DEFAULT PREDICTION USING NEURAL NETWORKS FOR ENTERPRISES FROM THE POST-SOVIET COUNTRY

Vendula Fialova[1,a] and Andrea Folvarcna[2,b,*]

[1]Institute of Entrepreneurship and Marketing, The College of Entrepreneurship and Law, Vltavská 585/14, 150 00 Praha 5, Czech Republic
[2]Institute of Entrepreneurship and Marketing, The College of Entrepreneurship and Law, Vltavská 585/14, 150 00 Praha 5, Czech Republic
[a]vendula.fialova@vspp.cz, [b]andrea.folvarcna@vspp.cz
*Corresponding author

Abstract: The application of neural networks for the classification of companies into a group of prosperous or non-prosperous companies is relatively new. Their use for this purpose has been the subject of research only since the early nineties of the last century. The work focuses on the possibilities of predicting the financial health of companies in the post-Soviet countries, especially in the Czech Republic, using neural networks. The aim is to create a model for predicting solvency or insolvency. The first part is a brief introduction to the issue and an explanation of the reason for the selection of neural networks for processing. Due to the rapid development of computer technology, neural networks are increasingly used in experimental and practice tasks. Application tasks such as information processing, classification of patterns or situations, optimization problems, and prediction tasks are used in industry, management, finance, telecommunications, military technology, healthcare, and elsewhere. The following methodology explains the procedure and the database processing method used. The results contain the output obtained from the SPP program in which the model was generated. We examined a total of 452 companies, some of which entered the training and testing phase. The results are interpreted graphically and in tables. The discussion and conclusion section summarizes the findings and possible possibilities for further processing using neural networks.

Keywords: default, solvency, prediction, neural networks

JEL Classification: C53, C45, G17

1. Introduction

Several prediction models have been created so far in the field of evaluation and prediction of the company's financial situation. In the recent period, we have witnessed many trends, the importance of which is intensifying in the economy. (Mihalovic, 2018, Faraglia et al., 2018)

The massive rise of information technology has brought innovative methods and possibilities for predicting stock exchange assets using mathematical models. From time immemorial, people have been fascinated by predicting quantities that seem chaotic at first glance. However, by using the knowledge of mathematical and statistical analysis, linear algebra, together with the right dose of human effort, regularities, among phenomena that act chaotically, can be correctly mathematically expressed and used as a basis for predicting states of the future. Linear

Default prediction using neural networks for
enterprises from the post-soviet country
Authors: Vendula Fialova, Andrea Folvarcna

models have become an advanced domain of investor decision-making for decades, mainly due to their relatively simple interpretation of the phenomena under study. Over time, however, linear models were not enough to meet the requirements of investors, and the era of development of models capable of capturing nonlinear relationships of the examined systems began.

Based on Stone-Weierstrass's theorem that any function can be approximated with an absolute precision by a suitably chosen higher order polynomial, it is possible in this direction to consider forward neural networks as a tool of so-called universal approximator, because they can approximate any continuous function with the desired accuracy, all with a suitable design of the network structure, i. the number of neurons in the input layer, the number of hidden layers and neurons and the number of neurons in the output layer, as well as the torture algorithm. (Starkloff, 2019; Auerbach, 2018)

Wilson and Sharda (1994) presented an exploratory study that compares the predictive capabilities of reliable neural network bankruptcy and classical multivariate discriminant analysis. They predicted the predictive accuracy of these two techniques in a complex, statistically reliable framework that suggests added value to the prediction problem of each method. The study suggests that neural networks have significantly better analysis than discriminatory analysis in predicting corporate bankruptcies.

