International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, Delaware, USA
Volume 13– No. 2, February 2026 – www.ijais.org

An Android-based Crime Data Collection and Analytics: An Integrated Framework

Sophie Uwho
Centre for Information and Telecommunication Engineering
University of Port Harcourt

Ifeoma B. Asianuba
Electrical/Electronic Engineering
University of Port Harcourt

ABSTRACT

In many developing countries, crime reporting is on the decline. It involves manual processes, and could be inaccessible, thereby reducing the responsiveness of law enforcement agencies and weakening public trust. Leveraging the widespread availability of mobile devices, the study presents an integrated Android-based intelligent crime reporting and analytic framework that incorporates geolocation, multimedia evidence captures, Naïve Bayes classification, and exponential smoothing forecasting for proactive public safety management. The application was developed using the Flutter framework to ensure a responsive and user-friendly interface, while backend functionalities were implemented using PHP and MySQL. The Naïve Bayes classifier achieved an accuracy of 88.6%, while exponential smoothing forecasting attained a Mean Absolute Percentage Error (MAPE) of 8.55%, indicating excellent predictive performance. Usability testing involving 130 participants demonstrated high acceptance, with 96% reporting ease of navigation and 92% highlighting the accuracy of GPS-based reporting. The system offers a scalable, user-friendly, and technologically solution for enhancing public security.

Keywords

Android application, Crime reporting, Exponential smoothing, GPS, Naïve Bayes classifier, Predictive analysis

1. INTRODUCTION

Crime incidents continue to pose a serious public safety challenge with the increasing cases of armed robbery, kidnapping, and vehicle accidents have diminished commuter confidence and disrupted regional mobility. A major barrier to timely intervention is the reliance on outdated, manual crime reporting mechanisms such as in-person visits to police stations or emergency hotlines, both of which are hindered by communication delays [1], [2]. These obstacles often contribute to slow reaction of the law enforcement, loss of important evidence, and under anticipated standard of crime data, thus compromising the state security.

The widespread adoption of smartphones and mobile internet presents an opportunity to modernize crime reporting through mobile applications. Some of the available mobile solutions offer low functionality, including simple crime notifications or one-way reporting but lack essential capabilities, including machine learning (ML) classification, geospatial intelligence, incident prediction, and system-wide analytics [3], [4]. Naïve Bayes classification models demonstrated great performance in identifying the types of incidents in real time in mobile contexts [5], and exponential smoothing with its classical time series analysis roots, provides a robust and adaptive predictor of trends, seasonality, and noise in time-related data [6]. Despite these advances, existing literature shows no distinct integrated system that combines instant reporting, machine-learning classification, and time-series forecasting within a single operational platform tailored to this context.

This study addresses the gap by developing an Android-based crime data collection and analysis system that combines: (a) crime reporting with geolocation and multimedia support; (b) Naïve Bayes classification for incident categorization; and (c) exponential smoothing forecasting to predict long-term crime trends. Finally, it strives to evaluate the system's performance in terms of high classification accuracy, strong forecasting reliability, and high user satisfaction across key usability metrics.

2. LITERATURE REVIEW

2.1 Mobile Crime Reporting Systems

Mobile applications represent a major shift in emergency communication by enabling direct, real-time citizen participation in public safety. Usman [7], noted the global shift towards mobile safety tools, yet in Nigeria, reporting remains largely manual and reactive [8]. As mentioned earlier [3] created a secure mobile crime reporting system that is resource-constrained, based on the findings of which the major concerns regarding usability are bandwidth and device constraints. Hawk Eye app in Nigeria [4], which offers real-time alerts, does not have inbuilt analytical features. C APP developed by [5] allows users to submit complaints and missing person reports but lacks real-time communication and GPS. Mahalakshmi [9] developed a facial recognition app for use only by the police, which excludes public participation and GPS features. Alameri in [10] created a geolocation-based app that aggregates reports but lacks real-time deployment and is limited to community crimes.

