NEURAL NETWORKS AS A POWERFUL TOOL IN THE HANDS OF THE MANAGER

Michal Fraňo
Žilinská univerzita v Žilině, Fakulta riadenia a informatiky, Katedra manažérskych teórií

Keywords: *backpropagation, estimation, external criteria of quality, forecasting, GMDH, mean prediction error, operational research, optimization, prediction*

1 Introduction

One of the key roles of every manager is planning. From the time point of view there are operative - short term planning and strategic - long term planning. The higher the level of management, the longer is the period that the plans have to cover. The managers on higher level of every organization make plans for months and years in order to ensure a benefit for the whole enterprise. There is an urgent need to predict the processes and their evolution in the future. Knowing the estimations of process future development helps a lot by making decisions. Above all by long-term plans and by huge amounts of products or services, that are regularly produced by an organization, good estimations of future values can save a lot of money. In the last decades an enormous progress came to information technologies. Computers and algorithms from the fields of statistics, operational research and prognostics deliver a significant help for managers in all kinds of organizations and in all levels of the organizations. Huge and sophisticated enterprise information systems like SAP, OLAP offer for managers many useful algorithms and tools for evaluating processes especially from statistical point of view. Statistical analysis is necessary for enterprise resource planning. The development of the process in the past can give useful information to the manager about the trend in the future. The more accurate the estimation of the process future trend, the higher the savings of internal resources and money could be. That is why it pays to invest more money and attention to good and exact prediction algorithms.

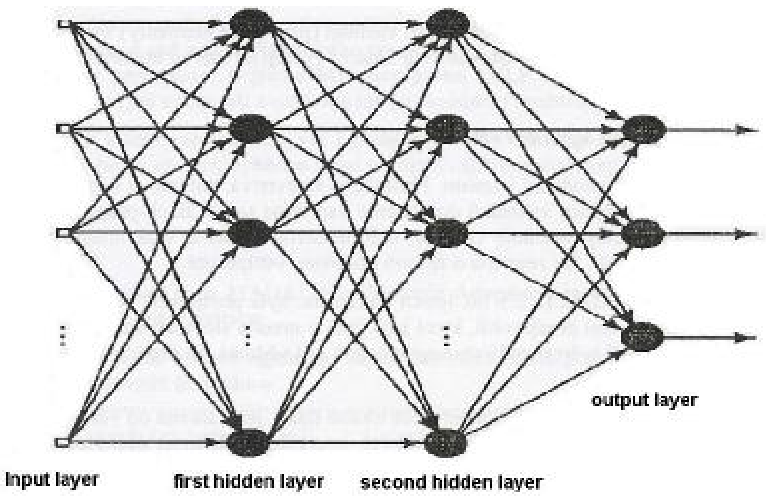
2 Neural networks in the field of prognostics

Besides the known and often used statistical algorithms like exponential smoothing, linear, quadratic and exponential regression there are many other more sophisticated algorithms nowadays. They use last state of knowledge for their work in branches like cybernetics, physics, microbiology and human behavior. Direct smoothing algorithm tries to approximate the process through composition of trigonometric functions. By enumerating of frequencies and magnitudes is able to predict future values. Seasonal algorithms like Winter´s models take into consideration not only trend of the process but also a seasonal factor, which helps in strong seasonal processes to predict the values with higher accuracy. Adaptive self-learning algorithm uses cybernetical principle and during the process of learning corrects its setting of model parametres.

New stage in development of prognostic algorithms are the neural networks. The principles taken from the microbiology try to immitate the way of human brain work. In the field of prognostics they are not only able to predict the future values, but also to see the internal relations among input values. Some of them also deliver the explicit mathematical description of the process at the end of the model development. In last two years the subject of my study and experiments are neural networks of two sorts: GMDH and Backpropagation. And because my study and work is concentrated in the branch of management, I am very excited about the possibilities and benefits that these two neural networks bring in the field of forecasting in strategical management.

2.1 Backpropagation Neural Network

Backpropagation algorithm belongs to the most used neural network algorithms in predictions. It is a neural network architecture of a perceptron type. It can have one layer, but mostly two or three layered backpropagation networks are used, which besides making simple classification tasks, can also solve more complicated problems. On the following picture there is a three layered backpropagation network.



Pic.1: The architecture of multilayer backpropagation network

Each neuron of the current layer is connected to the outputs of all neurons of the previous layer. The decision how many neurons should be in each layer expects some experience from the designer, but there are some useful suggestions.

