WHAT ARE ARTIFICIAL NEURAL NETWORKS AND WHAT THEY CAN DO?

Vlastimil Dohnal[a], Kamil Kuča[b], Daniel Jun[b]

[a] *Department of Food Technology, Faculty of Agronomy, Mendel University of Agriculture and Forestry, Brno, Czech Republic*
[b] *Department of Toxicology, Faculty of Military Health Sciences, University of Defence, Hradec Králové, Czech Republic*
e-mail: dohnal@mendelu.cz

The artificial neural networks (ANN) are very often applied in many areas of toxicology for the solving of complex problems, such as the prediction of chemical compound properties and quantitative structure-activity relationship. The aim of this contribution is to give the basic knowledge about conception of ANN, theirs division and finally, the typical application of ANN will be discussed. Due to the diversity of architectures and adaptation algorithms, the ANNs are used in the broad spectrum of applications from the environmental processes modeling, through the optimization to quantitative structure-activity relationship (QSAR) methods. In addition, especially ANNs with Kohonen learning are very effective classification tool. The ANNs are mostly applied in cases, where the commonly used methods does not work.

INTRODUCTION

Artificial neural networks (ANN), as one of the artificial intelligence (AI) methods, became to play important role in our scientific or personal life. They are applied in many areas of human work, such as in economy for prediction of stock market[1], electricity load forecasting[2], sheep milk production curve[3], prediction of toxicity of complex mixtures[4] or optimization of separation processes in chemistry[5]. They successfully penetrates into equipments of common use. Typical examples of this penetration are compact cameras, where AI is employed in auto focus function, respectively in washing machines for the on-line optimization of washing cycle.

The term "artificial neural network" (ANN) originates from its biological pattern – neural network (NN), which represents the network of interconnected neurons in living organisms. The ANNs are based on the same conception as the biological ones, they are considered as the collection of interconnected computing units called artificial neurons (AN). Both neural networks, the biological and artificial, can consist of a few to unlimited numbers of neurons. The function of NN is defined by many factors, for example by number and arrangement of neurons, their interconnections, etc. The ANNs are able to simulate these NN in both, architecture and function. The word "simulate" was used because until now there is not exactly known the function of biological neural network systems.
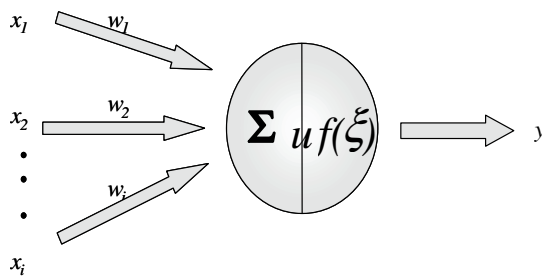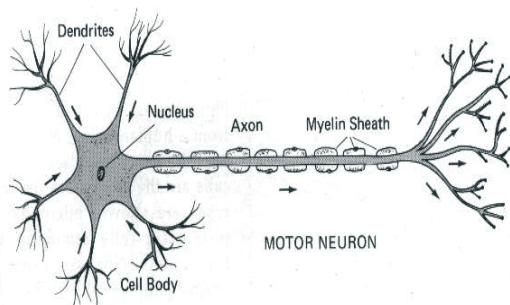
BIOLOGICAL VS. ARTIFICIAL NEURON

Human brain consists of many different neuron types, which were developed from one basic type of neuron by the specialization. In opposition to NN, the function and structure of basic neuron is relatively well known. It consists of three parts – the soma (cell body with nucleus), the dendrites (carry impulses toward the soma) and the axon (carry impulses away from the soma). The connection between neurons is provided through the gaps called synapses. The neuron receives the electrical signals from dendrites, passes them through soma, which performs some operation and then sends it to the axon. Axon distributes the signal to dendrites. The gap junctions exhibit various abilities to transfer the signal (Fig. 1).

The conception of AN fully originates from biological neuron. Each AN has certain number of inputs. These inputs have different transmission. Each of them has assigned its own weight, which indicates the importance of the input. In the neuron, the sum of weighted inputs is calculated and when this sum overcomes certain value, called bias (noise or threshold), the sum is processed using transfer function and the result is distributed through the output to the next AN. The comparison of function of both, biological and artificial neuron, is given on Fig. 1.

LEARNING (TRAINING OR ADAPTATION) OF ANN

From the AN theory it is evident, that there are many values (weights, biases) which have to be set. There were developed many adaptation algorithms for this purpose.

dendrites x_i Inputs

synaptic efficacy w_i weights

excitation level u noise (bias, threshold)

$f(x)$ activation function

signal y AN's output

a) Biological neuron b) Artificial neuron

Fig. 1. The comparison of the biological and the artificial neuron.

They construct two basic groups – supervised and unsupervised.

