

Visvesvaraya Technological University, Belagavi-590018



A PROJECT REPORT ON

“CRIME RATE PREDICTION”

Submitted in partial fulfillment of the requirements for the award of

**BACHELOR OF ENGINEERING IN
COMPUTER SCIENCE AND ENGINEERING**

For the Academic year

2022-2023

Submitted by:

**SHETTY DISHA ASHOK
SOWJANYA**

**4SH19CS062
4SH19CS068**

Under the Guidance of:

Mrs. Nishmitha M R

Assistant Professor
Department of CSE



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SHREE DEVI INSTITUTE OF TECHNOLOGY

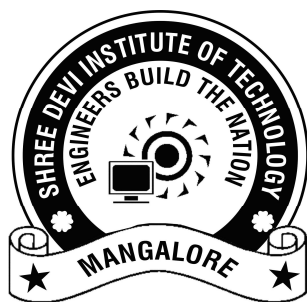
KENJAR, MANGALURU-574142

SHREE DEVI INSTITUTE OF TECHNOLOGY

(An Institution under VTU, Belagavi)

MANGALURU-574142

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

Certified that the project work entitled “**CRIME RATE PREDICTION**” is a bonafide work carried out by **SHETTY DISHA ASHOK, SOWJANYA** bearing USN’s **4SH19CS062, 4SH19CS068**, respectively in partial fulfillment of eight semester project, regards to the subject “**Major Project**” for the award of degree of **Bachelor of Engineering in Computer Science and Engineering** of the Visvesvaraya Technological University, Belagavi during year 2022-2023. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the degree of Bachelor of Engineering.

Signature of the Guide
Mrs. Nishmitha M R

Signature of the HOD
Prof. Anand S. Uppar

Signature of Principal
Dr. K E Prakash

EXTERNAL VIVA

Name of the Examiners

1.-----

2.-----

Signature with date

**SHREE DEVI INSTITUTE OF TECHNOLOGY KENJAR,
MANGALURU– 574142**

Department of Computer Science and Engineering



DECLARATION

We **Shetty Disha Ashok, Sowjanya** bearing USN **4SH19CS062, 4SH19CS068** students of Eighth semester Bachelor of Engineering, Computer Science and Engineering, Shree Devi Institute of Technology, Mangalore declare that the project entitled “**CRIME RATE PREDICTION**” has been duly executed by us under the guidance of **Mrs. Nishmitha M R**, Asst. Professor, Department of Computer Science and Engineering, Shree Devi Institute of Technology, Mangalore and submitted for the requirement for the eighth semester project of **Bachelor of Engineering in Computer Science and Engineering** during the year 2022-2023.

Date:
Place: Mangalore

SHETTY DISHA ASHOK [4SH19CS062]
SOWJANYA [4SH19CS068]

ACKNOWLEDGEMENT

A successful project is a fruitful culmination of the efforts of many people. Some directly involved and others who have quietly encouraged and extended their invaluable support throughout its progress.

We would like to convey my heartfelt thanks to **our Management** for providing the good infrastructure, laboratory facility, qualified and inspiring staff whose guidance was of great help in successful completion of this project.

We are extremely grateful and thankful to our beloved director and principal **Dr. K.E Prakash**, Shree Devi Institute of Technology, Kenjar for providing the congenial atmosphere and necessary facilities for achieving the cherished goal.

With heartiest gratitude, we would like to thank **Prof. Anand S Uppar**, HOD, Department of Computer Science and Engineering for his support, guidance and encouragement.

We are profoundly indebted to our Guide, **Asst. Prof. Nishmitha M R**, Department of Computer Science and Engineering, for their guidance throughout the seminar by innumerable acts of timely advice and encouragement.

We also thank all other teaching staff and non-teaching staff for allowing us to carry out the project work.

Finally we would like to thank our family for their support and understanding, to whom we owe so much.

