Rocznik Ochrona Środowiska

Volume 26 Year 2024 ISSN 2720-7501 pp. 603-615

https://doi.org/10.54740/ros.2024.053 open access

Received: November 2024 Accepted: November 2024 Published: November 2024

Predicting Road Accident Counts in Poland and the Czech Republic Using Neural Network Models

Piotr Gorzelańczyk[1], Lenka Ližbetinová[2], Jan Pečman[3]*

[1]Department of Transport, Stanislaw Staszic State University of Applied Sciences in Pila, Poland
https://orcid.org/0000-0001-9662-400X

[2]Department of Transport and Logistics, Institute of Technology and Business in České Budějovice, Czech Republic
https://orcid.org/0000-0001-8969-2071

[3]Department of Transport and Logistics, Institute of Technology and Business in České Budějovice, Czech Republic
https://orcid.org/0000-0002-4163-4881

**corresponding author's e-mail: lizbetinova@mail.vstecb.cz*

Abstract: Every year, there is a decline in the number of car accidents reported in Poland, the Czech Republic, and globally. While recent trends due to the pandemic have influenced these figures, the overall rate remains significant. Therefore, it is crucial to take measures aimed at reducing this number. The primary focus of this article is to analyze the traffic accident statistics for Poland and the Czech Republic. Annual data regarding traffic incidents in both countries has been scrutinized to achieve this. Projections for 2024 to 2030 have been developed based on police reports. Various neural network models were utilized to forecast the number of accidents. The findings indicate that the number of traffic incidents is likely to stabilize. This stabilization can be viewed in the context of the increasing number of vehicles on the roads and the expansion of new highways. Additionally, selecting sample sizes for training, testing, and validation is crucial in influencing the results. Forecasting the number of traffic accidents is important for environmental protection, as accidents can lead to air and water pollution and increase noise, negatively affecting human health and ecosystems.

Keywords: road accident, pandemic, forecasting, neural networks, Poland, Czech Republic

1. Literature Review

Road accidents involve incidents that result in injuries or fatalities to drivers, alongside causing property damage. According to estimates by the World Health Organization (WHO), approx. 1.3 million people die in traffic accidents each year. On a global scale, traffic accidents contribute to a 3% decrease in GDP for the average country. Traffic accidents are the primary cause of death for people aged 5 to 29 years. The United Nations General Assembly aims to reduce fatalities and injuries from traffic accidents by 50% by 2030 (World Health Organization 2018).

A key factor in evaluating the seriousness of a traffic incident is its overall scope. Assessing accident severity is vital for authorities to formulate effective traffic safety regulations with the goal of reducing accidents and mitigating hurts, fatalities, and ownership damage (Tambouratzis et al. 2014, Zhu et al. 2019). Before implementing measures to minimize accident severity, it is important to identify the primary factors that contribute to it (Arteaga et al. 2020). A multi-node Deep Neural Network (DNN) model, which was proposed by (Yang et al. 2022), predicts different levels of injury, fatality, and ownership damage, allowing for a detailed and accurate assessment of the seriousness of traffic accidents (Gorzelanczyk & Huk 2022).

Accident statistics are derived from various sources. Typically, government officials rely on relevant governmental agencies to gather and analyze this data. Key sources of information include police reports, databases from insurance companies, and hospital records. Consequently, there is an increasing trend in the transportation sector towards more comprehensive data analysis of traffic accidents (Chen 2017).

Intelligent transported systems are now the main data source for analyzing and predicting traffic events and MaaS (Dyczkowska et al. 2023). Information is gathered through GPS devices installed in vehicles in motion. Additionally, roadside microwave vehicle detection systems can continuously capture data regarding moving vehicles, including details such as vehicle type, speed, and traffic volume (Hudec & Cződörová 2022, Khaliq et al. 2019). In addition, significant volumes of traffic data can be gathered over a defined period using license plate recognition systems (Rajput et al. 2015). Social media also offers a potential source of information on traffic incidents, although the reliability of such reports may be limited by the inexperience of those providing the information (Zheng et al. 2018).

Utilizing a diverse array of data sources presents certain challenges before traffic accident information can be deemed valuable. Accurate analytical outcomes can be achieved by integrating various types of traffic accident data (Abdullah & Emam 2016).

Statistical analysis with goal to evaluate the seriousness of the issue and to determine the relationships between traffic participants and accidents, performed (Vilaca et al. 2017). The results of their study support the introduction of stricter traffic safety measures and improvements to traffic law standards.

Bąk et al. (2019) conducted a statistical study on traffic safety in a specific region of Poland, using the number of traffic accidents as a key indicator to investigate the causes of these incidents. This study utilized multivariate statistical analysis to examine the safety factors related to the causes of accidents. The specific traffic issue being examined determines the selection of data sources for accident analysis. Accident prediction and prevention accuracy is enhanced when statistical models are integrated with additional data from real driving conditions or insights derived from intelligent traffic systems (Chand et al. 2021).

Forecasting the number of road accidents is important for logistics (Dyczkowska et al. 2023a) and environmental protection (Cubranic-Dobrodolac et al. 2020, Čubranić-Dobrodolac et al. 2022), as accidents can lead to air and water pollution and increase noise, which negatively affects human health, ecosystems, sustainable urban mobility (Chamier-Gliszczynski 2016) and system mobility (Chamier-Gliszczyński 2012, Chamier-Gliszczyński 2012a). Road accidents often result in oil and chemical spills, which can contaminate the ground and groundwater and emit harmful substances into the atmosphere, contributing to the deterioration of air quality. In addition, accidents generate not only physical damage but also noise, which affects the quality of life of nearby residents, disrupting peace and leading to stress and health problems associated with prolonged noise exposure (Čubranić-Dobrodolac et al. 2022).

The selection of data sources for accident analysis depends on the specific traffic issue being investigated (combined with additional data from actual driving conditions or observations obtained from intelligent traffic systems). In 2023, it was found that Poland experienced 5.57 traffic accidents per 10,000 people. During the same period, the population of the Czech Republic was 10.67 million, with 20,768 reported traffic incidents. This indicates that the Czech Republic had 3.15 times more traffic accidents per 10,000 people compared to Poland:

$$NRA = \frac{NR}{NI} \cdot 10000 \tag{1}$$

where:
NR – quantity of road accidents,
NI – quantity of inhabitants.

The authors utilized the previously mentioned data to estimate the number of accidents occurring on roadways in Poland and the Czech Republic. They employed neural networks to predict the incidence rates of traffic road accidents in both countries (Chovancova et al. 2017).

2. Materials and Methods

A substantial number of traffic accidents take place on roads yearly. The expected figures have been affected by the recent decrease in traffic accidents due to the pandemic. However, even during the pandemic, road accidents remain prevalent. Therefore, it is essential to make every effort to reduce these numbers and identify the types of routes that contribute to the highest incidence of traffic accidents (shown in Fig. 1, Fig. 2).

