

Forecasting the Accident Frequency and Risk Factors: A Case Study of Erzurum, Turkey

Mohammad Ali SAHRAEI*, Merve Kayacı ÇODUR, Muhammed Yasin ÇODUR, Ahmet TORTUM

Abstract: Nowadays, life is intimately associated with transportation, generating several issues on it. Numerous works are available concerning accident prediction techniques depending on independent road and traffic features, while the mix parameters including time, geometry, traffic flow, and weather conditions are still rarely ever taken into consideration. This study aims to predict future accident frequency and the risk factors of traffic accidents. It utilizes the Generalized Linear Model (GLM) and Artificial Neural Networks (ANN) approaches to process and predict traffic data efficiently based on 21500 records of traffic accidents that occurred in Erzurum in Turkey from 2005 to 2019. The results of the comparative evaluation demonstrated that the ANN model outperformed the GLM model. The study revealed that the most effective variable was the number of horizontal curves. The annual average growth rates of accident occurrences based on the ANN's method are predicted to be 11.22% until 2030.

Keywords: accident frequency; artificial neural network; forecasting; generalized linear model; risk factors; traffic accident

1 INTRODUCTION

In the modern world, the quick growth of urbanization has led to the boom of automobiles, creating a range of issues including traffic accidents, air pollution, and traffic jam. All over the world, highway accidents create a great deal of economic burden for societies; and thus, highway safety improvement is regarded as a crucial issue for transportation engineering.

Due to their wide range of applications and important practical implications, traffic safety studies on highways have been attractive to engineers and planners for almost some decades. Although the studies which have been done on road safety have led to decreasing traffic accidents, highway traffic accidents are still considered to be one of the greatest challenging issues in Turkey both from the public health and socio-economic perspectives. The researches into Accident Prediction Models (APMs) that are done for this reason are improving rapidly and they have been preferred for predicting road safety compared to mathematical tools [1, 2].

Several independent factors impact accident frequency, and these are commonly associated with infrastructure geometric characteristics, section length, traffic flow, driver behavior, lighting, pavement surface conditions, and weather. The value of particular factors is occasionally tough to analyze, and the impact of such parameters on road crashes may not be similarly substantial. Therefore, from a broader set of independent parameters, experts commonly extract a diminished amount of factors for inclusion in the suggested model. Road geometry, traffic flow and seatbelt wearing behavior, for instance, have been identified as the main reasons for accidents [3, 4]. However, extensive literature [5-8] is available concerning accident prediction techniques depending on independent road and traffic features, while the mix factors including time, geometry, traffic flow and weather conditions are still rarely ever taken into consideration.

Based on the literature, prediction methods have been established by utilizing analysis techniques which may essentially be arranged into four major techniques, specifically multivariate analysis [9, 10], Empirical Bayes Method [11, 12], Fuzzy Logic [13-15] and Neural Network [16-18].

Roque and Cardoso [19] say that among these techniques, the multivariate analysis has a long and consolidated utilization in crash evaluation. The early techniques established with this particular method depended on multiple linear regression, while the presently utilized method is developed according to the Generalized Linear Model (GLM) approach, which enables to extend the linear modeling to stochastic factors that are generally not distributed with a constant variance. Wedderburn [20] presented the notion of quasi-likelihood technique, which prevents clearly indicating the distribution of the response, needing just to identify the link and variance functions. This particular notion was expanded to the idea of generalized calculating equations with specific application to longitudinal data by Liang and Zeger [21]. Lindsey [22] provides a broad variety of applications of GLMs method while Fahrmeir and Tutz [23] explain the GLMs method for multivariate responses.

Presently, GLM is utilized nearly specifically for the advancement of APMs by several experts such as Lord and Mannering [24], Lu et al. [25], Glavić et al. [26], Mohammadi et al. [27], Ozfirat et al. [28], Gianfranco et al. [29] and they have shown the specific regular circumstances under which conventional linear regression modeling is suitable.

Artificial Neural Network (ANN) approaches are functions that are identified utilizing multilevel network structures. They are made of a series of weights and nodes that connect several nodes together in a hierarchical manner including input, output, and hidden layers. A number of researches have examined (ANN) technique in traffic and transportation associated applications such as Yuan et al. [16], Çodur and Tortum [1, 30], Sahraei [31], Wang et al. [32], Singh et al. [33], Li et al. [34]. Abdelwahab and Abdel-Aty [35] utilized ANN to anticipate driver injury severity from several crash variables, i.e. driver, automobile, environment, and road features. In their particular research, the ANN technique established carried out much better compared to the ordered Probit technique. In another research, Delen et al. [36] implemented ANN to model injury severity of road crashes utilizing 17 considerable variables.

Several experts have refined prediction approaches, and they have also pointed out their particular operating limitations, and highlighted questions without any answer.

Nevertheless, these approaches have completely been established in countries where traffic characteristics, human behavior, infrastructure, and crash surveys vary compared to those in Turkey. In addition to variables such as traffic flow and geometry features, earlier researches did not often focus on factors such as seasonal data and weather conditions. Consequently, there are at least three major purposes for justifying this particular research. The first is, given the variety of the aforementioned approaches, a suitable option for modeling accidents needs to be recommended. The second is motivated by the requirement to evaluate the anticipated number of accidents for multilane highways of almost all the subsequent parameters such as traffic flow, road geometry, seasonal data and weather condition. The final purpose is for recommending countermeasures suitable for enhancing road safety. For instance, it may be much better comprehended whether or not geometry adjustment associated with horizontal or vertical curves decreases accidents, and specific season and weather has a beneficial impact on road safety.

The objective is to develop a particular prediction model to calculate future accident frequency as a function of traffic flow, road geometry, seasonal data and weather condition as well as to determine the importance and effectiveness of the risk factors in accidents. This research is carried out by applying both models of generalized linear model and ANN.

This paper is organized as follows. Section 2 provides a literature review related to the existing research about traffic accident prediction models with GLM and ANN. Section 3, methodology based on the GLM and ANN is presented. Results and discussion are presented to demonstrate the usefulness and benefits in Section 4. Concluding statements are also provided in Section 5.

2 LITERATURE REVIEW

As an initial phase in the review procedure, a summary of the existing literature concerning traffic accident prediction models with GLM and ANN is provided.

2.1 Generalized Linear Models

Abdel-Aty and Radwan [37] utilized Negative Binomial (NB) distribution to anticipate the number of accidents as a function of AADT, lane shoulder and median widths, section length, degree of horizontal curvature, and rural/urban status. Outcomes revealed that number of accidents increases with section length, degree of horizontal curvature and AADT. Hauer [38] established statistical road safety modeling by utilizing the NB distribution. The dependent parameter was the accident frequency per year, though the independent parameters were traffic flow and geometric characteristics. The findings indicated the fact that AADT and speed limit were important variables.

