##### A Project report on

**CRIME ANALYSIS AND PREDICTION**

###### A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

Submitted by

M.GURUSAI CHAWAN

(20H51A0517)

T.MANOHAR

(20H51A05D3)

M.MEGHANA

(20H51A05P5)

Under the esteemed guidance of

Mr. M. Shiva Kumar

(Assistant Professor)



**Department of Computer Science and Engineering**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(UGC Autonomous)

\*Approved by AICTE \*Affiliated to JNTUH \*NAAC Accredited with A+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

#### 2020- 2024

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



#### CERTIFICATE

This is to certify that the Major Project Phase I report entitled **"Crime Analysis and Prediction"** being submitted by M. Guru Sai Chawan (20H51A0517), T. Manohar (20H51A05D3), M. Meghana (20H51A05P5) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

###### The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Mr.M.Shiva Kumar Dr. Siva Skandha Sanagala**

**Assistant Professor Associate Professor and HOD**

**Dept. of CSE Dept. of CSE**

#### ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Mr.M.Shiva Kumar, Assistant Professor** , Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

M.GuruSai Chawan 20H51A0517

T.Manohar 20H51A05D3

M.Meghana 20H51A05P5

Crime Analysis and Prediction

**TABLE OF CONTENTS**

**CHAPTER**

**NO. TITLE PAGE NO.**

LIST OF FIGURES ii

ABSTRACT iii

**1** **INTRODUCTION** 1

1.1 Problem Statement 2

1.2 Research Objective 2

1.3 Project Scope and Limitations 2

**2** **BACKGROUND WORK** 3

2.1. Crime Analysis and prediction using data mining 4

2.1.1.Introduction 4

2.1.2. Merits, Demerits and Challenges 4

2.1.3. Implementation 5

2.2. Predicting and Preventing Crimes 6

2.2.1. Introduction 6

2.2.2. Merits, Demerits and Challenges 6

2.2.3. Implementation 7

2.3. Crime Predicting and other uses of neural network 8

2.3.1. Introduction 8

2.3.2. Merits, Demerits and Challenges 8

2.3.3. Implementation 9

**3 RESULTS AND DISCUSSION 10** 3.1. Performance metrics 11

4 **CONCLUSION** 12

4.1 Conclusion 13

**REFERENCES** **14-15**

**GitHub Link 16**

CMRCET B. Tech (CSE) Page No i

Crime Analysis and Prediction

**List of Figures**

**FIGURE**

**NO. TITLE PAGE NO.**

2.2 Architecture 7

CMRCET B. Tech (CSE) Page No ii

Crime Prediction and Analysis

# **ABSTRACT**

Crime analysis and prediction is a systematic approach for identifying the crime. This system can predict regions which have high probability for crime occurrences and visualize crime prone areas. Using the concept of data mining we can extract previously unknown, useful information from unstructured data. The extraction of new information is predicted using the existing datasets. Crimes are treacherous and common social problems faced worldwide. Crimes affect the quality of life, economic growth, and reputation of a nation. With the aim of securing the society from crimes, there is a need for advanced systems and new approaches for improving the crime analytics for protecting their communities. Propose a system which can analyze, detect, and predict various crime probability in a given region. Explains various types of criminal analysis and crime prediction using several data mining techniques.

CMRCET B. Tech (CSE) Page No iv

# **CHAPTER 1**

**INTRODUCTION**

**CHAPTER 1**

**INTRODUCTION**

**1.1.Problem Statement**

Day by day crime data rate is increasing because the modern technologies and hi-tech methods are helping the criminals to achieving the illegal activities .according to Crime Record Bureau crimes like burglary, arson etc have been increased while crimes like murder, sex, abuse, gang rap etc been increased.crime data will be collected from various blogs, news, and websites. The huge data is used as a record for creating a crime report database. The knowledge which is acquired from the data mining techniques will help in reducing crimes as it helps in finding the culprits faster and also the areas that are most affected by crime.

**1.2.Research Objective**

Machine learning is a process that is widely used for prediction. N number of algorithms are available in various libraries which can be used for prediction. In this article, we are going to build a prediction model on historic data using different machine learning algorithms and classifiers, plot the results and calculate the accuracy of the model on the testing data. Building/Training a model using various algorithms on a large dataset is one part of the data. But using these models within different applications is the second part of deploying machine learning in the real world.To put it to use in order to predict the new data, we have to deploy it over the internet so that the outside world can use it. In this article, we will talk about how we have trained a machine learning model, created a web application on it using Flask.

