previous work, KNN classification and decision tree models is shown to be giving high accuracy so we choose to use the same to predict crimes in Vancouver city.

ADVANTAGES:

- Handling multi-dimensional and multi-variety data.
- No human invention needed(automation).
- Continuous improvement.
- Wide application.
- Easily identifies trends and patterns.

DISADVANTAGES:

- Data acquisition.
- Time and resource.
- Interpretation of results.
- High error susceptibility.

CHAPTER-3

SOFTWARE REQUIREMENTS SPECIFICATIONS

SOFTWARE REQUIREMENTS:

Android sdk

Eclipse Ganymede IDE

OPERATING SYSTEM: window 11

CODING LANGUAGE:machine learning,python

EDITOR:vs code

HARDWARE REQUIREMENTS:

PROCESSOR:(min) processor p5

SYSTEM:intel i3 core.5^{th} gen

RAM:4GB

HARD DISK:8OGB

MONITOR:15”LED

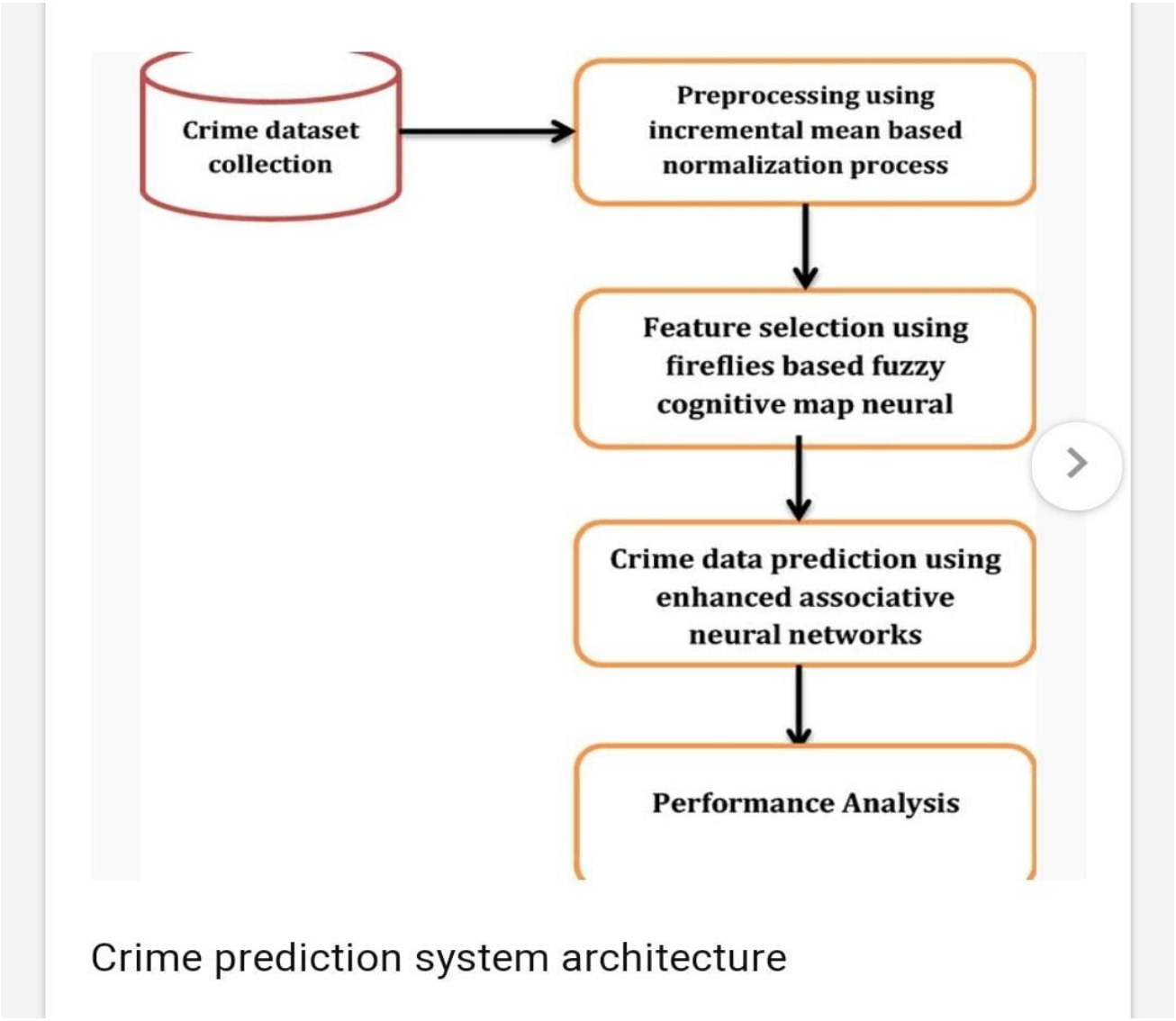
MOUSE:HP X1000

KEYBOARD:HP KEYBOARD

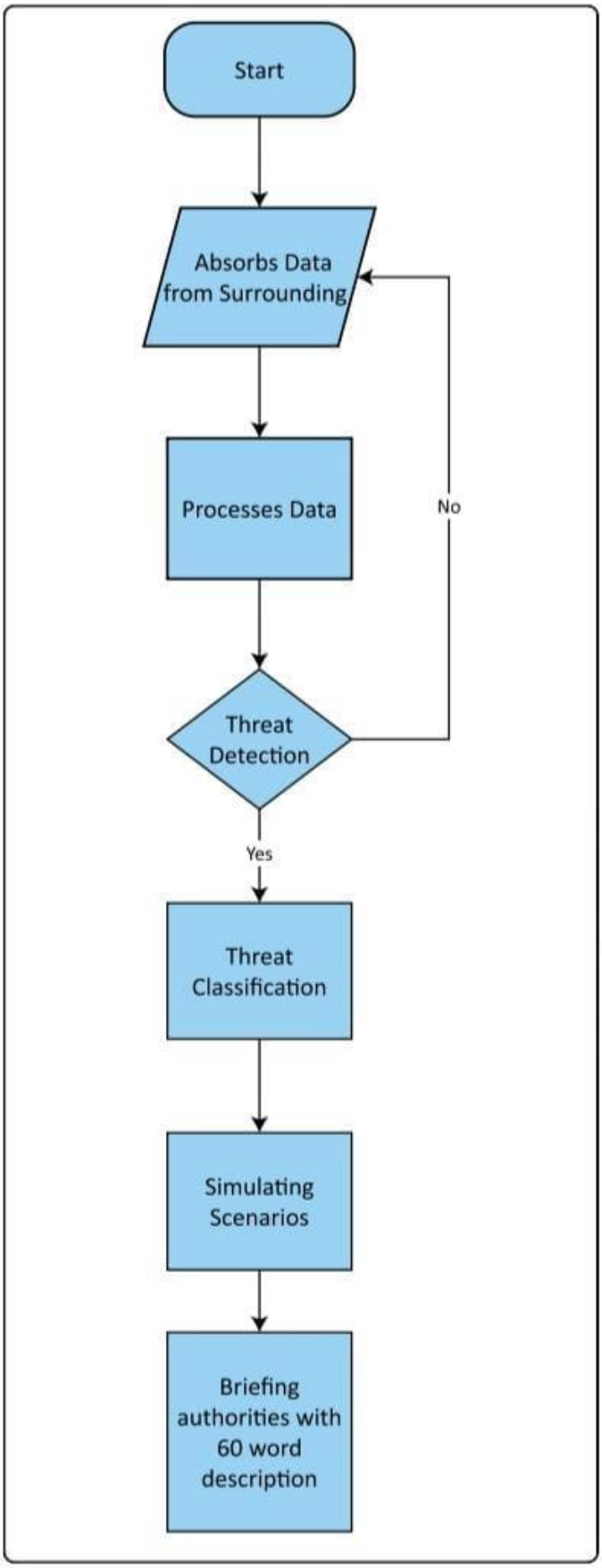
CHAPTER-4

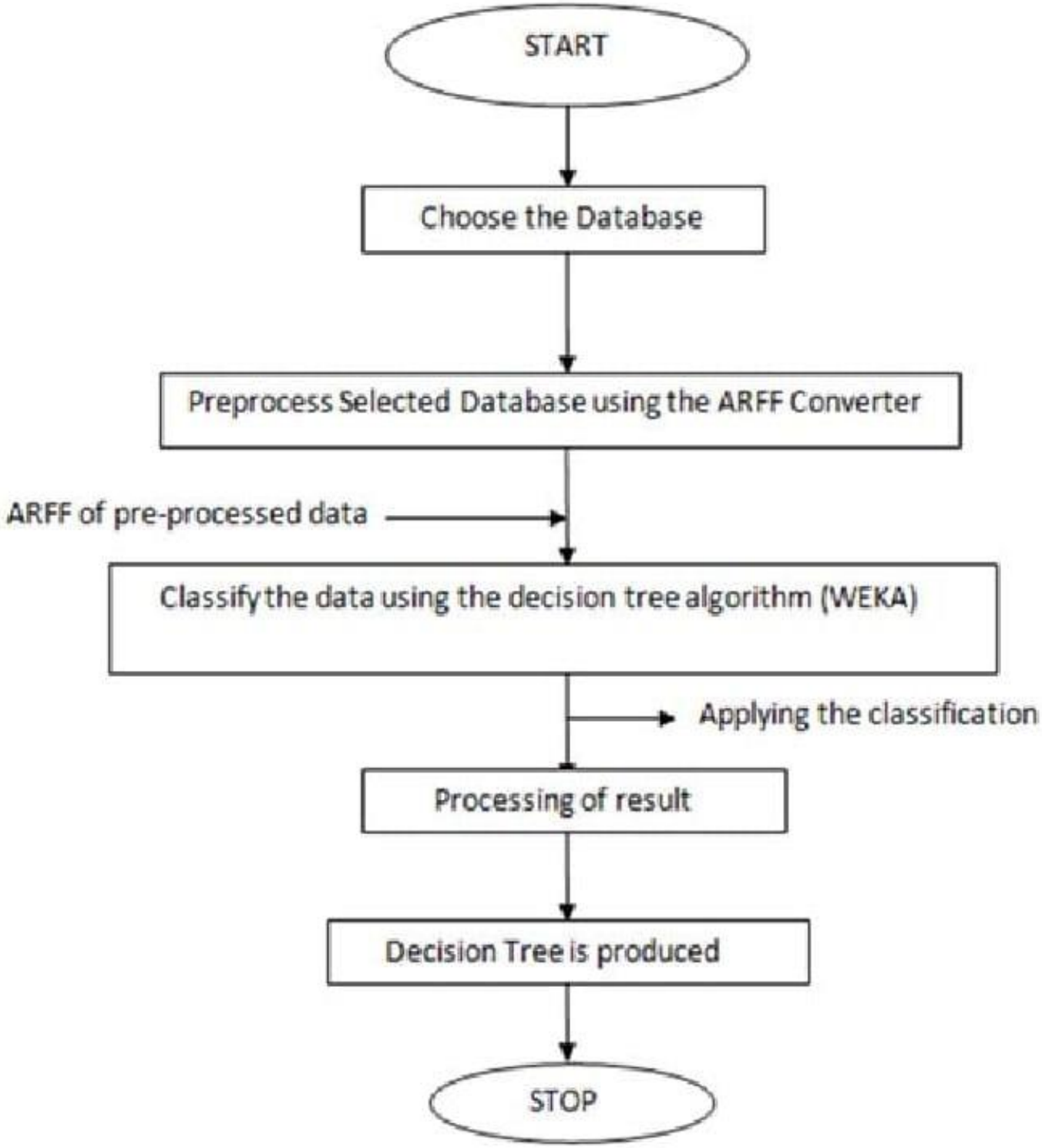
SYSTEM DESIGN

SYSTEM ARCHITECTURE:



Crime prediction system architecture


Start
Absorbs Data from Surrounding
Processes Data
Threat Detection
No
Yes
Threat Classification
Simulating Scenarios
Briefing authorities with 60 word description


START
Choose the Database
Preprocess Selected Database using the ARFF Converter
ARFF of pre-processed data
Classify the data using the decision tree algorithm (WEKA)
Applying the classification
Processing of result
Decision Tree is produced
STOP

CHAPTER-5

IMPLEMENTATION

The dataset used in this project is taken from Kaggle.com. The dataset obtained from kaggle is maintained and updated by the Chicago police department The Implementation of this project is divided into following steps

1.Data collection Crime dataset from kaggle is used in CSV format

2. Data Preprocessing10k entries are present in the dataset. The null values are removed using df = df. drop na () where df is the data frame. The categorical attributes (Location, Block, Crime Type, Community Area) are converted into numeric using Label Encoder. The date attribute is splitted into new attributes like month and hour which can be used as feature for the model.

3 Feature selection Features selection is done which can be used to build the model. The attributes used for feature selection are Block, Location, District, Community area, X coordinate, Y coordinate, Latitude, Longitude, Hour and month

4. Building and Traning Model After feature selection location and month attribute are used for training. The dataset is divided into pair of x train, y train and x test, y test. The algorithms model is imported form skleran. Building model is done using model. Fit (x train, y train).

5. Prediction After the model is build using the above process, prediction is done using model predict (x test). The accuracy is calculated using accuracy score imported from metrics – metrics accuracy score (y test, predicted).

ENVIRONMENTAL SETUP:

Crime Prevention Through Environmental Design (CPTED) is a set of design principles used to discourage crime and promote building security. The concept is simple: Buildings and properties can be designed to prevent damage from the force of the elements and natural disasters; they should also be designed to prevent crime.

