© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717355

# Time-Aware CrimeVec: Integrating Time2Vec for Spatio-Temporal Crime Embedding and Arrest Prediction

KANCHARLA PUNEETH CHOWDARY
*SRM University AP*

***Abstract-*** ***Crime analysis in urban areas is a critical component of community security and effective operation of law enforcement in cities today. Statistical summaries are used in most conventional methods of crime analysis. The episodes of crime are handled separately and do not indicate complex spatial and time associations among different types of crimes. Conventional data-driven approaches, e.g. criminal embedding techniques using Word2Vec, consider only the co-occurrence patterns, although generally not the temporal dynamics of the data, and as a result, are limited in prediction power. In order to address these limitations, in this paper, we present a novel Time-Aware CrimeVec (TA-CrimeVec) system that takes into consideration the temporal encoding and spatial-temporal crime embedding. The model is trained using Chicago Crime Dataset and uses Time2Vec to encode timestamps using linear and periodic functions. The criminal incidents are transformed to time-sensitive tokens and fed through a Skip-Gram based Word2Vec model to create hybrid embeddings that reflect crime type, time trends and spatial closeness. We then evaluate the learnt embeddings in predicting arrests with ML models like the RF and SVM. The results of the experiment demonstrate that the proposed method can attain a higher level of representation and prediction results, whereas the accuracy of the Random Forest is approximately 78.6 percent. The proposed solution is quite appropriate in capturing what happens when. It gives more information on the crime pattern and makes the decision making more informed in urban safety systems.***

***"Keywords— Crime analysis, Time-Aware Embedding, CrimeVec, Time2Vec, Spatial-Temporal Modeling, Machine Learning."***

## I. INTRODUCTION

The increase in urban crime is a colossal issue to the present cities, which has impacted on the safety of people, economic stability, and the quality of life at large. The constantly increasing urban population and the easily accessible large-scale data sets of crime necessitate a greater need of more sophisticated analytical methods to understand and model crime trends. The use of statistical tools such as autoregressive models has been used in the past to forecast time-series. The models are generally not representative of complicated and non-linear relationships in the crime data [1]. Similarly, conventional signal processing methods like fourier based methods can identify periodicities, but cannot model contextual relationships between events [2], [3].

The latest trends in ML and DL have enabled the ability to model sequential and relational data more effectively. Recurrent neural networks especially LSTM networks have shown a high ability in learning temporal patterns in sequential data [4]. Training efficiency and model convergence are improved with optimization techniques such as adaptive gradient methods [5]. Additionally, the attention-based architectures and transformer models have enhanced the sequence modeling and can model long-range associations without repetitions [6]. There are also other representation learning techniques like graph-based representations, which have proven effective to model structured data relationships [7].

However, despite these developments, numerous constraints exist to the current approaches to crime analysis. The criminal episodes are regarded as independent by many models or only geographical or temporal factors restrict the ability of the model to reflect the complex interrelations. Time-conscious models attempt to model temporal dynamics [8], [9] and point process-based methods model event sequences [10], although typically without simultaneously modeling crime type, spatio-temporal context. The lack of a single modeling makes it impossible to unearth the latent connections, and makes predictive systems less efficient.

To tackle these issues, this paper presents a Time-Aware CrimeVec (TA-CrimeVec) model, which is a framework that integrates geographical, temporal, and semantic aspects of crime data into a unified form. The proposed method incorporates continuous time encoding, and integrates it with embedding based learning to learn co-occurrence patterns as well as dynamics over time. The main contributions include construction of combined crime-time embedding, the modeling of the spatial-temporal connection using context-driven learning, and evaluation using ML classifiers. This method can lead to a complete understanding of criminal behavior and can be more effective in data-driven decisions in urban safety applications.

## II. RELATED WORK

Urban crime studies and forecasting has gained a lot of interest with the advancement of the spatio-temporal data-driven techniques. In the pre-study, statistical and rule-based methods were predominantly used in detection of crime hotspots and analysis of geographical distribution. Butt et al. also noted the change of paradigm driven by ML based frameworks and the necessity to integrate spatial and temporal dimension [12]. Correspondingly, Du and Ding demonstrated that multi-scale spatio-temporal models enhance the performance of prediction, but at the cost of complexity increment [13].

