Risk Zone Awareness: Detecting Black Spots and Alerting Users to Potential Hazard

Dr. Priya Karemore
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
pskaremore_it@jspmrscoe.edu.in

Prof. Rapali Umbare
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
rtumbare_it@jspmrscoe.edu.in

Viraj Sonawane
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
virajsonawane140@gmail.com

Atharva Kanase
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
atharvaskanase@gmail.com

Abhishek Gholap
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
imabhishekgholap@gmail.com

Sunil Naunde
Information Technology
Rajarshi Shahu College of Engineering
Tathawade, Pune, India
sunilnawande@gmail.com

***Abstract*—Road and traffic is an important issue not only for Indian government but also for common people.Mostly,it is found that Road Accident are more frequent at certain specific locations.The analysis of this spot will help us in identifying the area or the spot where the accident occurs frequently on that location.In this project we are trying to apply certain login and methods to reduce the number of accident on that particular spot or location. In this project we will develop one application which will include the login page and sign up page for the user and will consist another login page for the admin. Admin will be provided with separate login credential.After login of the user the google map will be displayed to the user. User can see the many spot on that map and the accident spot will be highlighted over the map. After the admin login, admin will be able to add the accident spot using the form. When the user will travel from one place to another as he/she will come near to accident spot they will receive the voice alert on their device. From this project we will try to reduce the count of the accident and contribute in reducing the fatality rate.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Road accidents have become a major global health concern due to the increase in car ownership, as they are a major cause of death and morbidity rates. These are multifactorial accidents, defined by the involvement of at least one moving vehicle and ending in injuries or fatalities. A number of reasons, including human error, driver weariness, speeding, bad road conditions, and traffic congestion, can lead to these kinds of occurrences. Significantly, the number of fatalities from traffic accidents increased in India from 37.5 in 2020 to 38.6 in 2021, as measured by the deaths per 100 collisions metric. There is an urgent need for creative ideas to improve road safety in response to this growing issue.In light of this, the suggestion of an application that would notify users as they approach locations that are prone to accidents is a proactive step toward reducing risks and raising road safety standards.

An extensive dataset of accident records has been obtained in order to investigate the underlying causes of traffic accidents in India. But as it is now, this data is disorganized, with information dispersed throughout an Excel database table, making it difficult to conduct insightful analysis and interpretation. By using data mining tools, this dataset offers the chance to reveal important insights and identify trends. Through the identification of relevant characteristics and trends, the research aims to provide insight into the reasons that lead to traffic accidents and determine which demographic groups are more susceptible to them. It is possible to develop well-informed interventions to address the root causes of traffic accidents and support road safety measures by raising awareness and understanding of the issue.

For both the Indian government and the general public, traffic congestion and road accidents are urgent problems. The consequences of traffic accidents go far beyond statistics; they have significant effects on public health and safety. The widespread desire to reduce risks and guarantee road safety emphasizes the necessity of coordinated action. The utilization of data analytics to examine traffic accident data can yield significant insights that can guide evidence-based interventions and promote safer driving behaviors. In the end, the pursuit of road safety goes beyond individual initiatives and requires group efforts to make a noticeable difference and lower the number of traffic accidents.

Data on traffic accidents can be analyzed to provide important insights that might guide evidence-based interventions and promote safer driving behaviors. However, there are obstacles to overcome when putting an Accident Spot Alert System into place, like ensuring that all system criteria are met,

This is an open access Journal

getting precise accident data, and streamlining data collection procedures.

II. Related Work

In the last few years, a lot of research has been done on Blackspot detection. In an effort to lower the accident rate, the Indian government is also concentrating on identifying blackspots. It has been shown in the past that finding BlackSpots can reduce accident rates by up to 28 percentage. Here are a few research publications.

[1] "Identification of Blackspots and junction improvements in Vishakhapatnam city," an analysis of the Andhra Pradesh city of Vishakhapatnam, India. With a total area of 550 km2, it is the second-largest city in Andhra Pradesh. Aside from being a port city, its main function is industrial. Approximately 59 Percentage of the district's overall traffic volume is generated by the city of Visakhapatnam (Gopala raju, 2011). Places with a greater average accident rate are referred to as "BlackSpots." One of the best methods for preventing traffic accidents is the identification, evaluation, and mitigation of high-risk areas. Typically, formal road safety audits are used to determine whether places are considered unsafe.