CPTED principles are based on anticipating the thought processes of a potential offender and creating an environment that discourages follow-through. CPTED has the added advantage of creating a sense of security and well-being among employees and tenants.When CPTED is put into practice, the resulting environment, including the building and its surroundings, will discourage or impede criminal behavior, and at the same time encourage honest citizens to keep a watchful eye.

The four main principles of CPTED are:

Natural surveillance

Natural access control

Territorial reinforcement

Maintenance

Although these principles were developed for the design and construction of new buildings, the concepts can be applied to existing businesses as well.

MODULE DESCRIPTION:

3.2.1 Module 1: Visualization of Crime Data Using Google Maps This module extracts the recent crime data from the dataset and based on longitude-latitude it tags the specific location of the city. This tagging also displays the crime location name, the type of crime that happened. This information is useful for an individual in knowing dangerous and risky areas and it thus can help them to avoid such areas. The picture can help the law enforcement to improve the security in the areas. Fig.1 shows that locations, where crimes occurred, are very near to each other. From this, we can analyze that if a location feasible to a criminal attack, then the nearby locations are also feasible for the crime to occur. This module also provides the facility to enquire about a specific location to show what type of crime is feasible to happen in that location.

Fig.1.Visulaization of Crime Data Using Google Maps.

3.2.2 Module 2: Visualization of Exact Location of Crime with 3D View This module visualizes the area where the crime has happened exactly. This helps the law enforcement to analyze the security

measures of an area. The module provides the interactive image which takes help of Google Maps to navigate around the crime location and it can help the analyst to analyze the security of an area, also what locations can be the target for next attack. This also helps the police for the clear understanding of the cause of crime and helps them to investigate the location by not visiting the location again and again. By just clicking the tag in Fig.1, it provides the realistic 3D interactive image of the location and helps in navigation around the location as shown in Fig.2.

Fig.2. Visualization of Exact Location of Crime with 3D View.

3.2.3 Module 3: Visualization based on type of Crime The type of crime is also an important factor as safety measures are majorly taken based on the type of crime. This module helps visualize the crimes that had happened based on category over different areas as shown in Fig.3. This helps the law enforcement to analyze what type of crimes are frequently happening in an area and helps them to improve security measures based on the type of crimes.

The type of crime is also an important factor as safety measures are majorly taken based on the type of crime. This module helps visualize the crimes that had happened based on category over different areas as shown in Fig.3. This helps the law enforcement to analyze what type of crimes are frequently happening in an area and helps them to improve security measures based on the type of crimes.

