

A Deep Learning Approach for Real-Time Crime Recognition and Prevention

¹Pooja Tiwari, ²Gayatry Sharma, ³Prakhar Sharma, ⁴Shradha Dobal

[*poojatiwari9557@gmail.com](mailto:poojatiwari9557@gmail.com), gayatrysharma31@gmail.com,

prakharsharma48@gmail.com, shraddhadobal@gmail.com

^{1,4}Assistant Professor, Amrapali University Lamachaur, Haldwani, 263139, India

Department of Computer Applications

²Research Scholar Uttarakhand Open University, ³Research Scholar Amrapali University

Abstract

Crime prediction has developed as one of the most significant applications of data science and machine learning in the field of public safety and law enforcement. With this rapid increase in urbanization and population growth, those crime rates have been increased in many regions which makes it difficult for law enforcement agencies to monitor, analyze and prevent criminal activities within the range on limited time. Traditional crime analysis methods based on manual investigation and historical observations which are consuming often time and after that not that much accurate. To overcome from these limitations, this study proposed a machine learning-based crime prediction model which is capable of analyzing historical crime records and forecasting future crime occurrences while improving accuracy and efficiency of the model. The primary objective of our research is used to develop a predictive system that can assist law enforcement agencies to take a proactive decision-making, efficient resource allocation and crime prevention strategies. The proposed system applies historical crime datasets which contains various attributes such as crime type, location and other relevant factors which will influence criminal activities.

Keywords: Machine Learning, Crime Prediction, Deep Learning, Artificial Intelligence, Crime Forecasting, Smart Policing, Pattern Recognition, Feature Selection

1. Introduction

Introduction

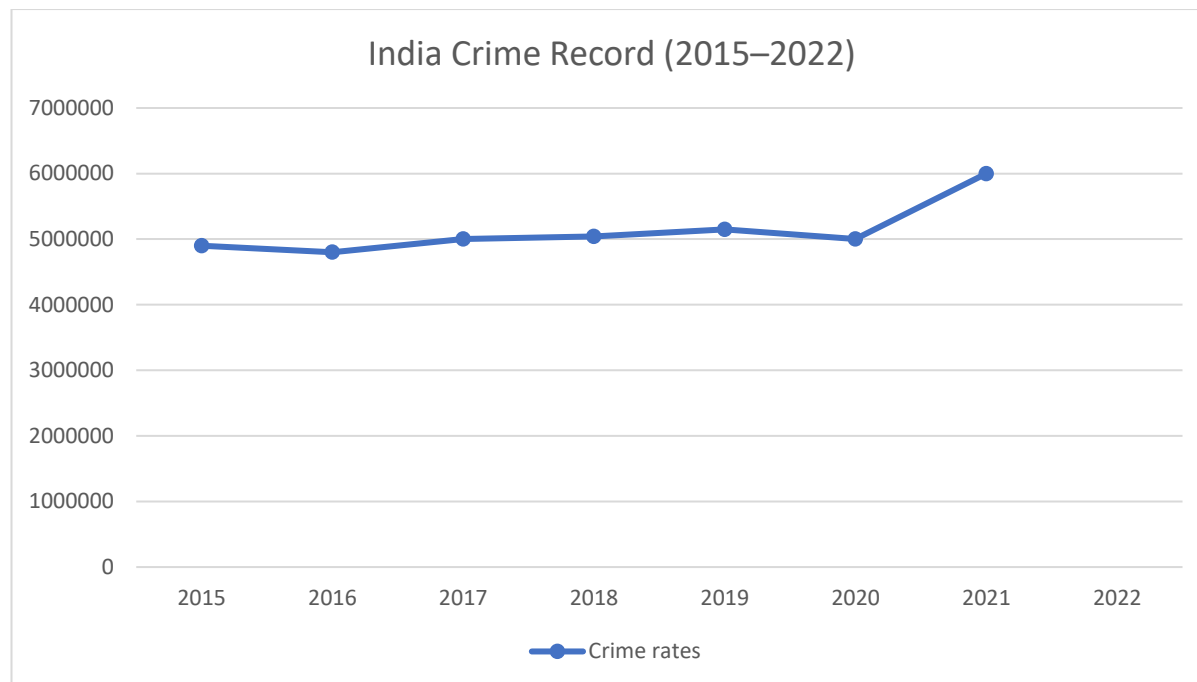
Crime is a major and serious problem which affects safety and security of our society. Due to the population increment and urbanization, crime rates are also rising in many places. Traditional methods which are using for crime analysis completely depend on manual investigation and past records which they had and this is really time-consuming and may not always give accurate predictions.

With the help of new technologies which occur and establishing day by day are really helpful to analyze the large amount of crime data and behalf of that data it identify and predict the actual result which we really need. These techniques will help us to finalize or predicting the upcoming crimes based on their factors like labels and confidences and suggest the actual crime. This will support all crime departments such as police for taking actions and implement multiple actions for preventing future crimes and improving public safety.

