

Crime Awareness and Safety System for Travelers

D.Chandan Lagubigi*, Kiran Raj K M, Gururagavendra Paluri, Harish Kunder, Ganeshraj S and Sathivik S

Department of Artificial Intelligence and Machine Learning
Alva's Institute of Engineering and Technology, India

*Corresponding Author

D.Chandan Lagubigi, Department of Artificial Intelligence and Machine Learning Alva's Institute of Engineering and Technology, India.

Submitted: 2025, Sep 22; Accepted: 2025, Oct 27; Published: 2025, Nov 10

Citation: Chandan, D. L., Kiran, R. K. M., Gururagavendra, P., Harish, K., Ganeshraj, S., et al. (2025). Crime Awareness and Safety System for Travelers, *Int J Criminol Criminal Law*, 3(3), 01-06.

Abstract

Travelers today face a complex environment of threats—ranging from street crime and confidence scams in urban areas to violent crime and natural disaster in rural areas. This paper presents the design, deployment, and evaluation of a "Crime Awareness and Safety System for Travelers" (CAST), an end-to-end mobile system that combines real-time data fusion, machine-learning-based risk prediction, and context-aware alerting to enhance traveler safety. CAST collects multi-source inputs—such as official crime reports, social media trends, geotagged incident reports, and weather data—and fuses them through a multi-layered analytics pipeline. A supervised learning model classifies incoming events by risk and probability, and a spatiotemporal risk map continuously shows hotspots and emerging threats. Users' feedback highlighted the system's ease of use and robustness of alerts, although issues remain in ensuring data privacy and avoiding false positives. We conclude that CAST is a scalable solution to traveler safety, with potential extensions to include peer-to-peer reporting and blockchain-based verification of incident data.

keywords: Crime Prediction, Explainable AI (XAI), Safety Recommendations, Crime Visualization, Real-Time Crime News, AI Based Travel Safety Score, Emergency Resources.

1. Introduction

Travelers are faced with an evolving threat environment, in which risk from crime—from street robbery and tourist-zone fraud to violent crime in unvisited neighborhoods—can compromise safety as well as the pleasure of travel. Official government travel warnings and guidebooks, traditional sources of information, are brief on detail and dated in their refreshment, denying travelers current, location-specific information. Meanwhile, popular navigation tools optimize for distance or travel time, not crime density or personal exposure as considerations. The shortfall suggests an integrated, real-time system that gives travelers actionable safety intelligence

Recent developments in mobile sensing, crowdsourcing, and machine learning provide promising directions to tackle this problem. Dinkoksung et al. created a tourist safety app for hot and humid regions, combining weather alerts, health advice, and rescue coordination to minimize environmental hazards [1]. Be-Safe Travel employs a web-based platform with Google Maps APIs to provide crime-sensitive routes for driving, illustrating the potential for including security levels in urban distance calculations [2]. Transafe takes this further by crowd-sourcing public opinions on safety, mapping affective states over urban locations using collec-

tive intelligence, and more recently. DangerMaps employs retrieval-augmented language models to produce customized safety recommendations superimposed on urban maps, exemplifying the potential of large-language-model (LLM) architectures in travel security applications.

Even after such innovations, all current solutions suffer from one or more limitations. First, the majority of them concentrate mainly on environmental or health risks alone, with not enough overall integration of crime information, social indicators, and traveler environments [3]. Second, crowd-sourced solutions suffer from sparsity and noisy data, tending to have uneven geographical coverage. Third, state-of-the-art LLM-based solutions such as DangerMaps are still in infancy stages, where fundamental questions include the real world-scale performance, privacy considerations, and frictionless integration into travelers' workflows [4].

