ORIGINAL ARTICLE

Exploratory data analysis, time series analysis, crime type prediction, and trend forecasting in crime data using machine learning, deep learning, and statistical methods

Esen Gül İlgün[1] · Murat Dener[1]

Received: 24 September 2024 / Accepted: 12 February 2025 / Published online: 20 March 2025

Abstract

Criminal activities are a critical obstacle to socioeconomic development and must be controlled. However, human surveillance-based control methods are prone to error, raise legal concerns, and necessitate the development of more robust alternatives. This study aims to contribute to the development of strategies for reducing and preventing crime by ensuring the optimal allocation of police resources to locations at the right time. To achieve this goal, crime datasets from three of the most metropolitan cities in the USA—San Francisco, Chicago, and Philadelphia—were subjected to comprehensive preprocessing and exploratory data analysis. The analysis identified the most reliable and dangerous months, days, and hours in terms of the frequency of criminal incidents, the most common types of crimes, and the police districts with the highest crime rates. Crime-type prediction models were developed using machine learning algorithms, including XGBoost, CatBoost, random forest (RF), decision tree (DT), multilayer perceptron (MLP), K-nearest neighbors (KNN), Gaussian Naive Bayes (GNB), and logistic regression (LR). Additionally, time series analyses were conducted in 10, 22, and 22 different police districts for the three datasets, respectively, using deep learning models such as long short-term memory (LSTM) and bidirectional long short-term memory (BLSTM) and statistical methods such as Holt–Winters exponential smoothing (HWES), Prophet, and seasonal autoregressive integrated moving average (SARIMA). The primary objective was to accurately predict future high-crime hot spots. Furthermore, crime trend forecasts for the next 5 years were made using the best models, based on the lowest root-mean-squared error (RMSE) values obtained through statistical methods. By combining traditional machine learning methods, deep learning approaches, and statistical techniques, this study analyzed criminal incidents from various perspectives, including crime-type prediction, regional crime prediction, trend forecasting, and exploratory data analysis. The results obtained are expected to contribute to the development of proactive policing strategies.

Keywords Crime data · Crime-type prediction · Trend forecasting · Time series analysis · Visualization

1 Introduction

Criminal activities, alongside their economic impact, undermine the fundamental sense of security, thereby reducing quality of life and societal well-being. This, in turn, constitutes a critical obstacle to the socioeconomic development of a society. Therefore, all strategies aimed at preventing or mitigating criminal activities are of

https://doi.org/10.1007/s00521-025-11094-9

utmost importance. In this context, crime prediction involves identifying variables related to crime, establishing connections among them, and forecasting where and when specific types of crimes might occur, whether criminal activities will recur, and even which individuals may be prone to committing or repeating crimes. This process serves as a foundation for developing strategies to prevent or reduce crime.

To delve deeper into the concept of crime and its associated variables, crime has two fundamental definitions from an etymological perspective: an act that can be prosecuted by the state and is punishable by law; and an act that, while not necessarily illegal, is considered immoral or shameful [1]. Three primary crime theories are often discussed: the routine activity theory, rational choice theory, and crime pattern theory. Routine activity theory posits that most crimes occur due to the convergence of three elements: motivated offenders, suitable targets, and the absence of capable guardians at a specific time and place [2]. Rational choice theory argues that offenders' decisions regarding location, targets, and methods can be explained by a rational balance of effort, risk, and reward [3]. The crime pattern theory integrates routine activity theory with rational choice theory to better explain the spatial distribution of crime events. According to this theory, individuals form "cognitive maps" and "activity spaces" through their daily routines. Potential offenders, using their cognitive maps, tend to select specific locations within relatively familiar areas to commit crimes [4]. In short, offenders are inclined to avoid unfamiliar places and choose locations where the "opportunity for crime overlaps with their cognitive domain," based on their rational capabilities. These locations often become crime hotspots due to their evident "crime-producing" or "crime-attracting" characteristics. Therefore, in predicting crime hotspots, it is essential to consider not only historical crime data, but also environmental factors [5]. Moreover, the occurrence of criminal activities is associated with various factors such as socioeconomic status, education level, ethnicity, age, gender inequality, child labor, housing stability, and the level of urbanization. The relationships among these factors are dynamic, shaped by the interaction of time, space, and individuals. Studies on crime analysis have also confirmed that crimes are unevenly distributed across time and space [6]. These complexities make it challenging to identify the variables influencing criminal activities and establish accurate connections between them using traditional methods, thereby complicating the development of strategies to prevent or mitigate crime. The success of the crime prediction process is critical for optimizing the allocation of police resources, reducing police response times, preventing the occurrence and recurrence of crimes, and ensuring fairness in decisions related to offenders. Crime prediction methods may involve police surveillance techniques such as monitoring suspects' phone conversations, using body cameras to record unusual and illegal activities, creating city maps, investigating crime scenes and accidents, managing traffic flow, and employing drones [7]. However, these methods based on human surveillance are prone to errors, making it challenging to establish accurate connections among the wide range of variables that could contribute to criminal activity, and consequently, to make precise predictions. Moreover, the use of these surveillance tools raises ethical and legal concerns. In summary, there is a need for a method that can simultaneously and effectively integrate various surveillance approaches, accurately identify crime-related variables, establish correct relationships among them, and incorporate the strengths of all methods without raising ethical or legal issues [8]. As data collection becomes easier and computer algorithms grow more complex, more advanced predictive technologies will be developed. Predictive policing will transform policing strategies across the country. Real-time reporting, professional crime analysts, and expanding computational capabilities are turning daily incidents, reports, and human tragedies into measurable and actionable data [9]. Big data analytics is a versatile technology that offers opportunities to control undesirable events that may occur in the future. Beyond its common applications in fields such as commerce, critical infrastructure, and industry, it has also been successfully applied in psychosocial domains like criminal activities, where establishing relationships among variables is relatively more complex. When a successful model is developed, it can have a rapid and broad positive impact. Promising results have been achieved in these areas. Crime analysis with big data analytics involves examining mathematical relationships among crime data to draw conclusions and serves as an essential auxiliary component in preventive policing [10]. Although crime analysis has been conducted for a long time, professional efforts began in the 1990s with the development of hardware and software technologies that enabled data collection and storage [11]. Today, crime analysis and prediction are conducted using crime data in several countries, including various states in the USA, Canada, India, and China. The San Francisco Police Department updates its crime prediction models with new data every eight hours and actively uses these models for preventive policing [12]. The cycle of increasing crime rates due to rapid urbanization and socioeconomic transformations, combined with the challenge of maintaining order with limited resources in such a complex environment, has made an effective solution essential. In recent years, there has been a surge in crime prediction studies utilizing the vast amount of crime data made available by big data analytics and advancing technology. Crime prediction, particularly in densely populated cities with high levels of urbanization and crime, can help prevent criminal activities by optimizing the allocation of police resources to specific areas at specific times [13]. However, several challenges hinder crime prediction and analysis from big data. These include the difficulty of structuring semistructured and unstructured crime data derived from police or social media reports, the necessity of updating crime prediction and analysis models due to the dynamic nature of crime [14], the requirement for extensive crime data to develop successful models, and the substantial computational cost associated with building models on large datasets and optimizing their hyperparameters.

The aim of this study is to understand which variables are associated with criminal acts and to observe changes in crime rates when these variables are controlled, enabling the optimal distribution of police resources across locations, reducing police response times, and contributing to the development of crime prevention strategies. To achieve this goal, exploratory data analysis, crime-type detection, general and regional time series analysis, and future trend forecasting were conducted using data preprocessing techniques, traditional machine learning and deep learning methods, and statistical methods on crime datasets from the three largest metropolitan cities in the USA: San Francisco, Chicago, and Philadelphia. The key contributions of this study can be summarized as follows:

(1) Three different and comprehensive datasets were used, containing 15 years of crime reports from San Francisco, 22 years of crime reports from Chicago, and 17 years of crime reports from Philadelphia. To contribute to the development of preventive policing strategies, a thorough preprocessing approach was applied to the three datasets, including the extraction of common attributes, removal of missing and duplicate values, creation of new attributes from timestamp and coordinate data that were predicted to have a dominant effect on the number of crime incidents, categorical data encoding, scaling, and feature selection to maximize performance from the models to be built.
(2) Visualizations were created for time, location, and crime type in all three datasets, allowing for the examination of crime trends by year, month, week, day, hour, and police district. The top ten most frequent types of crime in each of the three cities were also visualized.
(3) Crime-type prediction models were built using eight different machine learning algorithms—XGBoost, CatBoost, RF, DT, MLP, KNN, GNB, and LR—for 37 different crime types in the San Francisco dataset, 34 different crime types in the Chicago dataset, and 33 different crime types in the Philadelphia dataset.
(4) Monthly, weekly, and daily time series analysis was conducted using LSTM and BLSTM deep learning models, as well as statistical methods for time series analysis and trend forecasting until 2029 using HWES, Prophet, and SARIMA statistical methods.
(5) To identify high-crime hotspot areas, five different methods—three statistical (SARIMA, Prophet, HWES) and two deep learning (LSTM, BLSTM) methods—were applied to conduct regional time series analysis for ten police districts in San Francisco, 22 police districts in Chicago, and 22 police districts in Philadelphia. The performance of these methods was compared.

In Sect. 2 of the study, a literature review on crime analysis and forecasting is conducted, and relevant studies are summarized. In Sect. 3, the methodology proposed in the study is detailed, including data preprocessing applied to the dataset, exploratory data analysis, crime-type detection models, general and regional time series analysis, and statistical methods, machine learning, and deep learning techniques used for trend forecasting. The results obtained are then presented. In Sect. 4, the analysis and evaluation of the application results are carried out, and recommendations are provided.

2 Literature review

Crime analysis and forecasting involve identifying variables related to criminal acts, spatial and temporal crime forecasting, predicting the risk of reoffending, and forecasting the potential for individuals with certain characteristics to commit crimes. These activities contribute to the development of crime prevention strategies, making fair decisions in pre-trial, parole, probation, and even sentencing stages, reducing police response times by optimally distributing patrols and other resources across locations, and ultimately contributing to the formation of a healthy society. In recent years, the use of statistical methods, machine learning, and deep learning techniques to develop successful crime prediction models has been a significant research topic. Some studies in this area can be summarized as follows:

