sustainability

MDPI

Article

Predicting Financial Distress of Slovak Enterprises: Comparison of Selected Traditional and Learning Algorithms Methods

Elena Gregova [1], Katarina Valaskova [1,*], Peter Adamko [1], Milos Tumpach [2] and Jaroslav Jaros [3]

1 Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia; elena.gregova@fpedas.uniza.sk (E.G.); peter.adamko@fpedas.uniza.sk (P.A.)
2 Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemska cesta 1, 852 35 Bratislava, Slovakia; milos.tumpach@euba.sk
3 Center for Technology Transfer, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia; jaroslav.jaros@uvp.uniza.sk
* Correspondence: katarina.valaskova@fpedas.uniza.sk

Received: 26 March 2020; Accepted: 9 May 2020; Published: 12 May 2020

check for updates

Abstract: Predicting the risk of financial distress of enterprises is an inseparable part of financial-economic analysis, helping investors and creditors reveal the performance stability of any enterprise. The acceptance of national conditions, proper use of financial predictors and statistical methods enable achieving relevant results and predicting the future development of enterprises as accurately as possible. The aim of the paper is to compare models developed by using three different methods (logistic regression, random forest and neural network models) in order to identify a model with the highest predictive accuracy of financial distress when it comes to industrial enterprises operating in the specific Slovak environment. The results indicate that all models demonstrated high discrimination accuracy and similar performance; neural network models yielded better results measured by all performance characteristics. The outputs of the comparison may contribute to the development of a reputable prediction model for industrial enterprises, which has not been developed yet in the country, which is one of the world's largest car producers.

Keywords: default; enterprise in crisis; bankruptcy; financial distress; prediction models

1. Introduction

The economy is built on the successful functioning of enterprises. In current conditions, however, an increasing number of corporate defaults occur, which is caused by various factors. The financial distress of business entities is closely connected with unpleasant consequences, and these are the main motivation factors for managers or financial analysts to search for the methods that can predict possible financial problems or bankruptcy in advance. Financial analysis may help solve the problems, as it focuses on the determination of the factors (and their intensity); it forms the financial stability of enterprises and reveals corporate strengths and weaknesses and thus becomes a necessary and effective diagnostic tool of the corporate financial health prediction. Since the development of the first prediction models in 1930 [1], hundreds of bankruptcy prediction models have been introduced worldwide (e.g., Alaka et al. [2]). However, the results of many researches confirm that the reliability and predictive accuracy of the models decrease if they are used in different national environments and time horizons than those in which they were originally formed [3–5]. The development of prediction models in unique national conditions is of vital importance if the financial risks are to be estimated correctly. One-country studies play a significant role in the research of bankruptcy prediction.

Considering the eastern European countries, especially the Visegrad Group—which is the political and cultural alliance of countries for the purpose of the social, energy and economic cooperation—most of them predict the financial health of enterprises using the bankruptcy models that were formed in their national environment and are generally known. For instance, in Hungary, the most reputable model is the prediction model of Virag and Hajdu [6] developed for industrial enterprises; in the Czech Republic the models of the Neumaiers [7–10] focused either on industrial enterprises or all sectors of the economy; the widely used models in Poland are the bankruptcy models for industrial enterprises introduced by Maczynska [11], Gajdki and Stosa [12], Hamrol et al. [13]; the Poznanski model by Prusak [14] and the general model of Gruszynski [15]; in the conditions of Slovakia the most significant are the models of Chrastinova [16] and Gurcik [17], both developed for agricultural entities. All four countries are high-income industrial countries; thus, predicting the future financial situation with a high level of accuracy is essential. However, the situation in Slovakia—the world's largest per-capita car producer [18]—is not solved properly. Despite the fact that several models have been developed in the national conditions, e.g., Binkert [19], Hurtosova [20], Delina and Packova [21], Rohacova and Kral [22], Gulka [23] and Boda and Uradnicek [24], they specialize in the bankruptcy prediction of agricultural enterprises or do not have any sectoral orientation; none of them is focused on the industrial sector (the industrial sector includes the enterprises operating in sectors B–E according to the statistical classification of economic activities in the European Community).

The need for the development of a bankruptcy model for industrial enterprises in the Slovak environment is indisputable. The Slovak Republic has a small, open economy driven mainly by automobile and electronics exports, accounting for more than 80% of the GDP. It is also one of the fastest growing economies in Europe and the fifth largest car producer in the European Union. Slovakia continues exhibiting robust economic performance, with strong growth backed by a sound financial sector, low public debt and high international competitiveness drawing on large inward investment.

Several different statistical methods may be used to form the model; thus, the main aim of the paper is to compare selected traditional and machine learning methods (logistic regression, random forest and neural network) in the conditions of the Slovak environment when predicting the financial health of enterprises. Identification of the most relevant and accurate method is useful to form the model predicting the financial distress of industrial enterprises in the specific national environment, and the results may be applied also in other countries with a similar economic structure and business orientation. The originality of the paper is based on the presentation of different methods of bankruptcy prediction applied on a dataset of about 50,000 industrial enterprises (on average) in each analyzed period (2016–2018). This period was chosen due to new legislation being applied on entities in financial crisis into practice, which entered into force in 2016; the last period correspondents with the newest available data (2018). The importance of the study is underlined by the fact that the information about the stable development of industrial enterprises in the future can contribute to the elimination of potential financial risks and thus to the improvement of the decision-making process of investors and creditors.

Following the amended Slovak legislation (Law No. 513/1991 Coll. Commercial Code and Law no. 7/2005 Coll. On Bankruptcy and Restructuring as amended), a company is in default if it has liabilities to at least two entities and the value of the liabilities exceeds the value of its assets or if it is unable to pay at least two financial liabilities to at least two creditors 30 days after the due date. An enterprise is at risk of imminent default when it has a low equity-to-liability ratio, which is strictly limited to be less than 4 to 100 for the year 2016, 6 to 100 for the year 2017 and 8 to 100 for the year 2018 and any other following year [25]. According to available statistical data (financial reports of industrial enterprises were obtained from the register of the financial statements, www.registeruz.sk), if the law were in force between 2011 and 2015, 22,591, 38,413, 41,952, 41,905 and 40,636 enterprises would be in crisis each year, respectively. Moreover, an enterprise is in financial distress if the low value of the equity-to-liability ratio is accompanied by negative profit after taxes and the ratio of current-assets-to-current-liabilities (current ratio) is less than 1. Furthermore, the enterprises of which value of assets is lower than the

value of liabilities (negative equity) are also entities with financial problems [26,27]. Despite the fact that the determination of an enterprise in default is based on the Slovak legislation, the same limitation is relevant also in the context of different countries, as the negative equity, negative profit after taxes and low level of the equity-to-liability ratio are general indicators of non-prosperity. The optimal value of the current ratio is defined divergently in the literature; however, if its value is below 1, it shows that an enterprise may not be able to meet its obligations in the short run [28]. It should be emphasized that the equity-to-liabilities ratio varies across the industries [29]; however, on the other hand, the Slovak legislation does not consider the economic sector and sets the same limit value for all enterprises on the market. Based on the presented information, it can be assumed that the research on the use of traditional and learning algorithms for bankruptcy prediction in conditions of the Slovak Republic can produce interesting findings.

The paper is divided into the following sections. The literature review depicts the most important as well as recent studies and tries to connect the research aim and the literature's previous findings. The Data and Methodology section highlights the methods used and determines the data used for the analysis. The outputs of calculations, crucial findings and comparison of results with other studies are portrayed in the Results and Discussion section.