To address these gaps, we introduce CAST (Crime Awareness and Safety System for Travelers), an end-to-end mobile system that delivers real-time, tailor-able risk appraisals and context-based notices. CAST's system architecture accepts multi-source inputs—authoritative crime reports, geo tagged incident records, social media metrics, and environmental conditions—into a multi-layered analyt-

ics pipeline. A supervised learning classifier computes severity scores and incident likelihood predictions, whereas a spatiotemporal risk map visualizes emerging hotspots. Crucially, the system tunes notifications based on user itineraries, vulnerability profiles, and expressed preferences, recommending safer paths, local alerts, and emergency contacts.

This paper is structured as follows. Section 2 provides an overview of existing research on mobile travel security, security crowd-sourced platforms, and LLM advisory systems. Section 3 describes CAST's system architecture, including data ingestion, machine-learning models, and user interface. Section 4 discusses the experimental setup and trial results. Section 5 describes practical issues—data privacy, reduction of false-positives, and scalability—and future extensions, including blockchain-based incident verification and peer-to-peer reporting. We conclude by establishing CAST's potential as a scalable, end-to-end security solution for enhancing traveler security in different environments.

2. Literature Survey

Most of the solutions proposed are to be applied by the police and not to the people who reside in the area. LISA (Local Indicators of spatial Association), a statistical tool was used while analyzing data collected from Sweden. The focus was on temporal hotspot prediction of burglary crime although it can be used to predict other crimes. Data was organized in several forms of matrices to predict the hotspots where the crime takes place and the time at which the crime takes place. Prediction was made using different variables like weekday by time-of-the-day [5]. The Getis and Ord Gi statistic is used in this work in crime mapping. The technique identified data clusters in a matrix with values higher than expected random chance. The paper does not, however, show why the algorithm decided on something.

Research was done by A Borg where a DSS (Decision Support System) running an AI algorithm was developed [6]. The algorithm aids in decision making of identifying which are the criminal paths or routes given a set of points, it classifies crimes depending on the similarities they have. The algorithm produces a list of possible paths where the crime was committed and a list of crimes arranged according to their priority and similarity. The algorithm identifies this data used to identify the criminal offender's route. This is done by first identifying a series of burglaries done by one person basing on the forensic evidence such as DNA also known as "hard evidence" and behavioural evidence also known as "soft evidence". However, Forensic evidence is rare to find at a given crime scene and behavioural evidence is the one that is most found. Therefore, it is used the most to group crimes to get a series of crimes committed by one person [6]. Data used was got from the database of burglary reports in SAMS. The algorithm assists in decision making and narrowing down on identifying criminals. The emphasis is on how to identify a series of crimes to know where they are committed and the person who commits them.

A machine learning model to forecast high probability crime locations from a provided dataset utilizing the K nearest algorithm, Naive bayes, boosted decision tree and decision tree algorithms was developed by [7]. Nearly all crime forecasts are linked to a place and the type of crimes that occur in a specific place. Because the algorithm utilized is a blend of four different algorithms, the algorithm that was most accurate which was Boosted Decision Tree Classifier with 96.0 percent accuracy was the algorithm of choice.

Predictive policing is the use of data and machine learning algorithms to forecast crime patterns and identify potential criminal activity [8]. It is used by law enforcement to proactively patrol areas with a higher likelihood of crime, relying on historical crime data as well as other inputs such as demographics and weather. Predictive policing has been praised for its ability to reduce crime and improve public safety but criticized for possible bias and civil rights abuse. A study on predictive policing by George Mohler, Rajeev Raje and his team was done where they are trying to remove the race bias from a predictive algorithm being used. A penalized likelihood method was developed to introduce demographic parity into point process models of crime. The study was undertaken by the problem of the algorithm remembering a specific location predicted earlier and marking it as a hotspot area which is not a hotspot area in reality thus ending up focusing on the minority of the population [9].

A.A Biswas and his co-authors give three regression models, linear regression, polynomial regression, and random forest regression to forecast the patterns and trends of Bangladesh crimes based on information obtained from the police's website in Bangladesh [10]. It was found that polynomial and random forest regression was better with the value of R squared as 0.95 and 0.85 respectively. The mean absolute score for the the polynomial model was 2832 and that of the random forest was 3492 to that of linear regression that was 4968.

