A Machine Learning Analysis and Solution for Predicting Crime Rates

Anubhuti Srivastava

Research Scholar - RKDF University, Ranchi

Abstract

Predicting crime rates is essential to effectively managing resources and combating crime, which is a global problem. Machine learning algorithms can adapt to these changes and generate more accurate predictions as crime rates fluctuate over time in response to changes in social, economic, or political circumstances. Nevertheless, using machine learning to predict crime rates can provide moral dilemmas. Crime prediction is a complex issue that requires advanced analytical methods in order to successfully close the gaps in the detection systems that are already in place. Because of the abundance of crime data and the advancement of modern technology, researchers now have a unique opportunity to study crime detection using machine learning and deep learning methods. Statistics models and algorithms are used in machine learning, a subfield of artificial intelligence, to analyse and predict data. That being said, deep learning methods are a subset of machine learning that use multi-layered artificial neural networks to model complex input-output interactions. To solve the problem of crime prediction, machine learning and deep learning techniques could be applied in a variety of ways. Machine learning-based crime prediction is of interest to many academics. Law enforcement can create more effective programs to combat and deter criminal activity by having a better understanding of crime prediction techniques.

1. Introduction

Any intentional act that violates the law, results in property damage or loss, or causes bodily or psychological suffering is considered a crime. A deliberate action that breaks the law, causes harm to one's person or property, or both, is referred to as a criminal act. Everyone will most likely experience a crime at some point in their lives, and there are numerous types of crimes. People of all ages, places, and origins are impacted by crime. Traditional crime-solving techniques are slow and ineffective, which makes them unable to keep up with the current, rapidly increasing crime rate. Therefore, if we could develop ways to accurately predict crime before it occurs or build a "machine" to assist police officers, it would reduce the workload for police and help prevent crimes. Researchers have recently focused on identifying trends and patterns in the frequency of crimes, with an emphasis on applying machine learning and deep learning techniques to forecast crime. Preventing crimes before they occur can help prevent losses of life and property. Machine learning-based crime prediction has been the subject of extensive research. Gaining a better grasp of crime prediction techniques can help law enforcement agencies establish more effective programs to handle and deter criminal activity. In order to understand how these state-of-the-art techniques are being used to detect criminal activity, anticipate crime patterns, and prevent crime, this article will look at the latest advancements in deep learning and machine learning for crime prediction.

Objective

The primary goal is to forecast crime rates in advance. To quantify the possible decrease in criminal activity
The Police Department has the duty to control and decrease the crime activities, which have increased at an accelerated rate.

Due to the vast amount of crime data available, the Police Department's main issues are crime identification and prediction.

• Technology is required in order to solve cases more quickly.

• It was discovered through numerous documents and cases that data science and machine learning can simplify the task.

Review of Literature

We can better handle the massive amount of data we have these days and extract results by assessing machine learning and data mining approaches. Recently, law enforcement has employed machine learning, data mining, and detection tools for illicit operations in an attempt to lower crime.

By evaluating and assessing the probability of individual crimes as well as patterns and trends in the activity, law enforcement can more successfully detect criminal activity when the criteria for these operations are well selected.

The earlier studies that have been conducted on the subject are reviewed and discussed in this part. Numerous subfields of artificial intelligence (AI) are used in these studies; some concentrate on applying AI to crime data, while others forecast and anticipate violent crimes using machine learning or data mining. Spatial and temporal data are also used in several of these studies.

We found five surveys or overviews about data mining or machine learning and crime prediction during our survey. The first was based on a 2011 survey that looked at different methods for mapping spatial data into patterns. Known as spatial data mining (SDM) algorithms, these methods included spatially weighted regression, spatial auto-regression, conditional auto-regression, spatial clustering, spatial hot spots, spatial outliers, and co-location mining. The survey came to the conclusion that these SDM algorithms were efficient and would undoubtedly be applied in practical situations; nevertheless, further methods were needed to verify the theories that these algorithms produced.

In 2015, there were some research talks about the fields of data mining and machine learning. This study used a number of crime-related characteristics and discovered that the prediction of crime rates is unaffected by factors that affect the crime rate, such as age, alcohol consumption, hot locations, media, specific legislation, etc. Though the conclusion is weak, the research was helpful in sparking conversation. In 2016, a second poll was published. Over 100 data mining applications in criminal activities were evaluated. They came up with a concise summary by making a short table that contained the method used with a certain piece of software, the relevant study area, and the expected application and function.

To increase the benefits, developments, and usefulness of data mining approaches in crime data mining, they suggested expanding the training and educational domains for these tactics. Data mining approaches utilized in crime prediction and data mining studies completed between 2004 and 2018 were the basis for a 2019 thorough evaluation. Based on the challenges addressed and the quantity of research papers based on the technique employed, a gap was discovered in all 40 research publications. This discrepancy shows that the system's overall performance significantly declines as dataset sizes increase.

