iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** X **Month of publication:** October 2023

DOI: https://doi.org/10.22214/ijraset.2023.56227

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Spatio-Temporal Crime Predictions

Mrs Meera Sawalakar[1], Nilesh Rajput[2], Sakshi Patil[3], Tanvi Panhale[4], Shraddha Jadhav[5], Akash Pargavakar[6]
B.E. in Computer Engineering, JSPM Narhe Technical Campus, Narhe, Pune

Abstract: ***Crime and transgressions pose a formidable challenge to the principles of justice and require vigilant control. Precise crime prediction and forecasting of future trends hold the potential to significantly bolster urban safety through computational means. The inherent limitations of human capacity to process intricate information from vast datasets impede our ability to achieve early and precise crime prognostication. The accurate estimation of crime rates, categories, and focal points based on historical patterns presents a plethora of computational challenges and opportunities. Notwithstanding substantial research endeavors, there persists a pressing need for an enhanced predictive algorithm that can effectively guide law enforcement patrols in response to criminal activities.***

Previous scholarly investigations have fallen short in attaining the desired precision in crime forecasting and prediction by utilizing various machine learning algorithms, including logistic regression, support vector machine (SVM), Naïve Bayes, k-nearest neighbors (KNN), decision trees, multilayer perceptrons (MLP), random forests, and extreme Gradient Boosting (xgBoost). In addition to these machine learning methodologies, this study also harnessed time series analysis, particularly the long-short term memory (LSTM) and autoregressive integrated moving average (ARIMA) models, to more aptly model crime data.

The performance of LSTM in time series analysis exhibited a reasonably satisfactory level of accuracy, as evident from the magnitude of root mean square error (RMSE) and mean absolute error (MAE) on both datasets. A comprehensive exploration of the data unveiled the presence of over 35 distinct crime types and indicated an annual decline in the crime rate in Chicago, coupled with a marginal upturn in the crime rate in Los Angeles. Significantly, there was a lower incidence of reported crimes in February compared to other months. Projections suggest a moderate future increase in Chicago's overall crime rate, with a subsequent probable decline in the years ahead. In contrast, the ARIMA model implies a sharp decrease in the crime rate in Los Angeles. Furthermore, the study's crime forecasting results pinpointed specific high-crime regions in both cities.

I. INTRODUCTION

Criminality represents a pervasive global challenge, transcending borders and affecting nations both developed and underdeveloped. The deleterious impact of criminal activities extends beyond legal infractions; it can undermine economies and compromise the well-being and quality of life for residents, giving rise to a spectrum of social and societal issues. Moreover, criminal acts impose substantial costs on both the public and private sectors. The quest for public safety remains a paramount concern, particularly in the context of travel or relocation to new locales. The realm of criminality encompasses diverse offenses, each yielding distinct consequences. These offenses manifest under a complex interplay of factors, including underlying motives, human behavior, critical circumstances, and socioeconomic disparities such as poverty.

Furthermore, a host of socio-economic factors, including unemployment, gender inequality, high population density, child labor, and illiteracy, has been found to correlate with heightened rates of violent crimes. Notably, densely populated urban centers often exhibit a pronounced association with elevated crime rates, manifesting across a spectrum of environments, including commercial districts and residential areas. In essence, the sustainability of a community hinges on its ability to minimize criminality, enabling residents to live peacefully and actively. In stark contrast, societies plagued by corruption and insecurity struggle to realize both social and economic prosperity.

The critical importance of analyzing crime reports and statistics becomes evident when one considers the imperatives of enhancing safety, security, and sustainable development ,This study is an endeavor to enhance the accuracy of crime prediction compared to previous approaches, utilizing a variety of machine learning algorithms. Beyond evaluating crime prediction accuracy, this study also employs LSTM for time series analysis, employing different performance metrics. Furthermore, an exploratory data analysis is conducted to provide a visual overview of crime types and counts. Given that crime data often takes the form of time series data, exhibiting distinct temporal patterns and seasonality, the potential significance of crime activities evolving over time becomes apparent. Thus, time series analysis emerges as an essential tool for uncovering these temporal patterns.

