Business Failure Prediction for Slovak Small and Medium-Sized Companies

Lucia Svabova 1,*, Lucia Michalkova 1, Marek Durica 2 and Elvira Nica 3

1 Department of Economics, Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia; lucia.michalkova@fpedas.uniza.sk (L.M.)
2 Department of Quantitative Methods and Economic Informatics, Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia; marek.durica@fpedas.uniza.sk (M.D.)
3 Bucharest Academy of Economic Studies, Faculty of Administration and Public Management, 6 Piata Romana, 010374 Bucharest, Romania; popescu_elvira@yahoo.com (E.N.)
* Correspondence: lucia.svabova@fpedas.uniza.sk

Received: 3 May 2020; Accepted: 1 June 2020; Published: 3 June 2020

Abstract: Prediction of the financial difficulties of companies has been dealt with over the last years by scientists and economists worldwide. Several prediction models mostly focused on a particular sector of the national economy, have been created also in Slovakia. The main purpose of this paper is to create new prediction models for small and medium-sized companies in Slovakia, based on real data from the Amadeus database from the years 2016–2018. We created prediction models of financial difficulties of companies for 1 year in advance and also a model for 2 years prediction. These models are based on the combination of two methods, discriminant analysis and logistic regression that belong, among others, to the group of the most commonly used methods to derive prediction models of financial difficulties of the companies. The overall prediction power of the combined model are 90.6%, 93.8% and 90.4%. The results of this analysis can be used for early prediction of the financial difficulties of the company, that could be very useful for all the stakeholders.

Keywords: company’s non-prosperity; prediction models; bankruptcy prediction; discriminant analysis; logistic regression; failure prediction; SME

1. Introduction

The possibility to predict the development of the company’s financial situation for the future is based on the analysis and evaluation of the financial results achieved by the company at present. These are the symptoms of its further development [1]. Correct identification and subsequent interpretation of the indicators of future financial development of the company is the essence of prediction financial analysis. It aims to point out possible threats in the current development of the company in advance to enable it to take action to prevent serious problems. Financial analysis should, therefore, be carried out in every company regardless of its size, business activity or other specifics. These should, however, be taken into account in the content, extent and depth of the analysis carried out.

The issue of predicting the financial situation of companies is a relatively young field of the economic research. Its origin dates back to the 30s’ of the 20th century, but the first prediction models appeared not until the 60s’ of the 20th century. The first study focused on finding the main differences between companies with and without financial problems, based on the analysis of the financial ratios, is the work of Fitzpatrick [2]. Since then prediction financial analysis has undergone significant development, from one-dimensional and multidimensional discriminant analysis through logistic regression to artificial intelligence. The method of discriminant analysis was used for the first time
by Beaver [3], who also formed the basis for prediction models. Based on his research in 1968, Altman used multivariate discriminant analysis to develop probably the most famous bankruptcy prediction model [4]. Ohlson [5] was the first who used the method of logistic regression for creating the model to predict the probability of a company failure.

At present, experts’ have different views on various methods aimed at predicting the financial situation of the companies. Many of them strive to find a reliable tool that can predict a possible threat of the corporate bankruptcy [6]. Some authors deal with the possibility of using models developed in the last century for predicting bankruptcy of current companies. This results in different adjustments and recalculations in the original models. Other authors focus on creating new models using new ratios and modern methods.

As a result of the development of artificial intelligence, new methods such as machine learning techniques, neural networks and genetic algorithms are being introduced into prediction financial analysis [7]. These authors argue that artificial intelligence methods, in comparison with traditional mathematical and statistical methods, can more accurately classify companies as prosperous and non-prosperous [7,8]. For this reason, according to these authors, more researches should be oriented on the application of these methods into the issue of predicting the financial health of the companies [9]. On the other hand, some experts point out the complexity of artificial intelligence methods. They claim that traditional mathematical and statistical methods are comparable to artificial intelligence methods in terms of the accuracy of companies’ classifications. As a result, many prediction models based on traditional prediction techniques have still been developed around the world. Given the different opinions of experts on various prediction methods, it can be argued that every method has its advantages and disadvantages, and also limitations of its use [10]. The constant research in this area proves the currentness of this topic even today. In any case, the issue of predicting the financial situation of a company will always be up to date due to its great importance not only for the company itself but also for all the entities that come into contact with it [11].

The relevance of the topic, analyzed in this study, implies from the interest of scientists and economists in many countries worldwide in the issue of prediction of the financial difficulties of companies. History has shown that developed prediction models lose their classification ability with application in another time and especially in another country [12–14]. Models that were created under the conditions of some country’s economy may not be good enough to predict the financial difficulties of the companies in another country. Therefore, this study aims to create a prediction model able to capture accurately the economic situation of real Slovak companies. As a benefit of the study can be considered the use of the database of real companies operating in Slovakia in the years 2016–2018, the period of interest of this study. This study focuses on a group of small and medium-sized Slovak enterprises (SME), which are 98% of all enterprises in Slovakia. The main contribution of the article is the use of a combination of methods of logistic regression and discriminant analysis, taking into account not only the financial ratios of the companies but also their size and the economic activity in which they operate. The ambition of the combination of methods and using more detailed characteristics of companies is to achieve high prediction ability of the future financial troubles of the company. This good classification is achieved by the addition of the estimated value of the probability of financial difficulties of companies as a new variable, i.e., additional information on enterprises, to the discriminatory model. We consider the use of a combination of both methods in the two-step model creation process to be innovative in this area. At the same time, the benefit of the study is the creation of two prediction models for one year in advance and one model for two years in advance, all using a combination of the mentioned methods.

The rest of the paper is organized as follows. The literature review highlights some interesting studies in the field of prediction of bankruptcy or financial difficulties of companies in various countries in the world and also in Slovakia. The second part of the article describes the base of the statistical methods used and the main characteristics of the database. The third part of the article states the result of the prediction models for Slovak SMEs based on the combination of logistic regression and discriminant analysis. The discussion compares the prediction ability of the models
with other models created before and also states the main strengths and weaknesses and further possible continuation of the study. The conclusion part summarizes the results.

