APPLICATION OF FEEDFORWARD MULTILAYER NEURAL NETWORK AND RECURRENT NEURAL NETWORK IN THE PREDICTION OF BUSINESS INSOLVENCY - COMPARATIVE ANALYSIS

TOMASZ KOROL

Gdansk University of Technology – Faculty of Management and Economics - Poland
E-mail: tomasz.korol@zie.pg.gda.pl

Abstract - This article is about prediction of the bankruptcy of enterprises in Central Europe. The objective of the conducted research is to develop a predictive model using the feedforward multilayer neural network and recurrent neural network for the process of forecasting the risk of corporate bankruptcy. The effectiveness of the programmed artificial neural network models was verified with the use of two separate forecasting horizons – one year and two years before the bankruptcy of enterprises. The author carried out two research approaches characterized by different selection of input variables - minimization (with the use of only four financial ratios) and maximization (with the use of 28 different variables). The article presents also comparative analysis of the results of artificial neural networks from these two research approaches.

Index Terms - Artificial intelligence, bankruptcy, forecasting. The study has been prepared within the grant project No. 2015/19/B/HS4/00377, "Trajectories of life and the collapse of companies in Poland and in the world - identification, evaluation and forecast." Research funded by the National Science Centre in Poland (Narodowe Centrum Nauki).

I. INTRODUCTION

The importance of prognostic research has increased considerably after the global financial crisis that started in 2007. Today, the economic forecast - as stated at the Prague Forecast Symposium - is not a prophecy of the possibility of spontaneous development, but a scientific justification for the strategy of economy and ways of direct influence on economics (Zeliaś, 1997). Due to the steady, structural increase in the number of bankruptcies around the world - as a result of increased global competition - a precarious analysis of the risk of bankruptcy of businesses has become even more important today than it was in the past. The scale and negative economic and social consequences of bankruptcy induce a careful analysis of its causes and mechanisms, and force adequate precautions to be taken. One of the goals of business managers is to anticipate and counteract difficult situations and assess the financial condition of counterparties from the point of view of their solvency. In this situation, it is natural to forecast the corporate risk bankruptcy. The key issue in today's business environment is to identify areas of risk, monitor current financial situation, and effectively forecast bankruptcy risks to respond in advance.

This issue is of great importance, constantly seeking new and better solutions, also referring to other fields than economics such as physics or biology (artificial intelligence). Nowadays, analysts are no longer confronted with the dilemma of whether to forecast a risk of corporate bankruptcy, but what method to use to evaluate the financial situation to minimize forecast errors. In this article the author will try to answer this important question. That is why the objective of presented research in this paper is to develop feedforward multilayer neural network and recurrent neural network forecasting the bankruptcy of enterprises in Central Europe. Author also conducts comparative analysis of effectiveness of such two types of artificial neural networks in two research approaches – with minimization and maximization of number of the entry variables (financial ratios) to the models. The article consists of six sections. In the introduction, the author presents the justification for the topic and the objective of the research. In the second section brief characteristics of artificial neural networks is given. The author then presents the research assumptions and discusses the results of these studies. In the last section the comparative analysis of results is conducted and detailed conclusions are presented.

II. BRIEF CHARACTERISTICS OF ARTIFICIAL NEURAL NETWORKS

The concept of artificial neural networks is understood as mathematical models composed of networks of computing nodes called neurons and their connections which, simulating the action of biological systems can effectively solve specific problems (Chen et al., 2015; Min and Lee, 2005). In contrast to e.g. multivariate discriminant analysis, the essence of activity of neural networks is a purely mechanical approach to the analyzed phenomenon, without detection of internal relations and the strength of existing relationships (Baetge and Krause, 1993).