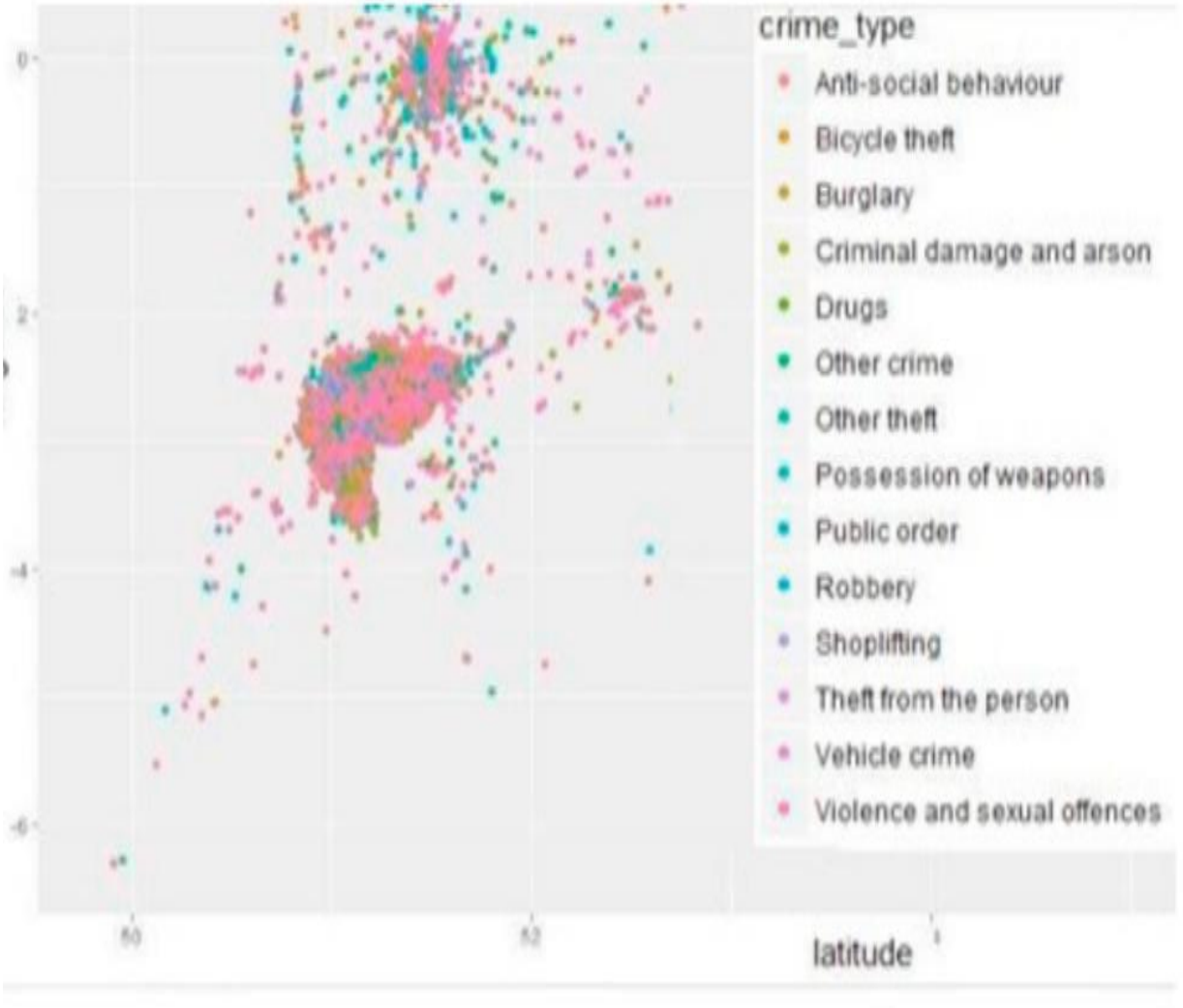


Fig.3. Visualization based on type of Crime.

3.2.4 Module 4: Visualization of Crime Hotspots The number of crimes happened in an area makes sense of how dangerous the area is. This module helps to visualize the crime hotspots as shown in the Fig.4. The areas on the map that have high crime density are called the crime hotspots [10]. Developing maps that contain hotspots are becoming a critical and influential tool for policing. These are used by the researchers and analysts to examine the occurrence of hotspots in certain areas and why they happen and help them to build the theories. This also allows researchers to explain why crime occurs in certain places and why crime does not in other places. Crime analysts can use these to make better decisions, target resources, formulate strategies and help the law agencies.

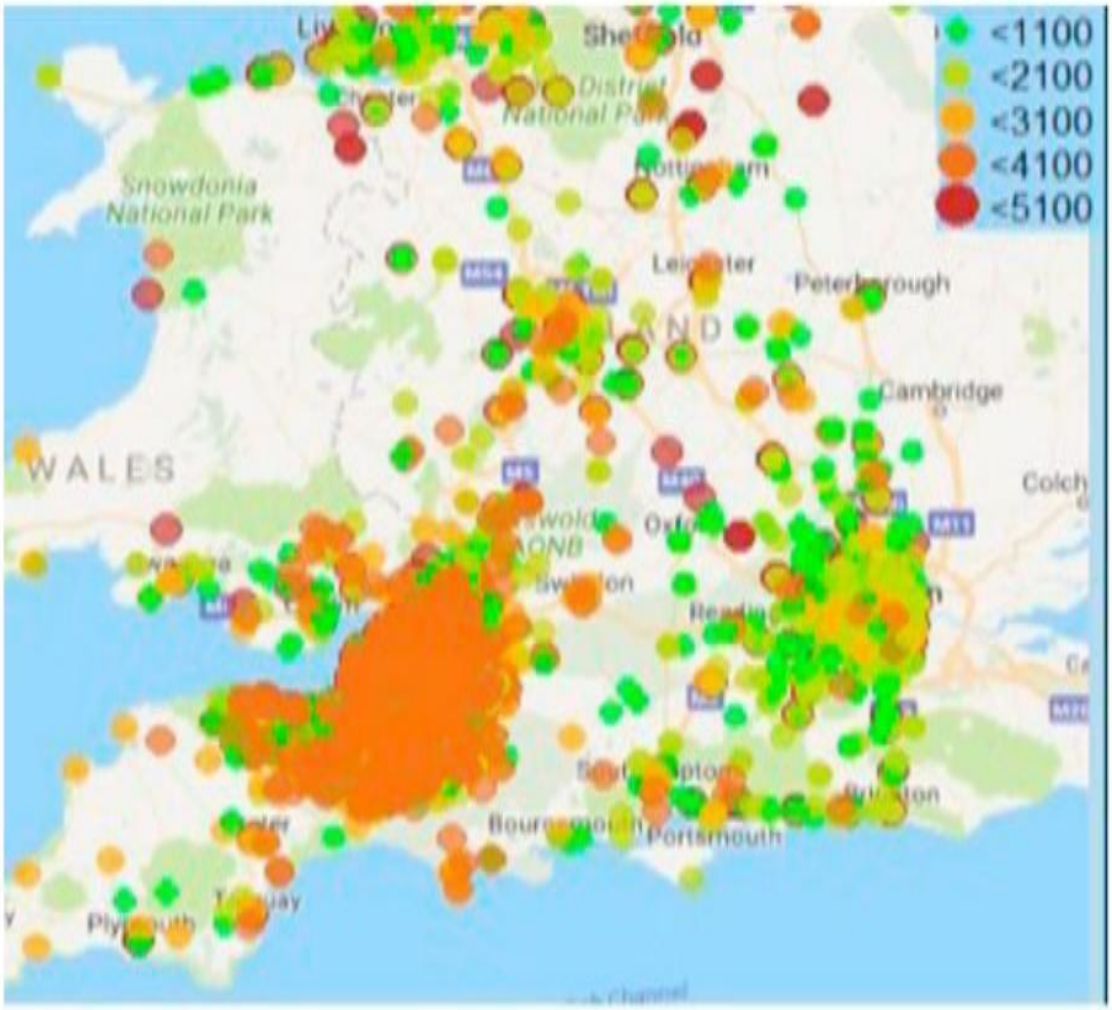


Fig.4. Visualization of Crime Hotspots

3.2.5 Module5:crime frequency report

This module helps to generate the crime report based on the number of crime that happened in every month and on different categories of crimes as show in fig5.this can help the public to take safety measures and help the crimes analysts to check which type of crimes are increased or decreased.

