Application of selected artificial intelligence methods in terms of transport and intelligent transport systems

Michael Štencl / *Viliam* Lendel
Received 2012-09-27

Abstract

In recent years, there has been increased interest among both transportation researchers and practitioners in exploring the feasibility of applying artificial intelligence (AI) paradigms to address some of the transport problems in order to improve the efficiency, safety, and environmental-compatibility of transportation systems. The contribution is to highlight the application of tools and methods in the field of artificial intelligence solutions to traffic problems, where using traditional solutions would be time consuming and more expensive. This article deals with the application of selected artificial intelligence methods in case of transport and intelligent transport systems.

Michael Štencl

Department of Informatics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic
e-mail: michael.stencl@mendelu.cz

Viliam Lendel

Department of Management Theories, University of Zilina, Univerzitná 8215/1, 010 26 Zilina, Slovak Republic
e-mail: viliam.lendel@fri.uniza.sk

1 Introduction

In recent years, artificial intelligence (AI) has attracted bigger attention of researches in many different branches from signal processing, pattern recognition or time series forecasting [14]. In global view, artificial intelligence is inspired by biological processes, generally learning from previous experience. Main tasks for artificial intelligence methods are based on learning from experimental data (including learning from patterns) and transferring human knowledge into analytical models [3]. Realization of those tasks belongs to two groups of methods. The first includes Artificial Neural Networks (ANN) or Support Vector Machines, which have mathematical models behind the learning process. The second group is focused on transferring human knowledge (experience) into workable systems and uses fuzzy logic (FL) systems. This classification is not the only way to categorize artificial intelligence methods. Other simple way to categorize artificial intelligence methods is by their "inspiration". From this point of view, also two groups of methods exist. Methods inspired by biological processes, such as ANN or genetic algorithms, and methods based on Machine Learning, such as SVM.

Our purpose here is to experimentally show the set up process of artificial neural network on approximation task. For the experiments, a multilayer feed-forward neural network was used. The article also discusses the applications of fuzzy logic systems on intelligent transport systems.

As a main topic of our paper is the testing of artificial neural network methods in the area of transport management systems with affiliation to further research. First, the time series approximation ability on real-world datasets in the Czech-Slovak conditions is tested. The length of the input datasets is the main problem (quality data sets are short) so the tests are performed for datasets with missing data.

2 Application of AI in terms of transport and intelligent transport systems

Currently, traffic engineers face the challenge of the increasing complexity of transport systems. Their main aim is to ensure safe, efficient and reliable transportation while minimiz-

ing its negative impact on the environment. Traffic engineers must solve several problems, in particular the unreliability and poor road safety, capacity problems, environmental pollution and wasted energy. The transport problems emerge by their very nature and structure of transport systems. These are complex systems involving many different components and participants who have different and often conflicting objectives.

Transport problems exhibit a number of features that allow the application of methods and tools of artificial intelligence. First, they include both quantitative and qualitative data. Transport systems can often be very difficult to simulate by the traditional approach, mainly because of interactions between different elements of the transport system. In case of transport problems often difficult optimization problems have to be solved that cannot be done fully by using traditional mathematical programming methods.

Artificial Intelligence provides a wide range of tools and methods for a solution to transport problems. Methods and tools of artificial intelligence are used primarily to predict the behaviour of transportation systems, transportation optimization problems in control systems and clustering, also in the transport planning process, decision making and pattern recognition. Each area uses artificial intelligence in case of transport systems concrete examples are clearly indicated in Table 1.

3 Selection methods

Artificial intelligence methods solve three kinds of tasks – regression, classification and optimization tasks. The regression task includes also time series forecasting.

The classical ANN could be described as black box. On input, no previous knowledge needed, but data pairs of input and target data are known. The weights are unknown for the first iteration of the learning process. The classical ANN works as universal approximation tool in most of the tasks. On the output computed data is presented; closest solution to target data. FL systems are totally opposite to ANN. With the FL model, the solution is known – experience, expertise; the knowledge about the process exists and is structured. No data required. The IF-RULES rules must be defined. With defined rules, the human knowledge is transformed into a workable system. The less previous knowledge exists, the more likely it is an ANN task. The more knowledge is available, the more suitable is the task for FL modelling. Advantages of both methods are listed in Table 2 and examples of disadvantages are listed in Table 3.

