ORIGINAL ARTICLE


Contact to corresponding author: Marek Durica, marek.durica@uniza.sk

Article history: Received: 1.04.2023; Accepted: 7.12.2023; Published online: 30.12.2023

Marek Durica
*University of Zilina, Slovakia*
[orcid.org/0000-0002-2118-166X](https://orcid.org/0000-0002-2118-166X)

Jaroslav Frnda
*University of Zilina, Slovakia*
[orcid.org/0000-0001-6065-3087](https://orcid.org/0000-0001-6065-3087)

Lucia Svabova
*University of Zilina, Slovakia*
[orcid.org/0000-0002-4722-4103](https://orcid.org/0000-0002-4722-4103)

Artificial neural network and decision tree-based modelling of non-prosperity of companies

JEL Classification: C38; G33

Keywords: artificial intelligence modelling; artificial neural network; ensemble; prediction model; financial ratios; non-prosperity; Slovak companies

Abstract

Research background: Financial distress or non-prosperity prediction has been a widely discussed topic for several decades. Early detection of impending financial problems of the company is crucial for effective risk management and important for all entities involved in the company’s business activities. In this way, it is possible to take the actions in the management of the company and eliminate possible undesirable consequences of these problems.

Purpose of the article: This article aims to innovate financial distress prediction through the creation of individual models and ensembles, combining machine learning techniques such as...
decision trees and neural networks. These models are developed using real data. Beyond serving as an autonomous and universal tool especially useful in the Slovak economic conditions, these models can also represent a benchmark for Central European economies confronting similar economic dynamics.

**Methods:** The prediction models are created using a dataset consisting of more than 20 financial ratios of more than 19 thousand real companies. Partial models are created employing machine learning algorithms, namely decision trees and neural networks. Finally, all models are compared based on a wide range of selected performance metrics. During this process, we strictly use a data mining methodology CRISP-DM.

**Findings & value added:** The research contributes to the evolution of financial prediction and reveals the effectiveness of ensemble modelling in predicting financial distress, achieving an overall predictive ability of nearly 90 percent. Beyond its Slovak origins, this study provides a framework for early financial distress prediction. Although the models are created for diverse industries within the Slovak economy, they could also be useful beyond national borders. Moreover, the CRISP-DM methodological framework enables its adaptability for companies in other countries.

**Introduction**

One of the risk management tasks is analysing the financial health and especially predicting possible financial problems that threaten a company in the following period. Therefore, this topic has been widely discussed for several decades. Financial problems can lead to insolvency and even bankruptcy of the company. Therefore, models that predict these situations are called bankruptcy models, sometimes just prediction models (Cheraghali & Molnár, 2023; Lohmann & Möllenhoff, 2023; Valaskova et al., 2023).

The traditional approaches to building prediction models are multivariate discriminant analysis (MDA) and logistic regression (LR), and these approaches are still in use nowadays (Maquieira, 2024; Powell, 2023). Later, other non-mathematical-statistical approaches were employed in this area. In particular, using various types of artificial neural networks (ANN) and decision trees (DT) leads to models with predictive ability at least comparable to traditional models (Gavurova et al., 2022; Kaczmarek et al., 2021).

The article presents the results of research aimed at predicting the financial problems of the companies. The prediction models are built as robust models applicable to all groups of companies operating in the conditions of the specific economy; in this study, the data on Slovak companies were employed for the creation of the models. Since the creation of prediction models is a typical data-mining task (Wang & Yu, 2022), it is possible to use various data-mining methodologies and procedures not only in the
creation of the model itself, but also in the preparation and cleaning of data, as we present in this article. In this study, we chose to use the CRISP-DM data-mining methodology, the main phases of which are business understanding, data understanding and preparation, the creation of the model and its evaluation (Cheng, 2023).

The models in this study are created using a precisely prepared dataset containing mainly financial data of real companies. Various machine learning tools were employed for the modelling, namely some common types of ANNs and DTs. We created individual models and ensembles as combinations of the mentioned models (Pavlicko et al., 2021).

The article is designed as a conceptual one contributing to financial distress prediction using modern machine learning techniques such as decision trees and neural networks and their ensembles. Its findings confirm the hypothesis about the success of using various machine learning tools in predicting financial problems. The adaptability of the models to the conditions of another country is ensured by using the CRISP-DM methodology mentioned above, making them valuable tools for early financial distress detection. First, of course, the created models must be detailed for further validation, and after that, they could be used in practice.

