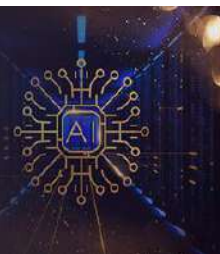


# International Journal of Computing and Artificial Intelligence



E-ISSN: 2707-658X

P-ISSN: 2707-6571

IJCAI 2022; 3(1): 47-49

Received: 18-11-2021

Accepted: 20-12-2021

**Arpita Shukla**

Research Scholar, Computer  
Science and Engineering  
Department, SSSUTMS,  
Sehore, Madhya Pradesh,  
India

**Dr. Jitendra Sheetlani**

Associate Professor, Computer  
Science and Engineering  
Department, SSSUTMS,  
Madhya Pradesh, India

**Corresponding Author:**

**Arpita Shukla**

Research Scholar, Computer  
Science and Engineering  
Department, SSSUTMS,  
Sehore, Madhya Pradesh,  
India

## A study on criminal behaviour and pattern prediction using machine learning techniques

**Arpita Shukla and Dr. Jitendra Sheetlani**

DOI: <https://doi.org/10.33545/27076571.2022.v3.i1.a.46>

### Abstract

Recent advances in the field of machine learning have enhanced the likelihood of applying these algorithms in real-time police situations. Machine learning advances have improved the likelihood of applying these algorithms in real-time policing. It's a must-have when it comes to AI and big data analysis (BDA). It has powerful algorithms that can classify data, recognise patterns, and learn to perform a task on its own. Despite its growing popularity in recent years, the general public and even most professionals are unfamiliar with this discipline. In this current era of the globe, crime and theft are on the rise at an alarming rate, and it has become an overwhelming problem that everyone is coping with. Finding a pattern that can be used to predict the future is critical.

**Keywords:** Crime, criminal behavior, machine learning, intelligence algorithm, big data analysis

### 1. Introduction

Everyone has been a victim of crime at some point in their lives. There is no universal definition of a crime, but I am, without a doubt, a loss and a danger to humanity. Murder, rape, child abuse, kidnapping, theft, physical abuse, imprisonment, battery, murder, and a slew of other crimes are just a few examples. The police, the FBI, and other agencies Abstract are battling with the adoption of various models, as well as solving crimes with a huge number of data sets, in all states and nations with much higher frequency. That is why operations based on pattern recognition and forecasting, such as the use of data methodologies and procedures, are required. Data mining and machine learning are two conditions that are related in some way and can be effective in criminal analysis and categorization. We can predict the types of crimes that might occur in this location using machine learning <sup>[1]</sup>. These are machine learning algorithms, which allow us to forecast some of the outcomes of our work and offer solutions to difficulties in Indian criminal cases. Classification and regression are the two most used forms of forecasting methodologies. This is the place where classification should be used in this crime-data prophesy. Classification is a regulated prediction strategy that has been applied in a variety of sectors, including stock forecasting and environmental performance, among others <sup>[2]</sup>. The major purpose of this study is to look at a variety of algorithms that may be used to forecast and analyse crime data, as well as to increase the accuracy and reliability of these models as a result of processing the data in order to get better results. The goal is to be the preferred data prediction model by validating the test dataset with a data gathering method <sup>[3]</sup>. The models that are used are as follows: Random forest, categorization, decision trees, and logistic regression.

### 2. Related work

Lawrence and Natarajan (2015) collected data on communities and crime using linear regression algorithms (regression and Trunk decision making with the same finite set of functions). As a result, the linear regression procedure is chosen from among the three options. The purpose of this study is to demonstrate the efficacy and accuracy of machine learning algorithms employed in data mining (the analysis involves forecasting violent criminal activity) <sup>[3]</sup>. VGGNet-19, a suggested system by Umadevi and Priyadharshini, was constructed using pre-trained deep learning models, and will allow you to detect a pistol and a knife in a person's hand to another person. The effects of two distinct pre-trained models, such as GoogLeNet InceptionV3-in-training, were compared <sup>[4]</sup>. Biswas and Basak's (2019) study is available to a wide audience and was collected by police in Bangladesh on the site.