2.2 Geospatial Crime Mapping and Hotspot Analysis

Geospatial technologies support hotspot identification, allowing law enforcement to allocate resources strategically. Global studies show extensive use of Geographic Information Systems (GIS) for crime visualization. For example, [11] used GIS to track the kidnapping hotspots on the Abuja-Kaduna corridor in Nigeria and have shown that spatial analytics can be beneficial in solving region-focused threats. GIS based spatial analysis was also implemented by [12], to map the locations of accident hotspots in Nigeria and they were also able to determine the spatial patterns of clustering which were found to be related to the environmental and infrastructural conditions. GIS-policing has developed significantly, using a special feature that cross-examine the crime patterns through predictive GIS layers that incorporate socioeconomic and time-related variables. A Turkish study by [13] created a mobile application that illustrated color-coded crime rate based on

Numbeo data; the study had no special features of direct user reporting, which showed the lack of continuity between visualization and active data collection.

2.3 Machine Learning for Crime Prediction and Classification

Machine learning emerged as one of the key strategies to crime analytics in the world. Algorithm like Random Forest, Neural Networks, Naïve Baye, etc. are always reported to be effective. Tamir et al. [14] applied ML to predict crime in major cities using the Chicago Police Department's CLEAR system with over 6 million records. Among the models tested; Random Forest, K-Nearest Neighbors, AdaBoost, and Neural Network, the Neural Network performed best with 90.77% accuracy. The study shows how predictive algorithms can reveal crime trends and support law enforcement in resource allocation and crime prevention. Raksha et al. [15] proposed a blockchain-based anonymous tip-off system, demonstrating the potential of combining anonymity and prediction for civic engagement.

2.4 Time-Series Forecasting

Forecasting crime trends is essential for proactive policing, resource planning, and strategic deployment. Due to inconsistent historical records, data loss, and non-digitized reporting channels, researchers often adopt classical forecasting models that perform reliably under data scarcity. Extensions such as Holt's linear model and the Holt-Winters method incorporate trend and seasonality components, making exponential smoothing applicable to a wider array of real-world data, including crime rates that often display periodicity and structural shifts [16]. Holt's exponential smoothing, has been demonstrated as a robust method for sparse or partially synthetic datasets, performing well in low-data environments [17]. This makes it appropriate for regions where police reporting systems are not yet fully digitized. The limited number of forecasting studies underscores a need for tools that can both generate and analyze crime trends collaboratively.

2.5 Research Gaps

The literature reveals several critical research gaps:

2.5.1 Lack of Integrated Systems

No prior solution unifies mobile crime reporting, machine learning–based classification, and time-series forecasting in a single framework.

2.5.2 Limited Mobile-optimized ML Deployment

ML models in African are often evaluated offline rather than deployed on mobile systems for real-time inference.

2.5.3 Weak GIS-ML Forecasting Integration

Existing research rarely connects spatial analytics with classification and temporal forecasting.

These gaps justify the development of a simple, integrated mobile data collection and crime-analytics system tailored for incorporating validated analytical techniques.

3. METHODOLOGY

The methodology adopted in this study integrates system design, mobile application development, machine learning classification, and time-series forecasting into a unified analytical workflow. The approach is structured into five stages: system requirements elicitation, system design and architecture modeling, mobile application implementation, machine learning model development, and forecasting analysis.

3.1 System Requirements

User and functional requirements were collected through a review of existing crime-reporting platforms and informal consultations with public safety stakeholders. Key requirements included instant crime reporting, GPS-based location capture, multimedia upload capability, automated crime-type classification, and administrative dashboard visualization. The resulting data flow begins with user-submitted incident reports, which are transmitted to the backend server, preprocessed, classified using Naïve Bayes, stored in the database, and visualized on the administrative dashboard. Historical data are then used to forecast future crime patterns.

3.2 System Architecture

The system architecture is a modular and service-oriented design that consists of three layers, the mobile-client and the application-server layer, and the analytics-engine layer. The mobile client manages the collection of data, geolocation and interaction. The server layer provides the functionality of authentication, storage of reports, and communication to the analytics engine. The analytics engine is used to classify crimes and make time-series forecasts automatically. The system architecture diagram (Fig. 1) shows how data flows between the various components, as well as how the different components can integrate with each other between the ML classifier, forecasting module, and the database management system.