An interactive map based on Google Maps was created by applying various preprocessing techniques to the crime datasets from San Francisco, Chicago, and Philadelphia, which enabled the prediction of streets with a high likelihood of criminal activity at specific dates and times. It was found that Prophet and LSTM outperformed traditional neural networks in predicting crime trends, with the most successful prediction made using a training dataset from a 3-year time period [15]. In a study conducted to predict the crime potential of drug-addicted youth, an anonymized dataset consisting of 502 records and 31 attributes collected from four different rehabilitation centers was used. The model, trained with a dataset containing 24 attributes selected using the Chi-square method and built using the MLP algorithm, achieved the highest accuracy and F1-score. Additionally, using Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME) algorithms, it was concluded that the most relevant attributes to criminal behavior included not following a regular routine, health deterioration, staying out all night, environment and bad friends, stealing money, and drug use [16]. In a study aimed at developing strategies to prevent white-collar crimes by predicting the likelihood of municipalities being exposed to corruption crimes, a dataset of white-collar crimes from 2008 to 2014 was used, achieving an approximately 80% prediction accuracy [17]. In a study that approached algorithmic risk assessment as a classification problem to predict the likelihood of reoffending during pre-trial, parole, probation, and even sentencing stages, models created using black-box and interpretable machine learning algorithms on datasets containing reoffending information from the states of Florida and Kentucky achieved more successful results compared to the Arnold PSA and COMPAS methods, which are used in the justice system to predict the likelihood of criminal behavior before trial. However, due to the low generalization success of the models across regions, it was concluded that updates need to be made for different locations [18]. A deep learning technique based on attention and sequential fusion, consisting of feature-level and decision-level fusion stages, was proposed for spatial–temporal crime forecasting. Attributes related to the type, time, location of the crime, and the responsible police department were selected from the San Francisco and Chicago crime datasets, with weather data also separately obtained. Models built with the San Francisco and Chicago datasets achieved mean absolute errors of 0.008 and 0.02, and symmetric mean absolute error percentages of 1.03% and 0.6%, respectively [19]. In a study using five different datasets and three different deep learning architectures to analyze spatial and temporal patterns that change based on precedence, succession, and parallelism, the deep learning architecture in which spatial attributes were extracted first, followed by temporal attributes, yielded the most successful prediction results [20]. In another study, in the first stage, a periodic integral mapping method was used to address irregularities in the temporal dimension of the data, and a cubic spline interpolation super-resolution technique was applied to address irregularities in the spatial dimension. In the second stage, models were created using ST-ResNet, and the predictive performance of the models was tested with real-time crime data. Due to the preprocessing, where spatial information is no longer isolated and crime hotspot transitions can be captured, the model predictions achieved high accuracy both in terms of location and time [21]. From the San Francisco dataset, key attributes such as time intervals, crime probability, and crime locations were calculated, and crime hotspots were predicted using the KNN algorithm, with vulnerable areas visualized for security vulnerability analysis. A model built with the Naive Bayes (NB) algorithm achieved an accuracy rate of 97.5% [22]. Seventy-nine studies using machine learning methods to predict the risk of reoffending were reviewed, and 12 studies that achieved successful crime predictions through model repetition with different datasets were selected. The average accuracy (ACC) and area under the curve (AUC) scores of these studies were calculated to be 0.81 and 0.74, respectively. Based on the findings of this review, the following conclusions were drawn: due to the inaccessibility of the models and algorithms used in judicial decisions, principles of transparency, impartiality, and justice may be compromised, potentially leading to discrimination between individuals or groups of individuals. To address these issues, transparent algorithms should be developed, or explainable artificial intelligence should be used, and a human operator should be placed at the head of the human–computer system, in line with an integrated cognitive system [23]. In a study conducted on a dataset of theft crimes collected between 2017 and 2020 from the XT Paichusuo region of ZG city in southeastern China, the XGBoost algorithm was used for crime prediction modeling, and the SHAP method was used to measure the contribution of attributes to criminal behavior. As a result, an accuracy of 91% was achieved, and the attribute most strongly associated with criminal behavior was found to be the proportion of the non-local population aged 25–44 [24]. In a study using the open datasets of Chicago and Los Angeles, machine learning algorithms such as support vector machine (SVM), Naive Bayes (NB), K-nearest neighbors (KNN), decision tree (DT), multilayer perceptron (MLP), random forest (RF), and XGBoost, along with the LSTM deep learning algorithm for time series analysis, and the ARIMA method for predicting areas with high crime rates, a crime-type prediction accuracy of 94% was achieved, along with an 8.78 mean square root error and 6% mean absolute error with LSTM [25]. In a study using the NCRB crime dataset, a stacking-based crime prediction method (SBCPM) based on SVM algorithms was proposed. Using the SBCPM method, which is a community learning model, a classification accuracy of 99.5% was achieved on test data [26]. In a study using a dataset containing public property crimes from a section of a large coastal city in southeastern China between 2015 and 2018, it was found that the model built with LSTM outperformed models built with KNN, RF, SVM, NB, and convolutional neural network (CNN). When structured environmental data for relevant locations and urban road network density were used as common variables, it was observed that the prediction results improved in the LSTM model compared to the original model based solely on historical crime data [5]. Using the San Francisco crime dataset and models built with NB, RF, gradient boosting, and DT algorithms, crime prediction accuracies of 65.82%, 63.43%, and 98.5% were achieved, respectively [27]. In a study using various types of crime data from the Chicago City Data Portal, American FactFinder, Weather Underground, and Google Street View, image data were used to extract environmental context information. The relationship between crime events and the collected data was analyzed through a statistical approach, and the model built using the dataset and DNN achieved an accuracy of 84.25, precision of 74.35, sensitivity of 80.55, and an AUC value of 0.8333 [28]. In a study using the Chicago crime dataset, an integrated model was proposed, consisting of three modules: a spatial–temporal feature extraction module (ST-GCN), a temporal feature extraction module (LSTM), and a feature integration module (GBDT). The advantage of the model is that the integration of ST-GCN and LSTM enables more accurate definition of the weighted relationship between the effect of the number of historical crime events and the effect of the transition of nearby crime events. The model achieved results of 0.39 MAPE, 1.03 RMSE, and 0.84 R^2 [29]. In a study aiming to utilize semantic information learned from textual data and historical crime data, and to transfer this information to a model-trained crime prediction, a data fusion technique was used with the ConvBiLSTM model to extract independent vectors from tweets and crime methods and combine them into a single representation that captures information from all methods. Using crime event data obtained from the Chicago Police Department, specifically covering the period from September 1–30, 2019, and data from tweets containing crime-related terminology specific to Chicago, the proposed multimodal data fusion using the ConvBiLSTM model outperformed SVM, LR, NAHC, feature-level data fusion DNN, CrimeTelescope, ANN + BERT, and BERT-based crime prediction models with an accuracy rate of 97.5% [30]. In a study conducted in Boston to identify high and low crime-prone locations, the most and least violent crime types, and annual and monthly crime rates through exploratory data analysis, it was found that the classification accuracy of the RF algorithm developed with principle component analysis (PCA) increased by 9% compared to the simple decision tree, and the classification accuracy of the decision tree algorithm developed with PCA increased by 5% compared to the regular decision tree [31]. In a study using crime datasets from Chicago, New York, and Lahore, monthly and weekly crime prediction were made using statistical techniques (SMA, WMA, EMA) and deep learning techniques (CNN-LSTM, LSTM, BiLSTM). By applying transfer learning with BiLSTM, the following results were achieved: for the Chicago dataset, 65.68 MAE, 58.66 MAD, and 81.154 MSE; for the New York dataset, 505.93 MAE, 373.41 MAD, and 621.82 MSE; and for the Lahore dataset, 87.53 MAE, 85.04 MAD, and 107.75 MSE [32]. In a study that used social media sentiment analysis to predict criminal activities with the Chicago crime dataset and datasets containing crime-related terms in tweets about Chicago, the issue of imbalanced sentiment classes was addressed by combining a reinforcement learning (RL) algorithm with a coverage loss function. Hyperparameter optimization was done using the artificial bee colony (ABC) technique, achieving accuracy values of 96.411% and 94.088%, respectively [14]. A summary of the relevant studies is provided in Table 1.

Upon reviewing the literature, it was found that crime prediction studies with highly dynamic data samples often do not use sufficiently up-to-date and comprehensive datasets, do not provide detailed attention to the dataset preprocessing stage, and only predict specific crime types. Additionally, there are very few studies that integrate machine learning, deep learning methods, and statistical methods for the analysis and prediction of crime events. Furthermore, no study was found that simultaneously addressed crime types, exploratory data analysis, time series analysis, trend prediction, and regional time series analysis. This study aims to make a contribution to the literature by applying comprehensive preprocessing to three different datasets, visualizing the time and location trends of crime events in these datasets, developing prediction models for all 37, 34, and 33 different crime types in the datasets, conducting monthly, weekly, and daily time series analysis, predicting crime trends, and performing regional crime predictions. By analyzing crime events from multiple perspectives, this study is expected to contribute to the development of strategies for preventive policing.

3 Methodology and data

3.1 Methodology

In this study, a six-stage methodology was proposed using publicly available crime data from the three most metropolitan cities in the USA—San Francisco, Chicago, and Philadelphia—with the aim of contributing to the development of strategies for preventive policing:

(1) Preprocessing of datasets;
(2) Exploratory data analysis;
(3) Development of crime-type prediction models using preprocessed datasets and traditional machine learning algorithms;
(4) Time series analysis using deep learning methods, including LSTM and BLSTM;
(5) Time series analysis and 5-year crime trend forecasting using statistical methods such as HWES, Prophet, and SARIMA;
(6) Regional time series analysis using statistical methods and deep learning techniques;
(7) Fig. 1 presents the flowchart of the proposed methodology.

As shown in Fig. 1, the first stage of the proposed methodology involved applying preprocessing steps such as data cleaning and formatting, categorical data encoding, and feature selection to the datasets. In the second stage, patterns and relationships among features in the datasets were explored by visualizing hourly, daily, monthly, and yearly crime rates, the distribution of the top 10 most committed crime types across datasets, and crime density by region. In the third stage, crime-type prediction models were developed using preprocessed datasets and machine learning algorithms such as XGBoost, CatBoost, RF, DT, MLP, KNN, GNB, and LR. The fourth stage involved conducting monthly, weekly, and daily time series analyses using LSTM and BLSTM deep learning architectures. In the fifth stage, future trends for the next 5 years were forecasted using statistical methods such as HWES,

Table 1 Summary of relevant studies

Year	Authors	Model	Dataset	Best results
2017	Wang B, D. Zhang, D. Zhang, P.J. Brantingham, and A.L. Bertoz.zi	ST-ResNet	Real-time Los Angeles crime dataset for the last six months of 2015	Accuracy: %84.78
2017	Kang HW, Kang HB	DNN	Chicago, American FactFinder, Weather Underground, Google Street View	Accuracy: %84.25, Precision: %74.35, Recall: %80.55, AUC: 0.8333
2019	Feng M, Zheng J, Ren J et al	LSTM, Prophet, neural network	San Francisco, Chicago, and Philadelphia crime datasets	San Francisco RMSE-LSTM: 45.65, Correlation-LSTM: 0.423 Chicago RMSE-Neural Network: 75.06, Correlation-Prophet: 0.658. Philadelphia RMSE-Neural Network: 63.68, Correlation-Prophet: 0.729
2020	Zhang X, Liu L, Xiao L, Ji J	LSTM KNN, RF, SVM, NB, CNN	China public domain crime dataset	HitRn: %59.9 HitRa: %57.6
2020	Han X, Hu X, Wu H, Shen B, J. Wu J	ST-GCN + LSTM	Chicago crime dataset	MAPE: 0.39, RMSE: 1.03, R2: 0.84
2021	Stalidis P, Semertzidis T, Daras P	SFTT, TFTS, ParB, ST-ResNet	Philadelphia, Seattle, Minneapolis, DC Metro, San Francisco crime datasets	F1 Skoru: 0.99, AUCPR:1.00, AUROC:0.99, AUC, PAI@5: 4.40
2021	Palanivinayagam A, Gopal SS, Bhattacharya S et al	NB, KNN, RF, SVM	San Francisco, Los Angeles crime datasets	Accuracy: %97.5
2021	Safat W, Asghar S, Gilan S	SVM, NB, KNN, DT MLP, RF, XGBoost, LSTM, ARIMA	Chicago, Los Angeles crime datasets	RMSE: 12.66, 8.78 MAE: 11.70, 6 Accuracy: %94, %89 F1-skor: %100, %100
2021	Kshatri SS, Singh D, Narain B et al	J48, SMO, NB, RF, bagging classifier, SBCPM	NCRB crime dataset	Accuracy: % 99.5
2021	. Sharma HK, Choudhury T, Kandwal A	DT, PCA + DT, RF, PCA + RF	Boston crime dataset	Accuracy: %60
2022	Nesa M, Shaha TR, Yoon Y	KNN, SVM, RF, XGBoost, MLP, SHAP, LIME	Data set obtained from Holy Life, Promises, BARACA and Sneho-Nir rehabilitation centers in Dhaka city	Accuracy: %95.36
2022	De Blasio G, D'ignazio A, Letta M	SMOTE, DT	SDI archive of the Italian Ministry of Internal Affairs	Accuracy: ∼ %80
2022	Wang C, Han B, Patel B, Rudin C	COMPAS, Arnold PSA, interpretable ML, black-box ML	Florida, Kentucky crime datasets	The recidivism prediction success of interpretable ML algorithms is better than COMPASS, Arnold PSA and black-box ML algorithms
2022	Tasnim N, Imam IT, Hashem MMA	Temporal-based attention LSTM, Spatiotemporal-based Stacked bidirectional LSTM, and fusion model	San Francisco and Chicago crime datasets	MAE: 0:008, 0:02 SMAPE: %1:03, %0:6
2022	Travaini GV, Pacchioni F, Bellumore S et al	LR-LogitBoost-RF-Glmnet-MLP-ensemble model, NBC, KNN, PNN, SVM, LDA-L1–LR-Penalized LDA	Thailand, MnSTARR + , LS/CMI, NCRA + , RisCanvi, FDJJ, RITA + , HCR-20 + ,YLS/CMI, SAVRY + , StatRec, DOI	Average for 12 different article studies: ACC: 0.81 AUC: 0.74

Table 1 (continued)

