BULETINUL Universităţii Petrol – Gaze din Ploieşti	Vol. LXI No. 3/2009	325 - 332	Seria Tehnică

A Neural Networks Application for Banking Decision Support

Irina Tudor

Petroleum – Gas University of Ploieşti, 39 Bucureşti Blvd., Ploieşti, ROMÂNIA
e-mail: tirinelle@yahoo.com

Abstract

Decisions making process represents a point of interest in various domains as financial, insurance, communication, industry etc. Manager's activities focused on good decisions evolving for increasing their institutions profitability. Mining data with different algorithms to find significant knowledge to help decision makers is a continuous concern. In this paper an application of neural networks, as a data mining technique is presented. A decision problem in banking domain is formulated: credit approval and the solution given is neural network technique as a decision support.

Key words: *data mining, neural networks, decision, banking.*

Introduction

In banking sector a bad decision can generate a disaster such as bankruptcy in a very short time. The quality of decision is important for decision making process and the managers should have definite answers in most cases. The current challenge for managers is to elaborate and take good decisions in short time, without their direct implication [1]. Automation is considered the base for industrial processes development, and not for the financial, communication, insurance fields. Automation involvement in banking domain comes with data mining techniques to illustrate the practicability of these concepts [5, 6, 7].

The decision to automate a bank function can be considered a decision to adopt an innovation. In this case the innovation involves three aspects: the use of a new product (the software package), if is necessary the use of additional hardware, and organizational changes [1, 2]. One way to measure performance effects of automation is to analyze the average production costs of the automated function. A practical managerial consequence is that managers requesting automation of a function should be very careful with promises of future lower average costs. Executives of banks which are leading in information technology use mention a number of reasons for deployment of information technology: to increase the share of revenues from low-risk insignificant activities, to help employees sell new services more effectively, to make the bank "information richer," to differentiate the bank in the market as a product innovator, to learn the technology before its maturation, and, more general, to gain competitive advantage or avoid competitive disadvantage [3, 4].

The authoress of this paper describes an application of neural networks, as a data mining technique, to create a decision support in credit approval problem.

Neural network solution for credit approval problem

In this paper is presented an example of banking function – credit approval – in the context of decision making, using neural network solution, through existing commercial module software.

NeuroIntelligence represents a data mining software package based on neural networks to solve prediction, classification and approximation problems. NeuroIntelligence is provided by Alyuda Research consortium [8] and consists in complete software for neural networks design and optimization. The authoress used the demo version in current paper.

NeuroIntelligence supports a various application in neural networks field, being used for:

1. analyzing and preprocessing data sets;
2. finding the best neural networks architecture;
3. testing and optimizing the selected network;
4. applying the network considered to solve a problem from various domains like education, finance, medical fields, assurance etc.

The following steps are necessary to design and use a new neural network:

- loading the data input set;
- selecting and marking the target column;
- designing the neural network (establishing the number of hidden layers and the number of neurons, as well as the type of training functions);
- training the network;
- testing the network.

NeuroIntelligence owns a user interface to rapid access the designing steps and the application form of the network generated. Selecting a simple option it can be activated the correspondent panel for each working step (fig.1).

