Contents lists available at http://qu.edu.iq

Al-Qadisiyah Journal for Engineering Sciences

Journal homepage: http://qu.edu.iq/journaleng/index.php/JQES

Analysis of rural road traffic crashes in Al-Diwaniyah province using Artificial Neural Network

Taghreed S. Khuzan[a] and Mohammed A. Al-Jumaili[b]

[a]*Roads and Transport Department, College of Engineering, University of Al-Qadisiyah, Iraq*
[b]*Civel Department, College of Engineering, University of Kufa, Iraq*

ARTICLE INFO

Article history:
Received 04 August 2022
Received in revised form 2 October 2022
Accepted 05 November 2022

Keywords:
Keywords:
Traffic crash
ANN
Rural roads
CSO

ABSTRACT

Traffic crash commonly referred to as a car crashes with another vehicle or other objects, such as pedestrians, animals, road barriers, or other immovable structures like a utility pole or wall. This can result in injuries, fatalities, vehicle and road damage. In latest years, the number of traffic crashes has increased in Iraq, particularly on rural roads. Because of the severity of the rural road crash problem in Iraq, it has become critical to investigate and understand the causes of rural road traffic crashes. Rural road traffic crashes in Al-Diwaniyah Province haven't been highlighted since 2003. According to the Planning Ministry's yearly statistical report, Iraq (2020) the number of crashes recorded in Iraq is 8186, 587 of them are in Al-Diwaniyah Governorate. The percentage of traffic crashes in 2020 inside the city was 47.01% and outside it was 52.99%. The goal of this research is to analyze traffic crash data to identify the crucial factors that impact the occurrence of vehicle crashes in Al-Diwaniyah Province.

1. Introduction

Due to the serious economic and social effects that traffic crashes have, traffic safety has recently attracted a lot of attention on a global scale. In addition, traffic crashes cause a large percentage of disability and human losses for young groups, and thus reduce the proportion of labor force. Adding to the monetary and emotional costs associated with traffic crashes, it is a fact that many people are killed or spend long periods of time in hospitals as a result of traffic crashes every year. After 2003, the amount of vehicles in Iraqi cities increased, as did the number of traffic crashes. There was a sharp increase in the total number of automobiles in Iraq after 2003, which has been linked to the economy's recovery after the economic blockade was lifted. Traffic crashes in Iraq are very dangerous, comparable to terrorist operations, and pose a great threat to the lives of the Iraqi people. It has been one of the issues that depletes material resources, causes social problems, and results in the loss of human resources, which are considered to be the most important component of society. Also, the individuals involved, their families, and their relatives all suffer negative psychological effects as a result of traffic crashes. Those who have been directly involved in traffic crashes face serious health consequences. Crashes in traffic systems cause long-term or permanent harm. Changes in an individual's quality of life, as well as changes in social, family, and professional lives following a traffic crash, including changes in views toward life, are all social consequences of crashes [1].

* *Corresponding author.*
E-mail address: taghreedkhuzan@gmail.com(Taghreed Shakir Khuzan)

https://doi.org/10.30772/qjes.v15i4.856

Road geometry (such as infrastructure quality and curvature), traffic conditions (such as speed and the traffic flow), and lighting conditions (such as night and day driving) can all affect traffic safety, as well as driver behavior [2]–[6].
The human errors is the main cause of traffic crashes in the city of Baghdad, according to the study of Al-Tamimi [7]. By investigating roadway elements, particularly roadway geometric design, human aspects can be controlled and predicted indirectly [8]. Also mechanical breakdown of a vehicle, such as tires, brakes and steering failure, lead to a slight percentage of crashes [9]. The condition and performance of the roadway, including the road surface, shoulders, crossing points and traffic control system, can influence a collision [10]. According to Garber having more lanes, heavier traffic volume, exceeding the posted speed limit, having a low lane width, and having a low shoulder width all increase the probability of a crash [11]. The installation of median barriers in the roads has a number of drawbacks, including a high cost and an increase in the number of traffic crashes reported as a result of a lack of recovery area. More collisions raise maintenance expenses and raise the possibility that maintenance staff will be exposed to traffic. Additionally constrained is the ability of maintenance and emergency vehicles to cross the median [12].
It is widely expected that deterioration of the road surface will lead to higher in crash rates. Because the state of the pavement is one of the road's environmental elements that influence both human and vehicle factors at the same time, it is regarded as a critical factor in highway safety. As a result, one of the main goals of pavement management systems is to improve highway safety by ensuring that pavements are well maintained. According to the findings of Lee et al., a deteriorating pavement reduces the severity of one-vehicle crashes on low-speed roadways but rises the severity of single-vehicle collisions on high-speed roadways in poor pavement condition. Otherwise, it worsens the severity of multiple-vehicle collisions on all roadways [13].
Al-Obaedi conducted a study to evaluation traffic crash rates in Al-Diwaniyah Governorate. According to the study findings, traffic crashes cause 33 deaths per 100,000 people in the city. And this is a significantly higher than the average global death rate, which is 18 deaths for every 100,000 people. Additionally, research has shown that 50% of traffic collisions are caused by driver errors, while vehicle malfunctions account for about 24 percent. In addition, the findings revealed that approximately 21% of crashes are not reported to the local traffic [1].

