Forecasting Crime: A Survey on Machine Learning Approaches

Shraddha J. Pawar
Research Scholar
MET BKC Institute of Engineering
Nashik, Maharashtra
✉ shraddhap113@gmail.com

Baisa L. Gunjal
Head, Information Technology
Amrutvahini College of Engineering
Sangamner, Maharashtra
✉ hod.it@avcoe.org

ABSTRACT

Crime prediction is crucial for improving public safety and allocating resources effectively. This paper presents an overview of "Crime Forecasting: A Machine Learning Approach," highlighting key challenges and the need for accurate predictions. Various machine learning techniques, including traditional statistical models and advanced deep learning methods, are examined alongside real-world case studies and datasets to demonstrate their practical application in predicting crime in different urban settings. The paper also addresses data preprocessing, feature engineering, and model evaluation specific to crime prediction tasks, while discussing ethical concerns, biases, and privacy issues associated with machine learning in crime forecasting. Furthermore, it explores the potential of predictive policing to optimize law enforcement strategies. In summary, this survey offers a concise yet comprehensive analysis of the role of machine learning in crime forecasting, serving as a valuable resource for researchers, policymakers, and practitioners interested in understanding its successes, limitations, and future directions.

KEYWORDS : *Crime forecasting, Machine learning approach, Public safety.*

INTRODUCTION

Crime poses a pervasive social challenge impacting societal well-being and economic development [1], influencing decisions on relocation and travel routes [2]. Since thieves often target familiar places, it is no surprise that police departments are looking for innovative data mining techniques and sophisticated geographic information systems to improve crime analytics and community safety [3, 4].

The timely and precise prediction of criminal activities is imperative for contemporary society. Technological advancements, particularly in machine learning, have revolutionized crime prediction, enabling scholars, policymakers, and law enforcement agencies to anticipate and prevent crimes more effectively. By harnessing data-driven insights, machine learning algorithms have the potential to enhance crime forecasting and response capabilities significantly.

This paper offers an in-depth exploration of crime forecasting within the framework of machine learning methodologies, discussing key concepts, challenges, and opportunities. It examines a spectrum of machine learning algorithms, from traditional statistical models to state-of-the-art deep learning techniques, supported by real-world case studies and datasets showcasing their applicability in diverse urban contexts. Methodological considerations such as data preprocessing, feature engineering, and model evaluation are also addressed, along with ethical implications and potential biases inherent in machine learning applications for crime forecasting.

Furthermore, the paper explores the emerging concept of predictive policing, which leverages machine learning insights to optimize law enforcement strategies and resource allocation. In summary, this survey aims to equip researchers, policymakers, and practitioners with a comprehensive understanding of crime forecasting through machine learning. It synthesizes current knowledge, discusses successes and challenges, and outlines future directions for the field.

LITERATURE SURVEY

Numerous scholars have addressed challenges in crime control, proposing diverse crime-prediction algorithms whose accuracy hinges on attribute selection and the reference dataset. Researchers in a study cited as [1] used data from mobile networks, demographic information, and actual crime statistics to identify areas in London, UK, most likely to have high crime rates. Also in [5], using the open-source information mining program WEKA and 10-overlay cross- approval, the creators thought about Choice Tree and Innocent Bayesian classification methods. The datasets utilized for this correlation came from the 1990 US Statistics, the 1990 US LEMAS review, and the 1995 FBI UCR. Furthermore, using an 18,288 event dataset, [6] investigated road accident trends in Ethiopia using KNN, Naïve Bayesian, and Decision tree algorithms, with prediction accuracy varying between 79% and 81%.

