IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 2, February 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

# Advanced Crime Prediction using Machine Learning with Crime Forecasting and Categorization

**S. Akash, A.P. Harrishsomesh, S. Shaik Safiudeen, Mrs.M. Viji**

UG Student, Dept. of CSBS., E.G.S Pillay Engineering College, Nagapattinam, Tamil Nadu, India

UG Student, Dept. of CSBS., E.G.S Pillay Engineering College, Nagapattinam, Tamil Nadu, India

UG Student, Dept. of CSBS., E.G.S Pillay Engineering College, Nagapattinam, Tamil Nadu, India

Assistant Professor, Dept. of CSBS., E.G.S Pillay Engineering College, Nagapattinam, Tamil Nadu, India

**ABSTRACT**: The "Advanced Crime Prediction Using Machine Learning with Crime Forecasting and Categorization," is a comprehensive project developed in Python that employs Machine Learning algorithms, specifically the Decision Tree Classifier and Bagging Classifier, to predict and classify various crime categories in Portland, Oregon, USA, from the years 2015 to 2023. The dataset utilized for this project consists of 505,063 data points, with a focus on 20 distinct crime classes, including 'Larceny Offenses,' 'Motor Vehicle Theft,' 'Assault Offenses,' 'Drug/Narcotic Offenses,' and others. The Decision Tree Classifier yielded impressive results, achieving a 98% accuracy on the training set and a 95% accuracy on the test set. Similarly, the Bagging Classifier demonstrated robust performance, achieving a 98% accuracy on the training set and maintaining a 95% accuracy on the test set. These high accuracies indicate the effectiveness of the machine learning models in predicting and classifying crimes.

The dataset encompasses 15 features, including address, case number, crime against category (Person, Property, or Society), neighbourhood, occur date, occur time, offense category, offense type, open data latitude/longitude, open data X/Y, and offense count. These features provide a comprehensive and diverse set of information, enabling the models to make accurate predictions. The project's significance lies in its potential application for law enforcement agencies, city planners, and policymakers. By accurately predicting and classifying crimes, it facilitates proactive decision-making and resource allocation, contributing to the enhancement of public safety and the efficient utilization of law enforcement resources.

The classification of crimes into specific categories allows for a more nuanced understanding of crime patterns, enabling stakeholders to implement targeted interventions and preventive measures. The high accuracy of the models demonstrates their reliability and effectiveness in handling real-world crime prediction scenarios. In summary, "Crime Prediction Using Machine Learning" presents a robust and accurate approach to crime prediction, leveraging advanced algorithms and a rich dataset. The project's success in accurately classifying diverse crime categories makes it a valuable tool for enhancing public safety and optimizing law enforcement strategies in urban environments.

## I. INTRODUCTION

In this project we will explore a Python project entitled &quot Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Direction. The project is based on a research paper published in IEEE 2023 Journal.

The paper focuses on the increasing crimes all over the world, especially in cities, and the need for scientific approaches to categorize and predict crimes.

The proposed system, named "Advanced Crime Prediction Using Machine Learning with Crime Forecasting and Categorization," signifies an evolution from the base paper's concept, incorporating advanced machine learning models and a more recent dataset to achieve higher accuracy in crime prediction. Given the growing accessibility of crime data

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

and advancements in technology, researchers are presented with a distinctive opportunity to delve into the realm of crime detection through the application of machine learning and deep learning methodologies.

This introductory section lays the groundwork for an exploration of the current landscape of machine learning and deep learning in crime prediction, shedding light on their roles in detecting criminal activities, forecasting crime patterns, and preventing unlawful acts.

The primary objective of this review is to furnish a comprehensive overview of recent advancements in crime prediction utilizing machine learning techniques, offering valuable insights for researchers in the field and assisting law enforcement agencies in refining their strategies for crime prevention and response.

Through the utilization of state-of-the-art technologies and the analysis of crime data, researchers aspire to refine crime prediction methodologies and contribute to the evolution of intelligent policing strategies.

