THE FUTURE OF CRIME PREVENTION: POLICE CASE ANALYSIS USING MACHINE LEARNING

(Clustering crimes against women and crime forecasting prediction)

Project Id: 2022-044

Final Draft Report

Anubama Logaratnam

IT20068196

BSc (Hons) in Information Technology Specializing in Software Engineering

Department of Computer Science & Software Engineering

Sri Lankan Institute of Information Technology

Sri Lanka

September 2023

DECLARATION

I declare that this is my own work, and this Thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my Thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
Anubama.L	IT20068196	

The above candidate has carried out this research thesis for the Degree of Bachelor of Science (honours) Information Technology (Specializing in Software Engineering) under my supervision.

Signature of the supervisor

(Ms. Hansika Mahaadikara) Date

Signature of the co-supervisor

(Ms. Sanjeevi Chandrasiri) Date

ABSTRACT

This report presents the development of a comprehensive Crime Prediction and Prevention System with a specific focus on enhancing women's safety. The problem at hand revolves around the critical need for proactive measures to combat crimes against women, given the existing gaps in traditional post-incident intervention approaches.

The primary purpose of this report is to introduce an innovative system that predicts future crimes against women and empowers law enforcement agencies and policymakers with actionable insights for prevention. The main areas addressed encompass the system's architecture, the utilization of machine learning algorithms, and the predictive analytics employed. The methodology involves the systematic analysis of historical crime data, training Random Forest regression models for crime prediction, and database storage of predictions for future reference. The general solution presented in this report is a data-driven approach to crime prevention, leveraging predictive analytics and pattern recognition to identify areas at risk. It equips authorities with the knowledge required for efficient resource allocation and targeted actions to improve public safety, especially for women.

The suggested system introduces a pioneering Crime Prediction and Prevention System tailored to address crimes against women. While findings are not discussed in this abstract, the system's methodology and approach promise to revolutionize law enforcement strategies and promote safer communities for all.

Key Words: Machine learning, Crime Prediction, Crime Prevention, Women's Safety, Random Forest Regression

ACKNOWLEDGEMENT

I sincerely convey my sincere thanks to our module coordinator Mr. Jayantha Amararachchi who helped us and gave us enough motivation and ideas to carry forward the project further and involve ourselves to our best with the project with much enthusiasm. And I would like to thank Ms. Hansika Mahaadikara and co supervisor Ms. Sanjeevi Chandrasiri for their valuable time, guidance and support that they provided throughout the project and helped me from the very start till the end and gave variety of ideas to develop the project in many aspects and also bearing up with all the mistakes that was made by and stood with me for the entire period of time with a lot of patience and care. Also, I thank the lecturers, assistant lecturers, instructors, my group members and academic and non-academic staffs of SLIIT who were always there to support me and help me to complete the requirements of the module. Finally, I thank my beloved family and friends who stood by me throughout the project period like pillars and provided a moral support to me at points where I felt like giving up on the project.

Contents

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS

Abbreviations	**Description**
ML	Machine Learning
IO	Input/output
CSV	Comma-Separated Values
SDLC	Software Development Life Cycle
MSE	Mean Squared Error
UAT	User Acceptance Test
IoT	Internet of Things

1 INTRODUCTION

1.1 Background Study and Literature Review

1.1.1 Background Study

Crimes against women is a critical and deeply troubling issue not only in Sri Lanka but worldwide. Sri Lanka's growing economic crisis and social inequality have increased the number of crimes against women, posing a significant threat to their physical, emotional and mental well-being. Unfortunately, many of these crimes go unreported, mainly due to social stigma, fear and lack of trust in existing legal and law enforcement mechanisms.

There have been years of efforts led by both governments and non-governmental organizations to combat crimes against women. These initiatives aim to support victims, raise awareness of women's rights and improve the legal framework. Despite these efforts, however, preventing and reducing such crimes remains a major challenge. The complexity of the problem, combined with the limited resources of law enforcement agencies, requires innovative and data-driven solutions.

The manual analysis of criminal cases, which is still prevalent in many fields, has several significant disadvantages. First, it leads to long investigation times, delays the administration of justice, and potentially gives criminals an opportunity to evade the law. Second, manually classifying and recording cases introduces a margin of error that can affect the accuracy of crime statistics. In addition, small and frequent cases may not receive the attention they deserve because resources are often stretched.

To close these gaps and fight crimes against women, this study proposes a new approach based on machine learning techniques. The main goal is to automate the analysis and prediction of such crimes, which will speed up response times and provide more accurate information about the nature of these crimes.

The proposed solution includes clustering of crimes against women, grouping based on various characteristics including location, type of crime and year of occurrence. These clusters are the basis for the development of a crime prediction model. Using historical crime data, this model predicts the probability of crime in a given area at a given time. This foresight allows law enforcement agencies to take proactive steps to prevent crimes against women, rather than simply reacting to incidents.

In addition to automating the analysis and prediction of crimes against women, this study recognizes the importance of addressing the underlying factors that contribute to such crimes. It is crucial to acknowledge that crimes against women are not isolated incidents; they are often deeply rooted in social, economic, and cultural contexts. Therefore, this research takes a holistic approach by considering both predictive modeling and prevention strategies.

In conclusion, this study is a significant step towards creating a more effective and data-driven approach to fighting crimes against women and ensuring their safety in society. Using the power of machine learning and data analytics, this solution aims to advance the broader goal of disproportionately reducing and preventing crimes against women and promoting a safer and more equal environment for all.

1.1.2 Literature Review

The issue of crimes against women and children has garnered significant scholarly attention in recent years, reflecting its pervasive nature and societal implications. In response to this pressing concern, researchers have been at the forefront of developing innovative approaches to mitigate and prevent such crimes. One noteworthy contribution to this evolving landscape is the work of Reza [15], who proposed the development of a sophisticated machine learning-based support system tailored to combat suppression against women and children. This research represents a pivotal stride in harnessing the capabilities of cutting-edge technology, aligning seamlessly with the growing trend of employing machine learning techniques to bolster crime prediction and prevention. This trend is further underscored by Adderley's comprehensive study [16], which offers an in-depth exploration of the various facets of crime predictive techniques, highlighting the transformative potential of data-driven methodologies.

Recent strides in wearable technology have ushered in a new era in the realm of women's safety. An illuminating example is the "Safeband," introduced by Islam et al. [17]. This wearable device, meticulously designed to enhance the security of women in Bangladesh, stands as a testament to the transformative potential of technology to empower and safeguard women, particularly in vulnerable situations. Notably, the "Safeband" seamlessly incorporates elements of the Internet of Things (IoT), marking a convergence of emerging technologies to address a critical societal issue.

Moreover, analytical approaches to crime prediction have undergone thorough exploration and refinement. Kiani and Keshavarzi [18] embarked on an insightful journey into the intricacies of crime analysis and prediction, leveraging the power of clustering and classification methods. Their findings unequivocally demonstrate the efficacy of data-driven techniques in deciphering complex crime patterns, shedding light on previously obscured nuances in criminal activities. Taking this concept a step further, Karmakar et al. [19] introduced "SafeBand," an IoT-based smart security band equipped with instantaneous SOS messaging capabilities. This innovative device exemplifies the seamless integration of technology in ensuring the safety and security of women. It serves as a beacon of hope for individuals facing threats to their well-being, offering a lifeline in times of distress.

Collectively, these studies underscore the paramount importance of technology and data-driven solutions in addressing the pressing issue of crimes against women and children. Beyond the realm of theoretical inquiry, they offer practical solutions that hold great promise in enhancing safety and security. By leveraging machine learning, wearable technology, and advanced analytical techniques, these contributions not only enable more accurate crime prediction but also empower individuals, particularly women, with tools to protect themselves in an increasingly complex and interconnected world. As technology continues to advance, these pioneering efforts provide a compelling blueprint for future endeavours aimed at creating safer communities and a more secure future for all.

1.2 Research Gap

Although many studies have explored wearable technologies, mobile apps and their role in combating crimes against women and children by providing immediate assistance to victims, there is a significant gap in crime prevention. Much of the existing research focuses primarily on post-incident interventions and victim-cantered support systems. In addition, there is a noticeable lack of standardized methods and approaches in the classification and analysis of crimes against women. This lack of uniformity leads to the use of different methods and algorithms in different studies, which makes it difficult to compare research results.

The proposed system aims to bridge this huge gap by adopting a holistic approach that not only involves predicting crimes against women, but also proactively helps governments and law enforcement agencies take preventive measures. The system's crime forecasting module goes beyond mere prediction, providing predictions of potential violent events and their likely locations, empowering authorities to take proactive measures to reduce crime. Furthermore, this research is characterized by the fact that it focuses on the analysis of state-specific and annual crime data. This analysis provides valuable information on levels and patterns of crimes against women, presented graphically, facilitating a deeper understanding of the problem based on various criteria, including temporal trends and geographic distribution.

Overall, the holistic approach aims to promote a more effective strategy to combat crimes against women. This emphasizes the importance of not only responding to events, but also preventing them. By providing valuable data-based insights to policymakers and law enforcement agencies, this research gives them the tools and information they need to

develop preventive interventions and develop policies that can make a difference in women's safety.

RESEARCH	FUTURE PREDICTION OF CRIME	FOCUS SPECIFIC FEATURES OF WOMEN CRIMES	POLICE RELATED	CRIME PREVENTION	AVAILABLE FOR SRI LANKAN DATA
COMPREHNSIVE STUDY ON CRIME PREDICTION TECHNIQUES [16]	✓	✗	✗	✗	✗
ANALYSIS AND CLUSTERING-BASED CRIME PREDICTION [20]	✓	✗	✗	✗	✗
MACHINE LEARNING BASED SYSTEM FOR MITIGATING THE SUPPRESSION AGAINST WOMEN AND CHILDREN[15]	✗	✗	✓	✗	✗
SAFE BAND [17]	✗	✓	✓	✗	✗
ANALYSIS AND CRIME PREDICTION BY CLUSTERING [18]	✓	✗	✗	✓	✗
CRIME CLUSTERING AND FORECASTING PREDICTION SYSTEM FOR CRIME AGAINST WOMEN	✓	✓	✓	✓	✓

Table 1.2.1: Comparison between existing research and the proposed component

1.3 Research Problem

Crimes against women, including domestic violence, sexual violence and human trafficking, continue to be major social challenges worldwide. Despite concerted efforts to combat these crimes, the need for robust and tailored crime-fighting strategies remains critical. Crime forecasting, a method that uses historical data analysis to predict future crime and identify underlying patterns, is a promising predictive solution. However, current crime prediction models focus mostly on overall crime rates and do not consider the specifics of crimes against women.

This limitation presents a significant challenge to law enforcement agencies seeking to implement specific preventive measures against such gender-based crimes. In addition, the inherent complexity of these crimes increases the challenge of prediction and prevention. Traditional approaches to policing tend to take a reactive stance that requires responding to crimes after they occur. Instead, clustering and forecasting models offer the ability to identify areas of increased prevalence of crimes against women and predict future crime. This paradigm shift can allow law enforcement agencies to allocate resources judiciously and intervene proactively, thus reducing the occurrence of such crimes.

This paradigm shift allows law enforcement agencies to allocate resources judiciously and act proactively, ultimately limiting the spread of these crimes. The main objective of the proposed research project is to develop an efficient cluster-based crime prediction model adapted specifically for crimes against women. The main purpose of this model is to predict these sexual crimes, classify their types and find likely locations. By providing law enforcement agencies with actionable intelligence, this study aims to improve resource allocation, facilitate timely

preventive action and accelerate the creation of effective crime prevention strategies. Therefore, the study maintains an unwavering commitment to combating crimes against women with the primary goal of creating a safer and fairer society.

The aim of the proposed research project is to develop an effective cluster-based crime prediction model for crimes against women, enabling proactive prevention. The purpose of this model is to predict crimes against women, mark their types and find likely locations. By providing law enforcement agencies with practical knowledge, the research aims to improve resource allocation, promote timely preventive measures and promote the development of effective crime prevention strategies, with a special focus on crimes against women.

1.4 Research Objectives

1.4.1 Main Objective

Develop a comprehensive crime analysis and prediction system to enhance the understanding of crime patterns related to crimes against women in the specified region. This system aims to assist law enforcement agencies, policymakers, and relevant stakeholders in formulating effective crime prevention and intervention strategies. The research seeks to empower stakeholders with data-driven insights, ultimately working towards the reduction of crimes against women and the creation of safer communities.

1.4.2 Specific Objectives

In order to accomplish the primary goal of this proposed system has consecutive sub-objectives need to be obtained.

- Develop a Predictive Model for Crimes Against Women
    - Data Preprocessing
    - Feature Engineering
    - Model Implementation
    - Model Training
    - Model Evaluation
    - Model Deployment

- Identify High Crime Areas and Crimes of Concern
    - High-Risk Areas
    - Priority Crimes
    - Tailored Prevention

- Analyze Historical Crime Trends
    - Data Visualization:

- Provide Data Integration and User Interaction a User-Friendly Interface
    - Scenario Input
    - Exploration
    - User-Friendly Interface

- Enhance Data Collection and Integration

- User Contribution.
- Validation and Integration

- Evaluate and Improve the System
    - Usability Testing
    - Quality Enhancement

  These research objectives form a comprehensive roadmap for developing a predictive model for crimes against women, deploying it in a user-friendly web application, and iteratively improving the system based on user feedback and data integration.

1.4.3 Business Objectives

- Develop a quality product.
- Develop an efficient system.
- Improve law enforcement agencies satisfaction.
- Data Quality and Quantity
- International Collaboration.

2 METHODOLOGIES

2.1 Introduction

This part illustrates the methodology for approach the proposed system's related functions. This is a methodological way of research. Follow the software lifecycle model to implement the system. Research has conducted more studies on the above research area.

This proposed system can be divided into several key stages, each contributing to the overarching goal of predicting and preventing crimes against women effectively. Hence, the gathered information will be used to achieve the main objectives and sub-objectives.

