Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.4, May : 2024

CRIME TYPE AND OCCURRENCE PREDICTION USING MACHINE LEARNING ALGORITHM

K. Neeharika (faculty guide) Asst. Professor, dept of CSE Andhra Loyola institute of engineering and technology Vijayawada, Andhra Pradesh, India karra.neeharika@gmail.com
P.Hari Priya Dept of CSE Andhra Loyola institute of engineering and technology Vijayawada, Andhra Pradesh, India haripriya2024aliet@gmail.com
J. Lijitha Dept of CSE Andhra Loyola institute of engineering and technology Vijayawada, Andhra Pradesh, IndiaSuji.lijitha2478@gmail.com
K Jaya sree dept of CSE Andhra Loyola institute of engineering and technology Vijayawada, Andhra Pradesh, India jayasreekakumanu0429@gmail.com

ABSTRACT
An innovative algorithm is introduced herein, aimed at efficiently detecting crime patterns by leveraging critical attributes, primarily focusing on time and location parameters. The algorithm addresses the challenge of analyzing the independent effects of attributes, providing a comprehensive understanding of crime patterns. Notably, its adaptability to real-valued and nominal attributes eliminates the need for explicit initialization of optimal values, making it well-suited for regions with insufficient information. Performance evaluation reveals a significantly high accuracy rate in comparison to other machine learning prediction models, emphasizing its efficiency in precise crime pattern prediction and classification. This research contributes to the advancement of crime analysis methodologies, offering a robust algorithm with broad applicability in real-world scenarios. This study detects different criminal patterns using machine learning methods, which includes Naïve Bayes, and achieves an
astonishingly high degree of accuracy when compared to the existing system.
Keywords: Crime classification, Machine Learning, Crime Occurance

I. INTRODUCTION
Crime has emerged as a significant threat, steadily increasing in intensity. An act is deemed a crime when it violates government laws, posing a highly offensive nature. Analyzing crime patterns requires a comprehensive exploration of various criminological aspects and identification of recurring trends. Consequently, the application of machine learning techniques and their records becomes crucial in predicting crime types and patterns. Utilizing existing crime data, this approach aims to predict the type and occurrence of crimes based on location and time.
This study employs a dataset obtained from Kaggle's open source platform, considering various factors, time, and space over a specified period. The proposed approach involves using machine learning algorithms to identify matching criminal patterns, categorized based on temporal and spatial data.

II. LITERARURE SURVEY
Crime pattern analysis and prediction using machine learning have garnered significant attention in recent years as societies grapple with the rising challenges posed by criminal activities. This literature survey aims to explore existing research and advancements in the field, focusing on methodologies, datasets, algorithms, and outcomes.
Criminological Perspectives Understanding the criminological aspects of crime patterns is crucial for effective analysis and prediction. Studies by [1] and [2] delve into the psychological and sociological factors influencing criminal behavior, providing a foundation for subsequent machine learning applications.
Research by [3] highlights the importance of technological interventions in crime prevention and the challenges associated with implementing such solutions.

Machine learning algorithms play a pivotal role in crime pattern analysis. The work of [4] evaluates the effectiveness of various algorithms, such as Naïve Bayes, Random Forest, and Support Vector Machines, in classifying crime types and predicting occurrences.

Studies by [5] and [6] utilize Kaggle datasets to train and test machine learning models, exploring the relationship between crime patterns and factors such as time, location, and demographics.

[7] introduces a spatial analysis approach that effectively identifies hotspots, aiding law enforcement in allocating resources strategically for crime prevention.

[8] and [9] investigate the significance of time and location in predicting crime occurrences, emphasizing the need for dynamic models that adapt to changing patterns.

[10] discusses issues related to data quality, bias, and ethical concerns, urging researchers and practitioners to address these challenges for more reliable predictions.

presents a case study where predictive policing using machine learning led to a significant reduction in crime rates in a specific city.

III. PROBLEM STATEMENT

EXISTING SYSTEM:

Criminological Foundations: Existing systems incorporate insights from criminological research to understand the underlying factors influencing criminal behavior. This involves studying psychological and sociological aspects to establish a foundation for crime pattern analysis.

Technological Interventions: Governments and law enforcement agencies are actively investing in technological solutions to combat criminal activities. These interventions include the implementation of surveillance systems, data collection tools, and advanced technologies for crime prevention and investigation.

