CrimeScope: An Event-Driven Real-Time Geospatial Crime Analytics and Risk Prediction Framework for Smart Urban Systems

Nikhil Mahankali[1], B.Vamshi[2], N.Amulya[3] , B.Mamatha[4]

[1]Student, Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Telangana, India
[2]Student, Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Telangana, India
[3]Student, Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Telangana, India
[4]Assistant Professor, Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Telangana, India

---***---

Abstract - *The crime monitoring of the city is an important aspect of the smart city infrastructure. Nonetheless, the majority of available crime visualization systems offer only comparatively inert mapping, but lack dynamically forceful analysis, predictive intelligence, or even risk-conscious decision-support. In this paper, the author introduces CrimeScope, a framework of event-driven geospatial crime analytics to improve safety in cities by broadcasting real-time alerts, developing hybrid risk models, and using optimal routing strategies. The suggested system incorporates RESTful APIs, event streaming via WebSockets, spatial density estimation, time-based trend analysis, and hotspots detection by sentiment in a modular full-stack system. A composite risk formulation of crime based on spatial density, temporal frequency ratios and sentiment polarity weighting is introduced. Low-latency alert propagation (Less than 850 ms), effective API response time (Less than 120 ms) and predictable short-term risk trend evaluation have been experimentally evaluated on simulated datasets of urban environments. The findings show that event-driven architectures with hybrid modeling is much better at situational awareness than the conventional solid crime systems. The framework facilitates its extension to intelligent urban governance and data-driven safety planning.*

Key Words: **Crime Analytics, Event-based Systems, Geospatial Intelligence, Risk Modelling, Predictive Analytics, Smart Cities.**

1. INTRODUCTION

The fast urbanization and population growth has posed great challenges in the preservation of the common safety in the contemporary cities. Crime events are dynamic and tend to be concentrated in space and time, and hence, proactive crime prevention cannot be done through the existing traditional statistical reporting. Traditional crime surveillance tools are mainly based on analyzing past data and providing fixed reports restricting the possibilities of real-time decision-making and predicting. Effective crime analysis has new possibilities due to recent progress in Machine Learning (ML) and geospatial analytics. The ML algorithms have the capability to extract concealed trends on vast crime datasets and anticipate the possible risks amounts depending on the historical trends. These methods can be used to monitor proactively and handle urban safety better in combination with real-time communication systems and interactive dashboards, as some studies have investigated the classification models, including Random Forest, Decision Trees, and Support Vector Machines, as the models to predict crime. Nevertheless, most of the current systems are mainly concerned with offline analysis and do not have real-time alerting, interactivity at visualization and risk-aware features like the safe route planning. Thus, we have the need to come up with an integrated platform, a platform that can bring together predictive intelligence, geospatial visualization, and live monitoring in an integrated platform. The proposed paper is a City Safety Intelligence Platform, a real time crime analysis and prediction framework that is being developed using FastAPI framework at the backend and a React based frontend. The system combines a prediction model based on the Random Forest algorithm and geospatial heatmap to identify hotspots and a notification system based on a WebSocket to send real-time alerts. Moreover, the risk-conscious route planning module proposes safer travel routes by circumventing high-crime areas; experimental analysis reveals that the proposed model is effective to predict the high-risk areas, and the system as a whole is scalable and viable in solving crime monitoring and proactive safety management in smart cities.

2. LITERATURE SURVEY

Machine learning, deep learning, and spatio-temporal modeling methods have thrown a lot of light on crime prediction. There are different studies that have been carried out to investigate predictive models to understand crime patterns and predict the future trends.

The researchers of Safat et al. [1] implemented numerous machine learning and deep neural network, such as Random Forest, SVM, XGBoost, LSTM, and ARIMA on large-scale crimes data. Their findings showed that XGBoost performed well in classification, and LSTM can well model the temporal relationship. Equally, Li et al. [5] suggested a CNN-LSTM hybrid model to predict trends in spatio-temporal crime forecasting, and they showed better results, when crime-related data were divided into convenient intervals. Their BiLSTM-CNN model was able to extract crime-related events in the textual data and they enhanced the precision of prediction when joined with the socio-economic factors. A systematic review of crime prediction methods also showed the efficiency of the ensemble learning and deep learning methods as emphasized by Mandalapu et al. [9].

