*Interdisciplinary Description of Complex Systems* 14(2), 116-124, 2016

# PREDICTING CUSTOMER CHURN IN BANKING INDUSTRY USING NEURAL NETWORKS

**Alisa Bilal Zorić***

University of Applied Sciences Baltazar Zaprešić
Zaprešić, Croatia

DOI: 10.7906/indecs.14.2.1
Regular article

*Received:* 2 November 2015.
*Accepted:* 10 March 2016.

## ABSTRACT

The aim of this article is to present a case study of usage of one of the data mining methods, neural network, in knowledge discovery from databases in the banking industry. Data mining is automated process of analysing, organization or grouping a large set of data from different perspectives and summarizing it into useful information using special algorithms. Data mining can help to resolve banking problems by finding some regularity, causality and correlation to business information which are not visible at first sight because they are hidden in large amounts of data. In this paper, we used one of the data mining methods, neural network, within the software package Alyuda NeuroInteligence to predict customer churn in bank. The focus on customer churn is to determinate the customers who are at risk of leaving and analysing whether those customers are worth retaining. Neural network is statistical learning model inspired by biological neural and it is used to estimate or approximate functions that can depend on a large number of inputs which are generally unknown. Although the method itself is complicated, there are tools that enable the use of neural networks without much prior knowledge of how they operate. The results show that clients who use more bank services (products) are more loyal, so bank should focus on those clients who use less than three products, and offer them products according to their needs. Similar results are obtained for different network topologies.

## KEY WORDS

data mining, neural network, banking, customer churn

## CLASSIFICATION

JEL: C45, G21

*Corresponding author, $\eta$: alisa.bilal.zoric@bak.hr; +385 98 842 314;
University of Applied Sciences Baltazar Zaprešić, Vladimira Novaka 23, 10290 Zaprešić, Croatia

## INTRODUCTION

With increased availability of data, inexpensive storage and processing power, the amount of raw data stored in banking databases is huge and constantly increasing. However, raw data by itself does not provide much information. Data mining is used to discover patterns and relationships in data in order to improve business decision processes. Its tools can answer business questions that in the past were too time consuming to resolve. We can define it as an interdisciplinary field that brings together techniques from machine learning, pattern recognition, statistics, database systems, data visualization, information theory, knowledge acquisition, artificial intelligence and neural networks [1]. Specific uses of data mining include: Market segmentation, Customer churn, Fraud detection, Direct marketing, Interactive marketing, Market basket analysis, Trend analysis, Credit analysis, Predicting payment default, etc [2].

In this paper, we will focus on Customer churn. Techniques that are most commonly used to predict customer churn are: neural networks, support vector machines and logistic regression models [3]. We want to make a model from stored customer data to predict churn and to prevent the customer’s turnover. Data mining research literature suggests that machine learning techniques, such as neural networks should be used for non-parametric datasets, because they often outperform traditional statistical techniques such as linear and quadratic discriminant analysis approaches [4].

In the era of globalization and intens competition in banking industry, banks are forced to fight more creatively and proactively to gain and maintain their clients. Questions data mining can answer are:

- What transactions does a customer do before shifting to a competitor bank?,
- Which bank products are often availed of together by which groups of customers?,
- What patterns in credit transactions show increased risk of fraud?,
- What is the profile of a high-risk borrower?, and
- What services and benefits would current customers likely desire [5]?

Banks have realized that customer relations are a very important factor for their success. The challenge banks face is how to retain most profitable customers. Literature suggests that a small change in the retention rate can result in significant impact on business [6]. Huge amount of costumer and transaction data are maintaine by banks, but because of size of these databases makes it impossible for the banks to analyze and to retrieve useful information for the decision makers. Data mining is a powerful tool that can find patterns and relationships within a data. Using data mining technique, it is possible to build a successful predictive model which transforms data into meaningful information [7].

This paper proposes a neural network based approach to predict customer churn in bank. Real-world data from one of the small Croatian banks was used for creating a model for Customer churn. The main hypothesis was that clients who use more bank services (products) are more loyal, and bank should focus on those clients who use less than three products, and offer them products according to their needs.

## METHODOLOGY

### DATA MINING PROCESS

The Data Mining Process is an iterative process which does not stop when a particular solution is deployed. There are four main phases in every data mining project.

