



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Special Issue 1, January 2026

Blackspot Detection using Machine Learning

Ajay Singh Rajpoot¹, Sahil Rajendra Joshi², Sanika Jagdish Patil³, Jitesh Jaysingh Saldur⁴

UG Student, Dept. of Information Technology, Indala College of Engineering, Kalyan, Maharashtra, India^{1,2,3,4}

ABSTRACT: Road traffic accidents are a major public safety issue in developing countries like India, where increasing urbanization and vehicle density lead to frequent fatalities and injuries. Many accidents repeatedly occur at specific locations known as accident blackspots. Traditional identification methods are manual, time-consuming, and prone to bias. This project proposes a Machine Learning-based Blackspot Detection System that uses accident records, traffic data, environmental factors, and geospatial information. After data preprocessing and feature engineering, clustering and classification algorithms such as K-Means, DBSCAN, and Random Forest are applied. GIS-based visualization highlights high-risk zones, supporting data-driven road safety planning and intelligent transportation systems.

KEYWORDS: Machine Learning, Accident Blackspot Detection, Road Safety, Geospatial Analysis, Intelligent Transportation Systems

I. INTRODUCTION

Road business accidents constitute a critical public safety challenge across the globe, particularly in developing countries similar as India, where rapid-fire urbanization, population growth, and a steady increase in vehicle power have significantly boosted business traffic. According to public road safety statistics, a large proportion of fatal and non-fatal accidents do constantly at specific road locales, generally appertained to as accident blackspots. These locales are characterized by a high frequency and inflexibility of crashes due to factors similar as poor road figure, shy signage, business violations, environmental conditions, and inadequate business operation structure.

Conventional approaches for relating accident blackspots are largely dependent on homemade checks, fixed statistical thresholds, and retrospective analysis of literal accident records. While these styles have been used for decades, they suffer from several limitations. They're time-consuming, labor-ferocious, reactive in nature, and frequently fail to capture complex nonlinear connections among multiple contributing factors. also, traditional ways warrant scalability and rigidity, making them infelicitous for assaying large and continuously growing datasets generated by ultramodern business systems. With the rapid-fire advancement of data collection technologies and the vacuity of large-scale accident and business datasets, machine literacy (ML) has surfaced as a promising result for intelligent road safety analysis. Machine literacy ways are able of automatically discovering retired patterns, correlations, and threat factors from multidimensional data, enabling more accurate and visionary identification of accident-prone zones. By incorporating geospatial information, temporal trends, business inflow characteristics, and environmental conditions, ML-grounded systems can significantly enhance the effectiveness of blackspot discovery.

This exploration focuses on the development of a Machine literacy-grounded Accident Blackspot Detection System aimed at perfecting the delicacy, effectiveness, and trustability of relating high-threat road parts. The proposed system utilizes miscellaneous datasets similar as police-reported accident records, road structure details, business viscosity data, and geographical equals. These datasets suffer methodical preprocessing, including data cleaning, normalization, missing value running, and geocoding, to insure data quality and thickness. point engineering ways are applied to prize meaningful attributes related to accident frequency, inflexibility, time-of-day patterns, and spatial distribution. The design employs both unsupervised and supervised literacy ways to dissect accident data. Clustering algorithms are used to descry spatial attention of accidents, while bracket models assess the threat situations associated with different locales. To support decision-timber and real-world perpetration, the linked blackspots are imaged using Civilians-grounded mapping tools. The issues of this exploration end to help business authorities, civic itineraries, and policymakers in enforcing targeted safety measures, optimizing resource allocation, and eventually reducing road business accidents and associated losses.