2. Literature Review

The bankruptcy prediction, i.e., the prediction of financial distress, has been a highly discussed topic for several decades. The first studies on bankruptcy forecasting date to the beginnings of the 20th century, with the most significant studies being those of Beaver [30], Altman [31] and Ohlson [32]. However, in Europe, this phenomenon assumed its importance in the 1990s when the economic systems started to be changed [33]. The complex list of models of European transition economies with the specification of their sample size, economic sector, type of statistical method used, and prediction accuracy are portrayed in the work of Kliestik et al. [34]. Different prediction models were developed worldwide, helping business entities to forecast their financial stability in the upcoming period, which is important not only for the enterprise itself but also for its business partners [35,36]. Kliestik et al. [37] in their research confirm that the issue of bankruptcy predictions ensures business continuity and sustainable and ethically responsible economic development.

Chou et al. [38] add that when forming the bankruptcy prediction model, the financial ratios selection and the classifier design play major roles. The importance of the financial ratios used as predictors of financial health is depicted in the renowned studies of Sharifabadi et al. [39], Tian et al. [40], Bellovary et al. [41], Ravi Kumar and Ravi [42], Calderon et al. [43], Dimitras et al. [44], O'Leary [45] and Scott [46]. However, two studies are especially important in the identification of crucial financial ratios. The first one is the research of Bellovary et al. [41], where the authors analyze 165 prediction models. They state that 752 different variables were used in the models, with up to 674 of these variables being used in only one or two models. At the conclusion of the study, they present 42 variables that were used in more than five models (the most frequently used were earnings-after-taxes-to-total-assets ratio, current-assets-to-current-liabilities ratio, working-capital-to-total-assets ratio, retained-earnings-to-total assets ratio and earnings-before-interest-and-taxes-to-total-assets ratio, appearing in more than 30 models). The authors of the second important study [42] analyzed 62 prediction models and put in order the 20 most relevant financial ratios based on their frequency of occurrence, i.e., earnings-after-taxes-to-total-assets ratio, retained-earnings-to-total assets ratio, sales-to-total-assets ratio, earnings-before-interest-and-taxes-to-total-assets ratio and current-assets-to-current-liabilities ratio. As declared in the study of Kovacova et al. [47], each country prefers different explanatory variables when developing a bankruptcy prediction model. Their results reveal that prediction models being developed in the Slovak Republic prefer the current ratio, liabilities-to-total-assets ratio, equity-to-total assets ratio, return on assets (ROA) and cash ratio. By contrast, the weakest predictors are macro-economic variables, analyst recommendations and industry variables [48], which is in contrast with other studies confirming the significance of macroeconomic variables in predicting

financial distress, e.g., Jacobson et al. [49], Bruneau et al. [50] and Nam et al. [51]. Nonetheless, the research of Zikovic [52] underlines that the probability of financial distress is influenced by both firm specificities and macroeconomic variables. The utility of combining accounting and macro-economic data in financial distress prediction is confirmed also in the research of Tinoco and Wilson [53] and Giriuniene et al. [54]. The study of Filipe et al. [55] concludes that the same firm-specific factors are essential in predicting the financial distress of small and medium-sized enterprises across Europe; however, considering the macroeconomic variables, they differ based on regional specifications. Kacer et al. [56] finds that the classification performance of the prediction models is improved when the non-financial variables are included in the model, but they do not recommend the use of macroeconomic variables. Wilson et al. [57] confirm the usefulness of non-financial information in the prediction of financial distress of enterprises and find that the transition process variables, along with financial and non-financial variables, influence the probability of failure. Du Jardin [58] highlights that time horizon also plays an important role in the bankruptcy prediction, and that the optimal forecasting horizon is usually one year.

Undoubtedly, an inevitable role is played by statistical methods and models used to predict the future development of enterprises [59]. Mattsson and Steinert [60] state that the quality of the model is given by the statistical method that is applied; the results of their research prove that in recent years artificial intelligence and machine learning methods have achieved promising results in corporate bankruptcy prediction settings compared to the traditional method used (logistic regression or multiple discriminant analysis). The assessment of the bankruptcy risk of large companies by Barbatu-Misu and Madaleno [61] confirms that the estimation of bankruptcy risk is important for managers in decision-making and in the process of the improvement of corporate financial performance. Their findings show that the principal component analysis based on discriminant analysis indices is more effective when used to determine the corporate financial risks. In addition, the methodological aspects of designing a scoring model for an early prediction of bankruptcy using ensemble classifiers are examined by Pisula [62]. Oliveira et al. [63] aim to develop a multiple criteria system to predict bankruptcy in small and medium-sized enterprises combining the cognitive mapping and categorical-based evaluation technique Macbeth. Tsai [64] demonstrates that assessing the credit risk and possibility of bankruptcy are important issues before investment, and moreover, the data mining and machine learning techniques are more frequently used to solve credit scoring problems. Le et al. [65] highlight the importance of bankruptcy prediction for financial institutions, fund managers, lenders, governments and economic stakeholders. However, they stress that the imbalance of bankruptcy companies and health companies may cause classification errors and advise the use of cluster-based boosting algorithms for effective bankruptcy prediction. In addition, their study on a sample of Korean companies proves that the use of oversampling methods to balance the dataset of analyzed enterprises enhances the performance of the bankruptcy prediction [66]. Le et al. [67] present a new machine-learning model (GPU-based extreme gradient boosting machine), which outperforms the current machine learning approaches for bankruptcy forecasting in terms of geometric mean and area under the receiver operating characteristic curve (AUC). The findings of the researches of Wang et al. [68] and of Mai et al. [69] confirm the superiority of learning machine algorithms (super vector machines and random forest) in terms of classification ability, type I and II errors and AUC curve. Hosaka [70] shows that convolutional neural networks show higher discrimination accuracies than conventional methods. Qu et al. [71] claim that the development of modern information technologies causes a decrease in the use of traditional prediction methods—logistic regression (LR) and multiple discriminant analysis (MDA)—and, by contrast, causes the evolution of the machine learning use to do the prediction. Moreover, several authors tried to compare the predictive ability of traditional prediction methods [72]. Affes and Hentati-Kaffel [73] found that the logit model outperforms the discriminant analysis model in terms of correct classification rate. Using data from 1985 to 2013 on North American enterprises, Barboza et al. [74] report that comparing the best prediction models, random forest led to 87% accuracy, while logistic regression and discriminant analysis led to 69% and 50% accuracy, respectively. A study of 236 enterprises operating in

Slovakia proved that the model based on a logit function outperforms the classification accuracy of the discriminant model [75]. As declared by the results of others researches focused on the comparison of traditional and machine learning models, e.g., Cho et al. [76], Van Gestel et al. [77], Kim [78], Chen [79] and Nyitrai and Virag [80], the models based on the principles of discriminant analyses achieve the weakest prediction ability (the linear regression models outperform the prediction accuracy almost in all cases); thus we decided not to include this method into the comparison of statistical methods in the conditions of Slovak industrial enterprises. However, the findings of all the cited researches indicate that the best predictive accuracy is achieved by learning algorithms. Altman et al. [81] also compare the predictive accuracy of different estimation methods used to assess the financial health of small and medium-sized enterprises up to 10 years before the default in an open European market. Their findings affirm that logistic regression and neural networks are superior to other approaches. The importance of the logistic regression in bankruptcy prediction modeling is underlined in the research of Ben Jabeur [82]. Olson et al. [83] found that decision trees are relatively more accurate compared to neural networks and support vector machines. The research of Klepac and Hampel [84] focuses on medium-sized enterprises in Europe that went bankrupt in 2014; and the importance of business risks of small and medium-sized enterprises for their operation is portrayed in the paper of Hudakova et al. [85]. They found that learning algorithms achieve much better results compared to other methods, especially in predictive accuracy. Garcia et al. [86] point out that both advanced statistical and machine learning models may demonstrate their effectiveness when assessing financial data, which are often specified by different imperfections. However, on the other hand, there are also researchers who do not recommend the use of machine learning in the field of business [87] because the prediction accuracy does not far exceed the statistical models and the results are not interpretable. As stated by Svabova and Durica [88], the proper use of statistical methods ensures the correct use of statistical tools and may lead to the creation of a strong prediction model with a statistically significant level of bankruptcy prediction.