[2]In the study "Identification of Accident Black Spots for National Highway Using GIS," the traffic in Meerut and Muzaffarnagar was examined. Haridwar and Meerut District to the south and west respectively enclose Muzaffarnagar District. northern district.Data were gathered from police stations, and a topographical map survey was examined. Then, using the Global Positioning System, the Ground Control Points were located, and the Critical Crash Rate Factor Method was used to identify the black spots. [3] "Black Spots Analysis On Pune - Bangalore National Highway" examined accident data from the National Highway Authority of India (NHAI) for the years 2014–2015 in order to identify unintentional Black Spots on a stretch of NH–4 (820 km–830 km). The Accidental Density Method (ADM) and the Weighted Severity Index (WSI) were utilized for recognizing the black spots.Through the use of the Accidental Density Method (ADM) with all of its parameters, they discovered dark spots at chainages of 821.2, 823, 824.1, 825.3, and 829.1 kilometers.

III. Methodology

The goals of this RSA activity are to raise road user safety, decrease accident rates, lessen traffic congestion, identify conflict sites, and improve the quality of the current roadways. The flow chart in Figure 1 illustrates the processes that were taken for the RSA.

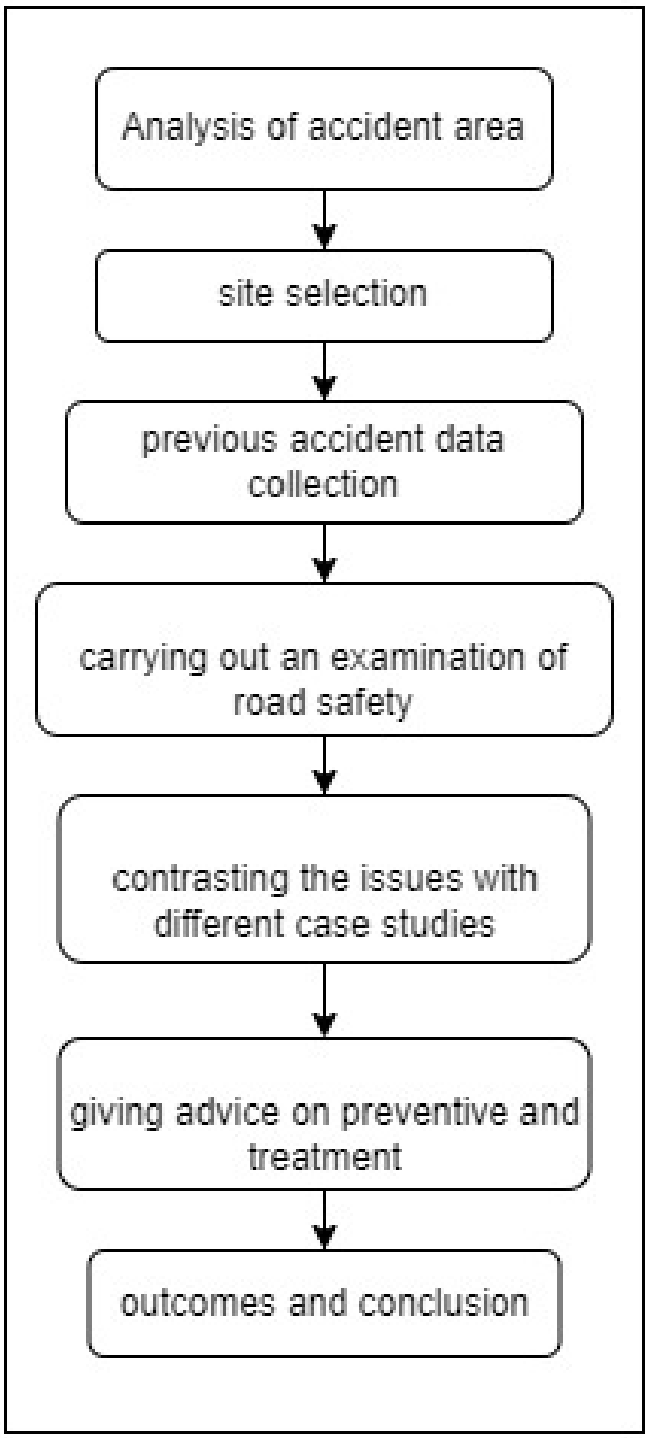

Fig.1 Road safety methodology

A. Analysis of Accident area

For Pune, India, here are some routes that are known to be vulnerable stretches where accident data could be collected:

1.Mumbai-Pune Expressway (NH 48)
2.Pune-Solapur Road (NH 65)
3.Pune-Ahmednagar Road (NH 60)
4.Pune-Nashik Road (NH 50)
5.Pune-Satara Road (NH 4).