SHETTY DISHA ASHOK [4SH19CS062]

SOWJANYA [4SH19CS068]

ABSTRACT

Crime is one of the biggest and dominating problem in our society. Daily there are huge number of crimes committed frequently. Here the dataset consists of the year and the crime rate that has taken place in the corresponding years for the particular crime. In this project the crime rate is based on the murder, kidnapping, crime against women, cyber-crime. We use linear regression algorithm to predict the rate of the crime in the future by using the previous data information. The year, crime rate of particular crimes, the district is given as an input to the algorithm for predicting the present crime rate and for the future prediction we select district and the year as the input.

CONTENTS

1. INTRODUCTION	1-2
1.1 Purpose	2
1.2 Scope	2
1.3 Overview	2
2. LITERATURE REVIEW	3-4
3. PROBLEM DEFINITION	5
4. OBJECTIVES	6
5. SOFTWARE REQUIREMENTS SPECIFICATION	7-9
5.1 Introduction	7
5.2 Purpose	7
5.3 User characteristics	7
5.4 Interfaces	8
5.4.1 Hardware interfaces	8
5.4.2 Software interfaces	8
5.5 Functional requirements	8
5.6 Non-functional requirements	9
6. SYSTEM DESIGN	10
6.1 Architecture design	10

7. IMPLEMENTATION	11
8. SYSTEM TESTING	12-15
8.1 Introduction	12
8.2 Testing objectives	13
8.3 Test approach	13
8.3.1 Black box testing	13
8.3.2 White box testing	13
8.4 Testing strategies	14
8.4.1 Unit testing	14
8.4.2 Integration testing	14
8.4.3 System testing	14
8.4.4 Validation testing	15
9. SNAPSHOTS	16
10. CONCLUSION AND FUTUREWORK	17
REFERENCES	18

LIST OF FIGURES

FIGURE NUMBER	TITLE	PAGE NUMBER
6.1	System Architecture Diagram	

CHAPTER 1

INTRODUCTION

Crimes are the significant threat to the humankind. There are many crimes that happen in regular intervals of time. Perhaps it is increasing and spreading at a fast and vast rate. Crimes happen from small village, town to big-cities. Crimes are different types- murder, kidnapping, crime against women cyber-crimes. Since crimes are increasing there is a need to solve, the cases in a much faster way. The crime activities have been increased at a faster rate and it is the responsibility of police department to control and reduce the crime. Crime prediction and criminal identification are the major problems to the police department as there are tremendous amount of crime data that exist. There is a need of technology through which the case solving could be faster. Through many documentations and cases, it came out that machine learning and data science can make the work easier and faster. The aim of this project is to make crime prediction using the features present in the dataset. The dataset is extracted from the official sites. With the help of machine learning algorithm, using python as core we can predict the crime rate which will occur in a particular area. The objective would be to train a model for prediction. The training would be done using Training data set which will be validated using the test dataset. The Linear Regression (LR) will be used for crime prediction. This work helps the law enforcement agencies to predict and detect the crime_percapita in an area and thus reduces the crime rate.

1.1 Purpose

The main purpose of this project is to predict crimes before they occur, and thus the importance of using crime forecasting methods is extremely clear. Furthermore, the prediction of crimes can sometimes be crucial because it may potentially save the life of a victim, prevent lifelong trauma, and avoid damage to private property

1.2 Scope

The main objectives of crime rate prediction include: 1. Extraction of crime patterns by analysis of available crime and criminal data. 2. Prediction of crime based on spatial distribution of existing data and anticipation of crime rate using different data mining techniques. 3. Predict rate of crime

1.3 Overview

Since, the numbers of crimes are growing at higher rate day by day, it is very essential to come up with a new model which can predict to handle large amount of crime data without reducing the performance of the model. The model implements advanced technologies of machine learning . This model will allow the user to input the district, year and rate of particular crime to predict the overall crime rate.