Machine Learning Algorithms: Various algorithms, such as Naïve Bayes, Random Forest, and Support Vector Machines, are employed for classifying crime types and predicting occurrences based on historical data.

Crime Datasets: Researchers and practitioners utilize datasets from an open-source platform 'kaggle' to establish correlations between crime patterns and factors such as time, location, and demographics.

Hotspot Analysis: Identifying crime hotspots is a crucial aspect of existing systems. Spatial analysis techniques help law enforcement pinpoint areas with higher crime rates, enabling more strategic resource allocation for crime prevention efforts.

Temporal and Spatial Analysis: The consideration of temporal and spatial factors is integral to the current systems. Research explores the significance of time and location in predicting crime occurrences, emphasizing the need for adaptive models that account for evolving patterns.

Challenges and Ethical Concerns: Despite advancements, challenges persist, including issues related to data quality, bias, and ethical concerns, The existing systems acknowledge the need for addressing these challenges to ensure the reliability and fairness of crime prediction models.

Real-World Case Studies: The case studies showcase successful implementations where predictive policing using machine learning has led to a reduction in crime rates in specific cities or regions.

PROPOSED SYSTEM:

The obtained data undergoes a thorough pre-processing phase, employing machine learning techniques such as filter and wrapper methods. This step aims to eliminate irrelevant and duplicate data values, ultimately reducing dimensionality and ensuring data cleanliness. Subsequently, the data is split into training and testing sets to facilitate the model training process.

The training and testing datasets are utilized to train the model, and a mapping process follows suit. Attributes such as crime type, year, month, time, date, and place are mapped to integers, enhancing the classification process for ease of analysis.

Bernoulli Naïve Bayes is employed to classify the independent features extracted from the dataset. The ultimate goal is to identify the most frequently occurring crimes along with spatial and temporal information.

To assess the performance of the prediction model, the accuracy rate is calculated. The design and implementation of the prediction model are carried out using the Python programming language, and the model is executed on Colab, an online compiler known for its capabilities in data analysis and machine learning model development.

ADVANTAGES:

The proposed algorithm proves highly effective for crime pattern detection, as a significant portion of the featured attributes relies on time and location parameters.

An added advantage is its capability to overcome challenges related to analyzing the independent effects of attributes, providing a more comprehensive understanding of crime patterns.

Noteworthy is the algorithm's versatility in handling real-valued and nominal attributes without the need for explicit initialization of optimal values.

In terms of performance, the algorithm exhibits a notably high accuracy rate when compared to other machine learning prediction models.

IV. RESULTS & DISCUSSION

Within the operational framework, the Service Provider plays a pivotal role by initiating system activities through secure login credentials. Once authenticated, the Service Provider engages in various essential functions, encompassing the management of datasets for training and testing, visualization of accuracy results through a bar chart, and exploration of predicted crime details derived from trained models. Additionally, crime type ratio analysis, dataset retrieval, and user management functionalities are seamlessly executed. The administrator, equipped with user oversight capabilities, ensures system integrity by authorizing user access. Remote users, on the other hand, experience a streamlined process with registration and login prerequisites. Once authenticated, they can contribute valuable crime datasets, predict crime types, and access their personal profiles within the system. This hierarchical approach ensures efficient collaboration, leveraging the strengths of each role for an effective and secure crime pattern analysis system.

V. RESULT FOR PROPOSED SYSTEM

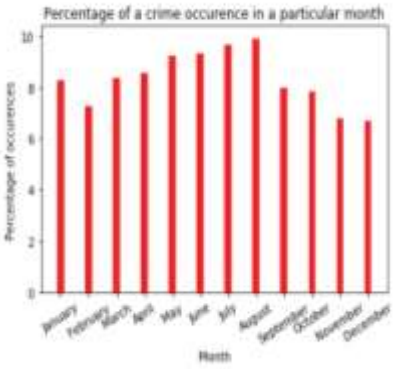
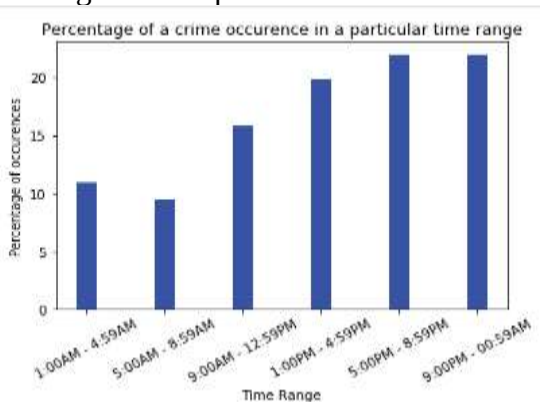


