



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 3, March 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# Crime Rate Prediction and Forecasting Using Machine Learning and Deep Learning

R.L.Indu Lekha M.E., Dharnesh Raja S, Imthiyas F, Jeeva Ganesan T

Assistant Professor, Department of Computer Science and Engineering, Adhiyamaan College of Engineering (Autonomous), Hosur, Tamil Nadu, India

Department of Computer Science and Engineering, Adhiyamaan College of Engineering (Autonomous), Hosur, Tamil Nadu, India

**ABSTRACT:** Crime Forecasting refers to the basic process of predicting crimes before they occur. Crimes are a common social problem affecting the quality of life and the economic growth of a society. A crime is a deliberate act that can cause physical or psychological harm, as well as property damage or loss, and can lead to punishment by a state or other authority according to the severity of the crime. For our daily purposes we have to go many places every day and many times in our daily lives we face numerous security issues such as hijacking, kidnapping, harassment, etc. Daily there are huge numbers of crimes occurring frequently. These require keeping track of all the crimes and maintaining a database for same which may be used for future reference.

## I. INTRODUCTION

Crimes are the significant threat to the humankind. There are many crimes that happens regular interval 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 of different type – robbery, murder, rape, assault, battery, false imprisonment, kidnapping, homicide. 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 activities. 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. The above problem made me to go for a research about how can solving a crime case made easier. Through many documentation 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 type of crime which will occur in a particular area. The objective would be to train a model for prediction. The training would be done using the training data set which will be validated using the test dataset.

Crime Forecasting refers to the basic process of predicting crimes before they occur. Crimes are a common social problem affecting the quality of life and the economic growth of a society. A crime is a deliberate act that can cause physical or psychological harm, as well as property damage or loss, and can lead to punishment by a state or other authority according to the severity of the crime. For our daily purposes we have to go many places every day and many times in our daily lives we face numerous security issues such as hijacking, kidnapping, harassment, etc. Daily there are huge numbers of crimes occurring frequently. These require keeping track of all the crimes and maintaining a database for same which may be used for future reference. The current problem faced are maintaining of proper dataset of crime and analyzing this data to help in predicting and solving crimes in future. The main objective of this project is to analyze dataset which consist of numerous crimes and predicting the type of crime which might occur in future depending upon various conditions. We will be using the technique of machine learning and data science for crime prediction of Chicago and Los Angeles crime data set. The K-Nearest Neighbor (KNN) classification and various other algorithms will be tested for crime prediction and one with better accuracy will be used for training. The main purpose of this project is to give a brief idea of how machine learning can be used by the law enforcement agencies to detect, predict and solve crimes at a much faster rate and thus reduce the crime rate. It is not restricted to Chicago and Los Angeles, this can be used in other states or countries depending upon the availability of the dataset.

## II. LITERATURE SURVEY

### **Empirical Analysis for Crime Prediction and Forecasting using Machine Learning and Deep Learning Techniques**

**Author – Dr. Suma T , Megha C**

**Year - 2021**

Crime data gets recorded for a long time, however not to do predictive policing, but to keep a record of crimes and to justify imprisonment. The focus is not the data itself but only the history and actions of a person. For a long time, crime was described as unique and random, in a lot of cases even as an act of opportunity. And thus crime forecasting made no sense, as predicting random events is impossible. This statement means it was not possible to find a general pattern or a connection between different crimes as there is none. However, due to collecting more and more data around crime and especially storing the location made some of these problems obsolete. Recent research showed that there are patterns in the data and that crimes are not unique and random in every case.

### **Comprehensive analysis of Classical Machine Learning models and Ensemble methods for predicting Crime**

**Author – S.R Divyasri ,R Saranya**

**Year – 2022**

The "routine activity approach" is one of the theories, which led to crime forecasting. Other than looking at the characteristics of the offender, the spotlight shifted to the circumstances around the crime. Not just the offender is the focus anymore, but the space and time of the motivated offenders, the targets and absence of capable guardians for the crime. The police records temporal connections between crimes quite well because they show if an individual committed not only the crime he or she was arrested for, but if there are more crimes with the same pattern as the crime of the individual. It means that there is a connection between the location of the crime and the crime itself.

