PREGLEDNI RAD – REVIEW SCIENTIFIC ARTICLE 10.7251/ZRPIM2101531B

TRAFFIC ACCIDENT ANALYSIS USING DATA MINING TECHNIQUES

Marija Blagojević, Željko Jovanović, Danijela Milošević, Snežana Dragićević

University of Kragujevac, Faculty of Technical Sciences in Čačak, Svetog Save 65, 32 000 Čačak, Serbia, marija.blagojevic@ftn.kg.ac.rs

ABSTRACT

The paper presents an example of the application of artificial neural networks for predicting the type of traffic accidents. Data for the period January-December 2019 in Čačak were analyzed. Data was collected from the open data portal https://data.gov.rs. Microsoft Visual Studio and Microsoft SQL Server Management Studio were used as tools. The model consists of an input layer (with parameters date, time and type), a hidden layer and an output layer with one neuron for predicting the type of accident. In order to evaluate the obtained model, 10-fold cross validation was used. Web-based application was created to display the results. The advantage is reflected in the Web-based application that allows users to use an artificial neural network while the limitations of the study require a larger number of measurements for more accurate results.

Keywords: traffic accidents, artificial neural network, web-based application.

INTRODUCTION

Traffic is a basic factor for the development of society. Having in mind the geographical position of Serbia at the crossroads of the Balkans, the traffic gained additional importance. The use of machine learning is very important in all areas of human creativity, especially in traffic. At the time when traffic accidents are present, measures are being taken to reduce them. To this end, machine learning techniques are used to classify or predict traffic accidents.

With the opening of data, portals are created from which data can be freely downloaded and further analyzed. The data for traffic accidents used for this research are also available on the portal. The aim of the research is to determine the possibility of using artificial neural networks to predict the type of accident.

There is a large amount of related research dealing with the application of artificial neural networks in traffic. Faghri, & Hua, (1993) is described an artificial neural network which has been applied "to dynamic traffic pattern classification to determine appropriate time intervals and the starting times for those intervals". Faghri, & Hua, (1993) concluded that the proposed approach has many advantages like feasibility and efficiently classification process. Nedic, Despotovic, Cvetanovic, Despotovic, & Babic, (2014) proposed a model of artificial neural networks for traffic noise detection. In addition, this paper presents a comparison of the results obtained via the network with statistical methods. The results are more accurate in favor of artificial neural networks. Cetiner, Sari, & Borat, (2010) used artificial neural networks to predict traffic-flow in Istanbul in order to obtain results that would be used in planning the further development of roads in the city. Abdelwahab, & Abdel-Aty, (2001) used the same technique, but in order to predict driver injury severity in traffic accidents at signalized intersections.

In comparison with related research, it can be noticed that artificial neural networks are used in the field of traffic and successfully solve research questions, but they have been used for different purposes in relation to the proposed research.

The goal of research is determination of possibility of using data mining techniques in predicting traffic accidents.

METHODOLOGY

Within the methodology, the usual procedures for each data mining process were followed. First, data was collected from the open data portal https://data.gov.rs. Like it was presented in Blagojevic, Blagojevic, & Licina, (2016) the usual steps in data mining process include:

- Data collection
- Data selection
- Data pre-processing
- Data transformation
- Artificial model development
- Model evaluation and testing
- Model interpretation

Data on traffic accidents were selected from open data portal. The portal provides data from many fields. The next step involves data selection. During the selection, the city of Čačak was chosen. Empty rows were removed in the preprocessing process. The transformation involved changing the date format to be analyzed in an appropriate form.
Microsoft Visual Studio and Microsoft SQL Server Management Studio were used as tools.

RESULTS AND DISCUSSION

To analyze the data obtained on traffic accidents, the model of artificial neural networks is developed. The model consists of an input layer (with parameters date, time and type), a hidden layer and an output layer with one neuron for predicting the type of accident.
Figure 1 shows the Microsoft Visual Studio environment in which visually selecting parameter values is used to "favor" a certain type of traffic accident.

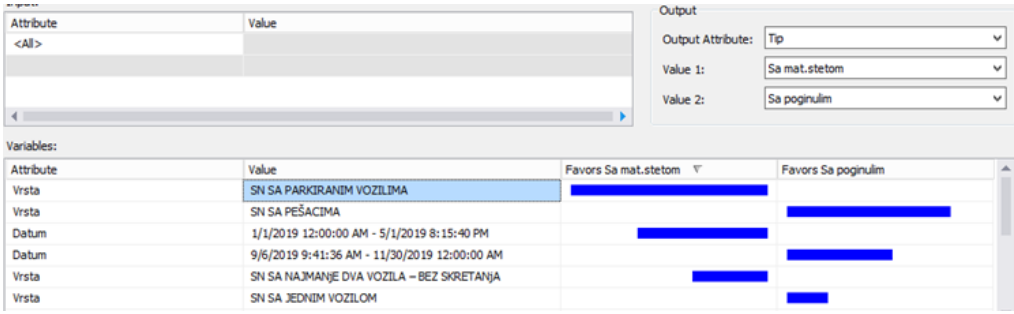

Figure 1. Parameter selection and accident type favoritism.

In order to evaluate the obtained model, 10-fold cross validation was used.

