Research Journal of Engineering & Technology. Vol., 1 (2), 53-60, 2024
Available online at http://www.rjetjournal.com
ISSN 2149-3340 ©2024

Application of Artificial Neural Network in Modeling Removal of Heavy Metals from Water through Phytoremediation

Abbas Khoshhal*, Fereshte Nazemi Harandi, Amir Hossein Malek

Department ofChemical Engineering, Payame Noor University (PNU), P.O. Box, 19395-3697, Tehran, Iran

*Corresponding Author Email: abbas_khoshhal@yahoo.com

Abstract: Heavy metals are one of the environmental pollutants and human's contact with some of them can cause chronic and sometimes acute poisoning. Among the methods used to clean up contaminated sites, phytoremediation is considered as a cost-effective and biocompatible option. Phytoremediation refers to a group of technologies that use plants to reduce, remove, analyze and consolidate environmental toxins. The modeling of phytoremediation processes can be used as a safe tool to save the economy and avoid repeated testing. Among the methods of modeling, artificial neural network is accurate and widely used in biotechnological processes. Results of the study showed that correlation coefficient reached 1 indicating that there is a good match between actual values and those predicted by modeling. From a total of 30 data, two-thirds of data was used to train a network and the remaining third was used to assess the modeling accuracy. The middle transition function tansig, output transfer functionpurelinand the number of neurons (one) were determined as the best parameters to train the network. The error rate of network training was estimated to be 0.0409 and error evaluation of the network accuracy was found to be 0.8992.

Keywords: phytoremediation, heavy metals, modeling, artificial neural networks, regression.

Introduction

Due to the discharge of urban and industrial wastewater, toxic heavy metals and improper management of waste, water pollution affected human health dangerously. The most common metals found in wastewater are lead, copper, cadmium,

chromium and nickel. Of these, lead is the most widespread heavy and toxic element in the environment in terms of emissions that has been widely distributed particularly throughout the world since its use in gasoline.

With regard to the above-mentioned discussion and due to the high consumption of copper and lead in today'sworld and consequently inevitable contamination of these metals in water resources and wastewater industries, it is crucial to make decisions to remove copper and water-soluble lead.

Various methods of filtration, reverse osmosis (RO), chemical oxidation, activated carbon, electrochemical treatment and other methods are used to remove heavy metals from wastewater of industries. However, the methods are inefficient and costly when metal concentration is low that may produce secondary waste that is difficult to treat. Therefore, it is necessary to find new technologies or new materials in order to remove heavy metals from wastewater.

Therefore, the use of physical and chemical methods to remove contaminants is both costly and inefficient due to contradiction with the laws of nature or even is harmful. However, the use of a new technique called phytoremediation is economically practical and environmentally very safe and can solve environmental pollution. Phytoremediation is a process that reduces pollutants in the water or wastewater using green plants (including, herbaceous and woody species). The process is used to remove environmental pollutants such as heavy metals, rare elements, organic compounds and radioactive materials. In fact, the plants reduce the level of pollutants based on adsorption mechanism using chemical, physical and biological methods. The process uses green plants and their relationship with soil microorganisms to reduce water pollution and wastewater. This technology can be used to remove both organic and inorganic contaminants.

Thus, in order to resolve the problem effectively, the study and modeling of biological systems require methods that are very close to human's thinking. Artificial neural network is one the methods used in this context. Artificial neural networks can model complex systems with nonlinear characteristics that have become the most popular tool for modeling biological processes.

A limited number of studies have been conducted on the application of modeling genetic algorithm- neural network to model and control micro-biological processes. However, some studies have been reported by Fernandez et al (2007) who used artificial neural networks to examine microbial growth and introduced the method as a convenient one for microbial prediction (Athar & vohora, 2007). Garcia et al (2009) modelledleuconostocmesenteroides growth using artificial neural network. Predictive models of the growth of the artificial neural network have been introduced as one of the proper tools to estimate growth parameters of leuconostocmesenteroides (Venkata, 2009). Sofo and Eckenjie (2007) studied estimation of the shelf life of yogurt with artificial intelligence model (Sharifi Arab et al., 2013). Another example of application of artificial neural network in biotechnology concerns with the presentation of a computerized model to predict the shelf life of lactobacillus acidophilus in probiotic yogurt (Nasiri et al., 2014).

This paper modelled removing heavy metals from water using artificial neural network and compared experimental results with data obtained from the modeling.

