AI-Driven Decision Support System for Accurate Traffic Accident Severity Prediction and Emergency Response Optimization

Narmatha A/P Kathiravan1, a), S Prabha Kumaresan 1, 2, b) and Mohammad Shadab Khan1, 3, c)

1Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

2Center of Natural Language Processing (NLP), CoE for Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

3Center of Advanced Analytics, CoE for Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Malaysia.

b) Corresponding author: prabha.kumaresan@mmu.edu.my   
a) narmatha.kathiravan@student.mmu.edu.my

c)shadab.khan@mmu.edu.my

**Abstract.** Traffic accidents pose a critical road safety challenge that requires proper traffic authorities and suitable prediction of accident severity for timely emergency response and resource deployment. Traditional statistical models often fail to capture complex, non-linear interactions among contributing factors, limiting their effectiveness in real-world applications. The current research uses the application of machine learning to develop decision-support for traffic authorities as well as emergency personnel to enable early, data-informed interventions. Three models-Random Forest, Gradient Boosting, as well as Logistic Regression, were utilized to forecast severity as minor, moderate, or serious in terms of causative influence from road conditions, weather, vehicle type, as well as driver characteristics. Logistic Regression offered the highest accuracy of 59.52%, while Random Forest offered better precision (64.66%) as well as F1-score (32.51%), indicating higher overall balance. Feature importance analysis indicated driver age, experience, as well as environmental factors to have consistently significant impacts across models. This research shows the potential for machine learning in providing emergency preparedness as well as optimizing road safety strategies which incorporate accurate assessment of severity.

# Introduction

Traffic accidents pose a significant issue globally, characterized by high fatalities, property loss, and losses in productivity. The ability to predict the potential severity of accidents based on contributing factors such as weather conditions, road conditions, type of vehicles, and the drivers themselves to enable the authorities to take steps to lower the risk and develop improved emergency responses. Here, we use machine learning approaches to evaluate prior traffic data to predict accident severity and provide useful data that can be utilized to improve road safety and management. The recent advancement of machine-learning approaches suggests their strength in identifying complex interrelationships between variables that enable us to better predict outcomes when compared to traditional methods. As a continuation of this work, the present research extended the methodology of using Random Forest, Gradient Boosting, and Logistic Regression models to classify accidents into minor, moderate, and serious injuries with improved accuracy and interpretability.

# Related works

Traffic accidents are a global problem and always cause human pain and economic loss. According to [1], every year, more than 1.35 million people are killed, and 20-50 million people are injured in traffic accidents, which translates into about 3,000 deaths every day on average and this high rate calls for predictive models that may reduce the level of accident severity. Machine learning usage would address relevant aspects of weather, road conditions, and driver behavior in improving emergencies and decision making. As stated in [2] knowing these contexts provides better safety practices and planning in the policy making space. Moreover, [3] points out that policymakers and transportation planners can mitigate accident severity through data-informed intervention.

Traffic accident severity prediction was initially based on statistical methods. For example [4], compare Artificial Neural Networks (ANN) and fuzzy Adaptive Resonance Theory (ART) for predicting crash severity and determined that both methods are effective. Research in [5] points out the dispersion of speed and reckless driving as both closely related with accidents and confirm road geometry as a related factor in rainy weather. Meanwhile [6] pointed out that while ordered logit and probit models are commonly used, they are limited by assumptions of data distribution. These early works laid a foundation for the acceptance of ML models, which provide better accuracy and the accommodation of accident severity complexity.

Recent developments in ML have significantly improved predictions of accident severity. Research work in [7] discuss ensemble models such as Random Forest (RF) and Gradient Boosting, which outperform traditional analytical methods due to their capacity to handle non-linear relationships. Similarly, [8] shows that hybrid models employing ensemble methods like stacking further enhance predictive capability. The study in [9] demonstrates that RF is robust in analyzing diverse features of weather, road conditions, etc. due to its high accuracy and recall. These studies offer insights into the potential of ML to revolutionize accident severity prediction.