Year	Authors	Model	Dataset	Best results
2022	Khan M, Ali A, Alharbi Y	NB, RF, gradient boosted decision tree	San Francisco crime dataset	Accuracy: %65.82, %63.43, %98.5
2023	Tam S, Tanriöver ÖÖ	Including SVM, LR, NAHC, DNN with feature-level data fusion, crime telescope, ANN + BERT, ConvBiLSTM	Chicago crime dataset, twitter data	Accuracy: %97.5
2024	Butt UM, Letchmunan S, Hassan FH et al	SMA, WMA, EMA, LSTM, BiLSTMs, CNN-LSTM	Chicago, New York, Lahore crime dataset	Chicago (MAE: 65.68, MAD: 58.66, MSE: 81.154): New York (MAE: 505.93, MAD: 373.41, MSE: 621.82) Lahore (MAE: 87.53, MAD: 85.04, MSE: 107.75)
2024	Yang L, Guofan J, Yixin Z et al	LSTM, GSCA,TwitterGuard, ConvBiLSTM, SocioCrimAnalytix, CNN-LSTM, DAC-BiNet, LSA, LSTMTwitter, MDL	Chicago crime and Chicago twitter (X) dataset	Accuracy: %96.411 and %94.088

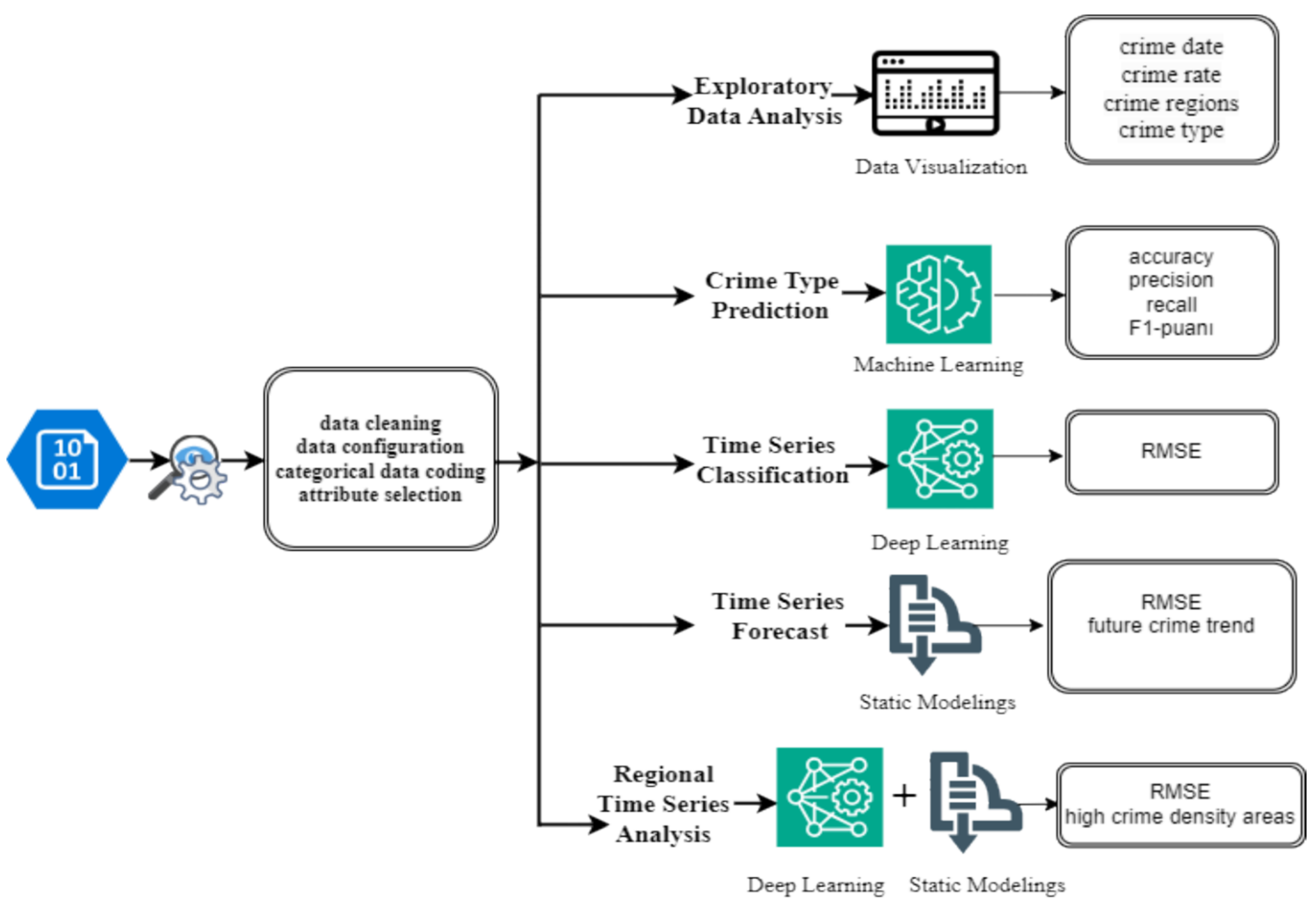


Fig. 1 Flowchart of the proposed methodology

Prophet, and SARIMA. Finally, in the sixth stage, regional time series analyses were performed for the ten police districts in the San Francisco dataset, the 22 police districts in the Chicago dataset, and the 22 police districts in the Philadelphia dataset. These analyses used both statistical methods (HWES, Prophet, and SARIMA) and deep learning methods (LSTM and BLSTM) to accurately predict high-crime hotspots. The RMSE values obtained from the analyses were compared.

3.2 Data preprocessing of datasets

In traditional machine learning and deep learning models built with big data, the data itself directly affects performance. As observed in the crime datasets used in our study, most large datasets consist of unstructured or semistructured data. To achieve optimal performance from the models, it is crucial to preprocess the datasets meticulously. In this study, publicly available crime datasets from San Francisco, Chicago, and Philadelphia—some of the largest metropolitan areas in the USA with high crime activity—were utilized. The San Francisco crime dataset contains 2,129,525 crime records and 35 different features spanning from 01/01/2003 to 12/31/2017 [12]. The Chicago crime dataset includes 7,754,190 crime records and 30 different features from 01/01/2001 to 12/31/2022 [33]. The Philadelphia dataset comprises 3,009,007 crime records and 18 different features from 01/01/2006 to 12/31/2022 [34]. Since the San Francisco crime dataset features changed after 2018, crime records from 2018 to 2022 could not be utilized. To ensure a more accurate and objective comparison among the datasets, only common features with similar functionality were used across all three datasets. These features and their descriptions are presented in Table 2.

The feature names listed in Table 2, which have similar counterparts across the datasets, were standardized to be the same in all three datasets as follows: "ID," "INumber," "ICode," "Date," "Time," "District," "*X*," "*Y*," and "Category" factors such as missing data, class imbalance issues, categorical variables, data values in varying ranges, and features that are ineffective in representing classes can directly reduce the performance of machine learning and deep learning models. These issues were addressed during the data preprocessing phase of this study. The preprocessing steps applied to the datasets included data cleaning and formatting, converting categorical variables into numerical representations, scaling data, and feature selection.

Data Cleaning and Organization After removing missing and duplicate values from the San Francisco, Chicago, and Philadelphia datasets, the record counts were updated to 2,129,524, 7,668,648, and 2,964,958, respectively. The "*X*" and "*Y*" coordinate attributes, once scaled with standard scaler, were grouped into 50 different regions based on Euclidean distances using the *K*-Means clustering method, and a new attribute named "Cluster" was added to the dataset. To extract more spatial information from the "*X*" and "*Y*" coordinates, these values were transformed from Cartesian space to polar space (rotated 30, 45, 60 degrees), thereby increasing the number of location-related attributes. This approach aimed to enhance the contribution of the crime location to

Table 2 Common attributes selected from San Francisco, Chicago, and Philadelphia datasets

Attribute Number	San Francisco	Chicago	Philadelphia	Attribute descriptions
1	PdId	ID	dc_key	Unique identifier for each crime record
2	IncidntNum	Case Number	Objectid	Unique record section number for the incident
3	Incident Code	IUCR VE FBI Code	ucr_general	Crime reporting code
4	Date	Date	Dispatch_date	Date of the incident occurrence
5	Time	Date	Dispatch_time	Time of the incident occurrence
6	PdDistrict	District	dc_dist	Police district where the incident occurred
7	X	X	X	Longitude of the crime scene
8	Y	Y	Y	Latitude of the crime scene
9	Category	Primary Type	Text_general_code	Type of crime

crime predictions. From the timestamp (“Date,” “Time”) attributes, nine different attributes were derived: year, month, day of the week, day of the month, day of the year, week of the month, week of the year, hour, and minute (“Year,” “Month,” “dayOfWeek,” “dayOfMonth,” “dayOfYear,” “weekOfMonth,” “weekOfYear,” “Hour,” and “Minute”). Attributes such as Hour_Zone, BusinessHour from the “Hour” attribute; Holiday from the “Date” attribute; Season from the “Month” attribute; and Weekend from the “dayOfWeek” attribute were derived to account for weekends, holidays, seasons, business hours, and different time intervals during the day, as they were considered influential on the number and location of crimes committed. In the final stage, as shown in Table 3, the number of attributes in all three datasets reached 29.

Categorical Data Encoding Categorical data, lacking mathematical or logical relationships, cannot be processed by many machine learning algorithms. Therefore, categorical variables in the datasets used in this study were numerically encoded using the commonly used label encoding method, which does not introduce a large number of variables and does not significantly increase computational costs.

Scaling In the datasets used in the study, Min–Max and Standard Scaler, which are commonly employed methods in the literature, were utilized at different stages of the proposed methodology. These methods were applied to ensure an objective comparison of data values across different ranges and to establish more accurate connections between them.

Feature Selection The Mutual Information (MI) method, one of the filtering-based feature selection methods, was used in the datasets employed in the study to eliminate attributes that were ineffective in representing classes and could reduce the classification performance of algorithms. MI was chosen due to its low computational cost, ability to measure arbitrary relationships between variables, and independence from transformations affecting different variables [35]. MI is calculated using Eq. 1:

$$I\,(X\,;\,Y)\;=\;H(X)-H(X\,|\,Y) \tag{1}$$

I = MI, X = Attributes, Y = Class Variable, H = Entropy.

The MI score ranges from 0 to 1, and the higher the score, the stronger the connection between the respective attribute and the class variable. This suggests that the attribute should be selected. The selection of attributes using the MI method is based on a predetermined threshold value. The threshold value that yields the best performance is determined through trial and error. The selected attributes from the San Francisco, Chicago, and Philadelphia datasets using the MI method, along with their importance coefficients, are presented in Table 4.

According to Table 4, the common attributes among the top 10 features with the highest importance coefficients in the datasets for all three cities are ICode, Radius, Rot30_X, Rot60_X, Rot45_Y, Rot60_Y, and Rot30_Y. The feature importance rankings in the three datasets are similar, and the datasets containing the selected features will be utilized in the analyses and prediction models to be developed.

