

**Tomasz Pisula, Grzegorz Mentel,  
Jacek Brożyna**

---

**Non-statistical methods of analysing  
of bankruptcy risk**

---

Folia Oeconomica Stetinensia 15(23)/1, 7-21

---

2015

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

## NON-STATISTICAL METHODS OF ANALYSING OF BANKRUPTCY RISK

---

Tomasz Pisula, Ph.D.<sup>1</sup>  
Grzegorz Mentel, Ph.D.<sup>2</sup>  
Jacek Brożyna, Ph.D.<sup>3</sup>

*Department of Quantitative Methods  
Faculty of Management  
Rzeszów University of Technology  
Powstańców Warszawy 8, 35-959 Rzeszów, Poland*

<sup>1</sup> e-mail: [tpisula@prz.edu.pl](mailto:tpisula@prz.edu.pl)

<sup>2</sup> e-mail: [gmentel@prz.edu.pl](mailto:gmentel@prz.edu.pl)

<sup>3</sup> e-mail: [jacek.brozyna@prz.edu.pl](mailto:jacek.brozyna@prz.edu.pl)

Received 7 October 2014, Accepted 27 April 2015

---

### Abstract

The article focuses on assessing the effectiveness of a non-statistical approach to bankruptcy modelling in enterprises operating in the logistics sector. In order to describe the issue more comprehensively, the aforementioned prediction of the possible negative results of business operations was carried out for companies functioning in the Polish region of Podkarpacie, and in Slovakia. The bankruptcy predictors selected for the assessment of companies operating in the logistics sector included 28 financial indicators characterizing these enterprises in terms of their financial standing and management effectiveness. The purpose of the study was to identify factors (models) describing the bankruptcy risk in enterprises in the context of their forecasting effectiveness in a one-year and two-year time horizon. In order to assess their practical applicability the models were carefully analysed and validated. The usefulness of the models was assessed in terms of their classification properties, and the capacity to accurately identify enterprises at risk of bankruptcy and healthy companies as well as proper calibration of the models to the data from training sample sets.

**Keywords:** forecast, modelling, risk, bankruptcy.

**JEL classification:** C01, C45, C53.

## **Introduction**

There is a wide range of applications for models designed for predicting a company's bankruptcy or survival. Yet, they are mainly needed for formulating an opinion whether or not an enterprise will continue its operations in the following year. Such a model can be used to verify the financial condition of a customer provided with a trade credit, or the financial standing of a company's partner, for instance in the process of a merger.

Bankruptcy forecasting models are particularly interesting for banks due to the necessity to verify the financial standing of an entity requesting a loan and to monitor loan repayment. On the other hand financial institutions operating within the capital market are interested in the condition of enterprises whose securities they purchase.

Even this short overview of economic entities which potentially may apply, and indeed do frequently use bankruptcy forecasting models, shows that such issues and the need to develop the models are of significance for the world of finance.

Given the above, this study aims at assessing the effectiveness of a non-statistical approach to analysing the risk of bankruptcy.

### **1. Review of selected literature investigating bankruptcy risk in business enterprises**

Non-statistical methods, most frequently used for classifying companies at risk of bankruptcy, are based on neural networks. Tam and Kiang (1992) compared the quality of bankrupt companies' classification by means of LDA models, logistic regressions, and k Nearest Neighbour with neural network models. The neural network models investigated by them achieved the best classification statistics for a one-year period forecast horizon. On the other hand in the case of data for 2 years before bankruptcy the LDA model was found to be most effective. Fletcher and Goss (1993) used neural network models to predict a company's bankruptcy as compared with the logit model. They applied the V-fold cross validation to choose the best model and they used three independent indicators (data related to 33 firms) as potential predictors of bankruptcy in the created and trained neural networks. The classification success rate of the estimated neural network model amounted to 82%, and for the alternative logit model it was only 77%.