II. RELATED WORK

Identification of road accident blackspots has been extensively studied to enhance road safety and reduce traffic fatalities. Initial research primarily relied on traditional statistical approaches, including accident frequency analysis, accident rate methods, and severity index calculations. Models such as Poisson and Negative Binomial regression were widely used to analyze historical crash data. Although effective for basic analysis, these methods are limited in handling nonlinear relationships among multiple contributing factors. With the advancement of spatial technologies, Geographic Information Systems (GIS) were introduced for blackspot analysis. GIS-based techniques such as Kernel Density Estimation (KDE) and spatial clustering methods enabled visualization of accident-prone regions by analyzing spatial distributions of crash data. However, these approaches are mostly descriptive in nature and lack predictive capability, restricting their usefulness for proactive accident prevention. Recent research has increasingly focused on machine learning-based techniques to overcome these limitations. Supervised learning algorithms such as Decision Trees, Random Forests, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) have been applied to classify road segments based on risk levels. Among these, Random Forest models have shown improved accuracy due to their robustness against noise and ability to handle complex feature interactions.

Furthermore, deep learning approaches, including Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN), have been explored for accident prediction and spatial feature extraction using large datasets and satellite imagery. These methods provide high prediction accuracy but require extensive computational resources and large labeled datasets. More recent studies integrate real-time data sources such as traffic sensors, GPS data, and weather information to dynamically identify blackspots. Despite these advancements, challenges such as data imbalance, model interpretability, and real-time deployment remain. The proposed system builds upon existing work by leveraging machine learning for accurate and scalable blackspot detection.

III. WORKING OF MODEL

Step 1: Data Collection

Collect accident-related data from authorized sources such as police FIR records, traffic databases, road infrastructure reports, and environmental datasets.

Step 2: Data Preprocessing

Clean the collected data by removing duplicate entries, handling missing values, and correcting inconsistencies. Normalize numerical attributes and encode categorical variables.

Step 3: Geocoding

Convert textual location information into geographic coordinates (latitude and longitude) to enable spatial analysis.

Step 4: Feature Engineering

Extract and derive relevant features such as accident frequency, severity index, temporal patterns, traffic density indicators, and spatial proximity measures.

Step 5: Model Training – Clustering

Apply unsupervised learning algorithms such as K-Means and DBSCAN to identify spatial clusters of accident occurrences.

Step 6: Model Training – Classification

Train supervised machine learning models such as Random Forest and Gradient Boosting to classify road locations into different risk levels.

Step 7: Model Evaluation

Evaluate model performance using accuracy, precision, recall, F1-score, and clustering validation metrics.

Step 8: Blackspot Identification

Identify high-risk locations exceeding predefined thresholds as accident blackspots.

Step 9: Visualization

Visualize detected blackspots on GIS-based maps and dashboards for easy interpretation.

Step 10: Output Generation

Generate the final output consisting of geospatially mapped blackspots with associated risk levels to support decision-making.