In this paper we describe the process of creating and learning time-series approximation task with feed-forward multilayer perceptron. The MLP network works on input – output mapping. The most complex part of artificial neural network usage is the set up process of an artificial neural network. The whole set up process consists of few steps. Most important steps are scaling of the inputs and building the network architecture (deciding on number of hidden layers and neurons).

Scaling of the inputs determines the effective scaling of the

weights in the bottom layer and it can have effect on the quality of the results. At the outset is best to standardise all inputs to have mean zero and standard deviation one [2]. Generally speaking it is better to have too many hidden neurons than too few. With less neurons in the hidden layer, the model might not have enough flexibility to capture the nonlinearities in the data. Hastie (2001) also says, that the choice of the number of hidden layers is guided by background knowledge and experimentations. Use of multiple hidden layers allows construction of hierarchical features at different levels of resolution [2]. These indicated above will be tested by practical experiment in following section.

4 Experiment and application

In following sections the experiment with set up artificial neural network architecture and learning process of feed-forward multilayer perceptron and FL application in intelligent transport systems is described.

MLP network is frequently used in time series forecasting. Structure of the selected network is shown in Fig. 1 which consists of an input layer, several hidden layers, and output layer.
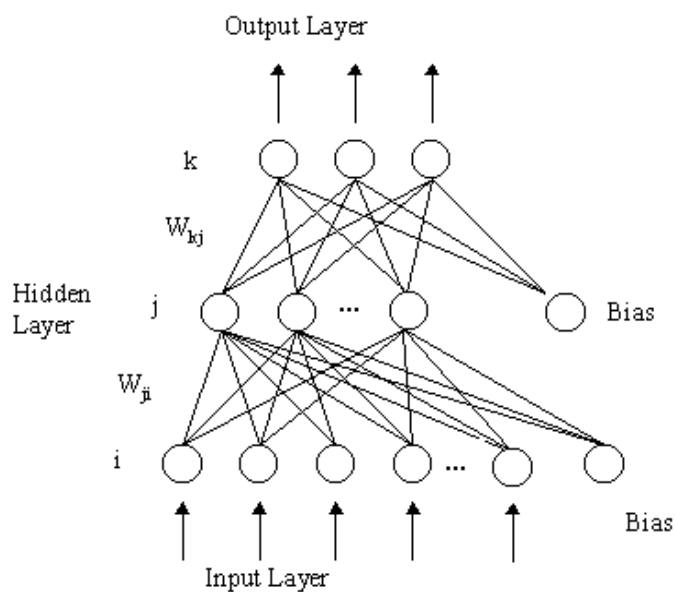

Fig. 1. Feed-forward multilayer perceptron (MLP)

Node with index i represent one neuron in input layer, analogy j is one neuron in hidden layer and k represents one neuron in output layer. The MLP network also includes weights (w) and an activation function. First are the inputs (xk, $k=1, \ldots, K$) to the neuron multiplied by weights (wki) and summed up together with the constant bias term (θi). The result (ni) is the input to the activation function (g). For the experiment a hyperbolic tangent sigmoid transfer function has been chosen and is defined as

$$\tan sig(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

The output of neuron becomes

$$y_i = g_i = g\left(\sum_{j=1}^{K} w_{ji} x_j + \theta_i\right) \tag{2}$$

Tab. 1. Areas and examples of artificial intelligence in case of transport systems

