IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Suraksha: Crime Awareness And Prevention By Analyzing Data

[1]George Lakra, [2]Hemanth Kumar L, [3]Mohammed Junaid Khan, [4]Ishaq Ahmed Patel, and [5]Amrita Sinha

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor

[1,2,3,4,5]Department of Information Science Engineering T.John Institute of Technology, Bangalore, India

Abstract: Safety of an individual is one of the most sensitive aspects of life. Non-locals and tourists often lack local crime-awareness and cannot take necessary precautions. This work presents Suraksha, an integrated system for crime awareness, prediction, and visualization that leverages historical crime data, predictive analytics, and interactive dashboards. The system fetches crime records, processes and normalizes them, predicts next-year crime trends, and presents structured outputs which the frontend consumes to display interactive visualizations and precaution messages.

***Index Terms* -** Crime Prediction, Data Normalization, Ridge Regression, Java Servlet, Python In- tegration, Data Visualization, Suraksha.

I. Introduction

Urban and tourist safety requires accessible, localized crime awareness. Suraksha is developed to address this need by providing a full-stack solution that fetches historical crime data, applies predictive analytics, and visualizes results on a web dashboard. The system integrates multiple components to ensure that users can quickly understand crime patterns in their area of interest, including top predicted crimes, historical trends, and demographic distributions of victims.

II. LITERATURE REVIEW

Crime analytics and visualization have gained significant attention in recent years, as urban plan- ning, public safety, and law enforcement increasingly rely on data-driven approaches. Several studies have contributed to this domain.

Prior research emphasizes the integration of interactive geographic and temporal visualizations to enhance situational awareness. For instance, visual analytics tools for COVID-19 and crime data demonstrate how combining multiple datasets with interactive dashboards can support informed decision- making [1].

Textual data mining and crime mapping have also been explored. Techniques such as focused crawling, named-entity recognition, and document classification allow the transformation of unstruc- tured news and reports into structured inputs suitable for crime trend analysis [2, 3].

Other approaches focus on predictive modeling using spatio-temporal and demographic features. Incorporating variables such as time of day, day of week, victim demographics, and geographic loca- tion improves forecasting accuracy [4, 5, 6, 7].

The analysis in this study was performed using the Indian Crimes Dataset [8], which contains com- prehensive crime records across various states and union territories of India.

This research utilizes a preprocessed version of the Indian Crimes Dataset [9], where we applied specific filtering criteria and feature engineering techniques to focus on relevant crime patterns for our analysis. The original dataset was extensively cleaned and transformed to suit the requirements of our prediction algorithm.

These studies collectively highlight three key themes relevant to our work:

- Integration of heterogeneous datasets (temporal, spatial, demographic) for actionable insights.
- Robust preprocessing and normalization of historical records, especially when data is partial or incomplete.
- Visualization and dashboard techniques that effectively communicate predictions to end-users.

III. SYSTEM OVERVIEW

Suraksha is designed as a modular system integrating database, backend, prediction engine, and frontend visualization to provide crime awareness and predictions. The system architecture consists of the following components:

- **Frontend:** A web-based interface where users enter the city of interest. It provides interactive charts and maps for visualizing historical and predicted crime trends.
- **Java Servlet Backend:** The servlet receives requests from the frontend, fetches historical crime data from the PostgreSQL database, and passes the data to the prediction module.
- **Python Prediction Module:** Normalizes the data, scales partial-year counts, predicts next-year crime counts using Ridge Regression, and formats data for charting and summaries.
- *Data Flow:*
  1. User inputs a city in the frontend.
  2. The servlet queries the database for historical records.
  3. Database rows are converted to JSON and sent to Python via a subprocess.
  4. Python processes the data and returns structured JSON to the servlet.
  5. The servlet merges the predictions with the original data and returns the final JSON to the frontend.
- Frontend updates charts and map visualizations accordingly.
- **Visualization:** Chart.js is used for displaying trends, probabilities, and distributions. OpenLay- ers is used for map-based visualizations including nearby police stations.

