Impact investigation of road characteristics on the accident rate based on the neural network modelling

E V Pechatnova¹, V N Kuznetsov²

¹Altai State University, 61, Lenina ave., Barnaul, 656049, Russia
²Altai State Agricultural University, 98, Krasnoarmeysky ave., Barnaul, 656049, Russia

E-mail: Phukcia@yandex.ru

Abstract. The use of neural networks promotes flexible non-linear modelling. Also, it is the perspective direction. The article investigates the influence of the main road characteristics on the accident rate. Twelve variables are selected as the main road characteristics. The result of modelling is the two-layer feed-forward neural network with hidden sigmoid nodes and linear output nodes. The operability of the model is estimated by calculating the index of regression by training, validation and testing. The results are acceptable for the intended purpose. The resulting model can be used for assessing and forecasting the safety rate on the two-lane roads with the intensity of 120 to 340 cars/hour. These include a significant part of the streets of CIS countries and other states. Besides, the neural network model can become the basis for the development of software products for assessing the accident rate based on road characteristics or can be implemented in existing software.

1. Introduction
Currently, increasing the number of vehicles all over the world leads to various negative consequences, one of which is the high traffic accident rate [1]. The road traffic safety evaluation is necessary for the development and practice of effective measures to reduce the accident rate.

At present, the road traffic safety evaluation on the road section in Russia is determined by its classification (or non-classification) to the place of accident concentration. The road section of motorways relates to the place of accident concentration if its length does not exceed 1000 m and there are three or more accidents of the same type occurred here or there are five or more accidents regardless of their type, as the result of which people have died or have been injured, occurred during the reporting year. This condition is the main for the destination and realization measures in the traffic police and road services.

Thus, the traditional approach to assessing road traffic risk based on crash history is realized. However, the [2] was emphasized that this approach had exhausted itself: lots of accidents occurred in places, where there were not accidents previously, or there were few of them. States with high road safety rate use the risk assessment based on the analysis of the potential risk of road elements. The [3] notes that local assessment of road traffic risk is required for the effective prevention of road accidents. Also, such assessment can be based on road elements hazard modelling. Besides, the assessing of the safety rate based on road characteristics is necessary for the design, planning and reconstruction of highways.
One of the mathematical modelling methods, which is widely used today, is the use of Artificial Neural Networks (ANNs) [4]. This method allows automatically detecting and modelling complex non-linear relationships between input and output parameters [5]. The results of practice the ANNs are used in adaptive management, forecasting, creating expert systems. Some scientists have used neural networks to solve various issues in the field of road traffic safety research.

The research of [6] presents the results of modelling the road traffic accidents rate using neural networks with backpropagation of error (BP neural network). The research was conducted through particular regions of China. The indicators of the system included road safety culture, administrative system, resource allocation and others. The accuracy of the modelling is not given. In the research work [3] BP neural network was used in the modelling; the authors used the “Matlab” tooling. Particular road data were used as the input variables. Frequency and severity of the accident were used as the output variables. However, activation function in neural networks, regression values in samples and usage restrictions are not indicated in work.

V.A. Korchagin and other Russian researchers used ANNs to analyze road traffic accidents. [7]. The model scheme was presented in the work; the activation functions were reflected. However, the primary purpose of the work was the classification of crash depending on the factors that accompany them. The research work [8] is devoted to the development of crash risk modelling based on fuzzy neural networks. The structure of the model is presented in the form of a four-layer neural network. Training is carried out based on the developed application program written by the authors in C ++ programming platform. The input parameters are the operational condition of roads, the motorization rate of the region and population density. The approximation error is 13%.

To summarize, neural networks are the widely spread tool for modelling the modern road traffic safety system. However, there are only a few researches, which are devoted to the assessment of the impact of road characteristics on the accidents risk. The sphere of use the neural network is not indicated. The basis of our study is the assumption that the influence of road elements on the safety rate differs in depend on the type of road. This work is aimed at development the neural network model of the influence the main road characteristics on the accident rate. The results of the work can be applied to the significant part of the highways of Russia and the CIS countries.

2. Theoretical basis of research
Artificial neural networks (ANNs) are used in two ways: supervised learning (it requires an output vector of target values) and unsupervised learning (the training set consists of only input values). The last option is used for clustering tasks.

The basis of the ANNs is the formal neuron, which consists of 3 elements: “synapses”, the adder and converter. Input signals \(x_1,...,x_n\) are converted with the use of “synapses” and using the weight coefficient \(w_1,...,w_n\), which reflects the size of influence the input value on the state of the formal neuron.

