Tumpach, Miloš; Surovičová, Adriana; Juhászová, Zuzana et al.

Article

Prediction of the bankruptcy of Slovak companies using neural networks with SMOTE

Ekonomický časopis

Provided in Cooperation with:
Slovak Academy of Sciences, Bratislava

Reference: Tumpach, Miloš/Surovičová, Adriana et. al. (2020). Prediction of the bankruptcy of Slovak companies using neural networks with SMOTE. In: Ekonomický časopis 68 (10), S. 1021 - 1039.
https://www.sav.sk/journals/uploads/1215135210%2020%20Tumpach-aut+SR.pdf.
doi:10.31577/ekoncas.2020.10.03.

Prediction of the Bankruptcy of Slovak Companies Using Neural Networks with SMOTE[1]

*Miloš TUMPACH – Adriana SUROVIČOVÁ – Zuzana JUHÁSZOVÁ – Anton MARCI – Zuzana KUBAŠČÍKOVÁ**

Abstract

Although the bankruptcy prediction models can be a stabilizing element on both macro and microeconomic levels, they are rather a domain of academic research than an instrument, widely applied in a business practice. It is especially true if the models are reflecting the conditions of countries of their origin, rather than countries of their intended uses. Besides, few of the models contain inherent flaws, including the absence of a methodical approach addressing this problem of the severely imbalanced representation of bankrupt companies in financial datasets. The article is focused on the use of oversampling with SMOTE (Synthetic Minority Oversampling Technique) algorithm under the condition of extremely imbalanced data sets of Slovak companies. While the model does not provide a single answer in many (if not most) of the situations, it still could be used for the selection of companies for which the more detailed (and expensive) analysis is not required.

Keywords: *artificial neural networks, bankruptcy model, oversampling, SMOTE, imbalanced data*

JEL Classification: M41, C45, G33

DOI: https://doi.org/10.31577/ekoncas.2020.10.03

* Miloš TUMPACH – Adriana SUROVIČOVÁ – Zuzana JUHÁSZOVÁ – Anton MARCI – Zuzana KUBAŠČÍKOVÁ, University of Economics in Bratislava, Faculty of Economic Informatics, Department of Accounting and Auditing, Dolnozemská cesta 1, 852 35 Bratislava, Slovak Republic; e-mail: milos.tumpach@euba.sk; adriana.surovicova@euba.sk; zuzana.juhaszova@euba.sk; anton.marci@euba.sk; zuzana.kubascikova@euba.sk

[1] This work was supported by the Slovak Research and Development Agency under the contract No. APW-16-0602 *Enhancement of the Relevance of the Accounting Data in the SR – from Expenses to Value.*

Introduction

The instability of the economic environment causes problems and increased costs not only for businesses but also for other parties that are directly or indirectly dependent on their activities or existence. Information about whether a business is exposed to the risk of bankruptcy is therefore important not only for its owners, prospective investors, or management but also for its employees, creditors, suppliers, government, and other stakeholders. Prediction of corporate bankruptcies is a relatively recent area of scientific research, as it required both the development of methods of advanced statistical analysis (with special regard to the businesses) and the availability of sufficient data as its precondition. Fitzpatrick (1932), Smith and Winakor (1935), and Merwin (1942) were among the pioneers of this research domain, paying interest primarily on the relation between the bankruptcy of companies and the single accounting data as the predictors of the bankruptcies. From another point of view, Chudson (1945) and Jackendoff (1962) aimed at the conditions that are essential for continuing business (going concern), rather than at the prediction of bankruptcy. An important milestone in the further development of this area of research was the paper of Beaver (1966), which proposed one-dimensional discriminant analysis for the prediction of corporate bankruptcies. Altman (1968) followed these findings with a well-known bankruptcy prediction study, using multivariate discriminant analysis (MDA). In the early 1970s, use of the multidimensional discriminant analysis became the principal approach for the construction of bankruptcy models (e.g. Deakin, 1972; Blum, 1974), followed by linear discriminant analysis (see Edmister, 1972; Altman 1973). The second half of the 1970s and the early 1980s marks the introduction of logistic regression analysis for bankruptcy prediction, with probit transformation function as proposed by Hanweck (1977), Martin (1977), or Ohlson (1980). Some authors (e.g. Mensah, 1983) developed models combining logistic regression and multidimensional discriminant analysis. At the beginning of the 1990s, Odom and Sharda (1990), Bell et al. (1990), Koster, Sondak and Bourbia (1990), Cadden, 1991) and many others proposed the concept of artificial neural networks (ANN) based bankruptcy prediction models for various industries.

Because the ANN-based bankruptcy models are using supervised training (rather than truly analytical approach), they are also subjects of a smaller number of assumptions and restrictions than other approaches (Coats and Fant, 1993) and thus less prone to fundamental errors. On the other hand, it is the very same nature of the training that makes the models inefficient if the processed data is severely imbalanced (i.e. there are only a few bankrupt companies present) or the population of the data for companies is extremely low. This could by one of the reasons, why the use of ANN models has not been significantly developed in

Slovakia, with a few notable exceptions (for example, Boďa, 2009; Cút, 2013; Mihalovič, 2016; Valášková et al, 2018; Zoričák et al. (2020). The claim is true despite the introduction of the national register of financial statements in 2013 by Slovak law, making financial statements for almost any company in the country publicly available for the external users. As of the end of October 2020, the register contained 178,246 individual financial statements for the year 2019 and 220,734 for the year 2018 (the decrease in the number of fillings in 2019 is deemed to be a result of postponements of fillings caused by COVID). It could be assumed, that with such an amount of data, the lack of data is not a problem anymore. Consequently, the disproportional representations of bankrupt companies could be considered as one of the remaining obstacles.