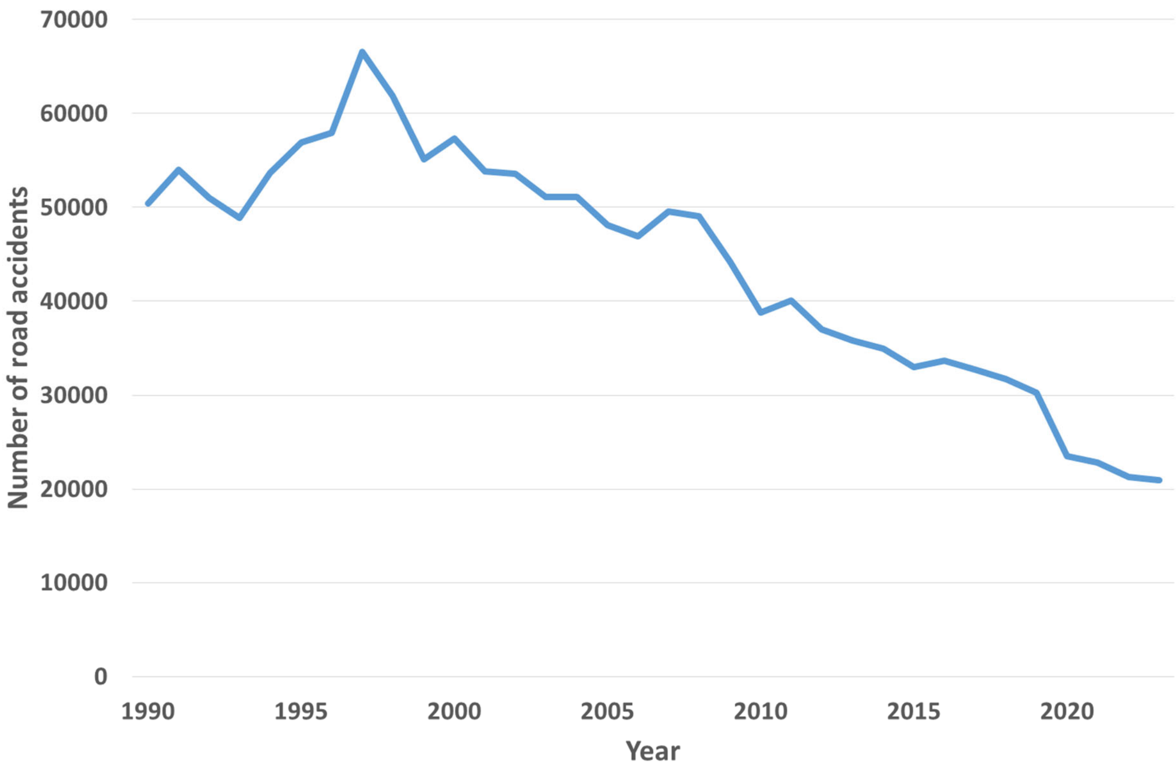

Fig. 1. Number of road accidents in Poland between 1990 and 2023 (Polish Police 2024)

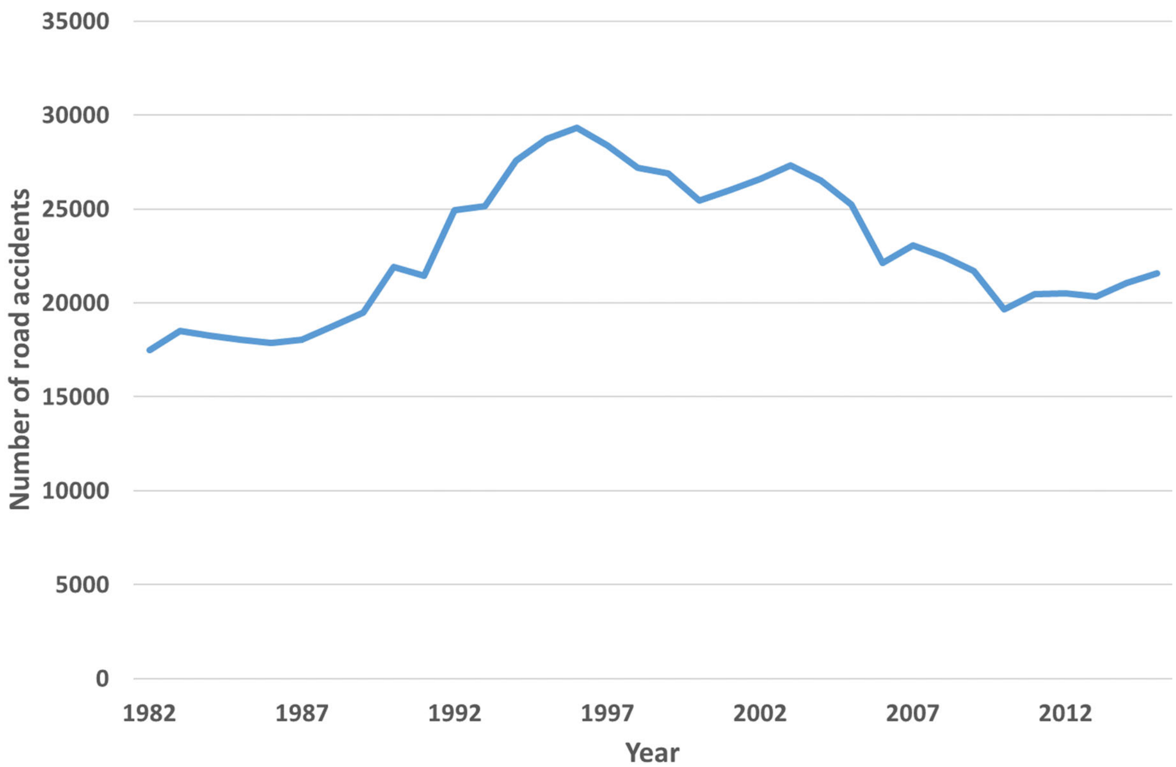

Fig. 2. Number of road accidents in the Czech Republic between 1990 and 2023 (Czech Statistics Office 2024)

Specific neural network models were employed to estimate the frequency of traffic accidents in Poland and the Czech Republic. This method is advantageous as it simulates the workings of the human brain. A neural network comprises nodes that process input data along with corresponding weights, biases, and output data. Statistica software was used to identify the optimal weights during the analysis. The accuracy of the predictions produced by this approach depends on the selected model and its parameters (Stopka 2022).

A neural network can be understood as a mathematical framework that functions like the nervous system. Typically, these networks consist of multiple layers that collectively form their architecture. The initial layer processes various data types, such as text, images, numbers, and audio, through a training process. Before reaching a final decision, the network can evaluate numerous inputs.

The essential elements of neural networks are artificial neurons, which function as mathematical models that replicate the behavior of biological neurons. These artificial neurons are akin to biological ones in that they accept multiple inputs but generate a single output value, much like the functioning of dendrites in real neurons. The development of artificial intelligence is heavily focused on neural networks, to create models that exhibit intelligent behavior, including the ability to establish a hierarchy of knowledge (Lake et al. 2017).

Neural networks find applications across a diverse array of fields. For instance, the power systems enable users to stream on-demand series by analyzing their viewing history to recommend films that align with their preferences. Additionally, neural networks facilitate text translation on platforms like Google Translate and help personalize product suggestions for bidders in online auctions. Moreover, neural network forecasting is employed to predict the frequency of traffic incidents (Marr 2019, Oronowicz-Jaśkowiak 2019) and production processes (Kielc et al. 2018).

A chosen neural network model is utilized to predict the occurrence of traffic accidents in the counties being studied. One of the key benefits of this technology is its ability to replicate the functioning of the human brain. A neural network consists of nodes that include inputs, weights, biases, and outputs (Wu et al. 2016, Yu 2019).

Models of the nervous system's operation are used to create mathematical structures known as neural network approaches. The network architecture is often composed of several levels. Through a process known as training, the first of these, the input layer, retains knowledge about text, numbers, pictures, and sound. Thousands of inputs may be used in this process, from which the network extracts certain conclusions. The concealed layer, sometimes known as the transition layer, is another layer that has been studied. Such layers may be many. The output layer (Fig. 3) is the final layer covered (Yadav & Rishi 2022).

