

OPTIMIZATION FORECAST OF HIGER EDUCATION AGRI BUSINESS COSTS

Lecturer Adriana BĂDESCU, Ph.D
Faculty of Economics, „IoanSlavici”
University, Timișoara, România



Lecturer Doina DARVASI, Ph.D
Faculty of Computer Sciences,
„IoanSlavici” University, Timișoara,
România



REZUMAT. În această lucrare este prezentată o structură a programului MATLAB pentru punerea în aplicare, formarea și utilizarea unei rețele neuronale folosite pentru a anticipa costurile din universitate. Utilizarea rețelelor neuronale artificiale pentru costurile din universitate este extrem de eficientă, predicția procentului de precizie fiind mai mare decât în cazul utilizării metodelor tradiționale.

Cuvinte cheie: management financiar, rețele neuronale artificiale, optimizare financiară, program MATLAB, previzionarea costurilor de universitate, metodele tradiționale, date de intrare.

ABSTRACT. In this paper is presented a Matlab program structure for the implementation, training and use of a neural network used to predict university costs. The use of artificial neural networks to university costs is extremely efficient, the prediction of accuracy percentage is higher than in the case of using traditional methods.

Key words: financial management, artificial neural networks, financial optimization, MATLAB program, university costs

1. THE POSSIBILITY OF USING ANN IN FINANCIAL MANAGEMENT

Generally, application features remain valid in this case too, beneficial exploitation of ANN is taking place within the economic processes in which mathematical model is difficult to achieve, too complex or the existing mathematical model has not the necessary accuracy (SUITABILITY low).

It is underlined once again the criteria of using ANN, in order to associate it with those issues of financial management where their application is efficient:

- mathematical process model is unknown, has too much complexity associated with insufficient accuracy (precision) and in some cases can not be determined;
- available data are incomplete in some cases, there are noise signals too, disturbance signals (noise term can be extrapolated from technical field and in other types of economic, genetic processes);
- there is a number of constraints (restrictions) applied to the process and must be simultaneously optimized.

The analysis and recommendations of the existing literature have outlined the following economic areas suitable for the use of ANN:

- verifying the authenticity of documents (including here for example verification of signature specimens);
- credit opportunities fund;
- predict of exchange rates and indices;
- assessment and diagnosis of certain elements of the firms' structure;

- predict university costs;
- credit card fraud detection system;
- market response for marketing problems, based on historical databases;
- predict firms' productivity;
- determining the optimum investment portfolio within financial institutions;
- class group operations (clustering), these types of operations fit within the unsupervised learning;
- customer segmentation;
- optimization issues (scheduling optimization, minimizing losses,...).

It is expected that ANN can be used successfully in the following areas:

- Analysis and prediction of university costs;
- Stock and bonds investment risk assessment;
- Credit card fraud detection;
- Determining the market response based on a knowledge ground accumulated over time;
- Assessing productivity;
- Investment management.

It should be highlighted an increase in the impact ANN had in economic processes, impact that can be also underlined by the presence in most specialty journals like: Management Science, IEEE Transactions on Systems, Man of Cybernetics, Decisions Science, Computers & Operations Research.

2. UNIVERSITY COSTS PREDICTION USING ANN. COMPARISON BETWEEN TRADITIONAL AND SPECIFIC METHODS TO ARTIFICIAL INTELLIGENCE

Within this subparagraph is proposed a comparison between methods called by the author as traditional (classical, conventional in the sense of their using and knowledge for several decades) and specific methods to artificial neural networks use. There are presented the theoretical grounds of the two methods, and gradually the results obtained after processing the same set of initial data, obviously it have been chosen the same initial data in order to make possible the comparison between the two methods, especially in terms of predicting accuracy.

