# **Crime Prediction Using Machine Learning**

## Authors

Rishav raj (22SCSE1010527) Md. Tariq Khan (22SCSE1011804)

Supervisor:

## Ms Km Ikra (Department of CSE, Galgotias University)

## ABSTRACT

Crime is a threat to national security and order. Therefore, crime analysis is gaining importance because it shows time and space through spatial data. Crime prediction has been studied by many researchers with the help of machine learning. In today's conditions where crime rates are rapidly increasing, problem-solving methods are slow and ineffective and do not yield results. Therefore, if we can develop a way to predict crimes before they occur or develop a "control" that can help the police, this will reduce the burden on the police and help prevent crime. In conclusion, the broad overview of research on crime prediction using machine learning and deep learning discussed in this article can provide valuable information to researchers in this field. By better understanding the crime prediction process, law enforcement agencies can develop strategies to prevent and respond to crimes.

## I. **INTRODUCTION**

Security is a very important part of life. Unless we are safe, most of us Basic needs cannot be met. Therefore, security is a human need A life that helps us achieve our goals, collectively or individually. sin is one Health problems have affected our lives in many ways. have the ability Identify areas that are not safe for crime and detect recent crimes Special places have become a growing concern for local (algorithms) Technologies that aim to prevent malicious attacks are as follows: The content of social media datasets is extracted using data mining techniques. Our important documents The place is social media. The main goal is to detect all the hidden areas of the data.

## And guess what?

Security is a very important part of life. However, using classical machine learning Unless we are safe, most of us Basic needs algorithms, text is processed at the level of cannot be met. Therefore, security is a human semantic comparison, a process based on need A life that helps us achieve our goals, semantic

Machine learning algorithms are generally divided into supervised algorithms and unsupervised algorithms.

Machine learning algorithms can apply past learning to

## **1.1 OBJECTIVE**

is extracted using data mining techniques. Using data analytics and predictive analytics Our important documents The place is social media. The main goal is to detect all the hidden areas of the data. Using data analytics and predictive analytics The system provides a platform for analyzing crime records. Crime data by location and time (time of day, day of week, and seasonality) prevent or control crime Reduce the occurrence of dangerous low levels Crime

## 1.2 **SCOPE**

The responsibility for the violation will lie with planning process. Control, the investigation and prevention. System Use time series, clustering and data mining techniques Make predictions about future violence. This will This is done through crime scene images. Show concentration data using geo-heat maps and instant access points.

### 1.3 DATASET

The data in our dataset covers the period from 2001-2016. Primary data age, area where the victim lives and year. Our data There are three forest,81.35%. categories: Incidents concerning women, IPC at child and state level. We use this **3.PROPOSED METHODOLOGY** information to estimate violations of state laws.

### 2. LITERATURE SURVEY

When the names of the defendants are hybrid model. we mean model at that time was: Analyze, prioritize and then deploy training Algorithms and collect flavor data. K-means the recovery rate of clustering algorithm exceeds 75% .

Then check, prepare and fit the model Used to **3.1 ALGORITHM** sample datasets and train the most common criminals Special crime. Author compares Naive Bayes Classifier and Decision Tree Crime Prediction method. investigate Identify crime scene [6]. Rajesh Kanna et al. CNN deep learning model based on long sequences Time frame required to investigate a crime is short.

to Future patterns of violence. Also, as Education Used to verify access using time and predicted How much security should be location feature. To improve the accuracy of measured and The need to control crime to feature selection, Adopt the Black echniques have been published Check crime records, crime predictions, crime statistics Identify and locate crime hotspots [8]. Abuse Stacked Sparse Autoencoder Networks for aive Bayes classifier takes less time to execute and A high accuracy rate of 78.05% was achieved. Checked. Many types of crimes are illegal and predictable. The probability of a criminal committing a crime Once again [10]

> Sivaranjani et al[13] conducted a crime study in six cities in the USA.In the Indian state of Tamil Nadu k-means clustering is analysis determines factors affecting serious crimes Price.The authors reported improved performance compared to random

The answer lies in statistics and machine learning the model uses classification. clustering and relationship between population, economy, society, victims and different areas. By analysis Patterns in crime records. Time Series Technology It has been proven that the above algorithm allows the model to: Accurately predict the nature of crime Characteristics of growth and change over time.