1. Review of Literature

Quite surprisingly, the idea of ANN-based bankruptcy prediction models appeared almost immediately with the emergence of the very concept of artificial neural networks. To provide an assessment of the efficiency, the first research studies on ANN-based models were rather comparative. Odom and Sharda (1990) experimentally confirmed the advantages of such models over rather traditional MDA, thus demonstrating the potential of artificial neural networks. The discourse, however, continued in the following periods. Tam and Kiang (1992) concluded, that while multi-layered neural networks were more suitable for the prediction of the bankruptcy one year before its declaration, in earlier periods they were outperformed by logistic regression models with the probit transformation function. These findings were, to some extent, challenged by Salchenberger, Cinar and Lash (1992) and Zhang et al. (1999) which claim, that ANN-based models using the backpropagation algorithm (BPNN) had higher reliability in predicting bankruptcy than logistic regression models. Many other studies confirm a better predictive ability of the future financial situation of the company using BPNN over the classical statistical methods (e.g. Fletcher and Goss, 1993; Boritz and Kennedy, 1995; Back, Laitinen and Sere, 1996; Lee, Han and Kwon, 1996; Carlos, 1996; Leshno and Spector, 1996; Pendharkar, 2005; Liang and Wu, 2005; Wu, Liang and Yang, 2008; Rafiei, Manzari and Bostanian, 2011). Based on analysis carried out in the construction industry, manufacturing, and retail, Lee and Choi (2013) confirmed both the improved prediction power of ANN-based models over MDA models and the advantages of industry-specific models over the general models. In subsequent periods, there was a shift of paradigm of research from comparative analysis to the increase of the prediction power of ANN-based models. Lee, Han and Kwon (1996) created a hybrid model, using MDA for the selection of predictors and self-organizing maps for the binary classification. Other studies (including Pendharkar and Rodger, 2004; and Sai, Zhong and Qu, 2007) draw attention to the use of genetic algorithms for training and validating ANN-based bankruptcy prediction models. Genetic algorithms were also used for the optimization of the MLP topology in studies of Ignizio and Soltys (1996), Wallrafen, Protzel and Popp (1996), and Abdelwahed and Amir (2005). The idea of reduction of sampling risk led to the creation of the concept of ensembles of artificial neural networks, mainly using the bagging (Shin, Lee and Kilic, 2006) or boosting methods (West, Dellana and Qian, 2005). Tsai and Hung (2014) compared results of prediction models using ensembles of artificial neural networks, hybrid neural networks, and single artificial neural networks with hybrid neural networks achieving purportedly better results than ensembles, with single ANNs at the last place. Blanco-Oliver et al. (2015) suggested both the size of companies and non-financial variables as relevant predictors for such models.

Generally, the prediction power of ANN-based bankruptcy models is a product of three aspects: relevance of the predictors, appropriateness of the architecture of the artificial neural network(s) used, and degree of compliance of datasets with requirements of supervised training (e. g. using the same financial reporting framework to achieve comparability of financial data). Historically, the researchers used samples in which the proportions of bankrupt and non-bankrupt companies were at par (50:50), as an easy and verifiable approach. Hoverer, in Slovakia, corporate fillings of bankrupt companies accounted on average for less than 0.4 percent of total corporate filings, making the previous approach untenable. To cope with this problem, researchers aimed their interest towards the severely imbalanced datasets. Manski and Lerman (1977), Zmijewski (1984) and others, initiated discourse on this problem. Veganzones and Séverin (2018) stated that predictive power significantly deteriorates with the proportion of bankrupt companies falling below 20% of the population, though samples exceeding 4,000 bankrupt companies will have almost the same predictive power.

Naturally, there are two approaches to the reduction of disparity in datasets. With *udersampling,* the proportion of the bankrupt companies in the sample is increased by the removal of systematically or randomly selected data of non-bankrupt companies. As an alternative, (quasi) random samples with the over-representation of data from bankrupt companies are used in *bagging* (Breiman, 1996) and *boosting* (Barboza et al., 2017) techniques. Additionally, replacement of data for non-bankrupt companies with *cluster-based vectors* (Kim and Kang, 2010; Li and Sun, 2012; Sanchez-Lasheras, 2012; Kim et al., 2016; Le, 2018; Onan, 2019) could serve the same purpose. *Oversampling* adds either exact replicas or synthetic derivates of the real data to the original financial dataset.

According to Szolno (2016), an imbalanced population is a significant contributor to noise, bias, and variance of bankruptcy prediction models. Given the nature of the supervised learning, Le (2018) empirically documented a significant trade-off between the type I and type II errors for severely imbalanced datasets. In Slovakia, from 2014 till 2019, the financial statements of bankrupt companies accounted on average for less than 0.4% of publicly available financial statements of all companies, thus making the datasets combining bankrupt and non-bankrupt entities severely imbalanced. As a result, instead of developing a perfect model with a single and clear-cut classification of companies, we have decided to focus on a less ambitious, but more realistic goal. We aim to determine whether the ANN-model could be used for scoring corporations either as low-risk or companies for which further, more detailed, and expensive analyses are not required.

Per the recommendation of Zhou (2013) we have selected oversampling as a tool for the reduction of bias, variance, and noise of the bankruptcy model. Following the conclusions of Le et al. (2018), Hardle et al. (2019), Shrivastava, Jeyanthi, and Singh, S. (2020), Smiti et al. (2020), and Faris et al. (2020), we used the SMOTE (synthetic minority oversampling technique) as the algorithm for oversampling the bankruptcy related entries.

2. Description of the Research Methodology and the Datasets

For the bankruptcy prediction model we choose to developed an ANN model with standard multi-layered perceptron architecture, a tangent activation function and a backpropagation learning algorithm. Its prediction power in terms of values of *accuracy*, *precision*, *recall*, *specification, F-score, retention,* and *random pick* (see over) are determined by processing the validation set of data from, not included in the training. For the training and validation, we used 213,931 financial records from Slovak companies operating in the G section (wholesale and retail trade; repair of motor vehicles and motorcycles) of the SK NACE rev. 2 classifications of industries. Data is covering the period 2014 – 2019, and for training and validation, it is split with a ratio of 80:20 to two datasets (the description of the data is provided in a later part of the article).

There are three layers of our ANN-based model used for its training and its results: *input layer* (vector of predictors), *hidden layer* (vector of nodes), and (binary) *output layer*. The input layer contains 8 predictors, six of them (X1, X2, X3, X4, X5, and REVENUES) are financial, two of them (LEGAL_FORM, AGE) are categorical. Predictors X1 through X5 are parameters used in the IN05 model (Neumaier and Neumaierová, 2005; Neumaier and Neumaierová 2013), for the prediction of bankruptcy of Czech companies. We selected them both because of the similarity of the Czech and Slovak accounting data used for their computation and the purported high prediction power of the model. The predictors capture measures of profitability, indebtedness, liquidity, activity, and interest coverage in the following way:

x_1 = *total assets/total liabilities,*

x_2 = *EBIT/interest expense,*

x_3 = *EBIT/total assets,*

x_4 = *revenues/total assets,*

x_5 = *current assets/(current liabilities).*

To reflect the conditions which are not addressed in IN05 model, we added three additional predictors (REVENUES, AGE, LEGAL_FORMS). Predictor REVENUES serves as a control variable, being a proxy of the size of a company. We assume, all other things being the same, the larger companies bear a different risk of bankruptcy than small and medium-sized companies. Predictor AGE is a result of stratification of companies into four bins by their age (expressed in years) from the date of their foundation till the balance-sheet date. The intervals for the bins are: (0; 5] for the first one, (5; 10] for the second one, (10; 20] for the third one, and (20; 100] for the fourth one. We assume (ceteris paribus), that the risk of bankruptcy is lower for the older companies. Though this claim is not scientifically proved in our research, it is supported by the very fact that age evidence ability to survive. The last additional predictor, LEGAL_FORMS, is used to capture the differences arising both from different protection of creditors and different modes of financing in various legal forms of companies.

