

## Research Article

# Analysis of the Severity of Accidents on Rural Roads Using Statistical and Artificial Neural Network Methods

Mohammad Habibzadeh <sup>1</sup>, Pooyan Ayar <sup>1</sup>, Mohammad Hassan Mirabimoghaddam,<sup>2</sup> Mahmoud Ameri,<sup>1</sup> and Seyede Mojde Sadat Haghighi<sup>1</sup>

<sup>1</sup>School of Civil Engineering, Iran University of Science & Technology (IUST), Tehran, Iran

<sup>2</sup>Department of Civil Engineering, University of Sistan and Baluchestan, Zahedan, Iran

Correspondence should be addressed to Mohammad Habibzadeh; [habibzadeh1991@gmail.com](mailto:habibzadeh1991@gmail.com)

Received 31 July 2023; Revised 12 December 2023; Accepted 20 December 2023; Published 29 December 2023

Academic Editor: Yajie Zou

Copyright © 2023 Mohammad Habibzadeh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study assesses the relationship that existed between various variables and their subvariables on rural roads in Qom, Iran, using statistical analysis and calculates the relationship between the considered factors and accident severity. A logit model was applied to determine the factors affecting the severity of accidents. In addition, two artificial neural network (ANN) models were developed using two kinds of learning methods to train neurons to select the best result. The results of modeling and analysis of accidents using various techniques revealed that each technique, depending on its purpose, examined the severity of accidents from a different point of view and represented various outcomes. Finally, the performance of the proposed models was validated utilizing other mathematical models. As a result, putting the output results together, the best measures can be suggested to increase the safety of people on rural roads. The outcomes of this study may aid these service providers in strategic planning and policy framework.

## 1. Introduction

The increase in road transportation increased damage dramatically, which resulted from accidents [1]. With the rapid growth of the economy, the transportation of rural roads plays a critical role in the system of country transportation [2–4]. Thus, it is necessary to make an accurate plan to reduce the destructive effects and evaluate the effectiveness of previously conducted activities [5, 6]. The examination and planning will not be carried out without collecting data and predicting the condition of the future [7, 8]. In order for optimal management and reducing accidents on rural roads, it is necessary that safety authorities collect comprehensive data related to the factors affecting accidents.

According to the latest statistics which are officially presented, the number of fatal accidents was 15923 in 2016 in Iran. Also, in this period, the estimation of fatal accidents on rural roads was 20.5 persons per 100,000 people,

which reveals the large number of accidents on rural and urban roads that impose a heavy burden on the government [9].

The unsuitable conditions of traffic in the current condition of the country have been increasing and its harmful effects on people's health are obvious. Daily waste of millions of hours of time as well as the lives of people in traffic accidents, pollution of cities, waste of facilities, community services, and national capitals, and finally the occurrence of accidents are some of the consequences of accidents. Thus, a wide range of transportation issues and challenges are required to be planned and considered for practical approaches to reduce accidents. Therefore, it is necessary to analyze the influence of variables on accident severity. Towards solving this issue, in this study, by collecting data and accident information on variables on rural roads in Iran, the variables which affect accidents were considered. The purposes of the research are mentioned below:

- (i) Extracting variables that can indicate various variables and their subvariables affecting accident severity.
- (ii) Developing a Friedman test to prioritize the effective factors.
- (iii) Developing a binary logistic regression model to predict the severity of accidents.
- (iv) Developing two types of artificial neural network (ANN) methods using various methods for training neurons and evaluating the performance of the developed models applying other mathematical models.

These analyses will assist safety authorities to provide insight and contribute to a more thorough knowledge of accidents. Additionally, it can aid in reducing accident severity, which enhances road traffic safety, as well as improves traffic efficiency for the traffic management department. In this study, since the purpose of the research is to detect the most influential parameters affecting accident severity, various methods, including statistical and mathematical models developed by MATLAB programming, were employed to detect the variables that have significant effects on accidents. Each method has its own unique results, and none of them can replace another one. It means that accident data are analyzed under different methods that can provide us with the best result to make a decision about safety issues. In this research, the severity of traffic accidents in Qom province, Iran, was evaluated using various methods of analysis and modeling such as the Friedman test, logit, and neural network models. Accident data were collected from 2017 to 2020 on rural roads to highlight the impact of traffic injuries on public health and enhance preventive efforts. In addition, the Friedman test is a reliable method used to prioritize and compare the effective factors. Also, other mathematical models, including ANN and logit models, can help us recognize the most significant factors as well as detect the power of the model to predict accident severity under different classes.

## 2. Literature Review

Due to the rising concern for driving safety, accident severity is currently receiving more and more research attention, and its contributing variables have also been thoroughly examined. In general, the primary viewpoints to investigate the influences of accidents are the driver's attributes and the vehicle's characteristics. Additionally, studies are done on extraneous elements like traffic fines and the surrounding area for drivers.

Moreover, the majority of the researchers considered accident severity from the perspective of various characteristics. Al-Ghamdi using a logistic regression model studied the effectiveness of accident factors on accident severity and found that two variables such as location and cause of the accident were considered as significant variables. Also, the results showed that the logistic regression model can be a robust tool to analyze accidents [10]. Pinto

et al. conducted research to reduce subjectivity in the evaluation of occupational accident severity. They proposed different functions to demonstrate biomechanical knowledge with the purpose of detecting the severity level of occupational accidents in the construction industry and, consequently, improving occupational risk assessment quality [11]. Petrović et al. analyzed accidents with autonomous vehicles that occurred in the US state of California. They considered the type of collision, maneuvers, and errors of the drivers of conventional vehicles causing accidents. The results showed that the type of collision "rear-end" had a significant effect on accidents. There were fewer "pedestrian" and "broadside" accidents in accidents, which involved autonomous vehicles [12]. Golob and Recker used linear and nonlinear multivariate statistical analyses to recognize how traffic volume, weather, and ambient lighting conditions are related to the different kinds of accidents that happen on the busiest freeways of Southern California. The results revealed that the type of collision had a significant relationship with median traffic speed and was also related to temporal variations in speed in the left and interior lanes [13]. Miaou and Lum employed four kinds of regression models, including two conventional linear regression models and two Poisson regression models, to examine the relationships between highway geometric design and accidents. The results indicated that these models are not able to predict accidents and their results were not accurate [14]. Elvik et al. used negative binomial regression models to evaluate characteristics that caused systematic variation in the number of injury accidents on road bridges in Norway. Annual Average Daily Traffic (AADT) was considered to be the most significant variable affecting accidents [15]. Al-Balbissi used statistical analysis to analyze the effectiveness of driver sex on accidents. In this research, the effect of public accidents, annual distance traveled, and social and economic participation was regarded. The results represented that males had a significant partition in accidents [16]. Rolison et al. analyzed the effect of inexperience, lack of skill, and risk-taking behaviors variables related to the collisions of young drivers. The major reasons for accidents were presented using multiple sources in this study. Official records of road accidents present the opinions of drivers and the professional opinions of police officials. The results of the study indicated that both lay views of the driving public and expert views of police officers closely estimated the typical factors related to collisions between young and older drivers. Their investigation demonstrated that there is a need for accident report forms to be continuously reviewed [17]. Beshah and Hill used data mining techniques to link observed road characteristics to the severity of accidents in Ethiopia and created a set of guidelines that the Ethiopian Traffic Agency might employ to increase safety [18]. Hammad et al. evaluated the relationship between accidents and the variables of weather conditions, including wind storms, rainfall, fog, and temperature. The results of this study showed that rainfall, severe coldness, fog, and heat conditions were directly associated with accidents [19]. Mirzahosseini et al. presented statistical and intelligent models to predict the likelihood of road traffic accidents.



FIGURE 1: Location of the study area.

They indicated that the variables of not paying attention to the front and then vehicle-motorcycle/bike accidents had the most influence on the occurrence of accidents [20].

By regarding different aims and a wide variety of unsafe behaviors, it is necessary to explore the relationship between the variables, including vehicle features, geometric conditions, human behaviors, and accident occurrence, which can provide insight for traffic managers to formulate targeted publicity and safety education as well as to take preventive measures to enhance safety which results in saving people's lives. Thus, an attempt was made in this study to use the variables and their subvariables that have been rarely investigated in previous studies. In this research, we made an effort to consider the mentioned variables using the Friedman test, as well as ANN and logit models. The Friedman test is a reliable approach applied to prioritize and compare the effective factors. Also, the reason for selecting ANN and logit models is that their characteristics are unique in complex problems. These models are capable of solving accident issues and have the ability to provide fast and reliable methods to consider the nonlinear relationships that exist among input variables. Moreover, there is no need to consider these assumptions in developing these models.