The introduction highlights the importance of understanding crime prediction techniques to improve law enforcement responses and ultimately create safer communities.

## II. RELATED WORK

Crime prediction is a crucial aspect of law enforcement, aiming to prevent and mitigate criminal activities. To enhance crime prediction capabilities, institutions need to upgrade their data and analytics capabilities, enabling the integration of new technologies such as Machine Learning into their systems.

The "related work" section in academic papers, plays a crucial role by critically reviewing and synthesizing existing literature and research pertinent to the subject under investigation. This section is designed to offer context, lay the groundwork for the current study, and showcase the advancements, methodologies, and discoveries from prior research in the field.

This involves summarizing and deliberating on past research, studies, and publications concerning crime prediction utilizing machine learning by contextualizing the current study within the broader research landscape, this review demonstrates a comprehensive understanding of the existing knowledge base.

### 2.1 Methodologies and Approaches:

Here, the section delineates the diverse methodologies, algorithms, and approaches employed in earlier research on crime prediction. It may delve into the strengths, limitations, and efficacy of various machine learning and deep learning models in forecasting criminal activity.

### 2.2 Datasets and Data Sources:

Researchers often provide insights into the datasets and data sources utilized in prior studies on crime prediction. This includes details on dataset characteristics, data preprocessing methods, and challenges encountered when working with real-world crime data.

### 2.3 Comparative Analysis:

The "related work" section may incorporate a comparative analysis of different studies, highlighting similarities, differences, and advancements in methodologies. This analysis serves to pinpoint trends, identify gaps, and suggest areas for further exploration in the domain of crime prediction using machine learning.

### 2.4 Findings and Contributions:

Researchers summarize the key findings, contributions, and insights from previous studies within the "related work" section. This summary aids in establishing the significance of the current study, building upon existing knowledge, and pinpointing avenues for future research.

Moreover, the automation brought about by big data analytics and ML systems in the criminal justice domain raises important ethical and legal questions. For example, a study on the automation in the criminal justice domain highlights the need to consider issues of privacy, human rights, and the potential impact on algorithmic discrimination. In light of these challenges, it is crucial for institutions to carefully assess the benefits and risks associated with the use of ML in crime prediction.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

Additionally, it is important for governments and law enforcement agencies to enhance public awareness and education about high-tech crimes, industry revolutions, and ML. Furthermore, there is a need for ethical considerations and safeguards to ensure the responsible use of ML in crime prediction. Based on the sources provided, it is evident that crime prediction using machine learning and artificial intelligence has gained significant attention in recent years. Researchers and law enforcement agencies are exploring the potential benefits of using ML in crime prevention and management. However, the implementation of ML tools for crime prediction must be carefully considered to mitigate risks such as biases and false alerts.

The use of machine learning and artificial intelligence in crime prediction has gained significant attention in recent years. Researchers and law enforcement agencies are exploring the benefits of using AI to improve crime prevention and management. However, there are important ethical and legal considerations that need to be addressed. Some of the key considerations include privacy concerns, algorithmic discrimination, and the potential for false alerts or biases in identification.

## III.PROPOSED SYSTEM

The proposed system, titled "Advanced Crime Prediction Using Machine Learning with Crime Forecasting and Categorization," represents an evolution from the base paper's concept.

This system aims to revolutionize crime prediction methodologies by integrating advanced machine learning models and leveraging a recent and extensive crime dataset from Portland, USA. Spanning from 2015 to 2023, the dataset comprises 5,063 records, enabling a thorough analysis of crime data and predictions across a diverse range of crime categories.

### 3.1 Machine Learning Models:

The system adopts Decision Tree Classifier and Bagging Classifier models, showcasing impressive training accuracy scores of 98% and test scores of 95%.

### 3.2 Dataset:

Utilizing a dataset from Portland, USA, the system benefits from more current data compared to the datasets utilized in the base paper.