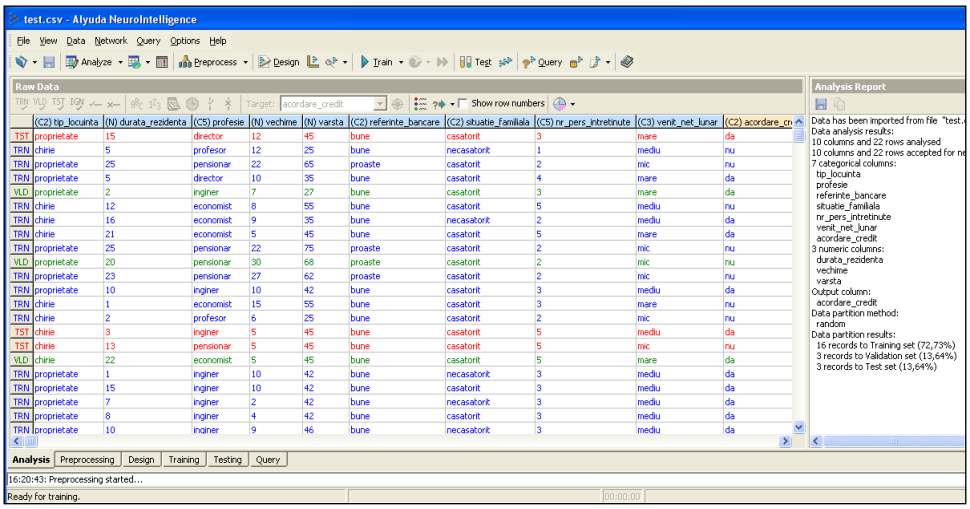


Fig.1. NeuroIntelligence GUI.

In the analyze phase are made some operations such as: columns identification, anomalies detection, separation of test data, validation data, defining the target variable etc. In preprocessing phase you can observe the intern representation of data input. In this phase the numeric values are modified referring to an accepted range and the columns with symbolic

value and date/hour value are codified. To the next level the designing of network is made. The best architecture can be generated by a series of searching algorithms, the network architecture can be modified, and the activation functions are established as well as the range of errors, classification model and/or accepting/excluding levels. In the training step, the neural network dynamic training process is traced by visualizing the training charts, error distribution and weights histograms. After training phase, the network is tested and the results are compressed in the result chart, the actual and compute values being compared. A better image of result is offered by the answer chart and confusion matrix. Finally, the designed network can be queried (manual or automat queries) to obtained answers of a new data set.

An objective of a bank called TCreditBank is to minimize the credit risk. A batch of economic analysis is made to discover knowledge regarding the bank range risk. The authoress considers that data mining techniques can simplify the financial experts work through decrease time of working, increase the decisions accurate, predicting the loan risk associated to the new clients etc., deploying the example below. The data set consists in economic and demographic data corresponding to the TCreditBank clients.

The variables used in this case are the following: *vârsta, timp reşedinţă, vechimea, venit net lunar, număr credite în derulare, număr persoane în întreţinere* with numerical values and *sex (feminin, masculin), stare civilă(căsătorit, necăsătorit), profesia de bază(doctor, inginer, economist, profesor, student), conturi curente(da, nu), depozite(da, nu), tip locuinţă(proprietate, chirie), posesor maşină(da, nu), referinţe bancare(bune, proaste), acordare credit (da, nu)* with nominal values.

The authoress considered as input variables the parameters $(1 \div 14)$ which represent the neurons in the input layer, and the target variable (15) which is marked for the neuron founded in the output layer. The input data set is recorded in the ***creditbank.CSV file*** (fig.2).

