

## Financial distress prediction in Slovakia: An application of the CART algorithm

### Marek Durica

*Department of Quantitative Methods and Economic Informatics,  
Faculty of Operation and Economics of Transport and  
Communications,  
University of Zilina,  
Slovakia*  
[marek.durica@fpedas.uniza.sk](mailto:marek.durica@fpedas.uniza.sk)

### Jaroslav Frnda

*Department of Quantitative Methods and Economic Informatics,  
Faculty of Operation and Economics of Transport and  
Communications,  
University of Zilina,  
Slovakia*  
[jaroslav.frnda@fpedas.uniza.sk](mailto:jaroslav.frnda@fpedas.uniza.sk)

### Lucia Svabova

*Department of Economics,  
Faculty of Operation and Economics of Transport and  
Communications,  
University of Zilina,  
Slovakia*  
[lucia.svabova@fpedas.uniza.sk](mailto:lucia.svabova@fpedas.uniza.sk)

**Abstract.** The topic of predicting financial distress situation has been of interest to many economists and scientists from around the world for several years. As it turned out in practice, the application of existing prediction models to predict the financial difficulties of Slovak companies brings lower prediction accuracy, as these models were created in the conditions of another country. Therefore, the main aim of the article is to create the model for the prediction of the financial distress of the Slovak companies, based on the real conditions of Slovak economics. For this analysis, a dataset of the most important financial ratios that may affect the financial health of the Slovak companies was obtained from the Amadeus database, containing the data on more than 100,000 real companies operating in the Slovak economy in the period 2016 to 2018. For the creation of the models for the prediction of the financial distress of companies one year and

**Received:**

July, 2020

**1st Revision:**

December, 2020

**Accepted:**

March, 2021

DOI:

10.14254/2071-  
8330.2021/14-1/14

two years in advance, the CART algorithm generating the binomial decision tree was used. The developed models achieve an overall accuracy of 87.3% and 91.9% and are very simple to the real application. The results from this prediction are important not only for companies themselves but also for all their stakeholders, as they could help the company to mitigate or eliminate the threat of financial distress and the other corporate risks related to such a situation in the company.

**Keywords:** CART, financial distress, financial ratios, prediction model, prediction ability.

**JEL Classification:** C38, G33

## 1. INTRODUCTION

During the last decades, the issue of the prediction of financial distress of the companies has been a very interesting topic, as it is very important mainly for the managers of the companies, but also for the employees, credit and business partners of the company, but overall for all other interested stakeholders. A functional and reliable prediction model can be a supporting tool for determining the financial situation of the company, which can subsequently lead to the introduction of the necessary measures to mitigate or eliminate the occurrence and threat of financial difficulties or even bankruptcy of the company. (Kliestikova et al., 2017). Investors' investments in the company will then be safer, or they can be adjusted by investors to reduce the risk of their investments. Ultimately, these operations will be for the company beneficial. (Kral et al., 2018)

Methodological approaches usually used to create financial distress prediction models include the group of statistical methods and the group of data mining methods. Multivariate discriminant analysis (MDA), binary logit or probit are the typical representatives in the first group. The pioneers in using these methods for the creation of financial distress prediction models were Altman and Ohlson (Kovacova et al., 2018). In 1968, Altman developed a well-known Z-score model based on MDA, and then in 1980, Ohlson published the first logit bankruptcy model. These methods are fully parametric (Kliestik et al., 2018). The group of data mining methods consists of many various methods: artificial neural networks (ANN), genetic algorithms (GE), decision trees (DT), etc. These methods are fully non-parametric and there are no assumptions about the distribution of the variables and statistical relations among them (Paliwal & Kumar, 2009).

At present, we can say that both groups of methods are used in relation to prediction models of financial difficulties. At the beginning of this issue development there was a boom in using mainly the discriminant analysis and logistic regression, while in recent years, many methods of data mining and artificial intelligence have come to the fore, including also the CART method generating binomial decision trees. Of course, all these methods have their supporters as well as opponents, and there are also certain advantages and disadvantages associated with each of these methods (Gregova et al., 2020).

The CART algorithm can handle any type of input variables without some statistical assumptions about their probability distribution. The variables may even be collinear, and this technique has no problems with outliers or missing values. Moreover, CART generates a simple decision tree that can easily be programmed, and the classification of new cases is then very fast (Affes & Hentati-Kaffel, 2018).

## 2. LITERATURE REVIEW

Financial distress is a notion covering more terms used in the case that company is confronted with financial difficulties. For the situations resulting in the problems of continuity of the company operations,

terms such as bankruptcy, failure, insolvency, and default are usually used. (Gandolfi et al., 2018) According to the Altman definition of failure (1993), failure is a situation where "the realized rate of return on invested capital, with allowances for risk consideration, is significantly and continually lower than prevailing rates of similar investments".

The models for the prediction of financial distress used to be created by various statistical techniques. Before the 1980s, the most widely used method in this field was the MDA. This method is sometimes criticised for its relatively strict assumptions and data requirements (Balcaen & Ooghe, 2006) because, in real-world applications, these assumptions are usually hard to satisfy (Karas & Srbova, 2019). The second approach widely used in the field of financial distress prediction is logistic regression, firstly used by Ohlson (1980). This method contains the assumptions of homogeneity of variances and independence of the predictors. Gradually, some other techniques were applied in the field of prediction of bankruptcy, for example, probit analysis firstly used by Zmijewski (1984) and later for example by Kasgari et al. (2012).

