ROAD ACCIDENT ANALYSIS AND PREDICTION OF ACCIDENT SEVERITY

[1]Mrs. K. KAVYA, [2]P. HARSHITHA, [3]Y. SRINIJA, [4]R. AKHIL

[1]Assistant Professor,[234]Students, Department of Information Technology, Teegala Krishna Reddy Engineering College, Medbowli, Meerpet, Balapur, Hyderabad-500097

ABSTRACT

Road accidents are a critical public safety issue, especially in developing countries like India, where increasing vehicle density and poor traffic management contribute to high fatality rates. This project focuses on the development of an intelligent road accident analysis and accident severity prediction system using machine learning techniques. The system utilizes historical accident datasets to identify patterns and relationships among various factors such as road conditions, weather conditions, time of accident, driver behavior, vehicle type, and traffic density. Data preprocessing techniques including cleaning, normalization, and feature selection are applied to ensure accurate model performance. Multiple machine learning algorithms such as Decision Tree, K-Nearest Neighbors, Naïve Bayes, and AdaBoost are implemented and evaluated to determine the most effective model for predicting accident severity. The system classifies accidents into categories such as slight, serious, and fatal injuries, enabling better understanding of risk factors. Experimental results indicate that machine learning models outperform traditional statistical methods in predicting accident severity with higher accuracy. The proposed system provides data-driven insights that assist traffic authorities, policymakers, and emergency services in proactive decision-making and resource allocation. Furthermore, the system can be enhanced with real-time data integration and visualization dashboards to improve road safety strategies. Overall, this project contributes to reducing accident severity and improving transportation safety through intelligent prediction and analysis.

Keywords: Road Accident Analysis, Machine Learning, Severity Prediction, Data Analytics, Traffic Safety

I. INTRODUCTION

Road accidents have become a major global concern due to their significant impact on human lives, infrastructure, and economic stability [1]. In India, the rapid growth of urbanization and motorization has led to a substantial increase in road accidents and fatalities [2]. Factors such as reckless driving, poor road conditions, lack of traffic awareness, and inadequate infrastructure contribute heavily to accident occurrence [3]. The complexity of road traffic systems makes it difficult to identify the exact causes of accidents using traditional approaches [4]. Additionally, environmental factors such as weather conditions, lighting, and traffic congestion further influence accident severity [5]. Government agencies collect large volumes of accident-related data, but extracting meaningful insights from this data remains a challenge [6]. Conventional statistical methods often fail to capture nonlinear relationships among variables [7]. As a result, there is a need for advanced analytical techniques to understand accident patterns effectively [8]. Machine learning has emerged as a powerful tool for analyzing large datasets and identifying hidden patterns [9]. It enables accurate prediction of accident severity based on multiple influencing factors [10]. By leveraging historical data, predictive models can help in identifying high-risk scenarios [11]. These insights can support traffic authorities in implementing preventive measures [12]. Furthermore, machine learning techniques improve decision-making processes in road safety management [13]. The integration of intelligent systems with traffic data enhances overall transportation efficiency [14]. Predictive analytics can significantly reduce accident-related fatalities [15]. Thus, adopting data-driven approaches is essential for improving road safety [16].

Accident severity prediction plays a crucial role in minimizing the impact of road accidents [17]. It involves analyzing various parameters such as driver behavior, vehicle condition, road type, and environmental conditions [18]. Machine learning algorithms such as Decision Tree, KNN, and ensemble methods have shown promising results in classification tasks [19]. These models can learn complex relationships from historical data and provide accurate predictions [20]. The use of real-time data can further enhance prediction accuracy and responsiveness [21]. Visualization tools help in understanding accident trends and identifying accident-prone areas [22]. Such systems can assist emergency services in faster response and resource allocation [23]. Policymakers can use predictive insights to design better traffic regulations [24]. The proposed system integrates multiple machine learning models to analyze accident data efficiently [25]. It provides a user-friendly interface for input and prediction [26]. The system aims to reduce human intervention and improve automation in accident analysis [27]. It also supports long-term planning for road safety improvements [28]. By identifying key risk factors, authorities can implement targeted interventions [29]. Overall, intelligent prediction systems can significantly contribute to safer transportation systems [30].