1.1. Literature Review

The prediction of bankruptcy is a topic that has been analyzed in recent years by economists in different countries of the world. Different methods of creating the prediction models are used, whether historically known methods of discriminant analysis and logistic regression, or even more modern methods of neural networks, genetic algorithms, classification trees, and random forests. For example, the study [15] compares three models of predicting corporate financial distress of French companies created by the discriminant analysis, logistic regression and random forests. The results show that the best classification results are given by the random forest method. In a study [16], the authors compare the logit model and data mining models in the field of prediction of bank failures in the USA. They found that the logit model predicts bank failures less precisely than data mining models, but on the other hand, produces fewer missed failures predictions. In [17], the author analyzes how time-varying variables and changes in the macroeconomic environment affect the probability of companies’ financial distress in Croatia. In his study, discrete-time hazard models using logit were applied to demonstrate that the probability of distress is influenced not only by firm-specific variables but also by macroeconomic variables. In [18], the author developed models of crisis diagnostics for Russian companies renting commercial real estate. These models are created by discriminant analysis and logistic regression. In a study [19], the author uses the bootstrapping method and the multivariate discriminant analysis to compare the prediction ability of models created for specific industry with general models created in Poland. Another aim of the research was to define the determinants of joint-stock company bankruptcy in three particular industries of the Polish economy. In [20], the authors apply a multivariate discriminant analysis to differentiate defaulted nations from the non-defaulted ones. Their results indicated discriminant analysis as the most suitable method for insolvency prediction compared to other methods like probit and logit model. In a study [21], the authors prepared a model to predict the risk of encountering financial difficulties in Estonia by using discriminant analysis and logit analysis. The authors of [22] focus their study on a credit risk modelling for prediction of default of companies operating in Romania. The estimations were performed first using logistic regression, and then by artificial neural networks method. A linear discriminant analysis, logistic regression, classification trees and the method of nearest neighbors are applied in [23] to analyze possible bankruptcy signals and to evaluate the financial condition of the selected sector consisting of transport, spedition, and logistics for entities from Poland and Slovakia. In their study [24], the authors deal with the issue of sovereign credit ratings (CR) and the development of the financial market in the European region and say that these serve two main purposes: it verifies the financial condition of the country and signals and the change in the prevailing financial condition. By using a non-linear panel cointegration methodology approach, the authors investigated whether there is a long-run relationship between sovereign CR and financial market development (FMD) in the European region and found an asymmetric effect; both in the long and short run. In [14], the authors verify five traditional bankruptcy models on the sample of Czech construction SME companies and then create a new model for construction companies. The comparison of the prediction accuracy of the verified models shows that the correct classification fluctuates between 50% and 94% for non-failed companies and between 10% and 86% for failed companies one year in advance. Their new prediction model achieves 77% accuracy for non-failed companies and 86% for failed ones. In a study [25], the authors compare 5 different methods for creation of the bankruptcy prediction model, considering up to 10 years before the event in an open, European economy. The authors found that logistic regression and neural networks achieve better results than other approaches. Several authors have published a systematic review of articles in recent years with the issue of models predicting the financial difficulties of companies, for example [13,26–29].

Some prediction models have also been created in Slovakia recently. In addition to already known models of Chrastinova [30], Gurcik [31], Hurtosova [32] and Gulka [33], several authors have
tried to create a prediction model with the best classification ability. In [34], the authors developed models for bankruptcy prediction of Slovak companies using logit and probit methods and provided the comparison of overall prediction ability of the two developed models. The logit model for prediction of non-prosperity of Slovak companies one year in advance has the overall prediction accuracy of 86.7%. The study [35] is focused on the design of bankruptcy models, specifically the selection of suitable predictors to verify whether bankruptcy predictors are industry-specific. Another objective of their research was to determine indicators that can detect signs of bankruptcy earlier than one period before bankruptcy. They found that the application of industry-specific determinants and indicators that can detect signs of bankruptcy for more than one year before can help to increase the prediction accuracy of bankruptcy models for Slovak companies. In [36], the authors analyze the impact of trend variables on the prediction ability of the models constructed using discriminant analysis and decision trees. They developed a new model for Slovak companies using the decision tree technique. The study [37] was also dedicated to the development of bankruptcy prediction models in the Slovak Republic. The study is focused on the comparison of the overall prediction performance of the two developed models: the first one is estimated via discriminant analysis, the other one is based on logistic regression. Also, the authors developed new prediction models for various economic sectors in the studies [36,38–41]. In [37], the author creates the discriminant analysis model with the overall prediction accuracy of 64.41% and the model of logistic regression with the overall prediction accuracy of 68.64%. The study [38], contains the data about Slovak companies from the years 2002–2005. One of the resulting models, logit model, achieved the overall correct classification of 90.24%. In a study [39], the authors created their model using discriminant analysis, covering the years 2002 and 2003 and achieved the overall prediction accuracy of 87.1%. The study [40] was based on the dataset of the companies from 2002–2004. Final model achieved 85.6% of correct classification for all companies in the test sample. The study [41] covers the years 2011 and 2012 and their model of discriminant analysis achieved 88% successfulness in the classification of all companies and the model of logistic regression has the prediction power of 82% for all companies. In this context, we consider the application of a combination of two methods to be a novelty of this study. Such an approach has not yet been applied in the studies of Slovak authors, and as it turned out, helps to improve the classification results of models. In this area, we see the scientific gap that this study seeks to fill. The use of an extensive database of real Slovak companies is another contribution to meet this goal.

Also, several authors in Slovakia have dealt with the evaluation and applicability of already created models in the conditions of Slovakia. The study [42] examines the application of the models of multivariate discriminant analysis for companies. In [10], the authors summarize the advantages and disadvantages of the most used methods in this field: Logistic regression, discriminant analysis, decision trees, and neural networks. The study [43] examines the relationship between the prediction ability of the models of various authors and the number of variables in them. The study [44] deals with the comparison between the results of prediction models and the results of the method for the determination of the value of the company. In [45], the authors apply the well-known Taffler model and the Springate model to validate companies from the electrical engineering industry and conclude that the sector appears to be financially healthy, although the models predict expected difficulties for about a quarter of the companies analyzed. In [46], the authors also created the prediction model for Slovak SME from the years 2008–2011 and achieved the overall prediction power of 88%.

