Front. Struct. Civ. Eng. 2022, 16(5): 657–666
https://doi.org/10.1007/s11709-022-0827-z

RESEARCH ARTICLE

Presentation of machine learning methods to determine the most important factors affecting road traffic accidents on rural roads

Hamid MIRZAHOSSEIN[a*], Milad SASHURPOUR[b], Seyed Mohsen HOSSEINIAN[a], Vahid Najafi Moghaddam GILANI[b]

[a] *Civil–Transportation Planning Department, Faculty of Technical and Engineering, Imam Khomeini International University (IKIU), Qazvin 34148-96818, Iran*

[b] *School of Civil Engineering, Iran University of Science and Technology (IUST), Tehran 13114-16846, Iran*

[*]*Corresponding author. E-mail: mirzahossein@eng.ikiu.ac.ir*

ABSTRACT The purpose of this research was to develop statistical and intelligent models for predicting the severity of road traffic accidents (RTAs) on rural roads. Multiple Logistic Regression (MLR) was used to predict the likelihood of RTAs. For more accurate prediction, Multi-Layer Perceptron (MLP) and Radius Basis Function (RBF) neural networks were applied. Results indicated that in MLR, the model obtained from the backward method with the correct percent of 84.7% and R^2 value of 0.893 was the best method for predicting the likelihood of RTAs. Also, MLR showed that the variables of not paying attention to the front not paying attention to the frontroad ahead, followed byand then vehicle-motorcycle/bike accidents were the greatest problems. Among the models, MLP had a better performance, so that the prediction accuracy of MLR, MLP, and RBF were 84.7%, 96.7%, and 92.1%, respectively. MLP model, due to higher accuracy, showed that the variable of reason of accident had the highest effect on the prediction of accidents, and considering MLR results, the variables of not paying attention to the front and then vehicle-motorcycle/bike accidents had the most influence on the occurrence of accidents. Therefore, motorcyclists and cyclists are more prone to accidents, and appropriate solutions should be adopted to enhance their safety.

KEYWORDS safety, rural accidents, multiple logistic regression, artificial neural networks

1 Introduction

Road Traffic Injuries (RTIs) and deaths are known as universal health concerns [1]. Thousands of people are injured or disabled annually. These losses present a considerable expense to families and societies in terms of health care, emergency response, and human grief [2]. Road traffic accidents (RTAs) are now globally the eighth main cause of death. By 2016, Road Traffic Fatalities (RTFs) had risen to 1.35 million a year, with up to 50 million injuries. Approximately 3700 persons die on roads every day around the world [3]. This amount varies in different parts of the world. On rural roads of Iran, the mortality rate from an accidental injury is nearly 4.33 per 1000 persons, and fatalities from unintentional road injuries are 5213 [4].

Many studies have been conducted on estimating and modeling RTAs and their consequences. They have identified the key factors affecting the increase in the severity of RTAs. Various techniques have also been used to explore the impact of these factors, among which, regression models are the most common methods used for RTA severity analysis, such as Multiple Logistic Regression (MLR) [5–7], Ordered Probit (OP) [8–11], ordered logit [12–15], mixed logit [16–19], etc. Moreover, Artificial Neural Network (ANN) methods have been widely used recently in traffic research. Compared to statistical methods, ANNs have been mostly applied to present more accurate prediction models because of their capability to handle more complex functions. ANNs use many different methods to deal with

Article history: Received Oct 30, 2021; Accepted Jan 24, 2022

classification problems and have been applied as predictive tools in groups of studies that used traditional methods such as regression models and then proposed ANNs to validate the results, which indicated that ANN is an effective alternative method with greater accuracy than traditional methods for predicting the severity of RTAs by comparing its performance with those of methods [20–22]. Some other studies have also studied other models of ANN in the investigation of accidents, such as Bayesian neural network [23–25], extreme learning machine [26,27], probabilistic neural network [28,29], and so forth. Omrani et al. [30] presented MLR, Multi-Layer Perceptron (MLP), Radius Basis Function (RBF) and Support Vector Machine (SVM) methods to forecast the travel mode of individuals in Luxembourg. They indicated that MLP was significantly the best model, followed by RBF, MLR, and SVM, respectively. Slimani et al. [31] investigated traffic forecasting in Morocco using ANN, including MLP, RBF, Adaline, and NoProp, to predict traffic flow. Results showed that MLP was the best method with the lowest error rate. Amin [32] applied MLP model to understand gender features of older drivers involved in RTAs and model the factors relating to older female and male driver accidents. Results showed that journey purpose was the greatest contributor to RTA likelihood for older drivers, and light condition was the second most important factor. Amiri et al. [33] had a comparison between MLP model (as ANN) and the hybrid intelligent genetic method to predict the severity of fixed-object RTAs in elderly drivers. In comparing these models, MLP revealed better performance than the other method and had more accuracy in low-severity RTAs. Finally, the light condition was recognized as the most critical factor, followed by the presence of left and right road shoulders.

