2023 International Conference on Emerging Research in Computational Science (ICERCS) | 979-8-3503-5976-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICERCS57948.2023.10434246

Enhanced Public Safety through LSTM-Based Spatio-Temporal Crime Analytics

1st Sairam Pimple,
Vellore Institute of Technology
Chennai Campus
Tamilnadu, India
*1st sairampimple003@gmail.com

2nd Mansoor Hussain D, 3rd Karthikeyan N
School of Computer Science and Engineering
Vellore Institute of Technolog,
Chennai Campus
Tamilnadu, India
1st mansoorhussain.d@vit.ac.in
2nd narenkarthikeyan.mecse@gmail.com

Abstract - **The study explores the application of Long Short-Term Memory (LSTM) neural networks in handling Space-Time Crime Analytics, with an aim to boost public security in city regions. Utilizing LSTM's potential for mapping intricate crime dynamics within both time and space dimensions, it seeks to deliver prompt and accurate forecasts of criminal activities. Main goals include marking danger zones prone to crimes, understanding temporal patterns & discerning emerging trends; aiding police departments by equipping them with active measures against future incidents & effective resource distribution strategies. This trailblazing advancement aims at transforming predictive policing methods towards secure urban societies while highlighting how deep learning can be pragmatically applied on a spatio-temporal basis regarding analyzing crimes that elevates tactics associated with digital-age safety.**

Keywords: Long Short-Term Memory ,Crime Predictions, spatio-temporal, Hotspots

I. INTRODUCTION

Modern studies are progressively focused on tackling city crime by harnessing sophisticated machine learning models and data-driven approaches. Recognizing the variety of obstacles encountered in policing efforts and public wellness, numerous research undertakings have been carried out in major cities like New York and Chicago to delve into crime prediction deeply. These probes not only unravel the complex web that constitutes urban criminal activity but underscore concerted strategies used to mitigate these issues too. Included within these strategic endeavors is an array of cutting-edge tools such as XGboost, Convolutional Neural Networks, Long Short-Term Memory networks (LSTMs), among other innovative methodologies [1]. The primary aim here isn't just about understanding spatio-temporal intricacies around crimes occurring within a metropolis; it's also about acknowledging unique features specific to each cityscape involved simultaneously while doing so effectively assists law enforcement for optimal resource distribution considerations. Henceforth studying algorithms alongside novel methodical proceedings regarding predicting possible felonious acts has shed light upon discernable patterns making certain areas hotspots across various kinds of urban environment. Indeed, the application of LSTM networks in understanding variations and progressions related to illicit activities has been significantly successful. Over time, employing XGBoost algorithms increased precision for crime forecasting due to their adaptability in managing vast amounts of data which further aid crafting practical law enforcement measures. On another note, Convolutional Neural Networks have displayed remarkable skills by pinpointing spatial patterns thus improving crime forecast accuracy even more. This illustrates how city-specific attributes are crucial when identifying high-risk areas and comprehending trends relating to criminal occurrences. Those frameworks were instrumental tools that offer substantial insights into resource management decisions taken by legal bodies aiming at enhancing strategies involving both temporal along with spatial facets; making it a holistic approach towards combating illegal operations Incorporation of machine learning techniques within predicting crimes showed noteworthy promise reinforcing law-enforcement competencies Furthermore this addresses ongoing challenges occurring from urban infractions across different cities laying groundwork useful as an innovative reference or benchmark aimed strategically improving decipher providing comprehensive evidence thereby amplifying efficacy associated respective initiatives.

The dynamic nature of urban crime needs adaptable and inventive solutions. Cutting-edge models, such as LSTMs, have showed promise in capturing the developing patterns of criminal behaviours across time, providing a detailed understanding of temporal sequences in crime. Simultaneously, XGBoost, a fast and resilient algorithm, has considerably enhanced prediction accuracy, assisting law enforcement organisations in more precise resource allocation and strategy design. Convolutional Neural Networks have emerged as effective tools for comprehending spatial patterns, which is critical for accurate crime prediction. These algorithms have proven beneficial in detecting high-risk regions and forecasting crime trends, providing insights that help law enforcement allocate resources more effectively. These predictive models' consideration of city-specific factors emphasises the significance of customising solutions to the particular peculiarities of each urban area.The incorporation of machine learning approaches in crime prediction is proving revolutionary in enabling law enforcement to successfully address the issues posed by urban crime. These models serve as cornerstones for strengthening law enforcement strategies, increasing public safety, and optimising resource allocation in cities. Their capacity to combine temporal and spatial variables provides a comprehensive perspective necessary for proactive crime prevention and management. In conclusion, the integration of cutting-edge machine learning models and data-centric approaches is critical in understanding and addressing the complex difficulties faced by urban crime. These research emphasise the importance of city-specific analysis, the possibilities of complex algorithms such as LSTMs and XGBoost, and the value of spatial comprehension using Convolutional Neural Networks. Leveraging machine learning in crime prediction is a promising technique for law enforcement, adding significantly to urban safety and the effective management of crime in various urban environments.