The Statistica software, featuring integrated modules for artificial neural networks, refined the weights during the testing process. A multilayer perceptron (MLP) neural network, which included layers of hidden neurons, was employed for the predictions. In the cases examined, the number of neurons in the hidden layer varied from two to eight. The output layer comprised a single neuron that provided the time series output values for the number of traffic incidents (Hudec et al. 2021). The success of the predictive techniques employed depends on the chosen model and its parameters (Witt 2023). Predictive accuracy was evaluated based on various prediction errors calculated using specific formulas (2-7):

- ME – mean error

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

- MAE – mean error

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

- MPE – mean percentage error

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

- $MAPE$ – mean absolute percentage error

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

- SSE – mean square error

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

- M^2 – Theila measure

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

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

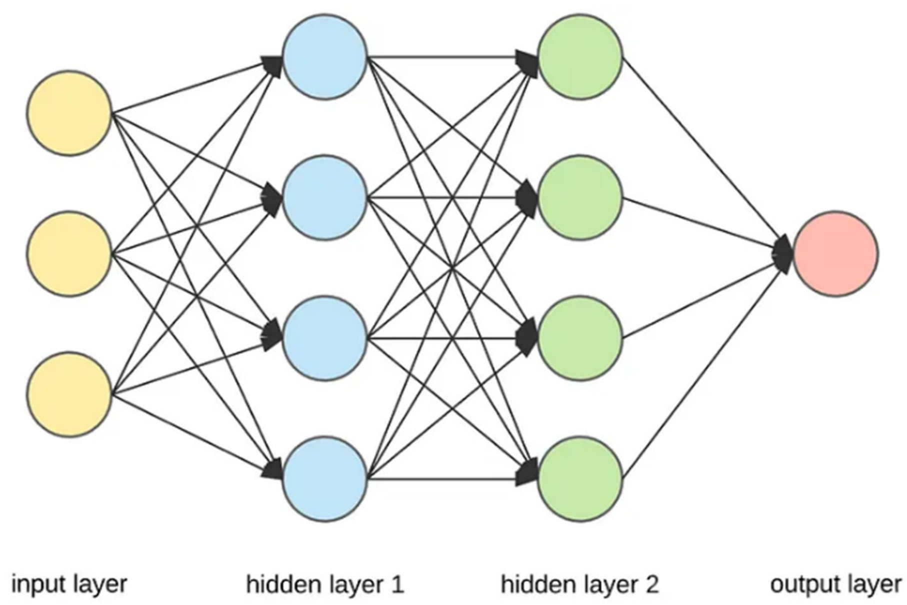

Fig. 3. Neural network models (Wójcik 2014)

To forecast the number of future traffic accidents, neural network models were utilized that demonstrated the lowest mean percentage error (MPE) and mean absolute percentage error (MAPE).

3. Results

Data for the Czech Republic were obtained from the Czech Statistical Office (Czech Statistics Office 2024), while data from the Polish Police covering the years 1990 to 2023 (Polish Police 2024) predicted the annual amount of traffic accidents on Polish roads. In every case, two random sample sizes were assumed when using Statistica software for research:

1. teaching 70%, test 15%, validation 15%.
2. teaching 80%, testing 10%, validation 10%,

using 20, 40, 60, 80, 100, and 200 learning networks, for which the MP error value was negligible (Table 1-4).

Study Results on Traffic Accidents in Poland (Fig. 4):

- Accident Trends:
  - The study suggests that there may be a slight increase in traffic occurrences on Polish roads.
  - Nonetheless, the total number of accidents is anticipated to stabilize in the next years.
- Impact of Sample Proportions:
  - The number of random samples used influences the results.
  - A larger proportion of the training group relative to the test and validation groups reduces the average percentage error.
- Error Rates:
  - First Exam (80-10-10):
    - Learning Group: 80%
    - Test Group: 10%
    - Validation Group: 10%
    - Average Percentage Error: 4.63%
  - Second Exam (70-15-15):
    - Learning Group: 70%
    - Test Group: 15%
    - Validation Group: 15%
    - Average Percentage Error: 5.68%
- Factors Influencing Findings:
  - The findings are affected by:
    - An increased number of cars on roads in Poland.
    - The impact of the Covid-19 pandemic.

