Acta Tecnología - International Scientific Journal about Technologies
Volume: 10 2024 Issue: 2 Pages: 51-56 ISSN 2453-675X

https://doi.org/10.22306/atec.v10i2.199 *Received: 05 Apr. 2024; Revised: 15 May 2024; Accepted: 27 June 2024*

Using neural networks to forecast the frequency of accidents on particular Polish road types

Piotr Gorzelanczyk
Stanislaw Staszic University of Applied Sciences in Pila, Podchorazych 10 Street, 64-920 Pila, Poland, EU, piotr.gorzelanczyk@ans.pila.pl ORCID 0000-0001-9662-400X (corresponding author)
Miriam Garbarova
Faculty of Operation and Economics of Transport and Communications, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, EU, miriam.garbarova@uniza.sk

Keywords: forecasting, road accident, road types, neural networks.
Abstract: Globally, the quantity of road accidents is steadily decreasing annually. Even if the epidemic has recently had an effect, this number is still quite high. Therefore, every attempt should be made to reduce this number. The purpose of this article is to forecast the number of accidents that will transpire on various kinds of roads in Poland. For this purpose, monthly data on the total number of traffic accidents in Poland broken down by type of road were examined. Based on police data, a forecast was created for the years 2022–2040. A few neural network models were used to predict the frequency of accidents in Poland. The results show that we should keep expecting a high degree of stability in terms of the quantity of traffic accidents. On the one hand, the construction of new highways and the increase in the number of cars on Polish roads have an effect on this. The number of random samples utilized in the processes of learning, testing, and validation affects the results. In the case of 70-15-15, the average projected number of accidents for 2022-2024 should be: motorway - 827, expressway - 1136, 2 one-way carriageway - 5451, single carriageway -3067 and 1 carriageway 2 directions - 28048, respectively. For 80-10-10, these values should be respectively: motorway - 1057, expressway - 1141, 2 one-way carriageway - 5486, single carriageway -2745 and 1 carriageway 2 directions - 28039. As you can see, the values are similar.

1 Introduction

A significant number of people lose their lives in traffic accidents each year. The WHO estimates that over 1.3 million people die in car crashes annually. For kids and young people between the ages of 5 and 29, traffic accidents are the leading cause of death [1]. The goal of reducing traffic-related fatalities and injuries by half by 2030 has been set by the UN General Assembly.

Many sources provide information regarding road accidents. One example is information that government organizations have obtained with the help of pertinent government entities. Data is gathered from police reports, insurance databases, and hospital records. The transportation sector's partial traffic accident data is then examined more broadly [2].

The literature has a variety of methods for accident frequency forecasting. The most often used techniques are time series methods [3,4]. However, their drawbacks include frequent residual component autocorrelation and the inability to evaluate prediction quality based on prior forecasts [5]. In contrast, Sunny et al. [6] utilized the Holt-Winters exponential smoothing method, while Procházka et al. [7] used the multiple seasonality model for forecasting. This method prevents the inclusion of exogenous variables in the model [8,9].

We can also use autoregressive models [10], regression models with curve fitting, and vector autoregressive models, which have the drawback of requiring a large number of variable observations in order to correctly estimate their parameters [11]. This is not always possible. Consequently, these just call for autoregressive order (assuming the series is already stationary) and simple linear relationships [12,13].

In their research, Chudy-Laskowska and Pisula [14] used the ANOVA approach to forecast the problem at hand. This strategy's drawback is that it necessitates the employment of additional assumptions, which when broken can lead to inaccurate findings [15]. Neural network models are also used to estimate the amount of traffic accidents. This strategy's drawback is that it necessitates a fundamental comprehension of the subject [16]. Furthermore, because SNF is frequently referred to as a "black box," in which input is provided and the model provides findings without knowledge of the analysis, the forecast result is dependent on the adoption of the network's beginning conditions as well as the inability of traditional interpretation [17].

Using the previously provided statistics, the author made estimates for the quantity of accidents that would occur on Polish roads. Neural networks were used to forecast how many accidents would occur.