Atiya (2001) published a review of literature discussing neural network applications for corporate bankruptcy prediction. He also estimated a neural bankruptcy prediction model where the investigated input variables were not based on primary financial indicators but on certain novel indicators which significantly improved the classification accuracy. The statistics for

classification accuracy achieved for his model amounted to 85% – for the training sample set and 82% – for the testing sample set.

Serrano-Cinca (1996) compared classification accuracy achieved by Kohonen's neural network (SOM), traditional BP NN type models and the LDA model. The classical financial ratios proposed by Altman were used as predictors. The network had the following architecture: 4 input nodes and 144 output nodes (organized in a  $12 \times 12$  square grid) to comprise 74 patterns for the input dataset. He proposed two hybrid neural networks combining either the LDA model with SOM network or classical BP NN with SOM network.

Kaski et al. (2001) introduced a new method of identifying nonlinear dependencies between bankruptcies and financial ratios used as bankruptcy predictors. They investigated the use of network models based on the Euclidean metric SOM-E and Fisher metric SOM-F for identifying bankruptcies in small and medium enterprises. They demonstrated that models based on SOM-F are more successful than SOM-E.

A bankruptcy risk analysis based upon neural networks models is applied to companies from all over the world, e.g. Tseng and Hu (2010) use four techniques (the logit model, the quadratic interval logit model, back propagation multi-layer perceptron and the radial basis function network) to predict bankrupt and non-bankrupt firms in England. Lee and Choi (2013) using the back-propagation neural network and multivariate discriminant analysis presents a multi-industry investigation of the bankruptcy of Korean companies. Fedorowa, Gilenko and Dovzhenko (2013) in their paper apply combinations of modern learning algorithms to identify the most effective approach to bankruptcy prediction for Russian manufacturing companies. Korol (2012) in his paper he compares the effectiveness of twelve models for forecasting the bankruptcy risk of Stock Exchange companies from Central Europe and Latin America.

More information about the use of various types of artificial neural networks for predicting corporate bankruptcy can also be found in other studies (Lee et al., 2005; Lam, 2004; Leshno, Spector, 1996; Wilson, Sharda, 1994; Kiviluoto, 1998).

## **2. Characteristics of financial ratios and research samples used in bankruptcy forecasting in logistics companies**

Data concerning the bankruptcies of Polish companies were obtained from *Ogólnopolski Informator Upadłościowy* (Polish Bankruptcy Information), of EMIS (Emerging Markets Information Service) database.

The bankruptcy predictors selected for the assessment of companies operating in the logistics sector included 28 financial indicators characterizing these enterprises in terms of their financial standing and management effectiveness. The indicators were divided into five groups: liquidity ratios, profitability indicators, indebtedness and financial leverage ratios, indicators of operational effectiveness (efficiency), and the remaining indicators describing the structure of corporate capital and assets.

Statistical data related to the financial indicators for the Polish companies were based on their financial statements. The financial indicators for the enterprises from Slovakia were derived from Company Screener, EMIS system database related to business companies.

Research samples were specified taking into account the acquired statistical data. The dependent variable was the qualitative-dichotomous variable  $Y$  defining whether a given enterprise had declared bankruptcy ( $Y = 1$  – bankrupt), or it did not face the risk of bankruptcy ( $Y = 0$  – non-bankrupt). The set of input variables (bankruptcy predictors) consisted of 28 financial indicators described above.

Two research samples were created. The first one included bankrupt companies operating in the logistics sector, and their healthy counterparts, with available statistical data for one year before bankruptcy (one-year forecast horizon). The second research sample consisted of bankrupt and healthy companies with available statistical data illustrating a period of two years before bankruptcy (two-year forecast horizon). Within the research groups, each bankrupt firm was paired with a matching company which was not at risk of bankruptcy. The selection of healthy enterprises was preceded with in-depth analysis of indicators, and only those logistics companies were selected whose indicators suggested a good financial condition and capacity to pay off their liabilities.

