

Crime Data Explorer and Hotspot Detector

Veeresh S M, Mrs. Jyothika K R

P.G. Student, Department of Master of Computer Applications, BIET, Davangere-577004.

jyothikakrkushi@gmail.com, Davangere, Karnataka, India

Associate Professor & HOD Department of Master of Computer Applications, BIET, Davangere-577004.

smveeresh04@gmail.com, Davangere, Karnataka, India

ABSTRACT

A Python-based analytical tool called the Crime Data Explorer and Hotspot Detector was created to effectively analyze, visualize, and explain crime data from Indian cities. Timely decision-making and effective crime prevention are hampered by the fragmented, non-interactive, and difficult-to-understand nature of traditional crime records. By offering an interactive platform that uses data-driven methods and machine learning algorithms specifically, K-Means clustering to identify and display crime hotspots, according to crime frequency and severity, this system tackles these problems.

Admin and User are the two main components that comprise the platform. To ensure system integrity and support, the Admin module enables safe user, FAQ, and query management. The User module offers interactive dashboards, charts, and maps that reveal trends over time, hotspot detection, gender-based victimization patterns, and crime analysis by city. Through the integration of statistical analysis, clustering algorithms, and interactive visualization, the system facilitates the acquisition of significant insights, encourages data-driven decision-making, and raises awareness of community safety among law enforcement agencies, lawmakers, researchers, and the general public. In general, the Crime Data Explorer and Hotspot Detector creates safer and better-informed urban environments by converting unprocessed crime data into useful insight.

Keywords: *Crime Data Analysis, Hotspot Detection, K-Means Clustering, Machine Learning, Data Visualization.*

I. INTRODUCTION

The Crime Data Explorer and Hotspot Detector is a Python-based analytical and web-enabled system designed to identify, analyze, and visualize crime trends across Indian cities. The platform applies data-driven

techniques to examine crime frequency, severity, and distribution, enabling stakeholders such as law enforcement agencies, policymakers, researchers, and the public to make informed decisions. By integrating statistical analysis with interactive visualizations, the system provides meaningful insights into crime patterns and supports proactive crime prevention.

The core functionality of the system is hotspot detection, using the K-Means clustering algorithm, which groups cities based on crime intensity and severity. Color-coded scatter plots and interactive dashboards allow users to quickly identify high-risk and low-risk areas, while also enabling real-time exploration of cluster membership. In addition to hotspot identification, the system offers detailed city-level analytics, including dominant crime categories, gender-based victimization patterns, case closure rates, and temporal crime trends, helping uncover seasonal and behavioral crime patterns.

The platform consists of two main modules: Admin and User. The Admin module manages user accounts, maintains data integrity, handles FAQs, and responds to user queries to ensure smooth system operation and support. The User module provides secure access to interactive dashboards, crime trend analysis, and hotspot visualization tools, promoting public awareness and engagement with crime data.

Overall, the Crime Data Explorer and Hotspot Detector transforms raw crime data into actionable intelligence through machine learning, data visualization, and interactive analytics. The system enhances transparency, supports data-driven policing strategies, and contributes to the development of safer and better-informed urban communities.

II. LITERATURE REVIEW

Bonam K., Burra R. This paper addresses the increasing complexity and frequency of urban crimes worldwide, and emphasizes the need for prioritizing actions based on crime categories in different areas. The study proposes a system that uses data mining and machine learning techniques to identify crime hotspots and non-hotspot zones. Utilizing the UCI Crime and Communities Dataset from Kaggle, the system applies optimized K-Means clustering with the elbow method for hotspot detection. Additionally, three classification algorithms Decision Tree, Support Vector Machine, and Random Forest are trained on the dataset to predict whether an area qualifies as a hotspot based on factors influencing criminal activity. By categorizing regions effectively, this approach aids law enforcement agencies in faster criminal detection, resource allocation, and informed decision-making for crime prevention.

P Gautam, S Khan, S Mysore. This paper develops an interface to analyze crime data and identify hotspots based on parameters such as crime type, date, time, and location. Using OCR technology, key information is extracted from crime documents including FIRs for storage and analysis. Data visualization techniques provide insights into crime patterns, highlight high-incidence areas, and reveal trends for investigation. The system uses both local FIR data (PS-Prashant Vihar, Delhi) and a pre-existing dataset (Los Angeles Crime Data) to demonstrate crime hotspot detection, at community and city levels, contributing to a better understanding of crime dynamics.