Theory

In neural network modeling that is derived from the human brain performance, knowledge or rules behind data is transferred to the network structure by processing experimental data. In other words, general laws are trained to the network by performing calculations on numerical data. The resulting network will be able to model and predict very complex processes with respect to the training provided with it. For instance, this method is used to detect hand-written notes and convert voice (speech) into handwriting (text).

The human brain contains a large number of (about 1011) interconnected components (104 connections per component) that is called neurons. The neurons include three main components: Dendrites, Cell body, and Axon.

Dendrites are tree-like receptor networks, including nerve fibers that transmit electrical signals into Cell body. Cell body gathers these input signals effectively and places them on thethreshold. Axon is a long sequence that transmits signal from Cell body to other neuron. Point of contact between Axon, a Cell body and another Dendrite is called Synapse. Efficiency of neural network is determined by arranging neurons and strength of Synapses, which are determined by a complex chemical process.

Each artificial neuron consists of one or more inputs, weight (for multi-input neuron, multiple-weight), bias or offset, collector, activation or transfer function) and one output. Scalar input p is multiplied by scalar weight w to form wp that is one of the terms that will be sent to the collector. Another input, 1, is multiplied by bias, b and sent to the collector. Collector output n that is usually called input net is sent to transfer function f that produces neuron scalar output a. If we compare this model with biological neuron, the weight w will relate to the strength of Synapse, Cell body will be expressed by collector and transfer function and neuron output will express the signal on the Axon.

An artificial neural network consists of several layers of neurons. Input data enters input layer. Outputs of the layer form the input of the next layer, and output of the last layer is the network output. Each layer is comprised of a number of neurons. The layer whose outputs are the network output is called output layer. Other layers are called hidden layers (Debska & Guzowska, 2011).

Test Description

Absorbent used in the study is the "platanus fruit" that was rinsed by distilled water in order to remove impurities and increase the uptake sites (Rezaei et al., 2013). Then, it was dried in oven at 80 °C for 72 hours, ground with electric mill after drying and turned into fine particles. The prepared ground fruits were classified by sieves into three groups: <120, 150 -250, > 300 mm. All chemicals used in the study were purchased from Merck

Company (Germany). Metal salts of lead (II), nitrate and copper (II) and sulphate with a purity of 99% were used to prepare solutions containing lead and copper. HCL and NaOH solutions were used to adjust pH.

In order to optimize the absorption process, four factors were evaluated at three levels. Therefore, 30 experiments were designed that are given in Table 1 (Rezaei et al., 2013). pH, time, adsorbent amount and particle size were optimized. Values of 0.2, 5.1 and 10 mg/L of platanusfruit were considered as absorbents. Then, the solution containing heavy metals of lead and copper (II) was produced by metallic salt of lead (II), nitrate, copper (II) and sulphate with a concentration of 50 mg. Amounts determined by absorbent were added to the tubes containing 40 ml of metal ions solution.In all experiments, composition of solution and absorbent was stirred on a vibrating device at around 150 rpm. In this experiment, pH was set 2, 5 and 8, adjusted using nitric acid and sodium hydroxide and measured by pH meter.Designated time in the study was 10, 95 and 180 minutes. After the elapse of the defined time, the solution was passed through filter paper and injected into atomic absorption spectrophotometry in order to determine the amount of metal (Debska & Guzowska, 2011).

Table 1. Experimental conditions and characteristics on the adsorbent.

No	pH	Time	Particle size	Adsorbent amount	Concentration of adsorbed lead	Concentration of adsorbed copper
1	5	95	5.1	2	3.9	6.6
2	8	180	5.1	2	4.5	5.7
3	8	10	5.1	2	7.1	6.1
4	2	10	5.1	2	26.2	11.3
5	5	95	10.0	1	3.4	7.3
6	5	95	0.2	3	20.2	8.3
7	5	95	0.2	1	18.5	7.7
8	5	95	10.0	3	4.2	7.2
9	2	180	5.1	2	23.2	9.6
10	5	95	5.1	2	4.1	7.0
11	5	10	10.0	2	4.3	7.4
12	8	95	5.1	1	3.9	6.0
13	5	10	0.2	2	21.7	8.5
14	5	180	0.2	2	17.4	7.2
15	8	95	5.1	3	5.2	6.5
16	5	95	5.1	2	4.3	7.1
17	2	95	5.1	1	28.8	10.3
18	5	95	5.1	2	4.3	7.6
19	5	180	10.0	2	3.6	7.1
20	2	95	5.1	3	27.1	11.7
21	2	95	0.2	2	27.0	10.7
22	2	95	10.0	2	21.0	9.7
23	5	10	5.1	3	6.8	7.5
24	5	95	5.1	2	4.3	7.4
25	5	10	5.1	1	4.4	7.5
26	5	180	5.1	3	4.8	7.5
27	5	180	5.1	1	4.9	7.5
28	5	95	5.1	2	4.9	7.5
29	8	95	10.0	2	5.3	6.2
30	8	95	0.2	2	3.9	2.1

