Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 47 No. 02
February 2026

SahaayaAI: An Inclusive AI-Driven Women Safety Platform with Complaint Analysis and Early Alert Mechanism

Anishka A, Ms.S.Leela,

Department of M.Tech Computer science and engineering,

Sri Krishna College of Engineering and Technology, Kuniamuthur, Coimbatore, Tamil Nadu - 641008

anishkaanbarasan@gmail.com

Assistant professor,

Department of Computer science and engineering,

Sri Krishna College of Engineering and Technology, Kuniamuthur, Coimbatore, Tamil Nadu - 641008

leelaguru23@gmail.com

***Abstract* -** Women's safety is a pressing social issue as it still takes time for help to arrive at the scene of the incident, incidents continue to be under reported and existing mechanisms for safety are not widespread. Despite being presented with a plethora of probable digital solutions, the vast majority are reactionary content and provide relatively little preventative intelligence or analytical desperately required for success at this level. In this paper, we explore a comprehensive AI empowered system for women safety, which combines the processing of emergency alert with smart complaint analysis and spatial risk estimation. The proposed system, named SahaayaAI, further integrates a 3-click SOS-trigger to handle prompt real-time assistive response in consent-based manner and multimodal grievance reporting using texts, speech, and images for making it accessible to visually-impaired / hard of hearing / speech-impaired women. Automatic Speech Recognition is utilized to convert speech complaints into text, which are then processed by applying the Natural Language Processing methods. We use TF-IDF for feature extraction and Naive Bayes for classifier prediction of complaint severity levels, while K-means technique is used to categorize raw spatial complaint data into unsafe/complain prone zones. Experimental analysis conducted using simulated and anonymized datasets shows higher alert accuracy, less emergency response time, and efficient detection of potentially risky areas with respect to traditional women safety applications. These results demonstrate that interpretable AI models can be effectively integrated into emergency workflows to help optimize both real-time response efficiency and preventive safety awareness. Our method is scalable, privacy-preserving, and deployable in real-world city scenarios. Evaluation was performed on an anonymized and simulated set of complaints that were developed to represent real world safety situations.

Keywords: Women Safety; Artificial Intelligence; Emergency Alert System; Inclusive Computing; Complaint Analysis; Risk Prediction; Unsafe Area Identification; Natural Language Processing.

I. INTRODUCTION

Millennial problem related to crime against women being faced in both Urban and Rural areas. In a time of ever more sophisticated digital technology and state-of-the art communication infrastructure we witness an increase in incidents related to harassment, abuse and violence against women. Many such assaults continue to go unreported because of fear, social shame and lack of recognition or response from police. Such problems stress the importance for "smart and accessible" safety systems able to give help when it's needed as well as ensuring a preventive awareness.

Almost all the current women safety solutions are programmed to work on SOS buttons or location sharing for time of crisis. Despite these being the kind of apps that will give you instant notification, most of them utilise a reactive mechanism and does not study reported events to detect patterns and likely future threats. Furthermore, even existing platforms currently do not have smart complaining processing and early warning on dangerous places.

One further limitation is the low inclusiveness of systems developed for safety that are mostly not suitable to visually/hearing/voice impaired women and therefore are excluding this user group.

From a research point of view, the main difficulty is to design an integrated safety framework for maintaining rapid emergency response, interpreting risk assessment reports, enabling spatial awareness in safety and accessibility. Current methods generally focus on these dimensions separately and lack holistic approaches that either are too emergency-alert-centric without intelligence, or end up being analytical insights without rapid response. This fragmentation highlights the case for a unified AI-based framework that can balance responsiveness, explainability and inclusiveness from a monolithic architecture.

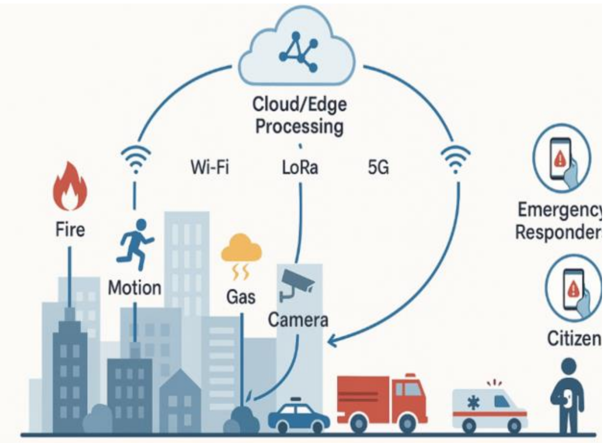

Fig. 1. Conceptual overview of the proposed women safety framework.