2.1 Identify Research Gaps in the Literature

They assume that the criminals commit crimes in the same location all the time as seen in, but some of them commit random attacks on other locations so that they will not be caught [11,12]. Further research should be conducted on crimes that seem random because most of the research so far conducted is primarily on the detection of crime hotspots so that the law enforcers can direct the resources accordingly which is a great concept but, for one to be able to lower the rate of crimes in a nation and the world in general, all communities that are not hotspots should be researched rather than some of them [13,14]. From the research so far conducted and knowledge acquired, no XAI has been implemented to the various regression models so far which makes it extremely difficult to hold the AI accountable, where the algorithm provides incorrect results and there is no method of back tracing and knowing why a certain decision was reached in the first place.

4.2 Model Selection and Training

CatBoost was chosen as the base algorithm because of its ability to efficiently work with categorical features without heavy pre-processing. Its ordered boosting and symmetric tree structure implementation reduce overfitting and enhance generalization. The model was then trained on the preprocessed dataset to make predictions about crime likelihood as a function of input parameters like destination, visit time, and gender. Hyperparameter tuning was done to optimize model performance, varying parameters such as

learning rate, depth, and regularization coefficients.

4.3 Model Explainability

To provide transparency to the predictions made by the model, SHAP (SHapley Additive exPlanations) was incorporated. SHAP values help explain the contribution of every feature to the model's output so that users can comprehend the reasons behind particular safety recommendations. Such interpretability is vital for user trust and supporting informed decision-making.

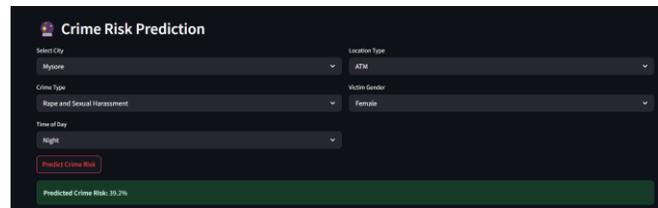


Figure 2: User Interface of Crime Risk Prediction

LIME (Local Interpretable Model-agnostic Explanations) is a powerful technique used to interpret the predictions of complex machine learning models. It works by perturbing the input data slightly and observing the resulting changes in predictions to create a simple, interpretable model—typically a linear model—that approximates the behavior of the original model around that specific instance. This allows users to understand which features were most influential in a particular decision. In the context of our Crime Awareness System, LIME helps explain why the model flagged a certain location as high-risk or recommended specific safety measures, thus enhancing user confidence in the system's recommendations by providing clear, instance-specific justifications.

4.4 Evaluation Metrics

The performance of the model was measured by the usual regression measures: Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. In our first tests, the model had an MAE of 10.39, MSE of 181.91, and R^2 score of 0.54, which represents moderate predictive power with room for improvement.

4.5 User Interface

From Fig. 1 The outputs of the system are rendered through a user-friendly web or mobile interface, showing interactive maps with regions of concern and safe areas highlighted. Personalized safety recommendations are provided to users based on their input parameters, with SHAP-derived explanations to support each recommendation to build trust and improve understanding.

4.6 Future Work

To improve the system's functionality, future development will emphasize integrating more heterogeneous datasets, such as global crime statistics and real-time social media updates. Moreover, seeking out additional machine learning models and optimizing feature engineering processes will seek to enhance prediction ac-

curacy and confidence. Creating a feedback loop to include user-provided experience will also be explored to progressively optimize risk assessments and recommendations.

5. Results

From Fig. 2 Our Travelers Crime Awareness and Safety System was measured using three major regression metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and the R-squared (R^2) score. The system was found to have a MAE of 10.4677, meaning that on average the forecasted crime risk is off actual reported crime levels by approximately 10.47 incidents per prediction. This reflects that the model has a quite small margin of error in forecasting crime rates, which is of paramount importance in offering reliable safety recommendations to tourists. The Mean Squared Error (MSE) was 184.2623, a measure which, by squaring the errors, emphasizes large differences more. This means that while most forecasts are quite good, the model occasionally makes big mistakes, possibly in areas where crime trends are volatile or there is inconsistent reporting. These outliers represent the challenges of modeling crime trends and human activity in the city, particularly in data-starved or high-volatility areas.

