FUZZY NEURAL NETWORKS FOR EVALUATING THE CREDITWORTHINESS OF THE BORROWERS

Natalia Shovgun

Abstract: *The problem of assessing the creditworthiness of the borrower is considered. The application of fuzzy neural networks for this problem solution, fuzzy neural networks TSK and Mamdani was suggested. The experimental investigations of application of these networks for our task were carried out and comparison with classical methods was performed. The modification of adaptation and learning algorithms of fuzzy neural networks was suggested.*

Keywords: *fuzzy neural networks, credit rating, fuzzy logic*

ACM Classification Keywords: *H.*4 Information systems applications; H.4.2. Types of Systems Decision Support

Introduction

The main activity of commercial bank is a credit activity. Lending provides almost half of bank profits; however, it is inextricably linked with risk. Credit risk is connected to possible misconduct by the borrower and it is one of the most significant risks of commercial banks. Consumer loans to individuals is a basic banking products. The bank assesses the creditworthiness of a potential borrower before lending. This is a method to minimize the losses of the bank. Careful selection of borrowers and effective assessment of creditworthiness is the main way of assessing and reducing credit risk. Information for decision making about lending may be inaccurate, incomplete, and information about the borrower may be such that it is difficult formalized. The analysis of existing methods of credit analysis showed the feasibility of using the methods based on fuzzy logic. These methods can work with both quantitative and qualitative characteristics and decision-making process is based on a comprehensible rules base.

For example, the technique of assessing the creditworthiness of individuals using the method of paired comparisons and fuzzy systems with Mamdani-type logical conclusion was considered in [Kuznetsov, 2007]. But the downside of fuzzy inference is that they can not learn automatically. Parameters and type of membership functions, which describe factors of creditworthiness, is given by an expert that's why it may be inadequate. Fuzzy neural networks (FNN) combine the advantages of fuzzy inference systems and neural networks - the ability to adapt and automatic learning, and the ability to interpret the process results.

To analyze the creditworthiness of borrowers applied fuzzy neural network with output Mamdani [Zaychenko, 2008] and fuzzy neural network with output Sugeno.

The work is devoted to the study of the FNN in the problem of assessing the creditworthiness of the borrower. A comparison of the results for the credit assessment by FNN TSK with classical methods such as logic model was performed and with popular in recent years Bayesian networks. Also provided are methods for setting up the rule base on which to base make a decision on lending.

Methods for assessing the creditworthiness of the borrower and their features

Based on questionnaires borrower the bank needs to decide - whether to grant credit. Each such form can be represented as a vector $\{X_1, X_2, ..., X_i, ..., X_M\}$, where X_i – some way the formalized data of borrower and the parameters of loan. This vector is the input of network. The decision on lending to the borrower is the output of network.

Most commonly used to solve this problem is using linear or logistic regression [Duffie, 2003]. Using linear regression we have function that determines the credit rating by approximating of a linear function wich argument is the vector characteristics of the borrower, i.e.:

$$p = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \ldots a_N \cdot x_N,$$

where a_0 – the free term; a_i, *i = 1,…, N* – weights of borrower characteristics; x_i - characteristics of the borrower.

All regression methods are sensitive to the correlation between the characteristics, so in the model should not be strongly correlated independent variables. In addition, the regression coefficients are not giving enough information about mechanism of influence characteristics of the borrower on the risk.

Bayesian networks (BN) are used in situations with some uncertainty. BN is a triple $N = < V, G, J >$, where V is a set of variables, G is a directed acyclic graph whose nodes correspond to random variables modeled process, J is a joint probability distribution of variables $V = \{X_1, X_2, ..., X_i, ..., X_N\}$. Bayesian networks, which may be presented with discrete and continuous variables is called hybrid Bayesian networks. Details about the BN can find at [Dawid, 2007].

In FNN results is obtained by using fuzzy logic, but the corresponding membership function is customized using learning algorithms FNN. Thus, the network uses a priori information for find new knowledge and it is logically transparent to the user.

