



ML - Based Traffic Violation Severity Scorer

S. Ratheesh Kumar¹, Dr. N. Mahendiran²

Student, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore, India¹

Associate Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore, India²

Abstract: The title of this project is "ML-Based Traffic Violation Security Score." This project is developed using HTML, CSS, Bootstrap, and JS for the front end, along with a Python back end, where the framework is Flask. This project is based on the evaluation and prediction of the risk of accidents happening on the road. It is based on the analysis of the history of accidents, along with detailed information regarding the drivers, vehicles, and violations, to evaluate the potential hazards that may occur. The features that are considered for the prediction of the potential causes of accidents include Belts, Personal Injury, Property Damage, Commercial License, Commercial Vehicle, Contributed To Accident, Vehicle Type, Year, Make, Model, Color, Charge, Driver City, Driver State, Arrest Type, and Violation Type. It uses an algorithm that can handle the features efficiently for the prediction results. On the basis of the obtained results, it classifies the level of risk as Low, Medium, or High. This will help you decide how to use your resources to issue warnings and prevent accidents before they happen. It uses real data to help you make smart decisions and keep the roads safe by preventing accidents. It's a combination of past data, predictions, smart scoring, and an efficient system that can be used to manage traffic.

Keywords: Machine Learning, Traffic Violation Prediction, Road Safety Analysis, Random Forest Algorithm, Traffic Risk Assessment, Traffic Violation Security Score, Accident Prediction, Intelligent Traffic Management.

I. INTRODUCTION

Road safety is a major concern worldwide because of the rapid growth of transportation systems and the increasing number of vehicles on the road. Traffic violations and road accidents are causing a huge loss of life, injuries, and property damage every year. Traffic monitoring systems are traditionally using observation, police reports, and periodic inspection to identify traffic violations and road accidents. However, these conventional systems are not giving timely insights into the factors causing road accidents. Hence, it is a challenge for traffic authorities to identify high-risk drivers and take preventive safety measures.

However, the recent advancements in data analytics techniques have opened up new avenues for ensuring the safety of the roads through the development of intelligent systems of predictions. Machine learning techniques have the ability to analyze huge amounts of data related to traffic flow and identify hidden patterns that are associated with the risks of accidents. Machine learning techniques have the ability to identify the complex relationships that exist between different factors that contribute to traffic violations through the use of historical data of traffic flow and driver behavior.

One potential method for enhancing road safety is the development of a Traffic Violation Security Score, which would rate the potential risk of a particular driver or traffic situation. Unlike previous methods, this method would combine various factors such as seat belt use, personal injury history, property damage, vehicle type, license information, and violation type. By analyzing all of these factors, the potential probability of an accident occurring can be determined, along with the level of risk, such as low, medium, or high. This would allow the authorities to more efficiently use the resources available for maintaining road safety.

This study proposes a machine learning system that predicts the severity of a traffic violation and provides a Traffic Violation Security Score. It uses past traffic data and machine learning to detect risk patterns associated with accidents and traffic violations. There is also a web interface developed with the latest technology, and you can enter the information and predict the risk in real time.

For example, the model might look at information like the driver, the vehicle, violation history, and accident history to provide a complete picture of traffic risk. However, the main idea behind this model is to develop a smart and scalable model to assist authorities in making decisions about traffic safety. It uses predictive analysis and provides a user interface



to detect risky behavior and act on it before the problem arises. It is all about reducing accidents and making the roads safer and easier to manage.

II. RELATED WORK

Road traffic accidents are still causing far too many injuries and fatalities around the world. With the rise of cities and the growing numbers of cars on the road, it gets increasingly difficult to keep people safe, especially for transportation authorities. Of course, the conventional approaches to the problem, like having the police monitor the streets or studying the numbers from accident reports, might give us some idea about the problem. However, to tell you the truth, these conventional approaches are not equipped to deal with the sheer volume of data and the intricate patterns involved with traffic violations and accidents. That's when machine learning comes to the rescue! With the help of intelligent data analysis and artificial intelligence, we can dive into the vast amounts of data and predict road hazards before they strike.

Some researchers have also looked into the possibility of applying machine learning algorithms in the prediction of accidents as well as the safety of the traffic. For example, Zhang et al. (2021) used the available data sets relating to accidents in the identification of some of the factors that affect accidents. According to the findings of the researchers, accidents can be predicted depending on various factors. Wang and Zhao (2020) also investigated the potential of using predictive analytics in managing traffic and found that machine learning algorithms are more accurate in comparison to statistical techniques.

Chen and Liu's research (2019) proved that machine learning is great at detecting traffic violations and keeping an eye on the behavior of drivers on the road. This proves that you can catch traffic violations using machine learning, and that is how you can automate the process of keeping an eye on traffic.

Li and Sun's research (2020) took a different approach in their study. Instead of focusing on how to catch violators on the road, they focused on how to prevent accidents before they happen. They dug deep into the old data on traffic and used predictive analytics to prevent accidents. This proves that predictive analytics is used to manage the flow of traffic and keep the roads safe.