II. RELATED WORK

Zhang and colleagues [1] examined machine learning techniques for forecasting crime in a coastal Chinese city between 2015 and 2018. According to the findings, the LSTM model outperformed KNN, random forest, SVM, naive Bayes, and CNN. Furthermore, incorporating environmental factors enhanced prediction accuracy, emphasising the necessity of examining historical crime data and associated covariates in crime prediction. Butt and colleagues [2] conducted a review of the literature on methodologies for finding and predicting spatio-temporal crime hotspots in their study, revealing a lack of comprehensive studies in this topic. The study included 49 papers, focusing on data mining, machine learning, time series analysis, deep learning approaches, and potential limitations within these procedures. Kshatri and colleagues [3] developed the "assemble-stacking based crime prediction method (SBCPM)" in 2021 using MATLAB and Support Vector Machine (SVM) methodologies. This technique outperformed previous studies with a classification accuracy of 99.5% on testing data. The study highlights the promise of ensemble learning techniques in the field of crime prediction. Forradellas et al. (2021) [4] used data from 2016 to 2019 to create a Python-based crime prediction model for Buenos Aires, Argentina. Using K-means clustering and the SEMMA paradigm, the model predicts numerous crimes in the urban setting, improving crime prediction. Shah and colleagues [5] emphasise the need for more effective crime prevention tactics in their 2021 study by emphasising the limitations of traditional methods. They advocate for the integration of machine learning and machine vision techniques in order to improve crime detection and resolution. By emphasising the inadequacies of existing methods, Shah and colleagues (2021) emphasise the need for efficient crime prevention strategies. Their concept entails combining machine learning and computer vision techniques in order to improve crime detection and resolution [6]. Safat et al. [7] use machine learning methods to anticipate yearly drops in Chicago and minor rises in Los Angeles, which enhances crime prediction in urban safety management. The ARIMA model anticipates a large reduction, which will guide law enforcement actions and strategies. Khan and colleagues [8] present a crime prediction model based on Random Forest, Naive Bayes, and Gradient Boosting Decision Tree algorithms. This model produced a 97% prediction rate and a 98.5% accuracy, demonstrating potential benefits for security organisations in allocating resources and developing proactive crime prevention tactics. In 2020, Han and colleagues [9] developed a daily crime prediction model that used LSTM and ST-GCN to identify high-risk locations prone to stealing offences. They evaluated the model's effectiveness using crime data from Chicago, giving useful insights that could help urban security and management. Baek et al. (2021) [10] created a machine learning system for predicting crime kinds and risk levels that outperformed standard algorithms by 7% to 8%. The system employs text-based KICS data and a simple graphical user interface. Li, X., Kang, X., and Wang, C. [11] provide a neural-network-based model for charge prediction in legal documents. This novel system merges fact descriptions and criminal interpretations into a low-dimensional space, efficiently incorporating judicial crime interpretation. Using this strategy, the model considerably improves charge prediction accuracy, successfully resolving difficulties associated to inconsistent data and ambiguous charges [11]. Tasnim, Imam, and Hashem [12] provide an innovative technique to improving crime prevention by combining machine learning and deep learning methodologies. This approach predicts future crime occurrences using data from the previous 24 hours, providing valuable insights for law enforcement. In terms of predicted accuracy, it outperforms other models in both San Francisco and Chicago. Big data analytics (BDA) is the methodical examination of huge information in order to uncover patterns, correlations, and trends, especially when applied to criminal data [13]. This includes data mining, deep learning techniques, and exploratory data analysis. Feng et al. (2019) studied BDA for visualising crime data and forecasting trends [13]. BDA has provided useful insights and patterns in places such as San Francisco, Chicago, and Philadelphia. According to the findings, a three-year data period is excellent for training neural network models. These findings have a substantial impact on law enforcement, boosting crime comprehension and resource allocation while also improving decision-making processes [13]. The study used LR, KNN, RF, and XGBoost to forecast long-term crime trends in Dallas by combining spatiotemporal lag variables for robbery occurrences from 2014 to 2018 [14]. Deng et al. (2023) focused on improving crime prediction accuracy by incorporating spatiotemporal dependencies into machine learning models, using linear regression and 5-fold crossvalidation to optimize model parameters [14]. The review paper employs machine learning and deep neural learning methodologies for crime foresting, offering valuable perspectives on dataset analysis and the evolving trends in this domain [15]. It pinpoints crucial deficiencies and recommends future paths to enhance predictive precision, serving as a valuable resource for researchers and law enforcement agencies striving to refine their crime prevention and response strategies [15]. The article by Liang, W. introduces TD-Crime, a novel Tensor Decomposition framework for crime prediction [16]. It addresses the challenge of forecasting crimes using historical data, even when the data is incomplete. TD-Crime uses tensor organization and nonnegative CP decomposition to implicitly capture spatial, temporal, and categorical correlations, improving forecasting accuracy. This approach outperforms previous methods in real-world datasets, providing a potential option for crime prediction in various contexts with missing data [16]. The study employs a Time Delay Neural Network for serial crime prediction, incorporating time, distance, and suspect biographies [17]. It introduces an updated NARX model with SiRBF activation functions, resulting in reliable crime predictions. This method beats existing techniques when tested across five time-series datasets [17]. The study by Hakyemez, T. C., and Badur, B. [18] explores the impact of park events on crime risk and introduces Park Event Density (PED) to assess event density across parks. PED is integrated into crime hotspot prediction models for robbery and theft incidents in Chicago from 2016-2018, resulting in a significant accuracy improvement of up to 25%.The study by B. Zhou, L addresses fine-grained road-level crime prediction by introducing a road-level framework using urban sensing data. It incorporates spatio-temporal features and addresses the crime near-repeat phenomenon to overcome data sparsity issues. Furthermore, the framework offers risk-aware route recommendations. Real-world data from New York City validates its accuracy in predicting road-level crime risk, thereby assisting in the development of effective public safety strategies [19]. Xu, W.'s research [20] unveils the Fusion Information Graph Attention Networks (FIGAT), which effectively categorizes individuals into high or low crime risk groups using their personal movement time series and location trajectories. This approach notably enhances classification accuracy in contrast to conventional machine learning, deep learning, and graph neural network techniques. The study by Zhang, X. [21] integrates XGBoost with 17 variables and SHAP interpretation for improved crime prediction, highlighting the impact of population demographics and reconciling machine learning's prediction capabilities with transparency issues. The study by Tam, S. integrates Twitter sentiments via ConvBiLSTM, achieving a 97.75% crime prediction accuracy, surpassing traditional models. Social media sentiment assists in public safety assessment, and data fusion enhances crime prediction [22]. The paper outlines the model architecture, results, and future research.