3. Data and Methodology

To compare the prediction accuracy of the models in the conditions of Slovak enterprises, we created logistic regression (LR), random forests (RF) and neural networks (NN) models to predict whether a company will be in financial crisis in the following year. Each of the models has both advantages and disadvantages, and our goal was to find the most suitable model [89–92].

3.1. Data and Conditions of Classification

The data were obtained from the register of the financial statements (www.registeruz.sk) using API (API is the acronym for Application Programming Interface, which allows an application to communicate with another application or an operating system) and C# (C-Sharp is a programming language). Financial statements from 2016, 2017 and 2018 were analyzed. To calculate the predictors for the year 2016, the financial data of 2015 were analyzed. To determine if an enterprise is in the financial crisis in 2016, the data of 2016 were used. The same procedure was used for the statements from the years 2017 and 2018. To distinguish the enterprises into the groups of financially sound enterprises and enterprises in financial crisis, the legislative criteria were applied. Financial reports of enterprises do not include the information about the number of creditors and payment delays; thus a different procedure was used to identify an enterprise in crisis and the following criteria were set: (i) the equity-to-liability ratio does not exceed the given value of 0.4; (ii) the current ratio is less than 1; and (iii) earnings after taxes are not positive. If an enterprise meets all three conditions, it is classified as an enterprise in financial crisis. These criteria treat potential and real indebtedness of a company (due to the accumulation of losses from previous years, which would indicate: (i) an inability to generate profit in the longer term; (ii) insolvency; and (iii) the current inability to make a profit).

The data of 2016 were divided in the ratio 75:25 for training and validation samples. Both parts preserve the same proportion of companies in the response class as the original data. All the financial

predictors from 2016 and 2017 were used to test the model. The comparison of training data and testing data is summarized in Table 1.

Table 1. Data classification.

Data	Total	Not in Crisis	In Crisis	Ratio
2016	47,414	41,266	6148	12.97%
2016 training	35,561	30,950	4611	12.97%
2016 validation	11,853	10,316	1537	12.97%
2017	64,757	54,835	9922	15.32%
2018	56,743	49,016	7727	13.62%

Financial stability of enterprises is also influenced by the development of the optimal values of crucial financial indicators. We decided to choose the most relevant financial ratios [90] of profitability, liquidity, activity, indebtedness and capital structure (Table 2). Measuring and assessing the financial ratios of profitability, activity, liquidity and indebtedness help create a competitive advantage for an enterprise [93]. However, symptoms of financial distress never occur at the same time but in certain phases. First, there is a decrease in output volume, a decrease in profitability, an increased need for working capital and a deterioration of the capital structure, and finally, it comes to persistent insolvency.

Table 2. Selected financial ratios and their algorithms.

	Profitability ratios (P)	**Algorithm**
R1	Return on capital (net)	EAT / total liabilities
R2	Return on capital (gross)	(EBIT + cost interests) / total liabilities
R3	Return on corporate revenues (net)	EAT / revenues
	Activity ratios (A)	**Algorithm**
A1	Asset turnover	Revenues / total assets
A2	Current assets turnover	Revenues / current assets
	Liquidity ratios (L)	**Algorithm**
L1	Cash ratio	Cash and cash equivalents / current liabilities
L2	Quick ratio	(Cash and cash equivalents + account receivables) / current liabilities
L3	Current ratio	Current assets / current liabilities
L4	Net working capital ratio	Net working capital / total assets
	Ratios of indebtedness and capital structure (Z)	**Algorithm**
Z1	RE–TA ratio	Retained earnings / total assets
Z2	Debt ratio	(Current + non-current liabilities) / total assets
Z3	Current debt ratio	Current liabilities / total assets
Z4	Financial debt ratio	(Bank loans + issued bonds) / total assets
Z5	Debt–equity ratio	(Current + non-current liabilities) / equity

Note: EAT—earnings after taxes; EBIT—earnings before interest and taxes; RE—retained earnings; TA—total assets

Summary statistics of these predictors are shown in Table 3.

Table 3. Summary statistics.

	L1	L2	L3	L4	R1	R2	R3	A1	A2	Z1	Z2	Z3	Z4	Z5
St. dev.	1.69	2.32	2.48	0.58	0.23	0.25	0.25	1.94	0.84	0.51	0.52	0.51	0.14	2.50
Var.	2.85	5.39	6.15	0.34	0.05	0.06	0.06	3.77	0.71	0.26	0.27	0.26	0.02	6.27
Min	0.00	0.00	0.00	−4.87	−1.61	−1.61	−1.87	0.00	0.00	−3.76	0.05	0.00	0.00	−0.82
1stQ	0.06	0.43	0.69	−0.22	−0.02	−0.02	−0.02	0.39	0.25	−0.08	0.44	0.32	0.00	0.05
Mean	0.87	1.64	1.94	0.02	0.02	0.04	0.00	1.27	0.65	−0.05	0.76	0.67	0.07	1.26
Median	0.24	0.91	1.12	0.07	0.02	0.03	0.01	0.66	0.41	0.01	0.73	0.61	0.00	0.37
3rdQ	0.84	1.71	2.03	0.38	0.09	0.12	0.06	1.22	0.71	0.16	0.95	0.89	0.07	1.30
Max	13.02	22.19	23.68	2.14	1.63	1.76	1.81	19.10	9.62	1.61	5.42	5.40	0.72	20.96

3.2. Methods Used for Bankruptcy Prediction

As we mentioned in the literature review, several methods may be used in the bankruptcy prediction models, but to analyze the industrial enterprises in Slovak conditions we selected three popular statistical methods—logistic regression, random forest and neural networks—which were proven to be the most accurate in other researches and studies.

Logistic regression is a method which tries to model the unilateral dependence of variables from which the examined dependent variables are binary, ordinal or categorical, and the explanatory variables can be of any type. It is suitable for modeling of the unilateral dependence between variables in a situation where the dependent variable is categorical, and the explanatory variables may be continuous or categorical. Logistic regression is often compared to multiple discriminant analysis; however, its fundamental restrictions are not so strict, e.g., it does not require the assumptions of normality of variables or homoscedasticity of individual groups. Additionally, the classification ability tends to be better than in the case of models based on discriminant analysis [94]. When modeling the financial distress of enterprises using logistic regression, two categories are recognized: prosperous and non-prosperous enterprises. Each enterprise belongs only to one category depending on the value of the dependent variable. The modeling of prosperity/ non-prosperity (conditional probability) is based on the conditional probability of the dependent variable (Y) depending on the independent variables, predictors (X). All used predictors should be independent of each other, as the existence of mutual dependence (multicollinearity) can affect the stability of the model [95]. The relationship between the probability and the vector of independent variables for non-prosperous enterprise is calculated using the following algorithm:

$$\pi = \frac{e^{\beta_0+\beta_1 X_1+\ldots+\beta_k X_k}}{1+e^{\beta_0+\beta_1 X_1+\ldots+\beta_k X_k}} = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\ldots+\beta_k X_k)}} \quad (1)$$

where π is the conditional probability that an enterprise is non-prosperous. Thus, the logistic regression assumes that an enterprise is non-prosperous if the predicted probability is greater than the limit value (most often it is the value 0.5), and vice versa, if the predicted value is below the determined limit values, an enterprise belongs to the group of prosperous enterprises.