3.3 Exploratory data analysis

Exploratory data analysis is a process that involves analyzing a dataset before forming a hypothesis, extracting data patterns, detecting errors, anomalies, or outliers, and understanding the relationships between variables. It is fundamentally a process driven by data visualization, which guides the work to be done. In this study,

Table 3 Attributes of preprocessed datasets

Attributes
“ICode”, “Radius”, “Rot30_X”, “Y”, “Rot45_X”, “Rot60_Y”, “Rot30_Y”, “Rot60_X”, “X”, “Rot45_Y”, “Angle”, “ID”, “INumber”, “Minute”, “Cluster”, “District”, “BusinessHour”, “Hour”, “Hour_Zone”, “Year”, “Weekend”, “Season”, “dayOfWeek”, “weekOfMonth”, “Month”, “dayOfYear”, “dayOfMonth”, “weekOfYear”, “Holiday”

Table 4 Selected attributes and their importance coefficients from the datasets using the MI method

No:	San Francisco		Chicago		Philadelphia	
	Attributes	Significance coefficients	Attributes	Significance coefficients	Attributes	Significance coefficients
1	ICode	2.625827	ICode	2.468959	ICode	2.463508
2	Radius	0.331433	FCode	2.445897	ID	1.087553
3	Rot30_X	0.331309	Angle	0.289704	Angle	0.458134
4	Y	0.331125	Rot30_Y	0.287521	Rot45_Y	0.455321
5	Rot45_X	0.330754	Rot45_Y	0.286572	Rot60_Y	0.454434
6	Rot60_Y	0.330730	Rot60_X	0.286064	Rot30_Y	0.453897
7	Rot30_Y	0.330568	Rot45_X	0.285376	Y	0.453518
8	Rot60_X	0.330376	Rot30_X	0.284664	Radius	0.451756
9	X	0.330338	Rot60_Y	0.284502	Rot30_X	0.451400
10	Rot45_Y	0.329896	Radius	0.283441	Rot60_X	0.451055
11	Angle	0.329594	Y	0.204331	X	0.451036
12	ID	0.255807	ID	0.174396	Rot45_X	0.450880
13	INumber	0.255241	X	0.170194	INumber	0.258116
14	Minute	0.119872	BusinessHour	0.126087	Hour	0.109021
15	Cluster	0.102388	Minute	0.098302	Hour_Zone	0.104245
16	District	0.063431	Cluster	0.077721	BusinessHour	0.098905
17	BusinessHour	0.050450	Hour_Zone	0.074456	Cluster	0.070812
18	Hour	0.034804	District	0.067808	District	0.059447

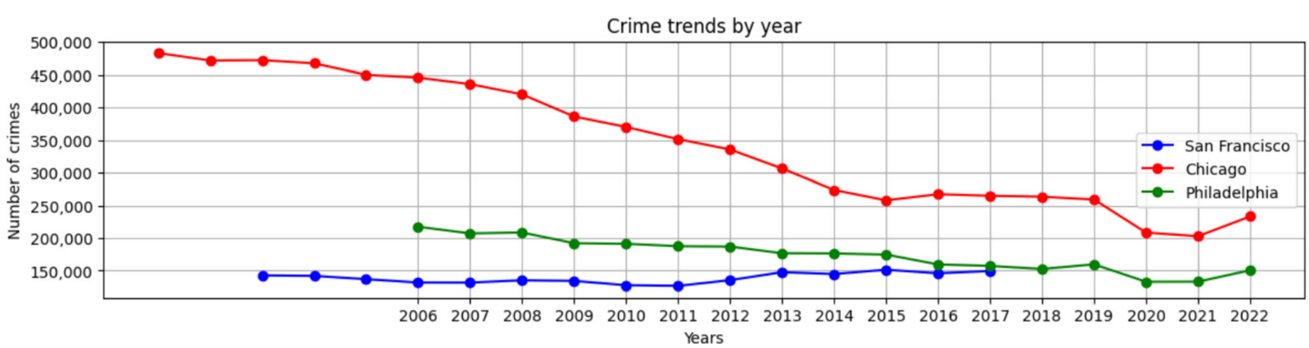


Fig. 2 Crime trends over the years in the three cities

visualizations related to time, location, and crime types have been performed on the three datasets used. Figure 2 shows the variation in crime trends over the years in the cities of San Francisco, Chicago, and Philadelphia.

As seen in Fig. 2, crime incidents in San Francisco were at their lowest in 2011. They reached a maximum number in 2015. Between 2013 and 2017, there were no significant increases or decreases. In Chicago, crime incidents were at their highest in 2001. Between 2001 and 2015, there was a steady, small increase, while between 2016 and 2019, there were no significant changes in the number of crime incidents. Between 2019 and 2021, there was a decrease of approximately 50,000 due to the pandemic. In 2022, with the end of the pandemic, there was an increase of approximately 30,000 crime incidents. In Philadelphia, there was an overall decrease in crime incidents between 2006 and 2021, with the largest decrease occurring between 2019 and 2020. Post-pandemic, between 2021 and 2022, crime incidents increased again by approximately 17,000. Figure 3 shows the variation in crime trends by month in the cities of San Francisco, Chicago, and Philadelphia.

As shown in Fig. 3, crime incidents in San Francisco are at their lowest in February and December, and at their highest in August and October. In Chicago, crime incidents are at their lowest in February and December, and at

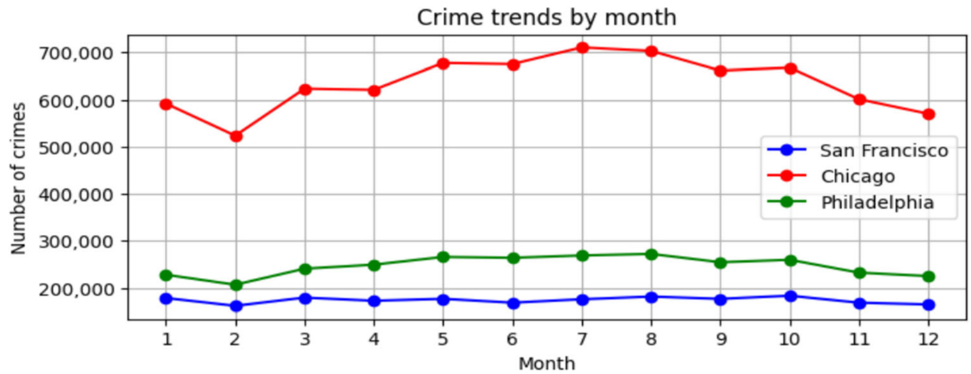


Fig. 3 Crime trends by months in the three cities

their highest in July. In Philadelphia, crime incidents are at their lowest in February and December, and at their highest in August. In all three cities, the least number of crime incidents occurred in February and December. The highest number of crime incidents in all three cities occurred during the summer months, indicating that seasons have an impact on the number of crime incidents, with summer being the season with the most crimes. Figure 4 shows the variation in crime trends by day in the cities of San Francisco, Chicago, and Philadelphia.

As shown in Fig. 4, in San Francisco and Chicago, crime incidents occur the least on Sundays and the most on Fridays. In Philadelphia, crime incidents occur the least on Sundays and the most on Tuesdays and Wednesdays. A general evaluation of crime trends by day of the week across the three cities shows that in San Francisco and Chicago, Sunday and Friday are the days with the least and most number of crime incidents, respectively. In Philadelphia, although Sunday has the least number of crime incidents, similar to the other two cities, Tuesday stands out as the day with the highest number of crime incidents, which is different from San Francisco and Chicago. Based on this evaluation, Sunday is the safest day in terms of crime occurrences. Figure 5 shows the variation in crime trends by hour in the cities of San Francisco, Chicago, and Philadelphia.

As shown in Fig. 5, the hour with the least number of crime incidents in all three cities is around 5:00 AM. The hours with the most number of crime incidents are 6:00 PM in San Francisco, 12:00 PM in Chicago, and 4:00 PM in Philadelphia. A general evaluation of crime trends by hour shows a rapid decline in crime numbers between 12:00 and 5:00 AM, a general increase between 5:00 and 4:00 PM, and a decline again between 8:00 and 11:00 PM. Figure 6 illustrates the variation in crime trends by police district in San Francisco, Chicago, and Philadelphia.

As shown in Fig. 6, the police district with the highest number of crime incidents is District 1 (Southern) in San Francisco, District 8 (Chicago Lawn) in Chicago, and District 15 (Harbison & Levick) in Philadelphia. The police district with the fewest crime incidents in San Francisco is District 10 (Richmond), while in Chicago, District 21, with only four crimes between 2001 and 2022, and District 31, with 229 crimes, have the least incidents. Excluding these, the district with the fewest crime incidents in Chicago is District 20 (Lincoln). In Philadelphia, the data for police districts 4 and 23 are available until 2016, and the data for District 92 are available until 2018. Therefore, the district with the least number of crimes among others is District 77 (77th International Airport).

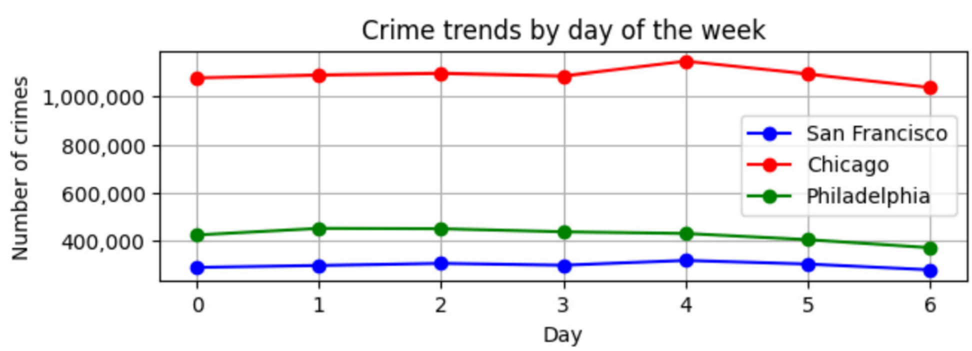


Fig. 4 Crime trends by days of the week in the three cities

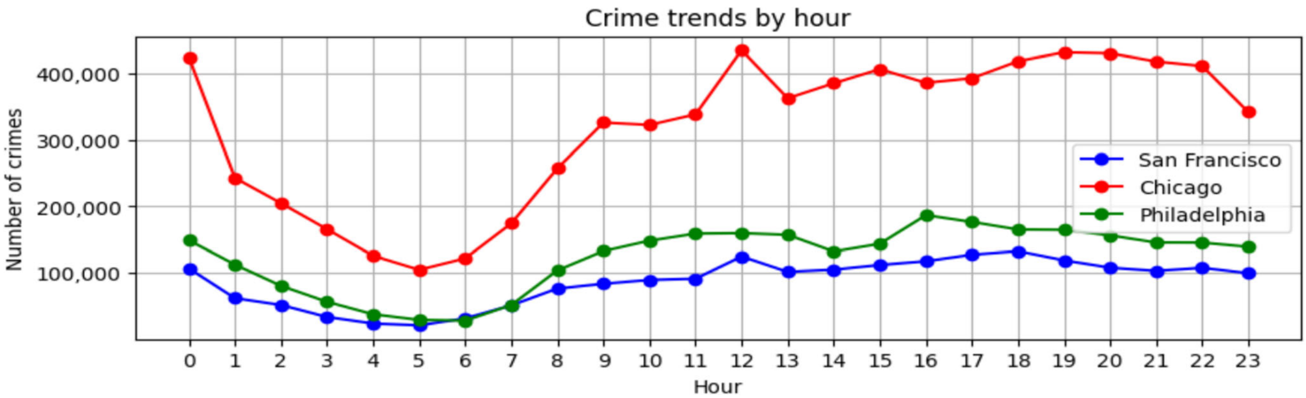


Fig. 5 Crime tendencies by hours in three cities

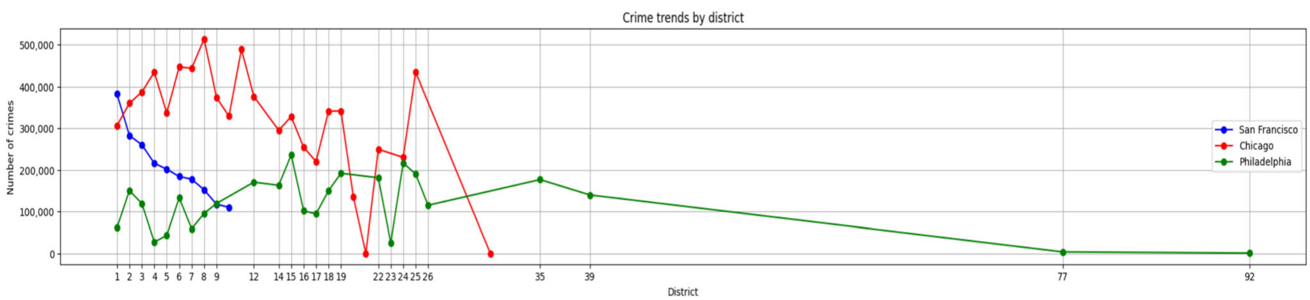


Fig. 6 Crime tendencies by police districts in three cities

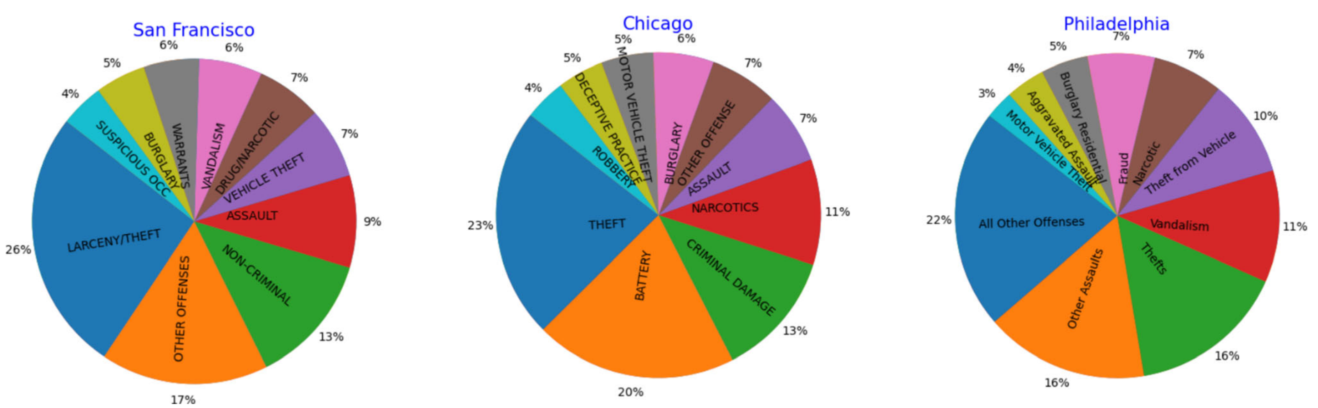


Fig. 7 Distribution of the top 10 most committed crime types in the cities of San Francisco, Chicago and Philadelphia

Figure 7 shows the distribution of the ten most common types of crimes in the datasets for San Francisco, Chicago, and Philadelphia.