The study by Yan, Z. proposes two XGBoost models, OVR-XGBoost and OVO-XGBoost, to improve theft crime prediction by addressing unbalanced class distributions through the SMOTENN algorithm. The results show better accuracy compared to baseline models, emphasizing the need to handle imbalanced datasets [23]. Butt, U. M.'s paper [24] introduces an innovative approach to forecast high-risk crime areas within smart cities. This method utilizes Hierarchical Density-Based Spatial Clustering and Seasonal Auto-Regressive Integrated Moving Average, surpassing prior methodologies by achieving an average Mean Absolute Error (MAE) of 27.03. In the investigation by Esquivel, N. [25], a fusion of CNN and LSTM network is harnessed to predict the likelihood of crime incidents in Baltimore. This research delves into historical crime matrices to evaluate the model's proficiency in projecting forthcoming events.

III. METHODOLOGY

The creation of LSTM aimed to address the challenges posed by the vanishing and exploding gradients in traditional RNNs, which hinder their ability to effectively preserve information over extended sequences. These difficulties commonly result in the omission or disproportionate emphasis on specific segments within the input sequence, particularly when dealing with long temporal dependencies. The distinct architecture of LSTM allows it to sustain and transmit information across numerous time steps, rendering it highly suitable for tasks cantered on sequential data, is shown in figure 1. LSTMs achieve this by employing a sophisticated structure that includes memory cells and gating mechanisms.

