 https://doi.org/10.3311/PPci.41167

Periodica Polytechnica Civil Engineering, 69(3), pp. 1034–1045, 2025

Forecasting the Number of Road Accidents in Poland and Hungary Using Neural Networks

Viktoria Ötvös[1,2*], Piotr Gorzelańczyk[3]

[1] Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary
[2] Directorate of Strategic, Research-development and Innovation, Institute for Transport Sciences (KTI), Than Karoly u. 3-5., H-1119 Budapest, Hungary
[3] Department of Transport, Stanislaw Staszic State University of Applied Sciences in Pila, St. Podchorazych 10, 64-920 Pila, Poland
* Corresponding author, e-mail: otvos.viktoria@edu.bme.hu

Received: 31 May 2025, Accepted: 29 May 2025, Published online: 12 August 2025

Abstract
The incidence of road accidents in Poland and Hungary has been on a downward trend annually, a pattern that is also evident globally. Although the recent pandemic has impacted these statistics, the total remains alarmingly high. Consequently, it is crucial to implement all possible strategies to reduce this figure further. This article aims to project future road accidents in Poland and Hungary. To achieve this, we analysed yearly statistics regarding road accidents in the two nations. Utilising data from the Polish Police and the Hungarian Central Statistical Office, we made forecasts from 2024 to 2030. Several selected neural network models were employed for this predictive analysis. The findings indicate that we can anticipate a continued stabilisation in the number of road accidents. Several factors, including the rising number of vehicles on the roads and the development of new highways, influence this trend. Additionally, the chosen sizes for the sample sets (learning, testing, and validation) play a significant role in the results obtained.
Keywords
road accident, pandemic, forecasting, neural networks, Poland, Hungary

1 Introduction

Road accidents are events that result in damage to property, as well as injury or loss of life to other motorists. According to the World Health Organisation (WHO), approximately 1.3 million individuals lose their lives in traffic accidents each year. On average, road incidents lead to a 3% loss in GDP for countries worldwide. These accidents are recognised as the primary cause of death among children and young adults aged 5 to 29 [1]. The United Nations General Assembly aims to halve fatalities and injuries from traffic collisions by 2030.

The extent of a traffic collision plays a crucial role in assessing its severity. Accurately estimating the severity of accidents is vital for authorities to develop effective traffic safety regulations to prevent incidents and mitigate injuries, fatalities, and property damage [2, 3]. Identifying the critical factors influencing accident severity is essential before implementing measures to reduce and prevent it [4]. Singh et al. [5] present a multi-node Deep Neural Network (DNN) framework designed to predict various levels of injury, death, and property damage, allowing for a thorough and precise evaluation of traffic accident severity.

Accident statistics can be sourced from several channels, typically compiled and analysed by government officials through relevant agencies. Data is collected from various outlets, including hospital records, insurance company databases, and police reports. The transportation sector then analyses this data on a broader scale [6].

Intelligent transportation systems are currently the most significant sources of information for analysing and forecasting traffic incidents. This data can be processed through vehicle GPS devices [7]. Additionally, roadside vehicle detection systems continuously monitor and record information about moving vehicles, including their speed, traffic volume, and vehicle type [8]. Another method to gather extensive traffic data over a specified period is through license plate recognition systems [9].

Cite this article as: Ötvös, V., Gorzelańczyk, P. "Forecasting the Number of Road Accidents in Poland and Hungary Using Neural Networks", Periodica Polytechnica Civil Engineering, 69(3), pp. 1034–1045, 2025. https://doi.org/10.3311/PPci.41167

Social media also serves as a potential source of information regarding traffic and accidents; however, the reliability of such data may be compromised due to the inexperience of the individuals reporting the information [10].

To ensure that accident data is useful, working with various data sources and validating this information appropriately is essential. The analytical outcomes can be significantly more precise by integrating multiple data sources and consolidating diverse traffic accident statistics [11].

Vilaca et al. [12] conducted a statistical analysis to evaluate the severity and establish a connection between traffic participants and accidents. The study's findings suggest enhancing the standards of traffic safety regulations and implementing additional traffic safety measures.

In another study, Bąk et al. [4] performed a statistical examination of traffic safety in a specific region of Poland, focusing on the number of traffic accidents as a key indicator for researching their causes. This study utilised multivariate statistical analysis to investigate the safety factors associated with individuals responsible for accidents.

The choice of accident data sources for analysis depends on the specific traffic issues being investigated. By integrating statistical models with supplementary data from real-world driving scenarios or information gathered from intelligent traffic systems, the effectiveness of accident prediction and prevention can be significantly improved [13].

Numerous methods for predicting the frequency of traffic accidents can be found in the literature. Time series approaches [14, 15] are among the most commonly utilised techniques for estimating accident frequency. However, they have limitations, including the inability to evaluate the accuracy of forecasts based on previous predictions and the frequent presence of autocorrelation in residuals [16]. For instance, Procházka et al. [17] employed a multi-seasonality model for their forecasting, while Sunny et al. [18] utilised the Holt-Winters exponential smoothing method. A notable drawback of these methods is that they do not allow exogenous variables in the model [19].

Another critical metric is the number of accidents per 10,000 inhabitants (NRA). In 2023, Poland recorded 20,936 road accidents with a population of 37.6 million. By applying the relevant data to the established equation, we find that there were 5.57 road accidents per 10,000 inhabitants in Poland that year. Conversely, Hungary experienced 14,452 road accidents with a population of 9.6 million, resulting in 15.05 road accidents per 10,000 inhabitants—a rate that is three times higher than that of Poland.

$$NRA = \frac{NR}{NI} \cdot 10000 \quad (1)$$

where:

- NR - number of road accidents
- NI - number of inhabitants.

Utilising the data mentioned above, the authors formulated predictions regarding the incidence of road accidents in both Poland and Hungary. To achieve this, neural networks were employed as a forecasting tool to estimate the number of accidents occurring.