©IJRASET: All Rights are Reserved

In this context, the Long Short-Term Memory (LSTM) algorithm is recognized for its ability to classify crimes over time effectively based on comprehensive historical data. Additionally, forecasting crime trends via the Autoregressive Integrated Moving Average (ARIMA) model is of paramount importance.

II. OBJECTIVES

1) Enhance Crime Prediction Accuracy: Develop and implement advanced machine learning algorithms to significantly improve the accuracy of crime prediction.
2) Utilize Historical Data: Leverage comprehensive historical crime data to accurately estimate crime rates, identify crime types, and pinpoint crime hotspots.
3) Comparative Analysis: Systematically compare various machine learning methods for their effectiveness in predicting crimes, with a focus on multiple cities, including underutilized datasets from Los Angeles and Chicago.
4) Temporal Pattern Recognition: Utilize time series analysis, particularly the Long Short-Term Memory (LSTM) algorithm, to classify and analyze crimes over time, taking into account temporal patterns and seasonality in crime data.
5) Forecasting Future Trends: Implement the Autoregressive Integrated Moving Average (ARIMA) model to forecast future crime trends, including crime rates and density areas for the next five years.
6) Enhance Public Safety: Contribute to the improvement of public safety and security by providing accurate crime predictions, aiding law enforcement in proactive crime prevention measures.
7) Exploratory Data Analysis: Conduct exploratory data analysis to visually present crime types and counts, offering valuable insights for law enforcement and policymakers.
8) Societal and Economic Impact: Address the socioeconomic impact of crime by promoting the sustainability of communities through crime reduction.
9) Facilitate Law Enforcement Practices: Develop predictive algorithms that can guide police patrols effectively, optimizing resource allocation for crime prevention.

III. LITERATURE SURVEY

1) Economic and Ecological Factors: Some studies emphasize the role of economic and ecological factors, such as education, income levels, and unemployment, in influencing criminal activities. These factors serve as significant determinants of crime rates, shedding light on the underlying causes of criminal behavior.
2) Spatial-Temporal Crime Events: Research has also delved into the realm of spatial-temporal crime events, exploring the geographical and temporal patterns of criminal activities. This approach provides a comprehensive understanding of when and where crimes are more likely to occur.
3) Integration of New Data Sources: The literature highlights the integration of new and unconventional data sources for crime prediction. This includes data extracted from online forums like Twitter and mobile phone data, offering novel insights into the dynamics of criminal activities in the digital age.
4) Cause and Consequence Analysis: Many studies focus on understanding the causes of crimes and their resulting consequences, offering valuable perspectives on the societal impact of criminal acts.
5) Emphasis on Accuracy: However, it's noteworthy that these studies often lay emphasis on enhancing the accuracy of crime prediction. Achieving substantial accuracy is a common objective shared by researchers, necessitating the implementation of various techniques and approaches.
6) *Studies on Chicago*

a) Kang et al. improved prediction models using environmental context information, achieving favorable results through data fusion and deep neural networks.
b) Stec and Klabjan merged convolutional and recurrent neural networks to predict crime types in Chicago, showcasing notable accuracy.
c) Another study employed the ARIMA model for forecasting future crimes in Chicago, introducing their own model, LFSNBC, and achieving impressive accuracy.
d) Utilizing satellite images, Najjar et al. inquired into crime rates and achieved a 79% prediction accuracy through convolutional neural networks.
e) These studies employ diverse methodologies, including neural networks and spatial-temporal approaches, to predict and understand crimes in Chicago comprehensively.

©IJRASET: All Rights are Reserved

7) Studies on Los Angeles

a) Research in Los Angeles has addressed various aspects, including technological changes in the newsroom, the impact of medical marijuana dispensaries on crime rates, and rearrest rate predictions.