2.2 System overview

Clustering crimes against women and crime forecasting prediction

Considering the outcome from the literature review, most important was to decide the technologies, software solution and most appropriate tools for implementation phase. The proposed system can be divided as following components:

- Comprehensive Data Gathering
- Thorough Data Cleaning and Transformation
- Random Forest Regressor for Future Crime Rate Prediction
- Model Training on Historical Data
- Prevention Strategies for Various Crime Categories

The methodology used in this study to analyze and visualize crimes against women in Sri Lanka is a comprehensive and systematic approach to address this critical social problem. This starts with a careful collection of crime statistics from reliable and authoritative

sources, ensuring the integrity of the data set. Subsequent data preprocessing steps handle missing values and maintain data consistency. The core of the system is a Random Forest Regressor model, chosen for its ability to handle complex, non-linear relationships between predictors and crime. This model is extensively trained using historical crime data to capture complex patterns and trends, and is usable with interactive user interfaces that accept inputs such as state, year, and crime type to predict future crime rates. The predicted results are then interpreted as categorical crime rates, which helps to understand the expected crime rate. The random forest regression model was chosen because of its ability to handle complex, nonlinear relationships between predictors and crime. This model is carefully trained using historical data to capture patterns and trends. To evaluate its predictive performance, user-friendly interfaces were created that accept user inputs, predict future crime based on a selected crime category, and interpret the results into different categories such as "low crime area" or "very high crime area". These predictions are then converted into categorical crime rates, making them easier to interpret. Data visualization techniques such as bar and pie charts are used to visually represent crime patterns to improve user understanding. Data visualization techniques such as bar and pie charts are used to visually represent crime patterns. In addition, they were used to graphically describe the development of crime. In addition, the system extends its usefulness by providing prevention strategies tailored to different crime groups that address the unique challenges associated with each crime category. In conclusion, this methodology provides a comprehensive framework for understanding, predicting and preventing crimes against women, providing valuable insights for crime prevention in Sri Lanka and promoting the safety and well-being of women in the region.

The study also presents a comprehensive method for analyzing and visualizing data on women's crime. The aim of this study is to develop a predictive model to analyze and

visualize crimes against women with a focus on predicting future crime in different states and federations of Sri Lanka.

In addition, the project offers tailored prevention strategies for different crime groups that address the unique challenges associated with each crime category. This research framework provides valuable information on trends in female crime and provides effective crime prevention information.

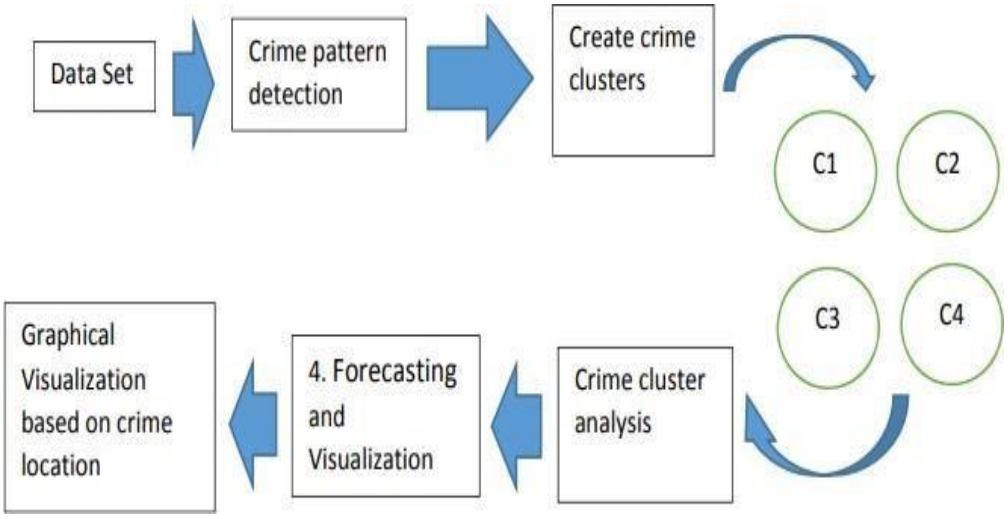


Figure 2.2.1: High level architecture diagram

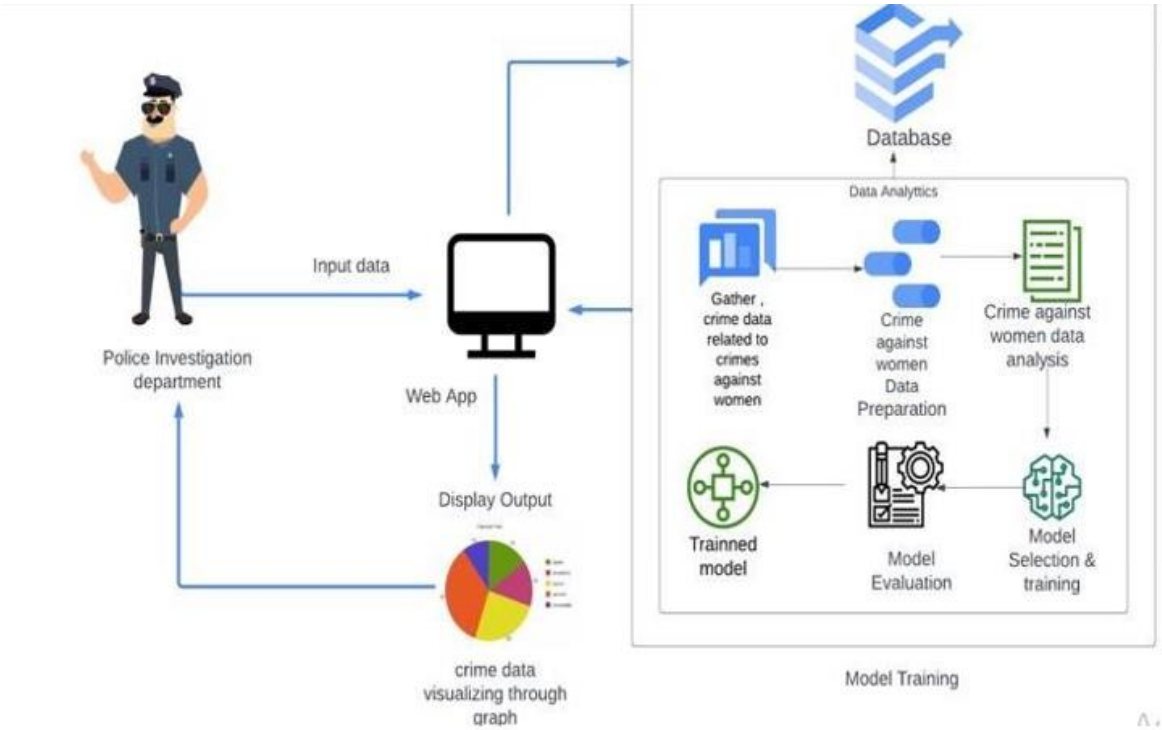


Figure 2.2.2: System diagram for clustering crimes against women and future crime forecasting prediction

Data Collection

The first critical step in our research was data collection, with great emphasis on data quality and reliability, which is essential to the validity of our analysis. The data is obtained from a reliable and authoritative source specializing in crime statistics related to crimes against women in various states and union territories of Sri Lanka. It should be noted that

the credibility of the information source was considered most important to ensure the integrity and accuracy of the information underlying our study.

Data pre-processing:

This was followed by data processing, a painstaking process that lays the foundation for all subsequent analysis and modeling. The steps involved are as follows.

- **Handling missing data**

  Addressing missing data was considered crucial to avoid possible errors in our analysis. A conservative approach was used to remove rows with missing values. This decision ensured that our dataset consisted exclusively of complete and reliable records. The reason for removing missing data was determined based on the proportion of missing values and their potential impact on the entire dataset.

- **Feature Selection**

  The material had several properties, although not all of them were relevant to our research objectives. Careful consideration led to the selection of characteristics considered critical for our analysis, including “STATE/UT”, “Year” and certain types of crimes against women. The selection process of the characteristics was justified, why these particular characteristics were chosen and their contribution to the achievement of the research objectives was explained.

- **Categorical Variable Encoding**

  The dataset contained categorical variables such as "STATE/UT", which cannot be directly exploited by many machine learning algorithms. To solve this, label coding was used to facilitate the conversion of these categorical values into

numerical forms. This process involved assigning unique numerical titles to each category. The choice of coding method and its effect on subsequent modeling was explained.

- **Numerical Feature Standardization**

  Standardization of numerical features was found essential to ensure a common scale for all features, preventing certain variables from potentially dominating the modeling process due to their larger scale. A standard scaler was used to convert the numerical properties to ensure that they had a mean of 0 and a standard deviation of 1. The reasons for this standardization process were explained and its importance in machine learning modeling was highlighted.

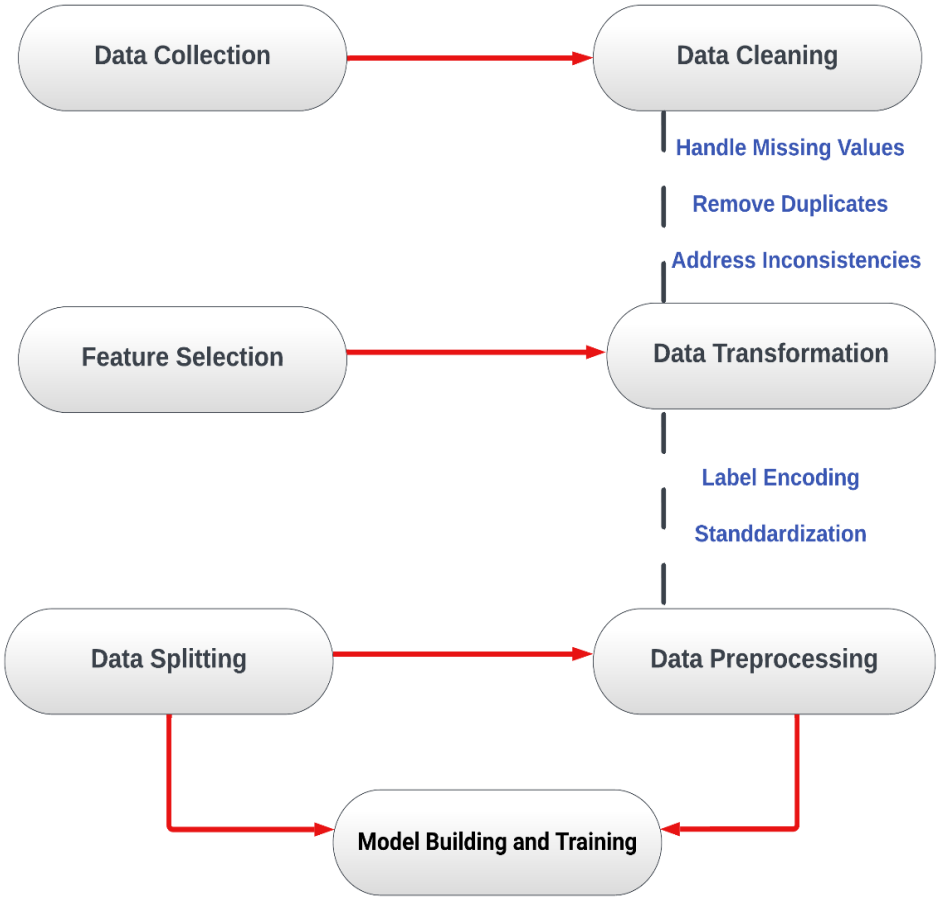


Figure 2.2.3 Data Pre-processing Diagram.

Crime Prediction Model

In relation to this final year thesis, the crime forecasting model is an important part of the research methodology. This section provides a detailed explanation of feature selection, model selection, training and testing processes, and key evaluation metrics used to evaluate model performance.

Feature Selection:

The first step in creating a crime prediction model involved selecting the subset of data that was considered most important in predicting crime. Selected features include:

- STATE/UT: This categorical attribute represents the different states and territories in the dataset and captures the geographic aspect of the crime data.
- Year: The temporal dimension is crucial to understanding how crime rates change over time, so "year" is an important feature.
- Crime types: Several crime types such as 'Rape,' 'Kidnapping and Abduction,' 'Dowry Deaths,' and 'Importation of Girls" were selected as predictor variables. These crime categories represent the specific crimes of interest in this study.

Model Selection:

The Random Forest Regressor algorithm was chosen to develop an effective crime prediction model. Here's why:

- Robustness: Random Forest is known for its robustness when dealing with complex data sets, making it suitable for analyzing crime data, which often has complex patterns and relationships.
- Compatibility with Mixed Types: Crime data often contains a mixture of categorical (e.g. STATE/UT) and numerical attributes (e.g. 'Year' and crime count). Random Forest can handle both types of data efficiently.

Training and testing:

The dataset was divided into two separate subsets: the training set and the test set. This distribution was crucial to accurately assess model performance. The distribution was as follows:

- Training set (80%): The majority (80%) of the data set was reserved for training the model. During training, the model learned patterns, relationships and trends from the data, which allows it to make predictions.
- Test set (20%): The remaining 20% of the data was reserved for models. This subset served as independent evaluation data used to evaluate the predictive accuracy of the model.

Evaluation metrics:

Several evaluation metrics were used to measure the effectiveness of the model:

- Mean Squared Error (MSE): MSE measures the mean squared difference between the predicted crime rate and the actual values in the test data set. A lower MSE means a more accurate model.
- R-squared (R^2) scores: R^2 scores estimate the proportion of variance in the dependent variable (crime rates) that is explained by the independent variables (features). A higher R^2 score indicates that the model fits the data better. The R^2 score quantifies the proportion of the variance in the dependent variable (crime counts) that is predictable from the independent variables (features). An R^2 score of 0.9604 suggests that approximately 96.04% of the variance in crime counts is explained by the model.

- Accuracy scores: Accuracy scores are a general measure of regression problems that represent the overall accuracy of the model in predicting crime rates. It shows the proportion of correctly predicted values in the test data set.

Interpretation results:

After the Random Forest Regression was trained and evaluated, it was ready to make predictions. This model is used to predict the number of specific crimes for future years and states based on user input. These predictions were then interpreted to provide insight into potential crime rates.

The crime prediction model developed for this thesis is an important analytical tool for understanding and predicting crimes against women in Sri Lanka. Its robustness, compatibility with mixed types and strong performance on evaluation metrics make it a valuable asset for research purposes. The predictions and interpretations of the model contribute significantly to the broader study of crime trends and prevention strategies.

