From Classical Models to Artificial Intelligence Models: Prospects for Crime Prediction in the Era of Big Data

Mohammed Elseidi
Assistant Professor of Statistics
Umm Al Quwain University, Umm Al Quwain, UAE
drmohammed.elseidi@uaqu.ac.ae

Abstract

Accurate crime prediction is crucial for effective law enforcement and security, enabling proactive resource allocation and risk reduction. Criminal behavior is influenced by complex, diverse socio-economic factors, necessitating advanced models capable of extracting intricate patterns from large datasets. This research presents a **methodological and applied comparison** of four primary categories of time series forecasting models: Statistical Models (*AutoARIMA*), Machine Learning models (*AutoLightGBM*), Deep Learning models (*N-HiTS*), and Foundation Models (*TimeGPT*). The study's **innovation** lies in (1) integrating these diverse categories in a single comparative framework tailored for security decision-makers, (2) explicitly applying cutting-edge AI, particularly **Foundation Models (*TimeGPT*)** with pre-training on vast, multi-domain time series, for crime prediction for the first time, and (3) demonstrating a comprehensive application using daily crime data from Chicago (2017–2019), with the final month serving as a challenging test set for assessing robustness against sudden fluctuations. Results indicate that Foundation (*TimeGPT*) and Deep Learning (*N-HiTS*) models outperform in accuracy, effectively capturing nonlinear relationships and complex seasonalities. Statistical (*ARIMA*) and traditional ML (*LightGBM*) models offer greater interpretability and faster training but are less adept at handling unexpected surges. This comparative, automated approach offers a *practical solution* for security agencies seeking AI adoption without significant programming complexity. The research underscores time series modeling's role in enhancing security operations and explores new avenues for AI-driven proactive crime prevention using big data.

Keywords: Crime Prediction, Time Series, *ARIMA*, Foundation Models, Artificial Intelligence in Policing, Big Data, Deep Learning.

1 Introduction and Literature Review

Effective crime prediction is increasingly vital for strategic security planning and optimal deployment of police resources. Law enforcement agencies rely on timely and accurate forecasts to enhance patrol planning, allocate personnel and technology proactively, and ultimately mitigate risks to public safety. Building security policies on measurable, scientific foundations is paramount, particularly in complex urban environments like large cities, which are characterized by dynamic populations and continuous social and economic shifts. These factors contribute to the inherent irregularity and volatility often observed in daily crime data. Predicting criminal patterns in such contexts poses significant challenges, as underlying behaviors can change rapidly due to demographic changes, economic fluctuations, political events, or even targeted media campaigns.

Consequently, *predictive models* are essential tools for understanding potential future scenarios and transforming large volumes of historical crime data into actionable intelligence. Beyond simply providing numerical forecasts, security agencies require models that can be updated swiftly with incoming daily data and offer sufficient interpretability to gain the trust of decision-makers. While traditional methods like *ARIMA* provide good interpretability, advancements in Machine Learning (ML) and Deep Learning (DL) enable the discovery of more complex and nonlinear patterns than conventional statistical models can capture. Furthermore, the recent advent of **Foundation Models for Time Series** represents a potentially revolutionary step towards a single, general model capable of adaptation (*Fine-Tuning*) to various domains, including crime prediction, without the need for extensive training from scratch.

This paper presents a critical review and comparative analysis of prominent approaches for forecasting crime counts, categorized into four major groups. **Statistical models**, such as the widely used Autoregressive Integrated Moving Average (*ARIMA*) family [1, 2, 3, 4, 5, 6, 7, 8, 9], form a foundational approach to time series analysis, often assuming linearity and stationarity (or requiring transformation to achieve it). While offering interpretability and suitability for stable data, they can struggle with complex, non-linear patterns and require separate modeling for each series, making scaling challenging without automation tools like *AutoARIMA*.

Machine Learning methods, including ensemble techniques like *Gradient Boosting* (*LightGBM*, *XGBoost*) and *Random Forest* [10, 11, 12, 13, 14, 15, 16, 17, 18, 19], offer greater flexibility to capture nonlinear relationships and effectively integrate numerous external features (e.g., weather, events, demographics). These methods often excel when rich data is available and can unify modeling across multiple related series. However, they typically require significant feature engineering and may offer less direct interpretability compared to statistical models. Automated versions, like *AutoLightGBM*, mitigate the burden of hyperparameter tuning.

Deep Learning techniques [20, 21, 22, 23, 24, 25, 26, 27, 28, 29], such as *LSTM*, *GRU*, or hierarchical models like *N-HiTS*, have demonstrated remarkable ability to automatically learn complex, hidden patterns and handle long-term dependencies without extensive manual feature engineering. These models are particularly suited for large datasets with intricate, multi-frequency seasonalities, common in high-volume daily crime logs. Despite their power, they often require substantial computational resources (*GPU/TPU*) and fine-tuning, and their ”black-box” nature can pose challenges for interpretability, though research into explainable DL is ongoing.

Most recently, the concept of **Foundation Models**, popularized by large language models (*GPT*), has extended to time series with models like *TimeGPT* [30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41]. These models are pre-trained on massive, diverse time series datasets, learning general temporal representations and dynamics based on architectures like the *Transformer* [42]. This pre-training enables capabilities like *Zero-Shot Inference* on new series or efficient *Fine-Tuning* with limited local data, potentially offering high performance even in data-scarce scenarios or those with significant domain shifts. While promising, they share computational and interpretability challenges with other large deep models.

This paper aims to provide a unique, comprehensive comparison of these four model categories – *AutoARIMA*, *AutoLightGBM*, *N-HiTS*, and *TimeGPT* – within the specific domain of crime prediction. By evaluating these diverse approaches on a real-world dataset of daily crime incidents from Chicago, the study offers practical insights into their performance, strengths, and limitations, particularly their ability to adapt to the inherent volatility and complex patterns of urban crime data in the era of big data. The automated nature of the chosen model representatives facilitates evaluation by security practitioners, highlighting avenues for adopting advanced AI without needing deep expertise in model tuning.

2 Methodology

This study conducts an empirical comparison of four distinct time series forecasting approaches applied to daily urban crime data. The objective is to evaluate their relative performance and suitability for security applications, considering their underlying mechanisms, automation capabilities, and data requirements. The chosen models represent key paradigms in time series analysis, ranging from classical statistical methods to state-of-the-art AI techniques.

2.1 Statistical Models: AutoARIMA

The *AutoRegressive Integrated Moving Average (ARIMA)* model is a classical statistical approach for time series forecasting, capturing linear temporal dependencies through autoregressive (AR) and moving average (MA) components applied to a differenced series. The integration (I) component involves differencing (d) the series to achieve stationarity, where the mean and variance stabilize over time. A general representation applied to the differenced series $\nabla^d y_t$ can be expressed using the backshift operator B as:

$$\Phi(B)\,(1-B)^d\,y_t \;=\; \Theta(B)\,\varepsilon_t,$$

where $\Phi(B)$ and $\Theta(B)$ are polynomials representing the AR(p) and MA(q) components, and ε_t is white noise. The seasonal extension, *SARIMA*, incorporates seasonal AR, I, and MA components [1, 2, 4]. We utilize *AutoARIMA* [6, 7, 8], an automated method that employs algorithms like the Hyndman-Khandakar approach to search for the optimal $(p, d, q)(P, D, Q)_s$ order by minimizing information criteria (e.g., AIC, BIC), thereby simplifying model selection.

