Nonlinear Analysis: Modelling and Control, 2008, Vol. 13, No. 1, 61–70

Interpretable Nonlinear Model for Enterprise Bankruptcy Prediction

O. Purvinis, R. Virbickaitė, P. Šukys

Kaunas University of Technology
Klaipėdos str. 3, Panevėžys, Lithuania
opurvi@panf.ktu.lt

Received: 26.09.2007 **Revised:** 01.02.2008 **Published online:** 06.03.2008

Abstract. The aim of this research was to model bankruptcy dependency of Lithuanian enterprises on their financial ratios and its dynamics over time by the integration of artificial neural networks and fuzzy logic technology using Adaptive Network – based Fuzzy Inference System (ANFIS). We used data from financial reports for three years' of 230 Lithuanian going and failed enterprises. Input variables used for the ANFIS model training and testing composed of 13 financial ratios of the last year before bankruptcy and 13 variables characterizing changes of that ratios over time. It was checked 1716 subsets of input variables, each subset containing five input variables. This way the ANFIS model and the best subset of predictive variables with minimal training errors was found. Test of that model showed that percentage of right failure and success predictions reached 80 %. Fuzzy rules of the ANFIS were used to construct interpretable rules base, which can be useful for enterprise managers as knowledge for the linguistic diagnosis of failure or financial problems.

Keywords: bankruptcy, prediction, ANFIS, knowledge.

1 Introduction

The financial state of an enterprise is estimated according to its various factors. Usually financial ratios X_i are calculated, when some factors, describing the size and performance of an enterprise, are divided by other factors. In this way ratios, not depending on the size of an enterprise, are obtained [1] and it lets compare the enterprises of different size.

The search of complex index or its system to estimate the possibility of enterprise bankruptcy was started in the 20th century and it hasn't been finished yet [2]. From methodological side W. Beaver's research in 1966 was particularly valuable as pioneering work, and based on the comparative analysis of indicators of profitable and insolvent enterprises [3]. Later E. Altman, A. Kovaliov, T.Poddig and others researchers used multivariate approaches and methods of discriminant functions [4–6].

J. Mackevičius considered the characteristics of Lithuanian financial accountability and pointed out main groups of indicators: profitability (there are about 12 indicators of it), short-term and long-term solvency (there are about 18 of it), efficiency of activity (about 36) and capital market (about 15) [1].

To identify the possibility of crisis or bankruptcy most often linear score function is used

$$Z = k_1X_1 + k_2X_2 + \ldots + k_mX_m, \tag{1}$$

where $X_1, X_2, \ldots, X_m$ are financial ratios.

Z value describes the possibility of a crisis or bankruptcy in an enterprise – a small possibility of a bankruptcy, a big possibility of a bankruptcy, and very big possibility of a bankruptcy.

Every method has different financial ratios X_i and different coefficients k_i. The most famous for bankruptcy prediction is Altman's function including five financial ratios [4]

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5,$$

where X_1 – working capital / total assets, what measures liquid assets in relation to an enterprise size; X_2 – retained earnings / total assets, what is a measure of cumulative profitability that reflects an enterprise age as well as an earning power; X_3 – earnings before income taxes / total assets, what is a measure of operating efficiency separated from any leverage effects; X_4 – market value of equity / book value of debt, this ratio expresses a market dimension; X_5 – sales / total assets, this is a standard turnover measure.

Altman's method was developed 39 years ago using the sample of 66 companies, 33 failed and 33 going. During this time things have changed and the economic environment differs from country to country [7]. Recent research by V. Boguslauskas and A. Stundžienė revealed that the Altman's method, when applied to 56 Lithuanian joint stock companies, produced meaning errors [8].

On the other hand, the score function Z (1) is limited to linear form, while dependencies between sophisticated economic factors and phenomena may be nonlinear. Vlachos and Tollias using the same financial ratios as were by Altman employed proposed an Adaptive Network-based Fuzzy Inference System that outperformed classical and modern methods of bankruptcy prediction [9].