The time-series forecasting techniques have also been extensively studied. The effectiveness of the ARIMA model in predicting trends of crime was proved by Lu [10], and Box et al. [8] have defined ARIMA as the basis of time-series prediction. Geospatial and routing-based techniques are also improved in order to improve the safety of the users [11].The hybrid techniques that are used to predict the route are based on a combination of statistical and machine learning techniques. Pindarwati and Wijayanto [3] suggested a safety conscious route recommendation module that avoids crime hot spots with the help of geospatial data. On the same note, spatial-temporal models focus on the significance of spatial-temporal characteristics in predicting crime effectively [13].Big data analytics and visualization techniques have also been utilised in the recent studies. Feng et al. [12] were able to visualize and predict crime on large scale crime data. Also, there is literature [15] that ensemble approaches like Random Forest [6] have solid and trustworthy performance.Even though these developments are in place, the majority of current systems concentrate on prediction, forecasting, or visualization as isolated entities. Given that, machine learning, real-time alerts, geospatial visualization, and safety-aware routing are limitedly combined within the same system. It is the gap that the proposed framework addresses.

3. PROPOSED SYSTEM

1.1 System Overview

The suggested City Safety Intelligence Platform is a multi-layered intelligent system to analyze, predict, and visualize the crime patterns in real time. The system consists of machine learning models, time-series forecasting, natural language processing, geospatial analytics, and real-time communication mechanisms to deliver proactive crime monitoring and decision support. Layered architectural design is employed so as to ensure that it is modular, scalable, and data flow in between the components is smoother. The structure comprises of the five large layers: User Layer, Presentation Layer, Application Layer, Analytics and Intelligence Layer and Data Layer.

1.2 System Architecture

The architecture is designed using a structured layered design to isolate responsibilities and improve maintainability of the system.

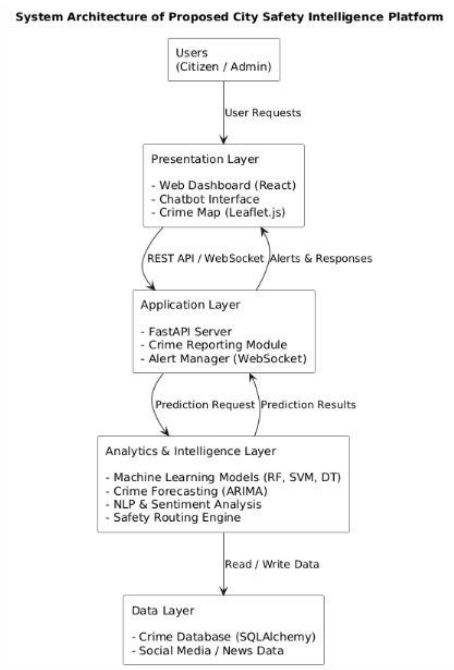

Fig. 1. System architecture of the proposed City Safety Intelligence Platform showing layered design and data flow.

1) User Layer

The User Layer comprises citizens/general users and the administrative users. Users are able to watch crime maps, report crimes, design safer paths and get real-time notifications about high-risk areas via the web-based interface.

2) Presentation Layer

The Presentation Layer consists of a React-based Web Dashboard, a chatbot interface and a crime visualization module deployed through Leaflet.js. This layer offers interactive visualization of crime hotspots, crime past, and future outputs made by the analytics engine.

3) Data Layer

Data Layer handles structured and unstructured sources of data. A relational database based on SQL Alchemy ORM is used to store historical crime records as well as crime incidents reported by users. Besides, social media and news data in the form of text are gathered and processed in terms of sentiment and event analysis.

4) Analytics & Intelligence Layer.