As shown in Fig. 7, the most common crime type in San Francisco is "Larceny/Theft," in Chicago it is "Theft," and in Philadelphia it is "All Other Offenses." The second most common crime types are "Other Assaults" and "Thefts." Table 5 provides a summary of the results from the exploratory data analysis.

A general assessment of the data visualization indicates that there is a decrease in the number of crime incidents over the years. The most reliable months for crime occurrences are February and December, the most reliable day is Sunday, the most reliable hours are 5–6 in the morning, and the most frequent crime type is theft. The police districts with the highest number of crimes are District 1, District 8 and District 15. The most dangerous months in terms of crime occurrences are the summer months, and the most dangerous hours are

Table 5 Results from the exploratory data analysis of San Francisco, Chicago, and Philadelphia crime datasets

Comparison	San Francisco(2003–2018)	Chicago (2001–2023)	Philadelphia(2006–2023)
Crime spike hour	18.00	12.00	16.00
Lowest crime hour	05.00	05.00	05.00
Crime spike day	Friday	Friday	Tuesday
Lowest crime day	Sunday	Sunday	Sunday
Crime spike month	October	July	August
Lowest crime month	February- December	February- December	February- December
Crime spike year	2015	2001	2006
Lowest crime year	2011	2021	2020
Crime spike area	Southern (1.)	Chicago Lawn (8.)	Harbison & Levick (15.)
Lowest crime area	Richmond (10.)	Lincolen (20)	77th International Airport (77.)
Top 3 Crimes	larceny/theft, other offenses, non-criminal	theft, battery, criminal damage	larceny/theft, other offenses, non-criminal

between 12:00 and 18:00. As a result, more police resources should be allocated during the summer months, between 12:00 and 18:00, and especially in Districts 1, 8, and 15, with a focus on theft crimes.

3.4 Crime-type prediction

The San Francisco, Chicago, and Philadelphia crime datasets, which underwent preprocessing steps including data cleaning and formatting, categorical data encoding, scaling, and feature selection, were divided into 70% training data and 30% testing data for crime-type prediction. Eight different machine learning algorithms—XGBoost, CatBoost, RF, DT, MLP, KNN, GNB, and LR—were used to create crime-type prediction models using the training data. The algorithms were run with default parameter values, and hyperparameter tuning was not performed. The models were tested with the test dataset using accuracy, precision, recall, and F1-score metrics. The results of the tests are presented in Table 6.

When evaluating the results in Table 6, it can be observed that the 37 different crime types in the San Francisco dataset were detected with 100% accuracy by the models built with XGBoost and DT. The model built with LR had the lowest detection accuracy, with a rate of 43.2%. In the Chicago dataset, the 34 different crime types were detected with 100% accuracy by the models built with XGBoost, DT, and CatBoost. The model built with KNN had the lowest detection accuracy, with a rate of 76.9%. In the Philadelphia dataset, the 33 different crime types were detected with 90.5% accuracy by the model built with XGBoost. The model built with LR had the lowest

Table 6 Test results of models built with San Francisco (S), Chicago (C), and Philadelphia (P) crime datasets

Algorithms	Accuracy			Precision			Recall			F1-score		
	S	C	P	S	C	P	S	C	P	S	C	P
XGBoost	1.000	1.000	0.905	1.000	1.000	0.901	1.000	1.000	0.905	1.000	1.000	0.899
CatBoost	0.991	1.000	0.902	0.991	1.000	0.898	0.991	1.000	0.902	0.991	1.000	0.895
RF	0.932	0.999	0.902	0.929	0.999	0.899	0.932	0.999	0.902	0.929	0.999	0.899
DT	1.000	1.000	0.890	1.000	1.000	0.890	1.000	1.000	0.890	1.000	1.000	0.890
MLP	0.892	0.998	0.886	0.888	0.998	0.882	0.892	0.998	0.886	0.885	0.998	0.869
KNN	0.546	0.769	0.751	0.523	0.748	0.738	0.546	0.769	0.751	0.525	0.748	0.738
GNB	0.687	0.978	0.870	0.710	0.979	0.864	0.687	0.978	0.870	0.681	0.978	0.866
LR	0.432	0.797	0.678	0.335	0.726	0.612	0.432	0.797	0.678	0.328	0.743	0.625

detection accuracy in Philadelphia as well, with a rate of 67.8%. Overall, the machine learning algorithms performed better in the San Francisco and Chicago datasets. However, in the Philadelphia dataset, which had more imbalanced class data, the classification performance was lower compared to the other datasets. The graph in Fig. 8 compares the crime-type detection accuracy rates of the models built with the algorithms across the three datasets used.

According to Fig. 8, the most successful models for crime-type detection across the three datasets were the models built with XGBoost, while the least successful models were those built with LR. Overall, for crime-type prediction across the three datasets, the highest accuracy rates were achieved as follows: 100%, 100%, and 90.5%, with F1-scores of 100%, 100%, and 89.9%, respectively. Considering that 37 different crime types were predicted for San Francisco, 34 for Chicago, and 33 for Philadelphia, these results are quite successful when compared to the literature.

3.5 Time series analysis with LSTM and BLSTM models

The LSTM architecture was proposed as a solution to the vanishing gradient problem in RNNs (recurrent neural networks). A typical LSTM unit consists of input, forget, and output gates, where the gates act as mechanisms for allowing or blocking the flow of information [36]. In the first section, the forget gate, the cell decides whether the information from the previous time step should be remembered or forgotten if it is deemed irrelevant. The second section, the input gate, tries to learn new information from the input to the cell. In the third section, the output gate, the cell transfers the updated information from the current time step to the next time step. In this way, valuable information is extracted from sequential data. LSTMs are widely used in solving many complex problems such as handwriting recognition and generation, language modeling and translation, acoustic modeling of speech, speech synthesis, and analysis of audio and video data due to their effectiveness in capturing long-term temporal dependencies. The fundamental idea behind bidirectional LSTMs (BLSTM) is that the key difference between LSTM and BLSTM models is that LSTM networks allow inputs in only one direction, whereas BLSTM networks allow the flow of information in both directions by adding a new LSTM layer that reverses the sequence. The outputs of both layers are combined through averaging, summing, multiplication, or concatenation. The bidirectional flow state means that the BLSTM will have complete sequential information about both the preceding and succeeding points, enabling a better learning process [37]. In this study, LSTM and BLSTM deep learning models were used for time series analysis on crime datasets. However, to achieve better performance from these architectures, the data were made stationary and scaled before application. Due to computational costs, search methods for parameter tuning could not be used in the LSTM and BLSTM models, and the parameters frequently chosen in the literature, listed in Table 7, were used.

RMSE (root-mean-square error) is the square root of the average of the squared differences between predicted and actual observations. It is used as a metric to evaluate model performance because it has the advantage of being interpretable in the same unit as the predicted target variable. Additionally, by penalizing large errors and highlighting poor performance, RMSE provides a useful measure of model accuracy.

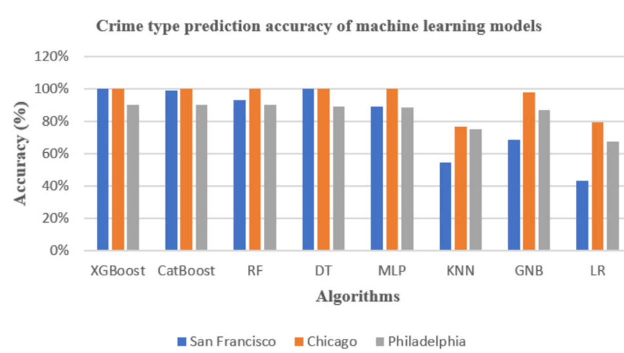


Fig. 8 Crime-type prediction accuracies of algorithms in San Francisco, Chicago, and Philadelphia datasets

Table 7 LSTM and BLSTM models' parameters

Parameter	Value
epoch	256
batch_size	80, 75, 128, 40, 55,
dropout	0.4, 0.2
optimizer	adam
activation	relu
loss	mse

The RMSE values obtained from monthly, weekly, and daily time series analysis conducted using LSTM and BLSTM deep learning models on the San Francisco, Chicago, and Philadelphia crime datasets are presented in Table 8.

When evaluating the results in Table 8, the RMSE values obtained from monthly and weekly time series analyses are close to each other and are considered successful compared to results found in the literature. In the daily analysis, however, the RMSE values obtained from the analysis with daily data are worse than those obtained from monthly and weekly time series analyses. This discrepancy can be attributed to the increased volume of data and the narrower interval with more variable numbers in daily analysis. For monthly and weekly data, time series analysis with BLSTM generally performs slightly better than that with LSTM.

3.6 Time series analysis and trend forecasting with statistical methods

In this section, time series analysis was performed for all three datasets using the HWES, Prophet, and SARIMA statistical methods, and trend forecasts for the next 5 years were made using the models with the best RMSE values obtained.

3.6.1 SARIMA model

Autoregressive integrated moving average (ARIMA) is a method used for univariate time series analysis and forecasting. In the ARIMA model, also known as the Box–Jenkins models, the values of the time series are modeled based on past values and predictions are made [38]. ARIMA consists of the following components:

AR (Autoregression) (p) This represents the lags of the variables in the forecasting equation (the number of lags included in the model).

I (d) (Integrated) This indicates the degree of differencing required to make the series stationary (the degree of differencing).

MA (Moving Average) (q) This shows the lags of the forecasting errors (the number of periods in the moving average window).

Model notation ARIMA (p, d, q).

The parameter d refers to the number of differencing steps taken to make the time series stationary, while the p and q parameters can be determined through various methods, with autocorrelation function (ACF) and partial autocorrelation function (PACF) plots being commonly used. These plots can also provide insights into whether the series contains trend and seasonality. In this study, ACF and PACF plots showed that the data used was

Table 8 RMSE values obtained from monthly, weekly, and daily time series predictions using LSTM and BLSTM deep learning models in three cities

Datasets	Monthly time series (RMSE)		Weekly time series (RMSE)		Daily time series (RMSE)	
	LSTM	BLSTM	LSTM	BLSTM	LSTM	BLSTM
San Francisco	11.1938	9.7033	10.9852	10.2895	25.8704	26.2750
Chicago	30.7510	28.9522	28.5377	27.9595	50.2144	62.3348
Philadelphia	18.4233	17.2525	17.7128	17.4941	32.1350	36.1578

influenced by both trend and seasonality. For time series with seasonality, a more suitable model is SARIMA (Seasonal AutoRegressive Integrated Moving Average).

The SARIMA model, in addition to the non-seasonal components (p, d, q) in the ARIMA model, includes seasonal components $(P, D, Q)_m$. The seasonal components represent the autoregressive order, differencing order, moving average order, and the number of periods, respectively.

Model notation SARIMA (p, d, q) $(P, D, Q)_m$.