We consider two different FNN. FNN with Mamdani-type fuzzy rules use next base of rules:

R_i: if x_1 is A_{i1} and x_2 is A_{i2} and … and x_n is A_{in} then y_i is C_i,

where x_i and y_i are input and output variables of the network, A_i and C_i are input and output fuzzy sets.

FNN with Sugeno-type fuzzy rules (FNN TSK) use next base of rules:

$$R_k : \text{if } x_1 \text{ is } A_1^{(k)};\ x_2 \text{ is } A_2^{(k)};\ \ldots;\ x_n \text{ is } A_N^{(k)}, \text{ then } y_i = p_{i0} + \sum_{j=1}^{N} p_{ij} x_j \ ,$$

where $A_i^{(k)}$ is fuzzy sets of variable x_i, i=1,2…N (data of borrower) for rule R_k with membership function

$$\mu_A^{(\kappa)}(x_i) = \frac{1}{1 + \left(\frac{x_i - c_i^{(k)}}{\sigma_i^{(k)}}\right)^{2b_i^{(k)}}}.$$

For network training is used back-propagation algorithm. To find the parameters of membership function is used a gradient or genetic method. In the gradient method for configuring the parameters of membership function can be used resilient propagation method [Riedmille, 1992] to reduce the learning process. Each of the considered algorithms has its drawbacks. Thus, the gradient algorithm is highly dependent on the initial conditions, and genetic frequently converge to local optima. The author proposed to use a hybrid algorithm in which the initial approximation (initial values of membership functions) is found by using genetic algorithm, and only then it is considered as the starting point for the gradient algorithm.

Scaling up the rules is the most efficient algorithm among the algorithms adapting FNN. The algorithm proposed in [Kruglov, 2002], based on an assessment of the accuracy of approximation. In this algorithm, we add a rule, if the existing knowledge base gives too large an error for the current point.

In [Juang, 1999] proposed an adaptation algorithm based on firing strength of rules, which is faster than algorithm based on approximation accuracy. In this algorithm, a new rule is added if the condition: $I = \arg\max_{1 \le k \le r} F^k(x) \le F_{in}$, where $F^k(x(t)) = w_k$ is firing strength of rule k for input vector t, and F_{in} is a pre-specified threshold that decays during the learning process. In new rule the parameters of membership functions are set as follows: $c_j^{(k)} = x_j(t)$, $b_j^{(k)} = 1$, $\sigma_j^{(k)} = \beta \cdot \prod_{j=1}^{N} \frac{1}{1 + (\frac{x_j - c_j^{(I)}}{\sigma_j^{(I)}})^{2b_j^{(I)}}}$. That is, we add a rule, if none of the existing rules do not describe well enough the current input vector. For the efficiency is proposed to make additional checks: if $f_o = \min(f_1, f_2, ..., f_m) > R$, where $f_k = \left\| \vec{c}^{(k)} - \vec{x}^i \right\|$, $k = 1, m$ is the distance between the centers of each membership function for each rules and the current point, $R = const$ then new rule is generated. Thus increasing the control of the number of rules. Accordingly the optimal network structure is building and thus the training requires less time.

Experimental results

The data sample of one of Ukrainian banks, which consists of 1,000 samples is used for credit analysis using the proposed methods. The feasibility of using FNN can be seen from Figure 1.