Table 1. Summary of neural network learning for the case of random sample sizes of 70% learning, 15% testing, and 15% validation 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	M L P 1-7-1	0.97	0.98	0.99	BFG S 12	S O S	Tanh	Logistic	932.03	2374.97	2.85%	6.69%	2747.76	4.56E-03
20	M L P 1-8-1	0.97	0.98	0.99	BFG S 4	S O S	Linear	Logistic	664.14	2092.80	3.19%	6.46%	2667.94	4.30E-03
20	M L P 1-2-1	0.97	0.98	0.99	BFG S 4	S O S	Exponential	Exponential	1168.74	2119.69	3.23%	5.69%	2657.54	4.26E-03
20	M L P 1-4-1	0.96	0.97	0.99	BFG S 4	S O S	Exponential	Linear	1815.06	2891.74	2.84%	7.52%	3551.21	7.61E-03
20	M L P 1-3-1	0.97	0.97	0.99	BFG S 42	S O S	Tanh	Logistic	1113.77	2387.70	3.25%	6.64%	2802.32	4.74E-03
40	M LP 1-7-1	0.97	0.98	0.99	BFG S 8	S O S	Exponential	Exponential	837.43	2001.24	2.77%	5.68%	2484.15	3.72E-03
40	M L P 1-3-1	0.97	0.98	0.99	BFG S 8	S O S	Exponential	Exponential	909.41	2052.58	2.80%	5.71%	2526.15	3.85E-03
40	M L P 1-2-1	0.97	0.96	0.99	BFG S 7	S O S	Logistic	Logistic	1110.51	2279.94	3.35%	6.40%	2764.05	4.61E-03
40	M L P 1-8-1	0.97	0.96	0.99	BFG S 5	S O S	Logistic	Exponential	1483.15	2363.70	4.57%	6.80%	2927.92	5.17E-03
40	M L P 1-3-1	0.96	0.95	0.99	BFG S 5	S O S	Logistic	Exponential	1035.60	2590.23	3.60%	7.73%	3048.65	5.61E-03
60	M L P 1-4-1	0.97	0.97	0.99	BFG S 7	S O S	Tanh	Logistic	1031.48	2377.94	3.10%	6.68%	2772.29	4.64E-03
60	M L P 1-7-1	0.97	0.97	0.99	BFG S 5	S O S	Tanh	Logistic	777.04	2415.30	2.17%	6.63%	2763.52	4.61E-03
60	M L P 1-2-1	0.97	0.97	0.99	BFG S 9	S O S	Exponential	Logistic	1109.93	2233.82	3.17%	6.12%	2715.75	4.45E-03
60	M L P 1-5-1	0.97	0.97	0.99	BFG S 12	S O S	Tanh	Exponential	1090.27	2283.83	3.28%	6.40%	2721.35	4.47E-03
60	M L P 1-6-1	0.97	0.98	0.99	BFG S 5	S O S	Exponential	Exponential	1071.39	1987.71	3.83%	5.94%	2590.07	4.05E-03
80	M L P 1-5-1	0.97	0.97	0.99	BFG S 7	S O S	Tanh	Logistic	1040.10	2404.10	3.15%	6.78%	2796.23	4.72E-03
80	M L P 1-8-1	0.97	0.98	0.99	BFG S 14	S O S	Exponential	Logistic	1023.90	2217.06	3.03%	6.15%	2661.45	4.27E-03
80	M L P 1-3-1	0.97	0.98	0.99	BFG S 7	S O S	Exponential	Logistic	801.65	2239.47	2.48%	6.27%	2623.54	4.15E-03
80	M L P 1-7-1	0.97	0.96	0.99	BFG S 7	S O S	Logistic	Logistic	978.70	2426.95	2.68%	6.63%	2851.96	4.91E-03
80	M L P 1-7-1	0.97	0.98	0.99	BFG S 13	S O S	Exponential	Logistic	873.08	2237.49	2.54%	6.16%	2638.93	4.20E-03
100	M L P 1-8-1	0.97	0.98	0.99	BFG S 18	S O S	Exponential	Logistic	1021.62	2260.14	2.99%	6.25%	2688.97	4.36E-03
100	M L P 1-5-1	0.97	0.97	0.99	BFG S 6	S O S	Logistic	Logistic	1108.50	2402.26	3.24%	6.69%	2819.42	4.80E-03
100	M L P 1-4-1	0.97	0.98	0.99	BFG S 11	S O S	Logistic	Exponential	909.58	2320.20	2.93%	6.65%	2707.86	4.43E-03
100	M L P 1-2-1	0.96	0.95	0.99	BFG S 7	S O S	Tanh	Logistic	1114.15	2426.31	4.15%	7.43%	3005.79	5.45E-03
100	M L P 1-2-1	0.97	0.97	0.99	BFG S 8	S O S	Tanh	Logistic	894.03	2347.67	2.83%	6.67%	2720.34	4.47E-03
200	M L P 1-6-1	0.96	0.96	0.99	BFG S 8	S O S	Tanh	Logistic	644.88	2480.32	2.22%	7.15%	2814.19	4.78E-03
200	M L P 1-6-1	0.97	0.96	0.99	BFG S 7	S O S	Tanh	Logistic	770.95	2330.09	2.51%	6.64%	2702.56	4.41E-03
200	M L P 1-3-1	0.97	0.97	0.99	BFG S 10	S O S	Logistic	Logistic	970.77	2347.01	2.97%	6.61%	2750.08	4.56E-03
200	M L P 1-2-1	0.95	0.92	0.99	BFG S 7	S O S	Logistic	Exponential	319.55	2657.15	0.32%	7.61%	3035.78	5.56E-03
200	M L P 1-4-1	0.97	0.97	0.99	BFG S 6	S O S	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. A summary of neural network learning for the scenario of random sample sizes, with 80% allocated for training, 10% for testing, and 10% for validation, 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	M L P 1-5-1	0.96	0.99	1.00	BFG S 8	S O S	Logistic	Linear	422.40	1830.32	0.90%	5.12%	2362.11	3.37E-03
20	M L P 1-5-1	0.96	0.99	1.00	BFG S 5	S O S	Linear	Tanh	420.05	2152.64	0.39%	6.51%	2773.07	4.64E-03
20	M L P 1-3-1	0.96	0.99	1.00	BFG S 63	S O S	Tanh	Logistic	702.37	1986.10	2.31%	5.57%	2455.03	3.64E-03
20	M L P 1-8-1	0.96	0.99	1.00	BFG S 6	S O S	Linear	Tanh	326.74	2130.77	0.17%	6.45%	2734.11	4.51E-03
20	M L P 1-8-1	0.96	0.99	1.00	BFG S 6	S O S	Logistic	Tanh	265.62	1759.88	0.80%	4.63%	2294.49	3.18E-03
40	M L P 1-5-1	0.96	0.99	1.00	BFG S 5	S O S	Tanh	Exponential	1544.71	2539.20	6.27%	8.23%	3300.44	6.57E-03
40	M L P 1-5-1	0.96	0.99	1.00	BFG S 6	S O S	Linear	Tanh	180.47	2355.28	0.72%	7.31%	2994.43	5.41E-03
40	M L P 1-2-1	0.96	0.99	1.00	BFG S 6	S O S	Linear	Tanh	184.52	2325.08	1.67%	7.28%	2934.77	5.20E-03
40	M L P 1-6-1	0.96	0.98	1.00	BFG S 4	S O S	Logistic	Logistic	725.12	2046.17	3.35%	6.03%	2699.40	4.40E-03
40	M L P 1-2-1	0.96	0.99	1.00	BFG S 10	S O S	Logistic	Tanh	397.20	1761.51	1.05%	4.76%	2339.08	3.30E-03
60	M L P 1-2-1	0.95	0.98	1.00	BFG S 5	S O S	Logistic	Exponential	46.12	2638.20	0.89%	7.75%	3021.27	5.51E-03
60	M L P 1-6-1	0.96	0.99	1.00	BFG S 5	S O S	Linear	Tanh	381.15	2625.38	2.79%	8.41%	3359.27	6.81E-03
60	M L P 1-6-1	0.95	0.98	1.00	BFG S 5	S O S	Logistic	Logistic	1436.54	2605.98	2.62%	6.58%	3107.29	5.83E-03
60	M L P 1-3-1	0.95	0.98	1.00	BFG S 7	S O S	Tanh	Tanh	225.51	2181.93	1.10%	6.66%	2827.63	4.83E-03
60	M L P 1-6-1	0.95	0.99	1.00	BFG S 7	S O S	Exponential	Logistic	231.31	2206.35	0.69%	5.98%	2657.70	4.26E-03
80	M L P 1-2-1	0.96	0.99	1.00	BFG S 11	S O S	Logistic	Tanh	63.00	2068.87	0.35%	6.24%	2669.34	4.30E-03
80	M L P 1-3-1	0.96	0.99	1.00	BFG S 4	S O S	Linear	Tanh	261.75	2325.06	0.42%	7.18%	2957.41	5.28E-03
80	M L P 1-2-1	0.96	0.98	1.00	BFG S 7	S O S	Logistic	Linear	553.25	2205.41	2.23%	6.74%	2759.02	4.59E-03
80	M L P 1-2-1	0.95	0.98	1.00	BFG S 6	S O S	Tanh	Logistic	81.51	2328.80	0.41%	6.55%	2719.89	4.46E-03
80	M L P 1-7-1	0.96	0.99	1.00	BFG S 5	S O S	Linear	Tanh	159.97	2374.17	0.82%	7.38%	3018.42	5.50E-03
100	M L P 1-7-1	0.96	0.99	1.00	BFG S 7	S O S	Linear	Tanh	573.15	2175.01	0.84%	6.54%	2792.33	4.71E-03
100	M L P 1-2-1	0.95	0.99	1.00	BFG S 9	S O S	Tanh	Logistic	334.46	2310.29	1.71%	6.79%	2726.36	4.49E-03
100	M L P 1-5-1	0.96	0.99	1.00	BFG S 5	S O S	Linear	Tanh	180.96	2441.21	1.90%	7.72%	3101.83	5.81E-03
100	M L P 1-2-1	0.96	0.99	1.00	BFG S 7	S O S	Linear	Tanh	573.25	2174.78	0.84%	6.54%	2791.98	4.70E-03
100	M L P 1-4-1	0.96	0.99	1.00	BFG S 5	S O S	Linear	Tanh	100.84	2331.11	0.91%	7.25%	2967.66	5.32E-03
200	M L P 1-8-1	0.96	0.99	1.00	BFG S 6	S O S	Tanh	Tanh	380.18	2350.66	2.18%	7.47%	3034.06	5.56E-03
200	M L P 1-2-1	0.96	0.98	1.00	BFG S 7	S O S	Tanh	Linear	265.66	2300.27	1.76%	7.12%	2877.84	5.00E-03
200	M L P 1-8-1	0.96	0.99	1.00	BFG S 2	S O S	Tanh	Tanh	1932.17	2744.39	4.08%	6.86%	3486.86	7.34E-03
200	M L P 1-3-1	0.96	0.98	1.00	BFG S 7	S O S	Logistic	Tanh	38.54	1969.49	0.44%	5.51%	2441.32	3.60E-03
200	M L P 1-4-1	0.95	0.98	1.00	BFG S 5	S O S	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. A summary of neural network learning for the scenario of random sample sizes, with 70% allocated for training, 15% for testing, and 15% for validation, for the Czech Republic