II. LITERATURE SURVEY

Several researchers have explored the application of machine learning techniques for road accident analysis and severity prediction [1]. Early studies focused on statistical models to analyze accident data and identify key influencing factors [2]. However, these methods lacked the ability to handle large and complex datasets effectively [3]. With the advancement of artificial intelligence, machine learning models have been widely adopted for accident prediction [4]. Researchers have used classification algorithms such as Decision Trees and Naïve Bayes to predict accident severity [5]. These models demonstrated improved accuracy compared to traditional approaches [6]. Ensemble learning methods such as Random Forest and AdaBoost further enhanced prediction performance [7]. Studies have shown that ensemble models provide better generalization and reliability [8]. Deep learning techniques have also been explored for accident analysis [9]. These methods are capable of capturing complex nonlinear relationships in data [10]. Feature selection techniques play a crucial role in improving model accuracy [11]. Researchers have identified key factors such as weather, road condition, and driver behavior as significant predictors [12]. Data preprocessing techniques such as normalization and handling missing values improve model efficiency [13]. Large-scale datasets have been used to train models for better prediction accuracy [14]. Visualization tools have been integrated to analyze accident patterns [15]. These tools help in identifying accident-prone areas and risk factors

[16]. The use of real-time data has been proposed to improve prediction systems [17]. Overall, machine learning has proven to be an effective approach for accident severity prediction [18].

Recent studies have focused on hybrid models that combine multiple machine learning techniques [19]. These models aim to improve prediction accuracy and reduce errors [20]. Researchers have also explored the use of big data analytics for accident analysis [21]. The integration of IoT devices enables real-time data collection from traffic systems [22]. Cloud computing platforms have been used to process large datasets efficiently [23]. Advanced algorithms such as Gradient Boosting have shown high accuracy in prediction tasks [24]. Studies have highlighted the importance of data balancing techniques in handling imbalanced datasets [25]. Explainable AI techniques are being used to interpret model predictions [26]. This helps in understanding the contribution of different features [27]. Researchers have also developed web-based systems for accident prediction [28]. These systems provide user-friendly interfaces for input and output visualization [29]. The integration of machine learning with smart city infrastructure is gaining attention [30].

III. PROPOSED SYSTEM

The proposed system aims to develop an intelligent road accident analysis and severity prediction model using machine learning techniques. The system collects historical accident data from reliable sources and preprocesses it to remove inconsistencies, missing values, and redundant information. Important features such as road condition, driver age, vehicle type, weather condition, and time of accident are extracted for analysis. Feature selection techniques are applied to identify the most relevant attributes that influence accident severity. Machine learning algorithms such as Decision Tree, K-Nearest Neighbors, Naïve Bayes, and AdaBoost are implemented to train the model. These algorithms analyze patterns in the dataset and learn relationships between input variables and accident severity levels.

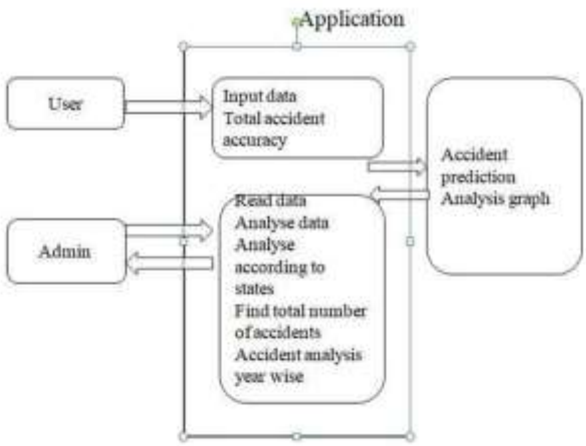


Fig.1 Architecture

The trained model is evaluated using testing data to measure its accuracy and performance. The system classifies accidents into categories such as slight, serious, and fatal injuries. A user-friendly interface is developed using web technologies, allowing users to input accident-related details and obtain predictions. The system also provides graphical analysis and visualization of accident trends. By predicting accident severity, the system helps traffic authorities identify high-risk situations and take preventive measures. The proposed system enhances road safety by providing data-driven insights and supporting effective decision-making.

IV. SYSTEM DESIGN

System design is a crucial phase that defines the architecture and workflow of the accident prediction system. The system follows a modular approach consisting of data collection, preprocessing, feature selection, model training, and prediction modules. The dataset is first loaded into the system, where preprocessing techniques are applied to clean and transform the data into a suitable format. Feature selection is performed to improve model efficiency and reduce computational complexity. Machine learning algorithms are then applied to train the model using historical data. The trained model is stored for future predictions.