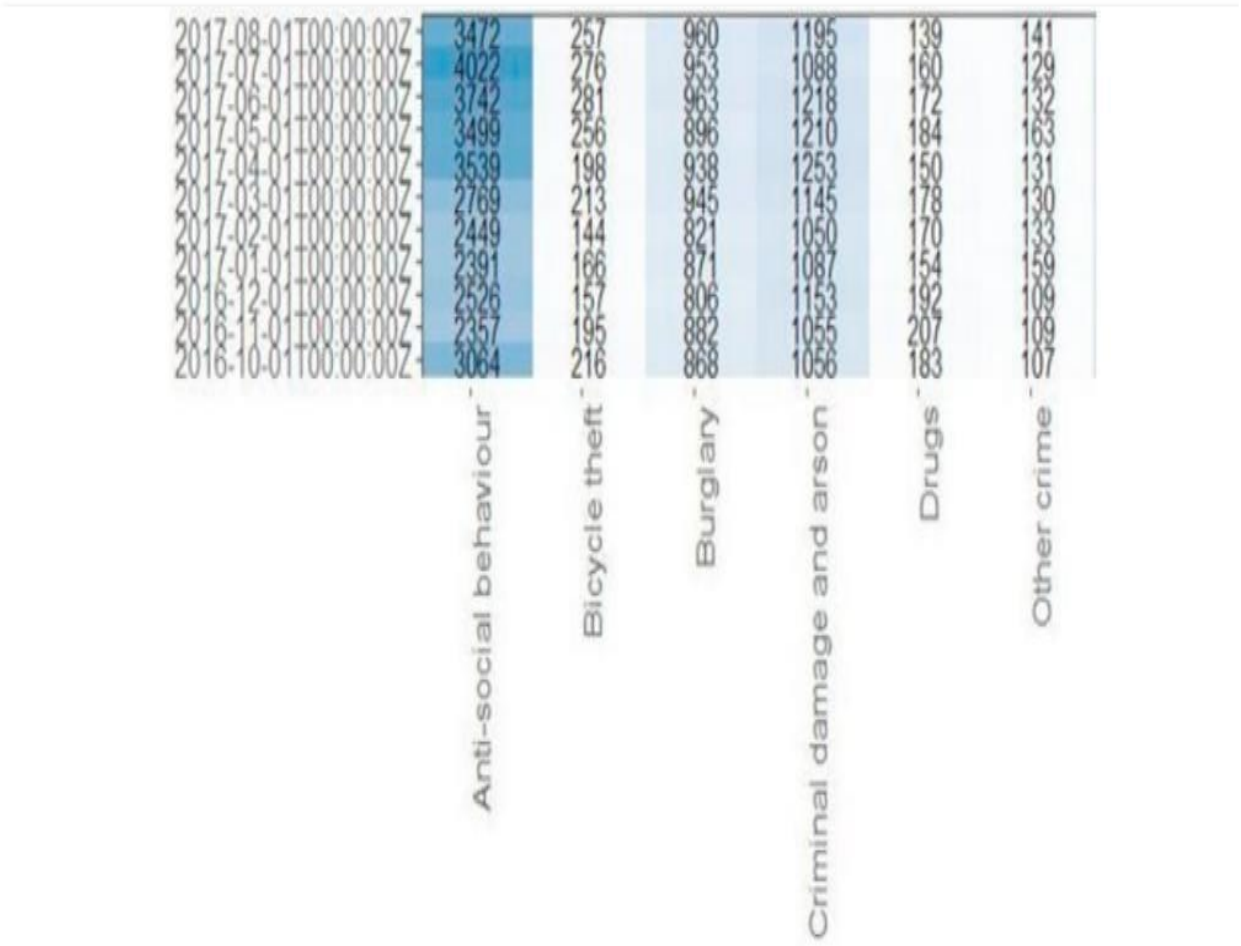


Fig.5. Crime Frequency Report

3.2.6 Module6: Interactive crime frequency report using and bar chart:
This module helps to generate a video representation of the trend of the type of crimes in every month that is extracted from the datasets. This helps to visualize that which type of crime has increased or decreased compared to previous months. Fig.6 represents the bar-chart, when the video is played it shows how the frequency of each crime is changing in every month (the bars move up and down depending on the crimes have increased or decreased) and Fig.7 represents the graphical representation in the bar chart. This module helps the analysts to understand the trend of every crime that has happened in an area. Fig.6 shows that there are 3523 Anti-social behaviour crimes and around 1000 burglary crimes are reported on 13-April-2017. Fig.7. represents the graphical representation of the bar chart. This module helps the analysts to understand the trend of every crime that has happened in an area. It shows that there are 3539 Anti-social behaviour crimes (the blue line) are reported on 1-January-2017. It also displays that Anti-social behaviour is the most frequent crime in every month. The Burglary cases (the red line) have increased majorly in every month, and remaining crimes are tending to be constant.

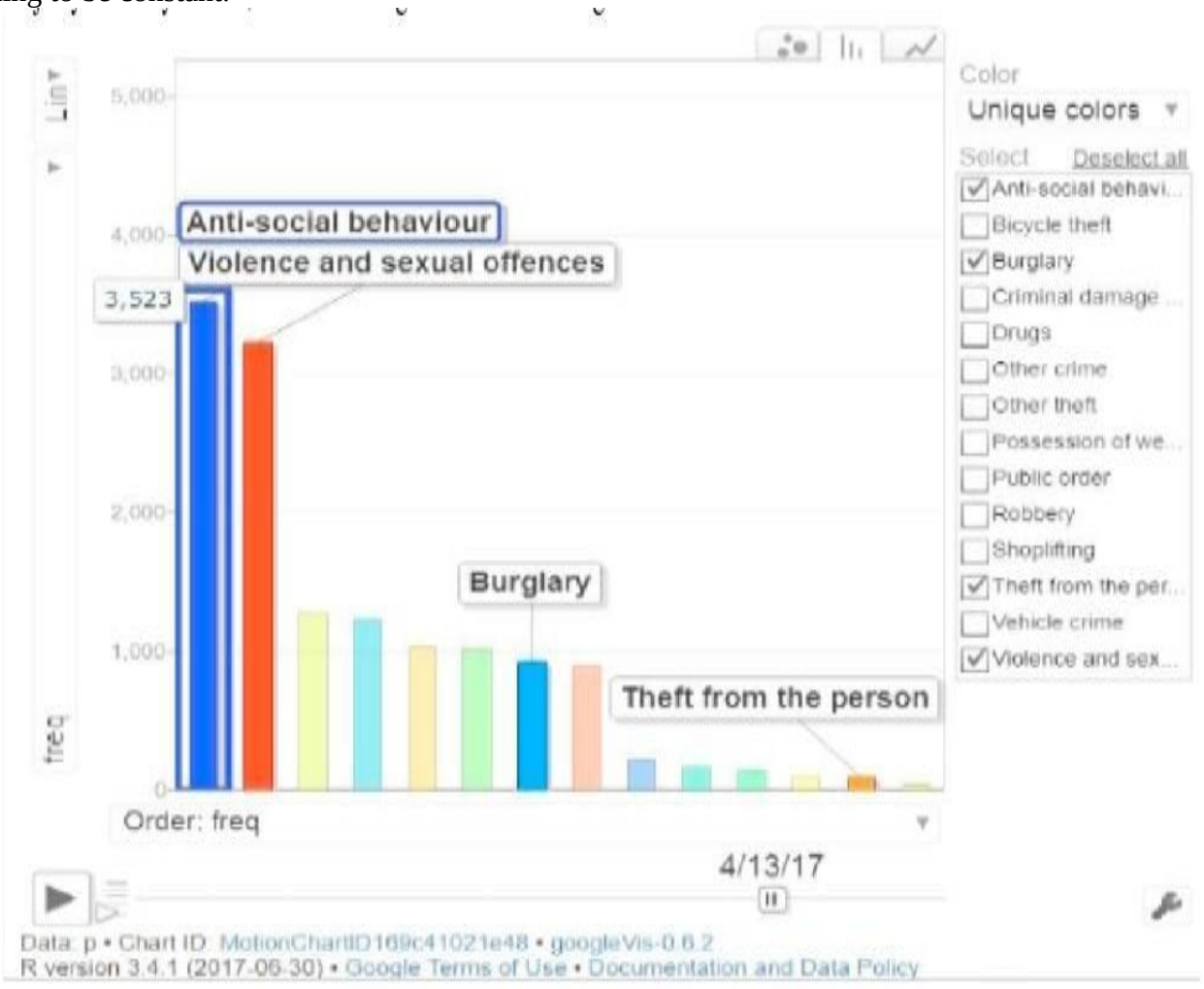


Fig.6. Bar chart between crime frequency and time (month).

