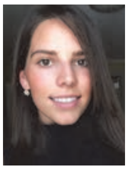
DOI: https://doi.org/10.14256/JCE.2273.2017

Građevinar 9/2018

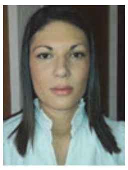
Primljen / Received: 8.11.2017.
Ispravljen / Corrected: 19.3.2018.
Prihvaćen / Accepted: 30.5.2018.
Dostupno online / Available online: 10.10.2018.

Estimation of recycling capacity using ANN and SVM

Authors:



[1]Aleksandra Vujkov, MCE
vujkova@uns.ac.rs



[1]Dragana Bibić, MCE
draganabibic@uns.ac.rs

[1]Assist.Prof. Igor Peško, PhD. CE
igorbp@uns.ac.rs

[1]Assist.Prof. Vladimir Mučenski, PhD. CE
mucenskiv@uns.ac.rs

[1]Prof. Jasmina Dražić, PhD. CE
dramina@uns.ac.rs

[1]Prof. Milan Trivunić, PhD. CE
trule@uns.ac.rs

[1]University of Novi Sad
Faculty of Technical Sciences
Department of Civil Engineering and Geodesy

Preliminary report

Aleksandra Vujkov, Dragana Bibić, Igor Peško, Vladimir Mučenski, Jasmina Dražić, Milan Trivunić

Estimation of recycling capacity using ANN and SVM

This paper presents estimation of the quantity of concrete and reinforcement that can be recycled for residential buildings constructed as skeleton structures. Models based on artificial intelligence, involving the use of Artificial Neural Networks (ANNs) and the Support Vector Machines (SVM) methods, were formed in order to estimate quantities of these materials. The results show that the application of ANNs and SVM methods is a good solution for the estimation of recycling capacity. The mean absolute percentage error (MAPE) for the selected ANNs for predicting quantity of concrete and reinforcement is 8.74 % and 12.58 %, respectively.

Key words:

buildings, recycling, concrete, reinforcement, artificial neural networks, support vector machine method

Prethodno priopćenje

Aleksandra Vujkov, Dragana Bibić, Igor Peško, Vladimir Mučenski, Jasmina Dražić, Milan Trivunić

Ocjena mogućnosti recikliranja pomoću metoda ANN i SVM

U ovom je radu prikazana analiza količina betona i armature koje se mogu reciklirati iz stambenih zgrada sa skeletnim konstrukcijama. Kako bi se procijenile količine tih materijala, izrađeni su modeli bazirani na umjetnoj inteligenciji, pri čemu su korištene metode ANN (umjetne neuronske mreže) i SVM (metoda potpornih vektora). Rezultati pokazuju da se primjenom metoda ANN i SVM postižu dobra rješenja za procjenu mogućnosti recikliranja. Srednja apsolutna postotna pogreška (MAPE) mreža ANN odabranih za predviđanje količine betona i armature iznosi 8,74 % za beton i 12,58 % za armaturu.

Ključne riječi:

zgrade, recikliranje, beton, armatura, umjetne neuronske mreže, metoda potpornih vektora

Vorherige Mitteilung

Aleksandra Vujkov, Dragana Bibić, Igor Peško, Vladimir Mučenski, Jasmina Dražić, Milan Trivunić

Beurteilung der Recyclingfähigkeit mithilfe der KNN und SVM-Methoden

In dieser Abhandlung wird die Analyse der Menge an Beton und Bewehrung von Wohngebäuden mit einer Skelettkonstruktion dargelegt, die wiederverwertet werden kann. Um die Menge dieser Materialien einzuschätzen, wurden Modelle angefertigt, die sich auf künstlicher Intelligenz begründen, wobei die KNN- (künstliches neuronales Netzwerk) und die SVM (Stützvektormethode) Methoden angewendet wurden. Die Ergebnisse zeigen, dass durch Anwendung der KNN- und der SVM-Methoden gute Lösungen für die Einschätzung der Recyclingfähigkeit erzielt werden. Der mittlere absolute prozentuale Fehler (MAPE) der für die Vorhersage der Beton- und Bewehrungsmenge ausgewählten Netze des KNN beträgt 8,74 % für den Beton, d. h. 12,58 % für die Bewehrung.

Schlüsselwörter:

Gebäude, Wiederverwertung, Beton, Bewehrung, künstliche neuronale Netzwerke, Stützvektormethode

1. Introduction

Building, as a result of a construction project, has its own lifetime that begins with the need for constructing it, and ends when the building is no longer needed and can not be changed, i.e. when it has to be demolished. Removal of buildings, at the end of their service life, produces large quantities of materials that can still be used to a greater or lesser extent. Defining ways for easy and simple use of materials produced in the process of building demolition, is aimed at reducing or completely eliminating the costs of building disposal. Ever increasing costs of waste disposal and growing emphasis on the need for higher awareness and concern for the environment, have led to an increased use of recycled materials in construction industry. Depending on the type of buildings planned for demolition, there are various kinds of materials that can be generated during the demolition process and then recycled. Considering that in the territory of the Republic of Serbia there are 22,272,500 square meters of flats that are older than 70 years, and 74,053,973 square meters of flats that are older than 45 years, it can easily be concluded that the country has a significant potential for the recycling of construction materials [1].

The reusability of materials obtained in the process of demolition of existing reinforced concrete structures, along with adequate preparation and processing, largely eliminates the need for disposal of large quantities of materials. The use of materials obtained by recycling can lead to reduction of costs of new building processes, by reduced use of new materials and reduced material disposal costs [2]. For this reason, the evaluation of the quantity of materials suitable for recycling, i.e. concrete and reinforcement in this particular case, and their use in the construction of new facilities, is a very important step forward.