2. Materials and Methods

The financial analysis enables the company to evaluate the financial health identifying the strengths that a company can use to its advantage and also looking for its weaknesses, which are currently not shown very much, but in the future could cause significant problems. The results of the financial analysis provide information to managers enabling them to make optimal choices [13]. Basically, the most sophisticated and the most used financial analysis tools are financial ratios that will be used as predictors in the model for Slovak SMEs also in this study. In [47], authors state that financial ratio models are one of four classes of models: financial ratio models, cash flow models,
market-adjusted returns models, and standard deviation models. These authors state that financial ratio models perform best in the year before bankruptcy. Financial indicators are generally considered to be the primary and the most important predictors of financial stability and economic activity in the area of predicting a company’s financial health, as well as models based on them. Other models are, in fact, also based on adjusted financial indicators, and thus in this respect, we consider them rather as models of a complementary nature. In addition, the models based on market-adjusted returns are not well applicable in (e.g.,) Slovakia due to a low number of publicly traded companies [1,3].

11 suitable predictors were selected, their names and algorithms of calculation are listed in Table 1.

Table 1. Financial ratios used as predictors.

| Title 1 | Title 2 |
|---------|---------|
| $x_1$ | Sales/Total assets |
| $x_2$ | Current assets/Current liabilities |
| $x_3$ | Net income/Shareholders equity |
| $x_4$ | Operating profit/Total assets |
| $x_5$ | (Non-current + current liabilities)/Total assets |
| $x_6$ | Cash & cash equivalents/Total assets |
| $x_7$ | Cash flow/Total assets |
| $x_8$ | Cash flow/(Non-current + current liabilities) |
| $x_9$ | Cash flow/Sales |
| $x_{10}$ | Net income/Sales |
| $x_{11}$ | Non-current liabilities/Total Assets |

Source: own elaboration.

When selecting appropriate predictors of financial health, we used the research of several authors who discussed this issue, for example, [13,26,27,34,48–50]. The predictors in Table 1 are selected according to these analyses, continuing with all logical controls and accuracy checks of the values of financial ratios and also economic controls and selection of variables from a mathematical-statistical point of view with emphasis on the economic importance of selected ratios. Table 2 shows the characteristics of the variables used in the analysis.

Table 2. Characteristics of the financial ratios.

| Variable/Characteristics | Min | Max | Mean | Std. Deviation | Mode | Range | Lower Quartile | Median | Upper Quartile |
|--------------------------|-----|-----|------|----------------|------|-------|---------------|--------|---------------|
| X1_2017                  | 0.00| 7.17| 1.55 | 1.34           | 0.00 | 7.17  | 0.58          | 1.24   | 2.19          |
| X2_2016                  | 0.00| 16.07| 2.62 | 2.73           | 9.81 | 16.07 | 0.91          | 1.53   | 3.29          |
| X3_2016                  | −1.15| 1.49| 0.16 | 0.35           | −0.26| 2.64  | 0.00          | 0.13   | 0.32          |
| X4_2017                  | −1.67| 0.76| 0.10 | 0.18           | 0.00 | 2.43  | 0.01          | 0.08   | 0.18          |
| X5_2017                  | 0.00| 2.98| 0.59 | 0.43           | 0.00 | 2.98  | 0.26          | 0.52   | 0.80          |
| X6_2017                  | 0.00| 2.49| 0.39 | 0.33           | 1.00 | 2.49  | 0.09          | 0.31   | 0.67          |
| X7_2017                  | −0.56| 0.80| 0.08 | 0.13           | 0.01 | 1.36  | 0.01          | 0.01   | 0.14          |
| X8_2017                  | −1.36| 1.89| 0.55 | 0.41           | 0.84 | 3.25  | 0.16          | 0.80   | 0.84          |
| X9_2017                  | −6.86| 0.89| −2.88| 3.46           | −8.68| 7.74  | 0.01          | 0.10   |               |
| X10_2017                 | −15.49| 0.63| −1.68| 4.91           | −15.49| 16.12 | −0.02         | 0.03   | 0.10          |
| X11_2017                 | 0.00| 0.15| 0.01 | 0.03           | 0.00 | 0.15  | 0.00          | 0.00   | 0.01          |

Source: own elaboration.
Undoubtedly, among the most widely used methods of predicting financial difficulties of the companies, also the method of multivariate discriminant analysis as well as the method of logistic regression have their place. Discriminant analysis is the oldest method used to discriminate objects [51]. Discriminant analysis generally identifies the ability of the variables to distinguish existing, previously known, groups of statistical units in the base file and to formulate the classification rule for their allocation to these groups. This results in the two functions of discriminant analysis—descriptive function and predictive function [52]. If we want to identify above-known differences in the groups and quantify these differences with a set of relevant explanatory variables, we talk about a descriptive discriminant analysis [53]. If we use discriminant analysis to create a classification rule by which we can predict the inclusion of a new object into one of the existing groups based on the values of its explanatory variables, we talk about a predictive discriminant analysis [52,54]. The principle of application of multivariate discriminant analysis in the area of analysis and assessment of the financial health of the company lies in the classification of the company into one of the groups of prosperous (financially stable, healthy) or non-prosperous (unstable, in financial troubles) companies. Classification of the company into one of these two groups is realized by estimating the value of variable Y as a function of selected financial and economic indicators in the role of explanatory variables [55]. Therefore, the main goal of applying discriminant analysis is to construct a discriminant function as a linear combination of discriminatory variables and to ensure the best separation of companies into the two groups. Despite the numerous assumptions that discriminatory analysis places on data, this method is very popular and widely used. This is due to the relatively simple interpretability of its findings, even though some statistical techniques can provide better results [54].