Another author dealing with bankruptcy prediction using different methods was du Jardin (2010), who evaluated the accuracy of forecasts of models designed using different classification methods. (Starkloff He pointed out that a neural network-based model leads to better results than a set selected with criteria used in the financial literature. It also shows that how the set of variables can represent the business profiles of healthy companies plays a role in reducing errors. (Klimenka and Wolter, 2019)

Odom and Shard are considered to be pioneers in the use of neural networks for financial prediction. (Shah and Murtaza, 2000) Odom and Sharda used ratio financial indicators from the calculation of Altman's Z-score as inputs to neural networks. The authors compared the reliability of neural networks with the assurance of multidimensional discriminant analysis. In the given research, neural networks in comparison with multidimensional discriminant analysis proved to be an analytical tool with four better predictive ability, which subsequently aroused increased interest in their use to predict business bankruptcy. Therefore, we chose neural networks in this case. (Fleiner et al., 2019)

Since the 1960s, many studies (Altman, 1968; Beaver, 1967, 1968; Blum, 1969; Deakin, 1972; Du Jardin, 2018) have shown that an appropriately weighted set of financial and economic indicators can be used effectively to assess the risk of failure in companies.

Such studies mainly used "traditional" statistical methods, such as multivariate discriminant analysis (MDA) and logistic regression (LRA). In addition, attention has almost always been focused on large and medium-sized enterprises. Only a small number of studies have shown that specific failure risk modeling systems are needed to assess the risk profiles of smaller firms (Edmister, 1972; Altman and Sabato, 2005, 2006; Ciampi and Gordini, 2008; Vallini et al., 2008; Clementi and Palazzo, 2019).

2. Methodology

When processing this issue, we had data on 452 companies in the Czech Republic. The database processing method we have chosen is a neural network using the SPSS program. Also based on the analysis of the literature, we decided to use neural networks in the prediction, as they have proven themselves in solving many problems and can have the most practical effect

in the fields of modeling, forecasting, and signal processing and expert systems (Lippmann, 1987). The predictive power of neural networks falls within the prediction area.

The basic structure of neural networks is formed by an oriented graph consisting of vertices (neurons), which are arranged in layers and are connected by edges (synapses) (Du and Swamy, 2019, Fritsch, 2019). The input layer consists of all input variables in separate neurons, and the output layer consists of dependent variables. Frisch et al. (2019) describe the function of the simplest multilayer perceptron (MLP):

$$o(x) = f\left(w_0 + \sum_{i=1}^{n} w_i x_i\right) = f(w_o + w^T x) \tag{1}$$

Where w_0 denotes a constant, $w = (w_1, ..., w_n)$ is a vector containing all synaptic weights without a constant and $x = (x_1, ..., x_n)$ is a vector of all input variables. To increase the flexibility of the modeling, hidden layers can also be attached. The function of the MLP with a hidden layer consisting of J hidden neurons is calculated as:

$$o(x) = f\left(w_0 + \sum_{j=1}^{J} w_j * f\left(w_{oj} + \sum_{i'1}^{n} w_{ij} x_i\right)\right)$$

$$= f\left(w_0 + \sum_{j=1}^{J} w_j * f\left(w_{oj} + \sum_{i'1}^{n} w_j^T x\right)\right) \tag{2}$$

where w_0 denotes the constant of the output neuron and w_{0j} the constant of the j-th hidden neuron, w_j denotes the synaptic weight corresponding to the synapse starting at the j-th hidden neuron and leading to the output neuron, and $w_j = (w_1, ..., w_n)$ is the vector of all synaptic weights synapses leading to the nth hidden neuron (Mihalovic, 2018). All hidden and output neurons count the output:

$$f\big(g(z_0, z_1, ..., z_k)\big) = f(g(z)) \tag{3}$$

from the outputs of all previous neurons $z_0, z_1, ..., z_k$, where the integration function g and the activation function f are:

$$g: R^{k+1} \to R$$
$$f: R \to R \tag{4}$$

The integration function can be defined as:

$$g(z) = w_0 z_0 + \sum_{i=1}^{k} w_i z_i = w_0 + w^T z \tag{5}$$

The logistic sigmoid function was used as an activation function due to the binary nature of our output variables. This function displays the outputs of each neuron in the interval [0,1]. (Mihalovic, 2018)