2 Materials and methods

Every year, a number of incidents occur on Polish roads. Recent years have seen a decrease in road accidents due to the epidemic, which affects the estimated value that was established. Despite the epidemic, there are still a lot of road accidents. As a result, every attempt should be made to reduce this figure and determine the types of routes where most accidents will take place (Figure 1). Most traffic accidents occur on highways having a single carriageway that runs in both directions.

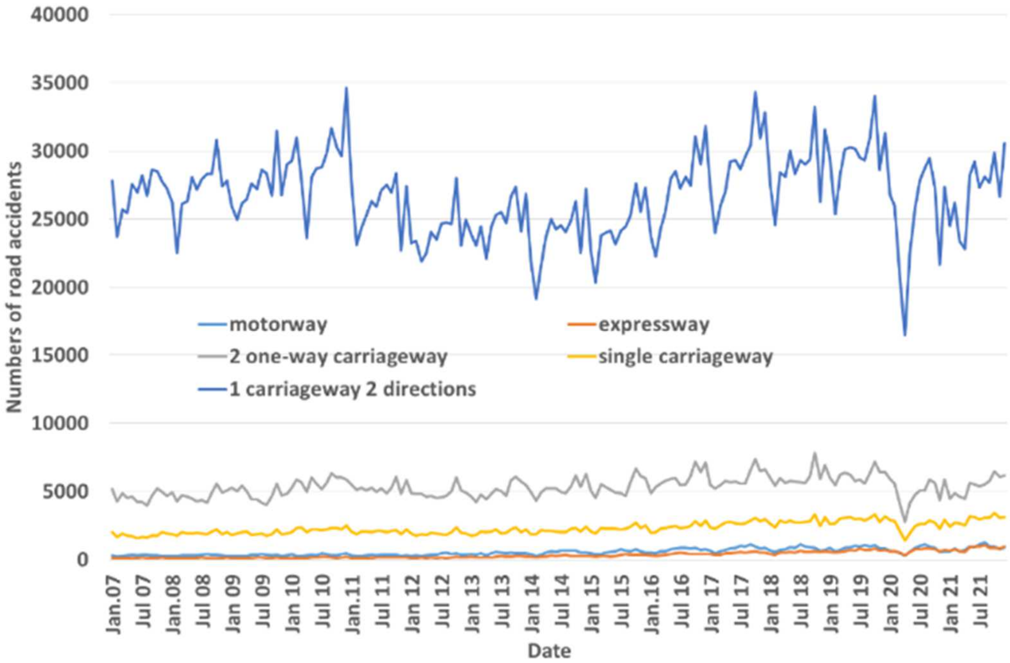

Figure 1 Number of road accidents in Poland by road type in 2007-2021 [18]

Different neural network models were used to predict the number of accidents that will occur on Polish roads based on the kind of road. The advantage of this method is that it mimics the way the human brain works. Neural networks are composed of nodes, each of which has inputs, weights, variances, and outputs. The optimal weights for the study were selected by the Statistica software. The result of predicting using the previously discussed method will depend on the model and its parameters that are chosen.

Using the forecasting mistakes obtained from equations (1-5), the following metrics of analytical forecasting brilliance were calculated:

- ME – mean error

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

- MAE – mean error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_p| \tag{2}$$

- MPE – mean percentage error

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i - Y_p}{Y_i} \tag{3}$$

- MAPE - mean absolute percentage error

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

- SSE – mean square error

$$SSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_p)^2} \tag{5}$$

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

Neural network models with the least mean absolute percentage error and mean percentage error were used to forecast the frequency of traffic accidents by type of road.

3 Results

To look into the temporal evolution of the accident rate on different kinds of roads, the Kruskall-Wallis test was used. 830 is the test statistic, while 0.000 is the test probability. The result shows that we do not take the mean number of traffic accidents to be equal (Figure 2).