Various methods have been adopted for effective prediction of accident severity. For instance, [10] highlighted Knowledge Discovery in Databases (KDD) as essential for data pre-processing, feature selection, and model evaluation. [4] integrate Recursive Feature Elimination (RFE) and SHAP values for feature selection, enhancing interpretability and performance. Meanwhile, [11] employs a Stacking Ensemble classifier and Select K Best to show how model combinations improve outcomes. Additionally [6] discuss challenges in managing missing values and imbalanced datasets, necessitating oversampling or under-sampling to boost model effectiveness. Many studies reiterate that systematic procedures and feature engineering are critical to good prediction.

Despite the significant advancements, certain limitations remain. As noted in [7], single classifiers often struggle with generalization, while ensemble methods tend to offer more reliable results. Further, [12] highlighted inconsistencies in RF performance across different accident types, suggesting the need for more context-aware models. The growing use of explainable AI tools such as SHAP and LIME, mentioned in [13], has improved transparency by clarifying feature importance, but larger datasets and more intensive modeling are still required to address ongoing challenges. Gaps in the literature continue to restrict the effectiveness of ML-based prediction models. Among the various ML algorithms tested, XG Boost and RF consistently emerge as top performers. In [14], XG Boost was praised for its superior accuracy, recall, and precision in multi-class classification tasks. Meanwhile, [15] reported that RF provides high accuracy and strong performance across individual classes. On the other hand, [16] showed that ANNs are particularly effective at capturing complex patterns in traffic data, though they are less commonly used due to issues with interpretability.

This review has discussed how predicting the severity of traffic accidents has progressed from statistical methods to advanced ML models. Early research laid the groundwork, while recent studies demonstrate the success of ensemble techniques and explainable AI. Yet, challenges like class imbalance, feature limitations, and lack of real-time data persist. This research seeks to address those challenges by developing an integrated ML model capable of accurately forecasting traffic accident severity to support safety improvements and emergency response optimization.

# Research methodology

## Data Collection

Historical traffic accident datasets were sourced from open-access platforms [17], to ensure accessibility and ethical compliance. The selected datasets contained a wide variety of features relevant to accident severity prediction, such as road surface conditions, weather factors, types of vehicles involved, traffic density, driver behaviors, and demographic information like driver age and experience. Preference was given to public datasets over real-time or proprietary data sources to overcome integration challenges and availability restrictions. The diversity presented in the datasets enabled the development of models that could generalize effectively across different accident scenarios and conditions.

## Data Exploration and Preprocessing

Once the data was acquired, EDA was performed to understand the structure, distribution, and quality of the datasets. Initial checks showed that the datasets had missing values in several features. Numerical missing values were handled by imputing the median value of the corresponding columns, while mode imputation was used to fill in missing categorical values so that they remained consistent. After imputation, validation checks were performed to confirm the absence of any remaining missing values. Categorical variables such as "Weather," "Road Type," and "Vehicle Type" were encoded into numerical values using Label Encoding, ensuring that each unique category was properly represented numerically without introducing artificial relationships. To normalize the numerical features and bring them onto a comparable scale, Standard Scaler was applied. This step was critical to prevent features with larger scales from dominating the learning process. Following preprocessing, the dataset was partitioned into training and testing subsets in an 80:20 ratio, using a fixed random seed to ensure reproducibility and consistency in model evaluation.

## Model Development

Three models were developed and trained to predict accident severity into three categories: minor, moderate, and serious. The models used were Random Forest, Gradient Boosting, and Logistic Regression. The models had to be ranked due to their resistance to overfitting, insensitivity to high-dimensional data, as well as intrinsic capability to carry out feature selection. The Gradient Boosting model was used to test the performance of an ensemble model with capability to learn complex non-linear relationships, whereas the model used as the baseline due to its simplicity as well as high interpretability was the Logistic Regression model. The preprocessing began with the development, followed by the models being trained separately using the training subset. The models then predicted the unseen test data to establish their generalizability. Extensive hyperparameter tunning was carried out where it was necessary to enhance performance while avoiding overfitting. The general process followed during this research, including all stages from data collection to model comparison as well as to ascertain the featured importance, is as shown in Figure 1.

A diagram of a data flow

AI-generated content may be incorrect.

**FIGURE 1.** Workflow diagram for traffic accident severity prediction

## Model Evaluation

Model performance was assessed using four core classification metrics: accuracy, precision, recall, and F1-Score. Accuracy quantified the proportion of correct predictions, precision measured the correctness of positive predictions, recall evaluated the model’s ability to capture actual positives, and the F1-Score offered a balanced measure between precision and recall.