2.2 Machine Learning Models: AutoLightGBM

Machine learning models, particularly ensemble methods like Gradient Boosting, provide a flexible framework for modeling nonlinear relationships and integrating multiple exogenous variables as features. *LightGBM* is an efficient implementation of gradient boosting that constructs an ensemble of decision trees by minimizing a loss function iteratively. The predicted output $F(\mathbf{x})$ is an additive combination of individual tree predictions $f_k(\mathbf{x})$:

$$F(\mathbf{x}) \;=\; \sum_{k=1}^{K} f_k(\mathbf{x}).$$

LightGBM employs techniques like Leaf-wise tree growth and Histogram-based binning to optimize training speed and handle large datasets effectively [10, 11, 12]. For time series forecasting, the problem is typically framed as a regression task where lagged values of the target variable and relevant external factors are used as predictors [13, 14, 16]. Our implementation uses *AutoLightGBM*, which automates the critical process of hyperparameter tuning using methods such as Bayesian Optimization to identify the most effective configuration for the specific time series task.

2.3 Deep Learning Models: N-HiTS

Deep learning models are capable of learning complex, hierarchical representations and long-term dependencies in sequential data without explicit feature engineering. *N-HiTS* (*Neural Hierarchical Interpolation for Time Series*) [39] is a deep architecture based on the *N-BEATS* model, designed specifically for time series forecasting. It adopts a multi-block structure where each block processes the output of the previous one (residuals), performing both a partial forecast and a 'backcast' to explain the historical input. The final forecast $\widehat{\mathbf{y}}$ is the sum of the partial forecasts from all blocks:

$$\widehat{\mathbf{y}} = \sum_{b=1}^{B} \widehat{\mathbf{y}}^{(b)}.$$

This hierarchical decomposition allows *N-HiTS* to capture patterns at different temporal frequencies simultaneously, making it particularly effective for series with complex or nested seasonalities and handling non-stationary dynamics [20, 21, 22, 23, 25]. While powerful, deep learning models generally require significant amounts of data and computational resources for training.

2.4 Foundation Models for Time Series: TimeGPT

Foundation models represent a new paradigm, leveraging large-scale pre-training to develop models with broad capabilities transferable across various downstream tasks. *TimeGPT* [37, 38, 39, 40, 41] is a pioneering Foundation Model for time series, built upon the *Transformer* architecture [42]. The Transformer's core innovation is the self-attention mechanism, which allows the model to weigh the importance of different time steps in the input sequence when making predictions. The standard scaled dot-product attention is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V,$$

where Q, K, and V are matrices representing queries, keys, and values derived from the input sequence, and d_k is the dimension of the keys. *TimeGPT* is pre-trained on a massive and diverse dataset of time series, enabling it to learn generalized temporal patterns. This pre-training allows for Zero-Shot Inference on new series or efficient adaptation through Fine-Tuning on smaller, domain-specific datasets like crime data, offering potential advantages in generalization and handling volatile data.

3 Study Data and Experimental Setup

The empirical evaluation of the four selected time series forecasting models – AutoARIMA, AutoLightGBM, N-HiTS, and TimeGPT – was conducted using a dataset of daily crime incidents from the city of Chicago. This section outlines the data source and the experimental setup including the temporal partitioning for training and testing.

3.1 Data Description and Partitioning

The dataset comprises daily crime incident counts obtained from the City of Chicago Data Portal (data.cityofchicago.org). The study period spans from January 1, 2017, to December 31, 2019, providing a granular time series with daily temporal resolution. This frequency is crucial for capturing short-term fluctuations and event-driven anomalies characteristic of urban crime patterns. For the purpose of evaluating forecasting performance on unseen data, a strict chronological partition was applied. The period from January 1, 2017, through November 30, 2019, served as the training set, utilized for model parameter estimation and tuning. The subsequent month, December 2019, was reserved as the test set. This out-of-sample evaluation protocol is designed to simulate real-world deployment scenarios and specifically assess the models' generalization capability and robustness when predicting potentially volatile conditions, such as those often observed towards the end of the year.

3.2 Evaluation Metrics

The performance of the forecasting models was quantitatively assessed using two common metrics for time series prediction: the Root Mean Squared Error (*RMSE*) and the Symmetric Mean Absolute Percentage Error (*SMAPE*).

RMSE measures the square root of the average of the squared differences between actual values (y_i) and predicted values ($\hat{y}_i$):

$$\text{RMSE} = \sqrt{(1/n)\ \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2},$$

where n is the number of observations in the test set. *RMSE* provides a measure of the magnitude of errors and penalizes large errors more heavily.

SMAPE is a percentage error metric defined as:

$$\text{SMAPE} = (200\%/n)\ \sum_{i=1}^{n} \left(\left|\hat{y}_i - y_i\right|\ /\ \left|\hat{y}_i\right| + \left|y_i\right|\right).$$

SMAPE provides a relative error measure that is symmetric with respect to over- and under-forecasting, making it suitable for data with zeroes or near-zero values, although daily crime counts are generally positive. Lower values for both *RMSE* and *SMAPE* indicate better forecasting performance.

4 Data Analysis

This section presents an exploratory analysis of the daily crime incident time series for the city of Chicago over the period 2017–2019. The objective of this analysis is to characterize the dataset's key statistical properties, identify prominent temporal patterns, and assess underlying assumptions relevant to time series forecasting methodologies. Understanding these data characteristics is fundamental for guiding the selection and configuration of appropriate predictive models.

4.1 Descriptive Statistics

An initial statistical summary of the daily crime counts is provided in Table 1. These descriptive statistics offer insights into the central tendency, dispersion, and range of the dataset over the 731-day study period.

Metric	**Value**
Count	*731.000000*
Mean	*720.837209*
Standard Deviation	*80.713307*
Minimum	*271.000000*
First Quartile (*25%*)	*672.000000*
Median (*50%*)	*724.000000*
Third Quartile (*75%*)	*777.000000*
Maximum	*939.000000*

Table 1: Descriptive statistics for the daily crime counts over the *2017–2019* period.

The mean daily crime count over the study period is approximately 721, with a median of 724, indicating a distribution that is relatively symmetric around the central value. However, the standard deviation of approximately 81 highlights considerable variability in daily crime figures. The wide range, spanning from a minimum of 271 to a maximum of 939 incidents per day, further emphasizes the presence of significant fluctuations and occasional extreme values or outliers within the series. The difference between the first and third quartiles (672 to 777) suggests that while the majority of daily counts fall within a roughly 100-incident range, the presence of minimum and maximum values considerably outside this interquartile range indicates substantial day-to-day volatility.

4.2 Temporal Visualization

Visual inspection of the time series provides crucial insights into underlying patterns such as trends, seasonality, and irregular components.

4.2.1 Overall Time Series Trend

Figure 1 displays the daily crime count time series across the entire 2017–2019 period.