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									MAE	**MAE**	**MPE**	**MAPE**	**SSE**	**Theil**
20	M L P 1-6-1	0.92	0.96	0.93	BFG S 9	SOS	Logistics	Linear	111.76	892.27	0.65%	4.11%	1127.53	2.55E-03
20	M L P 1-6-1	0.93	0.97	0.93	BFG S 16	SOS	Exponential	Linear	87.36	857.77	0.57%	3.94%	1075.19	2.32E-03
20	M L P 1-7-1	0.91	0.95	0.94	BFG S 8	S O S	Linear	Tanh	58.56	952.54	0.40%	4.42%	1210.53	2.94E-03
20	M L P 1-5-1	0.93	0.97	0.93	BFG S 5	S O S	Linear	Linear	56.86	878.44	0.41%	4.04%	1107.98	2.46E-03
20	M L P 1-6-1	0.93	0.97	0.93	BFG S 5	S O S	Exponential	Tanh	210.53	859.10	1.14%	3.97%	1092.52	2.39E-03
40	M L P 1-4-1	0.93	0.97	0.93	BFG S 5	S O S	Linear	Linear	71.95	877.99	0.48%	4.04%	1107.30	2.46E-03
40	M L P 1-6-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	87.28	949.18	0.53%	4.41%	1210.93	2.94E-03
40	M L P 1-2-1	0.93	0.97	0.93	BFG S 5	S O S	Linear	Linear	69.79	874.38	0.48%	4.02%	1104.26	2.44E-03
40	M L P 1-2-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	59.35	955.16	0.39%	4.44%	1213.80	2.95E-03
40	M L P 1-3-1	0.93	0.97	0.93	BFG S 14	S O S	Tanh	Linear	104.46	864.91	0.65%	3.98%	1091.84	2.39E-03
60	M L P 1-3-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	58.01	953.11	0.39%	4.42%	1211.13	2.94E-03
60	M L P 1-7-1	0.91	0.95	0.94	BFG S 5	S O S	Linear	Tanh	58.92	952.63	0.40%	4.42%	1210.68	2.94E-03
60	M L P 1-3-1	0.91	0.95	0.94	BFG S 3	S O S	Linear	Tanh	72.89	973.04	0.43%	4.53%	1237.19	3.07E-03
60	M L P 1-2-1	0.91	0.95	0.94	BFG S 4	S O S	Linear	Tanh	157.57	966.41	0.83%	4.50%	1239.91	3.08E-03
60	M L P 1-3-1	0.91	0.95	0.94	BFG S 8	S O S	Linear	Tanh	59.28	954.53	0.40%	4.43%	1213.02	2.95E-03
80	M L P 1-3-1	0.91	0.95	0.94	BFG S 11	S O S	Linear	Tanh	58.56	952.54	0.40%	4.42%	1210.53	2.94E-03
80	M L P 1-7-1	0.91	0.95	0.94	BFG S 5	S O S	Linear	Tanh	59.51	952.62	0.40%	4.42%	1210.76	2.94E-03
80	M L P 1-4-1	0.91	0.95	0.94	BFG S 3	S O S	Linear	Tanh	62.01	962.47	0.39%	4.48%	1223.42	3.00E-03
80	M L P 1-3-1	0.91	0.95	0.94	BFG S 7	S O S	Linear	Tanh	58.56	952.54	0.40%	4.42%	1210.53	2.94E-03
80	M L P 1-6-1	0.93	0.97	0.93	BFG S 6	S O S	Linear	Linear	30.26	879.93	0.29%	4.04%	1108.20	2.46E-03
100	M L P 1-3-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	64.09	952.92	0.42%	4.42%	1211.77	2.94E-03
100	M L P 1-7-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	61.94	955.02	0.41%	4.44%	1214.00	2.95E-03
100	M L P 1-8-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	58.04	953.11	0.39%	4.42%	1211.13	2.94E-03
100	M L P 1-4-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	62.92	954.83	0.41%	4.43%	1213.91	2.95E-03
100	M L P 1-2-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	59.45	953.21	0.40%	4.43%	1211.45	2.94E-03
200	M L P 1-4-1	0.91	0.95	0.94	BFG S 8	S O S	Linear	Tanh	58.14	953.01	0.39%	4.42%	1211.03	2.94E-03
200	M L P 1-5-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	58.24	953.65	0.39%	4.43%	1211.81	2.94E-03
200	M L P 1-5-1	0.91	0.95	0.94	BFG S 7	S O S	Linear	Tanh	62.93	960.21	0.40%	4.46%	1220.63	2.99E-03
200	M L P 1-2-1	0.91	0.95	0.94	BFG S 5	S O S	Linear	Tanh	109.92	944.09	0.64%	4.38%	1209.00	2.93E-03
200	M L P 1-7-1	0.91	0.95	0.94	BFG S 6	S O S	Linear	Tanh	58.13	953.72	0.39%	4.43%	1211.88	2.94E-03
								Minimal	**30.26**	**857.77**	**0.29%**	**3.94%**	**1075.19**	**2.32E-03**

Table 4. A summary of neural network learning for the scenario of random sample sizes, with 80% designated for training, 10% for testing, and 10% for validation, for the Czech Republic