The technique of *random forest* was developed for datasets containing a large number of predictors. Random forests can combine multiple categorical and numeric variables in one analysis [96]. This method consists of a set of simple trees $T_1, \ldots, T_N$ whose classification or regression function can be expressed as $h(X, \Theta_1), \ldots, h(X, \Theta_N)$, where h is a function, X is a predictor and $\Theta_1, \ldots, \Theta_N$ are independent, equally distributed random vectors. For a random forest method, CandRT binary trees are used. Similarly to the formation of individual trees, using the RF method, the dataset is split into test and training files. Training files for individual trees T_i are bootstrap selections from the L data files. Bootstrap selections are random repetitions of n size. Observations that are in the i-th bootstrap selection L_i are used to create the T_i tree (training set), and the observations that were not selected (test set) are used to estimate the error. Error estimates on a test set are called out-of-bag estimates. The total number of out-of-bag estimates is 1/3 of the data set. When using RF to classify enterprises, we get information from each tree about the classification of each observation into the resulting category. The result of forest classification is given by the majority decision of all the trees. The RF method increases the accuracy (reduces distortion) by letting trees grow, while maintaining a bearable variation by combining the results of individual trees (majority vote/averaging). Compared to other forests, however, there is an effort to ensure a low level of correlation between individual trees. Decreasing the correlation between the trees is achieved by a random selection of a certain number of predictors. The random forest method uses a random selection of observations and a random selection of predictors [97].

The *neural network* is a set of connected input–output units (artificial neurons), with each connection having a certain weight. Artificial neurons are based on the principle of biological neurons that make

up the human nervous system. The input information is weighted. The threshold value is subtracted from the sum of weighted input signals and the activation function transforms the signal into an output signal that is sent to the input of the neurons to which the given neuron is associated. There are several types of neural networks and algorithms. The most commonly used type of neural network, which was used to analyze the Slovak industrial enterprises, is a multilayer pre-implemented neural network. It consists of several layers of neurons—the input layer, several hidden layers and the output layer [98]. This type of neural networks is used for the classification and prediction of a continuous function (numerical prediction). Provided that they have enough hidden layers and training examples, they can approximate any function. The use of a neural network for bankruptcy prediction is recommended as this method is tolerant of data noise and the ability to model complex relationships between inputs and outputs. Algorithms of neural networks can be parallelized, which reduces the calculation time [99].

3.3. Metrics of the Prediction Models Comparison

The quality of the prediction model can be quantified by several measures. In this study, different types of models are compared using multiple metrics.

The *accuracy* is the ratio between the number of correct predictions and the total number of predictions. If the number of occurrences in classes varies greatly, accuracy is biased. The *per class accuracy* is the average accuracy for each class. It should be used when the classes are imbalanced.

The *error rate* is calculated as *1 − Accuracy*.

The *mean per class error* is an average of the error rate for each class.

The coefficient of determination (R-squared) is calculated as:

$$R^2 = 1 - \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 / \sum_{i=1}^{N} (y_i - \overline{y})^2 \tag{2}$$

where $y_i, \overline{y}, \hat{y}_i$ are the original data values, mean and predicted values, respectively. R-squared measures the percentage of variability of the dependent variable, which is explained by independent variables in the model.

The *logarithmic loss* (LogLoss) measures the uncertainty of the probabilities (p) of a model by comparing them with true labels. To calculate log loss the following algorithm is used:

$$logLoss = \frac{-1}{N} \sum_{i=1}^{N} (y_i \cdot \log(p_i) + (1 - y_i) \cdot log(1 - p_i)) \tag{3}$$

Thus, log loss quantifies the accuracy of a model by penalizing false classifications.

The *receiver operating characteristic* (ROC) is a curve with points [x,y], where x = 100 − Specificity and y = Sensitivity for different cut-off (threshold) points. The closer the ROC curve is to the upper left corner, the higher the overall accuracy of the model. The ROC curve reveals a trade-off between sensitivity and specificity—increasing of sensitivity implies decreasing of specificity and vice versa. The biggest benefit of using the ROC curve is that it is independent on the change in the proportion of outcomes.

The *area under curve* (AUC) is one of the most common and most frequently used metrics. An area under curve with a value in the range $\langle 0.97; 1.0 \rangle$ characterizes a perfect classification ability of a model. AUC values of $\langle 0.92; 0.97)$ present excellent results of prediction, $\langle 0.75; 0.92)$ good classification, $\langle 0.6; 0.75)$ acceptable classification ability and AUC bellow 0.6 indicates a model that is inappropriate for the prediction of a financial crisis [100]. Like other methods, a high AUC value does not necessarily guarantee the top quality of the model. For example, there are situations where the sensitivity is in the range of only a few hundred and the specificity is over 90%, and the AUC is still above 0.8.

The *Gini coefficient* is derived from the AUC and measures the inequality among the values of a frequency distribution. The Gini coefficient is calculated using the algorithm:

$$Gini = 2AUC - 1 \tag{4}$$

A Gini coefficient of more than 60% shows a good predictive ability of the model.

The *mean squared error* (MSE) measures the average squared difference between the estimated values and the actual values:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{5}$$

where y_i is the vector of observed values of the variable being predicted and $\hat{y}_i$ is the predicted value. MSE is highly affected by outlier values [101].

The *root mean squared error* (RMSE) is defined as the square root of the MSE. It aggregates the magnitudes of the errors in predictions and measures the accuracy, which is used to compare forecasting errors of different models for a particular dataset [102].

4. Results and Discussion

All the selected enterprises are considered using the logistic regression, random forests and neural networks models, which are assessed by the determined metrics. The result is the determination of the most accurate and relevant prediction model in the conditions of Slovak industrial enterprises, enabling the prediction of financial distress in the upcoming period.

We started the analysis with fourteen predictors; however, in the process of the models´ development, some of the predictors were removed. Models were created and tested in R with the H2O package. The H2O package allows setting the maximum number of predictors used when creating the LR model [103]. When developing the model, all predictors were used and their maximal allowed number was changed (max active predictors option). When switching from five to six predictors it was proven that the use of six or more predictors does not improve the resulting model (compared to increasing complexity; thus using a simpler model is a better model). However, max active predictors cannot be used in RF and NN models, though the process of the model development was similar but more time-consuming. Table 4 shows the final use of predictor variables in each considered model.

Table 4. Predictors used in models.

		LR	RF	NN
R1	ROAeat	+	+	+
R2	ROAebit	+	+	−
R3	Net profit margin	−	+	−
L1	Cash ratio	−	−	+
L2	Quick ratio	+	+	+
L3	Current ratio	+	+	+
L4	Net working capital / Total assets	−	+	+
Z1	Retained earnings /Total assets	−	−	+
Z2	Debt Ratio	−	+	+
Z3	Current liability / Total assets	−	+	−
Z4	Credit indebtedness	−	−	−
Z5	Equity / Total liabilities	+	+	+
A1	Total Asset Turnover	−	−	−
A2	Current Asset Turnover	−	−	−

Note: LR—logistic regression; NN—neural network; RF—random forest; ROA—return on assets.

A plus sign indicates that a financial predictor was used in the final model, and vice versa, predictors with a minus sign were not included. The combination of plus and minus signs indicate that for each model different financial predictors are important.

In Table 5, confusion matrices for the models are portrayed; real yes/no information about the corporate financial crisis is on the left side of the table, and predicted yes/no information is on the top. Max minimum per class accuracy specifies the threshold at which the class accuracy is the worst.

Table 5. Confusion matrices for maximal min_per_class accuracy threshold.