Suggested numbers of neurons for hidden layers are:

$$N_{first\_hidden} = N_{output} \cdot \left(\sqrt[3]{\frac{N_{input}}{N_{output}}} \right)^2 \qquad N_{second\_hidden} = N_{output} \cdot \left(\sqrt[3]{\frac{N_{input}}{N_{output}}} \right) \qquad 2.1$$

As a transfer function sigmoid is commonly used. All outputs are computed using sigmoid thresholding of the inner product of the corresponding weight and input vectors. The behaviour of neurons could be described by transfer formula:

$$y = S\left(\sum_{i=1}^{n} w_i \cdot x_i + \Theta \right) \quad \text{, where} \quad S(j) = \frac{1}{1+e^{-g\,j}} \qquad 2.2$$

During the work the backpropagation algorithm is minimizing the energetic function. This energetic function is also known as error of the network.

It is defined as

$$E = \frac{1}{2} \sum_{i=1}^{n} (y_i - d_i)^2 \qquad 2.3$$

, where n is the number of outputs of the network, y_i is output of neuron i, d_i is demanded output of neuron i.

After each comparison of computed and demanded outputs the weights are modified backwards in order to find the minimum of error function E.

At the beginning of the learning process it is recommended to set the weights w_{ij} in the range between -0.05 and +0.05.

An important task is to choose proper training and testing sets of inputs. These sets should be representative for the predicted process. Here is recommended to work on man-machine approach. The operator should appoint the training and testing set of data. The description of the whole work of the Back-propagation follows:

```
random initialising of weights

repeat

  repeat
      choose_pattern from the training set
      make data of the chosen pattern to inputs
      compute_outputs of the network
      compare_outputs with demanded values
      modify_weights
  until choosen_all_patterns from the training set
until Error_function < criteria
```

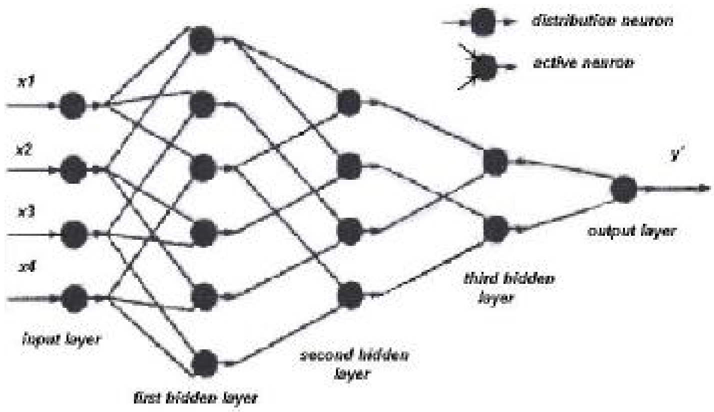
During the work of the backpropagation neural network a problem can appear that the demanded criteria of quality is not reached, but there in no more improvement at all. This is called a stagnation in local minimum. In order to prevent this situation it is recommended to choose the input vector not continually value by value, but to choose it randomly from the learning set of the data.

2.2 GMDH

Neural Network GMDH (Group Method of Data Handling) differs from other neural networks in some aspects. The most neural networks are seeking the optimal model of the system by deduction. The architecture of neural network and transfer functions are prescribed. Afterwards the network is trained in order to set all the parametres, so that the error of the output will be minimal.

The other approach is induction, which uses the GMDH algorithm. At the beginning, there are some building units - neurons. From these, the model should be derived during the training period. The transfer functions could change during the training process.

Seeking the optimum by GMDH can be compared to the natural choice of stronger element. eurons, whose combinations affect the result more, are chosen and others are left from the model.



Pic.2: Scheme of the GMDH neural network

At the beginning of the model creation all transfer functions of the neurons are linear. In case of GMDH neural network each neuron has only 2 inputs (*i*,*j*) and one output *y*.

The output of each neuron is quadratic combination of its inputs:

$$y = ai^2 + bij + cj^2 + di + ej + f \qquad 2.4$$

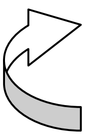
Transfer through each layer powers the output polynomial by 2. This polynomial is called Ivachnenko´s polynomial according its founder, Ukrainian scientist Prof. Ivachnenko. Other often used Ivachnenko´s polynomials are:

$$y = a + bi + cj \qquad 2.5$$

$$y = a + bi + cj + dij \qquad 2.6$$

The polynomial 2.5 is linear. The polynomial 2.6 is nonlinear, it is a small modificiation of the polynomial 2.5. In most processes the linear polynomial is sufficient to deliver exact predictions. If the process has a quadratic character, it is reasonable to take the polynomial 2.4. The learning process runs together with creation of the network structure. On each layer such number of neurons arise, which corresponds the number of composite combinations of pairs of previous layer outputs. Then the neurons are configured – in the case of using the 2.4 quadratic polynomial the values of all 6 coefficients of the polynom are computed. The best neurons in sense of external criteria are selected. The process of creating new layers continues until the criteria of quality is reached.