Network number	Network name	Quality (learning)	Quality (testing)	Quality (validation)	Learning algorithm	Error function	Activation (hidden)	Activation (output)	Errors					
									MAE	MAE	MPE	MAPE	SSE	Theil
20	M L P 1-4-1	0.93	0.97	1.00	BFG S 3	S O S	Exponential	Linear	487.53	1416.32	3.00%	6.54%	1669.17	5.58E-03
20	M L P 1-7-1	0.93	0.98	1.00	BFG S 3	S O S	Exponential	Tanh	23.35	1097.18	0.77%	4.94%	1426.59	4.08E-03
20	M L P 1-8-1	0.93	0.97	1.00	BFG S 4	S O S	Tanh	Exponential	2110.65	2189.33	10.24%	10.53%	2501.58	1.25E-02
20	M L P 1-3-1	0.93	0.97	1.00	BFG S 2	S O S	Linear	Logistics	1025.64	2435.73	5.93%	11.28%	2685.78	1.45E-02
20	M L P 1-8-1	0.93	0.97	1.00	BFG S 1	S O S	Tanh	Logistics	1083.86	2537.25	6.24%	11.75%	2789.78	1.56E-02
40	M L P 1-7-1	0.92	0.97	1.00	BFG S 3	S O S	Linear	Exponential	603.14	1168.53	3.35%	5.50%	1458.17	4.26E-03
40	M L P 1-3-1	0.93	0.97	1.00	BFG S 7	S O S	Exponential	Linear	116.51	816.45	0.81%	3.75%	1067.44	2.28E-03
40	M L P 1-6-1	0.93	0.98	1.00	BFG S 7	S O S	Tanh	Exponential	108.79	819.40	0.77%	3.75%	1067.58	2.28E-03
40	M L P 1-4-1	0.92	0.98	1.00	BFG S 5	S O S	Tanh	Exponential	437.36	867.66	2.15%	3.98%	1174.93	2.77E-03
40	M L P 1-8-1	0.92	0.97	1.00	BFG S 5	S O S	Exponential	Linear	616.46	961.42	2.82%	4.41%	1323.89	3.51E-03
60	M L P 1-8-1	0.93	0.97	1.00	BFG S 3	S O S	Tanh	Exponential	1959.70	1959.71	9.34%	9.34%	2254.53	1.02E-02
60	M L P 1-5-1	0.93	0.97	1.00	BFG S 3	S O S	Exponential	Linear	154.12	947.77	1.25%	4.33%	1278.74	3.28E-03
60	M L P 1-7-1	0.92	0.97	1.00	BFG S 3	S O S	Exponential	Linear	5.33	1300.86	0.78%	5.82%	1654.10	5.48E-03
60	M L P 1-2-1	0.92	0.97	1.00	BFG S 4	S O S	Exponential	Linear	1699.99	1699.99	8.08%	8.08%	2012.01	8.11E-03
60	M L P 1-6-1	0.93	0.97	1.00	BFG S 3	S O S	Exponential	Tanh	466.30	1482.88	2.95%	6.83%	1745.75	6.11E-03
80	M L P 1-2-1	0.92	0.97	1.00	BFG S 4	S O S	Tanh	Exponential	531.83	1128.50	3.02%	5.30%	1426.35	4.08E-03
80	M L P 1-5-1	0.92	0.97	1.00	BFG S 4	S O S	Exponential	Linear	1136.36	1191.13	5.37%	5.60%	1556.41	4.85E-03
80	M L P 1-7-1	0.93	0.97	1.00	BFG S 3	S O S	Linear	Exponential	745.95	1445.15	4.13%	6.78%	1695.72	5.76E-03
80	M L P 1-7-1	0.93	0.97	1.00	BFG S 3	S O S	Exponential	Linear	1931.41	1931.63	9.20%	9.20%	2226.76	9.94E-03
80	M L P 1-6-1	0.92	0.97	1.00	BFG S 3	S O S	Linear	Exponential	383.63	1052.33	2.33%	4.90%	1362.38	3.72E-03
100	M L P 1-6-1	0.93	0.97	1.00	BFG S 4	S O S	Linear	Exponential	1901.39	1987.15	9.21%	9.53%	2275.92	1.04E-02
100	M L P 1-2-1	0.92	0.97	1.00	BFG S 5	S O S	Linear	Exponential	1648.65	1665.92	7.88%	7.94%	1976.47	7.83E-03
100	M L P 1-6-1	0.93	0.97	1.00	BFG S 4	S O S	Linear	Exponential	1757.21	1884.75	8.56%	9.04%	2162.32	9.37E-03
100	M L P 1-3-1	0.92	0.97	1.00	BFG S 4	S O S	Linear	Exponential	465.61	1121.79	2.73%	5.24%	1420.30	4.04E-03
100	M L P 1-4-1	0.93	0.97	1.00	BFG S 3	S O S	Exponential	Linear	610.44	1559.61	3.62%	7.23%	1802.27	6.51E-03
200	M L P 1-7-1	0.93	0.97	1.00	BFG S 2	S O S	Linear	Exponential	1605.53	1815.74	7.93%	8.72%	2091.05	8.76E-03
200	M L P 1-8-1	0.93	0.98	1.00	BFG S 4	S O S	Logistics	Exponential	756.42	961.57	3.51%	4.42%	1371.58	3.77E-03
200	M L P 1-7-1	0.93	0.97	1.00	BFG S 3	S O S	Linear	Exponential	1046.13	1415.38	5.36%	6.76%	1680.94	5.66E-03
200	M L P 1-5-1	0.93	0.98	1.00	BFG S 7	S O S	Tanh	Exponential	126.32	817.91	0.85%	3.75%	1072.29	2.30E-03
200	M L P 1-2-1	0.93	0.98	1.00	BFG S 4	S O S	Tanh	Exponential	1147.94	1210.56	5.45%	5.71%	1567.71	4.93E-03
								Minimal	**5.33**	**816.45**	**0.77%**	**3.75%**	**1067.44**	**2.28E-03**

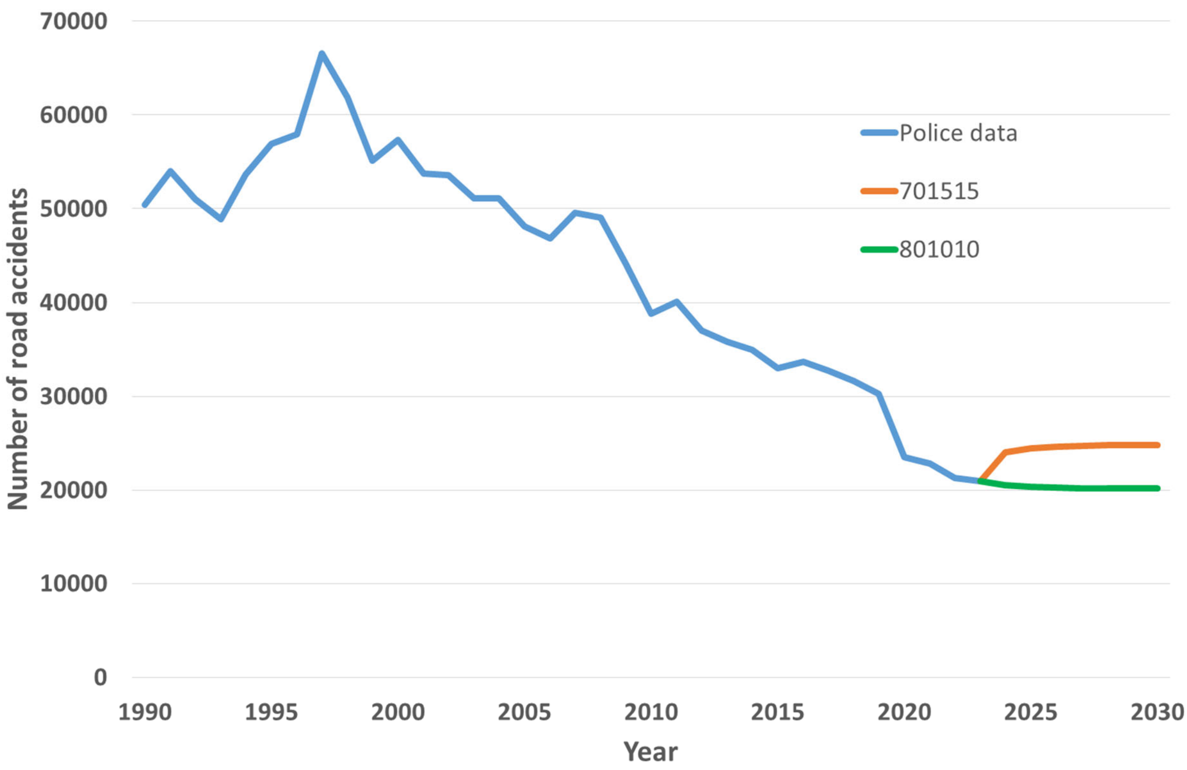

Fig. 4. Projected quantity of road accidents for 2022-2030 in Poland

Study Findings on Traffic Accidents in the Czech Republic (Fig. 5):

- Overall Trends:
  - In the next years there may be a slight increase in traffic accidents on Czech roads.
  - Ultimately, the amount of accidents is expected to stabilize in the nation.
- Influence of Sample Size:
  - The results are significantly affected by the choice of random sample size.
- An increased proportion of the training group in relation to the test and validation groups contributes to a reduction in the average percentage error.
- Error Rates:
  - Second Test (80-10-10):
    - Average Percentage Error: 3.75%
  - Learning Group (70%):
  - Test Group: 15%
  - Validation Group: 15%
  - Average Percentage Error: 3.94%
- Factors Influencing Findings (Šarkan et al. 2024):
  - The results are influenced by:
    - The impact of the recent epidemic.
    - The growing number of automobiles on Czech roads.