2. Study area

Al-Diwaniyah province is one of the provinces of the Middle Euphrates, It is about 180 km from Baghdad. Through which one of the branches of the Euphrates River passes, known as Shatt al-Diwaniyah. The latitude of Diwaniyah City is 31o57'50.91" N and the longitude 44o 54'23.94" E. Elevation of Diwaniyah City is 50 meters above sea level. As for the city's administrative boundaries, it is bordered to the north by the provinces of Babil and Wasit, to the east by the provinces of Dhi Qar and Wasit, to the south by the Muthanna province, and to the west by the Najaf province.
According to the report of the Diwaniyah Investment Commission, the area of the governorate is about 8153 km2, and thus it constitutes about (1.9%) of the total area of the country, and about (8.1%) of the total area of the provinces of the Middle Euphrates region. Out of the total population of Iraq, which exceeds 41 million for the year 2021, according to the estimates of Ministry of Planning - CSO reports, the population of Diwaniyah province is estimated at 1,325,031 distributed over 57% live in urban and 43% in rural area. Figure 1 shows the administrative boundaries of the city.

Figure1. Diwaniyah province boundaries

3. Time limits

The study's time frame was determined by the number of road traffic crashes recorded between 2016 and 2021 by the Diwaniyah Traffic Directorate.

4. Data collection

Data on traffic crashes is collected by responsible agencies and authorities in most developed countries according to well-defined standards and procedures. In these countries crash data comes from a variety of sources including insurance, police reports and hospital records. In Iraq, there are only two sources of traffic crash data. First, crash report forms, which are filled out by the traffic police officer. Second, by preparing a questionnaire and distributing it to a certain number of people who are often participants in a specific traffic crash, asking them a series of questions, then collecting and analyzing the data.
Questionnaires are usually prepared by researchers, and the questions of the questionnaire depend on the data that the researcher needs in his study.
Government sources, are more comprehensive. Also, the process of collecting traffic crash data is a difficult process in Iraq, because of the routine procedures in the country, as well as because of the Ministry of Interior's reservation on this data, as it is data that contains some drivers' details and is used for the purposes of legal cases.

4.1. Collecting crash reports from Diwaniyah traffic directorate

It was difficult for this study to obtain data from the traffic crash records from the Diwaniyah Traffic Directorate because of the security and routine procedures in Iraq, as obtaining the data required approval from the Ministry of Interior in Baghdad. The fact that this data, according to their claim, contains personal information of the participants in the crash and is only released for the purposes of legal cases. After submitting official requests to the General Traffic Directorate in the Ministry of Interior in Baghdad, and after a period of more than three months and many reviews to the Ministry of Interior in Baghdad, approval was obtained to provide this study with the necessary data. Crash reports are the data obtained from the Traffic Directorate. These reports contain the details of the crash, which include the type of crash, the type of road, the cause of the crash, the number of vehicles involved in the crash, the number of deaths and injuries, and some other details. This traffic data is more detailed for crashes than the data of the CSO.

5. Artificial neural networks (ANN)

An artificial neural network is a model of information processing that mimics how biological nerve systems, like the brain, process information. The impressive quality of artificial neural network models is that they can predict and display desired results even with small amounts of data [15]. A.O. Kurban looked into artificial neural networks, which have non-linear mapping algorithms and a structure that is based in part on concepts found in biological nervous systems.

Typical real neurons have branching dendritic trees that gather signals from numerous other neurons in a small space; and a lengthy branching axon that contacts the dendritic trees from several additional neurons to distribute the response. The response of each neuron is fundamentally a non - linear function of the its input and is strongly impacted by the connections among those inputs. Figure 2 shows the multilayer feedforward network.