## Feature Importance Analysis

Feature importance analysis was performed to determine the most important attributes that contribute to the predictions of accident severity. For the models Random Forest and Gradient Boosting, feature importance values were obtained from their respective feature importances attributes. For Logistic Regression, the absolute value of the model coefficients was used in determining the importance. The top features with the highest importance present in all the models were Driver Age, Driver Experience, Speed Limit, and Weather Conditions. The results conform to the existing literature since it reflects the importance of environmental as well as behavioral factors in the outcome of accidents. The relative importance of the features was visually represented graphically to encapsulate them in an effective manner.

## Visualization and Final Deliverables

Visualization played a crucial role throughout the study, assisting in the effective communication of results. Key graphical outputs included distributions of the target variable, per-class performance metrics, feature importance plots, and overall model comparisons. These visuals provide intuitive insights into model behavior and feature relationships. The outputs of this study comprised the fully developed Random Forest, Gradient Boosting, and Logistic Regression models, accompanied by detailed evaluation reports, visual analytics, and comprehensive documentation of the methodology and results. The findings are intended to contribute towards enhancing predictive capabilities in traffic accident severity classification and optimizing emergency response strategies.

# Results and discussions

## Model Implementation and Training

The analysis employed three classification algorithms: Random Forest, Gradient Boosting, and Logistic Regression to estimate the severity of traffic accidents. The Random Forest classifier was trained on a dataset encompassing road surface conditions, weather, vehicle type, traffic density, driver age, and driver experience. Prior to model fitting, missing numerical values were imputed with column medians and categorical entries with modes, label-encoded categorical features, and standardized numerical predictors. The cleaned data were partitioned into training and testing subsets in an 80:20 ratio. Training and prediction were carried out using the Scikit-learn library in Python. Gradient Boosting and Logistic Regression were trained similarly for comparison.

## Performance Evaluation

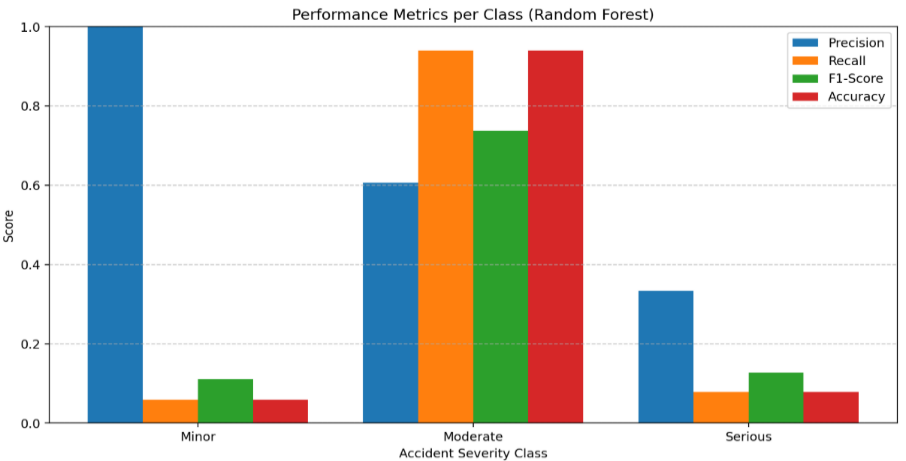
Predictive performance was assessed via accuracy, precision, recall, and F1-Score. Logistic Regression achieved 59.52% accuracy, with a precision of 21.00% and a recall of 33.33%. Gradient Boosting recorded 54.76% accuracy, with a precision of 41.22% and a recall of 34.56%. Random Forest achieved 58.93% accuracy, demonstrating balanced performance across severity classes with a precision of 64.66% and a recall of 35.91%. Figures 2 to 4 illustrate the metrics for each class, revealing that all models struggled with moderate-severity classes. This suggests future work might explore resampling or cost-sensitive learning to address class imbalance. The summarized performance metrics for all three models are presented in Table 1, with Figure 5 further illustrating the overall accuracy comparison, confirming Random Forest’s superior balance between predictive accuracy and consistency across classes.