El-Basyouny and Sayed [39] investigated two forms of regression techniques including the Modified Negative Binomial (MNB) and Traditional Negative Binomial (TNB) in Canada. Although the MNB method presents itself to fit the data much better compared to the TNB method, there was a small distinction in the outcomes of

the recognition and rating of accident-prone areas. This is probably because of the nature of the application and the data set utilized.

APMs for evaluation of the tangents and curves for a four-lane median-divided highway in Italian were established by Caliendo et al. [40] utilizing the Poisson, NB regression and negative multinomial regression methods. Distinct forecasted models for overall accidents, injury and fatal accidents indicate that substantial parameters are length, curvature, and AADT, while for tangents they are length, AADT and the presence of a junction.

Lord and Mannering [24] offered a comprehensive review of the crucial factors related to the accident rate information as well as the weaknesses and strengths of the several regression analyses including the Poisson regression model, negative binomial (Poisson-gamma) regression model and generalized estimating equation model. This can assist researchers to address and consider these factors in their works. A research was conducted by Lu et al. [25] depending on the Logistic regression method in China. This study considered the relationships among the traffic accident and the vehicle type, road type, the weather, the driving characteristics, and the date. The outcomes present that the position of the vehicle in the highway, the visual condition, the road surface condition, the road safety grade, the driver state, and the vehicle condition are the most substantial variables that may cause a traffic crash.

Glavić et al. [26] considered the impact of road and traffic characteristics on the event of traffic crashes. The regression evaluation was utilized to analyze the effect of the factors, such as AADT, the radius of horizontal curvature, the percentage of commercial vehicles, the pavement width and the gradient. As a result, two models of the multivariate regression evaluation were established. One of them involved the radius of horizontal curvature and AADT, while the other involved pavement width and AADT. The first model displays that the rise of AADT and reduction of the radius, raise the overall range of traffic crashes. The second model displays that rises of AADT and reduction of pavement width raise the overall range of traffic crashes.

Gianfranco et al. [29] established a predictive model for urban highways that is capable of calculating the range of crashes for three circumstances in an urban highway system including a three- or four-way intersection, roundabout and a straight stretch of highway. The models created depend on Poisson and NB regression and can be easily used for crash prediction or recognition of the hotspot area. The results indicate that the majority of representative independent parameters was ADT, particularly for roundabouts and junctions. No other independent parameter was discovered to be specifically essential for both situations.

2.2 Artificial Neural Network

Conventional research on road accidents is conducted by Mussone et al. [41] to evaluate the impact of factors including vehicular characteristics, road geometry, vehicular flows, and the computation of the range of crashes using ANN in Italy. As a result, junction

complexity may be identified as a greater crash index based on the regulation of junctions. The greatest index for running over of pedestrian takes place at unsignalized junctions at night-time. Chang [42] utilized the ANN and NB regression method based on the two years of accident information pertaining to the National Freeway in Taiwan. Through evaluating the prediction functionality between the ANN model and NB regression, this research displays that ANN is definitely a reliable technique for examining freeway crash frequency.

Akgüngör and Doğan [43] suggest a Genetic Algorithm (GA) as well as an ANN approach to calculate the range of crashes, injuries, and fatalities in Ankara, Turkey. The evaluation of the model outcomes pointed out that the efficiency of the ANN approach was much better compared to the GA model. In addition, Cansız [44] established a nonlinear crash model with Smeed formula and ANN technique to determine the range of deaths in crashes.

Yu et al. [45] utilized two techniques, i.e. Support Vector Machine (SVM) and ANN to predict the traffic crash period. The results indicate that both ANN and SVM approaches had the capability to anticipate the traffic crash period within appropriate limits. Although the ANN approach gets a much better outcome for long period incident circumstances, the extensive functionality of the SVM approach is significantly better compared to the ANN approach for the traffic crash period prediction. Huang et al. [46] created an improved Radial Basis Function Neural Network (RBFNN) and NB approach to estimate the nonlinear associations between the number of accidents and the appropriate risk variables in Hong Kong. The outcomes reveal that the RBFNN has much better functionality compared to the NB approaches.

Alkheder et al. [47] utilized an ANN to estimate the damage intensity of traffic crashes in Abu Dhabi. WEKA (Waikato Environment for Knowledge Analysis) data mining software was utilized to develop the ANN classifier. The experimental outcomes pointed out that the created ANN classifiers can easily anticipate crash intensity with affordable precision.

Sameen and Pradhan [48] utilized a deep learning approach utilizing a Recurrent Neural Network (RNN) to anticipate the damage intensity of traffic crashes in Malaysia. In comparison with traditional ANN, the RNN approach is much more efficient for continuous data. The suggested RNN approach was compared with Bayesian Logistic Regression (BLR) and Multilayer Perceptron (MLP) techniques to comprehend its benefits and restrictions. The outcomes of the comparison studies revealed that the RNN method outperformed the BLR and MLP approaches for forecasting the damage intensity of traffic crashes.

Yuan et al. [16] tried to forecast real-time accident risk by taking into consideration time series dependency along with utilizing an LSTM-RNN model, conditional logistic model, and synthetic minority over-sampling technique (SMOTE). The comparability outcomes revealed that the LSTM-RNN with SMOTE outperformed the conditional logistic approach. The approaches and results of this research effort confirm the feasibility of real-time accident risk prediction by utilizing LSTM-RNN along with an over-sampled dataset (SMOTE).

Singh et al. [33] utilized deep neural networks (DNN) technique for the prediction of highway crashes. To examine the efficiency of the DNN technique, Random Effect Negative Binomial (RENB) and Gene Expression Programming (GEP) approaches were utilized. The results of this research showed that the DNN approach outperformed both RENB and GEP technique on almost all performance indicators. It can be utilized when the goal of modeling is to forecast crash frequency precisely to calculate associated costs. GEP approach also carried out much better compared to the RENB approach, and it also has the potential of determining essential risk factors and quantifying their impact on crashes.

3 MATERIAL AND METHODS

3.1 Case Study

In this research, the gathered data on the number of accidents include a period of 14 years from 2005 to 2019 and are related with the road network of the Province of Erzurum. It is situated in the eastern part of Anatolia, on the northern side of Palandöken Mountain, and in the south-eastern part of Erzurum plain. The city of Erzurum is the biggest city in the eastern part of Turkey, where dramatic changes in land use and land cover have occurred. The region, in general, has hilly topography with high mountains. The city has an altitude of 1757 meters with a humid continental climate.