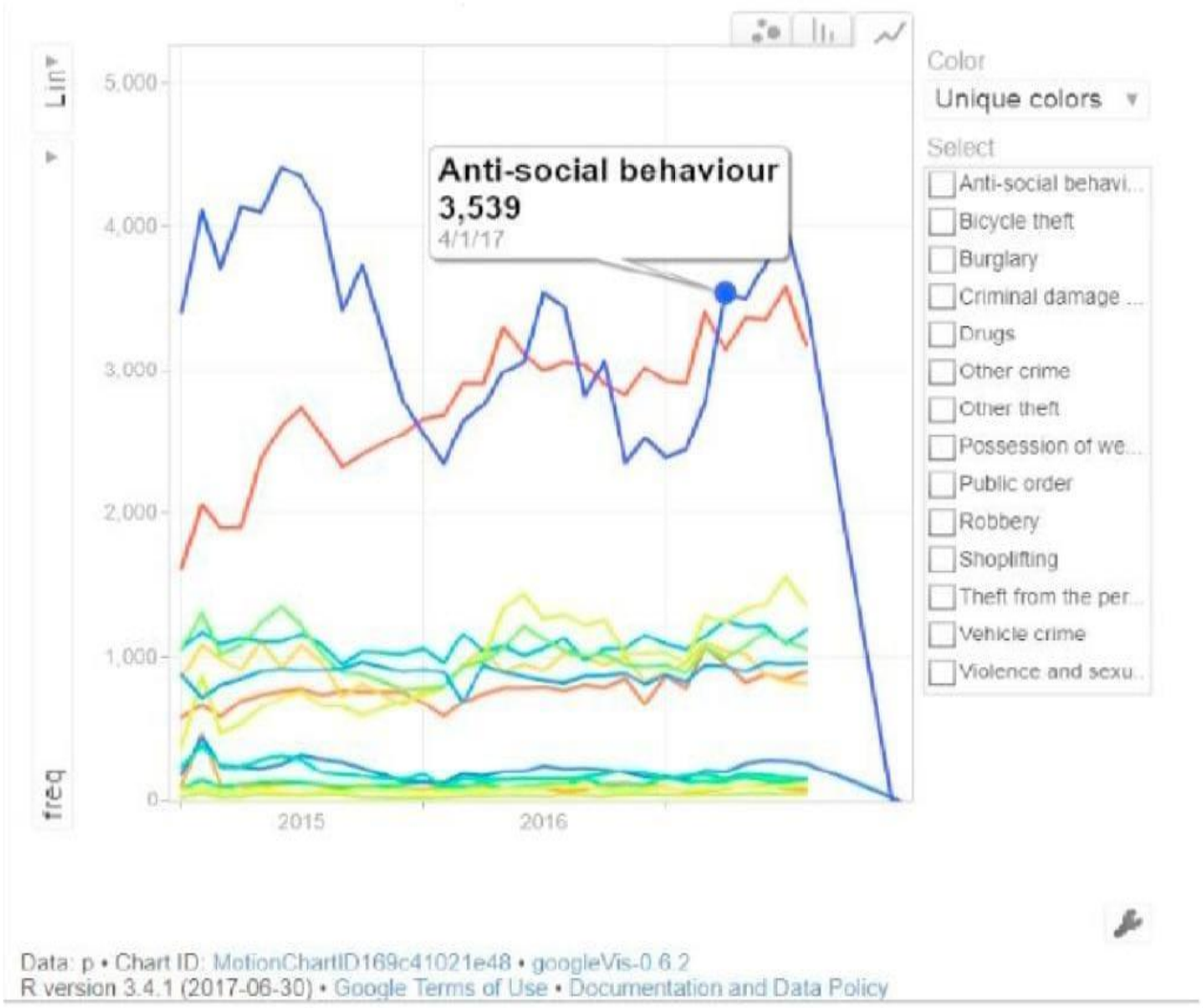


Fig.7. Month wise crime trend report

SOFTWARE DESCRITION:

- The human rights group said it had sent a total of 90 Freedom of Information requests out last year to discover which forces used the technology.

- It believes the programs involved can lead to biased policing strategies that unfairly focus on ethnic minorities and lower-income communities.And it said there had been a "severe lack of transparency" about the matter.

    - Defenders of the technology say it can provide new insights into gun and knife crime, sex trafficking and other potentially life-threatening offences at a time when police budgets are under pressure.

- One of the named forces - Avon and Somerset Police - said it had invited members of the press in to see the Qlik system it used in action, to raise public awareness.

- "We make every effort to prevent bias in data models," said a spokeswoman.

- "For this reason the data... does not include ethnicity, gender, address location or demographics."

- But Liberty said the technologies lacked proper oversight, and moreover there was no clear evidence that they had led to safer communities.

- "These opaque computer programs use algorithms to analyse hordes of biased police data, identifying patterns and embedding an approach to policing which relies on discriminatory profiling," its report said.

- "[They] entrench pre-existing inequalities while being disguised as cost-effective innovations." Analysis of criminal intelligence is a vital component of any law enforcement agency, demanding complex investigative efforts & tools for their cyber forensics.

- Extracting crucial insights from web data can be the turnkey to solving investigations, but the extraction process can demand tedious efforts. It is critical to reveal and discover links between threat actors, identity locations and generate full reports on the criminal investigation.

- Cobwebs' crime prediction software grants law enforcement agencies the power of artificial intelligence (AI) technology to effectively manage uncertainty and new challenges with advanced digital forensics.

SYSTEM ANALYSIS

System analysis may be defined as examination and processing of information that resultsdevelopment of recognizable patterns of criminal activity and in the identification of offenders. It is of two types: Modus Operandi (MO) Pattern Detection and Correlation Statistical Analysis Modus Operandi (MO) Pattern Detection and Correlation refers generally to using mapping techniques or searching the various [manual or automated] files in order to determine if similar offenses or a crime series are occurring or to correlating cases once an arrest is made, 12 Statistical analysis is done to mathematically establish likely times, locations and probabilities of future criminal events. It may also be used to determine relationships between events. Both types of analysis must be used if the analyst is to provide officers and investigators with accurate information.In this chapter we will focus on Modus Operandi (MO) Pattern Detection and Correlation analysis techniques. Part Two of this book presents a complete discussion of statistical analysis techniques. The use of both types of analysis techniques will not only enable you to comprehensively analyze your data, but will provide validity and integrity to your conclusions.Types of Crime Patterns:The task of analyzing data and drawing conclusions from them is much like putting a jig-saw puzzle together. The analyst is working with many data elements that have been gathered from a wide variety of sources. The goal is to combine the data so that a true "picture" of criminal activity can be determined. These pictures main fest: themselves in the form of crime patterns. As described below, they are of two types: Geographic concentration patterns refer to patterns identified on the basis of similarity of crime type (e.g.. commercial burglary) • multiple occurrences in well-defined geographic area. Identification of a geographic concentration pattern (known more simply as a "crime pattern") means only that certain types of similar crimes are frequently occurring in particular areas of the jurisdiction. Without any other similarities or relationships between these crimes. there is no reason to believe that the same person or persons are responsible for their commission.Specific and recurring MO patterns refer to patterns identified on the basis of similarity of suspect and/or suspect vehicle description(s)unique MO characteristics.Once a specific and recurring MO pattern has been identified, it is called crime series. A crime series is characterized by the presence of sufficient similarities to give the analyst reason tobelieve that the same person or persons are responsible for committing each crime in the series.To demonstrate the difference between a crime pattern and a crime series, assume that murders have been occurring in a particular section of town. In one, a transient was stabbed to death in an alley. In another, a suspect entered a residence and shot a victim. In still another, a victim was strangled in a motel room. Some murders. occurred during the day, others occurred at night; witnesses gave varying descriptions of suspects.

