Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 16(1), 47-57, 2026
Journal of the Institute of Science and Technology, 16(1), 47-57, 2026

ISSN: 2146-0574, eISSN: 2536-4618

Computer Engineering **DOI: 10.21597/jist.1749946**

Research Article

Received: 25.07.2025 **Accepted: 02.09.2025**

To Cite: Sancar, Y. & Öztaş, S. (2026). Application of Random Forest and SMOTE in Predicting Fatal Traffic Accidents: The Case of Erzurum. *Journal of the Institute of Science and Technology*, 16(1), 47-57.

Application of Random Forest and SMOTE in Predicting Fatal Traffic Accidents: The Case of Erzurum

Yasin Sancar [1], Sinan Öztaş [2*]

Highlights:

- Random Forest
- SMOTE
- Prediction of fatal traffic accidents

Keywords:

- Traffic accident
- Machine learning
- Random Forest
- SMOTE

ABSTRACT:

This study presents a machine learning model developed using the Random Forest algorithm to predict whether traffic accidents in Erzurum will be fatal. Data from 16793 traffic accidents that occurred between 2014 and 2023, provided by the General Directorate of Security, was used. This dataset includes various variables such as driver characteristics, weather conditions, road type, road condition, lighting, shoulder, etc. Due to the minority of fatal accidents in the dataset, class imbalance was addressed using the SMOTE (Synthetic Minority Over-sampling Technique) method. The model was tested on training and test data with high performance metrics such as 98% accuracy, sensitivity, and F1 score. The results obtained reveal the impact of variables such as accident type, driver age, and number of vehicles on fatal accidents, contributing to data-driven policy development processes aimed at improving traffic safety.

[1] Yasin SANCAR (**Orcid ID:** 0000-0002-4200-1293), Atatürk University, Open Education Faculty, Erzurum, Türkiye

[2*] Sinan ÖZTAŞ (**Orcid ID:** 0000-0002-9630-6586), Atatürk University, Faculty of Engineering, Department of Industrial Engineering, Erzurum, Türkiye

***Corresponding Author**: Sinan ÖZTAŞ, e-mail: sinanoztas@atauni.edu.tr

INTRODUCTION

Traffic accidents continue to be a serious public health and public safety issue in Turkey, as they are throughout the world. Fatal traffic accidents, in particular, have a profound social and economic impact not only on individuals but also on society as a whole. Research conducted worldwide has revealed that driver errors, environmental conditions, and road characteristics play a significant role in the occurrence of traffic accidents (Khan & Hussain, 2024; Reddy et al., 2022). In Turkey, studies on this subject have provided important analyses from both spatial and statistical perspectives. For example, Özden and Caci (2018) examined the relationship between weather conditions and accidents in Adana province, while Yıldız and colleagues (2020) highlighted the impact of seasonal distribution on the severity of accidents.

Studies conducted using Geographic Information Systems (GIS) have determined that accidents are clustered in urban areas and busy intersections (Özden et al., 2018; Yavuz et al., 2021). Additionally, the impact of factors such as driver behavior, speed violations, and day-night variability on accidents was examined by Karayel and colleagues (2019) using logistic regression analysis. Analyses conducted on highways have determined that morning hours are more risky, emphasizing the importance of time-based traffic policies (Çetin et al., 2021).

In recent years, artificial intelligence and machine learning methods have come to the fore in predicting traffic accidents. Algorithms such as Random Forest, XGBoost, and Logistic Regression have been successful in predicting fatal accidents by evaluating many variables such as driver age, vehicle type, road type, weather conditions, and accident time (Tang et al., 2025; Cicek et al., 2023; Ersöz & Ersöz, 2022). In some studies, sampling techniques such as SMOTE have been used to address class imbalance issues, thereby improving model performance (Infante et al., 2022; Bokaba et al., 2022).

In this study, a classification model was developed to predict whether traffic accidents in Turkey will be fatal or not. The Random Forest algorithm was preferred in the modeling process. This algorithm is frequently preferred in the literature due to its high classification success and the comprehensibility of variable importance levels (Bedane et al., 2011; Liu et al., 2021).

Based on this study, a classification model was developed to predict fatal accidents using traffic accident data provided by the General Directorate of Security for the province of Erzurum. During the modeling process, six different machine learning algorithms were considered, and the Random Forest classification algorithm, which achieved the highest accuracy, was selected for use. This method is frequently used in the literature due to its high classification accuracy and the comprehensibility of variable importance levels.