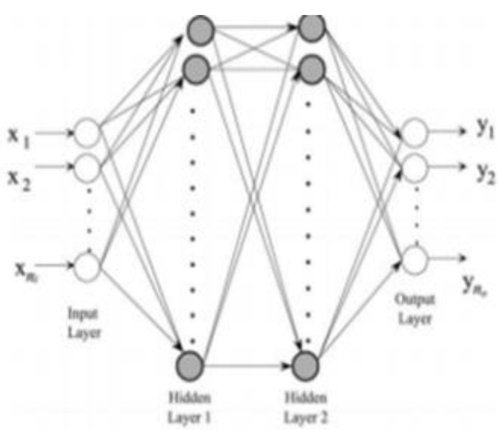

Figure2. Multilayer Feedforward Network [16]

Typically, an ANN is made up of a lot of simple processors linked by connection weights. The data that is available locally at each node, whether it is stored internally or comes in through weighted connections, is the only factor affecting each node's output. Each nodes send its output to a number of other nodes while also receiving input from a large number of other nodes. By changing the connection topology and connecting weight values, a network is specialized to perform various functions. Connecting units with the right weights allows for the implementation of complex functions [16].

6. Statistical analysis using ANN

After the process of filtering the data obtained from the Diwaniyah Traffic Directorate and after excluding the unclear and incomplete copies, the number of copies available was 200 copies of the crash report. SPSS Version 28 and a multilayer feedforward network with one hidden layer was used to unveiled some important factors that might contribute to a more severe crash. As an introduction to dealing with neural networks, the data must be filtered and not be excessive in including a large number of independent variables because this negatively affects the neural network [17].

6.1. ANN inputs and outputs

Selecting the input variables that have the biggest effects on model performance is a crucial stage in creating ANN models. A good subset of input variables can greatly enhance model performance [18]. The following kinds of data changes have been made: removal of invariant columns, removal of columns with a lot of variety, removal of redundant columns, elimination of useless data, and classification of some columns. After preprocessing the traffic crashes data, the 16 different attributes were reduced to 8 attributes.

The inputs were season, day of week, number of lanes, direction, crash type, number of vehicles involved, vehicle type, collision with, contributing factor, causes, road condition, road type, weather and vehicle manufacture. While the output was crash severity.

Table 1. Class labels selected 8 attributes with their data type

NO.	Attribute name	Data type
1	Season	Nominal
2	Day of week	Nominal
3	Time	Nominal
4		
5	Crash type	Nominal
6	Collision with	Nominal
7	Contributing factor	Nominal
	Causes	
8	Vehicle	Nominal
9	Manufacture	Nominal

Eight nodes are chosen in this study as the number of nodes in each layer, which is equal to the number of the input layer's nodes. Available data is divided into two sets: training and testing. The training set is used to modify the neural network's connection weights, and the testing set is used to evaluate the network's performance. In the current study, the optimal division was chosen using a trial-and-error procedure. The software's default and non-default parameters were used to create a number of networks with various divisions, and the results showed that 66.5% of the data were used for training and 33.5% for testing. Figure 3 illustrated the distribution of the data.

Classification

		Predicted			
Sample	Observed	.00	1.00	2.00	Percent Correct
Training	.00	9	1	0	90.0%
	1.00	0	86	1	98.9%
	2.00	0	0	36	100.0%
	Overall Percent	6.8%	65.4%	27.8%	98.5%
Testing	.00	6	1	0	85.7%
	1.00	0	32	0	100.0%
	2.00	0	0	28	100.0%
	Overall Percent	9.0%	49.3%	41.8%	98.5%

Dependent Variable:severity

Figure 3. Classification of data

6.2. Structure and operation of ANN

A typical ANN structure comprises of a few artificial neurons, or nodes, and the links that connect them. There is a weight parameter associated to each bond. Every processing element in a particular layer is connected to many other processing units via weighted connections, either entirely or in part. The scalar weights control how strongly linked neurons are connected to each other. When a weight is zero, there is no connection between any two neurons, and when a weight is negative, there is a prohibitive association[19]. One layer of input neurons, one or maybe more layers of hidden neurons, and one layer of output neurons typically make up ANN neural networks; however, the subsequent layers are completely connected, as illustrated in Figure 4.