2 Materials and methods

Each year, a significant number of road accidents take place on highways and streets. In recent years, the pandemic has led to a decrease in the overall number of road accidents, which influences the accuracy of the forecasted values. However, despite this temporary reduction, road incidents remain alarmingly high. Given this situation, it is crucial to implement all necessary measures to decrease these figures and identify the types of roads where the highest frequency of accidents is likely to occur (see Figs. 1 and 2).

Selected neural network models have been utilised to predict the number of road accidents in Poland [20, 21] and Hungary [22]. Here are some key points about this method [23].

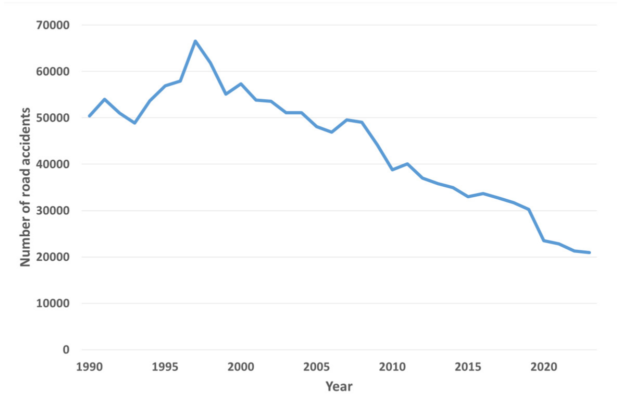

Fig. 1 Number of road accidents in Poland between 1990 and 2023 [20, 21]

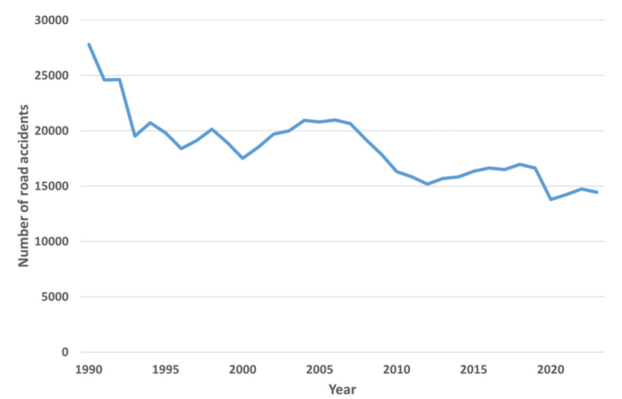

Fig. 2 Number of road accidents in Hungary between 1990 and 2023 [21]

- Imitation of Human Brain Behaviour: Neural networks are designed to mimic the way the human brain processes information, which can enhance their predictive capabilities.
- Structure of Neural Networks:
    - It comprises nodes that include inputs, weights, variances, and outputs.
    - Typically organised in multiple layers, the first layer contains data from various sources like images, numbers, text, or audio.
- Training Process:
    - It involves processing thousands of inputs to help the network learn and draw conclusions.
    - The optimal weights for the network are determined using software like Statistica, which influences the accuracy of predictions.
- Artificial Neurons:
    - The basic units of neural networks function similarly to biological neurons.
    - They have multiple inputs and produce a single output, similar to the operation of biological neurons' operation.
- Role in Artificial Intelligence:
    - Neural networks are fundamental to the advancement of AI, focusing on creating models that can exhibit intelligent behaviour, such as generalising knowledge.

This approach to accident prediction leverages the strengths of neural networks to analyse complex data and improve forecasting accuracy.

There are several uses for neural networks. This includes, for instance, systems that let users watch web series on demand and use past-watching data to identify a user's most comparable movie interests. Neural networks may also assist the Google Translate platform's text translation feature or show items customised for each bidder in online auctions. Forecasting, including the frequency of traffic accidents, is another use for neural networks [24–27].

A neural network model predicts the number of traffic accidents in the counties under study. The benefit of this method is that it simulates how the human brain works. A neural network comprises nodes with inputs, weights, variations, and outputs.

The Statistica software and its artificial neural network modules adjusted the ideal weights throughout the testing process. A multilayer perceptron (MLP) neural network including layers of hidden neurons was employed for prediction. In the examined situations, the middle layer's number of neurons varied from two to eight neutrons, whereas the output layer's single neutron represented the time series output values of the number of traffic incidents. The model and its parameters determine the predictive outcomes of the approaches discussed. The predictive quality measure was computed using the following prediction errors derived from Eqs (2)–(7).

- ME – Mean error

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Y_p\right) \tag{2}$$

- MAE – Mean absolute error

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - Y_p\right| \tag{3}$$

- MPE – Mean percentage error

$$\mathrm{MPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i - Y_p}{Y_i} \tag{4}$$

- MAPE - Mean absolute percentage error

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_i - Y_p\right|}{Y_i} \tag{5}$$

- SSE – Error Sum of Squares

$$\mathrm{SSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Y_p\right)^2} \tag{6}$$

- M^2 - Theila measure

$$\mathrm{M}^2 = \frac{\sum_{i=1}^{N}\left(Y_i - Y_p\right)^2}{\sum_{i=1}^{N}Y_i^2} \tag{7}$$

where:

- n – length of forecast horizon,
- Y – the observed value of road accidents,
- Y_p – projected value of road accidents.

Neural network models with the lowest mean percentage error and mean absolute percentage error were employed to forecast the frequency of traffic accidents in the future.

3 Results

To forecast the annual number of road accidents in Poland, data from the Polish Police from 1990 to 2023 were used [20, 21], while for Hungary, they were taken from the Hungarian Central Statistical Office [22]. In both cases, the research was carried out in Statistica software, assuming two random sample sizes:

1. teaching 70%, testing 15% and validation 15%.
2. teaching 80%, testing 10% and validation 10%, with the following number of learning networks: 20, 40, 60, 80, 100, and 200, for which the error value was minimal (Table 1–4).