First, there is initial phase of Problem definition in which specific business problem is translated into data mining problem.

Second phase is Data gathering and preparation phase. In this phase we transform data into prespecified format and we perform data cleansing, which is the process of detecting and correcting, or removing corrupt, inaccurate or irrelevant records. Data preparation tasks are likely to be performed multiple times, and not in any prescribed order. This phase can take up to 80 % of all analysis time. Data quality is a major challenge in data mining [8].

Then, there is Model building and evaluation phase. In this phase, various modelling techniques are selected and applied and parameters are calibrated to optimal values.

Four phase is Knowledge deployment, use of data mining within a target environment. In this phase we organize and present the results of data mining to the user [9]. The discovered knowledge is visually presented. Visualization techniques are more effective in understanding the output for end users [10].

## DATA ANALYSIS

The used database consist information on 1866 clients on the date of analysis. We wanted to show that there is much smaller possibility for client that uses two or more bank products to leave the bank, in comparison to clients with just one product. Based on the information that we got from the bank, we determined each client's likelihood to leave the bank, whether it is low or high. We designed neural network using Alyuda NeuroInteligence software package and we got a model in which we can determine likelihood of client leaving the bank on the basis of some data. Characteristics that we used are: sex, age, private status, average monthly income, usage of internet banking and usage of two or more bank products.

Bank products are currency account, credit, savings, internet banking, mobile banking, SMS, standing orders, etc. We grouped similar products together, so we have only one category Savings and not special savings like Open, Active, Currency, Foreign Currency, etc. We did the same thing with Credits. We did this because the bank has many different products and few customers using these products. We divided Private status into: employed, pensioners, students and unemployed. Average monthly income we divided into these categories: 0 to 5 000,00 kn, from 5 000,00 kn to 10 000,00 kn, from 10 000,00 kn to 20 000,00 kn and more than 20 000,00 kn. By age, we divided them from 0 to 25, from 26 to 35, from 36 to 50, from 51 to 60, and more than 61. We used one client as the basic unit. We achieved the uniqueness of the client by choosing them by registration number.

As we mentioned earlier, data preparation is the most time consuming phase. Problems that we had with data are: missing values (financial laws are changing constantly, so some data that did not exist in the past has now become obligatory). Sometimes we could add these data on the basis of other data (sex, by name and surname), sometimes we could not do that out of several reasons: we needed to contact the client in order to get the answer (average monthly income, place of birth), sometimes large amount of data was missing or it was incomplete, sometimes there was a big number of possible input data, nonlinear dependences, inconsistency (different names for the same attribute), contain errors or exceptions.

## DESCRIPTION AND APPLICATION OF THE CHOSEN METHOD

Neural networks are considered alternative statistical methods. Today, there are tools that enable analysts to use neural networks without the knowledge of how they operate [11]. A neural network is a system of programs and data structures that imitates the operation of the human brain. It is nonlinear predictive model that learns through training and resembles biological neural networks in structure [12]. The basic building block of a neural network is the neuron. Each neuron consists of two parts: the net function and the activation function. The net function determines how the network inputs are combined inside neuron. There are three types of neurons: input, hidden and output. The output of the neuron is related to the network input via linear or non-linear transformations called activation function [13]. The result of output neuron is called prediction. The difference between a classical approach and neural networks is that in the classical approach first a mathematical model of the measured data is developed and then a system based on the measured model is developed. Neural networks operate directly with the data and do not need to know the model of the measured data. The process works by analysing past events and making current decisions based upon past experience, it learns from examples. Neural networks are typically organized in layers. Layers are made up of a number of interconnected nodes which contain an activation function. Patterns are presented to the network via the input layer, which communicates to one or more hidden layers where the actual processing is done via a system of weighted connections. The hidden layers then link to an output layer where the answer is output. Each input is sent to every neuron in the hidden layer and then each hidden layer's neuron's output is connected to every neuron in the next layer.