In order to evaluate the traffic accidents on the highways, initially requires to choose highway(s) which have a broad range of traffic and geometric characteristics. The target of this data gathering is to separate these site studies into sections with homogenous attributes. After evaluating various highways around Erzurum, it was determined that D100-28, D950-03, D100-29, and D052-03, i.e. four-lane median-divided highways, were most suitable for this research.

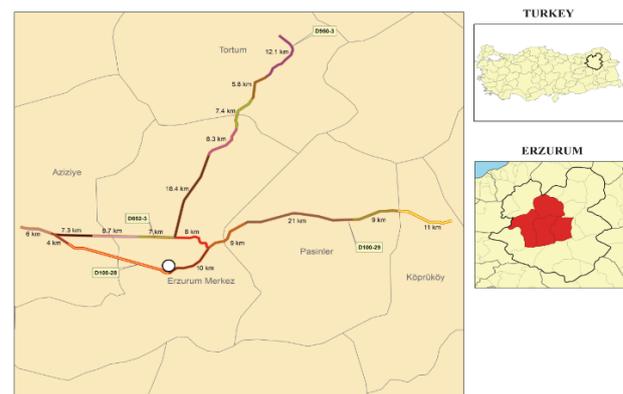


Figure 1 Erzurum's highway and sections

The case studies consist of 153 km arterial routes of the Erzurum region, as shown in Fig. 1. The arterial routes are long enough to produce an adequate number of segments to constitute the model. In the highway segments, a new section would be started if the median width changed from 2 to 4 m. Therefore, each case study is consistent with regard to all the traffic features and possible geometry. The information on highways includes geometric characteristics such as the gradient of the highways,

horizontal and vertical curves, shoulder widths, median widths, and traffic characteristics such as AADT.

Totally 21500 traffic accident reports were used in this research (a substantially larger data was used comparing the previous studies in the literature); all data were gathered from the 12th Highway Regional Directorate and Directory of City Traffic Region. Each crash record has several data including the accident location, date, vehicle type, pavement type, driver's age, driver's gender, the day and time, road surface condition, day or nighttime, weather condition, the number of injured persons, the number of deaths, the number of damaged vehicles, and the number of involved vehicles. After the preprocessing procedure and removing the erroneous and missing data, 18956 accident information was employed in this investigation.

In the case of input and output data, initially, there were 18 variables as an input such as years, highway sections, section length, the gender of the driver, day or nighttime, road surface condition, shoulder width, median width, highway gradients, the number of both horizontal and vertical curvature, AADT, accidents included with heavy and light vehicles, accidents occurring in winter, fall, summer, and spring. The number of traffic accidents was the output. After that, it has been agreed to establish the GLM and ANN models by utilizing eight input and one output variable.

Table 1 Properties of the highway sections

Name of Highway Section	Definition of the Highway Section	Number of Divided Highway Segments	Length of the Highway Section / Kilometer
Erzurum-Tortum	D-950-03	5	52
The South Ring Road	D-100-28	3	20
The North Ring Road	D-052-03	4	30
Erzurum-Köprüköy	D-100-29	4	50

3.2 Methods

There are two distinguished methods that belong to the most commonly appreciated ways of modeling the transportation data including statistics and Computational Intelligence (CI). The former one, statistics, is the mathematics of gathering, managing and understanding the numerical data, especially once these data revolve around the analysis of population characteristics through concluding from sampling [49]. Statistics have gained solid and commonly appreciated mathematical foundations and can enable understanding of the mechanisms that create the data. Nevertheless, they frequently collapse while handling complex and highly nonlinear data (curse of dimensionality).

The latter one, CI, combines elements of learning, adaptation, evolution and fuzzy logic to create models that are "intelligent" in that structure emerging from an unstructured beginning (the data) [50-52]. Nowadays, CI models, have been commonly applied to diverse transportation challenges, partially owing to the fact that they are very generic, exact and traditional mathematical models that can readily simulate numerical model components.

3.2.1 Generalized Linear Models

The term GLM's method identifies a greater class of models prominent by McCullagh and Nelder [53] and McCullagh [54]. GLM's method offers a common technique to a wide variety of response modeling problems. Poisson, Normal, and binomial responses are usually the most generally utilized; however other distributions can be utilized as well. Besides indicating the response, GLMs additionally require a link function to be set that enables more flexibility in the modeling [55]. There are varieties of probability distributions and link functions including Poisson distribution: log function, normal distribution: identity function and binomial distribution: logit function.

a) Poisson distribution: log function

Generally, there are 3 factors in GLM such as probability distribution, link function, and linear predictor. In the case of Poisson regression, it can easily be formulated like Eq. (1).

$$\ln y_i = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \rightarrow \lambda_i \sim \text{Poisson}(y_i) \quad (1)$$

Linear predictor is only a linear formula of the explanatory variable (x) and parameter (b). Link function actually "links" the parameter for probability distribution and the linear predictor. For Poisson regression, the common link function is actually the log link function. The last element is the probability distribution that creates the observed variable y_i . Once we utilize Poisson distribution below, the model is known as Poisson regression.

The prediction curve is exponential as the inverse of the log link function is an exponential function. Because of this fact, it is also obvious that the parameter for Poisson regression computed by the linear predictor assured to be positive. The final formula for GLM based on the Poisson distribution can be found in Eq. (2).

$$\ln y_i = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \rightarrow y_i = \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n) \quad (2)$$

where: y_i : dependent variable (output: accident numbers), x_i : independent variables (inputs), $b_0, b_1, b_2, \dots, b_n$: coefficients

b) Normal distribution: identity function

Linear regression is a case of GLM, where it just utilizes identity link function (the parameter for the probability distribution and the linear predictor are identical) and normal distribution as the probability distribution. Eq. (3) shows the GLM's formula based on the normal distribution.

$$y_i = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad (3)$$

where: y_i : dependent variable, x_i : independent variables, $b_0, b_1, b_2, \dots, b_n$: coefficients.

c) Binomial distribution: logit function

If we utilize logit function as the link function and Bernoulli distribution/binomial distribution as the probability distribution, the model is known as logistic regression, as shown in Eq. (4).

$$\log \frac{q_i}{1-q_i} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad (4)$$

If we symbolize the linear predictor with z_i , the previously mentioned formula is equal to Eq. (5).

$$z_i = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \rightarrow z_i = \log \frac{q_i}{1-q_i} \quad (5)$$

The final formula for GLM based on the binomial distribution can be found in Eq. (6).

$$q_i = \frac{1}{1 + \exp(-z_i)} \quad (6)$$

where: q_i : dependent variable, z_i : linear predictor.