A significant obstruction in wrongdoing expectation is the efficient and reliable analysis of big crime databases. Data mining uncovers hidden patterns within these datasets, enhancing prediction accuracy and reducing errors [7]. Most studies aim to locate “crime hotspots,” or locations with significantly higher than normal crime rates [8]. One example is the work of [8], which suggested area-specific prediction models utilizing sparse data after comparing the hotspot mapping techniques Kernel Density Estimation (KDE) and Risk Terrain Modeling (RTM). A similar spatial-temporal model was used for hotspot prediction in [9], which also made use of Linear Discriminant Analysis (LDA), KNN, and statistical approaches based on histograms. To gauge wrongdoing areas of interest in Bangladesh, a calculation that checks crime incidences trained an ANN augmented by the Gamma test [10]. analyzed Taiwanese drug-related criminal data and identified potential areas of interest utilizing an information driven AI strategy [11]. In order to forecast criminal activity in Nova Scotia, Canada, the authors of [12] used reverse-geocoding in conjunction with a density-based clustering method that made use of Open Street Map (OSM) and geographical data. Also, a Deep Neural Network (DNN) highlight level information combination approach was recommended for Chicago crime expectation in [13]. [14] examined several approaches to crime prediction and argued that KDD techniques—a combination of measurable demonstrating, AI, data set capacity, and artificial intelligence advances — were the most successful. A transfer-learning method that can capture trends across different urban datasets in terms of time and space was presented in [15]. Also, in order to model the dependence between environmental elements and crime statistics in New South Wales, Australia, [16] used a completely probabilistic calculation in view of Bayesian approaches. In order to forecast future crimes in Mississippi, researchers used WEKA to compare three different algorithms: decision stump, linear regression, and additive regression. Furthermore, a review of ANN, decision trees, rule induction, nearest-neighbor approaches, evolutionary algorithms, and criminal data mining was offered in a survey study [18]. A prediction model for urban crime trends was developed using the Auto-Regressive Integrated Moving Average (ARIMA) model [19]. One review utilized the irregular timberland strategy to assess the meaning of metropolitan markers in wrongdoing expectation in Brazil, while another review introduced a probabilistic model of the geological way of behaving of known lawbreakers [20]. [22] utilized planned approaches, the multi-bit strategy, and the Dempster-Shafer hypothesis of proof to make a wrongdoing expectation answer for Chilean huge towns. With the end goal of wrongdoing expectation in San Francisco, KNN, Parzen windows, and Brain Organizations were made and evaluated in [23]. [24] used the weighted page-rank way to deal with debilitate and disturb criminal organizations and Gradient Boosting Machine (GBM) to uncover hidden linkages in such networks.

Karie et al. [25] proposed a system to improve cyber forensics using deep learning algorithms; this system is called the DLCF framework. Initialization, data source identification, investigation with deep learning assistance, reporting, decision-making, and closure are the various layers that make up this framework. Deep learning is used to establish relationships between entities, which helps with incident analysis, pattern extraction, and improved predictive capabilities.

Bolger and Bolger [26] conducted a comprehensive overview on feeling of dread toward wrongdoing in a humble community, exploring this phenomenon by analyzing individual demographic factors, community-level effects, and their combined influence. Their analysis focused on two models: the vulnerability model, which emphasizes demographic factors' susceptibility to fear of crime, and the incivilities model, which explains the impact of neighborhood issue on dread, taking into account social confusion and disappointment with policing as contributing factors.

Mittal et al. [27] examined the relationship between India's economy and crime rates using various calculations, including choice trees, irregular woodlands, linear regression, and neural networks. The study found that linear regression produced the most accurate results, confirming a correlation between unemployment and robbery, when comparing subordinate factors like burglary, thievery, and theft to autonomous factors like Gross District Domestic Product (GDDP) and the joblessness rate.

Kadar and Pletikosa [28] gathered information from different sources, including as registration records, wrongdoing measurements, area based informal organizations, metro and taxi excursions, and that's only the tip of the iceberg, to make two kinds of crime forecast models: one that focuses on rates of crime over a longer period of time (one to five years), and another that is shorter.

