FINANCIAL DISTRESS CLASSIFICATION

Mária Stachová – Pavol Kráľ

Abstract

Financial problems can grow to enormous proportions and cause large losses or affect the operation of several entities in the near and distant surroundings of the company, it is why is important to identify, predict and then try to detect these problems with sufficient accuracy and preferential financial connection. Therefore, the main goal of this contribution is to create a mechanism that is able to notice the financial menace.

The paper deals with the possibility of predicting the financial health of selected companies of the Slovak Republic, and continues with comparison of the created mathematical-statistical model that are applied to multi-year real data. These data consist of ten financial indicators/predictors of companies, nine of them are ratio indicators and one is rate indicator. We chose to use two methods to build the prediction models. The first one is well-known classification tree model based on Breiman's CART algorithm from R package "rpart" and the second one is its modification, conditional inference tree model from R package "party".

Key words: financial distress, classification, prediction models, tree-based models

JEL Code: C38, G33

Introduction

The financial situation is the state of the company, which is observable from the outside and is the result of stability/instability of the business economy or the effectiveness of financial management. Bigger commotion is caused by negative events, and not by the positive ones. In case of business entities is the situation very similar. We are more likely to hear about companies that are experiencing financial difficulties, drowning in debt and in danger of bankruptcy, than about companies that are doing well and experiencing a period of growth or stabilization.

Financial distress or bankruptcy of a company brings affect not only the economic environment, but often also affect the social environment and the political environment. It is in the best interest of every company to prevent financial problems, which is why prediction options have been evolving and improving since the 1960s (Altman, 1968).

The trend of finding the proper and effective way to predict the financial distress of companies continues until nowadays. We can list many authors and many studies that deal with this topic. E.g. in (Boďa & Úradníček, 2016; Balcean & Ooghe, 2006) one can find financial prediction models constructed using various classification models such as discriminant analysis, logistic regression or decision trees. Our hypothesis is based on (Kráľ et al., 2014; Stachová et al. 2015) and we believe, that time gap incorporated into these well-known models can improve their predictive accuracy.

In our contribution, we restricted ourselves to an approach where predictors and a response variable in our model are from different years and the time gap between them is one or three years. We chose to estimate classification tree model using the CART algorithm from R (R Core Team, 2021) package "rpart" (Therneau & Atkinson, 2019) and its modification, conditional inference tree model from R package "party" (Hothorn et al., 2006).

1 Data and Methodology

The data set is created by the financial statements (balance sheet and income statement) of a selected group of business units accounting in the double-entry bookkeeping system in the periods 2016, 2018 and 2019 operating in the Slovak Republic.

The data set consists of financial information about 559 companies, for which we had 9 financial ratios and one rate indicator. We chose these indicators because of their linear independence and non-correlation, whereas we took into account previous results of similar research provided by Altmana (1968), Karas and Režňáková (2013), Úradníček et al. (2016) and Boďa and Úradníček (2021). The names, description, and method of calculating predictors are provided in Table 1.

As a predicted variable, we take financial status of the company in the 2019. This variable is called the default and takes two possible alternatives values: 0 or 1. An enterprise with default value 0 is assigned as a prosperous enterprise with no financial distress. The second group consists of companies with a default value 1 and they are considered to be non-prosperous companies at risk of financial distress. To label companies being and not being financially distressed we used so-called key that was inspired by the research of Boďa and Úradníček (2021). The key consists of the following three criteria:

- Earnings before interest, taxes, depreciation and amortization (EBITDA) is less than interest expense,
- earnings after taxes (EAT) is less than 0,
- equity (EQUITY) is less than 0.

All three criteria have to be meet to label the company as financially distressed.

Earnings before interest, taxes, depreciation and amortization is usually understood as free cash flow and is designed to signal financial difficulties if a company's cash flow is not sufficient to meet the price required to repay its debt. Earnings after tax indicates the financial distress of an enterprise if the enterprise is loss-making in all its activities. Finally, equity may point to excessive corporate debt.