Area of use	Examples of use
Nonlinear prediction	Predicting traffic demand, or predicting the deterioration of transportation, infrastructure as a function of traffic, construction, and environmental factors.
Transportation planning process	AI-based decision support systems for transportation planning.
Optimization problems in transportation	Designing an optimal transit network for a given community, developing an optimal shipping policy for a company, developing an optimal work plan for maintaining and rehabilitating a pavement network, and developing an optimal timing plan for a group of traffic signals.
Controlling a system	Signal control of traffic at road intersections, ramp metering on freeways, dynamic route guidance, positive train control on railroads, and air traffic control.
Clustering	Identification specific classes of drivers based on driver behaviour.
Design (Computer-Aided Design)	Geometric design of highways, interchange design, structural design of pavements and bridges, culvert design, retaining walls design, and guardrail design.
Decision making	Deciding whether to build a new road, how much money should be allocated to maintenance and rehabilitation activities and which road segments or bridges to maintain, and whether to divert traffic to an alternative route in an incident situation.
Pattern recognition or classification	Automatic incident detection, image processing for traffic data collection and for identifying cracks in pavements or bridge structures.

Tab. 2. Some advantages of ANN and FL (Source: Kecman, 2001)

ANN	FL system
– Can approximate any multivariate nonlinear function	– Can approximate any multivariate nonlinear function
– Do not require deep understanding of the process of the studied task	– Are applicable when mathematical model is unknown or impossible to obtain
– Are robust to presence a noisy data	– Operate successfully under lack of precise sensor information
– Same ANN can cover broad and different classes of tasks	– Same ANN can cover broad and different classes of tasks

Tab. 3. Some disadvantages of ANN and FL (Source: Kecman, 2001)

ANN	FL system
– Need long training or learning time, problems with local minima or multiple solutions	– Human solutions of the problem must exists, and this knowledge must be structured
– Do not uncover basic internal relations or physical variables, and do not increase the knowledge about the process	– Learning (changing membership functions' shapes and positions, or rules) is highly constrained (more complex than with ANN)
– No guidance is offered about ANN structure or optimization procedure, or the type of ANN to use for a particular problem	– Experts sway between extreme poles, too much aware in field of expertise, or tending to hide their knowledge

A MLP network is defined as several neurons in parallel and series. Regarding the MLP definition, the output of two layers MLP is then defined as

$$y_i = g_i = g\left(\sum_{j=1}^{3} w_{ji}^2 g\left(n_{j1}^1\right) + \theta_{j2}^2\right) =$$

$$g\left(\sum_{j=1}^{3} w_{ji}^2 g\left(\sum_{k=1}^{K} w_{kj}^1 x_k + \theta_{j1}^1\right) + \theta_{j2}^2\right) \quad (3)$$

Designing the network consist of several steps. In the first step, the number of hidden layers and number of neurons in each layer is set up. In the second step, activation functions for each layer are chosen. The last unknown parameters to be estimated are the weights and biases (*wkji*, *θkik*). To determine the network parameters training algorithm is chosen. The most know supervised training algorithm is the Back-propagation algorithm [16].

Back-propagation algorithm is a gradient-based algorithm and in a nutshell includes four steps:

1 The architecture of the network is defined; activation functions on each layer are chosen and the network parameters, weights and biases are initialized.

2 The training parameters – network error, learning goal or maximum number of epochs – are defined.

3 On each input an output value is counted and then compared with target value. Error of the counted value is measured and the training algorithm changes weights.

4 Step 3 is iteration process until one of the training parameters is fulfilled.

Time series that have been used for experimenting represents Goods transport by Rail. The data originate from the Czech Statistical Office and are measured quarterly. All data were scaled between [-1, 1] for better ANN performance. The length of the time series is 32 units and represents quarters of the years 2000 and 2008.were used to train networks and to test network generalization ability for each time series. Input data starts with the first quarter of 2000 and ends at the last quarter of 2007. Training data starts from the first quarter of the year 2001 and ends at the last quarter of 2008.