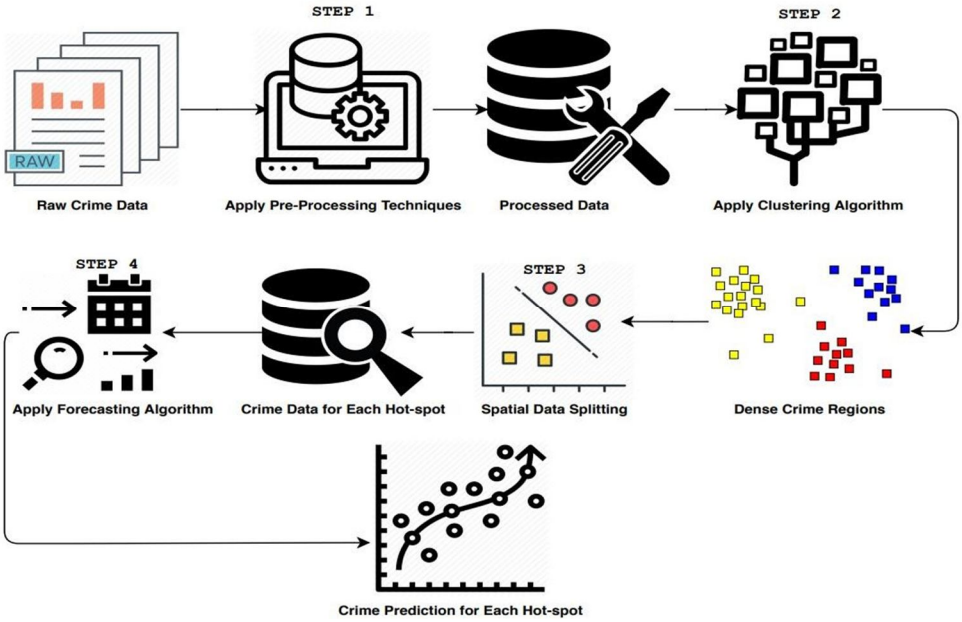
b) Studies have also explored racial biases in arrests and the impact of rail transit on crime in neighborhoods near transit stations.

c) Environmental risk factors have been analyzed to influence gang violence in East Los Angeles.

d) Predictive policing experiments have been carried out to evaluate the racial biases in arrests.

These studies offer unique insights into the dynamics of crime prediction and safety measures in Los Angeles.

IV. IMPLEMENTATION DETAILS OF MODULES



1) Module 1: Data Integration and Preprocessing

Objective: The first module focuses on gathering and preprocessing the data required for crime prediction and analysis. This involves collecting historical data from authoritative repositories, such as public datasets for Chicago and Los Angeles. The data will be carefully cleaned and structured to ensure its suitability for analysis.

2) Module 2: Advanced Machine Learning Models

Objective: The second module involves the implementation of advanced machine learning models for crime prediction. These models, such as deep neural networks, convolutional neural networks (CNN), and recurrent neural networks (RNN), will be employed to analyze and predict crime patterns. Specific attention will be given to feature-level data fusion and data-level integration to enhance the accuracy of predictions.

3) Module 3: Time Series Analysis

Objective: Module three is dedicated to time series analysis, focusing on identifying temporal patterns and seasonality in crime data. The primary tool for this task is the Long Short-Term Memory (LSTM) algorithm, which excels at classifying crimes over time. The goal is to extract valuable insights from the temporal aspects of crime data.

4) Module 4: Spatial-Temporal Analysis

Objective: This module seeks to investigate spatial-temporal crime events, emphasizing the identification of crime hotspots. It involves the application of clustering algorithms, such as DBSCAN and ARIMA, to detect high-risk crime regions. The aim is to forecast future crime trends using a spatial-temporal approach, ensuring a comprehensive understanding of when and where crimes are more likely to occur.

©IJRASET: All Rights are Reserved

5) *Module 5:* Socio-Economic Analysis

Objective: Module five delves into the socio-economic factors influencing crime. It employs linear regression analysis to explore the relationships between crime and indicators such as poverty rates, unemployment, and education. This analysis provides a nuanced understanding of the economic and ecological factors contributing to criminal behavior.

6) *Module 6:* Novel Data Sources

Objective: This module focuses on the integration of unconventional data sources, such as online forums and mobile phone data, to gain new insights into the dynamics of criminal activities in the digital age. The goal is to expand the scope of data used for crime prediction and analysis.