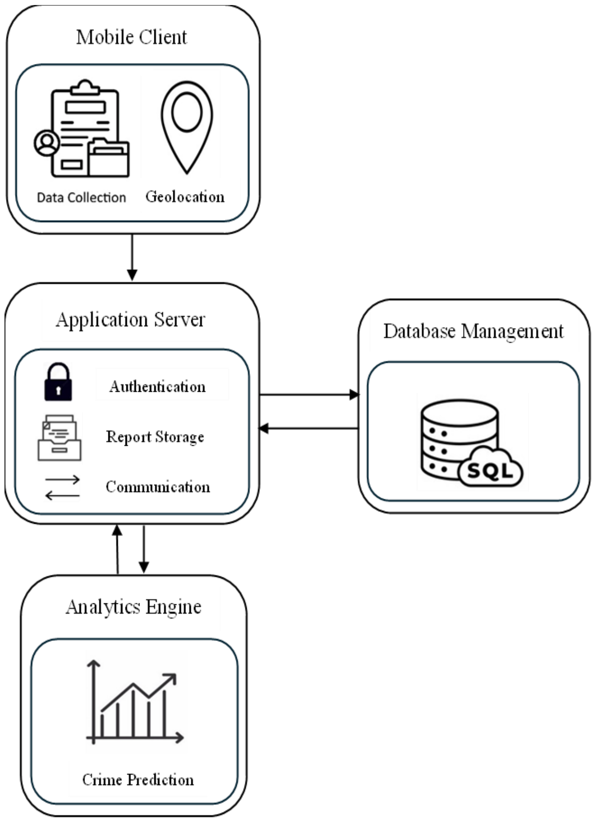

Fig 1: The system architecture

3.3 Naïve Bayes Algorithm

Naïve Bayes classifier was chosen because it is a computationally efficient algorithm that can work with small sets of data. The information consisted of textual descriptions, timestamps, and location properties. Preprocessing included tokenization of text descriptions, encoding categorical data and division of data into training (80%) and testing (20%) groups. The model training (Fig 2) was carried out by estimating the conditional probability of every type of crime after which the model was evaluated based on accuracy, precision, recall and F1-score.

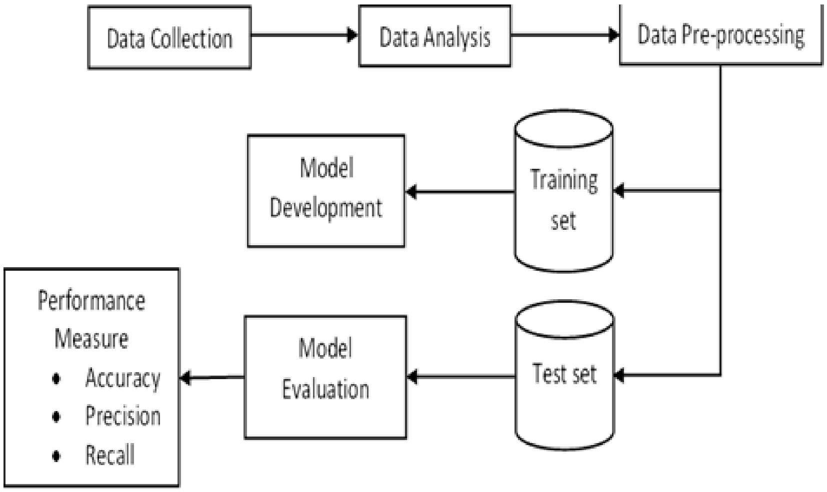

Fig 2: Naïve Bayes Pipeline & Data Processing Flow

3.4 Time-Series Forecasting Procedure

To enable true time-series forecasting, a synthetic multi-year Nigerian crime dataset spanning 2014–2024 was generated using a mixed methods approach. Data was collected through surveys, research, interviews, and documented reports. Exponential smoothing was used to estimate future crime trends. The future trends of crime were estimated by exponential smoothing. This method was selected due to its ability to work well in the environments where there is not much past data. The numbers of crimes were rolled up on monthly basis and model parameters were optimized through repeated trials to ensure the minimization of errors through evaluation measures like Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

3.5 Usability and System Performance Evaluation

Usability testing involved 70 participants, including commuters and security personnel. Participants performed reporting tasks, and their experiences were evaluated using questionnaire. System performance metrics processing latency, server response time, and dashboard load time were measured using server logs.