Neural networks work very well for problems like capturing associations or discovering regularities within a set of patterns, problems where the volume, number of variables or diversity of the data is very big, problems where the relationships between variables are vaguely understood or the relationships are difficult to describe adequately with conventional approaches [14]. Today, neural networks are applied in many areas of life due to its ability to capture complex patterns present in the data, such as medicine, banking, engineering, geology, physics, etc., usually for the following tasks: pattern recognition, image processing, speech processing, optimization problems, nonlinear control, processing inaccurate and incomplete data, simulation etc. Usually, it is combined with other methods because of the difficult interpretation of the results [15].

One of the disadvantages of neural networks is relatively slow and tedious process of learning model. Another disadvantage is that neural networks do not offer as final data model understandable relationship between important variables. The model is implicit, relationships between variables are hidden in the network structure [16].

## RESULTS

The paper used neural network method within the software package Alyuda NeuroInteligence to detect Customer churn. After selecting the database, software goes through all the above mentioned phases.

In data analysis phase, we are defining three types of characteristics: characteristics which we will reject (name, surname), characteristics we will use, and determine the target characteristics that we want to calculate. Alyuda NeuroInteligence divides Data into three sets: training, validation and testing set.

In preprocessing phase program adds some columns if data is marked as Categorical.

In the network design phase we are selecting a number of hidden layers. Program offers the best topology, which we can change. In our case, that is a neural network with three hidden layers with 8, 4 and 2 neurons.

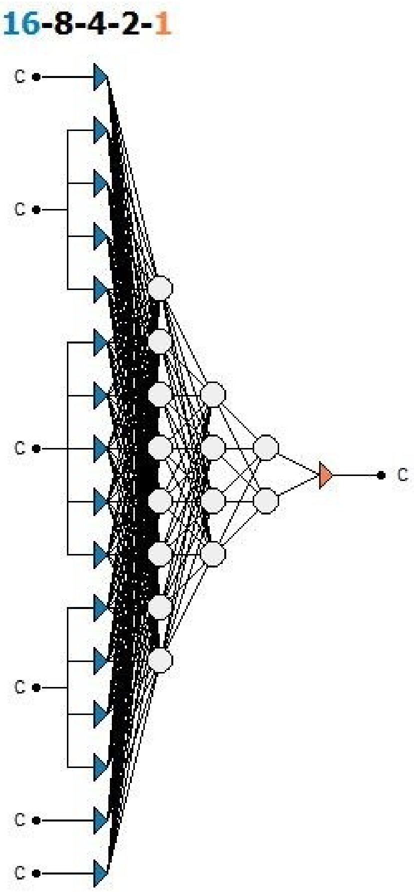


**Figure 1.** Alyuda – Network Topology.

After designing, there is training in which we can define different parameters as shown in Figure 2.