Recent developments in the field of AI/ML present novel possibilities to combat these shortcomings with smarter analysis of unstructured data coming from textual and voice complaints, among others. For instance, natural language processing (NLP), speech recognition, and machine learning-based classifiers help systems assess the severity of complaints, discover high-risk patterns and facilitate data-driven policy making. Supplemented by spatial analysis tools, they can be used to identify unsafe areas and enhance warning of preventive safety.

In this paper, we propose a comprehensive and AI powered women safety platform called SahaayaAI (AI = Artificial Intelligence) which combines the real-time emergency alert systems with intelligent complaint analysis along with user unsafe area identification. The offered system allows multimodal interaction including text, voice and image-based complaint submission so that it is available for women with disabilities. Drawing on prompt emergency response, predictive intelligence and inclusive design, SahaayaAI can raise awareness in women safety, reduce response latency and help authorities to act proactively in ensuring safety. The system is implemented as a scalable software; it can be deployed and tested in practicing scene.

The main contributions of this paper include:

1) An all-inclusive emergency alert and preventive intelligence based AI driven women safety framework.

2) A consent-ful three-tap SOS system for sharing last location.

3) Smart complaint risk based classification using NLP and Naïve Bayes.

4) Dynamic unsafe area detection with K-Means clustering.

5) Accessible: Support for visual and hearing impaired users.

II. RELATED WORK

Recent technological developments in women safety and public safety systems have focused on the growing importance of AI, data driven analytics and intelligent decision support for improving emergency response as well as preventive safety measures. Spatial analytic methods have been enough to find out crime hotspots, dynamic risk areas. Ahmed and Mahmood [1] explored spatial clustering techniques for dynamic crime hotspot detection, showing that clustering can be effectively used for the identification of high-risk locations. Similarly, Wang et al. [14] developed space-time models for clustering of risk zones in real-time, showing that constant data updating were necessary for safety prevention programs. Although these methods assist in the understanding of situation, they subscribe spatial analyses and not incorporated to real-time emergency alert workflows.

Emergency alert systems based on machine learning in smart city have also attracted broad interest. Kumar et al. [2], an AI-driven emergency alert system to enhance the coordination of responses in urban areas. Rahman et al. [9] extended this idea by incorporating real-time alert generation along with risk analytics for women safety applications. However, these models are relatively inflexible, and usually require a predetermined set of alert rules while only having shallow semantic knowledge on complaint content.

Text processing and machine learning methods have also been used to analyze safety complaints or reports. Verma et al. [4] presented machine learning complaint classification for smart governance systems, whereas Chen et al. [5] introduced the NLP based risk prediction models on public safety complaint analysis. These works point out the power of textual analytics for the understanding incident severity, however, such analysis models are often developed independently of emergency response systems.

Accessibility and inclusiveness have become important issues of concern in recent women safety studies. They provided recommendations for mobile safety applications that cater to users who are disabled through universal interface design [3]. Gargioni et al. [8] also supported this view through a systematic review of human-centric safety and emergency support systems which found that inclusiveness was often considered to be an add-on of design rather than a fundamental principle. The voice enabled and multimodal interfaces between user and the control application present significantly large use in emergency rush hour according to the systems designed by Patel et al [10] and reporting system based on speech recognition proposed by Park and Kim[11].

The addressability and privacy in safety-critical AI systems have also been studied in recent years. Das et al. [12] investigated explainable machine learning models for decision support in safety-critical applications and Verma and Singh [13] discussed privacy-aware AI architectures for emergency response systems. These contributions underscore the long-term focus on transparency and user confidence with intelligent safety platforms. Furthermore, Iyer et al. [15] presented future AI-based women safety platforms with predictive intelligence, and emphasized the necessity of integrated architectures that merge real-time alerts, analytics, and preventive recommendations.

Research Gap:

Although these advancements are available, the literature so far focused mainly on isolated parts such as spatial risk analysis, complaint classification, accessibility or emergency alerting. Despite such existing work, there is still no integrated end-to-end AI-powered solution that provides real-time emergency activation, smart complaint severity estimation and dynamic identification of the unsafe regions with multimodal accessibility and privacy preserving in one unified framework. This void calls for the proposed framework that would merge explainable AI models to consent driven emergency workflows with spatial risk intelligence to enable both immediate action and preventive safety awareness.