The adder realizes the summation of the weight signals. The converter is characterized by the activation function. The output signal of the neuron is calculated using the converter. The most common activation functions are binary step, identity, sigmoid, hyperbolic tangent. The mathematical model of the neuron can be expressed in such a way:

\[
y = f(\sum_{i=1}^{n} (w_i \cdot x_i + b))
\]

where \(b\) is the neuron bias.

The multi-layer neural network, which has excellent computing power, is most often used for the problems of mathematical modelling of various processes. The defining characteristics of the multi-layer feed-forward neural network are: neurons are arranged in layers, the output of one layer is the input for the next layer.

3 types of neurons are distinguished in the multi-layer neural network:
- input nodes, to which input values are supplied (it is not always considered as the separate layer);
- hidden nodes, which are intermediaries between input and output nodes (they do not have feed-forward with input signals and do not give out output ones);
- output nodes (output of such nodes is the result for the whole network).

The neural network, which is used for solving the issue, is considered as the multidimensional non-linear system. The purpose of this system is searching the optimal solution to some function that quantitatively determines the solution quality. The algorithms for tuning the set of weight coefficients are formed for neural networks as multidimensional non-linear control objects.

The choice of the neural network structure, the number of layers and neurons depends on the specificity of the tasks.

The training is carried out after determination of the neural network structure. Each input signal undergoes processing, and the output signal is calculated during the supervised learning. The output signal is compared with the corresponding value of the targeting vector. The error is calculated. The weight coefficients are changed based on the error. The most common method is the error backpropagation algorithm.

3. Materials and methods

The training is carried out based on the motorway section A–322 Barnaul – Rubtsovsk – the state border with the Republic of Kazakhstan: from 27 to 299 km. It was analyzed 273 kilometres outside large settlements. The road is federal and connects the regions of Western Siberia with the Republic of Kazakhstan, Kyrgyzstan, Uzbekistan. It is also the part of the Asian route AN-64. The investigated road section is two-lane, non-high-speed, the average traffic volumes is anywhere from 120 to 340 cars/hour. These indexes are typical for Russian roads.

Input data are the road characteristics of the kilometre of the road. The division of the road into kilometre sections is explained by the specificity of the accidents statistics in Russia. 12 main road characteristics were selected among all the possible: the sum of the central nodes, the generalized characteristic of the gradient and slope length, its maximum value, the number of lanes, the possibility of the animals' appearance, the length of the bridge (m), the length of the fence (m), the length of the zone of insufficient visibility (m), the characteristics of the intersection (points), the characteristics of the crosswalks and stops of public transport (points), the total power of the service facilities, the average annual traffic volume. These indicators are selected based on analysis of the researches of the leading international scientists and preliminary correlation analysis.

The information about the evaluation index of the road can be obtained from the road design data and the traffic flow data and accident data.

The output data is the average annual accident rate per kilometre of the road (accident/year). Information about fatal-and-injury crashes and property-damage-only accidents was used - analysis time period: 2014–2018 years.

The construction of the ANNs was carried out with the use of the Matlab Neural Network Toolbox. The automatic network search was selected for the selection of the optimal model parameters. Data were scaled.

Feed-forward ANNs were used in work: the signal is transmitted from input nodes to output nodes (no cycles or loops).

The whole sample is divided into three groups for modelling: training – 75%, validation – 15% and testing–15%.

Levenberg-Marquardt backpropagation algorithm method was chosen for training the neural network. This algorithm typically requires more memory, but less time. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.

4. Results and discussion

As a result of iterative enumeration of the model parameters, the network with the best performance was selected. The activation functions in the layers of the neural network are different.
As a result of modelling, two-layer feed-forward network with hidden sigmoid neurons and linear output neurons was received (without input layer).

The model has 7 neurons at the hidden layer and 1 neuron at the output layer. The structure of the Neural Network is presented in Figure 1.