The research sample representing the data for one year before bankruptcy consisted of: 33 – bankrupt firms and 33 – healthy companies (statistical data for the period of one year before bankruptcy were available only for this number of enterprises); data for two years before the time of bankruptcy were represented by 57 – healthy firms and 57 – bankrupt companies. The research samples were randomly divided into two groups: the training sample set, which was taken into account while estimating the parameters for prognostic models and the testing sample set verifying the effectiveness of the accuracy of the classification. The training sample set for a one-year forecast horizon consisted of 47 enterprises (23 bankrupts, and 24 non-bankrupts), and the testing sample set consisted of 19 companies (10 bankrupts and 9 non-bankrupts). On the other hand for the two-year forecast horizon the testing sample set comprised

86 logistics firms (43 bankrupts and 43 non-bankrupts), and the testing sample set consisted of 28 enterprises (14 bankrupts and 14 non-bankrupts).

In order to assess the impact of selected explanatory variables on the response variable identifying corporate bankruptcy, the predictors were subjected to a ranking analysis. This is a very important issue, because it is necessary to select only those predictors which have the best prognostic capacities for separating, or distinguishing bankrupt companies from healthy companies. An effectively performed ranking of predictors, relative to their classification strength, makes it possible to apply the following factors: Information Value (IV), Gini and Cramer's V.

### 3. Characteristics of the models used for bankruptcy prediction

The present study, designed to predict the risk of bankruptcy in firms operating in the logistics sector, used the following non-statistical models of classifying bankruptcy: models of neural networks based on multilayer perceptron's (MLP) and the support vector method (SVM).

#### Artificial neural networks – MLP

Artificial neural networks rank among the most frequently used techniques for solving problems connected with the accurate classification of companies at risk of bankruptcy.

There are many classifications of neural networks determining their structure and shape. One of the criteria is the number of layers occurring in the neural network. As a consequence the following types of networks are distinguished: single-layer – with only one layer of neurons, which is both input and output layer; two-layer – with clearly distinguished the input and output layers of the network; multi-layer – with intermediate layers containing hidden neurons.

The second important criterion of the classification is the type of connections between neurons occurring in the artificial neural network. Based on this criterion, neural networks are divided into: feed forward, recurrent, cellular as well as networks with radial basis functions (RBF).

The process of training (teaching) neural networks employed the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm, applying the Quasi-Newton method and two variants of error function, i.e. the sum of squares  $ERR_{SoS}$  and cross entropy (CE).

Table 1 presents a summary of the findings related to the training of four of the most successful networks (one per each variant of input variables and the time horizon assumed for the forecast).

Table 1. The most successful models of neural networks for various options of input variables as well as a one-year and two-year bankruptcy forecast horizon

Type of network (Network ID)	Training error function	Hidden neuron activation function	Output neuron activation function	Classification accuracy rate training sample set [%]	Classification accuracy rate testing sample set [%]
Forecast horizon – 1 year to bankruptcy					
Variant 1: input variables: $X_1$ – $X_{28}$ excluding: $X_{23}$ i $X_{28}$					
MLP 26-8-2	Cross entropy $ERR_{SoS}$	Linear	Linear	91.5	94.7
Variant 2: input variables: $X_1, X_2, X_6, X_{11}, X_{18}, X_{27}$					
MLP 6-3-2	Cross entropy $ERR_{CE}$	Tanh	Softmax	87.2	89.5
Forecast horizon – 2 years to bankruptcy					
Variant 1: input variables: $X_1$ – $X_{28}$ excluding: $X_{12}, X_{19}, X_{22}, X_{23}, X_{24}, X_{28}$					
MLP 22-17-2	Sum of squares $ERR_{CE}$	Logistic	Softmax	86.1	92.9
Variant 2: input variables: $X_2, X_5, X_{13}, X_{26}$					
MLP 4-8-2	Cross entropy $ERR_{SoS}$	Exponential	Tanh	73.3	82.1

Source: own study.