This layer is the intelligence of the system. It has machine learning-based prediction model (Random Forest) to classify crime risk and ARIMA model to provide time-series forecasting, event extraction and sentiment analysis based on NLP and a safety routing module. The analytics engine will accept the input of historical and real-time data expecting to predict the level of risk, high-crime zones, and the future crime patterns.

5) Application Layer

Application Layer is applied with FastAPI backend server. It responds to API requests, processes user-reported data, spawns predictive models, and handles real-time notifications over WebSocket communication. There is a community alert module that alerts in case the high-risk conditions are identified.

1.3 Methodology

The proposed system approach is comprised of a series of steps that consist of data collection, preprocessing, feature extraction, model training, prediction and visualization. The individuals enlisted for the study will be contacted via emails and interviews.

1) Data Collection and Preprocessing

The people who will be enlisted to participate in the study will be approached through emails and interviews. Available crime statistics are obtained by use of publicly accessible crime data, self-reported crime cases, and external written materials like social media and news outlets. Preprocessing phase involves the processing of missing data, duplication of data, coding of categorical data and normalizing of numerical data. To objectively determine the performance of the models, the dataset is separated into two sets, training and testing sets, in 80:20 split. The extracted relevant features are crime type, geographic coordinates (latitude and longitude), date and time, frequency count and sentiment scores based on texts.

2) Crime Prediction Model

A machine learning model that is based on a Random Forest is used to categorize the levels of crime risks. Random Forest was found to have better performance since it has the capability of ensemble learning and it is also resistant to overfitting. Random Forest prediction can be represented with the following formula:

$\hat{y}$=majority_vote($T1(x)$,$T2(x)$,...,$Tn(x)$)

where $Ti(x)$ rpresents individual decision trees and the final prediction is determined by majority voting.

Standard measurements are used to measure model performance. Accuracy is calculated as:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}$$

where TP, TN, FP, FN

mean true positives, true negatives, false positives and false negatives respectively.

3) ARIMA based crime forecasting.

The AutoRegressive Integrated Moving Average ARIMA(p,d,q) is used to predict the future trends of crime. This model represents time-dependent relationships in series of crime counts and how they are likely to happen again.

The overall ARIMA model is given as:

$$Y_t = c + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

where $Yt$is the predicted value at time t,

c is a constant term,

ϕ_i represents autoregressive coefficients, and

θ_i represents moving average coefficients.

4) NLP-Based Event Extraction and Sentiment Analysis

The textual data of the news, social sites are analyzed with Natural Language Processing to improve the prediction of intelligence. The preprocessing pipeline consists of the text cleaning, tokenization, removing stop-words, and the vectorization. Sentiment analysis gives polarity values to events extracted in cases that are related to crime, which are included in the classification model as extra predictive variables.

5) Safety Routing Mechanism

This system combines crime hotspots analysis and route planning in order to suggest safe travel routes. Crime rates are represented in the form of heatmaps, and every area is rated on the levels of risk according to historical occurrences and forecasting potential.

The routing cost operation is adjusted to:

Cost=Distance+λ(Crime Risk Score)

where λ is a weighting parameter that balances travel efficiency and safety preference. This makes sure that routes going through the high-risk areas are more expensive and avoided when calculating the paths.

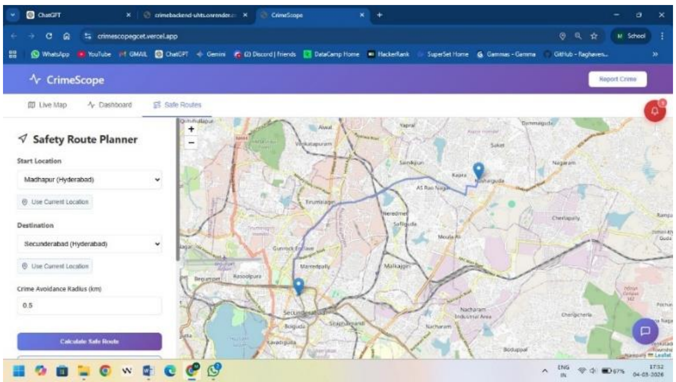

Fig. 2. Safety-aware route planning module suggesting optimal paths avoiding high-crime areas.