For each multilayer perception neural network, there are several hyperparameters that affect the performance of the prediction model. Apart from the number of layers (there are three in our case), other parameters include a number of nodes in a hidden layer, a number of learning cycles, a learning rate, and an activation function, to name just a few of them. While there is evidence of a positive correlation between the number of nodes and the prediction power of the bankruptcy model, it comes with a price. First, even a small increase of nodes could require significantly more time necessary for the training of the model. Second, the function of the prediction power with number of nodes as its parameter is a concave function for which a local maximum exists. Because there is no generally applicable rule for the determination of the number of nodes, we used scenarios with 100 and 200 nodes. Also, for our experiments, we used 1,000 and 2,000 learning cycles, and three variants of learning rate (0.001, 0.0001 and 0.0005). In addition, due to poor results of gaussian and min-max normalizers in our preliminary experiments, we used binning normalizers only.

For inference statistics of the prediction models, *accuracy*, *precision, recall, specificity,* and *F-score* are traditionally used as metrics for assessment of their prediction power. *Accuracy* measures a ratio between correctly predicted observations and the total number of observations. *Precision* measures a ratio between the total correctly predicted positive observations (i.e. the number of bankrupted companies labeled as bankrupted by the model) and the number of total observation, labeled as positives by the model (i.e. both true and false positives). *Recall* measures the ratio between the correctly predicted bankruptcies and the sum of both correctly predicted bankruptcies and false negatives. *Specificity* measures the ratio between the correctly identified non-bankrupt companies and the sum of all observed negatives (whether correctly identified or not). For formulas for precision, recall, accuracy, and F-score, see Table 1.

Table 1

Evaluation Metrics for Predictive Model

Accuracy	(true positives + true negatives)/total observations
Precision	true positives/(true positives + false positives)
Recall	true positives/(true positives + false negatives)
Specificity	true negatives/(true negatives + false positives)
F-score	2 × precision × recall / (precision + recall)

Source: Kaderová et al. (2020).

The ANN-based bankruptcy models are classifying companies in accordance with the associated probability of a bankruptcy computed for each company by the model. If the probability for any given company exceeds the predetermined threshold, the company is labeled as bankrupt, otherwise it is classified as non-bankrupt. Except for extreme cases, adjustment of the threshold has an impact on the proportion of true positives, false positives, true negatives, and false negatives, and the related measures of the prediction power. To provide a more dynamic overview, we decided to compute measures for various values of the threshold from 0.02 till 0.98, with a marginal increase of 0.02 (see Table 6, Table 7, and Figure 1), instead of rather traditional assumptions of the threshold being equal to 0.5. In our opinion, this dynamic approach better addresses the needs of the users and allow them to trade between false negatives and false positives in accordance with their risk strategy. *We aim to assess the merits of the ANN-based model for a binary classification of companies as either those for which there is a low risk of bankruptcy (and hence, no further analysis is required) or those for which the risk is above the accepted level.* For this purpose, we compute two additional measures. The first one, *retention* (see Table 6 and Table 7), is a ratio representing the proportion of the companies (of the total sample) that need to be further analyzed outside the ANN-based models. In other words, it represents a ratio between both true and false positives, and total observations. To control the numerator, we hold the number of false negatives at a maximum level of 1 (out 137 actually bankrupted companies), scenarios which results with higher number of negatives were not taken into consideration. Conversely, *1 – retention*, measures a decrease of further necessary analyses, as it measures the ratio between the true and negatives identified by the model (which need not to be analyzed) and the total number of observations. The second additional variable is the *random pick* (see Table 6 and Table 7) which express the probability of randomly picking the same number of true positives as the ANN-based model, with the same sizes of the samples and the total population. For emulation of hypergeometric distribution, we used spreadsheet function *hypgeom.dist*.

To train and validate the prediction model, we used datasets from Finstat, representing financial statements of Slovak companies for the years 2014 – 2019, with additional information about the declared bankruptcies, and years of foundation of respective companies. To regularize companies´ data, we removed the following entries from the datasets:

- records related to interim financial statements,
- records from financial statements prepared in accordance with the international financial reporting standards (because of the lack of comparability),
- deemed annual statements covering periods exceeding the interval of 364 – 366 days,
- records representing non-profit entities (e.g. foundations, political parties, non-investment funds, municipalities, organizations established in the public interest),
- erroneous records (e.g. containing negative amounts of liabilities or assets, financial statements for a given year filled in the register before the end of that year),
- records in a grey zone, for which the binary classification is not possible (e.g. companies which have been restructured, bankrupt companies for periods two year before the declaration of bankruptcy and earlier, bankrupt companies for which the date of declaration of bankruptcy is not available).

The final database contained 897,230 entries representing companies across all industries, originating, covering the period 2014 – 2019. Out of this amount, 3,023 entries were related to bankrupt companies, with financial records for the periods starting one year before the declaration of bankruptcy (earlier records were eliminated).

According to Valáškova, Klieštik and Kováčová (2018), in terms of SK NACE rev. 2 classification, highest number of bankruptcies and restructured companies in Slovakia were in sections G (wholesale and retail trade, repair of motor vehicles and motorcycles), C (manufacturing), and F (construction). We confirmed the same results for bankrupt companies only as well (see Table 2). For our research, we selected the section G, with the highest number of bankrupt companies.

Table 2

Share of Bankrupt Companies in Various Industries in Slovakia (2015 – 2019)

SK NACE	Bankrupt companies		All companies		SK NACE	Bankrupt companies		All companies	
	Volume	Share	Volume	Volume		Volume	Share	Volume	Volume
A	137	4.53%	21,728	2.43%	L	0	0.00%	0	0.00%
B	18	0.60%	936	0.10%	M	488	16.14%	210,489	23.58%
C	506	16.74%	68,630	7.69%	N	112	3.70%	57,487	6.44%
D	41	1.36%	4,281	0.48%	O	0	0.00%	1,277	0.14%
E	46	1.52%	4,232	0.47%	P	7	0.23%	10,223	1.15%
F	531	17.57%	96,413	10.80%	Q	16	0.53%	35,329	3.96%
G	687	22.73%	213,931	23.96%	R	24	0.79%	13,582	1.52%
H	118	3.90%	36,313	4.07%	S	31	1.03%	14,746	1.65%
I	114	3.77%	34,500	3.86%	T	0	0.00%	48	0.01%
J	65	2.15%	55,254	6.19%	U	2	0.07%	120	0.01%
K	80	2.65%	13,211	1.48%					

Source: Own computations, source data was provided by Finstat.

For all companies within the section G, we computed predictors X1, X2, X3, X4, X5, REVENUES, LEGAL_FORM, AGE, and a binary indicator of bankruptcy (0 if the company wasn´t bankrupt, 1 if it was). However, with 687 bankrupt companies and 213,244 non-bankrupt companies (see Table 2), the population is severely imbalanced. To solve this problem, we decided to use the oversampling, rather than undersampling, approach. For this purpose, additional data were artificially generated with SMOTE algorithm. Because the validation of the model is based on actual data, we randomly split the original datasets into training and validation sets (at a ratio of 80:20) and apply the SMOTE algorithm only to the training set. For the our research we have using three scenarios of oversampling (see Table 3).