		2017				2018		
		no	**yes**	**error**		**no**	**yes**	**error**
LR	**no**	44,716	10,119	0.185	**no**	40,024	8992	0.183
	yes	1852	8070	0.187	**yes**	1438	6289	0.186
	totals	46,568	18,189	0.185	**totals**	41,462	15,281	0.184
NN	**no**	44,874	9961	0.182	**no**	40,202	8814	0.180
	yes	1798	8124	0.181	**yes**	1397	6330	0.181
	totals	46,672	18,085	0.182	**totals**	41,599	15,144	0.180
RF	**no**	44,817	10,018	0.183	**no**	40,161	8855	0.181
	yes	1804	8118	0.182	**yes**	1399	6328	0.181
	totals	46,621	18,136	0.183	**totals**	41,560	15,183	0.181

Note: LR—logistic regression, NN—neural network, RF—random forest.

Performance characteristics are shown in Table 6. The symbols ↑ (↓) indicate that the larger (smaller) number is better.

Table 6. Performance characteristics.

	2017				2018		
	NN	**LR**	**RF**		**NN**	**LR**	**RF**
MSE	0.088	0.096	0.090	↓	0.081	0.087	0.082
RMSE	0.297	0.309	0.300	↓	0.284	0.296	0.287
Mean per class error	0.205	0.212	0.199	↓	0.211	0.208	0.207
LogLoss	0.331	0.405	0.339	↓	0.294	0.345	0.309
R^2	0.322	0.263	0.306	↑	0.315	0.257	0.300
Gini	0.758	0.742	0.755	↑	0.772	0.755	0.764
AUC	0.879	0.871	0.877	↑	0.886	0.877	0.882

Note: MSE—mean squared error; RMSE—root mean squared error; AUC—area under curve.

The comparison of the methods in 2017 reveals that better results were achieved by the new learning algorithms (NN and RF) following all selected metrics. Choosing the best method in conditions of the prediction accuracy, the neural network model has the best predictive ability measured almost by all performance characteristics (except for mean per class error).

The situation is similar in 2018, with slightly different results. The NN and RF model financial distress more accurately following the results of almost all performance characteristics. Neural networks outperform the other two methods in almost all analyzed metrics except for mean per class error, where the best result was achieved by RF (followed by LR).

Comparing the results of the neural network and logistic regression models in the analyzed period, NN models show 2%–22% better results (depending on the performance metrics); the biggest differences are in the results of LogLoss and R^2. These findings correspondent with the results of Barboza et al. [71], who conducted intensive research evaluating bankruptcy using traditional statistic techniques and early artificial intelligence models and found that machine learning models show 10% better accuracy in relation to tradition models (LR and MDA).

The difference between both learning algorithms, NN and RF, does not usually exceed 5%, and the differences in the metrics achieved are even smaller (about 2%). Naidu and Govinda [104] affirm that the use of artificial neural networks and random forest have proven to be more efficient over the traditional algorithms. Thus, the neural network models have the advantage of being able to detect non-linear relationships and show better performance, describing the blatant information in corporate failure prediction problems [105].

The results portrayed in Tables 5 and 6 indicate that the neural network model is the best, the random forest model is the second and the logistic regression is the third in order when measuring their strength by selected statistical performance characteristics. Additionally, the models work more accurately with data from 2018 than with the data from 2017, even though they are built on the data from 2016. The most likely reason is that the data from 2016 and 2018 have more similar ratios of the responding variables than data from 2016 and 2017.

Our findings correspond with the results of Lee et al. [106], who analyzed the performance of discriminant analysis, logistic regression and neural networks in the context of Korean enterprises and confirmed the importance of neural networks in predicting bankruptcy. Bagheri et al. [107] affirm, based on the dataset of 80 Tehran enterprises, that artificial neural networks have higher accuracy than logistic regression models used in the bankruptcy prediction. The comparative analysis of logit and probit models, random forests and artificial neural networks by Karminsky and Burekhin [108] revealed that neural networks outperform other methods in predictive power measured by Gini and AUC coefficients. As this study was conducted on a sample of Russian industrial companies, the confirmation of our findings by this study is crucial. Moreover, the authors added that there is no significant impact of non-financial indicators on the probability of bankruptcy. Chaudhuri and De [109] stress that artificial neural networks have become a dominant modeling paradigm. Their study of the 50 largest bankrupt organizations with capitalization of no less than $1 billion underlines the relevance of neural network models used for bankruptcy prediction. However, they claim that the choice of appropriate parameters plays an inevitable role in the performance of the model.

The results of our analysis are contrary to the study of Chen [110], who reported that traditional statistical methods are more relevant to handle large samples without sacrificing prediction performance, while learning algorithms achieve better predictive ability with a smaller dataset. However, our dataset of more than 50,000 enterprises shows opposite results.

Each model has its advantages and all models have the potential to be used in practice as decision support tools. Neural network models provide better results but logistic regression is more convenient to be used in practice. It is noteworthy that models have preferred slightly different predictors. It results from the fact that models have different abilities to detect relationships between predictors and outcome, as well as interactions among predictor variables.

5. Conclusions

The vast majority of enterprises accept that their lifetime is unlimited and will bring continuous benefits to their owners, creditors and stakeholders in the form of profit, rising market value of their enterprise and growing or, at least, non-declining number of employees. However, due to the entropy and turbulence of the economic and political environment in which companies are interacting, it may come to a loss of key customers, change of the crucial macroeconomic fundamentals, reduction of expected returns, increase of costs or the emergence of new, unexpected expenses. The cardinal question is if the financial distress of enterprises can be predicted with sufficient time in advance and with appropriate accuracy. This important phenomenon is solved by the models of bankruptcy prediction. The models differ in several aspects following the economic and legislative principles of the country in which they were formed. Moreover, they use various financial ratios as potential predictors of financial distress. The most important role is played by the model's statistical principle, which is used to predict the financial crisis of enterprises.

Until recently, the dominating bankruptcy prediction methods were based on statistical modeling; the most frequently used were multiple discriminant analysis and logistic regression models, having much better prediction accuracy. However, lately, models based on machine learning have been proposed and have been successfully used for many classification and regression problems. Moreover, machine learning models often outperform traditional classification methods. The purpose of bankruptcy prediction is to reveal the future financial development and perspective of enterprises. Therefore in this study, the comparison of traditional (logistic regression) and new learning algorithm methods (random forest and neural network) was conducted to reveal their prediction ability and accuracy in the condition of Slovak industrial enterprises. Comparing the methods on a scale of different metrics, the new machine learning models show higher predictive performance; particularly neural network model yielded better results measured by all performance characteristics. The accurate prediction of corporate bankruptcy for enterprises operating in specific industries is crucial for creditors and stakeholders as the reduction of potential risk. The results of Lee and Choi [111] declare that prediction using industry samples outperforms using the entire sample of enterprises and the best predictive accuracy is achieved by the neural network model. Our results underline the importance of the neural network model for the bankruptcy prediction and highlight its relevance to assess the financial distress of industrial enterprises.

Identification of the most relevant and accurate method is useful to form the model predicting the financial distress of industrial enterprises in the specific national environment of Slovakia, which has not been developed yet. Despite the huge extent of performance characteristics comparing the models, several more methods used for bankruptcy prediction can be included in the comparison, e.g., multiple discriminant analysis, probit regression, rough sets, linear programming, principal component analysis, data envelopment analysis and survival analysis. This limitation can be omitted by further research also including other models, not only the most frequently used, and investigating the prediction accuracy of the models in a longer time horizon.

