

# Crime Prediction Based on Multi Perspective Feature Fusion and Extraction

Min Yu

Shanghai Zhihe Law Firm, Shanghai, China

Email: ym12012023@163.com

**How to cite this paper:** Yu, M. (2024). Crime Prediction Based on Multi Perspective Feature Fusion and Extraction. *Open Journal of Social Sciences*, 12, 403-413. <https://doi.org/10.4236/jss.2024.1210027>

**Received:** September 27, 2024

**Accepted:** October 25, 2024

**Published:** October 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The threat posed by criminal activities to the personal and property security of citizens, as well as to the harmony and stability of society, cannot be underestimated. Preventing crime has become one of the challenges in social governance. Crime prediction technology can extract potential crime patterns from historical crime data and predict future crime situations. The results of prediction can provide data support for formulating anti-crime policies and optimizing police patrol routes. One of the current challenges in the field of crime prediction lies in the independent extraction of features from the three most important perspectives: crime time, crime space, and crime type. Ignoring the other two perspectives while processing one may lead to the loss of useful information. This paper proposes a multi-perspective feature extraction and fusion technique, breaking the sequence of feature extraction among the three perspectives of time, space, and type. Based on the type node and bridging time and space, a multi-perspective fusion graph is constructed. The proposed method is thoroughly experimented on a real-world crime dataset. Experimental results demonstrate that the multi-perspective feature extraction and fusion technique proposed in this paper can achieve up to a 11.1% and 5.1% improvement in the Macro-F1 metric compared to the more recent crime prediction models Mist and Crime Forecaster, respectively.

## Keywords

Deep Learning, Crime Prediction, Multi-View, GNNS

## 1. Introduction

Effectively combating crime is a well-known challenge in social governance. There are two common types of crimes: those targeting property security, such as theft, burglary, and robbery, and those targeting personal safety, such as murder, assault, and aggravated assault.

The number of criminal activities both domestically and internationally in recent years is significant. According to the FBI's Crime Report for 2018-2019 (see [FBI—Crime in the U.S. 2019](#)), the United States witnessed over 1.1 million cases of violent crimes in both 2018 and 2019, including homicides, rapes, robberies, and aggravated assaults. Notably, the number of homicides and aggravated assaults increased in 2019 compared to 2018. Apart from violent crimes, the United States also experienced over 6 million cases of property crimes annually, including burglaries, thefts, motor vehicle thefts, and arsons, with thefts alone accounting for over 4.5 million cases.

The presence of crime naturally prompts efforts to combat it. Effectively reducing crime rates has long been the goal of law enforcement agencies and scholars alike. Looking back to the early 20th century, scholars focused on the evolution of criminal behavior and the interplay between criminal activities and their social environment ([Bogomolov et al., 2014](#)). In recent years, with some cities (such as Chicago and Los Angeles) making anonymized crime records public, scholars have begun macroscopic studies on crime activities based on these data. The goal is to extract potential crime patterns from large volumes of crime data and predict future crime activities. The rapid advancement of mature data mining techniques has propelled the development of data-driven crime prediction research ([Silva et al., 2020](#); [Ayele, 2020](#)). Graph Neural Networks (GNNs) are commonly used to extract spatial features from crime data, while Recurrent Neural Network (RNN) models, including LSTM ([Hochreiter & Schmidhuber, 2010](#)), GRU ([Cho et al., 2014](#)) and Temporal Convolutional Network (TCN) ([Bai et al., 2018](#)) are often employed to extract temporal features from crime data. Compared to traditional crime analysis methods, these neural network models significantly improve the accuracy of crime prediction.

## 2. Literature Review

In recent years, with the opening of historical crime data by some cities and the increasing maturity of data mining technology, crime prediction has gradually attracted attention from scholars both domestically and internationally. The emergence of crime prediction in recent years can be attributed to three main factors: firstly, the rapid development of cities has generated a large amount of urban data, which can explain the causes and diffusion patterns of crime from multiple perspectives; secondly, the widespread application of deep learning algorithms in recent years has provided new tools for feature extraction of crime patterns; thirdly, the continuous development of spatiotemporal prediction algorithms has inspired new ideas in the field of crime prediction.