Table 1 Summary of neural network learning for the case of random sample size, teaching 70%, testing 15% and validation 15% for Poland

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
20	MLP 1-7-1	0.97	0.98	0.99	BFGS 12	SOS	Tanh	Logistic	932.03	2374.97	2.85%	6.69%	2747.76	4.56E-03
20	MLP 1-8-1	0.97	0.98	0.99	BFGS 4	SOS	Linear	Logistic	664.14	2092.80	3.19%	6.46%	2667.94	4.30E-03
20	MLP 1-2-1	0.97	0.98	0.99	BFGS 4	SOS	Exponential	Exponential	1168.74	2119.69	3.23%	5.69%	2657.54	4.26E-03
20	MLP 1-4-1	0.96	0.97	0.99	BFGS 4	SOS	Exponential	Linear	1815.06	2891.74	2.84%	7.52%	3551.21	7.61E-03
20	MLP 1-3-1	0.97	0.97	0.99	BFGS 42	SOS	Tanh	Logistic	1113.77	2387.70	3.25%	6.64%	2802.32	4.74E-03
40	MLP 1-7-1	0.97	0.98	0.99	BFGS 8	SOS	Exponential	Exponential	837.43	2001.24	2.77%	5.68%	2484.15	3.72E-03
40	MLP 1-3-1	0.97	0.98	0.99	BFGS 8	SOS	Exponential	Exponential	909.41	2052.58	2.80%	5.71%	2526.15	3.85E-03
40	MLP 1-2-1	0.97	0.96	0.99	BFGS 7	SOS	Logistic	Logistic	1110.51	2279.94	3.35%	6.40%	2764.05	4.61E-03
40	MLP 1-8-1	0.97	0.96	0.99	BFGS 5	SOS	Logistic	Exponential	1483.15	2363.70	4.57%	6.80%	2927.92	5.17E-03
40	MLP 1-3-1	0.96	0.95	0.99	BFGS 5	SOS	Logistic	Exponential	1035.60	2590.23	3.60%	7.73%	3048.65	5.61E-03
60	MLP 1-4-1	0.97	0.97	0.99	BFGS 7	SOS	Tanh	Logistic	1031.48	2377.94	3.10%	6.68%	2772.29	4.64E-03
60	MLP 1-7-1	0.97	0.97	0.99	BFGS 5	SOS	Tanh	Logistic	777.04	2415.30	2.17%	6.63%	2763.52	4.61E-03
60	MLP 1-2-1	0.97	0.97	0.99	BFGS 9	SOS	Exponential	Logistic	1109.93	2233.82	3.17%	6.12%	2715.75	4.45E-03
60	MLP 1-5-1	0.97	0.97	0.99	BFGS 12	SOS	Tanh	Exponential	1090.27	2283.83	3.28%	6.40%	2721.35	4.47E-03
60	MLP 1-6-1	0.97	0.98	0.99	BFGS 5	SOS	Exponential	Exponential	1071.39	1987.71	3.83%	5.94%	2590.07	4.05E-03
80	MLP 1-5-1	0.97	0.97	0.99	BFGS 7	SOS	Tanh	Logistic	1040.10	2404.10	3.15%	6.78%	2796.23	4.72E-03
80	MLP 1-8-1	0.97	0.98	0.99	BFGS 14	SOS	Exponential	Logistic	1023.90	2217.06	3.03%	6.15%	2661.45	4.27E-03
80	MLP 1-3-1	0.97	0.98	0.99	BFGS 7	SOS	Exponential	Logistic	801.65	2239.47	2.48%	6.27%	2623.54	4.15E-03
80	MLP 1-7-1	0.97	0.96	0.99	BFGS 7	SOS	Logistic	Logistic	978.70	2426.95	2.68%	6.63%	2851.96	4.91E-03
80	MLP 1-7-1	0.97	0.98	0.99	BFGS 13	SOS	Exponential	Logistic	873.08	2237.49	2.54%	6.16%	2638.93	4.20E-03
100	MLP 1-8-1	0.97	0.98	0.99	BFGS 18	SOS	Exponential	Logistic	1021.62	2260.14	2.99%	6.25%	2688.97	4.36E-03
100	MLP 1-5-1	0.97	0.97	0.99	BFGS 6	SOS	Logistic	Logistic	1108.50	2402.26	3.24%	6.69%	2819.42	4.80E-03

Table 1 Summary of neural network learning for the case of random sample size, teaching 70%, testing 15% and validation 15% for Poland (continued)

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
100	MLP 1-4-1	0.97	0.98	0.99	BFGS 11	SOS	Logistic	Exponential	909.58	2320.20	2.93%	6.65%	2707.86	4.43E-03
100	MLP 1-2-1	0.96	0.95	0.99	BFGS 7	SOS	Tanh	Logistic	1114.15	2426.31	4.15%	7.43%	3005.79	5.45E-03
100	MLP 1-2-1	0.97	0.97	0.99	BFGS 8	SOS	Tanh	Logistic	894.03	2347.67	2.83%	6.67%	2720.34	4.47E-03
200	MLP 1-6-1	0.96	0.96	0.99	BFGS 8	SOS	Tanh	Logistic	644.88	2480.32	2.22%	7.15%	2814.19	4.78E-03
200	MLP 1-6-1	0.97	0.96	0.99	BFGS 7	SOS	Tanh	Logistic	770.95	2330.09	2.51%	6.64%	2702.56	4.41E-03
200	MLP 1-3-1	0.97	0.97	0.99	BFGS 10	SOS	Logistic	Logistic	970.77	2347.01	2.97%	6.61%	2750.08	4.56E-03
200	MLP 1-2-1	0.95	0.92	0.99	BFGS 7	SOS	Logistic	Exponential	319.55	2657.15	0.32%	7.61%	3035.78	5.56E-03
200	MLP 1-4-1	0.97	0.97	0.99	BFGS 6	SOS	Tanh	Logistic	1200.19	2356.37	3.63%	6.66%	2816.40	4.79E-03
								Minimal	319.55	1987.71	0.32%	5.68%	2484.15	3.72E-03