Modeling

The modeling in the study included the following steps:

Step one: the input dataset (laboratory data) was included and recorded in Excel software and then was loaded and called for network training phase in the MATLAB environment (two-thirds of data). The data were chosen from the entire dataset of the main base so that it represented characteristics of the entire set

Step two: inputs and outputs were defined for this stage, and pH, time, amount of absorbent, and particle size were introduced as inputs. Moreover, concentration of lead absorbed and copper absorbed were introduced as output.In fact, the network had an input layer of neuron (node) as well as an output layer of neuron. Then, characteristics of the network were discussed.

Step three: transfer functions of hidden layers and output were determined in this step. We have two elements of weight w and transfer function f for a simple neuron. Input p is applied on neuron and is weighted by multiplying the weight w. The sum was added to the corresponding bias, applied as an input on the transfer function f and final output was obtained. Transfer function is a linear or non-linear function. There are 12 common transfer functions in neural networks. Different functions of each layer were tested in the study and the best results were observed for the transfer function of the hidden layer tansig and output layer purelin.

Step four: the number of nodes or hidden layer neurons were obtained in this step. The number of input layer neurons and the number of neurons in the output layer were equal to the number of output data. The number of hidden layer neurons was achieved by trial and error. In fact, the error rate was determined by trial and error through changes in the number of hidden layer neurons. One neuron had the lowest error in the study. Therefore, it was possible to reach the best answer with one neuron for hidden layer.

Levenberg-Marquardt (LM) is the best method of learning algorithm. Levenberg-Marquardt is a network training function that updates weight and bias values based on Levenberg-Marquardt optimization. Levenberg-Marquardt is often the fastest back propagation algorithm that is highly recommended as a first-choice supervised algorithm. Now, that the network has been trained, the remaining one-third of the original dataset was used to evaluate and test the network. Error rate was measured and network error was drawn. Figure 1 shows the structure of the neural network obtained in the study. Each input and output layer consists of one neuron.

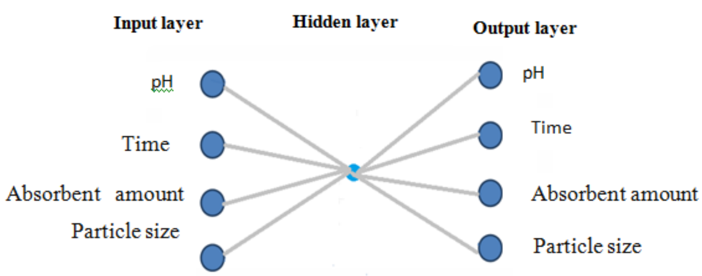

Figure 1. Structure of modeling neural network.

Discussion

Training in the neural network is based on the assumption that usually about two thirds of the available experimental data are randomly selected for training the network and the remaining one-third of the data is used to evaluate the model. Two important definitions that are used to evaluate network are mean square error (MSE) and regression coefficient (R^2), which are defined as follows:

$$MSE = \frac{\sum_{i=1}^{n}(b_i - b_i^{\exp})}{n} \quad , \qquad R^2 = 1 - \frac{\sum_{i=1}^{n}(b_i - b_i^{\exp})^2}{\sum_{i=1}^{n}(b_i^{\exp} - b_m)^2} \qquad (1)$$

In the above equation, , , and n respectively are the model output, laboratory output, experimental data mean and the number of data. If the network responses are drawn according to the expected outputs (laboratory values), a straight line of 45 degrees should be obtained. The linear regression coefficient close to 1 and MSE close to zero indicate accuracy of the model.

In order to determine the transfer functions of the middle and output layers as well as the number of neurons in the middle layers,the number of the neurons in the middle layer is assumed to be 3 and various transfer functions are tested for the output layer with constant transfer function of the middle layer as function tribas. Therefore, the function with the lowest error is selected as the transfer function of the output layer. Table 2 shows the results. According

to the results, the function purelin has the least error rate in the output layer. These two functions are defined as follows:

$$tribas(x) = \frac{1}{(1+e^{-x})} \quad purelin(x) = x \qquad (2)$$

Table 2. Determination of the best transfer function for the output layer.