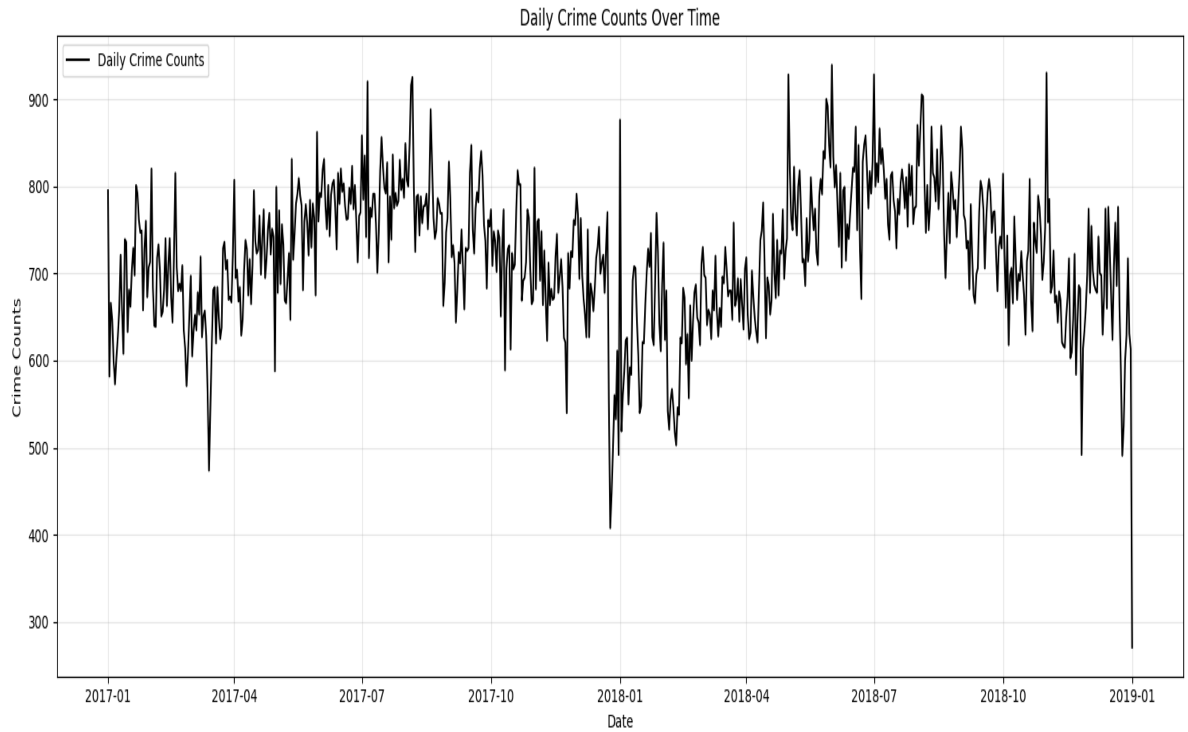

Figure 1: Evolution of daily crime counts in Chicago from 2017 to 2019.

The plot reveals notable temporal dynamics. Apparent peaks and troughs are visible throughout the series, suggesting the presence of cyclical or seasonal patterns. While a strong, consistent linear trend is not dominant across the entire three-year span, subtle shifts in the central tendency appear in different periods (e.g., a slightly higher baseline in parts of 2018 compared to late 2019). The pronounced spikes and dips underscore the data's volatile nature and the potential influence of specific events or time-dependent factors.

4.2.2 Rolling Statistics

To further investigate changes in the series' mean and variability over time, the rolling mean and rolling standard deviation were computed over a 30-day window. The results are presented in Figure 2.

The rolling mean (red line) exhibits clear fluctuations, confirming that the series' average level is not constant over time, which is indicative of non-stationarity. The rolling standard deviation (blue line) remains relatively stable during certain pe-

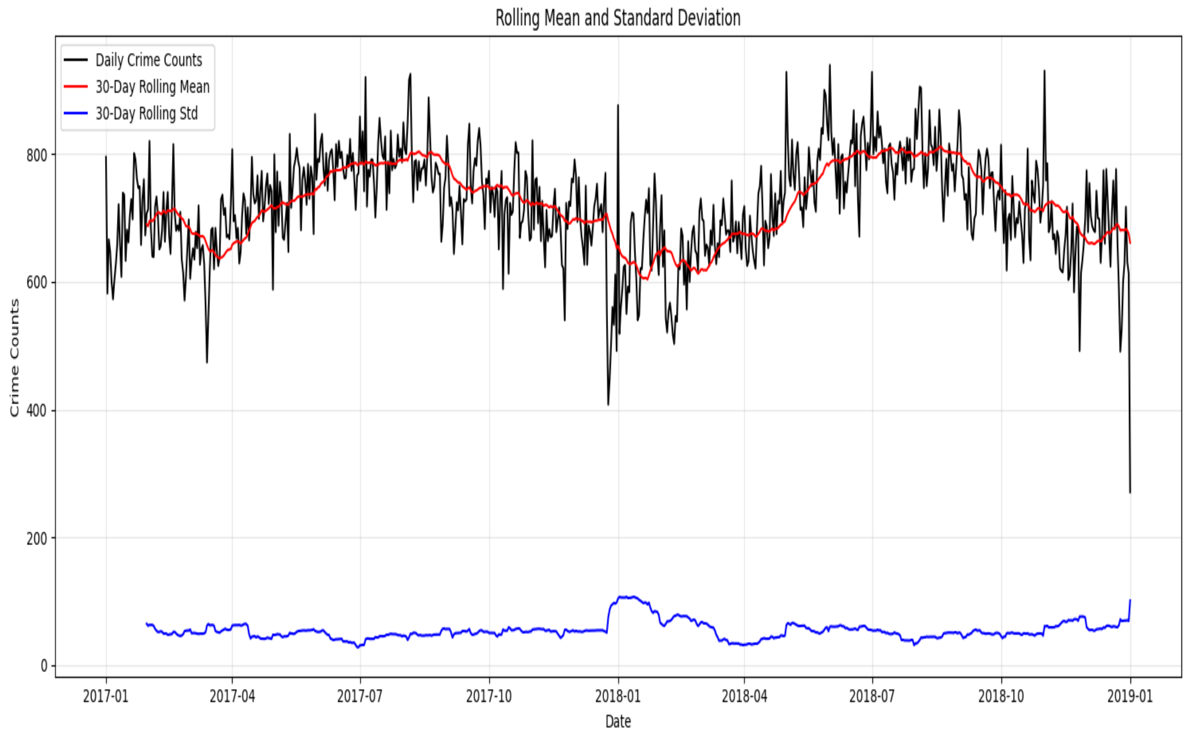

Figure 2: 30-day rolling mean (red) and rolling standard deviation (blue) of daily crime counts.

riods but shows noticeable increases in others (e.g., early 2018). These periods of elevated rolling standard deviation correspond to times of higher volatility in daily crime counts, suggesting that the degree of fluctuation also varies temporally.

4.2.3 Distributional Properties

The histogram and estimated probability density function provide insight into the overall distribution of the daily crime counts across the study period (Figure 3).