The alternative methods for modelling business failure – such as multi-logit analysis, survival analysis, dynamic event history analysis, multidimensional scaling, decision trees, expert systems and neural networks have been applied. However, as literature does not provide a clear overview of the application of alternative methods to the topic of business failure prediction, further research concerning these methods is necessary [10].

Neural networks were applied by the authors of this paper to predict Lithuanian enterprises failure and were reported in another paper [11]. They appeared to have good performance. But neural networks being good for modeling nonlinear systems have a disadvantage as well. They are lack of interpretability. Neural networks cannot explain how the specific value or forecast is obtained.

Identification of specific factors or their combination that leads to unfavorable forecast is very important to managers. That is, managers can make more use of conventional or fuzzy interpretable IF THEN rules rather than neural networks.

This research is an attempt to model bankruptcy dependency of Lithuanian enterprises on their financial ratios and its dynamics over time by the integration of artificial neural networks and fuzzy logic technology using Adaptive Network based Fuzzy Inference System (ANFIS) and to obtain interpretable rules.

The novelty of this research is the obtaining of linguistic rules from financial ratios of Lithuanian enterprises by ANFIS taking into account dynamics over time.

2 Sample

We had financial data of 30 going Lithuanian enterprises and 200 failed ones. The different number of going and failed enterprises is explained by the reason of meeting the difficulties trying to get the financial data of enterprises. The data were taken from the financial statements for three years' time. Some data were for 2002–2004 period, some for 2003–2005. The data of failed enterprises were used for the last three years' time until the bankruptcy. The data included: working capital, long-term assets, total assets, capital and reserves, retaining earnings, long-term liabilities, current liabilities, sales and service, gross profit.

Having used this data thirteen rates characterizing enterprise financial state were calculated. Twelve of them (X_1 = total asset turnover, X_2 = long-term asset turnover, X_3 = current asset turnover, X_4 = ratio of long-term liabilities (long-term liabilities / total asset), X_5 = ratio of long-term liabilities (long-term liabilities / long-term asset), X_6 = current ratio, X_7 = ratio of current liabilities, X_8 = ratio of total liabilities, X_9 = financial lever1, X_{10} = financial lever2, X_{11} = ratio of Golden Balance Rule, X_{12} = ratio of financial dependency) are well-known in economics financial ratios [1], but the thirteenth X_{13} used in this work assets ratio characterizing the ratio of long-term assets and total assets (LTA/TA) is not found in the economical literature:

$$X_{13} = \frac{long - term\, assets}{total\, assets}.$$

One of the research goals was to estimate the influence of factor dynamics over time on bankruptcy prediction, so the correlation coefficient r of the every ratio values during three years' with the years' numbers $j = 1, 2, 3$ was calculated. In the paper these correlation coefficients are termed as dynamics indicators $X_{14}, \ldots, X_{26}$.

A problem arose when the any ratio X_i of both kinds of enterprises was stable during three years' time period, i.e. $x_{i1} = c, \;\; x_{i2} = c, \;\; x_{i3} = c$. In this case it was impossible to calculate correlation coefficient applying a well known formula

$$r = \frac{mean(x_i j) - mean(x_i)mean(j)}{stdev(x_i)stdev(j)},$$

because $stdev(x_i) = stdev(const) = 0$.

But it is possible to obtain using this formula that correlation coefficient r between the first and the second terms of the sample $(c, 1),\ (c+\delta, 2),\ (c+\delta, 3)$ has these properties:

$$r = \frac{\sqrt{3}}{2}\,\mathrm{sign}\,\delta,$$

$$\lim_{\delta \to +0} r = \frac{\sqrt{3}}{2} \approx 0.87.$$

It lets in the mentioned case the correlation coefficient assume equal to 0.87. This value can be explained as turning point that separates stable factors from those increasing over time.

In this way 26 dimension vector characterized every enterprise. Moreover the variable – bankruptcy score o, was assigned two codes: when an enterprise is in bankruptcy – code is $o = 1$, when an enterprise is successfully going – code is $o = 0$.