4. SYSTEM ARCHITECTURE AND UML DIAGRAMS

This section presents the system's architecture and UML models to illustrate structural and behavioral components. The diagrams reflect the interactions between mobile users, backend services, analytics modules, and administrative stakeholders.

4.1 System Architecture

The system follows a modular architecture (Fig 3) that integrates mobile data collection, backend application processing, and dashboard displays on web browsers.

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, Delaware, USA
Volume 13– No. 2, February 2026 – www.ijais.org

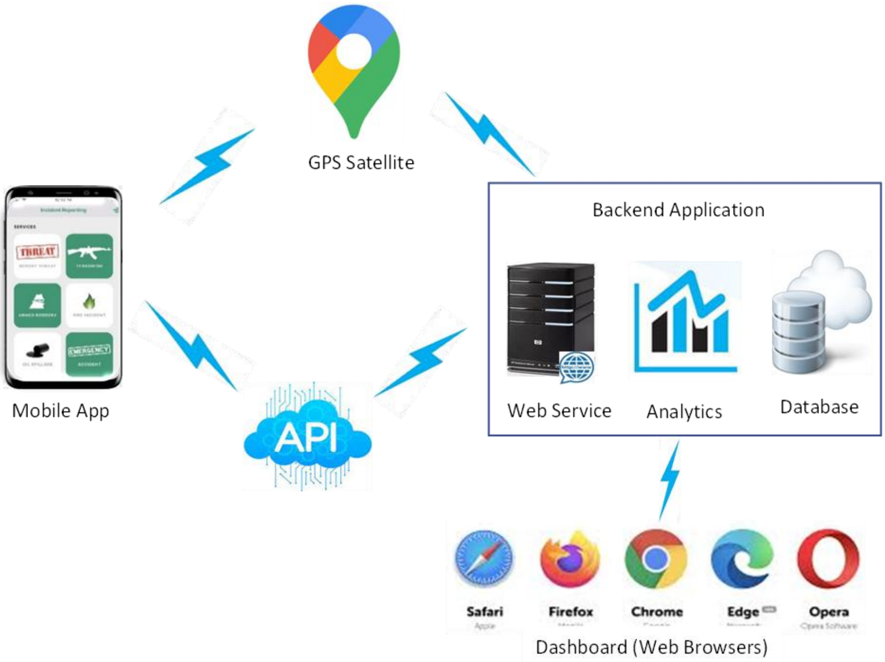

Fig 3: The workflow connection of the mobile app

4.2 Use-Case Diagram

The use-case diagram in Fig 4 illustrates primary interactions between system actors (mobile user, admin, and GPS) and respective activities.