Table 2 Summary of neural network learning for the random sample size, teaching 80%, testing 10%, and validation 10% for Poland

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
20	MLP 1-5-1	0.96	0.99	1.00	BFGS 8	SOS	Logistic	Linear	422.40	1830.32	0.90%	5.12%	2362.11	3.37E-03
20	MLP 1-5-1	0.96	0.99	1.00	BFGS 5	SOS	Linear	Tanh	420.05	2152.64	0.39%	6.51%	2773.07	4.64E-03
20	MLP 1-3-1	0.96	0.99	1.00	BFGS 63	SOS	Tanh	Logistic	702.37	1986.10	2.31%	5.57%	2455.03	3.64E-03
20	MLP 1-8-1	0.96	0.99	1.00	BFGS 6	SOS	Linear	Tanh	326.74	2130.77	0.17%	6.45%	2734.11	4.51E-03
20	MLP 1-8-1	0.96	0.99	1.00	BFGS 6	SOS	Logistic	Tanh	265.62	1759.88	0.80%	4.63%	2294.49	3.18E-03
40	MLP 1-5-1	0.96	0.99	1.00	BFGS 5	SOS	Tanh	Exponential	1544.71	2539.20	6.27%	8.23%	3300.44	6.57E-03
40	MLP 1-5-1	0.96	0.99	1.00	BFGS 6	SOS	Linear	Tanh	180.47	2355.28	0.72%	7.31%	2994.43	5.41E-03
40	MLP 1-2-1	0.96	0.99	1.00	BFGS 6	SOS	Linear	Tanh	184.52	2325.08	1.67%	7.28%	2934.77	5.20E-03
40	MLP 1-6-1	0.96	0.98	1.00	BFGS 4	SOS	Logistic	Logistic	725.12	2046.17	3.35%	6.03%	2699.40	4.40E-03
40	MLP 1-2-1	0.96	0.99	1.00	BFGS 10	SOS	Logistic	Tanh	397.20	1761.51	1.05%	4.76%	2339.08	3.30E-03
60	MLP 1-2-1	0.95	0.98	1.00	BFGS 5	SOS	Logistic	Exponential	46.12	2638.20	0.89%	7.75%	3021.27	5.51E-03
60	MLP 1-6-1	0.96	0.99	1.00	BFGS 5	SOS	Linear	Tanh	381.15	2625.38	2.79%	8.41%	3359.27	6.81E-03

Table 2 Summary of neural network learning for the random sample size, teaching 80%, testing 10%, and validation 10% for Poland (continued)

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
60	MLP 1-6-1	0.95	0.98	1.00	BFGS 5	SOS	Logistic	Logistic	1436.54	2605.98	2.62%	6.58%	3107.29	5.83E-03
60	MLP 1-3-1	0.95	0.98	1.00	BFGS 7	SOS	Tanh	Tanh	225.51	2181.93	1.10%	6.66%	2827.63	4.83E-03
60	MLP 1-6-1	0.95	0.99	1.00	BFGS 7	SOS	Exponential	Logistic	231.31	2206.35	0.69%	5.98%	2657.70	4.26E-03
80	MLP 1-2-1	0.96	0.99	1.00	BFGS 11	SOS	Logistic	Tanh	63.00	2068.87	0.35%	6.24%	2669.34	4.30E-03
80	MLP 1-3-1	0.96	0.99	1.00	BFGS 4	SOS	Linear	Tanh	261.75	2325.06	0.42%	7.18%	2957.41	5.28E-03
80	MLP 1-2-1	0.96	0.98	1.00	BFGS 7	SOS	Logistic	Linear	553.25	2205.41	2.23%	6.74%	2759.02	4.59E-03
80	MLP 1-2-1	0.95	0.98	1.00	BFGS 6	SOS	Tanh	Logistic	81.51	2328.80	0.41%	6.55%	2719.89	4.46E-03
80	MLP 1-7-1	0.96	0.99	1.00	BFGS 5	SOS	Linear	Tanh	159.97	2374.17	0.82%	7.38%	3018.42	5.50E-03
100	MLP 1-7-1	0.96	0.99	1.00	BFGS 7	SOS	Linear	Tanh	573.15	2175.01	0.84%	6.54%	2792.33	4.71E-03
100	MLP 1-2-1	0.95	0.99	1.00	BFGS 9	SOS	Tanh	Logistic	334.46	2310.29	1.71%	6.79%	2726.36	4.49E-03
100	MLP 1-5-1	0.96	0.99	1.00	BFGS 5	SOS	Linear	Tanh	180.96	2441.21	1.90%	7.72%	3101.83	5.81E-03
100	MLP 1-2-1	0.96	0.99	1.00	BFGS 7	SOS	Linear	Tanh	573.25	2174.78	0.84%	6.54%	2791.98	4.70E-03
100	MLP 1-4-1	0.96	0.99	1.00	BFGS 5	SOS	Linear	Tanh	100.84	2331.11	0.91%	7.25%	2967.66	5.32E-03
200	MLP 1-8-1	0.96	0.99	1.00	BFGS 6	SOS	Tanh	Tanh	380.18	2350.66	2.18%	7.47%	3034.06	5.56E-03
200	MLP 1-2-1	0.96	0.98	1.00	BFGS 7	SOS	Tanh	Linear	265.66	2300.27	1.76%	7.12%	2877.84	5.00E-03
200	MLP 1-8-1	0.96	0.99	1.00	BFGS 2	SOS	Tanh	Tanh	1932.17	2744.39	4.08%	6.86%	3486.86	7.34E-03
200	MLP 1-3-1	0.96	0.98	1.00	BFGS 7	SOS	Logistic	Tanh	38.54	1969.49	0.44%	5.51%	2441.32	3.60E-03
200	MLP 1-4-1	0.95	0.98	1.00	BFGS 5	SOS	Logistic	Logistic	704.60	2296.70	1.35%	6.17%	2731.82	4.50E-03
								Minimal	38.54	1759.88	0.17%	4.63%	2294.49	3.18E-03

