

**KARATINA UNIVERSITY**

**DEPARTMENT OF COMPUTER SCIENCE AND INFORMATICS**

**COMPUTER SCIENCE AND INFORMATION TECHNOLOGY PROJECT**

**BACHELOR OF SCIENCE IN INFORMATION TECHNOLOGY**

**RESEARCH PROPOSAL**

**CRIME RATE PREDICTION SYSTEM**

**KARIUKI RUTH NJERI**

**P100/0883G/19**

# DECLARATION

This report is my original work and has not been submitted neither is it currently being presented in any other university for a degree. No section of this project may be replicated without prior written permission from the author and/or Karatina University.

KARIUKI RUTH

P100/0883G/19

Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# 

# DEDICATION

Specially dedicated to my beloved parents, siblings and friends for their kind support over this period since I began this task.

# APPROVAL

This research proposal has been done under my supervision and is ready for submission to the school of pure and applied sciences, Department of computer science and informatics of KARATINA UNIVERSITY.

Signed…………………………………Date……………………………………

# 

# ACKNOWLEDGEMENT

I thank the Almighty God for making it possible for me to complete this paper work. Special thanks for the knowledge, wisdom, courage and determination He has granted, also I would like to thank everyone who had contributed to the successful completion of this project.

I extend my appreciation to my supervisor MR THOMAS NJOROGE for his guidance and enormous patience throughout the development of the project development.

In addition, I would also like to express my gratitude to my loving parents and friends who had helped and give me the encouragement.

# Abstract

To be better prepared to respond to criminal activity, it is important to understand patterns in crime. In our project, we analyze crime data from the city of Indore, scraped from publicly available website of Indore Police. At the outset, the task is to predict which category of crime is most likely to occur given a time and place in Indore. The use of AI and machine learning to detect crime via sound or cameras currently exists, is proven to work, and expected to continue to expand. The use of AI/ML in predicting crimes or an individual’s likelihood for committing a crime has promise but is still more of an unknown. The biggest challenge will probably be “proving” to politicians that it works. When a system is designed to stop something from happening, it is difficult to prove the negative. Companies that are directly involved in providing governments with AI tools to monitor areas or predict crime will likely benefit from a positive feedback loop. Improvements in crime prevention technology will likely spur increased total spending on this technology. We also attempt to make our classification task more meaningful by merging multiple classes into larger classes. Finally, we report and reflect on our results with different classifiers, and dwell on avenues for future work.

Table of Contents

[DECLARATION i](#_Toc128651775)

[DEDICATION ii](#_Toc128651776)

[APPROVAL iii](#_Toc128651777)

[ACKNOWLEDGEMENT iv](#_Toc128651778)

[Abstract v](#_Toc128651779)

[CHAPTER ONE 3](#_Toc128651780)

[1.1 Background of the study 3](#_Toc128651781)

[1.2 Statement of the Problem 4](#_Toc128651782)

[1.3 Key Terms 4](#_Toc128651783)

[1.4 Purpose of the Research 5](#_Toc128651784)

[1.5 Rationale 5](#_Toc128651785)

[1.6 Research Objectives/Questions 5](#_Toc128651786)

[1.6.1 General objectives 5](#_Toc128651787)

[1.6.2 Specific objectives 5](#_Toc128651788)

[1.7 Motivation of the Research 6](#_Toc128651789)

[1.8 Scope of the Research 6](#_Toc128651790)

[1.9 Significance of the Research. 6](#_Toc128651791)

[CHAPTER TWO 7](#_Toc128651792)

[2.0 LITERATURE REVIEW 7](#_Toc128651793)

[2.1 Introduction 7](#_Toc128651794)

[2.2.0 Historical development of the topic 7](#_Toc128651795)

[2.2.1 Criminal Law Principles 8](#_Toc128651796)

[2.2.2 Punishments & Its Purpose 8](#_Toc128651797)

[2.2.3 Crime Rate 9](#_Toc128651798)

[2.3 Theoretical framework/review 9](#_Toc128651799)

[2.4. Empirical studies review 12](#_Toc128651800)

[2.4.1 State of Security in Kenya 12](#_Toc128651801)

[2.4.2 Experiences Encountered when Reporting a Crime to the Police 14](#_Toc128651802)

[2.4.4 Unit Dispatch and Arrival at the Scene 14](#_Toc128651803)

[2.4.5 Review of other systems 16](#_Toc128651804)

[2.4.5.1 PredPol (Predictive Policing) Predictive System 16](#_Toc128651805)

[2.5 Knowledge gap 21](#_Toc128651806)

[2.6 Conceptual framework 22](#_Toc128651807)

[2.7 Summary 23](#_Toc128651808)

[CHAPTER THREE 24](#_Toc128651809)

[3.0 Introduction 24](#_Toc128651810)

[3.1 Research Methodology 24](#_Toc128651811)

[3.1.1Agile Methodology 24](#_Toc128651812)

[3.3.0 Research Philosophy 27](#_Toc128651813)

[3.3.1 Positivism 27](#_Toc128651814)

[3.3.2 Interpretivism 27](#_Toc128651815)

[3.4Research Paradigm 28](#_Toc128651816)

[3.5 Research Approaches 28](#_Toc128651817)

[3.5.1 Quantitative research approach 28](#_Toc128651818)

[3.5.2 Qualitative research approach 28](#_Toc128651819)

[3.6 Research Design 29](#_Toc128651820)

[3.6.1 Exploratory research 29](#_Toc128651821)

[3.6.2 Explanatory research 29](#_Toc128651822)

[3.7Research Methods 30](#_Toc128651823)

[3.7.1 Quantitative methods 30](#_Toc128651824)

[3.7.2 Qualitative methods 33](#_Toc128651825)

[References 35](#_Toc128651826)

# 

## CHAPTER ONE

## 1.1 Background of the study

Crimes, in many ways, have a negative impact to the society. Therefore, it is paramount to study the factors and relationship between different crimes and find a way to accurately and correctly predict the crimes and avoid these crimes. With the advent of the big data and the availability of fast, efficient algorithms for data analysis, understanding patterns in crime from data is an active and growing field of research.

Crimes are the significant threat to the people. Many crimes take place in regular intervals of time. It is increasing and spreading at a fast rate. Crimes happen from villages, towns and mostly in big cities. There are different types of crime including robbery, murder, rape, assaults battery, false imprisonment, kidnapping, homicide etc. Since crimes are increasing there is need to solve the case in a much faster way. Crimes have increased at a fast rate and it is the responsibility of the police department to control and reduce the vice. Crime rate prediction and criminal identification are the major problems to the police department as there are tremendous amount of crime data that exist. There is a need of technology through which the case solving could be faster.

Machine learning and data mining can make the work easier and faster. A suitable algorithm with inputs like time (hour, day, month and the year), place of occurrence and the type of the crime is to formulated. This algorithm is expected to output the type of crime that is most likely to occur at a certain place.