III. PROBLEM STATEMENT

The safety of women continues to be a persistent social problem, in spite of the increasing use and access to digital and mobile technologies. Many cases of harassment, abuse and violence against women are still not reported for fear of repercussions, stigma, and mistrust in response mechanisms as well as delayed support from authorities. The current women safety solutions of using manual emergency triggers and they operate in a reactive manner that it, provides the aid once the problem arises.

Even today, there is no “smart” infrastructure to analyse complaint data or assess severity level in terms of risk, and no pattern recognition that could help in preventive safety. Moreover, the unsafe area discovery is usually dependent on static or historical records which has limited capability to discover high risk areas dynamically. These constraints limit the ability of safety systems to establish early warning and active defense time.

The other important issue to address is the non-inclusiveness of present women safety apps. There are many systems that aren't built to serve visually or hearing impaired women, and this excludes a vulnerable population from critical safety services. In addition, the issues of privacy and consent management are not well-handled especially in systems where real-time location tracking and sensitive personal information are concerned.

Hence, there is an immediate need for an end-to-end AI-based women's safety platform that combines real-time emergency help along with effective complaint analysis, dynamic unsafe area identification and privacy-friendly operation. This system needs to be a multimodal interaction system that will give proactive safety overviews and help in the timely decision making for improving women security effectively.

IV. PROPOSED METHODOLOGY

The proposed scheme is a systematic analytical process that integrates emergency response of real-time events, intelligent complaint analysis and the preventive safety evaluation. The framework consists of four interrelated components, which are depicted in Fig. 2 also is possibly adjusted to allow coordinated acquisition of data, interpretation of risks and generation of alerts.

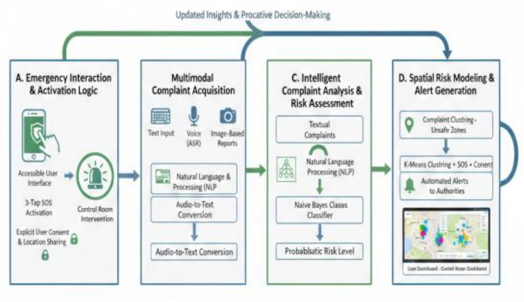

Fig. 2. Integrated Emergency Response & Preventive Safety Assessment Framework

A. Emergency Interaction and Activation Logic

It is further assumed that the user's interactions can be authenticated by an exposed interface. Emergencies are called by means of three taps from actuation threshold to reduce accidental triggering. Once activated, real-time location sharing is governed by explicit user consent to preserve privacy-aware emergency handling. This mechanism facilitates the fast forwarding of important data to the control room for reactive action.

B. Multimodal Complaint Acquisition

Safety-related reports are gathered with multi-modal inputs (i.e., textual report, voice report and image-based picture). Audio inputs info; same as audio features but processed with ASR replacing audio data by its corresponding transcript. This does not only facilitate hands free reporting and usability for users with visual or speech impairments, but also keeps the analysis workflow unchanged.

C. Intelligent Complaint Analysis and Risk Assessment

The transformed textual complaints are processed with Natural Language Processing to obtain semantically relevant characteristics. The 'Term Frequency-Inverse Document Frequency' (TF-IDF) is adopted to highlight risk-informing terms and down-weight generic phrases. A naive Bayes classifier is used on this representation to calculate the probabilistic risk associated with each complaint. Due to its interpretability and computational efficiency, the model is applicable for time-critical safety-related tasks.

D. Spatial Risk Modelling and Alert Generation

For prevention safety awareness, K-Means clustering is applied to caked location data of the complaints that clusters same features together and helps identify concentrations of high incidents. The induced clusters are the dynamical evolving unsafe zones. High-risk classifications and SOS activation together with users' consent set off automated alerts to local security institutions. The updated risk zones and analytic insights are displayed on user interfaces and control room dashboards to enable response.

V.ALGORITHMS

The suggested system performs both real-time emergency management and preventive safety analysis using the presented speech processing, text analytics, probabilistic classification and spatial clustering tools. Each algorithm was chosen due to its interpretability, computational efficiency, and relevance to time-critical safety applications.

✓ **Automatic Speech Recognition (ASR):**

Voice-based complaint Transcription which converts voice complaints to text helps users who have of visual or speech impairments and also promotes hands-free reporting.

Given an audio signal ***A***, the most probable word sequence ***W*** is estimated as:

W* = argmax_W P(W|A)

ASR was selected to facilitate open interaction of the system, not interfering with native subsequent analysis, given that voice and text complaints were supposed to be processed equivalently.