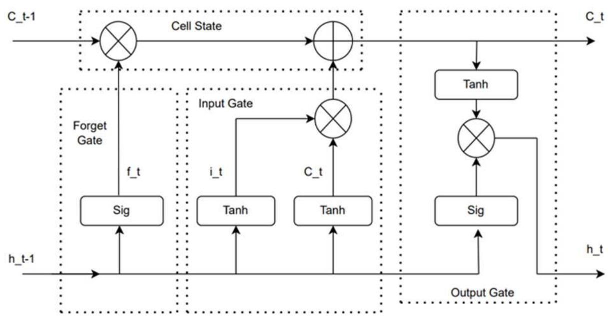

Figure 1. LSTM architecture [13]

The tanh function is an abbreviation for hyperbolic tangent, compresses input values within the range of -1 to 1, playing a significant role in different facets of LSTM functionality. Its key usage lies in producing potential cell state values within the input gate and computing the resultant hidden state output via the output gate within LSTM cells. This restricted range enables LSTMs to handle both positive and negative information, successfully counteracting the vanishing gradient problem and considerably increasing the network's capacity to store memory and process sequential input. The sigmoid function in the LSTM architecture is critical because it introduces non-linearity by squashing values between 0 and 1. This function allows gating systems to control information flow by determining what to retain, discard, or output at various stages in an LSTM cell. By compressing values in the range of 0 to 1, the sigmoid function fine-tunes information flow, improving the network in managing long-term dependencies in sequential input and enhancing its capacity to store critical information across longer sequences. The LSTM is a chain structure composed of memory blocks known as cells and four neural networks. Cells store data, and gates control memory.

A. Forget Gate

The forget gate in an LSTM cell is critical for regulating information flow. Using a sigmoid function, it determines whether information from the prior cell state should be maintained (indicated by values close to 1) or deleted (indicated by values close to 0), taking into account both the current input and the previous hidden state.This gate is critical for the LSTM's ability to keep relevant information over long sequences by filtering the cell state and determining what to remember or discard, assisting the network in learning and managing sequential input efficiently. [1] The forget gate equation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f \quad (1)$$

Where, f_t denotes Is the vector of forget gate values, and W_f: Is the weight matrix for the forget gate, h_{t-1}: is the Previous hidden state, and x_t: Current input, b_f: Bias for the forget gate,and σ: Bias for the forget gate.

B. *Input Gate*

The purpose of the input gate is to determine which new information is to be incorporated into the memory cell state within a given time step. This gate evaluates the relevance of fresh candidate values and generates prospective data for inclusion in the cell state by utilising both sigmoid and hyperbolic tangent functions. This method enables the LSTM to selectively absorb and retain relevant elements from the ongoing input, assisting the network in its learning process with sequential data. As a result, this dramatically improves the network's ability to store and process critical information across long durations. The input gate equation [1] is as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

Where, i_t : Vector of input gate values, W_i , W_c : Weight matrices for the input gate and candidate values, h_{t-1} : Previous hidden state, x_t : Current input. b_i,b_c : Biases for the input gate and candidate values, σ : Sigmoid activation function, tanh : Hyperbolic tangent activation function.

C. *Output Gate*

The output gate significantly influences the creation of the ultimate hidden state by managing the information to be disclosed from the cell state. Through the utilization of sigmoid functions, it governs the cell state and calculates the final hidden state. This gate ensures that only relevant data is included in the output, which affects the LSTM's prediction skills or decision-making based on processed sequential input. Its role is critical in improving the network's ability to generate valuable outputs for a variety of applications by extracting critical information from the cell state. The output gate equation [1] is written as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t \odot tanh(C_t) \quad (5)$$

Where, o_t: Vector of output gate values, W_o: Weight matrix for the output gate, h_{t-1}: Previous hidden state, x_t: Current input, b_o: Bias for the output gate, ⊙: Element-wise multiplication.

D. *Steps of Spatio-Temporal Crime Prediction*

Spatio-temporal crime prediction includes a sequence of processes for forecasting criminal activity in specific areas. It starts with gathering raw crime data, which includes information such as timestamps and geographical positions. After that, pre-processing procedures are used to clean and standardise the data. Clustering algorithms are employed to identify dense crime regions or 'hot-spots.' The spatial data is split into individual hot-spot regions, and specific crime data is prepared for each of these areas. Forecasting algorithms are used to generate predictive models for each hotspot, taking historical data, regional clustering, and temporal patterns into account. These models can forecast future crime episodes in each region. Law enforcement organisations can effectively allocate resources and undertake targeted preventive actions in high-risk regions by following the methods shown in Figure 2.