Supervised algorithm requires the knowledge of desired output. The algorithm than calculates the output with current weights and biases. The output is compared with targeted output and the weights and biases are adjusted by algorithm. This cycle is repeated until the difference between targeted and calculated values is not sufficient. The most applied supervised algorithms are based on gradient methods (for example back-propagation[6]) and genetics (genetic algorithms). The most used supervised ANNs are feed-forward ANN and radial basis function neural networks (RBF).

While the supervised learning algorithm requires the knowledge of output values, the unsupervised does not need them. It produces its own output, which needs further evaluation. Typical example of unsupervised algorithm is Kohonen learning used in self-organising maps[7].

ANN ARCHITECTURE

Similarly, as the neurons are differently interconnected in brain, the topology of ANNs is varying from case to case too. The most frequently the ANs are sorted into layers. There can be distinguished 3 types of layer – an input, hidden and output layer. The ANs in input layer take the inputs and distribute them to the following layers. Hidden layer does calculations and the output layer serves the calculated results to the user. The ANN consists of one input and one output layer. The absence or presence and number of hidden layers depends, such as the number of AN in each layer, on complexity of the solved problem. The examples of ANN structures are given on Fig. 2.

MULTILAYER FEED-FORWARD ANN (MLF-ANN)

The feed-forward ANN consists of neurons into 3 or more layers. The first one is input layer, the last one the output layer. Between these layers, the one or more hidden layers are implemented. The MLF-ANNs with the AN arranged in one hidden layer are the most used, because they are able to solve common problems. The architecture can vary in number of ANs in layers, the connections (fully or partially interconnected layers of ANs) and in the transfer function of AN. In the case of adaptation algorithm, the situation offers choice of different methods. The most common are gradient methods, based on the first derivation of sum of differences between targeted and predicted output values. This function is called "root-mean-square-error" (RMSE). The goal is to minimize the value of this function, sometimes very complex with local and global minima. For the minimization of RMSE value were employed algorithms based on crystallography, genetics etc. For example, it is known, that the arrangement of atoms in a crystal form has the lowest energy. Therefore the algorithm simulating the annealing and crystallization process during the cooling was developed (simulated annealing) to find the global minima of the function.

Genetic algorithms (GA) are inspired by Darwin evolution theory – the one with the best gene combination survives. GA introduces biological processes into the calculation. The ANN parameters (genes), such as weights and biases, form vector called "chromosome". At the start, the population of chromosomes is generated. During calculation process, similarly as in genetics, the crossover (genes interchange) and mutation (random gene change) take place and the chromosomes with the highest "error function" are excluded. The average value of error function of population is decreasing during evolution. On the end, the "best" chromosome is selected and used.
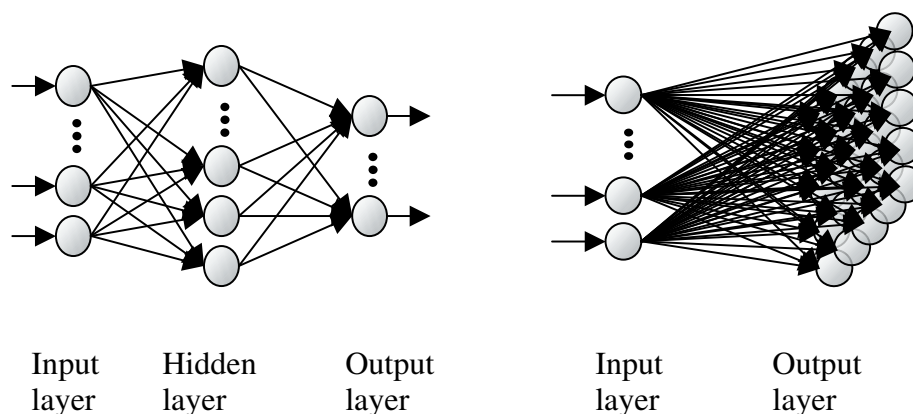
Fig. 2. The examples of ANN structures – multi-layer perceptron (left) and Kohonen self-organising map (right).

KOHONEN SELF-ORGANISING MAPS (SOM)

This type of ANN, due to its properties, is very often used as the clustering or classification tool. It provides projection of input data, arranged as the multidimensional vector, into lower dimensional space. SOM consists of two layers – the input and the output. In comparison with other types of ANN, the hidden layer is missing. The input layer contains the same number of AN as the input vector descriptors. The output layer contains ANs sorted in 1, 2 or more dimensions. The adaptation is provided by special learning algorithm. The AN with output the most closest to the targeted is selected as the "winning neuron" and the weight adjustment has the highest influence on it. The weights of ANs in neighborhood are adjusted too, but the change takes lower influence. The change of weights is indirectly proportional to the distance of ANs from the "winning neuron". Application of this algorithm causes, that the similar vectors are projected on ANs in the output layer into the same AN, respectively relatively closely themselves. There is possible to distinguish group of similar vectors based on input data and makes SOM alternative to the statistical methods, such as cluster analysis or principal component analysis.