The main objective of this study is to develop a machine learning model that can predict whether an accident will be fatal or not and to evaluate it in accordance with explainability principles. However, since the proportion of fatal accidents in traffic accident data is quite low, a serious class imbalance problem was encountered. To address this issue, the SMOTE method was applied to increase the number of minority class examples, and the class weighting (class_weight) strategy was also evaluated.

The model's performance was measured using metrics such as accuracy, recall, specificity, F1 score and ROC-AUC. The subsequent sections of the study include a literature review, materials and methods, findings, conclusions, and discussion.

A Short Summary of Literature

This section reviews studies conducted in recent years, particularly in the areas of traffic accidents and machine learning, and summarizes studies that may be relevant to this study.

Özden and Caci (2018) analyzed traffic accidents in Adana province between 2005 and 2014 in relation to meteorological data and found significant correlations between variables such as air temperature, precipitation, and wind speed and the number of accidents. Similarly, Yıldız et al. (2020) examined the seasonal distribution of traffic accidents and found that accidents were more serious in terms of both number and fatalities during the winter months. These studies emphasize the need for special measures to be taken in regions with severe climatic conditions.

Özden et al. (2018) conducted a spatial analysis of traffic accidents across Turkey using GIS, noting that accident density increases in urban areas and at intersections. Yavuz et al. (2021) used GIS-supported density maps to show that accidents in metropolitan areas occur more frequently during times close to working hours and on weekdays.

Karayel et al. (2019) analyzed traffic accidents across Turkey using logistic regression methods and showed that driver errors, speeding, night-day differences, and weather conditions were factors contributing to accidents. Çetin et al. (2021) examined highway accidents using decision trees and clustering techniques, determining that accidents were concentrated in the morning hours. Demir et al. (2021) identified weather conditions, road type, and speed as factors affecting the severity of accidents. Güler et al. (2022) emphasized that pedestrian accidents in urban areas occur more frequently in areas with high pedestrian traffic, such as school surroundings.

Ersöz and Ersöz (2022) used algorithms such as Random Forest, Naive Bayes, and Gradient Boosted Trees in their study conducted across Turkey and determined that weather conditions, driver age, and education are the most important factors affecting accidents. Cicek et al. (2023) showed that factors such as seatbelt use, alcohol influence, and speeding affect accident severity, with the highest accuracy achieved using the MLP model. Gündüz et al. (2020) analyzed the effects of driver type, road structure, and weather conditions on accident severity by integrating the AHP and TOPSIS methods.

Bedane et al. (2011) predicted traffic accidents in Addis Ababa using Logistic Regression, Decision Tree, and Random Forest algorithms and emphasized the effects of road type and traffic density. Infante et al. (2022) compared statistical and machine learning models using Portuguese data and noted that statistical models were more successful in cases of data imbalance. Bokaba et al. (2022) compared six different algorithms in South Africa and reported that Random Forest achieved the highest success rate.

Khan and Hussain (2024) analyzed the spatial distribution of traffic accidents in Pakistan using GIS and XGBoost models, identifying hotspots and recommending urban safety measures. Megnidio-Tchoukouegno and Adedeji (2023) evaluated vehicle safety systems and policy effects in the UK and found that the Decision Tree model was the most successful algorithm. Muktar and Fono (2024) used data from Montreal, Canada, to predict the effect of weather and road conditions on accidents with high accuracy using Random Forest.

Tang et al. (2025) demonstrated the effect of driver errors and speed variables on accident severity using the IXGB algorithm in China. Lin and Li (2020) indicated that neural networks can be successfully applied by making real-time traffic predictions using crowdsourced data in Beijing. Liu et al. (2021) developed AI-based models to predict the environmental impacts of traffic accidents and demonstrated the usefulness of these models in emergency response processes. Reddy et al. (2022) noted that accidents increase by 70–80% in rainy and snowy weather in the US.

Kumeda et al. (2019) revealed the relationship between the causes and consequences of traffic accidents using data-driven decision support systems and emphasized that such systems can be effectively used in traffic safety policies by making predictions using machine learning algorithms.

The most fundamental feature that distinguishes this study from similar studies in the literature is that the data used in this study are being used for the first time. In addition, the MANOVA method, which evaluates multiple dependent and independent variables simultaneously, was tested on real-time data, going beyond traditional statistical analysis methods.