**1.3.Project Scope**

This work in future can be extended to have better classification algorithms which can detect criminals more accurately. We can also increase privacy and some other security measures to protect data set that we are using. Along with this, this work can be further extended to predict who will commit a crime and this can be done using Face recognition. The system will detect if there is any suspicious change in the behavior or usual movements. For example, if a person is moving back and forth in same region over and over might indicate that he is a pickpocket and it will also track person over time.

**CHAPTER 2**

**BACKGROUND WORK**

**CHAPTER 2**

**BACKGROUND WORK**

**2.1 Crime analysis and prediction using data mining**

**2.1.1. Introduction**

Crime analysis and prediction using data mining is a multidisciplinary approach that combines the fields of data science, machine learning, and law enforcement to analyze historical crime data and make predictions about future criminal activities. It involves the use of various data mining techniques to uncover patterns, trends, and relationships within crime data, enabling law enforcement agencies to allocate resources more efficiently and proactively prevent and address criminal activities.

**2.1.2. Merits, Demerits and Challenges**

**Merits**

1.Crime Prevention: Data mining can help law enforcement agencies predict where and when crimes are likely to occur. This allows them to allocate resources and personnel to potential hotspots, thereby deterring criminal activity and improving public safety.

2.Resource Optimization: By analyzing crime data, law enforcement agencies can better allocate their limited resources, such as personnel and patrol cars, to areas with a higher probability of crime, reducing response times and increasing the overall efficiency of operations.

3.Pattern Recognition: Data mining can identify hidden patterns, trends, and relationships in crime data, even across multiple jurisdictions. This can help law enforcement understand criminal behaviors and modus operandi, aiding in investigations.

4.Enhanced Investigative Tools: Data mining can provide law enforcement with tools to link seemingly unrelated crimes and individuals, making it easier to solve complex cases.

**Demerits**

1.Data Quality Issues: The effectiveness of data mining relies on the quality of the available data. Incomplete or inaccurate data can lead to unreliable predictions.

2.Privacy Concerns: Collecting and analyzing data can raise concerns about individual privacy. Striking the right balance between public safety and privacy is challenging.

**Challenges**

Model Accuracy: Ensuring the accuracy of predictive models is crucial. A false-positive prediction can lead to the misuse of resources, while false negatives can have severe consequences.

Legal and Ethical Challenges: The use of data mining in law enforcement raises ethical questions about surveillance, profiling, and individual rights. Striking a balance between public safety and civil liberties is a constant challenge.

**2.1.3 Implementation**

Implementing crime analysis and prediction using data mining involves several steps:

1.Data Collection: Gather historical crime data, which may include incident reports, arrest records, and other relevant information. Ensure data quality and integrity.

2.Data Preprocessing: Clean, transform, and prepare the data for analysis. This may involve handling missing values, normalizing data, and selecting relevant features.

3.Model Selection: Choose appropriate data mining algorithms and techniques for crime prediction. Common choices include decision trees, random forests, and machine learning models.

4.Training: Train the chosen model on historical data, using a portion of the data for this purpose. Ensure the model's accuracy and adjust parameters as needed.

5.Testing and Validation: Evaluate the model's performance using separate testing data. This helps ensure that the model can make accurate predictions on new, unseen data.

6.Deployment: Implement the model into law enforcement operations, integrating it with real-time data streams for ongoing crime prediction.

7.Monitoring and Feedback: Continuously monitor the model's performance, and refine it as more data becomes available and as the predictive accuracy improves.

8.Ethical Considerations: Address ethical and legal issues, such as algorithmic bias, privacy, and transparency, in the deployment of data mining solutions.

**2.2 Predicting and Preventing crimes**

**2.2.1. Introduction**

It is critical in different fields of science to serve analytical purposes and plays an essential role in human life and fields such as education, business, medicine, health, and science. Data mining is an attractive process of discovering a valid, understandable, helpful pattern and valuable information in large amounts of data. The main goal of data mining is to find out fascinating and concealed knowledge in the data and summarize it in a significant form. Similarly, the results should be in the form that conveys the inside information effectively. Therefore, classification techniques are among the most important and commonly used techniques in data mining, and supervised class prediction techniques allow nominal class labels for predictions. The main goal is to propose a prediction model that predicts crime based on past criminal records.

**2.2.2. Merits, Demerits and Challenges**

**Merits**1.Resource Allocation: By analyzing crime data, agencies can optimize resource allocation, directing personnel and assets to areas with higher crime probabilities, thus reducing response times and improving overall effectiveness.