This dataset encompasses 20 distinct crime classes, including offenses like larceny, motor vehicle theft, assault, drug offenses, weapon law violations, and vandalism.

### 3.3 Accuracy Enhancement:

The primary goal of the proposed system is to elevate the overall accuracy of crime prediction beyond the capabilities of the existing system highlighted in the base paper.

Through these enhancements, the proposed system aims to equip law enforcement agencies with precise crime predictions and valuable insights into crime patterns, thereby facilitating more effective crime prevention and control measures

By analyzing various data sources such as crime records, demographics, socio-economic factors, and historical patterns, the system can identify potential hotspots and predict areas with high crime rates. Additionally, the system could consider factors such as time of day, seasonality, and weather conditions, as they may also impact criminal activities.

By using predictive analytics, the system can generate real-time crime forecasts and provide law enforcement agencies with actionable insights to allocate resources strategically. Through the implementation of this proposed system for crime prediction, law enforcement agencies can proactively intervene and prevent crimes before they occur.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

This system can revolutionize crime prevention and help law enforcement agencies in making informed decisions regarding resource allocation, patrolling, and investigation. However, it is essential to address potential challenges and risks associated with the implementation of such a system.

These challenges include concerns about privacy and data protection, as well as the potential for biases and unfair treatment. These concerns must be thoroughly addressed through transparency in algorithm design, regular audits, and ethical guidelines to ensure that the system is fair, unbiased, and respects individuals' rights.

## IV.THEORATICAL BACKGROUND

The theoretical background in the context of crime prediction using machine learning and deep learning techniques encompasses the foundational principles, concepts, and frameworks that underpin the development and application of predictive models in the field of criminology and law enforcement. Understanding the theoretical underpinnings is essential for researchers and practitioners to effectively leverage advanced technologies for crime detection and prevention. Here are some key theoretical aspects relevant to crime prediction:

### 4.1 Criminological Theories:

Criminological theories provide insights into the causes and motivations behind criminal behavior. Theoretical frameworks such as rational choice theory, routine activity theory, social disorganization theory, and strain theory help researchers understand why crimes occur in specific contexts. By integrating criminological theories into predictive modeling, researchers can identify risk factors and patterns associated with criminal activities.

### 4.2 Machine Learning Algorithms:

Machine learning algorithms form the basis of predictive modeling in crime prediction. These algorithms, including decision trees, random forests, support vector machines, and neural networks, learn patterns and relationships from historical crime data to make predictions about future criminal incidents. Understanding the theoretical foundations of these algorithms is crucial for selecting the most appropriate model for a given crime prediction task.

### 4.3 Deep Learning Architectures:

Deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are designed to process complex data types like images, text, and audio.

The theoretical background of deep learning involves understanding how neural networks learn hierarchical representations of data and extract features that are relevant for crime prediction tasks. Knowledge of deep learning principles is essential for building accurate and efficient crime prediction models.

### 4.4 Spatial Analysis:

Spatial analysis theories and techniques play a significant role in crime prediction, especially when considering the geographic distribution of criminal activities. Spatial theories like spatial autocorrelation, hotspot analysis, and spatial clustering help researchers identify spatial patterns of crime and allocate resources effectively for crime prevention strategies. Integrating spatial analysis theories with machine learning models enhances the spatial-temporal prediction capabilities in crime detection.

### 4.5 Ethical and Legal Considerations:

Theoretical frameworks related to ethics, privacy, and legal implications in crime prediction are essential for ensuring responsible and unbiased use of predictive models.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

Understanding the ethical considerations surrounding data collection, model interpretation, and decision-making processes is crucial for deploying crime prediction systems in real-world scenarios while upholding fairness and transparency.

By incorporating these theoretical foundations into the development and evaluation of crime prediction models, researchers can enhance the accuracy, interpretability, and ethical implications of predictive systems aimed at improving public safety and crime prevention efforts.