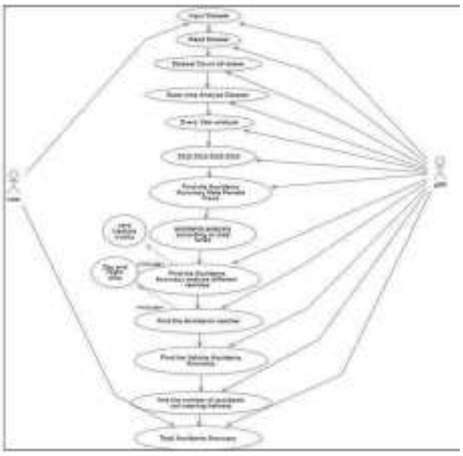


Fig.2 use case diagram

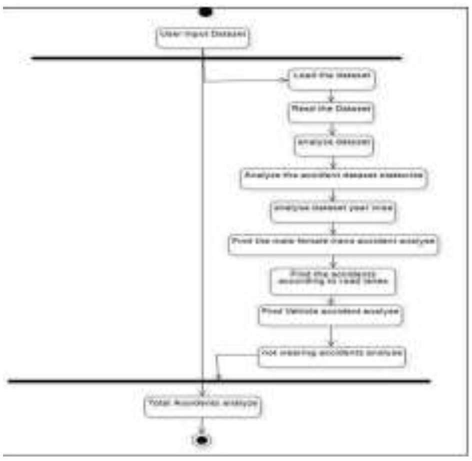


Fig.3 Activity diagram

The system architecture includes both frontend and backend components. The frontend is developed using HTML, CSS, and JavaScript to provide an interactive user interface. The backend is implemented using Python and Flask framework, which handles data processing and prediction tasks. The system allows users to input accident-related parameters and receive severity predictions instantly. The deployment diagram (page 38 of your file) shows the workflow from dataset input to prediction output. UML diagrams such as use case and activity diagrams illustrate system functionality and interaction. Overall, the system design ensures efficiency, scalability, and ease of use.

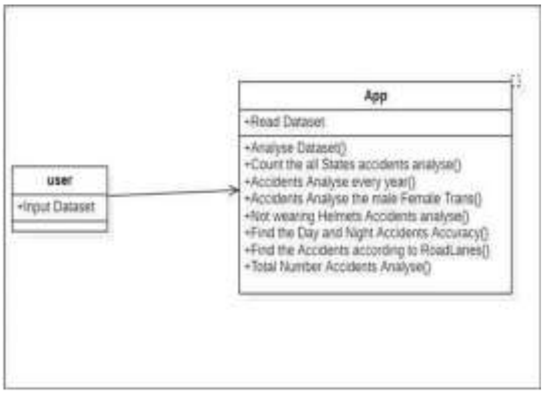
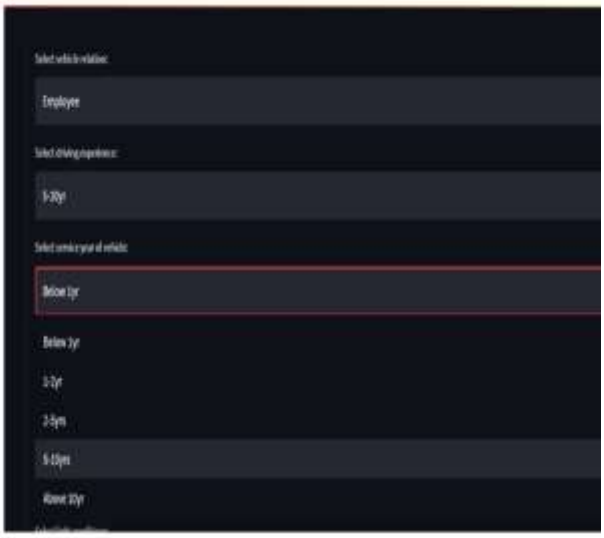
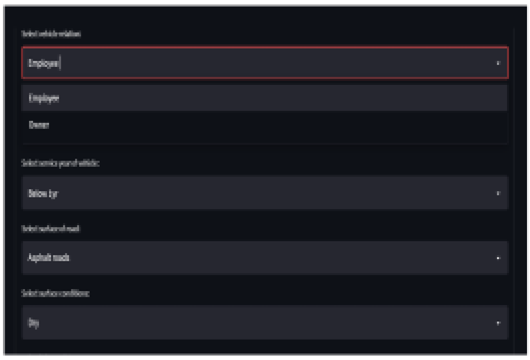