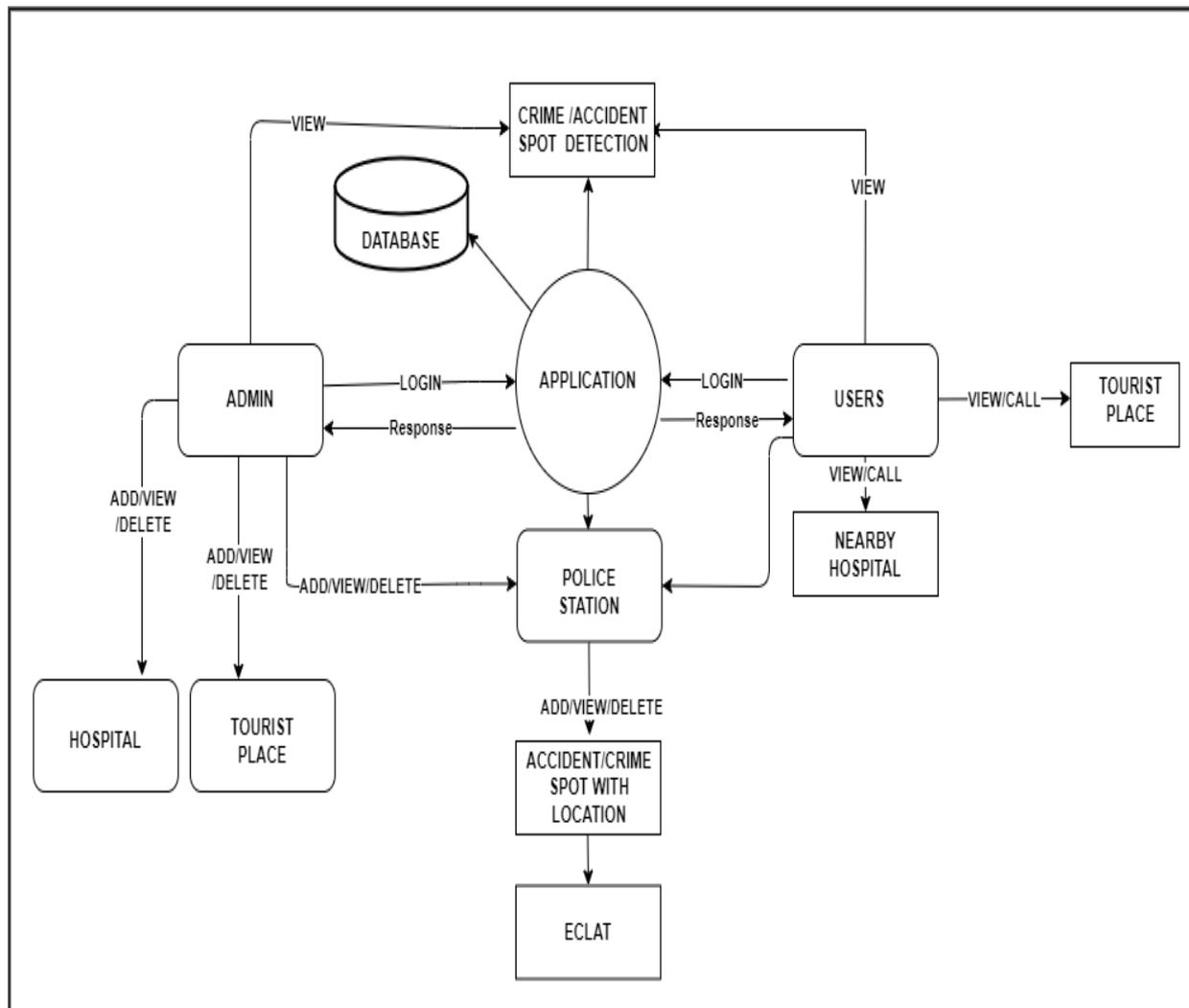


FIG 5.1 DFD of the BlackSpot Detection System

IV. RESULTS

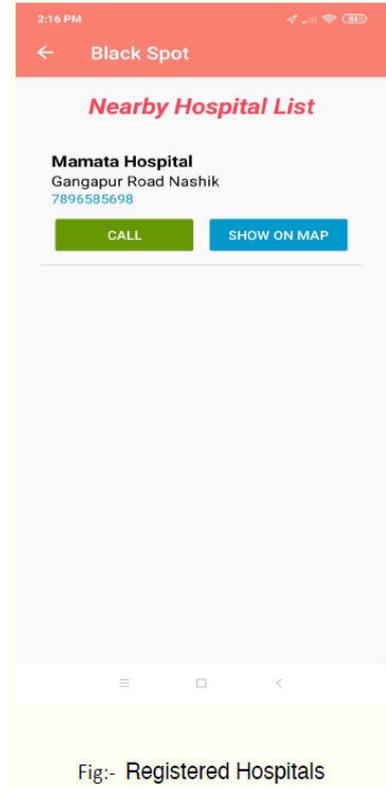
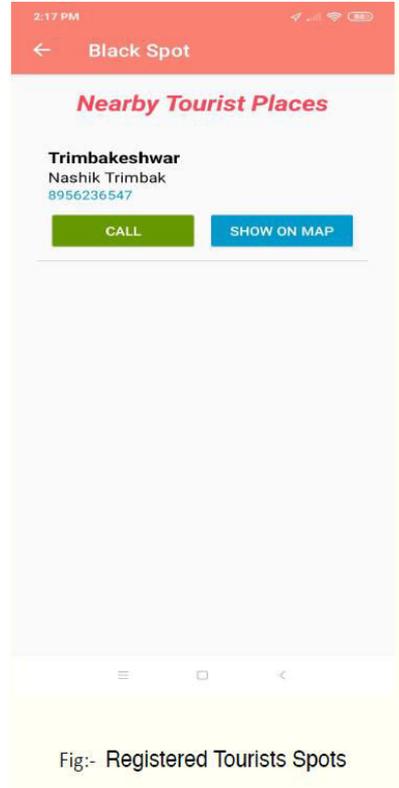
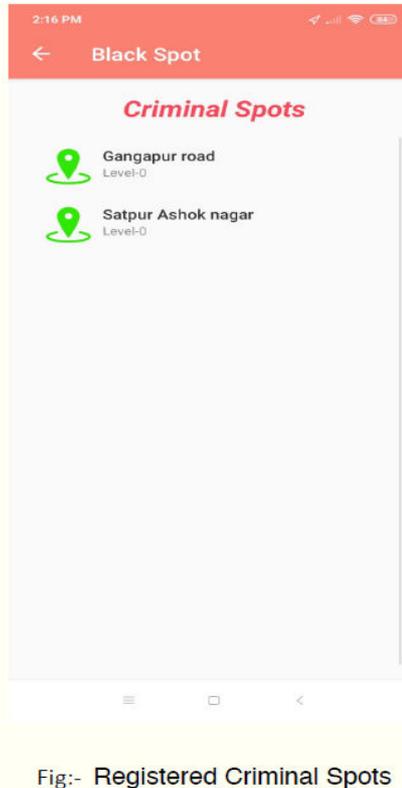