The method of logistic regression (or only "logit") belongs to a group of generalized linear models. It is appropriate in the case when the dependent variable is not continuous, but categorical. Its main aim is to find a model that describes the relationship between the explained variable and the group of explanatory variables. The simplest case of its use is a situation where the dependent variable is binary and thus it shall take only two values. Such a variable also corresponds with the situation of predicting the probability of non-prosperity of the company based on its financial and economic ratios or other characteristics. Logistic regression can be used in similar situations as discriminant analysis. Compared to discriminant analysis, it has the advantage that the input assumptions are less restrictive. Classification ability of the logistic regression model is usually better than a model of discriminant analysis. The role of logistic regression is to estimate the probability that the company will be non-prosperous in case of know values of its explanatory variables [52,54]. In this study, we applied the combination of both the abovementioned methods. That is, in the first step of the analysis, we used the logistic regression, and predicted the probability of the non-prosperity for each company, assuming the given values of financial indicators. Then, we use this probability as one explanatory variable in the discriminant model. By this combination of the methods, we supposed the improvement of classification results, because the probability from the first step is estimated using all relevant data about the company. All the relevant variables that we had at our disposal, contributed to the estimation of the probability of financial difficulties of the companies. Then, we did not used the threshold (usually 0.5) to classify the companies based on their probability. Instead, the estimated probabilities were added as other relevant variables to the discriminant model to improve its classification ability because the linear combination of all significant variables carries additional information on the enterprise. We supposed this additional information to be important and influential in the models with a significant impact on the resulting discriminant score, i.e., on the classification of the company into one of the groups. In both models, a stepwise method was used to search for a set of significant variables, with the significance of the contribution of a given variable being tested at a significance level of 0.05. Also, other tests were performed at a significance level of 0.05. Among the step-methods of the model creation, we consider the stepwise method to be the best choice. This method, unlike backward elimination and forward selection, tests the significance of previously included variables at each step, which can be changed by adding other variables. The inclusion of the variable is therefore not definitive, which is its advantage.
The classification ability of the models will be evaluated by a classification table and ROC curve and its AUC measure. Classification table contains absolute and relative frequencies of correctly and incorrectly classified companies into the groups of prosperous and non-prosperous companies. The ROC curve indicates the quality of the prediction model, the closer the AUC value of the area to 1, the better the prediction power the model [16,34].

To create prediction models for SMEs in Slovakia, we used data from the financial statements of real Slovak companies. The source of this data was the Amadeus database, focusing on the period of the years 2016–2018 in the sense that the financial indicators of companies are measured in 2016 and 2017 and the result variables, indicating the prosperity or non-prosperity of the company, are set in 2017 and 2018. The whole database contains (after the preparation of the dataset including the correctness check, missing data analysis and analysis of outliers) totally 75,652 companies. Among all, there were the following numbers and frequencies of prosperous and non-prosperous companies (Table 3).

| Year/Company Status | Number of Companies 2017 | Number of Companies 2018 |
|---------------------|--------------------------|--------------------------|
|                     | Frequency | Percent | Frequency | Percent |
| prosperous          | 67,432    | 89.1     | 66,155    | 87.4     |
| non-prosperous      | 8,220     | 10.9     | 9,497     | 12.6     |
| total               | 75,652    | 100      | 75,652    | 100      |

Source: own elaboration.

The company has the status “non-prosperous provided that it fulfils the following individually determined criteria, based on the valid legislation of the Slovak Republic. From 1 January 2016, the amendment to Act no. 513/1991 Coll. Commercial Code as amended introduced the institute “company in crisis” in the form of §67a to 67i. According to this amendment-a company is in crisis if it is in the state of bankruptcy or is threatened by bankruptcy. The bankruptcy of the company is governed by Act no. 7/2005 Coll. on Bankruptcy and Restructuring as amended. Under §3 of this Act, the company is in bankruptcy if it is indebted or unable to pay. This means that it has more than one creditor and the value of its liabilities exceeds the value of its assets (i.e., has negative equity) or is unable to pay at least two financial obligations to more than one creditor for 30 days after the due date. We approximated this fact by the variable \( x_2 \) (see Table 1), determining the breakpoint of this indicator at level 1. Companies are at risk of bankruptcy if the solvency ratio (Solvency ratio (liability based) = Shareholders funds/(Non-current liabilities + Current liabilities) - 100) is less than 0.06 in 2017, or less than 0.08 in 2018 (Act no. 513/1991 Coll.) [13]. If at least one of the three conditions were not met, the company was considered to be “prosperous”. So, the companies were classified into two groups: non-prosperous (unhealthy, in financial troubles) and prosperous (non-default, healthy) in the context of legislative adjustments [56,57]. We take into account all standardized legal forms of companies with non-consolidated financial statements. The sample does not contain self-employed persons and general partnerships, as these two forms the do not apply the legislative stipulations about a company in crisis.

Thus, more than 89% (in 2017) and 87% (in 2018) of Slovak SME companies are prosperous and the rest of them, nearly 11% or 13% respectively are in financial troubles. This status of the company is given in the database by the output variable \( Y \) that is binary and takes the value 0 for a prosperous company and 1 for an non-prosperous company.

3. Results

Together, three prediction models using the combination of methods were created:

1. 1-year prediction model using the data from 2016 (and prosperity of the company from 2017);
2. 1-year prediction model using the data from 2017 (and prosperity of the company from 2018);
3. 2-year prediction model using the data from 2016 (and prosperity of the company from 2018).
Firstly, we created the models of the logistic regression using the forward method (likelihood ratio based), with the significance level of the variables in the model of 0.05. As the predictors, we applied all the mentioned financial ratios and the type of the economic activity of the company given by the NACE code and the size of the company (small company or medium-sized company). All calculations in this study were performed using SPSS software. The generated model has the value of Nagelkerke R-Square, which is an alternative to R-Square from linear regression, at 0.393, 0.583, and 0.437 respectively (Table 4). This means the percentage of the “pseudo-variability” of the dependent variable is explained by this model. From the previous experience concerning the creation of bankruptcy prediction models, we know that such level of this characteristic of the model is common. Therefore, from this perspective we consider it to be good enough.

The classification power of the models was evaluated by the classification table and by the AUC criterion (Table 4).

| Model               | Nagelkerke R-Square | Percentage Correct | AUC  |
|---------------------|---------------------|---------------------|------|
| 1-year prediction y_2017 | 0.393               | 87.7                | 0.806|
| 1-year prediction y_2018 | 0.583               | 90.9                | 0.890|
| 2-year prediction y_2018 | 0.437               | 89.7                | 0.856|

Source: own elaboration.