Okutan et al. [29] utilized unconventional signals from a variety of sources, such as GDELT, OTX, and Twitter data, to forecast cyber-attacks. Used techniques such as prescient sign attribution (PSI), collecting signals with huge slacks (ASL), and SMOTE++ for imbalanced information,successfully tackling the problem of cyber-attack prediction.

Kim et al. [30] introduced a novel approach combining Gaussian mixture model (GMM) and convolutional neural networks (CNNs) for object detection, using GMM for background subtraction and object extraction, followed by CNN classification within the region of interest (ROI), demonstrating high accuracy, particularly in identifying small, distant moving objects.

Vomfell et al. [31] used the social disorganization theory to examine crime statistics, taking into account variables like population, POIs, taxi traffic, and resident tweets; analyzed crime from a social and structural viewpoint; used machine learning techniques, spatial linear regression, and Poisson generalized linear models to confirm correlations between crime and spatial factors; and helped with crime prediction and the development of policing strategies.

Seo et al. [32] introduced a partially generative neural network (PGNN) architecture to improve data classification and prediction by generating missing values effectively.Zhao and Tang [33] proposed a region-transfer approach, where a model trained in one region is transferred and tested in another region with common characteristics.

Mohd et al. [34] explored feature selection methods for crime data, focusing on factors such as race, pay class, age, family structure, instruction, populace, area, and joblessness rate, identifying relevant features for crime prediction as a complex process. The use of the Crime Anticipation System (CAS) by the Amsterdam Police Department to identify crime hotspots was introduced by Hardyns et al. [35] as an integrative approach to policing that integrates strategic systems and smart preplanned policing. The system effectively processes extensive data attributes using a blend of strategic relapse and brain organizations.

Vural and Gok [36] evaluated the capabilities of decision trees and Naive Bayes classifiers in crime prediction, with decision trees using a tree-like model to represent data attributes and Naive Bayes employing the Bayes hypothesis to evaluate the likelihood of events based on evidence.

Kouziokas [38] highlighted the scaled form slope calculation as a fast and time-compelling learning method, particularly for optimizing neural network models used in categorizing crime hotspots based on spatial data. Russell [39] emphasized the importance of evaluating predictive modeling effectiveness in terms of validity, equity, reliability, and usefulness, often measured through the receiver operating characteristic (ROC) curve..

DISCUSSION AND COMPARATIVE ANALYSIS

Criminology, among the oldest academic disciplines globally, has evolved in tandem with human civilization. As crimes have become more sophisticated over time, they present increasingly intricate challenges to society. In response, societal structures such as law enforcement and legal frameworks have been developed, relying on the analysis of historical crime episodes. The consideration of wrongdoing highlights in the grouping and forecast process holds paramount importance, as these features are intricately linked to socio-economic, cultural, and demographic factors. Specifically, socio-economic factors exert a substantial influence on criminal activity, their impact changing in light of the social construction and pecking order of a given society. Table 1 outlines various crime prediction methodologies: Spatial linear regression, Poisson GLMM, and CAR offer predictive capabilities tailored to urban areas but may be susceptible to changes. Behavioral-based techniques display potential. Transfer learning predicts crimes within specific blocks but is sensitive to layering. Filter methods and algorithms like Naive Bayes demonstrate high accuracy in feature selection but are narrow in scope. Logistic regression, neural networks, and ensemble models facilitate grid-level crime prediction but have inherent limitations.