However, the evaluation of quantities of concrete and reinforcement resulting from demolition of existing buildings is not always an easy process. The reason for this is in many cases the lack of design documentation, which complicates assessment of quantities of concrete and reinforcement. The development of simple estimating models diminishes influence of visual examination of design, or influence of the data that can not be collected. One of the ways for forming such models is through application of artificial intelligence.

In construction industry, artificial Intelligence can be applied at all stages of construction projects, from the earliest stages of planning and design, to the recycling of building materials during demolition. Models based on artificial intelligence can be used for quick and easy estimation of the quantities of concrete and reinforcement for recycling, in case the design and technical documentation for the building planned for demolition proves unavailable.

The basic prerequisite for formation of a model based on artificial intelligence is a suitable database, as it is considered essential for accurate estimation by the model [3]. This paper presents an overview of current work in this field as well as the application of artificial intelligence in solving similar problems. It also provides the analysis of the database specially made for solving this problem, and depicts application of ANN (Artificial Neural Networks) and SVM (Support Vector Machines) for the purpose of estimating the recycling capacity of residential buildings that are to be demolished at the end of their service life. The main goal of this research is to analyse and confirm that artificial intelligence can provide a great support in the construction waste management process. The primary hypothesis is that ANN and SVM models can offer sufficient estimation accuracy, which will diminish influence of building design when planning management of construction waste.

2. Literature review

Nowadays, ANN and SVM are tools that are mostly applied for solving problems of regression and classification by changing the parameters that control their training. The beginning of the neural network development can be traced back to the year of 1943 and the article of Warren McCulloch and Walter Pitts entitled "Logical calculations of the ideas typical of neural activity". A great contribution to artificial neural networks has been made due to the interest of the American Military Agency DARPA (Defense Advances Research Projects) [4]. The SVM was introduced by Vladimir V. Vapnik [5] in 1995 and the method was primarily used for solving classification problems. However, in recent times, the method has spread to the domain of solving regression problems. The SVM is a method for training and defining the function of separation in classification problems, or for making predictions in regression problems. This approach is based on the SLT (statistical learning theory) developed by Vapnik and Chervonenkis [6] in the late twentieth century.

Ginaydin and Gibson [7] analysed the models based on ANN where the output data, the cost of the building, were formed as per square meter of the building area. Thirty prestressed-concrete buildings were analysed, and the prediction accuracy was 7 %. Lazarevska et al. [8] presented application of neural networks in solving complex construction problems, in particular for determining fire resistance of structural elements. They made a prediction model for defining fire resistance of RC columns embedded into walls and exposed to standard fire from one side. The database included 398 cases and was divided into two groups: the training data – 318, and the testing data – 80 cases. They made a comparison of calculated (actual) and predicted fire resistance curves for an eccentrically loaded RC column. It was concluded that these curves and relevant results are similar. In their research, Lazarevska et al. [9] presented application of an artificial neural network and proposed a model for determining fire resistance of centrally loaded composite columns made of concrete and steel. They analysed 87 cases, 70 of which were chosen for the training of neural networks and 17 for prediction accuracy testing. Minor errors were observed by comparing predictions obtained by ANN models with real values. Mučenski et al. [10] published a paper about estimation of recycling capacity of concrete and rebars for residential buildings constructed as skeletal structures. 95 projects were found in the training set, with 9 input and 2 output data. The models were based on ANN in Matlab. Z-score normalisation of data was used. The resulting percentage error and the actual value in this paper was 9.10 %. By applying neural networks, Peško [3] proposed a model for estimating cost and time needed for the construction of urban roads. The database included 130 ongoing and realized construction/reconstruction projects involving urban roads. The data were divided into two groups: the training data 115 instances and the testing data 15 instances, and the pseudo random sampling method was applied. The normalisation of data was conducted using the z-score normalisation. The MAPE for the model that provided the best result was 7.93 % for cost and 5.87 % for construction period.

Kim et al. [11] published a study on the estimation of costs according to three models, and made a comparative analysis. The models were based on the multiple regression analysis MRA, artificial neural network ANN, and reasoning based on and case-based reasoning (CBR). The survey was conducted on 530 residential building projects constructed in the period from 1997 to 2000 in Seoul. The model accuracy measurement was performed using absolute error. The best prediction was obtained by applying ANNs. However, the disadvantage of this method compared to the CBR model was that MRA was a slow process of finding the optimal network using an iterative procedure. Sonmez [12] made a conceptual cost estimation of building projects in the USA in the period between 1975 and 1995. He compared the models made by regression analysis and ANN. The validation was performed on the basis of square error and absolute error. Additionally, the advantages and disadvantages of these models were presented, and comments regarding simultaneous use of these two models were given, which suggested combined use of the two models. Wang and Gibson [13] analysed two model types in order to predict success of various projects. The first model was based on simple linear regression, whereas the second one was based on artificial neural networks. They analysed data for 62 industrial facilities and 78 multi-storey buildings. The analysis of their model showed that the projects for which good planning was initially done exhibited better performance upon completion.