## 

## 1.2 Statement of the Problem

The population for major towns and cities is increasing and with it comes an upsurge in crimes. The officers entrusted with the enforcement of the law finds themselves with an enormous task of maintaining law and order by predicting where crime is likely to occur and thus preventing such occurrences. Since the officer’s focus on many issues, they may not predict the occurrence of crimes satisfactorily.

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.

Our study aims to find spatial and temporal criminal hotspots using a set of real-world datasets of crimes. We will try to locate the most likely crime locations and their frequent occurrence time. In addition, we will predict what type of crime might occur next in a specific location within a particular time. Finally, we intend to provide an analysis study by combining our findings of a particular crime’s dataset with its demographic’s information.

## 1.3 Key Terms

Algorithm

Law enforcers

Prediction

Crimes

Machine learning.

## 1.4 Purpose of the Research

This research will improve the ability of law enforcers to curb crime before it arises. It will help the law enforcers to know areas that are more prone to crime and which crime is likely to happen at what time.

## 1.5 Rationale

With the rapid urbanization and development of big cities and towns, the graph of crimes is also on the increase. This phenomenal rise in offences and crime in cities is a matter of great concern and alarm to all of us.

There are robberies, murders, rapes and what not. The frequent and repeated thefts, burglaries, robberies, murders, killings, rapes, shoplifting, pick pocketing, drug- abuse, illegal trafficking, smuggling, theft of vehicles etc., have made the common citizens to have sleepless nights and restless days.

They feel very insecure and vulnerable in the presence of anti-social and evil elements. The criminals have been operating in an organized way and sometimes even have nationwide and international connections and links.

## 1.6 Research Objectives/Questions

### 1.6.1 General objectives

The main objective of this project is to develop a crime detection and prediction system that will predict crime hotspot, understand crime pattern, classify crime based on location and classify crimes.

### 1.6.2 Specific objectives

1. To Establish the current literature on crime detection systems available.
2. To determine the requirements on a suitable crime detection system.
3. To develop and implement a crime detection and prediction prototype.
4. To test and evaluate the prototype.
5. To recommend for the system adoption.

## 1.7 Motivation of the Research

Government of Kenya employs very many law enforcers to ensure its citizens are safe and that they conduct business in a peaceful environment. However, the criminals appear to be ahead of the law enforces every time, they are always running after a criminal incidence has occurred. This study was motivated by the need to equip the law enforcers with a tool that nib the crime before it arises.

## 1.8 Scope of the Research

This study will be carried out in the city of Nairobi and it shall be limited to formulation of an algorithm that can predict the type of crime that is most likely to occur at a certain place.

## **1.9 Significance of the Research**.

The findings of this research will go a long way in bettering our preparedness to response to criminal activity, it is important to understand patterns in crime. In our project, we analyze crime data and make projections of possible crime insurgences in the future. Improvements in crime prevention technology will likely lower the crime rates and make society a peaceful place to live and do business. This research will therefore impact the law enforcement department positively. It will also advance the research in machine learning.

# 

# CHAPTER TWO

## 

## 2.0 LITERATURE REVIEW

## 2.1 Introduction

This chapter entails the overall literature review of the crime prediction, past studies on the crime, and research gap.

### 2.2.0 Historical development of the topic

Security, an essential aspect of strengthening the economy of any country needs to be attended to with maximum attention. Thus, it is the responsibility of law enforcement agencies of a given country to control crime to create a sustainable peaceful environment from which people can live and carry out their businesses. It is unfortunate that in every city in the modern world there exists crime ranging from small ones like pickpocketing to major ones like daylight robbery with violence, murder, drug abuse, rape cases and even terrorism. To combat this problem in cities there are plenty of tools that are being used and among them is the Geographical Information System (GIS) which has assisted researchers in presenting numerous crime detection and prediction models.

In general terms, a crime is an act or omission that is socially damaging and violates state/Country-protected principles. It is an incident forbidden by rule, and can be accompanied by an arrest in criminal proceedings and, consequently, imprisonment on conviction. Thanks to decades of academic research, criminality has become the most important multidisciplinary approach to modern ways of handling crime and terrorism. Crime scientists typically use a wide range of various disciplines and sciences, rather than typical criminologists, to achieve their goal of reducing crime. Crime research has been able to provide innovative ideas for the more complex problems that concern the health and safety of millions of people, using expertise from chemistry, geography, econometrics, physics, statistics, psychology, applied mathematics, and engineering. Statistics and machine learning are perhaps the most commonly used methods for crime scientists of all areas and disciplines used. Let's talk about crime now. Every crime has two main elements: action, "actus reus," and intent or "men’s rea." The crime of brass, for example, has two parts: fire in a building and doing it intentionally and knowingly. Fire can't be a crime by mistake. The actus reus and men’s rea must be proved in most criminal cases. There have been no crimes if any aspect is missing. Let us switch to the criminal law principle to better understand these elements

### 2.2.1 Criminal Law Principles

"Innocence presumption" is a criminal justice system principle. The accused is supposed to be innocent until the guilty has been demonstrated. "Burden of proof" means that the Prosecuting Attorney has to prove the culprit. There is no need for the defense counsel to prove the defendant is innocent. The word "beyond reasonable doubt" applies to the probability of a crime committed by the defendant. The Crown must show that the accused is guilty and that the judge or jury cannot have any fair doubts

### 2.2.2 Punishments & Its Purpose

It must be the goal of law to preserve society and to root out criminal tenderness, to be accomplished by enforcing a suitable judgment. Law should also address the challenges facing society as the cornerstone of the building of order. The sentencing authority must be able to determine and analyze the appropriate and effective circumstances surrounding the offense. The purpose of the criminal justice system is not only to deter crime but also to take into account the resources available to the criminal authorities and forms the perpetrator will be rehabilitated. Criminal damage is a hard change of deprivation of rights purposely forced

### 2.2.3 Crime Rate

An increase in crime rate is as a result of unpunished criminals and testifies to the countless problems facing society. The crime epidemic is unrelenting danger to the society. Crime remains a nightmare to the government of the day from a small threat to robbery with violence. The need to solve the majority of these problems prompted many policymakers to search for solutions with the help of technology. The higher crime rate discourages people from building up their properties, thus slowing down economic activity and failing to support growth. Unemployment is also a major cause of rising violence, because of the unemployment rate, the population continues to commit crimes to achieve the ends. Crimes also emerge from injustice and organized crime.

## 2.3 Theoretical framework/review

According to McClendon et al., 2005, data mining and machine learning have become a vital part of crime detection and prevention. In their research ‘WEKA’ an open-source data mining software was used to conduct a comparative study between the violent crime patterns from the Communities and Crime Unnormalized Dataset provided by the University of California-Irvine repository and actual crime statistical data for the state of Mississippi that has been provided by neighborhoodscout.com. A Linear Regression, Additive Regression, and Decision Stump algorithms were implemented using the same finite set of features, on the Communities and Crime Dataset. Overall, the linear regression algorithm performed the best among the three selected algorithms. The scope of this project was to prove how effective and accurate the machine learning algorithms used in data mining analysis can be at predicting violent crime patterns.

