www.sjm06.com

Serbian Journal of Management 4 (2) (2009) 143 - 155

Serbian Journal of Management

ARTIFICIAL NEURAL NETWORK METHOD APPLIED ON THE NONLINEAR MULTIVARIATE PROBLEMS

Živan Živković*, Ivan Mihajlović and Đorđe Nikolić

University of Belgrade, Technical Faculty in Bor, Department of Management, Vojske Jugoslavije 12, 19210 Bor, Serbia

(Received 5 August 2008; accepted 28 September 2008)

Abstract

This paper presents the research of the possibility to apply Artificial Neural Network (ANN) for solving nonlinear multivariate problems of the technological processes. After concurrent analysis facilitated using multivariate regression analysis (MRA) and ANN, on the same set of data, aiming to determinate the copper content in the waste slag depending on its chemical composition, following values of R^2 were obtained: 0.09 and 0.999. These results indicate that the ANN is more appropriate for solving the nonlinear problem, of this technological process at the industrial level, described in this paper.

*Keywords:*multiple linear regression, artificial neural networks, copper content in the slag

1. INTRODUCTION

For the mathematical modeling of the technological processes where influence of the input (X_i) on the output (Y_i) values is to be measured, multiple linear regression analysis (MLRA) and the nonlinear regression analysis (NRA) are most often applied. According to the obtained analytical models, with the certain probability, the Xi values could be used for managing and control of the Yi values. In the case of complex technical technological processes, R^2 of such models is relatively small, resulting with unreliable prediction of the output values. At the same time, if there is more than one output value, both MRA and the NRA are resulting with limited possibilities of use (Liu, D, et al, 2009; Shelgani and Jorjani, 2009; Aldrich et al., 1994).

In the last decade, artificial neural networks (ANN) have emerged as attractive tools for nonlinear process modeling,

* *Corresponding author: zzivkovic@tf.bor.ac.rs*

especially in the situation where the development of the phenomenological or conventional regression models becomes impractical or cumbersome. The ANN is a computer modeling approach that learns from examples through iterations without requiring a prior knowledge of the relationships of process parameters and is, consequently, capable of dealing with uncertainness, noisy data, and non-linear relationships (Baughman and Liu, 1995).

Neural networks are computer algorithms inspired by the way the information is processed in the nervous system. An ANN is a massively parallel-distributed processor that has a natural propensity for storing experimental knowledge and making it available. An important difference between neural networks and standard information technology solutions is their ability to learn (Dreyftus, 2004). This learning property has yielded a new generation algorithms. An ANN paradigm is composed of a large number of highly interconnected processing elements, analogous to biological neurons that are tied together with weighted connections that are analogous to synapses. Learning in biological systems involves adjustments to the synaptic connections between neurons. This is true for ANNs as well. Learning typically occurs through training or exposure to a true set of input/output data where the training algorithm iteratively adjusts connection weights. These connection weights represent the knowledge necessary to solve a specific problem (Eberhart and Dobbins, 2002).

2. ARTIFICIEL NEURAL NETWORK ARCHITECTURE

The ANN used in the model development is depicted in Fig.1. As shown, the network usually consists of three layers of nodes. The layers described as input, hidden and

