BULETINUL	Vol. LXI	225 222	Corio Tolorio×
Universității Petrol – Gaze din Ploiești	No. 3/2009	323 - 332	Seria Tehnică

A Neural Networks Application for Banking Decision Support

Irina Tudor

Petroleum – Gas University of Ploiești, 39 București Blvd., Ploiești, 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).

Ŵ	- 📙 📑 Analy	2e • 🔜 • 🔟 🚮	Preprocess +	<mark>⊵</mark> <u>D</u> esign	• اې ا	🕨 Irain 🔹 🍈 - 🕅	🕴 😼 Te <u>s</u> t 🥬 💡	Query 📄 📝 🔹 🧇			
Raw	Data										Analysis Report
TRN	VLD TST IGN 🚈	×- 180 123 🐼 (D ł ł 🗄	Target: acor	dare_credit		+ 🖂 Show row numb	ers 🕒 🗸			.
	(C2) tip_locuinta	(N) durata_rezidenta	(C5) profesie	(N) vechime	(N) varsta	(C2) referinte_bancare	(C2) situatie_famíliala	(C5) nr_pers_intretinute	(C3) venit_net_lunar	(C2) acordare_cn 🔨	Data has been imported from file "te
TST	proprietate	15	director	12	45	bune	casatorit	3	mare	da 🔳	Data analysis results: 10 columns and 22 rows analysed
TRN	chirie	5	profesor	12	25	bune	necasatorit	1	mediu	nu	10 columns and 22 rows analysed 10 columns and 22 rows accepted for
TRN	proprietate	25	pensionar	22	65	proaste	casatorit	2	mic	nu	7 categorical columns:
TRN	proprietate	5	director	10	35	bune	casatorit	4	mare	da	tip_loculta profesie
VLD	proprietate	2	inginer	7	27	bune	casatorit	3	mare	da	referinte bancare
IRN	chirie	12	economist	8	55	bune	casatorit	5	mediu	nu	situatie_familiala
		16	economist	9	35	bune		2	mediu	da	nr_pers_intretinute venit net lunar
RN	chirie	21	economist	5	45	bune	casatorit	5	mare	da	acordare credit
RN	proprietate	25	pensionar	22	75	proaste		2	mic	nu	3 numeric columns:
٨D	proprietate	20	pensionar	30	68	proaste	casatorit	2	mic	nu	durata_rezidenta
'RN	proprietate	23	pensionar	27	62	proaste	casatorit	2	mic	nu	vechime varsta
IRN	proprietate	10	inginer	10	42	bune	casatorit	3	mediu	da	Output column:
IRN	chirie	1	economist	15	55	bune	casatorit	3	mare	nu	acordare_credit
IRN	chirie	2	profesor	6	25	bune	casatorit	2	mic	nu	Data partition method: random
TST	chirie	3	inginer	5	45	bune	casatorit	5	mediu	da	Data partition results:
TST	chirie	13	pensionar	5	45	bune	casatorit	5	mic	nu	16 records to Training set (72,73%
٨D	chirie	22	economist	5	45	bune	casatorit	5	mare	da	3 records to Validation set (13,64% 3 records to Test set (13,64%)
RN	proprietate	1	inginer	10	42	bune	necasatorit	3	mediu	da	5 records to rest set (13,64%)
RN	proprietate	15	inginer	10	42	bune	casatorit	3	mediu	da	
IRN	proprietate	7	inginer	2	42	bune	necasatorit	3	mediu	da	
RN	proprietate	8	inginer	4	42	bune	casatorit	3	mediu	da	
RN	proprietate	10	inginer	9	46	bune	necasatorit	3	mediu	da 🗠	
¢ 🗉										>	<
	vsis Preprocess	ing Design Trainir	na Testina	Query							

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 reşedință, vechimea, venit net lunar, număr credite în derulare, număr persoane în întreținere 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 locuință(proprietate, chirie), posesor maşină(da, nu), referințe 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).