Recently, representation learning has gained popularity in research and embedding-based approaches have been explored to improve the prediction of crime data. To identify latent correlations, Crivellari and Ristea proposed CrimeVec, a model which learns vector representations of crime categories and areas, using spatial-temporal co-occurrence patterns [11]. Qian et al. in this line of work suggested GeST, a grid-based embedding model that considers spatial correlations and temporal interdependence [14]. These methods raise relational reasoning and typically implicitly suppose discrete spatial partitions and do not explicitly involve continuous temporal development.

Recent studies have employed DL and graph-based models to learn complex relationships in crime data. Hou et al. [15] suggested an attention-based graph model to learn the spatial-temporal connectivity and enhance the precision of the prediction by concentrating on the significant areas and periods. FIGAT, proposed by Xu et al., involves the combination of multi-source information through the graph attention method to predict crime risks [16]. Additionally, more complicated structures such as transformers and hybrids are also proposed. Butt et al. created START, a spatio-temporal autoregressive transformer to encode long-range dependencies in sequences of crimes [19], and Guo introduced a multimodal GCN-LSTM model that addresses urban dynamics [20]. Although these methods are useful in offering a considerable predictive performance, they often require extensive computing resources and involved feature engineering.

There are also some approaches aiming at enhancing interpretability and geographical context. Wang et al. [18] enhanced the model of urban areas by using building footprint data. Salcedo-Gonzalez et al. [17] proposed geo-visualization systems to visualize crime in real-time and aid decisions. But these methods are largely concerned with the spatial map, as opposed to the semantic connections underlying the map.

In spite of these developments, the existing approaches often fail to explain spatial and temporal interactions in a unified manner. Embedding based methods implicitly model time. DL models have time sequences but do not explicitly model time. In addition, the types of crimes are usually considered as unchanging objects, which do not consider the time variation. To address these limitations, we present a TA-CrimeVec structure, a hybrid of the continuous-temporal encoding and the spatial-temporal embedding to the joint representation of the criminal semantics and the time-varying dynamics.

## III. METHODOLOGY

### A) Proposed System

This paper provides a Time Aware CrimeVec (TA-CrimeVec) model to model and predict city crime trends using spatial-temporal embedding learning. The system is fed with raw crime data of Chicago Crime Dataset that contains information on the type of crime, date and location. The videos are pre-processed and turned into coded sequences based on the geographical locations and as per the chronological order. The system uses temporal encoding using Time2Vec, which converts timestamps into continuous vectors representation, capturing linear and periodic patterns. Such representations of time are combined with categories of crime to generate time-conscious tokens, which are used to establish contextual associations in a specified temporal window. An embedding model that is Word2Vec based is learned, obtaining dense vectors of these tokens with respect to semantic and temporal relationships. Lastly the learnt embeddings are used as an input to ML models as features to forecast crime related events like the likelihood of arrest.

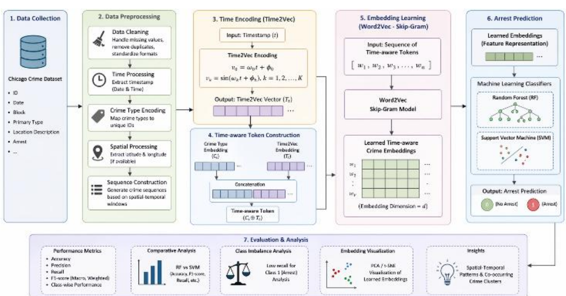

Fig.1 System Architecture

Crime data is collected followed by cleaning, normalizing timestamps and selecting features, as illustrated in Fig. 1. Time2Vec converts the time information and types of crime into time-sensitive tokens. A specific time window physically and temporally clusters these symbols into sequences of contexts. A Word2Vec model is trained with these sequences to encode the spatial-temporal relationships. Lastly, the generated embeddings are used to classify crime using classification algorithms such as RF and SVM.

B) Dataset Description
The data used in this paper is the Chicago Crime Dataset. This data set consists of vast amounts of documented crime rates in Chicago city since 2001 up to date. It is a massive dataset of millions of records and has a wide range of geographical and temporal data to research crime patterns. This research has a subset of significant attributes that are chosen, such as Primary Type, Date, Beat and Arrest. The type of crime is known as Primary Type, the date and time when the crime occurred is known as Date, the area where the crime occurred is known as Beat and the arrest information is known as Arrest. The geographical and temporal dimensions of the simulation of criminal behavior can be effectively performed using these features.

