

Harm Spot Detection using Artificial Intelligence & Machine Learning

Sariya Anjum¹, Rohan Gowda U², Spoorthi E S³, Srushti a M⁴, Dhanush N⁵

¹Asst Professor, Department of Computer Science & Engineering Mysuru Royal Institute of Technology, Mandya, Karnataka, India.

^{2,3,4,5}Student, Department of Computer Science & Engineering, Mysuru Royal Institute of Technology, Mandya, Karnataka, India.

Abstract- Crime is a major concern in urban and semi-urban areas, and identifying locations where crime risk may be higher can support better planning and resource allocation. This project presents an Artificial Intelligence and Machine Learning based Crime Hotspots Detection system that uses location-based crime-related features to estimate the risk level of a selected geographical point. The system combines a web-based map interface with a machine learning backend. A user selects a location on an interactive map, the frontend sends latitude and longitude to the backend API, and the backend extracts relevant features and applies a trained XGBoost model to predict the crime-risk category. The result is returned to the user as a risk score and a category such as Low, Medium, or High. The system is designed to provide an easy-to-understand visual representation of potential crime-risk areas and can support data-driven analysis and planning.

Keywords: Artificial Intelligence, Machine Learning, Crime Hotspots, XGBoost, FastAPI, React, Leaflet, Risk Prediction, Geospatial Data.

I. INTRODUCTION

Crime patterns can vary across geographical locations and time periods. Traditional analysis generally depends on historical records and manual examination, which can make it difficult to identify spatial patterns quickly. Recent developments in Artificial Intelligence, Machine Learning, and geospatial technologies provide methods for analysing historical crime-related data and estimating risk for selected locations. This project proposes a Crime Hotspots Detection system using AI and ML. An interactive map allows the user to select a location.

The selected latitude and longitude are sent through an API to the backend, where features are prepared and passed to a trained XGBoost model. The prediction is then displayed through the web interface.

Objectives

1. To develop an AI/ML-based system for detecting potential crime-risk hotspots.
2. To use geographical coordinates and relevant features for risk prediction.
3. To integrate a machine learning model with a web-based interactive map.
4. To classify predicted risk into understandable categories such as Low, Medium, and High.
5. To provide a simple interface for location-based crime-risk analysis.

II. METHODOLOGY

The proposed system follows a pipeline in which a geographical location is selected through an interactive map and processed by a backend prediction service. The frontend collects latitude and longitude, while the backend validates the request, prepares the model features, and produces the risk prediction.

- **Input Collection:** The user selects a geographical location on the interactive map.
- **Data Validation:** The backend checks whether the received latitude and longitude are valid.
- **Feature Extraction:** Relevant features required by the trained model are prepared from the input and available data.
- **Prediction:** The trained XGBoost machine learning model predicts the crime-risk level.
- **Risk Classification:** The prediction is converted into a clear risk category such as Low, Medium, or High.

- API Response: FastAPI returns the prediction and risk information to the frontend.
- Visualization: React and Leaflet display the result on the map interface.
- User Interaction: The user can select another location and request a new prediction.