The neural network calculates the output o (x) for the given inputs x, and the balance, and the learning algorithm calculates the error function. When modeling the network, the sum of squares of deviations (SSE) was used, which indicates the difference between the predicted and observed outputs, where $l = 1, ..., L$ represents the given input-output pairs and $h = 1, ..., H$ are output nodes. (Mihalovic, 2018)

Default prediction using neural networks for
enterprises from the post-soviet country
Authors: Vendula Fialova, Andrea Folvarcna

$$E = \frac{1}{2}\sum_{l=1}^{L}\sum_{h=1}^{H}(o_{lh} - y_{lh})^2 \tag{6}$$

The prediction model with the smallest selection error among all the applied combinations is selected as a representative training model. The results, accuracy, and performance of which will then be tested on a test data set. (Mihalovic, 2018; Todorov, 2020)

3. Results

In this research paper, we focused on analyzing solvent and bankrupt companies in the Czech Republic using neural networks. The aim is to create a model that would be able to predict companies' solvency or insolvency. Such a model could be useful for banks when deciding on lending. The total number of analyzed entities represents 452 companies from the Czech Republic. Of these, 91 companies defaulted.

In the position of the dependent variable, we chose the answer of the company yes/no to the question concerning the previous default. We chose liquidity as a factor and determining indicators in the covariance, such as the company's age, ROA, ROE, total indebtedness, and long-term indebtedness.

Table 1 of the Case Processing Summary shows how many cases were included in the training phase and how many were held and were part of the testing phase. In this case, 67% entered the training, and the remaining 33% were tested.

Table 1: Case Processing Summary

		N	Percent
Sample	Training	303	67.0%
	Testing	149	33.0%
Valid		452	100.0%
Excluded		0	
Total		452	

Source: Output from SPSS

The following Table 2 provides the necessary network information. The already mentioned inputs, which are ascertained in the input layer, are a factor of liquidity and covariates: company´s age, ROA (return on assets), ROE (return on equity), total indebtedness, long term indebtedness. There is only one hidden layer within which there are six units.

Table 2: Network Information

	Factors	1	Liquidity
InputLayer		1	AGE
	Covariates	2	LONG-TERM INDEBTEDNESS
		3	ROE
		4	ROA

		5	TOTAL INDEBTEDNESS
	Number of Units[a]		8
	RescalingMethodforCovariates		Standardized
	Number of HiddenLayers		1
HiddenLayer(s)	Number of Units in HiddenLayer 1[a]		6
	ActivationFunction		Hyperbolic tangent
	DependentVariables	1	DEFAULT
	Number of Units		2
Output Layer	ActivationFunction		Softmax
	ErrorFunction		Cross-entropy

a. Excludingthebiasunit

Source: Output from SPSS

The results from the training phase and the subsequent application of networks on the rest of the sample are shown in Table 3 Model Summary. Simultaneously, it testifies to the correctness of the model, as there is a relatively low error rate.

Table 3: Model Summary

	CrossEntropyError	102.707
	Percent IncorrectPredictions	14.2%
Training	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	TrainingTime	0:00:00.16
Testing	CrossEntropyError	49.634
	Percent IncorrectPredictions	14.1%

DependentVariable: DEFAULT

a. Errorcomputations are based on thetestingsample.

Source: Output from SPSS

The practical results of using the neural network are available for each case of yes/no answers based on a comparison of the obtained and predicted in percentage expression of a correct estimate (Table 4). From a training sample of 303 companies, he was able to correctly classify 260 companies. In the test sample, 128 companies out of 149 correctly classified.

Table 4: Classification

Sample	Observed	Predicted		
		no	yes	Percent Correct
Training	no	246	0	100.0%

	yes	43	14	24.6%
	Overall Percent	95.4%	4.6%	85.8%
	no	114	1	99.1%
Testing	yes	20	14	41.2%
	Overall Percent	89.9%	10.1%	85.9%

DependentVariable: DEFAULT

Source: Output from SPSS

Figure 1 shows clustered box plots predicted pseudo probability for a combination of training and testing samples.