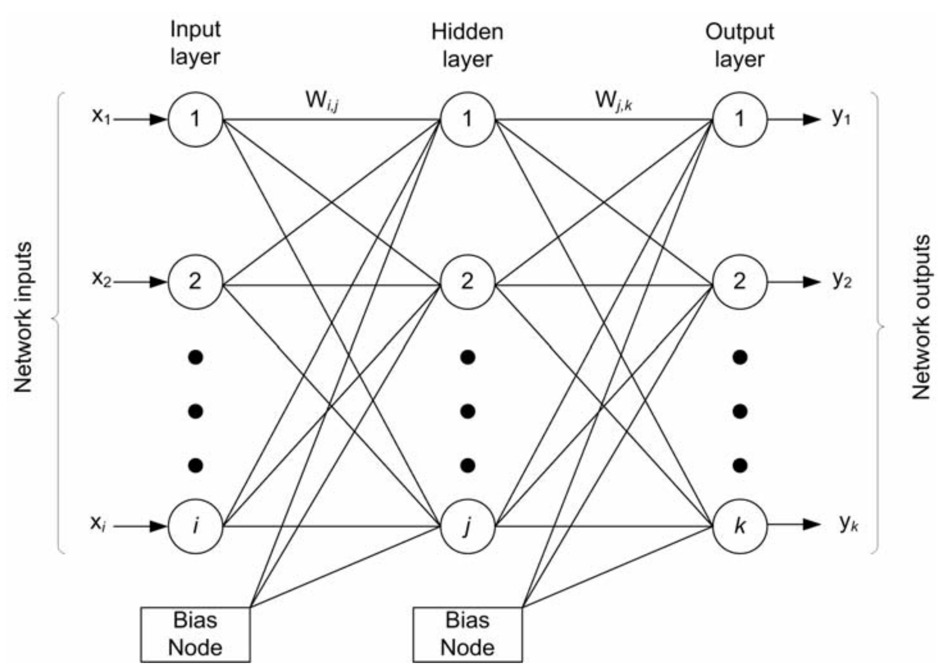

Figure 1. A schematic diagram of ANN with one hidden layer

output i, j and k number of processing nodes, respectively. Each node in the input (hidden) layer is linked to all the nodes in the hidden (output) layer using weighted connections. In addition to the i and j number of input and hidden nodes, the ANN architecture also houses a bias node (with fixed output + 1) in its input and hidden layers; the bias nodes are also connected to all the nodes in the subsequent layer and they provide additional adjustable parameters (weights) for the model fitting. The number of nodes (i) in the ANN network input layer is equal to the number of inputs in the process whereas the number of output nodes (k) equals the number of the process outputs. However, the number of hidden nodes (j) is an adjustable parameter magnitude of which is determined by issues, such as the desired approximation and generalization capabilities of the network model (Zeng and Chen, 1997; Dreyftus, 2004).

The back propagation algorithm modifies network weights to minimize the mean squared error between the desired and the actual outputs of the network. Back propagation uses supervised learning in which the network is trained using data for which the input, as well as desired outputs are controlled and selected (Eberhart and Dobbins, 2002).

The use of ANN usually comprises three phases. First is the training phase which is facilitated on 70 to 80% randomly selected data from the starting data set. During this phase the correction of the weighted parameters of the connections is achieved through necessary number of iterations, until the mean squared error between the calculated and the measured outputs of the network is minimal. During the second phase, the remaining 20 to 30% of the data is used for testing of the „trained“network. In this phase, the network is using the weighted parameters determined during the first phase. The new data, excluded during the learning of the network, are now incorporated in it as new input values Xi which is then transformed to the new outputs Yk. The evaluation of the network is obtained after calculation of the error between the network outputs and the real values obtained by the measurements. The third phase is the validation of the network on the new data set. This data set is consisting of already measured or data from new experimental measurements. The validation phase is presenting the final level of success or unsuccess in predictions using the network developed in the previous two stages, on future datasets.

3. APPLYING THE ANN FOR PREDICTION OF THE COPPER CONTENT IN THE SLAG – A CASE STUDY

During the pyrometallurgical procedure for copper extraction, the crucial phase of the process is smelting operation which results with two products: waste slag and the copper matte. The copper matte is further processed during the technological process of copper extraction (Habachi, 2007). Regardless the present phase in the life cycle of the technology applied (Živković, D., and Živković, Ž, 2007), for the global problem of the slag deposition, local solutions should be seek out (Parnel, 2006). Analytical dependences of the slag and matte composition influencing the copper content, permanently lost with the waste slag, are presented in the literature (Živković, et al., 2009). These dependences were defined using the MLRA methods.

The copper content in the waste slag, and also the slag and matte composition, resulting from the operation of the Bor copper smelter, for the time period 2003-2009, are presented in the Table 1.

During the previous investigations, multiple linear dependences were defined for the copper content in the slag on its composition and the copper content in the matte (concerning all the data from the table 1). The form of this dependence can be expressed as (Živković, et al. 2009):

$$(Cu) = a + b(SiO_2) + c(FeO) + d(Fe_3O_4) + e(CaO) + f(Al_2O_3) + g[Cu] \quad (1)$$

where: a,b,c,d,e,f and g are coefficients of the linear regression equation, while ( ) and [ ] are concentration of the components in wt. %.

Using the MLRA methodology the coefficients of the equation (1) were determined. This way the equation obtains the following form:

$$(Cu) = - 0.728 + 0.031(SiO_2) + 0.001(FeO) + 0.016(Fe_3O_4) - 0.031(CaO) + 0.034(Al_2O_3) + 0.002[Cu] \quad (2)$$

Above dependence is with the coefficient of determination equal to $R^2 = 0.09$, which indicates very low fitting result. Figure 2, presents the dependence between calculated and experimentally measured values, which illustrate high deviation from the ideal position. Considering this low level of the fitting obtained using MLRA, the data from the Table 1 was subjected to the ANN method.