In the group of data mining methods, support vector machines (SVM) and decision trees (especially CART and C4.5 trees) are the most commonly used techniques for identifying financial distress (Alaminos et al., 2016; Zhao et al., 2016; Liu & Wu, 2017; Tsai et al., 2014; Zieba et al., 2016). The models based on these techniques are commonly used in a practice. This is because these models are relatively easy to understand and to apply. Regarding the differences between these two methods, CART generates binary decision trees, while C4.5 does not. To select the best splitting criterion, CART employs the Gini index, while the information gain is used in the C4.5 algorithm. Another difference is in pruning. A cost-complexity pruning is used in the CART algorithm, whereas C4.5 employs a single-pass pruning algorithm.

In the field of identification of financial difficulties, research brings using various AI-based methods. Researchers use various machine learning techniques to solve this problem, but ANN is the most used one (Georgescu, 2017; Jones et al., 2016; Kumar & Ravi, 2007). Among the alternative techniques, the case-based reasoning approach was used by Li and Sun (2009) and genetic algorithms were used by Shin and Lee, 2002. In 2012, du Jardin and Severin used self-organising maps (or Kohonen maps), while simulation analysis was applied by Cohen et al. (2012). In addition to classical methods, Virag and Nyitrai (2014) used rough sets of financial distress prediction.

CART algorithm was introduced by Breiman et al. (1984). Right in the year 1985, Frydman et al. were the first to use the technique of DT in the field of the prediction of the failure of the companies. They found their DT achieved a better prediction accuracy of the failure of the companies compared with the MDA or Logit. Among the models created in recent years, we can mention several studies dealing with financial distress prediction. Several decision trees generating algorithms (CART, C5.0, and CHAID) together with the logistic regression (LR) approach were used to model the financial health of Taiwanese listed companies (Chen, 2011). In 2015, Irimia-Dieguez et al. published their study in which selected prediction models such as CART and Logit-based were compared.

Also, several studies discussing the creation of prediction models for some selected specific groups of companies were published. For example, in 2016, Brozyna et al. presented a model designed to predict the financial health of Polish and Slovak banks. This model was created based on a strong data sample from 2006-2012 containing the values of 25 financial ratios of 155 European banks. In 2017, Karas and Reznakova used the values of almost thirty ratios of construction companies operating in the Czech Republic and Slovakia to create a model by the CART technique.

In Slovakia, some national models have also been introduced in the last two decades. For agricultural companies, Gurcik (2002) and Chrastinova (1998) created a model using the classical MDA approach. These models are still used, not only for the agricultural sector. The first Slovak Logit models were published by Hurtosova (2009) and Gulka (2016). In 2017, Kovacova and Kliestik introduced the Logit model and the Probit model and compared the classification accuracy of these models. Gavurova et al. (2017) used the DT

technique to develop a new Slovak national model. Karas & Reznakova (2017) focused their study on the development of the CART-based model for construction companies operating in the Slovak economy. In the study of Mihalovic (2016), the author developed two national models created by the MDA and Logit approach. Using the MDA, Ferancova & Sabolova (2015) created a prediction model for the automotive industry companies. None of these prediction models used such a large database of companies as was used in this study. Moreover, a significant part of the mentioned models for Slovak companies was created for specific economic sectors. Many studies have shown that the developed prediction models have a much lower classification capability when used in different economic conditions or at different times than where or when they were developed.

According to this, our study is focused on the creation of a new Slovak national model using a real sample of Slovak companies and their financial data from the period of the years 2016 – 2018. (Kovacova et al., 2019) We apply a CART algorithm to create this model. The originality of our models lies in the use of real data on more than 100,000 Slovak companies. Therefore, the models created in this study reflect the real situation of the Slovak economy. Therefore, our prediction model could be an effective tool for business managers. In this article, we present two models. The first is designed to predict financial distress over a one-year horizon, and the second over a two-year horizon. Both models achieve pretty high prediction ability. We see its novelty in the use of such a large data sample of real Slovak companies, which come directly from their financial statements, and in the application of the data mining CART method, which is one of the modern, currently more used methods achieving very good results.

The rest of the paper is structured as follows. The Methodology describes the data used in this study and the base of the CART algorithm. Section 4 reports the results of estimated models, compares and discusses the results, lists the strengths and weaknesses of this research, and suggests possible future directions for the study. Finally, Section 6 concludes.

### 3.METHODOLOGY

This article presents a CART-based financial distress prediction model created for Slovak companies. As published studies have shown (Karas & Reznakova, 2017; Kumar & Ravi, 2007), the decision tree generating the CART algorithm is often and successfully used in this field. Also, this approach solves some of the problems associated with the use of the classic MDA and Logit methods, which have relatively strong preconditions for their use. CART can be used to model complex relationships between predictors regardless of their probability distribution. This nonparametric algorithm results in a list of simple if-then rules. These rules can be very clearly illustrated in the form of a dendrogram and, in particular, are easy to interpret and practical application. Moreover, the predictive power of these models is at least comparable to the classic MDA and Logit models (Li et al., 2010).

In this study, we used the following methodological procedure. We subjected the data set, containing the values of financial and other characteristics of companies, to a more thorough check of the values of individual variables. In the first step, the prosperity and non-prosperity of the company were determined, based on the criteria mentioned later. Those enterprises for which it was not possible to determine the (non)prosperity due to the missing value of some crucial variable. Then, the checks, edits, and verifications of the values of the variables were performed. The missing values of variables were checked and the zero values of some such variables, for which it is logically and economically justified based on their definition, were supplemented. The correctness of the values of variables, both in terms of logical but also economic correctness was checked. For example, some variables can only acquire positive values (variables  $X_1, X_2, X_{10}, X_{12}, X_{15}, X_{21}, X_{22}$ ). Those companies that do not meet this condition, are excluded from the database because they probably have poorly prepared financial statements (it contains, for example, logical

errors). After performing such a detailed preparation of the data file, we applied the CART method to create a model for classifying a company into a group of prosperous or non-prosperous. As we mention later, this model was created on training and verified on a test sample of companies. We will present a more detailed choice of CART method settings later. The output of this method is a graphical representation of the tree using decision rules. In this study, we present the binomial tree that was created using a test sample of companies. Subsequently, the predictive ability of the model is quantified using a classification table and its classification power is also visualised by the ROC curve and quantified by the AUC criterion. We applied this procedure both for the prediction model created for one year in advance and for the model for two years in advance. The result is relatively simple binomial trees with very good classification ability.