### 3. Study Route and Methodology

**3.1. Data Collection.** The current study was carried out in Qom province, Iran, and the rural roads of this province were considered. Qom province is one of the most important provinces in Iran. Figure 1 indicates the location of Qom province in Iran. It is located in the area in which major corridors connect important provinces of the country including south, southeast, and southwest provinces. The dataset was gathered from 2017 to 2020 on the monitored roads in Iran. It should be noted that in certain circumstances, even though police officers reported that the people were injured at the location of an accident, these people could have died while being transported to the hospital or

just after. Additionally, some of the collisions that cause damage typically resolve amicably without being recorded by police. However, for fatal accidents, the cases were completely and accurately recorded.

In this study, 403 accident data were gathered, of which 283 (70.2%) were damage accidents, and 120 (29.8%) were fatal/injury ones. The information includes the number of accidents, the location of the accident, and the type and the severity of accident. The target variable in the research was the severity of accident, split into three kinds of accidents, including fatal, injury, and damage. It should be noted that since the number of fatal accidents was few compared to the total ones, the goodness of fit and the significance of the developed models could not be provided given three kinds of the target variable, and thus fatal and injury accidents were merged with each other. Therefore, the dependent variable was divided into two levels, including damage accidents and fatal/injury ones. More details about the variables are represented in Table 1.

### 4. Statistical Analysis

**4.1. Friedman Test.** The Friedman test (FT) is usually employed to compare the classification of several datasets. This is a nonparametric method that can be applied to analyze the samples related to computational biology and other issues. The process of the evaluation of the FT is the analysis of the variance by ranks. In fact, this test analyzes the recognized ranks or rank scores created by numerical or ordinal results. This test is employed once a researcher does not want to have strong distributional assumptions. The simultaneous evaluations are represented as well as the ordinal common approximation used by the overall plan of covariance for the distribution of the amount of variance used in the form of ranking [21].

The rank of each variable in the paper was examined using the Friedman test. The rank equality relating to parameter levels was evaluated using the FT. Table 2 provides

TABLE 1: The variables of the study.

Variable	N	Variable levels	Percent
Accident severity (AS)	283	(1) Damage	70.2
	120	(2) Fatal/injury	29.8
Accident location (AL)	351	(1) Roadway	87.1
	16	(2) Shoulder	4.0
	6	(3) Median refuge	1.5
	12	(4) Roadside	3.0
	7	(5) Right-of-way	1.7
	11	(6) Other cases	2.7
The application of accident location (PL)	11	(1) Residual	2.7
	13	(2) Administrative-commercial	3.2
	9	(3) Industrial	2.2
	6	(4) Agriculture	1.5
	16	(5) Educational	4.0
	348	(6) Other cases	86.5
Type of vehicle (TV)	83	(1) Passenger vehicle	20.6
	30	(2) Pick up	7.4
	19	(3) Mini truck	4.7
	2	(4) Bus	0.5
	17	(5) Minibus	4.2
	125	(6) Truck	31.0
	106	(7) Trailer	26.3
	19	(8) Motorcycle	4.7
	2	(9) Other cases	0.5
Collision with (CW)	22	(1) Motorcycle	5.5
	7	(2) Bicycle	1.7
	213	(3) Single device	52.9
	21	(4) Multiple devices	5.2
	3	(5) Parked vehicle	0.7
	20	(6) Pedestrian	5.0
	3	(7) Animals/cattle	0.7
	33	(8) Fixed object	8.2
	38	(9) Overturning	9.4
	33	(10) Run-off-road	8.2
	6	(11) Fire	1.5
	4	(12) Other cases	1.0
Type of collision (TC)	68	(1) Head on	16.9
	64	(2) Head-side	15.9
	96	(3) Read-end	23.8
	12	(4) Side to side	3.0
	7	(5) Side end	1.7
	156	(6) Other cases	38.7
Vehicle facilities (VF)	397	(1) None	98.5
	4	(2) Airbag	1.0
	2	(3) ABS	0.5
The type of maneuver of the guilty vehicle (MV)	224	(1) Moving forward	55.6
	20	(2) Turning to left	5.0
	28	(3) Turning to right	6.9
	18	(4) Turning	4.5
	11	(5) Overtaking	2.7
	23	(6) Moving to reverse gear	5.7
	17	(7) Start moving suddenly	4.2
	15	(8) Sudden stop	3.7
	15	(9) Stop off the road	3.7
	6	(10) Stop on the roadway	1.5
	16	(11) Spiral movement	4.0
	10	(12) Other cases	2.5

TABLE 1: Continued.

Variable	N	Variable levels	Percent
Geometric conditions (GC)	275	(1) Straight flat	68.2
	7	(2) Climbing	1.7
	17	(3) Downhill	4.2
	26	(4) Flat curve	6.5
	28	(5) Climbing curve	6.9
	23	(6) Downhill curve	5.7
	13	(7) Tunnel	3.2
	8	(8) Bridge	2.0
	7	(9) Other cases	1.5
The type of shoulder (TS)	23	(1) None	5.7
	66	(2) Soil shoulder	16.4
	314	(3) Asphalt shoulder	77.9
Human factors in accidents (HF)	31	(1) None	7.7
	10	(2) Fatigue and drowsiness	2.5
	18	(3) Defect of effective organ	4.5
	22	(4) Weakness due to old age	5.5
	23	(5) Drug use	5.7
	12	(6) Consumption of alcoholic beverages	3.0
	15	(7) Not properly restraining the cargo	3.7
	228	(8) Unnecessary haste	56.6
	29	(9) Failure to recognize the share of pedestrians	7.2
	12	(10) Lack of familiarity with the road	3.0
	3	(11) Other cases	0.7
Road factors (RF)	187	(1) None	46.4
	5	(2) Deficiency of public signs	1.2
	3	(3) The width of the road is not enough	0.7
	11	(4) The presence of obstacles and bumps	2.7
	16	(5) Lack of soil shoulders and parking	4.0
	17	(6) The difference between asphalt curve and shoulder	4.2
	13	(7) Lack of roadside safety barriers	3.2
	12	(8) Road subsidence	3.0
	16	(9) Asphalt surface defect	4.0
	17	(10) Curve with a sharp angle	4.2
	14	(11) Nonstandard cross and longitudinal slope	3.5
	13	(12) Defective road lighting	3.2
	16	(13) Defective road markings	4.0
	63	(14) Other cases	15.6
Marking of accident location (ML)	10	(1) None	2.5
	388	(2) Disconnected	96.3
	5	(3) Continuous	1.2

information on the level of significance, chi-square value, degrees of freedom, and statistical significance of the statistical sample volume, demonstrated by Sig.

According to Table 3, a significance level of less than 5% represents that  $H_0$  should be rejected, and claims of equal rank for the four listed parameters were rejected, and ratings were therefore inconsistent. The average rating for each variable is shown in Table 3 along with its ranking condition, with a lower average rating indicating a greater influence of the variable.

Table 3 indicates that the variables of vehicle facilities, the location of accident, accident severity, and geometric condition were regarded as the most influential variables on accident severity with values of 3.01, 3.617, 3.91, and 5.11, respectively. In addition, the parameters of collision with, the application of accident location, and human factor in accidents had the least influence on accidents.

TABLE 2: Results of the Friedman test.

Number of data	Chi-square	Degrees of freedom	Asymp. sig.
403	1914.970	11	0.0

**4.2. Logistic Regression Analysis.** A logistic regression model was designed to evaluate the influence of different independent variables. The dependent parameter is separated into two groups in this method [22]. It was labeled 1 to 2, including damage and fatal/injury. The independent variables are the accident location, the application of accident location, type of vehicle, collision with, type of collision, vehicle facilities, the type of maneuver of the guilty vehicle, geometric condition, type of the shoulder of accident location, human factor in accidents, marking of accident location, and road factor. Each independent variable range is divided according to Table 1.

TABLE 3: Mean rank in the Friedman test.

Variables	Mean	Rank
Vehicle facilities	2.65	1
Accident location	3.24	2
Geometric conditions	5.56	4
Marking of accident location	5.41	5
The type of maneuver of the guilty vehicle	5.65	6
Type of the shoulder of accident location	6.57	7
Type of collision	7.01	8
Road factor	7.20	8
Type of vehicle	8.08	9
Collision with	8.48	10
The application of accident location	9.18	11
Human factor in accidents	9.97	12

The stepwise forward and backward procedures were frequently applied to process the data. After that, two  $R^2$  criteria and the right proportion were taken into account to select the best approach. The values of the right percentage and goodness of fit for the backward direction were higher than those for the forward direction, as shown in Table 4. The backward technique, with an accurate percentage of 74.9% and a  $R^2$  value of 0.421, was selected as the best method to develop the logistic model to predict traffic accidents on rural roads in the province of Qom because of its accuracy in predicting accident severity.

Table 5 depicts that the model successfully predicted 50 out of 120 fatal and injury accidents and 252 out of 283 damage accidents, respectively. According to measurements, the logit model's predictive values for damage and fatal/injury accidents were 89% and 41.7%, respectively. As a result, the classification and separation capabilities of the model for damage accidents were superior to those for death or injury accidents. Additionally, the total success rate of the model in determining accident severity was 74.9%.