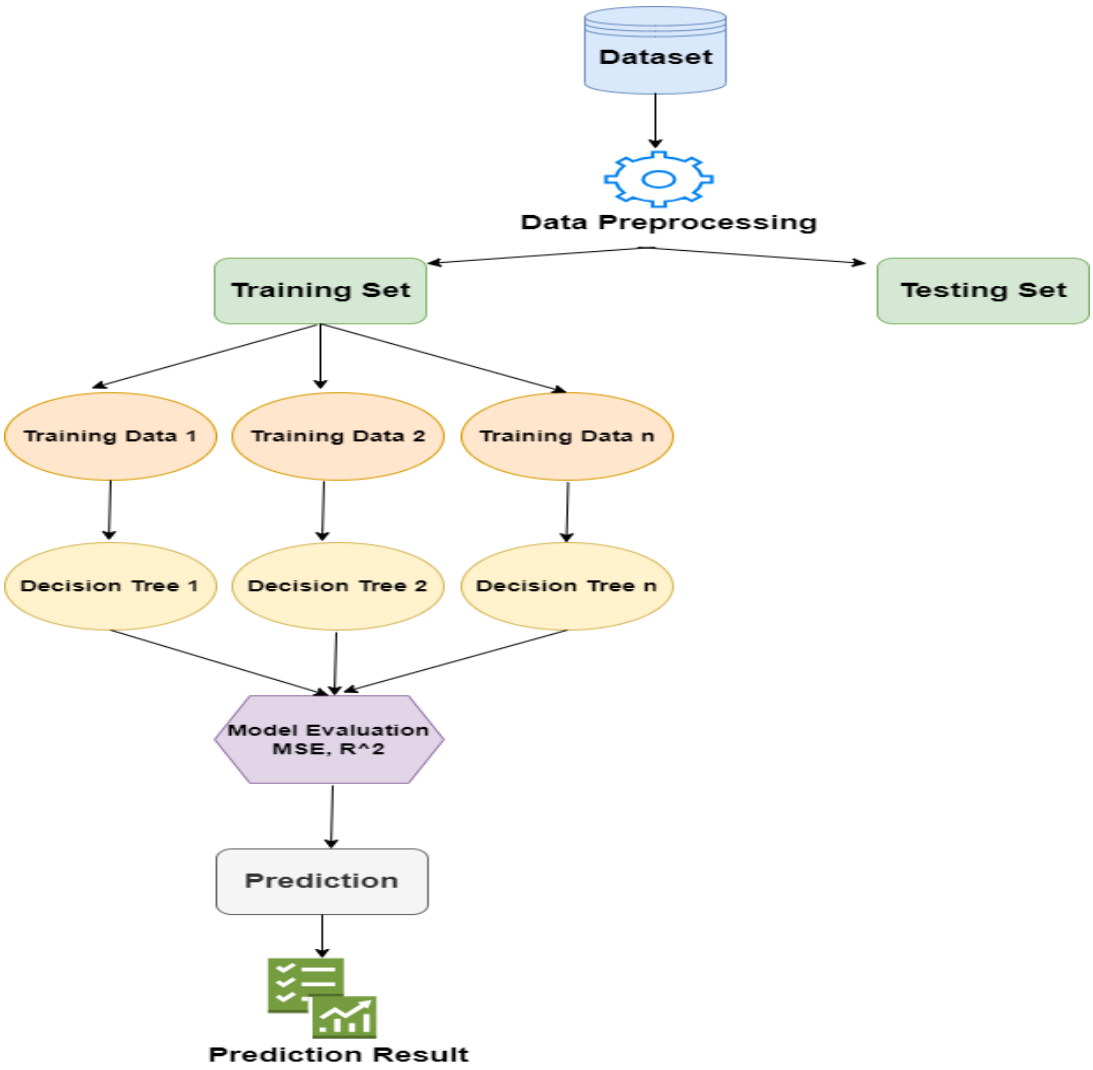


Figure 2.2.4: The training and prediction process of clusteringcrimes against women and future crime forecasting prediction.

Crime analysis and prevention:

To comprehensively understand women's crime knowledge and develop effective crime prevention strategies, two core modules are implemented in the system - Crime Analysis Module and Crime Prevention Strategies Module. Supported by robust methodologies, these modules are the cornerstone of our research and contribute significantly to the critical issue of crimes against women.

Crime Analysis Module:

The crime analysis module is the first critical part of our research. Its main purpose is to delve into historical crime data, extract valuable insights and present those results in an understandable way. Below the most important aspects of this module are explained:

- Data exploration
  This module allows users to interact with the system by selecting a specific crime category of interest. The system uses this input to retrieve relevant historical crime data.

- Visual representation
  Once the data is acquired, data visualization techniques are used to transform it into intuitive graphical representations. In particular, bar charts and pie charts are used to present crime trends over time and across states. These visualizations allow users to easily identify patterns, variations, and anomalies in data.

- Interpretation
  The graphical representations created by this module facilitate the interpretation of complex crime data. With a visual narrative, users can easily understand the

dynamics of crime trends and their evolution over the years. It helps in informed decision making and policy making.

Crime Prevention Strategies Module:

The crime analysis module complements the crime prevention strategy module, which is an important part of our research framework. This module aims to provide tailored and proactive strategies to mitigate crimes against women. Below the main elements of this module are explained:

- Prediction of highest crime
  Using predictive modeling techniques, this module identifies the crime category predicted to be highest in a given year and state. This foresight is important in proactively dealing with crime problems.

- Prevention strategies:
  The module generates a set of crime prevention strategies for the highest crime category identified. These strategies are carefully tailored to the nuances of each type of crime. This precision ensures that the recommendations are not one-size-fits-all, but rather address the unique challenges associated with each crime category.

The Crime Analysis and Crime Prevention Strategies modules form the analytical and strategic pillar of our research. The former provides users with a comprehensive view of historical crime data, while the latter empowers stakeholders to take proactive steps to

combat crimes against women. Together, these modules form a powerful framework that uses data-driven insights to improve public safety and protect women's rights. They reflect the synergy between data analysis, visualization and predictive decision-making that significantly contributes to our overall goal of a safer and fairer society.

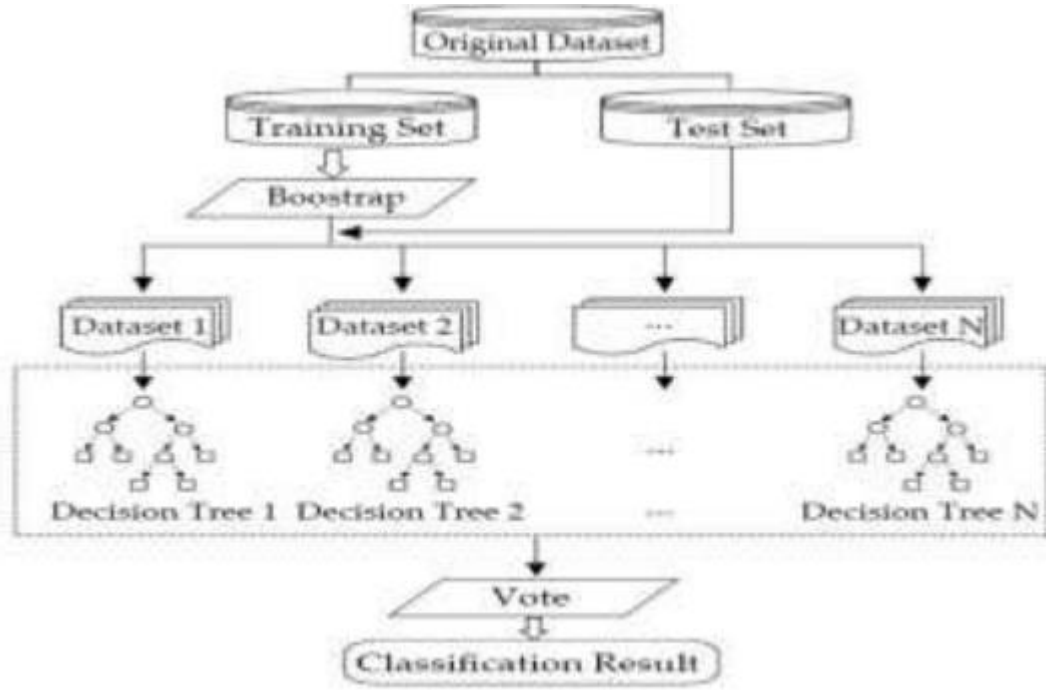


Figure 2.2.5: Random Forest Architecture

2.3 Diagrams

2.3.1 Work Breakdown Structure

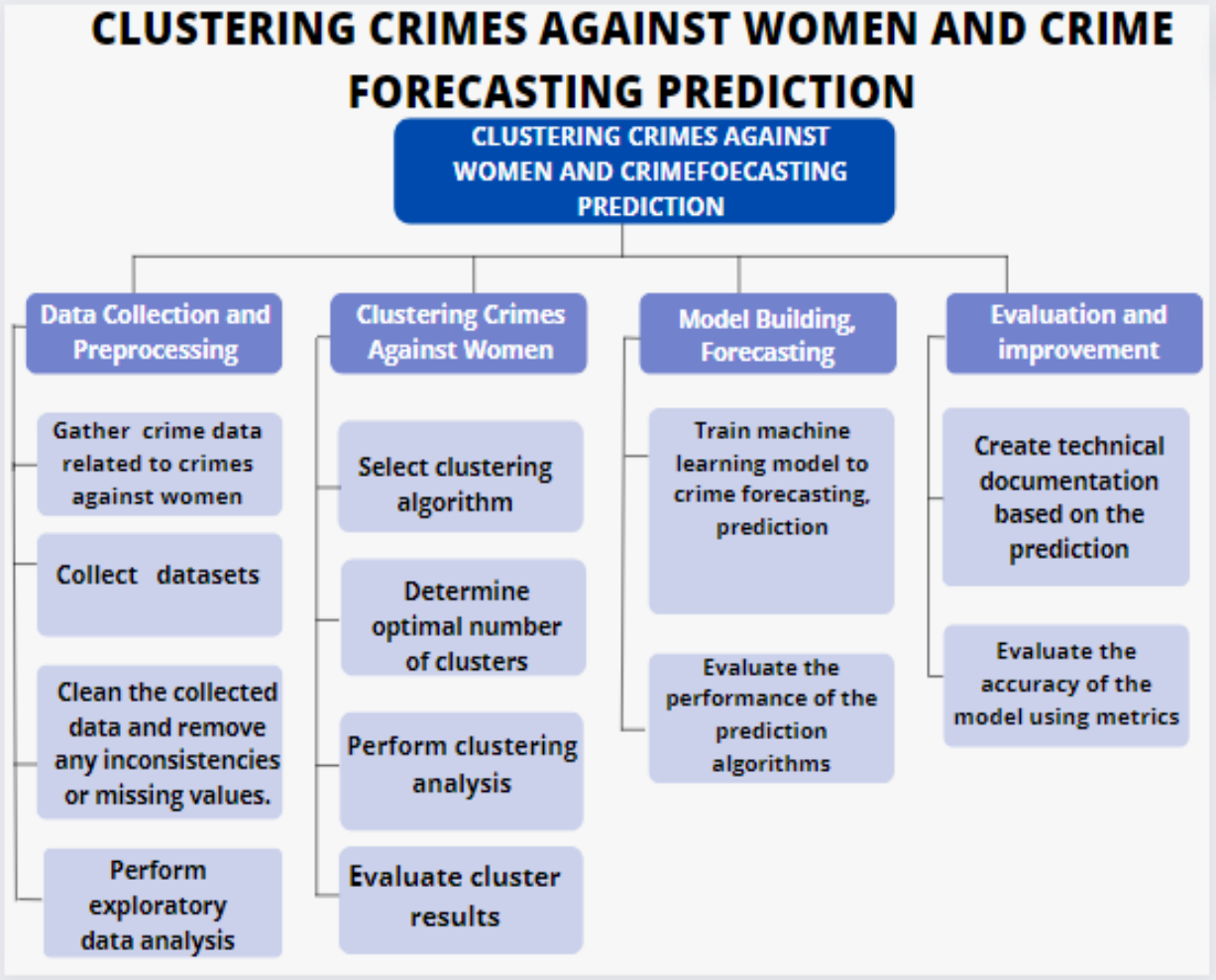


Figure 2.3.1: Work Breakdown Structure of the system

2.3.2 Flowchart

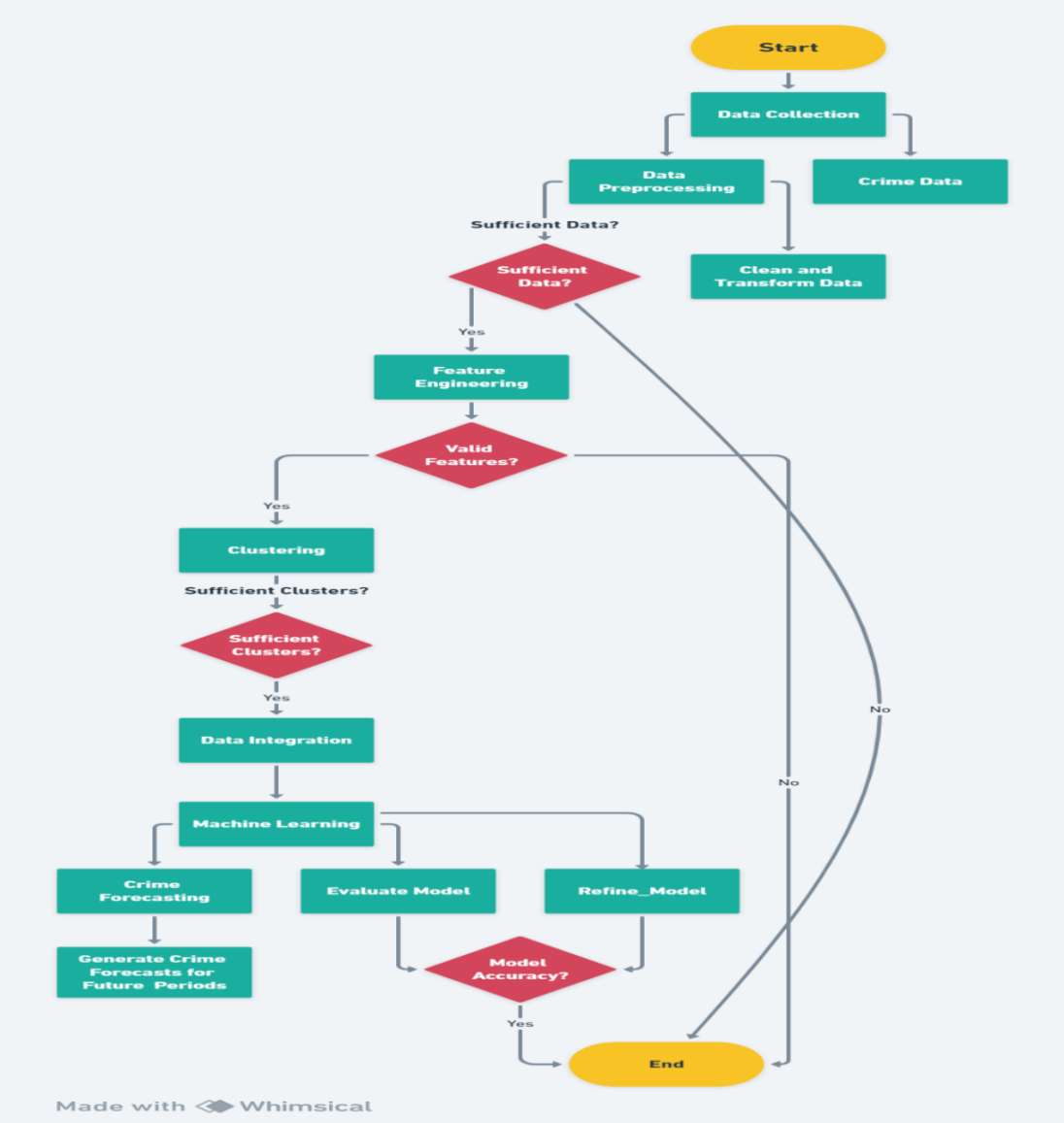


Figure 2.3.2: Flow Chart of the system

2.4 Development Process

The development process for the research component focused on the analysis and prevention of crimes against women closely followed the Waterfall model, which was chosen for its suitability for well-defined requirements and minimal expected change. Since the waterfall model is linear, it aids in the efficient development of the system, and the needs and specifications are clearly defined, it doesn't change according to the timetable. This approach is chosen after identifying and stating the system's needs. The research journey began with the recognition of a problem at hand - the rise in crimes against women. This was followed by a comprehensive needs assessment phase using extensive research and stakeholder consultation. This knowledge was then translated into carefully structured functional requirements, which were described in more detail by creating detailed use case diagrams. Clear expectations were set for each functional component, closely aligned with the overall goal of effectively predicting and preventing such crimes. To ensure proper allocation of time and resources during the year-long research period, a detailed project schedule with planned task orders was prepared. The development process went through distinct stages, starting with a thorough analysis of requirements, system design and careful implementation of functional components. This was followed by rigorous testing, deployment in a controlled environment and user training. Post-commissioning maintenance and upgrades have been carefully integrated, ensuring the continued effectiveness of the system. This approach, illustrated in Figure 2.4.1, involves breaking down functional requirements into manageable components. In short, it can be stated that the Waterfall

model enabled a clear, systematic and effective development path, which resulted in a solid tool for analysing and preventing crimes against women.