Acta Tecnología – International Scientific Journal about Technologies
Volume: 10 2024 Issue: 2 Pages: 51-56 ISSN 2453-675X

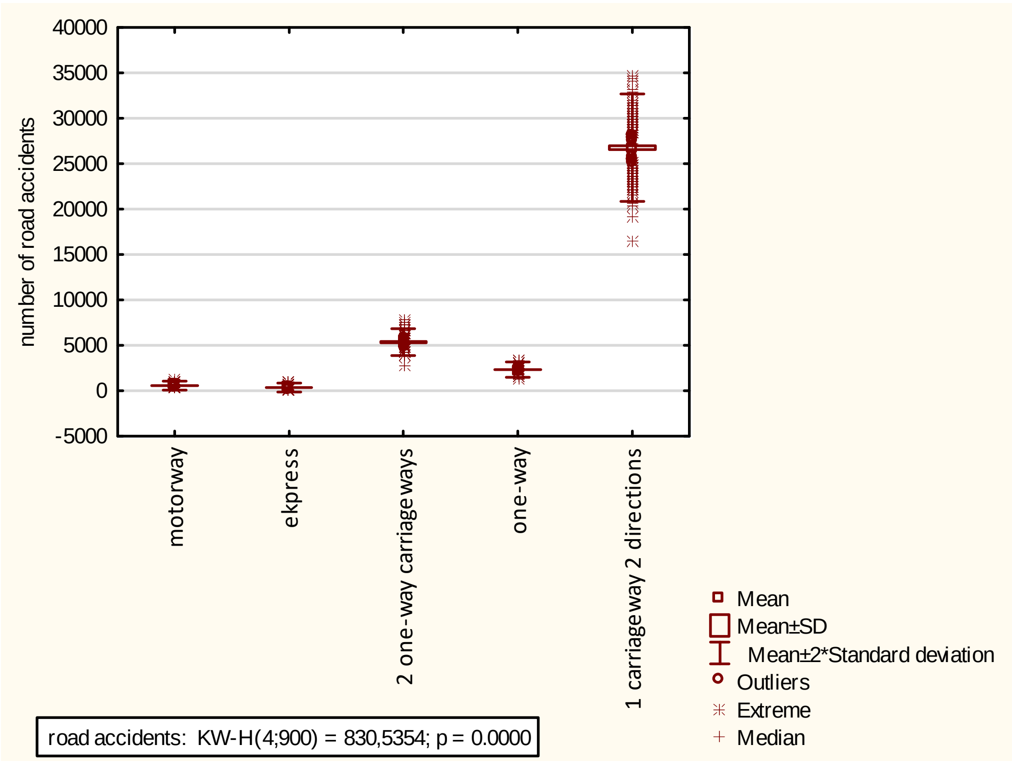

Figure 2 Comparison of the average number of road accidents in Poland by road type from 2007 to 2021 [18]

Using data from the Polish Police from 2007 to 2021, the annual number of traffic accidents in Poland was estimated based on the kind of route [18]. Two random sample sizes were assumed when the study was carried out using Statistica software:

- 70% is education, 15% testing, and 15% validation.
- 80% of instruction, 10% testing, and 10% validation.

With 20, 40, 60, 80, 100, 200 learning networks, for which the MP error value was negligible (Table 1, Table 2).

Table 1 Neural network learning summary for the random sample size case teaching 70%, testing 15% and validation 15%

Type of road	Network number	Network name	Quality (learning)	Quality (learning)	Quality (validation)	Learning algorithm	Activation (hidden)	Activation (output)	Errors				
									ME	MAE	MPE	MAPE	SSE
motorway	200	MLP 12-5-1	0.925071	0.937989	0.971657	BFGS 10	Logistics	Linear	9.5659 7754	63.361 67	0.46 %	11.83 %	88.491 27
expressway	20	MLP 12-7-1	0.951552	0.974197	0.975260	BFGS 7	Linear	Linear	8.0165 5915	44.850 12	0.33 %	13.65 %	67.759 09
2 single carriageways	20	MLP 12-5-1	0.689403	0.719414	0.873958	BFGS 9	Tanh	Logistics	53.456 7291	378.25 78	0.00 %	7.21%	521.13 92
one-way	40	MLP 12-4-1	0.883924	0.898429	0.938170	BFGS 61	Tanh	Logistics	13.570 2339	133.39 45	0.05 %	5.81%	187.79 56
1 carriageway 2 directions	60	MLP 12-8-1	0.869698	0.808519	0.700377	BFGS 61	Logistics	Linear	85.878 5093	1249.3 93	0.02 %	4.80%	1679.7 73

Table 2 Neural network learning summary for the random sample size case teaching 80%, testing 10% and validation 10%