According to Table 6, a total of 8 independent parameters, which include motorcycle, head-on, rear-end, moving forward, turning to the left, turning to the right, overtaking, and the presence of obstacles and bumps, were investigated. The sign of the  $B$  parameter can indicate a variation trend. The  $B$  value is the "estimated increase in the exp of the outcome per unit increase in the value of the exposure." Exp ( $B$ ) is calculated as  $e^B$ , which could be utilized to measure the magnitude of the special influence. The results represented that rear-end collision has a significant negative impact on accident severity ( $B = -1.111 < 0$ , Exp ( $B$ ) = 0.329). It means that if it is assumed that the impact of other variables remains, an increase in the rear-end collision variable is associated with a decrease in the odds of accident severity. In other words, the probability of accident severity will be reduced by the value of  $-1.111$ , meaning that the type of collision related to rear-end has less likelihood of being involved in accidents compared to other types of collision. An increase in motorcycles has been related to the reduction of the odds of accident severity ( $B = -4.529 < 0$ ), with an exp of 0.011 (95% CI, 0.0002–0.64). Thus, the probability of a motorcycle reduces with the severity of accidents. As for head-on collision, the higher the head-on collision, the lower

the accident level. For the type of maneuver of the guilty vehicle factors, as expected, all of them positively influence accident levels. It indicates that the increase in four types of maneuvers of the guilty vehicle labeled 1 to 4 corresponds to the rise in the probability of accident severity.

The significance (sig), degree of freedom (df), and chi-square values of the backward method are shown in Table 7, for the first step in the modeling. The logistic model associated with Step 1 has a chi-square value of 107.985 and a significant value of less than 5%, which reveals that the ability of the model to predict accidents was confirmed.

## 5. Modeling Using Artificial Neural Network

In this research, two kinds of artificial neural network (ANN) models developed using MATLAB programming were employed to predict accident severity. The first method is the scale conjugate gradient (SCG) with pattern recognition capability applied to develop the predictive model. Applications of neural networks in computer vision, speech recognition, and text classification heavily rely on pattern identification [23]. It functions by utilizing either unsupervised or supervised classification to divide incoming data into objects or groups based on essential characteristics. The machine learning method uses the same input and output labels mentioned in Table 1. It should be noted the target parameter includes different groups of accident severity. The target parameter was separated into two classes, including damage and fatal/injury. Afterwards, the ANN model was created using an algorithm that existed in the software. The work of the input data in the ANN model was split into three classes.

- (i) Training: during training for the learning process, these are given to the network, and the network is adjusted based on its error.
- (ii) Validation: the data considered for validation are used to assess the generalization of the network and to end training when generalization reaches a certain point.
- (iii) Testing: these do not affect training and offer a dependable indicator of network efficiency both before and after training. Understanding how closely the neural network's findings match the actual outcome is the primary criterion.

The number of accidents occurred during 2017–2020 and the sum of three years were enough to train the network. 70% of the dataset was applied for the training phase, 15% of the data was utilized for the validation phase, and the remaining 15% was considered a test of the developed model.

*5.1. Results of Confusion Matrix.* The confusion matrix of the three phases of training, testing, and validation used for developing the ANN models of traffic accidents is shown in Figure 2. This kind of matrix provides a contribution to analyzing the accuracy of the network in predicting accidents (damage, injury, and fatal). The squares (1.1) and (2.2)

TABLE 4: Summary of the forward and backward methods of accident severity.

Number	Logit regression methods	Correct percentage	Goodness of fit ( $R^2$ )
1	Forward stepwise	71.5	0.219
2	Backward stepwise	74.9	0.421

TABLE 5: Classification table in the logit model.

Observed		Predicted		Correct percentage
		Accident severity		
		Damage	Fatal and injury	
Accident severity	Damage	252	31	89.0
	Fatal and injury	70	50	41.7
Overall percentage		74.9		

shown in green squares are the cases correctly classified by the model and the squares (1.2) and (2.1) shown in pink squares are the cases presenting a false prediction of the model [24]. In addition, the gray square represents the total predictive power of the network. As shown in Figure 2, the confusion matrix represents the three modes of training, validation, and testing; out of 120 property damage accidents, only 2 cases, and out of 283 fatal/injury accidents, 279 cases were properly classified by the model. The predictive accurate percentage of property damage accidents in the model is 1.7%, and the predictive accurate percentage of fatal/injury accidents is 98.6%. To be more specific, square (1.1) represents 2 accidents that are correctly predicted as damage and square (1.2) indicates that 4 accidents, leading to fatal/injury, were inaccurately predicted as damage. Also, square (2.1) depicts that 118 fatal and injury accidents were mistakenly classified as damage and square (2.2) implies that 279 accidents were accurately classified as fatal/injury accidents. Finally, the gray square indicates the overall traffic accidents' predictive power of the model which is 69.7%.

**5.2. The Results of the Performance of Neural Network.** Figure 3 illustrates the neural network's performance for training the network. Figure 3 depicts the results of the artificial neural network (ANN) training performance. ANN training performance shows the amount of gradient, and the best validation performance based on the mean squared error (MSE) values is 0.29325 at epoch 4. The architecture of MLP developed using the SCG method is shown in Figure 3. The number of neurons was eight using MATLAB programming.

**5.2.1. Sensitivity and Specificity Analysis of the MLP Model.** Figure 4 displays the sensitivity analysis of the true positive rate of the created ANN model in comparison to the false positive rate for accidents. A method for displaying, arranging, and choosing classifiers according to their performance is the receiver operating characteristic (ROC) graph. Its popularity is attributed to a number of well-researched traits, including the intuitive visual interpretation of the curve and the simplicity of model comparisons. The performance of the multiclass classification problem is checked or visualized using the ROC curve.

In the ANN model, 15% of the data are used for testing, 15% are used for validation, and 70% are utilized for training. The more angled the top and left curves are, the more effective the network is at estimating and predicting more correctly, as illustrated in Figure 4. Class 1 shows the correctness of the network's prediction for current accidents, while Class 2 indicates its accuracy for future accidents. Figure 5 shows the gradient of the MLP model in terms of the number of epochs. As shown in Figure 5, after nine iterations, the gradient converges to the value equal to 0.016821; therefore, in the modeling process, the number of iterations was set to eleven. Also, the histogram of the datasets in the MLP model created by the SCG method is represented in Figure 6.

**5.2.2. The Comparison of the Performance of ANN Developed by the SCG Method and Logit Regression Model.** A comparison between the percentage of correct prediction in the neural networks and the logit model indicated that the logit regression performed better, which can provide better prediction in comparison to the MLP model. In fact, the percentage of correct prediction in the logit regression model was 74.9%; however, the prediction accuracy of the MLP model was 69.7%. In other words, the prediction error rate of the MLP model was 30.3%, while the logit prediction error was 25.1%, which indicates the logit model was capable of predicting traffic accidents more efficiently. In addition, as shown in Table 8, the percentage of correct prediction of the logit model for damage accidents was 89.0%, showing that this model can be considered a robust model in predicting damage accidents, while the percentage of correct prediction of the MLP model for damage accidents was 1.7%, indicating that this model was so poor in predicting damage accidents; however, its prediction percentage for fatal/injury accidents was 98.6%, which represents that this model was successful in recognizing fatal/injury accidents, in contrast to the performance of the logit model.

**5.3. The Performance of MLP Built by the LM Method.** The second ANN model developed in this research is the multilayer perceptron (MLP) built using the Levenberg-Marquardt (LM) method which is the fastest way to train the networks as shown in Figure 7. Convergence

TABLE 6: The estimation of logistic regression model variables in the first step.

Variables	Factors	B	Std. error (SE)	Exp (B)	Wald	Sig	Exp (B) (95% CI)
Collision with	Motorcycle = 1	-4.529	2.085	0.011	4.718	0.030	0.0002-0.64
Type of collision	Head on = 1	-0.902	0.510	0.406	3.124	0.077	0.149-1.103
Type of collision	Rear-end = 3	-1.111	0.401	0.329	7.693	0.006	0.150-0.72
The type of maneuver of the guilty vehicle	Moving forward = 1	1.580	0.954	4.855	2.738	0.098	0.747-31.5
The type of maneuver of the guilty vehicle	Turning to left = 2	1.990	1.120	7.313	3.156	0.076	0.815-65.69
The type of maneuver of the guilty vehicle	Turning to right = 3	1.779	1.075	5.924	2.739	0.098	0.72-48.67
The type of maneuver of the guilty vehicle	Overtaking = 5	2.363	1.205	10.623	3.843	0.050	1.001-112.73
Road factor	The presence of obstacles and bumps = 4	2.310	0.901	10.07	6.568	0.010	1.72-58.97



TABLE 7: The results of the backward stepwise model.