To apply the ANN method on the data

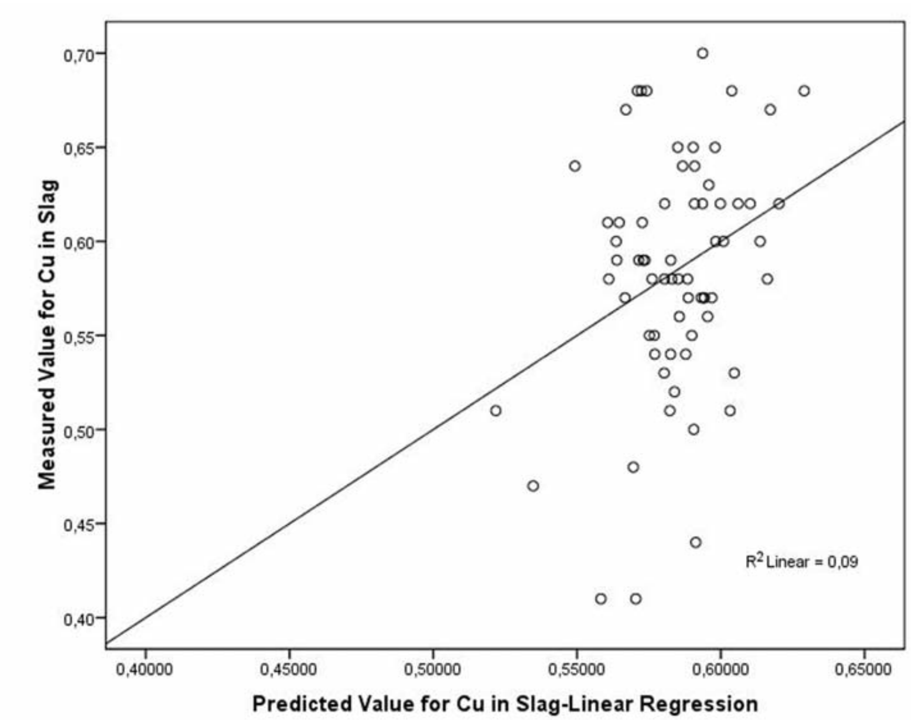

Figure 2. Dependence between calculated and measured values of the copper content in the waste slag (MLRA method)

Table 1. Composition of the slag and the matte for the time period of January 2004 to July 2009

Sample No.	Cu in the slag, %	SiO_2 in the slag, %	FeO in the slag, %	Fe_3O_4 in the slag, %	Al_2O_3 in the slag, %	CaO in the slag, %	Cu in the matte, %
1	.65	34.89	51.09	7.46	4.06	2.84	39.38
2	.59	33.85	50.58	7.87	4.48	4.03	41.62
3	.53	34.15	49.58	7.73	4.63	3.68	40.50
4	.51	34.83	49.94	7.69	4.77	3.68	41.16
5	.51	34.91	49.04	7.91	4.68	2.98	42.16
6	.57	33.77	52.02	8.17	5.34	2.56	42.80
7	.59	32.84	50.15	8.11	4.53	3.33	39.20
8	.62	32.95	47.56	8.72	5.43	3.26	42.25
9	.51	34.76	50.29	6.24	3.31	4.95	31.90
10	.61	32.79	51.37	7.20	3.31	2.85	33.08
11	.62	34.43	49.90	7.55	5.23	3.77	45.12
12	.58	34.70	50.29	7.59	4.57	3.29	44.58
13	.55	33.97	52.09	7.74	5.10	3.36	41.64
14	.60	32.37	53.67	7.80	7.56	3.01	39.88
15	.65	33.83	51.48	8.05	4.82	2.36	40.52
16	.64	35.20	50.04	7.02	5.52	2.64	40.91
17	.41	33.56	52.63	7.74	5.72	4.07	41.00
18	.61	32.70	49.93	7.85	6.38	3.70	40.80
19	.62	34.83	50.29	8.26	6.28	3.96	44.76
20	.59	35.25	47.24	6.89	5.62	4.15	33.82
21	.61	33.96	49.47	7.19	6.52	4.07	42.86
22	.58	34.10	44.91	6.84	6.61	4.14	29.28
23	.68	33.75	49.36	8.21	7.65	4.68	41.76
24	.47	34.35	48.71	6.32	7.94	4.70	36.48
25	.67	35.39	46.34	8.07	5.00	3.12	41.56
26	.60	34.02	47.42	8.14	3.74	3.47	40.82
27	.60	33.42	44.55	8.37	4.91	3.57	42.12
28	.68	32.27	43.72	8.33	4.12	2.49	42.96
29	.64	35.24	47.42	7.44	4.48	3.65	41.22
30	.64	34.42	50.12	6.94	4.19	4.61	40.48
31	.59	36.93	49.22	7.45	3.93	4.85	39.50
32	.50	34.81	48.57	7.73	5.80	3.61	42.24
33	.57	34.86	48.50	7.66	4.86	3.79	43.38
34	.56	34.97	49.39	8.19	5.01	3.88	44.32
35	.58	33.92	49.89	8.08	5.20	4.22	44.62
36	.58	34.33	49.58	8.79	3.40	3.24	46.16
37	.62	32.34	51.23	8.88	5.21	3.20	45.84
38	.62	33.75	52.81	8.62	5.76	2.57	46.38
39	.60	32.82	49.58	8.72	5.65	3.46	42.74
40	.58	34.88	51.34	7.36	4.99	3.11	44.44
41	.63	34.28	51.23	8.15	4.38	3.21	45.64
42	.62	35.20	47.46	8.54	4.06	3.49	45.94
43	.65	35.73	53.96	8.36	4.08	3.25	41.08
44	.53	39.68	51.37	7.07	3.78	2.82	41.70
45	.57	34.54	48.50	7.51	4.28	3.19	43.62
46	.57	35.05	50.83	7.94	6.20	3.41	45.52
47	.67	34.12	57.54	8.17	4.39	3.19	43.80
48	.70	35.06	48.93	7.68	4.61	3.37	42.10
49	.68	34.58	51.19	7.34	4.39	3.44	39.52
50	.68	34.83	52.92	7.44	4.44	2.97	39.28
51	.57	33.24	53.77	7.36	4.24	2.54	38.38
52	.52	34.62	50.94	7.48	4.63	3.17	43.60
53	.55	34.86	51.26	7.89	4.36	3.26	43.06
54	.58	34.93	53.10	7.84	3.61	3.20	45.06
55	.56	34.57	50.97	7.76	4.27	3.06	38.92
56	.44	34.92	50.58	7.54	4.92	2.67	39.44
57	.58	33.62	51.77	8.24	5.30	3.35	34.30
58	.62	33.30	51.70	8.77	4.66	3.44	42.72
59	.54	33.74	51.19	7.82	5.02	3.55	42.26
60	.54	33.08	54.46	8.86	4.58	3.53	43.90
61	.59	32.72	53.13	8.52	4.77	3.80	42.56
62	.68	37.60	51.11	8.19	4.20	3.38	38.48
63	.55	34.64	50.68	7.62	4.36	3.40	37.08
64	.48	33.99	51.80	7.70	4.89	3.48	37.86
65	.57	32.18	51.37	8.76	4.46	2.89	39.54
66	.54	34.18	51.58	8.37	4.62	3.24	37.16
67	.41	32.85	51.22	8.04	4.64	3.73	39.32