The right side of Eq. (6) is known as a logistic function. Thus, this model is known as logistic regression. Logistic regression is actually utilized mainly for binary classification issues. Because of this fact and nonlinear relationships among accident variables, the Poisson distribution, in this research, is utilized to predict the number of accidents.

3.2.2 Artificial Neural Network

ANN possesses great characteristic features because of its versatility to any system, learning and generalization capability, modeling flexibility, operating with significant levels of multi-dimensional information, excellent prediction capability. Due to these characteristics, ANN has a great prospect to utilize in data analytic methods for transportation investigation [56].

A neural network is a computing model whose layered structure resembles the network structure of neurons in the brain, with layers of connected nodes. It consists of an input layer, one or more hidden layers, and an output layer. The layers are interconnected via nodes, or neurons, with each layer using the output of the previous layer as its input. A neural network can learn data so it can be trained to forecast future events.

The ability of the network to learn complex relationships between input and output patterns makes the ANN approach appealing, which would be difficult to model with conventional methods [57].

The framework of the ANN utilized for this research is a three-layer ANN, as demonstrated in Fig. 2. The fundamental components are generally neurons, and each one is interconnected with all the neurons in the following layer. The initial layer is the input layer, where the data are introduced to the ANN. The input parameters can certainly be either categorical or numerical values. The hidden layer is created in the intermediate layer. The purpose of this layer is to calculate the complex pattern associations. An individual hidden layer has been discovered to be acceptable for the majority of purposes, though the quantities of neurons were chosen by trial and error throughout running the code by software in various times to obtain the greatest performing network and smallest

acquired error. The final layer is the output layer, representing the network result to the related input.

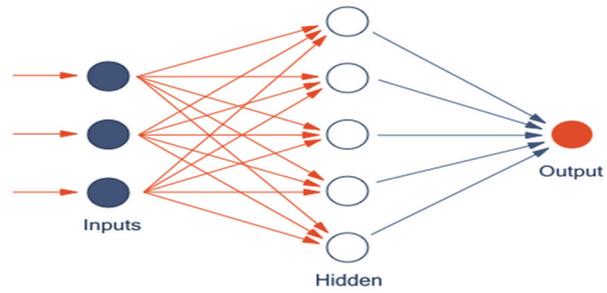


Figure 2 Artificial Neural Network Architecture

Then, the ANN can certainly be trained with a training algorithm where the Back-Propagation (BP) rule, which is actually one of the most generally utilized training algorithms, is used in the following research. The BP-ANN algorithm is a multi-layer feedforward network trained in accordance with the error backpropagation algorithm. BP network can be utilized to learn and store a great deal of mapping relations of the input-output model and it does not require to reveal beforehand the mathematical formula that explains these mapping relations. Its learning rule is to follow the steepest descent technique by which the backpropagation is utilized to control the threshold value and weight value of the network to obtain the lowest error sum of square [58]. The network input to a unit j is provided, as shown in Eq. (7).

$$X_j = \sum_i W_{ij} X_i + b_j \quad (7)$$

where: X_i - output from the previous layer, W_{ij} - weight of the connection between layer i and j , and b_j - bias

The concept of the training is to minimize the total output error explained in Eq. (8), where T_i is the target value, O_i is the model output value and MSE is the mean squares error.

$$MSE = \frac{1}{n} \sum_{i=1}^n (T_i - O_i)^2 \quad (8)$$

4 RESULTS AND DISCUSSION

4.1 Model Development and Results

In the GLM and ANN models, input variables are named as independent and output variables are called dependent. As described in the methodology, in the case of input and output data, initially there were 18 variables while it was agreed to establish the GLM and ANN models by utilizing eight input variables simply because parameters were found to be significant based on those criteria. In addition, one output variable was selected which was the number of traffic accidents. All variables are shown in detail in Tab. 2.

The significant parameters are years (i.e. 15 years from 2005 until 2019), highway sections (i.e. 4 segments), section lengths consist of 153 km arterial routes of Erzurum region which were determined as categorical values. In this regard, section lengths were included D 950-

03 (i.e. totally 52 km), D100-28 (i.e. totally 20 km), D052-03 (i.e. totally 31) and D100-29 (i.e. totally 50 km). Other parameters were specified as annual average daily traffic (AADT), the degree of horizontal curvature, the degree of vertical curvature, traffic accidents with heavy vehicles (percentage), and traffic accidents that occurred in summer (percentage). These input variables would represent the potential risk factors for accidents.

Table 2 Definition of the variables

Variables	Definition	Binary/ Numerical code	In/output layers
X_1	Years	Categorical value (2005-2006-2007-2008-2009-2010-2011-2012-2013-2014-2015-2016-2017-2018-2019)	15
X_2	Highway location	Categorical value (D 950-03, D100-28, D052-03, D100-29)	4
X_3	Section length	Categorical value (D 950-03(12.1-5.8-7.4-8.3-18.4), D100-28 (6-4-10), D052-03 (7.3-8.7-7-8), D100-29 (9-21-9-11))	16
X_4	Annual average daily traffic	Numerical Value	1
X_5	The number of horizontal curvature	Numerical Value	1
X_6	The number of vertical curvature	Numerical Value	1
X_7	Traffic accidents with heavy vehicles	Numerical Value	1
X_8	Traffic accidents occurred in summer	Numerical Value	1
Y	Number of accidents	Numerical Value	1

4.1.1 Generalized Linear Models

As described in methodology, logistic regression is actually utilized mainly for binary classification issues. In addition, because of nonlinear relationships among accident variables, the Poisson distribution, Eq. (2), is utilized to establish model accidents.

Totally 21500 traffic accident reports were used in this research; all data were gathered from the 12th Highway Regional Directorate and Directory of City Traffic Region. Each crash record has several data and after preprocessing procedure and removing the erroneous and missing data, 18956 accident information was employed in this investigation.

As described before, initially there were 18 variables as an input and the number of traffic accidents was the output. After that, it has been agreed to establish the GLM models by utilizing eight input and one output variables. The case studies consist of 153 km arterial routes of the Erzurum region. The arterial routes are long enough to produce an adequate number of segments to constitute the model.

Regression analysis involves numerical parameters. Therefore, if a set of data consists of categorical parameters in a regression model needed, supplementary actions are generally needed to create the outcomes interpretable. In

these procedures, the categorical parameters are usually recorded directly into a set of individual binary parameters. This recording is known as "dummy coding" and results in the generation of a table known as the contrast matrix. This is carried out automatically through statistical software. In this research, year, highway location and section length are categorical data which need to be additionally analysed before modeling.