### Support vector method – SVM

The method based on Support Vector Machines (SVM) (Yu et al., 2008) was first proposed by Vapnik (1995). Generally the method involves the non-linear mapping of input objects (input data vectors) belonging to various categories (e.g. bankrupts and non-bankrupts) onto a certain multidimensional hyperspace, using specific kernel functions, most frequently:

- linear kernel:  $K(x_i, x_j) = x_i^T \cdot x_j$ ,
- polynomial kernel:  $K(x_i, x_j) = (\gamma x_i^T x_j + c)^d$ ,
- kernel function with the so-called radial basis functions (RBF):  $K(x_i, x_j) = e^{-\gamma |x_i - x_j|^2}$ ,
- sigmoid kernel:  $K(x_i, x_j) = \text{Tanh}(\gamma x_i^T x_j + c)$ .

Bankruptcy related classification of logistics companies was performed with a module of *Statistica* package – Other methods of machine learning – Support Vector Machines SVM. *Statistica* software uses two different types of SVM models, depending on the adopted error function: *C-SVM* (type 1) or *ni-SVM* (type 2). The method based on the classification SVM type 1 is used for finding the most effective solution for the following problem of conditional optimization:

$$\min_{w,b,\xi} \left\{ \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i \right\} \tag{1}$$

subject to the constraints:

$$y_i (w^T \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, n,$$

where:  $C$  – model parameter called capacity,  $w$  – vector of weighting coefficients for input (independent) variables,  $y_i \in \{-1,1\}$  are class labels of the investigated objects,  $x_i$  – independent variables,  $\varphi$  – kernel function transforming independent variables to a new feature space,  $\xi_i$  – parameters for handling the so-called overlapping cases,  $b$  – a constant in the model, and  $i = 1, \dots, n$  numbers the indexes of  $n$  – training cases.

The capacity parameter  $C$  is of critical importance for the model, as it significantly impacts error value, therefore it must be chosen with care to avoid an over fitting of the model to the training data.

Table 2. The most successful SVM classification models for the prediction of bankruptcy in logistics companies for a one-year and two-year forecast horizon

Type of SVM model	Optimal values of parameters C/Ni	Number of support vectors	The applied kernel and its parameters	Classification accuracy rate training sample set	Classification accuracy rate testing sample
Forecast horizon – 1 year to bankruptcy					
Type 1 C – SVM	C = 3	25: (19 bound vectors) vectors per class: 12(+) B, 13(-) NB	linear	93.6	89.5
Type 2 Ni – SVM	Ni = 0.5	27: (20 bound vectors) vectors per class: 13(+) B, 14(-) NB	RBF gamma = 0.038	93.6	84.2
Forecast horizon – 2 years to bankruptcy					
Type 1 C – SVM	C = 9	75: (73 bound vectors) vectors per class: 37(+) B, 38(-) NB	Polynomial of 3rd degree gamma = 0.045; constant = 0	73.3	82.1
Type 2 Ni – SVM	Ni = 0.8	73: (66 z bound vectors) vectors per class: 36(+) B, 37(-) NB	RBF gamma=0.045	74.4	78.6

Source: own study.

Method based on classification SVM type 2 is applied to solve the problem of conditional optimization represented by:

$$\min_{w,b,\xi} \left\{ \frac{1}{2} w^T w - \nu p + \frac{1}{n} \sum_{i=1}^n \xi_i \right\} \quad (2)$$

subject to the constraints:

$$y_i (w^T \varphi(x_i) + b) \geq \rho - \xi_i, \xi_i \geq 0, \rho \geq 0, i = 1, \dots, n.$$

The parameter  $\nu \in (0,1)$  is interpreted both as the maximum rate of training errors and the minimum indicator of the number of applied support vectors in relation to the number of training cases.

The input variables adopted for the SVM models included 26 indicators for a one-year forecast horizon and 22 indicators for a two-year forecast horizon (the same as in the models of neural networks). The training algorithm for the SVM models employed two classes of models available in *Statistica*: *C*-SVM and *Ni*-SVM, as well as various forms of kernel function: linear, polynomial, RBF, and sigmoid. In order to select the most successfully trained model, just like other analyses, our study applied the *V*-fold cross validation, adopting a recurrent method of searching for the best values for capacity parameter *C*, within a grid from 1 to 20, with step 1, and for the parameter *Ni* within a grid in the range 0.1–0.9 with step 0.1.