The findings indicate a significant requirement for research and investments in rural infrastructures to enhance rural road safety. In other words, more research activities and analysis of RTAs and their causes are indispensable aspects of road management and design [34]. Therefore, the goals of this research are:

1) providing a better understanding of RTAs occurring on rural roads of Guilan province of Iran by frequency analysis of variables, including RTA severity and effective variables;

2) modeling various factors affecting RTA severity as a target variable;

3) comparison of RTA models and assessment of other related studies;

4) providing safety solutions to decrease the severity of traffic accidents on rural roads in Guilan province.

2 Methodology

2.1 Data collection

Data applied in the research were collected from the traffic police center of Guilan province, Iran. These data were related to 2602 RTAs, involving fatality, injury, or damage on rural roads of Guilan province from March 2017 to March 2020. The dependent variable in this study were various classes of severity of RTAs, and the target variables were presented in two classes of damage and injury/fatal RTAs. Table 1 describes the variables and their contribution to the occurrence of RTAs.

2.2 Multiple Logistic Regression

Regression analyses are statistical techniques for exploiting a set of predictor variables and modeling the relationships of a target and several predictor variables [35]. The logistic regression models are the improved regressions used for dichotomous dependent variables, and the sum of their probabilities will eventually be one. Let t be a linear function of the independent variable x_i [36]:

$$t = \beta_0 + \beta_1 x_1 + \cdots + \beta_M x_M + \varepsilon, \tag{1}$$

where β_i represents the regression coefficient, and ε represents the unobservable error. The binary logistic model will be in the form of Eq. (2) which represents the possibility of target variable [37]:

$$F(x) = \frac{1}{1 + \mathrm{e}^{-(\beta_0 + \beta_1 x_1 + \cdots + \beta_M x_M)}}. \tag{2}$$

2.3 Neural network

An ANN is based on the structure of the human brain and is made of a set of simple components called neurons in which parallel operations are performed [38–40]. The neurons receive the signals or inputs and, by summation function, determine the weighted average. Then, transfer functions process the network inputs and change them and determine the neurons’ outputs [41,42]. By comparing the predicted amounts with actual amounts of system, the error values are calculated and repeated with a series of stop conditions; this process is called training of the network [43,44]. The principal aim of the presentation of ANN methods is to describe the effects of predictor factors on the target factor [45]. The reason for prioritizing intelligent modeling methods in this study was their great performance in comparison with other models in identifying the importance of variables affecting the occurrence of RTAs. MLP and RBF models were applied in this study, and eventually, the modeling results were compared.

3 Results

3.1 Multiple Logistic Regression

The forward and backward stepwise methods were

Table 1 Specifications of variables in this research

description	values	percentage
accident severity	damage	15.76%
	injury/fatal	84.24%
accident time	00:00 to 06:00	4.92%
	06:00 to 12:00	21.87%
	12:00 to 18:00	43.7%
	18:00 to 24:00	29.52%
accident day	start of the week	25.78%
	middle of the week	40.55%
	weekends	33.67%
accident season	spring	28.71%
	summer	30.17%
	autumn	20.41%
	winter	20.71%
road surface condition	dry	87.32%
	wet	12.45%
	snowy	0.23%
daylight condition	day	71.02%
	night	27.82%
	sunset/sunrise	1.15%
type of vehicle accident	vehicle-vehicle	24.21%
	vehicle-heavy truck	0.54%
	vehicle-agricultural machinery	0.77%
	vehicle-motorcycle/bike	42.85%
	vehicle-animal	0.5%
	vehicle-object	6.57%
	getting out off the road	24.56%
driver age	less than or equal 18	2.92%
	18 to 30	36.4%
	30 to 45	38.28%
	45 to 60	16.48%
	60 and over	5.92%
driver gender	male	95.39%
	female	4.61%
weather condition	clear/sunny	79.01%
	cloudy	13.53%
	rainy	6.96%
	snowy	0.5%
reason of accident	not paying attention to the front	14.45%
	violation to left and right	34.74%
	violation of the speed limit	0.73%
	technical and safety tips defect	1.69%
	turning in a forbidden place	6.26%
	backward movement	2.23%
	failure to observe longitudinal and transverse spacing	2.31%
	right-of-way violation	17.06%
	inability to control	20.22%
	animal-caused accident	0.31%

applied to enter data into the logistic equation, a summary of which is presented in Table 2, which includes the criteria of goodness of fit (R^2) and correct percent.

According to Table 2, the backward stepwise method with the correct percent of 84.7% and R^2 value of 0.893 was selected as the best method to create the logit model of the severity of RTAs. More detailed results of the backward stepwise method are described in Table 3. The chi-square value is based on the ability to predict y values with and without x, and df is the number of independent variables in the regression model. Also, the significance level is the probability of rejecting the null hypothesis when it is true. Moreover, Table 4 shows the accuracy of the MLR model.