Tab. 1: Names, description and method of calculation of financial indicators used in the construction of predictive models for Slovak companies

Label	Description	Method of calculation
OROA	operating return on assets	earnings before income and taxes/ total assets
IC	interest coverage	earnings before income and taxes / interest expense
EM	equity multiplier	total liabilities / equity
RWC	relative working capital	(current assets – current liabilities / total assets
CTO	cash turnover	sales / cash and cash equivalents
RTO	receivables turnover	sales / short-term and long-term receivables
ITO	inventory turnover	sales / inventory
WCTO	working capital turnover	sales / (current assets – current liabilities)
RES	retained earnings share	retained earnings/ total assets
LASS	logarithm of assets	natural logarithm of total assets in 1000 €

Source: Boďa and Úradníček (2021)

We hypnotize, that in companies that find themselves in financial difficulties, it is possible to observe signals indicating unfavorable developments in the future. This assumption makes it possible to predict the financial condition of the company through data from previous periods. Therefore, the values of the predictors and the predicted variable are not from the same time. The difference between these time periods represents the length of time for which the financial health of the company is to be predicted. We have chosen a one-year and three-year as the forecast horizon. This means that the values of predictors come from 2016 and 2018 and the value of the predicted variable comes from year 2019. In our data set, there were 559 enterprises, 27 enterprises in financial distress. Which is about 4.83%.

At the beginning, we divided our data set into training group and testing group (70:30), then the two different algorithms were estimated on training data sets. The first one is classification tree model using the CART algorithm and the second one is its modification, conditional inference tree model. The test data were used to evaluate overall accuracy of these models.

2 Results of CART algorithm

The first method for estimating a prediction model was a decision tree. We used the rpart package (Therneau et.al., 2017), which includes the CART algorithm. The resulting trees are shown in Figure 1-2.

Figure 1 shows the classification tree model constructed on the predictors from 2016. The model divided the enterprises in the root node according to the EM indicator (equity multiplier). Enterprises that meet the top condition are labelled as not at financial at risk (0). Other companies are divided in the next node according to the condition of receivables turnover (RTO).

The predictive ability of the estimated classification tree is expressed in Table 2 in the form of a classification matrix.

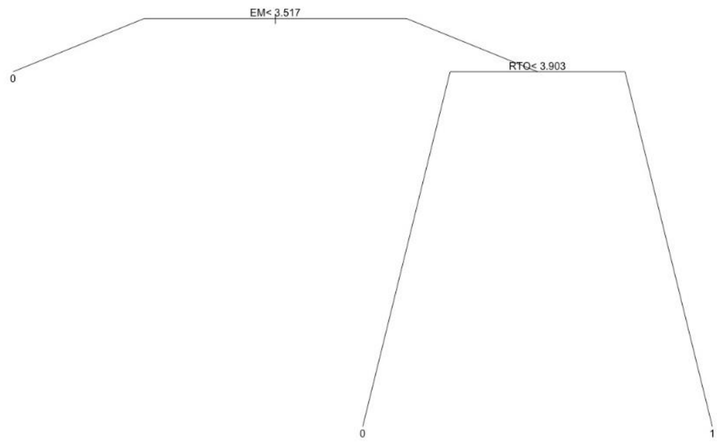
We can see that prediction accuracy of model on training data set is a little higher, that is expectable. The difference is remarkable in the case of classification of financially distressed companies. The low accuracy may be caused by fact, that we have lack of financially unhealthy companies in data set.

Tab. 2: Classification table of the decision tree estimated on the predictors from 2016 and financial status from 2019

Actual class	Predicted class					
	Train data set			Test data set		
	0	1	% correctly predicted comp.	0	1	% correctly predicted comp.
0	373	1	99,73	155	3	98,10
1	11	6	35,29	10	0	0
Overall accuracy	96,93 %			92,26 %		

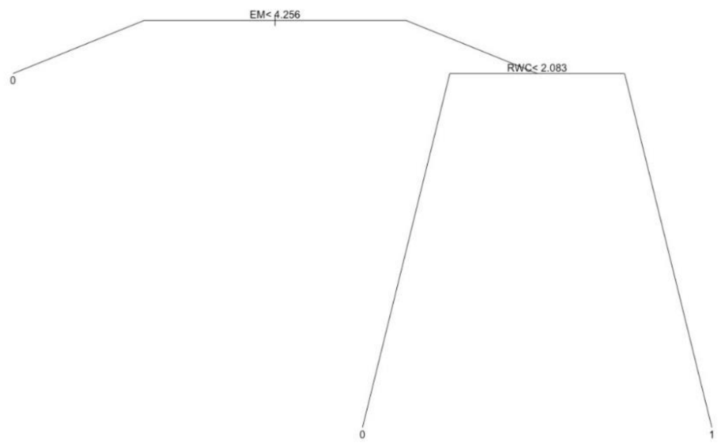
Source: The author's work

Fig. 1.: Classification tree estimated on the predictors from 2016 and financial status from 2019



Source: The author's work

Fig. 2.: Classification tree estimated on the predictors from 2018 and financial status from 2019



Source: The author's work

Figure 2 shows a model of the classification tree estimated on data from 2018, in this case it is a one-year prediction. Enterprises are divided in the root node on the basis of the EM indicator (equity multiplier), and if the enterprise does not meet the condition in this node, there is a further division, this time on the basis of relative working capital (RWC).

