Application of Neural Networks in Banking Sector

Irina Tudor\*

\* Petroleum-Gas University, Ploieşti

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. 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.

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 [4] 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.

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:

1. sex {male, female}
2. age (numeric)
3. time\_of\_residence (numeric)
4. marital\_status {married, single}
5. job\_title {doctor, engineering, economist, teacher, student}
6. work\_time (numeric)
7. mounth\_wage (numeric)
8. current\_account {yes, no}
9. banking\_deposit {yes, no}
10. number\_current\_loans (numeric)
11. house\_type {owner, lodger}
12. car {yes, no}
13. number\_family\_members (numeric)
14. banking\_references {good, bad}
15. ***loan\_approval*** {yes, no}

The authoress considered as input variables the parameters (114) 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***.

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.

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.).

A comparatively table is generated to visualize the current values and the computed values, regarding the target variable ***loan\_approval*** .

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.

The best architecture for the network is given using the facility offered by the NeuroIntelligence software.

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, marital\_status, job\_title, current\_accounts, banking\_deposits,house\_type, car, banking\_references, loan\_approval);*
  + 6 numerical variables (age*, time\_residence, work\_time, mounth\_wage, number\_current\_loanns, number\_family\_members).*

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 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.

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, November 1992, pp. 378-385
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/](http://dssresources.com/faq/.c.topica.com/maac15DabcXM5cgnx0nbafpOmV/)
4. \*\*\*, <http://www.alyuda.com>