Traditional crime prediction methods typically rely on demographic and geographic factors to estimate the crime rate in a given area. The rapid development of mobile devices and the internet has generated a large amount of urban data, which can offer new perspectives for crime prediction. Wang et al. demonstrated a significant improvement in prediction accuracy compared to traditional

methods by utilizing a large amount of point of interest (POI) data and taxi flow data (Wang et al., 2016). In addition to the aforementioned data sources, Zhao et al. introduced check-in data and 311 public service complaint data, proposing a spatiotemporal correlation-based crime prediction framework (Zhao & Tang, 2017). These urban data introduced can be classified into two categories. One category reflects the internal situation of the jurisdiction, such as POI data and 311 public service complaint data. The other category reflects the spatial correlation between different jurisdictions, such as taxi flow data. These data sources can provide explanatory evidence for the causes and diffusion of criminal activities.

Deep learning algorithms have achieved significant success in the field of crime prediction. Huang et al. have made pioneering efforts in both crime prediction and urban anomaly event prediction. In Huang et al. (2018), they developed a new deep learning framework called DeepCrime, which utilizes Gated Recurrent Unit (GRU) and attention mechanism to capture dynamic crime patterns.

In recent years, the emergence of graph neural network (GNN) algorithms has provided better spatial representation capabilities for crime prediction models. Zhang et al. proposed a multi-perspective joint learning model to obtain embeddings for urban areas, which was applied to downstream tasks such as land use classification and urban crime prediction (Zhang et al., 2021). Sun et al. (2021) designed and implemented an end-to-end spatiotemporal deep learning framework called Crime Forecaster, which captures both temporal repetitions and spatial dependencies. In these studies, recurrent neural networks (such as LSTM and GRU) are commonly used to extract temporal features, while graph neural networks (such as GCN) are employed to extract spatial features. Models and methods for prediction based on time and space, exemplified by the traffic prediction domain, have provided many new insights for crime prediction. Considering that existing methods overlook the heterogeneity of spatiotemporal data, Song et al. (2020) proposed a new model capable of effectively capturing complex local spatiotemporal relationships. Zheng et al. (2019) concentrated on spatiotemporal factors, presenting a graph-based multi-attention network to predict the traffic conditions at different locations in the road network overtime Addressing the issue of information loss in the feature extraction process from the perspectives of time, space, and type, this study considers how to simultaneously extract and integrate features from these three perspectives. Inspired by the work of Song et al. (2020), we take administrative units as nodes and construct a spatial graph based on spatial adjacency relationships. Subsequently, temporal edges are established between the same nodes at adjacent time points, forming a spatiotemporal graph. The nodes in the graph possess both temporal and spatial edges. We then expand upon this by enlarging the temporal perception range of the nodes. For the type perspective, we employ an attention mechanism on the initial crime count feature vectors. By utilizing a deep learning network, we learn the weights between different crime types and perform weighted summation, ensuring that the influence of crime types is adequately diffused before generating initial node features.

### 3. Research Design

Time and space have always been the two most prominent perspectives in the prediction domain, and in recent years, the rapid advancement of deep learning technology has enhanced the exploration of these two perspectives. Temporal Convolutional Networks (TCNs) and recurrent neural network (RNN) algorithms (such as LSTM, GRU, etc.) are commonly used for extracting time features. Graph Neural Network (GNN) algorithms (such as GCN) are often employed for extracting spatial features.

Regarding the methods for extracting time and space features, two approaches have gradually emerged. One approach is to first extract spatial features and then perform time feature extraction on top of that. The other approach involves separately extracting time and space features and then merging them (Zheng et al., 2019). Both of these spatiotemporal extraction methods suffer from a common issue: the separation of the time and space feature extraction processes. When extracting time features, spatial features are not considered, and vice versa. Time and space are strongly correlated, and this separation in the feature extraction process may lead to the loss of useful information specific to time and space.