Table 3

Impacts of the SMOTE Oversamplings on the Volume of Data

	Original data			Training set with oversampling by 300% (SMOTE)	Training set with oversampling by 20,000% (SMOTE)
	Total companies	Validation set (20%)	Training set (80%) without SMOTE		
Bankrupt companies	687	137	550	2,200	110,550
Non-bankrupt companies	213,244	42,649	170,595	170,595	170,595
Total	213,931	42,786	171,145	172,795	281,145
Bankrupt/Total	0.32%	0.32%	0.32%	1.27%	39.32%

Source: Data from our experiment, computations conducted in AZURE Machine Learning Studio.

For the first one, used for the control purposes, there is no oversampling, and the training dataset contains 550 bankruptcy records and 170,995 entries for non-bankrupt companies. In the second scenario, the number of bankruptcy records is synthetically increased by 300%. To 2,200 bankruptcy records and 170,995. For the third scenario, the number of bankruptcy records was increased by 20,000%.

3. Research Results

For the development of the model, we made 36 experiments, each one with different combinations of rates of oversampling (0%, 300%, and 20,000%), learning rates (0.00005, 0.0001, and 0.001), number of nodes (100 and 200) and learning cycles (100 and 2000). First, we started the experiments with the assumption, that even for a highly imbalanced set of data, the model would be able to provide reasonable results, without any need for oversampling. To confirm this claim, we conducted 12 experiments (labeled as NOSMOTE) with oversampling rate of 0%, and various learning rates, numbers of nodes, and learning cycles (see above). The results of all four scenarios were the same (see Table 4).

Table 4

Evaluation Metrics of the Model with No Oversampling

	TP	FN	TN	FP	Retention	Accuracy	Precision	Recall	Specificity	F-score
NOSMOTE	0	137	42,649	0	0.00%	99.68%	100.00%	0.00%	100.00%	0.00%

Note: TP stands for true positives, FN for false negatives, TN for true negatives, and FP for false positives.

Source: Data from our experiment, computations conducted in AZURE Machine Learning Studio.

Table 5

Best Values of "1 – Retention" Metrics Achieved with the Given Threshold

Threshold	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
1 – Retention	*0.106*	*0.015*	*0.056*	*0.088*	*0.123*	*0.149*	*0.009*	*0.014*	*0.022*	*0.029*
Threshold	0.22	0.24	0.26	0.28	0.30	0.32	0.34	0.36	0.38	0.40
1 – Retention	*0.029*	*0.047*	*0.057*	*0.066*	*0.077*	*0.088*	*0.100*	*0.112*	*0.126*	*0.139*
Threshold	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
1 – Retention	*0.152*	*0.166*	*0.179*	*0.193*	*0.206*	*0.222*	*0.237*	*0.254*	*0.273*	*0.294*
Threshold	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
1 – Retention	*0.314*	*0.183*	*0.201*	*0.119*	*0.128*	*0.138*	*0.150*	*0.162*	*0.175*	*0.189*
Threshold	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	–
1 – Retention	*0.189*	*0.224*	*0.247*	*0.273*	*0.307*	*0.351*	*0.134*	*0.199*	*0.191*	–

Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

Table 6

Impact of the Smote Oversampling Ratios, Learning Rates (LR), Nodes (ND), and the Learning Cycles (LC) on the Evaluation Metrics of the Model; Threshold Equal To 0.50

Threshold = 0.50			TP	FN	TN	FP	Retention	Accuracy	Precision	Recall	Specificity	F-score	Random pick
SMOTE 20,000%; LR:	.00005	ND 100, LC 1000	137	0	85	42,564	99.80%	0.52%	0.32%	100.00%	0.20%	0.64%	76.12%
		ND 100, LC 2000	137	0	2,407	40,242	94.37%	5.95%	0.34%	100.00%	5.64%	0.68%	0.04%
		ND 200, LC 1000	137	0	0	42,649	100.00%	0.32%	0.32%	100.00%	0.00%	0.64%	100.00%
		ND 200, LC 2000	137	0	26	42,623	99.94%	0.38%	0.32%	100.00%	0.06%	0.64%	92.00%
	.0001	ND 100, LC 1000	137	0	42	42,607	99.90%	0.42%	0.32%	100.00%	0.10%	0.64%	87.39%
		ND 100, LC 2000	137	0	2,851	39,798	93.34%	6.98%	0.34%	100.00%	6.68%	0.68%	0.01%
		ND 200, LC 1000	137	0	0	42,649	100.00%	0.32%	0.32%	100.00%	0.00%	0.64%	100.00%
		ND 200, LC 2000	137	0	40	42,609	99.91%	0.41%	0.32%	100.00%	0.09%	0.64%	87.96%
	.001	ND 100, LC 1000	137	0	49	42,600	99.89%	0.43%	0.11%	100.00%	0.11%	0.23%	85.45%
		ND 100, LC 2000	137	0	14	42,635	99.97%	0.35%	0.03%	100.00%	0.03%	0.07%	95.61%
		ND 200, LC 1000	137	0	86	42,563	99.80%	0.52%	0.20%	100.00%	0.20%	0.40%	75.87%
		ND 200, LC 2000	137	0	80	42,569	99.81%	0.51%	0.19%	100.00%	0.19%	0.37%	77.35%
SMOTE 300%; LR:	.00005	ND 100, LC 1000	136	1	8,832	33,817	79.36%	20.96%	20.71%	99.27%	20.71%	34.27%	0.00%
		ND 100, LC 2000	136	1	1,569	41,080	96.33%	3.98%	3.68%	99.27%	3.68%	7.09%	3.10%
		ND 200, LC 1000	137	0	121	42,528	99.72%	0.60%	0.28%	100.00%	0.28%	0.57%	67.80%
		ND 200, LC 2000	137	0	4,120	38,529	90.37%	9.95%	9.66%	100.00%	9.66%	17.62%	0.00%
	.0001	ND 100, LC 1000	45	92	39,861	2,788	6.84%	93.27%	93.46%	32.85%	93.46%	48.61%	0.00%
		ND 100, LC 2000	54	83	39,125	3,524	8.56%	91.57%	91.74%	39.42%	91.74%	55.14%	0.00%
		ND 200, LC 1000	116	21	27,563	15,086	35.58%	64.69%	64.63%	84.67%	64.63%	73.30%	0.00%
		ND 200, LC 2000	117	20	26,774	15,875	37.42%	62.85%	62.78%	85.40%	62.78%	72.36%	0.00%
	.001	ND 100, LC 1000	8	129	42,601	48	0.43%	99.59%	99.89%	5.84%	99.89%	11.03%	0.00%
		ND 100, LC 1000	8	129	42,583	66	0.47%	99.54%	99.85%	5.84%	99.85%	11.03%	0.00%
		ND 100, LC 2000	12	125	42,233	416	1.29%	98.74%	99.02%	8.76%	99.02%	16.09%	0.00%
		ND 200, LC 1000	52	85	39,317	3,332	8.11%	92.01%	92.19%	37.96%	92.19%	53.77%	0.00%

Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

Table 7

Impact of the Smote Oversampling Ratios, Learning Rates (LR), Nodes (ND), and the Learning Cycles (LC) on the Evaluation Metrics of the Model; Threshold Equal To 0.98