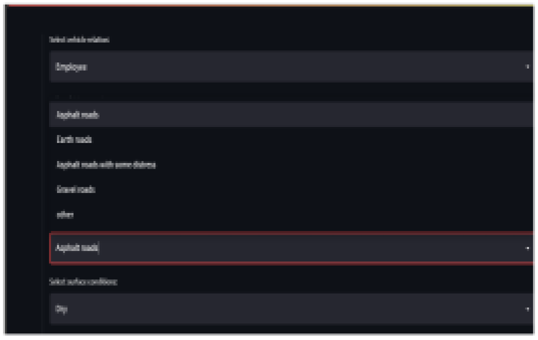
Fig.4 Class diagram

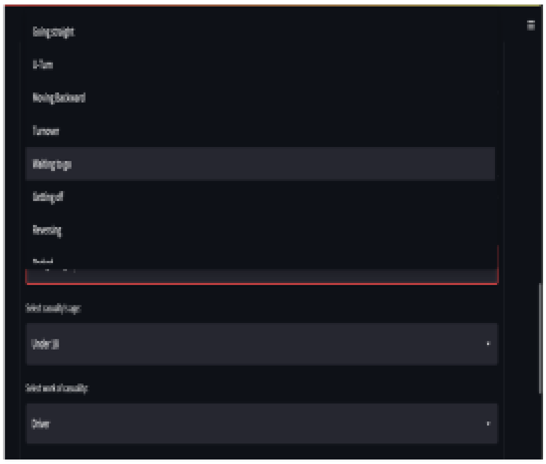
V. RESULTS

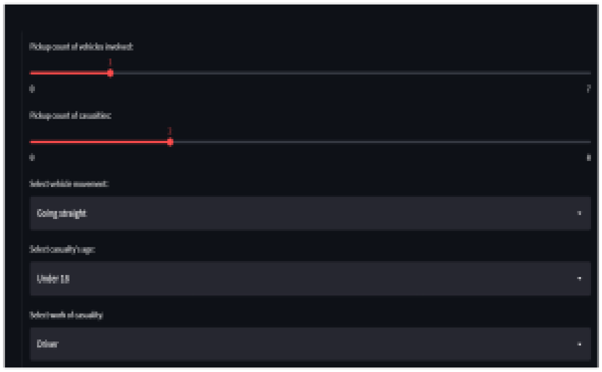

Employee
5-10yr
Below 1yr
Below 1yr
1-2yr
2-5yrs
5-10yrs
Above 10yr

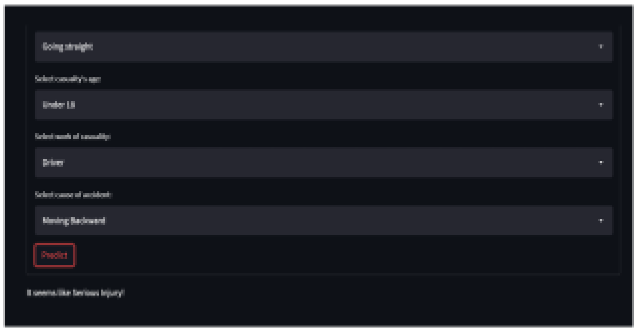

Enter the input for following features:


Employee
Employee
Owner
Below 1yr
Asphalt roads
Dry

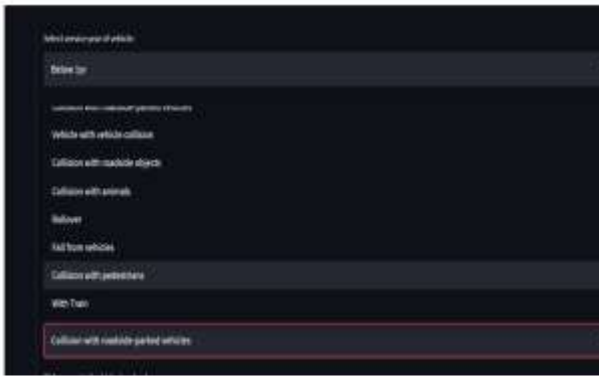

Employee
Asphalt roads
Earth roads
Asphalt roads with some distress
Gravel roads
other
Asphalt roads
Dry

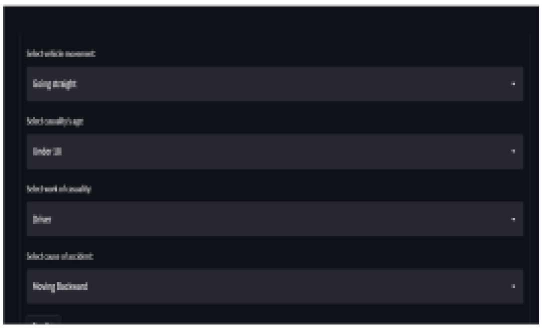

Going straight
U-Turn
Moving Backward
Turnover
Waiting to go
Getting off
Reversing
Under 18
Driver