Microsoft Excel - creditbank												
Eişier Editare Yizualizare Inserare Format Instrymente Date F	ereastră	Ajutor						Tasta	ți o întreba	re	• -	8
🚰 🖬 🚑 👌 🗠 - Σ - ી 📶 🕐 🐥 Arial	v 10	- B /	υ	E 8	= 🖬	9	% 000	18 408	€E €E	- 60	ð - 1	Α.
A1 • f> sex,varsta,timp resedinta,stare civila,profe												
A B locuinta, posesor masina, numar persoane							enite,ue	spozite,ii	umai cie	uite in	ueruiare	s,up
sex varsta timp resedinta, stare civila, profesia de baza, vechimea, ven							ito in c	lorularo t	n locuint	o noco	cor mo	eina
feminin ,34,3, casatorit, doctor,6,1200, da, da,2, proprietate, da,4, bune, o		ar,contur	curem	e,depoi	cite,numa	i cieu	ite in c	erulare,	piocum	a,pose	sor ma	Silia
feminin, 56,12, casatorit, inginer, 20, 1500, da, da, 1, proprietate, au, 4, bune, c			-									
feminin, 27, 2, casatorit, inginer, 2, 1580, da, nu, 1, proprietate, nu, 4, bune,			-									-
feminin 26,10, necasatorit, inginer 2,1290, da, da, 2, proprietate, nu, 2, pro			-									-
feminin, 36, 12, casatorit, inginer, 20, 1500, da, da, 1, proprietate, nu, 4, bun			-									-
feminin, 47, 13, necasatorit, inginer, 20, 2550, da, da, 1, proprietate, id, 4, ban			-									-
feminin, 20,0, necasatorit, student, 0,0, nu, nu, 0, proprietate, nu,0, bune, r			-						-			-
feminin, 56, 15, casatorit, profesor, 20, 580, da, da, 1, chirie, nu, 2, bune, nu												+
feminin, 22,2, necasatorit, student, 0,0, nu, da,0, chirie, nu,1, bune, nu									-			-
feminin 23.7. casatorit profesor.1.890.da.da.2. proprietate.nu.2. bune.	nu		-			-						-
masculin 28,10, casatorit, doctor 0,670, nu, nu, 0, chirie, nu, 2, bune, nu												
masculin ,38,12, casatorit, doctor, 3,990, nu, nu, 1, chirie, nu, 3, bune, nu												+
masculin 48.20 casatorit inginer 4.1070 nu da 1 proprietate da 2 pro	aste da		-									-
masculin.24.1.casatorit.profesor.1.840.nu.da.0.chirie.nu.2.bune.nu			-									
masculin ,28,7, casatorit, doctor ,2,1170, nu, nu, 0, proprietate, nu, 2, bune	s da											-
masculin 26.3 casatorit economist 3.790 da nu 0 chirie nu 2 bune n												+
masculin, 22, 12, necasatorit, student, 0, 0, nu, nu, 0, chirie, nu, 0, bune, nu												-
masculin,28,10,necasatorit,inginer,1,1267,nu,nu,0,proprietate,nu,4,	nroaste di	9										
masculin 58,10, casatorit, doctor, 29,1270, nu, da, 1, proprietate, nu, 2, pr		-										+
masculin .68 .40 .casatorit .doctor .30 .1570 .nu .nu .0 .proprietate .nu .4 .bu												-
feminin, 30, 10, necasatorit, economist, 6, 1600, da, da, 1, proprietate, nu, 2												-
feminin, 32, 12, casatorit, economist, 8, 1670, da, da, 1, proprietate, nu, 2, b												-
feminin 35,15, casatorit, economist, 9,1580, da, nu 0, proprietate, nu, 2, b												
feminin,40,11,necasatorit,economist,6,1600,da,da,1,proprietate,nu,2												
feminin 41,11, casatorit, doctor,8,1640, da, da,3, proprietate, nu,2, bune												
feminin 38,12, casatorit, inginer,5,1500, da, da, 1, proprietate, nu,2, bune												
feminin, 39, 12, casatorit, profesor, 6, 1600, da, da, 1, proprietate, nu, 2, bun												
feminin, 42, 13, casatorit, profesor, 7, 960, da, da, 1, proprietate, da, 2, bune												
feminin 44,14, casatorit doctor 16,2600, da, da, 1, proprietate, nu, 2, bun												
feminin 46,13, casatorit, economist, 16,2640, da, da, 1, proprietate, da, 2,												
feminin,51,14,casatorit,inginer,6,1600,da,da,1,proprietate,nu,2,bune												
feminin 20,15, necasatorit, student 0,0, nu, nu,0, proprietate, nu,0, bune												
↓ > H\creditbank /				4		1			1		1	F
												4