### **3.1. Background of the research, data and variables used**

Slovakia is an important industrial country. The most important industrial sectors include the automotive, electro-technical, engineering, and chemical industries. Slovakia is the world's largest producer of cars per capita. Approximately 85% of Slovak exports go to EU countries. Slovakia has been a member of the Eurozone since 2009, and its currency is the euro. In 2015, GDP per capita in purchasing power parity achieved 76% of the EU average. Slovakia has been one of the fastest-growing economies of the EU in recent years, the third most powerful economy in post-communist countries. In the last five years, real GDP grew by 1.5 to 3.9%, the unemployment rate fell annually on average by 2%, and the inflation rate grew annually by an average of 0.22%. For these reasons, we consider the Slovak economy to be specific and therefore consider it necessary to focus on the development of Slovak national prediction models.

As the main aim of this study is to create a model for predicting the financial distress of Slovak companies one or two years in advance, using the CART generating algorithm, we used the data on real Slovak companies obtained from the database Amadeus. Data contains financial and other characteristics of 104,452 Slovak companies in the year 2016 to 2018. The first reason why this period was chosen was that it did not differ significantly from other periods. In these years, there were no relevant disturbances in the Slovak economy (no economic difficulties, regulations, or other disruptive factors) that could have influenced the database and the trained model. At the same time, it is the first period after the crisis when the Slovak economy grew, and the last consequences of the crisis disappeared. It is why we did not include earlier periods in which the impact of the crisis was significant. (Valaskova et al., 2020)

As independent variables, values of 37 financial indicators (not only ratios) calculated from real financial statements from 2016 to 2018 were used. They are presented and characterized in more details for example in Durica et al. (2019). Of course, these indicators are not only the most frequently used (Korol, 2013) but also those less frequently used, which may consider the specificities of the Slovak economy. For financial distress prediction two years in advance, data from 2016 were used. Analogously, data from 2017 were used to create a one-year prediction model.

In 2016, the Slovak legislation has introduced the institute of "a company in crisis" in §3 of Act no. 7/2005 Coll. on Bankruptcy and Restructuring as amended. According to this Act, the company in crisis is either over-indebted or is unable to repay its obligations. In other words, it is a company with negative equity, or it is unable to pay at least two financial obligations for 30 days after the due date. The company is considered as financial-distressed in case the mention conditions are satisfied and conversely if they were not met, the company was considered a non-financial-distressed one. Therefore, based on data from 2018, companies were initially divided into two groups: ones. Table 1 shows the numbers of companies indicated as financial-distressed and non-financial-distressed in the year 2018.

Table 1

## Separation of companies in 2018

Financial Distress 2018	Frequency	Percentage
No	81,268	77.8
Yes	23,184	22.2
Total	104,452	100.0

Source: own elaboration

### 3.2 CART algorithm

For the creation of the financial distress prediction model, decision trees generating by the CART algorithm is used. The generated DT classifies the cases into two or more groups. CART generates DT by choosing the variable (financial ratios or other indicators of the company) that separate the companies into two sub-groups: financial-distressed companies and non-financial-distressed companies. This division into two groups is carried out with the aim that in each child node there are as many cases of companies from one group and a few from the other as possible. This operation is then repeated until the stop criterion is met. Finally, the model is formed by these final sub-groups and the list of rules on how to achieve all these final sub-groups.

The quality of the generated tree can be measured by the purity of the final sub-groups. CART algorithm usually employed the Gini index for measuring the impurity

$$Gini(\text{node}) = 1 - \sum_{j=1}^k f_j^2 = 1 - \sum_{j=1}^k \left(\frac{n_j}{n}\right)^2,$$

where  $f_i$  are the ratios of the classes (of prosperous or non-prosperous companies) in the node to be predicted. According to the binary classification task, the Gini index varies from the value of 0 (all companies belong to one class) to 0.5 (half of the companies belong to the group of prosperous and half belong to the group of non-prosperous).

The stop criterion often combines several rules. This approach is used in this study, too. The division does not continue if at least one of these conditions is fulfilled:

1. The number of divisions reaches a fixed limit of 5 levels.
2. The group is pure (all companies belong to only one class).
3. The group does not contain more than one hundred companies.
4. Some of the sub-groups do not contain more than fifty companies.
5. The next division does not significantly increase the quality of the tree (the minimum change in improvement is 0.0001).

Often, the CART tree can be overfitting or overtraining. Due to this problem, as a first, the CART algorithm generates a maximum tree (concerning the above-mentioned stop criterion). Then, the algorithm applies cost-complexity pruning of the maximum tree. According to this, it is sufficient to create a test sample, separate from a training sample to carry out this detection, where the ratio of these two samples is 80:20. On the test sample, each sub-tree of the maximum tree is testing and the best-pruned three is the one with the lowest error rate in testing.