MATERIALS AND METHODS

In this study, a machine learning model was developed to predict whether traffic accidents in Erzurum province were fatal. The modeling process was carried out using open source Python libraries. Data preprocessing, class imbalance handling, model training, evaluation metrics, and explainability analyses were systematically performed. At this point, the figure prepared for the materials and methods section is presented below in Figure 2.

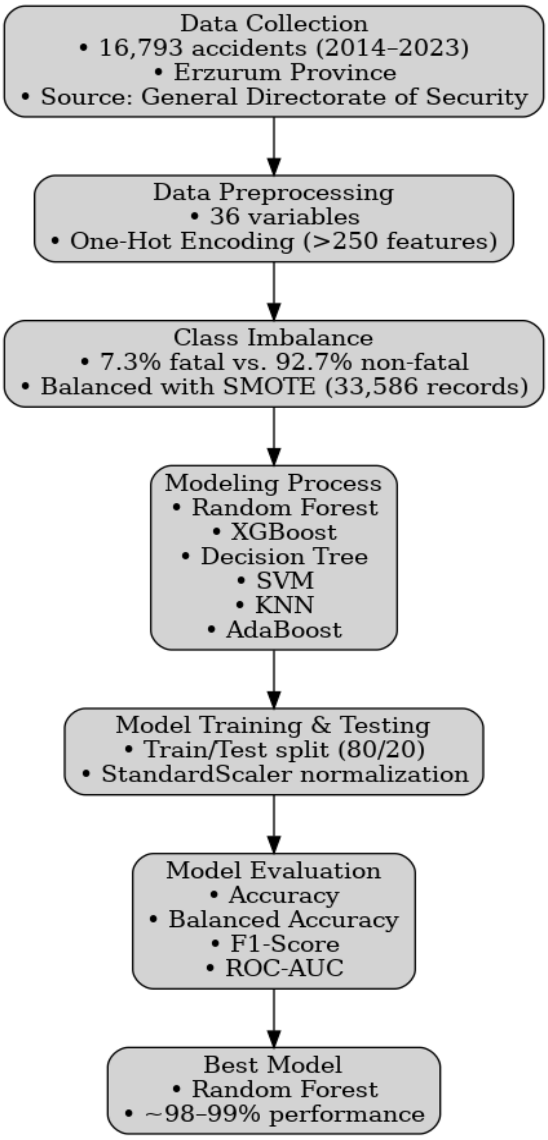

Figure 1. Study Flow Chart

Data Set

The dataset used in this study is public data provided by the General Directorate of Security, containing traffic accidents in Erzurum province between 2014 and 2023. The dataset contains information on a total of 16793 traffic accidents. These accidents are defined by a total of 36 variables across various categories, including physical conditions at the time of the accident, driver and vehicle information, and environmental factors.

The target variable in the dataset is "TotalFatalAccidents," which represents information on fatal accidents, and a binary classification is performed:

- 0 (Non-fatal accident): 15563 records
- 1 (Fatal accident): 1230 records
- Total: 16793 records

The information about the data is presented in groups in Table 1 below.

Table 1. Grouping of data

Type	Number of Variables	Sample Variables
Driver Characteristics	6	Driver age, gender, license status, alcohol test result, education status
Vehicle Characteristics	6	Vehicle type, model year, purpose of use, direction of movement, collision area
Road Conditions	8	Road type, class, gradient, intersection status, speed limit, ground conditions
Environmental Factors	6	Weather conditions, lighting, day/night, number of lanes
Accident Details	10	Accident type, number of vehicles, pedestrian crossing status, traffic signal status, point of first contact

The 36 independent variables in the dataset represent a total of approximately 150 different categorical values. For example, the "Accident Type" variable consists of 3 different classes (Single-Vehicle, Two-Vehicle, Multi-Vehicle), "Road Class" consists of 7 different categories, and "Weather Conditions" consists of 5 different conditions. Variables such as vehicle type, usage purpose, road type, ground condition, lighting, traffic light, and pedestrian crossing status also have between 2 and 10 different categories. To enable machine learning algorithms to process this categorical structure, the One-Hot Encoding method was applied, resulting in a total feature count exceeding 250. This extensive and multidimensional feature space enabled the model to classify accidents more accurately.