#### 4. Validation of the estimated bankruptcy models

In order to assess their practical applicability the models were carefully analysed and validated. The usefulness of the models was assessed in terms of their classification properties, and the capacity to accurately identify enterprises at risk of bankruptcy and healthy companies as well as the proper calibration of the models to the data from the training sample sets.

The basic tool designed for assessing classification models for their effectiveness and accuracy are matrixes of correct classification (see Table 3). The *TN* (True Negative) number in the table represents the number of healthy companies correctly identified by the model. Similarly, the *TP* (True Positive) number represents the number of bankrupt companies correctly classified by the model. If healthy firms are classified by the model as bankrupts, this classification error is called Type I error, and *FP* (False Positive) represents the number of such erroneous classifications. Type II error, presenting a significantly greater drawback, occurs when the model incorrectly identifies bankrupts and classifies them into the group of those not at risk of bankruptcy, and *FN* represents the number of such classification errors.

Table 3. Matrix of correct classifications for a bankruptcy forecasting model

Actual status of company	Predicted status of company	
	NB	B
NB	TN (True Negative)	FP (False Positive) Type I error
B	FN (False Negative) Type II error	TP (True Positive)

Source: own study.

The process of validating models designed to classify companies at risk of bankruptcy most frequently applies the following measures: Information Value (IV) factor, Gini coefficient and Divergence index, as well as the Kolmogorov-Smirnov test and Hosmer-Lemeshow test.

A model’s information impact, or the IV factor, represents the model’s capacity to divide the distribution of results into the populations of bankrupt and non-bankrupt. The factor is calculated after the objects constituting the sample have been sorted in descending order taking into account the probability values, based on the model, indicating that the objects belong to the negative class (probability of the company’s bankruptcy).

The Gini coefficient is used to assess the advantage of the estimated model over a random model – for arbitrarily taken decisions. Like in the previous case, the objects in the research samples must first be arranged in descending order taking into account the probability of bankruptcy.

The value of the Kolmogorov-Smirnov test (KS statistics) specifies the maximum distance between distribution functions for conditional probability distributions for bankruptcy in the population of healthy companies (NB) and bankrupts (B) and the value is calculated in accordance with the formula (Thomas, 2009):

$$KS = \max_x |F(x | B) - F(x | NB)| \tag{3}$$

Divergence is a measure of distance between the examined conditional probability distributions for bankruptcy in both groups of enterprises and is represented by the formula (Thomas, 2009):

$$D = \frac{1}{2} \left( \frac{1}{\sigma_G^2} + \frac{1}{\sigma_B^2} \right) (\mu_G - \mu_B)^2 + \frac{(\sigma_G^2 - \sigma_B^2)^2}{2\sigma_G^2 \sigma_B^2} \tag{4}$$

where:  $\mu_{NB} = \sum_x x \times f(x | NB)$  – average probability value for bankruptcy in the population of healthy companies (NB),  $\mu_B = \sum_x x \times f(x | B)$  – average probability value for bankruptcy in the

population of bankrupts (B),  $\sigma_{NB}^2 = \sum (x - \mu_{NB})^2 \times f(x | NB)$ ,  $\sigma_B^2 = \sum (x - \mu_B)^2 \times f(x | B)$  – variance in distribution of probability of bankruptcy in populations of healthy and bankrupt companies, respectively,  $f(x | NB)$ ,  $f(x | B)$  – percentage of healthy and bankrupt companies for a given category of probability of bankruptcy.

It is assumed that divergence should take on the values above 0.5, so that the relevant distributions are sufficiently far apart from each other and the investigated model has acceptable ability to correctly separate bankrupts from enterprises which are not at risk of bankruptcy.