C) Data Preprocessing
Preparation of data to be used in effective spatial-temporal modeling and embedding learning is required to guarantee data quality. To guarantee the integrity of the data, first, missing and wrong values are removed. Drop records that have a null or invalid crime type, date, and location. The step enhances the accuracy and stability of the model through the analysis of complete and reliable data points only. The Date attribute is turned into a timestamp to allow time analysis. The time stamps are converted to numbers and then scaled down to a range. This normalization is dependable at the time values needed to learn the Time2Vec model. Lastly, due to the size of the collection, 800,000 records were selected at random. Sampling makes the computer less complex, but maintains the variation and distribution of the crimes.

D) Time Representation using Time2Vec
In order to capture more variations in crime data over time, we model time series with a Time2Vec-based approach that transforms timestamps into continuous vectors. In contrast to standard scalar time representations, Time2Vec can learn complex time relationships: it models both linear flow and periodicity of time. The Time2Vec function of a particular timestamp, denoted as time, is:

$$t2v(\tau)[0] = w_0\tau + \emptyset_0$$

$$t2v(\tau)[i] = \sin(w_i\tau + \emptyset_i)\,, i = 1,2,\ldots,k$$

where the parameters of the periodic components are learnable, i.e., the parameters omega i and phi i, and k is the number of periodic components. The terms of time and the sine terms are a linear term and periodic behaviour respectively.

Inclusion of sinusoidal functions is especially important as criminal events are often periodic, i.e. they have daily or weekly cycles. Time2Vec can establish deeper time representations by integrating linear and periodic features, which are essential in modeling time-varying criminal dynamics in our model.

E) Spatial-Temporal Sequence Construction
Beat is used to aggregate crime records to form spatial-temporal relationships. A Beat is a special geographical location. Timestamps within each region are used to represent the time order of events in order to maintain the time order. Sequences of ordered crime-time tokens are then mined to form sequences that record the temporal history of crime in each location. These sequences can be modeled as inputs to form contextual relationships and are the foundation of learning the spatial-temporal patterns in the subsequent stages.

F) Temporal Context Window Generation
In short-term temporal relationships, there exists a frame of context of -15 minutes on each criminal incidence. All the events in the vicinity that exist in the same spatial order and in the same time frame are selected as context to a specific target event. This uncovers those crimes that are time-linked and can have concealed trends. The target-context pairings they find are used to generate the training samples, such that the model can learn the successful temporal co-occurrence correlations of criminal events.

G) Time-Aware Crime Embedding (TA-CrimeVec Model)
The most important finding of the paper is the creation of a Time-Aware CrimeVec (TA-CrimeVec) model that incorporates criminal semantics with temporal dynamics in one embedding representation. Under the proposed method, every criminal event is transformed into a composite token and consists of the combination of its type of crime and Time2Vec representation. This process can be formulated as:

$$Token = Crime\,Type + Time2Vec(\tau)$$

This method is a sure way of ensuring each event is not only recorded on what happened, but also when it happened. The proposed model learns a joint representation, unlike the current methods which consider crime type and crime time as independent characteristics. Embedding can be defined as:

$$Embedding = f(Crime, Time)$$

The system can be used to model the relationships between the types of crimes and their chronological context by modelling them together. By comparison, more traditional CrimeVec methods only focus on the spatial-temporal co-occurrence of crime categories, without necessarily representing continuous time information. Moreover, traditional ML models treat time as an independent numerical variable, which cannot well represent periodicity and temporal correlations.

The suggested method combines Time2Vec and CrimeVec to generate time-sensitive embeddings of which a single type of crime may be expressed differently at different times. This enables the model to learn fine time scale dynamics, like daily or seasonal variations in criminal behavior, and significantly enhances the expressiveness and forecasting capability of the learned embeddings.