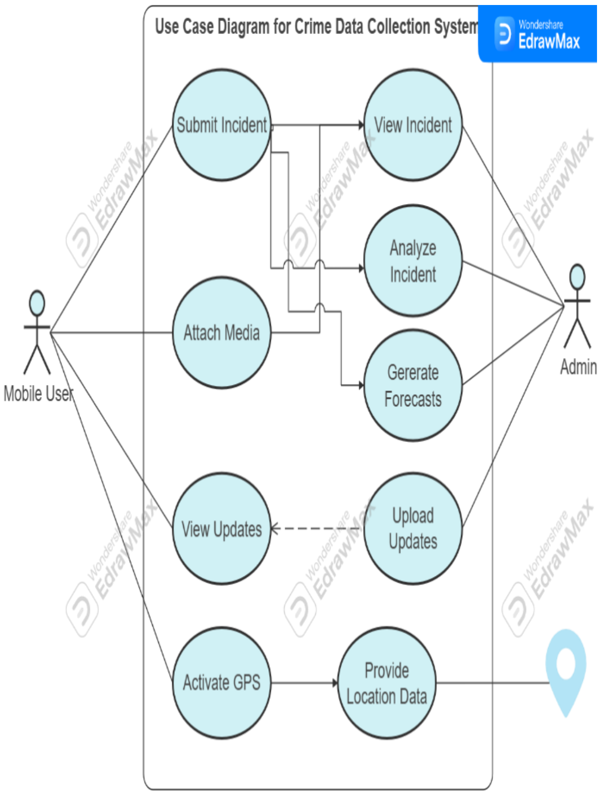

Fig 4: The use-case diagram of the mobile system

4.3 System Component Roles

System Component Roles describe the specific functions, responsibilities, and interactions of each major module within the system architecture. Table 1 outline what each component does and how it contributes to the overall workflow, this section clarifies how data flows through the platform; from mobile data capture to server processing, machine-learning analysis, and forecasting.

Table 1: System Component and Roles

Components	Roles
Mobile App	Collects incident reports, GPS, and media.
API Server	Receives requests, validates input, routes operations to backend services.
Database	Stores reports, classifications, and forecast outputs.
ML Classifier	Categorizes incidents automatically.
Forecasting Module	Generates monthly crime predictions.
Admin Dashboard	Displays analytics, visualizations, and system insights for authorities.

5. RESULTS AND DISCUSIONS

This section presents key findings from the deployed system, focusing on the performance of the ML classifier, the accuracy of the forecasting model, system usability outcomes, and overall operational efficiency.

5.1 Android Application Interface

Users can submit reports with GPS data and media attachments. Categories include robbery, kidnapping, accidents, fire incident and oil spillage. The app sends confirmation and allows users to track report status. Fig. 5 displays the user interface platform of the Android mobile app.

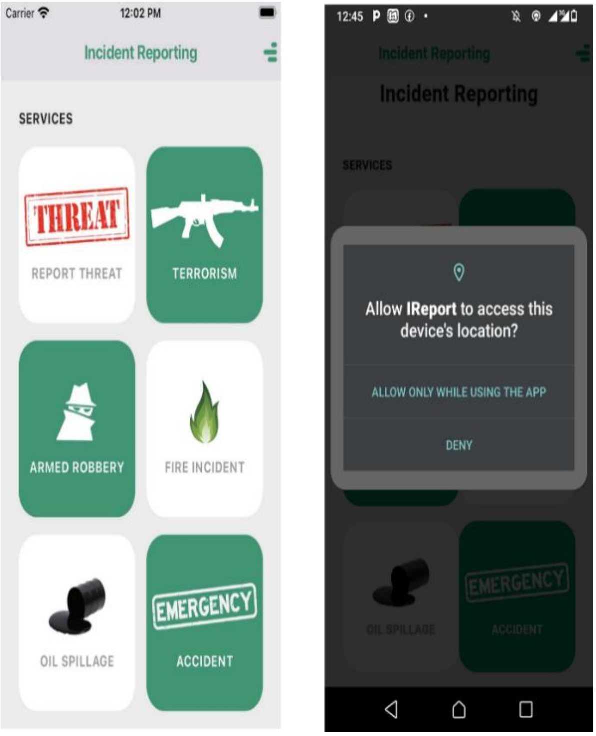

Fig 5: The user interface of the android app

5.2 Naïve Bayes Classification Results

The Naïve Bayes classifier was evaluated using a dataset of 2,048 incident records to assess its effectiveness in identifying high-risk zones (Table 2) and supporting response prioritization. Performance evaluation was conducted using standard classification metrics as illustrated in Fig. 6.