The article is divided into several parts. The first is a Literature review presenting significant studies in the researched field, from pioneering to current. The used methodological procedures, analytical approaches and data are described in the Research methodology section. Then, the created models are presented and evaluated in the Results section. In the Discussion section, these results are further discussed and compared with the results of other authors. Finally, everything is summarised in the Conclusions section, where the limitations of the created models and the direction of future research are also described.

**Literature review**

From the methodological point of view, the prediction of the prosperity of the companies could be considered a classification task. For solving this task, Altman’s model (Altman, 1968), Ohlson’s model (Ohlson, 1980) and Zmijewski’s model (Zmijewski, 1984) are considered pioneering. The Altman model was created using MDA, the Ohlson model using LR and the Zmijewski model using probit analysis. These three mathematical-
statistical methods are nowadays still used in this field. However, in recent years, various machine-learning techniques, mainly ANNs, DTs, the technique of support vector machines (SVMs) and combinations of these approaches could serve as popular tools for classifying companies into prosperous and non-prosperous ones, mainly based on the values of financial ratios explaining the company’s financial situation.

The mentioned methods of ANNs and DTs were used in several studies for solving the issue of predicting the financial situation of the companies. The high predictive ability of ANNs compared to other types of models was identified by Geng et al. (2015). Moreover, they found that majority voting accuracy is better than individual neural network accuracy.

In 2016, Zięba et al. (2016) proposed an innovative approach to financial problem detection by combining decision trees and extreme ranking boosting techniques. They designed the model for Polish companies based on data from the Emerging Markets Information Service (EMIS) from 2007 to 2013. Popescu and Dragotă (2018) modelled financial problems using the CHAID decision tree and several ANNs. The authors identified the CHAID model and specific variables as the most suitable for predicting the financial distress of companies operating in post-communist countries.

Mihalovič (2018) published various types of hybrid models based on data from the financial statements of Slovak companies from 2014 to 2017. It was a model combining a genetic algorithm with a neural network (GA-NN), a back-propagation neural network model (BP-NN) and a model based on MDA. Among them, the GA-NN model achieved the best results, as it correctly predicted the financial status of more than 91% of the companies in the sample.

Durica et al. (2019) proposed a prediction model for Polish companies using different types of decision trees based on relevant data from the Amadeus database. The overall accuracy of this model in predicting the financial situation of Polish companies was more than 83%. Korol (2019) explained that predicting the financial situation is one of the main company’s financial management tasks and employed, among others, the techniques of ANNs and DTs for constructing the prediction models.

Wieprow and Gawlik (2021) analyzed the impending bankruptcy risk of companies operating in the tourism sector and listed on the Warsaw Stock Exchange. In Poland, so-called Polish bankruptcy prediction models were designed, derived from the Altman model, e.g., Appenzeller and Szarzec (2004), Hamrol et al. (2004).
Manogna and Mishra (2021) used sensitivity analysis to select relevant input variables to create a prediction model. Data on more than 1,900 Indian companies were potential predictors of models based on decision tree algorithms such as CHAID, CART, C5.0 and QUEST (Quick Unbiased Efficient Statistical Tree). The results have showed that C5.0 and CHAID decision trees were the best algorithms, providing excellent model prediction characteristics. For Czech industrial enterprises, Horak et al. (2020) created several models based on SVM and ANN techniques. The authors compared these models and identified that ANN models achieved better results than SVM models. Contrary to many authors, Papana and Spyridou (2020) have found that the MDA-type model achieves the best overall performance compared to other models they developed. These were MDA, LR, DT, and ANN models created using 50 financial ratios. Surprisingly, the DT model achieved the worst classification results.

Saeedi (2021) deals with the prediction of financial problems based on the financial statements of 37,325 companies listed on the NYSE, AMEX, and NASDAQ between 2001 and 2017. The author compared four techniques — decision trees, support vector machines, k-nearest neighbors, and rough sets and all show a high predictive power of more than 84%.

The prediction of financial problems was dealt with by Halim et al. (2021), who compared multiple data mining approaches with a deep learning model in financial distress prediction. The results show that deep learning is better for achieving higher predictive power.

Kim and Upneja (2021) applied the decision tree technique in combination with the majority voting set method to create a prediction model. The research was based on US restaurant data from 1980 to 2017, and the models achieved a predictive ability of over 80%.

Companies listed on the Johannesburg Stock Exchange (JSE) were analyzed by Dube et al. (2021), who proposed an ANN-type prediction model. Their model correctly identified the financial status of more than 80% of financial services companies and almost 97% of manufacturing companies. Tong and Tong (2022) proposed a C4.5 decision tree to predict the financial problems of Chinese enterprises.