The Hosmer-Lemeshow statistic is based on a chi-squared test and is calculated in accordance with the formula (Thomas, 2009):

$$HL = \sum_{i=1}^N \frac{(n_i p_i - NB_i)^2}{n_i p_i (1 - p_i)} \quad (5)$$

where:  $p_i$  – average probability of belonging to the non-bankrupt class for a given assumed rating  $i$ -category,  $NB_i$  – number of healthy enterprises in a given rating category,  $N$  – a defined number of rating categories into which the range of variations in the probability of bankruptcy has been divided. The Hosmer-Lemeshow statistic has  $\chi^2$  distribution with  $df = N - 2$  degrees of freedom. Higher H-L statistic values correspond with the better ability of a model to differentiate distributions in both populations (B and NB) and more effective classification properties of the model.

The previously characterized measures assess the discriminative efficiency of models. The assessments of discriminative efficiency and accuracy of model calibration to training and testing data were performed using the Brier Score as well as model reliability rate LL (Likelihood of the model).

Calculation of the Brier Score (BS) is based on the formula (Löffler and Posch, 2007):

$$BS = \frac{1}{n} \sum_{i=1}^n (d_i - PD_i)^2 \quad (6)$$

where:  $n$  – number of observations in a sample,  $d_i$  – zero-one variable, assuming the value of 1 when an enterprise is considered to be bankrupt, and the value of 0 in the other situation,  $PD_i$  – probability of bankruptcy estimated using the model.

Lower values of the Brier Score correspond to the better calibration of the model to the data and potentially its higher prognostic accuracy.

Model reliability rate  $LL(\text{model})$  is defined by the formula (Prusak, 2005):

$$LL = \prod_{i=1}^n P(Y_i | X_i) = \prod_{i=1}^n PD_i(X_i)^{Y_i} \times (1 - PD_i(X_i))^{1-Y_i} \quad (7)$$

where:  $n$  – number of observations,  $PD_i(X_i)$  – estimated probability of bankruptcy for the values of input (independent) variables defined in the model,  $Y_i$  – zero-one variable, dependent, representing ( $Y = 1$  – bankrupts) and ( $Y = 0$  – non-bankrupts).

Higher reliability values achieved by the classification model for training a sample set correspond to better calibration based on input data. High values of the reliability rate for the testing sample set should indicate good classification capacities of the model for new, unknown cases, as well.

Tables 4 and 5 present validation statistics for all of the investigated models of forecasting bankruptcy in logistics companies.

Table 4. Validation parameters of the assessed models for a one-year forecast horizon

Model	Eff <sub>1</sub> NB	Eff <sub>2</sub> B	IV	K-S	Gini	Divergence	H-L	AUROC	Brier Score	LL(model)
MLP 26-8-2 Network	training sample set									
	92%	91%	4.0	0.83	0.89	5.3	17.9	0.95	0.152	$1.3 \times 10^{-10}$
	testing sample set									
	89%	100%	2.4	0.89	0.82	2.9	48.2	0.91	0.162	$2.1 \times 10^{-5}$
MLP 6-3-2 Network	training sample set									
	92%	83%	2.6	0.75	0.86	4.2	13.1	0.93	0.135	$1.5 \times 10^{-9}$
	testing sample set									
	89%	90%	2.8	0.79	0.91	7.0	3.6	0.96	0.111	$8.6 \times 10^{-4}$
C-SVM	training sample set									
	92%	96%	5.7	0.96	0.93	14.3	1.3	0.96	0.064	-(0)
	testing sample set									
	89%	90%	2.8	0.79	0.6	6.6	-(0)	0.8	0.105	-(0)
Ni-SVM	training sample set									
	92%	96%	5.7	0.96	0.93	14.3	1.3	0.96	0.064	-(0)
	testing sample set									
	78%	90%	2.3	0.68	0.49	4.1	-(0)	0.74	0.158	-(0)

Source: own study.