Cheng and Wu [14] linked two approaches to artificial intelligence (fast genetic algorithm-fmGa and SVM) for the purpose of solving problems in the area of construction management. The name of the new model is ESIM (evolutionary support vector machine inference model). The aim of developing a new model was to obtain a minimum prediction error, and to simultaneously find optimum parameters trade-off and kernel parameter. Based on the analysis and results obtained in the study, the authors concluded that the ESIM model could be used for solving various problems in construction management. Strobbe et al. [15] investigated whether it was possible to learn the architectural style from a set of cases, and whether there was a possibility to classify new styles as similar or different styles (designs) from the observed cases. Two methods were applied (SVM and graph kernels). In addition, the authors demonstrated feasibility of the proposed detection method on the example of "Malagueira houses". They concluded that their model, with an accuracy of 87.5 %, was able to generalize styles that were not taken into account during training. Cheng et al. [16] proposed the EFSIMT model (Evolutionary Fuzzy Support Vector Machine Model Inference for Time Series Data) for the prediction of HPC (High Performance Concrete) strength. The EFSIMT model was obtained by joining the multiple methods FL (Fuzzy Logic), wSVM (weighted Support Vector Machines) and fmGA (fast messy genetic algorithms). The database included 1030 samples of concrete, and the data were divided into two groups

- the training data - 90 % or 927 samples
- the testing data - 10 % or 103 samples.

The authors compared the EFSIMT model results with the results obtained by SVM and BNP, which were previously explored by other authors. It was noted that better results in prediction of high-performance were obtained by the EFSIMT model, compared to the SVM and BPN. It was concluded that EFSIMT is a good tool for defining the HPC strength. Zhang et al. [17] predicted profitability of construction companies in China using the PCA (Principal Component Analysis) method and SVM. Based on the PCA method, the authors obtained the index ("composite index"), and then applied the technique to predict profitability with this index by using SVM. The results showed that the well 'trained' SVM model could predict profitability with an accuracy of over 80 %. By applying the ESIM (Evolutionary Support Vector Machine Inference Model), Cheng et al. [18] predicted success of a number of construction projects. ESIM is a method that integrates two methods: SVM and fmGA. The research also included the application of CAPP (Continuous Assessment of Project Performance) in order to select factors that influence success of a project. The database included 46 construction projects. The authors concluded that the ESIM method provides good results in predicting project success.

By applying a set of ANN and SVM, Wang et al. [19] predicted the cost of construction and construction time schedule, i.e. they explored in which way project success is influenced by early planning of construction work. The database included 92 valid samples (of the project) collected in the period from 2007 to 2010, representing a total construction cost of about $ 1.1 billion. The collected information was used for construction and for testing the set of neural networks and SVM prediction models. The data were divided into two groups: 67 training sets and 25 testing sets. By comparing the results, the authors concluded that the SVM model provides a cost prediction accuracy of 92 %, whereas for predicting successfulness of a dynamic plan the use of ANN (Adaptive Boosting NN) ensures the prediction accuracy of 80 %.

The review of relevant literature has not provided any research that refers to application of ANN and SVM methods for the purpose of estimating quantity of construction material (concrete and reinforcement) required for the construction, recycling etc.

3. Proposed methodology

The research required collecting data on the quantities of material of skeletal buildings, which were in fact taken from the design for the building permit. ANN and SVM models were formed using the data collected in this way. The key part of the analysis in the application of artificial intelligence is the database that needs to be properly prepared. For a good model, a valid database is certainly a necessary precondition, but it is not sufficient. A good model has good generalization capabilities. Models were formed based on 9 input parameters describing characteristics of buildings and two output parameters, the quantity of concrete and steel reinforcement. ANN and SVM models were created, and models which provided the lowest MAPE were selected, using an iterative procedure via the software STATISTICA 8 [20]. 25 ANN models for estimating quantity of concrete and reinforcement in the case of two output data [21], and 5 SVM models for estimating the quantity of concrete, were formed in the previous research [22]. The z-score normalisation for data preparation was applied in both options.

In order to expand the research presented in the following text, the data were prepared using another type of min-max normalisation, as well as the data that were not normalized. Additionally, 120 ANNs models were formed, which enabled separate estimation of concrete and reinforcement. Thirty models were formed using the SVM method in order to estimate the quantity of concrete and reinforcement. At the end, a comparative analysis was conducted on the accuracy of ANN and SVM models that provided the most acceptable error.

3.1. Database for ANN and SVM Models

The database was based on the characteristics and quantities of the studied materials from 100 projects (residential buildings) located on the territory of Novi Sad, Republic of Serbia, with the purpose being to form an estimation model on the quantity of concrete and reinforcement embedded in existing structures i.e. to estimate the quantity of concrete and reinforcement that could be recycled (fundamental structures, columns, stiffening walls, beams and floor structures).

When forming the database, the aim was to keep it simple yet informative, i.e. to incorporate all data that are highly relevant to the accuracy of the model. The data contained in the database for formation of ANN and SVM models were divided into input and output data. Parameters that describe building characteristics were numerical, geometric and structural. Input data were related primarily to building characteristics that are in correlation with the quantity of concrete and reinforcement such as: complexity of the building, total gross area of the building, average gross floor area, height of the building, number of stiffening walls, longitudinal and transverse disposition of the structure, type of floor structure, and type of floor support structure. Output data were the quantity of concrete and reinforcement. All structures rest on the foundation slab, and so it was considered to be a constant for all buildings. The foundation work characteristics were not analysed as input data. Although a small number of buildings have other types of foundations, it was not advisable to consider this information due to the lack of data and possible distortion to ANN and SVM. The database included only buildings without dilatations or with just one dilatation.

The parameters chosen for describing characteristics of the structure are shown in Table 1. They include geometric and structural characteristics of the building.