#### Spacetime Fusion Topology Graph Construction

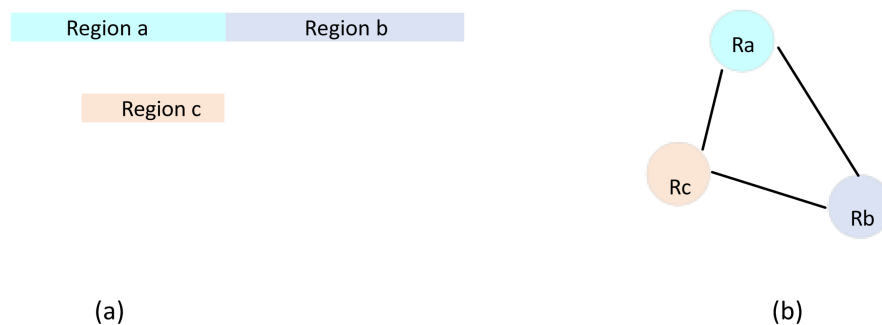
To break the sequential nature of feature extraction between time and space perspectives, this paper integrates time and space into one graph. The nodes in the graph represent administrative regions in the real world, and the edges represent the connections between these regions. We refer to this graph as the “Spacetime Fusion Topology Graph.” Each node in the graph has both temporal and spatial neighbors. In the subsequent feature extraction and fusion stages, temporal and spatial features can be aggregated without distinction, achieving synchronous extraction and fusion of spatiotemporal features.

Before constructing the Spacetime Fusion Topology Graph, it is necessary to construct the spatial graph. In the spatial graph, nodes represent actual administrative regions, and edges represent spatial adjacency between nodes. If two administrative regions are adjacent in geographic space, an edge is constructed between the corresponding nodes in the spatial graph.

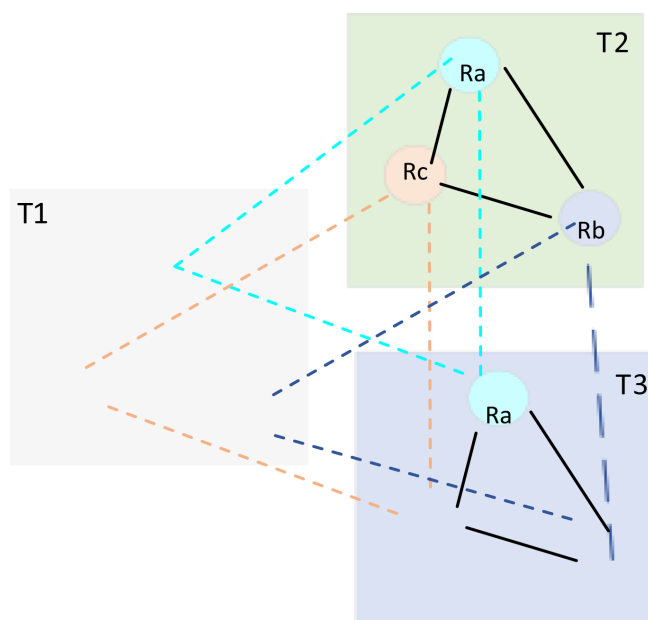
To visually demonstrate the construction process of the Spacetime Fusion Topology Graph, let's consider a hypothetical city with three administrative regions. We'll use this as an example for illustration. The geographic representation of these three regions is shown in **Figure 1(a)**, where each region is denoted as Region A, Region B, and Region C, respectively. Based on their spatial adjacency relationships, we construct their spatial topology graph, as shown in **Figure 1(b)**.

With the spatial graph in place, we extend it to the temporal dimension by creating temporal edges between pairs of the same region at different time intervals, as illustrated in **Figure 2**. In **Figure 2**, nodes of the same color represent the same administrative region, while edges of the same color as the nodes indicate temporal

edges established within the same region at different time intervals. Black edges represent spatial edges, indicating spatial adjacency relationships.



**Figure 1.** Spatial map construction.



**Figure 2.** Spatiotemporal fusion topology diagram.

Taking the Ra node at time slot T2 as an example, it is connected not only spatially to Rb and Rc at the same time slot but also temporally to Ra at time slot T1 and T3. In other words, the Ra node at T2 has both spatial and temporal neighbors, allowing it to perceive the influence of both space and time on crime patterns simultaneously.