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).

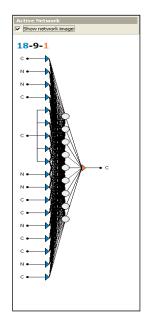


Fig.3. The [18:9:1] neural network architecture.

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).

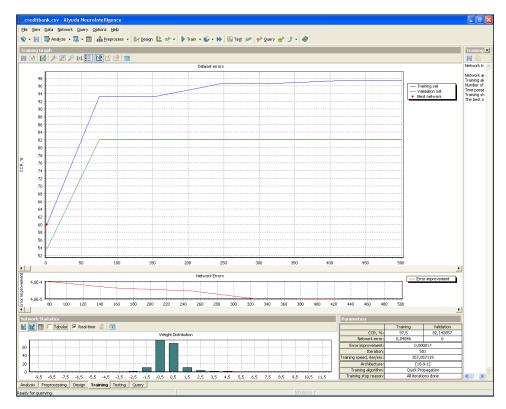


Fig. 4. The graphic results.

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

Act	ual vs	s. Outpu	t Table				
	Pa	TRN VL	D TST AL	.L 🕅 Ir	puts 🗌 Ac	tivations	
	Row	Target	Output	Match?	Probability		^
TRN	0	da	da	ОК	0,999944		
TRN	2	da	da	ОК	0,999999		
TRN	3	da	da	ОК	1		
TRN	4	da	da	ОК	1		
TRN	6	nu	nu	ОК	1		
TRN	7	nu	nu	ОК	0,999958		
TRN	8	nu	nu	ОК	0,999985		
TRN	10	nu	nu	ОК	0,998835		
TRN	12	da	da	ОК	0,998602		
TRN	13	nu	nu	ОК	0,995974		
TRN	14	da	da	ОК	0,980875		
TRN	15	nu	nu	ОК	1		
TRN	20	da	da	ОК	0,997763		
TRN	21	da	da	ок	0,999908		
TRN	22	nu	nu	ок	0,999733		
TRN	23	da	da	ок	0,999117		
TRN	24	nu	nu	ок	0,98984		
TRN	25	da	da	ОК	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).

達 credi	creditbank.csv - Alyuda NeuroIntelligence														
<u>F</u> ile ⊻je	Elle Yew Qata Network Query Options Help														
🔊 • F	🕥 + 🔡 🎝 Analyze + 😨 + 🔟 🍶 Preprocess + 🖻 Design 🖆 🛷 + 🕨 Train + 🎯 + 🕨 💱 Trejt 🥓 🥐 Query 💣 🔅 + 🛷														
Manual	Manual Query														
8 3	1/2														
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.