With regard to architecture, neural networks can be divided into (Gajda, 2001, p. 222):

1. Feedforward neural networks – i.e. networks with one-direction flow of signals:
- monolayer, containing a single layer of output neurons, with no hidden neurons,
- multilayered, containing more than one layer of neurons - beside output layer they contain at least one hidden layer;

2 .Recurrent neural networks - in which there are feedback loops between the input, output and hidden layers. Neurons of recurrent network can thus react to own stimulations in response to signals associated with the previous observation (bi-directional data flow);

3. Modular neural networks - in which mutual coupling between processing elements relates only to neighborhood.

The most common type of neural networks in predicting bankruptcy of enterprises is feedforward multilayer neural network (Ravi and Ravi, 2007; Lin et al., 2012; Jayanthi et al., 2011; Brabazon et al., 2012), in which the signal flow is only in one direction, i.e. from the input, where the network takes input data, through the hidden layer, where the main processing of neural signals takes place, to the output, where the network provides a solution. The network is also called a multilayer perceptron (MLP). Determining the number of hidden neurons is not an easy task. Although the literature gives formulas to determine the optimal hidden layer (formulas to determine the number of hidden neurons: n/2; n ; n+1; 2n+1; where n – number of input neurons), the authors of publications on neural networks postulate not accepting them a priori, but rather designation of the number of neurons in each individual case, depending on the problem being solved (Zhang et al., 1999, pp. 24; Rahimian and Singh, 1993).

III. RESEARCH ASSUMPTIONS

The use of artificial neural networks to assess the risk of corporate bankruptcy requires a well-prepared set of examples, on the basis of which the network will be taught and tested. Each of them must contain information describing the situation of the enterprise in the form of input variables representing measurable characteristics, such as financial ratios and non-financial attributes such as, for example - experts' scores assigned to certain numerical values (rank) at a fixed scale. They can act as input signals. In the research, the author of this article used data on 180 manufacturing companies from Central Europe that published their financial reports. Within this population of enterprises two samples are distinguished:
- Learning sample - consisting of 39 enterprises at high risk of bankruptcy and assigned to them 39 non-bankrupt companies. These companies came from different manufacturing industries, where they were matched in pairs, ie the potential bankrupt was assigned to a "healthy" company in the same industry. For the purpose of research, the learning sample was developed for one year and for two years before the bankruptcy.
- Testing sample - also consisting of 39 companies in danger of bankruptcy and 39 "healthy" companies. However, in 24 cases, in the analysis for one year before bankruptcy author used the financial statements of other bankrupt companies than in the analysis for two years before the insolvency of firms.

It should be noted that in these studies the companies were classified as being at risk of bankruptcy on the basis of two criteria:
- court judgments stating the opening of the bankruptcy proceedings,
- declaration of bankruptcy.

The author used 27 financial ratios and one non-financial variable - the operating region of a given company. The results of studies on the intensity maps of bankruptcy risk of companies in individual voivodships, conducted and published earlier by the author were used to develop this variable. A full list of input variables used in the study is presented in the Table 1. As mentioned earlier, each of the 180 enterprises analyzed was described by 28 numerical characteristics, the last of which represented a non-economic variable. In addition, each company was described by an output variable, ie a variable grouping populations into two groups of companies - at risk of bankruptcy and not threatened with bankruptcy. Additionally, the author has used two research approaches of selecting diagnostic variables as entry variables to created artificial neural networks:

- first research approach - T1: the set contains all the diagnostic variables presented in Table 1,
- second research approach - T2: variables are derived based on the correlation matrix analysis.

In the second approach, using the data from Table 1, correlation analysis of individual diagnostic variables was performed. The input data for the model was derived from the correlation matrix, selecting only those features that are poorly correlated with each other and strongly correlated with the grouping variable representing the information if the enterprise is at risk of bankruptcy or a "healthy" company. This

approach provides a selection of features that do not duplicate the information provided by other indicators .