Simulations were done with MLP networks that have the following architecture: the first set of tests were made on architecture of the network with one hidden layer. The hidden layer contains different number of neurons – the first network 25, the second network architecture contains 30 and the last architecture has 35 neurons in the hidden layer and one nonlinear output neuron. Second part of the experiment continues with adding of hidden layers with same number of neurons, precisely 20 and also one nonlinear output neuron. In all cases the *tansig* function has been used as the activation function for neurons. The experiment was done on same computer in constant conditions and was focused on iteration process of setting up the chosen ANN.

$$NMSE = \frac{1}{N\sigma^2} \sum_{i=1}^{N} (x_i - d_i)^2 \qquad (4)$$

MATLAB neural network toolbox training function *traingb* was used for training networks. Stopping criteria was defined with normalized mean square error, 1000 epochs and gradient value was 0.001. Normalized mean square error (NMSE) was used as performance measure. In Eq. (4) $\sigma2$ is the variance of the desired outputs di and N is the number of patterns.

The Fig. 2 indicates the result of the comparison of same network architecture with different number of neurons in the hidden layer. The upper line marked with crosses (X) shows the used time series. All the rest of the lines are errors of the simulated data calculated as the variation between the counted data by networks and target data. The dotted line shows the network results formed from 25 neurons in hidden layer. The dashed line is representing data simulated by network with 35 neurons in hidden layer. The last, solid line shows network with 30 neurons in the hidden layer. Goal for the lines was to copy the line in zero value

on y-axe in all 32 steps. The nearest result to this criterion brings the second architecture (30 neurons in hidden layer). This means that adding more neurons to the hidden layer do not bring better results. First network with 25 neurons in hidden layer stopped learning process on total number of epochs with NMSE 0.00122 (close to the ending value defined by 0.001). Second network (30 neurons in hidden layer) finished by achieving NMSE criterion. The learning process is shown on Fig. 3. It is shown, that still exists performance of the network, it means that better result can be achieved by prolonging the learning epochs.

The third network architecture finished, as the first one, at the total number of epochs set up for the learning experiment. The comparison shows still big errors on touching the zero line. Adding hidden layers could do improve the results of the tested networks. Also three network architectures were tested.

Fig. 4 compares the networks with different number of hidden layers. Best results were achieved by the architecture with 20 neurons in 4 hidden layers (solid line on the Fig. 4) and ended at stopping criteria defined by NMSE. The dashed line figures the architecture with 30 neurons in 4 hidden layers for visual example of the problem when defining and setting up neural networks. Dotted line represents the three layer neural network with 20 neurons in hidden layer and stopping criteria NMSE stopped the learning process. However, simulated data error shown on Fig. 4 practically shows why NMSE is not the only one stopping criterion. Dashdot line is showing overfitted neural network architecture. The architecture consisted from 5 hidden layers of 20 neurons in each hidden layer. When neural networks have too many weights they will overfit the data at the global minimums as is shown on Fig. 4. For improving the results, other learning algorithm could be tested or other type of the network should be selected.

5 Discussion

Experiments described in this article show main problems of the neural network. Setting up the right neural networks architecture is the key to good results. As is mentioned above, choice of the number of neurons in the hidden layer and number of hidden layers is guided by background knowledge and experience. Some of the researches use cross-validation to estimate the optimal number, but this seems unnecessary if cross-validation is used to estimate the regularization parameter [2]. Typically the number of hidden neurons is in the range of 5 to 100 neurones. We have experimentally tested the number of neurons tested in previous tests. The experiment also shows the typical process of setting up the network architecture by putting down a reasonably large number of neurons and training them with regularization.