H) Word2Vec Training
We train the Skip-gram version of the Word2Vec model to find meaningful representations of crime-time tokens. The Skip-gram model is used to predict context tokens in a pre-defined window of a target token. In this case, tokens are associated with a category of crimes along with its time embedding, which enables the model to acquire spatial-temporal associations through prediction of contexts.

$$\max \sum_{t=1}^{T} \sum -c \leq j \leq c, j \neq 0 \log P\left(w_{t+j} | w_t\right)$$

Where, wt is the target token, wt+j are the context tokens within a window size c and T is the total number of tokens. The model maximizes the probability of observing context tokens given the target token.

During the training process the model is trained to change the representations of vectors in such a way that tokens in similar spatial-temporal contexts are similar in their embeddings. In this approach, offenses that are likely to occur in the identical locality and time period are positioned nearer in the embedding space, allowing to understand latent associations between separate crime groups and their time series.

I) Feature Extraction
Once the Word2Vec model has been trained, the learnt embeddings are retrieved on the crime-time token to give the feature representations to classify the crime. The embedding vector of the trained model is taken to produce a dense numerical representation of every record in the dataset, capturing criminal semantics and time. These embedding vectors are then summed up to produce the feature matrix X, with each row being a criminal event and each column being a dimension of the learnt embedding space. The target variable y is determined with the help of the Arrest property, which is coded to the binary format of either an arrest took place (1) or not (0). The structured representation allows prediction tasks to be done using supervised ML models.

J) Data Splitting
The model is tested on the dataset having a ratio of 80: 20 to split it into training and testing sets. To be more precise, 80 percent of the data are used to train the ML models, and the other 20 percent are used to test and evaluate the performance. We also use stratified sampling in the split to maintain the number of classes of the target variable (Arrest) in training and testing sets. This is especially applicable due to the lopsidedness of the incidences of arrests and non-arrests. The results of the evaluation are more valid and more reflective in the real-life situations when the proportions of the classes are maintained.

K) Classification Models
In order to evaluate the performance of the learnt time-conscious criminal embeddings we employ two popular models of supervised ML RF and SVM.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717355

These models are selected because of their strength, generalization capacity and the suitability in the high dimensional feature space.

Random Forest: RF is a method of ensemble learning that creates a large number of decision trees and combines their results to improve the accuracy of prediction and reduce overfitting. It is extremely effective in the modeling of complex non-linear interactions that are present in the embedding space. Moreover, its ability to work with large data volumes and provide high-quality performance makes it a suitable choice to use in crime prediction.

Support Vector Machine (SVM): SVM is a powerful classifier that identifies the optimal hyperplane to classify data points of different classes. It also performs well in high dimensional spaces and it does well when the number of features hugely out-numbers the number of samples. We test the linear separability of the learnt embeddings using the SVM. It is a good base to compare with.

This set of models allows the complete evaluation of the suggested embedding framework in the linear and non-linear classification perspectives.

## IV. EXPERIMENTAL RESULTS

### A) Experimental Setup

To assess it computationally efficiently and to have the data diversity, the experimental assessment uses 800,000 criminal records of the Chicago criminal Dataset. The data is divided into 8020 train/test split, stratified to keep the distribution of classes. The learnt embeddings of 50 dimensions can be very descriptive of spatial-temporal associations. The context sequences are designed to embed learning within a time frame of +15 and -15 minutes. To evaluate the prediction accuracy of the time-aware criminal embeddings to arrests, we use the RF and SVM to classify.

### B) Classification Performance Evaluation

i) Accuracy Comparison: The overall accuracy of the categorization models is provided in Table.1. We find that the random forest model is superior to the support vector machine model, and has a higher prediction accuracy on the test data set.

Table.1 Accuracy Comparison of Classification Models

| Model | Accuracy |
|---|---|
| Random Forest | 0.7869 |
| SVM | 0.7574 |

In Figure 2, it can be seen that the accuracy of the Random Forest model is higher compared to that of SVM model as indicated by the bar chart, implying that the Random Forest model is more appropriate to the learnt embedding properties.

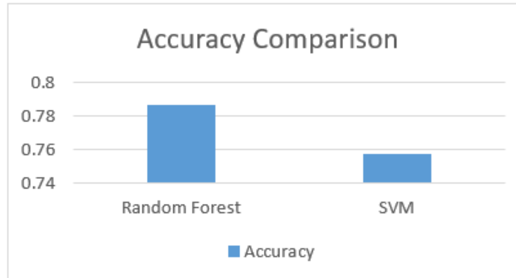


Fig.2 Accuracy Comparison

i) Class-wise Performance Analysis: Table.2 indicates the performance measures of all models i.e. accuracy, recall and F1-score per model per class. Class 0 is non-arrest cases whereas Class 1 is arrest cases. The results indicate that the two models get good recall on Class 0 but significantly worse recall on Class 1 and this brings to light the problems of the Class imbalance.