Testing of the artificial neural network was performed via DMX (Data Mining Extensions) query (Blokduk, 2017). The following is an example of such a query where the date is scheduled to enter the date from September 2020, 22 hours in time and a traffic accident with pedestrians as a type of accident. After executing the query, the type of accident was obtained, which is "with an injured person", which is shown in Figure 2.

```
SELECT [AP].tip,
PREDICT ([tip])
from [AP]
NATURAL PREDICTION JOIN
(SELECT
'2020-09-09' AS [Datum],
22 AS [Vreme],
'SN SA PEŠACIMA' AS [Vrsta]) as t
```

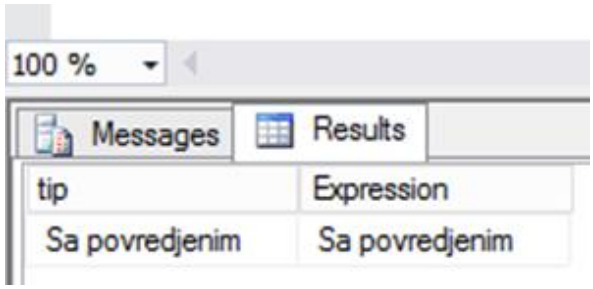

Figure 2. Query prediction result.

Given that a relatively small number of people have the skills to write DMX queries and that the prediction results are significant for a wider audience, a Web-based application was created to display the results.

The architecture of proposed system is shown in Figure 3.

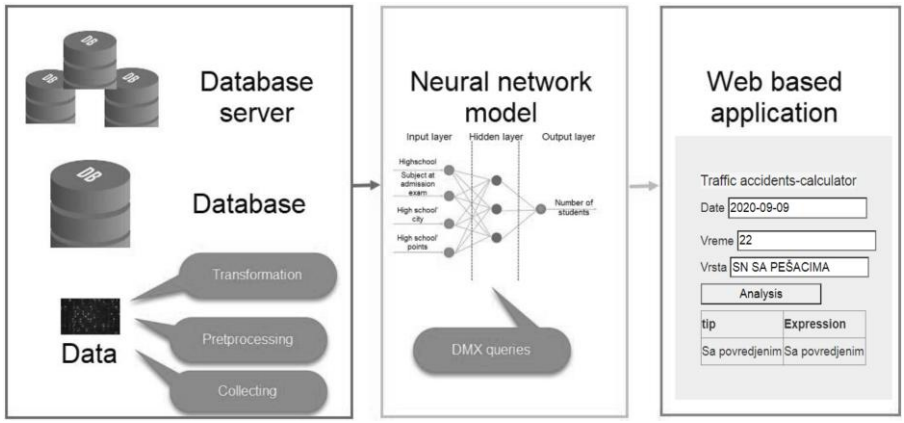

Figure 3. Architecture of proposed system (Blagojevic, & Micic, 2020).

Instead of predefined values, the user can enter the parameter for which he wants to obtain a predictive value for the type of traffic accident.

Figure 4 shows the graphical user interface of the application with one use case.

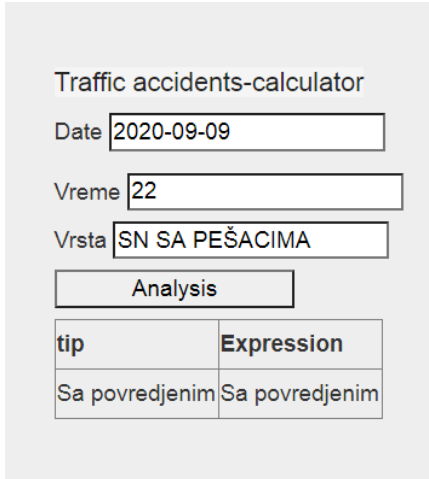

Figure 4. Graphical user interface.

In comparison with related research, the advantages and limitations of the proposed approach are noticed. The advantage is reflected in the Web-based application that allows users to use an artificial neural network while the limitations of the study require a larger number of measurements for more accurate results.

CONCLUSIONS

Having in mind the presented results, it can be concluded in several directions:

Artificial neural networks can be successfully applied in the field of transport in order to predict the traffic accidents with chosen input parameters.

A user interface can be created and people with basic IT skills can use neural networks and get results for their chose

Future work refers to the application of other data mining techniques in traffic.

ACKNOWLEDGMENT

This study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and these results are parts of the Grant No. 451-03-68/2020-14/200132 with University of Kragujevac – Faculty of Technical Sciences Čačak.

LITERATURE

Blagojević, M., Blagojević, M., & Ličina, V. (2016). Web-based intelligent system for predicting apricot yields using artificial neural networks. *Scientia Horticulturae*, *213*, 125-131.

Blagojević, M., & Mićić, Ž. (2020). The use of artificial neural networks in supporting education and curriculum planning. *International journal of engineering education*, *36*(6), 1752-1759.

Blokduk, G. (2018). Data Mining Extensions. *Brisbane, Australia: Emereo Pty. Ltd*.

Çetiner, B. G., Sari, M., & Borat, O. (2010). A neural network based traffic-flow prediction model. *Mathematical and Computational Applications*, *15*(2), 269-278. Retrieved March 19, 2020, from: https://www.mdpi.com/2297-8747/15/2/269

Abdelwahab, H. T., & Abdel-Aty, M. A. (2001). Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. *Transportation Research Record*, *1746*(1), 6-13.

Faghri, A. & Hua, J., (1993). Dynamic traffic pattern classification using artificial neural networks. *Transportation research record*, *1399*, 14-19. Rertrieved March 19, 2020. from: https://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds[]=citjournalarticle_604433_34

Nedić, V., Despotović, D., Cvetanović, S., Despotović, M., & Babić, S. (2014). Comparison of classical statistical methods and artificial neural network in traffic noise prediction. *Environmental Impact Assessment Review*, *49*, 24-30.