6 Conclusion

Finally we can say that we successfully carried out the set up process of artificial neural network on approximation task. Methods and tools of artificial intelligence have wide application in the field of transport systems. Neural networks have
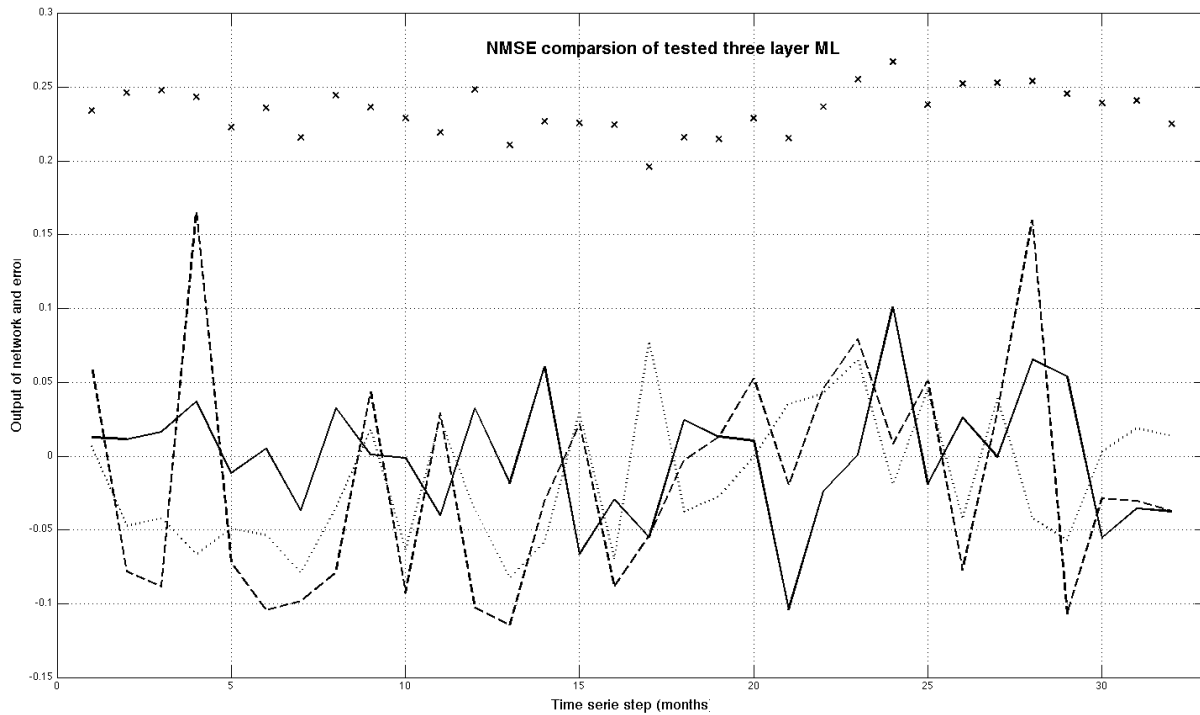
Fig. 2. NMSE comparsion of tested three layers MLP (Source: own experiment)
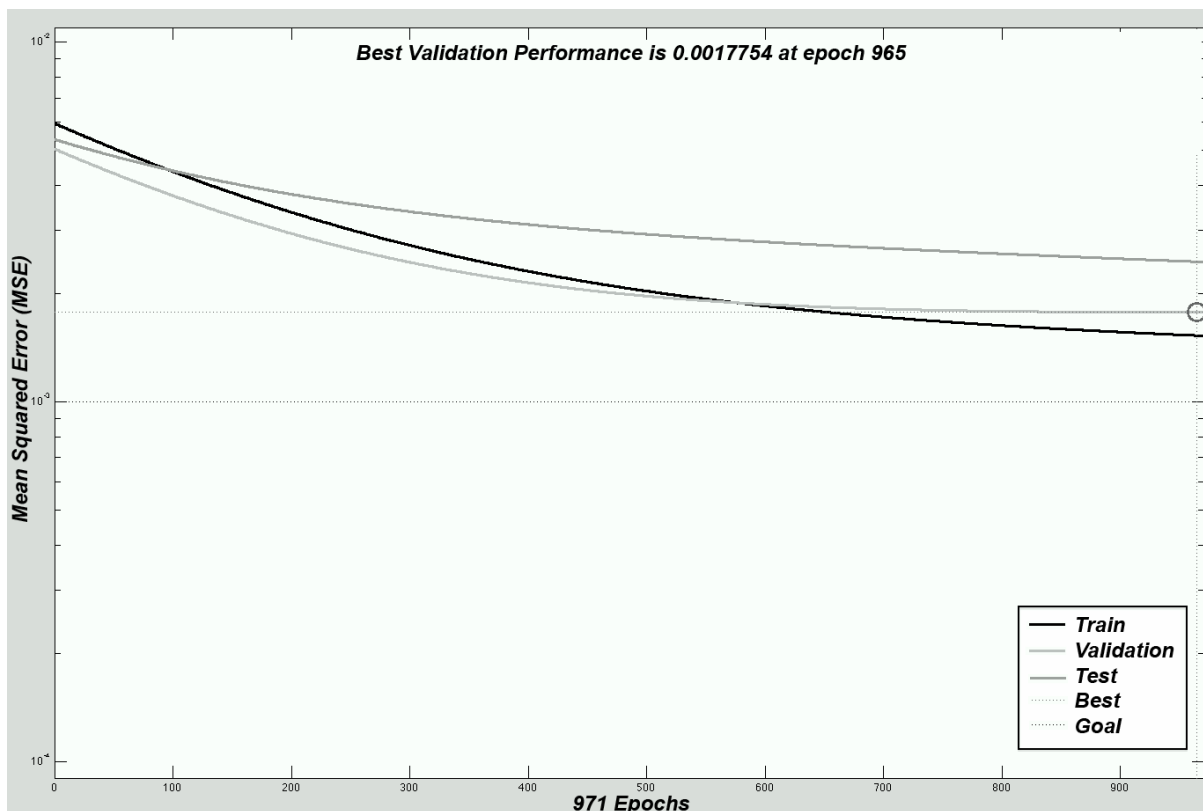