## Feature Importance Analysis

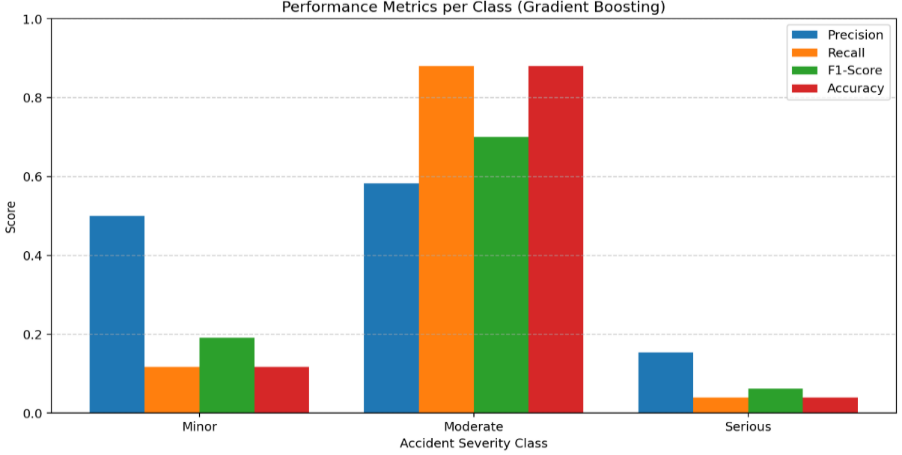
To interpret model decisions, feature importance was extracted from each algorithm. In Random Forest and Gradient Boosting, the feature importances attribute ranked the predictors, while in Logistic Regression, the absolute values of the coefficients were used. As illustrated in Figure 6, Driver Age and Driver Experience were consistently the most influential features across all models, followed by Speed Limit and Number of Vehicles. Environmental factors such as Weather Conditions and Road Surface Condition also contributed significantly to the predictions.

**TABLE 1.** Performance metrics of ML models for accident severity prediction

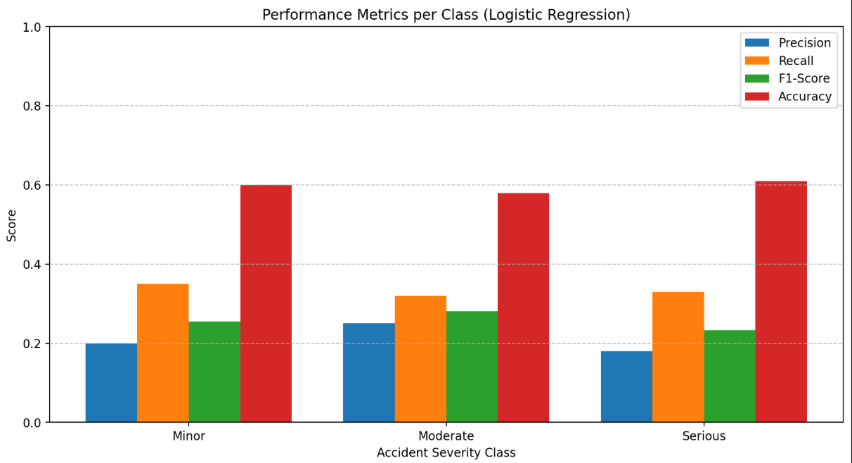
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | | **Accuracy (%)** | **Avg. Precision (%)** | **Avg. Recall (%)** | **Avg. F1-Score (%** |
| Random Forest | 58.93 | | 64.66 | 35.91 | 32.51 |
| Gradient Boosting | 54.76 | | 41.22 | 34.56 | 31.81 |
| Logistic Regression | 59.52 | | 21.00 | 33.33 | 25.61 |