![Figure 1. The structure of the Neural Network model](image)

The model allows determining the potential accident rate on the kilometre section of road based on 12 basic road characteristics. The value is obtained at the output of the neural network. The influence of road elements justifies this value. The critical feature of the model is the road traffic safety evaluation based on the combination of road characteristics. These can both compensate for the negative influence of each other and strengthen it. The matrix of the weights coefficients of the neural network hidden layer is presented in Table 1.

| Input variable | Neuron 1 | Neuron 2 | Neuron 3 | Neuron 4 | Neuron 5 | Neuron 6 | Neuron 7 |
|----------------|----------|----------|----------|----------|----------|----------|----------|
| 1              | -1.79    | -0.19    | -0.27    | -1.36    | -0.56    | 0.48     | 0.22     |
| 2              | 0.23     | -0.42    | -1.83    | -0.66    | 0.39     | 2.88     | 0.75     |
| 3              | 0.98     | 0.27     | -1.87    | 1.10     | 0.93     | 0.41     | 0.52     |
| 4              | -1.30    | -1.85    | 0.61     | -1.03    | -0.95    | 1.27     | 0.56     |
| 5              | 0.00     | 0.97     | -0.42    | 0.37     | -0.14    | -0.71    | 0.13     |
| 6              | 1.49     | -0.34    | 1.17     | -1.85    | 0.74     | 1.30     | 0.91     |
| 7              | -1.03    | 0.94     | 2.02     | -0.30    | 0.16     | 0.72     | -3.10    |
| 8              | -0.17    | 0.65     | -1.44    | -1.95    | 1.90     | -1.65    | 0.79     |
| 9              | 0.94     | -0.08    | 1.62     | -1.19    | 0.69     | 2.19     | -0.59    |
| 10             | -0.46    | 1.77     | -0.53    | 1.65     | 0.56     | -0.20    | -2.00    |
| 11             | 1.05     | 1.22     | -0.40    | -1.46    | -1.63    | -1.51    | -0.08    |
| 12             | -0.79    | 0.21     | 1.39     | -1.44    | 0.99     | -0.01    | 0.59     |

One of the ways for performance evaluation of the neural network is the determination of the regression values according to the samples. The resulting regression values according to the samples are presented in Figure 2.

The obtained regression values are acceptable for the intended purpose. The road characteristics explain the objective accident rate. The influence of subjective factors (vehicle malfunction, driving while intoxicated, and others) was not studied in this research.

The neural network model can be used for forecasting potential accident rates during the design, construction, reconstruction and roads repair, as well as for determining the effectiveness of measures to improve road traffic safety.
5. Conclusion
The mathematical model is obtained based on the use of neural networks. This model promotes the road traffic safety evaluation on the road section based on on-road characteristics. The model has acceptable accuracy. The results of the investigation can become the basis for the development of new software products for the evaluation crash risk or be implemented in existing application programs. The primary users of such programs are road management, road construction and design organizations, as well as government services responsible for road traffic safety. Besides, the neural network model can become an element of the intellectual transport system, which is implemented in Russia and other states.

The use of the model is recommended for evaluation of potential road traffic, safety hazard and its forecasting on the sections of two-lane roads outside settlements with the average traffic volume anywhere from 120 to 340 cars/hour. Such roads include most of the roads in Russia and the CIS countries. The authors recommend conducting additional research for the sections of roads with different characteristics.

References
[1] Pechatnoven EV and Kuznetsov VN 2019 J. Phys.: Conf. Ser.1333 032063
[2] Verma B and Stockwell D 2011 Neural Information Processing. ICONIP 2011. Lecture Notes in Computer Science7064pp 530–537
[3] Zhang J, Zhang W, Liu H and Zhao W 2010Road Safety Evaluation Based on BP Neural Network 2010 International Conference on Logistics Engineering and Intelligent Transportation Systems (LEITS2010 – Proceedings) ed Wenbing Huand Muchin V E (Wuhan: IEEE) pp 1–4
[4] Shen Y, Li T, Hermans E, Ruan D, Wets G, Vanhoof K and Brijs T 2009Soft Computing14(12) pp 1255–63
[5] Tangri N, Ansell D and Naimark D 2008 Nephrology Dialysis Transplantation 23(9)pp 2972–81
[6] Gong Li 2015 International Conference on Information Sciences, Machinery, Materials and Energy (ICISMME 2015) ed Xuebin Ching and Veclav Dvorik (Chongqing: Atlantis Press) pp 680–683

[7] Korchagin V A, Klyavin V E, Suvorov V A and Zelencov M V 2009 Mir transportaiteknologicheskii mashin4 (27) pp 42–46

[8] Skorobogatchenko D A and Erohin A V 2014 Bulletin of Volgograd State University of Architecture and Civil Engineering. Series: Construction and Architecture36 (55) pp 174–181