Table 2 Summary of the Multiple Logistic Regression methods

method	goodness of fit (R^2)	correct percent
forward stepwise	0.734	81.5
backward stepwise	0.893	84.7

Table 3 Backward stepwise model result

parameter		chi-square	df	significance level
step 15	step	–2.505	1	0.113
	block	539.423	21	0.000
	model	539.423	21	0.000

Table 4 Classification in MLR model

observed	predicted		
	accident severity		correct percent
	damage	injury/fatal	
accident severity			
damage	57	353	13.9
injury/fatal	44	2148	98.0
overall percent			84.7

Table 4 shows that among 410 damage RTAs, 57 cases, and among 2192 injury/fatal RTAs, 2148 cases, respectively, were predicted accurately by the MLR model. The prediction accuracy of MLR was 13.9% for damage RTAs and 98% for injury/fatal RTAs. The overall model accuracy in examining RTA severity was 84.7%.

The results of significant test of the parameters used and their effects on the regression model are presented in Table 5.

According to the outputs of the MLR model, for a unit of changing in the independent variables such as not paying attention to the front, vehicle-motorcycle/bike accident, violation to left and right, violation of the speed limit, male drivers, spring season, summer season, turning in a forbidden place, night, cloudy weather, technical and safety tips defect, failure to observe longitudinal and transverse spacing, right of way violation and inability to control, the likelihood of RTAs increased and vice versa for a unit of change in the variables with negative coefficients, including 00:00 to 06:00 time, vehicle-heavy truck accident and vehicle-object accident.

3.2 Multi-Layer Perceptron

For randomly splitting the dataset, the MLP model was created with 65% of the training dataset, and its accuracy was examined according to the rest of 35%. The number of independent variables as inputs was 10 in the model. Moreover, the number of neurons in the input and output

Table 5 Multiple logistic analysis result in the fifteenth step

variable	β	standard error	wald statistic	significance level	exp (β)
00:00 to 06:00	–0.707	0.256	7.627	0.006	0.493
spring	0.499	0.143	12.177	0.001	1.647
summer	0.436	0.084	26.941	0.000	1.547
night	0.282	0.077	13.413	0.000	1.326
vehicle-heavy truck	–0.602	0.298	4.081	0.044	0.548
vehicle-motorcycle/bike	0.766	0.068	126.894	0.000	2.151
vehicle-object	–0.129	0.032	16.251	0.000	0.879
male	0.528	0.258	4.188	0.041	1.696
cloudy	0.273	0.121	5.090	0.024	1.314
not paying attention to the front	0.894	0.340	6.914	0.008	2.445
violation to left and right	0.728	0.166	19.233	0.000	2.071
violation of the speed limit	0.564	0.244	5.343	0.021	1.758
technical and safety tips defect	0.264	0.127	4.321	0.037	1.302
turning in a forbidden place	0.375	0.094	15.915	0.000	1.455
failure to observe longitudinal and transverse spacing	0.244	0.085	8.240	0.004	1.276
right of way violation	0.182	0.045	16.358	0.000	1.200
inability to control	0.180	0.037	23.667	0.000	1.197
constant	–0.590	0.515	1.312	0.043	0.554

layers was 72 and 6, respectively. The automatic choice structure of the model selected 15 neurons in three hidden layers. Table 6 shows the correct percent of training and test samples.

Table 6 shows that 238 of 285 RTAs were correctly classified as damage ones, and 1502 of the 1531 RTAs were correctly predicted as injury/fatal ones. Overall, in 96.7% of cases, the training sample was shown to be accurate. Therefore, the accuracy of MLP method was 96.7%. Moreover, cross entropy error in training and test samples were 254.109 and 97.232, respectively.

Figure 1 represents the Receiver Operating Characteristic (ROC) curve. The higher the diagram to left and up, the power of the network is more valid in the prediction [46]. Moreover, the area under the curve for each category was 0.956, which indicates that the response from the model was positive.

Both Table 7 and Fig. 2 present the importance of independent variables on the dependent variable. As can be seen, the variables of reason of accident (25%), type of vehicle accident (13.5%), and accident season (13.1%) had the greatest impact on the severity of vehicle accidents, respectively. Therefore, this model showed that reason of accident had the highest effect on the occurrence of vehicle accidents.

Table 6 Classifications in MLP model

samples	observed	predicted		
		damage	injury/fatal	correct percent
training	damage	238	47	83.5%
	injury/fatal	29	1502	98.1%
	overall percent	74.3%	96.2%	96.7%
test	damage	113	12	90.4%
	injury/fatal	9	652	98.6%
	overall percent	85.4%	95.2%	97.3%

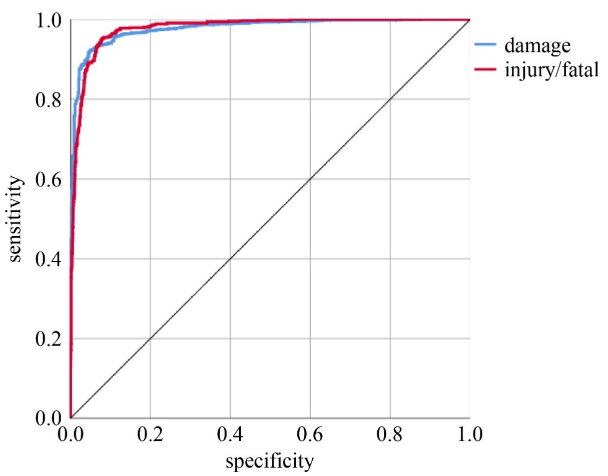

Fig. 1 ROC curve of MLP model.

