Exploring Crime Detection and Prediction Methods using Artificial Intelligence and Machine Learning

Giri Sharan Reddy Pusuluru
Master of Applied Computer Science,
Dalhousie University
gr861188@dal.ca

Abstract

This research paper presents a systematic literature review on the application of machine learning (ML) and artificial intelligence (AI) techniques in crime analysis and prediction. The study examines various research papers that utilize ML and deep learning techniques to predict crime trends and occurrence probability. The aim is to understand the effectiveness of these methods in optimizing law enforcement efforts and resources.

The study identifies several key findings. One research paper achieved an accuracy of 81.93% in predicting crime using the Adaboost algorithm and Spatio-Temporal Crime Prediction algorithm. Another study compares the performance of Naive Bayesian and Back Propagation algorithms, with Naive Bayesian outperforming in predicting crime categories. Hotspot detection approaches like Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Random Forest are considered effective in identifying crime hotspots. Time series analysis, particularly using the Auto Regressive Integrated Moving Averages (ARIMA) model, has been successful in forecasting crime patterns.

The paper proposes an integrated approach for law enforcement departments to combat crime effectively. This includes categorizing crimes, continuous monitoring of high-crime areas, tracking previous criminals, and utilizing forecasting software and models for crime prediction. By leveraging ML, AI, and data analysis techniques, law enforcement agencies can gain valuable insights into crime patterns, allocate resources more effectively, and implement preventive measures to ensure safer communities.

1. Introduction

In today's society, the top most priority of a government is the safety and security of individuals and communities but, crime rates continue to be a persistent concern, causing significant harm and distress. Despite advancements in technology, governments still grapple with the devastating impact of criminal activities that seem difficult to halt completely due to variety of crime types, motives, handling methods, and prevention techniques. Due to these complications and various attributes, crime prediction has become a powerful and widely used technique. As a result, law and order enforcement departments spend a great deal of time and resources detecting crime trends and predicting them. It is evident that attempting to predict crime manually would be difficult. However, amidst this ongoing struggle, there is a glimmer of hope. Usage of Machine Learning (ML) and Artificial Intelligence (AI) makes an innovative methods and approaches to lead and aid the process of crime predictability and criminology in eliminating this menace of crime from our society.

Harnessing the power of AI can provide a transformative approach to stopping crime.

By integrating ML, AI and leveraging advanced algorithms, we can proactively detect and prevent criminal activities. Through continuous monitoring of criminal's movements and high alert locations, AI can analyze vast amounts of data to identify patterns and anomalies associated with criminal intent. Moreover, AI systems can leverage historical criminal records and apply machine learning techniques to identify risk factors and predict the likelihood of future criminal behavior. By alerting law enforcement agencies when the probability of criminal activity surpasses a predefined threshold value, AI empowers them to take timely action, preventing crimes before they occur.

This research paper, provides an extensive systematic literature review to understand the research attempts and trends in crime analysis and prediction. Specifically, focusing on how research papers that use ML, and deep learning techniques leverage their capabilities in predicting crime trends and occurrence probability as this can help optimize law department efforts and resources.

2. Literature Review

The research papers collected in the study covered a wide range of methods and objectives in the field of crime prediction. In one particular research paper [1], the goal was to identify the key factors that influence crime by analyzing various statistical parameters. The paper focused on analyzing the category and location of crimes using machine learning models like K-nearest neighbors (KNN), Gradient Tree Boosting, and exploratory analysis techniques. The authors achieved an impressive accuracy of 81.93% using the Adaboost algorithm and Spatio-Temporal Crime Prediction algorithm [1], although it required more time for computation. The findings of the study highlighted the importance of factors such as time, hour, day, and location in predicting crime, which can aid law enforcement in targeting high-risk areas. However, the study also acknowledged the need for further investigation into the underlying causes of crime.