Fig. 3. Learning performance of second MLP (Source: own experiment)

wide application in the field of transport infrastructure and traffic management. Tools based on neural networks are used to predict the number of data required for traffic management. Often it is about the weather information, such as ice formation, which is signalled to the driver by the weather forecast. The outputs of neural networks are used in the diagnostic unit, which forms the basis for a system to monitor vibration of jet engines at Rolls-

Royce Company [15]. Implementation of neural networks in the production and management of cars is sponsored by the European Union under the 5th and 6 Framework Programme joined by several major car manufacturers. The principle of such control systems is the combination of detailed technical expertise with non-linear interpolation using neural networks.

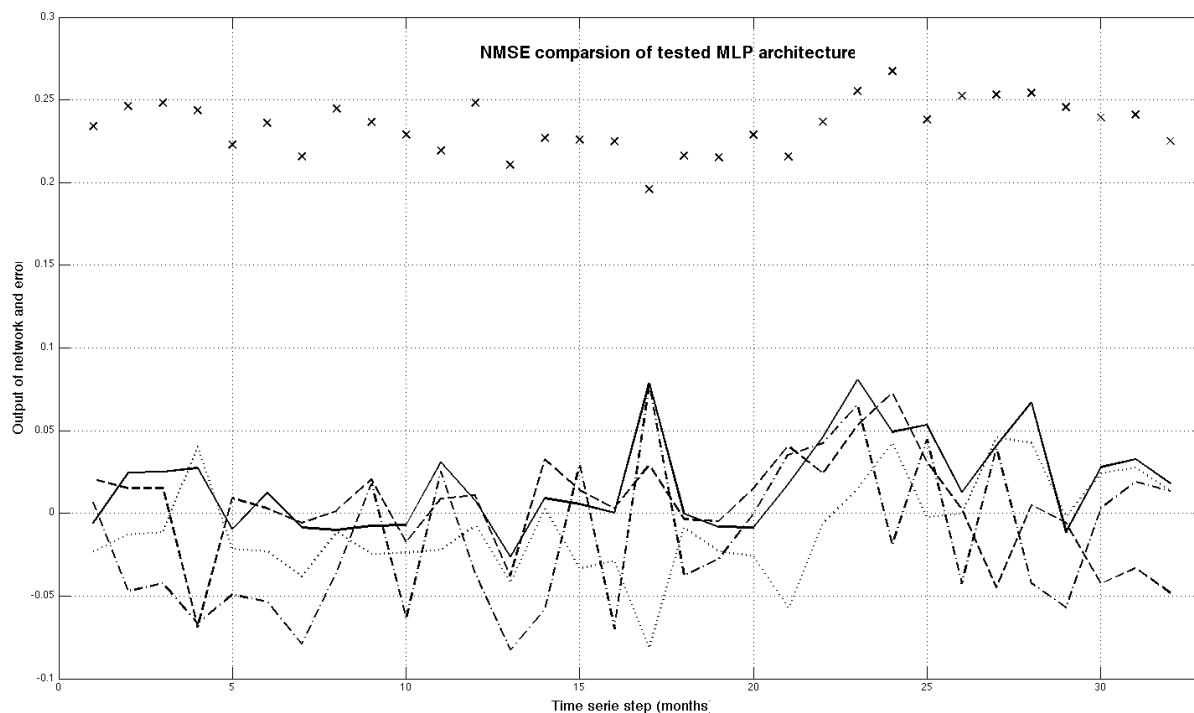Nonlinear and very complex management systems in road