Threshold = 0.98			TP	FN	TN	FP	Retention	Accuracy	Precision	Recall	Specificity	F-score	Random pick
SMOTE 20,000%; LR:	.00005	ND 100, LC 1000	135	2	9,030	33,619	78.89%	21.42%	0.40%	98.54%	21.17%	0.80%	0.00%
		ND 100, LC 2000	137	0	15,012	27,637	64.91%	35.41%	0.49%	100.00%	35.20%	0.98%	0.00%
		ND 200, LC 1000	137	0	2,559	40,090	94.02%	6.30%	0.34%	100.00%	6.00%	0.68%	0.02%
		ND 200, LC 2000	137	0	3,337	39,312	92.20%	8.12%	0.35%	100.00%	7.82%	0.69%	0.00%
	.0001	ND 100, LC 1000	137	0	3,955	38,694	90.76%	9.56%	0.35%	100.00%	9.27%	0.70%	0.00%
		ND 100, LC 2000	136	1	11,643	31,006	72.79%	27.53%	0.44%	99.27%	27.30%	0.87%	0.00%
		ND 200, LC 1000	137	0	456	42,193	98.93%	1.39%	0.32%	100.00%	1.07%	0.65%	22.99%
		ND 200, LC 2000	137	0	2,128	40,521	95.03%	5.29%	0.34%	100.00%	4.99%	0.67%	0.09%
	.001	ND 100, LC 1000	137	0	876	41,773	97.95%	2.37%	2.05%	100.00%	2.05%	4.03%	5.85%
		ND 100, LC 2000	137	0	139	42,510	99.68%	0.65%	0.33%	100.00%	0.33%	0.65%	63.99%
		ND 200, LC 1000	137	0	1,767	40,882	95.87%	4.45%	4.14%	100.00%	4.14%	7.96%	0.31%
		ND 200, LC 2000	137	0	981	41,668	97.71%	2.61%	2.30%	100.00%	2.30%	4.50%	4.15%
SMOTE 300%; LR:	.00005	ND 100, LC 1000	114	23	30,252	12,397	29.29%	70.97%	70.93%	83.21%	70.93%	76.58%	0.00%
		ND 100, LC 2000	133	4	14,391	28,258	66.37%	33.95%	33.74%	97.08%	33.74%	50.08%	0.00%
		ND 200, LC 1000	130	7	17,380	25,269	59.38%	40.92%	40.75%	94.89%	40.75%	57.02%	0.00%
		ND 200, LC 2000	125	12	24,871	17,778	41.87%	58.42%	58.32%	91.24%	58.32%	71.15%	0.00%
	.0001	ND 100, LC 1000	0	137	42,647	2	0.32%	99.68%	100.00%	0.00%	100.00%	0.00%	99.36%
		ND 100, LC 2000	1	136	42,646	3	0.33%	99.68%	99.99%	0.73%	99.99%	1.45%	1.27%
		ND 200, LC 1000	16	121	42,251	398	1.25%	98.79%	99.07%	11.68%	99.07%	20.89%	0.00%
		ND 200, LC 2000	17	120	42,157	492	1.47%	98.57%	98.85%	12.41%	98.85%	22.05%	0.00%
	.001	ND 100, LC 1000	0	137	42,649	0	0.32%	99.68%	100.00%	0.00%	100.00%	0.00%	100.00%
		ND 100, LC 1000	0	137	42,649	0	0.32%	99.68%	100.00%	0.00%	100.00%	0.00%	100.00%
		ND 100, LC 2000	0	137	42,649	0	0.32%	99.68%	100.00%	0.00%	100.00%	0.00%	100.00%
		ND 200, LC 1000	1	136	42,645	4	0.33%	99.67%	99.99%	0.73%	99.99%	1.45%	1.58%

Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

T a b l e 8

Thresholds Minimizing Retention under Various Scenarios; False Negatives Kept Less Than 2

Scenario |(oversampling rate as percentage, learning rate, number od nodes, number of learning cycles)	**Threshold from .02 to .98**									
20000, 0.00005, 100, 1000										
20000, 0.00005, 100, 2000								.68 – .92		
20000, 0.00005, 200, 1000										
20000, 0.00005, 200, 2000										
20000, 0.0001, 100, 1000									.94 – .96	
20000, 0.0001, 100, 2000					.14					
20000, 0.0001, 200, 1000										.98
20000, 0.0001, 200, 2000										
20000, 0.001, 100, 1000										
20000, 0.001, 200, 1000										
20000, 0.001, 100, 2000										
20000, 0.001, 200, 2000										
300, 0.001, 100, 1000						.16 – .62				
300, 0.001, 100, 2000										
300, 0.001, 200, 1000										
300, 0.001, 200, 2000							.64 – .66			
300, 0.0001, 100, 1000										
300, 0.0001, 100, 2000										
300, 0.0001, 200, 1000			.06 – .10							
300, 0.0001, 200, 2000		.04		.12						
300, 0.00005, 100, 1000										
300, 0.00005, 100, 2000										
300, 0.00005, 200, 1000										
300, 0.00005, 200, 2000	.02									

Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

While the accuracy and specificity of these experiments were high, other measures (not to mention the common sense) indicated, that the model did not work. In fact, it wasn't able to identify a single bankruptcy (out of 137 actual cases) in any of the of the 12 experiments. In other words, because the bankrupt companies account for less than 0.33% of all records, the model was unable to detect any pattern and instead labels all companies as non-bankrupt. As a result, we decided to conduct additional experiments with oversampling. Instead of simply replicating the exact copies of existing records, additional data is artificially derived from those records with the SMOTE algorithm. The results of the 12 experiments with the oversampling rate of 300% and 12 with the oversampling rate of 20,000% are summarized in the next tables. Table 5 represents the values of the descriptive measures for the threshold equal to 0.50 (a standard value), Table 6 for the threshold equal to 0.98. Because each of the 24 of experiments has different parameters, they react to adjustment of threshold differently as well. For the selection of the best combinations of hyperparameters for a given threshold, we first filtered only experiments with true positives equal to 137 or 136 (there have been 137 bankrupt companies in a training sample) and then selected the result with the *lowest* value of *retention*. See Table 8 for a summary of the optimal pairs of thresholds and various scenarios if the cap for the false negatives is less than 2, and Table 5 for a summary of the best values of *1 – retention* measure achieved with a given threshold.

Conclusions

Following the result of the 36 experiments and former studies, we were able to confirm two underlying ideas.

First, under of condition of extremely imbalanced datasets, it could be extremely difficult to achieve acceptable level of prediction power, even when artificial neural networks are employed in the prediction model. While this conclusion is not a result of the scientifically conducted analysis we proved, that it is true at least in some cases (thus rejecting the hypothesis of the absence of any problems).