No	The number of middle neurons	Transfer function of middle layer	Transfer function of output layer	Network training algorithm	Error of network training	Error of network testing
1	3	tribas	hardlim	trainlm	1.4847	2.0669
2	3	tribas	hardlims	trainlm	0.8976	1.5551
3	3	tribas	purelin	trainlm	0.0213	0.8221
4	3	tribas	satlin	trainlm	0.7879	2.1001
5	3	tribas	satlins	trainlm	0.6769	1.0034
6	3	tribas	logsig	trainlm	1.0006	0.9888
7	3	tribas	tansig	trainlm	1.1233	2.7879
8	3	tribas	poslin	trainlm	1.0067	1.0569
9	3	tribas	hardlim	trainlm	1.4847	2.0669
10	3	tribas	hardlims	trainlm	0.8976	1.5551

In the second phase, 10 neurons of middle layer and the transfer function purelin of the output layer were selected and different transfer functions were examined. According to the results in Table 3, the lowest error rate related to the transfer function tansig. The function is defined as follows:

$$\tan sig(x) = \frac{2}{(1+e^{-x})} - 1 \qquad (3)$$

Table 3. Determination of the best transfer function for the middle layer.

No	The number of middle neurons	Transfer function of middle layer	Transfer function of output layer	Network training algorithm	Error of network training	Error of network testing
1	3	hardlim	purelin	trainlm	2.5628	1.5657
2	3	hardlims	purelin	trainlm	0.3433	2.4446
3	3	satlin	purelin	trainlm	0.3331	2.5888
4	3	satlins	purelin	trainlm	0.9891	1.9887
5	3	logsig	purelin	trainlm	1.0034	1.6788
6	3	tansig	purelin	trainlm	0.0101	1.0055
7	3	poslin	purelin	trainlm	1.0003	2.4544
8	3	softmax	purelin	trainlm	2.2345	2.7790
9	3	radbas	purelin	trainlm	1.5433	3.7779
10	3	tribas	purelin	trainlm	0.4567	2.0001

The number of neurons in the middle layer was determined by identifying transfer functions of the middle and output layers in the final step. According to Table 4, 1 neuron has the lowest error rate.

Table 4. Determination of the number of appropriate neurons for the middle layer.

No	The number of middle neurons	Transfer function of middle layer	Transfer function of output layer	Network training algorithm	Error of network training	Error of network testing
1	1	tansig	purelin	trainlm	0.0409	0.8992
2	2	tansig	purelin	trainlm	0.4432	1.9936
3	3	tansig	purelin	trainlm	1.0056	2.0011
4	4	tansig	purelin	trainlm	1.9874	3.2390
5	5	tansig	purelin	trainlm	1.0023	2.0011
6	6	tansig	purelin	trainlm	0.0677	1.0067
7	7	tansig	purelin	trainlm	0.0556	0.9001
8	8	tansig	purelin	trainlm	1.6754	1.9899
9	9	tansig	purelin	trainlm	1.6079	2.6789
10	10	tansig	purelin	trainlm	2.8876	3.6768
11	11	tansig	purelin	trainlm	2.6668	3.6768
12	12	tansig	purelin	trainlm	1.0005	2.1222
13	13	tansig	purelin	trainlm	0.0890	3.2367
14	14	tansig	purelin	trainlm	0.2343	1.6760
15	15	tansig	purelin	trainlm	1.4027	3.7877
16	16	tansig	purelin	trainlm	1.6226	2.6768
17	17	tansig	purelin	trainlm	1.1044	2.0078
18	18	tansig	purelin	trainlm	1.2977	3.6768
19	19	tansig	purelin	trainlm	0.6738	2.9901
20	20	tansig	purelin	trainlm	0.2564	3.0012
21	21	tansig	purelin	trainlm	2.0485	3.6888
22	22	tansig	purelin	trainlm	0.5279	3.2344
23	23	tansig	purelin	trainlm	0.7944	0.6668
24	24	tansig	purelin	trainlm	1.8401	1.0023
25	25	tansig	purelin	trainlm	0.0787	1.8999
26	26	tansig	purelin	trainlm	0.9167	1.5657
27	27	tansig	purelin	trainlm	1.2442	2.7877
28	28	tansig	purelin	trainlm	0.3983	2.5655
29	29	tansig	purelin	trainlm	2.1678	3.6661
30	30	tansig	purelin	trainlm	1.3678	2.9901

In order to show the match between data from modeling and experimental data, the data was plotted on a graph. The graph that is also called linear regression evaluates the match between data from modeling (predicted) and actual or target data (experimental) and shows the result in the form of a straight line passing through the line of 45 degrees. If the line drawn on the line of 45 degrees is more cinsistent, it will indicate a better match between modeling and real data. All modeling and real data in the graph is plotted as dots around the line of 45 degrees. Figure 2 shows changes in regression coefficient for the trained network.