This is an open access Journal

B. Site selection

We have chosen the Pune-Satara Road (NH 4), Maharashtra, India, from the list of previously specified routes for our survey. This road is among the most susceptible to accidents that result in fatalities, according to the study of the accident zones. We covered a distance of 10 km with the audit, starting at Wakhad (0 km) and ending at Chennai's Katraj Tunnel (28 km). Important locations like Baner, Kothrud, Chandani Chowk, and Warje, Pune, are included in this stretch.

Figure 2 displays the location map with the terrain and traffic image.

C. Accident Data Collection:

1) From Police Station

The accident data for the chosen State highway segment were gathered from multiple police stations. Data on accidents were gathered throughout a number of prior years.

The accident site and specifics, such as whether the collision was fatal or not, were included in the accident data that was accessible from police records.

The data is collected from the following police station.

a. Warje Police Station

b. Sinhagad Police Station

c. Katraj Police Station

d. Wakad Police Station

2) Old news paper data using machine learning

a. Importing the necessary Python libraries and data sets.

b. The articles in the corpus are then vectorized. We employ Sci-Kit Learn's CountVectorizer for vectorization in order to produce a sparse matrix that contains the count of every word in an article.

c. Using Sci-Kit Learn's TfidfTransformers, we then determine the inverse term frequency for the terms for improved outcomes. We would use classification methods on this vectorized word matrix, which we obtained as a sparse matrix, to predict classes for the data in the test data set.

d. To generate a comparative report, compare the accuracy, training times, testing times, projections, etc.

D. Conduct of road safety audit data:

The RSA was carried out measures according to protocol and with all safety precautions. According to the results of the road safety audit, the following issues were identified and pointed out as the primary ones on that particular length of road:

1. There are no road markers for the whole length of the route.

2. Uncontrolled vegetation growth on both sides of the Wakad to Katraj Tunnel route is a major problem that causes the following miseries:

i. Vegetation development up to the end lane marker prevents the availability of the proper shoulder width.

ii. Because the required minimum shoulder width is not available, parking is not possible.

3. The entire length of the route is without street lighting.

4. Many places along the route lack appropriate signboards. The primary signage needed on the road are: i. School Zone v. Speed Limit vi; ii. U-turn or No U-turn iii. Bus stop Ahead is a dangerous bend.

5. Improper illumination at bus stations and significant intersections.

6. Retro reflectors are not installed or are not functional in several road areas.

7. There are no speed limiters on the nearby roads that link to this one.

8. Town name boards are not erected or kept up to date.

IV. Literature Survey

A literature review provides an overview of previously published publications on a certain topic. A academic paper can be a standalone piece or a component of a larger work, such as a book or essay. Its major goal is to offer the researcher/author and the audience with a broad comprehension of the available information on the topic under consideration. A thorough literature evaluation helps researchers identify the suitable study issue, theoretical framework, and methodology. Essentially, a literature review situates the current work within the relevant literature and offers context for the reader. As a result, the review is typically written before the technique and outcomes portions of the study.

The literature review focuses on text classification as a core topic in natural language processing and investigates different machine learning approaches for newspaper article classification. It contrasts convolutional neural networks (CNN), a deep learning technique, with more conventional methods like support vector machines (SVM), naive bayes, and K nearest neighbors (KNN). The study looks into preprocessing techniques like N-Grams and Bag-of-Words and assesses how they affect prediction time, training time, and classification accuracy. According to the results, SVM performs more accurately than KNN and Naive Bayes, whereas CNN gives competitive performance despite taking longer to train.