		Chi-square	Df	Sig.
Step 1	Step	107.985	82	0.029
	Block	107.985	82	0.029
	Model	107.985	82	0.029

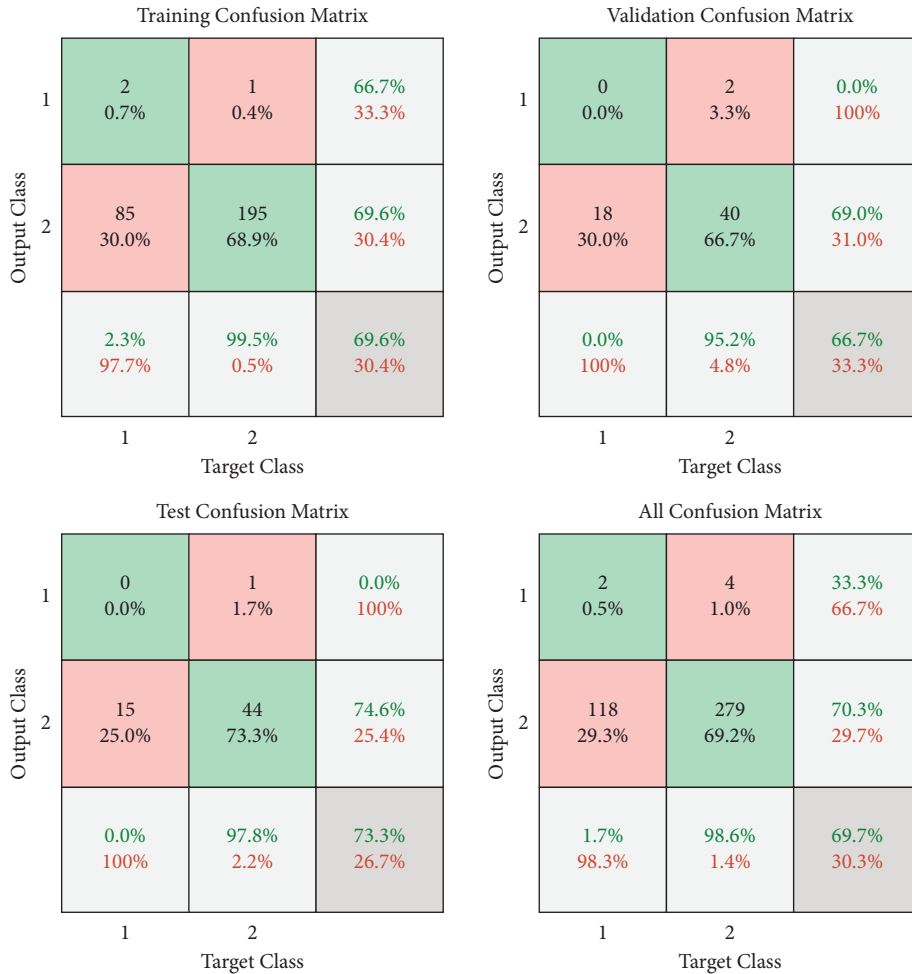


FIGURE 2: Confusion matrix of accidents.

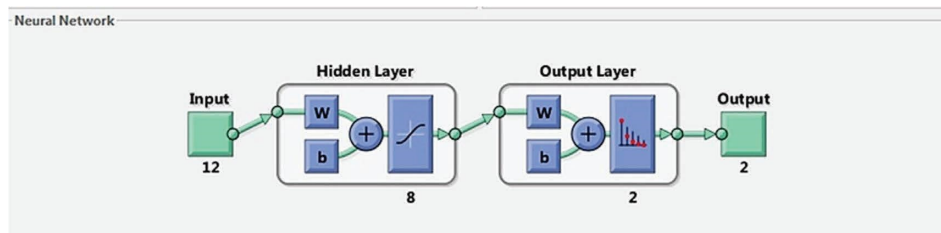


FIGURE 3: The developed MLP structure was trained using the SCG method.

quickness and model convergence assurance are its most crucial features. This algorithm’s purpose is to train neural networks. Data were randomly split into three groups: training, testing, and validating samples. The layer weight (LW) and input weight (IW) matrices were used in the MLP model. The MLP model had 12 inputs and 32 neurons in the

hidden layer. The output layer of the MLP model included one neuron. 70% of the data were applied in the training mode. The validation and testing datasets each contained 15% of the data. As shown in Figure 7, the number of neurons in the hidden layer was 32, detected based on trial and error in MATLAB software.

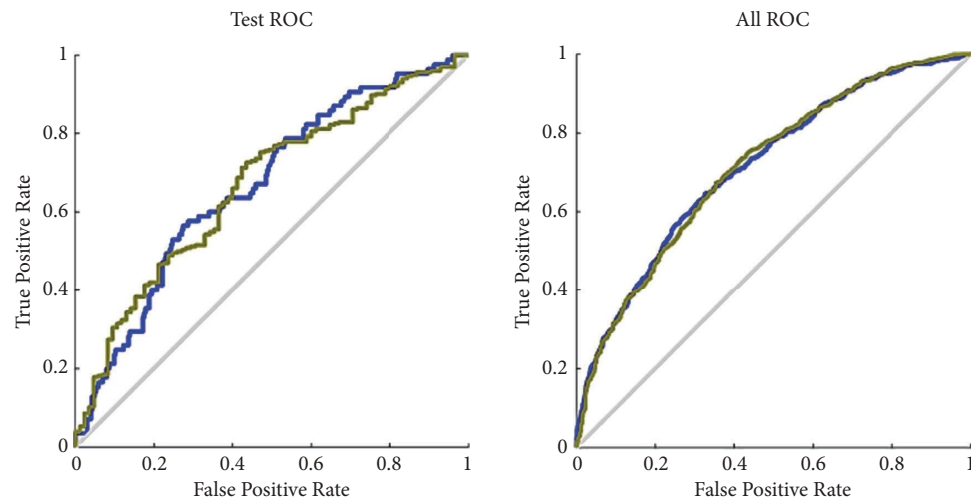
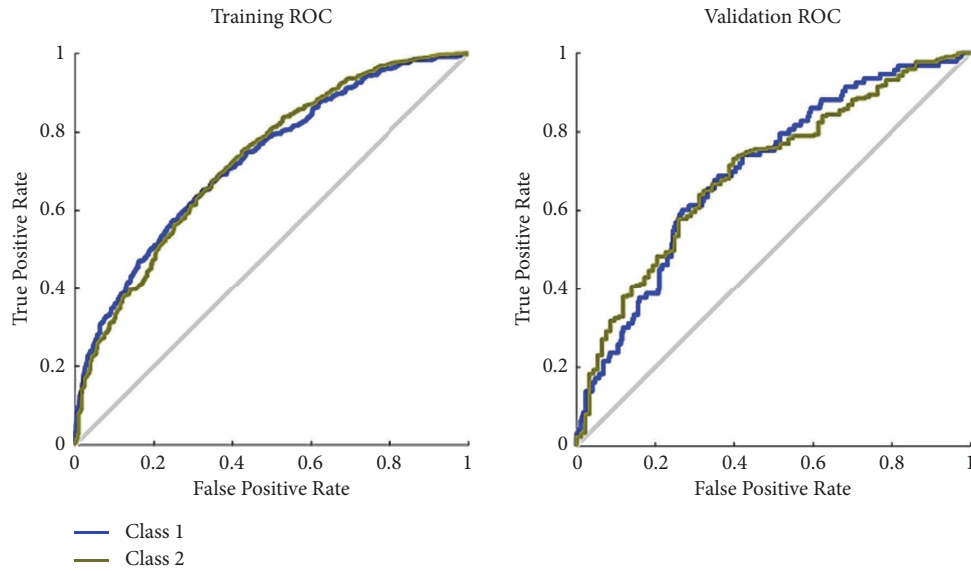


FIGURE 4: Sensitivity and specificity.

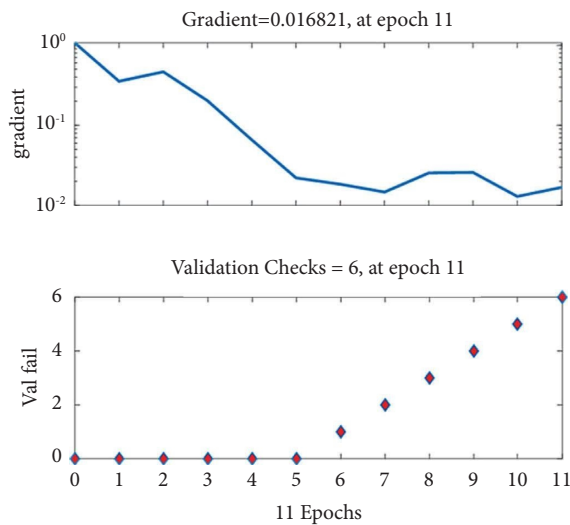


FIGURE 5: The amount of validation in terms of the number of epochs in the MLP model created by the SCG method.

Regression graphs for the output in relation to training, validating, and testing data are revealed in Figure 8. The correlation coefficient ( $R$ ) number was computed for each phase. The MLP model's overall response had an  $R$  value of roughly 0.46.