Table 1. Comparative Analysis of State of art system

Methodology	Pros	Cons
Spatial linear regression, Poisson GLMM, simultaneous autoregressive (CAR) model[31]	Utilizes behavioral-based crime prediction technique	Urbanized prediction model may be sensitive to changes
Logistic regression, SVM, decision tree, GVM, NN, PGNN[32]	Produces accurate values for missing features	Performance of PGNNs not assessed with other forms of crime
Transfer learning[33]	Predicts crimes in specific administrative blocks	Layer changes significantly affect results

As shown in Table 2, classification methods are examined. Naïve Bayes and Decision Tree are effective in classifying crimes. Artificial Neural Networks and K-Nearest Neighbor excel, with effectiveness dependent on the dataset. Real crime data from South Korea varies in predictive accuracy. Comparative studies reveal JRip as the best-performing technique. Ensembles like AdaBoost and stacked models offer improved accuracy. Bagging ensemble models, including random forest, prove effective for enhancing decision tree classifier performance.

Table 2. Classification algorithm analysis

Methodology and Tools	Datasets	Result Analysis
Naïve Bayes, Decision Tree [40]	UCI machine learning repository dataset "Crime and Communities" featuring 128 attributes and 1994 observations.	Compared to Naïve Bayes, Decision Tree outperforms it in predicting crime categories across several US states, with an accuracy of 83.95%.
J48, Neural Net, SVM, KNN, Naïve Bayes [41]	Iris, Liver Disorder, E-coil datasets with varying attributes and instances.	Artificial Neural Networks and K-Nearest Neighbor exhibit superior performance across datasets; algorithm effectiveness varies based on the dataset used for classification.
Neural Network, Decision Tree, Support Vector Machine, k-NN, and Naïve Bayes [42]	Crime statistics derived from South Korea	Comparison of different algorithms on crime data for predictive accuracy.

The analysis provided in Table 2 offers valuable insights into the performance of diverse classification algorithms across various datasets and methodologies. It's evident that algorithmic efficacy is strongly influenced by dataset characteristics and problem domain specifics. For example, Decision Tree consistently demonstrates robust predictive accuracy, especially in domains like crime prediction and accident analysis. However, algorithms like Neural Networks and K-Nearest Neighbor excel in specific contexts, such as datasets like Iris, Liver Disorder, and E-coil. Ensemble methods, including Bagging, Boosting, Stacking, and Random Forest, emerge as potent strategies for enhancing classification performa often outperforming individual classifiers, particularly in complex or diverse data settings. Among them, Bagging stands out for its robustness, notably when applied to decision tree-based models like J48 and CART.

CONCLUSION

In summary, this comprehensive exploration of machine learning methodologies for crime prediction underscores significant advancements achieved through data-driven approaches, aiming to enhance public safety and optimize resource allocation. The methodologies discussed encompass a spectrum from traditional statistical models to cutting-edge deep learning algorithms, showcasing the versatility of machine learning in this domain. Moreover, the paper highlights the ethical considerations and biases inherent in these methodologies, emphasizing the critical need for fair and transparent crime prediction systems.

Moving forward, it is imperative for future research efforts to prioritize refining fairness and transparency aspects within machine learning models. Furthermore, exploring innovative data sources such as social media and Internet of Things (IoT) devices holds promise for enhancing predictive capabilities. Additionally, fostering interdisciplinary collaboration among researchers, law enforcement agencies, and policymakers is essential to fully harnessing the potential of machine learning for fostering safer communities.