The work of GMDH network is described by following diagram:



1. Configuration of k-*th* layer (k is the number of actual created last layer)
   Creating of new neurons in the layer
   Computing of all six (or three, two or four) coefficients of the polynome of each neuron.
2. Selection of neurons of k-*th* layer
3. Termination of learning of the network

3 Impact on the economy of the enterprise

Management is an cross-disciplinary science. It uses tools and algorithms taken from many branches of the science in order to ensure the benefit of the enterprise. In the competitive environment it is not possible to raise the benefit by increasing the price for products and services. The way of increasing the benefit is to cut the expenses. Ill defined problems or fuzzy estimations bring extra costs, because the uncertainty of future evolution of the process make the management to prepare itself for unknown volumes of work, which have to be done. This brings two kinds of extra costs:

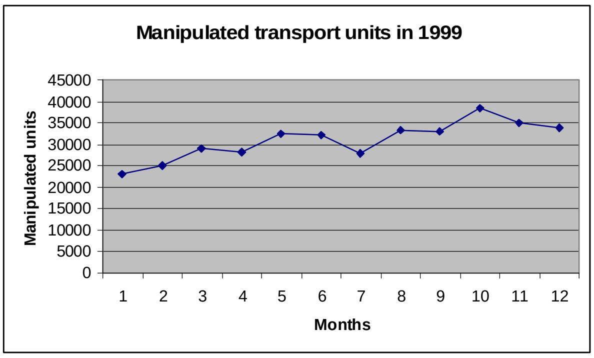
1. If the manager underestimates the future value, then there are fewer people and fewer equipment ready for the work as it should be. This leads to delays, because not all requests from the side of customers could be satisfied in time.
2. If the manager overestimates the future value, then there is too much potential ready for the work. More people were employed, but now they are in stand by status, because there is not as much work to do, as it was expected. Maybe some equipments were rented from other enterprises and this brings extra costs.

Exact forecast can prevent this situation. Even, if forecast is always overcast with certain error, using modern neural network based forecasting algorithms can lower this error to minimum and by this it saves a lot of extra costs.

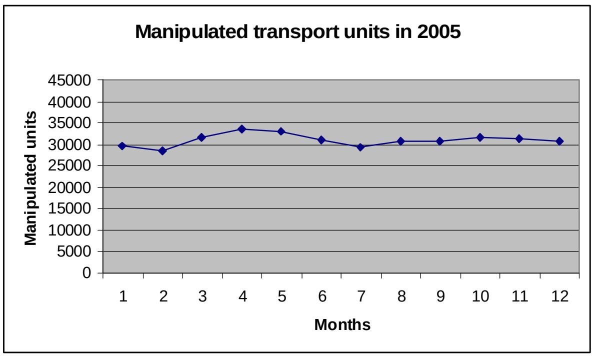
Let´s take an example from a transportation enterprise from a big town.

There are values from 192 months of manipulated transport units. These 192 values cover the years from 1990 to 2005. Following two graphs represent the evolution of the process in the the years 1999 and 2005. These graphs give initial information about the characteristics of the process.

Graph 1. Manipulated transport units in 1999



Graph 2. Manipulated transport units in 2005



In following text three estimations of the montly volumes of manipulated transport units in 1999 and in 2005 will be compared.

Firstly it is the expert estimation. It is made monthly by a manager of the transportation enterprise. He knows the technology of manipulation with the units. He also knows past values, he knows the general economical situation in the country. These volumes were put into the monthly plan.

Two other values will be obtained by forecasting using GMDH and Backpropagation neural networks..

Evaluating these three values consists of enumerating of mean prediction error, *which is defined as*

$$MPE = \frac{1}{d}\sum_{i=n+1}^{n+d}\frac{|x_i - pred_i|}{x_i} \qquad (3.1)$$

where d is length of forecast, n is index of last known value of time series and from the index $n+1$ begins the forecast, x_i is real value, $pred_i$ is prediction

The values of average MPE are shown in Table 1.

Table 1: Comparison of average MPE during the years 1999 and 2005

Year	Expert estimation	GMDH	BackPropagation
1999	34,85 %	6,77 %	8,87 %
2005	6,02 %	5,69 %	4,10 %

Mean prediction error in the year 2005 is in all three ways of estimation lower than in 1999. It is caused by the character of the process, which is in 2005 smoother.

To see the benefit of using neural networks in the estimation there should be made the financial expression of the extra costs of every individual way of estimating.

The manipulation costs are 20 Euro.

In the case of overestimation, where the estimated value is higher than the real value, the extra costs are enumerated as one fifth of the costs of manipulation. The extra costs are evaluated according following formula:

Extra costs = (estimation - real value) * 4 Euro

In the case of underestimation there will be the lack of crew and equipment and this will lead to delays and it will cause penalties for not delivering manipulated units in time. The cost per every underestimated unit will be two fifth of the manipulation costs. The formula will be following: Extra costs = (real value - estimation) * 8 Euro.