2.Reduction in Crime Rates: Proactive policing can lead to a decrease in crime rates, enhancing the safety and well-being of communities.

**Demerits**

1.Data Quality Issues: The effectiveness of data mining relies on the quality of the available data. Incomplete or inaccurate data can lead to unreliable predictions.

2.Algorithm Bias: Predictive models can inherit biases from historical crime data, potentially perpetuating disparities in law enforcement practices.

**Challenges**

Data Integration: Law enforcement agencies often use different systems and formats for recording crime data. Integrating and standardizing this data for analysis can be a significant challenge.

Data Volume: Managing and analyzing the vast amounts of data generated by law enforcement agencies can be overwhelming. Scalable solutions are needed to handle this volume effectively.

**2.2.3 Implementation**

The proposed model contains three techniques and performs evaluation through accuracy, precision, and recall evaluation matrices. The data is descriptively analyzed and statistical crime distribution over space and time is visualized to help attain potential patterns. The features are extracted from the original dataset, and the classification is performed. The experimental results show that the Gradient Boosting Decision Tree prediction model is better than the other two techniques for predicting crime, based on historical data from a city. The analysis and prediction model can help the security agencies utilize the resources efficiently, anticipate the crime at a specific time, and serve society well.

The Exploratory Data analysis is done i.e.,

Number of crimes in individual Category.

Top 10 crimes year-wise, month -wise, per day of week.

Normalizing by month reveals common pattern in data

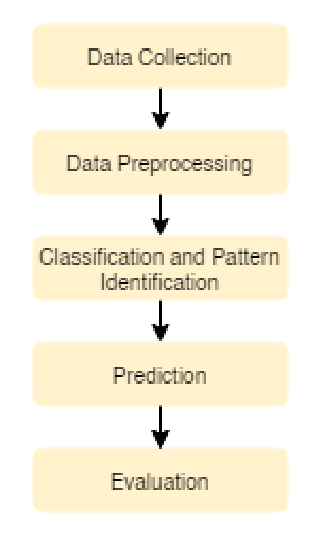
****

Fig 2.2 Architecture

**2.3. Crime Predicting and other uses of neural network**

**2.3.1 Introduction**

Neural network models to predict specific types of crime using location and time information and to predict a crime’s location when given the crime and time of day are developed to demonstrate the application of neural networks to police decision making. The neural network crime prediction models utilize geo-spatiality to provide immediate information on crimes to enhance law enforcement decision making. The neural network models are able to predict the type of crime being committed 16.4% of the time for 27 different types of crime or 27.1% of the time when similar crimes are grouped into seven categories of crime. The location prediction neural networks are able to predict the zip code location or adjacent location 31.2% of the time.

**2.3.2 Merits, Demerits and Challenges**

**Merits**

Effective Visual Analysis

Processing of Unorganized Data

Adaptive Structure

**Demerits**

Hardware Requirement

Incomplete Results

Data Suitability

Minimal Control

**Challenges**

Data Quality and Availability: Neural networks require large volumes of high-quality data to make accurate predictions. Crime data can be incomplete, inconsistent, and subject to errors, making it challenging to build reliable models.

Data Imbalance: Crime datasets often exhibit imbalances, where certain types of crimes are rare compared to others. Neural networks can struggle to learn patterns in imbalanced data, leading to biased or inaccurate predictions.

**2.3.3 Implementation**

To demonstrate NN model usage in police crime predicting and preparedness, two NN prediction models are developed. Both models utilize supervised learning NN architectures.

Two different training methods are used: backpropagation and radial basis function.Additionally, the backpropagation learning algorithm has been shown to be a reliable universal classifier. NN models may impose data limitations and hence the desire to also utilize RBF trained NN models. The best performing architecture of all attempted NN architectures is reported for each NN crime forecasting problem.

NN to Predict Type of Crime

The first NN model’s purpose is to predict what crime is occurring when given only location and time information. This is to simulate a 911 call response when the caller is unable to describe the crime that is occurring. It may also be viewed as a preparatory model for police officers as they are assigned to different neighborhoods on a specific day of the week and time and consequently hear a call for help. The goal of the NN is to predict the most likely crime that is occurring in order to better prepare police response.

Data were originally divided evenly into a training set and a validation set, with extra examples due to an odd number of samples for a specific crime cluster placed in the validation set. Backpropagation trained NNs with both a single hidden layer and two hidden layers, which would allow for greater nonlinearities in the solution surface are developed. Over represented clusters in the training data is a cause of poor classification and prediction performance by machine learning methods including NNs.