CHAPTER 2

LITERATURE REVIEW

[1] Shiju Sathyadevan, Devan M.S, proposed that Day by day the crime rate is increasing considerably. Crime cannot be predicted since it is neither systematic nor random. Also the modern technologies and hi-tech methods help criminals in achieving their misdeeds. According to Crime Records Bureau crimes like burglary, arson etc have been decreased while crimes like murder, sex abuse, gang rape etc have been increased. Even though we cannot predict who all may be the victims of crime but can predict the place that has probability for its occurrence. The predicted results cannot be assured of 100% accuracy but the results shows that our application helps in reducing crime rate to a certain extent by providing security in crime sensitive areas. So for building such a powerful crime analytics tool we have to collect crime records and evaluate it .It is only within the last few decades that the technology made spatial data mining a practical solution for wide audiences of Law enforcement officials which is affordable and available. Since the availability of criminal data or records is limited we are collecting crime data from various sources like web sites, news sites, blogs, social media, RSS feeds etc. This huge data is used as a record for creating a crime record database. So the main challenge in front of us is developing a better, efficient crime pattern detection tool to identify crime patterns effectively.

[2] Riya Rahul Shah proposed that Vancouver is most populated city in Canada. It is most ethnically diverse cities in Canada. Crime is one of the biggest and dominating problem in our society and its prevention is an important task. Even though Vancouver known to be the safest city it is observed that vehicle breaking and many more thefts is still a problem. There has been tremendous increase in machine learning algorithms that have made crime prediction feasible based on past data. The aim of this project is to perform analysis and prediction of crimes in states using www.jespublication.com Page No:695 Vol 11, Issue 5,May/2020 ISSN NO:0377-9254 machine learning models. It focuses on creating a model that can help to detect the number of crimes by its type in a particular state. In this project various machine learning models like K-NN, boosted decision trees will be used to predict crimes. Area Wise geographical analysis can be done to understand the pattern of crimes. Various visualization techniques and plots are used which can help law enforcement agencies to detect and predict crimes with higher accuracy. This will indirectly help reduce the rates of crimes and can help to improve securities in such required areas.

[3] Nishat Shama proposed that Criminal activities are present in every region of the world affecting quality of life and socio-economical development. As such, it is a major concern of many governments who are using different advanced technology to tackle such issues. Crime Analysis, a sub branch of criminology, studies the behavioral pattern of criminal activities and tries to identify the indicators of such events. Machine learning agents work with data and employ different techniques to find patterns in data making it very useful for predictive analysis. Law enforcement agencies use different patrolling strategies based on the information they get to keep an area secure. A machine learning agent can learn and analyze the pattern of occurrence of a crime based on the reports of previous criminal activities and can find hotspots based on time, type or any other factor. This technique is known as classification and it allows to predict nominal class labels. Classification has been used on many different domains such as financial market, business intelligence, healthcare, weather forecasting etc. In this research, a dataset from San-Francisco Open Data is used which contains the reported criminal activities in the neighbour hoods of the city San Francisco for a duration of 12 years. I used different classification techniques like Decision Tree, Naive Bayesian, Logistic Regression, k-Nearest Neighbour, Ensemble Methods to find hotspots of criminal activities based on the time of day. Results of different algorithms have been compared and most the effective approach has also been documented.

CHAPTER 3

PROBLEM DEFINITION

To develop a machine learning model that can accurately predict the crime rate in a given geographical area based on historical crime data, socio-economic indicators, and other relevant factors. The model should be able to provide insights into the factors that contribute to crime and help law enforcement agencies and policymakers make data-driven decisions to reduce crime and improve public safety.

CHAPTER 4

OBJECTIVES

The main objectives of this project are as follows:

1. The main objective of the project is to predict the crime rate and analyze the crime rate to be happened in future. Based on this information the official can take charge and try to reduce the crime rate.
2. The system will look at how to convert crime information into a regression problem, so that it will help detectives in solving crimes.
3. Crime prediction is based on available information to extract crime patterns. Using regression techniques, frequency of occurring crime can be predicted based on territorial distribution of existing data.