The predictive power of the created classification tree is expressed in the form of a classification matrix in Table 3.

Tab. 3: Classification table of the decision tree estimated on the predictors from 2018 and financial status from 2019

Actual class	Predicted class					
	Train data set			Test data set		
	0	1	% correctly predicted comp.	0	1	% correctly predicted comp.
0	372	2	99,47	155	3	98,10
1	12	5	29,41	9	1	0,10
Overall accuracy	96,42 %			92,86 %		

Source: The author's work

Table 3 shows that the model constructed on one-year data is very similar in its prediction capability to the three-year prediction model. The main difference is in the overall accuracy of the model performed on test data, where the model improved by 0,6 percentage points.

3 Results of conditional inference tree algorithm

Another method applied to construct a prediction model was a conditioned decision tree. We used the “party” package (Hothorn, Hornik & Zeileis 2006), which estimates the regression relationship using recursive partitioning in a conditional inference structure. The resulting trees are shown in Figures 3-4.

Figure 3 shows the estimated model on data from 2016. The model divided the enterprises in the root node based on the EM (equity multiplier Enterprises that meet the top condition fall into the left node and are classified as not at financial risk (0). In the next node, other companies are divided according to the criterion of asset logarithm (LASS) and receivables turnover (RTO). We can also see the node to which companies in financial distress will fall, namely Node 6, to which the company belongs if its equity multiplier is greater than 3,506, its logarithm of assets is smaller or equal to 1,932 and the receivables turnover is greater than 1,756.

The predictive ability of the created conditional classification tree is shown in Table 4 in the form of a classification matrix.

We can see that the predictive accuracy of the model is little bit higher in the case of training data set comparing to the CART model. Again, the model achieves a worse degree of classification quality on the test set.

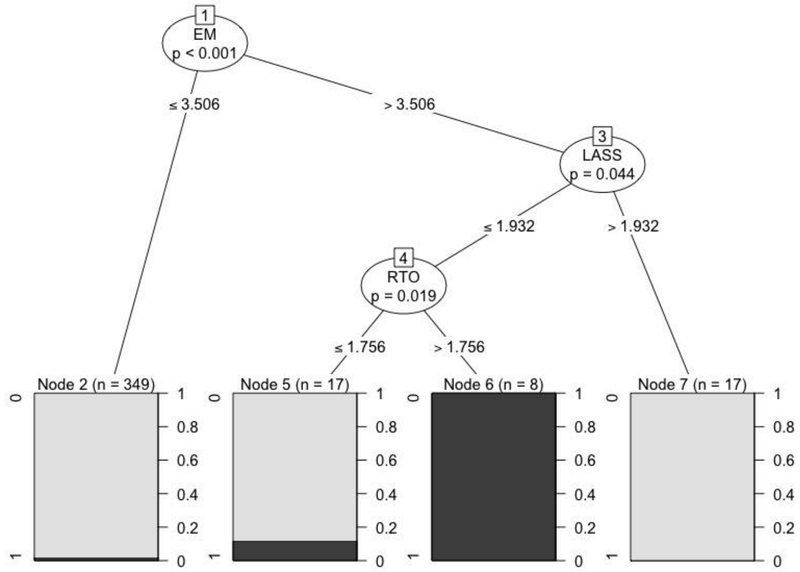
Tab. 4: Classification table of the decision tree estimated on the predictors from 2016 and financial status from 2019

Actual class	Predicted class					
	Train data set			Test data set		
	0	1	% correctly predicted comp.	0	1	% correctly predicted comp.
0	374	0	100	154	4	97,47

1	9	8	47,06	10	0	0
Overall accuracy	97,70 %			91,67 %		

Source: The author's work

Fig. 3.: Conditional inference tree estimated on the predictors from 2016 and financial status from 2019



Source: The author's work

Figure 4 shows a model of a conditional classification tree constructed on data from 2018, again predicting the financial health of the company in 2019. In this case, the model was not able to divide companies into nodes and marked all companies as financially healty.