Further, we construct the adjacency matrix for this example graph, as shown in **Figure 3**. The blue squares in the matrix represent spatial adjacency, while the yellow squares represent temporal adjacency. A value of 1 indicates the presence of an edge, while 0 indicates no edge. For example, the value of 1 in the first row and second column with a blue background indicates the presence of a spatial edge between regions Ra and Rb at time slot T1. Similarly, the value of 1 in the first row and fourth column with a yellow background indicates the presence of a temporal edge between region Ra at time slot T1 and region Ra at time slot T2.

		Time 1			Time 2			Time 3		
		Ra	Rb	Rc	Ra	Rb	Rc	Ra	Rb	Rc
Time 1	Ra	0	1	1	1	0	0	1	0	0
	Rb	1	0	1	0	1	0	0	1	0
	Rc	1	1	0	0	0	1	0	0	1
Time 2	Ra	1	0	0	0	1	1	1	0	0
	Rb	0	1	0	1	0	1	0	1	0
	Rc	0	0	1	1	1	0	0	0	1
Time 3	Ra	1	0	0	1	0	0	0	1	1
	Rb	0	1	0	0	1	0	1	0	1
	Rc	0	0	1	0	0	1	1	1	0

Figure 3. Adjacency matrix diagram of spatiotemporal fusion topology graph.

In the previous section, we constructed the spatiotemporal fusion topology graph and assigned corresponding features to the nodes and edges. However, we haven't yet considered the mutual influence between crime types and the influence of spatiotemporal changes.

Our initial approach is to further split the nodes of the spatiotemporal fusion graph, with each type corresponding to anode. After splitting, we establish type edges between the nodes representing different crime types and then connect them with the original spatiotemporal neighbors to establish temporal and spatial edges. Let's assume each node has three crime types denoted as  $C = \{C1, C2, C3\}$ . We split the node Ra at time T1 into three sub-nodes, as shown in Figure 4. Similarly, the nodes connected to the Ra node at the same time, such as Rb and Rc, and the nodes at time T2 and T3 are also divided into three sub-nodes.

The result of this approach is that anode becomes adjacent to temporal, spatial, and type nodes simultaneously, leading to two main issues: First, the introduction of type edges complicates the adjacency matrix, resulting in increased model parameter complexity and significantly elevated training difficulty. Second, due to the sparsity of crime data, splitting by type results in a large number of nodes with zero counts, causing severe sample imbalance.

Therefore, to address this issue, this study incorporates the influence relationships between different crime types into the node features, extending the crime type feature  $F_{type}$ . For the intra-type influence relationships, we employ an attention mechanism to learn the weights between a certain type and other types through a neural network, followed by weighted summation based on these weights, as illustrated in Figure 5.

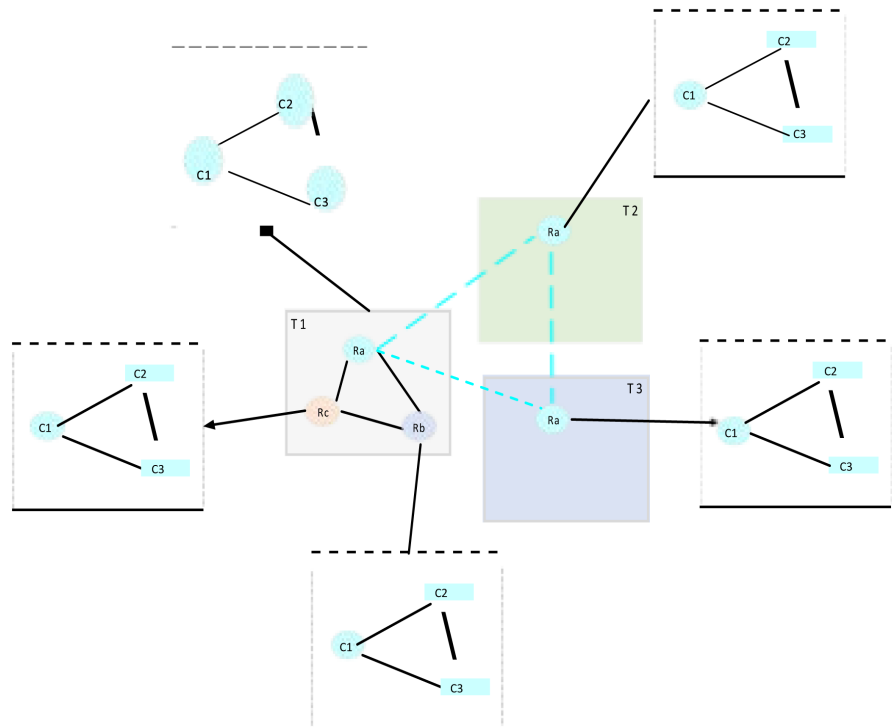


Figure 4. Schematic diagram of type nodes.