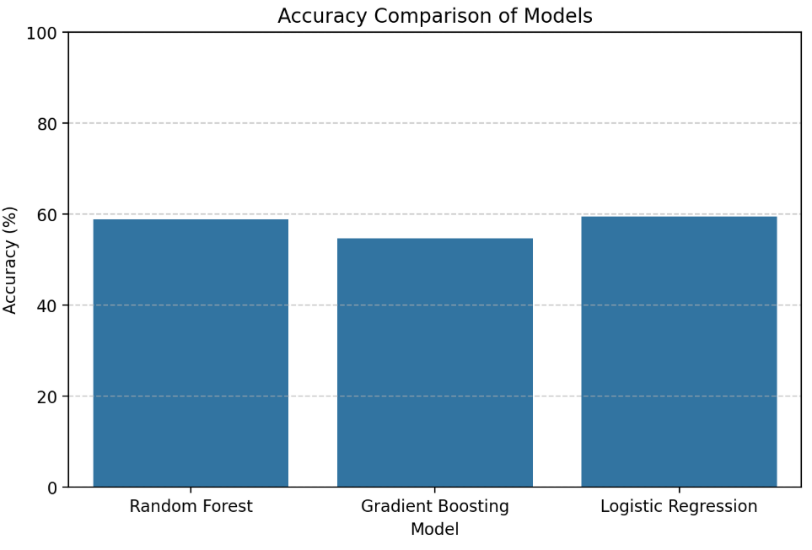
**FIGURE 2.** Performance metric**s** for each accident severity class as predicted by the Random Forest model



**FIGURE 3.** Performance metrics for each accident severity class as predicted by the Gradient Boosting model



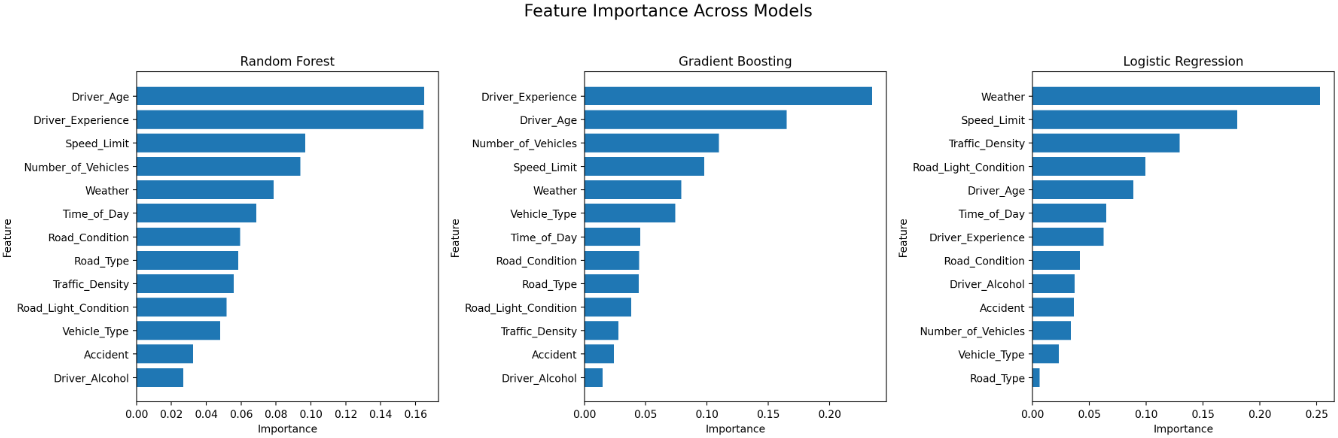
**FIGURE 4.** Performance metrics for each accident severity class as predicted by the Logistic Regression model.



**FIGURE 5.** Comparison of overall accuracy (%) achieved by Random Forest, Gradient Boosting, and Logistic Regression classifiers on the test set

**Overall Insights**

The superior accuracy and balanced performance of Random Forest confirm its robustness and suitability for real-world deployment. Logistic Regression had the benefit of being more interpretable, but also less flexible. The lower accuracy of Gradient Boosting suggests it may require more tuning, or enhancements in features. It remains a challenge to predict moderate-severity cases and indicates that data-level interventions need to be addressed. The key features identified illustrate the possibilities for decisions on safety policies, including the use of adaptive speed controls and driver awareness and engagement programs, for adverse weather.



**FIGURE 6.** Comparison of feature importance across three machine learning models

# CONCLUSION

The research effectively formulated the traffic accident severity prediction model using machine learning techniques. The significant contributions are effective data collection as well as preprocessing, complete model construction, as well as meaningful analysis in terms of features. Although the highest accuracy (59.52%) was provided by Logistic Regression, Random Forest provided enhanced overall performance with very high precision (64.66%) as well as F1-score (32.51%). These show a higher balance in predictive accuracy as well as dependability, particularly in the case of a scenario with imbalanced severity classes. Consequently, the Random Forest model was determined to be the best option for practical deployment. Future work will involve class imbalance reduction, real-time data stream integration, as well as improved exploration with the algorithm to further develop predictive capability as well as optimize emergency responses.