In 2020, 32 papers on spatial crime projections from 2000 to 2018 were analyzed in a final systematic review. In addition to several summaries of the top four suggested methods, best proposed methods, and baseline methods employed in the 32 chosen papers, this study also provides a surveying table with information about the research location and time, crime data, and forecasting information. This study concluded that future algorithm continuity should not be ignored after examining the benefits, drawbacks, dangers, and opportunities of the selected papers.

Methodology

Two steps make up the approach employed in this review study: first, relevant research on crime prediction utilizing data mining and machine learning studies is gathered and analyzed; second, a classification table is created. The literature examines data mining and machine learning techniques using a variety of datasets for different cities throughout the world. Additionally, the methodologies of machine learning and data mining to crime are compared using a border crime prediction system.

The information that we have included in this table for each chosen study will help future researchers identify the most potent categories of crime prediction approaches. Define and describe crime

Generally, crimes are classified in to three groups:

1. infractions

2. Felonies

3. misdemeanors

according to the degree, severity, and penalty of the offenses. Infractions include minor offenses including speeding, tailgating, and extended parking. Felonies are regarded as the most serious crimes, whereas misdemeanours are regarded as less serious offenses.

Aside from that, crimes are classified based on the time of day, week, month, and season. Machine learning and data mining techniques are used to find trends among various crime categories and predict them in the future. By using a dataset collected on a specific place for previous crimes, it is possible to forecast future crimes. Various sorts of crimes exist, contingent on the seriousness of the offense. Crimes are therefore divided into three categories: felonies, misdemeanours and infractions.

Table 1

Crime description

Crime type	Crimes
Felony	Murder/homicide
	Burglary
	Forcible rape
	Illegal drug selling
	Robbery
	Aggravated assault, battery
	Arson
	Forgery
Misdemeanour	Larceny-theft
	Fraud
	Embezzlement
	Stolen property
	Vandalism
	Gambling
	Drunkenness
Infraction and wobblers	Overtime parking
	Speeding ticket
	Tailgating
	Weapons violation

Prediction systems challenges

When it comes to predicting the place and time of crimes, researchers and government security personnel have challenges in selecting the most efficient approach. Researchers in computer science who use data mining, machine learning, and spatial-temporal data also encounter issues. The difficulties faced by developers of crime prediction systems

a. A significant amount of storage is needed due to the enormous volume of data.

b. Data pertaining to crime typically comes in a variety of formats, including relational data, unstructured data, semi-structured data, graphs, audio, text, and photographs. Therefore, it is also difficult to convert these data into a format that can be understood.

c. It might be difficult to assign the appropriate label (prediction or output) to an instance (context or input) in machine learning.

d. Applying the proper data mining method that outperforms the ones already in use

e. The weather and the absence of the law are two examples of environmental and surrounding elements that affect the possibility of crime and ultimately lead to serious mistakes in crime prediction algorithms. To prevent making mistakes and to attain high prediction accuracy, every crime forecast must take environmental and surrounding variables into account.

Crime datasets

Data about crimes is collected from a number of sources, such as criminal records, social media, news, and police reports. Collecting this lot of data is challenging. In many nations worldwide, the datasets are accessible online or collected from police departments.

Classification of prediction systems

The process of defining a model that can explain ideas or classes of information is known as classification in machine learning. This model's objective is to forecast the class of objects with an ambiguous class label. This high volume of criminal activity is too much for police forces to handle on their own in the real world. In the meanwhile, given the sharp increase in crime, data mining must be combined with police detection efforts to forecast and subsequently lower the number of criminal cases. As previously said, in order to fix this issue as quickly as possible, technology—particularly computer science technologies—is required. In the end, a variety of characteristics can be used to categorize prediction systems.

a. Based on methodologies, data mining and machine learning. Depending on the type of forecast, particular, and temporal

b. Image prediction and data prediction based on dataset

Conclusion

These days, crime prediction is a hot area of study because of its correlation benefits to the security of any society or country. In contrast, it is shown that supervised learning methods were employed in a significant percentage of crime prediction research. When it came to the accuracy of crime prediction, data mining approaches clearly surpassed machine learning techniques. However, when it comes to predicting crimes, machine learning typically performs better than data mining. However, when comparing the standard deviation of crime prediction accuracies, the machine learning algorithms perform better than the data mining techniques.

When comparing machine learning and data mining algorithms for crime prediction systems, the following conclusions are eventually drawn: some algorithms perform flawlessly on average, but may not perform well when applied to other datasets; the type of dataset may influence the algorithm selection (e.g., image, text, video, or voice dataset). Crime prediction methods that employ deep learning algorithms were not covered in this survey due to time limits. Crime has grown more complex as technology has advanced, creating difficult problems for police enforcement. With an emphasis on finding patterns and trends in criminal acts, researchers have recently become increasingly interested in using deep learning and machine learning to forecast crime.