Class imbalance problem

The target variable in the dataset, TotalFatalAccidents, indicates whether traffic accidents were fatal or not in a binary format. However, upon examining the distribution, a significant class imbalance is observed: Out of a total of 16793 records, only 1230 (7.3%) represent fatal accidents, while the remaining 15563 records (92.7%) represent non-fatal accidents. Such imbalanced structures, if modeled directly without balancing, cause machine learning algorithms to favor the dominant class (i.e., non-fatal accidents) and neglect the minority class (fatal accidents).

To address this problem, the SMOTE method was applied before splitting the dataset into training and test sets. SMOTE balances the data by creating new and similar synthetic examples between the minority class examples rather than randomly copying them. This increases the number of examples and prevents repetitions that could lead to overfitting.

As a result of the SMOTE process, the two classes in the dataset were balanced, and the total number of examples was increased to 33586. After this balancing, the model was trained with the goal of accurately classifying fatal accidents.

Modeling process

After balancing the data set, supervised machine learning algorithms were used to classify fatal traffic accidents. In this process, six different classification algorithms representing different learning paradigms and widely used in the literature were selected. The algorithms used are as follows:

- Random Forest Classifier
- XGBoost Classifier
- Decision Tree Classifier
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)
- AdaBoost Classifier

Each model was trained on balanced training data using SMOTE and evaluated on a 20% split test data. Categorical variables were converted using the One-Hot Encoding method, and the data was normalized using StandardScaler to eliminate scale differences that some algorithms are sensitive to.).

Model performance was comprehensively evaluated using the following metrics, particularly due to class imbalance, in addition to the classic accuracy rate:

- Accuracy
- Balanced Accuracy
- F1 Score
- ROC AUC (Area Under the Curve)

This multi-evaluation approach allowed us to measure not only the overall performance of the models, but also how accurately they identified fatal accidents, which are the minority class. The Random Forest Classifier, in particular, demonstrated the highest performance across all metrics and was accepted as the baseline model.

RESULTS AND DISCUSSION

In this study, six different machine learning algorithms, namely Random Forest, XGBoost, Decision Tree, SVM, K-Nearest Neighbors (KNN), and AdaBoost, were applied to classify fatal traffic accidents. The models were trained on training data balanced using the SMOTE method and evaluated on test data separated at a rate of 20%.

Since evaluating the models' performance solely based on accuracy could be misleading, additional metrics such as Balanced Accuracy, F1-Score, and ROC AUC, which are more meaningful in imbalanced datasets, were also used. Table 2 summarizes the basic performance metrics for each model.

Table 2. Performance Metrics

Model	Accuracy	Balanced Accuracy	F1 Score	ROC AUC
Random Forest Classifier	0.99	0.99	0.99	0.99
XGBoost Classifier	0.97	0.97	0.97	0.97
Decision Tree Classifier	0.95	0.95	0.95	0.95
SVM	0.95	0.96	0.95	0.96
KNN	0.91	0.91	0.91	0.91
AdaBoost Classifier	0.79	0.79	0.79	0.79

According to the results obtained, the Random Forest algorithm performed best in all metrics and was the most successful model in the study. It was followed by the XGBoost and Decision Tree algorithms, respectively. The performance of the AdaBoost algorithm was lower than that of the other models.

These experimental findings show that ensemble methods are quite effective in fatal accident classification and that multi-decision tree-based models such as Random Forest offer high performance. The Random Forest model, which achieved the highest success in the study, provided superior results to other algorithms in terms of both accuracy and overall balancing performance. One possible reason for this superiority is that Random Forest, as an ensemble method, reduces variance and overfitting by combining multiple decision trees, while some of the other algorithms, such as logistic regression and SVM, may have struggled to capture complex, non-linear relationships in the data. Therefore, the results were interpreted based on this algorithm in the continuation of the study. The classification report obtained using the Random Forest classification algorithm is provided in Table 3 below.

Table 3. Random forest classification report

Class	Precision	Recall	F1 Score	Support
0-Non-Fatal	0.98	0.98	0.98	3317
1 Fatal	0.98	0.98	0.98	3203
Overall Average	**0.98**	**0.98**	**0.98**	**6520**

The classification performance obtained with the Random Forest algorithm is quite high. The accuracy, sensitivity, and F1 score values for both classes are at the 0.98 level. This shows that the model can successfully distinguish between fatal and non-fatal accidents. The fact that such high performance is achieved even in fatal accidents, which are the minority class, demonstrates that the SMOTE method applied to address class imbalance is effective and contributes to the model's learning process. The F1 score obtained proves that the model exhibits balanced performance in terms of both precision and sensitivity. Figure 2 Confusion matrix of the Random Forest model, showing the comparison between actual and predicted classes. The high number of correctly classified fatal and non-fatal accidents demonstrates the model’s strong predictive performance and low error rates.