In paper [2] two ML classification algorithms Naive Bayesian and Back Propagation (BP) were compared in terms of their performance in predicting crime categories in terms of accuracy, precision, and recall using 10-fold cross-validation. The results demonstrate that the Naive Bayesian algorithm outperforms BP algorithm, achieving high accuracy rates of 90.2207% and 94.0822% for two distinct groups [2]. This study highlights the effectiveness of data mining techniques, specifically classification algorithms, in analyzing crime data and predicting crime categories. The findings suggest that Naive Bayesian and similar machine learning algorithms can be valuable tools for law enforcement agencies in combating crime and terrorism. Additionally, the analysis of crime records has the potential to uncover valuable insights into the social dynamics within communities [3]. This, in turn, enables government entities and decision-makers to gain a deeper understanding of specific demographic factors such as age groups, nationalities, and more, which require targeted attention for the prevention of related issues.

The research paper [4] aims to achieve two main objectives: identifying effective crime hotspot detection approaches and exploring crime prediction strategies. The study highlights that the Density-Based Spatial

Clustering of Applications with Noise (DBSCAN) and Random Forest are considered leading approaches for crime hotspot detection due to their high accuracy and efficiency. However, the paper acknowledges certain limitations, including the need for scalability, addressing sparse data, and considering underlying population and demographic factors.

In terms of crime prediction strategies, the study finds that time series analysis, particularly using the Auto Regressive Integrated Moving Averages (ARIMA) model [4], has been successful in forecasting various real-world events. However, further improvement is necessary for applying the ARIMA model to crime patterns that display seasonal and repetitive behaviors. The paper emphasizes the importance of generating new spatio-temporal datasets [4] and region-specific data to enhance the proposed approaches and facilitate the development of effective crime prevention strategies.

The paper [5] emphasizes the difficulties in predicting crimes due to limited data availability. It suggests using electronic media and data mining tools to extract useful information from crime data. The study compares two machine learning algorithms, KNN and Random Forest, to predict crimes. The results show that KNN is more accurate (92%) than Random Forest (62%) [5]. This is because KNN considers the class of the nearest neighbor samples, reducing classification errors.

The paper suggests using automated methods to remove duplicate data and advanced machine learning algorithms like reinforcement learning and deep learning to improve crime prediction. It also mentions the benefits of automatic methods for accurately identifying crime locations. By combining processed data with GIS-based software like ArcMap, important insights into crime patterns can be obtained [5].

Hotspot analysis is useful for identifying areas at high risk. KNN and Random Forest are effective in identifying areas prone to risks, assisting law enforcement in monitoring sensitive areas and maintaining security. In this method, historical crime data is overlaid onto a map, enabling law enforcement to allocate additional resources to those specific areas. However, it is important to note that this approach is just a reaction to what has happened in the past reactive rather than predictive.

The paper [6] explores the application of two forecasting models: ARIMA and Holt Linear in predicting crime rates and determine their accuracy for forecasting the Annual Crime Rate Time Series in India. The authors begin by applying the Box-Jenkins Methodology [6] to the Annual Crime Rate Time Series (ACRTS) data, conducting data analysis, testing for stationarity, model selection, and forecasting. The paper concludes that both ARIMA (212) and Holt Linear models are numerically significant for forecasting the ACRTS. However, the ARIMA (212) model may be preferred when a high level of accuracy is crucial, while the Holt Linear model can be a viable alternative when a simpler approach is desired without compromising the overall forecasting performance.

Figure 1: Crime hotspots in UK [7]

The paper [7] provides the importance of data visualization in understanding and enabling crime analysts to gain insights into crime patterns and trends based on attributes such as crime type, location, date, latitude, and longitude. Visualizing crime data on maps allows to identify dangerous areas and assists law enforcement in improving security measures as shown in Figure-1. By analyzing crime hotspots, which are areas with high crime density, law enforcement agencies can allocate resources more effectively and make informed decisions. The paper also discusses crime frequency reports generated based on the number of crimes occurring each month and different categories of crimes. These reports help the public in taking necessary safety measures and provide crime analysts with valuable information about changes in crime trends. In terms of crime prediction, apply data mining techniques and incorporate theories from criminology, such as the Rational Choice Theory and Routine Activity Theory. Machine learning algorithms, such as the KNN algorithm and Naïve Bayes classifier, are used to predict crimes based on factors such as location, date, and crime type.