FIG 6.1 RESULTS OF THE PROJECT

V. CONCLUSION

In conclusion, black spot detection using an Android app presents an innovative and practical solution for improving road safety by leveraging technology to identify and warn users about accident-prone areas. By integrating historical accident data, real-time inputs, and GPS-based tracking, the app provides drivers with timely alerts, helping to reduce the risk of accidents in high-risk zones. The ECLAT algorithm enhances this system by efficiently mining frequent accident patterns, allowing the app to pinpoint black spots accurately. Despite challenges such as data accuracy, user participation, and privacy concerns, this approach offers significant potential in making roads safer for all. Future developments, such as real-time data integration and predictive analytics, could further improve the system’s ability to prevent accidents and save lives.

REFERENCES

- 1) X. Wang, C. Zhu, and J. Jiang, "A deep learning and ensemble learning based architecture for metro passenger flow forecast," in IET Intelligent Transport Systems, vol. 17, pp. 487-502, 2023, doi: 10.1049/itr2.12274.
- 2) N. Pathik, R. K. Gupta, Y. Sahu, A. Sharma, M. Masud, and M. Baz, "AI Enabled Accident Detection and Alert System Using IoT and Deep Learning for Smart Cities," Sustainability, vol. 14, no. 13, p. 7701, Jun. 2022, doi: 10.3390/su14137701.
- 3) M. M. Rahman, M. K. Islam, A. Al-Shayeb, and M. Arifuzzaman, "Towards Sustainable Road Safety in Saudi Arabia: Exploring Traffic Accident Causes Associated with Driving Behavior Using a Bayesian Belief Network," Sustainability, vol. 14, no. 10, p. 6315, May 2022, doi: 10.3390/su14106315.
- 4) H. Hozhabr Pour et al., "A Machine Learning Framework for Automated Accident Detection Based on Multimodal Sensors in Cars," Sensors, vol. 22, no. 10, p. 3634, May 2022, doi: 10.3390/s22103634.