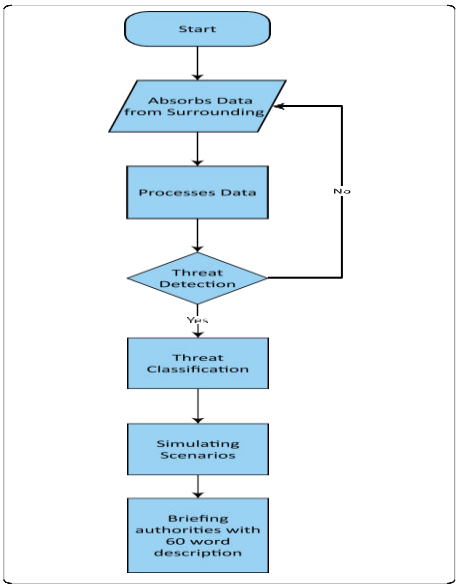


**Fig.1.Data flow diagram**

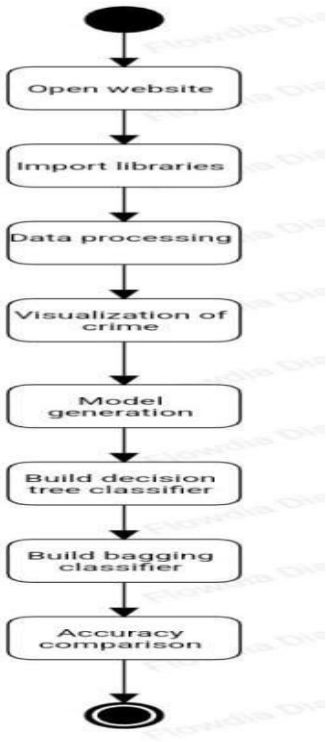


**Fig.2.Flow chart**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

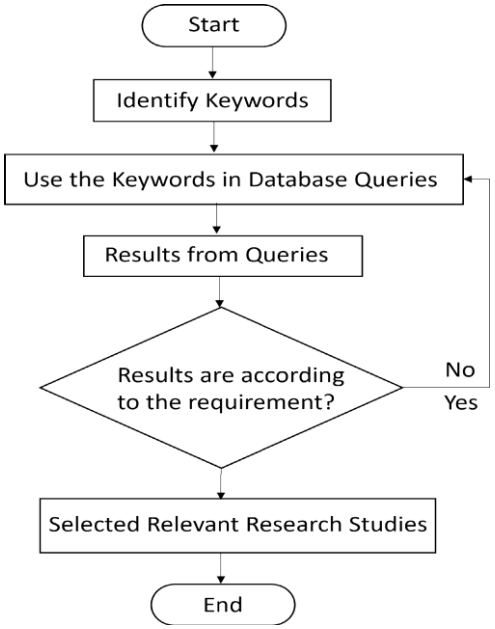


**Fig.3.Use case diagram**

## V. FUTURE ENHANCEMENTS

### 5.1 Ethical Considerations:

Further exploration is warranted into the ethical implications surrounding the utilization of machine learning and deep learning for crime prediction. This includes mitigating biases, addressing discrimination, safeguarding privacy, and preventing the misuse of personal data.

### 5.2 Real-World Validation:

There is a pressing need for comprehensive studies that assess the practical effectiveness and accuracy of machine learning and deep learning models in predicting criminal behaviour. Rigorous evaluations in real-world scenarios can validate the reliability of these predictive technologies.

### 5.3 Scalability Challenges:

Research efforts should focus on understanding and overcoming the scalability challenges associated with implementing machine learning and deep learning solutions in large-scale systems. This understanding is crucial for the successful deployment of these models in practical settings.

### 5.4 Interpretability:

Developing interpretable models is essential to elucidate the decision-making processes of machine learning and deep learning algorithms in crime prediction. Transparent models can foster trust and accountability in the context of predictive policing strategies.