3. MATLAB PROGRAM STRUCTURE FOR IMPLEMENTATION, TRAINING AND USE OF NEURAL NETWORKS USED TO PREDICT UNIVERSITY COSTS

In this paragraph is presented the source text of the program, as well as the meanings of instructions and their correlation with economic context of the issue.

```
% enter code
clear all
close all
clc
```

```
% input data for network training
```

```
X1=[1512 1595 1440 1722 1843 1655 1903
```

```
1552 1626 1525 1653 1498 1810 1701
1900 1605 1801 1689 1581 1781 1827
1738 1854 2034 1909 1853 1512 1595
1440 1722 1843 1655 1903 1552 1626
1525 1653 1498 1810 1701 1900 1605
1801 1689 1581 1781 1827 1738 1854
2034];
```

```
X2=[600 620 620 680 680 675 685
```

```
690 630 632 645 634 662 670
689 612 628 677 692 700 725
736 750 761 780 765 600 620
620 680 680 675 685 690 630
632 645 634 662 670 689 612
628 677 692 700 725 736 750
761];
```

```
X3=[0.8 0.8145 0.7 0.86 0.90.81 0.9211
```

```
0.68980.8412 0.78 0.8403 0.7208
0.9577 0.7955 0.8897 0.8415 0.9554
0.7888 0.6893 0.80.80.72 0.7547
0.8804 0.80.7398 0.80.8145 0.70.86 0.9
```

```
0.81 0.9211 0.6898 0.8412 0.78
0.8403 0.7208 0.9577 0.7955 0.8897
0.8415 0.9554 0.7888 0.6893 0.80.8
0.72 0.7547 0.8804];
```

```
X4=[0.9 0.9307 0.8894 0.7812 0.90.7915
```

```
0.9302 0.8109 0.8792 0.8113 0.8597
0.8905 0.8107 0.9343 0.9755 0.9281
0.955 0.8951 0.8637 0.93 0.90.8905
0.95 0.90.81 0.9222 0.90.9307 0.8894
0.7812 0.90.7915 0.9302 0.8109
0.8792 0.8113 0.8597 0.8905 0.8107
0.9343 0.9755 0.9281 0.955 0.8951
0.8637 0.93 0.90.8905 0.95 0.9];
```

```
X5=[66.14 62.69 69.44 58.07 54.25 60.42 52.55
```

```
64.43 61.5 65.57 60.49 66.75 55.25 58.78
52.63 62.3 55.52 59.2 63.25 56.14 54.73
57.53 53.93 49.16 52.38 43.17 52.91 50.16
55.55 46.46 43.41 48.34 42.09 51.55 49.2
52.46 48.4 53.4 44.2 47.03 42.11 49.84
44.42 47.36 50.6 44.92 43.79 46.03 43.15
39.33];
```

```
% target for input data values
```

```
T=[387.37 382.27 392.19 375.47 369.84
```

```
333.27 321.78 339.12 334.84 340.79
339.08 348.32 331.37 336.58 327.51
305.83 295.96 301.3 307.16 296.84
280.67 284.79 279.49 272.44 277.21
307.68 324.08 319.47 328.52 313.22
313.86 316.38 305.83 321.8 317.84
323.32 321.82 324.91 309.42 269.07
260.79 273.79 264.69 269.62 275.1
265.52 263.62 267.39 262.55 256.13];
```

```
P=[X1 ; X2 ; X3 ; X4 ; X5 ; X6 ; X7 ; X8 ; X9 ; X10
```

```
; X11 ; X12 ; X13 ; X14 ; X15 ; X16 ; X17];
```

```
pr=minmax(P)
```

```
% creation of a feedforward network type
```

```
net1=newff(pr,[1150 1],[ 'tansig', 'purelin'],'traincgf');
```

```
net1.trainParam.lr=20;
```

```
net1.trainParam.mc=0.5;
```

```
net1.trainParam.min_grad=1e-10;
```

```
net1.trainParam.show=1;
```

```
net1.trainParam.epochs=60;
```

```
net1.trainParam.goal=1;
```

```
% network training
```

```
[net1,tr1]=train(net1,P,T);
```

```
% network testing with data used in its practice
(X1...X17)
```

```
yantr=sim(net1,P);
```

```
% writing in exit_training.txt file of network targets
for
```

```
% practice values (Y) compared with desired target
in training
```

```

% (T), percentage errors for each case (Er), as well as
for % average errors.
ex_train=fopen('exit_practice.txt','w');
s_er=0;
fprintf(ex_train,' Y | T | Er\n');
for i=1:1:50
    er(i)=100-((yantr(i)/T(i))*100);

fprintf(ex_train,'%4.2f\t%4.2f\t%2.2f\n',yantr(i),T(i),
er(i))
    s_er=s_er+er(i);
end
if s_er<0 s_er=s_er*(-1);
end
fprintf(ex_train,'\nAverage error is %2.2f,s_er/50);
fclose(ex_train);

% data uploading from input.txt file
Q=load('input.txt')

% target calculation for the data from input.txt file
rez=sim(net1,Q);

% target writing in exit.txt file
exit=fopen('exit.txt','w');
for i=1:1:size(rez)
fprintf(exit,'%4.2f\n',rez);
end
fclose (exit)
% exit code