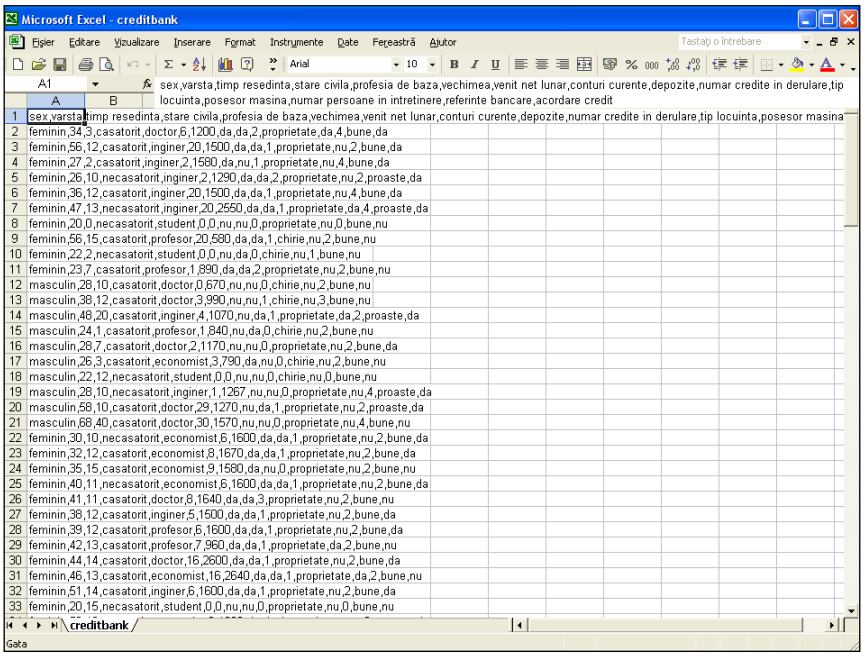


Fig.2. The input file *creditbank.csv.*

Various types of neural networks architecture were tested with one, two or three hidden layers, activation function being the sigmoid function. Number of iteration was: 50, 100, 150, 200, 500 and 1000.

An example of neural network tested in this case is [18:9:1], that means 18 neurons in the input layer, 9 neurons in the hidden layer and only one neuron in the output layer (fig.3).

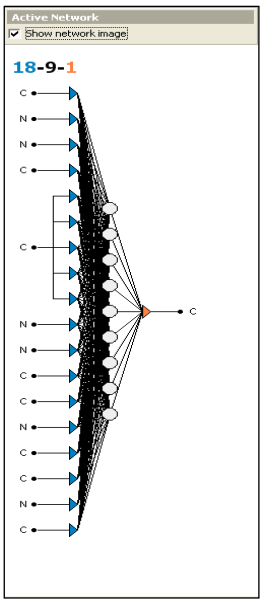


Fig.3. The [18:9:1] neural network architecture.

The analysis and preprocessing phases are followed by network training, an important step consisting in parameters assignment: training algorithm and its parameters, the stop criteria for training process. As a result of training, error, weights distribution and error distribution charts are provided, as well as the table with training data set (number of iterations, CCR etc.) (fig. 4).

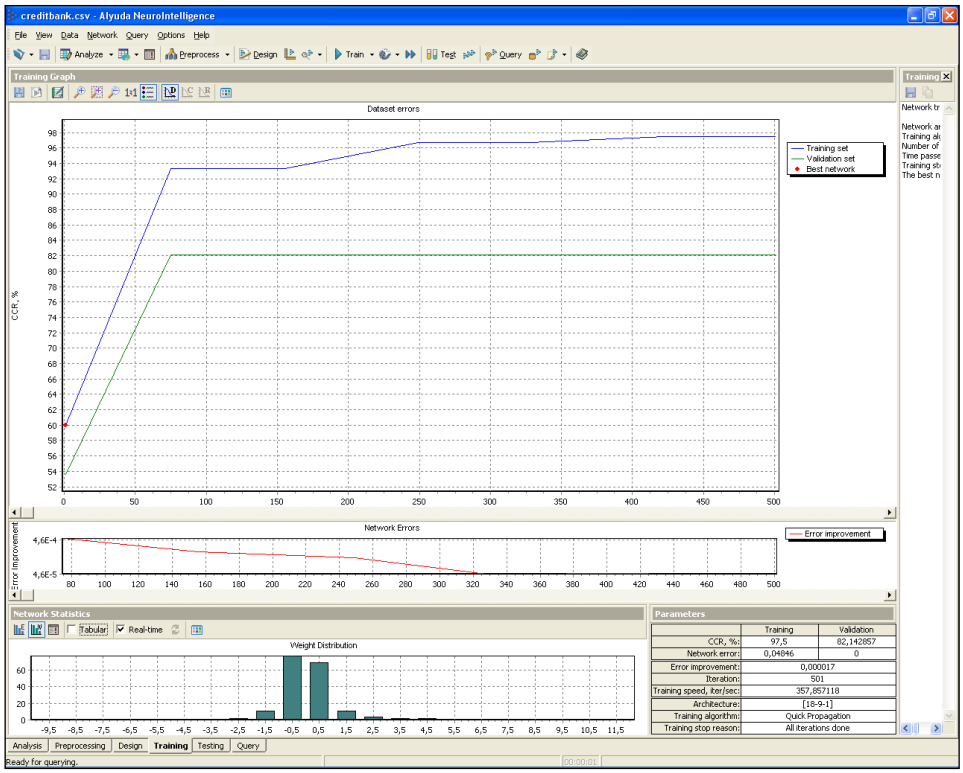
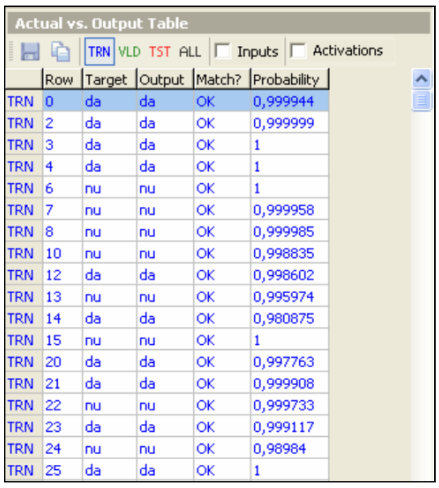