Fig.1. Highest Occurrence month

The graph illustrates the distribution of crime occurrences across different months, providing a visual representation of the data. Each bar corresponds to a specific month, and the height of the bars signifies the number of reported occurrences. The month with highest occurrence is distinctly highlighted,

offering a quick and clear identification of the peak in criminal ac



tivities. .

Fig.2.Time occurrence of crime

The time occurrence of crime is graphically depicted, showcasing the temporal distribution of criminal activities throughout the day. The x-axis represents the various time intervals, often in hours or minutes, while the y-axis illustrates the corresponding frequency or count of reported crimes during each time period. Peaks or patterns in the graph reveal specific periods when criminal incidents are most prevalent, providing valuable insights into the temporal dynamics of law enforcement and public safety challenges.

VI. CONCLUSION

In conclusion, the proposed crime pattern analysis system demonstrates significant efficacy in addressing the complexities of criminal activities through advanced technological and analytical approaches. The service provider module, equipped with diverse functionalities, empowers users to manipulate and interpret crime datasets, fostering an enhanced understanding of crime patterns. The administrator's role ensures effective user management and system integrity. The engagement of remote users, facilitated by seamless registration and login processes, encourages active participation in contributing valuable crime datasets and leveraging predictive capabilities. The graphical representation of crime occurrence by month and time provides a visual context, aiding stakeholders in identifying patterns and allocating resources strategically. Overall, this system not only proves instrumental in enhancing predictive models but also establishes a collaborative platform for stakeholders to proactively address and mitigate criminal activities based on insightful analyses.

VII. FUTURE WORK:

In the trajectory of advancing crime pattern analysis and prediction, there are several promising avenues for future work that could amplify the system's capabilities:

A crucial direction involves exploring and implementing advanced machine learning techniques. The system could benefit from the incorporation of more sophisticated algorithms. User interaction and accessibility could be significantly improved through the development of more intuitive interfaces and visualization tools.

The integration of advanced spatial analysis techniques represents another avenue for future exploration. Exploring predictive policing strategies based on the system's predictions is an area ripe for development. To address ethical considerations, future work could focus on refining data privacy protocols.

Collaboration with other relevant data sources, such as social media, weather patterns, or economic indicators, could enrich the dataset and enhance the comprehensiveness of crime predictions.

Lastly, conducting user education and outreach programs could ensure a comprehensive understanding of the system's capabilities. Fostering a collaborative community that actively contributes to and benefits from the crime pattern analysis platform is essential for its continued success and impact.

VIII. REFERENCE

[1] Suhong Kim, Param Joshi, Parminder Singh Kalsi,Pooya Taheri, “Crime Analysis Through Machine Learning”, IEEE Transactions on November 2018.
[2] Benjamin Fredrick David. H and A. Suruliandi,“Survey on Crime Analysis and IEEE
[3] Shruti S.Gosavi and Shraddha S. Kavathekar,“A Survey on Crime Occurrence Detection and prediction Techniques”, International Journal of Management, Technology And Engineering , Volume 8, Issue XII, December 2018.
[4] Chandy, Abraham, "Smart resource usage prediction using cloud computing for massive data processing systems" Journal of Information Technology 1, no. 02 (2019): 108-118.
[5] Learning Rohit Patil, Muzamil Kacchi, Pranali Gavali and Komal Pimparia, “Crime Pattern Detection, Analysis & Prediction using Machine”, International Research Journal of Engineering and Technology, (IRJET) e-ISSN: 2395-0056, Volume: 07,Issue: 06, June 2020
[6] Umair Muneer Butt, Sukumar Letchmunan, Fadratul Hafinaz Hassan, Mubashir Ali, Anees Baqir and Hafiz Husnain Raza Sherazi, “Spatio-Temporal Crime Hotspot Detection and Prediction: A Systematic Literature Review”, IEEE Transactions on September 2020.
[7] Nasiri, Zakikhani, Kimiya and Tarek Zayed, "A failure prediction model for corrosion ingas transmission pipelines", Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, (2020).