In an effort to come up with the most appropriate model that relates accident frequencies with their explanatory variables, a number of regression models referring to the Poisson distribution were tested in Stata Software. The model formulation of GLM is given in Eq. (9):

$$E(AN) = \exp \left(\begin{matrix} 1.982235 + 0.0315427 \cdot SL + \\ + 0.0001042 \cdot AADT + 0.0647847 \cdot NVC - \\ - 0.0642354 \cdot NHC - 0.3475845 \cdot TAOS - \\ - 0.398745 \cdot TAHV \end{matrix} \right) \quad (9)$$

where: $E(AN)$ - Total Number of Accidents, SL - Section Length, $AADT$ - Annual Average Daily Traffic, NVC - the Number of Vertical Curve, NHC - the Number of Horizontal Curve, $TAOS$ - Traffic Accidents that Occurred in Summer, $TAHV$ - Traffic Accident Heavy Vehicle.

4.1.2 Artificial Neural Network

The topological structure of the ANN, as described in the methodology, involved one input layer, one output layer, and one hidden layer (Fig. 2). In this regard, data sets were separated into three sections: the training set, the test set, and the verification set. Training algorithms never utilize the test or verification sets to modify network weights. The verification set may optionally be utilized to track the network's error performance, to recognize the greatest network and to prevent training if over-learning happens. The test set is not utilized in training at all, and it is developed to provide a distinct evaluation of the network's performance when the whole network design process is accomplished. Based on the 18956 records of traffic accidents, 70% of the records were randomly divided into a training set with the size of 13269, a validation set (15%) with a size of 2843, and a test set (15%) with the size of 2843. In order to create the ANN model, the Alyuda Neuro Intelligence software package was preferred in this research. Alyuda is an ANN-based predicting software and offers an intuitive, wizard-like interface that rapidly and simply acts via the procedure of making an ANN for estimation, prediction or optimization.

In the case of the training procedure, the more complicated the structure, the more time it requires to train the model. However, more complication can generate much better generalization performance and much less training error, complex networks can readily over-fit the training data, generating inadequate generalization capacity and testing performance, which is called over-fitting issue [59]. One of the methods to prevent this issue is to reduce the complexness of the model. To reduce the complexness, users can easily eliminate layers or decrease the quantity of neurons to create the network smaller. In addition, early stopping is actually a kind of regularization

during training a model with an iterative approach, for instance, gradient descent. Since all the ANN learn specifically by utilizing gradient descent, early stopping is a method appropriate to all the issues. This technique updates the model so as to create it to much better fit the training data with each iteration.

Prior to providing training on the system, a minimax algorithm was utilized for normalizing both output and input parameters throughout the variety between 0 and 1. Normalization of the data is important in increasing learning speed and also reducing the trained network error [60]. In addition, Alyuda Neuro Intelligence program instantly encoded the categorical columns in the data preprocessing part. One-of-N method through Alyuda Neuro Intelligence system was carried out which indicates a column with N different categories (values) is encoded directly into a set of N numeric columns, with one column for each category. For instance, for the Capacity column with values "Low", "Medium" and "High", "Low" will be displayed as (1, 0, 0), Medium as (0, 1, 0), and High as (0, 0, 1). The minimum and maximum of the data set were discovered and scaling elements chosen so that these were mapped to the desired minimum and maximum values.

To achieve the most precise model structure, various numbers of nodes were analyzed, among which the greatest one was recognized. Outcomes for an individual hidden layer with numerous neurons (i.e. from 2 to 12) are compared during modeling, as shown in Tab. 3. The performance of ANN models is validated referring to the fitness criteria.

Among all tested networks, the 41-10-1 structure (40-input neurons, 10-neurons in the hidden layer and 1-output neurons) led to the highest fitness criteria between the targeted and predicted output for training, testing and validation of data sets. Accordingly, the output of the traffic accident model fits the observed data well.

Table 3 The fitness criteria of the best five networks

Network	Fitness Criteria
40-2-1	0.672
40-6-1	0.712
40-7-1	0.737
40-10-1	0.902
40-12-1	0.840

4.2 Model Comparison and Discussion

In order to recognize the most acceptable model for the current research, the evaluation associated with the correlation coefficient, *R*-Squares, Root Mean Square Error (*RMSE*) and Mean Square Error (*MSE*) between data was performed, in which low value of *MSE* and *RMER* and great value of correlation would be selected as the most suitable model. These four tests were chosen to recognize the accuracy due to the fact that they were applied as a criterion for such an assessment.

The *RMSE* and *MSE*, Eq. (8), is frequently used to calculate the distinctions between actual data set and the values estimated by predicted models. A correlation coefficient is actually a numerical calculator, which means a statistical relationship between two parameters. They all presume values in the range from -1 to +1, where ±1 signifies the strongest possible agreement and 0 the strongest possible disagreement. *R*-squared recognized as

the coefficient of determination, is the proportion of the variance in the dependent parameter that is predictable from the independent parameters. In this regard, it can be estimated as follows:

$$\text{Mean of the observed data: } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \tag{10}$$

$$\text{Total sum of squares: } TTS = \sum_i (y_i - \bar{y})^2 \tag{11}$$

$$\text{Explained Sum of Squares: } ESS = \sum_i (f_i - \bar{y})^2 \tag{12}$$

$$R^2 = \frac{ESS}{TSS} \tag{13}$$

where: \bar{y} - the mean of the observed data, y_i - observed data, n - the number of observed data, f_i - the predicted values associated with new models.

The conclusion summary of the ANN's and GLM's models is given in Tab. 4. The values of correlation for GLM's and ANN' models are relatively close to 1 which was estimated to be around 0.86 and 0.98, respectively. It clearly shows that ANN's model provides a better result than GLM's model.

As can be seen from Tab. 4, the *R*-squared values for the ANN's and GLM's models are calculated to be around 0.97 and 0.75, respectively, which clearly shows that ANN's model generated reasonably more reliable results than GLM's model. In addition, the values of *MSE* and *RMSE* values for ANN's model were computed to be around 3.89 and 1.97, while it was estimated that for GLM's model they were around 16.59 and 4.07. In this case, the values closer to zero are much better. In general, the conclusion summary of the ANN in Tab. 4 contains high correlation coefficient and *R*-squared values as well as low *MSE* and *RMSE* of the tested network and the superiority of the ANN's model is understood.

Table 4 Comparison of GLM model to ANN model

Model	Correlation	<i>R</i> -Squares	<i>MSE</i>	<i>RMSE</i>
GLM	0.8664	0.7506	16.59	4.07
ANN	0.9874	0.9749	3.89	1.97

For more comparison between GLM's and ANN's model, Fig. 3 demonstrates the number of observations corresponding to the number of traffic accidents during analysis. This graph displays a line graph of the actual data and output values related to the ANN's and GLM models. Accordingly, the ANN's model is intimately related to the real traffic accident data while the outputs of the GLM's model were predicted in different trends. The results expressed the high accuracy of the ANN outputs in predicting the number of traffic accidents in different sections of the highway.