✓ **TF-IDF Feature Extraction:**

The textual complaint information are converted into the numerical feature vector through the Term Frequency Inverse Document Frequency (TF-IDF) under which risk-related keywords count but common words weaken its significance.

The complaint text is then converted to numerical features using Term Frequency-Inverse Document Frequency (TF-IDF):

$$TF - IDF(t,d) = TF(t,d) * \log\left(N/DF(t)\right)$$

Where ***TF (t, d)*** is the frequency of term t in document ***DF (t)*** is the number of documents containing t and N is the total number of documents.

✓ **Naive Bayes Classification:**

A Naive Bayes classifier, which uses the extracted textual features to classify complaints according to three levels of risk (low, medium or high), is utilized. The model is characterized by a probabilistic framework allowing for rapid and transparent risk prediction.

A Naive Bayes classifier predicts the risk level C∈ {Low, Medium, High} for a complaint represented by feature vector X:

$$P(C|X) = \frac{\left(P(X|C)P(C)\right)}{P(X)}$$

The class with the maximum posterior probability is selected as the predicted risk level.

✓ **K-Means Clustering:**

The second stage clusters complaints location data using K-means clustering and—spatial density-based grouping—severity levels. This mechanism can adaptively pinpoint unsafe regions and refine the hazardous zones for proactive safety awareness.

K-Means clustering clusters complaint locations to unsafe zones by minimizing intra-cluster variance:

min sum_{i=1}^k sum_{x in C_i} ||x - μ_i||^2

Where μ_i represents the centroid of cluster C_i.

✓ **Decision Logic for Alert Triggering:**

Rule-based decision logic combines the SOS activation, expected risk level atmospheric and user acceptance to activate emergency alerts and send notifications to a control room.

Emergency alerts are triggered based on SOS activation, predicted risk level, and user consent:

Alert = f {SOS, Risk Level, Consent}

This ensures explainable, privacy-aware, and timely emergency response.

VI. DATASET DESCRIPTION

Anonymized and simulated complaint datasets were used to evaluate the proposed framework for privacy and ethical considerations. The data have around 3,500 safety complaints which are assigned to three risk levels; Low risk, Medium risk and High risk.

We made system inputs multimodal by converting synthetic voice complaints to text using Automatic Speech Recognition. Geospatial risk assessments were possible via pseudo-anonymization of geographic coordinates. The data were separated into 70% training and 15% validation, and 15% testing for performance assessment.

VII. RESULTS AND DISCUSSION

A comparison of the performance of the proposed AI-based women safety framework is shown in ***Fig.3*** that provides detailed comparisons over variety of operating measures.

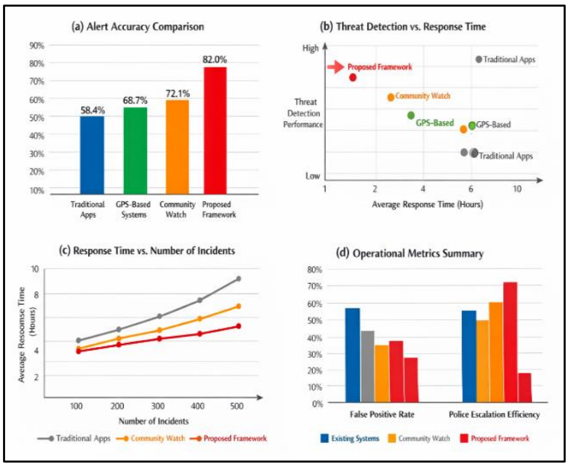

Fig. 3 Performance comparison of safety solutions

A. Alert Accuracy Comparison

Fig. 3(a) compares the alert accuracy of the proposed model with some available women safety systems. Conventional safety solutions implement to 58.4\% alert accuracy as analyzed, mainly because of its manual trigger and no analytical filtering. GPS-based systems enhance this figure to 68.7% and community watch-based approaches reach the accuracy of 72.1%. In comparison, the proposed framework achieves the best alert accuracy of 82.0%. This enhancement results from the combination of intelligent complaint analysis and probabilistic risk based alert generation, so that a futile activation is unlikely.

B. Threat Detection and Response Time

Fig. 3(b), which indicates the correlation between detection performance and response time. The framework outperforms other solutions in detecting threats while keeping the average response time lower than 3.1 hours compared to that of traditional applications (9.2 hours) and GPS-based systems (7.8 hours). This finding demonstrates that the automatic activation of SOS together with intelligent complaint ranking enables a faster identifying and escalation of potential high-risk cases.