In the study, the auto_arima() function was used to search for the optimal parameter values for the pmdarima model, and then SARIMAX ("X" indicates that the method supports external variables, which is optional to add) was used to obtain the best model with the lowest AIC. The hyperparameters that resulted in the lowest AIC values were determined and used as follows: For San Francisco dataset: (3, 1, 0)(1, 0, 1) 12; for Chicago dataset: (0, 1, 1)(1, 0, 1) 12; and for Philadelphia dataset: (4, 0, 0)(1, 0, 1) 12.

The original data were split into training and test data in a 70–30% ratio, and the trend forecast for the next 5 years was made using the SARIMA model. The RMSE values obtained from the time series analysis with SARIMA for the San Francisco, Chicago, and Philadelphia datasets are 11.2357, 38.73, and 30.78, respectively. The trend forecasting graphs for the next 5 years are shown in Fig. 9.

As shown in Fig. 9, the SARIMA model predicts a moderate increase in crime numbers in San Francisco, a clear increase in Chicago, and a clear decrease in Philadelphia by 2029.

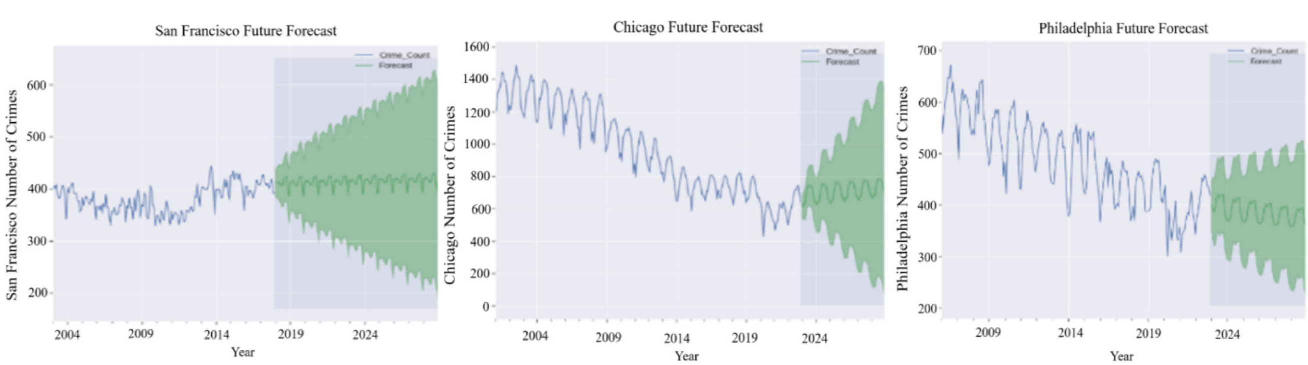
3.6.2 Prophet model

Prophet is an open-source library developed by Facebook for univariate time series forecasting. It applies a procedure to forecast time series data based on an additive model that accounts for non-linear trends, annual, weekly, daily seasonality, and holiday effects. The general formula that defines the time series in the Prophet model is as follows [39]:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \tag{2}$$

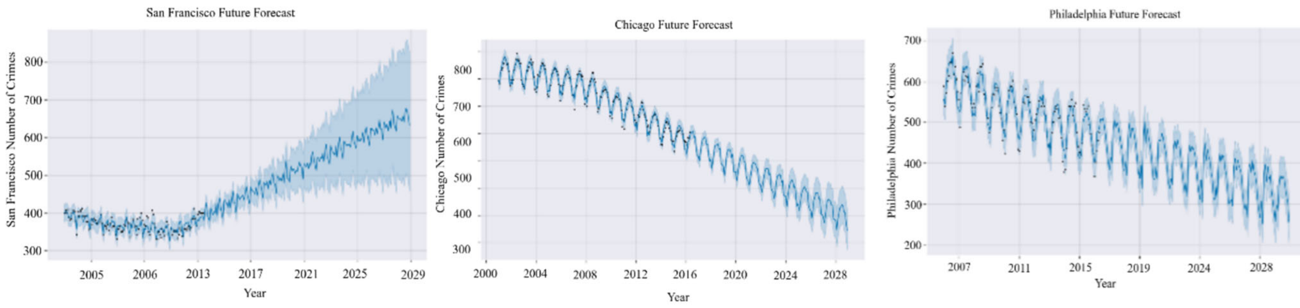
The formula is the modeling function for non-periodic variables in time series values. s(t) represents yearly, monthly, and weekly periodic changes. $h(t)$ represents the effects of holidays, usually irregularly occurring for one or more days. The error term $\varepsilon(t)$, often modeled as a normal distribution, accounts for any unique changes not captured by the model.

In the study, the original data was split into training and test sets, comprising 70% and 30% of the data, respectively. The Prophet model was then used to forecast trends for the next 5 years. The RMSE values for the Prophet predictions in the San Francisco, Chicago, and Philadelphia datasets are 22.59, 112.57, and 33.13, respectively. The trend forecasting plots for the next 5 years are presented in Fig. 10.



(a) (b) (c)

Fig. 9 Trend forecasting for the next 5 years in San Francisco (**a**), Chicago (**b**) and Philadelphia (**c**) with the ARIMA model



(a) (b) (c)

Fig. 10 Trend forecasting for the next 5 years in San Francisco (**a**), Chicago (**b**), and Philadelphia (**c**) with the Prophet model

As shown in Fig. 10, the Prophet model predicts a sharp increase in crime rates in San Francisco in the coming years, while a sharp decrease is forecasted for Chicago and Philadelphia.

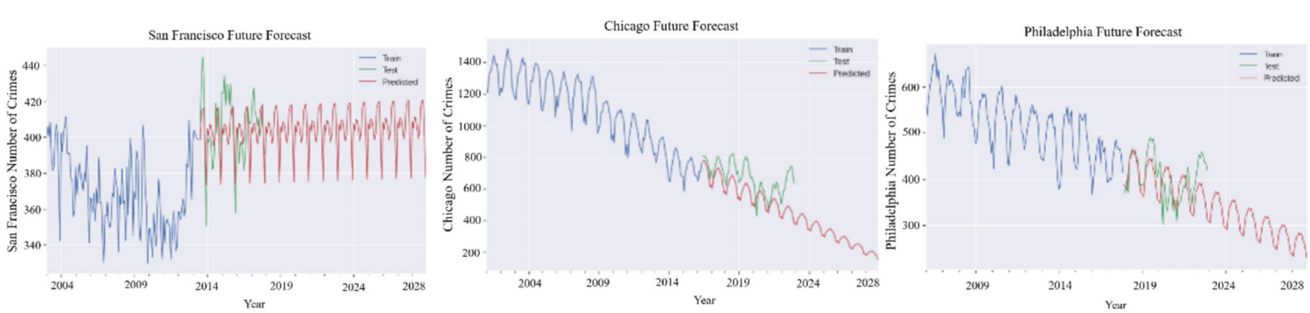
3.6.3 Holt–winters exponential smoothing (HWES) model

The Holt–Winters model, developed by Charles Holt and Peter Winters, is useful for time series forecasting, where users smooth time series data and then use it for predictions based on their areas of interest. Exponential smoothing is a method of smoothing time series data by assigning exponentially decreasing weights and values against historical data in order to reduce the weight of past observations [40]. The Holt–Winters exponential smoothing (ES) forecasting method uses exponential smoothing to encode many past values, and then uses these to predict typical values for the present and future [41]. The Holt–Winters model is categorized as either an additive model or a multiplicative model, depending on the seasonal pattern. The additive approach is used when changes within a season remain constant across periods, while the multiplicative method is used when changes within a season fluctuate according to the level of the period [42].

In the study, the original data were split into training and test datasets in a 70–30% ratio, and trend forecasting for the next 5 years was made using the HWES model. The RMSE values for the predictions made with the HWES model in the San Francisco, Chicago, and Philadelphia datasets are 14.91, 119.97, and 46.42, respectively. The trend forecasting graphs for the next 5 years are provided in Fig. 11.

As shown in Fig. 11, with the HWES model, a moderate increase is forecasted for San Francisco in the coming years, while a sharp decline is expected for Chicago and Philadelphia. The RMSE values for the predictions made with the Prophet, SARIMA, and HWES models for the three datasets are provided in Table 9.

As shown in Table 9, the success ranking of the models for the San Francisco, Chicago, and Philadelphia datasets is SARIMA > Prophet > HWES. Considering that the RMSE values in Table 9 will serve as a criterion for the prediction accuracy for the next 5 years, it is observed that for San Francisco, the predictions of SARIMA



(a) (b) (c)

Fig. 11 Trend forecasting for the next 5 years with HWES model in San Francisco (**a**), Chicago (**b**), and Philadelphia (**c**)

Table 9 RMSE values obtained from time series analysis with SARIMA, Prophet, and HWES models in San Francisco, Chicago, and Philadelphia datasets

Datasets	SARIMA (RMSE)	Prophet (RMSE)	HWES (RMSE)
San Francisco	11.2357	22.59	14.91
Chicago	38.73	112.57	119.97
Philadelphia	30.78	33.13	46.42

and HWES methods show similar trends, forecasting a moderate increase in crime rates in the coming years, while the Prophet model predicts a sharp rise. For Chicago, the predictions of the Prophet and HWES methods show similar trends, forecasting a sharp decline in crime rates in the coming years, while the SARIMA model forecasts a moderate increase. For Philadelphia, the predictions of the Prophet and HWES methods show similar trends, forecasting a sharp decline in crime incidents in the coming years, while the SARIMA model forecasts a moderate decrease.

The three datasets used in the study, as seen in the exploratory data analysis section of the study, contain sudden fluctuations in the data. The reason for the opposite results of the trend prediction by the SARIMA model compared to the Prophet and HWES methods for the Chicago and Philadelphia datasets can be attributed to the fact that SARIMA is sensitive to sudden fluctuations in the data. For the San Francisco dataset, the reason why the Prophet model gives different results compared to the SARIMA and HWES methods could be due to the model's performance varying based on the distinct characteristics of the dataset.

3.7 Regional time series analysis

Time series analysis was conducted using statistical methods (HWES, Prophet, and SARIMA) and deep learning models (LSTM and BLSTM) for the ten police districts in the San Francisco dataset, the 22 police districts in the Chicago dataset, and the 22 police districts in the Philadelphia dataset. For the SARIMA model, parameter tuning was performed using the grid search method for each police district. In the LSTM and BLSTM models, the rolling mean method was used to stabilize the time series, based on trial-and-error results for each district. Table 10 and Fig. 11 present the RMSE values obtained from time series analysis using five different methods for the police districts in San Francisco.

Table 10 RMSE values obtained from time series analysis with five different methods for San Francisco police districts

SF Police Districts	SH Police Districts No	Number of Crimes	HWES (RMSE)	PROPHET (RMSE)	SARIMA (RMSE)	LSTM (RMSE)	BLSTM (RMSE)
Southern	1	382,231	143.7	144.26	129.2	107.8959	93.2600
Mission	2	282,654	181.69	111.73	101.07	106.4044	102.9313
Northern	3	260,310	367.52	97.1	96.98	73.0621	70.6597
Central	4	215,901	380.42	146.87	94.17	74.2252	64.6733
Bayview	5	201,598	87.77	79.07	60.94	62.8655	56.0906
Tenderloın	6	184,114	185.82	135.22	60.84	67.3161	58.6584
Ingleside	7	177,561	134.3	88.01	68.61	54.4606	48.4679
Taraval	8	152,230	80.0	66.5	68.96	51.9559	50.4505
Park	9	117,347	292.87	140.5	69.02	44.2164	37.1170
Richmond	10	110,378	55.77	50.31	56.93	38.0304	31.7139

 https://doi.org/10.1007/s00521-025-11094-9

As seen in Table 10, in the time series analysis conducted with five different methods for the ten police districts in San Francisco, BLSTM was the most successful method for all ten districts, while HWES was the least successful method for nine districts. Additionally, Fig. 12 compares the RMSE values for the regional time series analysis of the San Francisco crime dataset using the five methods.

As depicted in Fig. 12, the ranking of methods from most successful to least successful for the six districts is as follows: BLSTM, LSTM, ARIMA, Prophet, and HWES. In three districts, SARIMA outperformed LSTM. Table 11 presents the RMSE values obtained from the time series analysis with five different methods for the police districts in Chicago.

As seen in Table 11, in the time series analysis conducted with five different methods for the 22 police districts in Chicago, BLSTM was the most successful method for 16 districts, while HWES was the least successful for 12 districts. Figure 13 compares the RMSE values of the five methods in the regional time series analysis for the Chicago crime dataset.