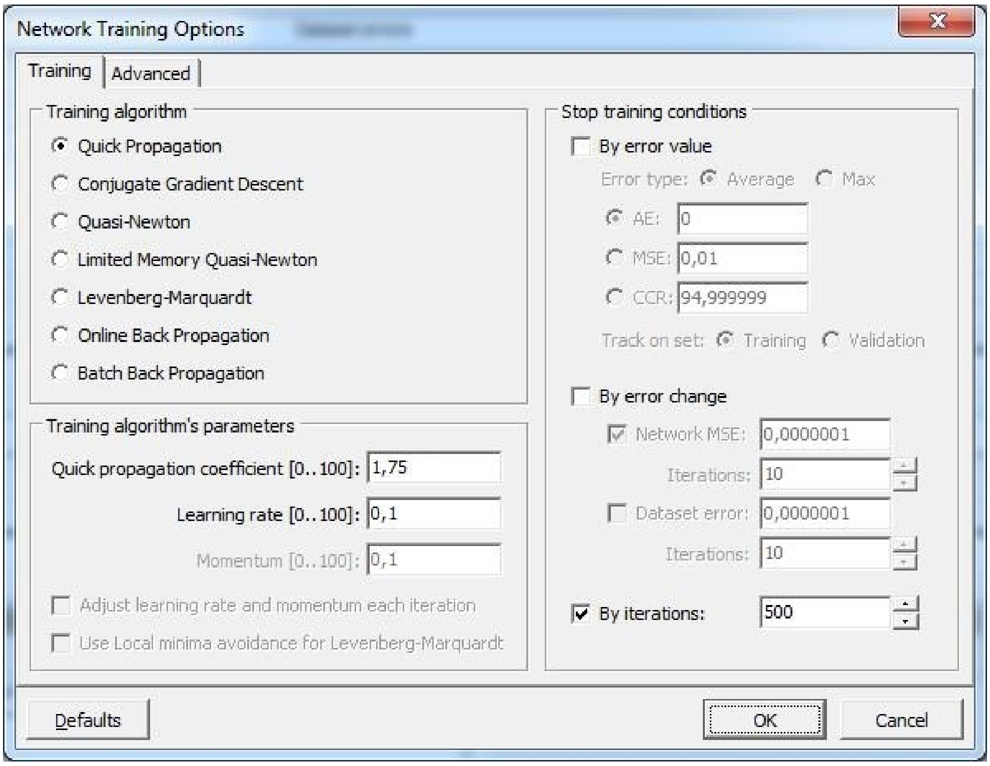


**Figure 2.** Alyuda – Network Training Options.

After training the network, we get results as shown in Figure 3.

| Parameters | | |
|---|---|---|
| | Training | Validation |
| CCR, %: | 95,984252 | 93,959732 |
| Network error: | 0,123164 | 0 |
| Error improvement: | 0,000009 | |
| Iteration: | 501 | |
| Training speed, iter/sec: | 62,625026 | |
| Architecture: | [16-8-4-2-1] | |
| Training algorithm: | Quick Propagation | |
| Training stop reason: | All iterations done | |

**Figure 3.** Alyuda – Network Training Results.

At the end we got model in which we can check the likelihood of client leaving the bank by entering some parameters.

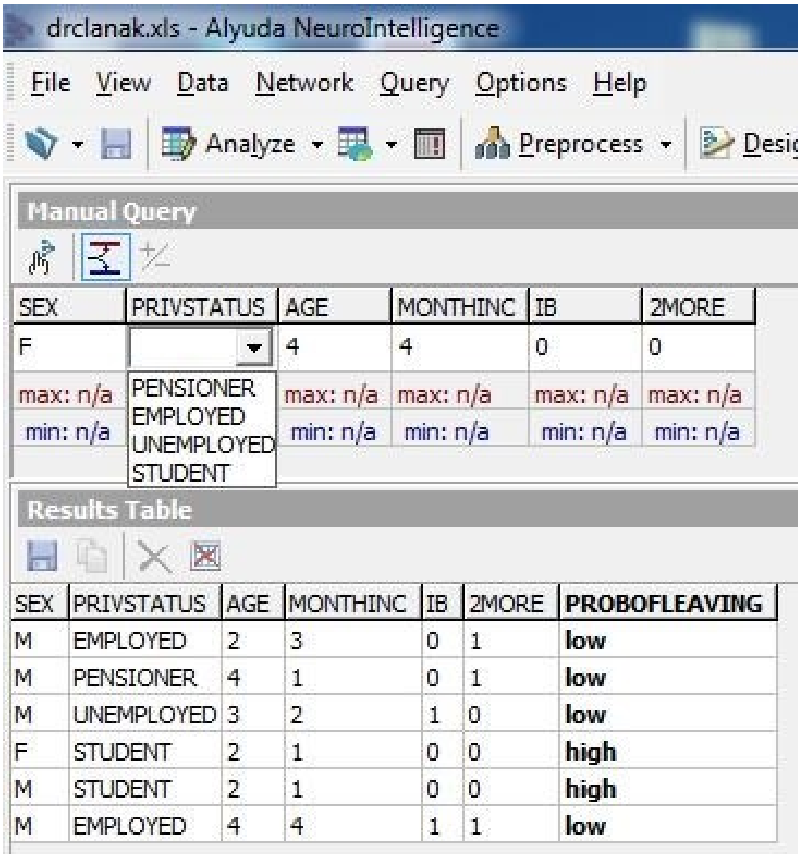


**Figure 4.** Alyuda – Query.

We can conclude that there is 'a problematic group' of young people (students) with less than three bank products, who in the future can become very important and very valuable clients. Bank should adjust its products to these clients. For example, Bank could introduce new products tailored to students' needs such as student loan, favourable interest rates, promotional use of internet banking, etc.

Also, we found out that by changing the topology of the neural network, we do not get better results. All topology we tried gave similar results.