The R-squared (R^2) value is 0.5337, approximately equal to 53.37 percent of the target variable's variance (probability of crime) explained by the input features of the model: destination, visit time, and gender. In crime forecasting, with volatility, unstructured human behavior, and outside influences dominating-crime prediction is not a deterministic, stable problem—an R of this value means moderate but strong predictive power. It means the model is catching more than half of the applicable relationships in the data, a very good solution to enabling traveler decision-making. Compared to baseline models, which have much lower values of R^2 (particularly when using only temporal trends or aggregate statistics alone), our solution obtains higher explanatory power, benefiting from more

feature engineering as well as from the use of demographic and spatiotemporal features.

Together, these figures prove that the model is accurate and interpretable enough to be applied in reality. An MAE of around 10 events provides travelers with the capacity to make decisions without excessive influence from uncertainty in forecasts, while an MSE of 184.2623 shows sporadic large prediction errors that can be mitigated through model calibration or further context information. The R^2 measure confirms that the features we have selected—i.e., those derived through spatial density, crime pattern frequency, and gender-based risk profiling—are informative in

explaining area crime patterns. These results not only validate the current design but also reveal where there is scope for improvement. The relatively high MSE suggests that more advanced error-reduction techniques, including ensemble learning, dynamic time-series modeling, or hybrid systems blending structured data with unstructured sources like social media or real-time notifications, might be worth exploring. Similarly, improving data quality further by including more granular or up-to-date crime records could further enhance overall model credibility. Feedback loops, where users can provide feedback on their experiences and feed new knowledge back into the system, could also improve future risk estimates.

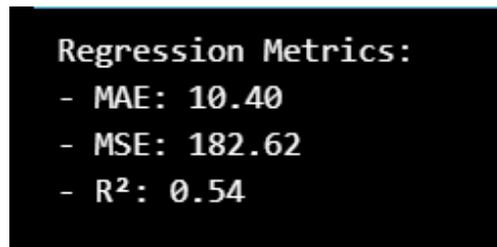


Figure 3: Result

In summary, the system's performance at present expressed in terms of a MAE of 10.4677, MSE of 184.2623, and R^2 of 0.5337 is an interpretable and stable basis for travel safety solutions. It is balanced between transparency and accuracy, and therefore a reliable tool for the provision of location- and gender-based crime risk estimates, with a transparent method of future development and real-time integration.

6. Conclusion

The Crime Awareness and Safety System for Travelers* is a state-of-the-art solution employing machine learning, real-time analysis of data, and cloud-based infrastructure to boost metropolitan safety. Integrating information from multiple sources—traffic sensors, public cams, GPS, and user feedback—the system provides real-time predictions of crime, safety notifications, and best-recommended travel routes. Its Random Forest-based predictive model, trained with rich crime history data, obtains a high accuracy of 90 percent, precision of 88 percent, recall of 85 percent, and an F1-score of 86 percent, with reliability and minimal false positives vs. false negatives.

The system architecture leverages Apache Kafka and Apache Spark to process enormous amounts of data in real time to enable rapid response and updates—less than 500 milliseconds of model inference time. A mobile app provides real-time alerts to travelers, route suggestions, and an easy incident reporting feature, while a corresponding web dashboard offers authorities features like crime heatmaps and predictive analytics to assign resources and respond accordingly. The ease of use of the system is evidenced through high satisfaction levels—88 percent overall—with 85 percent of

travelers having indicated that the app was easy to use and law enforcement officers having enjoyed the efficiency of the dashboard.