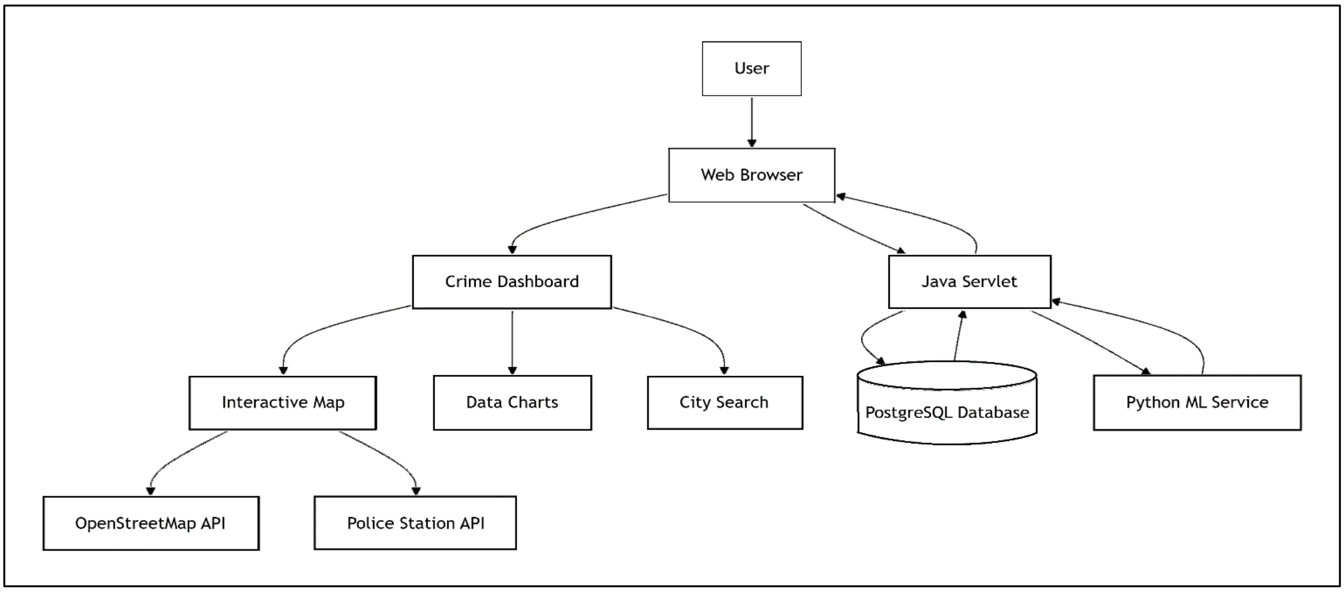

Figure 1: Suraksha system architecture (Frontend → Servlet → Database → Python → Frontend).

IV. METHODOLOGY

The methodology adopted in Suraksha focuses on structured data processing, predictive analytics, and interactive visualization:

1. **Data Collection:** Historical crime records are stored in PostgreSQL and fetched through a server-side function.
2. **Data Normalization:** Data fields such as dates, crime types, and time periods are standardized to ensure consistency across years and cities.
3. **Handling Partial-Year Data:** Counts are scaled proportionally to account for months with missing data, reducing bias in predictions.
4. **Prediction Algorithm:** Ridge Regression is applied to predict next-year crime counts per crime type. For crime types with insufficient historical data, the average of past occurrences is used.
5. **JSON Data Exchange:** The backend converts database rows into JSON, passes it to Python, and receives JSON predictions and summaries. This structured approach allows seamless integration between components.
6. **Visualization:** Frontend renders top predicted crimes, historical trends, time-of-day distribu- tions, victim demographics, and city maps. Chart.js and OpenLayers ensure interactivity and clarity.