Table.2 Class-wise Performance Metrics

| Model | Precision (0) | Recall (0) | F1-score (0) | Precision (1) | Recall (1) | F1-score (1) |
|---|---|---|---|---|---|---|
| RF | 0.78 | 1.00 | 0.87 | 0.99 | 0.18 | 0.31 |
| SVM | 0.75 | 1.00 | 0.86 | 0.91 | 0.08 | 0.14 |

In the comparison of the Class 0 and Class 1 measures in Fig. 3, it is possible to notice that there is a significant imbalance in the performance of the model. The two models perform very well in Class 0, and significantly worse in Class 1, in terms of performance, especially in recall.

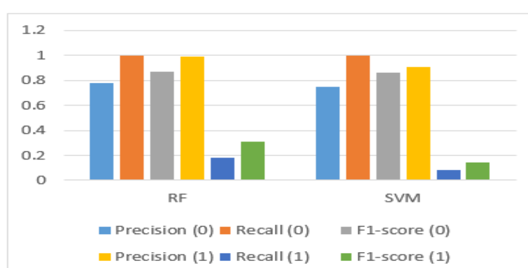


Fig.3 Class-wise Performance Comparison

iii) Overall Performance Metrics: The Table.3 displays the macro and weighted assessment measures of both models. Random Forest model demonstrates higher overall performance, compared to SVM, on all the evaluation criteria, which implies that it is useful when addressing the learnt embedding characteristics.

Table.3 Macro and Weighted Performance Metrics

| Model | Macro F1 | Weighted F1 | Macro Recall | Weighted Recall |
|---|---|---|---|---|
| RF | 0.59 | 0.73 | 0.59 | 0.79 |
| SVM | 0.50 | 0.67 | 0.54 | 0.76 |

This is further confirmed by the comparison of the macro and weighted measures in Fig. 4 which clearly shows that the Random Forest model is a consistent winner when compared to SVM. The variation in macro and weighted scores also reflects how the imbalance of classes influences a model of evaluation in general.

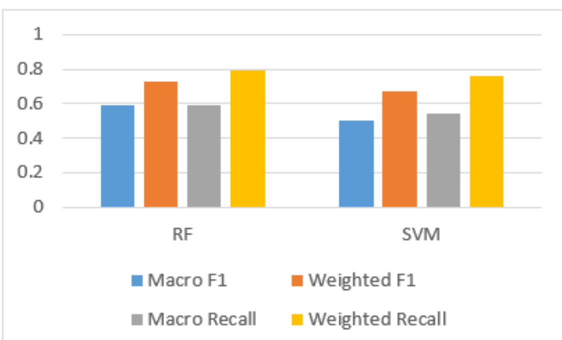


Fig.4 Macro vs Weighted Metrics Comparison

C) Comparative Analysis

The findings of the comparative analysis of the classification models indicate that the model of the SVM is never superior to the model of the RF in any of the performance measures. Based on Table 1-3, it is observed that RF is better to utilize the learnt time-conscious criminal embeddings, and achieves higher accuracy, F1-score and recall values compared to SVM.

The ability of the Random Forest to deal with non-linear relationships between the data is one of the primary causes of this difference in the performance. The proposed embeddings represent highly non-linear spatio-temporal associations between the nature of crime and time, therefore, the feature space is inherently non-linear. Since RF is a collection of decision trees, it is able to effectively characterize such non-linear patterns by dividing the feature space into a large number of areas, and thus enhances the classification performance.

The linear SVM model however, assumes linear decision boundary, and thus its ability to learn complex relationships in the embedding space is limited. Also SVM is more prone to the imbalance of the classes since it has a significantly lower recall of Class 1 (arrest cases). This implies that the model is biased towards the majority group and it does not identify the incidences of the minority group.