This is an example of a crime. pattern. The only similarities between the crimes is crime type (murder) and the geographical location in which they occurred.

ANALYZING PATTERS OF LOCATION:

Geographic analyses provide comparison which may not be easily determined from simply looking at addresses on crime reports.Further when a pin is used to indicate the location of one crime OR a map, the relationship of that crime location to other similarly illustrated crimes becomes apparent.Geographic analysis can be performed by mapping.And by graphical and statistical methods. Mapping techniques involve the use of a map to represent the actual geographical relationship between criminal events. A pin map displaying day or night residential burglaries over

a given period of time is one example of a mapping technique. Graphical methods include compare Dns of prescribed geographical areas with crime data.A series of bar graphs, e~{ch listing percentages for particular crimes in specific areas over a given time period, gives the analyst an overall picture of those crimes throughout a jurisdiction.Statistical methods of geographical data allow the analyst to identify crime patterns in beats, reporting areas or entire jurisdictions. Manual mapping techniques include pin maps, symbol maps, dot maps or any other means for depicting crimes by location. The number of data elements to be recorded on a map often dictates the most effective technique.If the crime report number is to be recorded with the location of the offense, a colored dot or flag pin (using colored heads) map can be utilized. The colors indicate the type of crime. The report number should be written on the dot or the flag. If particular MO factors are to be recorded, various color-coded and marked pins can be employed to identify specific MO patterns. The length of time a particular map or set of maps is to be maintained is also important. No absolute set of rules has been established regarding map maintenance time, but maps are best maintained for one to three months.Two maps may be maintained for each type of crime (one reflecting weekly or monthly information and the other year-to-date or quarterly information).Map data should be stored so that it is easily retrieved.Pin maps can be recorded by pbotograp1Jic means; using two))fojectors to superimpose one figure over another i" often effective. Color coded paper dots on ucetate overlays of area 1!1apS facilitate data rec:overy for law enforc:ement agencies whose access to equipment is limited. Computer produce,d and coded area maps can be easily stored and retrieved by today's microcomputer systems. Seven is the minimum number of maps most crime analysis units should maintain. Five crime types are well suited to geographical analysis: burglary, robbery, special theft classifications, auto theft and sex crimes. Two of these types (burglary and robb~ry) have logical sub classification(commercial and residential for burglary, and armed and unarmed for robbery). Three other informational maps have been found to be quite useful.One is designed to suppon career criminal ~apprehension programs by mapping the last known residences and specialties of known offenders. Another supports narcotics operations by indicating locations of crack and crash "pads", street corners where narcotics are sold and so on.A third map may be used to depict locations where stolen propeny is "fenced". Communities e plagued by gang activities may find a map relating to gang "turfs" helpful in deploying field resources, identifying criminal groups (and individuals), and anticipating future inter-gang conflict.

- A department's ability to "read" spray-painted graffiti provides the basic data for such maps.

Graphical analysis compares prescribed geographical areas to a body of compiled crime data.
Compiling raw data into information which can be plotted on graphs is extremely useful for identifying crime trends.
This method is well-suited to comparing criminal activity of a given type across different geographical areas.
Two examples of geographic-graphical analysis are illustrated in Figures 34 and 35. Figure 34 provides a cpmparison of differing types of burglaries in a 10 beat jurisdiction.
This type of analysis is useful (when performed on as short a time basis as possible-weekly, rather than monthly) for identifying crime trends promptly.
Figure 3S illustrates a graphic useful for determining the relationship of ...
the present year's crime to crime of previous years in a given geographical.

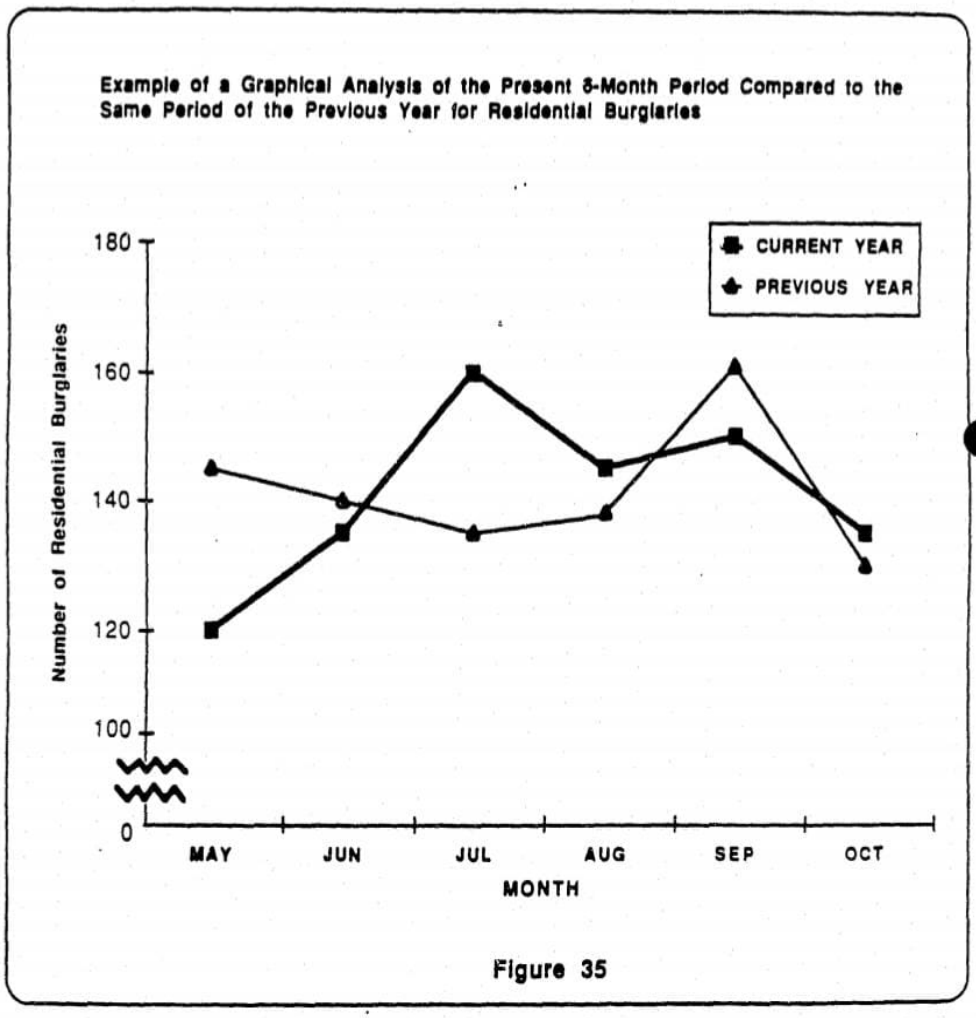


Figure 35

PROGRAM

```
import numpy as np
import pandas as pd
df=pd.read_csv(r'file name')
df.head()
df.shape
(9017,33)
a=df.groupby(['STATE/UT'])['MURDER'].sum().reset_index()
b=df.groupby(['STATE/UT'])['RAPE'].sum().reset_index()
import plotly.express as px
fig=px.pie(a,  values='murder',
                names='state/ut',
                title='STATE WISE TOTAL MURDERS 2001-2012',
                color_discrete_sequence=px.colors.sequential.viridis)
fig.show()
df[['STATE/UT','MURDER','RAPE','KIDNAPPING
&ABDUCTION']].groupby("STATE/UT").sum().plot.barh(stacked=true,figsize=(20,10));
df.innull()
df.innull().sum()
from sklearn.preprocessing import labelencoder
new=['STATE/UT']
le=labelencoder()
for I in new;
df[i]=le.fit_transform(df[i].astype(int)
x=df.drop(['DISTRICT','TOTAL IPC CRIMES'],axis=1)
y=df['TOTAL IPC CRIMES']
```