As seen in Fig. 13, the ranking of methods from most successful to least successful for six regions is BLSTM, LSTM, ARIMA, Prophet, HWES, and SARIMA outperformed LSTM in ten regions. Table 12 presents the RMSE values obtained from the time series analysis with five different methods for the police districts in Philadelphia.

As shown in Table 12, for the 22 police districts in Philadelphia, the most successful method for 21 districts is BLSTM, and the least successful method for 19 districts is HWES. Figure 14 compares the RMSE values from the regional time series analysis for five methods in the Philadelphia crime dataset.

As shown in Fig. 14, the ranking of methods from most successful to least successful for the 15 regions was BLSTM, LSTM, SARIMA, Prophet, and HWES. In one region, SARIMA achieved better results than LSTM.

For the time series analysis of the San Francisco, Chicago, and Philadelphia police districts, the best results were achieved with BLSTM in 10, 16, and 21 different police districts, respectively, while the worst results were obtained with HWES in 9, 12, and 19 different police districts, respectively. In these datasets, the methods were ranked from most successful to least successful as BLSTM, LSTM, SARIMA, Prophet, and HWES in 6, 6, and 15 different police districts, respectively, when time series analysis was performed with five different methods.

The fact that BLSTM achieved the best results in all three datasets suggests that deep learning methods are generally more successful in time series analysis compared to statistical methods. As there were fluctuations in the sequential data in all three datasets, it can be concluded that BLSTM performed better by processing input data bidirectionally (forward and backward) compared to LSTM, which processes data in one direction, resulting in better training and fewer errors. Again, for the time series analysis conducted in three, 10, and one different police districts according to the datasets, better results were achieved with SARIMA compared to LSTM. To improve the unexpected results obtained with the Chicago dataset, hyperparameter optimization could be performed for both SARIMA and LSTM models.

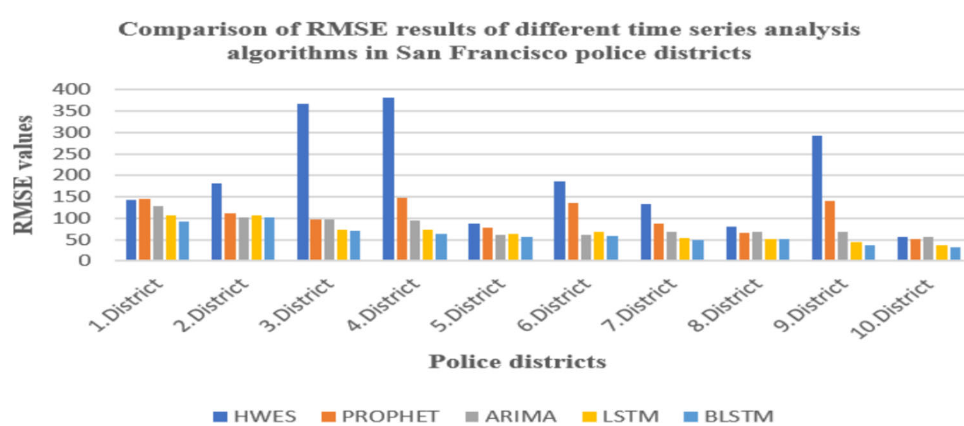


Fig. 12 Comparison of time series analysis results of five different methods in police districts in the San Francisco dataset

Table 11 RMSE values obtained by time series analysis conducted with five different methods in Chicago police districts

CH Police Districts	CH Police Districts No	Number of Crimes	HWES (RMSE)	PROPHET (RMSE)	SARIMA (RMSE)	LSTM (RMSE)	BLSTM (RMSE)
Central	1	310,671	348.15	168.0	132.36	89.7087	81.1353
Wentworth	2	364,614	291.42	238.89	63.22	87.8633	82.0699
Grand Crossing	3	390,961	285.14	230.46	86.2	83.7908	77.3915
South Chicago	4	437,620	223.58	188.96	88.88	100.1733	79.1936
Calumet	5	342,033	264.08	195.08	75.54	78.1419	62.8366
Gresham	6	450,516	177.61	224.85	89.62	96.0697	77.6745
Eaglewood	7	447,170	189.24	214.18	85.33	113.9078	87.3559
Chicago Lawn	8	518,381	225.16	293.58	79.2	72.3022	74.0838
Deering	9	377,205	189.71	172.53	60.49	73.3519	61.3378
Ogden	10	332,280	129.86	127.46	72.51	70.2685	60.0991
Harrison	11	495,537	155.87	138.03	115.64	101.4495	83.9250
Near West	12	381,120	208.57	237.62	94.16	89.3623	75.9482
Shakespeare	14	298,718	116.6	215.0	71.21	66.5006	58.6923
Austin	15	331,378	122.0	136.86	72.55	74.0730	64.5398
Jefferson Park	16	257,664	109.48	109.14	66.19	53.8775	50.9640
Albany Park	17	222,444	100.38	127.25	47.92	47.0326	43.7840
Near North	18	345,573	264.71	247.29	111.32	80.9705	79.1789
Town Hall	19	345,817	173.93	188.04	86.36	83.4122	77.0208
Lincolen	20	136,066	131.62	104.17	33.29	33.6973	31.2385
Morgan Park	22	252,659	143.04	144.6	54.44	60.2366	55.8196
Rogers Park	24	232,634	222.91	191.96	57.31	59.6527	54.8039
Grand Central	25	439,001	208.23	238.72	69.88	89.7892	76.0580

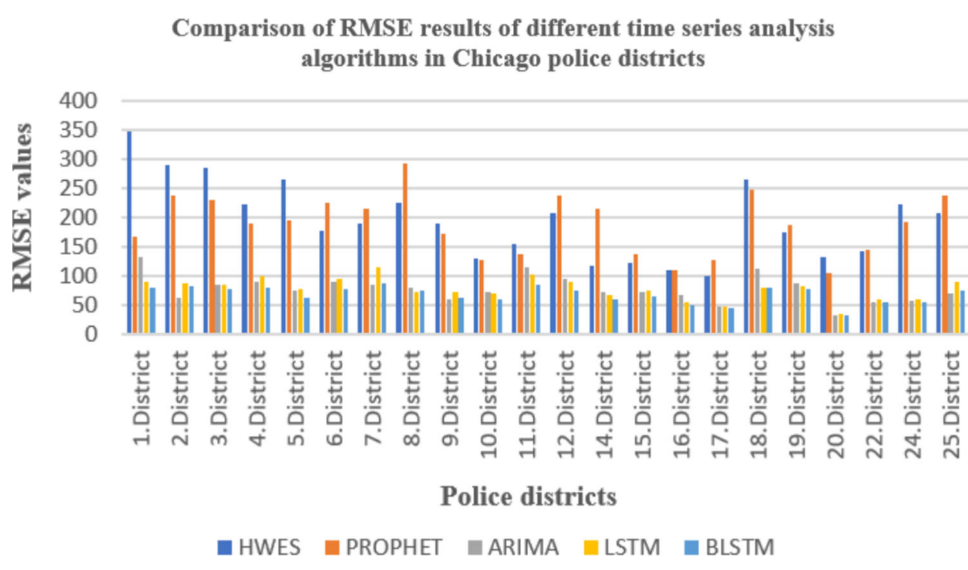


Fig. 13 Comparison of time series analysis results of five different methods in police districts in the Chicago dataset

4 Conclusion and future work

In this study, a six-stage methodology was proposed to contribute to the development of strategies for preventive policing. These stages include dataset preprocessing, exploratory data analysis, crime-type prediction model creation, time series analysis using LSTM and BLSTM deep learning methods, time series analysis and trend

Table 12 RMSE values obtained by time series analysis performed with five different methods in Philadelphia police districts

PH Police Districts	PH Police Districts No	Number of Crimes	HWES (RMSE)	PROPHET (RMSE)	SARIMA (RMSE)	LSTM (RMSE)	BLSTM (RMSE)
24th & Wolf	1	62,061	94.46	49.24	28.38	26.2814	18.5664
Harbison & Levick	2	151,617	101.28	104.75	60.39	54.4726	45.6055
11th & Wharton	3	121,197	576.53	128.67	60.18	59.1489	56.1789
Ridge & Cinnaminson	5	43,368	47.24	31.36	26.19	22.2040	16.6381
11th & Winter	6	135,550	148.33	85.58	73.9	54.2352	50.7995
Bustleton & Bowler	7	59,387	65.48	41.72	28.44	22.2775	20.7658
Red Lion & Academy	8	97,541	133.43	83.37	39.03	37.3895	33.4403
401 N. 21st	9	121,400	120.77	89.62	70.79	49.9913	47.9306
65th & Woodland	12	173,222	116.66	160.27	71.79	66.1291	61.4517
43 W. Haines Street	14	164,479	76.16	63.24	63.31	55.3669	53.3090
Harbison & Levick	15	238,452	131.83	113.33	77.09	85.2291	77.6910
39th & Lancaster	16	103,153	106.41	63.56	53.11	44.5641	40.8156
20th & Federal	17	95,621	65.46	80.76	38.76	41.2008	37.5169
55th & Pine	18	151,789	106.66	88.84	76.22	59.2123	53.5824
61st & Thompson	19	193,711	250.47	218.82	83.06	70.1530	68.5989
17th & Montgomery	22	183,940	141.37	116.36	73.86	62.1344	58.3424
3901Whitaker Ave	24	218,079	209.09	143.3	109.27	85.0851	77.7941
3901Whitaker Ave	25	191,906	166.56	74.03	63.9	55.7975	49.0268
615 E. Girard&Montgomery	26	116,066	115.45	67.26	51.26	41.4875	40.9897
5932 N. Broad & Champlost	35	178,280	114.56	89.44	77.9	57.1128	54.9785
22nd & Hunting Park	39	141,195	84.71	64.93	67.56	48.9137	46.9814
77th International Airport	77	11,237	13.24	10.63	9.8	7.3596	6.6420

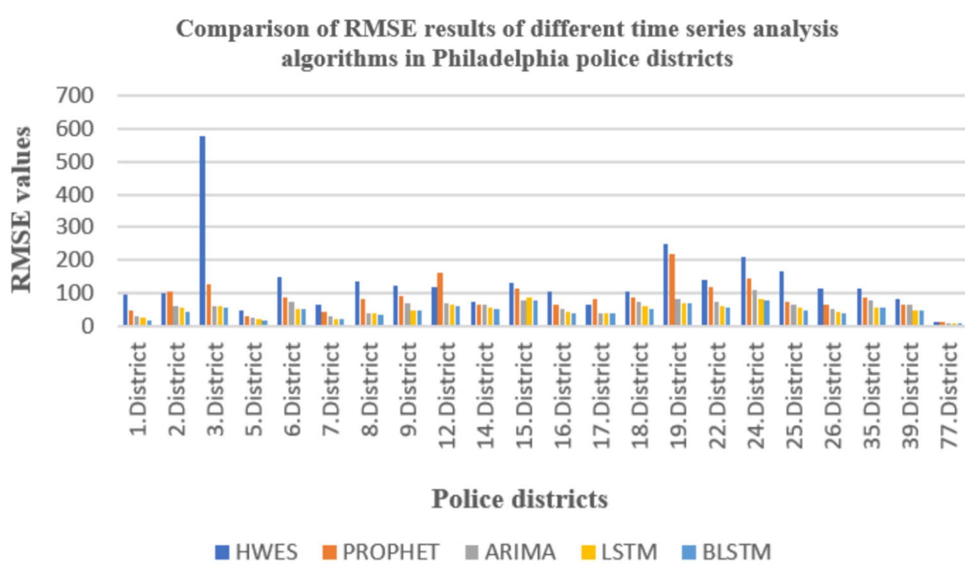


Fig. 14 Comparison of the results of time series analysis with five different methods in police districts of the Philadelphia dataset

forecasting for the next 5 years using HWES, Prophet, and SARIMA statistical methods, and regional time series analysis with statistical and deep learning methods.

As part of the proposed methodology, after preprocessing steps such as data cleaning, structuring attributes related to time and location, categorical data encoding, feature selection, and scaling, exploratory data analysis was conducted to identify crime density by time, location, and crime type. It was concluded that the most reliable month for crime occurrence is February, the most reliable day is Sunday, the most reliable hours are between 5–6 AM, the most common crime type is theft, and the police districts with the highest crime rates are districts 1, 8, and 15. The most dangerous months are the summer months, and the most dangerous hours are between 12:00 and 18:00. In summary, it was suggested that more police resources should be allocated to the 1st, 8th, and 15th districts, especially for theft crimes, during the summer months and between 12:00 and 18:00.