Type of road	Network number	Network name	Quality (learning)	Quality (learning)	Quality (validation)	Learning algorithm	Activation (hidden)	Activation (output)	Errors				
									ME	MAE	MPE	MAPE	SSE
motorway	40	MLP 12-7-1	0.898808	0.914107	0.950540	BFGS 4	Linear	Logistics	18.412 4803	79.69 277	0.05 %	14.77%	106.19 22
expressway	20	MLP 12-3-1	0.961444	0.977631	0.980714	BFGS 6	Linear	Linear	8.1032 5895	40.83 375	0.55 %	12.39%	61.321 51
2 single carriageways	20	MLP 12-5-1	0.681920	0.831358	0.873351	BFGS 6	Linear	Tanh	48.140 1442	379.1 432	0.06 %	7.26%	521.45 26
one-way	200	MLP 12-2-1	0.835982	0.897177	0.956450	BFGS 13	Logistics	Tanh	21.371 6664	159.9 512	0.03 %	6.95%	217.55 64
1 carriageway 2 directions	200	MLP 12-7-1	0.841886	0.797305	0.738706	BFGS 31	Tanh	Exponential	117.84 3275	1345. 519	0.06 %	5.12%	1713.8 32

Regardless of the type of road, Poland's annual rate of traffic accidents is likely to stabilize at a high level based on the research findings. The results are impacted by the random sample size selection. By increasing the proportion of the learning group compared to the test and validation group, the average percentage error is decreased.

During a learning group of 70%, test group of 15%, and validation group of 15% (70-15-15), the error was 13.65%; however, during the second test (80-10-10) the error was 14.7%. For highways, which are showing up on Poland's map more frequently every year. It is important to note that the outbreak affected the results (Figure 3, Figure 4).

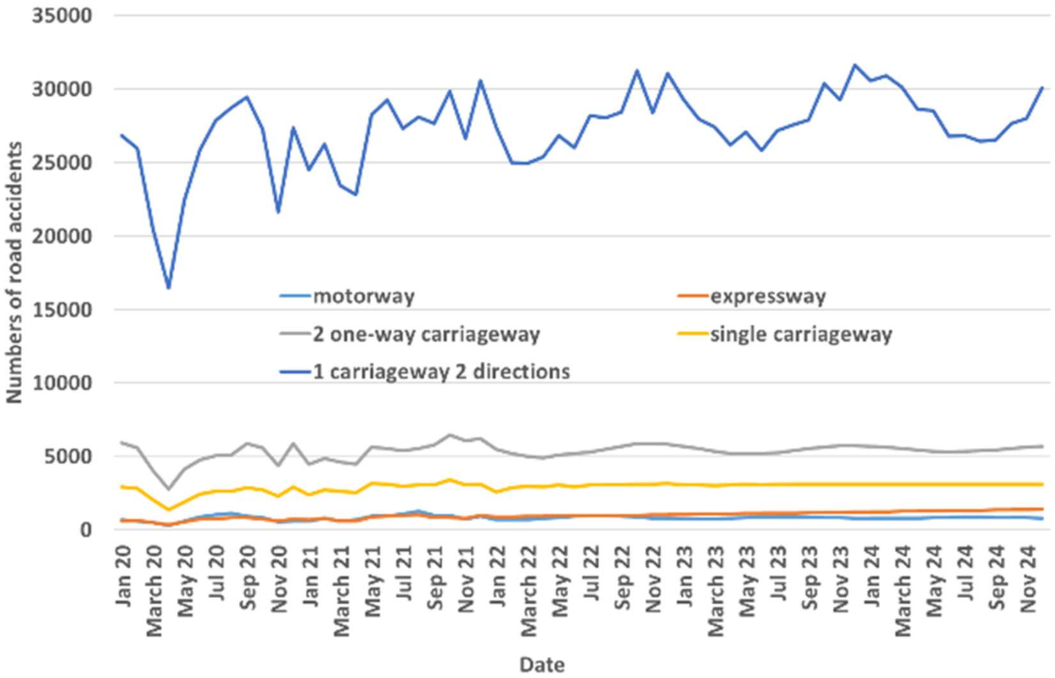

Figure 3 Forecasting number of road accidents for 2022-2024 for the 70-15-15 test group