### 3.3 Classification ability

The classification ability or accuracy of the financial distress models is evaluated by analysis of a classification table, containing the numbers and frequencies of companies that are classified correctly and incorrectly into considered two classes. In the case of data, used to create a model, the overall accuracy is

often overestimated. To improve this classification, the dataset is divided into the training sample (80% of all data) and test sample (the rest 20% of the data). The training sample is used for generating the model and the test sample is used for calculating the classification of the model. The classification ability is also checked using the ROC curve and quantified by the value of the AUC criterion, where the value of the good is close to the value of 1.

#### **4. EMPIRICAL RESULTS AND DISCUSSION**

Korol (2013) claims that financial distress prediction models generated as decision trees usually achieve higher classification capability compared to models created using other widely used methods, including classical MDA or Logit or even ANNs. These findings and the same CART advantages discussed hereabove represent the main motivation to create the Slovak national CART-based model that is presented in this study.

In this study, the sample of 104,542 Slovak companies is used, among them 81,268 (i.e., 77.80%) non-financial-distressed companies and 23,184 (i.e., 22.20%) companies in financial distress (Table 3). Their financial situation is considered in 2018, based on the mentioned criteria of a company in crisis. This dataset is randomly divided into two samples. The training sample is the first one and consists of 80% of all companies. The remaining companies form the test sample. The prediction model is developed on the training sample, while the test sample serves to quantification of its classification capability. The stop criterion for generating CART decision trees is described in the Methodology section. According to these conditions, as a first, the algorithm generates the maximum tree, and then this tree is pruned to avoid overfitting and achieve good generalizability of the final model. Overall accuracy. Gini index serves as an impurity function during the generating of the tree.

In the creation procedure, 37 financial indicators (Durica et al., 2019) calculated using real financial statements from 2016 and 2017, respectively, are used. Data from 2016 is used to create a two-year model, and the data from 2017 is used for one-year prediction.

##### **4.1 Two-years model**

Using the CART algorithm, the final CART-based model with five levels of sub-groups, and six final sub-groups was developed (Figure 1). During the generating of the maximum tree, values of 35 financial ratios (of all 37) were used. However, after the automatic pruning of this tree, the final model works with values of 3 ratios only. A closer look at the tree graph shows that the algorithm evaluated the variable *X10* as the "most important" variable in the first splitting criterion, where the value of this variable less than about 1 (exactly 1.00071) indicate the prosperity of the company. The value of the variable *X10* higher than 1 is further divided in the next step of the tree in the division point 1.28536, where enterprises with the value *X10* higher than this division point are definitively marked by the model as non-prosperous. For the second category of enterprises for which  $X10 \leq 1.28536$  the division proceeds in the next three steps, using the variable *X1*, *X4* and again *X10*. These three ratios, *X10* (Total Liabilities to Total Assets), *X01* (Sales to Total Assets) and *X04* (Net Income to Equity), are some of the most frequently used indicators in the prediction models (Korol, 2013).

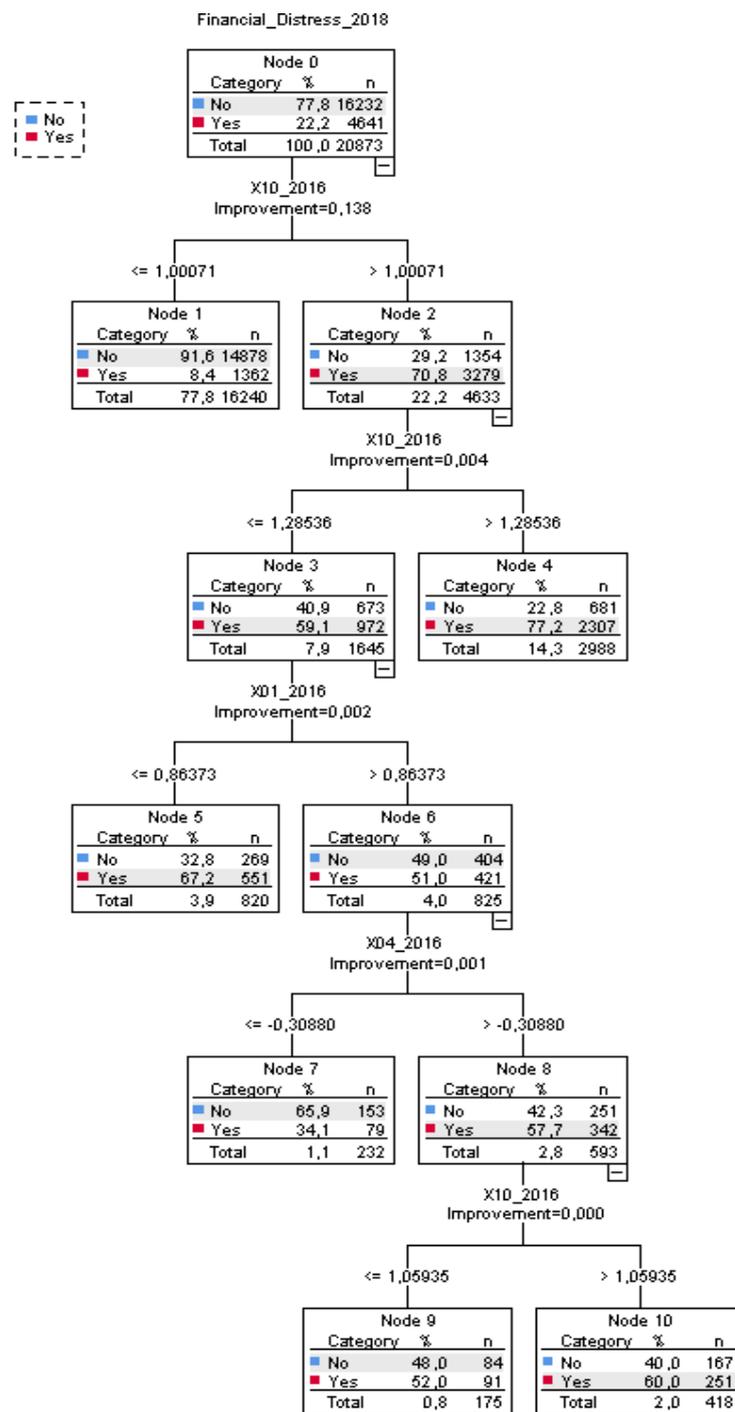


Figure 1. Decision tree for two-years prediction.