The dataset consists of records of all the different types of crimes (e.g. dacoits, robberies, kidnappings, murders of women and children, homicides, robberies, burglaries, weapons, explosives, drugs, and smuggling) from Canada [5]. Aditya Goyal *et al.* (2020) use artificial intelligence and the resulting methods to combat the detection of instant messaging fraud in the systems of cyber criminals. They describes the problem of studying the strategy of producing the cognitive and socio-class personality of the guilty party and approving individual information about the socio-class indicators of the party under study to be studied [6]. Pardo *et al.* (2020) play an important role in improving the accuracy of criminal investigation and detection outcomes, as well as their analysis and dissemination. Intelligent algorithms and intelligent algorithms are used to detect patterns in information concerning criminal episodes in order to identify, characterise, and meta-analyze the approaches. We used quantitative analysis, systematic review, and meta-analysis [10]. Md Amiruzzaman *et al.* (2021) uses up to two cities, all of which are in the United States, to predict regions with high crime rates. The results demonstrate that our model is not particularly good at predicting high and low crime locations (more than 98 percent and 95 percent in the first and second tests, cities and villages, respectively) [11].

### 3. Background

#### 3.1 Machine Learning

Machine learning is a subset of artificial intelligence that is linked to statistical methodologies, and machines will eventually be able to learn from their mistakes. There are several types of machine learning, including supervised, unsupervised, and reinforcement learning. Because of the nature of the data and the direction of export, this study employs supervised learning. We'll go over the three primary types of machine learning methods in detail: supervised learning, unsupervised learning, and reinforcement learning.

#### A. Supervised Learning

The most frequent paradigm for undertaking machine learning operations is the placebo-controlled experiment. It's commonly utilised for data with a clear relationship between the input and output. The algorithm for determining the function, in particular, forecasts the rating of modified ones using data, which in this case is a symbol. The system was able to create a relationship between the two variables as the training progressed, allowing us to forecast new outcomes. As a result, problem-oriented supervised learning algorithms exist. As we become more and more examples of this, he will be able to learn in the proper manner so that he can perform the job and influence us, and the results will be more accurate. Linear regression, random forest, support vector machine, and artificial intelligence [10, 6], for example, are some of the algorithms that are being trained under supervision. There are two basic categories of supervised learning tasks: categorization and strength.

- Classification is a supervised learning problem in which the student must anticipate a class label.
- Regression is a supervised learning task in which the student must predict a numerical label.

One or more input variables can be used in both classification and regression problems, and input variables can be of any data type, such as numerical or categorical.

#### B. Unsupervision Learning

Unsupervised machine learning has the advantage of allowing you to work with unlabeled data. This means that there's no need to make all of the data machine-readable, allowing you to work with much larger data sets in the application. The model discovers the data structure through observation. Once the model obtains a collection of data, it searches for patterns and connections in that data automatically forming groups [10, 11].

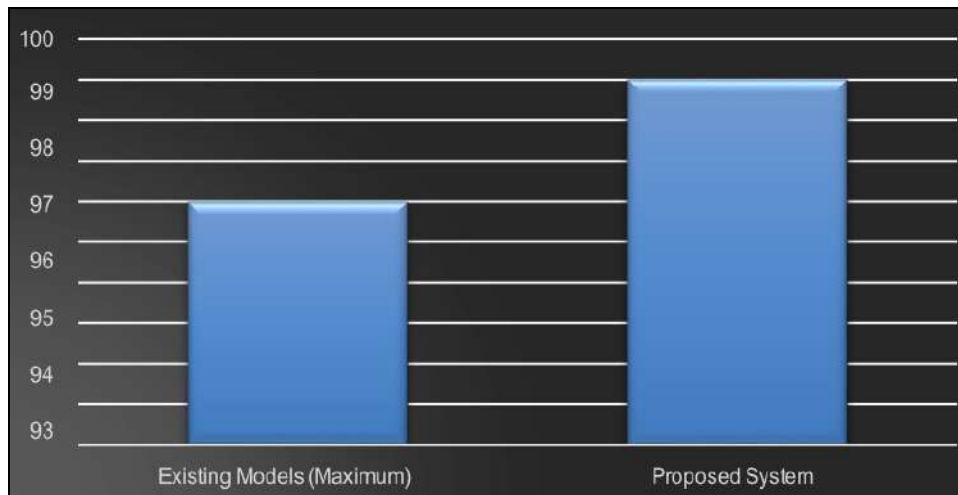
The algorithm will determine the exact nature of the relationship between any two points using controlled label study. Unsupervised learning, on the other hand, lacks tags to work with, resulting in the construction of hidden structures. Relationships between data points that an algorithm can find are abstract and require no human input. Unsupervised learning methods are so popular because they allow you to create these hidden structures. Unsupervised learning algorithms can adapt to data, which is connected with dynamic changes in hidden structures [18], rather than a fixed problem statement, allowing for more post-deployment development of supervised learning algorithms. What we can do is add linkages to our own cluster; for example, we won't be able to state that it's a collection of apples and mangoes since the mango fruit will be isolated. Assume we show models of apples, bananas, and mangoes, which represent actions to be performed based on formulas and relationships, and we generate clusters and divide the data into these groups. If the model receives fresh data, it will be added to the group. Unsupervised learning, clustering, k-means, principle components analysis, sniper rifle, FP-growth, and other techniques are used in this example.