Symbol of input variable	Variable type and calculation method
Liquidity ratios	
X1	Current assets / current liabilities
X2	[current assets - inventories] / current liabilities
X3	cash / current liabilities
X4	Working capital / total assets
Profitability ratios	
X5	Profit from sales / total assets
X6	Operating income / total assets
X7	Operating income / sales
X8	Net income / (stockholders' equity – net income)
X9	Income before taxes / total assets
X10	Net income / total assets
X11	Income before tax / total revenues
Debt ratios	
X12	Total liabilities / total assets
X13	Stockholders' equity / total liabilities
X14	(stockholders' equity + non-current liabilities) / fixed assets
X15	Current liabilities / total assets
X16	(net income + depreciation) / total liabilities
X17	Income before taxes / current liabilities
X18	Current liabilities / stockholders' equity
Activity ratios	
X19	sales / short term receivables
X20	Operating costs / inventories
X21	Operating costs / current liabilities
X22	sales / total assets
X23	Inventories / total revenues
Structural ratios	
X24	Stockholders' equity / total assets
X25	Fixed assets / total assets
X26	Cash / total assets
X27	Current assets / total assets
Non-financial information	
X28	voivodships – podkarpackie, świętokrzyskie, małopolskie - +5 voivodships – łódzkie, śląskie – (-5) other voivodship - 0

Tab. 1. Independent variables used in the research

The effectiveness of the created models was assessed on the basis of the overall effectiveness of the model, as well as on the type I and type II errors. The following formulas were used:

- Type I error - E1 = D1 / BR • 100%, where D1 – the number of bankrupted firms classified by the network as "healthy" firm, BR - number of bankrupted enterprises in the testing / learning sample;
- Type II error - E2 = D2 / NBR • 100%, where D2 - number of non-bankrupt companies classified by the network as insolvent companies, NBR - number of non-bankrupt firms in the testing / learning sample;
- Overall effectiveness of the network - S = {1 - [(D1 + D2) / (BR + NBR)]) • 100%.

The aim of the study is to compare the effectiveness of two types of artificial neural network models (recurrent network and feedforward multilayer network) for one year and two years before bankruptcy. The behavior of these models was also examined and compared in two different cases, ie, the maximization of input data (T1 research approach) and the minimization of input information to the model (T2 research approach). With the use of such two research approaches author will be able to answer important question that is still not resolved in literature. The key question is the appropriate choice of network structure for the topic of forecasting the risk of corporate bankruptcy, ie the presence or absence of feedback between neurons. The feedforward multilayer network is characterized by the absence of such feedback. Such models are always characterized by more stable behavior, which is its advantage. However, if the network has built-in feedback (it includes return connections between neurons), then it can perform more complex calculations (recurrent neural networks). A network with fewer neurons, which contains feedback, can perform as complex calculations as a network without feedback connections, containing significantly more neurons. The author's own research has allowed the empirical verification of the suitability of individual types of artificial neural networks in predicting bankruptcy of enterprises in Central Europe.

IV. COMPARATIVE ANALYSIS OF MODELS RESULTS FROM THE T1 RESEARCH APPROACH FOR ONE YEAR AND TWO YEARS BEFORE BANKRUPTCY

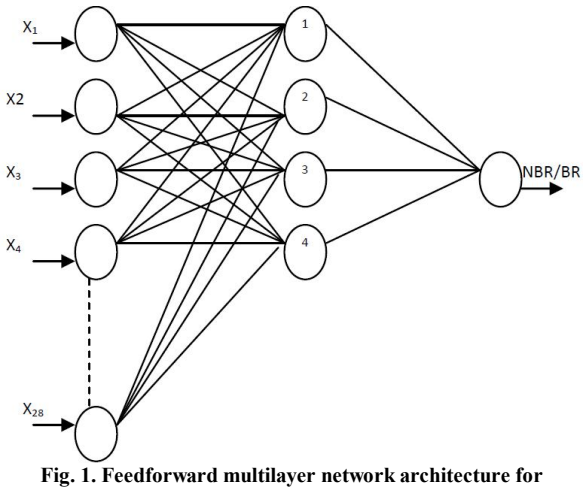


Fig. 1. Feedforward multilayer network architecture for forecasting the risk of corporate bankruptcy for one year and two years before (T1 research approach)

In the analysis of companies for one year and two years before bankruptcy, the author programmed a neural

network with architecture presented in Figure 1 for feedforward multilayer network and for the recurrent network in Figure 2. Both models were programmed with the use of maximization input data. Both models contain 28 entry neurons (input variables from Table 1 - from X1 to X28). Created models contain also 4 hidden neurons and 1 output neuron, representing the classification of companies at risk or not at risk of bankruptcy.