The distribution appears roughly unimodal, with a clear concentration of daily crime counts between 650 and 800 incidents. The peak of the distribution is centered around the mean/median value (720-730). The distribution exhibits a slight positive skew, characterized by a longer tail extending towards higher crime counts. This confirms the presence of infrequent days with significantly above-average crime figures, consistent with the high maximum value observed in the descriptive statistics.

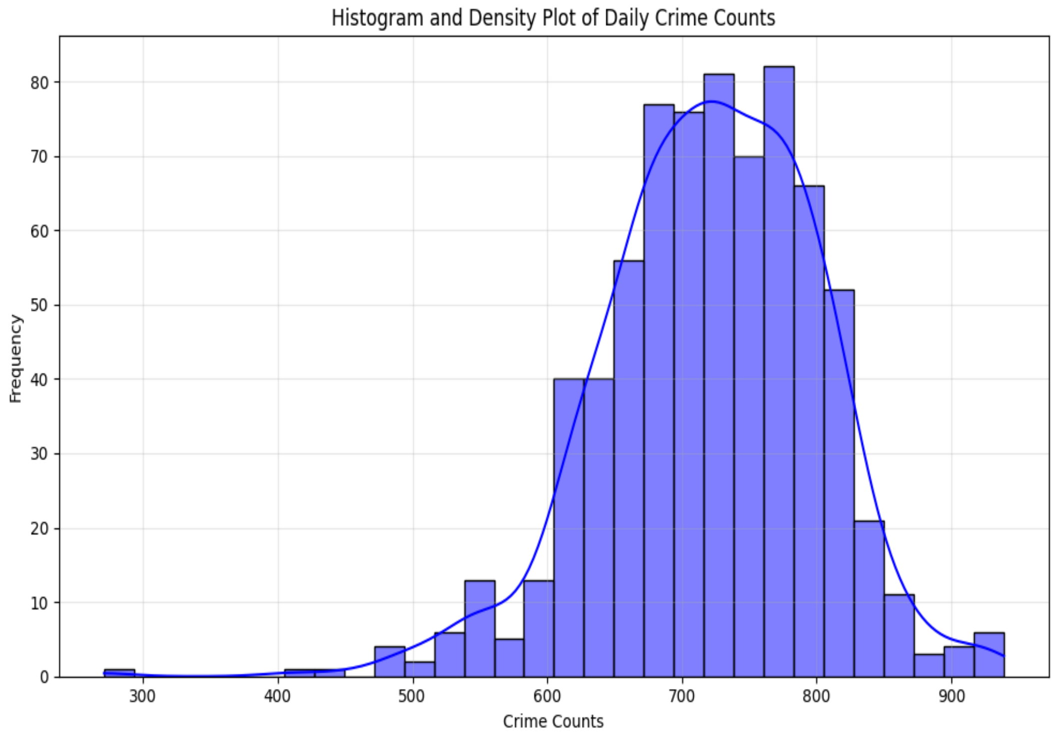

Figure 3: Histogram and estimated density function of daily crime counts (2017–2019).

4.2.4 Autocorrelation Analysis

The Autocorrelation Function (ACF) plot, shown in Figure 4, visualizes the correlation between the time series and lagged versions of itself, revealing the strength of temporal dependencies.

The ACF plot shows a high and statistically significant correlation at lag 1, indicating that today's crime count is strongly dependent on yesterday's count. The autocorrelation remains significant and decays gradually over several lags, characteristic of time series with trends or strong seasonal components. Notably, there appears to be elevated autocorrelation at lags corresponding to multiples of 7 (e.g., around lag 7, 14, etc.), suggesting a potential weekly seasonal pattern in the data. The persistence of significant autocorrelation over relatively long lags confirms that future crime counts are strongly influenced by past observations.

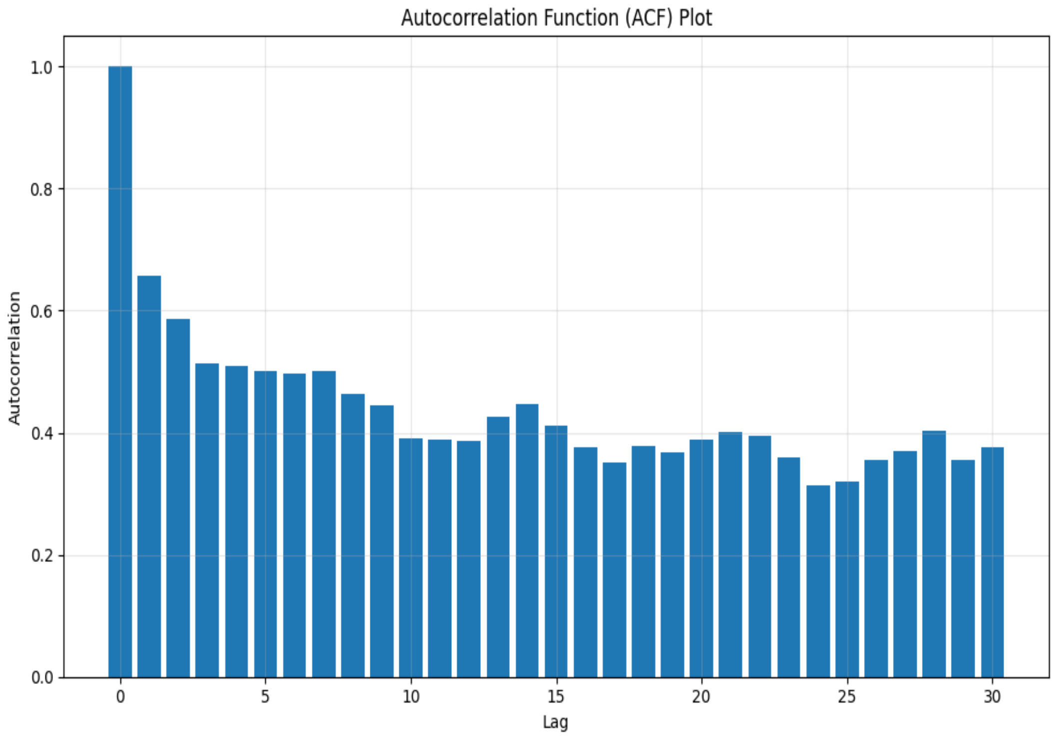

Figure 4: Autocorrelation Function (ACF) plot for the daily crime series.

4.3 Stationarity and Dependence Tests

Formal statistical tests were conducted to rigorously assess the stationarity and the overall presence of autocorrelation in the series.

The Augmented Dickey–Fuller (ADF) test was performed to test the null hypothesis that the time series has a unit root (i.e., is non-stationary). The test statistic was *-2.235* with a corresponding p-value of *0.194*. Since this p-value is greater than conventional significance levels (e.g., 0.05 or 0.10), we fail to reject the null hypothesis. This provides statistical evidence that the daily crime count time series is **non-stationary**, indicating that its statistical properties (mean, variance) change over time.

The Ljung–Box test was applied to test the null hypothesis that there is no overall autocorrelation in the series up to a specified number of lags. The results for selected lags are presented in Table 2.