## DISCUSSION

The Bank has very well-tailored services for pensioners, and this is the reason of high proportion of pensioners in the total number of clients (691/1866), and their likelihood of going to the competition is extremely low. The Bank offers them various benefits from lower fees for managing current accounts, loans for pensioners, free standing orders, etc. Bank should do something similar with the other groups of clients. For example, it should enable students as much as possible on line services.

We have shown that a simple analysis and application of neural networks can reach important results for the bank. It would be possible to include additional characteristics such as credit return, unauthorized overdraft, monthly consumption, the amount of savings, etc. to get different models for different problems.

## CONCLUSION

In order to be competitive in this market, banks have to be able to predict possible churners and take proactive actions to retain valuable loyal customers. Building an effective and accurate customer churn prediction model has become an important research problem for both academics and practitioners in recent years. Profiling enables a company to act in order to keep customers may leave (reducing churn or attrition), because it is usually far less expensive to keep a customer than to acquire a new one [17].

Neural network is a valuable forecast tool in financial economics due to the learning, generalization and nonlinear behaviour properties. It is powerful general-purpose software tool used for a number of data analysis tasks such as prediction, classification and clustering. Neural networks are used in finance such as portfolio management [18], credit rating [19] and predicting bankruptcy [20-22], forecasting exchange rates [23, 24], predicting stock values [25, 26], inflation [27] and cash forecasting [28] and others in order to achieve a reliable decision-making process through scientific approaches [29]. The ability of neural networks to discover nonlinear relationships in input data makes them ideal for modelling nonlinear dynamic systems such as banking industry.

The bank must operate on a long term customer strategy, young customers are recognized as being unprofitable in the early stage in lifecycle but will become profitable later on. In this paper we have shown that more and more young people use internet banking and that bank should offer different products/services which could be arranged without the client coming to the bank, such as savings that can be arranged and used only on the internet. It is necessary to develop new products that could be offered to such customers in order to keep them.

Cross-selling is one of the most important ways to increase the profitability of existing customers while increasing their loyalty. By selling additional products to customers we associate with them, thus increasing their loyalty (we have seen that more loyal customers are those who use more than two bank products). Analysing the data available we can determine what the next best offer for a particular client is. For example, bank could offer car insurance together with the car loan.

In this article a customer churn analysis on database of small Croatian bank was presented. The analysis focused on churn prediction based on only one method, Neural network. We could access other important information that could help banks to get competitive advantage by using other methods such as segmentation, decision trees, self-organizing maps. We wanted to show the simple usage of a complex method and to encourage others in similar research. Today, there are many very good software packages for data mining that do not require much pre-knowledge to use, and results can be very useful.

## REFERENCES


[1] Sumathi, S. and Sivanandam, S.N.: *Introduction to Data Mining Principles*. Studies in Computational Intelligence **29**(3), 1-20 2013,

[2] Ramageri B.M. and Desai, B.L.: *Role of data mining in retail sector.* International Journal on Computer Science and Engineering, **5**(1), 47-50, 2013,

[3] Hung, S.Y.; Yen, D.C. and Wang, H.Y.: *Applying data mining to telecomm churn management.* Expert Systems with Applications **31**(3), 515-524, 2006, http://dx.doi.org/10.1016/j.eswa.2005.09.080,

[4] Baesens, B.; Viaene, S.; Van den Poel, D.; Vanthienen, J. and Dedene, G.: *Bayesian neural network learning for repeat purchase modeling in direct marketing*. European Journal of Operational Research **138**(1), 191-211, 2002,

[5] Domingo, R.: *Applying data mining to banking.* http://www.rtdonline.com, accessed 18th November 2015,

[6] Van den Poel, D. and Larivie‘re, B.: *Customer attrition analysis for financial services using proportional hazard models*. European Journal of Operational Research **157**(1), 196-217, 2004, http://dx.doi.org/10.1016/S0377-2217(03)00069-9,

[7] Jackson, J.: *Data mining: a conceptual overview.* Communications of the Association for Information Systems **8**(19), 267-296, 2002,

[8] Blake, R. and Mangiameli, P.: *The effects and interactions of data quality and problem complexity on classification*. Journal of Data and Information Quality **2**(2), 160-175, 2011, http://dx.doi.org/10.1145/1891879.1891881,