Tab. 5 shows the importance percentages of traffic accident parameters based on the ANN's model. Outcomes exhibited that the number of horizontal curvature with 52 percentages is the most substantial factor which impacts the number of accidents that happened on the highways.

Table 5 Importance percentages of traffic accident parameters for ANN analysis

Input Column Name	Importance / %
The number of horizontal curvature	52
Annual average daily traffic	14
The number of vertical curvature	12
Section length	10
Highway location	6
Years	2.5
Traffic accidents occurred in summer	2.3
Traffic accidents with heavy vehicles	1.2

Additionally, the percentages of the AADT (14%), followed by the number of vertical curvature (12%) and section length (10%) have also impacted the traffic accidents remarkably. In this regard, heavy vehicles with

1.2% were determined as the lowest feature related to the traffic accident prediction on the highway.

In order to take into account the performance of the ANN's model for future prediction, ten years period from 2020 until 2030 was utilized. The outcome of the model demonstrates (Tab. 6) that, by 2030, the estimated number of accidents in Erzurum is predicted to be around 4591 which have annual growth rate of almost 11.22% from 2020. This is because the number of vehicles has been increasing; consequently, traffic accidents are expected to increase in the future. In addition, the values of growth rate are shown in Table 6 where the percentage for 2025, 2027, 2028, 2029 and 2030 was predicted to be much higher than for other years.

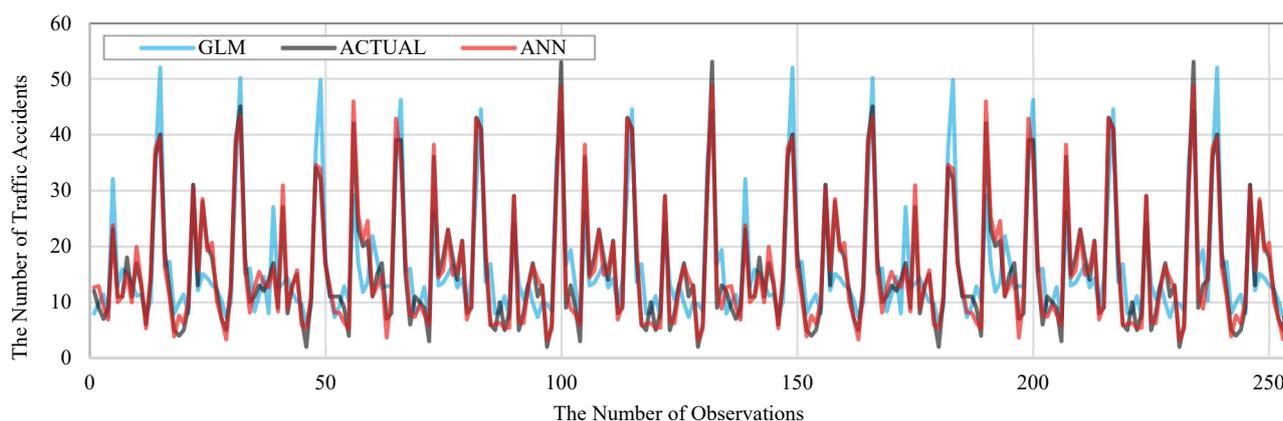

Figure 3 Relationship between target and output values of the models

Table 6 Number of accident prediction in Erzurum city from 2020 until 2030

Year	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Accident Prediction	1584	1687	1755	1854	1954	2345	2532	2853	3354	3824	4591
Growth rate / %		6.50	4.03	5.65	5.39	20.01	7.97	12.67	17.56	14.01	20.06

5 CONCLUSION

This research was carried out to create a robust model to predict the risk factors and accident frequency in Erzurum, Turkey using ANN and GLM methods. Data was gathered from the quantity of accidents included in a period of 14 years, from 2005 to 2019, from the 12th Highway Regional Directorate and Directory of the City Traffic Region. After the preprocessing procedure and removing the erroneous and missing data, information on 18956 accidents was employed in this research. In order to make the prediction, a variety of variables such as years, highway location, section length, AADT, the number of the horizontal curves, the number of the vertical curves, the number of traffic accidents caused by heavy vehicles, the number of traffic accidents in summer, were used. Compared to the ANN and GLM utilized in the current research, the suggested ANN model attained much better fitting and forecasting functionality compared to the GLM when modeling accident frequency was the greatest performing model with the high value of correlation and lowest *MSE* and *RMER*. To consider the functionality of the ANN model for future estimates, 10 years period from 2020 to 2030 was utilized. As the research outcomes signify, by 2030, the approximated quantity of crashes in Erzurum is predicted to be around 4591 with an average growth rate of almost 11.22% from 2020. The outcomes

demonstrated that the most important elements in forecasting the number of traffic crashes were the number of horizontal curvatures, followed by AADT, the number of vertical curvatures, and section length. Eventually, the results of this research are effective and strong modeling in road safety applications. In this regard, to overcome traffic crashes a much better-comprehended geometry adjustment associated with horizontal or vertical curves may be taken as an indicator of a significant safety issue in decreasing accidents.

6 REFERENCES

- [1] Çodur, M. Y. & Tortum, A. (2015). An artificial neural network model for highway accident prediction: A case study of Erzurum, Turkey. *PROMET - Traffic & Transportation*, 27(3), 217-225. <https://doi.org/10.7307/ptt.v27i3.1551>
- [2] Sahraei, M. A., Kuşkan, E., Çodur, M. Y., & Tortum, A. (2020). An Overview of Traffic Accident Prediction Models. *Nobel Bilimsel Eserler Sertifika*.
- [3] Shi, L., Huseynova, N., Yang, B., Li, C., & Gao, L. *A cask evaluation model to assess safety in Chinese rural roads. Sustainability*, 10(11), 3864. <https://doi.org/10.3390/su10113864>
- [4] Brătucu, G., Madar, A., Boşcor, D., Băltescu, C. A., & Neacşu, N. A. (2016). Road safety education in the context of the sustainable development of society: The Romanian case. *Sustainability*, 8(3), 278.