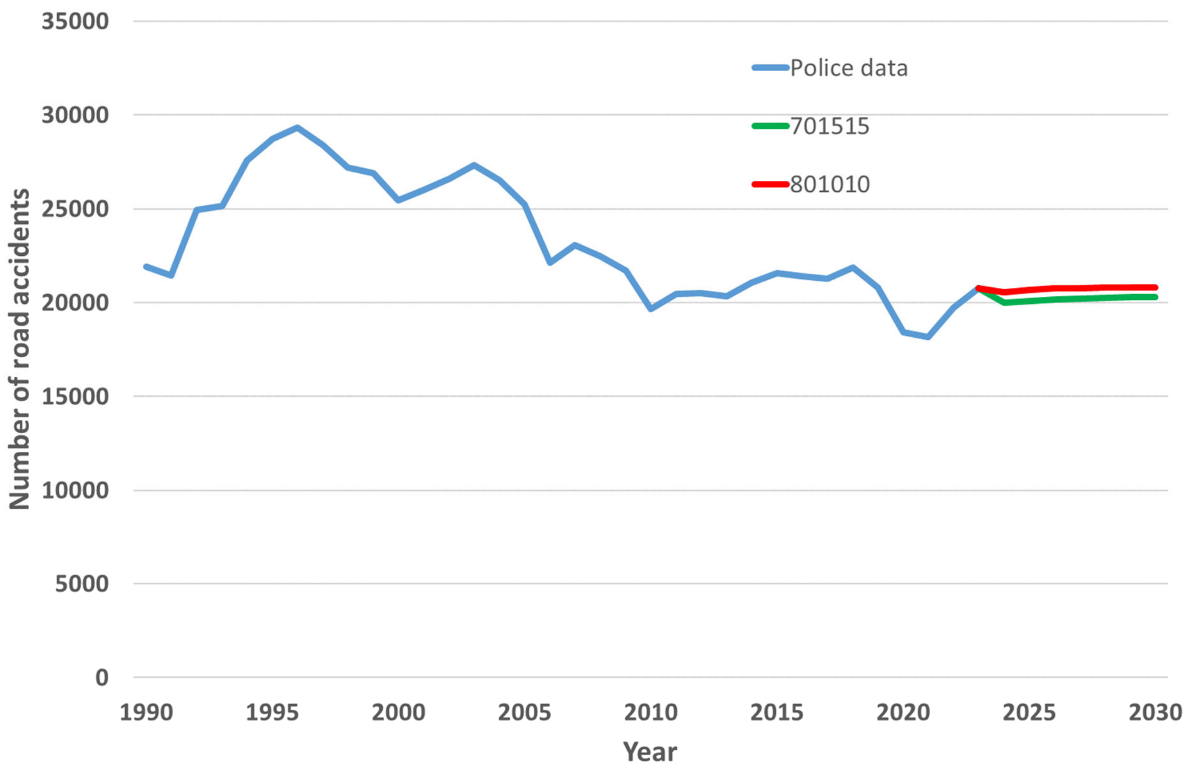

Fig. 5. Projected quantity of road accidents for 2022-2030 in the Czech Republic

4. Conclusions

Neural networks were utilized within the Statistica environment to forecast the occurrence of accidents in both Poland and the Czech Republic. The software assessed the weights used in the study to improve the accuracy of predictions, particularly regarding mean absolute error and mean absolute percentage error.

The collected data indicates that it may still be possible to anticipate a consistent trend in the number of traffic accidents, with a slight increase expected in each analyzed country. This observation can be contextualized by considering the ongoing pandemic and the rising number of vehicles on the roads. The projected forecast errors demonstrate the reliability of the models.

Given the forecasts produced, it is essential to implement measures aimed at further reducing traffic accidents. One potential strategy could be raising fines for traffic violations on Polish roads, which is set to commence on January 1, 2022. The pandemic's significant impact on the frequency of road accidents has clearly influenced the study's outcomes. For future research, the authors plan to explore additional statistical methods and consider various factors that might impact accident rates. These factors could include traffic volume, weather conditions, driver's age, and the application of exponential methods to assess the occurrence of traffic incidents.

This research was supported by the Project 05SVV2302 "Methodology proposal in the context of investigating the influence of the height profile of roads on the reduction of emissions from road transport".

References

Abdullah, E., Emam, A. (2016). *Traffic accidents analyzer using big data*. Proceedings - 2015 International Conference on Computational Science and Computational Intelligence, CSCI 2015, 392-397. https://doi.org/10.1109/CSCI.2015.187

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

Bąk, I., Cheba, K., Szczecińska, B. (2019). The statistical analysis of road traffic in cities of Poland. *Transportation Research Procedia*, *39*, 14-23. https://doi.org/10.1016/J.TRPRO.2019.06.003

Chamier-Gliszczyński, N. (2012). *Structure analysis of system mobility in urban areas*. Congress Proceedings, CLC 2012: Carpathian Logistics Congress, 509-515, 111467, Jesenik, Czech Republic, Tanger Ltd.

Chamier-Gliszczyński, N. (2012a). *Modeling system mobility in urban areas*. Congress Proceedings, CLC 2012: Carpathian Logistics Congress, 501-508, 111467, Jesenik, Czech Republic, Tanger Ltd.

Chamier-Gliszczynski, N. (2016). *City Logistics – Sustainable Urban Mobility*. CLC 2015: Carpathian Logistics Congress – Conference Proceedings, 263-268, Jesenik, Czech Republic, Tanger Ltd.

Chand, A., Jayesh, S., Bhasi, A.B. (2021). Road traffic accidents: An overview of data sources, analysis techniques and contributing factors. *Materials Today: Proceedings*, *47*, 5135-5141. https://doi.org/10.1016/J.MATPR.2021.05.415

Chen, C. (2017). Analysis and Forecast of Traffic Accident Big Data. *ITM Web of Conferences*, *12*, 04029. https://doi.org/10.1051/ITMCONF/20171204029

Chovancová, M., Stopka, O., Klapita, V. (2017). *Modeling the distribution network applying the principles of linear programming*. Proceedings – 21st International Scientific on Conference Transport Means 2017, Juodkrante; Lithuania; 20-22 September 2017, Code 135093, 73-77.