### 5.5 Data Utilization:

Leveraging diverse datasets to enhance the early detection of neighbourhood crimes is a promising avenue for future research. Accessing and effectively utilizing relevant data can significantly improve the accuracy and efficiency of crime detection models.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202058 |

### 5.6 Exploration of New Techniques:

Investigating novel methodologies such as reinforcement learning approaches, visual feature analysis for crime detection, and other innovative techniques can propel advancements in the field of crime prediction using machine learning and deep learning.

## VI. CONCLUSION

In conclusion, the project "Advanced Crime Prediction Using Machine Learning with Crime Forecasting and Categorization," introduces an innovative approach to crime prediction and categorization. By employing advanced machine learning models like the Decision Tree Classifier and Bagging Classifier, the project achieved notable training accuracy of 98% and a test score of 95%. The utilization of a recent dataset from Portland, USA, comprising 5,063 records spanning from 2015 to 2023, facilitated a thorough analysis and prediction of diverse crime categories.

In comparison to the existing system mentioned in the base paper, the proposed system offers significant improvements, including the adoption of contemporary machine learning models, a more up-to-date dataset, and an expanded range of crime classes. The user-friendly interface enables users to input crime cases and receive precise offense type predictions, while static charts provide insights into prevalent crime types and occurrence percentages. Overall, this project contributes to the advancement of crime prediction and has the potential to support law enforcement agencies in combating crime effectively.

the outlined project execution steps, users can experience firsthand the efficacy of machine learning and deep learning in predicting and categorizing crimes, leading to a more informed and proactive approach to crime prevention and law enforcement

## REFERENCES

[1] Lin, Y., Chen, T., Yu, L.: Using machine learning to assist crime prevention. In: 2017 sixth IIAI International Congress on Advanced Applied Science (IIAI-AAI) (2017)
[2] Munasinghe, M., Perera, H., Udeshini, S., Weerasinghe, R.: Machine learning based criminal short listing using modus operandi features (2015). https://doi.org/10.1109/icter.2015.7377669
[3] Chauhan, C., Sehgal, S.: A review: crime analysis exploitation data processing techniques and algorithms, pp. 21–25 (2017). https://doi.org/10.1109/ccaa.2017.8229823
[4] Anon: [online] Available at https://www.researchgate.net/publication/322000460_A_review_Crime_analysis_using_data_mining_techniques_and_algorithms. Accessed 30 Aug. 2019 (2019)
[5] Kerr, J.: 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/vancou
ver-police-go-high-tech-topredict-and-prevent-crime-before-it-happens-1.21295288. Accessed 09 Aug 2018
[6] Liang, W., Wang, Y., Tao, H. & Cao, J. Towards hour-level crime prediction: A neural attentive framework with spatial–temporal-categorical fusion. Neurocomputing 486, 286–297 (2022).
[7] Navalgund, U. V. & Priyadharshini, K. Crime intention detection system using deep learning, 1–6 (IEEE, 2018).
[8] Nakib, M., Khan, R. T., Hasan, M. S. & Uddin, J. Crime scene prediction by detecting threatening objects using convolutional neural network, 1–4 (IEEE, 2018).
[9] Chackravarthy, S., Schmitt, S. & Yang, L. Intelligent crime anomaly detection in smart cities using deep learning, 399–404 (IEEE, 2018).
[10] Yadav, R. & Sheeran, S. K. Crime prediction using auto regression techniques for time series data, 1–5 (IEEE, 2018).
[11] 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).
[12] Chun, S. A. et al. Crime prediction model using deep neural networks, 512–514 (2019).
[13] 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).
[14] Janisch, C., Zschech, P. & Heinrich, K. Machine learning and deep learning. Electronic Markets 31 (3), 685–695 (2021).
[15] Raza, D. M. & Victor, D. B. Data mining and region prediction based on crime using random forest, 980–987 (IEEE, 2021).

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details