Table 7 Independent variable importance for MLP model

variable	importance	normalized importance
accident time	0.085	33.9%
accident day	0.078	31.2%
accident season	0.131	52.5%
road surface condition	0.050	19.9%
daylight condition	0.086	34.4%
type of vehicle accident	0.135	53.9%
driver age	0.063	25.1%
driver gender	0.054	21.5%
weather condition	0.068	27.0%
cause of accident	0.250	100.0%

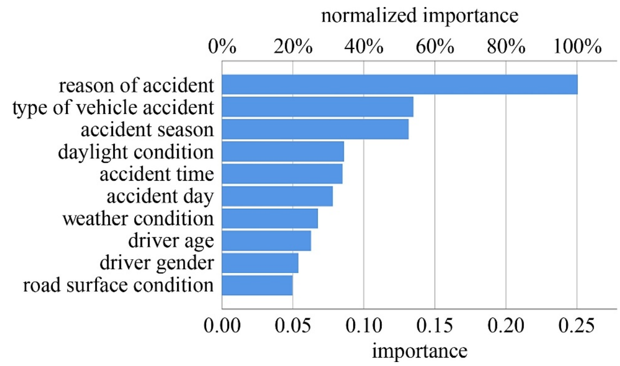

Fig. 2 Independent variable importance chart in MLP method.

3.3 Radius Basis Function

In this model, 65% of the data were used to train the network, and 35% of the data were applied for testing. The numbers of neurons in output and input layers were 4 and 67, respectively. Moreover, the automatic choice structure assigned 12 neurons in the hidden layer. The prediction accuracy of injury/fatal RTAs is presented in Table 8.

Results indicated that RBF model worked accurately in 92.1% of cases. This method was valid in identifying RTAs involving injuries/fatalities, but it performed poor in classifying RTAs resulting in damages. The sum-of-squares error for training and test samples of RBF model were 194.285 and 73.534, respectively.

Figure 3 indicates the ROC curve for the severity of RTAs. The area under the curve for each accident type was 0.929, which illustrates the power of the model in estimating the severity of accidents.

Table 9 shows the effect of the predictor variables on the target variable. Moreover, the importance of the variables affecting the severity of accidents is illustrated in Fig. 4.

Results of RBF model suggested that the most effective factor on the severity of RTAs in RBF method was reason of accident (27.8%). Then type of vehicle accident

Table 8 Classifications in RBF model

samples	observed	predicted		
		damage	injury/fatal	correct percent
training	damage	228	71	76.2%
	injury/fatal	96	1456	93.8%
	overall percent	71.2%	91.8%	92.1%
test	damage	94	17	84.6%
	injury/fatal	19	621	97.0%
	overall percent	81.2%	91.5%	95.4%

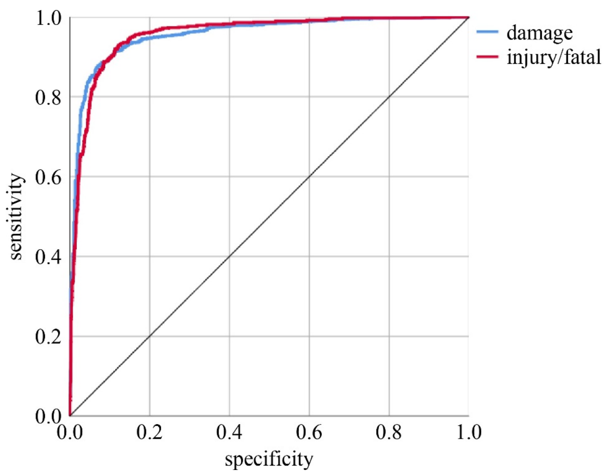

Fig. 3 ROC curve of RBF model.

Table 9 Independent variable importance for RBF model

variables	importance	normalized importance
accident time	0.084	30.1%
accident day	0.043	15.6%
accident season	0.075	27.0%
road surface condition	0.082	29.4%
daylight condition	0.085	30.5%
type of vehicle accident	0.116	41.8%
driver age	0.074	26.7%
driver gender	0.079	28.5%
weather condition	0.084	30.1%
reason of accident	0.278	100.0%

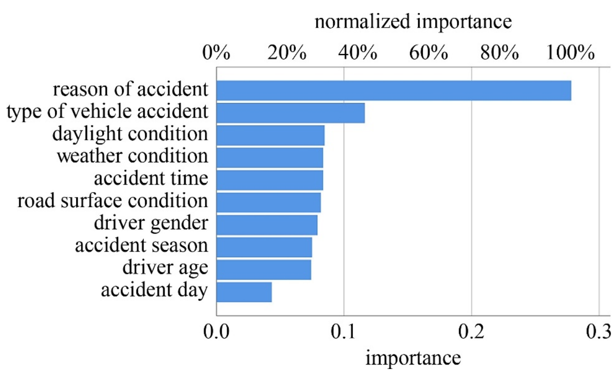

Fig. 4 Independent variable importance chart in RBF method.