For all tested lags, the Ljung-Box statistic is large and the p-values are exceed-

Lag	***LB Statistic***	***p-value***
10	*1931.856741*	< 0.001
20	*3089.880713*	< 0.001
30	*4106.061108*	< 0.001

Table 2: Ljung–Box test statistics and p-values for selected lags.

ingly small (<0.001). This strongly rejects the null hypothesis of no autocorrelation, providing significant evidence that the daily crime series exhibits substantial overall autocorrelation. This finding reinforces the observation from the ACF plot that values in the series are not independent over time and that past observations contain information predictive of future values.

4.4 Summary of Data Characteristics

In summary, the exploratory data analysis reveals several key characteristics of the Chicago daily crime time series (2017–2019). The series is characterized by considerable volatility and the presence of occasional high-magnitude events. Formal testing confirms that the series is non-stationary and exhibits significant temporal dependencies, including potential weekly seasonality as suggested by the ACF plot. These findings indicate that effective forecasting models must be capable of addressing non-stationarity, capturing complex temporal patterns, and ideally possessing some robustness to volatility and outliers.

4.5 Forecasting Performance Analysis

This subsection presents the forecasting results for each of the four evaluated models on the designated test set (December 2019). Visual comparisons of predicted versus actual values are shown, followed by a summary table of quantitative performance metrics.

4.5.1 AutoARIMA Results

Figure 5 displays the daily crime forecasts generated by the AutoARIMA model alongside the actual values for the test period (December 2019).

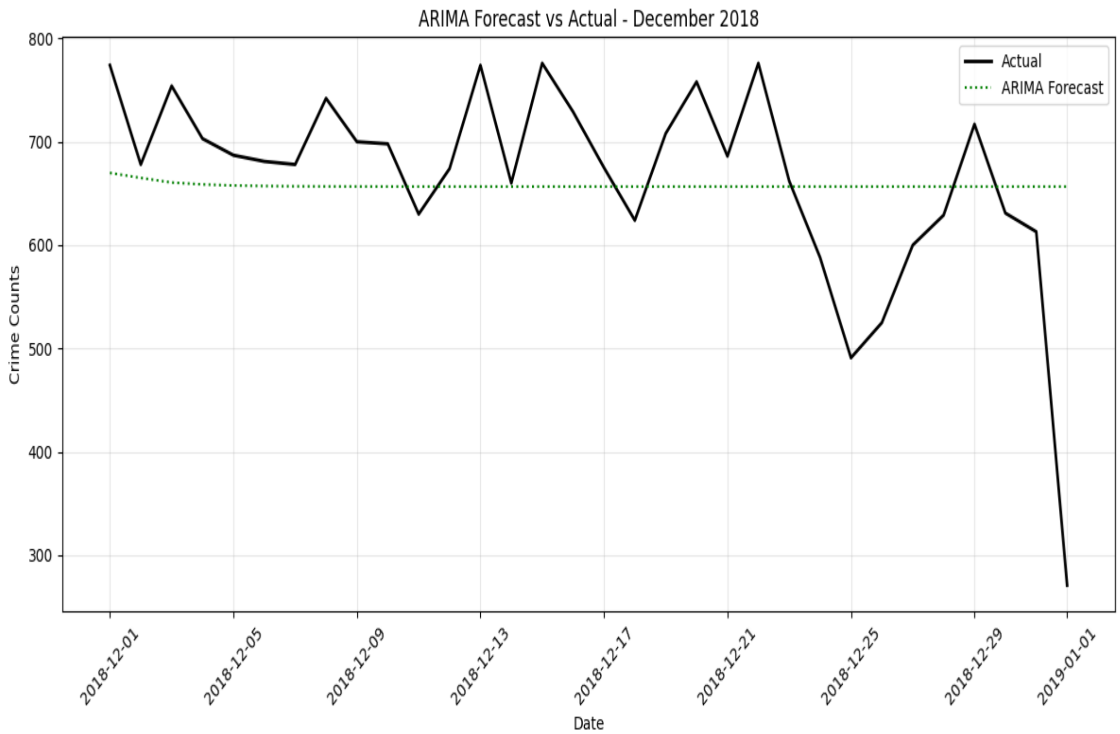

Figure 5: Comparison of AutoARIMA forecasts with actual crime values for December 2019.

The forecasts from AutoARIMA generally track the overall level of the series during stable periods. However, the model exhibits limitations in capturing the magnitude of sudden fluctuations and deviations from the average, as seen during periods of notable drops or increases in the actual crime count towards the end of the month. While AutoARIMA provides a clear and interpretable model structure, its inherent linearity and assumptions about stationarity after differencing may restrict its ability to fully adapt to the complex and potentially nonlinear dynamics present in volatile urban crime data.

4.5.2 AutoLightGBM Results

The forecasting performance of the AutoLightGBM model on the December 2019 test set is illustrated in Figure 6, comparing its predictions to the actual daily crime counts.

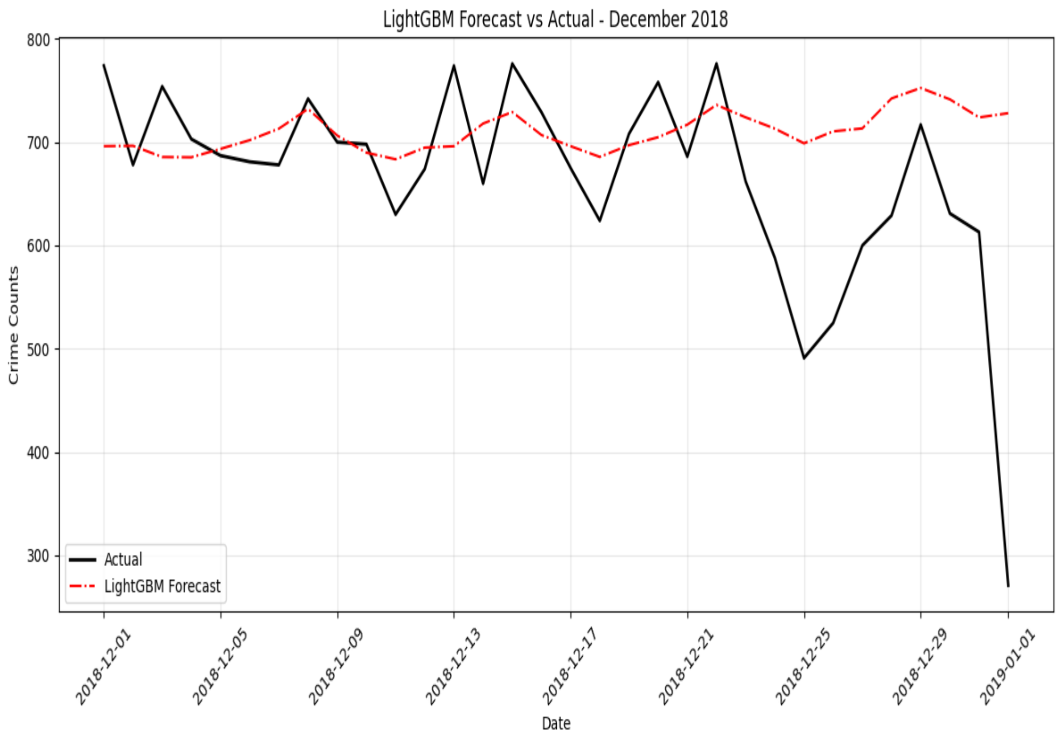

Figure 6: AutoLightGBM Forecasts vs. Actual Crime Values for December 2019.