CHAPTER 5

SOFTWARE REQUIREMENTS SPECIFICATION

5.1 Introduction

Software Requirement Specification totally defines how the projected software behaves without unfolding how the software will perform it. The elementary objective of the requirement stage is to yield the software requirement specification that designates the peripheral performance of the projected software. Software requirement can be well- defined as a condition of a capability required by a user to solve a problem or attain an objective.

5.2 Purpose

The main purpose of this project is to present a new model which predict the rate of crime that can happen in future with the help of existing data. The existing model does not provide the solution to the user. But, our model predict crime rate. The project is compatible the district, year and the rate of particular crime as the input for predicting the present and district and year as input for the future prediction. If the suggested solution is not satisfied, then it redirects to expert solutions.

5.3 User Characteristics

A web page is provided to the user with an input field where user must input a district, year and rate of particular crime as the input for predicting the present and district and year as input for the future prediction. The web page provides the facility to predict crime rate for future as well as past. Based on the previous year data, the future rate will be detected.

5.4 Interfaces

5.4.1 Hardware Interfaces

Processor: Any processor above 500 MHz

RAM: 8GB

Hard Disk: 1TB

Processor: i5

Output Device: Monitor

5.4.2 Software Interfaces

Operating system: Windows or Linux

Browser: Google Chrome or Firefox

Text Editor: Gedit, Notepad

Programming Language: Python

5.5 Functional Requirements

Functional Requirements defines a function of a software system and how the system must behave when presented with specific inputs and conditions. These may include calculations, data manipulation and processing and other specific functionality. Following are the functional requirements on the system:

- Collecting Data sets and data pre-processing is performed for the data set
- The data set will be subjected to various data mining techniques, Linear Regression will be performed on the given data set

5.6 Non-Functional Requirements

The Webpage must be efficient with very little lag in response time to users reply. The model must be reliable to predict rate with no fault. The Linear Regression is used for predicting the crime rate. Appropriate handling of unexpected input and correctly inform the user if it cannot provide solution.

CHAPTER 6

SYSTEM DESIGN

6.1 Architecture Design

The dataset to the model is prepared manually. The dataset has six attributes that are district, year, murder, kidnapping, cyber-crime, crime against women, rate is fed to the machine learning model. Before feeding the input to the model, all the sentences are pre-processed.

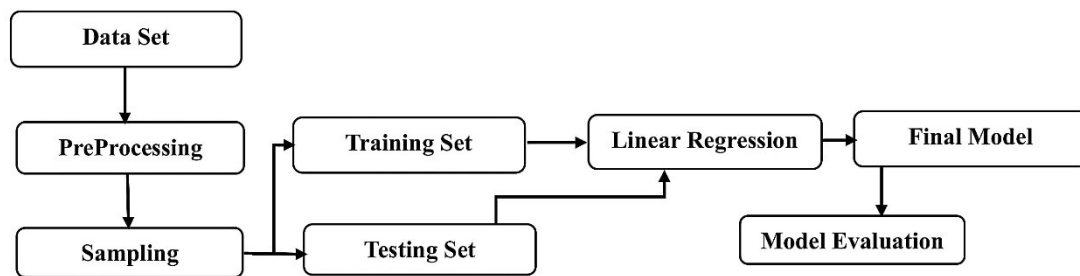


Figure 6.1: System Architecture Diagram

The data is split into training and testing data. Linear Regression is used to train the data. A sequential model is created to predict the overall crime rate of the particular district with the help of year and the rate of . A user friendly web page is made to get the status from the user. When the user gives the input, the data is fed to the machine learning model and it will predict the rate of crime that may happen in future. The result will be shown in a new web page.