Table 5. Validation parameters of the assessed models for a two-year forecast horizon

Model	Eff <sub>1</sub> NB	Eff <sub>2</sub> B	IV	K-S	Gini	Divergence	H-L	AUROC	Brier Score	LL (model)
MLP 22-17-2 Network	training sample set									
	88%	84%	3.6	0.74	0.87	5.9	9.6	0.94	0.103	$4.8 \times 10^{-13}$
	testing sample set									
	100%	86%	3.7	0.86	0.92	10.2	2.4	0.96	0.087	$3.7 \times 10^{-4}$
MLP 4-8-2 Network	training sample set									
	67%	79%	2.4	0.56	0.70	2.4	14.8	0.85	0.184	$1.8 \times 10^{-21}$
	testing sample set									
	100%	86%	3.7	0.86	0.94	6.8	11.4	0.97	0.167	$4.6 \times 10^{-7}$
C-SVM	training sample set									
	65%	81%	2.3	0.58	0.57	1.3	41.1	0.78	0.267	-(0)
	testing sample set									
	79%	86%	2.2	0.78	0.82	2.9	1.5	0.91	0.179	-(0)
Ni-SVM	training sample set									
	72%	77%	1.6	0.58	0.57	1.3	-(0)	0.78	0.256	-(0)
	testing sample set									
	79%	79%	2.2	0.71	0.70	1.9	-(0)	0.85	0.214	-(0)

Source: own study.

## 5. Bankruptcy forecasting

In order to predict potential bankruptcy by means of the investigated models it is a good idea to distinguish two groups, just like it was done earlier. One of these comprises forecasts estimated by the models for the period of one year before bankruptcy, and the other contains predictions made by the same set of models taking into account data for the period of two years before bankruptcy (Table 6). In the first case we can look at a sample of 125 “healthy” logistics enterprises, including 82 (65.6%) companies from Podkarpacie and 43 (34.4%) companies from Slovakia. The second variant comprises the total of 104 companies, including 61 (58.7%) firms from Podkarpacie and 43 (41.3%) from Slovakia.

Dividing the above probabilities of bankruptcy into two categories, up to the value of 0.5 and over 0.5, i.e. non-bankrupt or bankrupt, we must emphasize that within the group of logistics companies operating both in Podkarpacie and in Slovakia, there are no negative indications for the entire sector. The most negative prospects for the entire groups of enterprises are suggested by the models of MLP artificial neural networks since the scores exceed the value of 0.4. An exception in this case is model *MLP 22-17-2* which was assessed taking into account data for two years before bankruptcy.

Table 6. Mean values predicted by the investigated models for Podkarpacie and Slovakia

	Podkarpacie	Slovakia
	estimates based on data for one year before bankruptcy	
Model <i>MLP 26-8-2</i>	0.456908	0.473894
Model <i>MLP 6-3-2</i>	0.410757	0.388252
Model <i>C-SVM</i>	0.329268	0.372093
Model <i>Ni-SVM</i>	0.353659	0.372093
Mean one-year forecast	0.387648	0.401583
Mean two-year forecast	0.529848	0.529941
Mean three-year forecast	0.614493	0.601913
	estimates based on data for two years before bankruptcy	
Model <i>MLP 4-8-2</i>	0.407781	0.404069
Model <i>MLP 22-17-2</i>	0.298116	0.193183
Model <i>C-SVM</i>	0.360656	0.348837
Model <i>Ni-SVM</i>	0.295082	0.232558
Mean one-year forecast	0.340409	0.294662
Mean two-year forecast	0.473253	0.435436
Mean three-year forecast	0.551039	0.521934

Source: own study.

An assessment of the mean forecasts for the various periods from one to three years leads to a general conclusion that it is only two and three year indications which suggest possible bankruptcy. It was only in the case of three-year forecasts based on data related to two years before bankruptcy that the value of probability exceeded the threshold value of 0.5.

It is interesting to compare scores signalling a risk of bankruptcy, shown by the specific models (Table 7). Here we can make a collective assessment of the diversity of indications.

