



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Road Accident Severity Prediction using Machine Learning

Rajeshwari S, Varsha M L, Kavya M

Assistant Professor, Department of Artificial Intelligence and Data Science Engineering, CIT, Gubbi, Tumkur,
Karnataka, India

U.G. Student, Department of Artificial Intelligence and Data Science Engineering, CIT, Gubbi, Tumkur,
Karnataka, India

U.G. Student, Department of Artificial Intelligence and Data Science Engineering, CIT, Gubbi, Tumkur,
Karnataka, India

ABSTRACT This study develops a deep learning-based system to predict road accidents using factors like speed, traffic, and weather. It emphasizes the requirement of accurate predictions to enhance road safety, reduce accidents, and implement preventative measures. focuses on classifying accident severity into categories such as Fatal, Grievous, and Simple Injury using techniques like AdaBoost and deep neural networks [1]. evaluates many machine learning models, as well as Random Forest, Logistic Regression, and XGBoost, for predicting accident severity [2]. It emphasizes human factors (e.g., alcohol use, age) and environmental influences on accident severity [3]. The study found Random Forest and ensemble methods to be highly accurate in both binary and multiclass severity predictions, recommending their use in traffic safety and risk management efforts. The study highlights the global issue of road accidents and their impacts, aiming to reduce accident occurrences and mitigate risks by providing real-time risk information. Using data from various sources, including accident reports and weather conditions[4].

KEYWORDS: road accident, machine learning, K-Nearest Neighbours, Decision tree, XGBoost, AdaBoost

I. INTRODUCTION

Road traffic accidents (RTAs) are a major global concern, leading to significant fatalities, injuries, and economic losses annually. Despite numerous efforts to improve road safety, accurately predicting and mitigating the impact of RTAs remains challenging due to the multifaceted nature of influencing factors, such as road conditions, weather, driver behavior, and vehicle characteristics. [1][2]

Machine learning (ML) offers a viable strategy for improving road safety by enabling predictive models that analyze various accident-related data. These models leverage data from historical accident records, traffic conditions, and environmental factors to assess the likelihood and severity of accidents. Methods like Decision Trees, Random Forest, Naive Bayes, AdaBoost, Considering the widespread use of deep neural networks in these studies to classify accident severity and identify high-risk areas. Some studies also integrate real-time data sources, such as CCTV footage and weather sensors, to improve prediction accuracy.[1][2][3]

This research aims to employ ML algorithms to predict accident severity, aiding emergency responders, policymakers, and drivers in taking preventive measures. By evaluating the performance of different algorithms, This study offers insightful information about the factors impacting accident severity and supports the evolution of data-driven strategies to enhance road safety[1][2][3]

II. RELATED WORKS

Methodology

1. Machine Learning Algorithms

- Among the algorithms utilized are Tree of Decisions Tree, AdaBoost, K-Nearest Neighbors (KNN), Naïve Bayes, and Deep Neural Networks (DNN) Data Cleaning: Handle missing values, outliers, and inconsistencies in the

dataset. [1][2]

- Studies compared single models like Logistic Regression and KNN with group techniques such as random Forest and XGBoost, noting that ensemble models generally offer higher accuracy
2. Image Processing Techniques
 - Methods for accident detection involve using computer vision and processing of age, including the extraction of features
 - and vehicle tracking using techniques like Mask R-CNN and Centroid Tracking.[3]
 3. Clustering and Classification Algorithms
 - Clustering algorithms group accident data based on factors like location, time, and severity, while classification algorithms assign severity levels such as fatal or minor injury.[4]

Datasets Used

1. Government and Open Data Sources :The studies leverage public datasets from government sources, which include traffic flow, weather conditions, and accident histories. These datasets are valued for their reliability and comprehensiveness Model Training: Train models on labelled data, adjusting hyperparameters to optimize performance.
2. Social Media and Real-Time Data: Some approaches explore using unstructured social media data to supplement traditional sources. Real-time data from sensors, such as traffic cameras and GPS, are also incorporated for dynamic analysis
3. Video Surveillance Data: For computer vision-based accident detection, datasets consist of surveillance footage from intersections, capturing vehicle collisions under various conditions

Performance Evaluation Metrics

1. Accuracy: The accuracy of prediction models was commonly evaluated. Random Forest achieved the highest accuracy in predicting accident severity, followed by XGBoost and AdaBoost for both binary and multiclass classifications.
2. Detection and False Alarm Rates: For accident detection using video analysis, metrics like the detection rate and false alarm rate are used. A detection rate of 71% with a false alarm rate of 0.53% was achieved in one study.
3. Precision, Recall, and F1 Score: These metrics were used to assess classifier performance, especially in distinguishing between different severity levels. AdaBoost showed improvements in precision and F1 score when compared to other algorithms.