Second, while the bankruptcy prediction model achieves poor results in terms of signal to noise ratio (there are too many false positive cases predicted) it could be still used as a preliminary classification instrument. If the model is generating too many false positives, but is able to keep the level of false negative under control (in our case, with a tolerance of 1 company out of 137 bankrupted companies), all cases reported as negative would not require further, more detailed, time-consuming, and expensive analyses. For our experiments, the scenario with the threshold of 0.92, oversampling rate of 20,000%, the learning rate of 0.00005, 100 nodes, and 2,000 learning cycles (Table 8, Table 5, and Figure 1) achieved the best result, with the value of *1 – retention* equal to 0.351.

Table 9

Evaluation Metrics for the Combination of Parameters with Best Performance

Treshold = 0.92	**TP**	**FN**	**TN**	**FP**	**F-score**
SMOTE 20000, 0.00005, 100, 2000	137	0	15,012	27,637	0.98%
Random pick	**Retention**	**Accuracy**	**Precision**	**Recall**	**Specificity**
0.00%	64.91%	35.41%	0.49%	100.00%	35.20%

Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

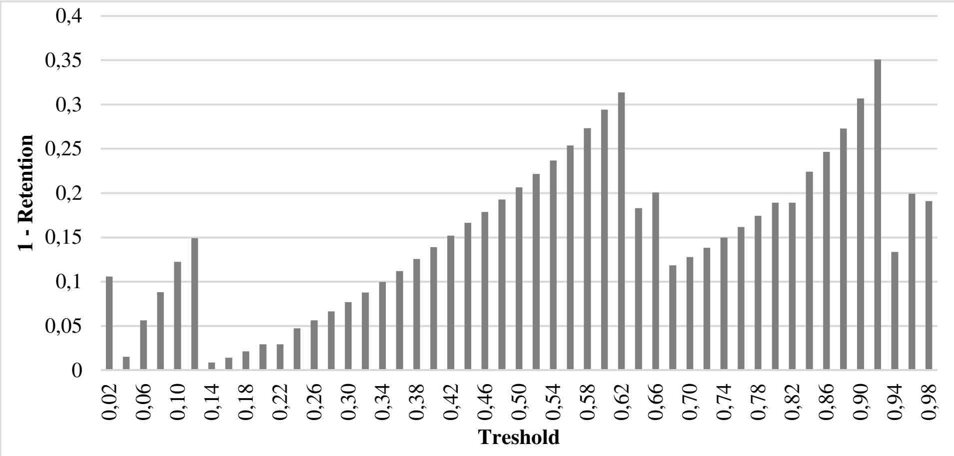
Under this scenario, of the sample of 42,786 Slovak companies with 137 bankruptcies and 42,649 non-bankrupt companies, the system was able to correctly identify 15,012 as non-bankrupt. Because there weren´t any false negatives, the number of companies which have to be subsequently assessed is equal to 27,774 (137 true positives + 27 637 false positives). Even though this number is still extremely high, the model results in significant decrease of the analyses which would be otherwise required by 35.1%. Value of the *Random pick* variable (see Table 9) represents the probability of randomly picking 137 bankrupted companies (out of 137) in a sample of 27,774 entries from the population of 42,649 records. In our case, the probability is almost zero, being equal to 1.73×10^{-26}.

There are few closing remarks. First, we intentionally rely on of the most rigid definition of bankruptcy. The company is flagged as actually bankrupt only if it has been declared as so by the court in accordance with the law. As a result, while the amount of available data for training was decreased, we have also reduced the risk of inconclusive results. Besides, we skipped the data that were corrupt (e.g. erroneous from the point of view of accounting). While the inclusion of such data could enhance the number of positive cases in a population, we believe, that such data are red flags of the financial problems on themselves (not to mention, that they could taint the overall learning process of the ANNs). Additionally, data from bankrupt companies covering the periods two years before the declaration of bankruptcy (or earlier) were not taken into consideration, because of their Schrödinger´s nature. Also, changing the parameters of our model (or any ANN-based model, for that matter) could result in finding a local, rather global, optima (see Figure1 and Table 8). Hence there is a risk of fast jumping to a conclusion, without taking in account combinations of other parameters and/or inputs which could further improve the predictive power of model.

Finally, the performance of the model could be increased by either taking more risky strategies (e.g. by allowing more false negatives) and/or by factoring in the estimated individual costs of false positives and false negatives.

F i g u r e 1

Impact of the Level of Thresholds on the Best Achieved Values of the "1 – Retention" Metric



Source: Summary of the results of our experiments conducted in AZURE Machine Learning Studio.

References

ABDELWAHED, T. – AMIR, E. M. (2005): New Evolutionary Bankruptcy Forecasting Model Based on Genetic Algorithms and Neural Networks. In: Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI05). IEEE Computer Society, pp. 1 – 5. DOI: 10.1109/ICTAI.2005.92.

ALTMAN, E. I. (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance, *23*, No. 4, pp. 589 – 609. DOI: 10.1111/j.1540-6261.1968.tb00843.x.

ALTMAN, E. I. (1973): Predicting Railroad Bankruptcies in America. Bell Journal of Economics and Management Science. *4*, No. 1, pp. 184 – 211. DOI: 10.2307/3003144.

BACK, B. – LAITINEN, T. – SERE, K. (1996): Neural Network and Genetic Algorithm for Bankruptcy Predictions. Expert Systems with Applications, *4*, No. 11, pp. 407 – 413. DOI: 10.1016/S0957-4174(96)00055-3.

BARBOZA, F. et al. (2017): Machine Learning Models and Bankruptcy Prediction. Expert Systems with Applications, *83*, pp. 405 – 407. DOI: https://doi.org/10.1016/j.eswa.2017.04.006.

BEAVER, W. H. (1966): Financial Ratios as Predictors of Failure. Journal of Accounting Research, *4*, No. 1, pp. 71 – 102. DOI: 10.2307/2490171.

BELL, T. et al. (1990): Neural Nets Versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures. Proceedings of the 1990 Deloitte Touche/University of Kansas Symposium on Auditing Problems, pp. 29 – 53.

BLANCO-OLIVER, A. – IRIMIA-DIEGUEZ, A. – OLIVER-ALFONSO, M. – WILSON, N. (2015): Improving Bankruptcy Prediction in Micro-Entities by Using Nonlinear Effects and Non-Financial Variables. Finance a Uver: Czech Journal of Economics & Finance, *65*, No. 2, pp. 144 – 166.

BLUM, M. (1974): Failing Company Discriminant Analysis. Journal of Accounting Research, *12,* No. 1, pp. 1 – 25. DOI: 10.2307/2490525.

BOĎA, M. (2009): Predicting Bankruptcy of Slovak Enterprises by an Artificial Neural Network. Forum Statisticum Slovacum, No. 6/2009, pp. 3 – 6.

BORITZ, J. – KENNEDY, D. (1995): Effectiveness of Neural Networks Types for Prediction of Business Failure. Expert Systems with Applications, *9,* No. 4, pp. 503 – 512. DOI: 10.1016/0957-4174(95)00020-8.

BREIMAN, L. (1996): Bagging Predictors. Machine Learning, *24*, No. 2, pp. 123 –140. DOI: 10.1007/bf00058655.