In order to recognize the validation error in the training window, Figure 9 indicates training, validation, and testing errors. The network at this iteration was returned because iteration 7 had the best validation efficiency. The network's mean square error is plotted in Figure 9 and is shown to be decreasing over time from a big value to a smaller value, which indicates that network learning is progressing. The network was trained using 70% of the vectors. The network's generalization was tested using 15% of those data points. As long as training lowers the network error on validation vectors, training vectors will persist. Figure 10 indicates the gradient of the MLP model in terms of the number of epochs. As shown in Figure 10, after eleven iterations, the gradient converges to the value equal to 0.23472; thus, in the modeling procedure, the number of

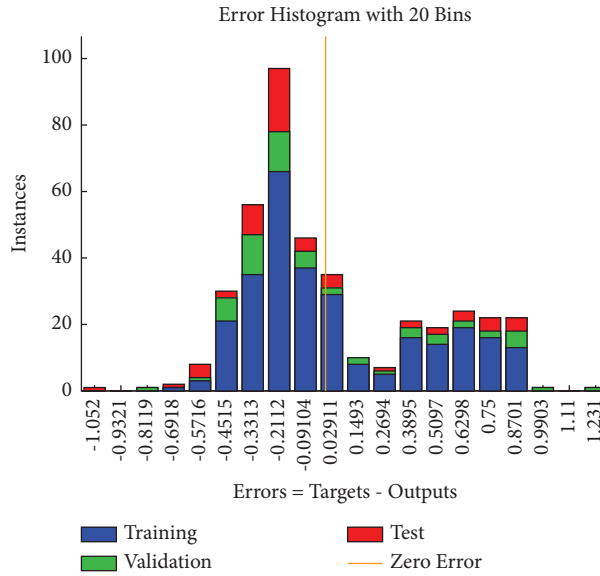


FIGURE 6: The histogram of train, validation, and test phases in the MLP model created by the SCG method.

TABLE 8: The comparison of the accurate percentage of the logit and the MLP models.

Model	Correct percentage		Overall
	Damage	Fatal and injury	
MLP	1.7	98.6	69.7
Logit regression	89.0	41.7	74.9

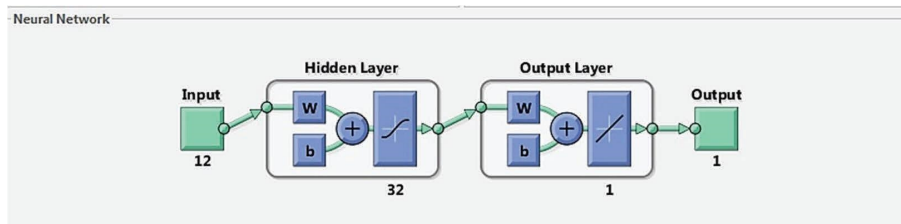


FIGURE 7: The architecture of the MLP model was developed using the LM method.

iterations was set to nine. It should be noted that each method can provide its result and none of them can be replaced with another one. In addition, the sigmoid was considered an active function for both MLP models. By comparing the values of MSE in both MLP models, it is indicated that the performance of the MLP built by the LM method was better in comparison to another one. Moreover, the histogram of the datasets in the MLP model created by the LM method is represented in Figure 11.

### 6. Validation of the Performance of Developed Models Using Other Mathematical Models

First of all, the primary aim of this section is to analyze the performance of the designed models using other kinds of models with different structures. In this regard, two kinds of MLP models were created using different methods and active functions which were considered to evaluate the developed models. The first model evaluated was the logistic

regression model. The error criteria of correct classification rate (CCR) and misclassification rate (MCR) were applied to verify the proposed models [25]. The MLP models were developed using the gradient descent (GD) and Mini-Batch methods. The CCR indicates the percentage of properly classified items. The number of individuals who are incorrectly classified as belonging to a group despite our knowledge that they do so is referred to as the misclassification rate or error. A lower value of root mean square error (RMSE) indicates that the developed model works better when comparing various models, whereas CCR exhibits the opposite pattern. The details of the models are shown in Table 9. The following algorithms display these indicators:

$$CCR = \left( \frac{\text{Number of correctly classified accidents}}{\text{Total number of observed accidents}} \right), \quad (1)$$

$$MCR = 1 - CCR.$$

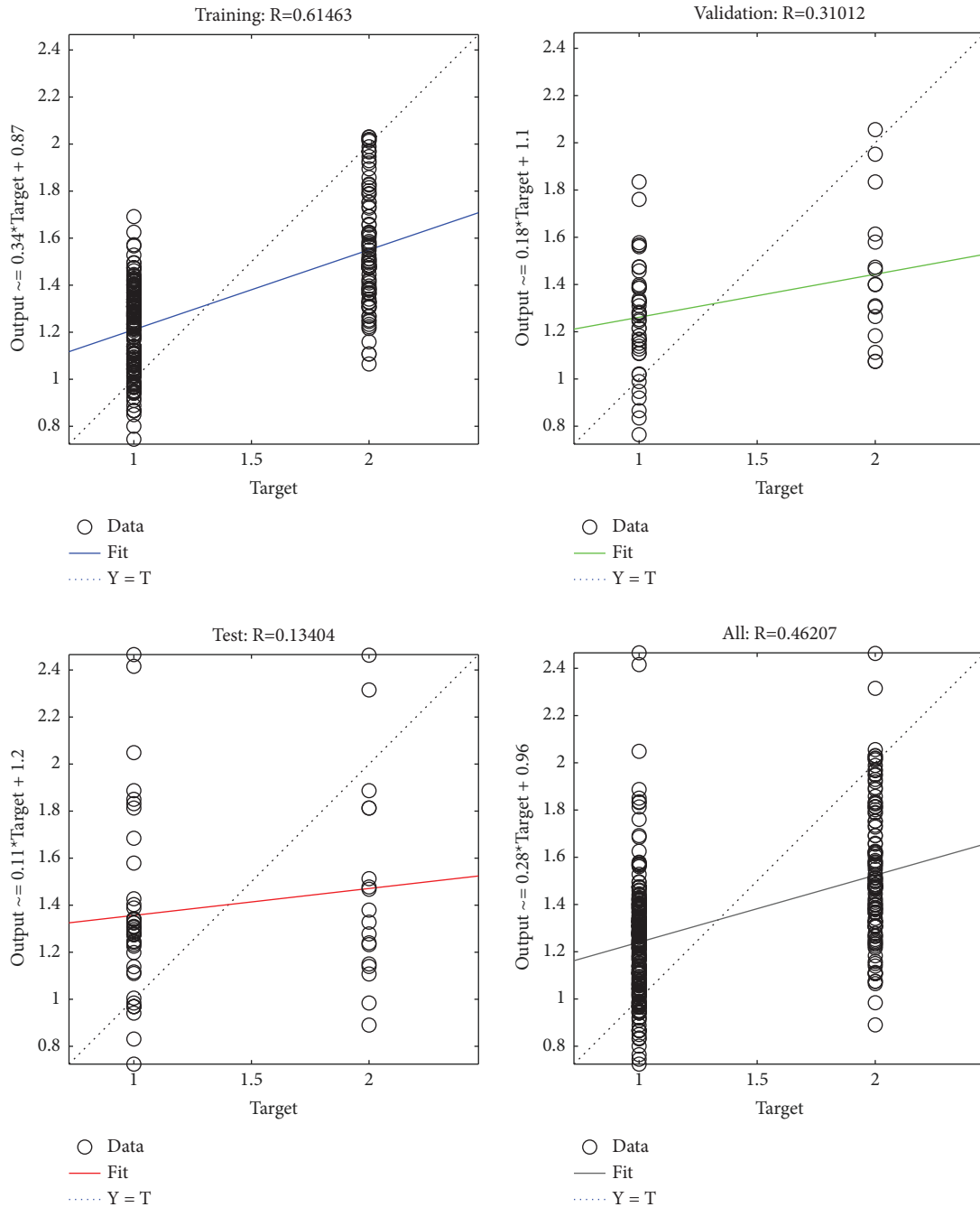


FIGURE 8: Regression plots for training, testing, and validation phases and the total response in the MLP model.