According to the percentage of successful classification of enterprises and according to the AUC criterion, the best one seems to be the 1-year model from 2018, which achieved 90.9% of the correct classification of all enterprises and its AUC is 0.890. However, the other two models also achieved high success.

The resulting logistic models were used to predict the probability of the non-prosperity of every company. Subsequently, we applied this probability as one of the predictors in the discriminant analysis model.

To verify the assumption of covariance matrices equality we used Box’s test with a null hypothesis of equal covariance matrices of the companies’ groups. The results of the Box’s test are in Table 5.

|                      | 1-Year Prediction y_2017 | 1-Year Prediction y_2018 | 2-Year Prediction y_2018 |
|----------------------|--------------------------|--------------------------|--------------------------|
| Box’s M              | 12757.97                 | 1259.56                  | 33833.31                 |
| Sig.                 | 0.00                     | 0.00                     | 0.00                     |

Source: own elaboration.

Based on the Box’s test result, we reject the zero hypothesis on the equality of covariance matrices of groups of prosperous and non-prosperous companies for all the models. In such a case, the use of a quadratic discriminant analysis instead of a linear one has to be considered, but this is not widely used in economic practice. It is usually mentioned more in theory as a more optimal variant in case of violation of this assumption as it is sensitive to deviations from normality and rarely gives better results than linear discriminant analysis. Its results are also more difficult to interpret [52]. Therefore, we had to treat this violation provided the use of discriminant analysis using this fact in the application of this method in SPSS, that means, to use separate groups covariance matrix instead of within-groups matrix. Table 6 shows the resulting discriminant equations for the Z-score (discriminant score) of each company. These models were created using the stepwise method of selecting significant variables from the set of financial ratios, NACE categories in the form of dummy variables, the size of the enterprises in the form of dummy variables and the probabilities obtained from the first step of the procedure. At a significance level of 0.05, the significance of the contribution of each variable to discrimination was verified by Wilk’s Lambda and the F-test of its significance. The stepwise discriminant analysis led to the following resulting equations.
The variables $pp_{1\text{year,2017}}$, $pp_{1\text{year,2018}}$ and $pp_{2\text{year,2018}}$ are the probabilities predicted from the model of logistic regression from the first step of the analysis. All these three probabilities were statistically significant in the discriminant models at the significance level of 0.05.

Using these equations, we can predict the classification of the company into the group of prosperous, or into the group of non-prosperous companies. As the constants are used in the models, the resulting discriminant score is always compared to the weighted average of the group’s centroid, which is zero in this case. So, the negative Z-score implies the classification of the company to the group of prosperous and positive value to the group of non-prosperous ones.

The quality of the models and the impact of the predictors on the value of the Z-score of the company is evaluated by the canonical correlation coefficients (referring about the correlation between the discriminant function and the explanatory variables used in the models), the significance of these correlations, the values of standardized canonical discriminant function coefficients and also correlation coefficients between the discriminant function and explanatory variables (used to assess discrimination ability of the variables in the resulting model). Table 7 shows these characteristics of the models and three variables with the highest correlation with the discriminant function for each model.

| Characteristic                  | Model          | 1-Year Prediction $y_{2017}$ | 1-Year Prediction $y_{2018}$ | 2-Year Prediction $y_{2018}$ |
|--------------------------------|----------------|------------------------------|------------------------------|-----------------------------|
| Canonical Correlation          | value          | 0.525                        | 0.663                        | 0.709                       |
|                                | significance   | 0.000                        | 0.000                        | 0.000                       |
| Standardized Canonical         | $PP_{1\text{year,2017}}$ | $x_5_{2017}$                 | $x_8_{2017}$                 | $PP_{2\text{year,2018}}$   |
| Discriminant Function          | 1st variable   | 1.001                        | 1.070                        | 0.988                       |
| Coefficients                   | coefficient    | 0.000                        | 0.000                        | 0.000                       |
|                                | 2nd variable   | $x_4_{2016}$                 | $x_8_{2017}$                 | $x_3_{2016}$                | -0.651                      | 0.354                      | -0.404                      |
|                                | coefficient    | -0.651                      | 0.354                       | -0.404                      |                             |                            |                            |
|                                | 3rd variable   | $x_7_{2016}$                 | $PP_{1\text{year,2018}}$   | $x_{10\text{2016}}$         | 0.620                       | 0.186                       | 0.037                       |
|                                | coefficient    | 0.620                       | 0.186                       | 0.037                       |                             |                            |                            |
| Correlation Coefficients       | 1st variable   | $PP_{1\text{year,2017}}$   | $x_5_{2017}$                 | $PP_{2\text{year,2018}}$   |
|                                | correlation    | 0.999                       | 0.896                       | 0.997                       |
|                                | 2nd variable   | $x_6_{2016}$                 | $PP_{1\text{year,2018}}$   | $x_5_{2016}$                | -0.162                      | 0.578                       | 0.610                       |
|                                | correlation    | -0.162                      | 0.578                       | 0.610                       |                             |                            |                            |
|                                | 3rd variable   | $size_{small}$              | $x_4_{2017}$                | $x_4_{2016}$                | 0.144                       | -0.300                      | -0.234                      |
|                                | correlation    | 0.144                       | -0.300                      | -0.234                      |                             |                            |                            |

Source: own elaboration.

Based on the past experience, the value of the canonical correlation coefficient, in the bankruptcy prediction models based on discriminant analysis, is usually just at such or very similar levels. On this basis, we consider that the discriminant function is good enough. Moreover, all of these
characteristics are statistically significant. The probabilities of the non-prosperity, predicted from the first step of the analysis, have the highest coefficients (the first higher in 1-year prediction model in 2017 and in the 2-year model and the third in the 1-year model in 2018). They are also strongly and positively correlated with the discriminant function. Among the other variables, \( x_4 \) and \( x_5 \) have a significant correlation with the discriminant function and are repeated in two of the models. The correlation of \( x_4 \) is negative and of \( x_5 \) is positive. This fact also has its logical interpretation, as \( x_4 \) is one of the rentability ratios and \( x_5 \) belongs to the group of indebtedness indicators.