Source: own elaboration

The classification table (Table 2) shows that the quality of the developed model is pretty high on the training and the test sample, too. The classification capability is more than 87%. More than 93% of the non-financial-distressed companies and almost 69% of the financial-distressed companies in the test sample are classified correctly.

Table 2

Classification table of a two-year model

Sample	Observed	Predicted		
		No	Yes	Percent Correct
Training	No	60,550	4,476	93.12%
	Yes	5,811	12,730	68.66%
	Overall Percentage	79.41%	20.59%	87.69%
Test	No	15,115	1,117	93.12%
	Yes	1,532	3,109	66.99%
	Overall Percentage	79.75%	20.25%	87.31%

Source: own elaboration

The classification capability of the model, expressed using the ROC curve, is shown in the following figure. The value of the AUC area for this model is 0.859.

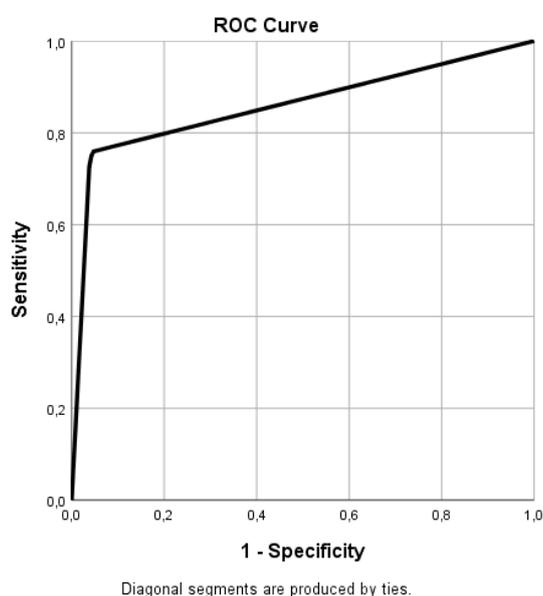


Figure 2. ROC curve for the two-years model

Source: own elaboration

## 4.2 One-year model

Again, using a CART algorithm with the same settings, a decision tree was grown. For growing of maximum tree, values of 33 ratios were used. After the pruning, the final DT with three levels of sub-groups, and three final sub-groups was generated (Figure 3). In this case, the final model uses only two variables as splitting criteria. This model, like the previous one, considers *X10* to be the first decision variable, also with a splitting criterion of about 1 (exactly 1.00011). Similarly, as in the previous model, a value of less than about 1 means the definitive classification of the company into the class of prosperous companies. A value of variable *X10* greater than 1 is further divided in the second step of the tree at the dividing point 1.14723 of the variable *X10*. Companies with a value of *X10* higher than this value will be classified by the model as non-prosperous. The others are further divided into two groups using the variable *X4*. This model is quite similar to the two-year model, but it is a bit simpler and does not work with the value of the variable *X1*. Instead, the values of *X10* (Total Liabilities to Total Assets) and *X04* (Net Income to Equity) are used.

Although the model predicting financial distress is simple, its accuracy is also high for non-financial-distressed companies, as well as expected. It is because the values of the financial ratios reflect the threatening financial distress one year in advance better than two years. Table 3 describes the classification capability of the one-year model. On the test sample, the classification capability is nearly 92%. Nearly 95% of non-financial-distressed companies in this sample were classified correctly. In the financial-distressed class, more than 81% of companies were correctly classified into this class.