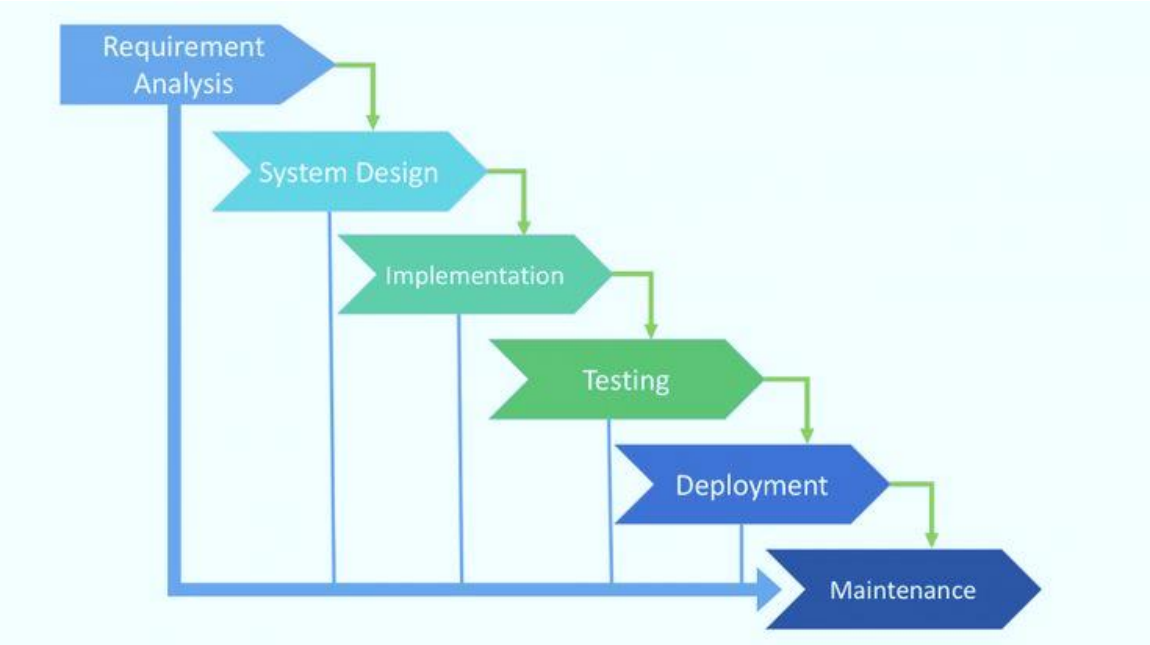


Figure 2.4.1: Development process of the system

Project Management

The primary objective of software project management is to offer a team of developers the tools they need to operate efficiently toward the timely and successful completion of a project. However, managing software projects is an extremely challenging process.

Therefore, a project management strategy was developed to carefully plan, monitor and direct each stage of the research. It involves a systematic approach to tasks such as data collection, pre-processing, model development, evaluation, and interpretation of results. Sharing roles and responsibilities, managing resources and sticking to a well-defined schedule are essential. In addition, risk assessment and mitigation strategies are integrated to address potential challenges. The ultimate goal is for the research team to be able to efficiently move through the complex stages of data analysis, modeling and forecasting, ensuring that the objectives of the thesis project are met within the agreed time frame. Effective project management not only streamlines the research process, but also improves the quality of the final results, which contributes to the overall success of the thesis.

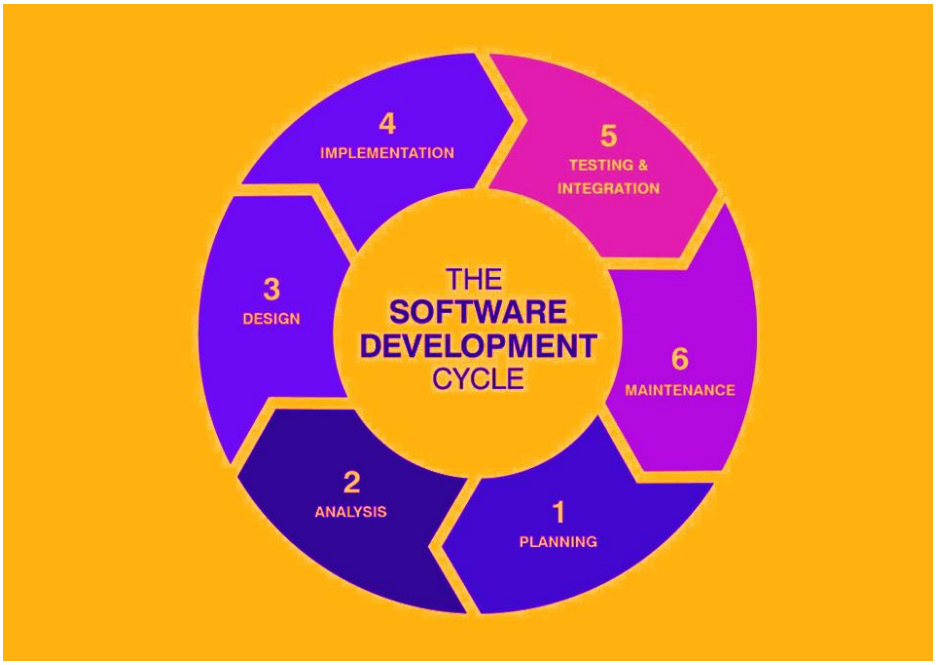


Figure 2.4.2: Software Project Management Process

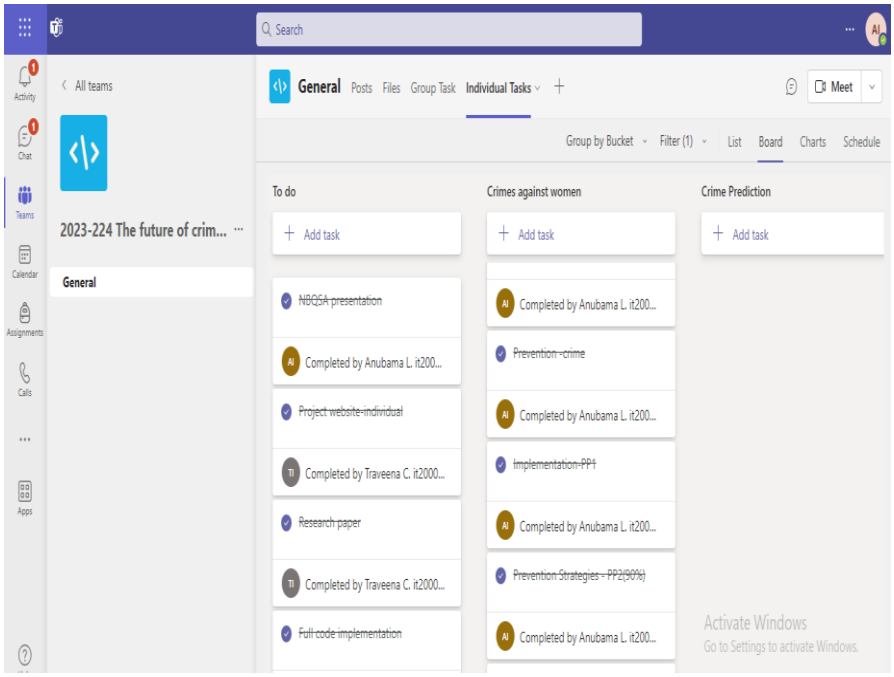


Figure 2.4.3: Software Project Management for the system

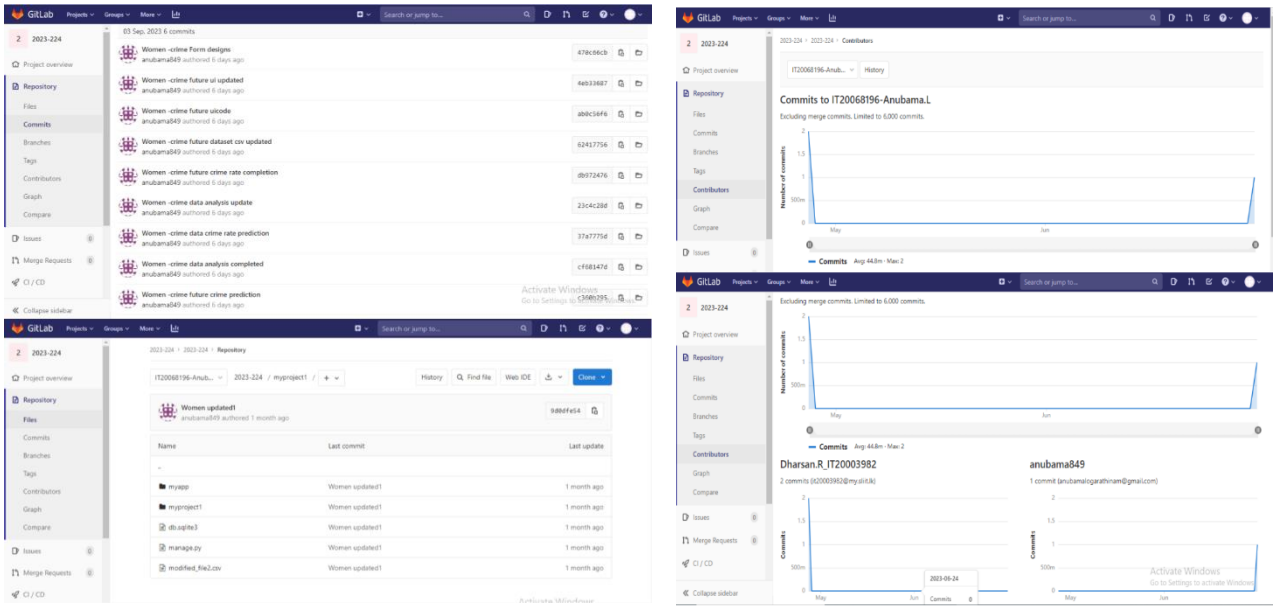


Figure 2.4.4: Project Code Management for the system

2.5 Requirements Gathering

The data collection process for the crime cluster and crime-against-women forecast involved working with key stakeholders and examining multiple data sources. Valuable insights were gained from the Deputy Inspector General (DIG) of Kurunegala Police on the practical needs of law enforcement and regional challenges. Mr. M.K.M Sajath legal expertise provided a perspective on human rights, evidence-based decision-making and victim support. Also government databases of historical crime data, partnered with NGOs and women's rights organizations are used to gain victim-focused insights, consulted academic experts to align research, and conducted community surveys to gauge public perceptions. These diverse requirements have carefully shaped the research to be data-driven, socially responsible and responsive to the needs of stakeholders to effectively combat crimes against women.

Functional Requirements:

- Crime Detection of patterns and trends in crimes against women.
- Clustering by factors like location, crime type, and time.
- Offering insights and recommendations for prevention.
- Forecasting future crimes using historical data.

Non-functional Requirements:

- Usability

  The system has user-friendly interface that is easy to use and navigate.
- Security

The system secures and protects sensitive data related to crimes against women.

- Availability

  The system is reliable and available 24/7 to ensure that law enforcement agencies and policy makers can access crime data and insights at any time.

- Scalability

  The system is scalable and able to handle an increasing amount of crime data as more data iscollected over time.

- Accuracy

  The system provides accurate and timely crime predictions and insights to ensure that law enforcement agencies and policymakers can take proactive measures to prevent crimes against women.

2.5.1 Resources Used

2.5.1.1 Software Boundaries

Backend - Python Language

Python serves as the backbone of the application. It's used to write the server-side logic for handling HTTP requests and responses. In this case, Python is used for processing user inputs, making predictions based on machine learning models, and handling data manipulation tasks. Libraries like Pandas, NumPy, Scikit-Learn, and Django are employed to manage data, perform calculations, and implement machine learning models.

Visual Studio Code Editor

Visual Studio Code (VS Code) is the integrated development environment (IDE) where Python scripts are written and developed. It provides features like code auto completion, debugging tools, and extensions for Python, making it a powerful tool for writing and managing Python code efficiently.

Frontend

HTML, CSS, JavaScript: The frontend of the web application is built using HTML, CSS, and JavaScript. HTML is used for creating the structure and content of web pages. CSS is used for styling the user interface, making it visually appealing. JavaScript is employed for adding interactivity to the web application, such as handling user inputs and triggering requests to the backend.

Web Application

The web application acts as the user interface through which users interact with the backend component. Users can input data, select options, and trigger predictions through web forms. The web application communicates with the Python backend by sending HTTP requests with user inputs and receives responses containing prediction results.

Libraries

1. **NumPy –** It is used for working with arrays. It supports for large, multi-dimensional arrays and matrices, including mathematical functions.

2. **Pandas -** An open-source toolkit called the Panda library was created to make data analysis simpler and is simple to use. It offers a data structure with a quick response time and a simple user interface.