The comparison was made between MLP models created by GD and Mini-Batch methods for training neurons and the model MLP (M1) developed by SCG. Additionally, the hyperbolic tangent was used as an active function for MLP models, while the log-sigmoid was the active function used for the M2 model. As shown in Table 9, there is a slight difference between the results of the proposed models and other mathematical models developed for validation. In addition, the performance of the MLP (M2) model created by the LM algorithm was also considered by two kinds of MLP models. In this regard, the error criteria of RMSE and MSE were employed to provide this goal. The RMSE and MSE represent

the differences between predicted and actual values. These indicators are indicated in the following formulas:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (P_i - A_i)^2}{N}}, \quad (2)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^n (P_i - A_i)^2,$$

where  $P$  and  $A$  are the actual and predicted values, respectively, and  $N$  is the total quantity of data for the training set. A similar finding may be made by comparing other

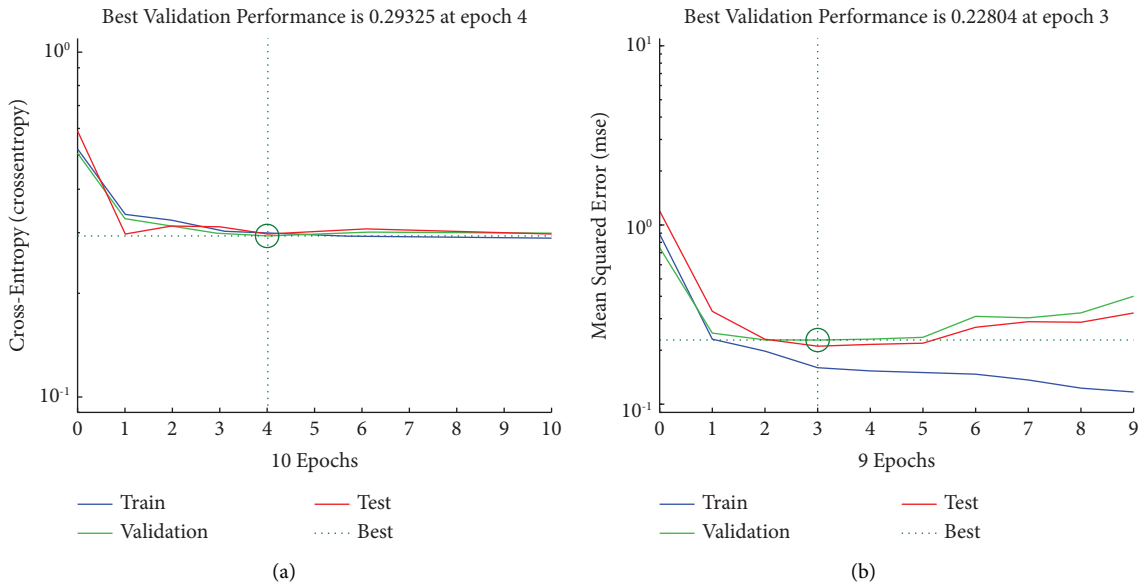


FIGURE 9: Neural network training performance. (a) The model was developed using the scale conjugate gradient (SCG) method. (b) The model was developed utilizing the Levenberg–Marquardt (LM) method.

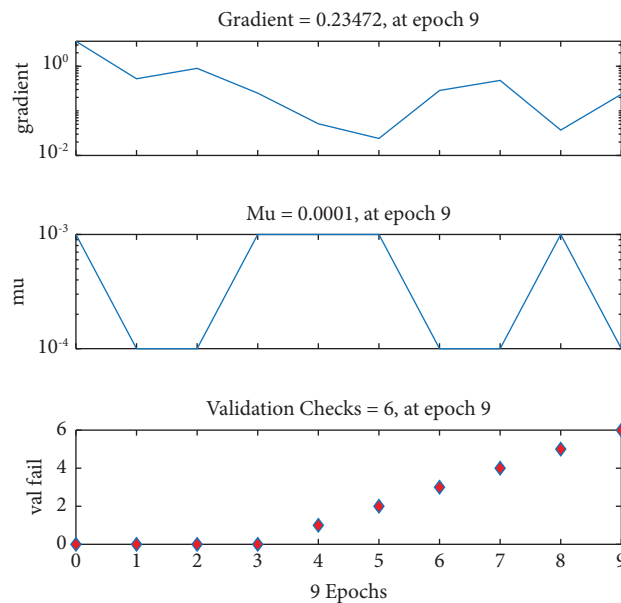


FIGURE 10: The amount of validation in terms of the number of epochs in the MLP model created by the LM method.

criteria such as RMSE and MSE. By comparing the results, it is found that the difference value between the proposed and created models was small. Thus, the performance of two kinds of MLP models designed using SCG and LM, as well as the logistic regression model, can be proven in terms of precision and efficiency, indicating that these models can be regarded as robust models in predicting accident severity.

### 7. Discussion

The major purpose of this research was to extract variables that could properly indicate accident location, the application of accident location, type of vehicle, collision with, type of

collision, vehicle facilities, the type of maneuver of the guilty vehicle, geometric condition, type of the shoulder of the accident location, human factor in accidents, marking of the accident location, and road factor. In addition, the logistic regression model was built to clarify the significant predictors of different accident levels. The results indicated that type of collision, collision with, and road factor variables are more likely to be involved in accidents, as emphasized in previous studies [26]. The type of maneuver of the guilty vehicle, including moving forward, turning to the left, turning to the right, and overtaking, has more likelihood to increase the severity of accidents, implying that the guilty vehicles are more likely to drive aggressively.



FIGURE 11: The histogram of train, validation, and test phases in the MLP model created by the LM method.

TABLE 9: Comparing the performance of mathematical models with proposed models.

Type of model	Learning method	CCR	Active function	MSE	RMSE	MCR
Logistic regression	—	74.9	—	—	—	25.1
MLP	Gradient descent (GD)	68.2	Tanh	—	—	31.8
MLP	Mini-Batch	70.5	Tanh	0.260	0.509	29.5
M1	Scale conjugate gradient (SCG)	69.7	Tanh	0.295	0.543	30.3
M2	Levenberg–Marquardt (LM)	—	Sigmoid	0.228	0.477	—

7.1. Variables Confirmed Vehicle Facilities, the Locations of Accidents, and Road Features. The Friedman test indicated five variables that have a significant effect on accident severity. These variables are related to vehicles, locations where accidents occurred, geometric conditions, and the marking of accident locations. Among the four derived variables, all were consistent with the vast majority of the earlier studies related to geometric aspects and vehicle facilities [27, 28]. Vehicle facilities indicate the ability to control as well as their brake system. Accident locations represent the necessity of the identification of black spots in terms of location and context conditions and proposing safety solutions. These two variables can be applied to evaluate vehicle characteristics and black spots.

Regarding the remaining two variables, the study emphasized more of the type of maneuver of the guilty vehicles under certain circumstances. Compared with the type of the shoulder of accident location extracted by shoulder inventory [29], the type of maneuver of the guilty vehicle has a wider range. This is because the type of shoulder of accident locations includes three types of common shoulder used in Iran; however, the type of maneuver of the guilty vehicle demonstrates various types of maneuver done by drivers related to the psychological actions and emotional states of drivers. Therefore, by taking into account both behavioral and emotional aspects, these two factors broaden the scope of the already-existing factors and aid in categorizing accident levels.

The research suggested that vehicle facilities can be considered as the cause of the type of maneuver of the guilty vehicles. When a vehicle has no special equipment like an Antilock Braking System (ABS) to control its ability, it could lead to dangerous maneuvers such as spirals. Both the location of accidents and geometric conditions can reflect road characteristics [30]. Road characteristics reflect visual characteristics in road design subjected to the perspective of drivers, important for road users and residents. The marking of accident locations and locations where accidents occurred refers more to the identification and improvement of black spots. The five factors, therefore, are interconnected and indicate the complexity of driving tasks. It is also challenging to develop a uniform definition of human factors due to the overlap of violations, vehicle facilities, the locations of accidents, and black spots.

7.2. Predictors of Accident Severity. In this study, a logistic regression model was implemented to analyze the capability of the application of accident locations, the type of vehicles, collision, vehicle facilities, geometric conditions, shoulder, and factors related to roads to predict the severity of accidents. It should be mentioned that the type of collision (rear-end) was the most influential factor in predicting the severity of accidents. Drivers, cars, and the environment are among the risk factor groupings that show strong evidence, and

together they all have a role in rear-end collisions [31]. Moreover, the type of maneuver of the guilty vehicle (moving forward and turning to right) has less effect in predicting the severity of accidents compared to other variables. However, the type of maneuver of the guilty vehicle, which includes turning to the left, as well as overtaking, is considered a significant factor for accident severity. Overall, it can be claimed that the type of maneuver of the guilty vehicle factor has the potential to predict the severity of accidents.

## 8. Conclusion

This study employed 12 variables to explore the relationship between accident severity and the considered factors. In order to provide the best result, several solutions can be suggested to improve security and lower the chance of accidents on certain roads. The factors influencing the severity of traffic accidents in Qom were looked into for this aim by applying statistical techniques such as the Friedman test, logit, and neural network models. The most significant outcomes are as follows:

- (i) The result of the Friedman test indicated that the most influential factors affecting traffic accidents were the reasons for vehicles, place of accidents, and geometry of accident location, indicating that the most influential factor was the reason of accident related to vehicle facilities and the second most important factor was accident location, affecting accidents.
- (ii) Based on the result of logistic regression, the best technique for creating the logit model of pedestrian accidents on rural roads was the model derived using the backward method with an accuracy percentage of 74.9 at step 1. According to the results of this model, the risk of traffic accidents increased for every unit of change in the independent variables of the guilty vehicle's type of maneuver (moving forward, turning to the left, turning to the right, and overtaking) and road factor (the presence of obstacles and bumps) and decreased for every unit of change in the variables with negative coefficients (collision with, type of collision, and type of collision).
- (iii) The results of MLP models indicated that the model developed using the LM method has less MSE value, indicating that it was a more accurate model in comparison to the model created by the SCG method. Also, the comparison of the MLP model created by the SCG method and the logit model indicated that the overall percentage of this model was better than the MLP model; however, the MLP model can be a reliable model for predicting fatal/injury accidents, in contrast to the result of the logit model. In addition, the performance of the developed models was validated using other mathematical models using CCR and MCR criteria, which showed that the performance of the models is reliable.