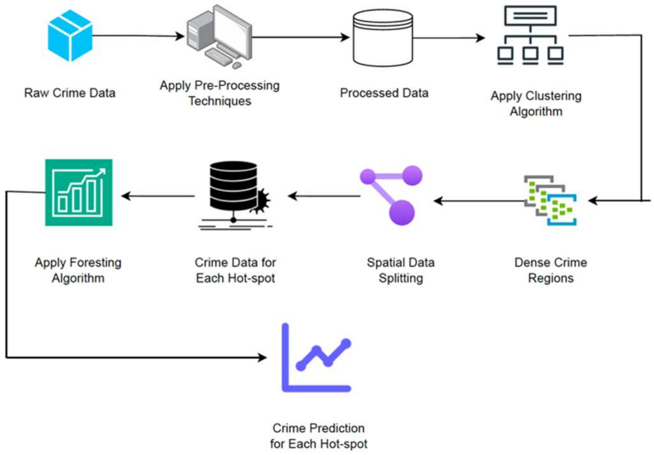

Figure 2: *Steps of Spatio-Temporal Crime Prediction*. [24]

E. *Advantages of LSTM*

Long-Term Dependencies: LSTMs excel in understanding and utilising long-term dependencies within sequential data by effectively retaining and processing information across long sequences. This skill is especially useful in situations that require understanding of temporal patterns. Avoiding Vanishing/Exploding Gradients: LSTMs reduce the vanishing and exploding gradient issues that plague regular RNNs. This enables for more effective sequence training by retaining and propagating information through the network. LSTMs are useful for many assignments, including natural language processing, speech recognition, time series forecasting, and similar tasks, due to their flexible properties and effective handling of sequences. They stand out in activities requiring sophisticated pattern identification.

IV. DISCUSSION

The LSTM model has the advantage of effectively processing and interpreting sequential data, capturing long-term interdependence in temporal sequences and allowing for a comprehensive comprehension of spatiotemporal crime patterns. Its ability to adapt and learn from many forms of data, as well as incorporate broader environmental aspects including demographic demographics and social media emotions, considerably enhances forecast accuracy. Overall, LSTM models have a stronger potential for robust crime prediction by taking into account comprehensive temporal, geographical, and contextual information. Although the Long Short-Term Memory (LSTM) model has outperformed traditional machine learning methods in crime prediction, it does have many limitations. Because it is based on previous data sequences, it suffers when the data is insufficient, missing, or biassed, decreasing its forecast accuracy. The model may suffer with excessively lengthy sequences, resulting in disappearing or ballooning gradient difficulties and restricting its capacity to capture long-term temporal trends. Because LSTMs involve extensive parameter tweaking, training is computationally and time-consuming. Overfitting is a risk, especially when the dataset is limited or biassed. The interpretability of this architecture can be challenging because comprehending how LSTMs arrive at specific predictions is not clear. Furthermore, these models may fail to account for exogenous factors that drive crime, limiting their application. Understanding these restrictions is crucial for designing more effective crime prediction systems. By addressing these challenges, future models will be able to combine varied data sources and external characteristics, potentially leading to more accurate and complete crime prediction approaches. By solving these obstacles, crime prediction algorithms could greatly improve, capturing a more comprehensive understanding of criminal activity in metropolitan contexts.

V. FUTURE WORK

The enhancement of public safety via LSTM-based spatiotemporal crime analytics has the potential to make considerable progress in a number of crucial areas. To begin, future research will concentrate on combining varied variables into LSTM models to improve forecast accuracy, taking into consideration social, economic, and environmental issues. Advances in spatiotemporal data mining will refine techniques designed specifically to handle complicated information, boosting the accuracy of hotspot prediction. The merging of urban and social media data will boost prediction capacities, enabling for the detection of crime hotspots in real time. Transparency and interpretability of models will be prioritised, and approaches such as SHAP and attention mechanisms will be employed to elucidate decision-making processes. Law enforcement will be able to respond to criminal activities more swiftly with real-time prediction systems based on LSTM models, allowing for proactive intervention.To ensure responsible and equitable model use in law enforcement, ethical considerations regarding biases and fairness in predictive analytics are critical. These advancements seek to improve model accuracy, interpretability, and applicability in real-time, thereby contributing significantly to more effective public safety and crime prevention measures.

REFERENCES

[1] Zhang, Xu, et al. "Comparison of machine learning algorithms for predicting crime hotspots." *IEEE access* 8 (2020): 181302-181310.