Classification ability of the model is quantified in Table 8 by the percentage of correct classification and by the AUC criterion.

| Model                  | Percentage Correct | AUC  |
|------------------------|--------------------|------|
| 1-year prediction \( y_{2017} \) | 90.6               | 0.933|
| 1-year prediction \( y_{2018} \) | 93.8               | 0.934|
| 2-year prediction \( y_{2018} \) | 90.4               | 0.907|

Source: own elaboration.

By comparing the percentage of correct classification and AUC criterion of the logistic regression model created in the first step of the analysis and the subsequent discriminant model, it is visible that the combination of methods brought improved prediction ability and better AUC quality of the model of non-prosperity of Slovak SMEs.

4. Discussion

Prediction models created using one of the two statistical methods, used in this study, are emerging worldwide over the last period. As aforementioned, new prediction models have been created also in Slovakia by several authors. Comparing our results with the other studies, we can say that we achieved very similar, or even a little bit better classification results. For example, in [34] the model achieved a little bit lower classification power than in our combined model in 2017. In a study [38], the model achieved a very similar result than our models, the model in [40] has a prediction power similar to the model of logistic regression that was the first step of our combined models. The two models in the study [46] have both lower classification results than our models in this study.

Combination of the methods of logistic regression and discriminant analysis brings very good classification ability of the resulting models with high AUC score. The classification result of the logistic regression, created in the first step of the analysis, was then improved by the consequent discriminant analysis that achieves much better classification results. All three models achieved more than 90% correct classification. We explain the improvement of classification results using the predicted probability of non-prosperity of the company by the fact that all the relevant data that we had at our disposal about the company at the given moment were used in this probability. All relevant variables contributed to the best possible estimation of the probability of financial difficulties. However, we did not proceed as usual in the classification of the enterprise based on the predicted probability, where the limit is usually considered to be the probability of 0.5. Enterprises with a probability higher than 0.5 are classified in the group of non-prosperous enterprises with a probability of non-prosperity below 0.5 are classified in the group of prosperous. However, this threshold is very "thin" because, for example, an enterprise with a probability of 0.495 will be classified as prosperous, but an enterprise with a probability of 0.505 will be classified as non-prosperous. So, we did not apply this classification but instead added the probability of every company as another relevant variable into the discriminatory model. This discriminant model is finally used to categorize companies into two groups. Probability, classified as one of its explanatory variables, improves its classification ability because it adds new additional information in the form of a linear combination of all the significant variables available in the enterprise. The individual variables themselves may prove insignificant in the discriminant model, but in a combination with several variables, their interactions may be significant. This is also shown in the analysis of the
resulting models, where the variable “probability” in all three models is included among the three most influential variables. The additional information is therefore important and influential and affects the resulting classification of the company into one of the groups.

The achieved classification result is, of course, appropriate, but the truth remains that not only correctly classified non-prosperous enterprises but also correctly classified prosperous enterprises signed on this total percentage. However, to predict financial difficulties, it is necessary to focus mainly on the correct classification of non-prosperous enterprises, so that the model can correctly predict the threat of problems. In this part of the classification, we see an even weaker side of the created models. However, a comparison with other studies shows that several authors achieved similar or even worse results in their, sometimes also well-known, models.

As a strength of this analysis, we consider the fact that we used an extensive database of real Slovak companies and their actual data from the financial statements from 2016–2018. The database was also subjected to thorough cleaning, preparation, and analysis of the accuracy of records. On the contrary, as a weakness of this study, we can consider the fact that many companies had to be removed from the database due to a large amount of missing data. This was also the case for several variables that could be very useful for the analysis, but because of the large percentage of missing data, they could not be included in the models. In the further continuation of this research, we will focus on the creation of prediction models for Slovak SMEs using data mining and artificial intelligence methods.

5. Conclusions

In the article, we focused on predicting the financial troubles of Slovak small and medium-sized companies, which are the largest groups of companies in Slovakia, one year or two years in advance. We applied the combination of logistic regression and the method of discriminant analysis to model the probability of the company’s failure and to model the classification into the groups of prosperous or non-prosperous companies. All the models match in the explanatory variable \( x_n \) and also in the variable describing the size of the company (in the model, the variable \( \text{size}_{\text{small}} \) was included, medium-sized companies were the reference category). Thus, we can say that the models have good prediction ability. However, it would be appropriate to verify this prediction power, for example by splitting the database into the training sample and test sample.

Author Contributions: Conceptualization, L.S. and L.M.; methodology, L.S.; software, M.D. and E.N.; validation, L.M., M.D.; formal analysis, L.S.; investigation, M.D.; resources, L.S.; data curation, M.D.; writing—original draft preparation, L.S and L.M.; writing—review and editing, L.S., L.M. and M.D.; visualization, M.D.; supervision, L.S. and E.N.; project administration, M.D.; funding acquisition, L.S. and M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Slovak Research and Development Agency-Grant NO. APVV-14-0841: Comprehensive Prediction Model of the Financial Health of Slovak Companies, and Institutional research of Faculty of Operation and Economics of Transport and Communications—University of Zilina, no. 11/PEDAS/2019.

Acknowledgements: This article has benefited from constructive feedback on earlier drafts by the editors, the chief editor, and the anonymous reviewers.

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. Weissova, I.; Siekelova, A.; Kramarova, K. Modeling of company’s default probability in relation to its credit risk. Glob. J. Bus. Econ. Manag. Curr. Issues 2016, 6, 130–137, doi:10.18844/gjbe.v6i2.1378.
2. FitzPatrick, P.J. A comparison of the ratios of successful industrial enterprises with those of failed companies. Financ. Ratio Anal. Hist. Perspect. 1978, 6, 727–731.
3. Beaver, W.H. Financial Ratios as Predictors of Failure. J. Acc. Res. 1966, 4, 71–111, doi:10.2307/2490171.
4. Altman, E.I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *J. Financ.* 1968, 23, 589–609, doi:10.2307/2978933.

5. Ohlson, J.A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *J. Acc. Res.* 1980, 18, 109–131, doi:10.2307/2490395.