Fig. 4. The graphic results.

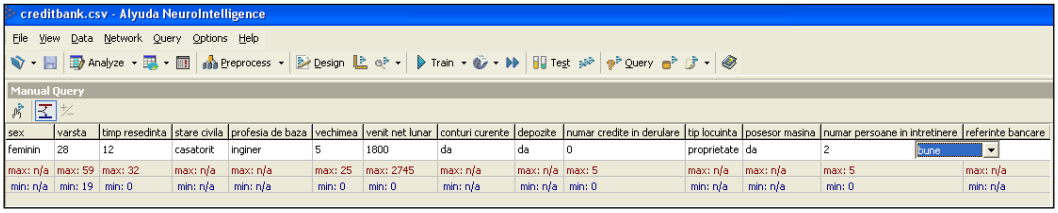
A comparatively table is generated to visualize the current values and the computed values, regarding the target variable ***loan_approval*** (fig.5)**.**



	Row	Target	Output	Match?	Probability
TRN	0	da	da	OK	0,999944
TRN	2	da	da	OK	0,999999
TRN	3	da	da	OK	1
TRN	4	da	da	OK	1
TRN	6	nu	nu	OK	1
TRN	7	nu	nu	OK	0,999958
TRN	8	nu	nu	OK	0,999985
TRN	10	nu	nu	OK	0,998835
TRN	12	da	da	OK	0,998602
TRN	13	nu	nu	OK	0,995974
TRN	14	da	da	OK	0,980875
TRN	15	nu	nu	OK	1
TRN	20	da	da	OK	0,997763
TRN	21	da	da	OK	0,999908
TRN	22	nu	nu	OK	0,999733
TRN	23	da	da	OK	0,999117
TRN	24	nu	nu	OK	0,98984
TRN	25	da	da	OK	1

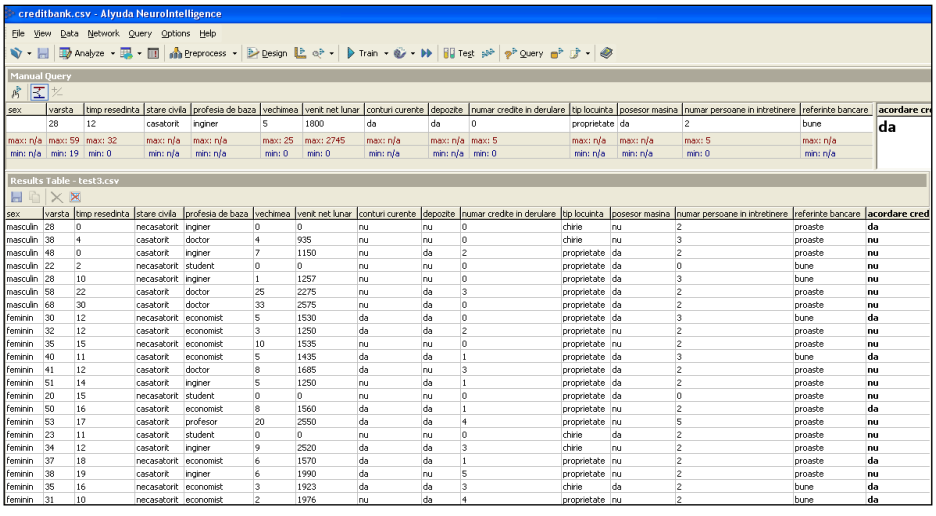
Fig. 5. Current values vs. Compute values.