Table 1. Description of input and output parameters of the building used for ANN and SVM training

Type	Definition	Design value or characteristic	Corresponding model value (value before normalisation)
Input data	Complexity of building	Simple	1
		Medium	2
		Complex	3
		Very complex	4
	Total gross area	from 1000 m^2 to 8000 m^2	from 1000 m^2 to 8000 m^2
	Average gross floor area	from 200 m^2 to 2000 m^2	from 200 m^2 to 2000 m^2
	Building height	from 13 m to 27 m	from 13 m to 27 m
	Number of stiffening walls	from 0 to 13	from 0 to 13
	Longitudinal and transverse disposition	1.00 m - 1.99 m	1
		2.00 m - 2.99 m	2
		3.00 m - 3.99 m	3
		4.00 m - 4.99 m	4
		5.00 m - 5.99 m	5
		6.00 m - 6.99 m	6
		7.00 m - 7.99 m	7
	Type of floor structure	Full RC slab	1
		Semi-prefabricated ceiling type "FERT"'	2
	Type of supporting floor structure	Direct support	1
		Girder support	2
Output data	Quantity of concrete	from 420 m^3 to 4500 m^3	from 420 m^3 to 4500 m^3
	Quantity of reinforcement	from 28500 kg to 310000 kg	from 28500 kg to 310000 kg

Table 2. Minimum and maximum values of subsets for training, validation and testing

Type		Subset for training		Subset for validation		Subset for testing	
		Min	Max	Min	Max	Min	Max
Input data	Total gross area [m^2]	1000	7500	1300	6800	1110	5500
	Average gross floor area [m^2]	200	1500	280	1,300	250	1200
	Building height [m]	13	27	15	25	16	23
	Number of stiffening walls [kom]	0	13	3	8	3	10
Output data	Quantity of reinforcement [kg]	28500	248000	42000	232000	31200	168000
	Quantity of concrete [m^3]	420	3,800	620	3200	600	2600

Buildings were classified into 4 categories according to their complexity. Buildings classified under category 1 (simple complex) were rectangular in base and were without any changes in the construction in height; buildings classified under category 2 (medium complex) were rectangular in base and were with minor changes in the construction in height; complex buildings classified under category 3 were those with an indented base while very complex buildings were those with a complex base and with significant changes in the floor structure. Category 4 is highly atypical, and buildings belonging to this category were not included in the database. Basic parameters which directly influence the output data were the total area of the building and the floor surfaces. When defining surfaces, gross surface data were taken for convenience of determination in the phase of analysing characteristics of the building when it was difficult to estimate the net area. The height of the building as a parameter was directly related to the total area of the building and in this way it affected the output data when processing the database. The height of buildings was measured from the ground surface to the highest point of the building. In aseismic design and construction of the skeletal system it was expected to build stiffening walls in order to amortize the impact of earthquakes. Stiffening walls were made of reinforced concrete, which explains their inclusion in the database. Ranges of main vertical elements were directly related to dimensions of horizontal supporting elements. Larger ranges require larger dimensions of the elements and thus a greater quantity of material. Buildings with two types of floor structure – solid reinforced concrete slab and semi-prefabricated FERT type floor structure, as well as two types of supporting floor structures, as shown in Table 1, were taken into consideration in the process of gathering material for this research.

For the purpose of forming a model based on artificial intelligence, i.e. on the application of ANN and SVM in this particular case, the entire database had to be divided into two subsets: the training data and the testing data. Additionally, extreme values (min and max) of all parameters (input and output) had to be included in the training subset. The scope of the newly established model was thus extended, which provided for higher accuracy of the estimation. Additionally, all projects that fell outside the scope of the database, i.e. that had an extreme value of some of the data compared to most of the data, were eliminated from further analysis. This was in fact the reason why the building with the largest area (9,500m2) was excluded from further analysis.

The data belonging to training set and to a testing set were not defined by random selection only. Namely, in the context of a database composed of 99 buildings (after exclusion of the project with the largest surface area), the value of output parameters was divided into 8 intervals for concrete (from 0 to 499, from 500 to 999, from 1000 to 1499, from 1500 to 1999, from 2000 to 2499, from 2500 to 2999, from 3000 to 3499, and from 3500 to 3999) and 9 intervals for reinforcement (from 25000 to 49999, from 50000 to 74999, from 75000 to 99999, from 100000 to 124999, from 125000 to 149999, from 150000 to 174999, from 175000 to 199999, from 200000 to 224999, and from 225000 to 249999). After counting, distribution of the database on training and testing data was carried out, but on condition that the structure of one set reflects the structure of the other, and vice versa. As described above, 10 projects were selected and they form a testing subset. In other words, a subset of training data consists of the remaining 89 projects.

When choosing data, care was taken that the data related to materials be harmonized, i.e. that selected projects with the quantity of concrete and reinforcement fit right into intervals with the highest number of repetitions. During selection of samples, minimum and maximum extremes were not taken into account to enable a more accurate prediction of the quantities of material. For the purpose of creating the ANN model, apart from the above mentioned divisions, another set of 10 samples (which was intended to avoid overfitting the network) was chosen within the selected training database. The above mentioned subset is the cross-validation subset, and the samples were chosen according to the same principle as the samples from the testing subset. Table 2 presents the minimum and maximum values of the subsets for training, validation and testing. Bearing in mind the input data, their characteristics, and their differences in the order of magnitude, it was necessary to actually prepare these data for proper use. For the model forming purposes, the data were normalised to reduce them to the same order of magnitude. This normalisation was performed using the z-score normalisation (performs transformation of data into the boundaries from -1 to 1) and the min-max normalisation (performs transformation of data into the boundaries from 0 to 1) [23]. The normalisation was carried out for input and output data within a training subset, as well as for input data from the testing subset. After definition of the model, it was necessary to reverse the data

(to return normalized data into original values) generated from the model, in order to carry out a comparative analysis with expected values.
The database preparation was followed by the model forming activity. The Statistica 8 software was applied for the purpose of forming the model. This software offered the possibility of data processing using the ANN and SVM methods.