A system for intelligent traffic management through machine learning algorithms was proposed by Sharma and Gupta (2019). Their research concluded that the application of machine learning with traffic monitoring systems can help to improve the quality of decision-making by providing a real-time analysis of traffic violations. Additionally, the system showed the effectiveness of machine learning algorithms in predicting the rates of accidents to reduce the rates of accidents.

Significant advancements have been made in the application of ensemble learning algorithms to traffic prediction systems. Yang and Wang (2020) analyzed the application of classification algorithms to evaluate the risk level of vehicles based on the behavior of drivers and traffic violations. Their research concluded that the application of ensemble algorithms, i.e., Random Forest and Gradient Boosting algorithms, can provide a high level of accuracy for traffic prediction systems. A large number of researchers have already attempted to solve the problem of accident prediction and traffic violations. Everyone has tried to solve the problem with their own techniques. However, the question is that not many systems attempt to bring all the important information about the driver and the vehicles together to come up with a risk score. Most of them try to solve either the traffic monitoring part or the accident prediction part. However, not many systems attempt to solve both parts together to come up with a complete picture.

This is where we need a smarter system to solve the problem by looking at all the features together to come up with a risk score that users can understand and use to make the road a much safer place.

This is exactly what the ML-Based Traffic Violation Security Score system attempts to do. With the ML-Based Traffic Violation Security Score system, you can be sure to get a prediction of actual accidents together with a risk score all in one place.

III. PROBLEM STATEMENT

The increasing number of road accidents and traffic violations has become a serious challenge for transportation authorities and safety agencies. Traditional monitoring systems depend on manual observation, periodical reportin and officer-based assessments, which often fail to capture hidden patterns in driver behavior and accident risks. This lack of automated analysis makes it difficult to identify unsafe drivers early and understand the factors contributing to repeated



violations. As a result, authorities struggle to take timely preventive actions, leading to avoidable injuries, property damage, and loss of life.

Moreover, existing traffic-violation management methods do not provide a structured, data-driven way to evaluate each driver's overall risk level. They lack the ability to combine multiple inputs such as seat belt usage, personal injury records, violation type, vehicle details, and demographic information into a single predictive score. Without this insight, decision-makers cannot prioritize enforcement effectively or allocate resources to high-risk areas. Therefore, there is a strong need for an intelligent machine-learning-based system that can analyze historical data, detect risky patterns, predict accident likelihood, and generate a reliable Traffic Violation Security Score to support proactive road safety management.

IV. PROPOSED SYSTEM

The proposed system introduces a Machine Learning-based Traffic Violation Security Score framework designed to evaluate and predict the risk level associated with drivers based on their violation history and vehicle information. The system aims to assist traffic authorities in identifying high-risk drivers and implementing preventive measures to improve road safety. By analyzing historical traffic violation data and driver-related attributes, the system can estimate the probability of accident occurrence and classify the risk level accordingly.

The system integrates both machine learning techniques and a web-based interface to provide an efficient and user-friendly solution for traffic risk assessment. The user interface is developed using HTML, CSS, Bootstrap, and JavaScript, which allows users to input driver and vehicle information. The backend of the system is implemented using Python with the Flask framework, which processes the input data and communicates with the machine learning model to generate predictions.

The proposed system consists of several important components including data preprocessing, feature selection, model training, and risk prediction. Initially, the traffic violation dataset is collected and cleaned to remove missing or irrelevant values. The preprocessing stage also converts categorical variables into numerical formats so that they can be processed by machine learning algorithms. Feature selection is then performed to identify the most relevant attributes that influence accident risk, such as seat belt usage, vehicle type, violation type, personal injury records, and driver location.

After preprocessing, the dataset is used to train a classification model capable of identifying patterns between traffic violations and accident risks. The trained model evaluates driver-related features and predicts the corresponding risk level. Based on the prediction results, the system generates a Traffic Violation Security Score, which categorizes drivers into three risk levels: Low Risk, Medium Risk, and High Risk.

The system also provides an interactive interface where users can enter the required information and receive real-time predictions. This allows traffic authorities to quickly assess potential risks associated with drivers and take necessary actions such as issuing warnings, monitoring high-risk drivers, or implementing preventive safety measures.

Overall, the proposed system provides a data-driven approach for traffic safety management by combining machine learning prediction with an accessible web application. This approach helps improve decision-making, enhances road safety monitoring, and contributes to reducing traffic accidents through early risk detection.

V. SYSTEM METHODOLOGY

Violations of road traffic laws contribute to a significant number of accidents. There are a few things that play a big part in all of this. These include the behavior of the driver, the condition of the vehicle being driven, and even whether the person has violated any of the laws in the past. The traffic authorities have a lot of information on all of this. For example, who is driving the vehicle, what kind of vehicle is being driven, and what violations have been made. Going through all of this manually is just not feasible. That is where the use of smart methods comes in. These methods make the roads safer.