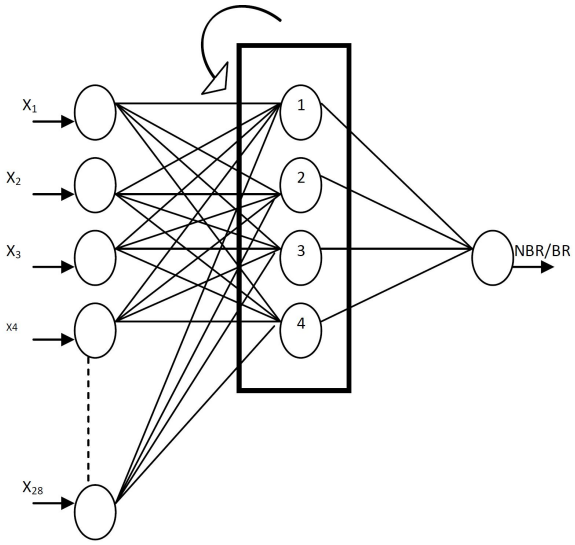


Fig. 2. Recurrent neural network architecture to study companies for one year and two years before bankruptcy (T1 research approach)

Comparing the results of the classification of enterprises based on testing sample (Table 2) obtained by two types of artificial neural networks, it can be seen that the overall effectiveness for both models (feedforward multilayer network and recurrent neural network) one year and two years before bankruptcy is very high - above 92%.

A more detailed analysis of Type I error and Type II error shows a difference in the quality of forecasts generated by these networks. Although one year before the bankruptcy of companies, the overall effectiveness of feedforward multilayer network (so called multilayer perceptron) and the recurrent network is the same, the cost of the error in the recurrent network is much higher than the error of a multilayer perceptron. The model without feedback in the hidden layer in the analysis one year before did not generate a single error of Type I. Type I error means the classification of future bankrupt company as a "healthy" enterprise without bankruptcy risk in future. The literature emphasizes that the cost of such an error is much greater than in the case of Type II error. The recurrent network generated Type I error at the level of 2.56%, without making the Type II error, which was generated by multilayer perceptron. Nevertheless, in each case these errors are very small. In the analysis for two years before the bankruptcy of enterprises, the overall effectiveness of two types of artificial neural networks differs by 1.28 percentage points in favor of a recurrent network. Both networks generated the same Type II errors (5.12%). Type I errors are lower by 2.56 percentage points when recurrent neural network is used in forecasting bankruptcy of companies. It can be concluded that as the forecast horizon increases, it is more beneficial to use a recurrent neural network.

Results	Multilayer feedforward network		Recurrent neural network	
	One year before	Two years before	One year before	Two years before
Type I error - E_1	0%	10.25% (4 firms)	2.56% (1 firm)	7.69% (3 firms)
Type II error – E_2	2.56% (1 firm)	5.12% (2 firms)	0%	5.12% (2 firms)
Overall effectivness – S	98.72%	92.31%	98.72%	93.59%

Table 2. Results of the classification of enterprises for one year and two years before the bankruptcy based on testing sample of models from the T1 approach

V. COMPARATIVE ANALYSIS OF MODEL RESULTS FROM THE T2 RESEARCH APPROACH FOR ONE YEAR AND TWO YEARS BEFORE THE BANKRUPTCY OF ENTERPRISES

While in the previous research approach T1 the structure of the two types of networks contained the same network architecture - the same number of hidden layers (1 hidden layer) and hidden neurons (4 hidden neurons), in the case of the T2 research approach with minimizing input variables to models, it was necessary to apply a completely different network architecture. In the case of feedforward multilayer network, in the forecasting horizon of two years before bankruptcy, the network architecture had to be used with three layers of hidden neurons (Figure 4). This ensured the best model performance for both learning and testing samples. The need to expand the network by two consecutive hidden layers shows the complexity of the problem of the bankruptcy of enterprises. However, in the analysis for one year before the fall of the companies, the network architecture with one hidden layer was used (Figure 3).