AutoLightGBM generally adheres closely to the short-term trends and average level of the series. As a gradient boosting model, it is capable of capturing nonlinear relationships if relevant features, such as appropriate lags and external variables (e.g., weather, holidays), are provided. However, similar to AutoARIMA, its ability to anticipate and match the amplitude of sharp, unexpected peaks or troughs appears limited without specific exogenous information explaining such events. Its performance indicates suitability for capturing typical patterns but potential challenges in volatile periods lacking clear feature-based explanations.

4.5.3 N-HiTS Results

Figure 7 shows the comparison between the forecasts generated by the N-HiTS model and the actual daily crime counts for December 2019.

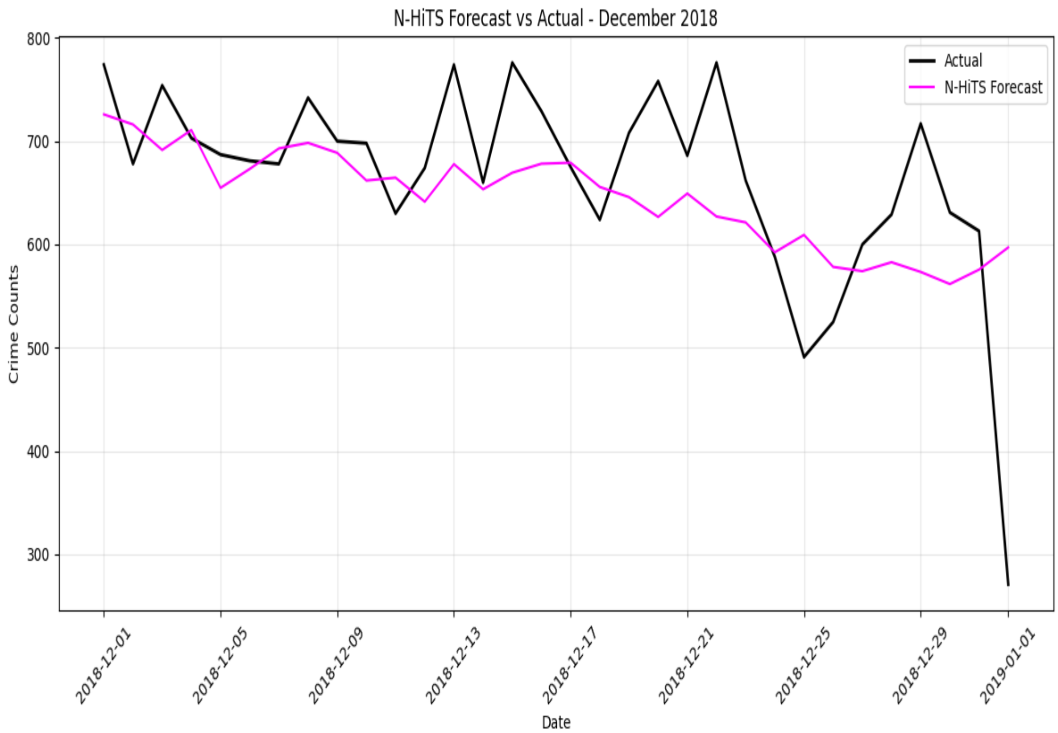

Figure 7: N-HiTS Forecasts vs. Actual Values in December 2019.

As a deep learning model with a hierarchical structure designed to capture multiple seasonalities and complex temporal dependencies, N-HiTS demonstrates a relatively strong ability to follow the overall trajectory of the series. It appears more flexible than the traditional models (AutoARIMA, AutoLightGBM) in responding to minor fluctuations. While it doesn't perfectly match the most extreme points, its predictions show better alignment with the general shape and shifts in the series compared to the simpler models, suggesting its effectiveness in capturing intricate patterns learned from the historical data. Its performance is promising for data with layered temporal structures.

4.5.4 TimeGPT Results

The forecasts from the Foundation Model, TimeGPT, for the December 2019 test period are presented in Figure 8, alongside the actual values and a 90% prediction interval (shaded blue area).

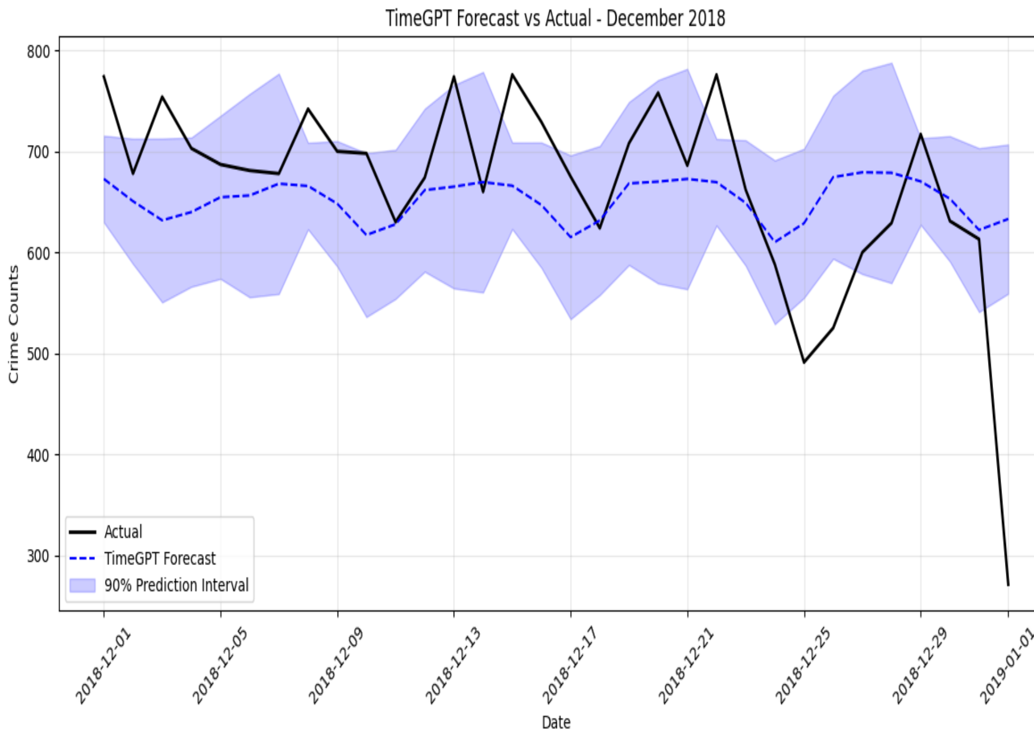

Figure 8: TimeGPT Forecasts vs. Actual Values (December 2019) with 90% Prediction Interval.