7) *Module 7:* Predictive Policing Experiments

Objective: Module seven entails predictive policing experiments using algorithmically predicted locations. This involves assessing the impact of rail transit on crime, analyzing the relationship between medical marijuana dispensaries and crime rates, and evaluating the influence of dispensaries for marijuana on criminal activities.

8) *Module 8:* Model Evaluation and Performance Metrics

Objective: Throughout the project, rigorous model evaluation will be conducted using various performance metrics. These metrics include accuracy, precision, recall, mean absolute error (MAE), and mean relative error (MRE). The goal is to ensure that the implemented models provide accurate and reliable crime predictions.

9) *Module 9:* Visualization and Reporting

Objective: The final module is dedicated to visualizing the results and reporting the findings. This involves creating visual summaries of crime types, counts, and trends. The project's outcomes will be documented in a comprehensive report, highlighting the key insights and contributions to the field of crime prediction and forecasting.

V. CONCLUSION

In the realm of smart city infrastructure, the provision of a secure and reliable environment is of paramount importance. Detecting crime hotspots and predicting crime rates within these regions is a pivotal component of this endeavor. Such information empowers stakeholders to create and maintain safe urban spaces for the citizens of a smart city. However, this pursuit is not without its challenges, particularly concerning the effective management and utilization of computational resources, which are increasingly strained by the ever-expanding data volumes in smart cities.

In response to these challenges, this paper has introduced a cost-efficient and pragmatic approach designed to accomplish the objectives of crime hotspot detection and crime rate prediction. The proposed system has been rigorously evaluated using a substantial dataset encompassing a decade of crime reports. The experimental results substantiate the effectiveness of the proposed system, showcasing its superior performance when compared to state-of- the-art systems. The system achieved an average Mean Absolute Error (MAE) of 11.47, highlighting its capacity to provide accurate predictions.

Looking ahead, we have identified opportunities for further enhancement. One promising avenue is the application of transfer learning, which involves leveraging the knowledge gained from an already established crime prediction model to address related crime regions.

This strategic utilization of pre-existing models promises to improve both cost- effectiveness and learning performance, ultimately leading to more accurate predictions.

Additionally, the adoption of clustering ensembles is a prospective enhancement that can bolster the robustness and precision of our crime detection and prediction model. This ensemble-based approach has the potential to refine our understanding of crime dynamics in smart cities, thus contributing to a safer and more secure urban environment.