File Vir	w Date	Network Qu	www. Online	e Helo										
	-													
🔊 - I	3 🗊 ·	Analyze 🝷 🔣	• 🔲 🧥	Preprocess 🔹 🚦	→ Design	🛓 📌 🔹 🕨	Train 👻 🦭 🖌	🕨 👬 T	ie <u>s</u> t 🕪 🦻 Query 💼	r 🗊 • 🝭	E			
Manual	Ouosu													
			_		_			_		_				_
βI	-													
sex	varsta		-									numar persoane in intretinere		
	28	12	casatorit	inginer	5	1800	da	da	0	proprietat	e da	2	bune	da
nax: n/a	max: 5	9 max: 32	max: n/a	max: n/a	max: 25	max: 2745	max: n/a	max: n/a	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	
📕 🖣	X	C .												
5ex		timp resedinta	tare civila	profesia de baza	vechimea	venit net lunar	conturi curente	depozite	putoar credite in deculare	tip locuipta	nosesor masina	numar persoane in intretinere	eferinte bancare	acordare cre
nasculin			necasatorit						0					da
nasculin				doctor					0					 nu
nasculin	48	D	asatorit	indiner	7	1150	nu	da	2	proprietate	da	2	proaste	nu
nasculin	22	2	necasatorit	student	0	0	nu	nu		proprietate		0	oune	nu
nasculin	28	10	necasatorit	inginer	1	1257	nu	nu	0	proprietate	da	3	oune	nu
nasculin	58	22	asatorit	doctor	25	2275	nu	da	3	proprietate	da	2	proaste	nu
masculin	68	30	asatorit	doctor	33	2575	nu	da	0	proprietate	da	2	proaste	nu
eminin	30	12	necasatorit	economist	5	1530	da	da	0	proprietate	da	3	oune	da
eminin	32	12	asatorit	economist	3	1250	da		2	proprietate			proaste	nu
eminin		15	necasatorit	economist				nu	0	proprietate			proaste	nu
eminin				economist					1	proprietate			oune	da
eminin				doctor					3	proprietate			proaste	nu
				inginer	-				1	proprietate				nu
eminin									0	proprietate				nu
eminin				economist					1	proprietate				da
				profesor					4	proprietate				nu
eminin				student	-				0					nu
eminin				inginer					3					nu J-
				economist	-				5	proprietate				da
eminin	30	19	:asatorit	inginer						proprietate				nu
eminin	35	16	necasatorit	economist	3	1923	da	da	3	chirie	da	2	hune	da

Fig. 7. Automated query.

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

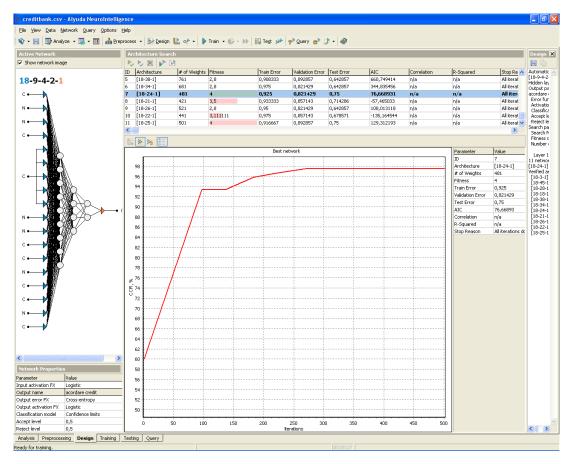


Fig. 8. The best neural network architecture generation.

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 locuință, posesor maşină, referințe bancare, acordare credit);*
- 6 numerical variables (vârsta, timp reședință, vechimea, venit net lunar, număr credite în derulare, număr persoane în întreținere).

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.

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%

 Table 1. CCR – first example of tests

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.

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%

 Table 2. CCR
 - the second example of tests

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://dssresourcescom.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.
- 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.
- 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 aplicație a rețelelor neuronale ca suport decizional bancar

Rezumat

Procesul de elaborare a deciziilor reprezintă un punct de interes în diverse domenii, spre exemplu domeniul financiar, asigurări, comunicații, industrie. În ultimii ani, activitățile managerilor s-au focalizat asupra luării unor bune decizii pentru instituțiilor lor în scopul profitabilității acestora. Minarea datelor cu diferiți algoritmi pentru a descoperi cunoștințe semnificative în scopul sprijinirii procesului decizional reprezintă o continuă preocupare. În această lucrare este prezentată o aplicație a rețelelor neuronale ca tehnici de data mining. Fiind formulată o problemă decizională în domeniul bancar, și anume aprobarea creditelor, este oferită o soluție prin intermediul rețelelor neuronale ca suport decizional.