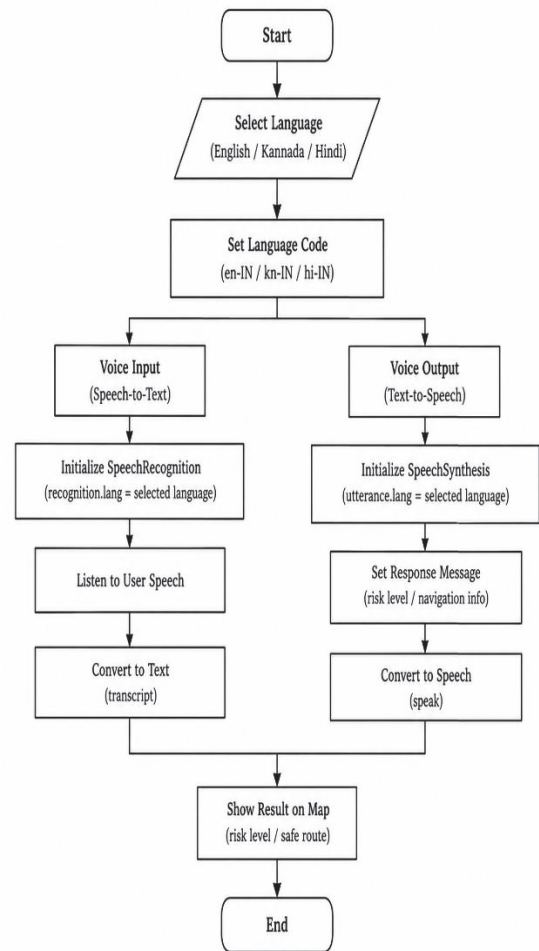


Figure 1: Flowchart

III. MODELING AND ANALYSIS

The proposed HarmSpot Detection using AI/ML system uses historical crime data to analyze crime patterns and predict the risk level of a selected location. The overall model combines data preprocessing, spatial analysis, feature extraction, machine learning, and visualization.

The processed data is then used to extract relevant features for machine learning. The XGBoost classification model is trained using these features to classify a location into Low, Medium, or High risk. The model learns patterns from historical crime occurrences and geographical and temporal information..

Table 1: System Modules and Their Functions

Module	Function
Map Interface	Allows the user to select a geographical location.
Input Validation	Checks latitude and longitude values.
Feature Extraction	Prepares the features required by the ML model.
Prediction API	Receives requests and returns prediction results.
XGBoost Model	Predicts crime-risk based on trained data.
Risk Classification	Maps the model output to Low/Medium/High.
Visualization	Displays the result using the interactive map.
Frontend	Provides the user interface using React.
Backend	Handles API logic and model integration using FastAPI.
Continuous Analysis	Allows repeated location selections and predictions.

For a new location, the system receives its geographical coordinates, extracts the required features, and passes them to the trained model. The model generates the corresponding risk prediction. The result is then displayed through an interactive map, allowing users to understand the estimated crime-risk level of the selected area.

The performance of the model can be analyzed using accuracy, precision, recall, F1-score, and confusion matrix. These measures help evaluate how effectively the model classifies different risk levels.

This integrated approach provides a systematic method for analyzing historical crime patterns and generating location-specific crime-risk predictions to support better safety awareness and decision-making.

IV. PROPOSED SYSTEM

The proposed Crime Hotspots Detection system combines an interactive map with a machine learning prediction service. The user selects a point on a Leaflet-based map. The frontend sends the latitude and longitude to the FastAPI backend through the prediction API. The backend validates the request and prepares the features required by the trained XGBoost model.

The XGBoost model processes the prepared features and generates a crime-risk prediction. The backend then returns the prediction to the React frontend. The result can be represented as a risk score and a category such as Low, Medium, or High. This architecture separates the user interface from the prediction logic and makes the system easier to maintain and extend.

The proposed HarmSpot Detection using AI/ML system uses historical crime data to analyze geographical and temporal crime patterns and estimate the risk level of a selected location.

The system uses DBSCAN to identify crime hotspots, KDE for crime-density analysis, and XGBoost to classify the selected location into Low, Medium, or High risk. The results are presented through an interactive map.

The system can also provide voice-based interaction in English, Kannada, and Hindi, depending on the implemented features.

V. EXISTING SYSTEM

The existing approach mainly depends on traditional crime reporting, static crime statistics, police records, and general online information. Users can obtain information about crime incidents or crime rates, but it may not provide a location-specific risk prediction for the exact area they want to visit.

In many existing systems, crime data is presented as reports, tables, or static maps. The user generally needs to interpret the information themselves. Such systems may not combine historical crime patterns, geographical information, and machine learning to estimate the risk of a selected location.

Conventional crime analysis often relies on historical crime records, reports, charts, and manual geographical analysis. Some systems may show previously recorded incidents but may not provide an integrated machine learning prediction workflow for a user-selected location.

Limitations commonly associated with such approaches include:

- Manual analysis can require more time when datasets become large.
- Static visualizations may not provide an interactive prediction for a new location.
- Separate tools may be required for data analysis, mapping, and prediction.
- Non-technical users may find raw crime datasets difficult to interpret.
- A unified frontend-to-ML prediction pipeline can simplify location-based analysis.