CADDEN, D. (1991): Neural Networks and the Mathematics of Chaos – An Investigation of These Methodologies as Accurate Predictions of Corporate Bankruptcy. The First International Conference on Artificial Intelligence Applications of Wall Street, pp. 52 – 57. DOI: 10.1109/AIAWS.1991.236574.

CARLOS, S. C. (1996): Self Organizing Neural Networks for Financial Diagnosis. Decision Support System, *17,* No. 3, pp. 227 – 238. DOI: 10.1016/0167-9236(95)00033-X.

COATS, P. – FANT, L. (1993): Recognizing Financial Distress Patterns Using A Neural Network Tool. Financial Management, *22*, No. 3, pp. 142 – 155. DOI: 10.2307/3665934.

CÚT, S. (2013): Aplikácia neurónových sietí vo finančnej analýze podniku s využitím SPSS. Forum Statisticum Slovacum, *9*, No. 7, pp. 38 – 43.

DEAKIN, E. (1972): A Discriminant Analysis of Predictors of Business Failure. Journal of Accounting Research, *10*, No. 1, pp. 167 – 179. DOI: 10.2307/2490225.

EDMISTER, R. (1972): An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. Journal of Financial and Quantitative Analysis, *7*, No. 2, pp. 1477 – 1493.

FARIS, H. et al. (2020): Improving Financial Bankruptcy Prediction in a Highly Imbalanced Class Distribution Using Oversampling and Ensemble Learning: A Case from the Spanish Market. Progress in Artificial Intelligence, *9*, No. 1, pp. 31 – 53. DOI: 10.1007/s13748-019-00197-9.

FITZPATRICK, P. J. (1932): A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Firm. Certified Public Accountant, No. 6, pp. 727 – 731.

FLETCHER, D. – GOSS, E. (1993): Forecasting with Neural Networks: an Application Using Bankruptcy Data. Information and Management, *24*, No. 3, pp. 159 – 167. DOI: 10.1016/0378-7206(93)90064-Z.

HANWECK, G. (1977): Predicting Bank Failures. [Research Papers.] Washington, DC: in Banking and Financial Economics, Financial Studies Section, Board of Governors of the Federal Reserve System.

HARDLE, W. et al. (2019): Variable Selection and Oversampling in the Use of Smooth Support Vector Machines for Predicting the Default Risk of Companies. Journal of Forecasting, *28*, No. 6, pp. 512 – 534. DOI: 10.1002/for.1109.

CHUDSON, W. (1945): The Pattern of Corporate Financial Structure. New York: National Bureau of Economic Research.

IGNIZIO, J. P. – SOLTYS, J. R. (1996): Simultaneous Design and Training of Ontogenic Neural Network Classifiers. Computers & Operations Research, *23*, No. 6, pp. 535 – 546. DOI: 10.1016/0305-0548(95)00058-5.

JACKENDOFF, N. (1962): A Study of Published Industry Financial and Operating Ratios. Philadelphia: Temple University, Bureau of Economic and Business Research.

KADEROVÁ, A. – KRÁTKA, Z. – KRČOVÁ, I. – MUCHA, V. – ŠOLTÉSOVÁ, T. (2020): Matematika pre ekonómov. Bratislava: Letra Edu. 333 s. ISBN 978-90-89962-62-4.

KIM, H. J. et al. (2016): Optimization of Cluster-Based Evolutionary Undersampling for the Artificial Neural Networks in Corporate Bankruptcy Prediction. Expert Systems with Applications, *59*, pp. 226 – 234. DOI: 10.1016/j.eswa.2016.04.027.

KIM, M. J. – KANG, D. K. (2010): Ensemble with Neural Networks for Bankruptcy Prediction. Expert Systems with Applications, *37*, No. 4, pp. 3373 – 3379. DOI: 10.1016/j.eswa.2009.10.012.

KOSTER, A. N. – SONDAK, E. – BOURBIA, W. (1990): A Business Application of Artificial Neural Network Systems. The Journal of Computer Information Systems, *31*, No. 2, pp. 3 – 9.

LE, T. et al. (2018): A Cluster-Based Boosting Algorithm for Bankruptcy Prediction in a Highly Imbalanced Dataset. Symmetry-Basel, *10*, No. 7, pp. 1 – 12, Article Number: 250. DOI: 10.3390/sym10070250.

LEE, K. C. – HAN, I. G. – KWON, Y. (1996): Hybrid Neural Network Models for Bankruptcy Predictions. Decision Support Systems, *18*, No. 1, pp. 63 – 72. DOI: 10.1016/0167-9236(96)00018-8.

LEE, S. – CHOI, W. S. (2013): A Multi-Industry Bankruptcy Prediction Model Using Back-propagation Neural Network and Multivariate Discriminant Analysis. Expert Systems with Applications, *40*, No. 8, pp. 2941 – 2946. DOI: 10.1016/j.eswa.2012.12.009.

LESHNO, M. – SPECTOR, Y. (1996): Neural Network Prediction Analysis: The Bankruptcy Case. Neurocomputing, *10*, No. 2, pp. 125 – 147. DOI: 10.1016/0925-2312(94)00060-3.

LI, H. – SUN, J. (2012): Forecasting Business Failure: The Use of Nearest-Neighbour Support Vectors and Correcting Imbalanced Samples – Evidence from the Chinese Hotel Industry. Tourism Management, *33*, No. 3, pp. 622 – 634. DOI: 10.1016/j.tourman.2011.07.004.

LIANG, L. – WU, D. (2005): An Application of Pattern Recognition on Scoring Chinese Corporations Financial Conditions Based on Backpropagation Neural Network. Computers & Operations Research, *32*, No. 5, pp. 1115 – 1129. DOI: 10.1016/j.cor.2003.09.015.

MANSKI, C. F. – LERMAN, S. R. (1977): The Estimation of Choice Probabilities from Choice Based Samples. Econometrica, *45,* No. 8, pp. 1977 – 1988. DOI: 10.2307/1914121.

MARTIN, D. (1977): Early Warning of Bank Failures: A Logit Regression Approach. Journal of Banking and Finance *1*, No. 3, pp. 249 – 276. DOI: 10.1016/0378-4266(77)90022-X.

MENSAH, Y. (1983): The Differential Bankruptcy Predictive Ability of Specific Price Level Adjustments: Some Empirical Evidence. The Accounting Review, *58*, No. 2, pp. 228 – 246.

MERWIN, C. (1942): Financing Small Corporations in Five Manufacturing Industries, 1926 – 1936. New York: National Bureau of Economic Research.

MIHALOVIČ, M. (2016): Performance Comparison of Multiple Discriminant Analysis and Logit Models in Bankruptcy Prediction. Economics & Sociology, *9*, No. 4, pp. 101 –118. DOI: 10.14254/2071-789X.2016/9-4/6.

NEUMAIER, I. – NEUMAIEROVÁ, I. (2005): Index IN 05. In: Sborník příspěvků mezinárodní vědecké konference „Evropské finanční systémy“. Brno: Masarykova univerzita v Brně, pp. 143 – 148.. ISBN 80-210-3753-9.