6. Czuba, M. Youth as a Recipient of Marketing Innovations in Activities of Companies Providing Municipal Services. *Forum Sci. Oeconomia* 2016, 4, 45–54.

7. Jones, S.; Johnstone, D.; Wilson, R. Predicting Corporate Bankruptcy: An Evaluation of Alternative Statistical Frameworks. *J. Bus. Financ. Acc.* 2017, 44, 3–34, doi:10.1111/jbfa.12218.

8. Jan, C. An Effective Financial Statements Fraud Detection Model for the Sustainable Development of Financial Markets: Evidence from Taiwan. *Sustainability* 2018, 10, 1–14.

9. Chromacakova, A.; Starzynska, H. People and Communication in CRM Process in Small and Medium Sized Czech Enterprises. *Forum Sci. Oeconomia* 2017, 59–70, doi:10.23762/FSO_vol5no4_17_5.

10. Svbova, L.; Durica, M. A Closer View of the Statistical Methods Globally Used in Bankruptcy Prediction of Companies. In *Globalization and its socio-economic consequences*, Proceedings of the 16th International Scientific Conference Globalization and its socio-economic consequences 2016, Part V, Rajecke Teplice, Slovak Republic, 5th–6th October 2016; Kliestik, T., Ed.; ZU – University of Zilina: Zilina, Slovak Republic, 2016; pp. 2174–2181, ISBN 978-80-8154-191-9.

11. Georgescu, V. Using genetic algorithms to evolve type-2 fuzzy logic systems for predicting bankruptcy. *Kybernetes* 2017, 46, 142–156, doi:10.1108/K-06-2016-0152.

12. Kliestikova, J.; Misankova, M.; Kliestik, T. Bankruptcy in Slovakia: International comparison of the creditor’s position. *Oeconomia Copernic.* 2017, 8, 221–237, doi:10.24136/oc.v8i2.14.

13. Kovacova, M.; Kliestik, T.; Valaskova, K.; Durana, P.; Juhaszova, Z. Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. *Oeconomia Copernic.* 2019, 10, 743–772, doi:10.24136/oc.2019.034.

14. Karas, M.; Srbová, P. Predicting bankruptcy in construction business: Traditional model validation and formulation of a new model. *J. Int. Stud.* 2019, 12, 283–296, doi:10.14254/2071-8330.2019/12-1/19.

15. Jabeur, S.B.; Fahmi, Y. Forecasting financial distress for French firms: A comparative study. *Empir. Econ.* 2018, 54, 1173–1186, doi:10.1007/s00181-017-1246-1.

16. Jing, Z.; Fang, Y. Predicting US bank failures: A comparison of logit and data mining models. *J. Forecast* 2018, 37, 235–256, doi:10.1002/for.2487.

17. Zikovic, I.T. Challenges in Predicting Financial Distress in Emerging Economies: The Case of Croatia. *East. Eur. Econ.* 2018, 56, 1–27, doi:10.1080/00128775.2017.1387059.

18. Barkar, A.A. Development of Instruments of Financial Diagnostics of Crisis. *Vestn. St. Petersb. Univ.-Ekonom.* 2017, 33, 658–672, doi:10.21638/11701/spbui05.2017.408.

19. Herman, S. Industry Specifics of Joint-Stock Companies in Poland and Their Bankruptcy Prediction. In Proceedings of the 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena, Zakopane, Poland, 9th–12th May 2017; Papiez, M., Śmiech, S. (Eds.); Foundation of the Cracow University of Economics: Cracow, Poland, 2017; pp. 93–102, ISBN: 978-83-65173-85-0.

20. Sztetla, B.; Mentel, G.; Brożyna, J. In search of insolvency among European countries. *Econ. Res.-Ekonom.* *Istraživanja* 2016, 29, 839–856, doi:10.1080/1331677X.2016.1237301.

21. Lohk, P.; Siimann, P. Predicting the Risk of Encountering Financial Difficulties by the Example of Estonian Municipalities. In Proceedings of the 5th International Conference on Accounting, Auditing, and Taxation (ICAAI 2016), Tallinn, Estonia, 8th–9th December 2016; Alver, J., (Ed.); Atlantis Press: Paris, France, 2016; pp. 297–306, doi:10.2991/icaa-16.2016.30.

22. Dima, A.M.; Vasilache, S. Companies default prediction using neural networks. *Rom. J. Econ.* 2016, 19, 127–143.

23. Brożyna, J.; Mentel, G.; Pisula, T. Statistical Methods of the Bankruptcy Prediction in the Logistic Sector in Poland and Slovakia. *Transform. Bus. Econ.* 2016, 15, 93–114.

24. Li, C.; Pervaz, K.; Asif Khan, M.; Ur Rehman, F.; Oláh, J. On the Asymmetries of Sovereign Credit Rating Announcements and Financial Market Development in the European Region. *Sustainability* 2019, 11, 6636, doi:10.3390/su11236636.