According to Bharati et al, 2018, the crime solving agencies can do a better work if they have a good idea of the pattern of criminal activities that are happening in a particular area. This can be done by using machine learning by employing different algorithms to find the patterns of the criminal activities in a particular area. In their work, crime data set for the past 18 years was used to predict the types of crimes in a particular area which helps in speeding up the classification of criminal cases and proceed accordingly. This research gives an efficient machine learning model for predicting the next criminal case. Various Machine learning models such as Logistic Regression, Decision Tree Classification, and Random Forest Classification were used to find the most efficient model to predict the type of crime at a particular location.

Boosting is a popular method for improving the accuracy of a classifier. In particular, AdaBoost is considered the most popular form of boosting and it has been shown to improve the performance of base learners both theoretically and empirically (Block eel et al. ,2013). The key idea behind AdaBoost is that it constructs a strong classifier using a set of weak classifiers. While AdaBoost is quite powerful, there are two major limitations: if the base classifier has a misclassification error of greater than 0.5, generalization decreases, and it suffers from overfitting with noisy data. AdaBoost requires weak classifiers to achieve an error rate less than 0.5, which can be problematic in multiclass classification.

The second limitation of overfitting occurs mainly because weak classifiers are unable to capture “correct” patterns inside noisy data. Noise can be introduced into data by two factors – mislabeled data, or limitation of the hypothesis space of the base classifier. During training, AdaBoost concentrates on learning difficult data patterns accurately. AdaBoost maximizes the “hard margin”, namely the smallest margin of those noisy data patterns and consequently the margin of other data points may decrease significantly. Different regularization strategies such as early stopping, shrinking the contribution of the individual weak classifiers, and soft margins, have been proposed to combat this issue. AdaBoost’s use of a hard margin increases the penalty exponentially for larger negative margins; these further increases error due to outliers.

Wang et al., (2013) introduced a novel, robust data-driven regularization strategy called Adaptive Regularized Boosting (AR-Boost), motivated by a desire to reduce overfitting. He replaced AdaBoost’s hard margin with a regularized soft margin that trades-off between a larger margin, at the expense of misclassification errors. Minimizing this regularized exponential loss results in a boosting algorithm that relaxes the weak learning assumption. Bounds for training and generalization errors are derived, and compared to AdaBoost. This approach combines early convergence with a soft margin by introducing a regularization term inside the exponential loss function. In every boosting round, the regularization term vanishes only if the weak classifier chosen at the current stage classifies the observations correctly.

According to Nath (2016), data mining can be used to model crime detection problems. Crimes are a social nuisance and cost our society dearly in several ways. Any research that can help in solving crimes faster will pay for itself. About 10% of the criminals commit about 50% of the crimes. Here we look at use of clustering algorithm for a data mining approach to help detect the crimes patterns and speed up the process of solving crime. We will look at k-means clustering with some enhancements to aid in the process of identification of crime patterns. We applied these techniques to real crime data from a sheriff's office and validated our results.

According to Sathya Devan et al., (2014)Crime analysis and prevention is a systematic approach for identifying and analyzing patterns and trends in crime. With the increasing advent of computerized systems, crime data analysts can help the Law enforcement officers to speed up the process of solving crimes. Using the concept of data mining we can extract previously unknown, useful information from an unstructured data. Here we have an approach between computer science and criminal justice to develop a data mining procedure that can help solve crimes faster. Instead of focusing on causes of crime occurrence like criminal background of offender, political enmity etc. we are focusing mainly on crime factors of each day.

Day by day the crime rate is increasing considerably. Crime cannot be predicted since it is neither systematic nor random. Also, the modem technologies and hi-tech methods help criminals in achieving their misdeeds. According to Crime Records Bureau crimes like burglary, arson etc. have been decreased while crimes like murder, sex abuse, gang rape etc. have been increased. Since the availability of criminal data or records is limited, we are collecting crime data from various sources like web sites, news sites, blogs, social media, RSS feeds etc. This huge data is used as a record for creating a crime record database. So, the main challenge in front of us is developing a better, efficient crime pattern detection tool to identify crime patterns effectively. The main challenges we are facing are: Increase in crime information that has to be stored and analyzed. Analysis of data is difficult since data is incomplete and inconsistent. Limitation in getting crime data records from Law Enforcement department.

## 2.4. Empirical studies review

### 2.4.1 State of Security in Kenya

Currently, the main source of criminal information and reporting is the reports developed from the police investigations and occurrence books at the police stations across the country (Baraka et al., 2021). The police often rely on the use of public information to piece together their investigation. The process of reporting crimes is generally slow in Kenya and ineffective. Getting through to the designated phone lines is often difficult and faces numerous technical barriers. Previous studies in the same region have revealed that corruption is one of the major challenges that have hindered crime-fighting within the country.

There is a weakness in the existing methodology that the security agencies use to receive information. In most cases, the reporting happens when it is already too late making it difficult for the relevant authorities to act on the information they have gathered. The use of the software system will increase the level of effectiveness and efficiency in the crimefighting activities within the city. This is because the implementation of real-time reporting would ensure an immediate response, hence the immediate reaction from security agencies. Police officers are not capable of patrolling the whole region because their ratio to the residents in the country is severely low (Mbogo et al., 2022). These are some of the challenges that are faced by the police officers when they try to fight crime in the city. When victims manage to get through to the emergency police numbers, the call handlers are not well-trained to handle the callers on the other end of the telephone call.

The government should create awareness of the existing emergency numbers and allocate well-trained personnel to become call handlers at the call centers. They should be trained on how to deal with emergencies they face. Victims often go through different situations that affect them in various ways. For this reason, the officers should be capable of accommodating the strangest type of calls they may receive from victims. Technologies and application would assist the situation. The first points of contact for the victims are the people around them and the doctors, for severe incidences. Therefore, there is a need for the victims to share this information more rapidly to the security agencies. This will reduce the response time for the security agencies to the crime scenes and consequently discourage criminal activities within the city (Stevenses al., 2022, p 12).

Currently, there is no known mobile-based technology that can be used to fight crime in the city of Nairobi. The level of mobile phone penetration in the city is relatively high. This indicates that a large percentage of the population within the city have and use mobile phones. Additionally, a large percentage of the people with mobile phones have smartphones. For this reason, it is very viable to use a mobile-based application to report criminal activities in the country. This will make the service available in other parts of the country as well. All this information will be relayed in real-time. This is made possible because of the internet. When the police respond in good time, there will be increased chances of apprehending the culprits.

The general idea is to reduce the time criminals have to escape and cover their tracks. Security agencies should respond to the crime scene within the shortest time possible to increase their chances of restoring peace in the region. Additionally, the citizens will gain confidence in the police and trust that they can to protect them against such incidences (Mwagut et al., 2022). Installing confidence in the region is of great significance since it ensures that citizens gain confidence in their government. The use of the mobile based application will shorten the time spent by the victims in reporting criminal activities that they experience. Once the police receive the notifications, they should respond to the crime scene as fast as possible. Currently, technology is lacking in the society that can be used to report a crime to the police officers that can allow them to respond to the scene of the crime immediately.