TimeGPT's forecasts generally capture the trend, although in this specific test period, the central forecast line appears to slightly underestimate the actual values on several days, clustering lower than the true counts observed around the 700-750 range. However, the model's pre-trained knowledge allows it to generate reasonable forecasts without extensive domain-specific fine-tuning. The wide prediction interval reflects the inherent uncertainty in forecasting a volatile series like daily crime data, especially during a potentially unusual period like year-end. Its performance is competitive, demonstrating the potential of large pre-trained models, although achieving peak accuracy may require more targeted fine-tuning on the specific dataset characteristics.

4.6 Model Performance Comparison

To quantitatively summarize the performance across the four models, Table 3 presents the Root Mean Squared Error (*RMSE*) and the Symmetric Mean Absolute Percentage Error (*SMAPE*) calculated on the December 2019 test set.

Model	***RMSE***	***SMAPE***
AutoARIMA	*97.870586*	*0.055060*
AutoLightGBM	*111.236106*	*0.056913*
N-HiTS	*86.245371*	*0.049266*
TimeGPT	*94.669446*	*0.053793*

Table 3: Quantitative performance comparison of the four models on the December 2019 test set.

The quantitative metrics confirm the visual observations. *N-HiTS* achieved the lowest *RMSE* (86.25) and *SMAPE* (0.049), indicating the highest accuracy among the evaluated models on this challenging test set. *TimeGPT* ranked second with an *RMSE* of 94.67 and *SMAPE* of 0.054, demonstrating competitive performance leveraging its pre-trained capabilities. AutoARIMA, while a strong baseline, performed slightly worse than N-HiTS and TimeGPT, with an *RMSE* of 97.87 and *SMAPE* of 0.055. AutoLightGBM recorded the highest errors (*RMSE* 111.24, *SMAPE* 0.057) on this particular test set, suggesting it might have struggled more than other models in capturing the specific dynamics or sudden shifts present in December 2019 data without the benefit of extensive manual feature engineering or deep temporal modeling inherent in N-HiTS and TimeGPT.

These results highlight the potential advantages of deep learning and foundation models in handling complex, non-linear time series with subtle patterns and volatility, while also demonstrating that simpler automated models like AutoARIMA can still offer reasonable performance and remain valuable for their interpretability and computational efficiency in different contexts.

5 Conclusion

The results of the comparison between statistical models (such as *AutoARIMA*), Machine Learning (*AutoLightGBM*), Deep Learning (*N-HiTS*), and Foundation Models (*TimeGPT*) showed that the *optimal choice* is determined by several key factors: data volume, complexity of seasonalities, frequency of sudden fluctuations, availability of computational resources, and the importance of interpretability for security agencies.

- **Statistical Models (*AutoARIMA*)**: Characterized by simplicity and interpretability, and show adequate performance when data is somewhat stable or when the security team needs clear knowledge of the influential coefficients.
- **Machine Learning Models (*AutoLightGBM*)**: Suitable when possessing multidimensional features (weather, events, social data...) with a desire for faster training than neural networks. However, they may fail to capture surges if not enhanced with abundant external data.
- **Deep Learning (*N-HiTS*)**: Showed a tangible advantage in cases of intertwined seasonalities and multiple cycles. It requires more historical data and computational resources, but achieves high accuracy in detecting complex curves.
- **Foundation Models (*TimeGPT*)**: Offer the possibility of *Zero-Shot* or *Fine-Tuning* when local data is available, and benefit from extensive ”prior knowledge.” However, their computational cost can be high and require strong infrastructure, especially in security institutions lacking high computing resources.

Consequently, there is no single ”one-size-fits-all” model for all security scenarios. Statistical models and traditional Machine Learning may be most suitable in environments with limited resources or a clear inclination towards **interpretability**, while deep and foundation models lead the way if **large data** is available, seasonal cycles are numerous, and external influences are intertwined.

References

[1] Li, J., Yang, X., Qian, T., and Xie, Q. (2023). The adaptive Fourier decomposition for financial time series. *Engineering Analysis with Boundary Elements*, 150, 139–153.
Scopus Link

[2] Wang, H., Al Tarawneh, L., Cheng, C., and Jin, Y. (2024). A decomposition-guided mechanism for nonstationary time series forecasting. *AIP Advances*, 14(1).
Scopus Link

[3] Ma, X., Hong, X., Lu, S., and Li, W. (2024). TS3Net: Triple Decomposition with Spectrum Gradient for Long-Term Time Series Analysis. In *Proceedings - International Conference on Data Engineering* (pp. 887–900).
Scopus Link

[4] Trull, O., García-Díaz, J.C., and Peiró-Signes, A. (2022). Multiple seasonal STL decomposition with discrete-interval moving seasonalities. *Applied Mathematics and Computation*, 433.
Scopus Link

[5] Dudek, G. (2023). STD: A Seasonal-Trend-Dispersion Decomposition of Time Series. *IEEE Transactions on Knowledge and Data Engineering*, 35(10), 10339–10350.
Scopus Link

[6] Mbuli, N., Mathonsi, M., Seitshiro, M., and Pretorius, J.-H.C. (2020). Decomposition forecasting methods: A review of applications in power systems. *Energy Reports*, 6, 298–306.
Scopus Link

[7] Wang, H., Cheng, C., and Jin, Y. (2022). Decomposition-based Ensemble Model for Non-stationary Time Series Forecast. In *IISE Annual Conference and*

Expo 2022.
Scopus Link

[8] Sebesta, V., Marsalek, R., and Pomenkova, J. (2013). The modified empirical mode decomposition method for analysing the cyclical behavior of time series. In *Proceedings - 27th European Conference on Modelling and Simulation, ECMS 2013* (pp. 288–292).
Scopus Link

[9] Das, S., Prusty, B.R., and Bingi, K. (2021). Review of adaptive decomposition-based data preprocessing for renewable generation rich power system applications. *Journal of Renewable and Sustainable Energy*, 13(6).
Scopus Link

[10] Thoiba Singh, N., Mehra, M., Verma, I., Singh, N., Gandhi, D., and Ahmad Alladin, M. (2024). Advancing Crime Analysis and Prediction: A Comprehensive Exploration of Machine Learning Applications in Criminal Justice. In *2nd International Conference on Intelligent Data Communication Technologies and Internet of Things, IDCIoT 2024* (pp. 1339–1343).
Scopus Link

[11] Adesola, F., Azeta, A., Misra, S., Oni, A., Ahuja, R., and Omolola, A. (2022). Spatial Analysis of Violent Crime Dataset Using Machine Learning. *Lecture Notes in Electrical Engineering*, 875, 183–191.
Scopus Link

[12] Sambasiva Rao, C., Gayathri, C.H., Gayathri, K., Venkata Ramaiah, J., and Prathyusha, O. (2024). Unveiling Patterns Using Machine Learning in Crime Prediction. *Lecture Notes in Networks and Systems*, 1095, 133–147.
Scopus Link

[13] Dhakksinesh, A., Katherine Olivia, R., Pooja, V.S., and Saravanan, S. (2023). Crime Analysis and Prediction Based on Machine Learning Algorithm. *Advances in Science and Technology*, 124, 549–555.
Scopus Link