Table 2: Comparison Between Existing and Proposed System

Feature	Existing System	Proposed System
Risk prediction	Usually not provided for a selected location	Low, Medium, High risk prediction
Machine Learning	Limited/Not integrated	XGBoost classification
Map visualization	Static or basic maps	Interactive map visualization
Risk explanation	Limited	Risk level with supporting information
Multilingual voice	Generally unavailable	English, Kannada and Hindi support can be included
Purpose	Reporting and viewing crime information	Crime-pattern analysis and location-specific risk assessment

VI. RESULTS AND DISCUSSION

The developed system is intended to demonstrate end-to-end communication between the interactive map, backend API, and machine learning model. A user selects a location on the map, and the request is processed by the FastAPI service. After feature preparation, the trained XGBoost model generates a risk prediction, which is returned to the frontend.

The output is designed to be presented in a simple form so that users can understand the predicted risk without reading raw model values. The interface can show the selected location together with a risk score and a Low, Medium, or High category.

The project architecture supports testing of individual components, including API validation, model prediction, and frontend visualization. Model performance should be reported using the actual evaluation results obtained from the final dataset.

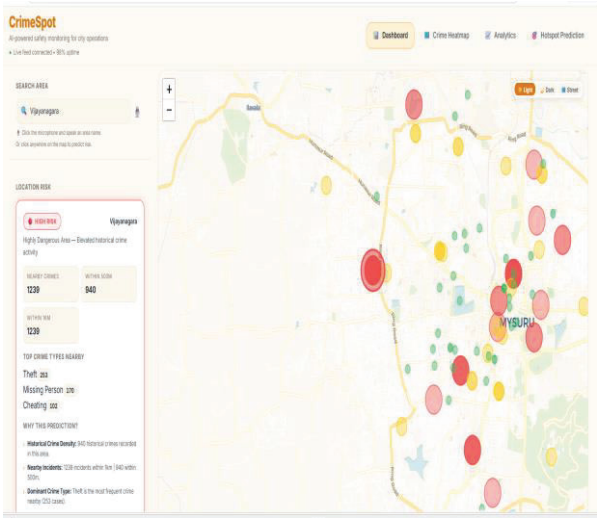


Figure 2: Risk Prediction



Figure 3: Crime Category Distribution

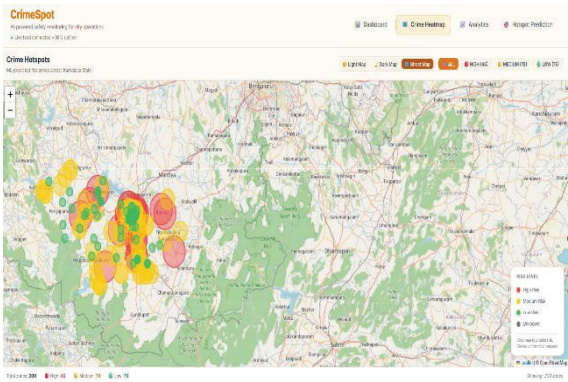


Figure 4: Crime HeatMap

CONCLUSION

The proposed Crime Hotspots Detection system demonstrates how Artificial Intelligence and Machine Learning can be integrated with geospatial visualization to support location-based crime-risk analysis. The system combines React and Leaflet for interactive location selection, FastAPI for backend API communication, feature extraction for model input preparation, and XGBoost for crime-risk prediction.

The system provides a clear workflow from selecting a point on the map to receiving a risk score and a risk category. This approach can help present complex model predictions in a simple and understandable manner. The modular architecture also allows individual components to be improved without changing the complete system.

Future enhancements may include larger and more diverse datasets, additional temporal and environmental features, improved model evaluation, real-time data integration, advanced geospatial visualizations, and deployment as a scalable web or mobile application.

REFERENCES

- [1] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.
- [2] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226–231, 1996.
- [3] X. Zhang, L. Liu, L. Xiao, and J. Ji, "Comparison of Machine Learning Algorithms for Predicting Crime Hotspots," IEEE Access, vol. 8, pp. 181302–181310, 2020.
- [4] "A Clustering Based Hotspot Identification Approach for Crime Prediction," Procedia Computer Science, vol. 167, pp. 1462–1470, 2020.
- [5] J. Alghamdi and T. Al-Dala'in, "Towards Spatio-Temporal Crime Events Prediction," Multimedia Tools and Applications, 2023.
- [6] "Crime Risk Prediction Incorporating Geographical Spatiotemporal Dependency into Machine Learning Models," Information Sciences, vol. 646, article 119414, 2023.