Classification Metrics:

- Accuracy: (892 + 924) / 2048 = 88.6%
- Precision: 892 / (892 + 106) = 87.4%
- Recall: 892 / (892 + 126) = 86.9%
- F1 Score ≈ 87.1%

Table 2 High-risk Zones Based on Reported Incidents (n=2,048)

Location Zones	Count of Reported Zones
East-West Road	812
Port Harcourt- Aba Road	476
Port Harcourt-Owerri Road	392
Other Identified Zones	368
Total	2,048

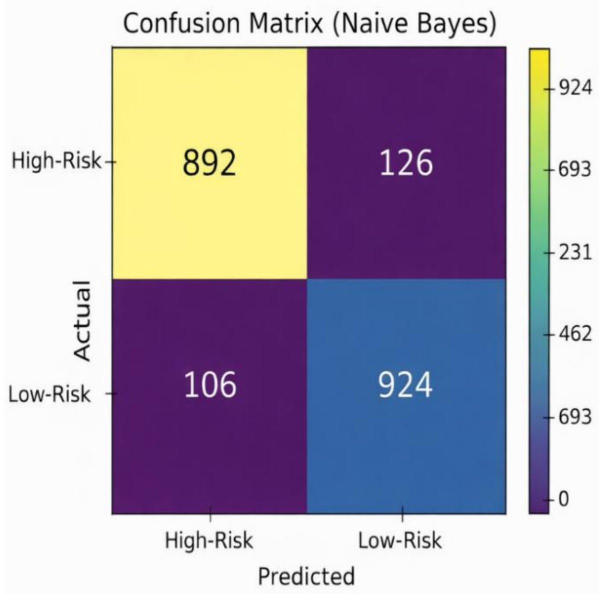

Fig 6: Confusion matrix (Naïve Bayes)

5.3 Time-series Performance

The Exponential smoothing was also used to stabilize the signals and also to make one step ahead predictions on the number of incidents per day. Fig 7 chart report indicates the number of crime incidences per day along with exponentially smoothed series (α=0.3). The raw data range is between 1 to 3 incidences in a day with most days having one incidence. The smooth curve levels off at an average of slightly over 1.0 incident per day and it is practically averting volatility in the short term. Low forecast errors (MAE ≈ 0.12, RMSE ≈ 0.29) are also evidence-based that exponential smoothing is an accurate and stable representation of the underlying trend in the crime in the annex dataset. The relative deviations between the predicted and observed counts are also modest since the MAPE of around 8.7% indicates relative deviation.

The chart in Fig 8 shows the relationship between the observed daily incident counts and the corresponding one-step-ahead predictions generated by the exponential smoothing model. The data points cluster closely around the central prediction band, indicating strong agreement between actual and predicted values. Observed incident counts range from 1 to 3, while predicted values are concentrated between approximately 1.0 and 1.7, reflecting the smoothing effect of the model.