Table 3 Summary of neural network learning for the case of random sample sizes, teaching 70%, testing 15% and validation 15% for Hungary

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
20	MLP 1-7-1	0.9177	0.5464	0.980911	BFGS 2	SOS	Logistics	Tanh	735.32	2020.68	0.06	0.12	2253.39	0.00345
20	MLP 1-8-1	0.916178	0.544947	0.980839	BFGS 1	SOS	Tanh	Linear	841	2105.03	0.07	0.13	2346.47	0.00389
20	MLP 1-2-1	0.914687	0.543461	0.98065	BFGS 1	SOS	Logistics	Linear	839.58	2133.13	0.07	0.13	2379.01	0.143
20	MLP 1-6-1	0.917913	0.546705	0.980944	BFGS 1	SOS	Tanh	Linear	848.28	2077.29	0.07	0.13	2314.38	0.149
20	MLP 1-7-1	0.913989	0.542873	0.980582	BFGS 1	SOS	Linear	Tanh	832.85	2099.73	0.07	0.13	2340.63	0.00439
40	MLP 1-5-1	0.91931	0.549112	0.981162	BFGS 2	SOS	Exponential	Logistics	441.78	1926.46	0.04	0.11	2175.96	0.00345
40	MLP 1-7-1	0.916846	0.545852	0.980982	BFGS 1	SOS	Exponential	Logistics	823.89	2129.56	0.06	0.13	2375.44	0.00389
40	MLP 1-4-1	0.916941	0.545805	0.980935	BFGS 1	SOS	Tanh	Tanh	823.74	2079.99	0.06	0.13	2318.05	0.143
40	MLP 1-7-1	0.918319	0.547222	0.980982	BFGS 2	SOS	Logistics	Tanh	501.73	1845.68	0.04	0.11	2073.12	0.149
40	MLP 1-5-1	0.9172	0.54604	0.98094	BFGS 2	SOS	Tanh	Tanh	121.61	1704.1	0.02	0.1	1975.69	0.00439
60	MLP 1-8-1	0.921527	0.553569	0.980878	BFGS 1	SOS	Exponential	Exponential	831.66	2130.07	0.07	0.13	2375.68	0.00345
60	MLP 1-7-1	0.918276	0.547683	0.981133	BFGS 1	SOS	Exponential	Logistics	800.67	2115.37	0.06	0.13	2359.95	0.00389
60	MLP 1-3-1	0.920849	0.552001	0.981076	BFGS 1	SOS	Exponential	Exponential	823.18	2126.11	0.06	0.13	2371.42	0.143
60	MLP 1-6-1	0.918361	0.547816	0.981142	BFGS 2	SOS	Exponential	Logistics	687.35	2056.46	0.06	0.12	2298.92	0.149
60	MLP 1-6-1	0.915678	0.544579	0.98084	BFGS 1	SOS	Exponential	Logistics	796.49	2117.05	0.06	0.13	2362.14	0.00439

Table 3 Summary of neural network learning for the case of random sample sizes, teaching 70%, testing 15% and validation 15% for Hungary (continued)

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Theil
80	MLP 1-3-1	0.916325	0.545234	0.980908	BFGS 1	SOS	Exponential	Exponential	815.79	2103.11	0.06	0.13	2345.05	0.00345
80	MLP 1-2-1	0.920072	0.550374	0.981154	BFGS 1	SOS	Exponential	Exponential	818.76	2095.63	0.06	0.13	2336.11	0.00389
80	MLP 1-5-1	0.9175	0.546603	0.981042	BFGS 1	SOS	Exponential	Exponential	824.95	2119.26	0.06	0.13	2363.41	0.143
80	MLP 1-7-1	0.916054	0.545032	0.98091	BFGS 1	SOS	Exponential	Logistics	832.26	2125.18	0.07	0.13	2370.03	0.149
80	MLP 1-7-1	0.918843	0.547922	0.981009	BFGS 2	SOS	Exponential	Logistics	150.09	1706.1	0.02	0.1	1971.93	0.00439
100	MLP 1-3-1	0.918253	0.54763	0.981125	BFGS 1	SOS	Exponential	Logistics	828.48	2129.92	0.07	0.13	2375.67	0.00345
100	MLP 1-7-1	0.918563	0.548162	0.981171	BFGS 2	SOS	Exponential	Logistics	644.53	2032.77	0.05	0.12	2275.28	0.00389
100	MLP 1-6-1	0.915421	0.54414	0.980733	BFGS 1	SOS	Tanh	Tanh	830.9	2117.81	0.07	0.13	2361.56	0.143
100	MLP 1-4-1	0.920386	0.549555	0.980895	BFGS 2	SOS	Exponential	Exponential	196.54	1607.15	0.03	0.09	1855.88	0.149
100	MLP 1-5-1	0.916093	0.544868	0.980833	BFGS 1	SOS	Tanh	Tanh	843.9	2122.56	0.07	0.13	2366.63	0.00439
200	MLP 1-5-1	0.916891	0.546228	0.981078	BFGS 2	SOS	Logistics	Linear	715.75	1952.53	0.06	0.12	2175.71	0.00345
200	MLP 1-6-1	0.918015	0.547247	0.981088	BFGS 1	SOS	Exponential	Logistics	844.2	2137.25	0.07	0.13	2383.63	0.00389
200	MLP 1-7-1	0.921302	0.553015	0.980961	BFGS 2	SOS	Exponential	Logistics	212.52	1767.86	0.03	0.1	2028.01	0.143
200	MLP 1-7-1	0.918004	0.547012	0.981016	BFGS 2	SOS	Linear	Tanh	239.41	1621.15	0.03	0.1	1863.05	0.149
200	MLP 1-5-1	0.920794	0.551922	0.981088	BFGS 1	SOS	Exponential	Logistics	820.08	2128.97	0.06	0.13	2374.89	0.00439
								Minimal	121.61	1607.15	0.02	0.09	1855.88	0.00345