To equalize the number of failed and going enterprises, the data of going enterprises was repeated for 6 to 7 times. 263 records obtained in such way were selected and these vectors were used for ANFIS model training, and the rest data, i.e. 137 records, was used for testing. It appeared, as it is discussed further in this article as is shown in the Table 2, that this approach enabled to avoid bias of the trained model towards all enterprises to be recognized as failed.

3 ANFIS model

The term neuro-fuzzy means hybrids of artificial neural networks and fuzzy logic. A hybrid intelligent system includes these two techniques by combining the human-like reasoning style of fuzzy systems with neural network ability to be trained by examples [12].

The Adaptive Network-based Fuzzy Inference System (ANFIS) is a set of fuzzy IF THEN rules

$$\text{IF } (X_1 \text{ is } mf_{1k}) \text{ and } (X_2 \text{ is } mf_{2j}) \text{ and } \dots \text{ THEN } O \text{ is } mfo \tag{2}$$

that models in general nonlinear dependencies between input variables X_i and output variable O. Parameters of the fuzzy system are tuned by neural network. This process is called training.

The mathematical software package Matlab was used to develop the ANFIS model. Calculations were done by personal computer with 1.4 GHz speed processor. Five factors $X_i,\ i = 1, 2, \dots, 26$, – financial ratios and their dynamics indicators were used in the model. This amount of the factors was chosen because the purpose of the paper was to obtain the rules and to interpret their factors as linguistic variables. But too big amount of variables become too complicate and too difficult for interpretation of the rules. Besides, to find more factors requires too much computing time. The method of selection five particular factors from all 26 is further discussed.

The interval of values of each factor X_i was divided into 3 fuzzy subsets by 3 Gaussian type membership functions (Fig. 1)

$$mf(x) = e^{-\frac{(x-c)^2}{2\sigma^2}}.$$

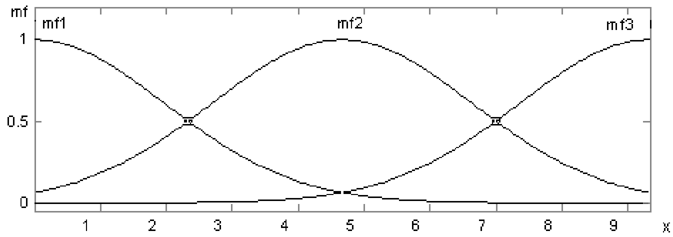

Fig. 1. The interval of the variable x values is divided into 3 fuzzy subsets by Gaussian membership functions mfi.

When increasing the number of membership functions, then the training time was essentially increasing. It was also happening when more than five fuzzy factors X_i, were used for ANFIS training. Besides, for human reasoning it is complicated to deal with more than five factors or linguistic variables.

For the sake of interpretability of the rules (2) the zero order Sugeno type fuzzy inference system was chosen. The main feature of the zero order Sugeno type fuzzy inference system is that the output membership functions are constant, i.e. parameters mfo of the rules (2) are constant but, in general case, may be different for each rule [13].

The training error e was defined as the square root of mean squared differences between the ANFIS outputs mfo_i and the assigned bankruptcy code o_i

$$e = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(mfo_i - o_i)^2},$$

where n is the number of enterprises selected for training.

It appeared that for most factor X_i subsets 10 epochs of training were sufficient to reach a small enough error value in comparison with those reached during 30 and 50 epochs and the training time within 10 epochs remained not very large. In Fig. 2 it is seen that from 10th to 20th epochs the training errors decreased just from 0.25 to 0.23, i.e. by 0.02.

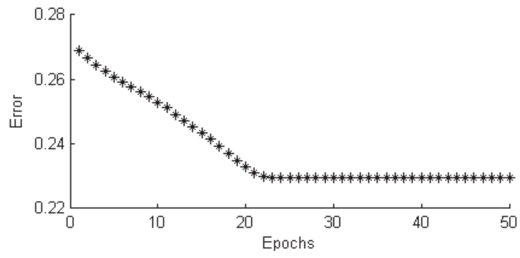

Fig. 2. Typical training errors.