Author Contributions: Conceptualization, E.G. and K.V.; methodology, E.G. and P.A.; software, P.A.; validation K.V., P.A. and J.J.; formal analysis, M.T.; investigation, P.A. and K.V.; resources, J.J.; writing—original draft preparation, P.A. and K.V.; writing—review and editing, K.V. and M.T.; visualization, J.J. and E.G.; supervision, M.T.; project administration, J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the Slovak Research and Development Agency—Grant No. APVV-14-0841: Comprehensive Prediction Model of the Financial Health of Slovak Companies and Vega 1/0121/20: Research of transfer pricing system as a tool to measure the performance of national and multinational companies in the context of earnings management in conditions of the Slovak Republic and V4 countries.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

References

1. Fitzpatrick, P. A comparison of ratios of successful industrial enterprises with those of failed firms. *Certif. Public Account.* **1932**, *2*, 598–605.
2. Alaka, H.A.; Oyedele, L.-O.; Owolabi, H.A.; Kumar, V.; Ajayi, S.O.; Akinade, O.O.; Bilal, M. Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Syst. Appl.* **2018**, *94*, 164–184. [CrossRef]
3. Kubickova, D.; Nulicek, V. Predictors of financial distress and bankruptcy model construction. *Int. J. Manag. Sci. Bus. Adm.* **2016**, *2*, 34–41.
4. Karas, M.; Reznakova, M. Creating a new bankruptcy prediction model: The grey zone problem. In Proceedings of the 24th IBIMA Conference: Crafting Global Competitive Economies: 2020 Vision Strategic Planning & Smart Implementation, Milan, Italy, 6–7 November 2014; Soliman, K.S., Ed.; International Business Information Management Assoc—IBIMA: Norristown, PA, USA, 2014; pp. 911–919.
5. Gavurova, B.; Packova, M.; Misankova, M.; Smrcka, L. Predictive potential and risks of selected bankruptcy prediction models in the Slovak business environment. *J. Bus. Econ. Manag.* **2017**, *18*, 1156–1173. [CrossRef]

6. Virag, M.; Hajdu, O. Pénzügyi mutatószámokon alapuló csõdmodellszámítások. *Bankszemle XV* **1996**, *5*, 42–53.
7. Neumaier, I.; Neumaierova, I. Try to count your Index IN 95. *Terno* **1995**, *5*, 7–10.
8. Neumaier, I.; Neumaierova, I. INFA Financial analysis—Application in energy sector. *Sekt. A Odvetv. Anal. Asp. Energ.* **1999**, *4*, 32–75.
9. Neumaier, I.; Neumaierova, I. Analysis of the value creation—Application of INFA financial analysis. *Sekt. A Odvetv. Anal. Asp. Invest. Strojir.* **2001**, *8*, 23–39.
10. Neumaier, I.; Neumaierova, I. Index IN 05. In Proceedings of the Conference European Financial Systems, Brno, Czech Republic, 21–23 June 2005; Masaryk University: Brno, Czech Republic, 2005; pp. 143–148.
11. Maczynska, E. Assessment of the conditions of the enterprise. Simplified methods. *Zycie Gospod.* **1994**, *38*, 42–45.
12. Gajdka, J.; Stos, D. *The Use of Discriminant Analysis in Assessing the Financial Condition of Enterprises*; Wydawnictvo Akademii Ekonomicznej v Krakowie: Krakow, Poland, 1996.
13. Hamrol, M.; Czajka, B.; Piechocki, M. Company bankruptcy—A discriminant analysis model. *Przegląd Organ.* **2004**, *6*, 35–39. [CrossRef]
14. Prusak, B. *Nowoczesne Metody Prognozowania Zagrozenia Finansowego Predsiebiorst*; DiFin: Krakow, Poland, 2005.
15. Gruszczynski, M. Financial distress of companies in Poland. *Int. Adv. Econ. Res.* **2004**, *10*, 249–256. [CrossRef]
16. Chrastinova, Z. *Methods of Assessment of Economic Solvency and Prediction of Financial Situation of Agricultural Enterprises*; VUEPP: Bratislava, Slovakia, 1998.
17. Gurcik, L. G-index—The financial situation prognosis method of agricultural enterprises. *Agric. Econ.* **2002**, *48*, 373–378.
18. SARIO. *Automotive Sector in Slovakia*; Slovak Investment and Trade Development Agency: Bratislava, Slovakia, 2020.
19. Binkert, C.H. Fruherkennung von Unternehmenskrisen mit Hilfe Geeigneter Methoden im deutschen und Slowakischen Wirtschaftsraum. Ph.D. Thesis, University of Economics in Bratislava, Bratislava, Slovakia, 1999.
20. Hurtosova, J. Development of Rating Model as a Tool to Assess the Enterprise Credibility. Ph.D. Thesis, University of Economics in Bratislava, Bratislava, Slovakia, 2009.
21. Delina, R.; Packova, M. Prediction bankruptcy models validation in Slovak business environment. *Ekon. Manag.* **2013**, *16*, 101–112.
22. Rohacova, V.; Kral, P. Corporate failure prediction using DEA: An application to companies in the Slovak republic. In Proceedings of the Applications of Mathematics and Statistics in Economics, Jindrichuv Hradec, Czech Republic, 2–6 September 2015; University of Economics: Prague, Czech Republic, 2015; pp. 1–8.
23. Gulka, M. The prediction model of financial distress of enterprises operating in conditions of SR. *Biatec* **2016**, *24*, 5–10.
24. Boda, M.; Uradnicek, V. Inclusion of weights and their uncertainty into quantification within a pyramid decomposition of a financial indicator. *Ekon. Cas.* **2016**, *64*, 70–92.
25. Svabova, L.; Durica, M. Being an outlier: A company non-prosperity sign? *Equilib.—Q. J. Econ. Econ. Policy* **2019**, *14*, 359–375. [CrossRef]
26. Belas, J.; Smrcka, L.; Gavurova, B.; Dvorsky, J. The impact of social and economic factors in the credit risk management of SME. *Technol. Econ. Dev. Econ.* **2018**, *24*, 1215–1230. [CrossRef]
27. Budiarto, D.S.; Rahmawati, B.; Prabowo, M.A. Accounting information system and non-financial performance in small firm: Empirical research based on ethnicity. *J. Int. Stud.* **2019**, *12*, 338–351. [CrossRef]
28. Bartosova, V.; Kral, P. Methodological framework of financial analysis results objectification in Slovak republic. *J. Mod. Account. Audit.* **2017**, *13*, 394–400.
29. Toth, Z.; Mura, L. Support for small and medium enterprises in the economic crisis in selected EU countries. In Proceedings of the 12th International Conference on Hradec Economic Days: Economic Development and Management of Regions, Hradec Kralove, Czech Republic, 4–5 February 2014.
30. Beaver, W.H. Financial ratios as predictors of failure. *J. Account. Res.* **1966**, *4*, 71–111. [CrossRef]
31. Altman, E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Financ.* **1968**, *23*, 589–609. [CrossRef]
32. Ohlson, J.A. Financial ratios and the probabilistic prediction of bankruptcy. *J. Account. Res.* **1980**, *18*, 109–131. [CrossRef]