The queries in manual or automatic way can be launch to the designed neural network, having new records for the variables used to solve the problem of loan approval (fig. 6, fig. 7).



sex	varsta	timp resedinta	stare civila	profesia de baza	vechimea	venit net lunar	conturi curente	depozite	numar credite in derulare	tip locuinta	posesor masina	numar persoane in intretinere	referinte bancare
feminin	28	12	casatorit	inginer	5	1800	da	da	0	proprietate	da	2	bune
max: n/a	max: 59	max: 32	max: n/a	max: n/a	max: 25	max: 2745	max: n/a	max: n/a	max: 5	max: n/a	max: n/a	max: 5	max: n/a
min: n/a	min: 19	min: 0	min: n/a	min: n/a	min: 0	min: 0	min: n/a	min: n/a	min: 0	min: n/a	min: n/a	min: 0	min: n/a

Fig.6. Manual query.



sex	varsta	timp resedinta	stare civila	profesia de baza	vechimea	venit net lunar	conturi curente	depozite	numar credite in derulare	tip locuinta	posesor masina	numar persoane in intretinere	referinte bancare	acordare cred
masculin	28	0	necasatorit	inginer	0	0	nu	nu	0	chirie	nu	2	proaste	da
masculin	38	4	casatorit	doctor	4	935	nu	nu	0	chirie	nu	3	proaste	nu
masculin	48	0	casatorit	inginer	7	1150	nu	da	2	proprietate	da	2	proaste	nu
masculin	22	2	necasatorit	student	0	0	nu	nu	0	proprietate	da	0	bune	nu
masculin	28	10	necasatorit	inginer	1	1257	nu	nu	0	proprietate	da	3	bune	nu
masculin	58	22	casatorit	doctor	25	2275	nu	da	3	proprietate	da	2	proaste	nu
masculin	68	30	casatorit	doctor	33	2575	nu	da	0	proprietate	da	2	proaste	nu
feminin	30	12	necasatorit	economist	5	1530	da	da	0	proprietate	da	3	bune	da
feminin	32	12	casatorit	economist	3	1250	da	da	2	proprietate	nu	2	proaste	nu
feminin	35	15	necasatorit	economist	10	1535	nu	nu	0	proprietate	nu	2	proaste	nu
feminin	40	11	casatorit	economist	5	1435	da	da	1	proprietate	da	3	bune	da
feminin	41	12	casatorit	doctor	8	1685	da	nu	3	proprietate	da	2	proaste	nu
feminin	51	14	casatorit	inginer	5	1250	nu	da	1	proprietate	da	2	proaste	nu
feminin	20	15	necasatorit	student	0	0	nu	nu	0	proprietate	da	0	proaste	nu
feminin	50	16	casatorit	economist	8	1560	da	da	1	proprietate	nu	2	proaste	da
feminin	53	17	casatorit	profesor	20	2650	da	da	4	proprietate	nu	5	proaste	nu
feminin	23	11	casatorit	student	0	0	nu	nu	0	chirie	da	2	proaste	nu
feminin	34	12	casatorit	inginer	9	2520	da	da	3	chirie	nu	2	proaste	nu
feminin	37	18	necasatorit	economist	6	1570	da	da	1	proprietate	nu	2	proaste	da
feminin	38	19	casatorit	inginer	6	1990	da	nu	5	proprietate	nu	2	proaste	nu
feminin	35	16	necasatorit	economist	3	1923	da	da	3	chirie	da	2	bune	da
feminin	31	10	necasatorit	economist	2	1976	nu	da	4	proprietate	nu	2	bune	da

Fig. 7. Automated query.