However, all the existing software or system have some limitations for finding an actual result and predict the actual output. They had lack of real time data usage and really difficult to handle complex pattern in crime data. These traditional approaches are not that much affected and sometime fail to provide efficient and reliable prediction. The related study of crime has completely evolving them and significantly over the last two centuries. Early researchers such as Adolphe Quetelet and André-Michel Guerry who focused on analyze crime patterns of frames at larger regional levels such as departments and countries. With time, research shifted toward smaller and more detailed regional units based including towns, neighbourhoods and individual street phases. But modern crime analysis how to focus on micro-level crime hotspots to understand better criminal behavior and improve crime prevention strategies. This technique transition toward best regional analysis which greatly supported the development of advanced crime prediction systems and machine learning-based crime mapping techniques [1].

While overcoming from these issues, that study proposing a best prediction model while using machine learning techniques or algorithms which will use to find the actual result in real time data base such as Logistic Regression,

Decision tree and Random Forest. This model aim is very clear that it is designed for improving prediction based on the dataset and used to help in best decision making for preventing crimes.



2. Literature Review

Lots of research paper worked on real time crime prediction based on multiple algorithm which are follows:

T. M. MITCHELL, (1997): The paper “Machine Learning” which explains the basic fundamental concepts and techniques of machine learning which are used to predict, analysis and classify crime problems. The study focused on algorithms such as Decision Trees, Logistic Regression, and probabilistic learning methods which can analyze past datasets to identify hidden patterns and make accurate predictions behalf of them. The paper highlights importance for training data, feature selection and model evaluation which helps to improve prediction, accuracy and developing intelligent decision-making systems.

VARUN MANDALAPU, LAVANYA ELLURI, (Member, IEEE), PIYUSH VYAS, NIRMALYA ROY (2023): The paper ” *Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions*” focused on machine learning and deep learning techniques which will demonstrated good job in crime prediction, issues with complete real-time implementation, data availability and legal considerations which still exist. The authors claim more accurate and useful systems for forecasting crime which require standardized datasets, interdisciplinary collaboration, and more context-aware models.

JIANMING ZHOU, ZHENG LI, JACK J. MA, FEIFENG JIANG (2020): The paper “*Exploration of the Hidden Influential Factors on Crime Activities: A Big Data Approach*” explains big data methodology for crime rate evaluation was done in order to identify the features, by joining the Gradient Boost Decision Tree (GBDT), Recursive Feature Elimination (RFE) and Geographic Information System (GIS) techniques.

PRIYANKA DAS, (Student Member, IEEE), ASIT KUMAR DAS, JANMENJOY NAYAK, DANILO PELUSI, AND WEIPING DING, (Senior Member, IEEE) (2019): The paper “*A Graph Based Clustering Approach for Relation Extraction From Crime Data*” proposed a novel cluster technique in order to predict the pattern of crime. In their work they propose a method for using criminological data to mine relations from the newspaper.

Authors & Year	Paper Title	Short Description
T. M. Mitchell (1997)	Machine Learning	Explains basic machine learning concepts and algorithms
Varun Mandalapu, Lavanya Elluri, Piyush Vyas, Nirmalya Roy (2023)	Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions	Focus on machine learning and deep learning methods for crime prediction.
Jianming Zhou, Zheng Li, Jack J. Ma, Feifeng Jiang (2020)	Exploration of the Hidden Influential Factors on Crime Activities: A Big Data Approach	Worked on big data techniques such as GBDT, RFE, and GIS to find hidden factors.
Priyanka Das, Asit Kumar Das, Janmenjoy Nayak, Danilo Pelusi, Weiping Ding (2019)	A Graph Based Clustering Approach for Relation Extraction From Crime Data	Worked on a graph-based clustering technique to extract.