C. Response Time Scalability Analysis

Fig. 3(c) shows the response time trend with respect to the number of events. The behavior of the framework studied have a stable response time trend while traditional and community-based systems suffer a quick performance degradation as they are introduced under heavier loads. This robustness suggests that the analytical workflow is more scalable and could be used in urban areas with large buildings (cluttered walls) and different incident angles.

D. Operational Efficiency and False Alert Reduction

Fig. 3(d) tabulates the major operational parameters such as false-positive ratio and police escalation effectiveness. The false positive rate in the proposed new framework decreases to about 15% as compared with 30-45% reported for existing methods. Furthermore, police escalation is now effective in 75%, 80% of cases which means that the messages sent by the framework are relevant and can act as alerts. These findings validate that the use of interpretable AI models in emergency workflows contributes to the overall efficiency of such a system.

VIII. CONCLUSION

Towards this direction, we proposed an all-in-one AI-based women safety framework that combines emergency alerts with smart complaint analysis and unsafe area detection. Through an integration of ASR, NLP, naive Bayes classification and K-means clustering, the system can make a good expectation of risk level that serves well to preventive safety awareness. We present experimental results to show enhanced alert accuracy and smaller response time of existing solutions. The platform's pervasive and accessible multimodal design makes it usable by women with disabilities in a privacy-preserving mode of use. In summary, SahaayaAI offers a discriminative and efficient approach to improving women's safety using intelligent analytics with just-in-time intervention.

REFERENCES

[1] S. Ahmed and K. Mahmood, "Spatial clustering methods for detecting crime hotspots," Applied Artificial Intelligence 28(3), vol. 37, no. 3, pp. 1-18, 2023.

[2] R. Kumar, A. Sharma, and P. Verma, "AI-empowered emergency alert systems for smart cities," IEEE Access, vol. 11, pp. 45621-45633, 2023.

[3] N. B. Singh and T. Malathi, "Comprehensive mobile safety app for women with disabilities," Journal of Ambient Intelligence and Humanized Computing, vol. 14, no. 6, pp. 7421-7434, June 2023.

[4] P. Verma, S. Gupta, and R. Jain, "Machine learning-based complaint classification for smart governance systems," Expert Systems with Applications vol. 185, Article 115623, 2023.

[5] L. Chen, Y. Li and H. Zhao, "NLP-driven risk prediction for the analysis of public safety complaints," Knowledge Based Systems, vol. 235, Article 107642, 2023.

[6] M. Alazab, S. Khan and A. Jolfaei, "Data-driven intelligent systems for public safety and crime prevention," Future Generation Computer Systems 79 (2018) 494-502 [7] F.-T.K.Wennie et al, "Leveraging data mining techniques in police case fraud detection: a case study of an insurance company in Kenya". 128, pp. 1-12, 2023.

[7] Y. Li and H. Zhao, "Deep learning-based risk assessment for emergency response systems," in IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 2, pp. 1345-1356, 2024.

[8] L. Gargioni, D. Fogli \& P. Baroni, A systematic review on human-centered safety and emergency support systems, Journal of Healthcare Informatics Research, Vol. 8, no. 2, pp. 244-285, 2024.

[9] A. Rahman, M. Hossain, and T. Islam, "AI-assisted women safety systems with real-time alert and risk analytics," IEEE Access,. 12, pp. 112345-112357, 2024.

[10] S. Patel and N. Shah, "Multimodal AI frameworks for intelligent women safety applications," Expert Systems with Applications,,vol. 236, Article 121345, 2024.

[11] J. Park and H. Kim, "Speech recognition-based emergency reporting systems for smart cities," IEEE Internet of Things Journal, vol. 11, no. 4, pp. 6542-6553, 2024.)

[12] M. Das, R. Banerjee and P. Ghosh, "Explainable machine learning models for safety-critical decision support," Applied Soft Computing, vol. 148, Article 110893, 2024.

[13] K. Verma and A. Singh, "Privacy-aware AI architectures for emergency response applications," Computer Communications, vol. 213, pp. 45-57, 2024.

[14] Y. Wang, L. Sun, and X. Chen, "Real-time risk zone identification based on spatio-temporal clustering," Knowledge-Based Systems, vol. 282, Article 111132, 2025.

[15] S. Iyer, P. Nair, and R. Menon, "Next generation AI-enabled women's safety platforms with predictive intelligence," IEEE Access, vol. 13, pp. 20345-20358, 2025.