- (iv) This study, however, has some limitations. ANN models and statistical methods have some limitations to evaluate safety and prevent accidents. These methods might not be able to consider all the details of the safety problem. Therefore, it is recommended for future research to use connected and autonomous vehicles (CAVs), which are among the innovative technologies of intelligent transportation system (ITS) methods, successfully applied to assess mixed traffic flow [32, 33]. In addition, the estimation of the severity of conflicts by examining the vehicle paths, which might arise from the introduction of CAVs, will be accurately performed in the continuation of this research [34]. For future research works, various statistical analysis and modeling methods can be incorporated with the proposed approach [35–39]. Deep learning methods can be applied in the continuation of these studies [40–43]. Moreover, optimization algorithms are also recommended in the future [44–47]. Various validation methods are suggested in this regard, such as experimental tests, numerical simulations, analytical solutions, or comparative studies [48, 49]. Furthermore, it is recommended to use the Bayesian model averaging approach to overcome the model uncertainty [50, 51]. Traffic safety can be greatly compromised by pavement distress and surface characteristics, which can affect the drivers' lane-changing behavior and cause accidents [52, 53]. Traffic safety can also be affected by drivers' fatigue and performance, which can be explored in future studies [54–57]. Potential safety hazards can be identified on the road to improve transportation safety [58]. Road infrastructure and traffic congestion caused by new constructions can play a vital role in traffic-related accidents [59]. The presented approaches can also consider various factors related to the environment, roadside constructions, lighting conditions, climate change, and weather conditions [60–64].

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Disclosure

In this study, Iranian governmental organizations have not been partners and sponsors, and this study is purely studios.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Mohammad Habibzadeh was responsible for supervision, conceptualization, and methodology. Pooyan Ayar was

responsible for investigation and software. Mohammad Hassan Mirabimoghaddam was responsible for formal analysis and review and editing. Mahmoud Ameri was responsible for project administration and supervision. Seyede Mojde Sadat Haghghi was responsible for validation, visualization, and review and editing. All authors have contributed to the manuscript.

## Acknowledgments

We thank the Traffic Police for their cooperation in granting us access to accident data for the purpose of the present research. Also, the text of the manuscript has been checked by a native expert, and Grammarly software was used in this regard. Dr. Seyed Mohsen Hosseinian is thanked for his help in copyediting this paper.

## References

- [1] X. Zhang, S. Wen, L. Yan, J. Feng, and Y. Xia, "A hybrid-convolution spatial-temporal recurrent network for traffic flow prediction," *The Computer Journal*, vol. 171, 2022.
- [2] V. Najafi Moghaddam Gilani, M. Sashurpour, S. Hassanjani, and S. M. Hosseinian, "Analysis of the effect of the speed factor on highway safety using the machine learning method," *Slovak Journal of Civil Engineering*, vol. 29, no. 4, pp. 19–28, 2021.
- [3] J. Chen, Q. Wang, W. Peng, H. Xu, X. Li, and W. Xu, "Disparity-based multiscale fusion network for transportation detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18855–18863, 2022.
- [4] Q. Do, M. A. Moriyani, C. Le, and T. Le, "Cost-Weighted TF-IDF: a novel approach for measuring highway project similarity based on pay items' cost composition and term frequency," *Journal of Construction Engineering and Management*, vol. 149, no. 8, pp. 4023069–4023075, 2023.
- [5] A. Dehghani, M. Alidadi, and A. Sharifi, "Compact development policy and urban resilience: a critical review," *Sustainability*, vol. 14, no. 19, Article ID 11798, 2022.
- [6] A. Dehghani and A. Soltani, "Site selection of car parking with the GIS-based fuzzy multi-criteria decision making," *International Journal of Information Technology and Decision Making*, vol. 2023, pp. 1–26, 2023.
- [7] J. Chen, Q. Wang, H. H. Cheng, W. Peng, and W. Xu, "A review of vision-based traffic semantic understanding in ITSs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 19954–19979, 2022.
- [8] J. Chen, M. Xu, W. Xu, D. Li, W. Peng, and H. Xu, "A flow feedback traffic prediction based on visual quantified features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 10067–10075, 2023.
- [9] S. M. Hosseinian, V. Najafi Moghaddam Gilani, B. Mirbaha, and A. Abdi Kordani, "Statistical analysis for study of the effect of dark clothing color of female pedestrians on the severity of accident using machine learning methods," *Mathematical Problems in Engineering*, vol. 2021, Article ID 5567638, 21 pages, 2021.
- [10] A. S. Al-Ghamdi, "Using logistic regression to estimate the influence of accident factors on accident severity," *Accident Analysis & Prevention*, vol. 34, no. 6, pp. 729–741, 2002.
- [11] A. Pinto, R. A. Ribeiro, and I. L. Nunes, "Fuzzy approach for reducing subjectivity in estimating occupational accident severity," *Accident Analysis & Prevention*, vol. 45, pp. 281–290, 2012.
- [12] Đ. Petrović, R. Mijailović, and D. Pešić, "Traffic accidents with autonomous vehicles: type of collisions, manoeuvres and errors of conventional vehicles' drivers," *Transportation Research Procedia*, vol. 45, pp. 161–168, 2020.
- [13] T. F. Golob and W. W. Recker, "Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions," *Journal of Transportation Engineering*, vol. 129, no. 4, pp. 342–353, 2003.
- [14] S. P. Miaou and H. Lum, "Modeling vehicle accidents and highway geometric design relationships," *Accident Analysis & Prevention*, vol. 25, no. 6, pp. 689–709, 1993.
- [15] R. Elvik, F. Sagberg, and P. A. Langeland, "An analysis of factors influencing accidents on road bridges in Norway," *Accident Analysis & Prevention*, vol. 129, pp. 1–6, 2019.
- [16] A. H. Al-Balbissi, "Role of gender in road accidents," *Traffic Injury Prevention*, vol. 4, no. 1, pp. 64–73, 2003.
- [17] J. J. Rolison, S. Regev, S. Moutari, and A. Feeney, "What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records," *Accident Analysis & Prevention*, vol. 115, pp. 11–24, 2018.
- [18] T. Beshah and S. Hill, "Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia," in *Proceedings of the 2010 AAAI Spring Symposium Series*, New York, NY, USA, March 2010.
- [19] H. M. Hammad, M. Ashraf, F. Abbas et al., "Environmental factors affecting the frequency of road traffic accidents: a case study of sub-urban area of Pakistan," *Environmental Science and Pollution Research*, vol. 26, no. 12, pp. 11674–11685, 2019.
- [20] H. Mirzahosseini, M. Sashurpour, S. M. Hosseinian, and V. N. M. Gilani, "Presentation of machine learning methods to determine the most important factors affecting road traffic accidents on rural roads," *Frontiers of Structural and Civil Engineering*, vol. 16, no. 5, pp. 657–666, 2022.
- [21] M. R. Sheldon, M. J. Fillyaw, and W. D. Thompson, "The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs," *Physiotherapy Research International*, vol. 1, no. 4, pp. 221–228, 1996.
- [22] S. Ourang, "Evaluation of inter-organizational coordination of housing services in rural Alaska through social network analysis," 2022, <https://dr.lib.iastate.edu/server/api/core/bitstreams/7c3aa7ff-96b3-4874-8759-4240821e503c/content>.
- [23] M. G. Zamani, M. R. Nikoo, S. Jahanshahi, R. Barzegar, and A. Meydani, "Forecasting water quality variable using deep learning and weighted averaging ensemble models," *Environmental Science and Pollution Research*, vol. 30, no. 59, pp. 124316–124340, 2023.
- [24] K. Tehranian, "Can machine learning catch economic recessions using economic and market sentiments?" 2023, <https://arxiv.org/abs/2308.16200>.
- [25] A. M. Amiri, A. Sadri, N. Nadimi, and M. Shams, "A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers," *Accident Analysis & Prevention*, vol. 138, Article ID 105468, 2020.
- [26] L. Eboli, C. Forciniti, and G. Mazzulla, "Factors influencing accident severity: an analysis by road accident type," *Transportation Research Procedia*, vol. 47, pp. 449–456, 2020.
- [27] M. G. Karlaftis and I. Golias, "Effects of road geometry and traffic volumes on rural roadway accident rates," *Accident Analysis & Prevention*, vol. 34, no. 3, pp. 357–365, 2002.
- [28] M. M. Abdul Manan, A. Várhelyi, A. K. Çelik, and H. H. Hashim, "Road characteristics and environment factors