5. Methodology

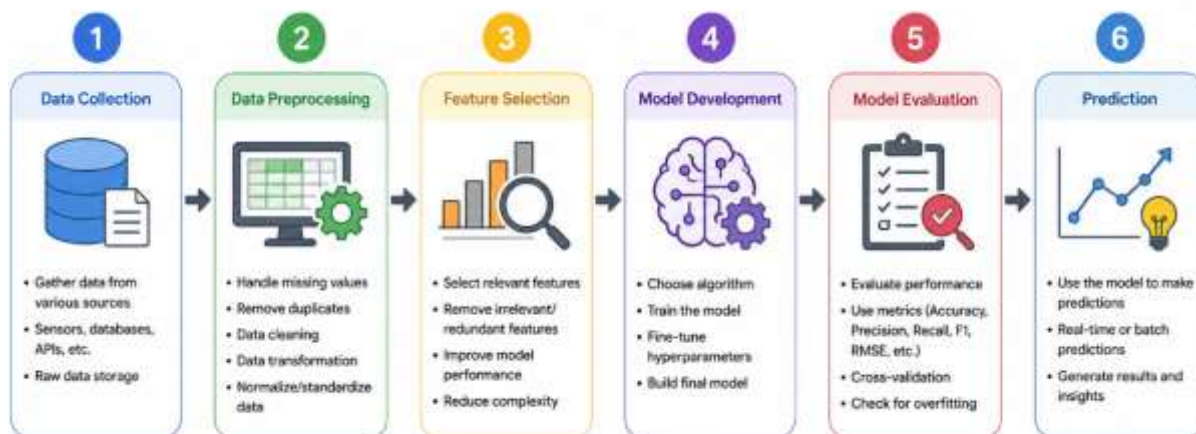
5.1 Introduction to Methodology

This study completely focuses on enhancing and developing a best system while using machine learning techniques for predicting crime. This methodology is used to design a complete structured way so that all crime frames patterns will be analyze effectively and also predicting future crime occurrences with better accuracy. This entire process has multiple stages including data collection, preprocessing, feature extraction, model building, evaluation and last but not the least prediction.

The goal of our proposed system is not only to build a model but also build to ensure that this model will be reliable, accurate and application in all real-world scenarios.

This methodology follows a simple and complete pipeline approach:

Data Collection → Data Preprocessing → Feature Selection → Model Development → Model Evaluation → Prediction



We will learn each step on by one by understanding all kind of features on each step which will determine the final prediction of proposed system.

5.2 Data Collection

This is our first step for this research, collecting crime-related data from different source such as social media(youtube, Instagram and so on), National Crime Records Bureau and publicly available datasets and so on which will be reliable

and authentic sources. That dataset is used for the study of model which contains historical crime records and really very essentials for identifying the patterns and trends.

The quality of the dataset will directly affects the performance of our model. Therefore, selecting a well-structured and complete dataset is really very important.

5.3 Data Preprocessing

After collecting the dataset from upper layer, preprocessing is used to improve data quality. Raw data may contain missing values, duplicate entries and inconsistent records that can affect model prediction accuracy. So, In this process first data cleaning techniques are use to remove errors and handle missing information. Categorical data is converted into numerical format using encoding techniques and after that normalization is performed to prepare the dataset for machine learning algorithms.

The preprocessing steps include:

- Data Cleaning
- Data Transformation
- Data Integration
- Data Reduction

5.4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis is used to study about crime patterns and trends within the dataset. Statistical analysis and visualization methods such as graphs, charts, and heat maps are used to identify the actual crime frequency based on their location, time, and crime type. This phase helps in understanding the major factors which are responsible for criminal activities and identifying crime related areas.

This analysis includes:

- Crime frequency by location
- Crime occurrence based on time and date
- Identification of high-crime areas
- Correlation between features

5.5 Feature Extraction

Feature selection is used to identify the most important variables which are affecting crime prediction rate. Selecting relevant attributes will improves the efficiency and accuracy of the models. It extract the feature regarding the crime and suggest real time prediction based on the feature extraction.

5.6 Model Development

Model development is one of the important phase in a machine learning-based crime prediction system. In this stage, A suitable machine learning pattern selected and also trained using the used dataset to identify hidden crime patterns and generate accurate predictions. Lots of algorithms such as Logistic Regression, Decision Tree and Random Forest used and it's rely on the complexity and nature of the data. The dataset genuinely divided into training and testing sets where training set should have 70% data and testing should have 30% data. So that the model will learn from past crime records and later predict real-time data. After this stage, the model parameters are adjusted and generate an optimized solution to improve prediction accuracy and minimize errors. Fine-tuning techniques and hyperparameter optimization are also used to achieve better performance.

5.7 Model Evaluation

Model evaluation is the next process of finding performance and effectiveness of our developed machine learning model. After trained our model, it is used to test data to identity how accurately it predicts crime occurrences. Major evaluation metrics such as Accuracy, Precision, Recall, F1-Score etc are used to analyze the model's performance. Accuracy measures the complete correctness of predictions, while Precision and Recall helps to evaluate the model's ability to correctly identify crime-related events.

5.8 Prediction

This is the last phase of the overall crime prediction system, where future criminal incidents are predicted using our trained and assessed model. To produce its predictions about potential criminal activity, this machine learning model examines real-time input data which includes location, crime history and other. Depending on the needs of our system predictions can build in real-time surveillance camera. The generated data assist law enforcement organizations for identifying high-risk regions, effectively allocating police resources and preventing prior criminal activity.

6. Conclusion

This study presented a model which is use to predict road accidents using machine learning techniques. It analyzed historical accident data to forecast the likelihood of future incidents. The research identified key factors that significantly impact accidents, including weather conditions, road type, traffic volume, time of day, and location.