   Pandas is based on NumPy, a Python library used for mathematical computations, and Matplotlib, a Python library used for data visualization. Pandas, which serves as a wrapper for matplotlib and NumPy, makes it possible to use their methods with a minimum of code. For instance, the pandas plot () method combines several matplotlib methods into a single method, allowing for the quick and easy plotting of charts in only a few lines of code. Commonly used for data manipulation and data analysis.
3. **Django Libraries-** (django.shortcuts, django.http, django.contrib.messages): These are fundamental Django libraries used for handling web requests, rendering templates, managing HTTP responses, and displaying messages to users.

4. **Pandas (import pandas as pd)-** Pandas is used extensively for data manipulation and analysis. It's used to read and process the crime dataset, perform data preprocessing, and manipulate data structures like Data Frames.
5. **Scikit-Learn-(**from sklearn.preprocessing, from sklearn.ensemble, from sklearn.model_selection, from sklearn.metrics): Scikit-Learn is a powerful library for machine learning tasks. It's used for preprocessing data (e.g., standardization and encoding), building predictive models (Random Forest Regressor), splitting data into training and testing sets, and evaluating model performance.

6. **Matplotlib (import matplotlib.pyplot as plt) -** Matplotlib is used for creating various data visualizations, including bar charts and pie charts. It's crucial for presenting the analysis results visually.

7. **Seaborn (import seaborn as sns) -** Seaborn is built on top of Matplotlib and is used to enhance the aesthetics of Matplotlib plots. It's often used to set plot styles and color palettes.

8. **IO (io, BytesIO) -**These are Python's built-in libraries for handling input and output operations. In this project, they're used to create and manipulate in-memory byte streams, which are particularly useful for generating and storing images in web applications.

9. **Base64 (import base64) -** Base64 is used to encode binary image data into a format that can be embedded directly into HTML and displayed as images in the web application.

10. **CSV (import csv) -** The CSV module is used for reading and writing CSV (Comma-Separated Values) files. It's used to add new data entries to the dataset.

11. **Random (import random) -** The Random module is used for generating random values. In this project, it's used to generate random colors for data visualization.

12. **Plotly Express (import plotly.express as px) -** Plotly Express is a library for interactive data visualization. It's used for creating interactive charts and graphs, although it's not used extensively in this code.

2.6 Commercialization of the Product

The women crime cluster and forecasting system offers significant opportunities for service oriented implementation and provides a transformative solution for various stakeholders. The system is mainly aimed at law enforcement agencies, authorities and policy makers who are focused on improving women's safety. The system has a user-friendly interface that ensures users with different technical backgrounds. Its exceptional predictive accuracy, based on advanced machine learning and data analysis techniques, enables stakeholders to proactively allocate resources and address potential issues. Especially in places where machine learning-based crime analysis tools are scarce, such as Sri Lanka, this system opens up opportunities for data-driven insight and predictive crime prevention.

In addition to the financial benefits for government agencies in particular, its applicability to different categories of women crime makes it a versatile tool for solving crime-related challenges worldwide. Future expansion possibilities for the system are promising as it paves the way for real-time crime reporting, automatic reporting and international cooperation between law enforcement agencies. Basically, the crime group and prediction system is not only a research project, but also a practical service that is ready to make significant progress in crime prevention and public safety. Aimed primarily at law enforcement agencies, government agencies and policy makers focused on women's safety,

Service Accessibility: The women crime cluster and forecasting system is strongly designed for service and not for profit. It offers an intuitive and user-friendly interface that ensures its features can be easily accessed and used by users from a variety of technical backgrounds, including law enforcement agencies, government agencies and decision makers committed to improving women's safety. Primary goal is to positively impact women safety by providing easy-to-use tools for data-driven decision-making.

Advanced Public Safety: Based on advanced machine learning and data analysis techniques, the system excels in predictive accuracy. This allows stakeholders to proactively allocate resources and effectively address potential security issues. Especially in places like Sri Lanka, where access to machine learning-based crime analysis tools is limited, this system acts as a catalyst for informed decision-making and women crime prevention. This system is focused on improving women safety.

Breaking Information Barriers: This web-based application is an important resource in areas with limited access to cutting-edge technology for women crime analysis and prevention. By breaking down information barriers, this system aims to facilitate a wider reach of users.

Effective use of resources: Automating the processes of systematic data analysis and crime forecasting significantly increases time and resource efficiency. Law enforcement agencies and decision makers often struggle with massive amounts of data. By streamlining these processes, this system enables more efficient allocation of resources and decision-making, ultimately contributing to the greater good of society.

Public Investment: This system is not driven by the pursuit of profit, but by commitment to public services. It offers a cost-effective solution for government agencies and institutions. By reducing operational costs associated with manual data analysis, these

resources can be directed to other key areas of crime prevention and public safety, promoting community well-being.

2.7 TESTING & IMPLEMENTATION

2.7.1 Testing

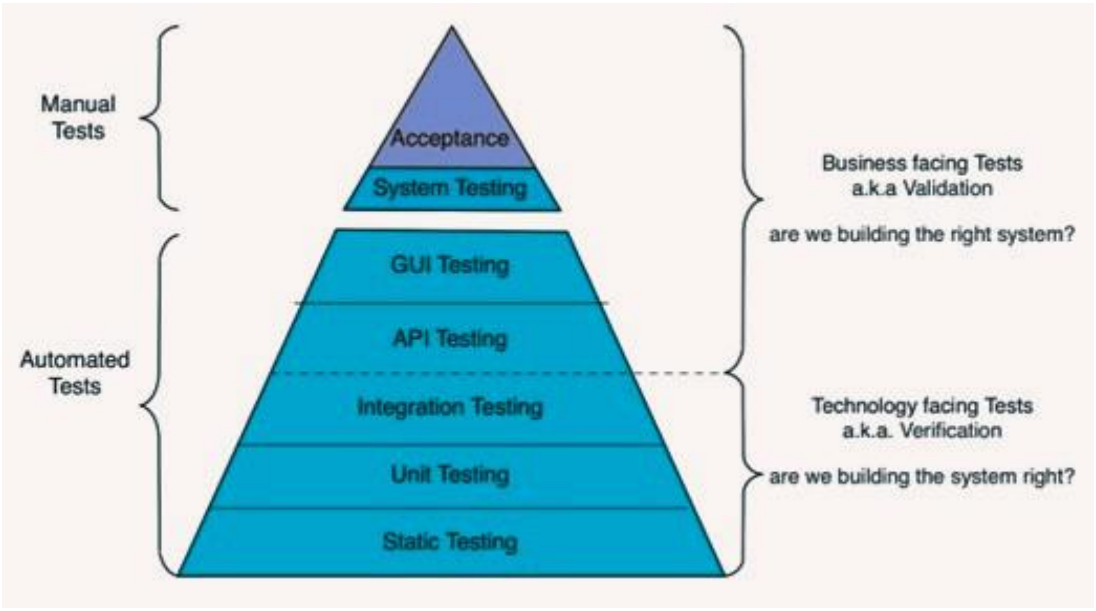


Figure 2.7.1.1: Levels of testing in software testing

Above figure 2.7.1.1 shows a prominent way to test the software before release at each development phase, testing the application, and fix the bugs. The successful development of a system for clustering crimes against women and forecasting crime rates necessitates a meticulous testing and implementation phase. This section outlines the comprehensive process undertaken to ensure the system's reliability, accuracy, and real-world applicability. Through testing, applications mostly ended up sort of more quality and the most excellent deliverable item. Software testing can generally be exceptionally basic in the development life cycle to recognize defects and mistakes that have been subtly made during the development stages. And it is very essential to make sure of the quality of the item. It is required for the effective performance of a product or computer program application, contrary to popular belief. It is required each stage that are being taken after in SDLC has underneath subtly testing categories.

1. **Unit Testing:**

   Unit testing involves scrutinizing individual components of the system in isolation. During this phase, the crime prevention system's clustering algorithms are tested to ensure that they accurately group similar crimes against women cases based on various factors such as location, crime type, and time of day. Additionally, unit tests verify that predictive models produce precise forecasts using historical crimes against women data. The testing process also assesses the correctness of data processing and transformation functions, critical for accurate analysis.

2. **Integration Testing:**

   Integration testing is essential to determine how different components of the crimes against women prevention system work together seamlessly. It ensures that data flows effectively between modules such as clustering, pattern recognition, and predictive analytics. Additionally, this phase assesses the system's ability to handle large datasets efficiently, as real-world crime data can be extensive and complex**.**

3. **System Testing**

   It is carried out once the system as a whole has been integrated. This testing evaluates each of the system's functional requirements. This kind of testing is to assess the end-toned specifications. The team members will troubleshoot the problem after it has been found during this testing phase.

4. **Regression Testing:**

   Regression testing is crucial for maintaining system stability over time. This phase involves re-running previous tests to verify that recent code changes have not introduced any new bugs or regressions. A diverse set of historical crime data is used to confirm the system's consistency and reliability across different scenarios.

5. **Module Testing**

   As an extension of integration testing, unit testing focuses on individual components or modules of the system. The purpose of this phase is to ensure the functionality, efficiency and reliability of each module individually. In this context, modules include crime analysis module, crime prevention strategies and others.

6. **Performance Testing:**

   Performance testing evaluates the system's responsiveness and resource utilization under various loads. It measures response times and resource consumption to ensure that the system can handle concurrent user requests without significant slowdowns. Furthermore, scalability is assessed to determine if the system can manage increased data volumes without compromising performance.

7. **User Acceptance Testing (UAT):**

   User Acceptance Testing involves engaging with law enforcement agencies and end-users to validate the system from a practical standpoint. Feedback is collected on user interface design, ease of use, and the relevance of insights and recommendations. This phase ensures that the system aligns with the needs and workflows of the intended users, optimizing its utility in real-world scenarios.

Maintenance

The maintenance phase of an advanced crimes against women prevention system is an important part of its life cycle. This includes continuous monitoring, bug fixes and updates to ensure system stability and security. Regular performance evaluations and scalability checks are performed to adapt to changing data volumes and user needs. Compliance with changing laws and data protection regulations is carefully followed. User feedback and usability improvements are considered and potential vulnerabilities are proactively patched. This step is important to ensure the long-term effectiveness and sustainability of the crime prevention system to support law enforcement and ensure public safety.

Test cases that are done for each testing method is shown below.

Test Case No	Test Case 01
Description	Verify whether empty validation message is fired for each empty mandatory field.
Test Steps	1. Login to the system. 2. Navigate to the 'Women-crime' option. 3. Click on the 'Add data' option. 4. Submit the form by clicking on the 'Upload' button without providing values to the fields.
Test Data	String, Integer
Expected Result	The form should not be submitted with missing data. Validation messages should get fired for each empty mandatory field
Actual Result	Pass
User Role	Police team.

Table 2.7.1. 1 Ensure the form is not submitted with missing data Manual Test Case 01

Test Case No	Test Case 02
Description	Verify whether error message is fired when user enters invalid data in one or more fields (e.g., text in the "Year" field).
Test Steps	1. Login to the system. 2. Navigate to the 'Women-crime' option. 3. Click on the 'Add data' option. 4. Enter text in 'Year' field or crime fields(e.g, Rape, Kidnapping and Abduction etc) 5. Submit the form by clicking on the 'Upload' button.
Test Data	Integer
Expected Result	Verify that appropriate error messages are displayed. Ensure the form is not submitted with invalid data.

Actual Result	Pass
User Role	Police team.

Table 2.7.1. 2 Form Submission with Invalid Data Manual Test Case 02

Test Case No	Test Case 03
Description	Verify whether after a successful form submission, the prediction results are displayed correctly.
Test Steps	1. Login to the system. 2. Navigate to the 'Women-crime' option. 3. Click on the 'Highest Crime Prediction' option. 4. Select 'Division', 'Year' options from the dropdown. 5. Click on the 'Predict' button.
Test Data	String, Integer

Expected Result	The predicted counts for each crime type should be shown. Highest predicted crime and its count should be displayed.
Actual Result	Pass
User Role	Police team.

Table 2.7.1. 3 Display of Prediction Results Manual Test Case 03

Test Case No	Test Case 04
Description	Verify that you are redirected to the correct page, possibly with information related to preventing the highest predicted crime.
Test Steps	1. Login to the system. 2. Navigate to the 'Women-crime' option. 3. Click on the 'Highest Crime Prediction' option. 4. Select 'Division', 'Year' options from the dropdown. 5. Click on the 'Predict' button.

	6. Predicted counts for each crime type are shown
Test Data	String
Expected Result	'Prevention Strategies' option should be displayed. User should be redirected to the correct page, possibly with information related to preventing the highest predicted crime.
Actual Result	Pass
User Role	Police team.

Table 2.7.1. 4 Functionality of Prevention Strategies Link Manual Test Case 04

Test Case No	Test Case 05
Description	Verify whether Analysis page loads successfully.
Test Steps	1. Login to the system. 2. Navigate to the 'Women-crime' option.

	3. Click on the ‘Analysis’ option.
Test Data	
Expected Result	Analysis page should be loaded successfully. Form for selecting a crime should be displayed.
Actual Result	Pass
User Role	Police Team

Table 2.6.1. 5 Functionality of Analysis page Link Manual Test Case 05

2.7.2 Implementation

To predict the product, should have to follow some tasks that need to be followed to implement. A well-designed web application can attract user and they will repeat a share with their community. But implementing a web application and train the model without any issues is very difficult. Applying crime clustering and forecasting to the research framework is an important step in the efforts to combat crimes against women. This section describes the practical aspects of how these components are implemented.

This culminated in a user-friendly, data-driven system that can provide valuable information about crime patterns, predict future crime levels, and provide reasonable crime prevention strategies. This implementation is a powerful tool in ongoing efforts to combat crimes against women and has the potential for broader applications to improve public safety and decision-making.