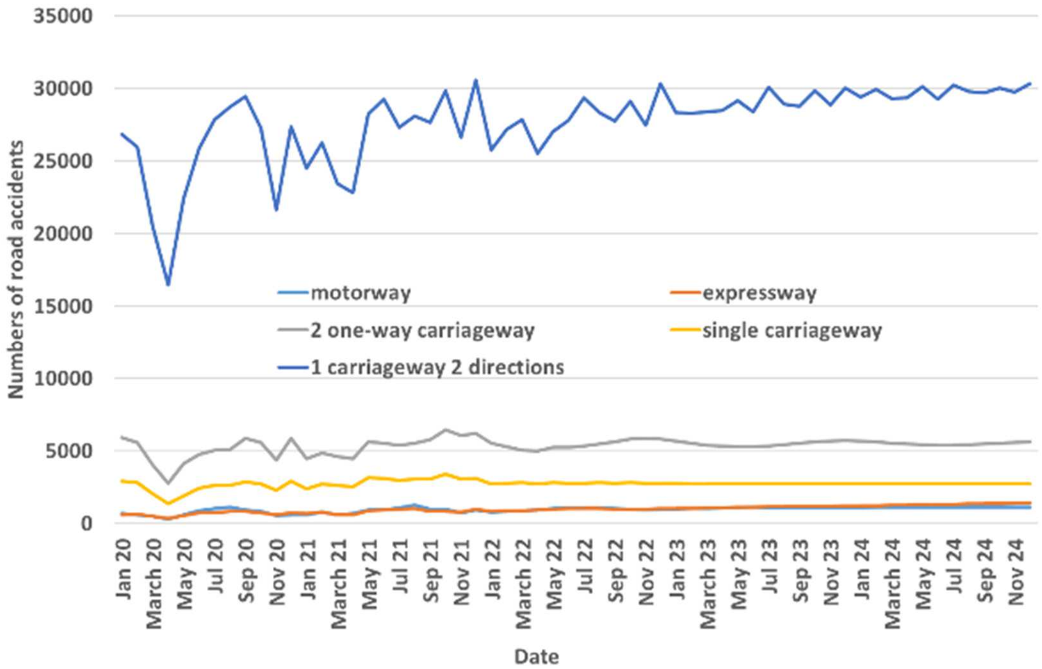

Figure 4 Forecasting number of road accidents for 2022-2024 for the 80-10-10 test group

4 Conclusion

The research utilized neural networks in the Statistica environment to predict the frequency of accidents in Poland based on the kind of road. The research's weights were assessed by the computer in order to reduce both the mean absolute error and the mean absolute percentage error.

The statistics makes it evident that there is still a chance for the number of traffic accidents to stabilize. This is affected, on the one hand, by the growth of motorways and the increase in the quantity of automobiles on Polish roads. The estimated forecasting errors show that the models that were utilized are valid.

Based on the derived projections, steps should be taken to further minimize the number of traffic accidents. Among these modifications might be the imposition of higher fines for moving offenses starting on January 1, 2022, on Polish roads. The epidemic surely affected the study's conclusions because it drastically decreased the number of accidents on the roads.

In the case of 70-15-15, the average projected number of accidents for 2022-2024 should be: motorway - 827, expressway - 1136, 2 one-way carriageway - 5451, single carriageway - 3067 and 1 carriageway 2 directions - 28048, respectively. For 80-10-10, these values should be respectively: motorway - 1057, expressway – 1141 , 2 one-way carriageway - 5486, single carriageway -2745 and 1 carriageway 2 directions - 28039. As you can see, the values are similar.

The authors want to include other variables that influence accident frequency and use alternative statistical methods to determine the total number of collisions in their follow-up study. These could include factors like the volume of traffic, the kind of weather, the age of the driver, and methods that compute accident frequency using exponential growth.

References

[1] WHO, The Global status on road safety, [Online], Available: https://www.who.int/publications/i/item/9789241565684 [02 Apr 2024], 2018.