- <https://doi.org/10.3390/su8030278>
- [5] Chen, Z. & Qin, X. (2019). A novel method for imminent crash prediction and prevention. *Accident Analysis & Prevention*, 125, 320-329. <https://doi.org/10.1016/j.aap.2018.07.011>
- [6] Basso, F., Basso, L. J., & Pezoa, R. (2020). The importance of flow composition in real-time crash prediction. *Accident Analysis & Prevention*, 137, 105436. <https://doi.org/10.1016/j.aap.2020.105436>
- [7] Useche, S., Montoro, L., Alonso, F., & Oviedo-Trespalacios, O. (2018). Infrastructural and human factors affecting safety outcomes of cyclists. *Sustainability*, 10(2), 299. <https://doi.org/10.3390/su10020299>
- [8] Zakharov, D., Magaril, E., & Rada, E. C. (2018). Sustainability of the urban transport system under changes in weather and road conditions affecting vehicle operation. *Sustainability*, 10(6), 2052. <https://doi.org/10.3390/su10062052>
- [9] Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis & Prevention*, 33(6), 799-808. [https://doi.org/10.1016/S0001-4575\(00\)00094-4](https://doi.org/10.1016/S0001-4575(00)00094-4)
- [10] La Torre, F., Meocci, M., Domenichini, L., Branzi, V., & Paliotto, A. (2019). Development of an accident prediction model for Italian freeways. *Accident Analysis & Prevention*, 124, 1-11. <https://doi.org/10.1016/j.aap.2018.12.023>
- [11] Miaou, S.-P. & Song, J. J. (2005). Bayesian ranking of sites for engineering safety improvements: decision parameter, treatability concept, statistical criterion, and spatial dependence. *Accident Analysis & Prevention*, 37(4), 699-720. <https://doi.org/10.1016/j.aap.2005.03.012>
- [12] Ozbay, K. & Noyan, N. (2006). Estimation of incident clearance times using Bayesian Networks approach. *Accident Analysis & Prevention*, 38(3), 542-555. <https://doi.org/10.1016/j.aap.2005.11.012>
- [13] Adeli, H. & Karim, A. (2000). Fuzzy-wavelet RBFNN model for freeway incident detection. *Journal of transportation engineering*, 126(6), 464-471. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2000\)126:6\(464\)](https://doi.org/10.1061/(ASCE)0733-947X(2000)126:6(464))
- [14] Hsiao, C.-H., Lin, C.-T., & Cassidy, M. (1994). Application of fuzzy logic and neural networks to automatically detect freeway traffic incidents. *Journal of transportation engineering*, 120(5), 753-772. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1994\)120:5\(753\)](https://doi.org/10.1061/(ASCE)0733-947X(1994)120:5(753))
- [15] Sayed, T., Abdelwahab, W., & Navin, F. (1995). Identifying accident-prone locations using fuzzy pattern recognition. *Journal of transportation engineering*, 121(4), 352-358. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1995\)121:4\(352\)](https://doi.org/10.1061/(ASCE)0733-947X(1995)121:4(352))
- [16] Yuan, J., Abdel-Aty, M., Gong, Y., & Cai, Q. (2019). Real-time crash risk prediction using long short-term memory recurrent neural network. *Transportation Research Record*, 2673(4), 314-326. <https://doi.org/10.1177/0361198119840611>
- [17] Yu, L., Du, B., Hu, X., Sun, L., Han, L., & Lv, W. (2021). Deep spatio-temporal graph convolutional network for traffic accident prediction. *Neurocomputing*, 423, 135-147. <https://doi.org/10.1016/j.neucom.2020.09.043>
- [18] Pu, Z., Liu, C., Shi, X., Cui, Z., & Wang, Y. (2020). Road surface friction prediction using long short-term memory neural network based on historical data. *Journal of Intelligent Transportation Systems*, 1-12. <https://doi.org/10.1080/15472450.2020.1780922>
- [19] Roque, C. & Cardoso, J. L. (2014). Investigating the relationship between run-off-the-road crash frequency and traffic flow through different functional forms. *Accident Analysis & Prevention*, 63, 121-132. <https://doi.org/10.1016/j.aap.2013.10.034>
- [20] Wedderburn, R. W. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss Newton method. *Biometrika*, 61(3), 439-447. <https://doi.org/10.1093/biomet/61.3.439>
- [21] Liang, K.-Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13-22. <https://doi.org/10.1093/biomet/73.1.13>
- [22] Lindsey, A. R. (1997). Wavelet packet modulation for orthogonally multiplexed communication. *IEEE Transactions on signal processing*, 45(5), 1336-1339.
- [23] Fahrmeir, L. & Tutz, G. (2001). *Models for multivariate responses: Multivariate extensions of generalized linear models*. Springer, 69-137.
- [24] Lord, D. & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291-305. <https://doi.org/10.1016/j.tra.2010.02.001>
- [25] Lu, T., Dunyao, Z., Lixin, Y., & Pan, Z. (2015). The traffic accident hotspot prediction: Based on the logistic regression method. *International Conference on Transportation Information and Safety (ICTIS)*, IEEE, 107-110.
- [26] Glavić, D., Mladenović, M., Stevanovic, A., Tubić, V., Milenković, M., & Vidas, M. (2016). Contribution to accident prediction models development for rural two-lane roads in Serbia. *PROMET - Traffic & Transportation*, 28(4), 415-424.
- [27] Mohammadi, M., Shafabakhsh, G., & Naderan, A. (2018). Effects of modal shares on crash frequencies at aggregate level. *Accident Analysis & Prevention*, 120, 295-303. <https://doi.org/10.1016/j.aap.2018.08.019>
- [28] Ozfirat, M., Ozfirat, P., Aksoy, C., Pamukcu, C., & Tatar, C. (2005). Modelling the risk factors of occupational accidents in an underground coal mine in Turkey using regression analysis. *CIM Bulletin*, 98.
- [29] Gianfranco, F., Soddu, S., & Fadda, P. (2018). An accident prediction model for urban road networks. *Journal of Transportation Safety & Security*, 10(4), 387-405. <https://doi.org/10.1080/19439962.2016.1268659>
- [30] Çodur, M. Y. & Tortum, A. (2009). Modelling car ownership in Turkey using neural networks. *Proceedings of the Institution of Civil Engineers-Transport*, 97-106.
- [31] Sahraei, M. A. & Puan, O. (2018). Traffic Delay Estimation Using Artificial Neural Network (ANN) at Unsignalized Intersections. *Proceedings of the 3rd International Conference on Civil, Structural and Transportation Engineering (ICCSTE'18)*, 10-12.06.
- [32] Wang, J., Kong, Y., & Fu, T. (2019). Expressway crash risk prediction using back propagation neural network: A brief investigation on safety resilience. *Accident Analysis & Prevention*, 124, 180-192. <https://doi.org/10.1016/j.aap.2019.01.007>
- [33] Singh, G., Pal, M., Yadav, Y., & Singla, T. (2020). Deep neural network-based predictive modeling of road accidents. *Neural Computing and Applications*, 1-10. <https://doi.org/10.1007/s00521-019-04695-8>
- [34] Li, P., Abdel-Aty, M., & Yuan, J. (2020). Real-time crash risk prediction on arterials based on LSTM-CNN. *Accident Analysis & Prevention*, 135, 105371. <https://doi.org/10.1016/j.aap.2019.105371>
- [35] Abdelwahab, H. T. & Abdel-Aty, M. A. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record*, 1746(1), 6-13. <https://doi.org/10.3141/1746-02>
- [36] Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, 38(3), 434-444. <https://doi.org/10.1016/j.aap.2005.06.024>
- [37] Abdel-Aty, M. A. & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5), 633-642. [https://doi.org/10.1016/S0001-4575\(99\)00094-9](https://doi.org/10.1016/S0001-4575(99)00094-9)