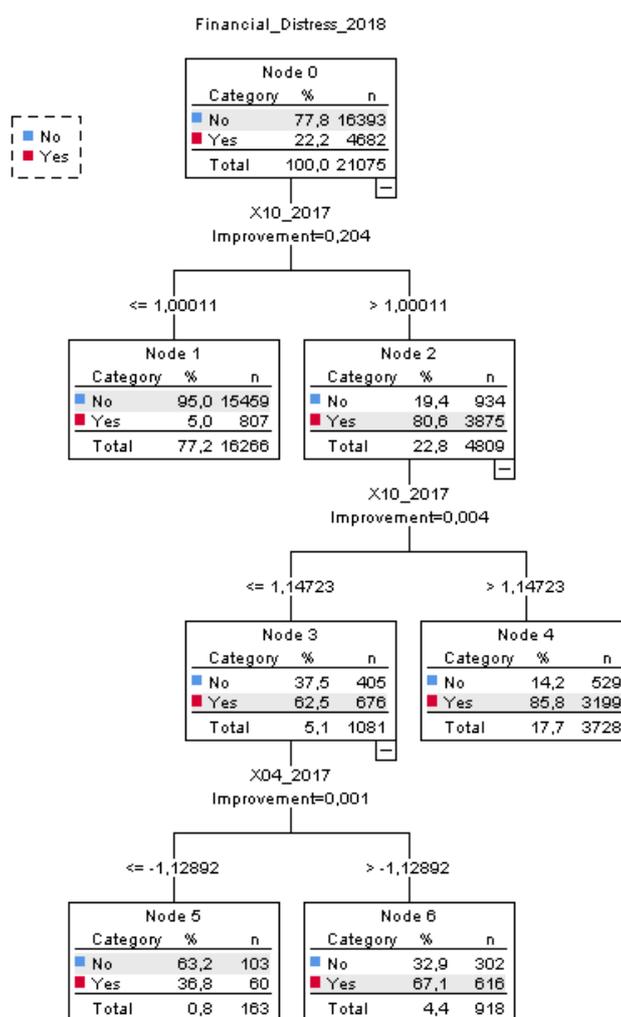


Figure 3. Decision tree of the one-year model

Source: own elaboration

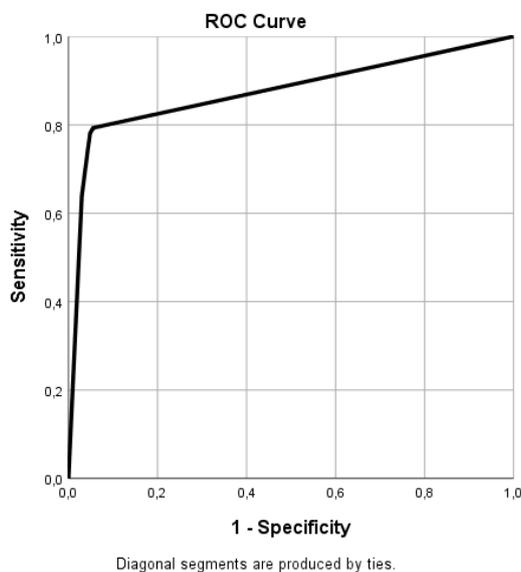
Table 3

Classification table of the one-year model.

Sample	Observed	Predicted		
		No	Yes	Percent Correct
Training	No	61,765	3,110	95.21%
	Yes	3,418	15,084	81.53%
	Overall Percentage	78.18%	21.82%	92.17%
Test	No	15,562	831	94.93%
	Yes	867	3,815	81.48%
	Overall Percentage	77.95%	22.05%	91.94%

Source: own elaboration

Again, the classification ability of the model was checked also using the value of AUC. Figure 4 shows the ROC curve for this model. The AUC criterion has a value of 0.875 for this model.



**Figure 4. ROC curve of the one-year model**

*Source:* own elaboration

### 4.3 Discussion

Based on the evaluation of the developed models, it can be concluded that both achieved pretty high accuracy, especially in the class of non-financial-distressed companies. As it can be expected, the two-years model achieved a little bit lower classification capability of 87.3%, while the one-year model achieved a higher classification capability of 91.9%.

In Slovakia, several prediction models were created using various methods recently. Among the most used ones belong to the Logit model created by Gulka (2016) with an accuracy of almost 80%. To model financial distress, Kovacova & Kliestik (2017) also used logistic regression and achieved an overall prediction accuracy of 86.6%. Mihalovic (2016) developed two financial distress prediction models using MDA and Logit that have an overall accuracy of 64.4% and 68.6%, respectively. The abovementioned models were created for the prediction of the financial distress situation one year in advance. Nevertheless, they achieve a lower accuracy than our model two-years model, with an overall accuracy of 87.3%. Moreover, the overall accuracy of these models mentioned above is much lower than the accuracy of both of our new models.

The quality of the models presented in this study is indicated by the ratio of correct classified companies, but also by the AUC value. It is clear, that the two-years model reached slightly lower values of these criteria. In general, however, the CART method is one of the modern machine learning methods that are based on self-learning from data able to detect mainly linear relationships very well in the data and derive classification rules that with a high percentage of success classify companies into one of two groups. This is its advantage over older methods such as MDA or Logit. Several authors in Slovakia have also been created prediction models based on the DT. In 2017, Gavurova et al. presented the two Slovak national models for the one-year prediction and also for two years prediction using MDA and the CHAID decision tree, both with the prediction accuracy of almost 87%. The CHAID algorithm-generated models with higher overall accuracy. Our model of predicting the financial distress two years in advance achieve approximately the same overall accuracy. However, the classification capability of our one-year model is higher by 5%.

Karas and Reznakova (2017) created two CART-based models for Slovak companies. Using these models, the financial status of financial-distressed companies is predicted with an accuracy of 94.9% and 91.5%, respectively. But the prediction ability in the class of financially healthy companies is only 61% or 62.6%, respectively. Thus, the situation is reversed for our one-year prediction model. This is because our model achieved a better classification of financially healthy companies (almost 95%), but that was also its goal. For financial-distressed companies, our model achieved a classification accuracy of 81.5%. However, the overall accuracy of our model is 91.9%, which is higher by more than 26% in comparison with the abovementioned model of Karas and Reznakova.

## 5. CONCLUSION

Although the issue of the creation of the prediction model was in recent decades in the focus of many economists from Slovakia and also from other countries, we can claim that a widely used Slovak national prediction model still does not exist. More models have been created recently, but none of them has yet become widely recognized and practically used in Slovakia. To fill this gap, the main goal of this article was to present a new model or models created using the CART algorithm. Thus, the created models are decision trees designed for the one-year and two-year prediction of the impending financial distress situation. The database was the data of more than one hundred thousand companies accounting in the conditions of the Slovak Republic. These were data from the years 2016 to 2018. The models predict the financial status of the company in 2018, taking into account the currently valid Slovak legislation.