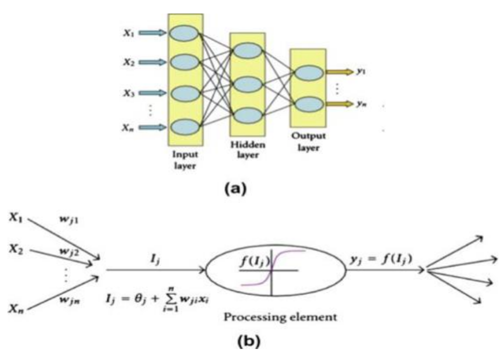

Figure 4. Typical structure and operation of ANN [19]

6.3. ANN prediction accuracy

Many trainings were performed on the artificial neural network to achieve the highest accuracy and lowest error rate. And that is by controlling the percentages of the training and testing samples, the change of weights, the hidden layers and the number of nodes of the hidden layers.
The total sample size of 200 was used to make the artificial neural network. The sample was divided into a training sample and a testing sample. 66.5% was used as a training sample and 33.5% as a test sample, which gave the highest accuracy rate. Figure 1 shows the details of the artificial neural network used. The correctness rate of the model was 98.5% for the training sample and 98.5% for the testing sample as well. This means that the model was able to correctly classify most of the data, except for 1.5%.

Figure 3 shows the process of classifying data from SPSS. The model classified crashes with PDO with an accuracy of 90%, injury crashes with 98.9%, and fatal crashes with 100%. This is for the training sample. These findings show that the prediction accuracy for fatal crashes is highest, followed by injury crashes and PDO crashes, in that order.
The original percentages of crashes were as follows: crashes of property damage were 8.5%, crashes of injury 59.6%, and crashes of death 32%. As for the testing sample, it was classified by 85.7%, 100%, and 100%, respectively, with regard to crashes of property damage, injury, and fatality. The prediction efficiency, however, cannot be fully described by accuracy alone, hence additional methods of assessing the predictive models are required. The total ability of a test to discriminate is measured by the area under the curve (AUC). An ideal test has an AUC of 1.00, while a completely random test has an AUC of 0.5.

For the current study, the created models were further assessed using AUC curves. The percentage of the area under the curve for each category of the dependent variable was as follows, as shown in figure 5.

Area Under the Curve

		Area
severity	.00	.997
	1.00	.999
	2.00	1.000

Figure 5. Area Under the Curve

The AUCs were all noticeably higher than 0.5 in each example. These findings show that all models performed well in predicting new cases.

7. Results

The importance analysis of the model is illustrated in figure 6 and indicates the cause of the crash has the greatest impact on the severity of the crash. Followed by the season in which the crash occurred, and then the time of the crash. Where the importance of these variables was 0.313, 0.181 and 0.158, respectively.

Independent Variable Importance

	Importance	Normalized Importance
season	.181	57.9%
Day	.104	33.3%
Time	.158	50.5%
Crash type	.052	16.7%
Collision with	.083	26.5%
Contributing factors	.050	16.0%
Causes	.313	100.0%
Veh. manufacture	.059	18.7%

Figure 6. independent variable importance analysis

The importance analysis results also indicate that the day of the crash, the object to which the vehicle hits, the vehicle brand, the type of crash, and the factors contributing to the crash have less influence on the expected result, having values for relative importance of 0.104, 0.083, 0.059, 0.052 and 0.05 respectively.

8. Conclusions

1. Artificial neural networks can be used to analyze traffic crash severity . It can analyze traffic crashes with high accuracy and does not require a very large sample size.
2. The accuracy rate of the model was 98.5% for the training sample and 98.5% for the testing sample.
3. The model classified crashes with PDO with an accuracy of 90%, injury crashes with 98.9%, and fatal crashes with 100%. This is

for the training sample. As for the testing sample, it was classified by 85.7%, 100%, and 100%, respectively.

4. The importance analysis of the model indicates the cause of the crash has the greatest impact on the severity of the crash. Followed by the season in which the crash occurred, and then the time of the crash.
5. The importance analysis results also indicate that the day of the crash, the object to which the vehicle hits, the vehicle brand, the type of crash, and the factors contributing to the crash have less influence.
6. There is a difficulty in obtaining traffic crash data in Iraq because of the routine procedures in the country, and even when obtained, these data are for legal purposes only and do not contain many details necessary for scientific research purposes. Therefore, it is necessary to provide forms of traffic crash reports that take into account scientific and engineering purposes.

Authors' contribution

All authors contributed equally to the preparation of this article.

Declaration of competing interest

The authors declare no conflicts of interest.

Funding source

This study didn't receive any specific funds.

REFERENCES

[1] M. Másilková, "ScienceDirect Review article Health and social consequences of road traffic accidents," *KONTAKT*, pp. 1–5, 2017, doi: 10.1016/j.kontakt.2017.01.007.