It was necessary to choose five factors from mentioned 26 to minimize the training error e.

If you want to make a full-scale search of these five factors, you have to test $C_{26}^{5} = 65780$ ANFIS models. To decrease the number of model search, some factors and dynamics indicators were joined into two pairs X_i, X_{i+13} and X_j, X_{j+13} with added any fifth factor. In this way the number of ANFIS models was decreased to 1716. The training of all those models with 10 training epochs lasted about 40 hours. The quality of each model was estimated according to the error e and according the number of wrong relative forecasts γ_b of failed and wrong relative forecasts γ_s of going enterprises which were used for training

$$\gamma_b = \frac{n_{be}}{n_b}, \quad \gamma_s = \frac{n_{se}}{n_s}, \tag{3}$$

where: n_b – the number of failed enterprises used for model training; n_{be} – the number of failed enterprises in the training sample, which were recognized as going by ANFIS model; n_s – the number of going enterprises used for training; n_{se} – the number of going enterprises from the training sample, which were recognized as failed.

4 Results

In all 1716 ANFIS models only six models had forecast errors γ_b and γ_s in training sample less than 0.2 (they are shown in Table 1).

As shown in Table 1, the different sets of factors enable to make ANFIS models with small and similar training errors.

For next analysis the last model with the least training error shown in Table 1 with the following factors was chosen: ratio of Golden Balance Rule, dynamics indicator of Golden Balance Rule, ratio of assets X_{13} and dynamics indicator of ratio of assets, current assets turnover. Further this ANFIS model was repeatedly trained with 50 epochs (Fig. 2) and tested with test data subset. The results showed that more than 80 per cent of predictions for failed and going enterprises were right (see Table 2).

The results also show that repeating the data of going enterprises for so many times to make the number of them equal to the amount of failed enterprises answered the purpose. The numbers of error and correct forecasts for going and failed enterprises are almost equal. Generally, when the amounts of both types of objects in the training sample are different, the threat appears to the model to be biased and as a result major part of the test sample objects will be recognized as belonging to the type which dominated in the training sample. As it is seen in Table 2, it appeared that the model had no bias.

The forecasts of ANFIS was compared with the results of Altman Z-score function (see Table 3).

Therefore, the incorrect forecasts percentage for failed enterprises according to the Altman's model appeared much often than according to ANFIS model. Furthermore Altman's model gives indeterminate forecasts, so called ignorance zone. These forecasts are not included in Table 3.

The software package Matlab enables to represent the ANFIS forecast in diagram form. The plot of ANFIS forecast dependency on two factors when other three are fixed is given in the Fig. 3 and there is seen that the relationship between bankruptcy forecast and financial ratios is complicated nonlinear.

In Fig. 3 it is seen that when the ratio of Golden Balance Rule is greater than 5, bankruptcy possibility depends only on this ratio and does not depend on the ratio of assets X_{13} = LTA/TA. Whereas when the ratio of Golden Balance Rule is less than 5, bankruptcy score nonlinearly depends on the ratio of long term assets LTA/TA. Similarly it is possible to set other factors in the horizontal axis and analyze their influence on bankruptcy forecast.