(11.6%) and daylight condition (8.5%) had the most influence on the prediction, respectively.

4 Discussion

In this study, several models were used to determine more accurately the factors influencing the severity of RTAs. The comparison of correct percent in MLR, MLP, and RBF indicated that MLP method performed better and its prediction error was lower, which confirms the studies of Abdelwahab and Abdel-Aty [47], Chimba and Sando [48], Omrani [30], and Slimani et al. [31] and contradicts with the studies of Abdelwahab and Abdel-Aty [49] and Delen et al. [50]. MLR showed that not paying attention to the front, vehicle-motorcycle/bike accidents, violation to left and right, violation of the speed limit, male drivers, spring, summer, turning in a forbidden place, night, cloudy weather, technical and safety tips defect, failure to observe longitudinal and transverse spacing, right of way violation and inability to control respectively had the highest effect on increasing the likelihood of accidents, whereas accident time 12 p.m. to 6 a.m., vehicle-heavy truck accidents and vehicle-object accidents had a significant influence on reducing the probability of RTAs. While in MLP method as a superior method, reason of accident, type of vehicle accident and accident season were the most influential factors in accident occurrence. Also, in RBF model, reason of accident, type of vehicle accident, and daylight condition were the most critical factors. Consequently, reason of accident was the common point determined by these three models of accident prediction as the most influential factor.

Violation of the speed limit and road rule violation (i.e., violation to left and right, turning in a forbidden place and right of way violation) as behavioral factors were one of the powerful predictors of accident severity and confirms the study of Siskind et al. [51], Yan et al. [52] and Shrestha et al. [53]. Results showed that 00:00 to 06:00 reduced the likelihood of RTAs, which confirms the finding of Shrestha et al. [53]. However, night time had a positive effect on the occurrence of accidents, and this is contradictory with the studies of Yan et al. [52] and Shrestha et al. [53]. Male drivers and vehicle-motorcycle/bike accidents were found in this study to increase the likelihood of accidents, which is in line with the studies of Sherafati et al. [54] and Shrestha et al. [53] but is in contradiction with the research of Casado-Sanz et al. [55], which indicated that motorbike and female drivers had a negative effect on the likelihood of RTAs. Spring and then summer had a positive influence on the occurrence of accidents, which is contrary to the conclusion of Sherafati et al. [54], which showed autumn and then winter and summer had a negative impact on the likelihood of RTAs and supported the study of Intini et al.

[56], which indicated that summer had a positive effect for inexperienced drivers. Moreover, the impact of weather conditions on the occurrence of RTAs was very low, which contradicts previous studies that reported that snow and heavy rain conditions are the dominant causes of RTAs in rural areas [57,58]. On the other hand, vehicle-heavy truck accidents had a negative influence on the occurrence of RTAs, which contradicts the studies of Intini et al. [56] for inexperienced drivers and Yan et al. [52] for vehicle/light truck RTAs. Although in this research the effect of road surface conditions on the severity of RTAs was shown to be very small, some studies have shown a high impact of this factor [59–64]. Despite the studies of Amin [32] and Amiri et al. [33], using the MLP model, which respectively showed that journey purpose and lighting condition were significant contributing factors in the RTAs, the MLP model as used this study showed that reason of accident and type of vehicle accident were the most critical factors affecting accidents, respectively. Briefly, the performance of each of the models used in this study is compared in Table 10. Table 10 shows the MLP model had the best performance in predicting the severity of RTAs on rural roads in Guilan province. However, other machine learning methods and optimization techniques can also be applied in future studies to be incorporated into the results in order to represent more accurate results [65–67].

5 Conclusions

In this study, by investigating and modeling the factors influencing the severity of RTAs on rural roads of Guilan province, Iran, the most critical factor in the occurrence of vehicle accidents was determined using the MLR, MLP, and RBF models. The main results of this study are as follows.

1) In the MLR method, the impact of independent variables on the severity of RTAs was examined, and the model obtained from the backward method with 84.7 correct percent at step 15 was the best method for predicting the likelihood of RTAs on rural roads of Guilan province.

2) According to the MLR results, the variables of not paying attention to the front, vehicle-motorcycle/bike accident, violation to left and right, violation of the speed limit, male drivers, spring season, summer season, turning in a forbidden place, night, cloudy weather, technical and safety tips defect, failure to observe longitudinal and transverse spacing, right of way violation and inability to control increased the probability of vehicle RTAs, respectively. Also, the variable of 00:00 to 06:00 accident time, vehicle-heavy truck accident, and vehicle-object accident reduced the probability of vehicle RTAs, respectively.