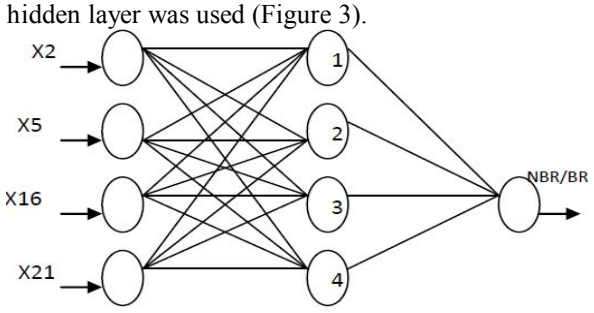


Figure 3. Architecture of the feedforward multilayer network of the T2 research approach for the forecasting horizon of one year

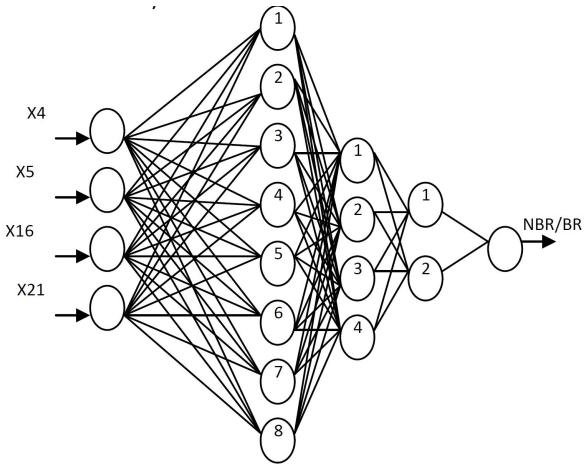


Figure 4. Architecture of feedforward multilayer network of the T2 research approach for the forecasting horizon of two years

In the case of a recurrent network, one hidden layer with feedback was enough to apply in the analysis of one year and two years before bankruptcy of companies (Figure 5).

In all created artificial neural networks at the entrance of the models there were 4 neurons constituting input variables (differing depending on the forecasting horizon) and 1 exit neuron presenting the classification of enterprises at risk (BR) and not at risk of bankruptcy (NBR).

The author obtained 100% effectiveness of the above models in case of learning sample. Comparing the accuracy of the classification of enterprises based on testing sample between the recurrent network and the feedforward multilayer network, it can be seen that, similarly to the models from the T1 approach in forecasting horizon of one year, both types of networks showed the same overall effectiveness – 97.44% (Table 3). As in the previous case, the multilayer perceptron did not generate any errors of Type I. The recurrent neural network did not generate any errors of Type II, thus making only the more costly errors of Type I.

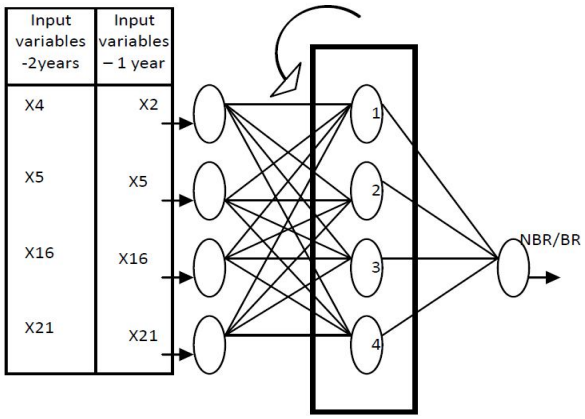


Figure 5. Architecture of recurrent neural network of the T2 research approach for forecasting horizon of one year and two years

In the analysis for two years before bankruptcy, in the case of a multilayer perceptron, significant errors of Type I occurred - amounting to as much as 15.38% (Table 3). Also Type II errors increased by 7.7 percentage points. In the case of a recurrent network, the overall effectiveness was clearly higher than that of the multilayer perceptron (92.31% versus 85.9%). The recurrent network generated over 50% smaller Type II errors and 5.1 percentage points smaller Type I errors than feedforward neural network. The need to expand hidden layers (from one layer to 3 layers), smaller overall effectiveness by 6.41 percentage points and higher Type I and Type II errors in case of multilayer perceptron compared to the recurrent network clearly demonstrate the high computational potential of the created recurrent network.