### 2.4.2 Experiences Encountered when Reporting a Crime to the Police

Currently, the crime is majorly reported in person by walking into a police station. A member of the public walks into the police station to record a statement. The items that were reported through this manner were the loss of personal items such as an identification card, and ATM cards, among others. This occupied a lot of time that could have been used by the victims to go about their businesses (Hassana et al., 2021). Moreover, the police officers have a reputation of being unfriendly. This discouraged most people from visiting a police station to record a statement. A lengthy process that discouraged people from visiting the premises characterized visiting the police station.

Upon reporting the incidences at the police station, one would obtain an abstract form that will be stamped at the police station. However, other more valuable items will require a report to be submitted. The report will allow the relevant authorities to launch investigations to determine how they were lost and apprehend the culprits. This provides a lengthy process that often discourages the victims from seeking police intervention. Some victims are afraid of being discriminated against based on the experiences they have gone through. For example, a victim who has been raped will find it difficult to repeat the ordeal to the police station. This is because they fear being discriminated against for their experience. Stigmatization is often very common within the city.

### 2.4.4 Unit Dispatch and Arrival at the Scene

#### 2.4.4.1 Nature of the Initial Police Response

In the majority of cases where the police attend a volume crime scene, uniformed officers undertake initial attendance. According to Tapia (2021) over 80 percent of 765 nonresidential burglaries examined were responded to by vehicle patrols; although in less urgent incidents foot patrols were frequently dispatched. A study by Tapia on residential burglaries revealed that, in most cases, the closest response vehicle unit staffed by uniformed officers attended the scene (with either local beat officers or detectives attending only occasionally as the first response).

#### 2.4.4.2 Response Time

A number of factors influences the time taken for officers to respond to incidents. These include the time taken for the incident to be reported; the grade allocated to the incident; the availability of officers and their distance from the scene. Ronoh, & Hamasi, (2021) found clear differences in the response times between the two emergency grades used in the force studied; cases graded as immediate were attended on average in 4.5 minutes, whereas the police took an average of 9.6 minutes to attend cases graded as requiring an early response. Burglaries that were reported ‘in progress” were more likely to be graded as emergencies/immediate and hence were usually attended more quickly than others. Further analysis of the data indicated that the response times were dependent on both pre-travel times (which reflected what officers were doing at the time of dispatch) and travel times (duration taken to travel to the scene). There are big differences in the times taken to respond to robberies.

#### 2.4.4.3 Response Times and Outcomes

The relationship between the likelihood of identifying an offender and response times has received considerable attention in the research literature. A study by Mwangi et al., (2022)., which focused on burglaries in six policing areas examined factors that account for differences between high and low clear-up (including the effects in response times). Overall, no evidence was found to indicate that quicker response times were linked with higher detection rates. However, this study looked at the impact of response times on all burglaries rather than focusing only on burglaries reported in progress or reported soon after execution – the outcome of these cases would be more likely to be influenced by the speed of response.

In addition, the sample of burglaries studied included offenses that had been detected through the case being taken into consideration – these cases would not have been affected by response times. The author acknowledges that aiming for quick response times may be a waste of resources unless the offense is reported immediately. It illustrates a central feature of the relationship between the ability of the police to detect and respond to crime, and the critical influence of the context of individual cases on this relationship. The relationship between police response times and case outcomes/focused on arrests rather than detections. The research aimed to assess the effectiveness of police response regarding increasing the likelihood of 12 arrest. The study analyzed 2,532 cases from one force. Response times were measured by three variables: dispatch time, travel time and response time. Possibility of apprehending the culprits were significantly associated with both response and travel times, although not with dispatch times. The likelihood of apprehending the offender increased as the travel and response times decreased. This pattern was evident in burglaries, thefts, vehicle thefts and robberies.

The study also included analysis using data from previous research on the relationship between response times and arrests made during follow-up investigations. Although the results were mainly inconclusive, there was some indication that quick response times were also related to increases in the number of arrests from follow-up investigations.

Network failures sometimes will hinder any successful call. Reporting in person has several disadvantages too: By the time the victim reaches the police station, the crime has already been committed especially where the nearby police station is a bit far. A victim will in most cases report after the crime is committed which imply that the person had to suffer first then seek justice. Therefore, this method is very reactive as opposed to proactive. This method of crime reporting is very slow because a lot of time is consumed to access the police station and give a statement. Might not dedicate their time to go to the police station and make a report probably because of fear, apathy, attending to personal matters or the nearby police station is very far.

### 2.4.5 Review of other systems

### 2.4.5.1 PredPol (Predictive Policing) Predictive System

PredPol (Predictive Policing) is a commercially available crime prediction software that uses historical crime data and algorithms to predict future crime patterns (PredPol, 2016). PredPol was developed by a team of criminologists and computer scientists at the University of California, Los Angeles (UCLA).

The software analyzes crime data from a specific area to identify the locations and times when crimes are most likely to occur. It then generates a 500x500 foot prediction map that displays the predicted hotspots of criminal activity. This information is used by law enforcement agencies to allocate their resources more effectively, so they can be in the right place at the right time to prevent or respond to crimes.

One of the key advantages of PredPol is its ability to provide real-time, actionable intelligence to law enforcement agencies. The software is designed to be easy to use, and it integrates with existing police systems, such as computer-aided dispatch (CAD) systems and records management systems (RMS). This enables officers to access the information they need to make informed decisions in the field.

PredPol has been adopted by law enforcement agencies in several countries, including the United States, the United Kingdom, and the Netherlands. Evaluations of the software have shown that it can significantly reduce crime in the areas where it is used, and it has been praised for its ability to help police departments allocate their resources more efficiently.

Despite its success, PredPol has some potential drawbacks, including:

1. Relying on past data: PredPol's predictions are based on historical crime data. This means that the system may struggle to predict new types of crime or criminal activity that have not been seen before.
2. Bias in data: The accuracy of PredPol's predictions depends on the quality and accuracy of the data it is fed.

#### 2.4.5.2 SpotCrime Predictive System

SpotCrime is a crime mapping and alert system that provides real-time crime data and alerts to communities SpotCrime (2023). It collects crime data from various sources such as police reports, news sources, and user submissions to provide a comprehensive view of crime in a specific area.

SpotCrime provides a platform for individuals to view crime data and crime maps for their local area. The crime data is displayed on an interactive map, allowing users to see the type, location, and frequency of crimes in their area. The platform also allows users to sign up for email alerts for crimes in a specific area, keeping them informed of any recent crime activity.

The platform collects crime data from a variety of sources, including police reports, news sources, and user submissions. This data is then verified and processed to provide a comprehensive view of crime in the area. The platform covers a wide range of crimes, including theft, assault, robbery, burglary, and more.

SpotCrime also provides a "Heat Map" feature that displays the density of crimes in a specific area. This feature allows users to see where crime is more concentrated in their area and make informed decisions about their personal safety.