NEUMAIER, I. – NEUMAIEROVÁ, I. (2013): Vypovídací schopnost indexu IN05. In: Sborník z mezinárodní vědecké konference „Ekonomika v pohybu“. Praha: Vysoká škola ekonomická v Praze, pp. 169 – 176. ISBN 978-80-245-1943-2.

ODOM, M. D. – SHARDA, R. (1990): A Neural Network Model for Bankruptcy Prediction. In: Proceedings from IJCNN International Joint Conference on Neural Networks. San Diego: IEEE, pp. 163 – 168. DOI: 10.1109/IJCNN.1990.137710.

OHLSON, J. A. (1980): Financial Ratios and Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, *18*, No. 1, pp. 109 – 131. DOI: 10.2307/2490395.

ONAN, A. (2019): Consensus Clustering-Based Undersampling Approach to Imbalanced Learning. Scientific Programming, No. 3, pp. 1 – 14, (Art. No. 5901087). DOI: 10.1155/2019/5901087.

PENDHARKAR, P. C. – RODGER, J. A. (2004): An Empirical Study of Impact of Crossover Operators on The Performance of Non-Binary Genetic Algorithm Based Neural Approaches for Classification. Computers & Operations Research, *31*, No. 4, pp. 481 – 498. DOI: 10.1016/S0305-0548(02)00229-0.

PENDHARKAR, P. C. (2005): A Threshold Varying Artificial Neural Network Approach for Classification and Its Application to Bankruptcy Prediction Problem. Computers & Operations Research, *32*, No. 10, pp. 2561 – 2582. DOI: 10.1016/j.cor.2004.06.023.

RAFIEI, F. M. – MANZARI, S. M. – BOSTANIAN, S. (2011): Financial Health Prediction Models Using Artificial Neural Networks, Genetic Algorithm and Multivariate Discriminant Analysis: Iranian Evidence. Expert Systems with Applications, *38*, pp. 10210 – 10217. DOI: 10.1016/j.eswa.2011.02.082.

SAI, Y. – ZHONG, C. J. – QU, L. H. (2007): A Hybrid GA-BP Model for Bankruptcy Prediction. In: Proceedings of the International Symposium on Autonomous Decentralized Systems (ISADS 2007), 21 – 23 March 2007, Sedona, pp. 473 – 477. DOI: 10.1109/ISADS.2007.3.

SALCHENBERGER, L. M. – CINAR, E. M. – LASH, N. A. (1992): Neural Networks: A New Tool for Predicting Thrift Failures. Decision Sciences, *23*, No. 4, pp. 899 – 916. DOI: 10.1111/j.1540-5915.1992.tb00425.x.

SANCHEZ-LASHERAS, F. et al. (2012): A Hybrid Device for the Solution of Sampling Bias Problems in the Forecasting of Firms' Bankruptcy. Expert Systems with Applications, *39*, No. 8, pp. 7512 – 7523. DOI: 10.1016/j.eswa.2012.01.135.

SHIN, S. W. – LEE, K. C. – KILIC, S. B. (2006): Ensemble Prediction of Commercial Bank Failure through Diversification of Input Features. In: Lecture Notes in Computer Science, *4304*, pp. 887 – 896. DOI: 10.1007/11941439_93.

SHRIVASTAVA, S. – JEYANTHI, P. M. – SINGH, S. (2020): Failure Prediction of Indian Banks Using Smote, Lasso Regression, Bagging and Boosting. Cogent Economics & Finance, *8*, No. 1, pp. 1 – 17, Article Number: 1729569.

SMITH, R. – WINAKOR, A. (1935): Changes in Financial Structure of Unsuccessful Industrial Corporations. [Bureau of Business Research, Bulletin, No. 51.] Urbana: University of Illinois Press.

SMITI, S. et al. (2020): Bankruptcy Prediction Using Deep Learning Approach Based on Borderline SMOTE. Information Systems Frontiers, *22*, No. 5, pp. 1067 – 1083. DOI: 10.1007/s10796-020-10031-6.

SZOLNO, O. (2016): Wybrane aspekty pomiaru efektywnosci i skutecznosci jednostki samorzadu terytorialnego. Finanse, Rynki Finansowe, Ubezpieczenia, *6/1*, No. 84, pp. 91 – 99.

TAM, K. Y. – KIANG, M. Y. (1992): Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. Management science, *38*, No. 7, p. 926 – 947. DOI: 10.1287/mnsc.38.7.926.

TSAI, CH. F. – HUNG, CH. (2014): Modeling Credit Scoring Using Neural Network Ensembles. Kybernetes, *43*, No. 7, pp. 1114 – 1123. DOI: 10.1108/K-01-2014-0016.

VALÁŠKOVA, K. – KLIEŠTIK, T. – KOVÁČOVÁ, M. (2018): Management of Financial Risks in Slovak Enterprises Using Regression Analysis. Oeconomia Copernicana, *9*, No. 1, pp. 105 – 121. Available at: <https://doi.org/10.24136/oc.2018.006>.

VEGANZONES, D. – SEVERIN, E. (2018): An Investigation of Bankruptcy Prediction in Imbalanced Datasets. Decision Support Systems, *112*, pp. 111 – 124.DOI: 10.1016/j.dss.2018.06.011.

WALLRAFEN, J. – PROTZEL, P. – POPP, H. (1996): Genetically Optimized Neural Network Classifiers For Bankruptcy Prediction. In: Proceedings of the 29th Annual Hawaii International Conference on System Sciences, Wailea, pp. 419 – 426. DOI: 10.1109/HICSS.1996.495427.

WEST, D. – DELLANA, S. – QIAN, J. (2005): Neural Network Ensemble Strategies for Financial Decision Applications. Computers and Operations Research, *32*, No. 10, pp. 2543 – 2559. DOI: 10.1016/j.cor.2004.03.017.

WU, D. – LIANG, L. – YANG, Z. (2008): Analyzing the Financial Distress of Chinese Public Companies Using Probabilistic Neural Networks and Multivariate Discriminate Analysis. Socio--Economic Planning Sciences, *42*, No. 3, pp. 206 – 220. DOI: 10.1016/j.seps.2006.11.002.

ZHANG, G. – HU, M. Y. – PATUWO, B. E. – INDRO, D. C. (1999): Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis. European Journal of Operational Research, *116,* No. 6, pp. 16 – 32. DOI: 10.1016/S0377-2217(98)00051-4.

ZHOU, L. (2013): Performance of Corporate Bankruptcy Prediction Models on Imbalanced Dataset: The Effect of Sampling Methods. Knowledge-Based Systems, No. 41 pp. 16 – 25. DOI: 10.1016/j.knosys.2012.12.007.

ZMIJEWSKI, M. E. (1984): Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research, *22*, No. 1, pp. 59 – 82. DOI: 10.2307/2490859.

ZORIČÁK, M. et al. (2020): Bankruptcy Prediction for Small- and Medium-sized Companies Using Severely Imbalanced Datasets. Economic Modelling, *84*, pp. 165 – 176. DOI: 10.1016/j.econmod.2019.04.003.