3) One of the most significant results of the MLR model was that not paying attention to the front had the highest impact on increasing the likelihood of RTAs, followed by vehicle-motorcycle/bike accidents. Therefore, more police presence is necessary to reduce the extent of RTAs (due to violation of the speed limit and violation to left and right), especially in spring and summer season and using helmets may have a direct influence on decreasing vehicle-motorcycle/bike accidents, which makes it more significant at nights (due to reduced driver visibility and not paying attention to the front) and male drivers which had the highest effect on increasing the RTA rates.

4) The MLR method also indicated that the interactive influence of the darkness of weather and vehicle-motorcycle/bike accidentsincreased the likelihood of RTAs. This increase may be because motorcyclists/cyclists in Iran tend to not wear helmets and drivers are less likely to see them, especially at night time, which increases the likelihood of the motorcyclist/cyclist accidents.

5) In ANN models for predicting RTAs and determining the effect of each variable, the result of the algorithms applied indicated that the MLP model was more accurate compared to the RBF model, showing that MLP was better at the prediction of RTA severity.

6) Sensitivity analysis of the two ANN algorithms revealed that MLP had greater power in the prediction of accidents in comparison with the RBF method.

7) In MLP, the variables of reason of accident, type of vehicle accident, and accident season had the most critical effect on the severity of RTAs, while in RBF, the variables of reason of accident, type of vehicle accident, and daylight condition had the greatest effect, respectively.

8) In general, MLP results due to higher accuracy showed that reason of accident had the most significant impact on the prediction of accidents and considering MLR model, not paying attention to the front, followed by vehicle-motorcycle/bike accident had the most influence on the occurrence of RTAs.

Table 10 The comparison of the models used in this research

methods	model prediction accuracy (%)	three most important factors
MLR	84.7	1. not paying attention to the front
		2. vehicle-motorcycle/bike
		3. violation to left and right
MLP	96.7	1. reason of accident
		2. type of vehicle accident
		3. accident season
RBF	92.1	1. reason of accident
		2. type of vehicle accident
		3. daylight condition

References

1. Banstola A, Kigozi J, Barton P, Mytton J. Economic burden of road traffic injuries in Nepal. International Journal of Environmental Research and Public Health, 2020, 17(12): 4571
2. Zou Y, Zhang Y, Cheng K. Exploring the impact of climate and extreme weather on fatal traffic accidents. Sustainability (Basel), 2021, 13(1): 390
3. World Health Organization. Global Status Report on Road Safety 2018. 2018
4. Heydari S, Hoseinzadeh A, Ghaffarpasand F, Hedjazi A, Zarenezhad M, Moafian G, Aghabeigi M R, Foroutan A, Sarikhani Y, Peymani P, Ahmadi S M, Joulaei H, Dehghankhalili M, Lankarani K B. Epidemiological characteristics of fatal traffic accidents in Fars province, Iran: A community-based survey. Public Health, 2013, 127(8): 704−709
5. K Abledu J, Abledu G. Multiple logistic regression analysis of predictors of musculoskeletal disorders and disability among bank workers in Kumasi. Journal of Ergonomics, 2012, 2(4): 2
6. Wang H Y, Peng J, Wang B K, Lu X, Zheng J Z, Wang K J, Tu X M, Feng C Y. Inconsistency between univariate and multiple logistic regressions. Shanghai Archives of Psychiatry, 2017, 29(2): 124
7. Zhang H, Xing H, Yao D, Liu L, Xue D, Guo F. The multiple logistic regression recognition model for mine water inrush source based on cluster analysis. Environmental Earth Sciences, 2019, 78(20): 612
8. Kwigizile V, Sando T, Chimba D. Inconsistencies of ordered and unordered probability models for pedestrian injury severity. Transportation Research Record: Journal of the Transportation Research Board, 2011, 2264(1): 110−118