Finally (fig. 8), the best architecture for the network is given using the facility offered by the NeuroIntelligence software.

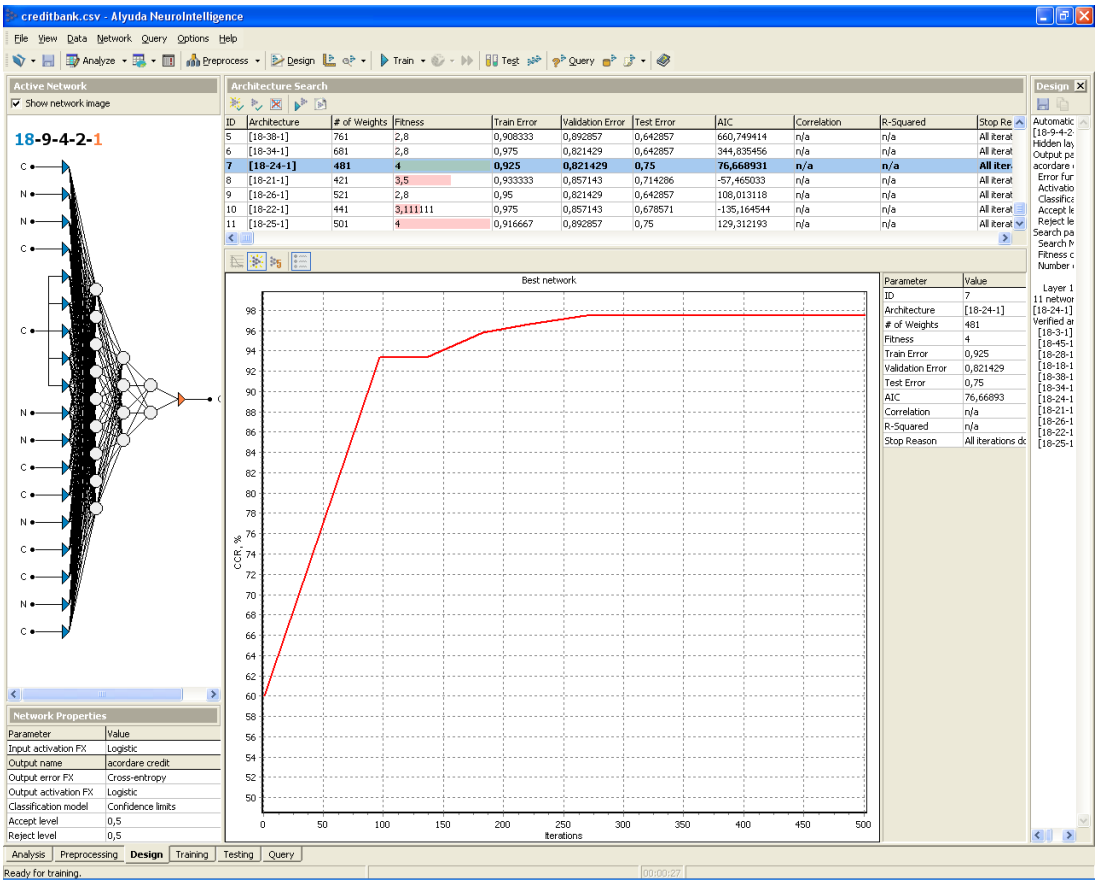


Fig. 8. The best neural network architecture generation.

Results and comments

The input data set consists in 180 records, stored in the *creditbank.csv file*. Only 176 records were validated for network training process. Input variables were classified in two classes:

- 9 nominal variables (*sex, stare civilă, profesia de bază,conturi curente, depozite, tip locuinţă, posesor maşină, referinţe bancare, acordare credit);*
- 6 numerical variables (*vârsta, timp reşedinţă, vechimea, venit net lunar, număr credite în derulare, număr persoane în întreţinere)*.

Incorrectness of records determines to eliminate 4 of the training records.

Partition method used in this case is random, with the following results:

- 120 record for the training set (68,18%);
- 28 records for validation set (15,91%);
- 28 records for testing set (15,91%).