Sr no.	Year	Method Used	Limitations	Practical Implications	Results
1	2022	OPTICS and HDBSCAN clustering algorithms, Cluster density, and Log Likelihood Ratio methods for evaluating significance	Ambiguity in determining the shape of clusters in road safety, Possibility of spurious patterns in the data	Identification of high-risk road accident zones, Application of machine learning algorithms and statistical significance to eliminate spurious patterns	Proposed approach identifies high-risk road accident zones, Comprehensive analysis of seasonal and temporal variations
2	2023	Integration of space-time cubes and machine learning techniques, Use of eXtreme Gradient Boosting (XGBoost) classifier for prediction	Ignoring spatio-temporal features surrounding hot spots, Higher false discovery rates when applying local spatial indices	Policy recommendations to enhance urban road traffic safety in the future, Importance of considering features of neighboring spots	Lower false discovery rates and higher prediction accuracy
3	2021	Deep learning for hotspot zone detection, Template matching for marking specific areas	The real-time performance of the system, especially in marking specific areas using the template matching algorithm, may be impacted by computational requirements and processing speed, potentially affecting the system's responsiveness in dynamic environments.	Identification of hotspot zones for Covid19 disinfection, Utilization of YOLO-v3 for real-time person detection accuracy	YOLO-v3 achieved maximum accuracy for person detection, 86.7% accuracy in detecting hotspot areas using confusion matrix
4	2023	Sorting, grouping, and accident prediction methods, Machine learning algorithms (classification, regression, clustering, dimensionality reduction)	Limitations of the sorting, grouping, and accident prediction methods, Limitations of statistical analysis methods for processing the data provided by ELSTAT	Identification of black spots is crucial for reducing traffic accidents, Machine learning algorithms can effectively classify data and predict accidents	Binary logistic regression and machine learning algorithms were compared for performance, Three metrics (Accuracy, Recall, and F1 Score) were used for evaluation
5	2020	ML-based technique for tentative decisions, Viterbi algorithm with HMM for final decisions	The framework relies on mobile sensing data from commercial-use vehicles, which may not always be comprehensive or accurate, potentially leading to false detections or missed hazardous spots.	Enhances hazardous spot detection accuracy using V2V cooperative framework, Combines mobile sensing and opportunistic networking for road safety improvement	Viterbi algorithm improved hazardous spot detection accuracy over support vector machines, V2V cooperative data processing enhanced accuracy for all vehicles
6	2022	EcRD: Edge-cloud-based Road Damage detection and warning framework, FedRD: Federated Road Danger detection framework with feature fusion and privacy protection	Challenging to recognize road dangers due to cluttered backgrounds and high intra-class differences, Limited data for one party and high privacy leakage risk	The proposed frameworks improve real-time road danger prediction and notification, They address challenges such as latency, communication efficiency, and privacy	EcRD achieves 91.96% accuracy with 0.0043s latency, FedMRD achieves 96.42% higher accuracy with 0.0351s latency and up to 250 times less communication cost
7	2020	Identification of jump-out danger spots, Utilization of past and surrounding information	The device's capacity to identify danger places depends on the availability and accuracy of facility identification information, as well as information about jump-out objects.	Identifies jump-out danger spots effectively in navigation devices, Enhances safety by identifying potential jump-out accidents based on past data	Identifies jump-out danger spots effectively, Navigation device can identify danger spots due to planimetric features
8	2021	Generating spatial zones using locating and electronic management means, Determining contour shape and dimensions based on acquired spatial coordinates	The performance of the system may differ based on the region or particular features of the area in which it is installed.	Enhances safety by creating and detecting danger zones accurately, Efficiently manages spatial zones using electronic and locating systems	Method for generating spatial zones, particularly danger zones, in space, Uses locating and electronic management means for defining and recording zones

This is an open access Journal

Sr no.	Year	Method Used	Limitations	Practical Implications	Results
9	2013	Confirm buffer area resolution, generate danger zone buffer, merge buffer, Query result buffer area, calculate ship resolution coordinates, query values	Unexpected occurrences or quickly shifting circumstances may affect how well the system identifies danger zones and issues warnings.	High-efficiency detection of ships entering complex danger zones, Reduced computation complexity in danger zone border detection	Strong adaptivity, reduced computation complexity in danger zone border detection, Supports high-efficiency detection of ships entering complex
10	2021	statistical techniques such as models based on hazards and regression analysis. Random Forest is one type of machine learning model used to forecast traffic incidents. Convolutional neural networks and Random Forests are combined in this ensemble learning model.	There are no explicit restrictions stated in the settings that are given.	RF and CNN are used in the RFCNN model to predict the severity of traffic accidents. determines important characteristics that affect the severity of accidents, such as visibility, temperature, and distance. improves the F-score for accident severity as well as forecast accuracy, precision, and recall.	With an accuracy of 0.991, the RFCNN model outperformed all other models. All models classification performance were improved by using significant features. At 0.974, the RF classifier provided the second-highest accuracy value. discusses difficulty and future plans for the application of multi-domain datasets.
11	2023	Unsupervised learning for the classification of driving behavior. Matching graphs to analyze similarities for pattern recognition and classification. Using sensor data and deep learning, driving habits are recognized.	Costs of communication and privacy are challenged by centralized data storage. Federated learning runs the risk of model divergence and consumes a large amount of computational resources.	Federated learning makes machine learning more robust, scalable, secure, and private. Federated learning overcomes centralized machine learning limitations in road traffic analysis.	Federated learning overcomes the limits of centralized machine learning to improve road safety. Machine learning techniques forecast obstacles to prevent mishaps and fatalities.
12	2016	Integrative thinking that addresses uncertainty by combining qualitative and probabilistic methods.trial-and-error-based adaptive analysis for making decisions.	Absence of an exhaustive summary of the contributions in the risk field. Not a thorough examination of risk assessments' shortcomings.	clarifies fundamental problems in risk management and assessment. places a focus on integrative research to examine risks from a wider angle.stimulates curiosity in the fundamental difficulties and uncertainties of the risk field.	highlights fundamental problems in risk management and assessment. covers decision-maker review, knowledge base, and risk assessment.