Čubranić-Dobrodolac, M., Švadlenka, L., Čičević, S., Dobrodolac, M. (2020). Modelling driver propensity for traffic accidents: a comparison of multiple regression analysis and fuzzy approach. *International Journal of Injury Control and Safety Promotion*, *27*(2), 156-167. https://doi.org/10.1080/17457300.2019.1690002

Cubranic-Dobrodolac, M., Svadlenka, L., Cicevic, S., Trifunovic, A., Dobrodolac, M. (2020). Using the Interval Type-2 Fuzzy Inference Systems to Compare the Impact of Speed and Space Perception on the Occurrence of Road Traffic Accidents. *Mathematics*, *8*(9). https://doi.org/10.3390/math8091548

Čubranić-Dobrodolac, M., Švadlenka, L., Čičević, S., Trifunović, A., Dobrodolac, M. (2022). A bee colony optimization (BCO) and type-2 fuzzy approach to measuring the impact of speed perception on motor vehicle crash involvement. *Soft Computing*, *26*(9), 4463-4486. https://doi.org/10.1007/s00500-021-06516-4

Czech statistics office. (2024). Road Traffic accidents in Czech Republic – time series. https://csu.gov.cz/produkty/nehody_v_doprave_casove_rady

Dyczkowska, J., Olkiewicz, M., Chamier-Gliczynski, N., Królikowski, T. (2023). Mobility-as-a-Service (MaaS) as a solution platform for the city and the region: case study. *Procedia Computer Science*, *225*, 4092-4100, 196245. https://doi.org/10.1016/j.procs.2023.10.405

Dyczkowska, J., Chamier-Gliszczynski, N., Olkiewicz, M., Królikowski, T. (2023a). Decision support in the area of Logistics 4.0. *Procedia Computer Science, 225*, 4758-4765, 196245. https://doi.org/10.1016/j.procs.2023.10.475

Gorzelanczyk, P., Huk, A. (2022). Road traffic safety: A case study of the Pila poviat in Poland. *Scientific Journal of Silesian University of Technology. Series Transport*, *114*, 31-42. https://doi.org/10.20858/SJSUTST.2022.114.3

Hudec, J., Cződörová, R. (2022). Analysis of the system of driving schools in selected EU countries and the Slovak Republic. *Perner's Contacts*, *17*(1). https://doi.org/10.46585/pc.2022.1.2115

Hudec, J., Sarkan, B., Caban, J., Stopka, O. (2021). The Impact of Driving Schools' Training on Fatal Traffic Accidents in the Slovak Republic. *Scientific Journal of Silesian University of Technology-Series Transport*, *110*, 45-57. https://doi.org/10.20858/sjsutst.2021.110.4

Khaliq, K.A., Chughtai, O., Shahwani, A., Qayyum, A., Pannek, J. (2019). Road Accidents Detection, Data Collection and Data Analysis Using V2X Communication and Edge/Cloud Computing. *Electronics*, *8*(8), 896. https://doi.org/10.3390/electronics8080896

Kielc, R., Sąsiadek, M., Woźniak, W. (2018). *Adoption of the evolutionary algorithm to automate the scheduling of the production processes*. Proceedings of 31st International Business Information Management Association Conference, IBIMA 2018, International Business Information Management Association Conference, Milan 2018, 5039-5046, 143853.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B., Gershman, S.J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253. https://doi.org/10.1017/S0140525X16001837

Marr, B. (2019). The Amazing Ways eBay Is Using Artificial Intelligence To Boost Business Success. Forbes: Enterprise Tech. https://www.forbes.com/sites/bernardmarr/2019/04/26/the-amazing-ways-ebay-is-using-artificial-intelligence-to-boost-business-success/

Oronowicz-Jaśkowiak, W. (2019). The application of neural networks in the work of forensic experts in child abuse cases. *Advances in Psychiatry and Neurology/Postępy Psychiatrii i Neurologii*, *28*(4), 273-282. https://doi.org/10.5114/PPN.2019.92489

Polish Police. (2024). Statistic Road Accident. https://statystyka.policja.pl/

Rajput, H., Som, T., Kar, S. (2015). An automated vehicle license plate recognition system. *Computer*, *48*(8), 56-61. https://doi.org/10.1109/MC.2015.244

Šarkan B., Pal'o J., Loman M., Stopka O., Caban J., Čeháková K., Gołębiowski W., Pečman J. (2024). Research on the Quantification of Exhaust Emission Volumes in an Opted Road Section. *Acta Polytechnica Hungarica*, *21*(7), 9-30. https://doi.org/10.12700/APH.21.7.2024.7.2

Stopka, O. (2022). Modelling Distribution Routes in City Logistics by Applying Operations Research Methods. *Promet – Traffic & Transportation*, *34*(5), 739-754. https://doi.org/10.7307/ptt.v34i5.4103

Tambouratzis, T., Souliou, D., Chalikias, M., Gregoriades, A. (2014). Maximising Accuracy and Efficiency of Traffic Accident Prediction Combining Information Mining with Computational Intelligence Approaches and Decision Trees. *Journal of Artificial Intelligence and Soft Computing Research*, *4*(1), 31-42. https://doi.org/10.2478/JAISCR-2014-0023

Vilaça, M., Silva, N., Coelho, M.C. (2017). Statistical Analysis of the Occurrence and Severity of Crashes Involving Vulnerable Road Users. *Transportation Research Procedia*, *27*, 1113-1120. https://doi.org/10.1016/J.TRPRO.2017.12.113

Witt, A. (2023). Determination of the Number of Required Charging Stations on a German Motorway Based on Real Traffic Data and Discrete Event-Based Simulation. *LOGI – Scientific Journal on Transport and Logistics*, *14*(1), 1-11. https://doi.org/10.2478/logi-2023-0001

Wójcik, A. (2014). *Autoregressive vector models as a response to the critique of multi-equation structural econometric models*. Publishing House of the University of Economics in Katowice, 193.

World Health Organization. (2018). Global status report on road safety 2018 Summary. In Global report (Issue 1). http://apps.who.int/bookorders

Wozniak, W., Walkowiak, J., Sasiadek, M., Stryjski, R. (2018). *Organisation of the Research Process into an Innovative, Anti-Clogging Assembly for Heavy Vehicles in the Interests of Increased Road Safety*. 32nd Conference of the International-Business-Information-Management-Association (IBIMA), Seville, Spain 2018, Vision 2020: Sustainable Economic Development and Application of Innovation Management, 4772-4784.

Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., … Dean, J. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. 1–23. http://arxiv.org/abs/1609.08144

Yadav, S., Rishi, R. (2022). Algorithm for Creating Optimized Green Corridor for Emergency Vehicles with Minimum Possible Disturbance in Traffic. *LOGI – Scientific Journal on Transport and Logistics*, *13*(1), 84-95. https://doi.org/10.2478/logi-2022-0008

Yang, Z., Zhang, W., Feng, J. (2022). Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework. *Safety Science*, *146*, 105522. https://doi.org/10.1016/J.SSCI.2021.105522

Yu, A. (2019). How Netflix Uses AI, Data Science, and Machine Learning – From A Product Perspective. Becoming Human Exploring Artificial Intelligence & What It Means to Be Human. https://becominghuman.ai/how-netflix-uses-ai-and-machine-learning-a087614630fe

Zheng, Z., Wang, C., Wang, P., Xiong, Y., Zhang, F., Lv, Y. (2018). Framework for fusing traffic information from social and physical transportation data. *PLOS ONE*, *13*(8), e0201531. https://doi.org/10.1371/JOURNAL.PONE.0201531

Zhu, L., Lu, L., Zhang, W., Zhao, Y., Song, M. (2019). Analysis of Accident Severity for Curved Roadways Based on Bayesian Networks. *Sustainability*, *11*(8), 2223. https://doi.org/10.3390/SU11082223