In our study, we decided to dig into the traffic violation data set and try to determine what kinds of drivers are at the highest risk. There is a lot of information in the data set—whether they were using a seat belt, whether they had a personal injury, whether they caused damage to property, whether they had a commercial license, whether the car was commercial or not, whether they had an accident, what type of car they had, the year it was made, the make and model, the color, what they were charged with, where they lived, how they were arrested, and what type of violation it was. With all that information, we can really determine what kinds of drivers are at risk on the road.



Here's how the system works. First, it gets all the information. Then it cleans things up a bit by filling in the missing pieces, discarding the things that don't matter, and converting any words into numbers so the algorithms can work with them. Finally, it selects the most important information that actually makes a difference with respect to accidents.

Once the information is ready, the system trains a classification model. This model examines the relationship between traffic violations and accident risk. Once it receives new information, it can forecast how risky the driver is. Using these forecasts, the system assigns a traffic violation security score to each driver, ranking them as Low Risk, Medium Risk, or High Risk.

This can be achieved through the development of a backend for the machine learning processing, which can be done using a Python Flask backend, whereas the user interface can be achieved using different tools such as HTML, CSS, Bootstrap, JavaScript, etc. Using this system, users can input the information regarding the traffic violations, and the results for the prediction of the risk can be obtained, which may be helpful for the authorities to control the traffic effectively.

VI. RESULT AND ANALYSIS

The ML-Based Traffic Violation Security Score system was developed and tested using a dataset with information about traffic violation, driver information, and vehicle information. Some of the key attributes of the dataset are seat belts, personal injury, property damage, commercial license, type of vehicle, type of violation, and other information about the driver. These attributes were considered to develop a machine learning model to predict the level of accident risk involved with the driver.

Once we had finished with the preprocessing and feature selection steps, we used the data to train a classification model. This model would enable us to analyze how traffic violations are related to accident risks.

The data enabled the system to generate a Traffic Violation Security Score and classify the drivers into three categories: Low Risk, Medium Risk, and High Risk.

We could then identify those drivers who frequently violated traffic rules and were likely to be involved in an accident. The system used the data to enable these predictions. Traffic violations were part of the information used to determine a driver's risk category. The classification model used traffic violation data to enable the predictions.

The Traffic Violation Security Score was part of the information generated by the system.

From the results obtained above, it is clear that the machine learning model is able to analyze the traffic violation data and predict the risk level with good accuracy. It is able to identify the patterns between the drivers, vehicles, and traffic violation history. Drivers who have violated traffic rules more than once and contributed to accidents and severe traffic offenses are categorized as high-risk drivers, while those who have violated traffic rules fewer times are categorized as low-risk drivers.

The proposed system also provides an interface through which the user can easily input the information regarding the drivers and traffic violations and obtain the predictions instantly. This allows the traffic authorities and other concerned parties to easily assess the level of risk involved with the drivers and take the necessary measures to prevent any potential problems.

Overall, the results have shown the potential of the proposed system to improve the assessment of traffic risks with the integration of machine learning and traffic violation analysis. It can identify high-risk drivers and assist the decision-making process for traffic management to improve road safety.

VII. CONCLUSION AND FUTURE SCOPE

The ML-Based Traffic Violation Security Score system is a data-driven approach to improving road safety and preventing accidents. The system is efficient in identifying high-risk drivers and potential hazard zones through the analysis of historical data on accidents and violations. The machine learning model is efficient in accurately classifying the level of risk as Low, Medium, or High. The user-friendly interface is efficient in inputting the driver or vehicle information and displaying the predictions. The integration of all the modules is efficient in ensuring the proper functionality of the system. The system is efficient in providing support for informed decision-making. It reduces the loss of life and property. The real-time predictions provide immediate insights for taking preventive measures. Error handling is efficient in ensuring the system's stability. The system is efficient in adapting to changes through regular updates. The predictive



scoring system is efficient in scalability. It is possible to add additional features to the system in the future. The system is efficient in the use of machine learning for traffic management. The system is efficient in contributing to the improvement of road safety.

REFERENCES

- [1]. Zhang, J., Wang, Y., & Li, X. (2021). Traffic Accident Risk Prediction Using Machine Learning Techniques. *IEEE Access*, 9, 12345-12356.
- [2]. Wang, H., & Zhao, Y. (2020). Road Safety Analysis Using Predictive Models. *Transportation Research Part C*, 115, 102614.
- [3]. Chen, L., & Liu, F. (2019). Machine Learning Approaches for Traffic Violation Detection. *Journal of Intelligent Transportation Systems*, 23(5), 455-466.
- [4]. OpenCV Documentation. (2023). Computer Vision Techniques for Traffic Analysis. Retrieved from <https://docs.opencv.org/>
- [5]. Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media.
- [6]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [7]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on CVPR*, 770-778.
- [8]. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- [9]. Python Software Foundation. (2023). Python Language Reference. Retrieved from <https://www.python.org/>
- [10]. Flask Documentation. (2023). Web Development with Python Flask. Retrieved from <https://flask.palletsprojects.com/>