```

Thus, it is defined gradually the 17th – dimensional vectors (X1, X2, X3, X4X17) each with 50 components, corresponding to 50 students used to network training; those 17th vectors are corresponding to the four input variables used and namely: the 17th vectors were included in a matrix type structure (pattern) P = [X1 X2 X3 X4.....X17];

- X1 – total number of students;
- X2 – number of students enrolled in the 1st year.
- X3 – the average percentage of graduated students from the first year to the second year of study;
- X4 – the average percentage of graduated students from the second year to the third year of study;
- X5 – the average remuneration for professor job;
- X6 – the average remuneration for the lecturer job;
- X7 - the average remuneration for the lecturer job;
- X8 – the average remuneration for assistant job;
- X9 – the percentage share of professor jobs;
- X10 – the percentage share of lecturer jobs;
- X11- the percentage share of lecturer jobs;
- X12- the percentage share of assistant jobs;
- X13 – the average remuneration for the staff;
- X14 – cost level for material base (per student)
- X15– the average cost for electricity;
- X16 – the average cost per cubic meter for water and heating;
- X17 – the average cost for Internet.

Further, T (target) vector is defined, obviously used, only in case of trending phase, vector which memorize the real situation in case of 50 students under discussion. Newff instruction is used in order to create a feedforward ANN; its syntax is the following:

Net1 = newff [pr, [2 25 1], {'tansig' 'purelin' 'purelin'}, 'traincgf'); in which:

Pr = min max (P), represents encoding matrix for minimum and maximum values for input variables; [2 25 1] – dimensions (numbers of neurons) for hidden layers (in this case two with 2 and respectively 25 neurons) and obviously a neuron for output layer, corresponding to the actual cost.

Tansig and purelin are functions used in transferring.

Trainlm and traincg are functions used for training; in the framework of the train function can be set parameters specific to training session as follows:

- Epochs – the maximum number of training epochs (successive resumption of training session with the same input data for several times);
- Goal – the precision with which the 0.01 target result is reached, representing, for example, an error of 1%.
- Time – maximum time of training in seconds;
- Min_grad – minimum performance gradient;
- Show – the number of epochs between two successive views of training session development;
- By using train and sum functions is achieved the network training itself, for the purpose of learning its features (basically is determined the weight coefficients w_{ij}).

In the second phase is proceeded to the actual operation of the program; so it is defined and initialized the 17th input vectors, X1, X2, X3, X4X17, each containing data of the second set of skills, considered for use, namely network check. Because within this period is done and checked the ANN data prediction accuracy, is also defined Y vector and the real values of each student's situation.

Determined network data will be memorized in network prediction.txt file opened with often function.

Finally, the network use results are tabulated, the rightmost columns of the table having the following meaning:

- Ynetwork – the prediction made by every network;
- Yreal – representing the real situation.

Accuracy – by comparing Ynetwork and Yreal columns is displayed in the last column the accuracy of prediction. In the end it is presented the percentage accuracy of prediction, being in fact the main parameter of assessing prediction results.

4. ANN STRUCTURE AND PARAMETERS OPTIMIZATION TRAINING IN ORDER TO INCREASE PREDICTION ACCURACY OF UNIVERSITY COSTS. QUASI-EMPIRICAL METHODS

The purpose of this subparagraph is that of improving prediction accuracy by optimizing ANN structures of the various parameters specific to ANN and to parameters of training.