- Data Collection and Pre-processing:
- Clustering Crimes:
- Crime Forecasting:
- Evaluation and Validation:
- User Interaction and Visualization:
- Crime Prevention Strategies:
- Deployment and Accessibility:

Code

Install the necessary libraries:

views.py

policeAnalysis > myapp > views.py

```
from django.shortcuts import render
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
from .models import CrimePrediction
import matplotlib.pyplot as plt
from django.http import HttpResponse
import seaborn as sns
import io
import base64
from django.shortcuts import render, redirect
from django.contrib import messages
import csv
import random
import plotly.express as px
from io import BytesIO
```

Figure 2.7.2.1: Import libraries for Clustering crimes against women and crime forecasting prediction

Several important external libraries have been added to this Django-based website to improve various aspects of Clustering crimes against women and crime forecasting prediction project. These libraries include pandas and numpy for efficient data processing and manipulation, scikit-learn for machine learning tasks like feature scaling and regression, and matplotlib with seaborn for creating informative data visualizations. In addition, the io and base64 libraries handle file I/O operations, while Django's built-in modules such as django.shortcuts and .models handle web rendering and database interactions. Interactive data visualization is realized with plotly.express, and binary data processing is smoother with BytesIO. Finally, the csv and random libraries play a

role in file management and random value generation. Together, these libraries enable an application to process data, build predictive models, visualize insights, and deliver a robust website.

views.py

policeAnalysis > myapp > views.py

```
# Read the CSV file into a pandas DataFrame
data = pd.read_csv("modified_file2.csv")

# Select the desired features
selected_features = ['STATE/UT', 'Year', 'Rape', 'Kidnapping and Abduction', 'Dowry Deaths',
                     'Assault on women with intent to outrage her modesty', 'Importation of Girls']
data = data[selected_features]

# Handling Missing Values
# Check for missing values
missing_values = data.isnull().sum()
data = data.dropna()  # Remove rows with missing values

# Select the desired features for clustering
selected_features = data[["STATE/UT", "Year", "Rape", "Kidnapping and Abduction", "Dowry Deaths",
                          "Assault on women with intent to outrage her modesty", "Importation of Girls"]]

# Encode categorical variables
label_encoder = LabelEncoder()
selected_features['STATE/UT'] = label_encoder.fit_transform(data['STATE/UT'])

# Normalize the numerical features
numerical_features = ["Rape", "Kidnapping and Abduction", "Dowry Deaths",
                      "Assault on women with intent to outrage her modesty", "Importation of Girls"]
scaler = StandardScaler()
selected_features[numerical_features] = scaler.fit_transform(selected_features[numerical_features])

# Separate the features and target variable
X = selected_features.drop(["Assault on women with intent to outrage her modesty"], axis=1)
y = selected_features["Assault on women with intent to outrage her modesty"]
```


Figure 2.7.2.2: Data Pre-processing for Clustering and Prediction

It begins by loading the CSV dataset into a pandas Data Frame, focusing on specific columns relevant to the project's goals, such as crime statistics for different states and years.

The next step involves handling missing values, which is essential to maintain data quality. It detects and counts the missing values in the dataset and then deletes the rows with missing values.

After cleaning the data, the code encodes categorical variables using label encoding, a necessary transformation to convert non-numerical data into a format suitable for machine learning models.

To ensure consistency and comparability of numeric properties, the code standardizes them using a StandardScaler, which scales the values to have a mean of 0 and a standard deviation of 1. Finally, the code separates the dataset into characteristic variables (X) and target variables (y), where X contains all the independent variables and y represents the dependent variable, which is the number of such attacks. This pre-processing prepares the data for subsequent machine learning tasks such as model training and evaluation.

```
53
54  # Split the data into training and testing sets
55  X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
56
57  # Create and train the random forest regression model
58  model = RandomForestRegressor(n_estimators=100, random_state=0)
59  model.fit(X_train, y_train)
60
61  # Define the list of crimes
62  crimes = ['Rape', 'Kidnapping and Abduction', 'Dowry Deaths',
63            'Assault on women with intent to outrage her modesty', 'Importation of Girls']
64
65
```

Figure 2.7.2.3: Building a predictive model

This code segment is an important step in building a predictive model. It first divides the dataset into training and testing sets and reserves 80% of the data for training and 20% for testing. This separation is necessary to accurately assess model performance.

Next, it creates a random forest regression model, a powerful machine learning technique used for both classification and regression tasks. In this case, it is used for regression, specifically to predict "attack on women to outrage her modesty" based on other selected characteristics. The model configuration includes specifying 100 decision trees (n_estimators) in a Random Forest, which ensures robustness and reduces overfitting. The "random_state" parameter ensures repeatability.

3 RESULTS & DISCUSSION

This chapter discuss the results obtained from the development of the personality prediction model component.

3.1 Results

This system will be used as a web application for forecasting prediction for crimes against women.

Figure 3.1.1: - Crime Rate Prediction and Classification

Moreover in this system Random Forest Regressor model is leveraged to predict future crime rates based on user input. Users specify a future state, year, and crime type through a form. The code processes this input, making predictions for the chosen crime in the selected state and year. Predictions are categorized into four levels of crime rates for user-friendly interpretation as ‘Low’, ’High’, ’Very high’, ’Medium’.

This which makes them easier to interpret. In addition, the function stores these predictions in a database for later use and analysis. This feature is in line with the system's goal of helping law enforcement agencies and policymakers proactively combat crimes against

women. Now, in the testing phase, this function becomes central. The system is put through a rigorous testing process following the test plan. This plan covers various scenarios, input combinations, and real-world use cases to ensure system reliability and dependability.

Once the testing is completed, the results obtained by the system will be critically evaluated. The actual results is compared with the expected results defined in the test plan. This step is crucial to evaluate the accuracy and efficiency of the system.

The results of this testing phase will be tests of the capabilities of the system. They provide valuable information about the performance of the Random Forest Regressor model, the accuracy of the predictions and the usability of the classification system. By matching actual results to the expectations, the functionality of the system is confirmed, and identified potential areas for improvement.

Essentially, this later part of the testing phase represents the culmination of the system's journey from computation to actionable knowledge. The system strengthens the capacity to empower law enforcement and policymakers to proactively combat crimes against women.

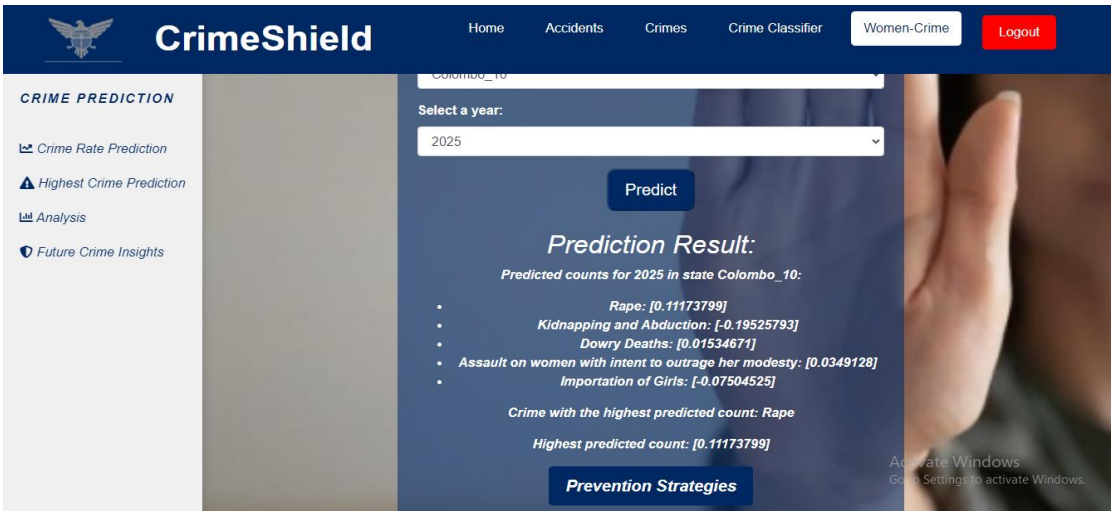


Figure 3.1.2: - Crime Rate Prediction and High Crime Analysis

Highest crime count prediction feature that predicts the type of crime expected to have the highest count in a specified future year and state. It utilizes a Random Forest regression model trained on historical crime data. Random Forest Regression, a powerful machine learning model to predict what crimes are expected to be highest in the next year and state. When the user enters these parameters, the function systematically analyses different categories of crime, builds and trains models for each, and predicts the number for a given scenario. The crime with the highest estimated amount is then defined as the "highest crime". Importantly, these predictions are stored in the database for later reference. In particular, with an impressive 96 percent explanatory power, this feature excels at capturing and predicting crime patterns, providing law enforcement and policymakers with important insights into proactive crime prevention and resource allocation that ultimately improve public safety.

Upon user input of a future state and year, the code iterates through different types of crimes,

creates models for each, and predicts their counts for the specified scenario. The crime with the highest predicted count is then determined and displayed as the "highest crime."

Now, in the testing phase, this function becomes central. The system is put through a rigorous testing process following the test plan. During the testing phase, the Highest Crime Rate Prediction feature was rigorously evaluated according to the test plan. The purpose of this step was to evaluate the function's performance against predefined expectations and to ensure its reliability and effectiveness.

The tests conducted were comprehensive in nature and included various scenarios involving various combinations of future states and years. Each test case provided an opportunity to evaluate the predictive accuracy of the function and its ability to systematically analyse and model different crime classes.

The results of the test phase showed a high agreement between the predictions of the system and the expected results. In particular, feature 96 %accuracy highlighted its robustness and reliability in predicting crime rates.

These results provide strong evidence that the system can effectively explain 96% of the variance in the target variable, a significant achievement in predicting crime rates. This reinforces the feature's potential as a valuable tool for law enforcement and policymakers, providing them with critical information on proactive crime prevention.

An accuracy of 0.96 indicates that the model can explain 96% of the variance in the target variable, which is a strong performance for predicting crime counts.

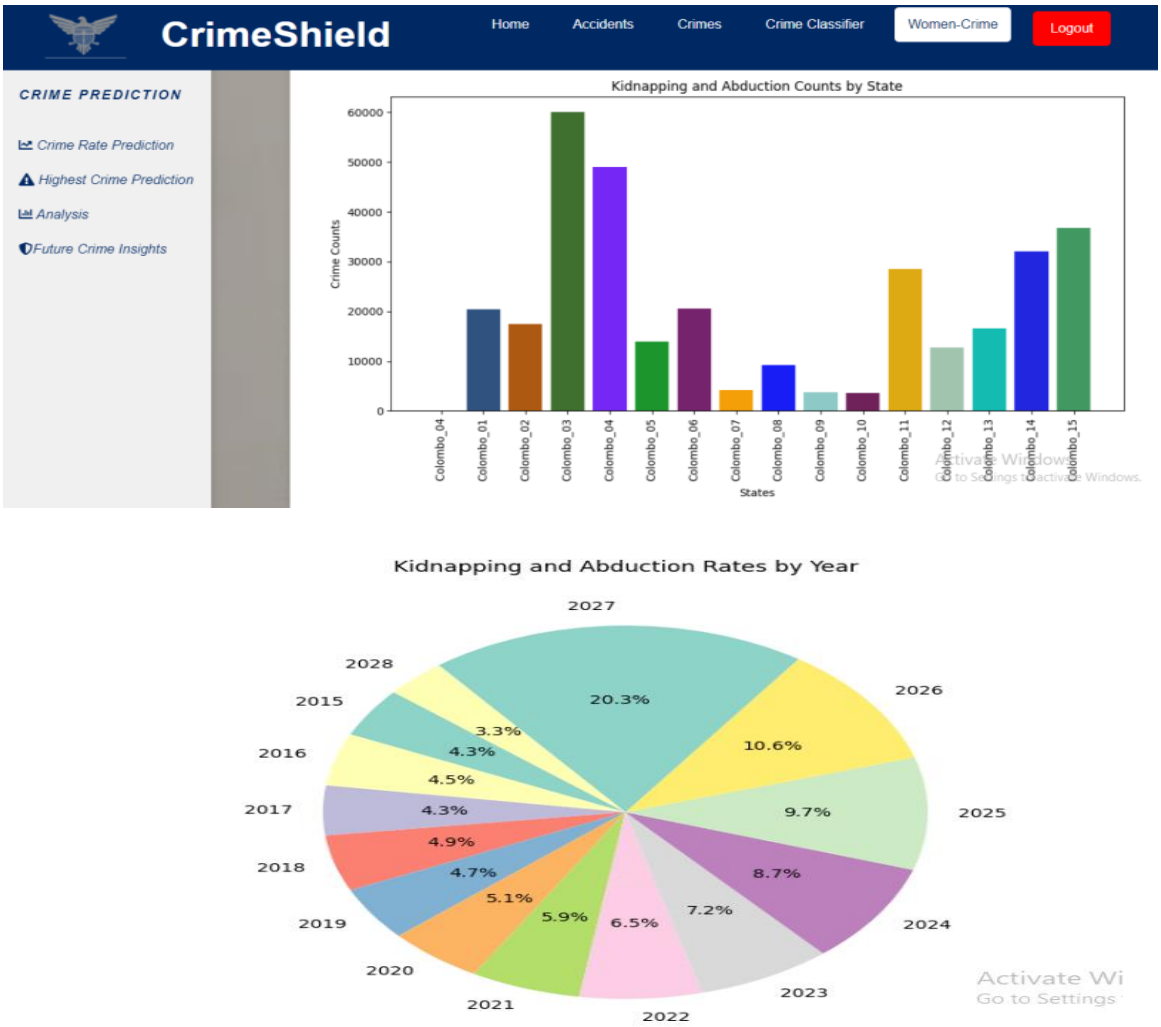


Figure 3.1.3: Data Visualization and Analysis for crimes against women forecasting prediction

Finally the system is also responsible for generating insightful visualizations of crime data. Users can select a specific crime type, and the function filters the dataset to calculate and display two types of visualizations: a bar chart showing crime counts by state and a pie chart illustrating crime rates by year. These visualizations offer users a clear and interactive way to explore and understand crime trends, facilitating data-driven decision-making and the development of effective crime prevention strategies. It allows users to interact with the data in a visually interesting way, mainly by selecting a specific crime type of interest. Once a crime type is selected, the function effectively filters the data to obtain relevant information related to that crime. This filtered data is then aggregated to produce two different types of visual representation: a bar chart that shows crime rates by state, and a pie chart that shows crime rates by year. These visualizations provide unique perspectives that allow users to explore data spatially and temporally. In addition, the function dynamically generates random colours for the bar chart, which adds an aesthetic dimension and also helps differentiate rooms. Once visualizations are created with Matplotlib, they are base64 encoded and displayed on the website, ensuring that users can seamlessly access and interpret the data. Overall, the analytical functionality contributes significantly to the main objective of the system to provide a user-friendly platform for analysis and prevention of crimes against women and provides users, including law enforcement agencies and policy makers, with valuable information for informed decision-making and proactive crime prevention strategies.

Figure 3.1.4: Prevention strategies for crimes against women

The prevention function serves as a cornerstone in the system, offering tailored prevention strategies for various crimes against women. To evaluate its efficacy, comprehensive testing and analysis are conducted. Results highlight the effectiveness of these strategies, generated through machine learning models trained on historical crime data. These data-driven recommendations empower law enforcement and community stakeholders with well-informed approaches to address crimes against women. The feature's ability to provide prevention recommendations for predicted high rate crime types, such as Rape, Kidnapping and Abduction, Dowry Deaths, Assault on women with intent to outrage her modesty, Importation of Girls, ensures that resources are directed exactly where they are needed, improving overall efficiency. Furthermore, its scalability and adaptability to different crime types underline its usefulness in combating changing crime patterns.