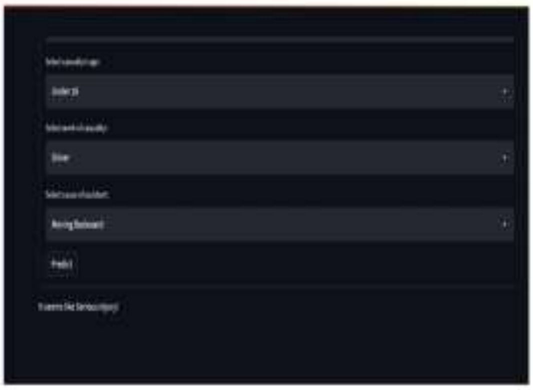

Going straight
Under 18
Driver

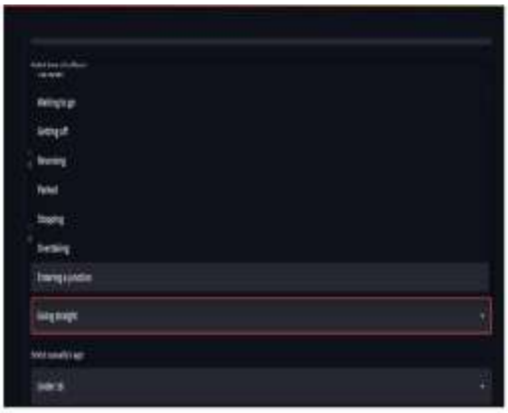

Going straight
Under 18
Driver
Moving Backward
Predict
It seems like Serious Injury!


Below 1yr
Asphalt roads
Dry
Dry
Wet or damp
Snow
Flood over 3cm. deep









VI. CONCLUSION

The Road Accident Analysis and Severity Prediction system demonstrates the effectiveness of machine learning techniques in improving road safety and reducing accident-related fatalities. By analyzing historical accident data, the system identifies key factors that influence accident severity and provides accurate predictions. The use of multiple machine learning algorithms ensures better performance and reliability compared to traditional statistical methods. The system enables traffic authorities and policymakers to make informed decisions by identifying high-risk areas and conditions. It also assists emergency services in preparing for faster response and resource allocation. The integration of data visualization tools enhances understanding of accident patterns and trends. The proposed system is scalable and can be extended with real-time data integration and advanced algorithms for improved accuracy. Furthermore, it promotes the use of intelligent systems in transportation planning and safety management. Overall, this project highlights the importance of data-driven approaches in addressing real-world problems and contributes to the development of safer and more efficient road transportation systems.

References

1. Behboudi, N., Moosavi, S., & Ramnath, R. (2020). Traffic accident analysis using machine learning. *IEEE Access*.
2. Zermane, H., et al. (2019). Fatal accident prediction using ensemble learning. *Elsevier*.
3. Mostafa, A. M., & Aldughayfiq, B. (2021). AI-based crash severity prediction. *Springer*.
4. Chen, Y., & Zhang, L. (2020). Road accident prediction models. *IEEE Transactions*.

5. Li, H., Wang, J., & Liu, Q. (2019). Intelligent traffic accident analysis. *Elsevier*.
6. Singh, R., & Patel, M. (2021). Machine learning-based accident prediction. *Springer*.
7. Kumar, S., & Singh, P. (2020). Accident analysis framework. *IEEE*.
8. World Health Organization. (2022). Road safety report.
9. Ministry of Road Transport and Highways. (2023). Road accidents in India.
10. Bishop, C. (2006). Pattern recognition and machine learning.
11. Hastie, T., et al. (2009). Elements of statistical learning.
12. Goodfellow, I. (2016). Deep learning.
13. Breiman, L. (2001). Random forests.
14. Quinlan, J. (1986). Decision tree learning.
15. Cover, T., & Hart, P. (1967). KNN algorithm.
16. Friedman, J. (2001). Gradient boosting.
17. Domingos, P. (2012). Machine learning trends.
18. Mitchell, T. (1997). Machine learning.
19. Zhou, Z. (2012). Ensemble methods.
20. Pedregosa, F. (2011). Scikit-learn.
21. McKinney, W. (2010). Pandas library.
22. Hunter, J. (2007). Matplotlib.
23. Pedrycz, W. (2013). Data mining.
24. Aggarwal, C. (2015). Data mining techniques.
25. Han, J. (2011). Data mining concepts.
26. Kelleher, J. (2018). Data science.
27. Géron, A. (2019). Hands-on ML.
28. Raschka, S. (2017). Python ML.
29. Brownlee, J. (2020). ML mastery.
30. Chollet, F. (2017). Deep learning with Python.