It is estimated that it is one of the most difficult problems in the implementation and use of ANN, in many works being considered an art; increased difficulty of this problem is the large number of parameters that can be changed, some of them can not even be quantified, so the choice of the network structure (topology) and chosen training method is about the experience in the field of user, experience that have to be gained by applying the ANN to the most varied cases.

Further will be presented in the author's opinion the main parameters that could compete to optimize the prediction, being removed from the beginning ("in common sense") those considered irrelevant. Obviously it requires a statistical study, using specific methods of removing the relevant parameters:

1. network topology (spread before perception, radial)
2. learning method;
3. non-linearity adopted function;
4. number of hidden layers;
5. number of neurons in each layer;
6. number of training epochs.

If each of the six factors considered aprioristic determinants would be considered only four different levels and would not make additional replicas for a given level of required parameters would be $46 = 4096$ experiments, and therefore would be required the use of statistical methods for planning the experiment.

Within this paragraph will be presented a first organizing of the experiment, being elaborated a quasi-empirical planning, followed by a developing of the first set of preliminary conclusions. There are displayed in the tables, obtained experimental results, indicating that generally have been made three replicas (the successive experiments for each line were kept the same with input parameter values).

5. (SCIENTIFIC) ANALYSIS AND SYNTHESIS OF EXPERIMENTAL DATA OBTAINED USING ARTIFICIAL NEURAL NETWORKS

After processing the results, according to previously mentioned algorithms were achieved the following drawings.

Fig. 1 and Fig. 2 present graphically and in an objective way the function variation (percentage accuracy of prediction), according to the two variables considered in this first phase the number of training sessions and number of layers, more than even applying the method of least squares for the experimental data was obtained a second degree model to predict the following type:

$$\text{Prec} = 43.543 + 11.501 * s + 0.182 * e - 1.542*s^2 + 0.002s*e$$

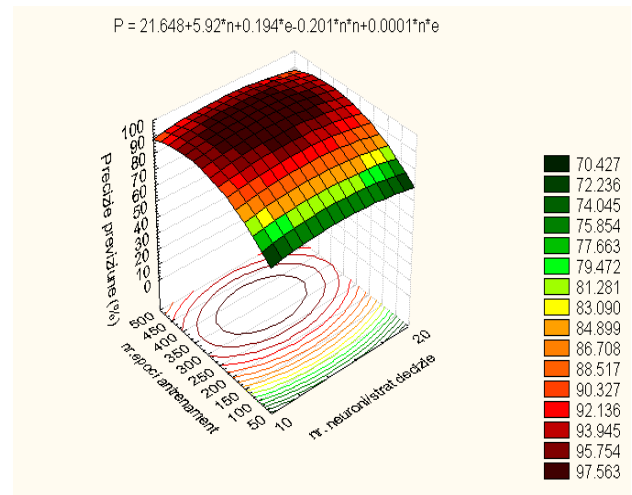


Figure 1. Prediction accuracy

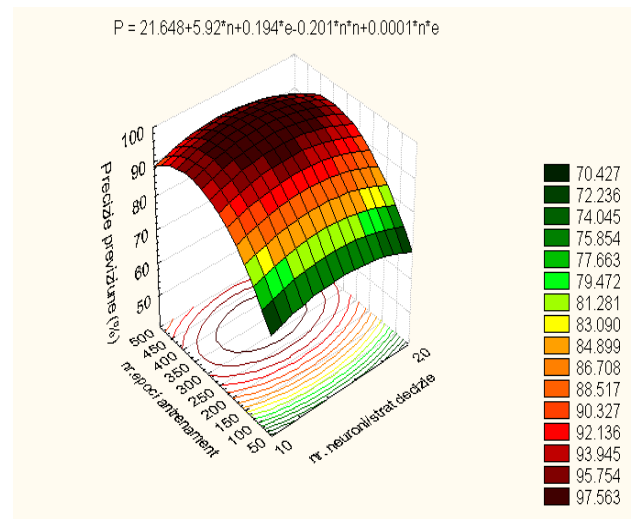


Figure 2. Prediction accuracy

Mainly Fig. 1 and Fig. 2 are identical, but differ in the reference z axis, corresponding to predict accuracy, in Fig. 1 the origin is considered at 0% and in Fig. 2 at 50% level.