- [38] Hauer, E. (2004). Statistical road safety modeling. *Transportation Research Record*, 1897(1), 81-87. <https://doi.org/10.3141/1897-11>
- [39] El-Basyouny, K. & Sayed, T. (2006). Comparison of two negative binomial regression techniques in developing accident prediction models. *Transportation Research Record*, 1950(1), 9-16. <https://doi.org/10.1177/0361198106195000102>
- [40] Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. *Accident Analysis & Prevention*, 39(4), 657-670. <https://doi.org/10.1016/j.aap.2006.10.012>
- [41] Mussone, L., Ferrari, A., & Oneta, M. (1999). An analysis of urban collisions using an artificial intelligence model. *Accident Analysis & Prevention*, 31(6), 705-718. [https://doi.org/10.1016/S0001-4575\(99\)00031-7](https://doi.org/10.1016/S0001-4575(99)00031-7)
- [42] Chang, L.-Y. (2005). Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. *Safety science*, 43(8), 541-557. <https://doi.org/10.1016/j.ssci.2005.04.004>
- [43] Akgüngör, A. P. & Doğan, E. (2009). An artificial intelligent approach to traffic accident estimation: Model development and application. *Transport*, 24(2), 135-142. <https://doi.org/10.3846/1648-4142.2009.24.135-142>
- [44] Cansiz, O. F. (2011). Improvements in estimating a fatal accidents model formed by an artificial neural network. *Simulation*, 87(6), 512-522. <https://doi.org/10.1177/0037549710370842>
- [45] Yu, B., Wang, Y., Yao, J., & Wang, J. (2016). A comparison of the performance of ANN and SVM for the prediction of traffic accident duration. *Neural Network World*, 26(3), 271.
- [46] Huang, H., Zeng, Q., Pei, X., Wong, S., & Xu, P. (2016). Predicting crash frequency using an optimised radial basis function neural network model. *Transportmetrica A: transport science*, 12(4), 330-345. <https://doi.org/10.1080/23249935.2015.1136008>
- [47] Alkheder, S., Taamneh, M., & Taamneh, S. (2017). Severity prediction of traffic accident using an artificial neural network. *Journal of Forecasting*, 36(1), 100-108. <https://doi.org/10.1002/for.2425>
- [48] Sameen, M. I. & Pradhan, B. (2017). Severity prediction of traffic accidents with recurrent neural networks. *Applied Sciences*, 7(6), 476. <https://doi.org/10.3390/app7060476>
- [49] Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1997). Statistical themes and lessons for data mining. *Data mining and knowledge discovery*, 1(1), 11-28. <https://doi.org/10.1023/A:1009773905005>
- [50] Sadek, A. W., Spring, G., & Smith, B. L. (2003). Toward more effective transportation applications of computational intelligence paradigms. *Transportation Research Record*, 1836(1), 57-63. <https://doi.org/10.3141/1836-08>
- [51] Engelbrecht, A. P. (2007). *Computational intelligence: an introduction*. John Wiley & Sons.
- [52] Vlahogianni, E. I. (2015). Computational intelligence and optimization for transportation big data: challenges and opportunities. *Engineering and Applied Sciences Optimization*, 107-128.
- [53] McCullagh, P. & Nelder, J. (1989). *Generalized Linear Models* Chapman and Hall. London. <https://doi.org/10.2307/2344614>
- [54] McCullagh, P. (2019). *Generalized linear models*. Routledge.
- [55] Faraway, J. J. (2010). *Generalized Linear Models. International Encyclopedia of Education (Third Edition)*. Elsevier: Oxford, 178-183.
- [56] Karlaftis, M. G. & Vlahogianni, E. I. (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, 19(3), 387-399. <https://doi.org/10.1016/j.trc.2010.10.004>
- [57] Murat, Y. S. & Ceylan, H. (2006). Use of artificial neural networks for transport energy demand modeling. *Energy policy*, 34(17), 3165-3172. <https://doi.org/10.1016/j.enpol.2005.02.010>
- [58] Li, J., Cheng, J.-H., Shi, J.-Y., & Huang, F. (2012). Brief introduction of back propagation (BP) neural network algorithm and its improvement. *Advances in computer science and information engineering*, 553-558.
- [59] Amiri, A. M., Sadri, A., Nadimi, N., & Shams, M. (2020). 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*, 138, 105468. <https://doi.org/10.1016/j.aap.2020.105468>
- [60] Alyuda neuro Intelligence manual (2004). Retrieved from <http://www.alyuda.com>

Contact information:

Mohammad Ali SAHRAEI, Dr

(Corresponding author)

1) Erzurum Technical University, Faculty of Engineering and Architecture, Civil Engineering Department, 25050, Erzurum, Turkey
E-mail: ali.sahraei@erzurum.edu.tr

2) Civil Engineering Department, Faculty of Engineering, Girne American University, Girne, N. Cyprus Via Mersin 10, Turkey
E-mail: mohammadalisahraei@gau.edu.tr

Merve Kayaci ÇODUR, Dr

Erzurum Technical University

Faculty of Engineering and Architecture, Industrial Engineering Department, 25050, Erzurum, Turkey
E-mail: merve.codur@erzurum.edu.tr

Muhammed Yasin ÇODUR, Associate Professor

1) Erzurum Technical University

Faculty of Engineering and Architecture, Civil Engineering Department, 25050, Erzurum, Turkey

2) American University of Middle East,

College of Engineering and Technology, Civil Engineering Department, Kuwait

E-mail: mycodur@erzurum.edu.tr

Ahmet TORTUM, Professor Dr.

Atatürk University

Engineering Faculty, Civil Engineering Department,

25240, Erzurum, Turkey

E-mail: atortum@atauni.edu.tr