Table 4 Summary of neural network learning for the case of random sample sizes, teaching 80%, testing 10% and validation 10% for Hungary

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Thail
20	MLP 1-6-1	0.874289	0.996897	0.942829	BFGS 5	SOS	Exponential	Tanh	311.89	706.44	0.02	0.04	988.45	0.00345
20	MLP 1-2-1	0.876307	0.996786	0.948805	BFGS 3	SOS	Tanh	Tanh	52.96	748.92	0.01	0.04	962.06	0.00389
20	MLP 1-2-1	0.87556	0.996824	0.947205	BFGS 6	SOS	Linear	Tanh	22.45	753.76	0.01	0.04	977	0.143
20	MLP 1-8-1	0.876835	0.996664	0.953241	BFGS 5	SOS	Logistics	Tanh	671.71	858.51	0.04	0.05	1196.22	0.149
20	MLP 1-6-1	0.87716	0.996789	0.947159	BFGS 13	SOS	Logistics	Linear	28.71	742.29	0.01	0.04	951.79	0.00439
40	MLP 1-5-1	0.874536	0.996596	0.957242	BFGS 4	SOS	Exponential	Logistics	196.91	707.87	0.01	0.04	970.65	0.00345
40	MLP 1-7-1	0.876775	0.996745	0.950236	BFGS 6	SOS	Linear	Tanh	51.80	771.3	0	0.05	989.16	0.00389
40	MLP 1-3-1	0.872934	0.996476	0.959214	BFGS 4	SOS	Logistics	Linear	839.34	970.6	0.05	0.06	1344.76	0.143
40	MLP 1-2-1	0.876637	0.996571	0.957499	BFGS 4	SOS	Logistics	Linear	282.46	750.82	0.02	0.04	1016.12	0.149
40	MLP 1-7-1	0.875931	0.996576	0.956499	BFGS 6	SOS	Logistics	Tanh	28.25	858.05	0	0.05	1070.45	0.00439
60	MLP 1-5-1	0.867988	0.996225	0.967635	BFGS 5	SOS	Logistics	Tanh	661.15	940.19	0.04	0.06	1295.86	0.00345
60	MLP 1-2-1	0.875055	0.996513	0.958619	BFGS 5	SOS	Logistics	Tanh	399.06	740.67	0.03	0.04	1051.09	0.00389
60	MLP 1-3-1	0.869451	0.996311	0.964911	BFGS 5	SOS	Logistics	Tanh	822.14	978.12	0.05	0.06	1344.45	0.143
60	MLP 1-5-1	0.875211	0.996527	0.958405	BFGS 4	SOS	Tanh	Tanh	479.95	825.38	0.03	0.05	1138.19	0.149
60	MLP 1-4-1	0.875556	0.99656	0.956983	BFGS 4	SOS	Logistics	Tanh	454.73	880.79	0.03	0.05	1173.19	0.00439

Table 4 Summary of neural network learning for the case of random sample sizes, teaching 80%, testing 10% and validation 10% for Hungary (continued)

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									ME	MAE	MPE	MAPE	SSE	Thail
80	MLP 1-8-1	0.870623	0.996392	0.962112	BFGS 5	SOS	Logistics	Tanh	550.85	840.2	0.03	0.05	1170.91	0.00345
80	MLP 1-4-1	0.837374	0.995634	0.977876	BFGS 4	SOS	Logistics	Tanh	577.29	1019.11	0.03	0.06	1387.03	0.00389
80	MLP 1-4-1	0.875917	0.996562	0.957034	BFGS 6	SOS	Logistics	Tanh	2.22	847.55	0	0.05	1063.59	0.143
80	MLP 1-5-1	0.853203	0.996208	0.964819	BFGS 3	SOS	Logistics	Linear	853.88	999.58	0.05	0.06	1376.97	0.149
80	MLP 1-4-1	0.868821	0.996768	0.951952	BFGS 4	SOS	Logistics	Logistics	506.44	782.67	0.03	0.05	1088.17	0.00439
100	MLP 1-4-1	0.871726	0.996382	0.962893	BFGS 5	SOS	Logistics	Tanh	446.26	789.52	0.03	0.05	1111.78	0.00345
100	MLP 1-5-1	0.874064	0.996445	0.96137	BFGS 5	SOS	Tanh	Tanh	627.76	854.45	0.04	0.05	1193.77	0.00389
100	MLP 1-8-1	0.872415	0.996428	0.961366	BFGS 5	SOS	Logistics	Tanh	52.38	783.75	0	0.05	1005.42	0.143
100	MLP 1-2-1	0.868535	0.996285	0.9655	BFGS 6	SOS	Logistics	Tanh	312.77	937.52	0.02	0.06	1202.23	0.149
100	MLP 1-7-1	0.866288	0.996416	0.960127	BFGS 5	SOS	Logistics	Linear	373.37	831.58	0.02	0.05	1120.59	0.00439
200	MLP 1-3-1	0.875243	0.99653	0.958366	BFGS 6	SOS	Tanh	Tanh	59.97	791.69	0	0.05	1016.92	0.00345
200	MLP 1-8-1	0.873761	0.996464	0.960253	BFGS 4	SOS	Logistics	Tanh	637.07	912.94	0.04	0.05	1254.17	0.00389
200	MLP 1-8-1	0.876175	0.996564	0.95767	BFGS 4	SOS	Tanh	Tanh	331.26	750.3	0.02	0.05	1030.26	0.143
200	MLP 1-4-1	0.87217	0.996412	0.961881	BFGS 4	SOS	Logistics	Tanh	675.4	950.02	0.04	0.06	1303.89	0.149
200	MLP 1-2-1	0.869323	0.99631	0.96485	BFGS 4	SOS	Tanh	Tanh	569.68	875.36	0.03	0.05	1213.63	0.00439
								Minimal	2.22	706.44	0	0.04	951.79	0.00345