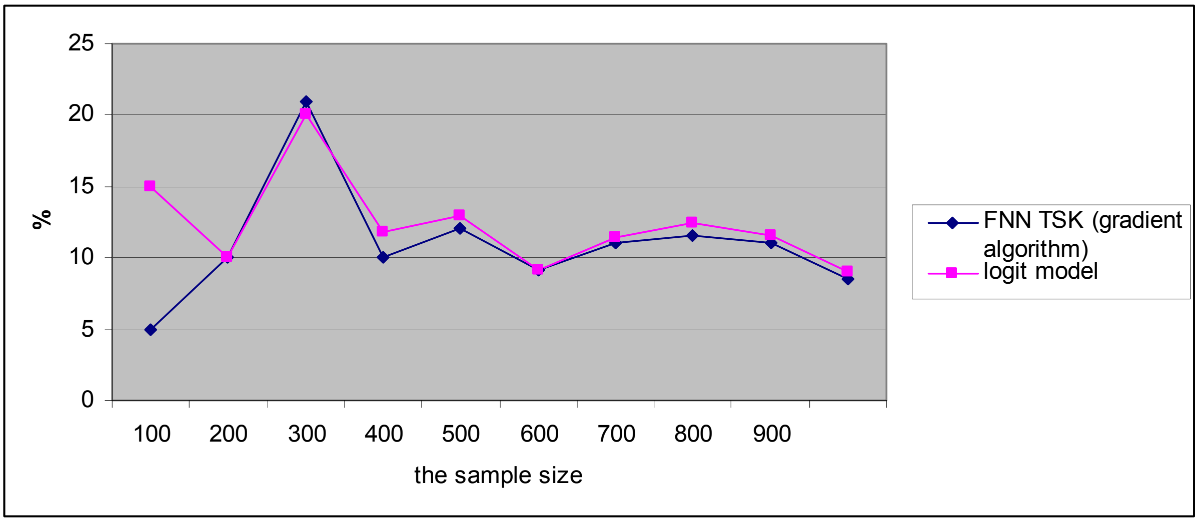

Figure 1. The errors dependence on volume of a sample

As you can see from the figure the percentage of incorrect classifications decreases with increasing training set. Errors are smaller when was using FNN TSK. To prevent re-training FNN TSK should be edit the complexity of network structure by adapting the algorithm parameters in accordance with the size of the training set. Building such an algorithm may be the subject of further research.

The data set consists of values: the borrower's age, sex, marital status, number of dependent, income, work experience, realty, monthly payment and reply. After the correlation analysis with using programs Netica (http://www.norsys.com/netica.html) was constructed the Bayesian network - Figure 1.

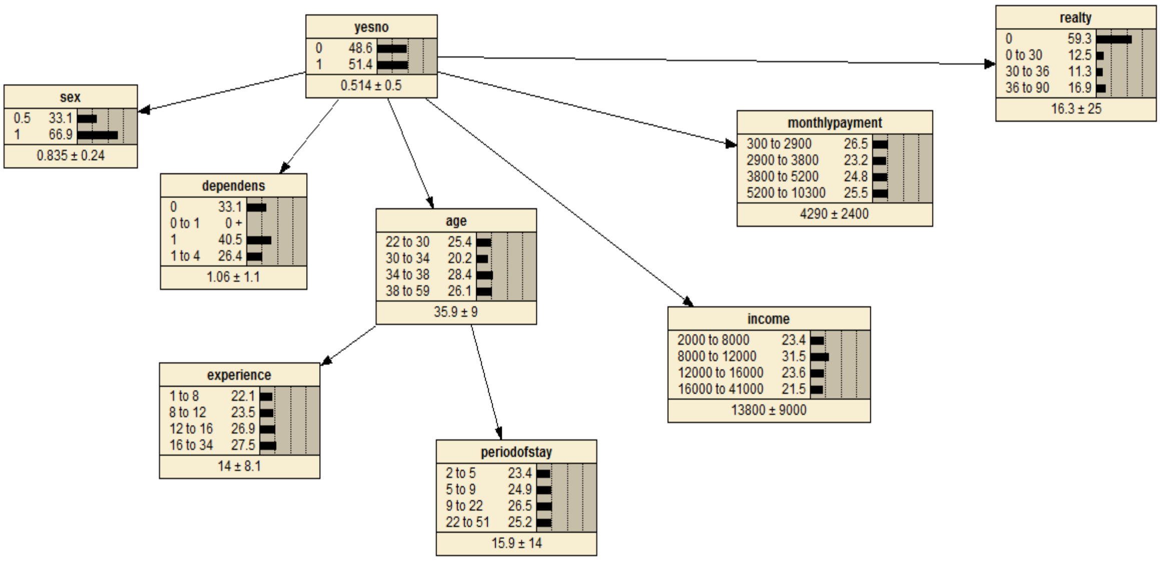

Figure 2. The structure of Bayesian network

The experemental results on figure 3.