SELECTED TOXICOLOGICAL APPLICATION OF ANN

Toxicological effect of compounds depends on molecular properties, such as solubility in pure or mixed media, partitioning between different solvents, pharmacokinetics etc.

Very important properties of drugs are their absorption, distribution, metabolism and excretion (ADME). Traditional widely used approach, when potential drugs are tested in vitro, was compared with computational methods, such as PLS or ANN in 2D and 3D-QSAR on UDP-glucuronosyltransferase in review of Smith et al.[8].

A three layer feed-forward ANN with back-propagation (BP) adaptation algorithm was used for QSAR investigation of organophosphorus compounds toxicity measured against midge larvae (*Chironomus riparius*) under varying temperature (11, 18, and 25 °C) and pH (6, 7, and 8) conditions and with or without sediment[9]. It was reached high discrimination between senzitizers and non-senzitizers.

BP-ANN was also used for toxicological values calculation of 140 organic compounds and the results were compared with multilinear regression (MLR) (ref.[10]). For higher accuracy authors included atom-type E-state indicies as one of molecular descriptors. It allowed building of model for toxicity prediction for diverse structures in data set used.

The modeling of drug solubility using BP-ANN was done by Jouyban et al.[11]. They used data collected from literature, totally 35 data sets – for different water-cosolvent mixtures. Solvents acetonitrile, dimethylformamide, dioxane, ethanol, methanol, ethylenglycol, propylenglycol and glycerine were tested. They evaluated ANN ability to predict solubility of compound in different cosolvents, structurally different compounds in one cosolvent and finally the possibility to predict solubility of new structurally similar drug in concrete solute. The results of ANN were more precise in comparison with standard statistical method MLR.

Mutagenicity is one of the important toxicological endpoints. There is certain similarity between cancerogenic and mutagenic effects. Due to the existence of quality databases they were calculated QSAR models including ANN methods. Valkova with coworkers[12] studied mutagenecity of 95 aromatic and heteroaromatic amines using counter-propagation ANN. This method combines an unsupervised classification by Kohonen SOM with supervised learning algorithm. The chemical structures studied were transformed into 275 descriptors including topology, topochemistry, geometry and quantum chemistry. In prediction mode, the compound is firstly projected into Kohonen map (closely to structurally similar compounds) and then to the output layer where the value of

mutagenecity is calculated. Three different divisions of data set (learning/test set ratio 1 : 2 to 2 : 1) were tested. The results showed satisfactory prediction power of model, where the test set was two times larger than learning set when rational division of data set in descriptor space was performed.

CONCLUSION

Due to the diversity of architectures and adaptation algorithms, the ANNs are used in the broad spectrum of applications from the environmental processes modeling, through the optimization to quantitative structure-activity relationship (QSAR) methods. In addition, especially ANNs with Kohonen learning are very effective classification tool. The ANNs are mostly applied in cases, where the commonly used methods does not work.

REFERENCES

1. Cao Q, Leggio KB, Schniederjans MJ. (2005) A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market. Computers & Operations Research *32*, 2499–2512.
2. Pai PF, Hong WC. (2005) Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy Conversion and Management *46,* 2669–2688.
3. Torres M, Hervás C, Amador F. (2005) Approximating the sheep milk production curve through the use of artificial neural networks and genetic algorithms. Computers & Operations Research *32*, 2653–2670.
4. Gagné F, Blaise C. (1997) Predicting the toxicity of complex mixtures using artificial neural networks. Chemosphere *35*, 1343–1363.
5. Dohnal V, Li H, Farková M, Havel J. (2002) Quantitative analysis of chiral compounds from unresolved peaks in capillary electrophoresis using multivariate calibration with experimental design and artificial neural networks. Chirality *14*, 509–518.
6. McClelland JL, Rumelhart DE. Explorations in parallel distributed processing. Cambridge, MA: MIT Press, 1988.
7. Kohonen T. Self-organising maps. New York: Springer, 2001.
8. Smith PA, Sorich MJ, Low LSC, McKinnon RA, Miners JO. (2004) Towards integrated ADME prediction: past, present and future directions for modelling metabolism by UDP-glucuronosyl-transferases. J Mol Graph Model *22*, 507–517.
9. Devillers J. (2000) Prediction of toxicity of organophosphorus insecticides against the midge, chironomus riparius, via a QSAR neural network model integrating environmental variables. Toxicology Methods *10*, 69–79.
10. Huuskonen J. (2003) QSAR modeling with the electrotopological state indices: predicting the toxicity of organic chemicals. Chemosphere *50*, 949–953.
11. Jouyban A, Majidi M, Jalilzadeh H, Asadpour-Zeynali K. (2004) Modeling drug solubility in water–cosolvent mixtures using an artificial neural network. Il Farmaco *59*, 505–512.
12. Valkova I, Vračko M, Basak SC. (2004) Modeling of structure–mutagenicity relationships: counter propagation neural network approach using calculated structural descriptors. Anal Chim Acta *509*, 179–186.