Table 7. Indications of risk for the specific models

	Bankruptcy	
	yes	no
	estimates based on data for one year before bankruptcy	
Model <i>C-SVM</i>	43 (34.4%)	82 (65.6%)
Model <i>Ni-SVM</i>	45 (36.0%)	80 (64.0%)
Model <i>MLP 26-8-2</i>	53 (42.4%)	72 (57.6%)
Model <i>MLP 6-3-2</i>	43 (34.4%)	82 (65.6%)
	estimates based on data for two years before bankruptcy	
Model <i>C-SVM</i>	37 (35.6%)	67 (64.4%)
Model <i>Ni-SVM</i>	28 (26.9%)	76 (73.1%)
Model <i>MLP 4-8-2</i>	14 (13.5%)	90 (86.5%)
Model <i>MLP 22-17-2</i>	22 (21.2%)	82 (78.8%)

Source: own study.

It seems that in the case of estimates based on data for one year before bankruptcy the differences in the indications are relatively small. Some differences can only be noticed in the variant where the calculations were based on data for two years before bankruptcy.

## Conclusions

The empirical verification of classification accuracy for the specific groups of methods of a non-statistical analysis of bankruptcy, from the point of view of their effectiveness, showed that these methods are characterized with high capacity for predicting bankruptcy. The presented approaches enable relatively easy assessment of a specific group of entities from the point of view of bankruptcy risk. It is particularly important to regularly estimate the probability of possible negative effects of operations in consecutive years and to pay attention to the occurring tendencies and changes related to these aspects. It is necessary to emphasize the fact that findings acquired with the help of the methods described above should not be treated as absolutely conclusive. This is because adequately applied methods can only provide assistance in evaluating the actual financial standing of the investigated companies. Therefore, they constitute a signal for taking reasonable decisions and not a mechanism designed as a definitive solution to the problem.

## References

---

- Atiya, A.F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks*, 12 (4): 929–935.
- Choi, W.S. & Lee, S. (2013). A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, 40 (8): 2941–2946.
- Fedorova, E., Gilenko, E. & Dovzhenko, S. (2013). Bankruptcy prediction for Russian companies: Application of combined classifiers. *Expert Systems with Applications*, 40: 7285–7293.
- Fletcher, D. & Goss, E. (1993). Application forecasting with neural networks an application using bankruptcy data. *Information and Management*, 24: 159–167.
- Kaski, S., Sinkkonen, J. & Peltonen, J. (2001). Bankruptcy analysis with self-organizing maps in learning metrics. *IEEE Transactions on Neural Networks*, 12 (4): 936–947.

- Kiviluoto, K. (1998). Predicting bankruptcies with self organizing map. *Neurocomputing*, 21: 191–201.
- Korol, T. (2012). Early warning models against bankruptcy risk for Central European and Latin American enterprises. *Economic Modelling*, 31: 22–30.
- Lam, M. (2004). Neural networks techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*, 37: 567–581.
- 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*, 29: 1–16.
- Leshno, M. & Spector, Y. (1996). Neural network prediction analysis: The bankruptcy case. *Neurocomputing*, 10: 125–147.
- Löffler, G. & Posch, P.N. (2007). *Credit risk modeling using Excel and VBA*. Chichester, West Sussex: Wiley (pp. 156).
- Prusak, B. (2005). *Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstw*. Warszawa: Difin (pp. 50).
- Serrano-Cinca, C. (1996). Self organizing neural networks for financial diagnosis. *Decision Support Systems*, 17: 227–238.
- Tam, K.Y. & Kiang, M. (1992). Predicting bank failures: A neural network approach. *Decision Sciences*, 23: 926–947.
- Thomas, L.C. (2009). *Consumer credit models. Pricing, Profit and Portfolios*. Oxford: Oxford University Press (pp. 111).
- Tseng, F.M. & Hu, Y.C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37 (3): 1846–1853.
- Wilson, R.L. & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11: 545–557.
- Witkowska, D. (2002). *Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe*. Warszawa: C.H. Beck (pp. 86–87).
- Yu, L., Wang, S., Lai, K.K. & Zhou, L. (2008). *Bio-Inspired Credit Risk Analysis. Computational Intelligence with Support Vector Machines*. Berlin Heidelberg: Springer-Verlag (pp. 14–15).