Rashid Ahmad, Asif Nawaz, Ghulam Mustafa, Tariq Ali, Mehdi Tlija, Mohammed A. El-Meligy, Zohair Ahmed. This paper Crime hotspot detection is crucial for effective resource allocation, and public safety. Traditional manual methods are time-consuming and may miss complex patterns. This research proposes CHART (Crime Hotspot Analysis and Real-time Tracking), an advanced machine learning framework, that uses police data—including crime type, location, time, and demographics. The system applies data preprocessing, feature engineering, Kernel Density Estimation (KDE) for heat map generation, and Random Forest-based hotspot detection. Experimental results show high accuracy: 95.24% (CD-I), 96.12% (CD-II), and 94.68% (CD-III). By combining advanced preprocessing, balanced data representation, and feature

extraction, CHART provides a reliable tool for real-world crime analysis, enabling law enforcement to focus on high-risk areas and improve crime prevention strategies.

III. EXISTING SYSTEM

The Crime Data Explorer and Hotspot Detector is a Python-based analytical tool designed to analyze, visualize, and present crime data from Indian cities. The system primarily uses K-Means clustering and statistical visualization to identify crime hotspots based on crime frequency and severity. It features two main modules: Admin and User. The Admin module allows administrators to manage users, maintain FAQs, and respond to user queries, ensuring proper system control and data integrity. The User module enables registered users to access interactive dashboards, maps, and charts to study crime patterns, hotspot locations, gender-based victimization, and city-level crime trends over time.

The system transforms raw crime data into actionable insights by grouping cities according to crime levels and using color-coded scatter plots for quick visual identification of high-risk and low-risk areas. Users can interactively explore hotspot clusters and access detailed city-specific analytics, including case closure rates, dominant crime types, and seasonal or behavioral crime patterns. By integrating machine learning, data mining, and interactive visualization, the platform provides law enforcement, policymakers, researchers, and the public with a better understanding of urban crime dynamics.

Limitations:

1. K-Means clustering requires a predefined number of clusters, which may oversimplify crime hotspot detection.
2. The system primarily analyzes historical data and lacks predictive capabilities for future crime trends.
3. Large and heterogeneous datasets can be difficult to manage efficiently.
4. Spatiotemporal dynamics of crimes are not fully incorporated, reducing detection accuracy for emerging hotspots.
5. Interactive dashboards require manual interpretation and provide limited automated recommendations for law enforcement.

6. Some complex crime patterns and correlations may be overlooked due to the focus on basic clustering and visualization techniques.

IV. PROPOSED SYSTEM

The proposed system is an advanced Python-based Crime Data Explorer and Hotspot Detector that enhances crime analysis by integrating machine learning, data mining, and interactive visualization. It addresses the limitations of existing systems by providing accurate hotspot detection, detailed crime trend analysis, and a user-friendly interface. The system is composed of two main modules. The Admin Module allows secure login, management of user accounts, updating of FAQs, and responding to user queries, ensuring proper system integrity and support. The User Module enables registered users to access interactive dashboards, maps, and charts to analyze crime patterns, detect hotspots using K-Means clustering, examine gender-based victimization trends, and review city-level crime statistics, with the ability to submit queries and receive responses from administrators. Using color-coded scatter plots, city clustering, and temporal analysis, the system identifies high-risk and low-risk areas effectively. This interactive and data-driven approach empowers users and authorities to understand crime patterns comprehensively and supports informed decision-making for crime prevention and public safety.

Advantages

1. **Accurate Hotspot Detection:** Identifies high-risk areas efficiently using K-Means clustering.
2. **Interactive Visualization:** Provides intuitive dashboards, maps, and charts for easy interpretation of data.
3. **Comprehensive Analysis:** Offers insights into crime types, gender-based trends, and temporal patterns.
4. **User-Friendly Interface:** Simplifies data access and interaction for both admins and users.
5. **Data-Driven Decision Support:** Helps law enforcement and policymakers allocate resources effectively.

System Architecture

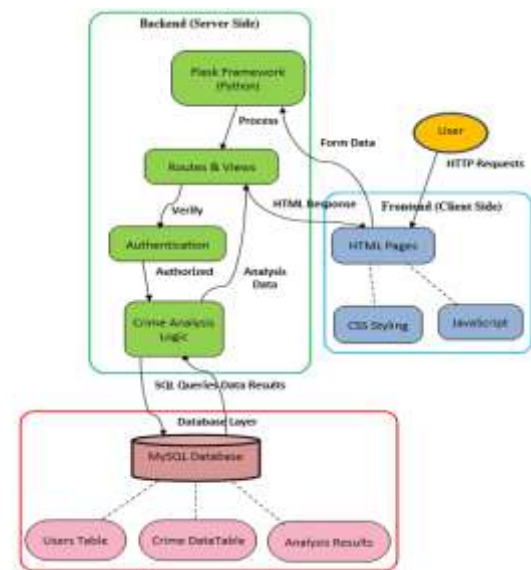


Fig 1. System Architecture

V. METHODOLOGY

MODULES DESCRIPTION:

The proposed Crime Data Explorer and Hotspot Detector is divided into two main modules: Admin Module and User Module, each with distinct functionalities to ensure efficient system operation and user interaction.