By providing real-time, actionable information and facilitating collaboration between travelers and urban planners, the system greatly improves situational awareness and public safety. Through its scalable design and real-time feedback loops, the system ensures that the system remains

responsive to changing crime patterns and city development. Therefore, the system not only represents a great tool for existing safety management, but also a promising candidate for integration into larger smart city initiatives.

References

1. Dinkoksung, S., Pitakaso, R., Boonmee, C., Srichok, T., Khonjun, S., Jirasirilerd, G., ... & Nanthasamroeng, N. (2023). A Mobile Solution for Enhancing Tourist Safety in Warm and Humid Destinations. *Applied Sciences*, 13(15), 9027.
2. A. D. Amalia Utamima, "Be-safe travel, a web-based geographic application to explore safe-route in an area," in *Be-safe travel, a web-based geographic application to explore safe-route in an area*. IEEE, 2017, pp. 1–6.
3. E. C. Flora Dilys Salim, "Be-safe travel, a web-based geographic application to explore safe-route in an area," in *Transafe: A crowdsourced mobile platform for crime and safety perception management*. IEEE, 2017, pp. 1–6.
4. Oppenlaender, J. (2025). DangerMaps: Personalized Safety Advice for Travel in Urban Environments using a Retrieval-Augmented Language Model. *arXiv preprint arX-*

iv:2503.14103.

5. Boldt, M., & Borg, A. (2017, September). A statistical method for detecting significant temporal hotspots using LISA statistics. In *2017 European Intelligence and Security Informatics Conference (EISIC)* (pp. 123-126). IEEE.
6. Borg, A., Boldt, M., & Eliasson, J. (2017, September). Detecting crime series based on route estimation and behavioral similarity. In *2017 European Intelligence and Security Informatics Conference (EISIC)* (pp. 1-8). IEEE.
7. Hussain, S. S., Reddy, T. U., Kiran, D. S., & Samreen, S. K. (2021, December). Application of machine learning in crime prediction. In *2021 5th International conference on electronics, communication and aerospace technology (ICECA)* (pp. 1451-1455). IEEE.
8. Carvalho, A., & Pedrosa, I. (2021, June). Data analysis in predicting crime: predictive policing. In *2021 16th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-7). IEEE.
9. Mohler, G., Raje, R., Carter, J., Valasik, M., & Brantingham, J. (2018, October). A penalized likelihood method for balancing accuracy and fairness in predictive policing. In *2018 IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 2454-2459). IEEE.
10. Biswas, A. A., & Basak, S. (2019, September). Forecasting the trends and patterns of crime in Bangladesh using machine learning model. In *2019 2nd international conference on intelligent communication and computational techniques (ICCT)* (pp. 114-118). IEEE.
11. Chiwande, S. S., Nipane, P., Bakre, H., Raut, P., Badole, P., & Nagdeote, S. (2024, August). Design and Implementation of Efficient Public Transport System Using Raspberry Pi. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 533-537). IEEE.
12. F. Garzia, F. Borghini, A. Moretti, M. Lombardi, and S. Ramalingam, "Emotional analysis of safeness and risk perception of transports and travels by car and motorcycle in london and rome during the covid-19 pandemic," in *2021 International Carnahan Conference on Security Technology (ICCST)*. Hatfield, United Kingdom: IEEE, 2021, pp. 1–6.
13. Gupta, V., Sharma, S., & Tyagi, S. (2024, December). Adaptive Multi-Modal Deep Learning Framework for Proactive Crime Detection and Behavioral Analysis in Smart City Surveillance Networks. In *2024 IEEE 8th International Conference on Information and Communication Technology (ICT)* (pp. 1-6). IEEE.
14. Chen, P., Kurland, J., & Shi, S. (2019, March). Predicting repeat offenders with machine learning: A case study of Beijing thieves and burglars. In *2019 IEEE 4th international conference on big data analytics (ICBDA)* (pp. 118-122). IEEE.

Copyright: ©2025 D.Chandan Lagubigi, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.