Overall, SpotCrime provides a valuable resource for individuals to stay informed about crime in their area and make informed decisions about their personal safety.

SpotCrime Predictive System also has some potential drawbacks, including:

1. Relying on incomplete data: The accuracy of SpotCrime's predictions depends on the quality and completeness of the crime data it is fed. If the data is incomplete or outdated, this can negatively impact the reliability of the predictions.
2. Bias in data: Like PredPol, SpotCrime's predictions may be influenced by biases in the crime data it is fed.

#### 2.4.5.3 Gunshot detection system.

A gunshot detection system is a technology that uses acoustic sensors to detect and locate the sound of gunfire in real-time. The system is designed to detect the sound of a gunshot, analyze the acoustic signature of the sound, and determine the location of the gunshot using triangulation. The system then sends an alert to law enforcement, emergency services, or other designated authorities, providing them with the location of the gunfire so that they can respond quickly and appropriately.

Gunshot detection systems can be installed in various locations, such as urban areas, college campuses, and transportation hubs, and can be integrated with other security systems to provide a comprehensive security solution. These systems can also be used to collect data on gun violence and help law enforcement agencies identify patterns and trends in gun violence.

There are two types of gunshot detection systems: indoor and outdoor. Indoor gunshot detection systems are designed to detect gunshots within a confined space, such as a building or room, while outdoor gunshot detection systems are designed to detect gunshots over a wider area, such as a neighborhood or city block.

The technology used in gunshot detection systems varies, but most systems use a combination of acoustic sensors, machine learning algorithms, and cloud-based processing to analyze the sound of gunshots and determine their location. Some systems also incorporate video cameras and other sensors to provide additional information about the situation.

#### 2.4.5.4 Hot Spot Policing.

Hot spot policing is a policing strategy that involves targeting areas with high levels of crime or disorder in order to prevent and reduce crime. This strategy is based on the idea that crime is not evenly distributed throughout a community, but rather tends to cluster in certain "hot spots" or areas with high levels of criminal activity.

Hot spot policing involves using data analysis and mapping to identify these high-crime areas, and then deploying police resources to these areas in a targeted and strategic manner. This can involve increased patrols, undercover operations, and other tactics designed to deter criminal activity and improve public safety in these areas.

The effectiveness of hot spot policing has been the subject of extensive research, with some studies showing that the strategy can be effective in reducing crime in targeted areas. However, other studies have raised concerns about the potential for unintended consequences, such as displacement of crime to nearby areas or increased scrutiny of certain communities by law enforcement.

Critics of hot spot policing argue that it can lead to over-policing of certain communities and may perpetuate systemic inequalities in the criminal justice system. It is important to consider these concerns and to ensure that hot spot policing is implemented in a fair and equitable manner.

#### 2.4.5.5 Risk terrain modelling

Risk terrain modeling (RTM) is a geographic information system (GIS) based approach that seeks to identify the underlying environmental factors that are associated with crime and other negative outcomes. RTM is based on the idea that certain environmental factors can increase the likelihood of crime and other negative outcomes, and that by identifying these factors, policymakers and law enforcement agencies can develop more effective strategies for preventing crime and improving public safety.

RTM involves analyzing data on crime, demographic information, and environmental factors such as land use, transportation networks, and social and economic factors, in order to identify the factors that are most strongly associated with crime and other negative outcomes. These factors are then mapped using GIS software to create a risk terrain map, which can be used to guide the deployment of police resources and inform other crime prevention strategies.

The goal of RTM is to identify the environmental factors that are most strongly associated with crime and other negative outcomes, and to use this information to develop more effective strategies for preventing crime and improving public safety. RTM can be used in a variety of settings, including urban and rural areas, and can be adapted to address a wide range of crime and other negative outcomes, including violence, drug use, and gang activity.

One advantage of RTM is that it allows policymakers and law enforcement agencies to identify the environmental factors that are most strongly associated with crime and other negative outcomes, and to develop targeted strategies for addressing these factors. However, it is important to ensure that RTM is implemented in a way that is fair and equitable, and that does not perpetuate systemic inequalities in the criminal justice system.