This field includes a group of algorithms. Kmeans segmentation is widely used. accept. Unlike the K-means method, this linear Batch data Feedback is used in interviews to identify transformation Use batch algorithms and then customers However, the Navy's cost number by category Bayesian algorithm provides the initial results. Achieve higher accuracy. Linear and multilinear Regression analysis shows the relationship between related data or data. Education and Training in the Field of different (such as age, gender, etc.) and scene. In this method, the ages of the victims are calculated according to: According to the input criteria specified in the metadata column. When looking at a crime scene, linear regression is used Crime prediction to determine the most crimes This is illegal. Analysis of historical data shows that The ratio of female to male victims is increasing. wake up. Statistics show how often people are victimized It is suitable for both men and women. Due to the large number of criminal records and similar connections between these different texts Criminology is a discipline suitable for application in literature. Data mining technology.

Figure 1 shows the proposed approach. Information Use the algorithm to convert this into a group. Then conduct a survey to identify areas where violence occurs. These categories describe various criminal practices. Overlapped on a map of the police jurisdiction. Combination Location of the According to the type of crime, Use protective measures in crime prone areas this is a problem. The easiest and most popular 3.2 OUTCOME category The technique used in research and business economics is K-means. This perform better than other monitoring methods. For example, distribution. Then check the samples, Sampling datasets and design and use for training algorithm. More than 75% provided by K-Means Clustering technology. Author used Broken window

Then firstly the model is analyzed and Put it into practice, train the algorithm, taste the collection information. The image models Verification, prior knowledge confirmation Crime centers. The most accurate of 0.87% Deep learning model. An overview of upcoming criminal records and

Independent variables found at the crime Creating a connection between two successful projects and using independent variables, multiple linear regression. For one class and many different classes, female, and those who do not know their gender. Age can be excluded There are three groups: young, old and young. K-recently Neighbor classifier is used for the following purposes: Different targets



## Figure 1

"India Crime Analysis" software system Currently available and designed Criminal investigations can perform tasks that cannot be done by other means. So, even some the problem was answered and the best solution emerged. Maps have been created for every city, state and country. All types of users. The system is accurate and viewable. Analyze and predict in animated form If the system cannot provide, the violation is real If such a situation occurs, it will be reported as unavailable. The and information is either direct or indirect.

Illegal betting is possible. Classification and the basics of algorithmic violations can be regression using machine learning machine. found in this section. them Criminal charges are assessed based on a number of factors 4.TEST CASES including: Age, gender, location, and monthly cost. A range of documents Sources and methods used to make the predictions include: Data analysis, co-analysis, and Statistical models for predicting future crimes. Because Some minorities are also included The violation causes data imbalance, model estimation There is an important thing to remember. So to solve this problem, We use random oversampling. Thinking Perform periodic analysis of current violations to predict future violations Crime. Now is an opportunity to study crime. Inference algorithms using time series research Current crime, predict future crimes. This The behavior of previously recorded data can be used to: Predict future crime patterns. Each model predicts target performance The relationship between the predictor variable and the dependent variable Can be changed. To make these models more accurate We must be able to analyze and predict many things Future events that may also affect victimization crime. This study predicts future crimes A clear and concise way.



Figure 2 Visualization of the data set from 2001 to 2015



**Figure 3 Predicted graph of child Marriage** in Haryana



## Figure 4 Predicted graph of murder case in Tamil nadu



Figure 5 Predicted graph of arms act 2020 in kerala



# Figure 6 Crime prediction for the state with accuracy



# Figure 7 Crime prediction for overall India and it's accuracy

## **5 CONCLUSION**

It is difficult to use a price algorithms have been used to create and we are testing and linear regression, Figures 1 and 2 show that Crime is . Although some linear systems work well, Full scene models are used to provide higher exposure K-Neeg Neighbors are our crime action It also provides the necessary truth. We can do this Use these predictions to identify the stronger machine. This algorithm will also show superiority Identity and location accuracy Crime rate is high. Finally used CNN algorithm analyzes image data and Google API to identify the heating zone.