presented in the Table 1, this complete set of data was divided into two groups. First group contained first 46 (68.7%) values and it was subjected to the training of the network, while the second group contained the remaining 21 (31.3%) which were used for testing of the network.

During creating of the artificial neural network, using the data from the table 1, the input parameters (X_i) were (SiO_2), (FeO), (CaO), (Fe_3O_4), (Al_2O_3) and [Cu] values, while the output parameter (Y_k) is the copper content in the waste slag (Cu). After selecting the network architecture with only one hidden layer, resulting ANN was developed for this case, which is presented in Figure 3.

The ANN presented in the Figure 3 is consisted of three layers: the input layer, the output layer and hidden layer. The neurons in the input layer take the information on slag composition and the Cu content in the mat, X_i (independent variables), and the output layer generates the outcomes of copper content in slag Y_k (dependent variable)

The input to any neuron i without its bias is given by:

$$I_j = \Sigma W_{ij} \cdot X_j \tag{3}$$

Where $W_{i,j}$ are the weights of interconnects between neuron *i* and *j*, X_j represent the signal at the connection concerned.

Important component of the ANN is its activation functions appearing after the input layer. Each hidden node and output node applies the activation function to its net input. Five types of the activation function, reported in literature, are shown in Table 2.

For the case which is concerned in this paper, as the activation function the log sigmoid one was chosen. This function is most often used for modelling of the similar systems (Aldrich, et al., 1994; Zeng and Chen, 1997; Meradi et.al., 2006):

$$f(x) = 1 /(1 + e^{-x}) \tag{4}$$

The overall transfer function of a neuron is thus structured as follows:

Table 2. Different activation functions

Case	Name of the activation function	Equation
1	Log sigmoid function (*logsig*)	$Y_i = 1/ (1 + \exp(- \textit{net } i)$
2	Tan hyperbolic function (*tansig*)	$Y_i = \tanh(\textit{neti})$
3	Linear function (*purelin*)	$Y_i = (\textit{neti})$
4	Radial basis function (*radbas*)	$Y_i = \exp[-(\textit{neti})^2]$
5	Triangular basis function (*tribas*)	$Y_i = 1 - \text{abs}(\textit{neti})$ if $-1 \leq (\textit{neti}) \leq 1$ otherwise, $Y_i = 0$

$$O_j = A_j = f(\Sigma\, W_{i,j} \cdot X_j) \quad (5)$$

In equation (5) O_j is the output of the neuron, Aj is its activation, X_j is input to the hiden layer neuron which is identical to the output of the preceding neuron, with index j of the observed element.

The aim of the learning process is to minimize the global network error:

$$E = ½\ \Sigma\, (y_j - O_j)^2 \quad (6)$$

where y_i are the target output values.

Adaptation of the weights is effected according to the equation as follows (Meradi, et al., 2006):

$$\Delta W_{ij} = W_{ij} \cdot (t + 1) - W_{ij}(t) = - \alpha \cdot \partial E / \partial W_{i,j} \quad (7)$$

Where the a is defined as the learning rate, this results in:

$$\Delta W_{ij} = \alpha \cdot \beta j \cdot X_j \quad (8)$$

Where the local error of a hidden element is calculated via:

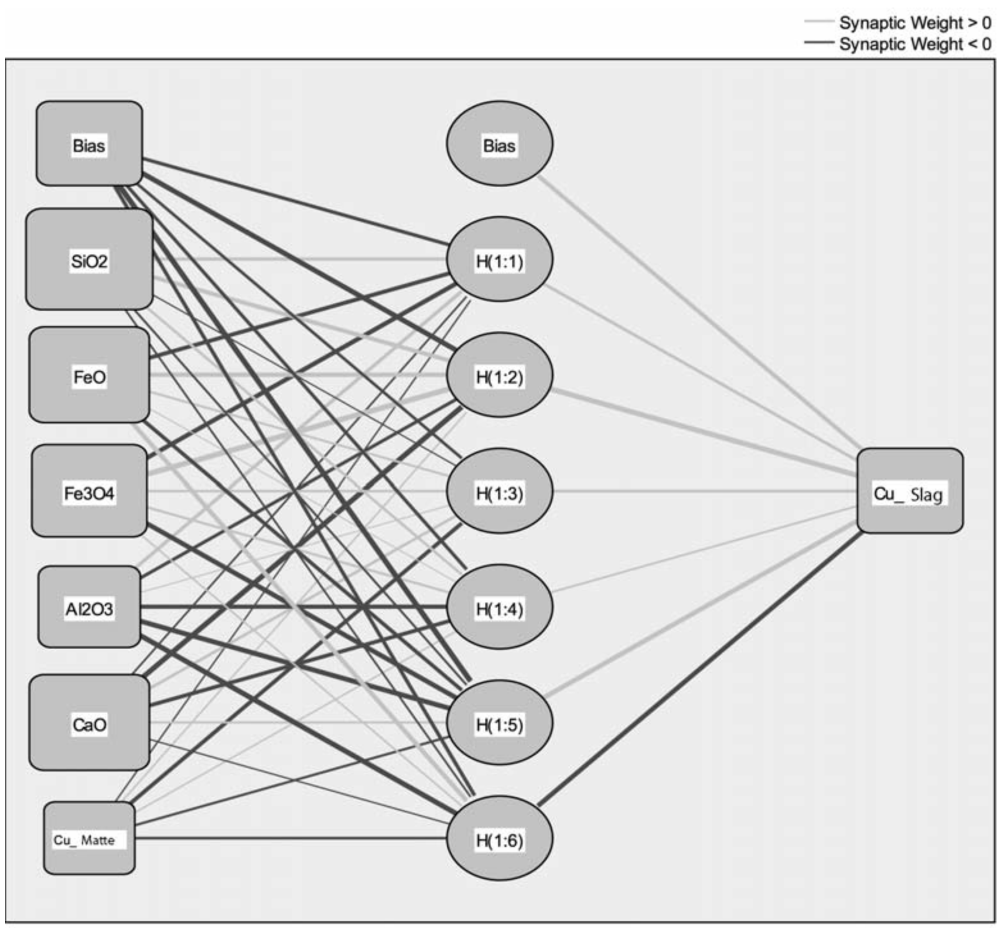

Figure 3. The ANN architecture for determination of the copper content in the smelting operation slag

$$\beta_j = f(I_j) \cdot \Sigma \beta_k \cdot W_{j,k} \quad (9)$$

The β_k components represent the errors of the elements in the following layer, while Wij represent the connection weights for these elements. The error of a neuron of the output layer is obtained via:

$$\beta_k = f(I_k) \cdot (y_k - O_k) \quad (10)$$

The error is first of all calculated and then back propagated into the hidden layer located before the output layer.

The connection weights can then be modified according to the calculated ΔW_{ij} in the concluding stage of this process.

The resolve a problem of local minimum of the error space, we have introduced a *momentum* term. The equation for adoption of the weight is modified as follows:

$$\Delta W_{ij}(t) = \alpha \cdot \beta_j \cdot X_j + \mu \cdot \Delta W_{ij} \cdot (t - 1) \quad (11)$$

μ is defined as the momentum, t is the current learning step and $(t - 1)$ the previous learning step.

The backpropagation training algorithm is an iterative gradient algorithm designed to minimize the main square error between the predicted and the desired output. The algorithm of backpropagation training, used in this study, is summarized in Fig.4 for the