NN to Predict Crime Location

The second NN model’s purpose is to predict the location of a crime when only the general type of crime and time of day and day of week are known. This simulates the situation when an emergency call comes in and the crime reporter, victim or otherwise, is unsure of their location. Such could occur if the reporter were a visitor to the city or was in some other way unable to determine their location, such as a kidnap victim in the trunk of a vehicle. The goal of this NN is to predict the most likely zip code. The choice of the zip code area for crime location prediction, where an area of prediction is used instead of spot identification of a precise geographic location (i.e., longitude and latitude).

**CHAPTER 3 RESULTS AND DISCUSSION**

**CHAPTER 3**

**RESULTS AND DISCUSSION**

The algorithms were run to predict each of the following features in the datasets: murders, murd Per Pop, rapes, rapes Per Pop, robberies, rob Per Pop, assaults, assault Per Pop, and Violent Crimes Per Pop. Note that per Pop refers to for every 100K of people. The algorithm that gives the lowest error values for each feature and the highest correlation coefficient is highlighted in the results present. The application displays the results of the crime type prediction based on the input crime summary. The middle part of the application displays the results of the crime type prediction based on the input crime summary.

The prediction results show the probability values for 21 crime types as a bar graph. the types of crimes with the highest probability values are displayed. Based on the predicted results, field personnel can quickly identify the type of crime. The predicted CRS is displayed at the bottom of the application. This prediction result is also displayed in real-time. The platform can predict crime type and CRS and display the prediction results in real-time therefore, field staff, such as police officers, can easily check predictive information about crimes received through the platform. Crime solving is very difficult work which requires experience and intelligence of human along with Artificial intelligence approaches which assist them in problems of crime detection.

Other than that predicting future patterns of crimes basically involves the changes in the rate of crime in next year and applied methods of prediction to help discover those changes in upcoming years. Prediction of crime with the help of artificial neural network is usually better in accuracy and evaluate the target test-time of different models function much faster. As proposed by the problem lies in finding out techniques which can analyse efficiently the growing data set of crime. The accuracy for predicting crime is basically depends upon on the crime data set used. If used training data set is very large, then model will be trained with very good accuracy while if the data set used for training purpose is having less size, then small degree of training is attained. Also,the prediction accuracy also dependent on dimension of training data set. This will give more right results if the model is highly trained and will not give good results if the model is not trained properly.

CHAPTER 4

**CONCLUSION**

**CHAPTER 4**

**CONCLUSION**

The purpose of this project is to identify the locations with highest particular crime so that the police can be given an act of hint to take proactive measures before the situation could worsen up. This model helps to analyze and predict crime. Using machine learning approaches the areas and hotspots can be predicted based on the type of crime and give the overall prediction of any crime. The project also focuses on building this approach by importing machine learning modules, Police API and calculations of supervised learning to give an inch of advancement for the same.The proposed model is very useful for both the investigating agencies and the police officials in taking necessary steps to reduce crime. The project helps the crime analysis to analyze these crime networks by means of various interactive visualizations. Future enhancement of this research work on training bots to predict the crime prone areas by using machine learning techniques. Since, machine learning is similar to data mining, advanced concepts of machine learning can be used for better prediction. The data privacy, reliability, accuracy can be improved for enhanced prediction.

**REFERENCES**

## REFERENCES

* <https://www.researchgate.net/publication/343750790_Crime_Prediction_and_Analysis>
* <https://www.ijcrt.org/papers/IJCRT2102057.pdf>
* <https://ieeexplore.ieee.org/document/9245120>
* <https://www.ijraset.com/research-paper/crime-prediction-and-analysis>
* <https://www.academia.edu/42277756/Crime_Analysis_and_Prediction_using_Data_Mining_Technique>
* <https://towardsdatascience.com/crime-location-analysis-and-prediction-using-python-and-machine-learning-1d8db9c8b6e6>

## 

**GitHub Link**

1. [https://github.com/manohar-3112/crime analysis and prediction.git](https://github.com/manohar-3112/crime%20analysis%20and%20prediction.git)
2. [https://github.com/GURUSAICHAWAN2863 /crime analysis and prediction.git](https://github.com/GURUSAICHAWAN2863%20/crime%20analysis%20and%20prediction.git)
3. [https://github.com/meghanamaroju /crime analysis and prediction.git](https://github.com/meghanamaroju%20/crime%20analysis%20and%20prediction.git)