25. Altman, E.I.; Iwanicz-Drozdowska, M.; Laitinen, E.K.; Suvas, A. A Race for Long Horizon Bankruptcy Prediction. *Appl. Econ.* 2020, 0, 1–20, doi:10.1080/00036846.2020.1730762.
26. Ravi Kumar, P.; Ravi, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *Eur. J. Oper. Res.* 2007, 180, 1–28, doi:10.1016/j.ejor.2006.08.043.
27. Bellovary, J.L.; Giacomino, D.E.; Akers, M.D. A Review of Bankruptcy Prediction Studies: 1930 to Present. *J. Financ. Educ.* 2007, 33, 1–42.
28. Prusak, B. Review of Research into Enterprise Bankruptcy Prediction in Selected Central and Eastern European Countries. *Int. J. Finance Stud.* 2018, 6, 60, doi:10.3390/ijfs6030060.
29. 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, doi:10.1016/j.eswa.2017.10.040.
30. Chrastinova, Z. Metódy hodnotenia ekonomickej bonity a predikcie finančnej situácie poňhospodárských podnikov. VUEPP: Bratislava, Slovak Republic, 1998; ISBN 978-80-8058-022-3.
31. Gurečík, L. G-index - the financial situation prognosis method of agricultural enterprises. *Agric. Econ. Zemědělská Ekonomická Revue.* 2012, 48, 373–378, doi:10.17221/5338-AGRICECON.
32. Hurtosova, J. Konštrukcia ratingového modelu, nástroja hodnotenia úverovej spôsobilosti podniku. Dissertation Thesis, Economic University in Bratislava: Bratislava, Slovakia, 2009.
33. Gulka, M. Model predikcie úpadku obchodných spoločnosti podnikajúcich v podmienkach SR. *Forum Stat.* *Slovakia* 2016, 12, 16–22.
34. Kovacova, M.; Kliestik, T. Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilib. Q. J. Econ. Policy* 2017, 12, 775–791, doi:10.24136/eq.v124.40.
35. Karas, M.; Reznakova, M. The stability of bankruptcy predictors in the construction and manufacturing industries at various times before bankruptcy. *Em Econ. Manag.* 2017, 20, 116–133, doi:10.15240/tul/001/2017-2-009.
36. Gavurova, B; Janke, F.; Packova, M.; Pridavok, M. Analysis of Impact of Using the Trend Variables on Bankruptcy Prediction Models Performance. *Ekon. Cas.* 2017, 65, 370–383.
37. Mihalovic, M. Performance Comparison of Multiple Discriminant Analysis and Logit Models in Bankruptcy Prediction. *Econ. Sociol.* 2016, 9, 101–118.
38. Gavlak, R. Bankruptcy prediction in Slovak companies using linear probability models. *Forum Stat. Slovakia* 2006, 5, 65–69.
39. Hiadlovsky, V.; Kral, P. Využitie diskrimináčnej analýzy na predikovanie finančnej situácie podnikov v SR. *Forum Stat. Slovakia* 2005, 1, 44–49.
40. Kral, P.; Hiadlovsky, V.; Gavlak, R. Použitie faktorovej analýzy pri predikcii finančnej situácie podnikov v SR s využitím SPSS. *Forum Stat. Slovakia* 2007, 4, 74–80.
41. Bieliková, T.; Bánynová, T.; Píterková, A. Prediction Techniques of Agriculture Enterprises Failure. *Procedia Econ. Finance.* 2014, 12, 48–56, doi:10.1016/S2212-5671(14)00319-0.
42. Durica, M.; Adamko, P. Verification of MDA Bankruptcy Prediction Models for Enterprises in Slovak Republic. In Proceedings of the 10th International Days of Statistics and Economics, Prague, Czech Republic, 8th – 10th September 2016; Löster, T., Pavelka, T. (Eds.); Melandrium: Prague, Czech Republic, 2016; pp. 400–407, ISBN 978-80-87990-10-0.
43. Kral, P.; Janoskova, K. Evaluation of prediction ability of bankruptcy prediction models applying logistic regression (LOGIT). *Lecture notes in management science* 2016, 66, 21–26.
44. Podhorska, I. Prediction models: Predictive ability in slovak conditions. *Podnik. Ekonom. Manaž.* 2018, 2018, 75–85.
45. Stefkov, R.; Jencova, S.; Vasanicova, P.; Litavcova, E. An Evaluation of Financial Health in the Electrical Engineering Industry. *J. Compet.* 2019, 11, 144–160, doi:10.7441/joc.2019.04.10.
46. Harumova, A.; Janisova, M. Rating Slovak Enterprises by Scoring Functions. *Ekon. Cas.* 2014, 62, 522–539.
47. Mossman, C.E.; Bell, G.G.; Swartz, L.M.; Turtle, H. An empirical comparison of bankruptcy models. *Finance Rev.* 1998, 33, 35–54, doi:10.1111/j.1540-6288.1998.tb01367.x.
48. Dimitras, A.I.; Zanakis, S.H.; Zopounidis, C. A survey of business failures with an emphasis on prediction methods and industrial applications. *Eur. J. Oper. Res.* 1996, 90, 487–513, doi:10.1016/0377-2217(95)00070-4.
49. Liu, J.; Wu, C. A gradient-boosting decision-tree approach for firm failure prediction: An empirical model evaluation of Chinese listed companies. *J. Risk Model. Valid.* 2017, 11, 43–64.
50. Zoricak, M.; Gnip, P.; Drotar, P.; Gazda, V. Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Econ. Model.* 2020, 84, 165–176, doi:10.1016/j.econmod.2019.04.003.
51. Kliestik, T.; Vrbka, J.; Rowland, Z. Bankruptcy prediction in Visegrad group countries using multiple discriminant analysis. *Equilib. Q. J. Econ. Econ. Policy* **2018**, *13*, 569–593, doi:10.24136/eq.2018.028.

52. Tuffery, S. An Outline of Data Mining Methods. In *Data Mining and Statistics for Decision Making*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2011; pp. 167–174, ISBN 978-0-470-97917-4.

53. Kral, P.; Musa, H.; Lazaroiu, G.; Misankova, M.; Vrbka, J. Comprehensive assessment of the selected indicators of financial analysis in the context of failing companies. *J. Int. Stud.* **2018**, *11*, 282–294, doi:10.14254/2071-8330.2018/11-4/20.

54. Tinsley, H.E.A.; Brown, S.D. *Handbook of Applied Multivariate Statistics and Mathematical Modeling*; Elsevier: Amsterdam, The Netherlands, 2000; ISBN 978-0-12-691360-6.

55. Islam, M.A.; Khan, M.A.; Popp, J.; Sroka, W.; Oláh, J. Financial Development and Foreign Direct Investment—The Moderating Role of Quality Institutions. *Sustainability* **2020**, *12*, 3556, doi:10.3390/su12093556.

56. Kliestik, T.; Kliestikova, J.; Kovacova, M.; Svabova, L.; Valaskova, K.; Vochozka, M.; Olah, J. *Prediction of Financial Health of Business Entities in Transition Economies*; Addleton Academic Publishers: New York, USA, 2018; ISBN 978-1-942585-39-8.

57. Kliestik, T.; Valaskova, K.; Lazaroiu, G.; Kovacova, M.; Vrbka, J. Remaining Financially Healthy and Competitive: The Role of Financial Predictors. *J. Compet.* **2020**, *12*, 74–92, doi:10.7441/joc.2020.01.05.11.

© 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/).