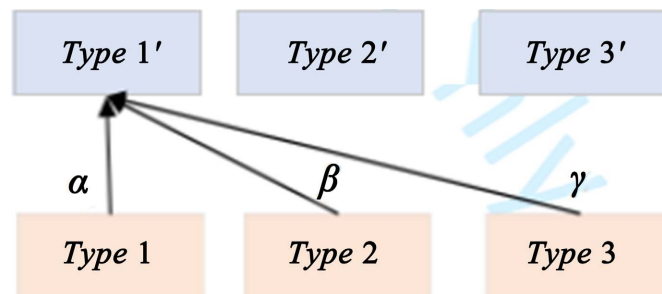
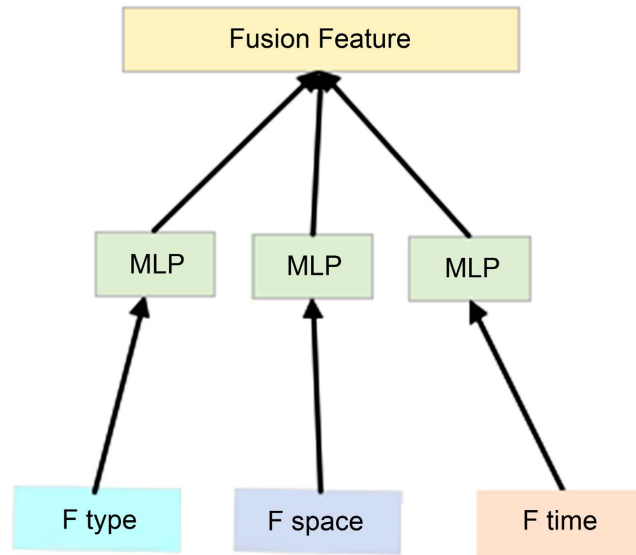


Figure 5. Type attention mechanism.

In this way, the diffusion of influence between crime types within nodes is fully achieved. In the subsequent convolution process, the type information within the nodes, along with the temporal and spatial edges, simultaneously participate in the convolution operation, achieving synchronous extraction and fusion of time, space, and type features. We refer to the graph constructed at this stage as the multi-view fusion graph.

To enhance the representational capacity of node features, as shown in Figure 6, we employ Multilayer Perceptron (MLP). Considering that the node features consist of three parts: crime patterns  $F_{type}$ , local spatial features  $F_{space}$ , and temporal cycle features  $F_{time}$ , we enhance these three parts separately. Subsequently, the enhanced features are dimensionally reduced using MLP to obtain the final node features. The fusion process is represented as shown in Equation (1).

$$X = \text{MLP}(\alpha \text{MLP}(F_{\text{pattern}}) \mid \beta \text{MLP}(F_{\text{local}}) \mid \gamma \text{MLP}(F_{\text{time}})) \quad (1)$$



**Figure 6.** Multi feature fusion.

## 4. Experiment and Conclusion

Experimental validation of the effectiveness of the proposed multi-perspective feature extraction and fusion technique, as well as the importance of each component, is conducted. Firstly, the datasets used in the experiments and the evaluation metrics to measure model performance are introduced. Then, a brief overview of the experimental setup is provided. Next, the comparative models used in the experiments are introduced, followed by experiments conducted on the datasets according to the experimental settings, and a brief analysis of the experimental results.

### 4.1. Dataset

The study utilized real-world crime datasets for experimentation. Specifically, the datasets were obtained from the Chicago Crime Dataset. For this study, crime data from Chicago spanning from January 1, 2022, to December 31, 2022, was selected. The dataset includes eight types of crimes: theft, criminal damage, narcotics-related offenses, robbery, assault, deceptive practices, burglary, and battery.