- 5) K. Athiappan, C. Karthik, M. Rajalaskshmi, C. Subrata, H. R. Rabiei Dastjerdi, Y. Liu, C. Fernández-Campusano, and M. Gheisari, "Identifying Influencing Factors of Road Accidents in Emerging Road Accident Blackspots," *Advances in Civil Engineering*, vol. 2022, article ID 9474323, pp. 1-10, 2022, doi: 10.1155/2022/9474323.
- 6) A. Comi, A. Polimeni, and C. Balsamo, "Road Accident Analysis with Data Mining Approach: Evidence from Rome," in *Transportation Research Procedia*, vol. 62, pp. 798-805, 2022, ISSN 2352-1465, doi: 10.1016/j.trpro.2022.02.099.
- 7) A. W. Siddiqui, S. A. Raza, M. A. Elahi, K. S. Minhas, and F. M. Butt, "Temporal impacts of road safety interventions: A structural-shifts-based method for road accident mortality analysis," in *Accident Analysis & Prevention*, vol. 174, p. 106767, 2022, ISSN 0001-4575, doi: 10.1016/j.aap.2022.106767.
- 8) J. P. Montero-Salgado, J. Muñoz-Sanz, B. Arenas-Ramírez, and C. Alén-Cordero, "Identification of the Mechanical Failure Factors with Potential Influencing Road Accidents in Ecuador," *International Journal of Environmental Research and Public Health*, vol. 19, no. 13, p. 7787, Jun. 2022, doi: 10.3390/ijerph19137787.
- 9) T. Bokaba, W. Doorsamy, and B. S. Paul, "Comparative Study of Machine Learning Classifiers for Modelling Road Traffic Accidents," *Applied Sciences*, vol. 12, no. 2, p. 828, Jan. 2022, doi: 10.3390/app12020828.
- 10) M. Sobhana, V. K. Rohith, T. Avinash and N. Malathi, "A Hybrid Machine Learning Approach for Performing Predictive Analytics on Road Accidents," 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026404.
- 11) D. Santos, J. Saias, P. Quaresma, and V. B. Nogueira, "Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction," *Computers*, vol. 10, no. 12, p. 157, Nov. 2021, doi: 10.3390/computers10120157.
- 12) E. K. Adanu, W. Agyemang, R. Islam, and S. Jones, "A comprehensive analysis of factors that influence interstate highway crash severity in Alabama," in *Journal of Transportation Safety & Security*, vol. 14, no. 9, pp. 1552-1576, 2022, doi: 10.1080/19439962.2021.1949414.
- 13) W. Tao et al., "An Advanced Machine Learning Approach to Predicting Pedestrian Fatality Caused by Road Crashes: A Step toward Sustainable Pedestrian Safety," *Sustainability*, vol. 14, no. 4, p. 2436, Feb. 2022, doi: 10.3390/su14042436.
- 14) Z. Yang, W. Zhang, and J. Feng, "Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework," in *Safety Science*, vol. 146, pp. 105522, 2022, ISSN 0925-7535, doi: 10.1016/j.ssci.2021.105522.
- 15) D. Devaraj, D. Chandrasekaran, B. Pandian, B. Binitha, M. Dipikhasre, and R. Anil, "Road accident analysis in Kerala and location based severity level classification using decision tree algorithm," *Paid. J.*, vol. 14, no. XIV, pp. 41-50, 2021.
- 16) A. Kashevnik, I. Lashkov and A. Gurtov, "Methodology and Mobile Application for Driver Behavior Analysis and Accident Prevention," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2427-2436, June 2020, doi: 10.1109/TITS.2019.2918328.
- 17) J. Shen, T. Wang, C. Zheng, and M. Yu, "Determinants of Bicyclist Injury Severity Resulting from Crashes at Roundabouts, Crossroads, and T-Junctions," *Journal of Advanced Transportation*, vol. 2020, article ID 6513128, pp. 1-12, 2020, doi: 10.1155/2020/6513128. R. Chen et al., "Toward Secure Distributed Spectrum Sensing in Cognitive Radio Networks," *IEEE Commun. Mag.*, vol. 46, pp. 50–55, Apr. 2008.
- 18) H. Khalife, N. Malouch, S. Fdida, "Multihop cognitive radio networks: to route or not to route," *IEEE Network*, vol. 23, no. 4, pp. 20-25, 2009.
- 19) Y.-C. Liang et al., "Sensing-Throughput Trade-off for Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1326–37, April 2008.
- 20) P. K. Visscher, "How Self-Organization Evolves," *Nature*, vol. 421, pp. 799–800 Feb. 2003.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details