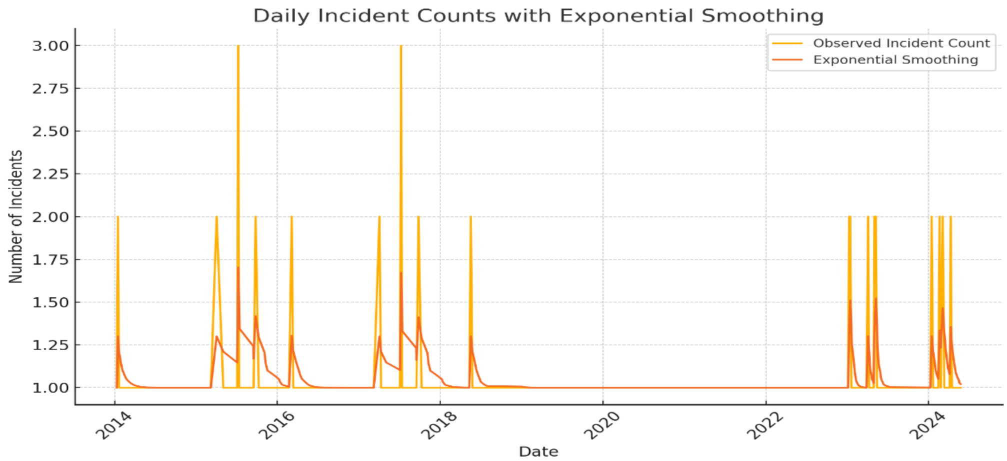

Fig 7: Time-series forecasting graph

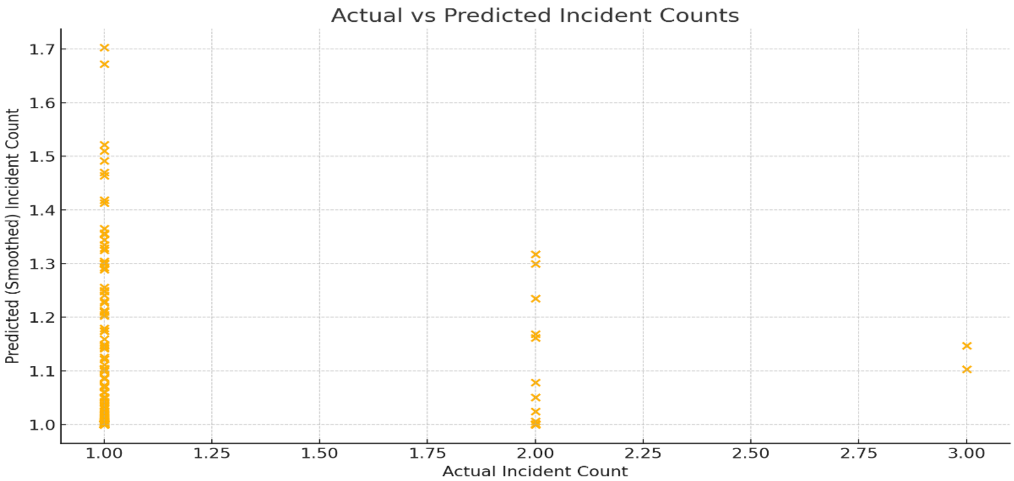

Fig 8: Observed daily incident count graph

5.4 Usability Testing

Feedback from 103 commuters and 27 officers revealed high satisfaction with the user interface, response time, and ease of use, presented in Table 3. Most users found the reporting process intuitive and appreciated real-time updates.

Table 3. User Satisfaction Survey Summary

Survey Theme	Key User Feedback	Response Frequency (%)
Ease of navigation	Most users reported successful navigation	96
GPS accuracy	GPS auto-location was found accurate	92
System responses	Satisfied with report submission speed and confirmation messages	90
Reliability of the system	Users felt confident that the app could assist in real emergency reporting.	88
Suggestions for improvement	Suggestions included offline functionality, and voice reporting	74

6. CONCLUSION

The Android-based incident reporting system significantly contributes to public safety by facilitating quick communication, precise geolocation, and predictive analytics through machine learning and time-series forecasting. By enabling users to report incidents promptly and accurately, the application enhances law enforcement responsiveness and supports data-driven decision-making. The integration of GPS, Naïve Bayes classifier and exponential smoothing further enables the identification of high-risk zones, promoting proactive intervention, and provide a quantitative forecast baseline for future incident trends by modelling daily crime dynamics. The model achieved an overall accuracy of 88.6%, indicating strong classification capability. Precision and recall values of 87.4% and 86.9%, respectively, demonstrating the model's reliability in correctly identifying high-risk incidents while minimizing false alarms. The resulting F1-score of 87.1% confirms a balanced trade-off between precision and recall. The system thus addresses critical gaps in conventional reporting mechanisms, offering a scalable, user-friendly, and technologically solution for enhancing public security.

Recommendations:

1. Integrate the app with national emergency systems.
2. Add features like facial recognition and vehicle plate scanning for enhanced verification.

3. Explore the use of deep learning models to analyze text and predict crime.

7. REFERENCES

[1] R. A. Okunola, "Commuters' Fear of Crime and the Road Transport Industry: A Study of Ibadan-Sokoto Transport Corridor," The Nigerian Journal of Sociology and Anthropology, vol. 7, no. 1, pp. 98–120, 2009, doi: 10.36108/njsa/9002/70(0120).

[2] V. Peccavi, (2019, May 12). Crime on Nigeria's Highways. Vox Peccavi. [Online]. Available: https://peccaviconsulting.wordpress.com/2019/05/12/crime-on-nigerias-highways/ [Accessed: Dec. 08, 2022].

[3] A. B. Sakpere, A. V. D. M. Kayem, and T. Ndlovu, "A Usable and Secure Crime Reporting System for Technology Resource Constrained Context," Proceedings - IEEE 29th International Conference on Advanced Information Networking and Applications Workshops, WAINA 2015, Gwangju, South Korea, Mar. 24-27, 2015, pp. 424–429, doi: 10.1109/WAINA.2015.97.