The values for CCR parameter after testing three types of neural networks with one, two and three hidden layers are given by the table below.

Table 1. CCR – first example of tests

Number of iterations / Network type	[18:9:1]	[18:9:4:1]	[19:9:4:2:1]
50	43.33%	57.5%	42.5%
100	97.5%	93.33%	90.33%
150	96.67%	95.83%	94.16%
200	97.5%	95.83%	87.5%
500	98.33%	96.67%	98.33%
1000	98.33%	98.33%	94.16%

The mean value for CCR is 95%.

The clients number correct classified increases with the iteration number. On the other hand, after a number of iteration, CCR remains at the same value (98,33%), that implies a network trained.

Complexity of the network determines more iteration for the training process. If the number on input variable decreases, the value of CCR increases.

Table 2. CCR - the second example of tests

Number of iterations / Network type	[17:8:1]	[17:8:4:1]	[17:8:4:2:1]
50	57.5%	57.5%	57.5%
100	94.16%	88.33%	86.67%
150	95%	96.67%	92.5%
200	96.67%	96.67%	95.83%
500	98.33%	98.33%	95%
1000	98.33%	96.67%	98.33%

In order to optimize the network architecture, the authoress' future work will consist in finding a method to assign the best parameters to a network to increase the number of correct classification rate. Using the principal component analysis (PCA) the best architecture can be found for the approval loan problem.

Conclusions

Decision making process in banking field (credit approval problem) by applying data mining techniques can improve the bank activities by:

- reducing the decision time and minimizing the labor cost;
- increasing decision resources to improve decision precision;
- improving speed, quality and consistency across manual decision processes;
- using available data resources effectively;
- achieving significant returns;
- increasing approval rates;
- offering clarity in risk management;
- reducing application processing costs.

The neural network approach demonstrates that automation can be applied with success in other domains, not only industrial fields, to assist the decision making process in order to supply decision tools. In other words, data mining can provide major premises to adopt automation concept in banking sector or other domains such as assurance, financial, education, telecommunication etc.

References

1. Alpar, P. – *Automation of Banking Functions and Its Managerial Implications*. IEEE Transaction On Management, Vol. 39, no. 4, pp. 378-385, 1992.
2. Bucklin, R.E., et al – *Decision Support to Decision Automation: A 2020 Vision*. MSI working paper, 1998.
3. Davenport, T. – *Decision Evolution.* CIO Magazine, 2004, http://dssresources-com.c.topica.com/maac15DabcXM5cgnx0nbafpOmV/
4. Gan, C., et al – *Consumer Choice Prediction: Artificial Neural Networks versus Logistic Models.* Journal of Social Sciences 1 (4), pp. 211-219, 2005.
5. James, R.C., Carol, E.B. – *Artificial neural networks in accounting and finance: modeling issues*. Intl. J. Intelligent Systems in Accounting, Finance and Management, 9, pp.119-144, 2000.
6. Koh, H.C., Low, C.K. – *Going concern prediction using data mining techniques.* Managerial Auditing Journal 19 (3), pp.462-476, 2004.
7. Walczak, S. – *An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks*. Journal of Management Information Systems, (17.4), pp. 203-222, 2001.
8. http://www.alyuda.com.

O aplicaţie a reţelelor neuronale ca suport decizional bancar

Rezumat

Procesul de elaborare a deciziilor reprezintă un punct de interes în diverse domenii, spre exemplu domeniul financiar, asigurări, comunicaţii, industrie. În ultimii ani, activităţile managerilor s-au focalizat asupra luării unor bune decizii pentru instituţiilor lor în scopul profitabilităţii acestora. Minarea datelor cu diferiţi algoritmi pentru a descoperi cunoştinţe semnificative în scopul sprijinirii procesului decizional reprezintă o continuă preocupare. În această lucrare este prezentată o aplicaţie a reţelelor neuronale ca tehnici de data mining. Fiind formulată o problemă decizională în domeniul bancar, şi anume aprobarea creditelor, este oferită o soluţie prin intermediul reţelelor neuronale ca suport decizional.