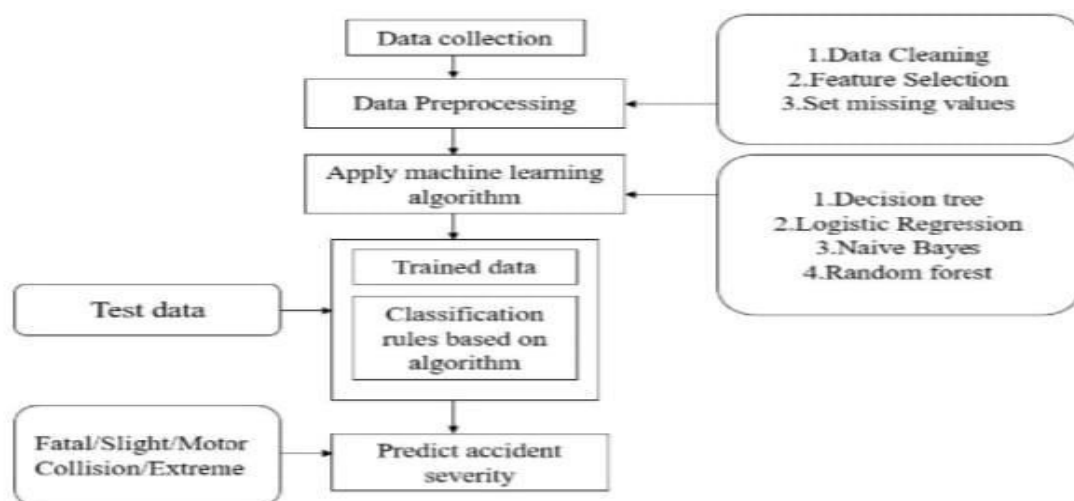


Fig 2.1: Proposed Model

III. CHALLENGES AND LIMITATIONS

1. Data Complexity and Quality
 - Road accident prediction models require large, diverse, and high-quality datasets, including details on driver behavior, road conditions, vehicle specifics, and environmental factors.
 - However, data often suffer from inconsistencies, missing values, or incomplete entries, making accurate prediction challenging.
2. Feature Selection and Engineering
 - Selecting relevant features is crucial for model accuracy. Factors like road geometry, driver age, and weather conditions are vital but challenging to model accurately due to their variability and non-linearity.
 - The inclusion of human factors like alcohol or drug use, often underreported, is essential for reliable prediction but difficult to obtain.
3. Algorithm Limitations
 - While ensemble models like Random Forest and AdaBoost perform well, they still face issues with high-dimensional data and may overfit if not carefully tuned.
 - Additionally, balancing prediction accuracy across different severity classes (e.g., fatal vs. minor) is complex, often requiring model-specific adjustments
4. Real-Time Prediction Constraints
 - Integrating real-time data, such as traffic conditions and weather updates, is vital for dynamic prediction it necessitates a large amount of computing power and infrastructure, making it challenging to implement on a large scale.
5. Ethical and Privacy Concerns
 - models for machine learning that use sensitive data including driver identity or personal behavior, face privacy concerns and potential misuse, which require strict adherence to data privacy laws and ethical guidelines.
6. Generalizability
 - Models trained on data from specific regions may not generalize well to different geographic areas due to variations in traffic laws, road conditions, and driving behavior. This limits the scalability of a single model for diverse regions.
7. Implementation and Cost
 - Deploying machine learning solutions at a national or regional level involves substantial costs for data collection, processing, and infrastructure maintenance, which can be prohibitive, especially for developing regions

IV. CONCLUSION

In conclusion, these studies on road accident prediction and severity analysis emphasize the critical role of machine learning in enhancing road safety. By employing a range of machine learning models, such as Decision Trees, AdaBoost, Deep Neural Networks, and Random Forest, researchers aim to accurately predict the likelihood and severity of accidents based on factors including road conditions, weather, traffic flow, and driver behavior. The research highlights the value of ensemble methods, such as AdaBoost and Random Forest, which often outperform single-model approaches by leveraging multiple weak classifiers for improved accuracy.

Predictive models enable more effective allocation of emergency resources and facilitate the identification of high-risk areas, allowing authorities to take preventative actions. The findings underscore that integrating real-time data sources, such as traffic cameras and weather sensors, could further enhance predictive accuracy, enabling more proactive accident prevention. Overall, these studies demonstrate the potential for machine learning to drive impactful improvements in traffic management and public safety by providing data-driven insights for reducing accident risks.

REFERENCES

- [1] Hemalatha, M., & Dhuwaraganath, S. "Road Accident Prediction Using Machine Learning." International Journal of Scientific Research in Science and Technology, 2024. This paper describes using deep learning to predict road accidents, analyzing various factors like speed, traffic conditions, and weather
- [2] Jadhav, A., Jadhav, S., Jalke, A., & Suryavanshi, K. "Road Accident Analysis and Prediction of Accident Severity Using Machine Learning." International Engineering and Technology Journal, 2020. The study explores AdaBoost and neural networks to classify accident severity into categories such as fatal and simple injury.
- [3] Arvind & Dr. M. Charles Arockiaraj. "Road Traffic Accident Severity Prediction Using Machine Learning." The International Journal of Advanced Research in Innovative Ideas in Education, 2024. This paper evaluates algorithms like Random Forest, XGBoost, and AdaBoost, emphasizing the impact of human factors such as alcohol use and age on accident severity
- [4] R. Arunachalam, S. Peararulsevi, M. Saraswathi, M. Subasri. "Road Accident Severity Prediction using Machine Learning." The International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), vol. 3, issue 8, April 2023, doi:10.48175/IJARSCT-9629, Impact Factor: 7.30
- [5] V. Manisha, S. Bharti and C. Naveen K, "Comparative Study of Machine Learning Algorithms for Traffic Accident Prediction and Prevention," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-5, doi: 10.1109/ICCCNT61001.2024.10723916.
- [6] V. B., V. S. Prakash and S. B. N., "Road Traffic Accident Prediction in India Using Machine Learning Algorithm Techniques," 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC), Bengaluru, India, 2024, pp. 1-5, doi: 10.1109/ICECCC61767.2024.105939



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