REFERENCES

1. A. Bogomolov, B. Lepri, J. Staiano, N. Oliver, F. Pianesi, and A. Pentland, "Once upon a crime: towards crime prediction from demographics and mobile data," Proc. of the 16th Intl. Conf. on Multimodal Interaction, pp. 427-434, 2014.
2. H. Adel, M. Salheen, and R. Mahmoud, "Crime in relation to urban design. Case study: the greater Cairo region," Ain Shams Eng. J., vol. 7, no. 3, pp. 925-938, 2016.
3. "Overall crime rate in Vancouver went down in 2017, VPD says," CBC News, Feb. 15, 2018. [Online] Available: https://www.cbc.ca/news/canada/british-columbia/crime-rate-vancouver2017-1.4537831. [Accessed: 09- Aug- 2018].
4. J. Kerr, "Vancouver police go high tech to predict and prevent crime before it happens," Vancouver Courier, July 23, 2017. [Online] Available: https://www.vancourier.com/news/vancouver-police-go-high-tech-topredict-and-prevent-crime-before-it-happens-1.21295288. [Accessed: 09- Aug- 2018]
5. R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. Shariat Panahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," Indian J. of Sci. and Technol., vol. 6, no. 3, pp. 4219-4225, Mar. 2013.
6. T. Beshah and S. Hill, "Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia," Proc. of Artificial Intell. for Develop. (AID 2010), pp. 14-19, 2010.
7. H. Chen, W. Chung, J. J. Xu, G. Wang, Y. Qin, and M. Chau, "Crime data mining: a general framework and some examples," IEEE Computer, vol. 37, no. 4, pp. 50-56, Apr. 2004.
8. M. Al Boni and M. S. Gerber, "Area-specific crime prediction models," 15th IEEE Intl. Conf. on Mach. Learn. and Appl., Anaheim, CA, USA, Dec. 2016.
9. Q. Zhang, P. Yuan, Q. Zhou, and Z. Yang, "Mixed spatial-temporal characteristics based crime hot spots prediction," IEEE 20th Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD), Nanchang, China, May 2016.
10. N. Mahmud, K. Ibn Zinnah, Y. Ar Rahman, and N. Ahmed, "CRIMECAST: a crime prediction and strategy direction service," IEEE 19th Intl. Conf. on Comput. and Inform. Technol., Dhaka, Bangladesh, Dec. 2016.
11. Y. L. Lin, L. C. Yu, and T. Y. Chen, "Using machine learning to assist crime prevention," IEEE 6th Intl. Congr. on Advanced Appl. Inform. (IIAI[1]AAI), Hamamatsu, Japan, Jul. 2017.
12. F. K. Bappee, A. S. Júnior, and S. Matwin, "Predicting crime using spatial features," Can. AI 2018: Advances in Artificial Intel.-Lecture Notes in Comput. Sci., vol. 10832, pp. 367-373, Springer, Mar. 2018.
13. H. W. Kang, H. B. Kang, "Prediction of crime occurrence from multi[1]modal data using deep learning," PLoS ONE, vol. 12, no. 4, Apr. 2017.
14. V. Grover, R. Adderley, and M. Bramer, "Review of current crime prediction techniques," Intl. Conf. on

Innovative Techn. and Appl. of Artificial Intel., pp. 233-237, Springer, London, 2007.

15. X. Zhao and J. Tang, "Exploring transfer learning for crime prediction," IEEE Intl. Conf. on Data Mining Workshop (ICDMW), New Orleans, LA, USA, Nov. 2017.

16. R. Marchant, S. Haan, G. Clancey, and S. Cripps, "Applying machine learning to criminology: semi-parametric spatial-demographic Bayesian regression," Security Inform., vol. 7, no. 1, Dec. 2018.

17. L. McClendon and N. Meghanathan, "Using machine learning algorithms to analyze crime data," Mach. Learn. and Appl.: an Intl. J. (MLAIJ), vol.2, no.1, Mar. 2015.

18. S. Prabakaran and S. Mitra, "Survey of analysis of crime detection techniques using data mining and machine learning," Nat. Conf. on Math. Techn. and its Appl. (NCMTA 2018), IOP J. of Physics: Conf. Series, vol. 1000, 2018.

19. E. Cesario, C. Catlett, and D. Talia, "Forecasting crimes using autoregressive models," IEEE 14th Intl. Conf. on Dependable, Auton. and Secure Comput., Auckland, New Zealand, Aug. 2016.

20. M. A. Tayebi, U. Glässer, and P. L. Brantingham, "Learning where to inspect: location learning for crime prediction," IEEE Intl. Conf. on Intel. and Security Inform. (ISI), Baltimore, MD, USA, May 2015,.