CHAPTER 7

IMPLEMENTATION

Implementation plays the most important part in any project report. Implementation is used to execute, or practice of a plan, a method, design, model, specification, idea standard or policy for doing something. The implementation phase is when the end user interacts with the product.

Our model mainly works on the input given by the user such as district, year and rate of particular crime. It predicts the output as overall rate for the crime of past and future by analyzing previous years data and present data. The first phase in implementation consists of collection and pre-processing of data. The collection of data is done from manually which consists of various district, the years and the rate of particular crime to predict the rate of the particular crime for the past and future.

CHAPTER 8

SYSTEM TESTING

8.1 Introduction

Testing is a procedure of executing the program with unequivocal intension of discovering mistakes, assuming any, which makes the program, fall flat. This stage is an essential piece of the product improvement.

It plays out an exceptionally basic part for quality affirmation and for guaranteeing unwavering quality of programming. It is the way toward finding the mistakes and missing operation and furthermore an entire confirmation to decide if the targets are met the client prerequisites are fulfilled.

The objective of testing is to reveal prerequisites, outline or coding blunders in the projects. Therefore, unique levels of testing are utilized in programming frameworks. The testing results are utilized amid upkeep.

This area manages the points of interest in the various classes of the test which should be directed to approve capacities, imperatives and execution. This can be accomplished fundamentally by using the methods for testing, which assumes a crucial part in the improvement of a product.

The structure of the program is not being considered in useful testing. Test cases are exclusively chosen on the premise of the prerequisites or particulars of a program or module of program but the internals of the module or the program are not considered for determination of experiments.

The program to be tried is executed with an arrangement of experiments and the yield of the program for the experiments is assessed to decide whether the program is executing not surprisingly. The accomplishment of testing in uncovering mistakes in projects depends basically on the experiments. There are two fundamental ways to deal with testing:

8.2 Testing Objectives

1. To verify the interaction between objects.
2. To verify proper integration of all the components of the software.
3. To verify all the requirements have been correctly implemented.
4. To identify and ensure defects are detected before the implementation of software.

8.3 Test Approach

8.3.1 Black Box Testing

Black box testing or functional testing allows testing the errors that are conducted at the software interface. It is approaches where the tests are derived from the program specification. In black box testing the examination of fundamental aspects of a system is with regard to the internal logical structure of the system. The black box system, whose behavior can be found out by studying its inputs with a related output, is equally applicable to systems that are organized as functions.

8.3.2 White Box Testing

White box testing strategy works with structure and internal logic of the code. White box testing is also recognized as glass, structural, open box or clear box testing. It mainly focuses on validating the flow of inputs and outputs through the application, improving design and usability, strengthening security.

8.4 Testing Strategies

8.4.1 Unit testing

Unit testing can be used to test a unit as a whole. This would test the interaction of many functions but combine the test within one unit. The scope of a unit is left to interpretation. Supporting test code, sometimes called scaffolding, may be necessary to support an individual test. This type of testing is driven by architecture and implementation teams.

8.4.2 Integration Testing

One of the important phases of the software development project is the integration strategy. Integration can be done all at once, top-down, bottom-up critical piece first, or else by first integrating functioning sub-systems and then integrating the subsystems in separate phases using any one of the basic strategies. Usually, the larger the project the more important is the integration strategy.

8.4.3 System Testing

System testing is performed on complete system, or on the integrated system to evaluate whether the system meets the specific requirement. System testing falls under the scope of black, so it should not require knowledge of the inner design of the code or logic.

8.4.4 Security Testing

The process of determining that an Information System can be used to protect the data and maintains functionality as intended. The six basic concepts in security testing are: confidentiality, integrity, authentication, authorization, availability, and non-repudiation.

8.4.5 Validation Testing

This testing method focuses on the need for consistent and thorough quality assurance processes and its standards or its guidelines. It is Implemented through test planning and management, and it can also be independently executed software testing process becomes a important element in ensuring quality in mission-critical projects and the long-term success of any organization.