Based on the research findings, the incidence of road accidents in Poland is expected to stabilise over the next few years. However, there may also be a slight rise in accidents occurring on Polish roads. The results are contingent upon the selection of the random sample size. The average percentage error can be reduced by increasing the proportion of the training set relative to the test and validation sets. Specifically, when using a training set comprising 70%, a test set of 15%, and a validation set of 15% (in a 70-15-15 ratio), the error rate was recorded at 5.68%. In contrast, when the proportions were adjusted to 80-10-10, the error decreased to 4.63%. These results are influenced by various factors, notably the growing number of vehicles on Polish roads and the impact of the pandemic in recent years (see Fig. 3).

From the research findings, it can be inferred that the frequency of road accidents in Hungary is likely to stabilise in the upcoming years, although there could also be a slight uptick in the number of accidents on Hungarian roads. The chosen random sample size influences the outcomes. Increasing the training set's proportion relative to the test and validation sets can minimise the average percentage error. For instance, with a training group consisting of 70%, a test group of 15%, and a validation group of 15% (in a 70-15-15 distribution), the recorded error was 0.095%. Conversely, when the proportions were set to 80-10-10, the error dropped to 0.042%. These findings are affected by several factors, particularly the rising number of vehicles on Hungarian roads and the ongoing impact of the pandemic in recent years (refer to Fig. 4)

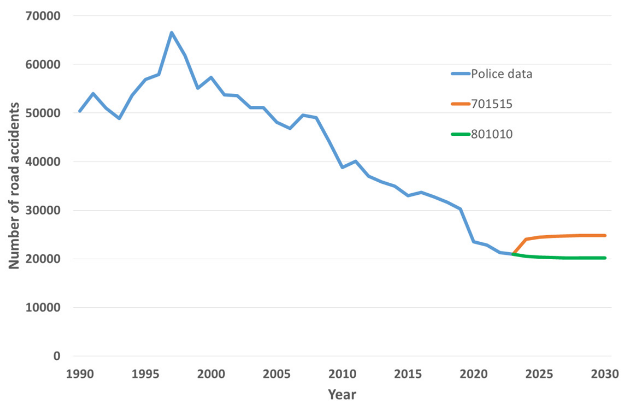

Fig. 3 Projected number of road accidents for 2022-2030 in Poland

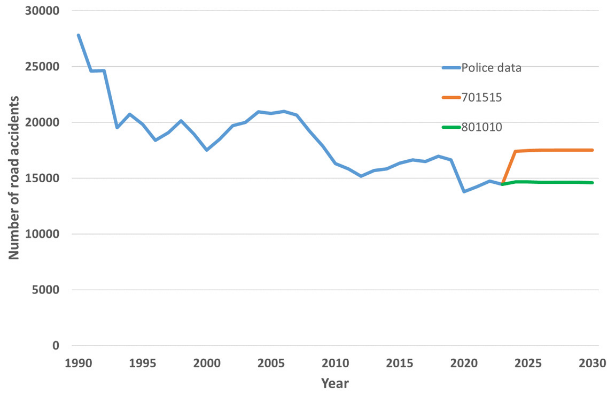

Fig. 4 Projected number of road accidents for 2022-2030 in Hungary

4 Conclusions

Neural networks were employed to forecast the number of accidents in Poland and Hungary, with the research conducted within the Statistica environment. The program optimised the weights utilised in the analysis to minimise both the mean absolute error and the mean absolute percentage error.

The findings suggest that we can anticipate stabilising the number of road accidents, with a slight increase in both examined countries. This trend is primarily influenced by the growing number of vehicles on the roads and the ongoing pandemic. The calculated forecast errors indicate the reliability of the models applied.

Based on the forecasts generated, it is essential to implement measures to reduce the incidence of road accidents further. Such initiatives could include enforcing stricter fines for traffic violations on Polish roads, effective January 1, 2022. The pandemic, which significantly altered accident statistics, has undoubtedly impacted the research results obtained.

In future studies, the authors intend to consider additional factors that influence accident rates and employ various statistical methods to estimate the number of road accidents. Potential factors for consideration may include traffic volume, weather conditions, the age of the accident perpetrator, and exponential methods for predicting accident numbers.

Acknowledgement

"Project No. Kdp-Ikt-2023-900-I1-00000957/0000003 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the Kdp-2023 Funding Scheme."

References

[1] Péter, B., Sütheö, G., Lakatos, I. "Accident Risk Analysis of Road Accidents Involving Personal Injury", Periodica Polytechnica Transportation Engineering, 53(3), pp. 301–307, 2025. https://doi.org/10.3311/PPtr.38004

[2] Al-Madani, H. M. N. "Global road fatality trends'estimations based on country-wise microlevel data", Accident Analysis & Prevention, 111, pp. 297–310, 2018. https://doi.org/10.1016/j.aap.2017.11.035

[3] Arteaga, C., Paz, A., Park, J. "Injury severity on traffic crashes: A text mining with an interpretable machine-learning approach", Safety Science, 132, 104988, 2020. https://doi.org/10.1016/j.ssci.2020.104988

[4] Bąk, I., Cheba, K., Szczecińska, B. "The statistical analysis of road traffic in cities of Poland", Transportation Research Procedia, 39, pp. 14–23, 2019. https://doi.org/10.1016/j.trpro.2019.06.003

[5] Singh, G., Pal, M., Yadav, Y., Singla, T. "Deep neural network-based predictive modeling of road accidents", Neural Computing and Applications, 32(16), pp. 12417–12426, 2020. https://doi.org/10.1007/s00521-019-04695-8

[6] Rabbani, M. B. A., Musarat, M. A., Alaloul, W. S., Ayub, S., Bukhari, H., Altaf, M. "Road Accident Data Collection Systems in Developing and Developed Countries: A Review", International Journal of Integrated Engineering, 14(1), pp. 336–352, 2022. https://doi.org/10.30880/ijie.2022.14.01.031

[7] Chand, A., Jayesh, S., Bhasi, A. B. "Road traffic accidents: An overview of data sources, analysis techniques and contributing factors", Materials Today: Proceedings, 47, pp. 5135–5141, 2021. https://doi.org/10.1016/j.matpr.2021.05.415.