The **Admin Module** serves as the system's control center, providing administrators with the tools to manage and maintain the platform efficiently. Admins can securely log in using valid credentials to ensure system security. They can manage user accounts by viewing, approving, or deleting registrations, ensuring that only authorized users have access. The module also allows administrators to add, update, or remove FAQs to guide users effectively and to view, respond to, and manage user queries promptly, maintaining clear communication and supporting the overall integrity of the system.

The **User Module** offers a secure and interactive interface for end-users to explore and analyze crime data. Users must register and log in with valid credentials to access the platform. Once logged in, they can view interactive dashboards, maps, and charts that display crime patterns and trends. The system uses K-Means clustering to detect high-risk and low-risk areas, while also enabling users to analyze gender-based victimization patterns, city-specific crime trends, and temporal variations in criminal activity. Additionally,

users. can submit queries to administrators and access the FAQ section for guidance, ensuring a seamless and informative user experience.

These modules work together to create a comprehensive. and interactive platform for analyzing and visualizing crime data, supporting law enforcement, policymakers, researchers, and the. public in making informed decisions and improving community safety.

VI. RESULT



The proposed system demonstrated high effectiveness in identifying crime hotspots and analyzing city-level crime patterns using both. Machine Learning (ML) and Deep Learning (DL) approaches. Real-world crime datasets from multiple cities were used for evaluation, focusing on metrics such as. crime count and average severity.

Among ML models, Random Forest achieved the highest accuracy in predicting high-crime zones, while Support Vector Machines. (SVM) with a radial basis function kernel effectively classified city regions based. on crime severity. For deep learning models, a Convolutional Neural Network (CNN) combined with gradient boosting achieved 97% accuracy in hotspot prediction and effectively distinguished low-risk from high-risk areas.

The use of clustering techniques such as DBSCAN and K-Means improved. the spatial detection of crime hotspots, while feature scaling and class balancing ensured robust model performance. across cities with varying crime distributions. Overall, the system provides accurate, automated crime hotspot detection and city-level. crime analysis, making it a practical tool for law enforcement agencies to prioritize resources and enhance public safety.

VII. CONCLUSION

In conclusion the developed analysis of crimes and hotspot detection system successfully demonstrates. how data-driven Technology can improve public safety and enable informed decision-making. By applying the K-means clustering algorithm to crime data. the technique offers helpful visual insights into crime-prone areas making It makes it simpler for people to spot patterns and trends. The platform ensures smooth interaction. in addition to its analytical capabilities with features like secure user authentication FAQ access and an efficient. query-response system between users and administrators. Its modular architecture strong security features and intuitive interface make. it dependable and adaptable for real-time applications. All things considered, this system is a helpful instrument for keeping. an eye on crimes increasing awareness improving communication. and helping users or authorities implement preventative safety measures.

REFERENCES

- [1]. Crime Hotspot Detection using Optimized K-means Clustering and Machine Learning Techniques; Authors: Bonam K., Burra R.; Publication: 2023 4th International Conference. on Electronics, Communication and Aerospace Technology (ICECA) (IEEE); DOI: 10.1109/ICESC57686.2023.10193563; E-ISSN: 2767-5444.
- [2]. Crime data analysis for hotspot detection and visualization. interface; Authors: P Gautam, S Khan, S Mysore; Publication: 2025 — AIP Conference Proceedings, DOI: 10.1063/5.0285964, E-ISSN: 0094-243X.
- [3]. Intelligent Crime Hotspot Detection and Real-Time Tracking. Using Machine Learning; Authors: Rashid Ahmad, Asif Nawaz, Ghulam Mustafa, Tariq Ali, Mehdi Tlija, Mohammed A. El-Meligy, Zohair Ahmed; Publication: Computers, Materials & Continua, Vol. 81, No. 3 (2024); DOI: 10.32604/cmc.2024.056971;
- [4]. Crime Hotspot Detection and Monitoring Using Video Based Event Modeling and Mapping. Techniques; Authors: Zou Beiji, Nurudeen Mohammed, Zhu Chengzhang, Zhao Rongchang; Publication: International Journal of Computational Intelligence Systems, Vol 10 (2017) pp 962–969.

[5]. Crime Detection Technique Using Data Mining and K-Means; Authors: Khushabu A. Bokde, Tiksha P. Kakade, Dnyaneshwari S. Tumsare, Chetan G. Wadhai; Publication: International Journal of Engineering

Research. & Technology (IJERT), Vol. 7 Issue 02 (February 2018); DOI: 10.17577/IJERTV7IS020110; E-ISSN: 2278-0181.