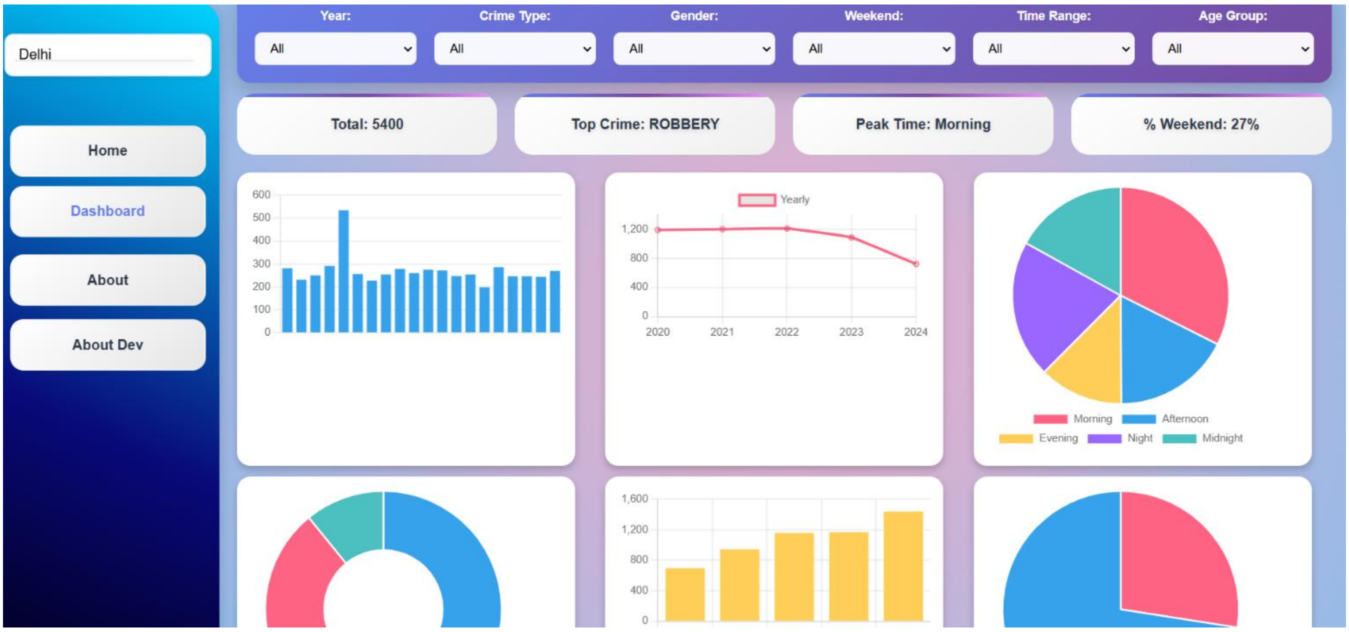

Figure 2: Example of Suraksha dashboard outputs.

V. RESULTS AND DISCUSSION

Suraksha delivers the following outputs:

- Top-3 predicted crime types per city with counts and probability percentages.
- Chart-ready historical rows for several charts including crime type frequency, year-over-year trends, and time-of-day distribution.
- Robust handling for missing or partial-year data by scaling counts to 12 months..

Figure 3: Top-3 predicted crime types per city.

VI. Limitations

- Prediction quality depends on historical depth per crime type; Ridge requires at least two yearly data points for accurate modeling.
- Scaling sparse months assumes uniform distribution across months, which may misrepresent seasonality.
- JSON parsing relies on clean output from the Python module; errors in formatting could affect data exchange.

VII. SECURITY AND DEPLOYMENT NOTES

- Keep database credentials and paths secure; do not commit sensitive information to public repositories.
- In production, consider REST endpoints instead of process-based stdin/stdout bridging for performance improvements.

VIII. CONCLUSION

Suraksha integrates historical crime data, predictive analytics, and interactive visualization to pro- vide users with actionable crime awareness. The system is modular and ensures seamless integration between database, backend, prediction module, and frontend. By presenting top predictions, chart- ready historical trends, and detailed insights, Suraksha enhances the safety and situational awareness of users in urban and tourist environments.

IX. ACKNOWLEDGMENT

We express our sincere gratitude to our guide, Amrita Sinha, from the Department of Information Science Engineering at T. John Institute of Technology for their exceptional guidance, constant encour- agement, and expert supervision throughout this research work. Their valuable suggestions and critical feedback significantly enhanced the quality of this paper. We also thank the institution for providing the necessary infrastructure and support.

X. References

[1] J. Smith and J. Doe, “Visual analytics for crime data and covid-19,” *Proceedings of the IEEE Computer Society Conference on Computer Science*, 2021.

[2] A. Wang and B. Li, “Textual data mining and crime mapping,” in *IEEE International Conference on Computational Techniques in Information and Data Engineering*, 2016.

[3] W. Zhang, “Web intelligence and agent technology for crime analysis,” in *IEEE/WIC/ACM International Conference*, 2006.

[4] R. Kumar and P. Sharma, “Spatio-temporal crime prediction using machine learning,” in *IEEE NCG 2018*, 2018.

[5] S. Lee and J. Park, “Predictive modeling in crime analytics,” in *IEEE ICEEICT 2022*, 2022.

[6] M. Patel and N. Singh, “Crime forecasting with data integration,” *IEEE Access*, 2019.

[7] L. Chen and Y. Wang, “Enhanced crime prediction using spatio-temporal features,” in *IEEE ICECA 2022*, 2022.

[8] S. HG, “Indian crimes dataset,” 2021, accessed: 2023. [Online]. Available: https:

//www.kaggle.com/datasets/sudhanvahg/indian-crimes-dataset

[9] G. Lakra, H. Kumar L, M. J. Khan, I. A. Patel, and A. Sinha, “Preprocessed indian crimes dataset,” 2023, enhanced version of [8] with custom filtering and feature engineering.