RANDOM FOREST REGRESSOR

```
From sklearn.ensemble import randomforestregressor
regressor=randomforestregressor(n_estimators=110,random_state=10)
regressor.fit(x_train,y_train)
Randomforestregressor(n_estimators=110, random_state=10)
Y_pred1=regressor.predict(x_test)
from sklearn import metrics
print("mean absolute Error:, metrics.mean absolute_error (y_test, y_pred1))

print('Root Mean Squared Error:', np.sqrt(metrics.mean_squared_error(y_test, y pred1)))) print('Mean
Squared Error: metrics.mean_squared_error(y_test, y pred1)))
```

Mean Absolute Error: 234.99115097762547

Root Mean Squared Error: 926.3706132406733

Mean Squared Error: 858162.513075901

K-NEAREST NEIGHBOR REGRESSOR

```
 from sklearn.neighbors import KNeighbors Regressor
knr = KNeighbors Regressor (n_neighbors=2, p=2, metric='minkowski') knr. Fit (X_train, y_train)
KNeighbors Regressor (n_neighbors=2)
 y pred = knr.predict(X_test)
 print('Mean Absolute Errar:', metrics.mean_absolute_error(y_test, y_pred))

print('Root Mean Squared Error:', np.sqrt(metrics.mean_squared_error(y_test, y_pred))) print("Mean
Squared Error:', metrics.mean_squared_error(y_test, y pred))
Mean Absolute Error: 217.89717294900223
Root Mean Squared Error: 809.3113902655302
Mean Squared Error: 654984.9264135255

ADABOOST
from sklearn.ensemble import AdaBoostRegressor
regr = AdaBoost Regressor(random_state=0, n_estimators=10)
regr.fit(x_train, y_train)
AdaBoostRegressor(n_estimators=10, random_state=0)
y pred2 =Type here to search
regr.predict(x_test)
y_pred2=regr.predict(X_test)
print('Mean Absolute Error:,metrics.mean_absolute_error(y_test,y_pred2))

print('Root Mean Squared Error:', np.sqrt(metrics.mean_squared_error(y_test, y_pred2))) print("Mean
Squared Error:', metrics.mean_squared_error(y_test, y_pred2))

Mean Absolute Error: 1998.2403802594522 Root Mean Squared Error: 2877.9308432576668 Mean
Squared Error: 8282485.938573785
```

OUTPUT:

	STATE/UT	DISTRICT	YEAR	MURDER	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	RAPE	CUSTODIAL RAPE	OTHER RAPE	KIDNAPPING & ABDUCTION	...	ARSON	HURT/GREVIOUS HURT	DOWRY DEATHS	ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY	INSULT TO MODESTY OF WOMEN	C HU REI
0	ANDHRA PRADESH	ADILABAD	2001	101	60	17	50	0	50	46	...	30	1131	16	149	34	
1	ANDHRA PRADESH	ANANTAPUR	2001	151	125	1	23	0	23	53	...	69	1543	7	118	24	
2	ANDHRA PRADESH	CHITTOOR	2001	101	57	2	27	0	27	59	...	38	2088	14	112	83	
3	ANDHRA PRADESH	CUDDAPAH	2001	80	53	1	20	0	20	25	...	23	795	17	126	38	
4	ANDHRA PRADESH	EAST GODAVARI	2001	82	67	1	23	0	23	49	...	41	1244	12	109	58	

[5]:

	STATE/UT	MURDER
0	A & N ISLANDS	320
1	ANDHRA PRADESH	63512
2	ARUNACHAL PRADESH	1682
3	ASSAM	30864
4	BIHAR	82490
5	CHANDIGARH	486
6	CHHATTISGARH	24120
7	D & N HAVELI	216
8	DAMAN & DIU	136
9	DELHI UT	12410
10	GOA	974
11	GUJARAT	27550
12	HARYANA	20942
13	HIMACHAL PRADESH	3908

ed

[6]:

	STATE/UT	RAPE
0	A & N ISLANDS	218
1	ANDHRA PRADESH	26958
2	ARUNACHAL PRADESH	1000
3	ASSAM	32356
4	BIHAR	26248
5	CHANDIGARH	562
6	CHHATTISGARH	23676
7	D & N HAVELI	112
8	DAMAN & DIU	40
9	DELHI UT	12848
10	GOA	700
11	GUJARAT	8498
12	HARYANA	12820
13	HIMACHAL PRADESH	3608
14	JAMMU & KASHMIR	5620
15	JHARKHAND	18318
16	KARNATAKA	10348
17	KERALA	14894

CHAPTER-6

SYSTEM TESTING

Our experiment choose the algorithm are
- Instance based algorithm
- Decision tree
- Linear regression
- K-means algorithm

1. Instance Based Algorithm

-The instance based algorithm is also called as tge machine based learning is a family of learning algorithm that, instead of performing explicit generalization, compares new problems instances with instance seen in training, which have been stored in memory. These stored their training set when predicting a value or class for a new instances, they compute distance training instances to make a decision.

The algorithm in this category for numerical prediction can divided into two types: similarity- based, e.g., Euclidean or entropy based and regression-based e.g., LWL Since regression is one of the most popular methods for numerical prediction[1].

The advantages of the Instances based Algorithm is it over other methods of machine learning is its ability to adapt its model of machine learning is its ability to adapt its model to previously unseen data. Instance based learners may simply store a new instance or throw an old instance away. The Disadvantages of the instances based Algorithm are its need more storage and computational complexity.

2. Linear Regression

-It is simple form of regression. Linear regression attempts to model the relationship between the two variables by fitting a linear equation to observe the data. this is widely used in statistics. For this purpose ,linear functions are used for which the unknown parameter i.e., weight of the independent variables, are estimated from the training data[1].this can be used to predict the values One of the most common estimating method is least mean square.

Linear regression algorithms for predicting include simple regression multiple regression and pace regression, which is suitable for data of high dimensionality and only accepts binary nominal attributes.[1].

The main advantages of the linear regressions is gain a far greater understanding of the variables that can impact its success in the coming weeks,months and years into the future.The disadvantages of the regression Is its linearity.If the data has non linear dependencies, a linear regression model will ouput the best fitting line which may not fit very well.

3. Decision Tree

Decision tree is used for both the prediction and classification. for the classification purpose a function can
be learned this is intervals defined by splits on the individuals attributes e decision trees algorithm for classification have been adapted to output a
numerical value the main differenceIs that the leaves of the tree have numerical values, unlike classification trees have class labels.

Advantages of the decision trees are It is very simple to understand and help determine worst,best and expected values for different scenarios.it can be combined with other decision techniques.Some of the Disadvantages of the Decision tree are They are unstable, They are often relatively inaccurate, Calculation
can get vey complex.

4. K-Means Algorithm

:K –means is the simplest and most commonly used portioning algorithm among the clustering algorithm in
scientific and industrial software[3].Acceptance of k means is mainly due to its being simple .This algorithm
is also suitable for clustering of a large datasets since it has much less computational complexity grows linearly by increasing of the data points.

Advantages of the k-means algorithm are relatively simple to implement, Scales to large dataset, Guarantees
convergence, easily adapts to new examples. Disadvantages of the k-means algorithm are Choosing manually, Being dependent on initial values, clustering data of varying sizes and density.