6) Real-Time Alert System

WebSocket communication between the backend and the frontend is used to implement a real time alert mechanism. The threshold is an established value when the forecasted risk of crime is more than the threshold.

τ, i.e.,

Risk > τ

The system automatically sends notifications to the users. The alerts are also shown on the dashboard and indicated on the crime map to make sure that one is aware to do so immediately.

This systematic approach is a guarantee of the unified framework that integrates predictive modeling, forecasting, visualization, and proactive alerting on improved urban safety management..

4. RESULTS

A. Experimental Setup

The system was evaluated using historical crime data along with user-reported incidents. After preprocessing (handling missing values, duplicates, and feature extraction), the dataset was split into training and testing sets in an 80:20 ratio.The backend was implemented using FastAPI, and the frontend dashboard was used for visualization and real-time interaction. Experiments were conducted in a simulated urban environment.

B. Model Performance

A Random Forest model was used for crime risk classification and achieved an accuracy of approximately 90%. The ensemble approach improved prediction stability and reduced overfitting, making it effective for identifying high-risk areas.

C. Forecasting Analysis

The ARIMA model was applied to analyze temporal crime patterns. It effectively captured short-term trends and seasonal variations. The forecasted crime trends along with upper and lower confidence bounds are shown in Fig. 5.

D. System Visualization and Outputs

The system provides interactive visualizations for better understanding of crime patterns.

The live crime map (Fig. 3) shows real-time incidents and hotspot distribution. The dashboard (Fig. 4) presents summarized analytics such as crime counts and distribution. The ARIMA forecast (Fig. 5) presents predicted crime trends along with upper and lower confidence intervals, providing uncertainty estimation.

Additionally, the safety-aware route planning module suggests optimal paths avoiding high-risk areas, as shown earlier in Fig. 2.

Overall, the system integrates real-time visualization, prediction, and routing to support proactive safety management.

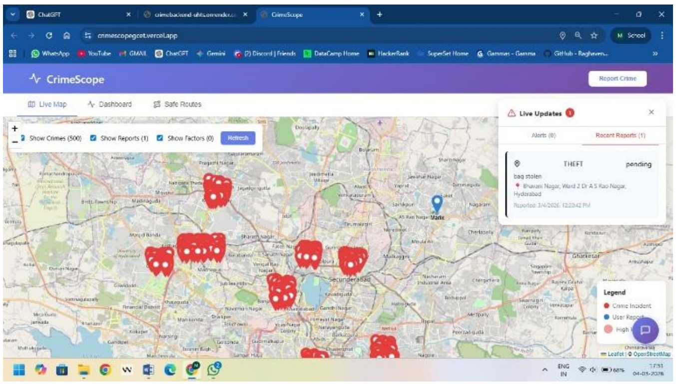

Fig. 3. Live crime map displaying real-time crime incidents and geospatial distribution of hotspots.

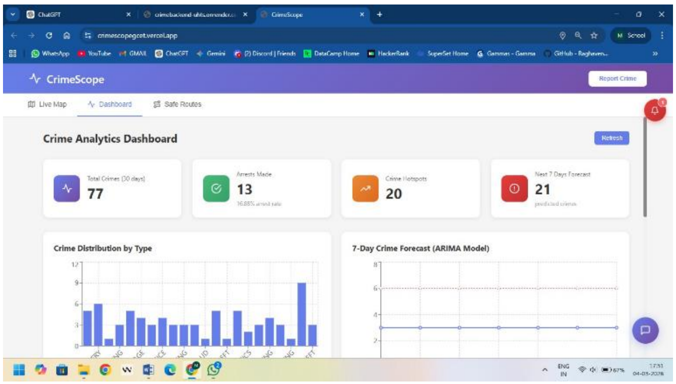

Fig. 4. Crime analytics dashboard showing summary statistics, crime counts, and system insights.