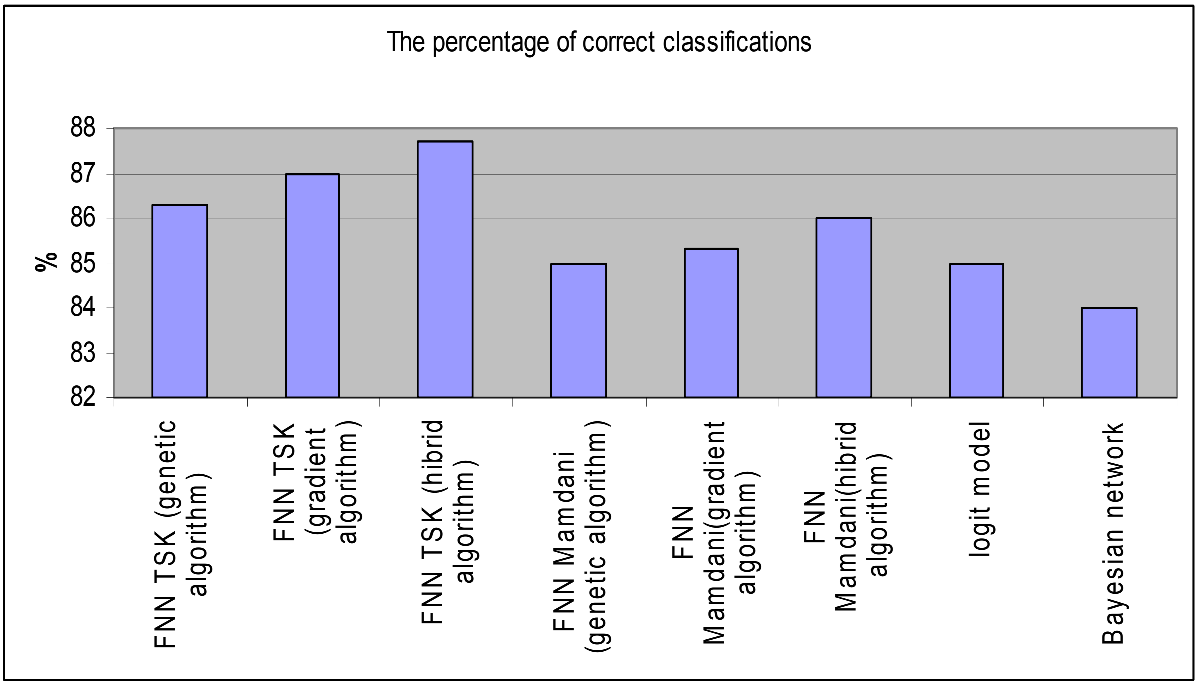

Figure 3. The results of the credit assessment of different methods

The best results has the FNN TSK (hibrid algorithm). Since the percentage of true classification isn't dependent on the type of membership function (Figure 4) we can speak of an automated construction of rules base on which a decision-making on lending is done. For example, if we trained the rule base with only two rule:

Rule 1: If the "Age" is "Low", "Sex" is "Female", "Married" is "No", "Number of dependents" is "High", "Income" is "Low", "Experience" is "High", "Residence time" is " High "," Monthly Payment " is "Low ", "Answer" is "No".

Rule 2: If the "Age" is "High", "Sex" is "Male", "Married" is "Yes", "Number of dependents" is "Low", "Income " is "High", "Experience" is "Low" , "Residence time" is "Low", "Monthly Payment" is "High", "Answe" is "Yes".