Although it is clear how the classification matrix will look like in this case, we present it here for completeness in Table 5.

Fig. 4.: Conditional inference tree estimated on the predictors from 2018 and financial status from 2019

Source: The author's work

Tab. 5: Classification table of the decision tree estimated on the predictors from 2018 and financial status from 2019

Actual class	Predicted class					
	Train data set			Test data set		
	0	1	% correctly predicted comp.	0	1	% correctly predicted comp.
0	374	0	100	158	0	100
1	17	0	0,00	10	0	0
Overall accuracy	95,65 %			94,05 %		

Source: The author's work

Conclusion

In the paper, we follow our previous analysis on financial distress prediction (Kráľ et al., 2016) and (Stachová & Kráľ, 2019). We use more recent financial data of Slovak enterprises covering different economic activities. We hypothesized that knowledge of the time evolution of financial indicators can help predict future changes in the financial health of the company. Thus, we compared the predictive power of selected financial indicators to estimate the financial distress of Slovak companies, which can potentially occur in a time horizon of one or three years. We were also interested in comparison of predictive ability of two different classification methods, namely classification tree model based on CART algorithm and its modification conditional inference tree model.

Both models performed quite very well, their overall accuracy overrun 90% and they can be considered as very useful tools in decision-making process in identification of financially distressed companies.

Acknowledgment

The work has been supported by project VEGA No. 1/0150/21 (Profit testing of pension insurance products).

References

Altman, E. I., (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, 23(4), 583-609.

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.

Boďa, M., & Úradníček, V. (2016). The portability of Altman's Z-score model to predicting corporate financial distress of Slovak companies. In Technological and Economic Development of Economy, 22(4), 532-553.

Boďa, M., Úradníček, V. (2021). Definition of financial distress supported by data in Slovak corporate conditions. European Journal of International Management, Forthcoming, doi: 10.1504/EJIM.2021.10033923.

Brezigar - Masten, A., & Masten, I. (2012). CART-based selection of bankruptcy predictors for the logit model. Expert Systems with Applications, 39(11), 10153–10159.

Hothorn T, Hornik K, Zeileis A (2006). "Unbiased Recursive Partitioning: A Conditional Inference Framework." *Journal of Computational and Graphical Statistics*, 15(3), 651–674.

Karas, M., Režňáková, M. (2013). Bankruptcy Prediction Model of Industrial Enterprises in the Czech Republic. In International Journal of Mathematical Models and Methods in Applied Sciences, 7(5).

Kráľ, P., Fleicher, M., Stachová, M., Nedelová, G., Sobíšek., L. (2016) Corporate financial distress prediction of Slovak companies: Z-score models vs. Alternatives. In the 19th AMSE, international scientific conference, conference proceedings, Slovakia, 224-231.

Kráľ, P., Stachová, M. & Sobíšek, L. (2014). Utilization of repeatedly measured financial ratios in corporate financial distress prediction in Slovakia. In the 17th AMSE, international scientific conference, conference proceedings, Poland, 156-163.

Stachová, M., Kráľ, P., Sobíšek, L., & Kakaščík, M. (2015). Analysis of financial distress of Slovak companies using repeated measurement. In 18th AMSE, International Scientific Conference Proceedings, Czech Republic.

Stachová, M., Kráľ, P. (2019), Statistical Learning Methods in Corporate Financial Distress Prediction of Slovak Enterprises: Comparison of Alternative Models, In The 13th international days of statistics and economics, conference proceedings, Slaný : Melandrium, 1438-1446.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Terry Therneau and Beth Atkinson (2019). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-15. https://CRAN.R-project.org/package=rpart

Úradníček, V. et al. (2016). Variantné metódy predikcie finančného zdravia podnikov v podmienkach dynamického ekonomického prostredia. Banská Bystrica : Belianum. Vydavateľstvo Univerzity Mateja Bela v Banskej Bystrici. Ekonomická fakulta, 136 p.

Contact

Mária Stachová

Faculty of Economics, Matej Bel University

Tajovského 10, 975 90 Banska Bystrica, Slovakia

maria.stachova@umb.sk

Pavol Kráľ

Faculty of Economics, Matej Bel University

Tajovského 10, 975 90 Banska Bystrica, Slovakia

pavol.kral@umb.sk