[2] Butt, Umair Muneer, et al. "Spatio-temporal crime hotspot detection and prediction: a systematic literature review." *IEEE access* 8 (2020): 166553-166574.

[3] Kshatri, Sapna Singh, et al. "An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: an ensemble approach." *Ieee Access* 9 (2021): 67488-67500

[4] Reier Forradellas, Ricardo Francisco, et al. "Applied machine learning in social sciences: neural networks and crime prediction." *Social Sciences* 10.1 (2020): 4.

[5] Shah, Neil, Nandish Bhagat, and Manan Shah. "Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention." *Visual Computing for Industry, Biomedicine, and Art* 4 (2021): 1-14.

[6] Yang, Dingqi, et al. "CrimeTelescope: crime hotspot prediction based on urban and social media data fusion." *World Wide Web* 21 (2018): 1323-1347.

[7] Safat, Wajiha, Sohail Asghar, and Saira Andleeb Gillani. "Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques." *IEEE access* 9 (2021): 70080-70094.

[8] Khan, Muzammil, Azmat Ali, and Yasser Alharbi. "Predicting and preventing crime: a crime prediction model using san francisco crime data by classification techniques." *Complexity* 2022 (2022).

[9] Han, Xinge, et al. "Risk prediction of theft crimes in urban communities: An integrated model of LSTM and ST-GCN." *IEEE Access* 8 (2020): 217222-217230.

[10] Baek, Myung-Sun, et al. "Smart policing technique with crime type and risk score prediction based on machine learning for early awareness of risk situation." *IEEE Access* 9 (2021): 131906-131915.

[11] Hakyemez, Tugrul Cabir, and Bertan Badur. "Incorporating park events into crime hotspot prediction on street networks: A spatiotemporal graph learning approach." *Applied Soft Computing* 148 (2023): 110886.

[12] Tasnim, Nowshin, Iftekher Toufique Imam, and M. M. A. Hashem. "A novel multi-module approach to predict crime based on multivariate spatio-temporal data using attention and sequential fusion model." *IEEE Access* 10 (2022): 48009-48030.

[13] Feng, Mingchen, et al. "Big data analytics and mining for effective visualization and trends forecasting of crime data." *IEEE Access* 7 (2019): 106111-106123.

[14] Deng, Yue, Rixing He, and Yang Liu. "Crime risk prediction incorporating geographical spatiotemporal dependency into machine learning models." *Information Sciences* 646 (2023): 119414.

[15] Mandalapu, Varun, et al. "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions." *IEEE Access* (2023).

[16] Liang, Weichao, et al. "Crime Prediction With Missing Data Via Spatiotemporal Regularized Tensor Decomposition." *IEEE Transactions on Big Data* (2023).

[17] Ghazvini, Anahita, et al. "Crime spatiotemporal prediction with fused objective function in time delay neural network." *IEEE Access* 8 (2020): 115167-115183.

[18] Hakyemez, Tugrul Cabir, and Bertan Badur. "Incorporating park events into crime hotspot prediction on street networks: A spatiotemporal graph learning approach." *Applied Soft Computing* 148 (2023): 110886.

[19] Zhou, Binbin, et al. "Escort: Fine-grained urban crime risk inference leveraging heterogeneous open data." *IEEE Systems Journal* 15.3 (2020): 4656-4667.

[20] [20] Xu, Wenbo, et al. "FIGAT: Accurately Classify Individual Crime Risks with Multi-information Fusion." *IEEE Transactions on Services Computing* (2022).

[21] Zhang, Xu, et al. "Interpretable machine learning models for crime prediction." *Computers, Environment and Urban Systems* 94 (2022): 101789.

[22] Tam, Sakirin, and Ömer ÖzgürTanrıöver. "Multimodal Deep Learning Crime Prediction Using Crime and Tweets." *IEEE Access* (2023).

[23] Yan, Zhongzhen, et al. "Research on prediction of multi-class theft crimes by an optimized decomposition and fusion method based on XGBoost." *Expert Systems with Applications* 207 (2022): 117943.

[24] Butt, Umair Muneer, et al. "Spatio-temporal crime predictions by leveraging artificial intelligence for citizens security in smart cities." *IEEE Access* 9 (2021): 47516-47529.

[25] Esquivel, Nicolás, et al. "Spatio-temporal prediction of Baltimore crime events using CLSTM neural networks." *IEEE Access* 8 (2020): 209101-209112.