Fig. 3 represents the result distributions.

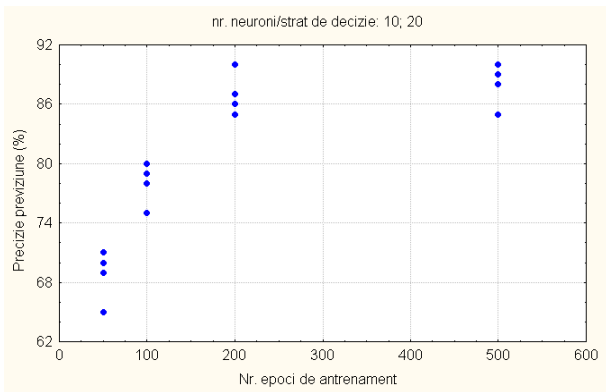


Figure 3 The result distributions.

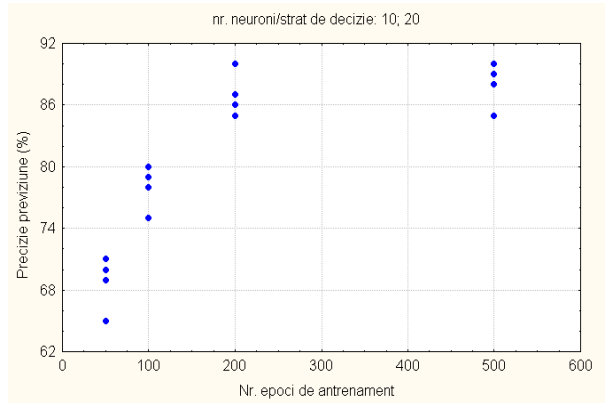


Fig. 4 illustrates graphically and comparative the performance objective function evolution depending on the number of training epochs, the number of layers is considered as the parameter.

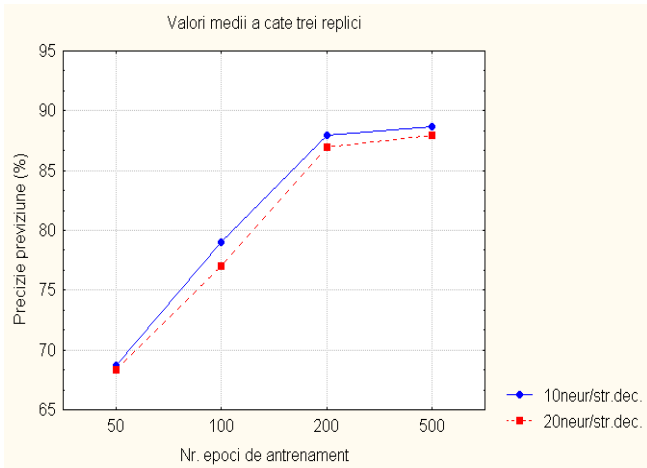


Figure 4 The performance objective function evolution.

Fig. 5, Fig. 6, Fig. 7 are graphically presented the most remarkable statistical quantities used in undertaking research using the following conventions of representation:

- point marks the average position;
- rectangle marks the standard error;
- the difference between the two horizontal segments marks the standard deviation.

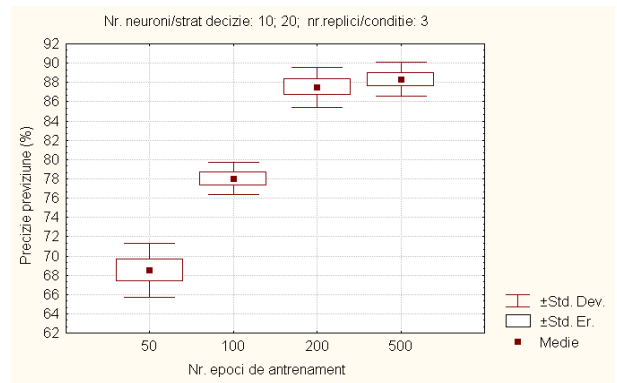


Figure 5 Statistical quantities used in undertaking research.