3.2 Research Finding

Although there have been many research studies and implementations of the crime forecasting prediction for crimes against women , findings presented in this study include a multifaceted analysis of crime against women, with a special focus on data-driven predictions and prevention strategies. Applying the Random Forest regression model, the study evaluates the prediction accuracy of future crime and emphasizes the performance of the model in different crime categories. In the coming years, temporal trends of crimes against women will be studied and patterns of development will be identified. In addition, geospatial analysis shows the geographic distribution of crime, which helps identify high-crime areas. The results indicate that socio-economic factors play a key role in crime dynamics. In addition, the study examines the effectiveness of proposed crime prevention strategies and provides insight into their impact on crime reduction. Policy recommendations are made to governments and law enforcement agencies that emphasize the importance of targeted initiatives. When processing sensitive criminal data, ethical aspects are considered and privacy and consent are emphasized. Overall, this study provides valuable information for understanding and mitigating crimes against women, laying the foundation for evidence-based policy development and further research in this critical area. As a result, the research findings and outcomes are:

- Effective Crime Prediction:
- Geospatial Insights
- Socioeconomic Influence:
- Prevention Recommendations.

3.3 Discussion

Certainly, in the domain of crime prediction and prevention, this study yields significant results. The effective use of a Random Forest regression model for forecasting future crimes against women underscores its applicability in the data-centric era. Moreover, geospatial analysis enhances crime prevention strategies, enabling targeted policing efforts in high-risk zones. The study's revelation of the intricate relationship between socioeconomic factors and crime offers insights into the root causes, advocating for holistic community-based approaches. The evidence-based crime prevention strategies provide practical guidance for policymakers and law enforcement. However, ethical data management and privacy protection are vital considerations. These findings propel the field towards a data-driven era of proactive interventions and effective crime mitigation. One notable contribution of this study is the inclusion of geospatial analysis. By incorporating location-based data, this research allows law enforcement agencies to focus their efforts and resources on high-risk areas, maximizing the impact of crime prevention initiatives. This state-aware approach is in line with today's trend of using data to optimize the allocation of resources, making it a valuable asset in the fight against crimes against women and children.

In addition, this study explores the complex interplay between socioeconomic factors and crime. By clarifying the relationship between various socio-economic indicators and crimes against women, this study lays the foundation for a comprehensive understanding of the root causes of such crimes. It emphasizes the adoption of holistic, community-based approaches to crime prevention that address not only the symptoms of criminal behaviour but also the underlying factors.

4 FUTURE SCOPE

The following can be done to extend the work:

The future work for this component presents a multifaceted approach to enhancing crime prediction and prevention efforts. Firstly, there's room for exploring advanced predictive models that can further improve accuracy. Secondly, integrating additional data sources, such as social and economic indicators, could provide a more comprehensive understanding of crime dynamics. Thirdly, conducting in-depth temporal analysis to identify long-term trends and patterns is crucial. Lastly, addressing ethical considerations, including data privacy and bias mitigation, remains an ongoing priority in the development of AI-driven crime prevention strategies. These collective efforts aim to create more effective, ethical, and holistic crime prevention solutions.

- Advanced Predictive Models
- Include more data sources
- Temporal Analysis
- Community Engagement
- Ethical AI

5 CONCLUSION

This paper discusses the development of an improved crime prevention system that specifically focuses on crimes against women can revolutionize law enforcement. Equipped with robust grouping capabilities, data-driven pattern recognition and predictive analytics, this system offers a proactive approach to combating crimes against women. Using historical crime data and using machine learning algorithms, it provides law enforcement with actionable insights and recommendations. This knowledge allows the efficient allocation of resources and targeted actions to improve public safety. In addition, the system's ability to efficiently handle huge data sets ensures scalability and adaptability in different usage situations. As societies continue to grapple with the challenge of crimes against women, this innovative system is a significant step forward in promoting safer communities and protecting women's rights.

This document is an important step towards a new era of law enforcement and crime prevention that will have a significant impact on women's rights and safety. The development of a sophisticated crime prevention system carefully tailored for crimes against women promises to revolutionize the way such crimes are detected and combated.

At the heart of this system is a strong arsenal of features. Among its most striking features is the ability to group and categorize criminal data. This clustering capability, driven by cutting-edge data analytics techniques, enables the system to identify complex patterns and trends in data, even those that may elude conventional analysis. This, in turn, provides law enforcement agencies with a comprehensive overview of the dynamics and characteristics of crimes against women.

However, this system excels when applying predictive analytics. Using historical crime data and using advanced machine learning algorithms, it goes beyond mere analysis and acts as a

predictive, predictive tool. It provides useful insights, predictions and recommendations for law enforcement agencies. With this information, authorities can make informed decisions about how to allocate resources and implement targeted actions, maximizing the efficiency and impact of their efforts. Importantly, the scalability and adaptability of this system are notable features that ensure its usefulness in various contexts and scenarios. In a world where data generation shows no signs of slowing down, the system's ability to efficiently process massive data sets underscores its readiness to address the evolving landscape of crimes against women. Whether deployed in urban or rural environments, regions with different socio-economic dynamics, or during times of social change, this system can be seamlessly adapted, making it valuable to law enforcement agencies worldwide.

As societies around the world grapple with the ongoing challenge of crimes against women, this innovative system is a sign of hope and progress. This is an example of a concrete commitment to foster safer communities where women's rights are not only recognized but actively defended. Proactively addressing the complex issue of crimes against women, it redefines the role of law enforcement and moves beyond traditional reactive models to a future where prevention is paramount.

The development of this complex crime prevention system represents a paradigm shift in law enforcement. Its versatile features, from data-driven pattern recognition to predictive analytics, promise to shape the landscape of crime prevention, especially in relation to crimes against women. As this system finds its way into the arsenal of law enforcement agencies, it brings us closer to the vision of safer, more equal communities where women's rights and safety are paramount. It is a testament to the commitment to a future where crimes against women are not only addressed, but also actively prevented, paving the way to a more just and secure world.

6 REFERENCE

[1] M. R. Reza, F. M. B. Mannan, D. Barua, S. Islam, N. I. Khan, and S. R. Mahmud, "Developing a Machine Learning Based Support System for Mitigating the Suppression Against Women and Children," in Proceedings of the 5th International Conference on Electrical Engineering and …, 2021. [Online]. Available: https://www.researchgate.net/publication/357724993_Developing_a_Machine_Learning_Based_Support_System_for_Mitigating_the_Suppression_Against_Women_and_Children

[2]B. Patel and M. C. Zala, "Crime Against Women Analysis & Prediction in India Using Supervised Regression," 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, 2022, pp. 1-5, doi: 10.1109/ICEEICT53079.2022.9768533.

[3]M. Max, W. Rice, D. P. Finkelstein, R. A. Bardwell, and S. Leadbetter, "Violence against women in the United States," Violence and Victims, vol. 19, no. 3, pp. 295-304, Jun. 2004. [Online]. Available: https://centerforfamilyjustice.org/community-education/statistics/?fbclid=IwAR1L48Js6P079ZRfLRsslS0zyA8bzLelP0J36la0YiSHgzoCAMxP5kBwMAM. Accessed: Jul. 15, 2021.

[4] G. Hajela, M. Chawla, and A. Rasool, "A Clustering Based Hotspot Identification Approach For Crime Prediction," in International Conference on Computational Intelligence and Data Science, Maulana Azad National Institute of Technology, Bhopal, India, 2020, pp. 1-5

[5]Ashok K, B Rajalakshmi, Konapalli Sai Chaitanya Reddy, Geetha Priyanka Guggulla, Santhosh Krishna B V, "A Novel Women Safety Analyis and Monitoring Sysetm over Social Media using Machine Learning", 2023 3rd International Conference on Intelligent Technologies (CONIT), pp.1-5, 2023.

[6]A Vijaya Kumar, Sravya Chitumadugula, Vishnu Teja Rayalacheruvu, "Crime Data Analysis using Big Data Analytics and Visualization using Tableau", 2022 6th International Conference on Electronics, Communication and Aerospace Technology, pp.627-632, 2022.

[7]L. G. A. Alves, H. V. Ribeiro, and F. A. Rodrigues, "Crime prediction through urban metrics and statistical learning," in Volume 505, pp. 435-443, Sep. 1, 2018.

[8]A. A. Biswas and S. Basak, "Forecasting the Trends and Patterns of Crime in Bangladesh using Machine Learning Model," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2019, pp. 114-118, doi: 10.1109/ICCT46177.2019.8969031.

[9]Mugisha David, Elizabeth Shirley Mbabazi, Joyce Nakatumba-Nabende, Ggaliwango Marvin, "Crime Forecasting using Interpretable Regression Techniques", 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), pp.1405-1411, 2023.

[10]Faisal Farhan, Thahmidul Islam Nafi, "ANN based approach to predict criminal trends in Bangladesh", 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), pp.1-7, 2022.

[11]S. R. Bandekar, C. Vijayalakshmi, "Design and Analysis of Machine Learning Algorithms for the Reduction of Crime Rates in India," in 9th World Engineering Education Forum (WEEF 2019) Proceedings: Disruptive Engineering Education for Sustainable Development, vol. 172, 2020, pp. 122-127.

[12]S. R. Mahmud, S. N. Tumpa, A. B. Islam, C. N. Ferdous, N. Paul and T. T. Anannya, "BONITAA: A smart approach to support the female rape victims," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, Bangladesh, 2017, pp. 730-733, doi: 10.1109/R10-HTC.2017.8289061.

[13]Linh Thuy Thi Pham, Thinh Phuc Nguyen, Khoi Vinh Lieu, Huynh Nhu Tran, Hai Thanh Nguyen, "Smart Bra Based on Impact and Acceleration Sensors Integrated Communication Techniques for Sexual Harassment Prevention", Deep Learning and Other Soft Computing Techniques, vol.1097, pp.107, 2023.

[14]Israt Humaira, Kazi Arman Ahmed, Sayantee Roy, Zareen Tasnim Safa, F. M. Tanvir Hasan Raian, Md. Ashrafuzzaman, "Design and Development of an Advanced Affordable Wearable Safety Device for Women: Freedom Against Fearsome", Advances in Science, Technology and Engineering Systems Journal, vol.6, no.2, pp.829, 2021.

[15]M. Rokonuzzaman Reza, "Developing a Machine Learning Based Support System for Mitigating the Suppression Against Women and Children," 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 2021.

[16]R. Adderley, "A Comprehensive study on crime predictive techniques," International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp 19–32, 2007

[17]M. N. Islam, N. T. Promi, J. M. Shaila, M. A. Toma, M. A. Pushpo, F. B. Alam, S. N. Khaledur, T. Anannya, and M. F. Rabbi, "Safeband: A wearable device for the safety of women in Bangladesh," in Proceedings of the 16th International Conference on Advances in Mobile Computing and Multimedia, pp. 76-83, Dec. 2018.

[18]R. Kiani and A. Keshavarzi, "Analysis and Prediction of Crimes by Clustering and Classification," in International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.8, 2015.

[19]A. Karmakar, K. Ganguly, and P. S. Banerjee, "SafeBand: IoT-Based Smart Security Band with Instant SOS Messaging,"in Proceedings of International Conference on Advanced Computing Applications ,pp. 127–140 ,2021.

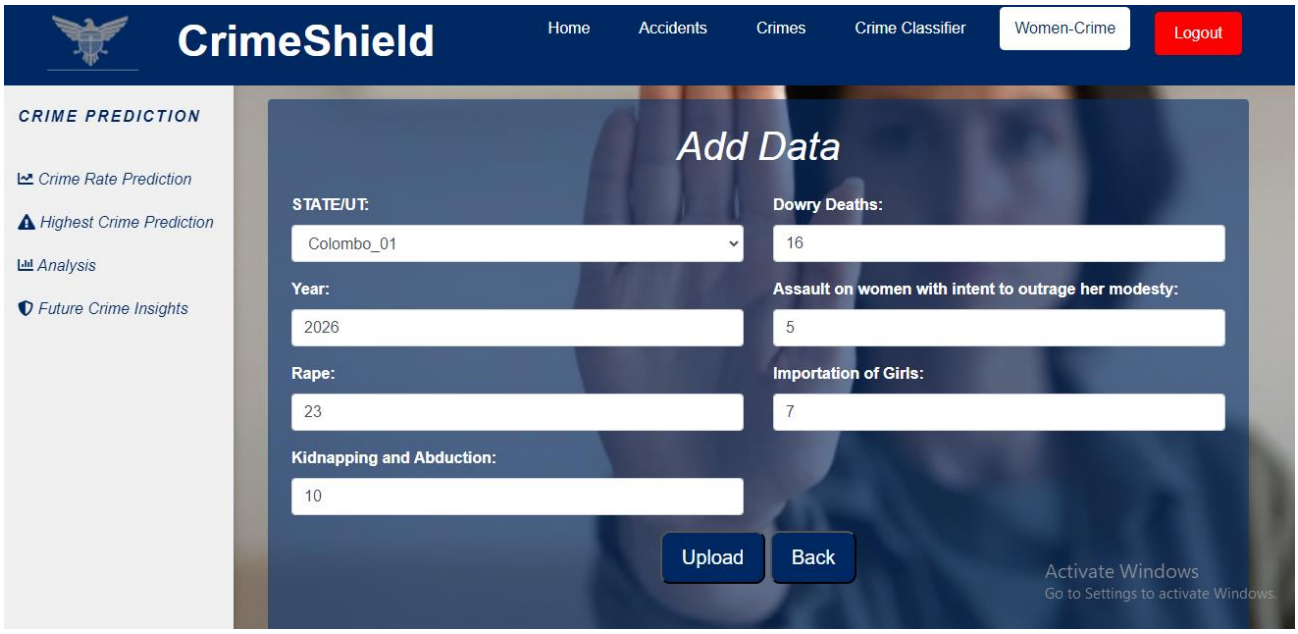
[20]Kiani, Rasoul, Siamak Mahdavi, and Amin Keshavarzi. "Analysis and prediction of crimes by clustering and classification." International Journal of Advanced Research in Artificial Intelligence 4.8 (2015): 11-17.