### 4.2. The Evaluation Metrics

For macro-level evaluation, Macro-F1 score is chosen to measure the overall prediction accuracy of the model across all crime types. For micro-level evaluation, Micro-F1 score is selected to assess the model's prediction accuracy for individual crime types, as it accounts for class imbalances by considering each class equally. The formulas to calculate Macro-F1 and Micro-F1 scores are as follows:

$$\text{Macro-F1} = \frac{1}{L} \sum_{i=1}^L \frac{2 \cdot \text{TP}_i}{2 \cdot \text{TP}_i + \text{FN}_i + \text{FP}_i} \quad (2)$$

$$\text{Macro-F1} = \frac{2 \cdot \sum_{i=1}^L \text{TP}_i}{2 \cdot \sum_{i=1}^L \text{TP}_i + \sum_{i=1}^L \text{FN}_i + \sum_{i=1}^L \text{FP}_i} \quad (3)$$

In the given formulas, “True Positive (TP)” refers to instances correctly predicted as positive (crime occurrence) by the model, while “False Positive (FP)” represents instances incorrectly predicted as positive (crime occurrence). “True Negative (TN)” denotes instances correctly predicted as negative (no crime occurrence), and “False Negative (FN)” indicates instances incorrectly predicted as negative (no crime occurrence) by the model.

### 4.3. Experimental Details

In our experiment, we divided the dataset into training, validation, and test sets. The first 6.5 months were used for training, 0.5 months for validation, and the remaining 5 months for testing. For the model parameters, we set the learning rate to 0.001, the number of training epochs to 50, and the batch size to 8. We chose Adam as the optimizer for the model.

We introduce several important hyperparameters used in the model. The Multilayer Perceptron (MLP) used for enhancing node features consists of three layers, with 256 neurons in the hidden layers. The activation function is relu. After the multi-view synchronous convolution algorithm, we use an MLP to enhance the representation capability of the node features after convolution. The activation function in the hidden layers is also relu, while the activation function in the output layer is sigmoid.

We conducted experiments on the Chicago crime dataset, employing the seven models categorized as mentioned earlier, in addition to the proposed model (MFF-GCN), which stands for Multi-Feature Fusion Graph Convolutional Network. The experimental outcomes on the Chicago dataset are summarized in **Table 1**.

**Table 1.** The experimental results on the Chicago dataset.

	ARIMA	SVR	LR	LSTM	GRU	MiST	CF	MFF-GCN
Macro-F1	0.421	0.389	0.568	0.436	0.512	0.586	0.643	0.697
Micro-F1	0.438	0.412	0.579	0.452	0.531	0.601	0.661	0.721

### 4.4. Comparison of Experimental Results

This study selects four types of model methods as comparison benchmarks: 1) Autoregressive Integrated Moving Average (ARIMA) (Khashei et al., 2009), a traditional time forecasting model based on historical data; 2) Support Vector Regression (SVR) [2] and Logistic Regression (LR) (Chang et al., 2011), representing traditional machine learning algorithms; 3) Long Short-Term Memory (LSTM) (Hosmer Jr et al., 2013) and Gated Recurrent Unit (GRU) (Gers et al., 2000), both recurrent neural network models commonly used for sequence processing; and 4) Crime Forecaster (Sun et al., 2021) serving as spatiotemporal prediction algorithms.

We conducted experiments on the Chicago crime dataset using the seven

models from the four categories mentioned above, along with the proposed model (MFF-GCN). The experimental results on the Chicago dataset are summarized in **Table 1**.

#### 4.5. Conclusion

This paper proposes a multi-perspective feature extraction and fusion technique. Firstly, it introduces how to construct a spatiotemporal topology graph, enabling the nodes in the graph to have both temporal and spatial neighbors simultaneously. Then, it captures the influence among different types of crimes and integrates this influence into node features to form a multi-perspective fusion graph, which is subsequently used for predictive tasks. Experiments are conducted on real-world crime datasets. The proposed model is compared and analyzed against baseline models to validate its effectiveness.