V. Conclusion

The project's analysis focuses on the crucial topic of road safety, with a specific emphasis on locating and reducing accident-prone areas. The project suggests creating an all-inclusive programme with integrated user-friendly interfaces for administrators and regular users. Through Google Maps, users will have access to up-to-date information regarding accident-prone locations, allowing them to navigate with awareness and caution. In contrast, administrators will be able to add new accident data to the system for ongoing updates and observation.

This research is important because it has the potential to lower the number of traffic accidents and fatalities by alerting and advising drivers in advance of potentially dangerous locations. The project intends to find patterns and trends in accident data by utilising data analytics and machine learning techniques, providing insights into the underlying causes of accidents and potentially vulnerable demographic groups. These insights have the potential to enhance road safety measures through the provision of evidence-based interventions and support.

The project's methodology calls for a methodical approach to data collecting, analysis, and intervention. Every action, from identifying locations that are prone to accidents to conducting road safety audits, is intended to advance the overall objective of raising the bar for road safety. The focus on using technology—such as GIS and machine learning—highlights the value of innovation in tackling difficult societal issues like road safety.

The project's conclusion emphasises the critical need for proactive approaches to address issues related to road safety, particularly in nations like India where there is a substantial danger to the public's health from traffic accidents. Through the application of technology and data-driven insights, the

This is an open access Journal

initiative hopes to significantly lower the number of traffic accidents and guarantee safer roads for all.

REFERENCES

[1] R. Jain and A. Bhat, "Determining Statistically Significant Road Accident Spatial Hotspots using Machine Learning Approaches," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 214-221

[2] Peijie Wu, Tianyi Chen, Yiik Diew Wong, Xianghai Meng, Xueqin Wang, Wei Liu,Exploring key spatio-temporal features of crash risk hot spots on urban road network: A machine learning approach,Transportation Research Part A: Policy and Practice,Volume 173,2023,103717,ISSN 0965-8564, https://doi.org/10.1016/j.tra.2023.103717.

[3] M. Z. Khan, M. U. G. Khan, T. Saba, I. Razzak, A. Rehman and S. A. Bahaj, "Hot-Spot Zone Detection to Tackle Covid19 Spread by Fusing the Traditional Machine Learning and Deep Learning Approaches of Computer Vision," in IEEE Access, vol. 9, pp. 100040-100049, 2021, doi: 10.1109/ACCESS.2021.3094720

[4] Ioannis, Karamanlis., Alexandros, Kokkalis., V., A., Profillidis., George, Botzoris., Athanasios, Galanis. (2023). Identifying Road Accident Black Spots using Classical and Modern Approaches. WSEAS transactions on systems, doi: 10.37394/23202.2023.22.56

[5] Yoshito, Watanabe., Wei, Liu., Yozo, Shoji. (2020). Machine-Learning-Based Hazardous Spot Detection Framework by Mobile Sensing and Opportunistic Networks. IEEE Transactions on Vehicular Technology, doi: 10.1109/TVT.2020.3021411

[6] (2022). Smart Road Danger Detection and Warning. doi: 10.53846/goediss-8965

[7] Shimizu, Akira., Matsumaru, Makoto., Noguchi, Ryoji., Takemura, Itaru. (2020). Dangerous place identification device, map data, dangerous place identification method, and program.

[8] Panaget, Jean-Philippe. (2021). Methods and system for generating and detecting at least one danger zone.

[9] Chen, Houwu., Liu, Song., Ma, Lin., Liu, Peng. (2013). Danger zone detection method.

[10] M. Manzoor et al., "RFCNN: Traffic Accident Severity Prediction Based on Decision Level Fusion of Machine and Deep Learning Model," in IEEE Access, vol. 9, pp. 128359-128371, 2021, doi: 10.1109/ACCESS.2021.3112546.

[11] Philippe Barbosa Silva, Michelle Andrade, Sara Ferreira, Machine learning applied to road safety modeling: Asystematic literature review, Journal of Traffic and Transportation Engineering (English Edition), Volume 7, Issue 6, 2020, Pages 775-790, ISSN 2095-7564, https://doi.org/10.1016/j.jtte.2020.07.004.

[12] Terje Aven, Risk assessment and risk management: Review of recent advances on their foundation, European Journal of Operational Research, Volume 253, Issue 1, 2016, Pages 1-13, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2015.12.023.

This is an open access Journal