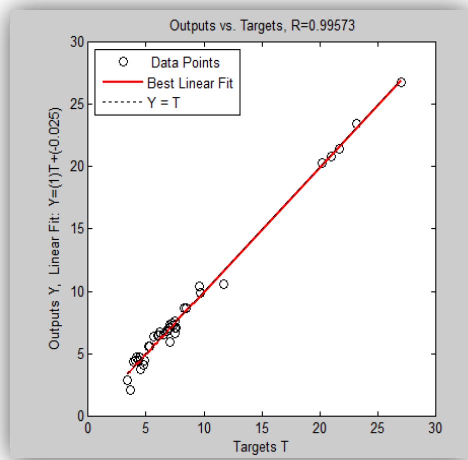

Figure 2. Changes in regression coefficients for the trained network.

As seen in Figure 2, the network outputs for data used for training were compared with experimental data. The straight line passing through the graph is well coincided with the line of 45 degrees. Moreover, these data points are very close to the straight line. Therefore, the developed model has good accuracy.

In order to evaluate whether the network can well generalize trained rules to other data, evaluation data (the remaining one-third of the data) were used. Figure 3 shows a similar comparison for the evaluation of the model accuracy. As shown, the straight line passing through the graph also coincides with the line of 45 degrees. These data points are also very close to the straight line. Therefore, the trained network can generalize rules to new data.

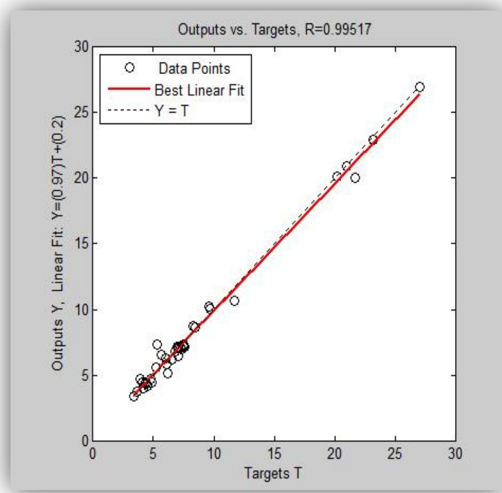

Figure 3. Changes in regression coefficient for the evaluation of the accuracy of modeling.

Conclusion

- Results of linear regression showed that correlation coefficient obtained (R-Value) between the measured and predicted output variables is close to 1 indicating that the network response is satisfactory.
- The results also showed that training artificial neural network with the process data to remove heavy metals with phytoremediation has been successful.
- With proper training, ANN can successfully predict the system output for new conditions.

References

Athar , M., vohora, S. B. (2007). Heavy metals & envirnment. Islamic Azad university of sanandaj, 10 – 11. [Publisher]

Debska, B., & Guzowska, B. S. (2011). Application of artificial neural network in food classification. Analytica Chimica Acta, 705, 283– 291. [Google Scholar] [Publisher] https://doi.org/10.1016/j.aca.2011.06.033

Nasiri, M., Mohammadi, A., & Tavakkoli, M. (2014). Prediction of the survival of Lactobacillus acidophilus in yogurt condensed with an artificial neural network. The 3rd National Conference on Food Industry. [Publisher]

Rezaei, M., Khaje Kazemi, R., & Alizade, R. (2013). Modeling and optimization of copper removal from water with plantain fruit using design of experiments. Applied research in chemical engineering, 7(1), 39-46. [Google Scholar] [Publisher]

Sharifi Arab, G. H., Rafiei, R., Jalali, H., & Ameri, M. (2013). Presentation of a computerized model to predict the shelf life of Lactobacillus in probiotic yogurt. Journal of Food Technology, Shahrood, Iran. [Google Scholar] [Publisher]

Venkata, K. K. (2009). Upadhyayula, l; science of the total environment. 4081- 13. [Publisher]