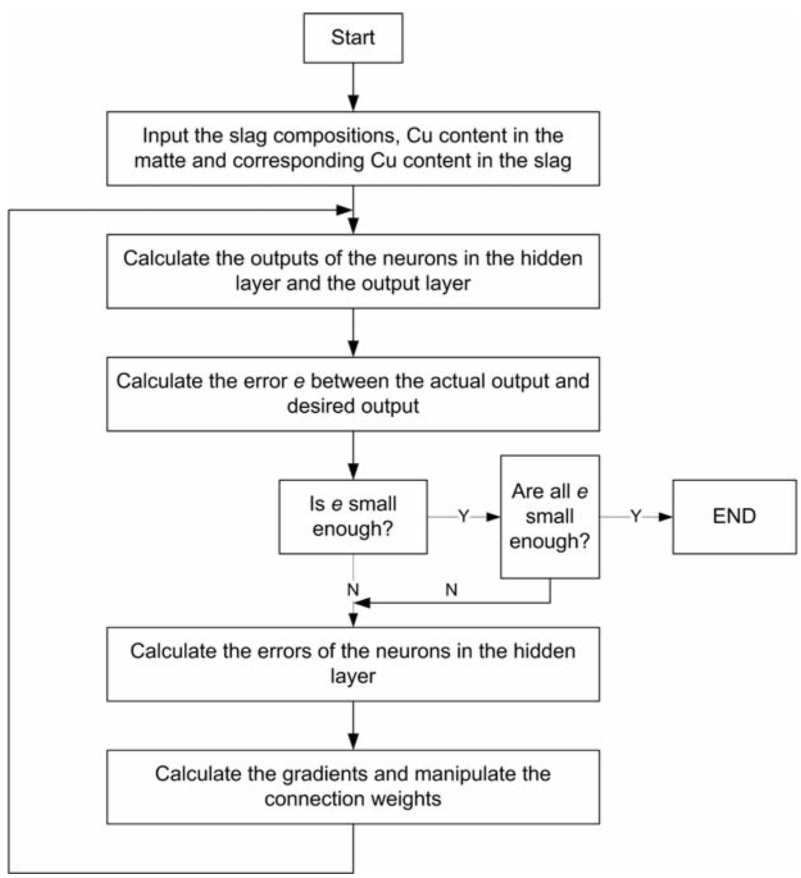

Figure 4. Flow chart of the backpropagation learning algorithm

benefit of the copper losses in slag during the sulfide smelt process.

4. DISCUSSION OF THE RESULTS

As the first stage of the ANN development, the training of the network was performed on the data from the table 1, from which as the randomly selected sample 46 vectors were included (68.7% of total sample). For the testing stage of the network, remaining 21 vectors (31.3%) were used. In the phase of training of the network, necessary number of iterations was performed, until the error between the measured copper content in the slag and calculated values wasn't minimized and remained constant. Obtained results from the training stage can be evaluated by comparison of the calculated values of copper content in the slag with measured ones, Figure 5.

Obtained coefficient of determination (R^2 = 0,997) is presenting large scale of fitting among calculated and measured values, which indicates that the ANN network was well prepared during the training phase and can be used in subsequent testing and validation.

After developing of this kind of "trained" ANN network, testing stage was performed using the data from the table 1 (total of 21 vectors). During the ANN testing phase, calculated values of the copper content in the, smelting operation, waste slag were compared to the measured ones. Statistical value of the correlation coefficient (R^2) was slightly increased in comparison to the training phase and now it equals 0.999.

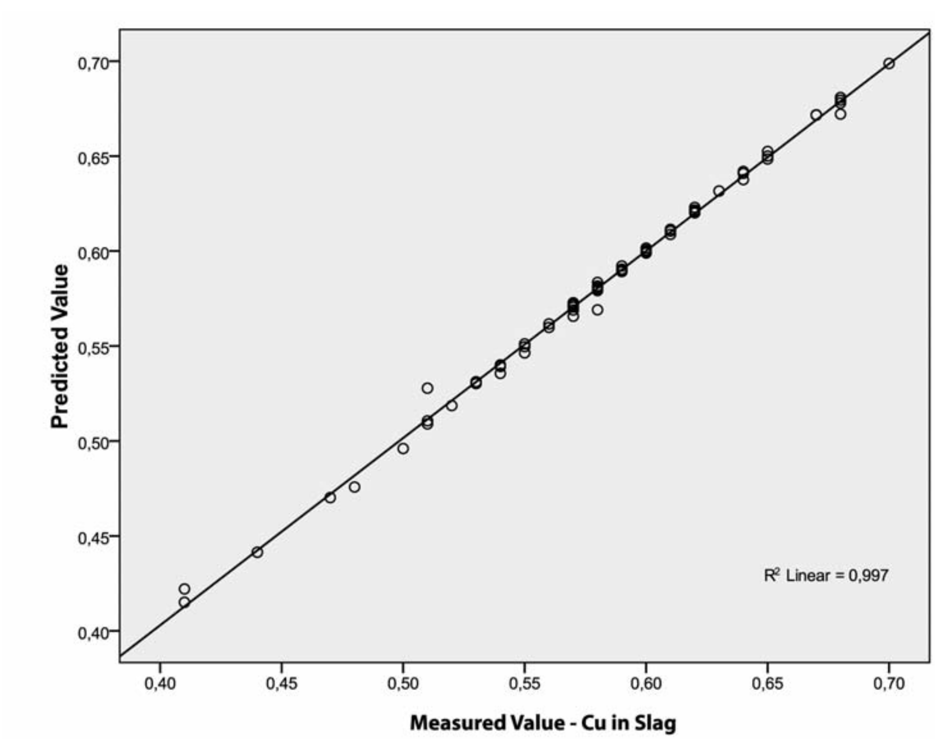

Figure 5. Relationship among measured values of copper content in the slag and values obtained after ANN calculations during the training stage

Figure 6 presents comparative presentation of the measured and the values calculated using both the MLRA and the ANN approach. Obtained values of the R^2 coefficients are 0.09 and 0.999, respectively. The results presented in the Figure 6, reveals that use of the ANN is much more appropriate for prediction of the copper content in the smelting operation slag, depending on its chemical composition and content of the copper in the matte.