Table 1. ANFIS models that had forecast errors γ_b and γ_s in the training sample less than 0.2

Fuzzy factors					Training error e
Current assets turnover	Dynamics indicator of current assets turnover	Current ratio	Dynamics indicator of current ratio	Ratio of Golden Balance Rule	0.27
Rate of total liabilities	Dynamics indicator of total liabilities	Current ratio	Dynamics indicator of current ratio	Ratio of long-term liabilities	0.32
Rate of total liabilities	Dynamics indicator of total liabilities	Current ratio	Dynamics indicator of current ratio	Long-term liabilities / long-term assets	0.32
Rate of total liabilities	Dynamics indicator of total liabilities	Current ratio	Dynamics indicator of current ratio	Ratio of current liabilities	0.33
Rate of total liabilities	Dynamics indicator of total liabilities	Ratio of current liabilities	Dynamics indicator of current liabilities	Current ratio	0.32
Ratio of Golden Balance Rule	Dynamics indicator of Golden Balance Rule	Ratio of assets X_{13}	Dynamics indicator of assets	Current assets turnover	0.25

Table 2. Results of ANFIS forecast - per cent and (actual)

	Going enterprises	Failed enterprises
Bankruptcy is forecasted	19.40 % (13)	81.43 % (57)
Bankruptcy possibility is very low	80.60 % (54)	18.57 % (13)

Table 3. Results of forecasts with Altman's Z-score function

	Going enterprises	Failed enterprises
Bankruptcy is forecasted	18.18	74.07
Bankruptcy possibility is very low	81.82	25.93

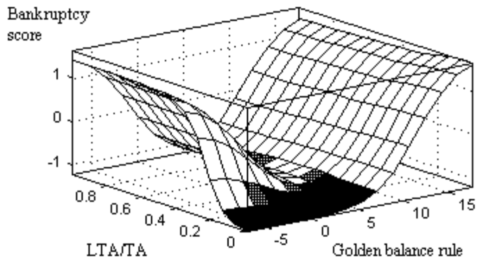

Fig. 3. A plot of the prediction surface of the ANFIS model using the ratio of Golden Balance Rule and the ratio of long-term assets LTA/TA as inputs.

Program Matlab also has a capability to represent in graphical form the information which fuzzy rules of ANFIS fire when the specific input data are entered. The forecast can be obtained by several rules. That is a specific feature of the fuzzy inference system, determined by the fact that the particular input value can belong to several membership functions at the same time (Fig. 1). Therefore this feature lets interpret the obtained forecast and its reasons in several ways.

The formed during the ANFIS training fuzzy rules IF . . . THEN were saved in the text format. This file was edited additionally. Only these rules which conclusions part THEN O is mfo had mfo value greater than 0.5 were chosen as they showed that bankruptcy is possible. There were 97 such rules obtained. But some factor's membership functions were adjacent. Such rules were joined into one rule. For example, two rules

IF *(Ratio of Golden Balance Rule is large) and (Dynamics indicator of Golden Balance Rule is small) and (Ratio of assets is medium) and (Dynamics indicator of assets is large) and (Current assets turnover is small)* THEN *bankruptcy is possible*

IF *Ratio of Golden Balance Rule is large) and (Dynamics indicator of Golden Balance Rule is small) and (Ratio of assets is medium) and (Dynamics indicator of assets is large) and (Current assets turnover is medium)* THEN *bankruptcy is possible*

differ just by adjacent membership functions of the fuzzy factor Current assets turnover, so they both can be joined into a single rule

IF *(Ratio of Golden Balance Rule is large) and (Dynamics indicator of Golden Balance Rule is small) and (Ratio of assets is medium) and (Dynamics indicator of assets is large) and (Current assets turnover is small or medium)* THEN *bankruptcy is possible.*

The membership functions of the factor current assets turnover are given in Fig. 4. The form of these functions and the overlapping extent depends on their parameters that are adjusted by neural networks.

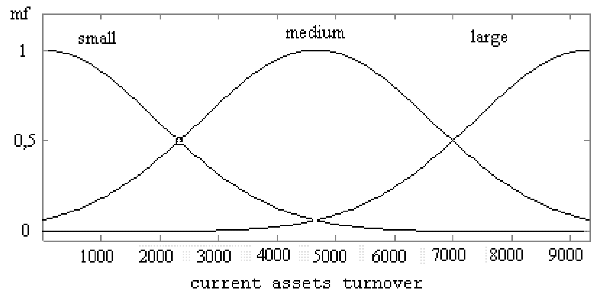

Fig. 4. Membership functions of the input factor current assets turnover.