Figure 1: Predicted Pseudo Probability
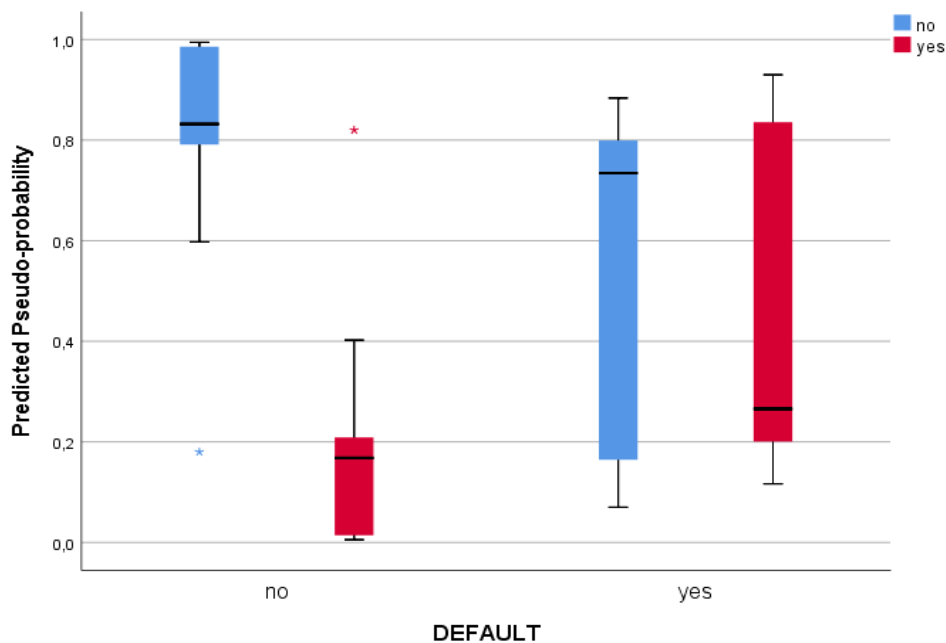
Source: Output from SPSS

The following figure 2 shows the interdependence between sensitivity and specificity for all possible cut-offs in a single plot. And in Graph 3, the number of total cases obtained from case responses is expressed as a percentage.

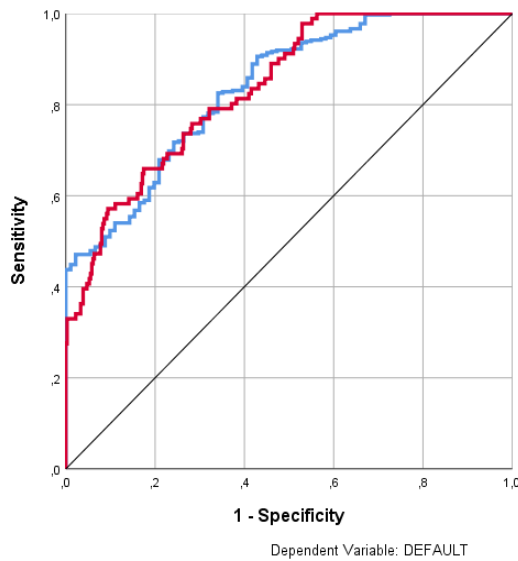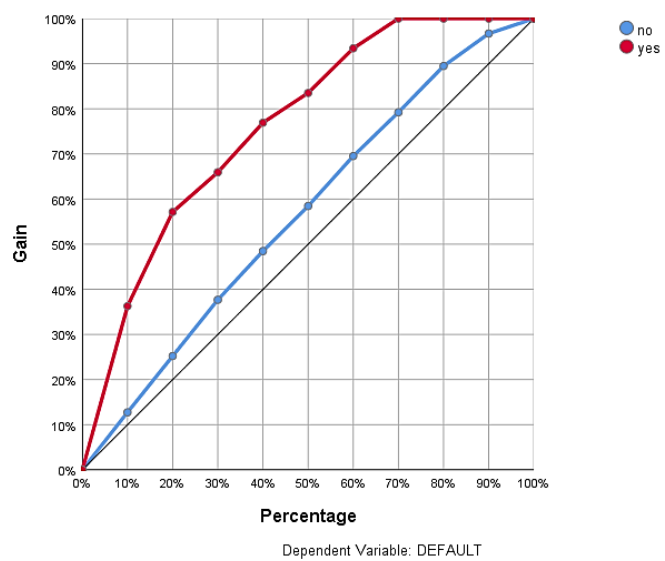Figure 2: Chart of Sensitivity & Specificity   Figure 3: Cumulative Gain´s Chart