[14] Sabakhtarishvili, Z., Panday, S., Jensen, C., and Ghosh, R. (2023). Crime Prediction Using Machine Learning: The Case of the City of Little Rock. In *HORA 2023 - 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*.
Scopus Link

[15] Biswas, A.A., and Basak, S. (2019). Forecasting the Trends and Patterns of Crime in Bangladesh using Machine Learning Model. In *2019 2nd International Conference on Intelligent Communication and Computational Techniques, ICCT 2019* (pp. 114–118).
Scopus Link

[16] Khatun, S., Banoth, K., Dilli, A., Kakarlapudi, S., Karrola, S.V., and Babu, G.C. (2023). Machine Learning based Advanced Crime Prediction and Analysis. In *2nd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2023 - Proceedings* (pp. 90–96).
Scopus Link

[17] Kim, S., Joshi, P., Kalsi, P.S., and Taheri, P. (2018). Crime Analysis Through Machine Learning. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2018* (pp. 415–420).
Scopus Link

[18] Mukherjee, A., and Ghosh, A. (2020). Heterogeneous Decomposition of Predictive Modeling Approach on Crime Dataset Using Machine Learning. *Learning and Analytics in Intelligent Systems*, 12, 1004–1012.
Scopus Link

[19] Saini, I.S., and Kaur, N. (2023). The Power of Predictive Analytics: Forecasting Crime Trends in High-Risk Areas for Crime Prevention using Machine Learning. In *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*.
Scopus Link

[20] Zhao, D., Du, P., Liu, T., and Ling, Z. (2023). Spatio-temporal Distribution Prediction Model of Urban Theft by Fusing Graph Autoencoder and GRU. *Journal of Geo-Information Science*, 25(7), 1448–1463.
Scopus Link

[21] Mishra, R.K., Ansari, A.R., Jothi, J.A.A., and Mishra, V. (2024). Analysis of Criminal Landscape by Utilizing Statistical Analysis and Deep Learning Techniques. *Journal of Applied Security Research*, 19(4), 560–585.
Scopus Link

[22] Tamilselvi, M., Ravi, C.N., Jayasudha, S., Shahad, P., Karthikeyan, S., and Samuthira Pandi, V. (2024). HDLCP: Experimental Analysis and Development of Hybrid Deep Learning Methodology for Crime Scenario Assesment and Prediction. In *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing, IC-CGU 2024*.
Scopus Link

[23] Sun, J., Yue, M., Lin, Z., Yang, X., Nocera, L., Kahn, G., and Shahabi, C. (2021). CrimeForecaster: Crime Prediction by Exploiting the Geographical Neighborhoods' Spatiotemporal Dependencies. *Lecture Notes in Computer Science*, 12461 LNAI, 52–67.
Scopus Link

[24] Tasnim, N., Imam, I.T., and Hashem, M.M.A. (2022). A Novel Multi-Module Approach to Predict Crime Based on Multivariate Spatio-Temporal Data Using Attention and Sequential Fusion Model. *IEEE Access*, 10, 48009–48030.
Scopus Link

[25] Ilhan, F., Tekin, S.F., and Aksoy, B. (2020). Spatio-Temporal Crime Prediction with Temporally Hierarchical Convolutional Neural Networks. In *2020 28th Signal Processing and Communications Applications Conference, SIU 2020 - Proceedings*.
Scopus Link

[26] Butt, U.M., Letchmunan, S., Hassan, F.H., and Koh, T.W. (2024). Leveraging transfer learning with deep learning for crime prediction. *PLoS ONE*, 19(4 April).
Scopus Link

[27] Pirajén, F., Fajardo, A., and Melgarejo, M. (2019). Towards a Deep Learning Approach for Urban Crime Forecasting. *Communications in Computer and Information Science*, 1052, 179–189.
Scopus Link

[28] Isafiade, O., Ndingindwayo, B., and Bagula, A. (2021). Predictive Policing Using Deep Learning: A Community Policing Practical Case Study. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST)*, 361, 269–286.
Scopus Link

[29] Jin, G., Wang, Q., Zhu, C., Feng, Y., Huang, J., and Zhou, J. (2020). Addressing crime situation forecasting task with temporal graph convolutional neural network approach. In *Proceedings - 2020 12th International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2020* (pp. 474–478).
Scopus Link

[30] Liu, Y., Qin, G., Huang, X., Wang, J., and Long, M. (2024). AutoTimes: Autoregressive Time Series Forecasters via Large Language Models. In *Advances in Neural Information Processing Systems*, 37.
Scopus Link

[31] Liu, Q., Liu, X., Liu, C., Wen, Q., and Liang, Y. (2024). TIME-FFM: Towards LM-Empowered Federated Foundation Model for Time Series Forecasting. In *Advances in Neural Information Processing Systems*, 37.
Scopus Link

[32] Mulayim, O.B., Quan, P., Han, L., Ouyang, X., Hong, D., Bergés, M., and Srivastava, M. (2024). Are Time Series Foundation Models Ready to Revolutionize Predictive Building Analytics? In *BuildSys 2024 - Proceedings of the 2024 11th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (pp. 169–173).
Scopus Link

[33] Jin, M., Wang, S., Ma, L., Chu, Z., Zhang, J.Y., Shi, X., Chen, P.-Y., Liang, Y., Li, Y.-F., Pan, S., and Wen, Q. (2024). TIME-LLM: TIME SERIES FORECASTING BY REPROGRAMMING LARGE LANGUAGE MODELS. In *12th International Conference on Learning Representations, ICLR 2024*.
Scopus Link

[34] Chen, H., Rossi, R.A., Kim, S., Mahadik, K., and Eldardiry, H. (2024). A Study of Foundation Models for Large-scale Time-series Forecasting. In *Proceedings - 2024 IEEE International Conference on Big Data, BigData 2024* (pp. 1255–1259).
Scopus Link

[35] Xue, H., and Salim, F.D. (2024). PromptCast: A New Prompt-Based Learning Paradigm for Time Series Forecasting. *IEEE Transactions on Knowledge and Data Engineering*, 36(11), 6851–6864.
Scopus Link

[36] Saravanan, H.K., Dwivedi, S., and Arjunan, P. (2024). Poster abstract: Analyzing the Performance of Time Series Foundation Models for Short-term Load Forecasting. In *BuildSys 2024 - Proceedings of the 2024 11th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (pp. 237–238).
Scopus Link

[37] Garza, A., Challu, C., and Mergenthaler-Canseco, M. (2024). *TimeGPT: The first foundation model for time series*. arXiv:2310.03589v3 [cs.LG].

[38] Oreshkin, B.N., Carpov, D., Chapados, N., and Bengio, Y. (2021). *Meta-learning framework with applications to zero-shot time-series forecasting*. In AAAI.

[39] Challu, C., Olivares, K.G., et al. (2023). *Neural Hierarchical Interpolation for Time Series Forecasting*. In AAAI.

[40] Wu, W., Xu, J., et al. (2022). *TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis*. arXiv preprint.

[41] Zhou, Z., Ma, Z., Wen, Q., Wang, X., Sun, L., and Jin, R. (2022). *FEDformer: Frequency Enhanced Decomposed Transformer for Long-Term Series Forecasting*. In ICML.

[42] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.