3.2. Building ANN Models

Artificial neural networks created using Statistica 8 allow selection of one or more parameters as output data, which is different and very useful compared to other forms of artificial intelligence that do not have this feature. One model can predict the quantity of several different materials. The networks that provide estimation in case of just one output data (concrete or reinforcement) were formed for the purposes of this research and to ensure comparability of data with the SVM method. Finding the most appropriate model of artificial neural network was an iterative process that was conducted in several phases. The first phase of model creation involved input of normalized data into the software, and arrangement of samples in three different subsets - training, validation and testing, as discussed in previous section. Input values (nine parameters shown in Table 1) and output values (quantity of concrete and reinforcement) were defined in the second phase. The range of hidden and output neurons, and the type of neural network, were selected in the third phase, which was the most important phase in the creation of the ANN model. The network type MLP (Multilayer Perceptron) was used for the purposes of this research. In addition to its use for solving classification problems, the MLP is also utilised for regression analysis [24].
The activation functions for the hidden and output neurons were set in the fourth phase. As to activation functions of hidden neurons, unipolar and bipolar sigmoid functions are the functions that are most frequently used for this type of problem [25]. According to that, logistic activation function and hyperbolic tangent were used for hidden neurons, whereas the activation function identity was used for output neurons.
Besides the mentioned functions that were selected in previous iterations, software provides the possibility for creating the networks in such a way that software chooses the best activation functions by itself from a wide range of offered functions, which was the last phase of model creation. The activation functions of ANN models that provided the most accurate prediction are shown in Table 3.

Table 3. Activation functions of ANN models

Function	Form of the function
Identity	a
Exponential	e^{-a}
Logistic	$\frac{1}{1+e^{-a}}$
Hyperbolic tangent	$\frac{e^{a}-e^{-a}}{e^{a}+e^{-a}}$

Each iteration for creating the artificial neural network involved setting a number of training networks that ranges from 50 to 5000, as well as the number of networks out of which only the first ranked network from those 50 to 5000 was taken into consideration.
Sixty models were created for estimating the quantity of concrete, and 60 models for the quantity of reinforcement. These models were obtained by analysing and processing the real data (without normalisation) as well as by processing the data using the z-score and min-max normalisation. The obtained output values were compared with the testing subset, and the models that provided the most accurate predictions were shown in the results. A comparative analysis was based on APE (absolute percentage error) and MAPE (mean absolute percentage error).

3.3. Building the SVM Models

The first step towards formation of the SVM model involved definition of the database and division of the data into input and output data, as explained in previous section. After adequate database preparation, the model forming activity was initiated. After defining the training and validation subset, the software offers a choice of error function. By using this software, two types of error functions can be distinguished (Table 4).

Table 4. Error function of SVM model

Function	Form of the function	Minimize subject to
Type 1	$\frac{1}{2}w^{T}w+C\sum_{i=1}^{N}\xi_i+C\sum_{i=1}^{N}\xi_i^{*}$	$w^{T}\phi(xi)+b-yi\leq\varepsilon+\xi^{*}_{i}$ $yi-w^{T}\phi(xi)-bi\leq\varepsilon+\xi_i$ $\xi_i,\xi^{*}_{i}\geq0$, $i=1, N$
Type 2	$\frac{1}{2}w^{T}w-C\left(v\varepsilon+\frac{1}{N}\sum_{i=1}^{N}\left(\xi_i+\xi_i^{*}\right)\right)$	$(w^{T}\phi(xi)+b)-yi\leq\varepsilon+\xi^{*}_{i}$ $yi-(w^{T}\phi(xi)-bi)\leq\varepsilon+\xi_i$ $\xi_i,\xi^{*}_{i}\geq0$, $i=1, N$, $\varepsilon\geq0$

Type 1 function is defined by the parameters of trade-off (C) and insensitivity zone (ε), and the type 2 function is defined by the parameters of trade-off (C) and Nu (ν) – upper bound on the fraction of errors and a lower bound on the fraction of support vectors (SVs) [26]. The values of parameters C and ε range from 0 to ∞, and the value of parameter ν ranges from 0 to 1. The Kernel RBF function, in which the parameter γ was varied, was also defined:

$$K(x_i,x_j) = \exp(-1/(2\sigma\cdot 2||x-x_i||\cdot 2), \quad \sigma - \textit{devijacija funkcije RBF}\ \text{(Gauss)} \quad (1)$$

Based on the entered data and selection of the above mentioned parameters, the training model was implemented and a subsequent validation of the model was conducted. More precisely, the accuracy of the generated and reversed data was tested i.e. deviations from expected values were checked. A comparative analysis was based on APE (absolute percentage error) and MAPE (mean absolute percentage error).
30 models in total were formed to estimate the quantity of material, i.e. 15 SVM models to estimate the required quantity of

concrete, and 15 SVM models to estimate the required quantity of reinforcement. Out of 15 SVM models for estimating the required quantity of concrete, 5 models were obtained from the analysis and processing of real values, 5 models were obtained from the z-score normalisation, and 5 models were obtained from the min-max normalisation. The same approach was applied for the remaining 15 SVM models, which were selected for estimating the required quantity of reinforcement.

4. Results and discussion

The ANN models, which provided the most accurate prediction of targeted materials, and a comparative overview of results by types of normalisation, are shown in Table 5 and Table 6.