[4] B. Orija, (2016, Dec. 9) Firm launches mobile app for national security, crime reporting. The Guardian Nigeria News – Nigeria and World News. [Online]. Available: https://guardian.ng/technology/firm-launches-mobile-app-for-national-security-crime-reporting/ [Accessed: Apr. 11, 2023].

[5] M. M. Mohanan, D. Loganathan, and M. Nisha, "C App- A Mobile App for Crime Report," IJCSN-International Journal of Computer Science and Network, vol. 8, no. 3, pp. 294–299, 2019, [Online]. Available: www.IJCSN.org

[6] N. Mathonsi, and T. van Zyl, "Hybrid exponential smoothing and LSTM models for multivariate time-series forecasting." South African Journal of Data Science, vol.5, no. 2, pp. 73–89, Dec. 2021.

[7] U. S. Usman, "THE SOCIO-ECONOMIC CAUSES AND EFFECTS OF ARMED," International Journal of Economics, Commerce and Management, vol. 3, no. 5, pp. 886–897, May, 2015.

[8] O. Ologunagbe, (2022, Dec. 31) Road accidents in Nigeria and their common causes in 2022. Business Day [Online]. Available: https://businessday.ng/news/article/road-accidents-in-nigeria-and-their-common-causes-in-2022/ [Accessed: Jan. 22, 2023].

[9] T. Mahalakshmi, E. S. Mary, M. Stella, and A. Theepak, "Crime Reporting and Management Using Mobile Application," International Journal of Advanced Research Trends in Engineering and Technology, vol. 7, no. 8, pp. 1–5, 2020.

[10] T. Alameri, A. H. Alhilali, N. S. Ali, and J. K. Mezaal, "Crime reporting and police controlling: Mobile and web-based approach for information-sharing in Iraq," Journal of Intelligent Systems, vol. 31, no. 1, pp. 726–738, 2022, doi: 10.1515/jisys-2022-0034.

[11] S. D. Abdullahi and A. Aishetu, "Intelligence Gathering on Kidnapping Incidence along Abuja – Kaduna Express Road of Nigeria Using Geospatial Technologies," African Scholar Publications & Research International, vol. 21, no. 2, pp. 65–74, 2021.

[12] A. Afolayan, S. M. Easa, O. S. Abiola, and F. M. Alayaki, "GIS-Based Spatial Analysis of Accident Hotspots: A Nigerian," Infrastructures (Basel), Aug. 2022, doi: 10.3390/infrastructures7080103.

[13] C. Cubukcu and M. Gurun, "A Mobile Application Development that Shows the Crime Rate in Istanbul," In Contemporary Management Practices Xi Smart Specialisation in the Decade of Business Connectivity to Locomotions, Jun. 2021. [Online]. Available: https://www.researchgate.net/publication/360843235 [Accessed: Jun. 23, 2025].

[14] A. Tamir, E. Watson, B. Willett, Q. Hasan, and J. S. Yuan, "Crime Prediction and Forecasting using Machine Learning Algorithms," Article in International Journal of Computer Science and Information Technology, pp. 26–33, Mar. 2021, [Online]. Available: https://www.researchgate.net/publication/355872171

[15] S. Raksha, R. Shreyaa, and M. S. Pranav, "Survey on Blockchain-based Anonymous Tip-off System," International Journal of Innovative Science and Research Technology, vol. 7, no. 4, pp. 909–913, Apr. 2022.

[16] A. Abrami, A. Aravkin, and Y. Kim, "Time Series Using Exponential Smoothing Cells." arXiv preprint, 2017, arXiv:1706.02829v4. [Online] Available: http://arxiv.org/pdf/1706.02829v4.

[17] M. Bernardi, G. Rossi, and L. Caruso, "Evaluating Exponential Smoothing Methods for Crime Trend Forecasting in Low-data Environments." Applied Forecasting Review, vol. 9, no. 1, pp. 21–34, 2024.