[8] Yijing, H., Wei, W., He, Y., Qihong, W., Kaiming, X. "Intelligent algorithms for incident detection and management in smart transportation systems", Computers and Electrical Engineering, 110, 108839, 2023. https://doi.org/10.1016/j.compeleceng.2023.108839

[9] Wei, C., Han, F., Fan, Z., Shi, L., Peng, C. "Efficient license plate recognition in unconstrained scenarios", Journal of Visual Communication and Image Representation, 104, 104314, 2024. https://doi.org/10.1016/j.jvcir.2024.104314

[10] Ghandour, A. J., Hammoud, H., Dimassi, M., Krayem, H., Haydar, J., Issa, A. "Allometric scaling of road accidents using social media crowd-sourced data", Physica A: Statistical Mechanics and its Applications, 545, 123534, 2020. https:/doi.org/10.1016/j.physa.2019.123534

[11] Berhanu, Y., Schröder, D., Wodajo, B. T., Alemayehu, E. "Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks", Results in Engineering, 23, 102737, 2024. https://doi.org/10.1016/j.rineng.2024.102737

[12] Vilaça, M., Silva, N., Coelho, M. C. "Statistical analysis of the occurrence and severity of crashes involving vulnerable road users", Transportation Research Procedia, 27, pp. 1113–1120, 2017. https://doi.org/10.1016/j.trpro.2017.12.113

[13] Chudy-Laskowska, K., Pisula, T. "Prognozowanie liczby wypadków drogowych na Podkarpaciu" (Forecast of the number of road accidents in Poland), Logistyka, 2015(4), pp. 2782–2796, 2015. (in Polish)

[14] Dudek, G. "Forecasting Time Series with Multiple Seasonal Cycles Using Neural Networks with Local Learning", In: Artificial Intelligence and Soft Computing, 12th International Conference, ICAISC 2013, Zakopane, Poland, 2013, pp. 52–63. https://doi.org/10.1007/978-3-642-38658-9_5

[15] Dutta, B., Barman, M. P., Patowary, A. N. "Application of Arima model for forecasting road accident deaths in India", International Journal of Agricultural and Statistical Sciences, 16(2), pp. 607–615, 2020.

[16] Chen, C. "Analysis and Forecast of Traffic Accident Big Data", ITM Web Conferences, 12, 04029, 2017. https:/doi.org/10.1051/itmconf/20171204029

[17] Procházka, J., Flimmel, S., Čamaj, M., Bašta, M. "Modelling the Number of Road Accidents", In: 20th AMSE, Applications of Mathematics and Statistics in Economics, International Scientific Conference, Szklarska Poręba, Poland, 2017, pp. 355–364. https://doi.org/10.15611/amse.2017.20.29

[18] Sunny, C. M., Nithya, S., Sinshi, K. S., Vinodini, V., Lakshmi, A., Anjana, S., Manojkumar, T. K. "Forecasting of Road Accident in Kerala: A Case Study", In: 2018 International Conference on Data Science and Engineering (ICDSE), Kochi, India, 2018, pp. 1–5. ISBN 978-1-5386-4855-1 https://doi.org/10.1109/ICDSE.2018.8527825

[19] Szmuksta-Zawadzka, M., Zawadzki, J. "Forecasting on the basis of Holt-Winters models for complete and incomplete data", Ekonometria, 24(38), pp. 85–99, 2009. (in Polish)

[20] Statistics Poland "Central Statistical Office of Poland (GUS)", [online] Available at: https://stat.gov.pl/ [Accessed: 17 April 2022]

[21] Statystyka "The Police of the Republic of Poland", [online] Available at: https://statystyka.policja.pl/ [Accessed: 17 April 2022]

[22] Hungarian Central Statistical Office "Hungarian Central Statistical Office (KSH)" [online] Available at: https://www.ksh.hu/stadat_files/ege/hu/ege0061.html; [Accessed: 17 April 2022]

[23] Gorzelańczyk, P., Pyszewska, D., Kalina, T., Jurkovič, M. "Analysis of road traffic safety in the Pila poviat", Scientific Journal of Silesian University of Technology. Series Transport, 107, pp. 33–52, 2020. https://doi.org/10.20858/sjsutst.2020.107.3.

[24] StatSoft "StatSoft (Data mining techniques)", [online] Available at: https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fstdatmin.htm [Accessed: 17 April 2022]

[25] Hamdan, N., Sipos, T. "Advancements in Machine Learning for Traffic Accident Severity Prediction: A Comprehensive Review", Periodica Polytechnica Transportation Engineering, 53(3), pp. 347–355, 2025. https://doi.org/10.3311/PPtr.40369

[26] Lake. B. M., Ullman, T. D., Tenenbaum, J. B., Gershman, S. J. "Building machines that learn and think like people", Behavioral and Brain Sciences, 40, e253, 2017. http://doi.org/10.1017/S0140525X16001837

[27] Oronowicz-Jaśkowiak, W. "The application of neural networks in the work of forensic experts in child abuse cases", Advances in Psychiatry and Neurology, 28(4), pp. 273–282, 2019. https://doi.org/10.5114/ppn.2019.92489