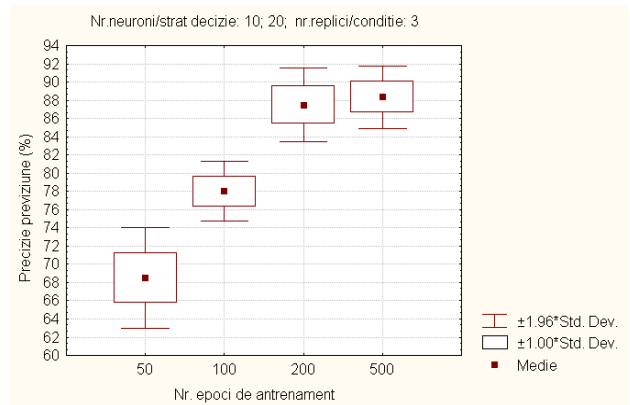


Figure 6 Statistical quantities used in undertaking research.

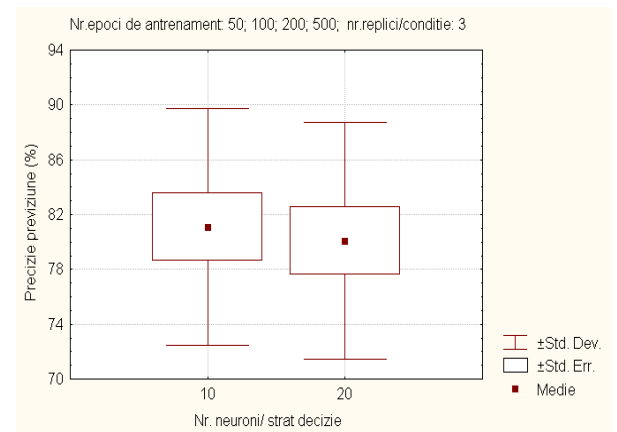


Figure 7 Statistical quantities used in undertaking research.

6. GENERAL CONCLUSION

The use of artificial neural networks for university costs prediction is extremely efficient, the prediction accuracy percentage is higher than in the case of using traditional methods; it also can be considered the existence of a data volume higher than the network training; (it is underlined that the training sessions were made for a set of 50 students).

REFERENCES

- [1] C.D. Căleanu, L. Petropoulakis (2000) *Improved Training of Multilayer Feedforward Neural Networks for Large Input Vectors*, Proc.8-th IEEE Med. Conf. Control and Aut., Rio, Patras ,Greece;
- [2] A. Cichocki, R. Unbehauen, (1993) *Neural Networks for Optimization and Signal Processing*, John Wiley and Sons,;
- [3] H. Demuth, M. Beale, (1998) *Neural Network Toolbox. User's Guide*, ver. 3.0,
- [4] D. Bivolaru (Darvasi), D. Cotleț (2007) *RNA and the possibility of applying them in the financial-accounting field*, Timisoara;
- [5] D. Darvasi, A. Badescu, C. Dobritoiu, F. Molnar, T. Slavici (2011) *Accounting Software Using Expert Systems*, in Proceedings of the fifth Wseas International Conference on Business Administration (ICBA '11) Puerto Morelos, Mexico, pp 97- 102.
- [6] D.Darvasi, (2010) *Contabilitatea fundatiilor si asociatilor prin utilizarea metodelor traditionale sau clasice si a sistemelor inteligentei artificiale*, Fundatiei pentru cultura si invatamant "Ioan Slavici" Publishing House, Timisoara.
- [7] T. Slavici, (2009) *Inteligenta artificiala*, Fundatiei pentru cultura si invatamant "Ioan Slavici" Publishing House, Timisoara.
- [8] T. Slavici, (2006) *Optimizarea management financiar cu ajutorul metodelor inteligentei artificiale*, PhD Thesis, Timisoara.