Source: Output from SPSS

The total number of analyzed entities represents 452 companies from the Czech Republic. Of these, 91 companies defaulted.

In the position of the dependent variable, we chose the answer of the company yes/no to the question concerning the previous default. We chose liquidity as a factor and determining indicators in the covariance, such as the company's age, ROA, ROE, total indebtedness, and long-term indebtedness. To determine the optimal variables for compiling the final ANN model, a sensitivity analysis (also known as an independent Variable Importance) was performed in Table 5. The importance of an independent variable is the extent to which the value of the predicted network's modelchanges for different values of the independent variable. Sensitivity analysis is used to calculate the significance of each predictor. The importance (see Table 5) shows that liquidity dominates the results.

Table 5: Independent Variable Importance

	Importance	Normalized Importance
LIQUIDITY	.445	100.0%
AGE	.051	11.4%
LONG TERM INDEBTEDNESS	.046	10.3%
RETURN ON EQUITY	.053	11.9%
RETURN ON ASSETS	.239	53.6%
TOTAL INDEBTEDNESS	.167	37.4%

Source: Output from SPSS

4. Discussion

The study presents a significant contribution, especially in terms of the use of technically artificial plastic neural networks on the practical example of 155 companies in the Czech Republic. The results indicate the results of the interrelationships between inputs and outputs. Inputs, in a way, represent the cost of innovation, and the output is the net turnover of

*Default prediction using neural networks for
enterprises from the post-soviet country
Authors: Vendula Fialova, Andrea Folvarcna*

companies. In the future, these inputs could be extended by others or specified in more detail already established. The advantage of neural networks is that they can work with both quantitative data and qualitative data. Currently, neural networks are used in almost every field, and this technique has great potential. Neural networks are currently being used more and more in terms of their reliability. As early as 1999, Wang et al. (1999) used neural networks for the first time to assess credit risk in China, and the results pointed to the effectiveness of the neural network. At the same time, they have shown that it is better than discriminatory analysis. Zhang et al. (2003) also investigated neural networks and demonstrated that they perform highly accurate credit risk assessments. Neural networks also have their limitations. One of the inches is the lack of a backward methodology and an understanding of how the result was derived and accurate. Another limitation stems from the fact that neural networks were flooded designed for classification purposes. Thus, these should be interconnected values, but we often use this tool to look for connections between things that differ. In the case of deforming basic logic, it is necessary to spend more time developing an algorithm to solve the problem. It is an algorithm that needs to be trained for a considerable time to achieve its result. Often, however, it is a time when it is necessary to wait weeks for useful results, which may be unacceptable in certain situations.

5. Conclusion

Since the conditions of the Czech Republic, the prediction models do not have a sufficient degree of predictive accuracy. In the present study, a prediction model based on neural networks, the process of data learning and adaptation of the functional relationship between the input variables was proposed.

The proposed model can be updated, re-evaluated, and parameters changed in the future. A higher degree of adaptation of the model architecture to the conditions of the economic environment may even increase its predictive power. We consider its narrower specification as a possible direction of model development in the future. It could find greater application, for example, in detecting a more specific sample in terms of size of enterprises, age of enterprises or industry.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), pp. 589-609.