Fig. 4. Comparison of the number of hidden layers MLP (Source: own experiment)

tunnel ventilation is a typical example of fuzzy inference system in practice. It allows describing the uncertainty and vagueness of statements with formalized rules instead of binary yes / no [15]. Ventilation in the road tunnel is characterized by non-linear systems with large time delay. It is suitable for the use of fuzzy logic control.

The concept of using the artificial neural network was successfully confirmed by testing different architectures of MLP network. The connected further research will be focused on other datasets and also further experiments with different Artificial Intelligence methods on approximations (together with possible time series predictions) and optimizations tasks.

Transport management includes also the consideration of costs. Transport costing can be effectively supported by calculation tools based on the modeling of technology performance relations [1]. It is intended to be examined how the results or outcomes of these costing methods can be integrated into intelligent transport systems relying on Artificial Intelligence procedures.

Acknowledgement

References

1 **Bokor Z**, *Improving Transport Costing by using Operation Modeling*, Transport **26/2** (2011), 128–132, DOI 10.3846/16484142.2011.586111.

2 **Hastie T, Tibshirani R, Friedman J**, *The elements of statistical learning*, New York, 2001. Springer series in statistic.

3 **Kecman V**, *Learning and Soft Computing: support vector machines, neural networks, and fuzzy logic models.*, The MIT Press, Cambridge, 2001. "A Bradford Book".

4 **Lendel V, Štencl M**, *Possibility of Applying Artificial Intelligence in Terms of Intelligent Transport Systems*, Journal of Information, Control and Management systems **8** (2010), no. 1, 37 – 46. Faculty of Management Science and Informatics. University of Žilina.

5 **Miles J C, Walker A J**, *The potential application of artificial intelligence in transport*, Journal Intelligent Transport Systems **153** (2006), no. 3. Institution of Engineering and Technology. ISSN 17480248.

6 **Mitchell T M**, *Machine Learning*, McGraw-Hill, 1997. 1st edition.

7 **Nedeliaková E, Dolinayová A, Gašparík J**, *Methodology of transport regulation in the Slovak Republic*, Periodica Polytechnica Transportation Engineering **38/1** (2010), 37–43.

8 **Official Web Site of Czech Statistical Office.**, http://www.czso.cz/eng/redakce.nsf/i/goods_transport_time_series.

9 **Ramos C, Augusto C J, Shapiro D**, *Ambient Intelligence – the Next Step for Artificial Intelligence*, Intelligent Systems (March/April 2008), 15–18.

10 **Ripley B D**, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, 1996.

11 **Sadek A W**, *Artificial Intelligence Applications in Transportation*.

12 **Sima J, Neruda R**, *Theoretical Issues of Neural Networks*, MATFYZPRESS, Prague, 1996.

13 **Smith B L**, *Artificial intelligence and intelligent transportation systems*, Journal Computing in Civil Engineering, New York, 1996.

14 **Smith K, Gupta J**, *Neural Networks in Business: Techniques and Application*, Hershey: IRM Pres., 2002.

15 **Spalek J, Janota A, balažovičová M, Přibyl P**, *Rozhodovanie a riadenie s podporou umelej inteligencie*, Žilina: EDIS., 2005.