Overall, the findings indicate that time-aware criminal embedding models such as Random Forest are more well-equipped to utilize time-aware criminal embeddings, especially when non-linear feature interactions are involved and unbalanced datasets are used.

D) Class Imbalance Analysis

According to the classification results, two classes are much imbalanced in the model performance. Both models are very good in recalling Class 0 (non-arrest cases) but the recall of Class 1 (arrest cases) is very poor particularly in SVM model. This implies that there are numerous real arrest cases which are falsely identified as non-arrest cases. This is attributed to the unequal proportion of the sample with the non-arrest cases being more.

The practical effects of such an imbalance are far reaching. Practically, in cases where positive identification of arrest related cases is not possible, it could hinder law enforcement decision-making and resource allocation in crime prediction applications. A model that would be biased towards the majority group may do well in general, but not minority outcomes of interest. Therefore, it is essential to reduce the imbalance of classes to enhance the fairness of models and the ability of predictive algorithms to deliver significant and practical insights in practice.

E) Time-Aware Crime Embedding Analysis

i) Embedding Visualization: To visualize the structure of the learnt embeddings, Principal Component Analysis (PCA) is used to reduce the high dimensional embedding space to a two dimensional space so that it can be visualized. As Fig. 5 shows, there are significant patterns of distribution of the projected embedding space, i.e. similar patterns of crime-time events are more likely to be close to one another. This demonstrates that the proposed model will effectively be able to extract the latent relationships in the data.

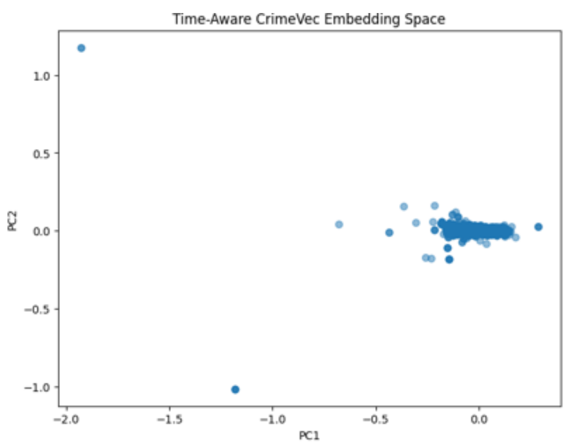


Fig.5 PCA Visualization of Time-Aware Crime Embedding Space

ii) Effect of Time2Vec: A significant finding using the formed embeddings is that the same form of crime would have different representations as a vector depending on the time when it occurred. To illustrate, the offense of burglary occurs on many occasions at different time-stamps, the various occurrences have varying embedding values. This difference can be attributed to the Time2Vec component that attempts to encode temporal information with the help of linear and periodic transformations. Thus, the model determines that the same crimes took place at various time through the dynamics of time which is absent in the traditional approach.

iii) Spatial-Temporal Insights: The learnt embeddings also capture significant spatial-temporal correlations of crimes. The embedding space is likely to cluster events that occur within close geographical distances, and time interval. This tendency of clustering behavior shows that the model can learn pertinent links among various types of crime and can also achieve co-occurrence patterns of crimes. The results are applicable in the research of crime dynamics, and can be used to make more informed decisions in predictive policing and city-safety research.

F) Discussion

The criminal embedding time-conscious strategy suggested can be utilized in a number of ways to observe the complex spatial-temporal relationships among the crime data. The model, Time2Vec and CrimeVec all build joint representations, capable of exploiting the nature of the crime and the time setting. This allows more efficient patterns detection and enhances the work of downstream classification models, which can be observed in the improved results of the RF classifier. Also, the method is generalized and can be applied to other event-based data. Nonetheless, these have some limits. The model predicts situations of the minority class very poorly, largely due to imbalance in the data set amongst the classes. In addition, fixed time frames might not be capable of detecting long-run relationships between criminal incidents. The embedding quality is also dependent on the parameter selection e.g. the window size as well as the embedding dimension. In the practical sense, the proposed solution can assist law enforcement agencies to detect crime trends, allocate resources effectively, and improve decision-making. Conversely, being able to capture both what and when aspects of crime provides a more holistic analytical approach to the real-life use of urban safety.