[2] V. Shankar, F. Mannering, and W. Barfield, "Effect of roadway geometrics and environmental factors on rural freeway accident frequencies," *Accid. Anal. Prev.*, vol. 27, no. 3, pp. 371–389, Jun. 1995, doi: 10.1016/0001-4575(94)00078-Z.

[3] F. Sagberg, "Road accidents caused by drivers falling asleep," *Accid. Anal. Prev.*, vol. 31, no. 6, pp. 639–649, Nov. 1999, doi: 10.1016/S0001-4575(99)00023-8.

[4] T. F. Golob, "Structural equation modeling for travel behavior research," *Transp. Res. Part B Methodol.*, vol. 37, no. 1, pp. 1–25, Jan. 2003, doi: 10.1016/S0191-2615(01)00046-7.

[5] L. Aarts and I. van Schagen, "Driving speed and the risk of road crashes: A review," *Accid. Anal. Prev.*, vol. 38, no. 2, pp. 215–224, Mar. 2006, doi: 10.1016/j.aap.2005.07.004.

[6] A. P. Jones, R. Haynes, I. M. Harvey, and T. Jewell, "Road traffic crashes and the protective effect of road curvature over small areas," *Health Place*, vol. 18, no. 2, pp. 315–320, Mar. 2012, doi: 10.1016/j.healthplace.2011.10.008.

[7] H. H. Joni, A. M. Al-Dahawi, and O. J. Al-Tamimi, "Analysis of Road Accidents in Baghdad City," *Assoc. Arab Univ. J. Eng. Sci.*, vol. 26, no. 1, pp. 136–144, 2019, doi: 10.33261/jaaru.2019.26.1.018.

[8] M. H. Islam, L. Teik Hua, H. Hamid, and A. Azarkerdar, "Relationship of Accident Rates and Road Geometric Design," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 357, no. 1, 2019, doi: 10.1088/1755-1315/357/1/012040.

[9] J. R. Aworemi, I. A. Abdul-Azeez, and S. O. Olabode, "Analytical study of the causal factors of road traffic crashes in southwestern Nigeria," *Educ. Res.*, vol. 1, no. 4, pp. 118–124, 2010.

[10] N. J. Garber and L. a Hoel, *Traffic and Highway*. 2009.

[11] "Tracking laser Doppler measurement for velocity of moving target," *Proc. - 2012 Int. Conf. Comput. Sci. Inf. Process. CSIP 2012*, vol. 32, pp. 426–431, 2012, doi: 10.1109/CSIP.2012.6308884.

[12] K. R. A. and J. G. Pigman, "Kentucky Transportation Center - Evaluation of Median Barrier Safety Issues," 2008.

[13] J. Lee, M. Asce, B. Nam, A. M. Asce, M. Abdel-aty, and M. Asce, "Effects of Pavement Surface Conditions on Traffic Crash Severity," pp. 1–11, 1996, doi: 10.1061/(ASCE)TE.1943-5436.0000785.

[14] J. Shaker and Al-Obaedi, "Evaluation of Traffic Accidents Rates in Al-Diwaniya City," *J. Babylon Univ. Sci.*, no. 3, p. 2015, 2015.

[15] F. R. Moghaddam, S. Afandizadeh, and M. Ziyadi, "Prediction of accident severity using artificial neural networks," *Int. J. Civ. Eng.*, vol. 9, no. 1, pp. 41–49, 2011.

[16] A. O. Kurban, "Analysis of shafts surface pressures using neural networks," *Ind. Lubr. Tribol.*, vol. 56, no. 4, pp. 217–225, 2004, doi: 10.1108/00368790410541561.

[17] S. Alkheder, M. Taamneh, and S. Taamneh, "Severity Prediction of Traffic Accident Using an Artificial Neural Network," *J. Forecast.*, vol. 36, no. 1, pp. 100–108, 2017, doi: 10.1002/for.2425.

[18] H. Al-zubaidi, A. Hadi, and A. Alhaddad, "Trip generation modeling for selected zone in Diwania city Republic of Iraq TRIP GENERATION MODELING FOR SELECTED ZONES IN DYWANIA CITY A Thesis Submitted to the Highway and Transportation Engineering Department College of Engineering / Al-Mustansiriyah ," no. July, 2020.

[19] M. A. Shahin, M. B. Jaksa, and H. R. Maier, "State of the art of artificial neural networks in xengineering," *Electron. J. Geotech. Eng.*, vol. 13, no. I, pp. 1–25, 2008.