[2] GORZELANCZYK, P., PYSZEWSKA, D., KALINA, T., JURKOVIC, M.: Analysis of road traffic safety in the Pila poviat, *Scientific Journal of Silesian University of Technology*, Series Transport, Vol. 107, pp. 33-52, 2020. https://doi.org/10.20858/sjsutst.2020.107.3

[3] HELGASON, A.F.: Fractional integration methods and short Time series: evidence from asimulation study, Political Analysis, Vol. 24, No. 1, pp. 59-68, 2016. https://doi.org/10.1093/pan/mpv026

[4] LAVRENZ, S.M., VLAHOGIANNI, E.I., GKRITZA, K., KE, Y.: Time series modeling in traffic safety research, *Accident Analysis & Prevention*, Vol. 117, pp. 368-380, 2018. https://doi.org/10.1016/j.aap.2017.11.030

[5] Forecasting based on time series, [Online], Available: http://pis.rezolwenta.eu.org/Materialy/PiS-W-5.pdf [02 Apr 2024], 2022. (Original in Polish)

[6] SUNNY, C.M., NITHYA, S., SINSHI, K.S., VINODINI, V.M.D., LAKSHMI, A.K.G., ANJANA, S., MANOJKUMAR, T.K.: *Forecasting of Road Accident in Kerala: A Case Study*, 2018 International Conference on Data Science and Engineering (ICDSE), Kochi, India, pp. 1-5, 2018. https://doi.org/10.1109/ICDSE.2018.8527825

[7] PROCHÁZKA, J., FLIMMEL, S., ČAMAJ, M., BAŠTA, M.: *Modelling the Number of Road Accidents*, 20th AMSE, Applications of Mathematics and Statistics in Economics, International Scientific Conference: Szklarska Poręba, 30 August - 3 September 2017, Conference Proceedings Full Text Papers, Publishing house of the University of Economics in Wrocław, Wrocław, pp. 355-364, 2017. https://doi.org/10.15611/amse.2017.20.29

[8] DUDEK, G.: *Forecasting Time Series with Multiple Seasonal Cycles Using Neural Networks with Local Learning*, In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds) Artificial Intelligence and Soft Computing, ICAISC 2013, Lecture Notes in Computer Science, Vol. 7894, Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-38658-9_5

[9] SZMUKSTA-ZAWADZKA, M., ZAWADZKI, J.: Forecasting on the basis of Holt-Winters models for complete and incomplete data, *Ekonometria*, Vol. 24, No. 38, pp. 85-99, 2009. (Original in Polish)

[10] MONEDEROA, B.D., GIL-ALANAA, L.A., MARTÍNEZAA, M.C.V.: Road accidents in Spain: Are they persistent?, *IATSS Research*, Vol. 45, No. 3, pp. 317-325, 2021. https://doi.org/10.1016/j.iatssr.2021.01.002

[11] WÓJCIK, A.: *Autoregressive vector models as a response to the critique of multi-equation structural econometric models*, Publishing house of the University of Economics in Katowice, Katowice, 2014.

[12] MAMCZUR, M.: Machine learning How does linear regression work? And is it worth using?, [Online], Available: https://miroslawmamczur.pl/jak-dziala-regresja-Linear-i-czy-warto-ja-stosowac [02 Apr 2024], 2024.

[13] PIŁATOWSKA, M.: The choice of the order of autoregression depending on the parameters of the generating model, *Econometrics*, Vol. 4, No 38, 2012.

[14] CHUDY-LASKOWSKA, K., PISULA, T.: Forecasting the number of road accidents in Subcarpathia, *Logistics*, Vol. 4, pp. 2782-2796, 2015. (Original in Polish)

[15] GREGORCZYK, A., SWARCEWIC, M.: Analysis of variance in a system of repeated measurements to determine the effects of factors influencing linuron

residues in soil, *Polish Journal of Agronomy*, Vol. 11, pp. 15-20, 2012. (Original in Polish)

[16] CHUDY-LASKOWSKA, K., PISULA, T.: Forecast of the number of road accidents in Poland, *Logistics*, Vol. 6, pp. 2710-2721, 2014. (Original in Polish)

[17] Data mining techniques StatSoft, [Online], Available: https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fstdatmin.html [02 Apr 2024], 2024.

[18] Statistic Road Accident, [Online], Available: https://statystyka.policja.pl [02 Apr 2024], 2023. (Original in Polish)

Review process

Single-blind peer review process.