# References

1. Adefabi, S. Olisah, C. Obunadike, O. Oyetubo, E. Taiwo, and E. Tella, “Predicting Accident Severity: An Analysis of Factors Affecting Accident Severity Using Random Forest Model,” IJCI **12**(6), 107–121 (2023).
2. I. Aldhari, M. Almoshaogeh, A. Jamal, F. Alharbi, M. Alinizzi, and H. Haider, “Severity Prediction of Highway Crashes in Saudi Arabia Using Machine Learning Techniques,” Applied Sciences **13**(1), 233 (2022).
3. X. Hu, “Using machine learning models to forecast future prices – an exploration of investment strategies,” in 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), (IEEE, Chongqing, China, 2023), pp. 468–474.
4. T. Baykal, F. Ergezer, E. Eriskin, and S. Terzi, “Accident Severity Prediction in Big Data Using Auto-Machine Learning,” Scientia Iranica 0(0), 0–0 (2023).
5. M. Ghasedi, M. Sarfjoo, and I. Bargegol, “Prediction and Analysis of the Severity and Number of Suburban Accidents Using Logit Model, Factor Analysis and Machine Learning: A case study in a developing country,” SN Appl. Sci. **3**(1), 13 (2021).
6. Md.K. Islam, I. Reza, U. Gazder, R. Akter, M. Arifuzzaman, and M.M. Rahman, “Predicting Road Crash Severity Using Classifier Models and Crash Hotspots,” Applied Sciences **12**(22), 11354 (2022).
7. M. S. I. Mohd Zubil, Z. Che Embi, and K. I. Ghauth, “Assessing the Efficiency of Deep Learning Methods for Automated Vehicle Registration Recognition for University Entrance,” Journal of Informatics and Web Engineering **3**(2), 57–69 (2024).
8. T. Ahmed Khan, R. Sadiq, Z. Shahid, M. M. Alam, and M. Mohd Su’ud, “Sentiment Analysis using Support Vector Machine and Random Forest,” Journal of Informatics and Web Engineering **3**(1), 67–75 (2024).
9. S. K. Reddy Koduru, “Prediction of Severity of an Accident Based on the Extent of Injury using Machine Learning,” IJCTT **70**(9), 43–49 (2022).
10. K. Kodepogu, V. B. Manjeti, and A. B. Siriki, “Machine Learning for Road Accident Severity Prediction,” MITS **2**(4), (2023).
11. A. P. Kumar, and D. Teja Santosh, “Road Accident Severity Prediction Using Machine Learning Algorithms,” IJCERT **9**(9), 175–183 (2022).
12. J. Lee, T. Yoon, S. Kwon, and J. Lee, “Model Evaluation for Forecasting Traffic Accident Severity in Rainy Seasons Using Machine Learning Algorithms: Seoul City Study,” Applied Sciences **10**(1), 129 (2019).
13. J. P. S. S. Madushani, R. M. K. Sandamal, D. P. P. Meddage, H. R. Pasindu, and P. I. A. Gomes, “Evaluating expressway traffic crash severity by using logistic regression and explainable & supervised machine learning classifiers,” Transportation Engineering **13**, 100190 (2023).
14. Muktar, and V. Fono, “Toward Safer Roads: Predicting the Severity of Traffic Accidents in Montreal Using Machine Learning,” Electronics **13**(15), 3036 (2024).
15. M. Garba, U. Sharif, M. Danjumma, S. U. S.Noma, M. A. Usman, and M. A. Giro, “Analysis of Sokoto State Road Accident and Prediction of Accident Severity using Machine Learning Technique,” International Advanced Research Journal in Science, Engineering and Technology **11**(9), (2024).
16. J. J. Ng, K. O. M. Goh, and C. Tee, “Traffic Impact Assessment System using Yolov5 and ByteTrack,” Journal of Informatics and Web Engineering **2**(2), 168–188 (2023).
17. T. Almanie, “Quantitative Study of Traffic Accident Prediction Models: A Case Study of Virginia Accidents,” IJANA **14**(05), 5582–5589 (2023).