Crime-type prediction models were built using machine learning algorithms including XGBoost, CatBoost, RF, DT, MLP, KNN, GNB, and LR, with XGBoost models achieving 100% accuracy for 37, 34, and 33 different crime types in the San Francisco, Chicago, and Philadelphia datasets, respectively. Time series analysis was performed using LSTM and BLSTM for the San Francisco, Chicago, and Philadelphia datasets. The RMSE values for monthly time series analysis were 9.70, 28.95, and 17.25 for BLSTM; weekly time series analysis results were 10.28, 27.95, and 17.49 for BLSTM; and daily time series analysis results were 25.87, 50.21, and 32.13 for LSTM. These results are successful when compared to the literature.

In trend forecasting for the next 5 years using SARIMA, Prophet, and HWES statistical methods, it was predicted that there would be a moderate increase in crime numbers in San Francisco, and a sharp decrease in Chicago and Philadelphia. Based on the RMSE values of the models, the performance ranking was SARIMA > Prophet > HWES. By optimally distributing police resources across regions to reduce response times to crime incidents and consequently prevent or reduce crime, time series analysis for crime hotspots was conducted for 10, 22, and 22 different police districts in each of the three datasets using LSTM, BLSTM, SARIMA, Prophet, and HWES methods. The success of the methods was compared based on RMSE values. In general, the most successful method was BLSTM, and the least successful was HWES. Through regional time series analysis and the RMSE values obtained, it is aimed to contribute to the development of preventive policing strategies.

In the conducted study, according to the methodology followed, time series analysis using LSTM and BLSTM deep learning methods, time series analysis using the SARIMA statistical method, and crime trend forecasting for the next 5 years were performed. During regional time series analysis using both statistical and deep learning methods, parameter tuning was done either by using a restricted range (0–5) and grid search as done in the SARIMA model due to computational cost (three different datasets were used) or by using commonly used parameters in the literature as done in the LSTM and BLSTM models. Due to computational costs, a limited number of hidden layers was used in deep learning models. Thus, the use of various methods, including traditional machine learning, deep learning methods, and statistical methods, was made possible in the same study.

For future studies, it is aimed to use transfer learning with larger datasets, compare different deep learning methods, perform hyperparameter optimization to increase model performance, experiment with more diverse data stationary and scaling methods in SARIMA, LSTM, and BLSTM models, as these directly affect model success, and conduct more comprehensive parameter search. Additionally, it is recommended to try more diverse methods and models for time series analysis, test the obtained models with real-time datasets, and consider the issues of accountability, ethics, and reliability when using machine learning and especially deep learning models, which can be black-box models, in the legal system, as the decisions made could present challenges. Therefore, the inclusion of explainable AI in machine learning and deep learning methods used in the legal and preventive policing systems, as well as conducting explainable AI-focused crime prediction and analysis studies, is a key objective.

Author contributions Conceptualization was done by Esen Gül Ilgün and Murat Dener; methodology was done by Esen Gül Ilgün and Murat Dener; formal analysis was done by Esen Gül Ilgün and Murat Dener; investigation was done by Esen Gül Ilgün; resources were done by Esen Gül Ilgün; writing—original draft preparation was done by Esen Gül Ilgün; writing—

review and editing was done by Esen Gül Ilgün and Murat Dener; visualization was done by Esen Gül Ilgün; supervision was done by Murat Dener. All authors have read and agreed to the published version of the manuscript.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). The authors did not receive support from any organization for the submitted work.

Data availability Publicly available datasets were used.

Code availability Codes will be available at GitHub soon.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval We do not contain ethics issues.

Consent for publication We, all authors, consent publication of everything mentioned in the paper.

References

1. Lamond G (2007) What is a crime? Oxf J Leg Stud 27(4):609–632. https://doi.org/10.1093/ojls/gqm018
2. Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. Am Sociol Rev 44(4):588–608. https://doi.org/10.2307/2094589
3. Gudjonsson G (1988) The reasoning criminal. Rational choice perspectives on offending. Behav Res Therapy 26(3):246–287
4. Brantingham PL, Brantingham PJ (1995) Criminality of place: crime generators and Crime Attractors. Eur J Crim Policy Res 13(3):5–26. https://doi.org/10.1007/BF02242925
5. Zhang X, Liu L, Xiao L, Ji J (2020) Comparison of machine learning algorithms for predicting crime hotspots. IEEE Access 8:181302–181310. https://doi.org/10.1109/ACCESS.2020.3028420
6. Miethe TD, Meier RF (1994) Crime and its social context: toward an integrated theory of offenders, victims, and situations. Book News Inc, Portland
7. McNeal G (2014) Drones and aerial surveillance: considerations for legislatures. Pepperdine University School of Law. https://www.brookings.edu/articles/drones-and-aerial-surveillance-considerations-for-legislatures/ Accessed 2 March 2023.
8. Shah N, Bhagat N, Shah M (2021) Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. Vis Comput Ind, Biomed, Art. https://doi.org/10.1186/s42492-021-00075-z
9. Ferguson AG (2017) Policing predictive policing. Washington University Law Review 94(5). https://openscholarship.wustl.edu/law_lawreview/vol94/iss5/5
10. Santos RB (2014) The effectiveness of crime analysis for crime reduction: cure or diagnosis? J Contemp Crim Justice 30(2):147–168. https://doi.org/10.1177/1043986214525080
11. Pattavina A (2005) Information technology and the criminal justice system. Sage, Thousand Oaks
12. City and County of San Francisco (2003) Police department incident reports: historical 2003 to May 2018. https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-Historical-2003/tmnf-yvry/about_data. Accessed 12 September 2023
13. Catlett C, Cesario E, Domenico T, Andrea V (2019) Spatio-temporal crime predictions in smart cities: a data-driven approach and experiments. Pervasive Mob Comput 53:62–74. https://doi.org/10.1016/j.pmcj.2019.01.003
14. Yang L, Guofan J, Yixin Z et al (2024) A reinforcement learning approach combined with scope loss function for crime prediction on twitter (X). IEEE Access 12:149502–149527. https://doi.org/10.1109/ACCESS.2024.3473296

15. Feng M, Zheng J, Ren J et al (2019) Big data analytics and mining for effective visualization and trends forecasting of crime data. IEEE Access 7:106111–106123. https://doi.org/10.1109/ACCESS.2019.2930410
16. Nesa M, Shaha TR, Yoon Y (2022) Prediction of juvenile crime in Bangladesh due to drug addiction using machine learning and explainable AI techniques. J Comput Soc Sci 5(9):1467–1487. https://doi.org/10.1007/s42001-022-00175-7
17. De Blasio G, D'ignazio A, Letta M (2020) Predicting corruption crimes with machine learning. A study for the Italian municipalities. https://doi.org/10.13140/RG.2.2.21886.82245.
18. Wang C, Han B, Patel B, Rudin C (2022) In pursuit of interpretable, fair and accurate machine learning for criminal recidivism prediction. J Quant Criminol 39:1–63. https://doi.org/10.48550/arXiv.2005.04176
19. Tasnim N, Imam IT, Hashem MMA (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. https://doi.org/10.1109/ACCESS.2022.3171843
20. Stalidis P, Semertzidis T, Daras P (2021) Examining deep learning architectures for crime classification and prediction. Forecasting 3(4):741–762. https://doi.org/10.3390/forecast3040046
21. Wang B, Zhang D, Zhang D, Brantingham PJ, Bertozzi AL (2017) Deep learning for real time crime forecasting. *ArXiv abs/1707.03340*. https://doi.org/10.48550/arXiv.1707.03340
22. Palanivinayagam A, Gopal SS, Bhattacharya S, Anumbe N, Ibeke E, Biamba C (2021) An optimized machine learning and big data approach to crime detection. Wirel Commun Mobile Comput 2021(1):10. https://doi.org/10.1155/2021/5291528
23. Travaini GV, Pacchioni F, Bellumore S, Bosia M, De Micco F (2022) Machine learning and criminal justice: a systematic review of advanced methodology for recidivism risk prediction. Int J Environ Res Public Health 19(17):10594. https://doi.org/10.3390/ijerph191710594
24. Zhang X, Liu L, Lan M, Song G, Xiao L, Chen J (2022) Interpretable machine learning models for crime prediction. Comput Environ Urban Syst 94(2):101789. https://doi.org/10.1016/j.compenvurbsys.2022.101789
25. Safat W, Asghar S, Gilan S (2021) Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques. IEEE Access 9:70080–70094. https://doi.org/10.1109/ACCESS.2021.3078117
26. Kshatri SS, Singh D, Narain B, Bhatia S, Quasim MT, Sinha GR (2021) An empirical analysis of machine learning algorithms for crime prediction using stacked generalization: an ensemble approach. IEEE Access 9:67488–67500. https://doi.org/10.1109/ACCESS.2021.3075140
27. Khan M, Ali A, Alharbi Y (2022) Predicting and preventing crime: a crime prediction model using San Francisco crime data by classification techniques. Complexity 2022(3):1–13. https://doi.org/10.1155/2022/4830411
28. Kang HW, Kang HB (2017) Prediction of crime occurrence from multi-modal data using deep learning. PLoS ONE 12(4):e0176244. https://doi.org/10.1371/journal.pone.0176244
29. Han X, Hu X, Wu H, Shen B, Wu J (2020) risk prediction of theft crimes in urban communities: an integrated model of LSTM and ST-GCN. IEEE Access 8:217222–217230. https://doi.org/10.1109/ACCESS.2020.3041924
30. Tam S, Tanriöver ÖÖ (2023) Multimodal deep learning crime prediction using tweets. IEEE Access 11:93204–93214. https://doi.org/10.1109/ACCESS.2023.3308967
31. Sharma HK, Choudhury T, Kandwal A (2021) Machine learning based analytical approach for geographical analysis and prediction of Boston City crime using geospatial dataset. GeoJournal 88(S1):15–27. https://doi.org/10.1007/s10708-021-10485-4
32. Butt UM, Letchmunan S, Hassan FH, Koh TW (2024) Leveraging transfer learning with deep learning for crime prediction. PLoS ONE 19(4):e0296486. https://doi.org/10.1371/journal.pone.0296486
33. Chicago Data Portal (2001) Crimes: 2001 to present.https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2/about_data. Accessed 2 September 2023
34. Philadelphia Data Portal (2006) Crime incidents. https://opendataphilly.org/datasets/crime-incidents/. Accessed 30 September 2023
35. Battiti R (1994) Using mutual information for selecting features in supervised neural net learning. IEEE Trans Neural Netw 5(4):537–550. https://doi.org/10.1109/72.298224
36. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
37. Sharfuddin AA, Tihami MN, Islam MS (2018) A deep recurrent neural network with BLSTM model for sentiment classification. In: Proceedings of the international conference on Bangla speech and language processing (ICBSLP), IEEE: 1–4. https://doi.org/10.1109/ICBSLP.2018.8554396.
38. Box G, Jenkins G (1970) Time series analysis: forecasting and control. Holden-Day, San Francisco
39. Taylor SJ, Letham B (2018) Forecasting at scale. Am Stat 72(1):37–45. https://doi.org/10.1080/00031305.2017.1380080
40. Panda M (2020) Application of ARIMA and holt-winters forecasting model to predict the spreading of COVID-19 for India and its states. medRxiv. https://doi.org/10.1101/2020.07.14.20153908
41. Kalekar PS (2004) Time series forecasting using holt-winters exponential smoothing

42. Nur Şebnem EN, Pınar G, Ayhan A, Burcu BG (2022) Comparatıve performance analysis of Arima, prophet and holt-winters forecasting methods on European Covıd-19 data. Int J 3D Print Technol Digit Ind 6(3):556–565. https://doi.org/10.46519/ij3dptdi.1120718

Authors and Affiliations

Esen Gül İlgün[1] · **Murat Dener**[1]

✉ Murat Dener
muratdener@gazi.edu.tr

Esen Gül İlgün
egul.ilgun@gazi.edu.tr

[1] Department of Information Security Engineering, Graduate School of Natural and Applied Sciences, Gazi University, Ankara, Turkey