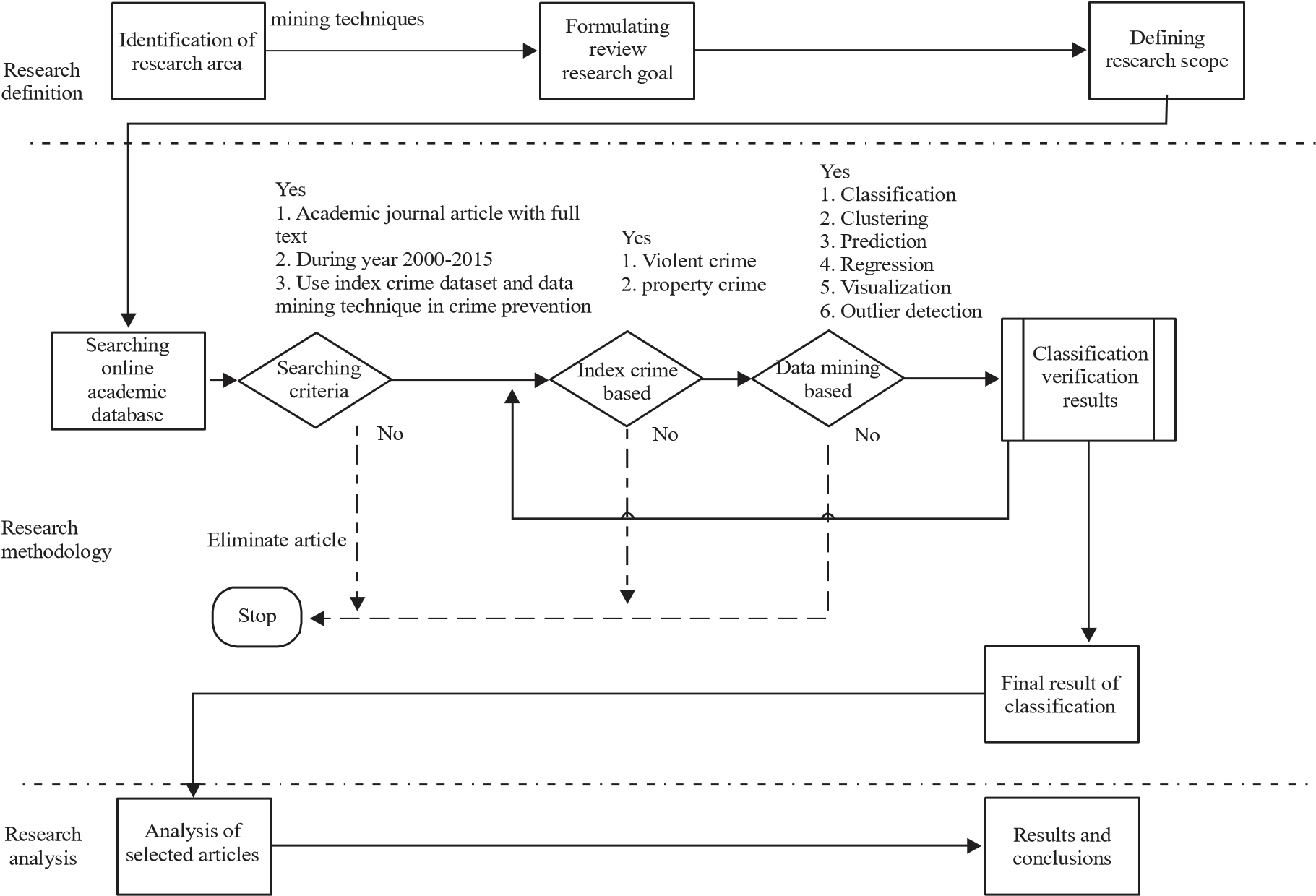
## 2.5 Knowledge gap

We present a systematic review study on AI strategies for crime prediction, which differs from existing work in this literature. The following represents a small summary of other surveys and the difference between their contributions and my study. My research differs from previous research in a number of ways, including the following: Presents the main focus of research in crime prediction with respect to the type and category of the crime, the study time, and which type of crime has been addressed by most researchers. Clarifies the frequency of what algorithms were used in training models and what factors were considered when choosing them. Analyzes the estimation accuracy of the models by focusing on four factors: performance metric, evaluation value, the dataset utilized, and the validation approaches for the model. Demonstrates the tools applied in the selected research papers, the strengths and weaknesses of the prediction models, and the limitations and future direction of crime prediction. Covers the period from 2022- 2023, which addresses recent work**.**

## 

## 

## 2.6 Conceptual framework



## 

## 2.7 Summary

The chapter covered the literature on crime prediction models the emphasis on how they assist in providing security solutions. Global context and problems of crime prediction in cases of major cities like New York, London and Paris have been discussed in depth. The Kenyan topic and issues faced by the Kenya police and citizens in reporting crimes are also discussed. The literature has sections on security force response to crime reports, experiences of the general public when reporting the crime within the city and their challenges. Finally, the conceptual framework of the system.

# CHAPTER THREE

## 3.0 Introduction

This chapter will outline the research methodology to be used in this project. The data collection methods. The data will then be analyzed using a variety of statistical and machine learning techniques to make predictions about crime rates in the area of interest. The results of these analyses will be used to make recommendations for improving the current system. The methodology described in this chapter will be used to ensure that the research is conducted in a rigorous and unbiased manner by using advanced data analysis and machine learning techniques. In order to achieve this goal, a thorough analysis of the current system and its limitations will be conducted, followed by the collection and cleaning of relevant data.

## 3.1 Research Methodology

### 3.1.1Agile Methodology

The software development model intended to be used in guiding all through during the development process is the Agile development model. This model employs iterative and interactive approach while developing modules of a system which can function well both independently and when integrated. The agile methodology breaks the developmental process into iterative steps, allowing for flexibility, testing, and change throughout the lifecycle of the project. Also,the Agile methodology would be useful because it is iterative and incremental, which would allow for regular reassessment and adjustment of the project as new information and insights are gained. Additionally, the Agile methodology emphasizes collaboration and flexibility, which would be beneficial for a project that is likely to encounter a high level of uncertainty.

#### 3.2.1.1 Phases of Agile method.

1. Requirements gathering: In this phase, is defining the requirements. It has explained business opportunities and plan the time and effort needed to build the project. Based on this information, you can evaluate technical and economic feasibility.

2. Design the requirements: When the project have been identified, it has to work to define requirements. It has the use the user flow diagram or the high-level UML diagram to show the work of new features and show how it will apply to your existing system.

3. Construction/ iteration: When have defined the requirements, the work begins.

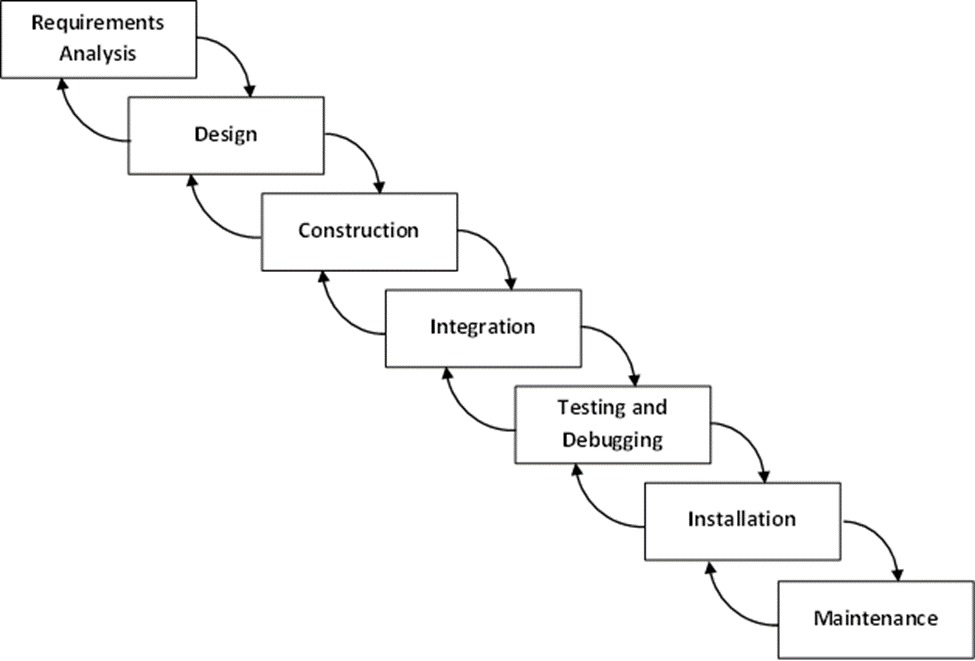
4. Testing: In this phase, it examines the Quality Assurance of the system performance and look for the bug.

5. Recommendation: In this phase, it have recommended a system to organizations for the user's work environment.

6. Feedback: After releasing the product, the last step is feedback. In this, it is receiving feedback about the product and works through the feedback.

Here's how agile methodology can be applied to a crime rate prediction system project:

1. Iterative Development: Agile methodology involves iterative development, with each iteration consisting of a set of requirements, design, development, testing, and deployment activities. In the context of a crime rate prediction system, this could mean developing a basic version of the system and gradually improving it with additional features and enhancements.
2. Continuous Feedback: Agile methodology emphasizes continuous feedback and collaboration between the development team and the stakeholders. In the context of a crime rate prediction system, this could mean working closely with law enforcement agencies to ensure that the system meets their needs and provides useful insights for crime prevention.
3. User-Centric Design: Agile methodology focuses on creating a user-centric design that prioritizes the needs of the end-users. In the context of a crime rate prediction system, this could mean designing a user interface that is intuitive and easy to use, and that provides law enforcement officers with the information they need to make informed decisions.
4. Adaptability: Agile methodology emphasizes adaptability and the ability to respond quickly to changes in the project requirements. In the context of a crime rate prediction system, this could mean being able to adapt the system to changing crime patterns, new types of criminal activity, and new sources of data.
5. Continuous Testing: Agile methodology involves continuous testing and quality assurance activities to ensure that the system meets the required standards of performance, security, and reliability. In the context of a crime rate prediction system, this could mean conducting regular tests to ensure that the system is accurate and reliable, and that it is not subject to issues like bias or discrimination.