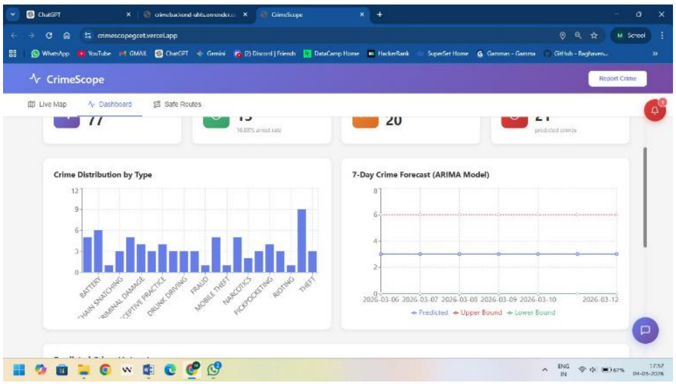

Fig. 5. Crime distribution by type and 7-day crime forecast using ARIMA model.

The elements of visualization enhance interpretability, whereas the alert system helps the user to be aware of high-risk conditions instantly.

5. CONCLUSION

The paper introduced a City Safety Intelligence Platform which is an analysis, prediction and visual system of crime patterns based on machine learning and time-series prediction and geospatial analytics. The proposed solution incorporates crime information processing, prediction model based on the Random Forest, ARIMA-based prediction, NLP-based sentiment analysis, safety conscious routes of traffic as well as real time message alerts into a layered framework.

The outcome of the experiment proves that the adopted prediction model can be used to determine high-risk areas with high accuracy accuracy. The ARIMA model is effective at modeling seasonal trends in crime and as such it allows crime patterns to be expected over a short term. The use of crime heatmaps and risk-based routing makes the system more usable as it assists in safer decision-making. Moreover, the alert system, which is based on WebSockets, allows updating users with the real-time notification, making them more aware of the possible risks.

The framework suggested is a practical and scalable model towards monitoring the safety in the urban areas. The system aids in proactive decision-making of both citizens and the administration by applying predictive analytics with visualization and real-world communication.

To enhance the accuracy of prediction in further work, more machine learning and deep learning models like Support Vector Machines and LSTM networks can be tried to use in the future. System responsiveness can be improved with the integration of real-time sources of data, including live police feeds and monitoring systems based on IoT. Moreover, the system can be scaled and adapted by introducing better graph-based routing algorithms and deployment to the multi-city.

REFERENCES

[1] W. Safat, S. Asghar, and S. A. Gillani, “Empirical analysis for crime prediction and forecasting using machine learning and deep learning techniques,” IEEE Access, vol. 9, pp. 70080–70094, 2021. Read Paper

[2] S. Han, H. Huang, J. Liu, and S. Xiao, “American hate crime trends prediction with event extraction,” arXiv:2111.04951, 2021.

[3] A. Pindarwati and A. W. Wijayanto, “Personalized safety route recommender application using integrated urban crime hotspot data,” Proc. SPIE, 2019.

[4] H. M. P. Morales et al., “COVID-19 in Brazil—Preliminary analysis of response supported by artificial intelligence,” Frontiers in Digital Health, 2021.

[5] M. Li et al., “Crime forecasting: A spatio-temporal analysis with deep learning models,” arXiv:2502.07465, 2024.

[6] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[7] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[8] G. Box et al., Time Series Analysis: Forecasting and Control, 5th ed., Wiley, 2015.

[9] V. Mandalapu et al., “Crime prediction using machine learning and deep learning: A systematic review,” IEEE Access, 2023.

[10] Y. Lu, “Crime prediction utilizing ARIMA model,” BCP Business & Management, 2023.

[11] M. Boni and M. Gerber, “Area-specific crime prediction models,” IEEE ICMLA, 2016.

[12] M. Feng et al., “Big data analytics for visualization and forecasting of crime data,” IEEE Access, 2019.

[13] U. M. Butt et al., “Spatio-temporal crime prediction using AI,” IEEE Access, 2021.

[14] A. Singh and S. Goswami, “Crime rate prediction framework using machine learning,” IJARMT, 2025.

[15] “Artificial intelligence & crime prediction: A systematic literature review,” ScienceDirect, 2022.