FUTURE SCOPE:

This paper presented the techniques and methods that can be used to predict crime and help law agencies.
The scope of using different methods for crime prediction. and prevention can change the scenario of law enforcement agencies.
Using a combination of ML and computer vision can substantially impact the overall functionality of law enforcement agencies.
In the near future, by combining ML and computer vision, along with security equipment such as surveillance cameras and spotting scopes, a machine can learn the pattern of previous crimes, understand what crime actually is, and predict future crimes accurately without human intervention.
A possible automation would be to create a system that can predict and anticipate the zones of crime hotspots in a city.
Law enforcement agencies can be warned and prevent crime from occurring by implementing more surveillance within the prediction zone.
This complete automation can overcome the drawbacks of the current system, and law enforcement agencies can teck depend more on these techniques in the near future.
De signing a machine to anticipate and identify patterns of such crimes will be the starting point of our future study.
Although the current systems have a large impact on crime prevention, this could be the next big approach and bring about a revolutionary change in the crime rate, prediction, detection, and prevention, i.e., a "universal police officer".

CHALLENGES

Although this paper has been implemented with high ac- curacy and detailed research, there are certain challenge that can pose a problem in the future.

First, the correct and complete building of the whole system has to be done in the near future, allowing its implementation to take place immediately and properly.

Furthermore, the implementation itself is a significant concern, as such technologies cannot be directly implemented in the open world.

The system must first be tested in a small part of a metropolitan area, and only then with constant improvements (revisions of the first model) can its usage be scaled up.

Hence, the challenge are more of a help in perfecting the model and thus gradually providing a perfect model that can be applied to the real world.

Moreover, there are a few hurdles in the technological aspects of the model, as the size of the learning data will be enormous, and thus processing it will take days and maybe even weeks.

Although these are challenges that need to be addressed, they are aspects that a collective team of experts can overcome after due diligence, and if so, the end product will be worth the hard work and persistence.

Although this paper has been implemented with high ac- curacy and detailed research, there are certain challenge that can pose a problem in the future.

First, the correct and complete building of the whole system has to be done in the near future, allowing its implementation to take place immediately and properly.

Furthermore, the implementation itself is a significant concern, as such technologies cannot be directly implemented in the open world.

The system must first be tested in a small part of a metropolitan area, and only then with constant improvements (revisions of the first model) can its usage be scaled up.

Hence, the challenge are more of a help in perfecting the model and thus gradually providing a perfect model that can be applied to the real world.

Moreover, there are a few hurdles in the technological aspects of the model, as the size of the learning data will be enormous, and thus processing it will take days and maybe even weeks.

Although these are challenges that need to be addressed, they are aspects that a collective team of experts can overcome after due diligence, and if so, the end product will be worth the hard work and persistence.

CHAPTER-8

CONCLUSION

The prediction analysis done in this paper provides patterns of crime in a particular area by analysing and using certain machine learning algorithms. In this paper Random Forest Classification is proven to be a better model for predicting the results which can be observed by comparing the accuracies with the other prediction algorithms. The relatively poor algorithm is Decision Tree classification which is having a low accuracy of 51.068%. As seen in the model fitting curve in the above section. In the future more data can be collected and there will be enhancement on the computer capabilities so, more efficient models can be developed. Throughout the research It is clear that basic details of criminal activities in a neighbourhood contain indicators that will be employed by machine learning agents to classify a criminal activity given a location and date. The training agent suffers from imbalanced categories of the dataset, it had been ready to overcome the problem by oversampling and under-sampling the dataset. This paper presents a crime data prediction by taking the types of crimes as input and giving are in which these crimes are committed as output using Colab notebook having python as a core language and python provide inbuilt libraries such as Pandas and Numpy through which the work will be completed faster and Scikit provides all the processes of how to use different libraries providing by the python. Results of prediction are different for different algorithms and the accuracy of Random Forest Classifier found to be good with the accuracy of 95.122%

CHAPTER-9

BIBLIOGRAPHY

In the first section of the bibliography, the literature listed portrays the development of crime-related forecasting, with most of the documents abstracted. Of the 55 citations, 34 are discussions of empirical forecasting studies, while the remaining work deals primarily with theoretical and methodological issues associated with crime forecasting. Section II lists citations for literature that deals with prison population forecasting, including current literature and many unpublished documents from State governments. Section III traces the development of the Uniform Crime Report System in the United States and presents publications that deal with issues in the use of such statistics and the use ofcriminal statistics in other countries. An illustrative but not exhaustive list of literature on the history of crime indicators is provided. In Section IV, an illustrative listing of crime-causation literature deals with theoretical and empirical work pertaining to explanations of the level of crime rather than individual criminality. The listing is divided according to the major sets of theories that have influenced crime forecasting works, including economic ecology, deterrence, opportunity, social disorganization, and demographic theories. Section V, which lists literature that deals with social forecasting methodology, is included as a sampling of forecasting basics. This section is intended to serve as a guide for those practitioners interested in the technicalities of conducting forecasting studies.

REFERENCE:

1. McClendon, Lawrence, and Natarajan Meghanathan. "Using machine learning algorithms to analyze crime data." Machine Learning and Applications: An International Journal (MLAIJ) 2.1 (2015): 1-12.
2. Alkesh Bharati, Dr Sarvanaguru RA. K," Crime Prediction and Analysis Using Machine Learning"
in International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 09 | September 2018
3. McClendon, L., & Meghanathan, N. (2015). Using machine learning algorithms to analyze crime data. Machine Learning and Applications: An International Journal (MLAIJ), 2(1), 1-12.
4. Chen, Hsinchun, et al. "Crime data mining: a general framework and some examples." computer 37.4 (2004): 50-56.
5. Chen, H., Chungcdvsfdf, W., Xu, J. J., Wangsac, G., Qin, Y., & Chauascas, M. (2004). Crime data
mining: a general framework and some examples. computer, 37(4), 50-56.
6. Chen, Hsinchun, Wingyan Chung, Jennifer Jie Xu, Gang Wang, Yi Qin, and Michael Chau. "Crime data mining: a general framework and some examples." computer 37, no. 4 (2004): 50-56.
7. Chen, H., Chungsasda, W., Xu, J.J., Wang, G., Qin, Y. and Chau, M., 2004. Crime data mining: a
general framework and some examples. computer, 37(4), pp.50-56.
8. Sathyadevan, Shiju. "Crime analysis and prediction using data mining." 2014 First International Conference on Networks & Soft Computing (ICNSC2014). IEEE, 2014.
9. Sathyadevan, S. (2014, August). Crime analysis and prediction using data mining. In 2014 First International Conference on Networks & Soft Computing (ICNSC2014) (pp. 406-412). IEEE.
10. Sathyadevan, Shiju. "Crime analysis and prediction using data mining." In 2014 First International
Conference on Networks & Soft Computing (ICNSC2014), pp. 406-412. IEEE, 2014.
11. Sathyadevan, S., 2014, August. Crime analysis and prediction using data mining. In 2014 First International Conference on Networks & Soft Computing (ICNSC2014) (pp. 406-412). IEEE.
12. Sathyadevan S. Crime analysis and prediction using data mining. In2014 First International Conference on Networks & Soft Computing (ICNSC2014) 2014 Aug 19 (pp. 406-412). IEEE.
13. Nath, Shyam Varan. "Crime pattern detection using data mining." 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops. IEEE, 2006.
14. Nath, S. V. (2006, December). Crime pattern detection using data mining. In 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops (pp. 41-44). IEEE.
15. Nath, Shyam Varan. "Crime pattern detection using data mining." In 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops, pp. 41-44. IEEE, 2006.
16. Nath, S.V., 2006, December. Crime pattern detection using data mining. In 2006 IEEE/WIC/ACM
International Conference on Web Intelligence and Intelligent Agent Technology Workshops (pp. 41-44). IEEE.
17. Nath SV. Crime pattern detection using data mining. In2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops 2006 Dec 18 (pp.

41-44). IEEE.
18. Wang, Tong, et al. "Learning to detect patterns of crime." Joint European conference on machine
learning and knowledge discovery in databases. Springer, Berlin, Heidelberg, 2013.
19. Wang, T., Rudin, C., Wagner, D., & Sevieri, R. (2013, September). Learning to detect patterns of
crime. In Joint European conference on machine learning and knowledge discovery in
databases (pp. 515-530). Springer, Berlin, Heidelberg.