33. Prusak, B. Review of research into enterprise bankruptcy prediction in selected central and European countries. *Int. J. Financ. Stud.* **2018**, *6*, 60. [CrossRef]
34. Kliestik, T.; Valaskova, K.; Kliestikova, J.; Kovacova, M.; Svabova, L. *Prediction of Financial Health of Enterprises in Transition Economies*; EDIS: Zilina, Slovakia, 2019.
35. Antunes, F.; Ribiero, B.; Pereira, F. Probabilistic modeling and visualization for bankruptcy prediction. *Appl. Soft Comput.* **2017**, *60*, 831–843. [CrossRef]
36. Stefko, R.; Gavurova, B.; Rigelsky, M.; Ivankova, V. Evaluation of selected indicators of patient satisfaction and economic indices in OECD country. *Econ. Sociol.* **2019**, *12*, 149–165. [CrossRef] [PubMed]
37. Kliestik, T.; Misankova, M.; Valaskova, K.; Svabova, L. Bankruptcy prevention: New effort to reflect on legal and social changes. *Sci. Eng. Ethics* **2018**, *24*, 791–803. [CrossRef]
38. Chou, C.; Hsieh, S.; Qiu, C. Hybrid genetic algorithm and fuzzy clustering for bankruptcy prediction. *Appl. Soft Comput.* **2017**, *56*, 298–316. [CrossRef]
39. Sharifabadi, M.R.; Mirhaj, M.; Izadinia, N. The impact of financial ratios on the prediction of bankruptcy of small and medium companies. *Quid* **2017**, *1*, 164–173.
40. Tian, S.; Yu, Y.; Guo, H. Variable selection and corporate bankruptcy forecasts. *J. Bank. Financ.* **2015**, *52*, 89–100. [CrossRef]
41. Bellovary, J.; Giacomino, D.; Akers, M. A review of bankruptcy prediction studies: 1930 to present. *J. Financ. Educ.* **2007**, *33*, 1–43.
42. Kumar, P.P.; Ravi, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *Eur. J. Oper. Res.* **2007**, *180*, 1–28. [CrossRef]
43. Calderon, T.G.; Cheh, J.J. A roadmap for future neural networks research in auditing and risk assessment. *Int. J. Account. Inf. Syst.* **2002**, *3*, 203–236. [CrossRef]
44. Dimitras, A.I.; Zanakis, S.H.; Zopoundis, C. A survey of business failure with an emphasis on prediction method and industrial applications. *Eur. J. Oper. Res.* **1996**, *90*, 487–513. [CrossRef]
45. O'Leary, D.E. Using neural network to predict corporate failure. *Int. J. Intell. Syst. Account. Financ. Manag.* **1998**, *7*, 187–197. [CrossRef]
46. Scott, J. The probability of bankruptcy: A comparison of empirical predictions and theoretical models. *J. Bank. Financ.* **1981**, *5*, 317–344. [CrossRef]
47. Kovacova, M.; Kliestik, T.; Valaskova, K.; Durana, P.; Juhaszova, Z. Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. *Oecon. Copernic.* **2019**, *10*, 743–772. [CrossRef]
48. Jones, S. Corporate bankruptcy prediction: A high dimensional analysis. *Rev. Account. Stud.* **2017**, *22*, 1366–1422. [CrossRef]
49. Jacobson, T.; Linde, J.; Roszbach, K. Firm default and aggregate fluctuations. *J. Eur. Econ. Assoc.* **2013**, *11*, 945–972. [CrossRef]
50. Bruneau, C.; De Bandt, O.; El Amri, W. Macroeconomic fluctuations and corporate financial fragility. *J. Financ. Stab.* **2012**, *8*, 219–235. [CrossRef]
51. Nam, C.W.; Kim, T.S.; Park, N.J.; Lee, H.K. Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *J. Forecast.* **2008**, *27*, 493–506. [CrossRef]
52. Tomas Zikovic, I. Challenges in predicting financial distress in emerging economies: The case of Croatia. *East. Eur. Econ.* **2018**, *56*, 1–27. [CrossRef]
53. Tinoco, M.H.; Wilson, N. Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *Int. Rev. Financ. Anal.* **2013**, *30*, 394–419. [CrossRef]
54. Giriuniene, G.; Giriunas, L.; Morkunas, M.; Brucaite, L. A comparison on leading methodologies for bankruptcy prediction: The case of the construction sector in Lithuania. *Economies* **2019**, *7*, 82. [CrossRef]
55. Filipe, S.F.; Grammatikos, T.; Michala, D. Forecasting distress in European SME portfolios. *J. Bank. Financ.* **2016**, *64*, 112–135. [CrossRef]
56. Kacer, M.; Ochotnicky, P.; Alexy, M. The Altman's revised Z'-Score model, non-financial information and macroeconomic variables: Case of Slovak SMEs. *Ekon. Cas.* **2019**, *67*, 335–366.
57. Wilson, N.; Ochotnicky, P.; Kacer, M. Creation and destruction in transition economies: The SME sector in Slovakia. *Int. Small Bus. J.—Res. Entrep.* **2016**, *34*, 579–600. [CrossRef]
58. Du Jardin, P. Dynamics of firm financial evolution and bankruptcy prediction. *Expert Syst. Appl.* **2017**, *75*, 25–43. [CrossRef]

59. Tuffnell, C.; Kral, P.; Siekelova, A.; Horak, J. Cyber-physical smart manufacturing systems: Sustainable industrial networks, cognitive automation, and data-centric business models. *Econ. Manag. Financ. Mark.* **2019**, *14*, 58–63.
60. Mattsson, B.; Steinert, O. Corporate Bankruptcy Prediction Using Machine Learning Techniques. Bachelor's Thesis, University of Gothenburg, Gothenburg, Sweden, 2017.
61. Barbuta-Misu, N.; Madaleno, M. Assessment of bankruptcy risk of large companies: European countries evolution analysis. *J. Risk Financ. Manag.* **2020**, *13*, 58. [CrossRef]
62. Pisula, T. An ensemble classifier-based scoring model for predicting bankruptcy of Polish companies in the Podkapackie Voivodeship. *J. Risk Financ. Manag.* **2020**, *13*, 37. [CrossRef]
63. Oliveira, M.D.N.T.; Ferriera, F.A.F.; Peréz-Bustamante Ilander, B.O.; Jalali, M.S. Integrating cognitive mapping and MDCA for bankruptcy prediction in small-and medium-sized enterprises. *J. Oper. Res. Soc.* **2017**, *68*, 985–997. [CrossRef]
64. Tsai, C. Feature selection in bankruptcy prediction. *Knowl. Based Syst.* **2009**, *22*, 120–127. [CrossRef]
65. Le, T.; Le, H.S.; Vo, M.T.; Lee, M.Y.; Baik, S.W. A cluster-based boosting algorithm for bankruptcy prediction in a highly imbalanced dataset. *Symmetry* **2018**, *10*, 250. [CrossRef]
66. Le, T.; Lee, M.Y.; Park, J.R.; Baik, S.W. Oversampling technique for bankruptcy prediction: Novel features from a transaction dataset. *Symmetry* **2018**, *10*, 79. [CrossRef]
67. Le, T.; Vo, B.; Fujita, H.; Nguyen, N.; Baik, W.S. A fast and accurate approach for bankruptcy forecasting using squared logistics loss with GPU-based extreme gradient boosting. *Inf. Sci.* **2019**, *494*, 294–310. [CrossRef]
68. Wang, M.; Chen, H.; Li, H.; Cai, Z.; Zhao, X.; Tong, C.; Li, J.; Xu, X. Grey wolf optimization evolving kernel extreme learning machine: Application to bankruptcy prediction. *Eng. Appl. Artif. Intell.* **2017**, *63*, 54–68. [CrossRef]
69. Mai, F.; Shaonan, T.; Chihoon, L.; Ling, M. Deep learning models for bankruptcy prediction using textile disclosures. *Eur. J. Oper. Res.* **2019**, *274*, 743–758. [CrossRef]
70. Hosaka, T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Syst. Appl.* **2019**, *13*, 287–299. [CrossRef]
71. Qu, Y.; Quan, P.; Lei, M.; Shi, Y. Review of bankruptcy prediction using machine learning and deep learning techniques. *Procedia Comput. Sci.* **2019**, *162*, 895–899. [CrossRef]
72. Kovacova, M.; Kliestik, T. Logit and probit application for the prediction of bankruptcy in Slovak companies. *Equilib. Q. J. Econ. Econ. Policy* **2017**, *12*, 775–791. [CrossRef]
73. Affes, Z.; Hentati-Kaffel, R. Predicting US banks bankruptcy: Logit versus canonical discriminant analysis. *Comput. Econ.* **2019**, *54*, 199–244. [CrossRef]
74. Barboza, F.; Kimura, H.; Altman, E. Machine learning models and bankruptcy prediction. *Expert Syst. Appl.* **2017**, *83*, 405–417. [CrossRef]
75. Mihalovic, M. Performance comparison of multiple discriminant analysis and logit models in bankruptcy prediction. *Econ. Sociol.* **2016**, *9*, 101–118. [CrossRef] [PubMed]
76. Cho, S.; Kim, J.; Bae, J.K. An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Syst. Appl.* **2009**, *36*, 403–410. [CrossRef]
77. Van Gestel, T.; Baesens, B.; Martens, D. From linear to non-linear kernel based classifiers for bankruptcy prediction. *Neurocomputing* **2010**, *73*, 2955–2970. [CrossRef]
78. Kim, S.Y. Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *Serv. Ind. J.* **2011**, *31*, 441–468. [CrossRef]
79. Chen, M.Y. Comparing traditional statistics, decision tree classification and support vector machine technique for financial bankruptcy prediction. *Intell. Autom. Soft Comput.* **2012**, *18*, 65–73. [CrossRef]
80. Nyitrai, T.; Virag, M. The effect of handling outliers on the performance of bankruptcy prediction models. *Socio-Econ. Plan. Sci.* **2019**, *67*, 34–42. [CrossRef]
81. Altman, E.I.; Iwanicz-Drozdowska, M.; Laitinen, E.K.; Suvas, A. A race for long horizon bankruptcy prediction. *Appl. Econ.* **2020**. early access. [CrossRef]
82. Ben Jabeur, S. Bankruptcy prediction using partial least squares logistic regression. *J. Retail. Consum. Serv.* **2017**, *36*, 197–202. [CrossRef]
83. Olson, D.L.; Delen, D.; Meng, Y. Comparative analysis of data mining methods for bankruptcy prediction. *Decis. Support Syst.* **2012**, *52*, 464–473. [CrossRef]