### **Crime pattern detection, Analysis & prediction. In Electronics, Communication and Aerospace Technology**

**Author – Yadav, S., Timbadia, M**

**Year – 2017**

Crime Forecasting refers to the basic process of predicting crimes before they occur. Crimes are a common social problem affecting the quality of life and the economic growth of a society. A crime is a deliberate act that can cause physical or psychological harm, as well as property damage or loss, and can lead to punishment by a state or other authority according to the severity of the crime. For our daily purposes we have to go many places every day and many times in our daily lives we face numerous security issues such as hijacking, kidnapping, harassment, etc. Daily there are huge numbers of crimes occurring frequently. These require keeping track of all the crimes and maintaining a database for same which may be used for future reference. The current problem faced are maintaining of proper dataset of crime and analysing this data to help in predicting and solving crimes in future. The main objective of this project is to analyse dataset which consist of numerous crimes and predicting the type of crime which might occur in future depending upon various conditions. We will be using the technique of machine learning and data science for crime prediction of Chicago and Los Angeles crime data set. The K-Nearest Neighbour (KNN) classification and various other algorithms will be tested for crime prediction and one with better accuracy will be used for training. The main purpose of this project is to give a brief idea of how machine learning can be used by the law enforcement agencies to detect, predict and solve crimes at a much faster rate and thus reduce the crime rate. It is not restricted to Chicago and Los Angeles, this can be used in other states or countries depending upon the availability of the dataset.

## III. PROJECT DESCRIPTION

### **EXISTING SYSTEM**

Classification and Decision Tree A decision tree is an algorithm that uses a tree shaped graph or model of decisions including chance event outcomes, costs, and utility. It is one way to display an algorithm. Naive Bayes -In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with independence assumptions between the features. The technique constructs classifier models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. Linear Regression – The analysis is a statistical process for estimating the relationships among variables. Linear regression is an approach for modelling the relationship between a scalar dependent variable Y and one or more explanatory variables denoted X.



#### DISADVANTAGES

- In this step we need prepare data into right format for analysis
- Data cleaning Analyse and Transform Variable We may need to transform the variables using one of the approaches
- Normalization or standardization
- Missing Value Treatment

#### IV. PROPOSED SYSTEM

Predictive modelling is the way of building a model that is capable of making predictions. The process includes a machine learning algorithm that learns certain properties from a training dataset in order to make those predictions. Predictive modelling can be divided further into two areas: Regression and pattern classification. Regression models are based on the analysis of relationships between variables and trends in order to make predictions about continuous variables. In contrast to regression models, the task of pattern classification is to assign discrete class labels to particular data value as output of a prediction.

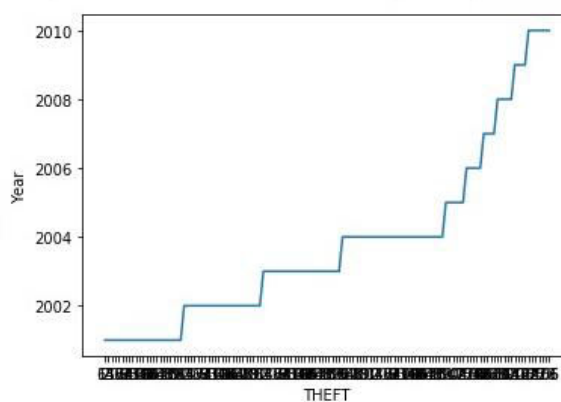
#### ADVANTAGES

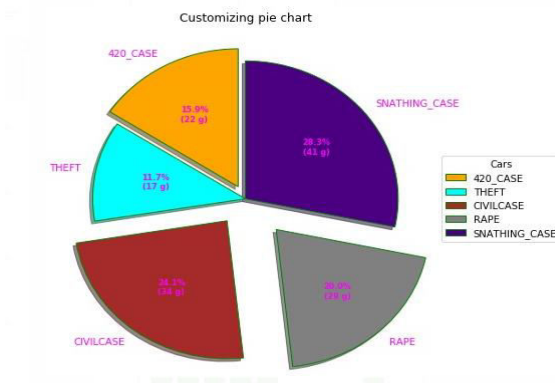
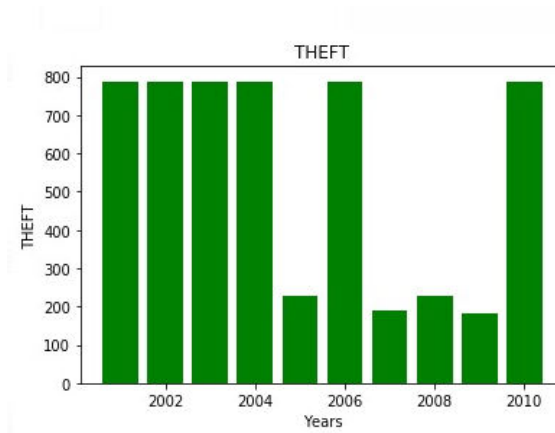
- Descriptive analysis on the Data
- Data treatment (Missing value and outlier fixing)
- Data Modelling
- Estimation of performance

#### V. RESULT

Crimes are serious threats to human society, safety, and sustainable development and are thus meant to be controlled. Investigation authorities often demand computational predictions and predictive systems that improve crime analytics to further enhance the safety and security of cities and help to prevent crimes. We achieved an improved predictive accuracy for crimes by implementing different machine learning algorithms on Chicago and Los Angeles crime datasets. Using the concept of machine learning we have built a model using training data set that has undergone data cleaning. Data visualization helps in analysis of data set. The graphs include bar, pie, line and scatter graphs each having its own characteristics. We generated many graphs and found interesting statistics that helped in understanding Chicago crimes datasets and Los Angeles crimes datasets that can help in capturing the factors that might help in keeping society safe.