9. Garrido R, Bastos A, de Almeida A, Elvas J P. Prediction of road accident severity using the ordered probit model. Transportation Research Procedia, 2014, 3: 214−223
10. Fountas G, Anastasopoulos P C. Analysis of accident injury-severity outcomes: The zero-inflated hierarchical ordered probit model with correlated disturbances. Analytic Methods in Accident Research, 2018, 20: 30−45
11. Piedra-Bonilla E B, da Cunha D A, Braga M J. Climate variability and crop diversification in Brazil: An ordered probit analysis. Journal of Cleaner Production, 2020, 256: 120252
12. Barua U, Tay R. Severity of urban transit bus crashes in Bangladesh. Journal of Advanced Transportation, 2010, 44(1): 34−41
13. Yasmin S, Eluru N, Ukkusuri S V. Alternative ordered response frameworks for examining pedestrian injury severity in New York City. Journal of Transportation Safety & Security, 2014, 6(4): 275−300
14. Zeng Q, Gu W, Zhang X, Wen H, Lee J, Hao W. Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors. Accident Analysis and Prevention, 2019, 127: 87−95
15. Azimi G, Rahimi A, Asgari H, Jin X. Severity analysis for large truck rollover crashes using a random parameter ordered logit model. Accident Analysis and Prevention, 2020, 135: 105355
16. Ye F, Lord D. Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. Analytic Methods in Accident Research, 2014, 1: 72−85
17. Haleem K, Alluri P, Gan A. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. Accident Analysis and Prevention, 2015, 81: 14−23
18. Chen F, Chen S, Ma X. Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data. Journal of Safety Research, 2018, 65: 153−159
19. Liu P, Fan W. Exploring injury severity in head-on crashes using latent class clustering analysis and mixed logit model: A case study of North Carolina. Accident Analysis and Prevention, 2020, 135: 105388
20. Chang L Y. Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network. Safety Science, 2005, 43(8): 541−557
21. Abdulhafedh A. Crash frequency analysis. Journal of Transportation Technologies, 2016, 6(4): 169−180
22. Mokhtarimousavi S, Anderson J C, Azizinamini A, Hadi M. Factors affecting injury severity in vehicle-pedestrian crashes: A day-of-week analysis using random parameter ordered response models and artificial neural networks. International Journal of Transportation Science and Technology, 2020, 9(2): 100−115
23. Riviere C, Lauret P, Ramsamy J F M, Page Y. A Bayesian neural network approach to estimating the energy equivalent speed. Accident Analysis and Prevention, 2006, 38(2): 248−259
24. Xie Y, Lord D, Zhang Y. Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. Accident Analysis and Prevention, 2007, 39(5): 922−933
25. de Oña J, Mujalli R O, Calvo F J. Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks. Accident Analysis and Prevention, 2011, 43(1): 402−411
26. Behbahani H, Amiri A M, Imaninasab R, Alizamir M. Forecasting accident frequency of an urban road network: A comparison of four artificial neural network techniques. Journal of Forecasting, 2018, 37(7): 767−780
27. Zhu W, Miao J, Hu J, Qing L. Vehicle detection in driving simulation using extreme learning machine. Neurocomputing, 2014, 128: 160−165
28. Parsa A B, Taghipour H, Derrible S, Mohammadian A K. Real-time accident detection: Coping with imbalanced data. Accident Analysis and Prevention, 2019, 129: 202−210
29. Yu B, Chen Y, Bao S, Xu D. Quantifying drivers' visual perception to analyze accident-prone locations on two-lane mountain highways. Accident Analysis and Prevention, 2018, 119: 122−130
30. Omrani H. Predicting travel mode of individuals by machine learning. Transportation Research Procedia, 2015, 10: 840−849
31. Slimani N, Slimani I, Sbiti N, Amghar M. Traffic forecasting in Morocco using artificial neural networks. Procedia Computer Science, 2019, 151: 471−476
32. Amin S. Backpropagation—Artificial Neural Network (BP-ANN): Understanding gender characteristics of older driver accidents in