Various machine learning algorithms were used, including Logistic Regression, Decision Tree, and Random Forest. These algorithms were compared to see how effective they were. Among them, Random Forest showed the highest accuracy and better ability to generalize and makes it the best choice for predicting road accidents. The findings suggest that machine learning can spot hidden patterns in accident data and offer valuable insights for improving road safety. The proposed model can help traffic authorities and urban planners find accidental areas, optimize traffic management strategies, and implement preventive measures like better road designs, improved signage and stricter law enforcement. However, the model's performance depends on the quality and completeness of the data. Limitations like missing data, lack of real-time information, and unpredictable human behavior can affect the accuracy of the predictions.

In conclusion, the study shows that systems based on machine learning can significantly lower accident rates and improve public safety. Future enhancements, such as integrating real-time traffic and weather data, can further boost the reliability and effectiveness of the system.

Additionally, the scope of our project model can be extended while extracting also road accident prediction which include the detection and prediction of different public safety incidents such as crimes, violent activities, fires and explosions by using advanced machine learning techniques like deep learning and video analysis techniques. With the growth of smart cities and surveillance technologies, A huge amounts of data are continuously generated through real time CCTV cameras, drones, traffic monitoring systems and social media platforms. By integrating these data sources with intelligent algorithms that proposed system can become a complete public safety monitoring and prediction framework.

For crime detection and prediction, machine learning models are analyzing historical crime records with geographical patterns, population density, time-based trends and social media factors to identify high-risk areas and predict the probability of criminal activities within the range. Algorithms such as Random Forest, Support Vector Machine and Neural Networks can help law enforcement agencies deploy police resources more effectively and improve crime prevention strategies with the range. Real-time surveillance systems combined with computer vision techniques that also detect suspicious behavior of that object, unauthorized access, theft or violent activities in crowded places such as railway stations, malls, schools and public events.

The system can also use to detect violent activities such as fighting, gunshot, explosion and any crime through video analysis. Deep learning models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks can analyze real time video streams which is used to recognize abnormal human behavior, aggressive movements or unusual crowd activities. Once suspicious actions are detected than automated alerts can be sent to nearby authorities for quick response and emergency handling with the specific time. Such systems can significantly reduce response time and help prevent crime rates. By using image preprocessing and sensor-based data this proposed model can identify smokes, flames, sudden flashes or any abnormal temperature changes in real time world. AI-enabled surveillance systems will provide too early warnings before a fire spreads or an explosion causes which will majorly effect. These technologies can be highly beneficial in industries, public buildings, shopping malls, fuel stations and transportation systems where the risk of accidents is high. Integration of IoT devices and smart sensors with our proposed model can further improve the accuracy and speed of detection.

References

- [1] D. Weisburd, W. Bernasco, and G. J. N. Bruinsma, *Putting Crime in its Place: Units of Analysis in Geographic Criminology*. New York, NY, USA: Springer, 2009.
- [2] Machine Learning, T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [3] Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2019.
- [4] K. B. Kale, V. D. Shinde, and S. T. Patil, "Crime Prediction Using Machine Learning and Data Mining," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 8, no. 4, pp. 234–239, 2019.
- [5] N. S. R. Prasad, P. R. Kumar, and B. Ramesh, "Crime Prediction Using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, pp. 222–227, 2019.
- [6] S. Wang, W. Ding, H. Chen, and C. Yang, "Crime Prediction Using Data Mining Techniques," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 2017, pp. 626–631.
- [7] 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," in *Proceedings of the 16th International Conference on Multimodal Interaction*, 2014, pp. 427–434.
- [8] H. Yu, Z. Yang, and R. O. Sinnott, "Predicting Crime Using Spatial Features," *Procedia Computer Science*, vol. 129, pp. 383–388, 2018.
- [9] M. Kadar, A. Pletzer, and J. Golbeck, "Predicting Crime Using Twitter and Kernel Density Estimation," *Decision Support Systems*, vol. 61, pp. 115–125, 2014.
- [10] Y. Huang, L. Zhang, and P. Zhang, "Crime Forecasting Using Machine Learning and Deep Learning Techniques," *IEEE Access*, vol. 8, pp. 193746–193757, 2020.
- [11] Crime Modeling and Mapping Using Geospatial Technologies, M. Leitner, *Crime Modeling and Mapping Using Geospatial Technologies*. Dordrecht, Netherlands: Springer, 2013.
- [12] Machine Learning, T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [13] Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2019.
- [14] K. B. Kale, V. D. Shinde, and S. T. Patil, "Crime Prediction Using Machine Learning and Data Mining," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 8, no. 4, pp. 234–239, 2019.
- [15] N. S. R. Prasad, P. R. Kumar, and B. Ramesh, "Crime Prediction Using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, pp. 222–227, 2019.