The samples of obtained rules are shown in the Table 4. In this way 36 linguistic rules of possible bankruptcy were obtained. Such linguistic rules may be convenient for managers in qualitative analysis and decision making.

Table 4. Examples of linguistic rules when ANFIS output prediction was "bankruptcy is possible"

Ratio of Golden Balance Rule	Dynamics indicator of Golden Balance Rule	Long-term assets ratio	Dynamics indicator of long-term assets ratio	Current assets turnover
large	small	medium	large	small or medium
medium	medium	large	any	medium
medium	large	large	medium	small
large	small	small or medium	small	small or medium
large	small	medium	large	small or medium

5 Conclusions

The research has revealed that computer potential which Lithuanian companies and state institutions usually have in dispose lets develop ANFIS model using five financial ratios.

It has been noticed that the almost equal training error appears in several different subsets of financial ratios. It can be explained that analyzed ratios are connected with complex and not evident links. This testifies the nonlinear ANFIS prediction surface (Fig. 3).

The test results using testing data and ratio of Golden Balance Rule, dynamics indicator of Golden Balance Rule, ratio of assets, dynamics indicator of assets and current assets turnover showed that forecast error was less than 20 per cent.

The developed ANFIS model lets estimate any enterprise in the aspect of bankruptcy. Besides, the set of fuzzy rules of this model can be used as interpretable information about the reasons determining bankruptcy. If the software used for ANFIS enables to trace which fuzzy rules fire, when the data of specific enterprise is entered into the trained ANFIS, then these rules may give for manager's quite concrete information in the linguistic form about the reasons of the problem.

References

1. J. Mackevičius, *Analysis of enterprise perfomance*, TEV, Vilnius, 2005, ISBN 9955-680-13-X (in Lithuanian).

2. A. H. Amershi, E. H. Feroz, The occurrence of Finobacci numbers in time series of financial accounting ratios: Anomalies or indicators of firm survival, bankruptcy and fraud? An exploratory study, *Managerial Finance*, **26**(11), pp. 5–20, 2000.

3. W. Beaver, Financial ratios as predictors of failure, *Journal of Accounting Research*, **4**, pp. 71–127, 1966.

4. E. Altman, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, **23**(4), pp. 589–609, 1968.

5. A. P. Kovaliov, *Diagnosis of bankruptcy*, Finstatinform Moscow, 1995 (in Russian).

6. Thorsen Poddig, Bankruptcy Prediction: A Comparision with Didcriminant Analysis, in: *Neural Networks in the Capital Markets*, John Wiley & Sons, Chichester, pp. 311–323, 1995.

7. T. N. Batova, N. V. Katunkina, Analysis of enterprise bankruptcy forecast methods, *Problems of economical and human education in technical higher school*, **211**(17), pp. 281–284, 2001 (in Russian).

8. V. Boguslauskas, A. Stundžienė, Valuation of Bankruptcy Risk for Lithuanian Companies, *Engineering Economics*, **49**(4), pp. 29–36, 2006.

9. D. Vlachos, Y. A. Tollias, Neuro-Fuzzy Modelling in Bankruptcy Prediction, *Yugoslav Journal of Operations Research*, **13**(2), pp. 165–174, 2003.

10. S. Balcaen, H. Ooghe, 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems, *British Accounting Review*, **38**, pp. 63–93, 2006.

11. O. Purvinis, P. Šukys, R. Virbickaitė, Research of possibility of bankruptcy diagnostics applying neural network, *Engineering Economics*, **41**(1), pp. 16–22, 2005.

12. J.-S. R. Jang, ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Transactions on Systems, Man, and Cybernetics*, **23**(3), pp. 665–685, 1993.

13. A. Leonenkov, *Fuzzy modelling in the Matlab and fuzzyTECH environment*, BHV-Peterburg, Sankt-Peterburg, 2003, ISBN 5-94157-087-2 (in Russian).