Although both developed models classified companies based on two financial ratios (Total Liabilities to Total Assets and Net Income to Equity) and three ratios (Total Liabilities to Total Assets, Sales to Total Assets, and Net Income to Equity), respectively, they have high overall accuracy. Anyway, the simplicity of the models is their advantage. This is because our models are easy to interpret and they are very applicable for the real prediction of financial distress for companies with incomplete accounting data. Finally, the models achieved a high classification accuracy of 87.3% for prediction two years in advance and even 91.9% for one-year prediction.

A further direction of this research lies in the verification and, where appropriate, adjustment of the model on the newer accounting data of real Slovak companies with the ambition of constructing a generally acceptable Slovak national prediction model. Another further direction of this study should also be the creation of partial prediction models for companies that operated in some selected economic sectors or operated in some selected Slovak regions.

## ACKNOWLEDGEMENT

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 the Faculty of Operation and Economics of Transport and Communications, University of Zilina, Slovakia, Institutional Grant No. 11/PEDAS/2019: Using selected data mining methods in business failure prediction.

## REFERENCES

- Affes, Z., & Hentati-Kaffel, R. (2018). Forecast bankruptcy using a blend of clustering and MARS model: case of US banks. *Annals of Operations Research*, 281(1-2), 27–64. doi:10.1007/s10479-018-2845-8
- Alaminos, D., Castillo, A. D., & Fernández, M. Á. (2016). A Global Model for Bankruptcy Prediction. *Plos One*, 11(11). doi: 10.1371/journal.pone.0166693
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy. *The Journal of Finance*, 23(4), 589–609. doi: 10.1111/j.1540-6261.1968.tb00843.x

- Altman, E. I. (1993). *Corporate financial distress and bankruptcy: a complete guide to predicting and avoiding distress and profiting from bankruptcy*. New York: Wiley.
- Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting Review*, 38(1), 63–93. doi: 10.1016/j.bar.2005.09.001
- Boettcher, J., Cavanagh, G., & Xu, M. (2014). Ethical Issues that arise in Bankruptcy. *Business and Society Review*, 119(4), 473–496. doi: 10.1111/basr.12042
- Breiman, L., Friedman, J., Stone, C. J., & Olsen, R. A. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.
- Brozyna, J., Mentel, G., & Pisula, T. (2016). Statistical methods of the bankruptcy prediction in the logistics sector in Poland and Slovakia. *Transformations in Business & Economics*, 15(1), 93–114.
- Chen, M.-Y. (2011). Predicting corporate financial distress based on the integration of decision tree classification and logistic regression. *Expert Systems with Applications*, 38(9), 11261–11272. doi: 10.1016/j.eswa.2011.02.173
- Chrastinova, Z. (1998). *Metody hodnotenia ekonomickej bonity a predikcie financnej situacie polnohospodarskych podnikov*. Bratislava, Slovakia: VEUPP.
- Cohen, S., Doumpos, M., Neofytou, E., & Zopounidis, C. (2012). Assessing financial distress where bankruptcy is not an option: An alternative approach for local municipalities. *European Journal of Operational Research*, 218(1), 270–279. doi: 10.1016/j.ejor.2011.10.021
- Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, 90(3), 487–513. doi: 10.1016/0377-2217(95)00070-4
- du Jardin, P., & Severin, E. (2012). Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research*, 221(2), 378–396. doi: 10.1016/j.ejor.2012.04.006
- Durica, M., Frnda, J., & Svabova, L. (2019). Decision tree based model of business failure prediction for Polish companies. *Oeconomia Copernicana*, 10(3), 453–469. doi: 10.24136/oc.2019.022
- Frydman, H., Altman, E. I., & Kao, D.-L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *The Journal of Finance*, 40(1), 269–291. doi: 10.1111/j.1540-6261.1985.tb04949.x
- Gandolfi, G. C., Regalli, M. C., Soana, M.-G. C., & Arcuri, M. C. (2018). Underpricing And Long-Term Performance Of Ipos: Evidence From European Intermediary-Oriented Markets. *Economics, Management, and Financial Markets*, 13(3), 11–36. doi: 10.22381/emfm13320181
- Gavurova, B., Janke, F., Packova, M., & Pridavok, M. (2017). Analysis of Impact of Using the Trend Variables on Bankruptcy Prediction Models Performance. *Ekonomicky casopis*, 65(4), 370–383.
- Georgescu, V. (2017). Using genetic algorithms to evolve type-2 fuzzy logic systems for predicting bankruptcy. *Kybernetes*, 46(1), 142–156. doi: 10.1108/k-06-2016-0152
- Gregova, E., Valaskova, K., Adamko, P., Tumpach, M., & Jaros, J. (2020). Predicting Financial Distress of Slovak Enterprises: Comparison of Selected Traditional and Learning Algorithms Methods. *Sustainability*, 12(10), 3954. doi:10.3390/su12103954
- Gulka, M. (2016). Predictive Model of Corporate Failure in the Slovak Business Environment. *Forum Statisticum Slovacum*, 12(1), 16–22.
- Gurcik, L. (2012). G-index – the financial situation prognosis method of agricultural enterprises. *Agricultural Economics*, 48(8), 373–378. doi: 10.17221/5338-agricecon
- Hurtosova, J. (2009). *Konstrukcia ratingoveho modelu, nastrojabadnotenia uverovej sposobnosti podniku* (dissertation). University of Economics in Bratislava, Slovakia.
- Irimia-Dieguez, A. I., Blanco-Oliver, A., & Vazquez-Cueto, M. J. (2015). A Comparison of Classification/Regression Trees and Logistic Regression in Failure Models. *Procedia Economics and Finance*, 23, 9–14. doi: 10.1016/s2212-5671(15)00493-1
- Jones, S., Johnstone, D., & Wilson, R. (2016). Predicting Corporate Bankruptcy: An Evaluation of Alternative Statistical Frameworks. *Journal of Business Finance & Accounting*, 44(1-2), 3–34. doi: 10.1111/jbfa.12218