Crime datasets can include attributes that pertain to two distinct aspects: the location-specific details concerning the crime occurrence and the neighborhood-related information, such as the rate of unemployment, household income, population, and so on. The paper [8] primarily focuses on exploring the relationship between crime and the economy specifically in India. It analyzes crime data involving theft, burglary, and robbery using machine learning techniques and investigates how economic indicators, specifically Gross District Domestic Product (GDDP) and the unemployment rate, impact crime rates. The ML algorithms used in the analysis include Decision Trees, Random Forest, Linear Regression, and Neural Network. The performance of these algorithms is evaluated based on correlation coefficients, coefficients of determination, absolute mean error, and accuracy. The results indicate that Linear Regression performs the best among the algorithms tested.

Furthermore, the paper explores the concept of causality between crime rates and economic indicators using Granger causality tests [8]. The findings reveal a unidirectional causation between the unemployment rate and robbery. This suggests that an increase in the unemployment rate is associated with an increase in robbery rates. Based on the results, the paper discusses possible actions to control crime by focusing on economic factors. For example, it suggests that reducing unemployment through the creation of more employment and education opportunities can help decrease crime rates. Additionally, the paper highlights the potential of incorporating economic indicators into crime forecasting software, such as Crime Mapping, Analytics, and Predictive Systems (CMAPS), used by law enforcement agencies. This integration could enhance the accuracy and effectiveness of crime prediction and prevention strategies [8].

3. Proposed Solution

In my observation, I propose that law enforcement departments should utilize previous crime data to categorize crimes into specific categories such as chain snatching, robbery, murder, kidnapping, and drug dealings. By employing ML and AI algorithms like K-nearest neighbors (KNN), they can analyze the data and determine the areas by using any visualization methods like heat maps as shown in Figure-2 where each type of crime is more prevalent. This division of crime zones based on categories would provide valuable insights for targeted policing and resource allocation.

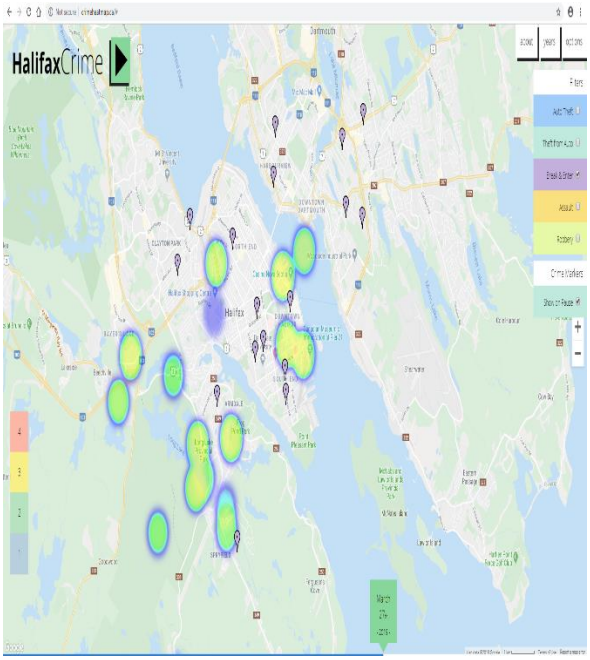

Figure 2: Heat Map for crimes in Halifax [9]

To effectively combat crime, the identified high-crime areas should be monitored continuously using various methods such as satellite surveillance and increased police patrolling. By keeping a close eye on these areas and promptly investigating any suspicious activities, law enforcement can take preventive measures to stop crimes from occurring.