84. Klepac, V.; Hampel, D. Prediction of bankruptcy with SVM classifier among retail business companies in EU. *Acta Univ.* **2016**, *64*, 627–634.
85. Hudakova, M.; Masar, M.; Luskova, M.; Patak, M.R. The dependence of perceived business risks on the size of SMEs. *J. Compet.* **2018**, *10*, 54–69. [CrossRef]
86. Garcia, V.; Marques, A.I.; Sanchez, S.J. Exploring the synergetic effects of samples types in the performance of ensembles for credit risk and corporate bankruptcy prediction. *Inf. Fusion* **2019**, *47*, 88–101. [CrossRef]
87. Son, H.; Hyun, H.; Du Phan, H.; Hwang, H.J. Data analytical approach for bankruptcy prediction. *Expert Syst. Appl.* **2019**, *138*, 112816. [CrossRef]
88. Svabova, L.; Durica, M. A closer view of the statistical methods globally used in bankruptcy prediction of companies. In *Proceeding of the 16th International Scientific Conference on Globalization and its Socio Economic Consequences, Rajecke Teplice, Slovakia, 5–6 October 2016*; Kliestik, T., Ed.; University of Zilina: Zilina, Slovakia, 2016; pp. 2174–2181.
89. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
90. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
91. Nielsen, M. *Neural Networks and Deep Learning*; Determination Press: New Haven, CT, USA, 2015.
92. Svabova, L.; Kral, P. Selection of predictors in bankruptcy prediction models for Slovak enterprises. In Proceedings of the 10th International Days of Statistics and Economics, Prague, Czech Republic, 8–10 September 2016; Loster, T., Pavelka, T., Eds.; Melandrium: Slany, Czech Republic, 2016; pp. 1759–1768.
93. Eysenck, G.; Kovalova, E.; Machova, V.; Konecny, V. Big data analytics processes in industrial internet of things systems: Sensing and computing technologies, machine learning techniques, and autonomous decision-making algorithms. *J. Self-Gov. Manag. Econ.* **2019**, *7*, 28–34.
94. Kral, P.; Kanderova, M.; Kascakova, A.; Nedelova, G.; Valencakova, V. *Multivariate Statistical Methods Focused on the Solution of Problems of Economic Practice*; Matej Bel University: Banska Bystrica, Slovakia, 2009.
95. Das, S.; Chatterjee, S. Multicollinearity Problem—Root Cause, Diagnostics and Way Outs. SSRN Library. Available online: https://ssrn.com/abstract=1830043 (accessed on 11 March 2020).
96. Hafezi, M.H.; Liu, L.; Millward, H. Learning daily activity sequences of population groups using random forest theory. *Transp. Res. Rec.* **2018**, *47*, 194–207. [CrossRef]
97. Komprdova, K. *Decision Trees and Forests*; IBA: Brno, Czech Republic, 2012.
98. Choudhary, A.K.; Harding, J.A.; Tiwari, M.K. Data mining in manufacturing: A review based on the kind of knowledge. *J. Intell. Manuf.* **2009**, *20*, 501–521. [CrossRef]
99. Williams, G.J.; Simoff, S.J. *Data Mining—Theory, Methodology, Techniques and Applications*; Springer: Berlin, Germany, 2006.
100. Klepac, V.; Hampel, D. Predicting bankruptcy of manufacturing companies in EU. *Econ. Manag.* **2018**, *21*, 159–174.
101. Lehman, E.L.; Casella, G. *Theory of Point Estimation*; Springer: New York, NY, USA, 1998.
102. Hyndman, R.J.; Koehler, A.B. Another look at measured of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]
103. Bien, J.; Friedman, J.; Hastie, T.; Simon, N.; Taylor, J.; Tibshirani, R.; Tibshirani, R.J. Strong rules for discarding predictors in lasso-type problems. *J. R. Stat. Soc.* **2012**, *74*, 1–22.
104. Naidu, P.; Govinda, I. Bankruptcy prediction using neural networks. In Proceedings of the 2nd International Conference on Inventive Systems and Control, Coimbatora, India, 19–20 January 2018.
105. Alfaro, E.; Garcia, N.; Gamez, M.; Elizondo, D. Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decis. Support Syst.* **2008**, *45*, 110–122. [CrossRef]
106. Lee, K.; Booth, D.; Alam, P. A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Syst. Appl.* **2005**, *29*, 1–16. [CrossRef]
107. Bagheri, M.; Valipour, M.; Amin, V. The bankruptcy prediction in Tehran share holding using neural network and its comparison with logistic regression. *J. Math. Comput. Sci.* **2012**, *5*, 219–228. [CrossRef]
108. Karminsky, A.M.; Burekhin, R.N. Comparative analysis of methods for forecasting bankruptcies of Russian construction companies. *Bizn. Inform.* **2019**, *13*, 52–66. [CrossRef]
109. Chaudhuri, A.; De, K. Fuzzy support vector machine for bankruptcy prediction. *Appl. Soft Comput.* **2011**, *11*, 2472–2486. [CrossRef]

110. Chen, M.Y. Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Comput. Math. Appl.* **2011**, *62*, 4514–4524. [CrossRef]
111. Lee, S.; Choi, W.S. A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Syst. Appl.* **2013**, *40*, 2941–2946. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).