## **6 REFERENCES**

[1] Zakir Hussain, K., Durairaj, M., Farzana, G.R.J.: Using data mining to investigate crime March 30-31, 2012. Technology, Nagapattinam, India Proceedings of Conference International Advance in Engineering Science and Management, IB, 2012, pp. 1–8.

[2] Kavitha, M, Roobini, S. System perspective and impact. Artificial Intelligence in Smart Healthcare Systems, Principles, Competition and Usage, Technology Learning and Artificial Intelligence in Healthcare system. 2023; 25– 56.

[3] Sathya, R, Ananthi S, Vaidehi K, Position Dependency Based Hybrid Hyperconvolutional Neural Networks Authorization Approval Method Intelligent transportation, concurrency and Informatics: Application and Knowledge, 2023; 35:1-25.

[4] Tayebi, M.A, Gla, U, Brantingham, P. L. Learning Check the Location: Crime Scene Investigation Weather, May 27-29, 2015, Baltimore, Maryland, USA, Proceedings of IEEE International Conference Intelligence and Security Informatics (ISI), IEEE, 2015, International pp. 25-30

[5] Akash, S., Prabhakaran Purnachandran, Vijay Krishna. Menon, Soman, Systematic Data, Springer Cham, 2019, Visualization Malware family identification Page 1158-1159. technology, Network Security and Safety Information Systems, pp. 24-46: i.

[6] Rajesh Kanna, P, Santhi, P. Hybrid Invasion Detection using Black Widow as Mitigation Guide Optimizing Convolutional professional applications, Vol. 2022, 194: (116545).

[7] Rajesh Kanna, P, Santhi, P, Combined Deep Learning High efficiency intrusion detection system method Using the features of a mixed-use site, 2021, Knowledge Based Systems Volume 226: (107132).

Investigative Data for Crime Investigation IEEE, 2016, pp. 1-6. and Weather Forecasting, Meeting Date: 19-India. Proceedings of Conference Network and Software Informatics, IEEE, 2016, page 406 412.

[9] Nath, S. V.: Use of data for crime investigation Mining, 18-22 December 2006, Hong Kong, China Proceedings of

Conference on Network Intelligence and Intelligent Agent Technology, IEEE, 2007, p. 41-44 : I.

K.P. [10] Zhao Xin and Tang Jianjun: Exploring Transformation Crime Scene, November 18-Chapter Issue: 12, Detailed research and 21, 2017, New Orleans, Louisiana, USA Data analysis Deep Learning Architecture and Mining Workshop Conference, IEEE, 2017,

> [11] Privatharishini, M, Nirmala Devi, M.: Deep Learning Malicious module stacking reported as Sparse Autocoder Networks for VLSI Circuits Belief, Measurement, 2022, Vol. 194 (111055).

Long Term Memory Neural networks, [12] Shamsuddin, N.H.M., Ali, N.A., Alwee, R.: Overview of Crime Prediction Methods, Tsib Hlis Ntuj 23-2 2017, Johor, Malaysia, 2017 6th ICT International Student Project Conference, IEEE, 2017, pp. 100 - 100. 1–5

[13] Sivaranjani, S, Sivakumari, S, Aasha, M.: Crime Tamil Nadu estimation and prediction Clustering Methods, 21-22 October 2016, Kollam, India, Proceedings of International [8] Sathyadevan, S., Gangadharan, S.: Use of Conference Trends in Emerging Technologies,

20 August 2014, meeting place: Guntur, [14] Kansara, Chirag, E.: Using Twitter to International reduce crime Big Data Analysis and Modeling, 23-25 Jaipur, India, December 2016 IEEE International Conference on Recent Advances and Engineering Innovation, IEEE, 2017, pp. 1-8