Results	ANN multilayer feedforward		ANN recurrent	
	One year before	Two years before	One year before	Two years before
Type I error - E_1	0%	15.38% (6 firms)	5.12% (2 firms)	10.25% (4 firms)
Type II error – E_2	5.12% (2 firms)	12.82% (5 firms)	0%	5.12% (2 firms)
Overall effectiveness	97.44%	85.90%	97.44%	92.31%
– S				

Table 3. Results of the classification of enterprises in forecasting horizon of one year and two years before the bankruptcy based on testing sample (T2 research approach)

CONCLUSION

The conclusions of these studies are significant. The research conducted by the author showed that artificial neural network models are characterized by high effectiveness in forecasting the risk of corporate bankruptcy. It has been also proved that the recurrent neural network can perform complex calculations containing much less hidden neurons than the feedforward multilayer network. The minimization of the input data had significant negative impact on the multilayer perceptron causing the reduction of overall effectiveness by even 6.41 percentage points. In case of maximization of input data it can be noted that both types of artificial neural networks exhibit similar high predictive properties. Additionally, in these studies author has also proved that in the process of forecasting the risk of corporate bankruptcy in Central Europe, it is possible to use not only financial ratios but also nonfinancial variables. The author of this paper is going to continue the research on using wide variety of nonfinancial variables that can help improving the process of recognition of risk of bankruptcy of enterprises.

REFERENCES

[1] Baetge J., Krause C. (1993), The Classification of Companies by Means of Neural Networks, „Journal of Information Science and Technology", Vol. 3.

[2] Brabazon A., Dang J., Dempsey I., O'Neill M., Edelman D. (2012), Natural computing in finance: a review. In: Rozenberg G., Back T., Kok J. (eds), Handbook of natural computing. Springer, Berlin, pp. 1707–1735.

[3] Chen N., Ribeiro B., Chen A. (2015), Financial credit risk assessment – a recent review, "Journal of Artificial Intelligence Review", no 45, pp. 1-23.

[4] Gajda J. (2001), Prognozowanie i symulacja a decyzje gospodarcze [Forecasting and simulation and economic decisions], Publishing House of C.H.Beck, Warsaw, p. 222.

[5] Jayanthi J., Joseph K., Vaishnavi J. (2011), Bankruptcy prediction using SVM and hybrid SVM survey, "International Journal of Computer Applications, no. 33(7), pp. 39–45.

[6] Lee K., Booth D., Alam P. (2005), A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms, „Expert Systems with Applications", no 29, pp. 1–16.

[7] Lin W., Hu Y., Tsai C. (2012), Machine learning in financial crisis prediction: a survey, "IEEE Transactions on Systems, Man and Cybernetics part C", no. 42, pp.421–436.

[8] Min J., Lee Y. (2005), Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters, „Expert Systems with Applications", no 28, pp. 603–614.

[9] Rahimian E., Singh S. (1993), Bankruptcy Prediction by Neural Networks, in: Trippi R., Neural Networks in Finance and Investing, Probus Publishing Company, Chicago.

[10] Ravi K.P., Ravi V. (2007), Bankruptcy prediction in banks and firms via statistical and inteligent techniques - a review, "European Journal of Operational Research", no. 180, pp. 1–28.

[11] Zeliaś A. (1997), Teoria prognozy [The theory of forecast], PWE, Warsaw.

[12] Zhang G., Hu M., Patuwo B., Indro D. (1999), Artificial neural networks in bankruptcy prediction – general framework and cross-validation analysis, "European Journal of Operational Research", no. 116, pp. 24.

★★★