- Karas, M., & Reznakova, M. (2017). Predicting the Bankruptcy of Construction Companies: A CART-Based Model. *Engineering Economics*, 28(2), 145–154. doi: 10.5755/j01.ee.28.2.16353
- Karas, M., & Srbova, P. (2019). Predicting bankruptcy in construction business: Traditional model validation and formulation of a new model. *Journal of International Studies*, 12(1), 283–296. doi: 10.14254/2071-8330.2019/12-1/19
- Kasgari, A. A., Divsalar, M., Javid, M. R., & Ebrahimian, S. J. (2012). Prediction of bankruptcy Iranian corporations through artificial neural network and Probit-based analyses. *Neural Computing and Applications*, 23(3-4), 927–936. doi: 10.1007/s00521-012-1017-z
- Kliestik, T., Vrbka, J., & Rowland, Z. (2018). Bankruptcy prediction in Visegrad group countries using multiple discriminant analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 13(3), 569–593. doi: 10.24136/eq.2018.028
- Kliestikova, J., Misankova, M., & Kliestik, T. (2017). Bankruptcy in Slovakia: international comparison of the creditor's position. *Oeconomia Copernicana*, 8(2), 221–237. doi: 10.24136/oc.v8i2.14
- Korol, T. (2013). Early warning models against bankruptcy risk for Central European and Latin American enterprises. *Economic Modelling*, 31, 22–30. doi: 10.1016/j.econmod.2012.11.017
- Kovacova, M., Valaskova, K., Durana, P., & Kliestikova, J. (2019). Innovation Management of the Bankruptcy: Case Study of Visegrad Group Countries. *Marketing and Management of Innovations*, (4), 241–251. doi: 10.21272/mmi.2019.4-19
- Kovacova, M., & Kliestik, T. (2017). Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilibrium*, 12(4), 775–791. doi: 10.24136/eq.v12i4.40
- Kovacova, M., Kliestik, T., Kubala, P., Valaskova, K., Radisic, M., & Borocki, J. (2018). Bankruptcy Models: Verifying Their Validity As A Predictor Of Corporate Failure. *Polish Journal of Management Studies*, 18(1), 167–179. doi: 10.17512/pjms.2018.18.1.13
- Koyuncugil, A. S., & Ozgulbas, N. (2012). Financial early warning system model and data mining application for risk detection. *Expert Systems with Applications*, 39(6), 6238–6253. doi: 10.1016/j.eswa.2011.12.021
- Kral, P., Musa, H., Lazaroiu, G., Misankova, M., & Vrbka, J. (2018). Comprehensive assessment of the selected indicators of financial analysis in the context of failing companies. *Journal of International Studies*, 11(4), 282–294. doi: 10.14254/2071-8330.2018/11-4/20
- Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. *European Journal of Operational Research*, 180(1), 1–28. doi: 10.1016/j.ejor.2006.08.043
- Li, H., & Sun, J. (2009). Gaussian case-based reasoning for business failure prediction with empirical data in China. *Information Sciences*, 179(1-2), 89–108. doi: 10.1016/j.ins.2008.09.003
- Li, H., Sun, J., & Wu, J. (2010). Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*, 37(8), 5895–5904. doi: 10.1016/j.eswa.2010.02.016
- Liu, J., & Wu, C. (2017). A gradient-boosting decision-tree approach for firm failure prediction: an empirical model evaluation of Chinese listed companies. *The Journal of Risk Model Validation*, 11(2), 43–64. doi: 10.21314/jrmv.2017.170
- Mihalovic, M. (2016). Performance Comparison of Multiple Discriminant Analysis and Logit Models in Bankruptcy Prediction. *Economics & Sociology*, 9(4), 101–118. doi: 10.14254/2071-789x.2016/9-4/6
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. doi: 10.2307/2490395
- Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications*, 36(1), 2–17. doi: 10.1016/j.eswa.2007.10.005
- Shin, K.-S., & Lee, Y.-J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321–328. doi: 10.1016/s0957-4174(02)00051-9
- Tsai, C.-F., Hsu, Y.-F., & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24, 977–984. doi: 10.1016/j.asoc.2014.08.047

- Valaskova, K., Durana, P., Adamko, P., & Jaros, J. (2020). Financial Compass for Slovak Enterprises: Modeling Economic Stability of Agricultural Entities. *Journal of Risk and Financial Management*, 13(5), 92. doi:10.3390/jrfm13050092
- Valaskova, K., Kliestik, T., & Kovacova, M. (2018). Management of financial risks in Slovak enterprises using regression analysis. *Oeconomia Copernicana*, 9(1), 105–121. doi: 10.24136/oc.2018.006
- Virag, M., & Nyitrai, T. (2014). Is there a trade-off between the predictive power and the interpretability of bankruptcy models? The case of the first Hungarian bankruptcy prediction model. *Acta Oeconomica*, 64(4), 419–440. doi: 10.1556/aoecon.64.2014.4.2
- Zhao, D., Huang, C., Wei, Y., Yu, F., Wang, M., & Chen, H. (2016). An Effective Computational Model for Bankruptcy Prediction Using Kernel Extreme Learning Machine Approach. *Computational Economics*, 49(2), 325–341. doi: 10.1007/s10614-016-9562-7
- Zieba, M. K., Tomczak, S. M., & Tomczak, J. undefined. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, 58, 93–101. doi: 10.1016/j.eswa.2016.04.001
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22, 59–82. doi: 10.2307/2490859