#### 3.2.1.2 Justification of using agile model

The Agile development model is appropriate for the development of this project due to its interactive and iterative nature. Development is based on stages, and with the use of the model, it always come back to a stage It had already finished up if there occurs a constantly changes in the user requirements. Agile development is focused on quick responses to change and continuous development. (Flora & Chande, 2014)

## 3.3.0 Research Philosophy

The research philosophy for this project is a combination of positivism and interpretivism.

### 3.3.1 Positivism

Positivism is a philosophy that holds that the natural world can be studied objectively and that knowledge can be gained through the scientific method. In this project, the positivist approach will be used to study the existing crime rate prediction system and its limitations through a thorough analysis of the data and existing research. This approach will provide a systematic and rigorous way to understand the crime rate prediction system and how it works.

### 3.3.2 Interpretivism

Interpretivism, on the other hand, emphasizes the subjectivity of social phenomena and the importance of understanding the meaning and context behind the data. In this project, interpretivism will be used to analyze the collected data and make predictions about crime rates in the area of interest. This approach will help us to understand the complexity of the crime rate prediction system, and how it is affected by different social and economic factors.

By combining these two philosophies, the project aims to gain a more comprehensive understanding of the crime rate prediction system and the factors that influence crime rates, and to develop improved methods for predicting and preventing crime. The choice of this philosophy is based on the fact that the project aims to understand both the objective and subjective aspects of crime rate prediction system, and how to improve it.

## 3.4Research Paradigm

The research paradigm for this project is a combination of positivism and interpretivism. Positivism is a philosophy that holds that the natural world can be studied objectively and that knowledge can be gained through the scientific method. In this project, the positivist approach will be used to study the existing crime rate prediction system and its limitations through a thorough analysis of the data and existing research.

Interpretivism, on the other hand, emphasizes the subjectivity of social phenomena and the importance of understanding the meaning and context behind the data. In this project, interpretivism will be used to analyze the collected data and make predictions about crime rates in the area of interest. This approach will also be used to make recommendations for improving the current system.

## 3.5 Research Approaches

The research approaches for this project will be both qualitative and quantitative research.

### 3.5.1 Quantitative research approach

The project will use a quantitative research approach to analyze the collected data and make predictions about crime rates in the area of interest. This will involve using statistical and machine learning techniques to analyze the data and identify patterns and trends. The goal of this approach is to develop a more accurate and reliable method for predicting crime rates.

### 3.5.2 Qualitative research approach

The project will use a qualitative research approach to gain a better understanding of the social and economic factors that influence crime rates in the area of interest. This will involve conducting a case study of the current system, conducting interviews with relevant stakeholders, and conducting a content analysis of news articles, reports, and other relevant literature. The goal of this approach is to gain a more comprehensive understanding of the problem and to identify potential areas for improvement.

By using a combination of both qualitative and quantitative research approaches, we will be able to gain a comprehensive understanding of the crime rate prediction system and the factors that influence crime rates, and to develop improved methods for predicting and preventing crime.

## 3.6 Research Design

The research design for this project is a combination of exploratory and explanatory research.

|  |  |  |
| --- | --- | --- |
| **EXPLORATORY DESIGN** | **STEPS** | **EXPLANATORY DESIGN** |
| Identify the problem | 1 | Design the research questions |
| Hypothesize the solution | 2 | Formulate a hypothesis |
| Design your methodology | 3 | Design the methodology and collect data |
| Collect and analyze data | 4 | Analyze data and report results |
| Avenue for the future research | 5 | Interpret the results and suggest solution for the future |

### 3.6.1 Exploratory research

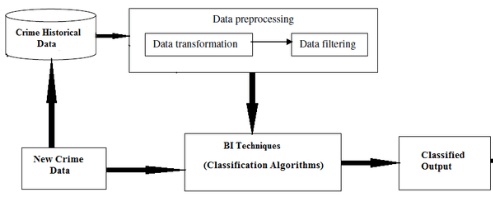
The project will use exploratory research to study the existing crime rate prediction system and its limitations. This will involve collecting and analyzing data, reviewing the literature, and conducting a case study of the current system. The goal of this phase is to gain a better understanding of the problem and to identify potential areas for improvement.

### 3.6.2 Explanatory research

The project will use explanatory research to develop and test new methods for predicting and preventing crime. This will involve collecting and analyzing additional data, using statistical and machine learning techniques to make predictions, and testing the effectiveness of the new methods. The goal of this phase is to explain the relationships between crime rates and various factors such as demographics, socioeconomic status, and other relevant variables.

The research design is a combination of both exploratory and explanatory research, where we will start with exploratory research to study the current system and its limitations, and then move to explanatory research to develop and test new methods for predicting and preventing crime. This approach will help us to gain a comprehensive understanding of the crime rate prediction system and to develop improved methods for predicting and preventing crime.

The DFD diagram



## 3.7Research Methods

The research methods for this project include both quantitative and qualitative methods.

### 3.7.1 Quantitative methods

#### 3.7.1.1 Data collection

Data collection, according to (Peck et al., 2019), is the process of acquiring and measuring information on specific variables in a system, allowing one to answer relevant questions and evaluate consequences. In research, a variety of data collection tools can be used. Questionnaires, interviews (structured, phone, group), observation, prototyping, and JAD are all examples of this. The questionnaire is a straightforward, easy-to-use data collection tool. Respondents are asked a series of open-ended or closed-ended questions about the topic at hand (Muhammad & Kabir, 2018).

The project will use a variety of data sources including crime statistics, demographic data, and socioeconomic data. These data will be collected from government agencies like police stations, citizens and other publicly available sources.

##### **3.7.1.1.1 Questionnaires**

A questionnaire can be used to collect data in several ways:

1. Gathering information on past crime: A questionnaire can be used to gather information on past crime incidents, such as location, type of crime, and time of occurrence. This information can be used to train a model that can make predictions about where and when crimes are likely to occur in the future.

2. Understanding community perceptions: A questionnaire can be used to gather information on community perceptions of crime and safety. This information can be used to identify areas of concern and areas where the existing crime rate prediction system may not be meeting the needs of the community.

3. Identifying factors that influence crime: A questionnaire can be used to gather information on factors that may influence crime in a given area, such as poverty, unemployment, and access to education. This information can be used to develop a more comprehensive understanding of crime and to identify potential intervention points.

4. Evaluating the effectiveness of current crime-prevention strategies: A questionnaire can be used to gather information on the effectiveness of current crime-prevention strategies, such as community policing programs and crime prevention through environmental design. This information can be used to identify strategies that are working well and those that may need to be improved or replaced.