It should be noted that out of 12 selected ANN models, the best results were obtained by the MLP 9-6-1 network, which had 6 hidden neurons and 9 inputs formed on the basis of the z-score normalisation. Model 5 had MAPE of 8.74 % for estimation of the quantity of concrete. As presented, Tanh activation function for hidden neurons provided better results than the exponential function. Interestingly, the min-max normalisation showed lower accuracy than the ANN with unmodified input data (without normalisation).

As can be seen in Table 6, model 18 provided the lowest MAPE of 12.58 % for the reinforcement quantity estimation. Model MLP 9-13-1 has 13 hidden neurons and 9 inputs that were formed on the basis of the z-score normalisation. Again, the z-score normalisation provided the best accuracy, while the min-max

Table 5. MAPE of ANN models – estimation of required quantity of concrete

Type of normalisation	Model No.	Network characteristics			MAPE prediction of concrete quantity [%]
		Network	Activation function for hidden neurons	Activation function for output neurons	
min-max	Model 1	MLP 9-2-1	Tanh	Identity	13.37
	Model 2	MLP 9-2-1	Tanh	Identity	15.26
	Model 3	MLP 9-2-1	Tanh	Identity	13.89
	Model 4	MLP 9-5-1	Exponential	Identity	12.26
z-score	Model 5	MLP 9-6-1	Tanh	Identity	8.74
	Model 6	MLP 9-3-1	Tanh	Identity	11.30
	Model 7	MLP 9-2-1	Tanh	Identity	14.50
	Model 8	MLP 9-2-1	Exponential	Exponential	11.25
Without normalisation	Model 9	MLP 9-4-1	Tanh	Identity	10.20
	Model 10	MLP 9-1-1	Tanh	Identity	14.25
	Model 11	MLP 9-2-1	Tanh	Identity	14.87
	Model 12	MLP 9-2-1	Exponential	Identity	12.05

Table 6. MAPE of ANN models – estimation of the required quantity of reinforcement

Type of normalisation	Model No.	Network characteristics			MAPE prediction of reinforcementquantity [%]
		Network	Activation function for hidden neurons	Activation function for output neurons	
min-max	Model 13	MLP 9-17-1	Tanh	Identity	15.94
	Model 14	MLP 9-12-1	Tanh	Identity	21.41
	Model 15	MLP 9-17-1	Tanh	Identity	19.08
	Model 16	MLP 9-14-1	Exponential	Logistic	19.98
z-score	Model 17	MLP 9-16-1	Tanh	Identity	15.26
	Model 18	MLP 9-13-1	Tanh	Identity	12.58
	Model 19	MLP 9-13-1	Tanh	Identity	14.45
	Model 20	MLP 9-10-1	Exponential	Identity	15.29
Without normalisation	Model 21	MLP 9-2-1	Tanh	Identity	14.71
	Model 22	MLP 9-2-1	Tanh	Identity	14.65
	Model 23	MLP 9-2-1	Tanh	Identity	12.97
	Model 24	MLP 9-2-1	Exponential	Identity	14.57

normalisation exhibited the lowest accuracy. Generally, it can be concluded that the ANN had better results when predicting the quantities of concrete, possibly because differences between structure designs were lower for concrete quantities. Designers of residential buildings often have vast experience and can easily predict dimensions of elements. Furthermore, they often rely on previous experience when adopting these dimensions. On the other hand, the quantity of reinforcement mostly depends on load parameters and is generated by an appropriate structural design software. Normally, dimensions of elements also play a very important role in this process, but small changes in concrete element dimensions can have a great impact on the quantity of reinforcement required for the structure. After analysing accuracy of ANN, the same experiment was applied on the SVM. In order to obtain a valid comparison, the same normalisations were used separately for concrete and reinforcement. Table 7 and Table 8 present MAPE for all 30 models of SVM, separately for concrete and reinforcement.

Out of 15 analysed models for predicting the quantity of concrete, the best results were obtained for model 6 with the lowest MAPE 9.28 % and with parameters C = 10, ε = 0.1 and $\gamma = 1/[(2\sigma)] \cdot 2 = 0.1$ formed on the z-score normalisation.

In the prediction of the quantity of reinforcement, the best results were obtained using the model 17 with MAPE 15.81 %

Table 7. MAPE of SVM models – estimation of required quantity of concrete

Type of normalisation	Model No.	C	ε	ν	γ	MAPE prediction of concrete quantity [%]
min-max	Model 1	10	0.7	-	0.1	9.30
	Model 2	10	0.1	-	0.05	9.75
	Model 3	10	0.1	-	0.07	9.31
	Model 4	10	-	0.5	0.01	12.63
	Model 5	20	-	0.7	0.01	13.38
z-score	Model 6	10	0.1	-	0.1	9.28
	Model 7	10	0.1	-	0.05	9.75
	Model 8	10	0.1	-	0.07	9.32
	Model 9	10	-	0.5	0.01	12.64
	Model 10	20	-	0.7	0.01	13.40
Without normalisation	Model 11	10	0.1	-	0.1	9.29
	Model 12	10	0.1	-	0.05	9.75
	Model 13	10	0.001	-	0.05	12.22
	Model 14	10	0.1	-	0.07	9.32
	Model 15	20	-	0.7	0.01	13.40