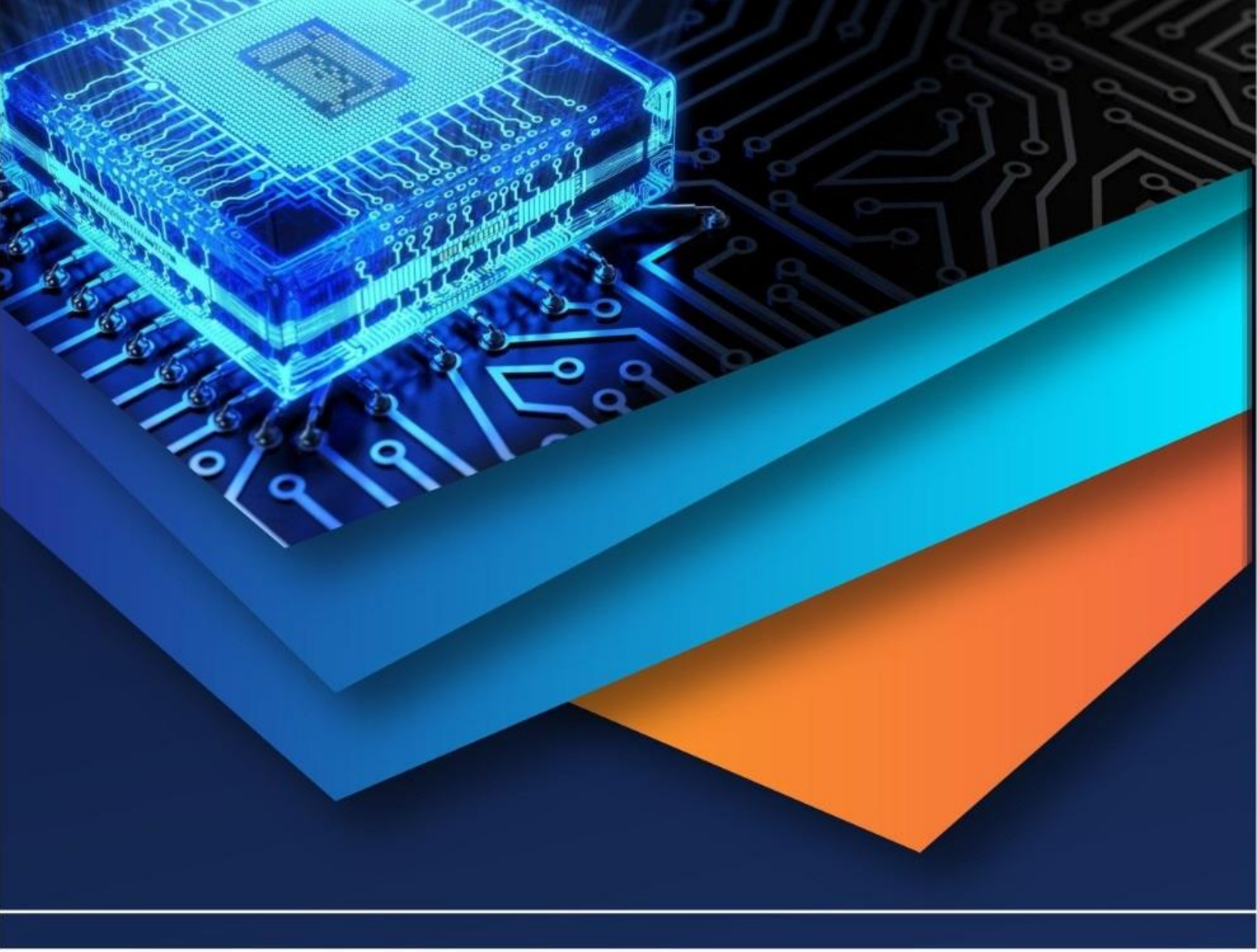
In summation, our work represents a significant step forward in the pursuit of reliable and secure smart city infrastructure. By delivering a practical and cost-efficient approach, we aim to empower stakeholders and decision- makers with the tools and insights required to ensure the safety and well-being of citizens in the evolving landscape of smart cities. The continuous refinement and expansion of our methodology, including the integration of transfer learning and clustering ensembles, promise even greater advancements in the field of smart city crime prediction and prevention.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] F. Cicirelli, A. Guerrieri, G. Spezzano, and A. Vinci, "An edge-based platform for dynamic smart city applications," Future Generation Computer Systems, vol. 76, pp. 106–118, 2017.
[2] H. H. R. Sherazi, R. Iqbal, F. Ahmad, Z. A. Khan, and M. H. Chaudary, "Ddos attack detection: A key enabler for sustainable communication in internet of vehicles," Sustainable Computing: Informatics and Systems, vol. 23, pp. 13–20, 2019.
[3] R. Iqbal, T. A. Butt, M. Afzaal, and K. Salah, "Trust management in social internet of vehicles: Factors, challenges, blockchain, and fog solutions," International Journal of Distributed Sensor Networks, vol. 15, no. 1, p. 1550147719825820, 2019.
[4] Z. Dar, A. Ahmad, F. A. Khan, F. Zeshan, R. Iqbal, H. H. R. Sherazi, and A. K. Bashir, "A context-aware encryption protocol suite for edge computing-based IoT devices," The Journal of Supercomputing, pp. 1–20, 2019.
[5] 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. 166 553–166 574, 2020.
[6] A. Baqir, S. ul Rehman, S. Malik, F. ul Mustafa, and U. Ahmad, "Evaluating the performance of hierarchical clustering algorithms to detect spatio-temporal crime hot-spots," in 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). IEEE, 2020, pp. 1–5.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)