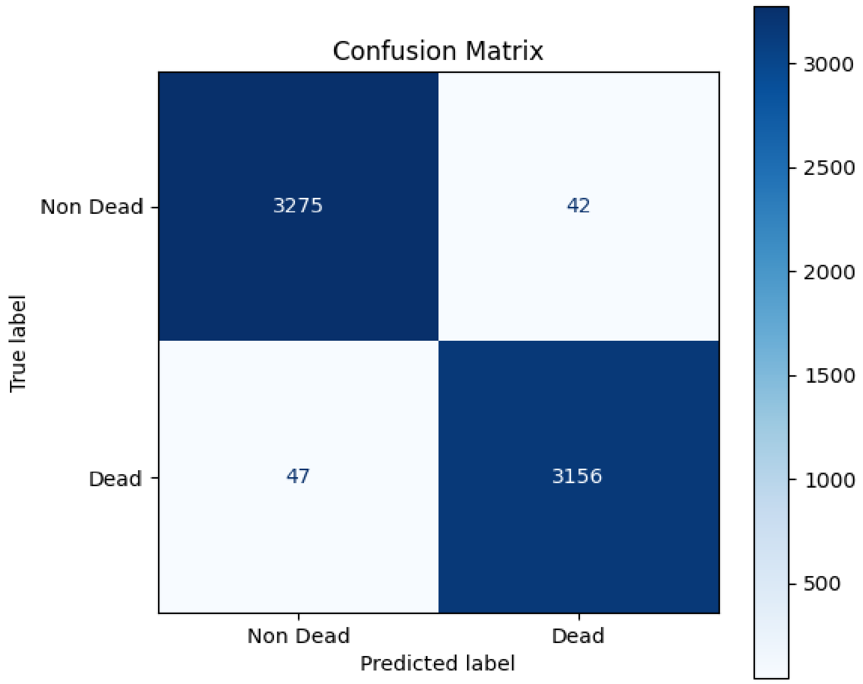

Figure 2. Model confusion matrix

This matrix shows the comparison of the model's predictions on the test data with the actual values. 3275 non-fatal accidents were correctly classified, with only 42 incorrectly labeled as fatal accidents. Similarly, 3156 fatal accidents were correctly predicted, with only 47 records incorrectly classified as non-fatal accidents. These results show that the model operates with very low error rates in both classes and achieves a high overall accuracy level. Additionally, the fact that the false positive and false negative rates are close to each other indicates that the model makes unbiased and balanced decisions.

The ROC curve obtained for the Random Forest model demonstrates the classifier's ability to distinguish the positive (fatal) class. Figure 3 ROC curve of the Random Forest model, illustrating the trade-off between sensitivity (true positive rate) and specificity (false positive rate). The AUC value of 0.99 indicates excellent discrimination ability in distinguishing fatal from non-fatal accidents.