Table 8. MAPE of SVM models – estimation of required quantity of reinforcement

Type of normalisation	Model No.	C	ε	ν	γ	MAPE prediction of reinforcement quantity [%]
min-max	Model 16	10	0.1	-	0.05	18.10
	Model 17	10	0.1	-	0.01	15.81
	Model 18	20	0.1	-	0.007	16.25
	Model 19	20	-	0.1	0.01	16.64
	Model 20	10	-	0.5	0.009	16.28
z-score	Model 21	10	0.1	-	0.01	15.82
	Model 22	20	0.1	-	0.007	16.25
	Model 23	15	-	0.5	0.03	19.08
	Model 24	20	-	0.1	0.01	16.66
	Model 25	10	-	0.5	0.009	16.29
Without normalisation	Model 26	10	0.1	-	0.05	18.10
	Model 27	10	0.1	-	0.01	15.81
	Model 28	20	0.1	-	0.007	16.27
	Model 29	20	-	0.1	0.01	16.66
	Model 30	10	-	0.5	0.009	16.28

and parameters C = 10, ε = 0.1 and γ=1/[(2σ)] · 2 = 0.01 formed on the min-max normalisation, and the model 27 with the same values of 15.81 % with parameters C = 10, ε = 0.1 and γ = 1/[(2σ)] · 2 = 0.01 whose data were not normalized.

The SVM showed greater precision stability as the magnitude of difference between the best and the worst prediction results was smaller, compared to the value obtained using ANN.

Table 9. MAPE of ANN and SVM models - comparative overview

MAPE [%]	Type of normalisation	ANN	SVM
Prediction of concrete quantity	min-max	12.26	9.30
	z-score	8.74	9.28
	without normalisation	10.20	9.29
Prediction of reinforcement quantity	min-max	15.94	15.81
	z-score	12.58	15.82
	without normalisation	12.97	15.81

Table 9 shows comparison of results obtained using the ANN and SVM models. When analysing the use of ANN with regard to this specific problem, it can be concluded that the type of normalisation can have a significant impact on the accuracy of ANN. At the same time, it was established that SVM was immune to the use of the normalisation process.

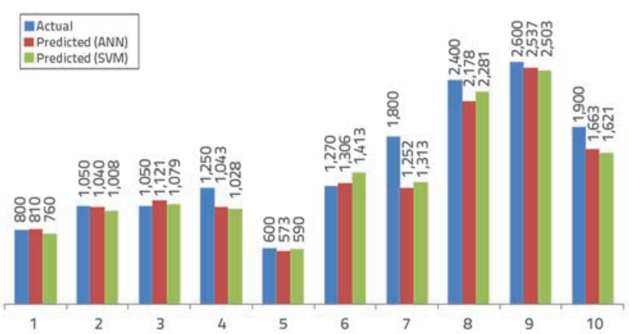


Figure 1. Comparison between predicted and the actual values for concrete of the best ANN and SVM models

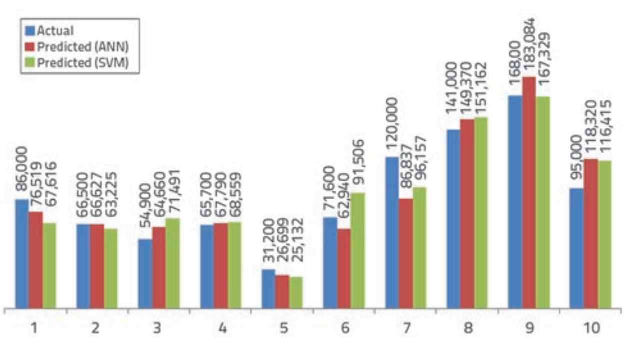


Figure 1. Comparison between predicted and the actual values for reinforcement of the best ANN and SVM models

When analysing the impact of normalisation, it can be concluded that the z-score normalisation exhibits the highest level of precision. ANN showed higher precision for prediction of quantities of both concrete and reinforcement. Figure 1 and Figure 2 show the comparison between the obtained (predicted) results and the real values for the concrete and reinforcement, for the models that give the best MAPE.

5. Conclusion

This paper shows how the ANN and SVM methods can be applied to predict the quantity of concrete and reinforcement based on the database of 100 projects. Several models with different parameters were formed, and the models with the most accurate results were selected. Models were formed based on nine input parameters describing characteristics of buildings, and two output parameters, the quantity of concrete and the quantity of steel reinforcement. Different parameters were chosen for building the ANN models: number of input and hidden neurons (from 1 to 20), type of activation functions, number of training networks (from 50 to 5000) and type of network (MLP). Two types of functions were chosen for building the SVM models: Type 1 in which parameters C and ε (from 0 to ∞) were varied, and Type 2, where parameters C and ν (from 0 to 1) were varied.

In the case of estimating the required quantity of concrete and reinforcement based on historical data, the best results were obtained for models with the lowest mean absolute prediction error. MAPE of the ANN model MLP 9-6-1 with 6 hidden neurons and 9 inputs formed on the basis of the data that are normalized by using z-score normalisation for the prediction of the quantities of concrete amounted to 8.74 %. For the prediction of the quantities of reinforcement, MAPE of the ANN model MLP 9-13-1 with 13 hidden neurons and 9 inputs formed on the basis of z-score normalisation amounted to 12.58 %.

After analysis, it was concluded that all models have the mean absolute error below 20 %, which is acceptable for this type of problem, i.e. for estimating recycling capacity of multi-storey residential buildings.

Future research should focus on analysing the database structure in full detail, as well as on increasing the number of parameters for the database. Adding new parameters will significantly improve accuracy of results regarding quantity of concrete and reinforcement. Also, it is possible to analyse the significance of input data on the accuracy of prediction, which may enable reduction in the number of input data, with potential increase in accuracy for the same quantity of data.

In addition, it is possible to expand the analysis with the application of other types of normalisation on historical data, which can significantly influence the accuracy of results. Although this paper analyses the reinforced concrete skeleton systems, this does not exclude the possibility of applying the analysis to other types of systems such as load bearing masonry structures, and prefabricated construction systems, as well as structures made of steel or wood.