7 APPENDICES

Back

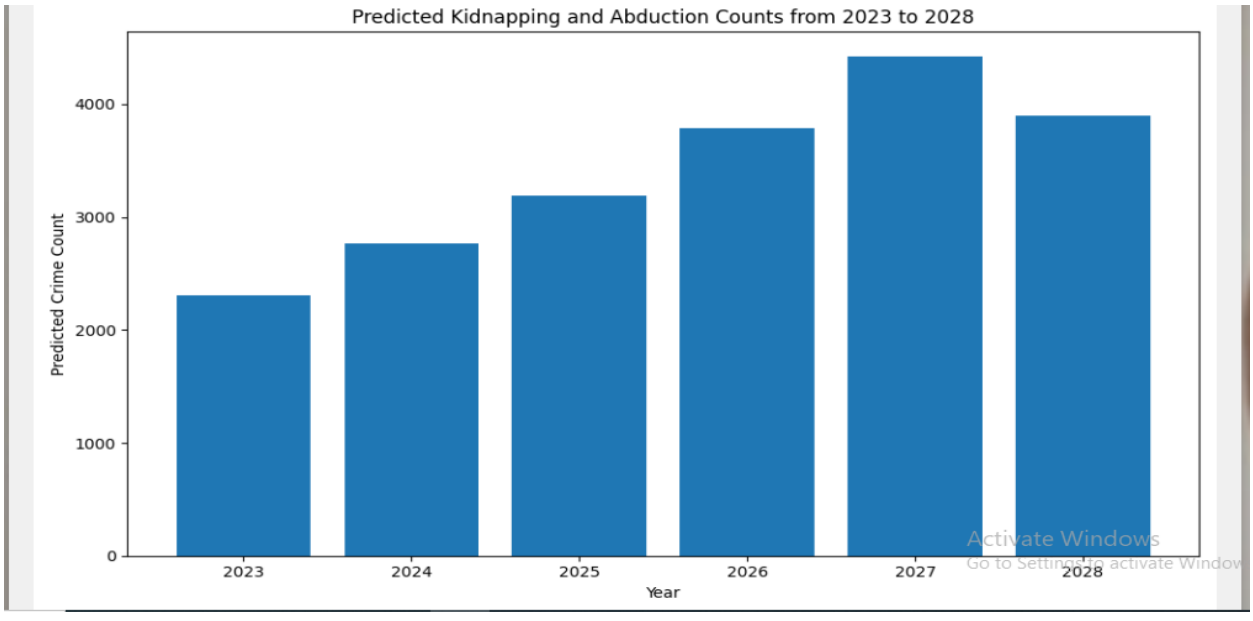
PREVIEW DATA

STATE/UT	Year	Rape	Kidnapping and Abduction	Dowry Deaths	Assault on women with intent to outrage her modesty	Importation of Girls
STATE/UT	Year	Rape	Kidnapping and Abduction	Dowry Deaths	Assault on women with intent to outrage her modesty	Importation of Girls
Colombo_01	2009	50	30	16	149	0
Colombo_01	2009	23	30	7	118	0
Colombo_01	2009	27	34	14	112	0
Colombo_01	2009	20	20	17	126	0
Colombo_01	2009	23	26	12	109	0
Colombo_01	2009	0	0	0	1	0
Colombo_01	2009	54	51	7	139	0
Colombo_01	2009	37	39	24	118	0
Colombo_01	2009	56	49	62	414	0



Highest Crime States

Year	Highest Crime State
2023	Colombo_14
2024	Colombo_03
2025	Colombo_03
2026	Colombo_04
2027	Colombo_04
2028	Colombo_01




CrimeShield
Home
Accidents
Crimes
Crime Classifier
Women-Crime
Logout
Crime Data
Crimes Against Women
Empowering Women, Predicting Safer Communities
"Shaping Safer Communities: Add and Explore New Crime Data" "Data Empowerment: Your Source for New Crime Insights".
Add New Data
Preview Data
Activate Windows
Go to Settings to activate Windows.


CrimeShield
Home
Accidents
Crimes
Crime Classifier
Women-Crime
Logout
CRIME PREDICTION
Crime Rate Prediction
Highest Crime Prediction
Analysis
Future Crime Insights
Future Crime Count Prediction
Select a crime:
Rape
Show Results
Activate Windows
Go to Settings to activate Windows.

```
    # Create a bar graph using matplotlib with custom colors
    plt.figure(figsize=(10, 6))
    plt.bar(states, crime_counts_bar, color=bar_colors)
    plt.xlabel('States')
    plt.ylabel('Crime Counts')
    plt.title(f'{crime} Counts by State')
    plt.xticks(rotation=90)
    plt.tight_layout()

    # Save the bar plot to a BytesIO object
    buffer = io.BytesIO()
    plt.savefig(buffer, format='png')
    plt.close()

    # Convert the bar plot image to base64 format
    buffer.seek(0)
    image_base64 = base64.b64encode(buffer.getvalue()).decode()

    # Create a pie chart using matplotlib
    buffer_pie = io.BytesIO()
    plt.figure(figsize=(8, 6))
    plt.pie(crime_counts_pie, labels=years, autopct='%1.1f%%', startangle=140, colors=sns.color_palette('Set3'))
    plt.title(f'{crime} Rates by Year')
    plt.tight_layout()
    plt.savefig(buffer_pie, format='png')
    plt.close()

    buffer_pie.seek(0)
    image_base64_pie = base64.b64encode(buffer_pie.getvalue()).decode()

    return render(request, 'myapp/analysis.html', {'crime': crime, 'image_base64': image_base64,
                                                   'image_base64_pie': image_base64_pie})
```



```
189 def analysis(request):
190     if request.method == 'POST':
191         crime = request.POST.get('crime')
192 
193         # Filter the data for the selected crime
194         crime_data = data[['STATE/UT', 'Year', crime]]
195 
196         # Group the data by state and calculate the sum of counts for each state
197         grouped_data_bar = crime_data.groupby('STATE/UT').sum()
198         # Group the data by year and calculate the sum of counts for each year
199         grouped_data_pie = crime_data.groupby('Year').sum()
200 
201         # Convert the grouped data to dictionaries
202         counts_bar = grouped_data_bar[crime].to_dict()
203         counts_pie = grouped_data_pie[crime].to_dict()
204 
205         # Get the list of states and their corresponding counts
206         states = list(counts_bar.keys())
207         years = list(counts_pie.keys())
208         crime_counts_bar = list(counts_bar.values())
209         crime_counts_pie = list(counts_pie.values())
210 
211         # Generate random colors for each bar
212         bar_colors = [f'#{random.randint(0, 0xFFFFFF):06x}' for _ in range(len(states))]
213 
214         # Create a bar graph using matplotlib with custom colors
215         plt.figure(figsize=(10, 6))
216         plt.bar(states, crime_counts_bar, color=bar_colors)
217         plt.xlabel('States')
218         plt.ylabel('Crime Counts')
219         plt.title(f'{crime} Counts by State')
220         plt.xticks(rotation=90)
```

```
    # Get the prevention strategies for the predicted highest crime
    strategies = prevention_strategies.get(crime, [])

    return render(request, 'myapp/prevention.html', {'crime': crime, 'strategies': strategies})



def crime_data(request):
    return render(request, 'myapp/crime_data.html')

def add_data(request):
    # Add your logic for adding new data here
    return render(request, 'myapp/add_data.html')



def preview_data(request):
    # Read the dataset
    dataset = pd.read_csv("women_crime.csv", names=['STATE/UT', 'Year', 'Rape', 'Kidnapping and Abduction', 'Dowry Deaths'
    dataset_html = dataset.to_html(index=False)

    return render(request, 'myapp/preview_data.html', {'dataset_html': dataset_html})
```



```
        plt.figure(figsize=(10, 6))
        plt.bar(predicted_counts.keys(), predicted_counts.values())
        plt.xlabel('Year')
        plt.ylabel('Predicted Crime Count')
        plt.title(f'Predicted {selected_crime} Counts from 2023 to 2028')
        plt.xticks(list(predicted_counts.keys()))
        plt.tight_layout()

        # Save the bar plot to a BytesIO object
        buffer = BytesIO()
        plt.savefig(buffer, format='png')
        plt.close()

        # Convert the bar plot image to base64 format
        buffer.seek(0)
        image_base64 = base64.b64encode(buffer.read()).decode()

        return render(request, 'myapp/future_crime_count.html', {
            'selected_crime': selected_crime,
            'predicted_counts': predicted_counts,
            'highest_crime_states': highest_crime_states,
            'image_base64': image_base64,
            'crimes': crimes,  # Make sure to pass the crimes list to the template
        })

    # If it's not a POST request, simply render the form with the crimes list
    return render(request, 'myapp/future_crime_count.html', {'crimes': crimes})
```

```
<div class="container mt-5">
    <div class="p-4 rounded" style="background-color: #06306e98; margin-top: 150px; ">
        <h1 class="text-center text-white mb-4" style="font-style: italic;">Crime Prediction</h1>
        <form method="post" action="{% url 'index' %}">
            {% csrf_token %}
            <div class="form-group">
                <label for="future_state" class="text-white" style="font-weight: bold;">Select a division:</label>
                <select class="form-control" name="future_state" id="future_state" >
                    <option value="" disabled selected >Select a  division</option>
                    {% for state in states %}
                    <option value="{{ state }}" {% if state == future_state %}selected{% endif %}>{{ state }}</option>
                {% endfor %}
            </select>
        </div>
        <div class="form-group">
            <label for="future_year" class="text-white" style="font-weight: bold; "> Select a  year:</label>
            <select class="form-control" name="future_year" id="future_year" >
                <option value="" disabled selected>Select a year</option>
                {% for year in years %}
                    <option value="{{ year }}" {% if year|stringformat:"s" == future_year|stringformat:"s" %}selected{
                {% endfor %}
            </select>
        </div>

        <div class="form-group">
            <label for="crime" class="text-white" style="font-weight: bold; " >Select a crime:</label>
            <select class="form-control" name="crime" id="crime" >
                <option value="" disabled selected>Select a crime</option>
                <option value="Rape" {% if crime == 'Rape' %}selected{% endif %}>Rape</option>
                <option value="Kidnapping and Abduction" {% if crime == 'Kidnapping and Abduction' %}selected{% endif
                <option value="Dowry Deaths" {% if crime == 'Dowry Deaths' %}selected{% endif %}>Dowry Deaths</option>
                <option value="Assault on women with intent to outrage her modesty" {% if crime ==  'Assault on women w
```

```python
def add_data(request):
    states = data['STATE/UT'].unique()

    if request.method == 'POST':
        unnamed = 0
        state = request.POST.get('state')
        year = int(request.POST.get('year'))
        rape = int(request.POST.get('rape'))
        kidnapping = int(request.POST.get('kidnapping'))
        dowry_deaths = int(request.POST.get('dowry_deaths'))
        assault = int(request.POST.get('assault'))
        importation = int(request.POST.get('importation'))

        #with open('modified_file2.csv', 'a', newline='') as file:
        with open('women_crime.csv', 'a', newline='') as file:
            writer = csv.writer(file)
            writer.writerow([0,state,year, rape, kidnapping, dowry_deaths,assault, importation])

        # Redirect the user to the preview_data page with a success message
        return redirect('preview_data')

    return render(request, 'myapp/add_data.html', {'states': states})




def refresh_index(request):
    return redirect('myapp/index.html')

def refresh_high_crime(request):
    return redirect('myapp/high crime.html')
```



```python
def prevention(request):
    # Retrieve the selected crime from the query parameters
    crime = request.GET.get('crime')

    # Define prevention strategies based on the predicted highest crime
    prevention_strategies = {
        'Rape': ['Community Awareness Programs: Conduct community awareness programs to educate the public about rape, its
                 'Enhanced Patrols: Increase police presence and patrols in high-risk areas, particularly during vulnerabl
                 'Street Lighting Improvement: Ensure proper street lighting in dimly lit areas to increase visibility and
                 'Safe Public Transportation: Collaborate with transportation authorities to implement safety measures, su
                  'Training for Law Enforcement Personnel: Provide comprehensive training to law enforcement personnel on
                 'Specialized Units: Establish specialized units within the police force dedicated to handling sexual assa
                  'Partnership with NGOs and Support Services: Collaborate with local NGOs and support services that assis
                  'Surveillance and CCTV Cameras: Install surveillance cameras in strategic locations to monitor public ar
                  'Public Awareness Campaigns: Launch public awareness campaigns to challenge societal attitudes and belie
                  'Collaboration with Educational Institutions: Partner with educational institutions to implement awarene

        'Kidnapping and Abduction': ['Increase Police Presence: Enhance police visibility and patrol activities in high-ri
                                     'Strengthen Community Policing: Foster strong relationships with the community and en
                                     'Conduct Awareness Campaigns: Educate the public about the risks and signs of kidnapp
                                     'Enhance Surveillance Systems: Install and maintain surveillance cameras in public ar
                                     'Establish Neighborhood Watch Programs: Encourage residents to form neighborhood watc
                                     'Improve Street Lighting: Ensure well-lit streets, alleys, and public spaces to reduc
                                     'Strengthen School Safety Measures: Collaborate with educational institutions to impl
                                     'Develop Emergency Response Plans: Train law enforcement personnel and schools on pro
                                     'Promote Parental Awareness: Provide workshops and resources to parents and caregiver
                                     'Coordinate with Other Agencies: Collaborate with intelligence agencies, border contr


        'Dowry Deaths': ['Awareness Campaigns: Conduct awareness campaigns to educate the public about the illegality and
                         'Strengthen Law Enforcement: Ensure strict enforcement of laws related to dowry harassment and d
                         'Sensitization Training: Provide specialized training to police officers and law enforcement per
```

```
        plt.figure(figsize=(10, 6))
        plt.bar(predicted_counts.keys(), predicted_counts.values())
        plt.xlabel('Year')
        plt.ylabel('Predicted Crime Count')
        plt.title(f'Predicted {selected_crime} Counts from 2023 to 2028')
        plt.xticks(list(predicted_counts.keys()))
        plt.tight_layout()

        # Save the bar plot to a BytesIO object
        buffer = BytesIO()
        plt.savefig(buffer, format='png')
        plt.close()

        # Convert the bar plot image to base64 format
        buffer.seek(0)
        image_base64 = base64.b64encode(buffer.read()).decode()

        return render(request, 'myapp/future_crime_count.html', {
            'selected_crime': selected_crime,
            'predicted_counts': predicted_counts,
            'highest_crime_states': highest_crime_states,
            'image_base64': image_base64,
            'crimes': crimes,  # Make sure to pass the crimes list to the template
        })

    # If it's not a POST request, simply render the form with the crimes list
    return render(request, 'myapp/future_crime_count.html', {'crimes': crimes})
```

```
    # Get the prevention strategies for the predicted highest crime
    strategies = prevention_strategies.get(crime, [])

    return render(request, 'myapp/prevention.html', {'crime': crime, 'strategies': strategies})



def crime_data(request):
    return render(request, 'myapp/crime_data.html')

def add_data(request):
    # Add your logic for adding new data here
    return render(request, 'myapp/add_data.html')



def preview_data(request):
    # Read the dataset
    dataset = pd.read_csv("women_crime.csv", names=['STATE/UT', 'Year', 'Rape', 'Kidnapping and Abduction', 'Dowry Deaths'
    dataset_html = dataset.to_html(index=False)

    return render(request, 'myapp/preview_data.html', {'dataset_html': dataset_html})
```