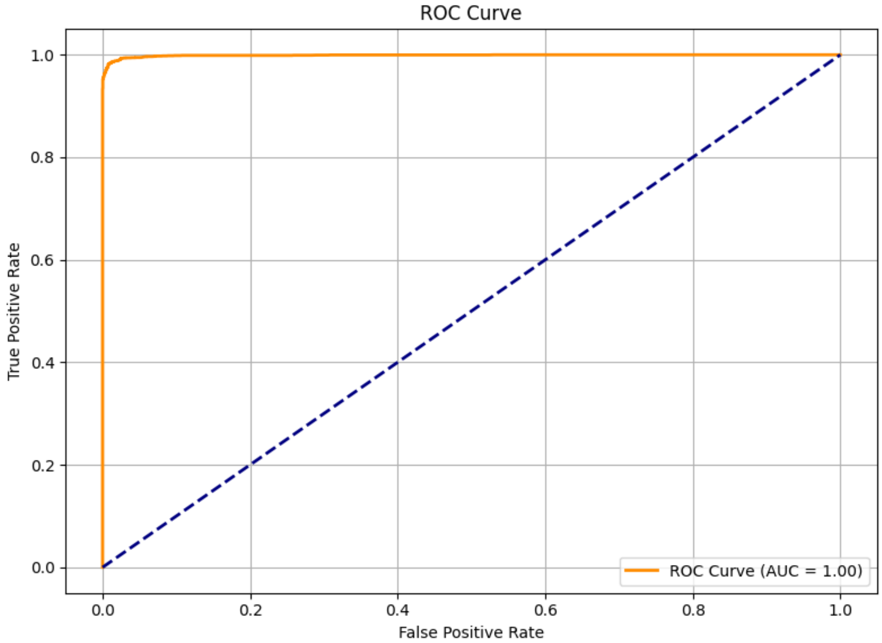

Figure 3. ROC curve for the proposed model

The area under the ROC curve (AUC) of the model was calculated as 0.99. This value shows that the model performs much better than a random classifier and can make a strong distinction between classes. Especially in imbalanced data sets, the AUC value is a critical measure of how successful the model is at classification at all threshold values. This high score demonstrates that the model has a very strong capacity to accurately detect fatal accidents.

CONCLUSION

In this study, a machine learning-based classification model was developed to predict fatal accidents using data from the General Directorate of Security on traffic accidents in Erzurum province. The class imbalance issue in the dataset was addressed using the SMOTE (Synthetic Minority Over-sampling Technique) method, and then six different classification algorithms (Random Forest, XGBoost, Decision Tree, SVM, KNN, and AdaBoost) were tested.

As a result of the experiments conducted, the Random Forest algorithm demonstrated the highest success in all metrics and was determined to be the most suitable model. The model's accuracy, F1 score, balanced accuracy, and ROC AUC values are all above 0.98. Additionally, the confusion matrix and ROC curve created also demonstrate that the model performs stable and balanced classification.

According to the feature importance analysis provided by the model, variables such as accident type (AccidentType), driver age (DriverAge), number of vehicles involved in the accident (NumberOfVehicles), and the presence of a pedestrian way (PedestrianWay) are among the decisive factors in predicting fatal accidents. Figure 4 Ranking of the 20 most influential variables in predicting fatal accidents using the Random Forest model. Accident type, driver age, number of vehicles, and pedestrian crossings are among the most significant factors affecting accident severity.:

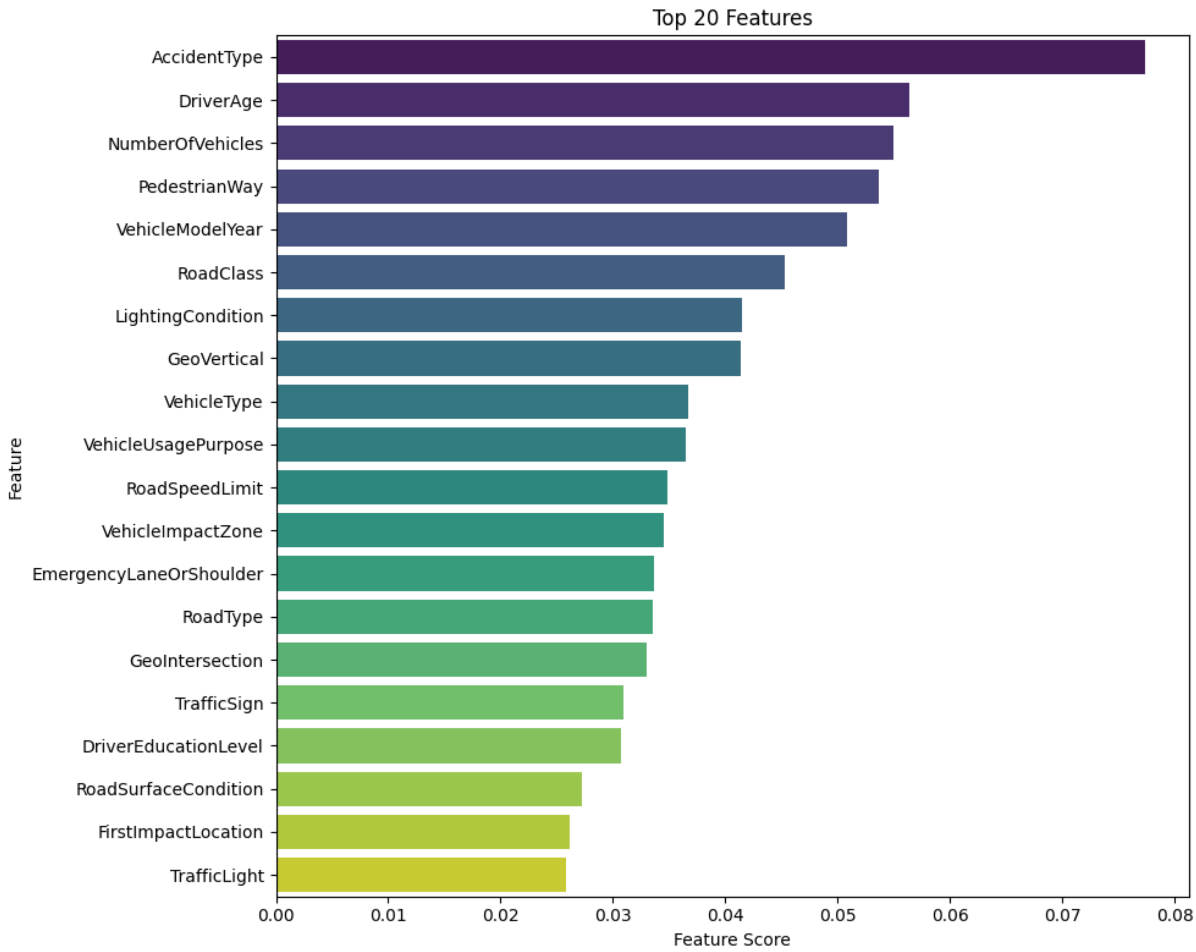

Figure 4. The 20 most important variables

The findings of the study reveal that, despite class imbalance issues, the Random Forest algorithm demonstrated high success in predicting fatal traffic accidents. In particular, the fact that basic performance metrics such as accuracy, F1 score, and ROC AUC were above 0.98 shows that the model was able to effectively learn both the majority and minority classes. These results confirm previous studies that support the strong generalization capabilities of ensemble-based algorithms such as Random Forest in complex and multidimensional datasets.

Feature importance ranking shows that fatal accidents are particularly influenced by structural and environmental factors such as accident type, driver age, number of vehicles, and pedestrian crossings. This highlights the need to develop preventive measures not only based on driver errors but also on road and environmental conditions. Similarly, the finding that urban infrastructure elements such as the presence of pedestrian crossings can influence fatal accidents is a consideration that should be taken into account in urban planning and traffic engineering.

One of the most important contributions of this study is that it demonstrates how machine learning can be used in traffic safety policies by providing a comparative analysis of numerous algorithms on a publicly available national dataset. In addition, the successful resolution of class imbalance using SMOTE provides a practical solution for researchers working on datasets with similar imbalanced structures.

Finally, the feature importance graph, which explains which variables are influential in the model's decision-making process, is also valuable in terms of interpretable artificial intelligence applications. This will enable decision-makers to focus not only on the results but also on the reasons behind them.

Based on the findings of this study, future research could develop comprehensive models that take into account different geographical regions, seasonal variables, or time series analyses. Additionally, conducting multivariate causality analyses that model the interaction of factors such as driver behavior, vehicle type, road infrastructure, and weather conditions using existing data could contribute to more accurate predictions of fatal accidents. Testing the performance of deep learning architectures, particularly in predicting time-dependent accidents, is another potential area for expanding this study.

Conflict of Interest

The article authors declare that there is no conflict of interest between them.

Author's Contributions

The authors declare that they have contributed equally to the article.

REFERENCES

Bedane, T. T., Assefa, B. G., & Mohapatra, S. K. (2011) Preventing traffic accidents through machine learning predictive models. 36-41. 10.1109/ICT4DA53266.2021.9672249. 2021 International Conference on Information and Communication Technology for Development for Africa (ICT4DA).

Bokaba, T., Doorsamy, W., & Paul, B. S. (2022). Comparative study of machine learning classifiers for modelling road traffic accidents. *Applied Sciences*, 12(13), 6655. https://doi.org/10.3390/app12136655

Cetin, A., Erkan, H., & Yilmaz, O. (2021). Otoyol trafik kazalarının veri madenciliği yöntemleriyle analizi. *Politeknik Dergisi*, 24(1), 55–63.

Cicek, O., Aydin, M., & Bulut, E. (2023). Analysis and prediction of traffic accident injury severity using machine learning algorithms. *Computational Intelligence and Neuroscience*, 2023, 1–13. https://doi.org/10.1155/2023/6679860

Demir, F., Caliskan, E., & Dede, T. (2021). Trafik kazası şiddet tahmini için makine öğrenmesi yöntemlerinin karşılaştırılması. *Gazi Üniversitesi Fen Bilimleri Dergisi* Part C: Tasarım ve Teknoloji, 9(1), 98–107.