Acknowledgement

The work reported in this paper is a part of the investigation within the research project TR 36017 ''Utilization of by-products and recycled waste materials in concrete composites in the scope of sustainable construction development in Serbia: investigation and environmental assessment of possible applications''. The project is supported by the ministry for Science and Technology, Republic of Serbia. This support is gratefully acknowledged.

REFERENCES

[1] Statistical Office of the Republic of Serbia, May (2017)

[2] The Economic Benefits of Recycling and Waste Reduction – WasteWise Case Studies from the Private and Public Sectors, Case Study, New Jersey, (2015), http://www.nj.gov/dep/dshw/recycling/wastewise/njwwcasestudy.pdf, March (2018)

[3] Peško, I., Trivunić, M., Ćirović, G., Mučenski, V.: A preliminary estimate of time and cost in urban road construction using neural networks, Technical Gazette, 20 (2013) 3, pp. 563-570

[4] DARPA, Neural Network Study, Massachusetts Institute of Technology (MIT), Lincoln Laboratory, USA, (1988)

[5] Vapnik, V., Golowich, S., and Smola, A.: Support vector method for function approximation, regression estimation and signal processing (eds. M. Mozer, M. Jordan, T. Petsche), Advances in Neural Information Processing Systems 9, Cambridge, MA, MIT Press, (1997), pp. 281-287

[6] Vapnik, V: ;The Nature of Statistical learning Theory, 2nd ed., Springer, New York, USA, (1999)

[7] Gunaydin, M., Dogan, Z.: A neural network approach for early cost estimation of structural systems of buildings, International Journal of Project Management, 22 (2004) 7, pp. 595-602

[8] Lazarevska, M., Knezevic, M., Cvetkovska, M., Trombeva-Gavriloska, A.: Application of artificial neural networks in civil engineering, Technical Gazette, 21 (2014) 6, pp. 1353-1359

[9] Lazarevska, M., Cvetkovska, M., Gavriloska, A.T., Kneževic, M., Milanovic, M.: Neural-network-based approach for prediction of the fire resistance of centrically loaded composite columns, Technical Gazette, 23 (2016) 5, pp. 1475-1480

[10] Mučenski, V., Trivunić, M., Ćirović, G., Peško, I., Dražić, J.: Estimation of recycling capacity of multi-storey buildings structures using artificial neural networks, Acta Polytechnica Hungarica, 10 (2013) 4, pp. 175-192

[11] Kim, G., Yoon, J., An, S., Cho, H., Kang, K., Kim, G.H., An, S.H., Kang, K.I.: Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning, Building and Environment, 39 (2004) 10, pp. 1235-1242

[12] Sonmez, R.: Conceptual cost estimation of building projects with regression analysis and neural networks, Can. J. Civ. Eng., 31 (2004), pp. 677-683

[13] Wang, Y., Gibson, E.: A study of pre project planning and project success using ANN and regression models, Automation in Construction, 19 (2009) 3, pp. 341-346

[14] Cheng, M., Wu, Y.: Evolutionary support vector machine inference system for construction management, Automation in Construction, 18 (2009) 5, pp. 597-604

[15] Strobbe, T., Wyffels, F., Verstraeten, R., De Meyer, R., Campenhout, J.: Automatic architectural style detection using one-class support vector machines and graph kernels, Automation in Construction, 69 (2016), pp. 1-10

[16] Cheng, M., Chou, J., Roy, A., Wu, Y.: High-performance concrete compressive strength prediction using time-weighted evolutionary fuzzy support vector machines inference model, Automation in Construction, 28 (2012), pp. 106-115

[17] Zhang, H., Yang, F., Li, Y., Li, H.: Predicting profitability of listed construction companies based on principal component analysis and support vector machine-Evidence from China, Automation in Construction, 53, pp. 22-28

[18] Cheng, M., Wu, Y., Wu, C.: Project success prediction using an evolutionary support vector machine inference model, Automation in Construction, 19 (2010) 3, pp. 302-307

[19] Wang, Y., Yu, C., Chan, H.: Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models, International Journal of Project Management, 30 (2012) 4, pp. 470-478

[20] Dell Statistica, http://en.community.dell.com/techcenter/information-management/statistica, November (2016)

[21] Bibić, D., Vujkov, A., Peško, I., Mučenski, V., Ćirović, G.: Estimation of recycling capacity of multi-storey building structures using ANN, Savremena dostignuća u građevinarstvu, (2017), pp. 921-929, 10.14415/konferencijaGFS2017.096

[22] Vujkov, A., Bibić, D., Peško, I., Mučenski, V., Ćirović, G.: Estimation of recycling capacity of multi-storey building structures using support vector machines, 1st International Conference on Construction materials for sustainable future - CoMS 2017, pp. 816-820, Zadar, 19-21 April 2017.

[23] Venugopal, V., Sundaram, S.: An online writer identification system using regression-based feature normalisation and codebook descriptors, Expert system with application, 72 (2017), pp. 196-206.

[24] Matignon, R.: Neural Networks Modeling using SAS Enterprise MIner, (2005), ISBN: 1-4184-2341-6

[25] Kecman, V.; Learning and Soft Computing, Support Vector Machines, Neural Networks and Fuzzy Logic Models, Massachusetts Institute of Technology (MIT), USA, (2001)

[26] Schölkopf, B., Smola, A., Williamson, R.C., Bartlett, P.L.: New support vector algorithms, Neural Computation, 12 (2000), pp. 1207-1245.