## V. CONCLUSION

In this paper, we introduce a new crime pattern analysis/arrest prediction framework based on the time-sensitive embedding solutions. The suggested model, TA-CrimeVec model uses CrimeVec and Time2Vec to co-train the representations that capture contextual associations between the crime of different types and the time-related features. We trained Word2Vec Skip-gram model spatial-temporal crime sequences and learnt embeddings on the Chicago Crime Dataset. The performance of these embeddings was evaluated on ML classifiers, and the RF model performed better than SVM. The findings indicate that temporal addition is highly beneficial in depicting criminal events, and enhances the classification accuracy. Besides, the embedded visualization identified significant grouping patterns that justify the ability of the model to acquire spatial-temporal correlations. To sum up, the proposed approach provides a strong base to examine criminal behavior and makes informed decision making on urban safety and law enforcement practices based on the data.

Future research could be on how to enhance the model performance on prediction of minority class by addressing class imbalance with advanced methods such as SMOTE or cost sensitive learning. The framework can be adapted to other contextual factors, including geographic coordinates, weather and socio-economic factors, which can enhance the quality of the embedding. In addition, more advanced DL models can be investigated, including Graph Neural Networks and models based on Transformers, to model more complex spatial-temporal relationships. The proposed embedding method can also be applied to the real-time crime monitoring systems that could be used to better analyze crime trends and improve decision making within law enforcement agencies.

REFERENCES

[1] H. Akaike, "Fitting autoregressive models for prediction," Annals of the Institute of Statistical Mathematics, vol. 21, no. 1, pp. 243–247, 1969.

[2] R. N. Bracewell, The Fourier Transform and Its Applications. New York, NY, USA: McGraw-Hill, 1986.

[3] L. Cohen, Time-Frequency Analysis. Upper Saddle River, NJ, USA: Prentice Hall, 1995.

[4] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[5] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[6] A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), 2017.

[7] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in Proc. ACM SIGKDD, 2016, pp. 855–864.

[8] I. M. Baytas et al., "Patient subtyping via time-aware LSTM networks," in Proc. ACM SIGKDD, 2017, pp. 65–74.

[9] Y. Li, N. Du, and S. Bengio, "Time-dependent representation for neural event sequence prediction," arXiv preprint arXiv:1708.00065, 2017.

[10] H. Mei and J. M. Eisner, "The neural Hawkes process: A neurally self-modulating multivariate point process," in NeurIPS, 2017.

[11] A. Crivellari and A. Ristea, "CrimeVec—Exploring spatial-temporal based vector representations of urban crime types and crime-related urban regions," ISPRS Int. J. Geo-Inf., vol. 10, no. 4, p. 210, 2021.

[12] U. M. Butt et al., "Spatio-temporal crime hotspot detection and prediction: A systematic literature review," IEEE Access, vol. 8, pp. 166553–166574, 2020.

[13] Y. Du and N. Ding, "A systematic review of multi-scale spatio-temporal crime prediction methods," ISPRS Int. J. Geo-Inf., vol. 12, no. 6, p. 209, 2023.

[14] Y. Qian et al., "GeST: A grid embedding based spatio-temporal correlation model for crime prediction," in Proc. IEEE DSC, 2020, pp. 1–7.

[15] M. Hou et al., "An integrated graph model for spatial–temporal urban crime prediction based on attention mechanism," ISPRS Int. J. Geo-Inf., vol. 11, no. 5, p. 294, 2022.

[16] W. Xu et al., "FIGAT: Accurately classify individual crime risks with multi-information fusion," IEEE Trans. Services Computing, vol. 16, no. 3, pp. 1890–1903, 2022.

[17] M. Salcedo-Gonzalez et al., "Spatiotemporal predictive geo-visualization of criminal activity for real-time systems," ISPRS Int. J. Geo-Inf., vol. 12, no. 7, p. 291, 2023.

[18] T. Wang, P. Chen, and M. Shan, "Crime spatiotemporal prediction through urban

region representation using building footprints," Big Data Cogn. Comput., vol. 9, no. 12, p. 301, 2025.

[19] U. M. Butt et al., "START: A spatiotemporal autoregressive transformer for enhancing crime prediction accuracy," IEEE Trans. Computational Social Systems, 2025.

[20] Y. Guo, "Multimodal spatio-temporal fusion: A generalizable GCN-LSTM with attention framework for urban application," Information Fusion, 2026.