This article independently extracts features from the three most important perspectives of crime time, crime space and crime type, providing effective theoretical support for the extraction of useful information from multiple perspectives, as well as suggestions for practical use, which breaks the order of feature extraction from the three perspectives of time, space and type, and improves the accuracy of predicting crime, and hopefully will help the cases to be solved more quickly and accurately in the future practical use.

#### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

#### References

- (2019). *FBI—Crime in the U.S. 2019*.  
<https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019>
- Ayele, F. (2020). Applying Data Mining Technique for Crime Prevention: The Case of Hos-saena Town Police Office. *International Journal of Advanced Engineering Research and Science*, 7, 136-140. <https://doi.org/10.22161/ijaers.71.17>
- Bai, S. J., Zico Kolter, J., & Koltun, V. (2018). *An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling*. arXiv: 1803.01271.
- Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., & Pentland, A. (2014). Once Upon a Crime. In *Proceedings of the 16th International Conference on Multimodal Interaction* (pp. 427-434). ACM. <https://doi.org/10.1145/2663204.2663254>
- Chang, C., & Lin, C. (2011). LIBSVM. *ACM Transactions on Intelligent Systems and Technology*, 2, 1-27. <https://doi.org/10.1145/1961189.1961199>
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. et al. (2014). Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724-1734). Association for Computational Linguistics. <https://doi.org/10.3115/v1/d14-1179>
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 12, 2451-2471.  
<https://doi.org/10.1162/089976600300015015>

- Hochreiter, S., & Schmidhuber, J. (2010). Long Short-Term Memory. *Neural Computation*, 9, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hosmer Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley. <https://doi.org/10.1002/9781118548387>
- Huang, C., Zhang, J., Zheng, Y., & Chawla, N. V. (2018). DeepCrime. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (pp. 1423-1432). ACM. <https://doi.org/10.1145/3269206.3271793>
- Khashei, M., Bijari, M., & Raissi Ardali, G. A. (2009). Improvement of Auto-Regressive Integrated Moving Average Models Using Fuzzy Logic and Artificial Neural Networks (ANNS). *Neurocomputing*, 72, 956-967. <https://doi.org/10.1016/j.neucom.2008.04.017>
- Silva, J., Romero, L., González, R. J., Larios, O., Barrantes, F., Lezama, O. B. P. et al. (2020). Algorithms for Crime Prediction in Smart Cities through Data Mining. In Á. Rocha, M. Paredes-Calderón, & T. Guarda (Eds.), *Developments and Advances in Defense and Security*. (pp. 519-527). Springer Nature Singapore. [https://doi.org/10.1007/978-981-15-4875-8\\_45](https://doi.org/10.1007/978-981-15-4875-8_45)
- Song, C., Lin, Y., Guo, S., & Wan, H. (2020). Spatial-Temporal Synchronous Graph Convolutional Networks: A New Framework for Spatial-Temporal Network Data Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 914-921. <https://doi.org/10.1609/aaai.v34i01.5438>
- Sun, J., Yue, M., Lin, Z., Yang, X., Nocera, L., Kahn, G. et al. (2021). Crimeforecaster: Crime Prediction by Exploiting the Geographical Neighborhoods' Spatiotemporal Dependencies. In Y. Dong, G. Ifrim, D. Mladenić, C. Saunders, & S. Van Hoecke (Eds.), *Machine Learning and Knowledge Discovery in Databases. Applied Data Science and Demo Track* (pp. 52-67). Springer International Publishing. [https://doi.org/10.1007/978-3-030-67670-4\\_4](https://doi.org/10.1007/978-3-030-67670-4_4)
- Wang, H., Kifer, D., Graif, C., & Li, Z. (2016). Crime Rate Inference with Big Data. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 635-644). ACM. <https://doi.org/10.1145/2939672.2939736>
- Zhang, M., Li, T., Li, Y., & Hui, P. (2021). Multi-View Joint Graph Representation Learning for Urban Region Embedding. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (pp. 4431-4437). International Joint Conferences on Artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2020/611>
- Zhao, X., & Tang, J. (2017). Modeling Temporal-Spatial Correlations for Crime Prediction. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 497-506). ACM. <https://doi.org/10.1145/3132847.3133024>
- Zheng, C. et al. (2019). *A Graph Multi-Attention Network for Traffic Prediction*. arXiv: 1911.08415.