##### **3.7.1.1.2 Interviews**

Interviews can be used to collect data in several ways:

1. Gathering information from experts: Interviews can be used to gather information from experts in the field of crime prediction, such as law enforcement professionals, criminologists, and statisticians. These experts can provide valuable insights into the strengths and weaknesses of existing crime prediction systems and suggest ways to improve them.

2. Understanding community perspectives: Interviews can be used to gather information from community members about their perceptions of crime and safety in their area

3. Identifying factors that influence crime: Interviews can be used to gather information from community members and experts about factors that may influence crime in a given area, such as poverty, unemployment, and access to education.

4. Evaluating the effectiveness of current crime-prevention strategies: Interviews can be used to gather information from law enforcement professionals, community members, and experts about the effectiveness of current crime-prevention strategies, such as community policing programs and crime prevention through environmental design

##### **3.7.1.1.3** **Case study**

A case study can be used to collect data in several ways:

1. Examining specific crime incidents: A case study can be used to examine specific crime incidents in depth, including information on location, type of crime, and time of occurrence. This information can be used to identify patterns and trends that can be used to improve the crime rate prediction system.

2. Understanding the context of crime: A case study can be used to gain a deeper understanding of the social and economic factors that contribute to crime in a specific area. This information can be used to develop a more comprehensive understanding of crime and to identify potential intervention points.

3. Evaluating the effectiveness of current crime-prevention strategies: A case study can be used to evaluate the effectiveness of current crime-prevention strategies in a specific area. This information can be used to identify strategies that are working well and those that may need to be improved or replaced.

4. Examining the experience of other jurisdictions: A case study can be used to compare the experience of other jurisdictions in terms of crime rate prediction system, this can help in identifying best practices and lessons learned that can be applied to the local context

##### **3.7.1.1.3 Observation**

Observations can be used to collect data in several ways:

1. Identifying crime hotspots: Observations can be used to identify areas where crime is more likely to occur, such as parks, public transportation stops, and shopping centers. This information can be used to improve the accuracy of crime prediction models.

2. Assessing the effectiveness of current crime-prevention strategies: Observations can be used to assess the effectiveness of current crime-prevention strategies, such as increased patrols or surveillance cameras, in reducing crime in specific areas. This information can be used to identify strategies that are working well and those that may need to be improved or replaced.

3. Understanding the context of crime: Observations can be used to gain a deeper understanding of the social and economic factors that contribute to crime in a specific area. This information can be used to develop a more comprehensive understanding of crime and to identify potential intervention points.

4. Gathering information on the physical environment: Observations can be used to gather information on the physical environment, such as lighting, architecture and layout, which can influence the risk of crime.

#### 3.7.1.2 Data analysis

The collected data will be analyzed using statistical and machine learning techniques such as regression analysis, decision tree analysis, and neural network analysis. These techniques will be used to make predictions about crime rates in the area of interest and to identify patterns and trends in the data.

### 3.7.2 Qualitative methods

#### 3.7.2.1 Case study

The project will conduct a case study of the current crime rate prediction system and its limitations. The case study will involve analyzing the data, reviewing the literature, and conducting interviews with relevant stakeholders such as law enforcement agencies and community organizations.

#### 3.7.2.2 Content analysis

The project will also conduct a content analysis of news articles, reports, and other relevant literature to gain a better understanding of the social and economic factors that influence crime rates in the area of interest.

By using a combination of quantitative and qualitative methods, the project aims to gain a more comprehensive understanding of the crime rate prediction system and the factors that influence crime rates, and to develop improved methods for predicting and preventing crime.

# References

Andersen, S. H., Richmond-Rakerd, L. S., Moffitt, T. E., & Caspi, A. (2021). *Nationwide evidence that education disrupts the intergenerational transmission of disadvantage.* Proceedings of the National Academy of Sciences, 118(31), e2103896118.

Baraka, G. E., & Murimi, S. K. (2021). *Why geographic information systems in spatiotemporal crime analysis? Attitude of Kenyan police officers*. Police Practice and Research, 22(5), 1453-1468.

Hassana, M. K., Hassanb, A., Khodayerc, A. M., & Khodayerd, O. M. *Internet Security Impact on E-Banking Users*.

Jia, S., Qiu, Y., & Yang, C. (2021). *Sustainable Development Goals, Financial Inclusion, and Grain Security Efficiency*. Agronomy, *11*(12), 2542.

Mazuri, M. H., Mwaeke, P., & Bor, E. (2022). *Factors Impeding Government Security Agencies Responses in Combating Cattle Rustling in Baringo County-Kenya.* European Journal of Humanities and Social Sciences, 2(2), 6-14.

Mbogo, A. M., & Wambua, P. (2022). *Assessing The Sociopolitical Security Determinants Of Crime Incidents In Kenya: A Case Of Kibera Informal Settlement.* Nairobi City County

Alkesh Bharati, & Sarvanaguru RA. K, (2018). *Crime Prediction and Analysis Using Machine Learning.* International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 09 .

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.

Mutonyi, G. P., Sirera, M. A., & Mwai, W. (2022). *The Limitations of Commercialized Security on National Security in Nairobi County, Kenya.* Path of Science, 8(1), 1009-1022.

Mwagut, M. M., & Minja, D. (2022). *Influence of community policing initiatives on youth radicalization in Mombasa County in Kenya*. International Academic Journal of Arts and Humanities, 1(3), 1-29.

Nath, S. V. (2016). *Crime pattern detection using data mining*.IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops (pp. 41-44). IEEE.

Ronoh, V., & Hamasi, L. H. (2021). *Community Policing and its Effect on Crime Management in Mombasa County, Kenya*. Journal of African Interdisciplinary Studies, 5(3), 68-80.

Sathish, C., & Ramchander, M. CRIME PREDICTION AND ANALYSIS USING MACHINE LEARNING.

Sathyadevan, S., Devan, M. S., & Gangadharan, S. S. (2014). *Crime analysis and prediction using data mining*. First international conference on networks & soft computing (ICNSC2014) (pp. 406-412). IEEE.

Stevens, L. M., Reid, E., Kanja, W., Rockowitz, S., Davies, K., Dosanjh, S., ... & Flowe, H. D. (2022). The Kenyan survivors of sexual violence network: Preserving memory evidence with a bespoke mobile application to increase access to vital services and justice. *Societies*, *12*(1), 12.

Tapia, B. A. B. (2022). *The Rise, Endurance, and Fall of Migrant Camps on the US-Mexico Border: A Sociology of Border Violence* .Doctoral dissertation, University of Colorado.

Wang, T., Rudin, C., Wagner, D., & Sevieri, R. (2013). *Learning to detect patterns of crime*. Joint European conference on machine learning and knowledge discovery in databases (pp. 515-530). Springer, Berlin, Heidelberg.

H. Block eel et al. (Eds.): ECML PKDD 2013, Part III, LNAI 8190, pp. 1–16, 2013. c Springer-Verlag Berlin Heidelberg 2013 2 B.N. Saha et al. and potentially distorts the optimal decision boundary.

"Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations" by PredPol, 2016.