#### OUTPUT





```

DISTRICT Year 420_CASE ... THEFT CIVILCASE SNATCHING_CASE
0 chennai 2001 20.0 ... 65 3 3.0
1 chengalpattu 2001 0.0 ... 3 147 147.0
2 villupuram 2001 5.0 ... 147 208 208.0
3 kalakurichi 2001 0.0 ... 208 3 3.0
4 salem 2001 0.0 ... 3 63 63.0
.. ..
125 puthukkottai 2010 NaN ... 75 1 208.0
126 vellur 2010 NaN ... 41 142 3.0
127 mayavaram 2010 NaN ... 228 208 41.0
128 ooty 2010 NaN ... 106 3 2.0
129 Puducherry 2010 NaN ... 66 41 75.0

[130 rows x 8 columns]
    
```

```

# Importing libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Loading data
df = pd.read_csv('data.csv')

# Cleaning data
df.dropna(inplace=True)

# Grouping data
df.groupby('Year').sum()

# Plotting pie chart
df.groupby('Category').sum().plot(kind='pie')
    
```

## VI. CONCLUSION

With the help of machine learning technology, it has become easy to find out relation and patterns among various data's. The work in this project mainly revolves around predicting the type of crime which may happen if we know the location of where it has occurred. Using the concept of machine learning we have built a model using training data set that have undergone data cleaning and data transformation. The model predicts the type of crime with accuracy of 0.789. Data visualization helps in analysis of data set. The graphs include bar, pie, line and scatter graphs each having its own characteristics. We generated many graphs and found interesting statistics that helped in understanding Chicago crimes datasets that can help in capturing the factors that can help in keeping society safe.

## VII. FUTURE ENHANCEMENT

The proposed systems build a baseline for crime forecasting with deep learning methods and heatmap generation. Especially, as forecasting in most cases is not done using unsupervised and semi-supervised generative models, the systems show what is possible using a form of artificial intelligence to forecast crime. However, there are many possibilities to improve the systems and methods. Improve Architecture One way to improve the systems would be to revise the architecture itself further and add further extensions and layers to it. It is always possible to change the general layers, by adding or removing fully connected or convolution layers.

## REFERENCES

- [1]. Yadav, S., Timbadia, M., Yadav, A., Vishwakarma, R.,& Yadav, N. (2017, April). Crime pattern detection, Analysis & prediction. In Electronics, Communication and Aerospace Technology (ICECA),2017 International conference of (Vol. 1, pp. 225-230). IEEE.
- [2] G. Vimal Raja, K. K. Sharma (2014). Analysis and Processing of Climatic data using data mining techniques. *Envirogeochimica Acta* 1 (8):460-467.
- [3] Mohit, Mittal (2013). The Rise of Software Defined Networking (SDN): A Paradigm Shift in Cloud Data Centers. *International Journal of Innovative Research in Science, Engineering and Technology* 2 (8):4150-4160.
- [4]. Al Boni, M., & Gerber, M. S. (2016, December). Area Specific Crime Prediction Models. In Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on (pp. 671-676). IEEE.
- [5]. A. J. Hussain, P. Liatsis, M. Khalaf, H. Tawfik, and H. Al-Asker, "A dynamic neural network architecture with immunology inspired optimization for weather data forecasting," *Big Data Res.*, vol. 14, pp. 81–92, Dec. 2018.
- [6]. Nath, S. V. (2006, December). Crime pattern detection using data mining. In Web intelligence and intelligent agent technology workshops, 2006.wi-iat 2006 workshops. 2006 ieeewic/acm international conference on (pp. 41-44). IEEE.
- [7]. M. Alsharif, M. Younes, and J. Kim, "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea," *Symmetry*, vol. 11, no. 2, p. 240, Feb. 2019
- [8]. Chicago Data Portal. Accessed: Nov. 2, 2019.[Online].Available:<https://data.cityofchicago.org/Public-Safety/Crimes-2001-topresent-Dashboard/5cd6ry5g>.
- [9]G. Vimal Raja, K. K. Sharma (2015). Applying Clustering technique on Climatic Data. *Envirogeochimica Acta* 2 (1):21-27.
- [10]Naga Ramesh Palakurti 2022. Empowering Rules Engines: AI and ML Enhancements in BRMS for Agile Business Strategies, *International Journal of Sustainable Development through AI, ML and IoT*, 1(2), 1-20. <https://ijsdai.com/index.php/IJSDAI/article/view/36>
- [11]. Los Angeles County GIS Data Portal. Accessed: Nov. 2,2019.[Online]. Available:<http://egis3.lacounty.gov/data/portal/?s=crime>



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details