[9] Oracle: *Data Mining Concepts*. http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm#DMCON002, accessed 18th December 2015,

[10] Herawan, T. and Deris, M.M.: *A soft set approach for association rules mining*. Knowledge-Based Systems **24**(1), 186-195, 2011,

[11] Batini, C.; Cappiello, C.; Francalanci, C. and Maurino, A.: *Methodologies for data quality assessment and improvement*. ACM Computer Surveys **41**(3), 16-27, 2009, http://dx.doi.org/10.1145/1541880.1541883,

[12] Han, J.; Kamber, M. and Pie, J.: *Data Mining Concepts and Techniques*. Elsevier, Burlington, 2011,

[13] Schölkopf, B. and Smola, A.J.: *Learning with kernels: Support vector machines, regularization, optimization, and beyond.* MIT press, 2002,

[14] Bar, M.V.: *The Computational Intelligence Techniques For Predictions-Artificial Neural Networks.* Annals of Computational Economics **2**(42), 184-190, 2014

[15] Cherkassky, V.; Friedman, J.H. and Wechsler, H.: *From statistics to neural networks: theory and pattern recognition applications.* Springer Science & Business Media, 2012,

[16] Frenkel, D. and Smit, B.*: Understanding molecular simulation: from algorithms to applications.* Academic Press **1**(1), 2001,

[17] Berry, M.J. and Linoff, G.S.: *Mastering Data Mining: The Art and Science of Customer Relationship Management*. Wiley Computer Publishing, New York, 2000,

[18] Ashwood, A.J.*: Portfolio selection using artificial intelligence*. http://eprints.qut.edu.au/66229/1/Andrew_Ashwood_Thesis.pdf, accessed 19th February 2016,

[19] Nazari, M. and Alidadi, M.: *Measuring credit risk of bank customers using artificial neural network.*
Journal of Management Research **5**(2), 17-27, 2013,
http://dx.doi.org/10.5296/jmr.v5i2.2899,
[20] Wilson, R.L. and Sharda, R.: *Bankruptcy prediction using neural networks*.
Decision Support Systems **11**(5), 545-557, 1994,
[21] Tsai, C.F. and Wu, J.W.: *Using neural network ensembles for bankruptcy prediction and credit scoring*.
Expert Systems with Applications **34**(4), 2639-2649, 2008,
[22] Bredart, X.: *Bankruptcy Prediction Model Using Neural Networks*.
Accounting and Finance Research **3**(2), 124-126, 2014,
http://dx.doi.org/10.5430/afr.v3n2p124,
[23] Panda, C. and Narasimhan, V.: *Forecasting exchange rate better with artificial neural network*.
Journal of Policy Modeling **29**(2), 227-236, 2007,
[24] Huang, W.; Lai, K.K.; Nakamori, Y. and Wang, S.: *Forecasting foreign exchange rates with artificial neural networks: a review.*
International Journal of Information Technology and Decision Making **3**(1), 145-165, 2004,
http://dx.doi.org/10.1142/S0219622004000969,
[25] Naeini, M.P.; Taremian, H. and Hashemi, H.B.: *Stock market value prediction using neural networks*.
Computer Information Systems and Industrial Management Applications, 132-136, 2010,
http://dx.doi.org/10.1109/CISIM.2010.5643675,
[26] Devadoss, A.V. and Ligori, T.A.A.: *Stock Prediction Using Artificial Neural Networks*.
International Journal of Data Mining Techniques and Applications **2**, 283-291, 2013,
[27] Haider, A. and Hanif, M.N.: *Inflation forecasting in Pakistan using artificial neural networks*.
Pakistan Economic and Social Review, 123-138, 2009,
[28] Simutis, R.; Dilijonas, D. and Bastina, L.: *Cash demand forecasting for ATM using neural networks and support vector regression algorithms.*
20th EURO Mini Conference Continuous Optimization and Knowledge-Based Technologies, Neringa, 20-23 May, 2008,
[29] Li, Y. and Ma, W.: *Applications of artificial neural networks in financial economics: a survey.*
Computational Intelligence and Design **1**(1), 211-214, 2010.