#### 4. Result and Discussion

We use criminal database to identity the behaviour and patterns of crime, predict it using regression and measure the performance. The simulation analysis of the proposed approach is done using the widely used machine learning tool python. This is user friendly and platform independent and consume less time in execution. For this research we use kaggle domain. We choose particular one segment i.e crime against women like rape, kidnapping, abduction, dowry death, molestation, sexual harassment.

**Table 1:** Accuracy comparison of Proposed and Existing models

| Model                     | Accuracy (%) |
|---------------------------|--------------|
| Existing Models (Maximum) | 96           |
| Proposed System           | 99           |



**Fig 1:** Accuracy (%)

The analysis of proposed methodology is done with existing methodology and it is found that the accuracy rate of exiting models for prediction of crime in India is 96% while the crime prediction rate of proposed methodology is about 99% which is more than the existing. The proposed approached is much better in the prediction of criminal behaviour and their patterns if crime than the others.

### 5. Conclusion and Future Scope

Machine learning is a strong field when it comes to AI, and the amount of precision that some algorithms can achieve will be nothing short of amazing. Without a doubt, intelligent data systems, machine learning, and big data analysis are the present and future of intelligent data systems. As can be observed, cutting-edge algorithms and artificial neural networks are used to solve the challenge of crime classification. Although ANN was superior to KNN, there is nothing equivalent in our database. This could be due to a variety of factors, including poor model design, selection, training, and adaptation, or a combination of these. Increasing the number of layers in the network and expanding the programme will almost certainly improve the output. It's a crime-infested format that might be useful in academia, but it'll be useless in real life. A possible expansion of this study would be a crime predictor that would have been projected to be the region of the city; this would be the point of crimes on any given day. The police will be alerted, and they will be able to increase security in this area, for example. The fundamental aim of this effort in the future is to develop a machine learning method (most likely neural networks) for recognising data trends (unsupervised learning).

### 6. References

- Jha G, Ahuja L, Ajay Rana. Criminal Behaviour Analysis and Segmentation using K-Means Clustering, in the proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Amity University, Noida, India, 2020, 83-88.
- Bharati A, Dr. Sarvanaguru RAK. Crime Prediction and Analysis Using Machine Learning, IRJET. 2018;5(9):230-325.
- McClendon L, Meghanathan N. Using Machine Learning Algorithms to Analyze Crime Data, International Journal MLAIJ. 2015;2(1):465-469.
- Navalgund UV, Priyadharshini K. Crime Intention Detection System Using Deep Learning. In proceeding of IEE Explore, 2018.
- Al-Biswas A, Basak S. Forecasting the Trends and Patterns of Crime in Bangladesh using Machine Learning Model, in the proceeding of the 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), Manipal University Jaipur, India, 2019, 28-29.
- Goyal A, *et al.* Criminal Profiling using Machine Learning. International Research Journal of Engineering and Technology. 2020;7(6):230-234.
- Mohammed AF, Wadhah Baiee R. Analysis of Criminal Spatial Events in GI for predicting hotspots, International Scientific Conference of Al-Ayen University, Narsiriyah, Iraq, 2020, 928-931.
- Maddileti T, *et al.* Crime Data Analysis Using Machine Learning Models. International Journal of Advanced Science and Technology. 2020;29(9):3260-3268.
- Patil R, *et al.* Crime Pattern Detection, Analysis & Prediction using Machine Learning. International Research Journal of Engineering and Technology (IRJET). 2020;7(6):2035-2038.
- de Jesus Prado KH, *et al.* Applied Intelligent Data Analysis to Government Data Related to Criminal Incident: A Systematic Review. Journal of Applied Security Research. <https://doi.org/10.1080/19361610.2020.1716511>.
- Md. Amiruzzaman, *et al.* Classifying crime places by neighborhood visual appearance and police geonarratives: a machine learning approach. Journal of Computational Social Science. <https://doi.org/10.1007/s4200-021-00107-x>.