21. L. G. A. Alves, H. V. Ribeiro, and F. A. Rodrigues, "Crime prediction through urban metrics and statistical learning," Physica A, vol. 505, pp. 435-443, 2018.

22. N. Baloian, E. Bassaletti, M. Fernández, O. Figueroa, P. Fuentes. R. Manasevich, M. Orchard. S. Peñafiel, J. A. Pino, and M. Vergara, "Crime prediction using patterns and context," IEEE 21st Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD), Wellington, New Zealand, Apr. 2017.

23. M. V. Barnadas, Machine learning applied to crime prediction, Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, Sep. 2016.

24. E. Budur, S. Lee, and V. S. Kong, "Structural analysis of criminal network and predicting hidden links using machine learning," arXiv:1507.05739, Sep. 2015.

25. Karie NM, Kebande VR, Venter HS (2019) Diverging deep learning cognitive computing techniques into cyber forensics. Forens Sci Int Syn 1:61–67

26. Bolger MA, Colin Bolger P (2019) Predicting fear of crime: results from a community survey of a small city. Am J Crim Just. 44(2):334–351

27. Mittal M, Goyal LM, Sethi JK, Hemanth DJ (2019) Monitoring the impact of economic crisis on crime in India using machine learning. Comput Econ 53(4):1467–1485

28. Kadar C, Pletikosa I (2018) Mining large-scale human mobility data for long-term crime prediction. EPJ Data Sci 7(1):26

29. Okutan A, Werner G, Yang SJ, McConky K (2018) Forecasting cyber-attacks with incomplete, imbalanced, and insignificant data. Cyber Secur 1(1):15 Big Data 5:22

30. Kim C, Lee J, Han T, Kim YM (2018) A hybrid framework combining background subtraction and deep neural networks for rapid person detection.

31. Vomfella L, Hardle WK, Lessmann S (2018) Improving crime count forecasts using twitter and taxi data. Dec Supp Syst 113:73–85

32. Seo S, Chan H, Brantingham PJ, Leap J, Vayanos P, Tambe M, Liu Y (2018) Partially generative neural networks for gang crime classification with partial information. In: Proceedings of the 2018 AAAI/ACM conference on AI, ethics, and society, pp 257–263

33. Zhao X, Tang J (2017) Exploring transfer learning for crime prediction. In: IEEE international conference on data mining workshops, pp 1158–1159

34. Jalil MA, Mohd F, Noor NMMZ (2017) A comparative study to evaluate filtering methods for crime data feature selection. Proced Comput Sci 116:113–120

35. Rummens A, Hardyns W, Pauwels L (2017) The use of predictive analysis in spatiotemporal crime forecasting. Appl Geogr 86:255–261

36. Vural MS, Gok M (2017) Criminal prediction using Naive Bayes theory. Neur Comput Appl 28(9):2581–2592

37. Zhuang Y, Almeida M, Morabito M, Ding W (2017) Crime hot spot forecasting: a recur- rent model with spatial and temporal information. In: IEEE, International Conference on Big Knowledge, pp 143–150

38. Kouziokas GN (2016) The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. Transp Res Proced 24:467–473

39. ussell J (2015) Predictive analytics and child protection: constraints and opportunities. Child Abuse Negl 46:182–189

40. Sathyadevan, Shiju, and Surya Gangadharan. “Crime analysis and prediction using data mining.” Networks & Soft Computing (ICNSC), 2014 First International Conference on. IEEE, 2014.

41. Bogomolov A, Lepri B, Staiano J, Oliver N, Pianesi F, Pentland A. Once upon a crime: towards crime prediction from demographics and mobile data. 2014 Nov 12 (pp. 427-434). ACM.

42. Zhang Q, Yuan P, Zhou Q, Yang Z. Mixed spatial-temporal characteristics based Crime Hot Spots Prediction. InComputer Supported Cooperative Work in Design (CSCWD, 2016 May 4 (pp. 97-1