Another crucial approach involves tracking previous criminals and continuously monitoring their activities and current locations. By studying the mindset of criminals and understanding the motivations behind their actions, law enforcement can gain valuable insights into the factors that lead to different types of crimes. This knowledge can aid in making predictions about the motives behind future criminal activities.

To integrate these approaches, law enforcement departments can leverage forecasting software like Crime Mapping, Analytics, and Predictive Systems (CMAPS), as well as employ models like Auto Regressive Integrated Moving Averages (ARIMA). These tools enable accurate prediction of potential crimes within specific regions and at specific intervals of time, facilitating proactive measures to prevent criminal activities. By combining the categorization of crimes, continuous monitoring of high-crime areas, tracking previous criminals, and utilizing advanced forecasting techniques, law enforcement agencies can optimize their efforts in crime prevention. This integrated approach, supported by harnessing the power of ML, AI, and data analysis techniques, enhances their ability to anticipate and prevent crimes and proactively combat criminal activities, ensuring safer communities.

4. Conclusion

In conclusion, the integrated utilization of machine learning and artificial intelligence techniques in crime analysis and prediction offers significant potential in improving public safety. By leveraging advanced algorithms and data analysis, law enforcement agencies can gain valuable

insights into crime patterns, high-risk areas, and potential criminal activities. This knowledge allows for proactive resource allocation, targeted policing, and the implementation of preventive measures to mitigate crime effectively.

5. References

[1] A. Shukla, A. Katal, S. Raghuvanshi, and S. Sharma, "Criminal combat: Crime analysis and prediction using machine learning," in 2021 International Conference on Intelligent Technologies (CONIT), 2021, pp. 1–5, doi: 10.1109/CONIT51480.2021.9498397.

[2] A. Babakura, M. N. Sulaiman, and M. A. Yusuf, "Improved method of classification algorithms for crime prediction," in 2014 International Symposium on Biometrics and Security Technologies (ISBAST), Kuala Lumpur, Malaysia, 2014, pp. 250–255, doi: 10.1109/ISBAST.2014.7013130.

[3] W. Little and W. Little, "Chapter 7. Deviance, crime, and social control," in Introduction to Sociology - 2nd Canadian Edition, BCcampus, 2016.

[4] U. M. Butt, S. Letchmunan, F. H. Hassan, M. Ali, A. Baqir, and H. H. R. Sherazi, "Spatio-temporal crime HotSpot detection and prediction: A systematic literature review," IEEE Access, vol. 8, pp. 166553–166574, 2020, doi: 10.1109/ACCESS.2020.3022808.

[5] A. Umair, M. S. Sarfraz, M. Ahmad, U. Habib, M. H. Ullah, and M. Mazzara, "Spatiotemporal analysis of web news archives for crime prediction," Appl. Sci. (Basel), vol. 10, no. 22, p. 8220, 2020, doi: 10.1007/s42979-023-01816-y

[6] M. Kumar et al., "Forecasting of annual crime rate in India: A case study," in 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 2087–2092, doi: 10.1109/ICACCI.2018.8554422.

[7] H. K. R. ToppiReddy, B. Saini, and G. Mahajan, "Crime prediction & monitoring framework based on spatial analysis," Procedia Comput. Sci., vol. 132, pp. 696–705, 2018, doi: 10.1016/j.procs.2018.05.075

[8] M. Mittal, L. M. Goyal, J. K. Sethi, and D. J. Hemanth, "Monitoring the impact of economic crisis on crime in India using machine learning," Comput. Econ., vol. 53, no. 4, pp. 1467–1485, 2019.

[9] "Halifax Regional Municipality Crime Mapping," *Canadian GIS & Geomatics*, 18-Jul-2013. [Online]. Available: https://canadiangis.com/halifax-regional-municipality-crime-mapping.php. [Accessed: 28-Jun-2023].