Altman, E. I. and Sabato, G. (2005). Effects of the new Basel capital accord on bank capital requirements for SMEs. *Journal of Financial Services Research*, *28*(1-3), pp. 15-42.

Altman, E. I. and Sabato, G. (2006). Modeling credit risk for SMEs: Evidence form the US market. *Abacus*, *19*(6), pp. 716-723.

Auerbach, A. J. (2018). Measuring the effects of corporate tax cuts. *Journal of Economic Perspectives*, *32*(4), pp. 97-119.

Beaver, W. (1967). Financial ratios predictors of failure. *Journal of Accounting Research*, Supplement to volume 4.

Beaver, W. (1968). Alternative accounting measures as predictors of failure. *The Accounting Review*, *43*(1), pp. 113-122.

Ciampi F. and Gordini, N. (2008). *Using economic-financial ratios for small enterprise default prediction modeling: An empirical analysis.* Proceedings of the 2008 Oxford Business & Economics Conference (Oxford, UK).

Clementi, G. L. and Palazzo, B. (2019). Investment and the cross-section of equity returns. *Journal of Finance, 74*(1), pp. 281-321.

Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research*, *10*(1), pp. 167-179.

Du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, *73*(10-12), pp. 2047-2060.

Du, K. L. and Swamy, M. N. S. (2019). Recurrent neural networks. *Neural Networks and Statistical Learning*, pp. 351-371, Springer, London.

Edmister, R. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis*, *7*(2), pp. 1477-1493.

Faraglia, E., Marcet, A., Oikonomou, R. et al. (2019). Government debt management: The long and the short of it. *Review of economic studies, 86*(6), pp. 2554-2604.

Fleiner, T., Jagadeesan, R. Janko, Z. et al. (2019). Trading networks with frictions. *Econometrica, 87*(5), pp. 1633-1661.

Fritsch, S., Guenther, F. and Guenther, M. F. (2019). Package 'neuralnet'. Training of Neural Networks. Recuperado de https://cran. r-project. org/web/packages/neuralnet/neuralnet. pdf.

Klimenka, F. and Wolter, J. L. (2019). Multiple regression model averaging and the focused information criterion with an application to portfolio choice. *Journal of Business and Economic Statistic, 37*(3), pp. 506-516.

Lippmann, R. P. (1987) An introduction to computing with neural nets. *IEEE ASSP Magazine*, pp. 4-22.

Mihalovic, M. (2018). Využitie skóringových modelov pri predikcii default ekonomický chsubjektov v Slovenskej republike. *Politická ekonomie*, *66*(6), pp. 689-708.

Shah, J. R. and Murtaza, M. B. (2000). A neural network based clustering procedure for bankruptcy prediction. *American Business Review*, *18*(2), pp. 80.

Starkloff, H. J. (2019). Stone-Weierstrass theorems for random functions. *Studia Universitatis Babes-Bolyai, Mathematica*, *64*(2).

Todorov, K. (2020). Quantify the quantitative easing: Impact on bonds and corporate debt issuance. *Journal of Financial Economics*, *135*(2), pp. 340-358.

Vallini, C., Ciampi, F., Gordini, N. and Benvenuti, M. (2008). *Can credit scoring models effectively predict small enterprise default? Statistical evidence from Italian firms.* Proceedings of the 8th Global Conference on Business & Economics, Florence, Italy.

Zhang, Q., Wei, X. and Xu, J. (2003). Global exponential stability of Hopfield neural networks with continuously distributed delays. *Physics Letters A*, *315*(6), pp. 431-436.

Wang, C. H., Liu, J. F., Hong, T. P. and Tseng, S. S. (1999). A fuzzy inductive learning strategy for modular rules. *Fuzzy Sets and Systems*, *103*(1), pp. 91-105.

Wilson, R. L. and Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, *11*(5), pp. 545-557.