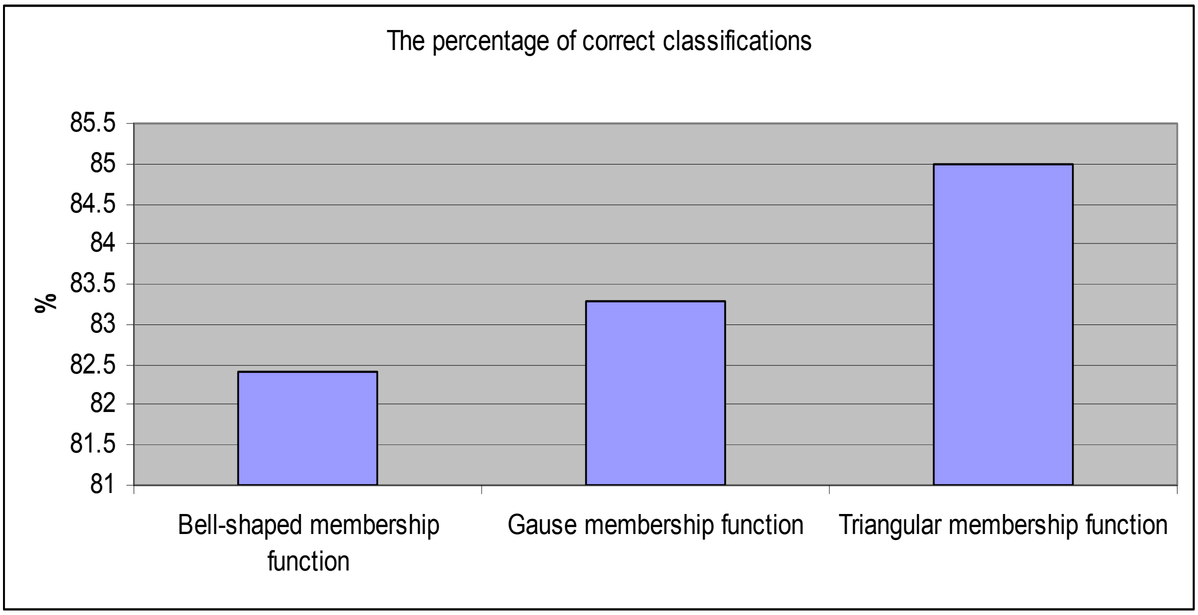

Figure 4. The results of the credit assessment for different membership function

We have a clear interpretation of the process of obtaining the decision making, credit institutions are given the opportunity to evaluate and adjust credit terms to offer the borrower an alternative parameters of lending.

Conclusion

The article considers the practical application of fuzzy neural networks to the problem of assessing the creditworthiness of the borrower. The results are compared with the classical method such as a logit model and some new as bayesian network. The best percentage of true classifications showed FNN TSK with hybrid (combination of gradient and genetic methods) learning algorithm. A new adaptation algorithm for fuzzy neural network was proposed, so we can build base rules automaticaly. As a result, there is an optimal network structure construction in accordance with the training set.

Acknowledgement

"The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua)."

Bibliography

[Dawid, 2007] Philip Dawid, Steffen L. Lauritzen, David J. Spiegelhalter Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks; Springer Science, 2007, 325 p.

[Duffie, 2003] Darrell Duffie, Kenneth J. Singleton Credit Risk: Pricing, Measurement, and Management, Princeton University Press, 2003, 396 p.

[HÁJEK,2008] Petr HÁJEK, Vladimír OLEJ MUNICIPAL CREDITWORTHINESS MODELLING BY NEURAL NETWORKS // Acta Electrotechnica et Informatica Vol. 8, No. 4, 2008, pp. 10 - 15

[Juang, 1999] CF Juang and CT Lin A recurrent self-organizing neural fuzzy inference network// IEEE transaction of neural networks, Volume 10, Number 4, pp. 828–845, July 1999

[Kruglov, 2002] Kruglov VV, Borisov VV, Artificial Neural Networks; Theory and Practice 2nd ed., Stereotype M: Hotline Telecom, 2002. 382 p.

[Kuznetsov, 2007] Kuznetsov L., Evaluation of credit history of individuals on the basis of fuzzy models // Managing large systems, 2007, № 21, pp. 84-106

[Riedmille, 1992] Riedmiller M., Brawn H. RPROP - a fast adaptive learning algorithms, Technacal Report // Karlsruhe: University Karlsruhe, 1992.

[Zaychenko Yu, 2008] Zaychenko Yu.P., Credit risk estimation on the base of fuzzy logic // System Research and Information Technologies, 2010, № 2, pp. 37–54

Author’s Information

***Shovgun Natalia**, PhD student, The National Technical University of Ukraine “Kyiv Polytechnic Institute”, Kyiv, email:* shovgun@gmail.com

Major Fields of Scientific Research: Neural network and fuzzy set