In table 2 are values of monthly extra costs of every kind of estimation in Euro.

Table 2.

Month	Real value	**Expert estimation**	Extra costs	**GMDH**	Extra costs	**BP**	Extra costs
I.99	23069	**36271**	52808	**27727**	18632	**28612**	22172
II.99	24989	**39732**	58972	**25246**	1028	**29521**	18128
III.99	28995	**43989**	59976	**30752**	7028	**29916**	3684
IV.99	28215	**40300**	48340	**29828**	6452	**31505**	13160
V.99	32361	**43989**	46512	**32269**	736	**30601**	14080
VI.99	32206	**40470**	33056	**32284**	312	**32497**	1164
VII.99	27816	**41819**	56012	**30773**	11828	**32040**	16896
VIII.99	33438	**41819**	33524	**28208**	41840	**30987**	19608
IX.99	33071	**40470**	29596	**34457**	5544	**33711**	2560
X.99	38340	**41819**	13916	**34057**	34264	**33749**	36728
XI.99	35057	**40470**	21652	**34086**	7768	**37055**	7992
XII.99	33846	**41819**	31892	**34960**	4456	**33499**	2776
Total:			**486256**		**139888**		**158948**
I.05	29547	**29347**	1600	**33775**	16912	**31998**	9804
II.05	28588	**26764**	14592	**28127**	3688	**31552**	11856

III.05	31593	**29845**	13984	**27286**	34456	**32663**	4280
IV.05	33729	**28818**	39288	**31142**	20696	**33416**	2504
V.05	32962	**29347**	28920	**33998**	4144	**33531**	2276
VI.05	30975	**28818**	17256	**34183**	12832	**33925**	11800
VII.05	29263	**29347**	336	**31440**	8708	**30622**	5436
VIII.05	30815	**29845**	7760	**29966**	6792	**30785**	240
IX.05	30810	**28818**	15936	**30605**	1640	**31644**	3336
X.05	31576	**29347**	17832	**30468**	8864	**33494**	7672
XI.05	31199	**28818**	19048	**31346**	588	**31470**	1084
XII.05	30621	**29845**	6208	**31449**	3312	**30445**	1408
Total :			**182760**		**122632**		**61696**

From the table it is evident that there are huge savings in year 1999 by using GMDH (486256-139888 = 346368 Euro) and Backpropagation neural network (486256 -158948 = 327308 Euro). In the year 2005 the savings are smaller, but the saved amount is still very high and it surely pays to use the neural networks prediction algorithms.

4 Conclusion

Neural networks offer very strong tool for efficient work of the manager. The higher the level of the management is, the more important is the use of prediction algorithms. The effectivity of management and the profit of the enterprise can be increased. Therefore it is advisable to invest in sophisticated decision making support systems, where neural network algorithms are implemented. The money invested in such system will surely bring their reward in the future.

Summary

Managers in their work use sophisticated procedures from many scientific branches. In order to make good decisions in the stretegic level of management, they can use for predictions relatively new and effective instruments like neural network algorithms GMDH and Backpropagation. Their contribution is shown in an example from the transportation enterprise. There is a time series with 192 values of manipulated transport units per month. The expert estimations made by the manager working in the transportation enterprise, which were made on the basis of technological knowledge and experience, were compared with the estimations made by above mentioned neural network prediction algorithms. In this comparison the benefit of using neural networks is evident. In the hands of a manager the neural networks are a powerful tool to increase the benefit of the enterprise.

Literature:

[1] Ivachnenko, A.G., Ivachnenko, G.A., Muller, J.A., Self-organization of optimal physical data sampling clusterization for weakened description and undefinite object forecasting. Pattern recognition and image analysis (1993) 3, no. 4

[2] Lorence, E. N. (1969). Atmospheric predictability is revealed by naturally occurring analogues . J. Atmos. Sci., 4, P. 636-646

[3] Muller, J.A., Analysis and Prediction of Ecological Systems. SAMS, 1996. Vol. 25 pp. 209-243

[4] Kolekt.: Adaptive and self learning systems, SNTL Praha 1980

[5] Montgomery, D., Johnson, L.,Gardiner, J. Forecasting and Time Series Analysis. McGraw-Hill 1990.

[6] Šnorek, M: Neural networks and neurocomputers, ČVUT Praha, 2004.
[7] Kosko, B: Neural Networks and Fuzzy Systems, Prentice Hall, 1992.
[8] www.math.com

Contact:

Ing. Michal Fraňo
Žilinská univerzita v Žilině
Fakulta radenia a informatiky, Katedra manažérskych teórií
Fándlyho 44
010 11 Žilina, SR
tel.: 00421/908860260
email: michal_frano@yahoo.com