West Midlands of United Kingdom. Safety Science, 2020, 122: 104539

33. Amiri A M, Sadri A, Nadimi N, Shams M. A comparison between Artificial Neural Network and Hybrid Intelligent Genetic Algorithm in predicting the severity of fixed object crashes among elderly drivers. Accident Analysis and Prevention, 2020, 138: 105468
34. Richter T, Ruhl S, Ortlepp J, Bakaba E. Prevention of overtaking accidents on two-lane rural roads. Transportation Research Procedia, 2016, 14: 4140−4149
35. Lombardo L, Cama M, Conoscenti C, Märker M, Rotigliano E. Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, southern Italy). Natural Hazards, 2015, 79(3): 1621−1648
36. Saeidi S, Mohammadzadeh M, Salmanmahiny A, Mirkarimi S H. Performance evaluation of multiple methods for landscape aesthetic suitability mapping: A comparative study between Multi-Criteria Evaluation, Logistic Regression and Multi-Layer Perceptron neural network. Land Use Policy, 2017, 67: 1−12
37. Wang S H, Zhan T M, Chen Y, Zhang Y, Yang M, Lu H M, Wang H N, Liu B, Phillips P. Multiple sclerosis detection based on biorthogonal wavelet transform, RBF kernel principal component analysis, and logistic regression. IEEE Access: Practical Innovations, Open Solutions, 2016, 4: 7567−7576
38. Yu H, Tao J, Qin C, Liu M, Xiao D, Sun H, Liu C. A novel constrained dense convolutional autoencoder and DNN-based semi-supervised method for shield machine tunnel geological formation recognition. Mechanical Systems and Signal Processing, 2022, 165: 108353
39. Jin Y, Qin C, Tao J, Liu C. An accurate and adaptative cutterhead torque prediction method for shield tunneling machines via adaptative residual long-short term memory network. Mechanical Systems and Signal Processing, 2022, 165: 108312
40. Zhou W, Liu J, Lei J, Yu L, Hwang J N. GMNet: Graded-feature multilabel-learning network for RGB-thermal urban scene semantic segmentation. IEEE Transactions on Image Processing, 2021, 30: 7790−7802
41. Li Y, Che P, Liu C, Wu D, Du Y. Cross-scene pavement distress detection by a novel transfer learning framework. Computer-Aided Civil and Infrastructure Engineering, 2021, 36(11): 1398−1415
42. Lv Z, Li Y, Feng H, Lv H. Deep Learning for security in digital twins of cooperative intelligent transportation systems. IEEE Transactions on Intelligent Transportation Systems, 2021: 1−10
43. Alizadeh Kharazi B, Behzadan A H. Flood depth mapping in street photos with image processing and deep neural networks. Computers, Environment and Urban Systems, 2021, 88: 101628
44. Alizadeh B, Li D, Zhang Z, Behzadan A H. Feasibility study of urban flood mapping using traffic signs for route optimization. 2021, arXiv:2109.11712
45. Yamany M S, Saeed T U, Volovski M, Ahmed A. Characterizing the performance of interstate flexible pavements using artificial neural networks and random parameters regression. Journal of Infrastructure Systems, 2020, 26(2): 04020010
46. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters, 2006, 27(8): 861−874
47. Abdelwahab H T, Abdel-Aty M A. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. Transportation Research Record: Journal of the Transportation Research Board, 2001, 1746(1): 6−13
48. Chimba D, Sando T. The prediction of highway traffic accident injury severity with neuromorphic techniques. Advances in Transportation Studies, 2009, 19: 17−26
49. Abdelwahab H T, Abdel-Aty M A. Artificial neural networks and logit models for traffic safety analysis of toll plazas. Transportation Research Record: Journal of the Transportation Research Board, 2002, 1784(1): 115−125
50. Delen D, Sharda R, Bessonov M. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. Accident Analysis and Prevention, 2006, 38(3): 434−444
51. Siskind V, Steinhardt D, Sheehan M, O’Connor T, Hanks H. Risk factors for fatal crashes in rural Australia. Accident Analysis and Prevention, 2011, 43(3): 1082−1088
52. Yan X, Radwan E, Abdel-Aty M. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. Accident Analysis and Prevention, 2005, 37(6): 983−995
53. Shrestha P P, Shrestha K J. Factors associated with crash severities in built-up areas along rural highways of Nevada: A case study of 11 towns. Journal of Traffic and Transportation Engineering (English Edition), 2017, 4(1): 96−102
54. Sherafati F, Rad E H, Afkar A, Gholampoor-Sigaroodi R, Sirusbakht S. Risk factors of road traffic accidents associated mortality in northern iran: A single center experience utilizing oaxaca blinder decomposition. Bulletin of Emergency and Trauma, 2017, 5(2): 116
55. Casado-Sanz N, Guirao B, Gálvez-Pérez D. Population ageing and rural road accidents: Analysis of accident severity in traffic crashes with older pedestrians on Spanish crosstown roads. Research in Transportation Business & Management, 2019, 30: 100377
56. Intini P, Berloco N, Colonna P, Ranieri V, Ryeng E. Exploring the relationships between drivers’ familiarity and two-lane rural road accidents: A multi-level study. Accident Analysis and Prevention, 2018, 111: 280−296
57. Pourfalatoun S, Miller E E. User perceptions of automated Truck-Mounted attenuators: Implications on work zone safety. Traffic Injury Prevention, 2021, 22(5): 413
58. Miller E E, Pourfalatoun S. Evaluating the Human-Automated Maintenance Vehicle Interaction for Improved Safety and Facilitating Long-Term Trust. Denver: Applied Research and Innovation Branch, 2021
59. Waseem M, Ahmed A, Saeed T U. Factors affecting motorcyclists’ injury severities: An empirical assessment using random parameters logit model with heterogeneity in means and variances. Accident Analysis and Prevention, 2019, 123: 12−19
60. Chen S, Saeed T U, Alinizzi M, Lavrenz S, Labi S. Safety sensitivity to roadway characteristics: A comparison across highway classes. Accident Analysis and Prevention, 2019, 123: 39−50

61. Ahmad N, Ahmed A, Wali B, Saeed T U. Exploring factors associated with crash severity on motorways in Pakistan. Proceedings of the Institution of Civil Engineers—Transport, 2020
62. Chen S, Saeed T U, Alqadhi S D, Labi S. Safety impacts of pavement surface roughness at two-lane and multi-lane highways: Accounting for heterogeneity and seemingly unrelated correlation across crash severities. Transportmetrica A: Transport Science, 2019, 15(1): 18−33
63. Chen S, Saeed T U, Labi S. Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. Analytic Methods in Accident Research, 2017, 16: 75−89
64. Saeed T U, Hall T, Baroud H, Volovski M J. Analyzing road crash frequencies with uncorrelated and correlated random-parameters count models: An empirical assessment of multilane highways. Analytic Methods in Accident Research, 2019, 23: 100101
65. Tong D, Yuan Y, Wang X. The coupled relationships between land development and land ownership at China's urban fringe: A structural equation modeling approach. Land Use Policy, 2021, 100: 104925
66. Dong H, Zhao B, Deng Y. Instability phenomenon associated with two typical high speed railway vehicles. International Journal of Non-linear Mechanics, 2018, 105: 130−145
67. Bie Y, Ji J, Wang X, Qu X. Optimization of electric bus scheduling considering stochastic volatilities in trip travel time and energy consumption. Computer-Aided Civil and Infrastructure Engineering, 2021, 36(12): 1530−1548