The use of machine learning and deep learning algorithms to predict or identify criminal behavior has shown a lot of promise in the crime detection challenge. These advancements may help increase the precision and efficacy of crime prediction models by leveraging large datasets and sophisticated algorithms. Despite the progress made in this area, little research has been done on how these technologies might be used to solve the problem of crime prediction. Consequently, our study contributes to the comprehension of the impacts of various ML and DL methodologies. Additionally, the datasets and future initiatives we described will help the current research community continue their work in the subject of crime prediction.

References

1. Shah, N., Bhagat, N. & Shah, M. Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. Visual Computing for Industry, Biomedicine, and Art 4, 1–14 (2021) .
2. Chun, S. A. et al. Crime prediction model using deep neural networks, 512–514 (2019).
3. Kshatri, S. S. et al. An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: An ensemble approach. IEEE Access 9, 67488–67500 (2021) .
4. Janiesch, C., Zschech, P. & Heinrich, K. Machine learning and deep learning. Electronic Markets 31 (3), 685–695 (2021) .
5. Raza, D. M. & Victor, D. B. Data mining and region prediction based on crime using random forest, 980–987 (IEEE, 2021).
6. Elluri, L., Mandalapu, V. & Roy, N. Developing machine learning based predictive models for smart policing, 198–204 (IEEE, 2019).
7. Meijer, A. & Wessels, M. Predictive policing: Review of benefits and drawbacks. International Journal of Public Administration 42 (12), 1031– 1039 (2019) .
8. Hossain, S., Abtahee, A., Kashem, I., Hoque, M. M. & Sarker, I. H. Crime prediction using spatio-temporal data, 277–289 (Springer, 2020).
9. Saraiva, M., Matijoˇsaitienˉe, I., Mishra, S. & Amante, A. Crime prediction and monitoring in porto, portugal, using machine learning, spatial and text analytics. ISPRS International Journal of Geo-Information 11 (7), 400 (2022) .
10. Kounadi, O., Ristea, A., Araujo, A. & Leitner, M. A systematic review on spatial crime forecasting. Crime science 9, 1–22 (2020) .
11. Morrisey, L. J. in Bibliometric and bibliographic analysis in an era of electronic scholarly communication 149–160 (Routledge, 2013).

12. Hofmann, M. & Chisholm, A. Text mining and visualization: Case studies using open-source tools Vol. 40 (CRC Press, 2016).

13. Tamilarasi, P. & Rani, R. U. Diagnosis of crime rate against women using k-fold cross validation through machine learning, 1034–1038 (IEEE, 2020).

14. Kumar, A., Verma, A., Shinde, G., Sukhdeve, Y. & Lal, N. Crime prediction using k-nearest neighboring algorithm, 1–4 (IEEE, 2020).

15. Agarwal, S., Yadav, L. & Thakur, M. K. Crime prediction based on statistical models, 1–3 (IEEE, 2018).

16. Bandekar, S. R. & Vijayalakshmi, C. Design and analysis of machine learning algorithms for the reduction of crime rates in india. Procedia Computer Science 172, 122–127 (2020) .

17. Sangani, A., Sampat, C. & Pinjarkar, V. Crime prediction and analysis (2019).

18. Goel A, Singh B. White collar crimes: A study in the context of classification, causation and preventive measures. Contemp Soc Sci. 2018;27:84–92.10.29070/27/58311Search in Google Scholar

19. Sardana D, Marwaha S, Bhatnagar R. Supervised and unsupervised machine learning methodologies for crime pattern analysis. Int J Artif Intell Appl. 2021;12(1):43–58.10.5121/ijaia.2021.12106Search in Google Scholar

20. Okeke OC. An overview of crime analysis, prevention and predicton using data mining based on real time and location data. Int J Recent Technol Eng. 2022;5(10):99–103.10.33564/IJEAST.2021.v05i10.015Search in Google Scholar

21. Sardana D, Marwaha S, Bhatnagar R. Supervised and unsupervised machine learning methodologies for crime pattern analysis. Int J Artif Intell Appl. 2021;12(1):43–58.10.5121/ijaia.2021.12106Search in Google Scholar

22. Jonas p, Paul E, Stijn V, Marc MVH, Guido D. Gaining insight in domestic violence with emergent self organizing maps. Expert Syst Appl. 2009;36(9):11864–74.10.1016/j.eswa.2009.04.027Search in Google Scholar

23. Kang HW, Kang HB. Prediction of crime occurrence from multi-modal data using deep learning. PLoS One. 2017;12(4):e0176244.10.1371/journal.pone.0176244Search in Google Scholar

24. Holst A, Bjurling B. A Bayesian parametric statistical anomaly detection method for finding trends and patterns in criminal behavior. In 2013 European Intelligence and Security Informatics Conference. IEEE; 2013.10.1109/EISIC.2013.19Search in Google Scholar