Obtained results for the nonlinear correlation, defined using ANN methodology, enables ranking of each individual parameters degree of significance on influencing the output result (copper content in the waste slag), Figure 7. The significances of influence of the waste slag components are: SiO_2 – 0.254; CaO – 0.228; FeO – 0.226; Fe_3O_4 – 0.182; Al_2O_3 – 0.105 and Cu_{matte} – 0.04. Obtained results could be quite important for managing the smelting operation, concerning minimizing or optimization of the copper content in the waste slag. Obtained results indicate that the degree of significance, of individual slag compounds influencing its copper content, is decreasing according to the following string: (SiO_2) → (CaO) → (FeO) → (Fe_3O_4)→ (Al_2O_3) → [Cu].

5. CONCLUSIONS

This paper presents comparative analyses of two statistical techniques applied on the same problem. One of the techniques is artificial neural network and the other is multiple regression analysis. Both were applied for predicting of the copper content

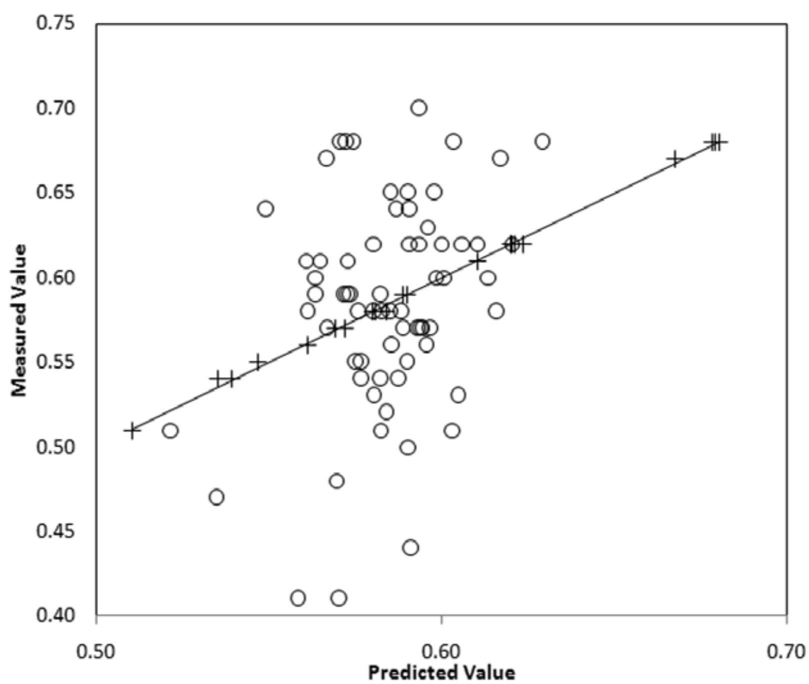

Figure 6. Comparison of the measured and the values calculated using ANN and MLRA for prediction of the copper content in the slag (o - MLRA; x – ANN ; — ideal)

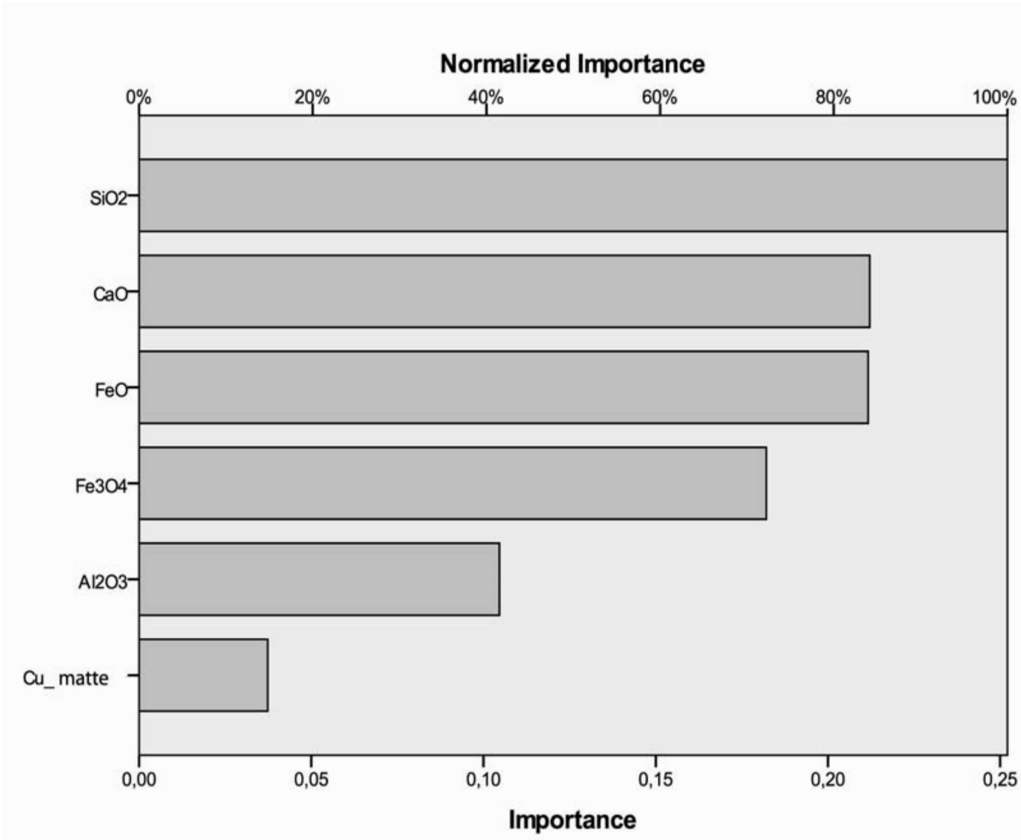

Figure 7. Degree of significance of the individual compounds influence on the copper content in the slag