- associated with motorcycle fatal crashes in Malaysia,” *IATSS Research*, vol. 42, no. 4, pp. 207–220, 2018.
- [29] P. Zhao, J. Ma, C. Xu, C. Zhao, and Z. Ni, “Research on the safety of the left hard shoulder in a multi-lane highway based on safety performance function,” *Sustainability*, vol. 14, no. 22, Article ID 15114, 2022.
- [30] A. Sharf Aldeen, A. Abdi Kordani, A. Fallah, and S. M. Hosseinian, “Safety comparison of simple and spiral horizontal curves based on side friction factor dynamic modeling,” *Journal of Advanced Transportation*, vol. 2023, Article ID 7954346, 23 pages, 2023.
- [31] F. S. Didin and H. Iridiastadi, “Risk factors for rear-end collision: a systematic literature review IOP Conference Series: materials Science and Engineering,” *IOP Publishing*, vol. 909, no. 1, 2020.
- [32] X. Yang, Y. Zou, and L. Chen, “Operation analysis of freeway mixed traffic flow based on catch-up coordination platoon,” *Accident Analysis & Prevention*, vol. 175, Article ID 106780, 2022.
- [33] Y. Fang, H. Min, X. Wu, W. Wang, X. Zhao, and G. Mao, “On-ramp merging strategies of connected and automated vehicles considering communication delay,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15298–15312, 2022.
- [34] T. Miqdady, R. de Oña, and J. de Oña, “In search of severity dimensions of traffic conflicts for different simulated mixed fleets involving connected and autonomous vehicles,” *Journal of Advanced Transportation*, vol. 2023, Article ID 4116108, 21 pages, 2023.
- [35] D. Tong, Y. Sun, J. Tang, Z. Luo, J. Lu, and X. Liu, “Modeling the interaction of internal and external systems of rural settlements: the case of Guangdong, China,” *Land Use Policy*, vol. 132, Article ID 106830, 2023.
- [36] Y. Xiao and A. Konak, “The heterogeneous green vehicle routing and scheduling problem with time-varying traffic congestion,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 88, pp. 146–166, 2016.
- [37] G. Shen, C. Han, B. Chen, L. Dong, and P. Cao, “Fault analysis of machine tools based on grey relational analysis and main factor analysis,” *Journal of Physics: Conference Series*, vol. 1069, no. 1, Article ID 012112, 2018.
- [38] Z. Li, C. Han, and D. W. Coit, “System reliability models with dependent degradation processes,” in *Advances in Reliability and Maintainability Methods and Engineering Applications: Essays in Honor of Professor Hong-Zhong Huang on His 60th Birthday*, pp. 475–497, Springer Nature Switzerland, Cham, Switzerland, 2023.
- [39] M. Khodayari, J. Razmi, and R. Babazadeh, “An integrated fuzzy analytical network process for prioritisation of new technology-based firms in Iran,” *International Journal of Industrial and Systems Engineering*, vol. 32, no. 3, pp. 424–442, 2019.
- [40] L. Sun, J. Liang, C. Zhang, D. Wu, and Y. Zhang, “Meta-transfer metric learning for time series classification in 6G-supported intelligent transportation systems,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 2023, pp. 1–11, 2023.
- [41] H. Yin, Q. Wu, S. Yin, S. Dong, Z. Dai, and M. R. Soltanian, “Predicting mine water inrush accidents based on water level anomalies of borehole groups using long short-term memory and isolation forest,” *Journal of Hydrology*, vol. 616, Article ID 128813, 2023.
- [42] H. Yin, G. Zhang, Q. Wu et al., “A deep learning-based data-driven approach for predicting mining water inrush from coal seam floor using micro-seismic monitoring data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.
- [43] C. Han and X. Fu, “Challenge and opportunity: deep learning-based stock price prediction by using Bi-directional LSTM model,” *Frontiers in Business, Economics and Management*, vol. 8, no. 2, pp. 51–54, 2023.
- [44] T. Lin and H. Pham, “Reliability and cost-benefit analysis for two-stage intervened decision-making systems with interdependent decision units,” *International Journal of Mathematical, Engineering and Management Sciences*, vol. 4, no. 3, pp. 531–541, 2019.
- [45] T. Lin and H. Pham, “A two-stage intervened decision system with state-dependent random inspection mechanisms,” *IEEE Transactions on Computational Social Systems*, vol. 6, no. 2, pp. 365–376, 2019.
- [46] G. Shen, W. Zeng, Y. Zhang, C. Han, and P. Liu, “Determination of the average maintenance time of CNC machine tools based on type II failure correlation,” *Eksplatacja i Niezawodnosc Maintenance and Reliability*, vol. 19, no. 4, pp. 604–614, 2017.
- [47] T. Lin and H. Pham, “Modeling security surveillance systems with state dependent inspection-maintenance strategy,” *IEEE Transactions on Computational Social Systems*, vol. 10, no. 5, pp. 2467–2478, 2023.
- [48] A. Mehrjerd, H. Rezaei, S. Eslami, M. B. Ratna, and N. Khadem Ghaebi, “Internal validation and comparison of predictive models to determine success rate of infertility treatments: a retrospective study of 2485 cycles,” *Scientific Reports*, vol. 12, no. 1, p. 7216, 2022.
- [49] Y. Zhang, L. Mu, G. Shen, Y. Yu, and C. Han, “Fault diagnosis strategy of CNC machine tools based on cascading failure,” *Journal of Intelligent Manufacturing*, vol. 30, no. 5, pp. 2193–2202, 2019.
- [50] S. Wu, Y. Zou, L. Wu, and Y. Zhang, “Application of Bayesian model averaging for modeling time headway distribution,” *Physica A: Statistical Mechanics and its Applications*, vol. 620, Article ID 128747, 2023.
- [51] M. Achite, F. B. Banadkooki, M. Ehteram, A. Bouharira, A. N. Ahmed, and A. Elshafie, “Exploring Bayesian model averaging with multiple ANNs for meteorological drought forecasts,” *Stochastic Environmental Research and Risk Assessment*, vol. 36, no. 7, pp. 1835–1860, 2022.
- [52] Z. Luo, H. Wang, and S. Li, “Prediction of international roughness index based on stacking fusion model,” *Sustainability*, vol. 14, no. 12, p. 6949, 2022.
- [53] Z. Lin, H. Wang, and S. Li, “Pavement anomaly detection based on transformer and self-supervised learning,” *Automation in Construction*, vol. 143, Article ID 104544, 2022.
- [54] J. Xu, X. Zhang, S. H. Park, and K. Guo, “The alleviation of perceptual blindness during driving in urban areas guided by saccades recommendation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16386–16396, 2022.
- [55] J. Xu, S. H. Park, X. Zhang, and J. Hu, “The improvement of road driving safety guided by visual inattention blindness,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4972–4981, 2022.
- [56] J. Xu, K. Guo, X. Zhang, and P. Z. Sun, “Left Gaze Bias between LHT and RHT: a recommendation strategy to mitigate human errors in left-and right-hand driving,” *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 10, pp. 4406–4417, 2023.
- [57] J. Xu, K. Guo, and P. Z. Sun, “Driving performance under violations of traffic rules: novice vs. experienced drivers,”

- IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 4, pp. 908–917, 2022.
- [58] X. Ma, Z. Dong, W. Quan, Y. Dong, and Y. Tan, “Real-time assessment of asphalt pavement moduli and traffic loads using monitoring data from Built-in Sensors: optimal sensor placement and identification algorithm,” *Mechanical Systems and Signal Processing*, vol. 187, Article ID 109930, 2023.
- [59] H. Yuan and B. Yang, “System dynamics approach for evaluating the interconnection performance of cross-border transport infrastructure,” *Journal of Management in Engineering*, vol. 38, no. 3, Article ID 04022008, 2022.
- [60] S. Yu, C. Zhao, L. Song, Y. Li, and Y. Du, “Understanding traffic bottlenecks of long freeway tunnels based on a novel location-dependent lighting-related car-following model,” *Tunnelling and Underground Space Technology*, vol. 136, Article ID 105098, 2023.
- [61] F. Zhao, H. Wu, S. Zhu et al., “Material stock analysis of urban road from nighttime light data based on a bottom-up approach,” *Environmental Research*, vol. 228, Article ID 115902, 2023.
- [62] Z. Yin, Z. Liu, X. Liu, W. Zheng, and L. Yin, “Urban heat islands and their effects on thermal comfort in the US: New York and New Jersey,” *Ecological Indicators*, vol. 154, Article ID 110765, 2023.
- [63] X. Xie, B. Xie, J. Cheng, Q. Chu, and T. Dooling, “A simple Monte Carlo method for estimating the chance of a cyclone impact,” *Natural Hazards*, vol. 107, no. 3, pp. 2573–2582, 2021.
- [64] Y. Han, B. Wang, T. Guan et al., “Research on road environmental sense method of intelligent vehicle based on tracking check,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 1, pp. 1261–1275, 2023.