CHAPTER 9

SNAPSHOTS

Predicting Crime Rate:

1. Past Prediction



The screenshot shows a web browser window titled "Crime Prediction". The main heading is "CRIME RATE PREDICTION". Below the heading, there are several input fields and a button:

- Select District:** A dropdown menu with "Shimoga" selected.
- Select Year:** A dropdown menu with "2020" selected.
- Enter Murder Count:** A text input field with "57" entered.
- Enter Crimes Against Women Count:** A text input field with "57" entered.
- Enter Cyber Crime Count:** A text input field with "55" entered.
- Enter Kidnapping Count:** A text input field with "51" entered.
- Predict** button.
- Predicted crime rate: 339.00** (displayed below the button).
- Future Prediction** button (located at the bottom).

2. Future Prediction



The screenshot shows a web browser window titled "Future crime rate prediction". The main heading is "FUTURE CRIME RATE PREDICTION". Below the heading, there are several input fields and a button:

- Select District:** A dropdown menu with "Banglore" selected.
- Select Year:** A dropdown menu with "2023" selected.
- Predict** button.
- Future crime rate: 463.44** (displayed below the button).

CHAPTER 10

CONCLUSION AND FUTURE WORK

With the help of machine learning technology, it has become easy to find out relation and patterns among the various data's. The work in this project mainly revolves around predicting the types of crime and the crime rate which may happen in future. Using the concept of machine learning we have built a model using training data set that have undergo data transformation using Linear Regression Algorithm. This system presents a crime data prediction by taking the types of crimes and year as input and giving rate in which these crimes are committed as output using Jupyter notebook having python as a core language and python provides inbuilt libraries such as Pandas and Numpy through which the work will be completed faster and scikit provides all the process of how to use different libraries providing by the python. From the encouraging results, we believe that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis. Visual and intuitive criminal and intelligence investigation techniques can be developed for crime pattern. As we applied regression technique of data mining for crime prediction also we can perform other techniques of data mining such as classification.

We would analyze the current report and identify any areas where we can think more data could be collected to create additional graphs. Explore the possibility of using predictive models, such as regression analysis, to create graphs that can forecast future crime rates based on these variables. Consider inco-operating data from new sources, such as social media or crime maps, to enhance the accuracy of our predictive models. Consider collaborating with experts in the field of criminology or data analysis to identify new trends or insights that can be represented in graphs. This can help to refine our predictive models and improve the accuracy of our predictions. Finally, provide a road map for future updates to the report, including plans to incorporate new graphs as new data becomes available. This will demonstrate a commitment to ongoing analysis and improvement and help to keep report relevant over time. Additionally, it is important to continuously monitor the accuracy of the predictive models and update them as necessary to ensure that they remain relevant and reliable.

REFERENCES

- [1] Alkesh Bharathi, Dr Sarvanaguru RA. K,” Crime Prediction and Analysis Using Machine Learning” in International Research Journal of Engineering and Technology(IRJET), Volume: 05 Issue: 09 | September 2018
- [2] Python : The Complete Reference – Martin C. Brown
- [3] SharvilKatariya, Saurabh Jain – Stock Price Trend Forecasting using Supervised learning Methods.
- [4] Sathyadevan, S.(2014, August). Crime analysis and prediction using data mining. In 2014 First International Conference on Networks & Soft Computing(ICNSC2014) (pp.406-412). IEEE
- [5] Jangra M, Kalsi S. Naïve Bayes approach for the crime prediction in Data Mining. Int J Comput Appl. 2019;178(4):33–7.
- [6] R.Iqbal, M.A.A,Murad, A.Mustapha, P.H.S.Panahy, and N.Khanahmadliravi, “An experimental study of classification algorithms for crime prediction”, Indian Journal of Science and Technology, vol.6, no. 3, pp. 4219-4225, 2013
- [7] Safat W, Asghar S, Gillani SA. Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques. IEEE Access J. 2021;9:70080–94