in the reverberatory furnace slag, depending on its chemical composition and the copper content in the matte. The investigations presented in this paper revealed that the artificial neural networks can be successfully applied for prediction of the copper content in the slag, if its composition and the copper content in the matte are known. Comparing ANN technique to the traditional multiple linear regression analysis indicated that the value of the resulting analytical error is much smaller in the case of ANNs and, in some cases, even could be eliminated. This paper shows that copper content in the smelting slag, depending on its chemical composition and content of the copper in the matte, can be predicted very precisely using ANNs. Considering that the ANNs have great capacity and relatively simple working algorithm, they can be applied for solving large number of nonlinear problems. These results also indicated that this tool can be used for managing the copper content in the slag, since the slag composition can be controlled in the stage of smelting charge preparation as well as degree of desulfurization obtained during oxidative roasting of the copper concentrate. At the same time, ANN could be very useful tool for training on other data sets obtained from industrial production practice in the metallurgical facilities - in general, as well as for large number of the processes in the chemical technology.

ПРИМЕНА ВЕШТАЧКИХ НЕУРОНСКИХ МРЕЖА НА НЕЛИНЕАРНЕ МУЛТИВАРИЈАНТНЕ ПРОБЛЕМЕ

Живан Живковић, Иван Михајловић и Ђорђе Николић
Универзитет у Београду, Технички факултет у Бору, Одсек за менаџмент, Војске Југославије 12, 19210 Bor Serbia

Абстракт

У раду се презентирају могућности примене артифицираних неуронских мрежа - ННА као алата за решавање нелинеарних мултиваријантних проблема у технолошким процесима. Упоредном анализом примене мултиваријантне регресионе анализе- МРА и ННА на исти скуп података за случај одређивања садржаја бакра у отпадној шљаци у зависности од њеног хемијског састава добијене су следеће вредности за R^2: 0,09 и 0,999 . Ови подаци указују на веће погодности примене ННА за решавање нелинеарних проблема у технолошким процесима у индустријским условима.

Кључне речи: мултилинерна регресија, артифициране неуронске мреже, сдржај бакра у шљаци

References

Aldrich, C., Van Deventer, J.S.J., Reuter, M.A.,(1994) The application of neural nets in the metallurgical industry, Minerals Engineering, 7(5/6) : 793 – 809.

Baughman, D.R., Liu, Y.A., (1995) Neural networks in bioprocesing and chemical engineering , Academic Press, New York.

Chelgani, S.C., Jorjani, E., (2009) Artificial neural network prediction of Al_2O_3 leaching recovery in the Bayer process – Jajarm alumina plant (Iran), Hydrometallurgy, 97 105 – 110.

Cilek, E.C.,(2002) Application of neural network to predict locked cycle flotation testresults, Minerals Engineering, 15 : 1095 – 1104.

Dreyfus, G. (Edd.) (2004)Neural Networks, Methodology and Applications, Springer- Verlag, Berlin , 2004.

Eberhart, R.C., Dobbins, R.W., (2002) Neural network PC tools, A Practical guide, Academic, New York, 2002.

Habashi, F (2007) Copper metallurgy at the crossroads, Journal of mining and metallurgy, Section B: Metallurgy, 43(1)B: 1 – 19.

Liu, D., Yuan., Y., Liao, S.,(2009) Artificial neural network vs. Nonlinear regression for gold content estimation in pyrometallurgy , Expert Systems with Applications, 36 : 10397-10400

Meradi, H.,S. Bouhousche, S., Lahreche, M., (2006) Prediction of bath temperature using neural networks, Proceedings of world academy of science, Engineering and Technology, 17 : 319 – 323.

Parnell, J., (2006) Reassessing the ''think global, act local'' mandate: evaluation and synthesis, Serbian Journal of Management, 1(1): 21 – 28.

Zeng, W., Chen, N., (1997) Artificial neural network method applied to enthalpy of fusion of transition metals, Journal of Alloys and Compounds, 257: 266 – 267.

Živković, Ž.,Mitevska, N., Mihajlović, I., Nikolić, Dj., (2009) The influence of the silicate slag compositionon copper losses during smelting of the sulfide concentrates, Journal of Mining and Metallurgy, Section B: Metallurgy, 45(1)B: 23-34.

Živković, D., Živković, Ž., (2007) Investigation of the influence of technology life cycle on company life cycle Case study: Metallurgical production of copper in RTB Bor (Serbia), Serbian Journal of Management, 2(1): 57 – 65.