Ersöz, T., & Ersöz, F. (2022). Data mining and machine learning approaches in data science: Predictive modeling of traffic accident causes. *International Journal of 3D Printing Technologies and Digital Industry*, 6(3), 530–539.

Guler, F., Arslan, Z., & Bilgin, Y. (2022). Şehir içi yaya kazalarının incelenmesi ve çözüm önerileri. *İnönü Üniversitesi Mühendislik Bilimleri Dergisi*, 13(2), 115–124.

Gunduz, M., Yildiz, M., & Cetinkaya, C. (2020). Trafik kazalarında şiddet derecesinin çok kriterli karar verme yöntemleriyle tahmini. *Gazi Üniversitesi Fen Bilimleri Dergisi* Part C: Tasarım ve Teknoloji, 8(3), 636–649.

Infante, P., Ribeiro, P., & Pais, C. (2022). Comparison of statistical and machine-learning models on road traffic accident severity classification. *Computers*, 11(10), 142. https://doi.org/10.3390/computers11100142

Karayel, M., Bayrakdar, M., & Kaya, M. (2019). Lojistik regresyon modeli ile trafik kazalarının analizi. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 9(2), 229–236.

Khan, A. A., & Hussain, J. (2024). Utilizing GIS and machine learning for traffic accident prediction in urban environment. *Civil Engineering Journal*, 10(6), 1922–1936.

Kumeda, B., Zhang, F., Zhou, F., Hussain, S., Almasri, A., & Assefa, M. (2019). Classification of Road Traffic Accident Data Using Machine Learning Algorithms 2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN), Chongqing, China, 2019, pp. 682-687, doi: 10.1109/ICCSN.2019.8905362.

Lin, X., & Li, W. (2020). Real-time traffic congestion impact prediction based on crowdsourcing accident data and deep learning. *IEEE Access*, 8, 130515–130524.

Liu, Y., Zhang, J., Chen, M., & Wu, G. (2021). Prediction of traffic accident impact duration using artificial intelligence methods. *Accident Analysis & Prevention*, 151, 105964. https://doi.org/10.1016/j.aap.2021.105964

Megnidio-Tchoukouegno, M., & Adedeji, J. A. (2023). Machine learning for road traffic accident improvement and environmental resource management in the transportation sector. Sustainability, 15(6), 5225. https://doi.org/10.3390/su15065225

Muktar, B., & Fono, V. (2024). Toward safer roads: Predicting the severity of traffic accidents in Montreal using machine learning. *Electronics*, 13, 3036. https://doi.org/10.3390/electronics13153036

Ozden, E., & Caci, M. (2018). Meteorolojik verilere dayalı olarak yaralanmalı trafik kazalarının tahmini: Adana ili örneği. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 24(2), 266–275.

Ozden, S., Karakas, M., & Sahin, A. (2018). Trafik kazalarının coğrafi bilgi sistemleriyle analizi: Türkiye örneği. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 18(1), 145–155.

Reddy, K. M., Joseph, R., & Hamza, T. (2022). Machine learning models for predicting traffic accident severity based on environmental factors. *Journal of Intelligent Transportation Systems*, 26(4), 456–469.

Sysoev, A., Klyavin, V., Dvurechenskaya, A., Mamedov, A., & Shushunov, V. (2022). Applying machine learning methods and models to explore the structure of traffic accident data. *Computation*, 10(4), 57. https://doi.org/10.3390/computation10040057

Tang, J., Huang, Y., Liu, D., Xiong, L., & Bu, R. (2025). Research on traffic accident severity level prediction model based on improved machine learning. *Systems*, 13(1), 31. https://doi.org/10.3390/systems13010031

Yavuz, O., Kaya, I., & Turan, G. (2021). Trafik kazalarının mekânsal analizi: Büyükşehir örneği. *Türkiye Coğrafya Dergisi*, 77, 45–60. https://doi.org/10.17211/tcd.898765

Yildiz, E., Arslan, F., & Korkmaz, B. (2020). Türkiye'de mevsimsel etkilerle trafik kazalarının değerlendirilmesi. *Trafik ve Yol Güvenliği Dergisi*, 3(1), 13–21.