In addition to newly created models, some authors have published review studies focused on predictions of financial problems. For example, Ravi Kumar and Ravi (2007) collected and compared prediction models published from 1968 to 2005. These were models based on traditional mathematical-statistical techniques MDA and LR, but also models based on
various modern data mining techniques, e.g., ANN, DT, Case Based Reasoning, Genetic Algorithms, Rough Set Technique, and others. Perez (2006) also analyzed 30 research studies on neural networks for estimating company health. Based on a detailed analysis in the Google Scholar and Research Gate databases, a similar study was published by Prusak (2018). However, he focused on models from the Central and Eastern European regions.

As has been shown in numerous studies, the modern techniques of artificial intelligence modelling are suitable for creating prediction models of the financial problems of the companies. Anyway, all these methods have their pros and cons. As the above-listed authors mentioned in their studies, the ANNs usually achieve a pretty high prediction accuracy, but on the other hand, the resulting models are considered black boxes, as they are so complex that they cannot be interpreted. Besides the ANNs, the DT-based models are very clear to understand but sometimes have quite lower classification ability (Gabrikova et al., 2023). However, as we intend to demonstrate in this study, this can be strengthened by employing ensemble modelling using various techniques for enhancing the accuracy or stability of the models.

Research methods

We consider the creation of a prediction model of the financial problems of Slovak companies as a classic data-mining task; therefore, we use the CRISP-DM data-mining methodology consisting of six phases — Business understanding, Data understanding, Data preparation, Modeling, Evaluation and Deployment.

In the first phase of Business understanding, we define the overall problem, define goals, and propose a plan to achieve them. While studying the literature, we used various collection and comparative methods. As a result, we identified a significantly dynamic state of the Slovak economic environment and a significant potential for the use of ex-ante financial instruments created by data mining tools based on the principles of artificial intelligence. Therefore, we deduce the possibility of using these financial analysis tools to create a financial health prediction model in the specific conditions of the Slovak economy, which is the main research objective of
the research. Considering this goal, we have formulated the following research hypothesis:

H1: The application of selected data mining techniques may be able to predict with high quality the one-year financial development of companies operating in the conditions of the Slovak economy.

H2: The combination of selected individual models optimizes the prediction ability of the individual models in a one-year time horizon.

We are aware that the problem being solved is very complex, and therefore, despite the effort to make our model more complex, we must abstract from the possible influence of some factors. As part of the literature study, we have identified 25 financial ratio indicators used in this area. We supplemented this list of potential predictors together with an identifier of the company’s size and an identifier of the company’s economic activity based on the official Statistical Classification of Economic Activities NACE rev. 2. The Amadeus database is a source database for us.

We combined the Data Understanding and Data Preparation phases into one phase, the goal of which was to create a final data set for modelling. Since the main aim is to create a one-year prediction model, we must work with data from two consecutive accounting periods. We chose data calculated from the financial statements of companies for the years 2018 and 2019 because these are the last years unaffected by the COVID-19 pandemic. In the beginning, the database contained more than 660 thousand real Slovak companies. Some of them were excluded, as they did not perform economic activities within the Slovak Republic or their financial data from both monitored periods were unavailable.

As part of data preparation, it was also necessary to treat missing, invalid and outlier values of variables, as some of the modelling methods are not robust to handle them. Due to enough potential predictors (25 financial ratios, size and NACE indicator), we excluded nine of these ratios from the database because of a very high proportion of missing or invalid values. Missing data within the remaining variables were replaced with either zero or the arithmetic mean. The zero value was used in five variables in case it was determined, based on the analyses, that zero is the logically correct value. In other cases, for a specific company and a specific variable, the
missing value was replaced by the arithmetic mean of this variable in the size and NACE category to which the company belongs.

We subsequently identified outlier values using an interquartile range-based method (Larose & Larose, 2017). A value is considered an outlier if it is less than or equal to $Q_1 - 1.5 \cdot IQR$ or greater than or equal to $Q_3 + 1.5 \cdot IQR$, where $IQR$ is the interquartile range defined as the difference between the upper quartile $Q_3$ and the lower quartile $Q_1$. We excluded companies that met these conditions for any of the variables. The analysis of the reasons for their exclusion showed that the calculation of some used financial ratio indicators was based on erroneous data from the items of individual financial statements.

The previously mentioned nine indicators were excluded from the original set of 25 financial ratios (in Table 1) due to more than half of the missing or invalid data. Thus, 16 ratios remained in the set of potential predictors together with the company size identifier and the NACE category identifier.

Since some modelling algorithms are sensitive to multicollinearity among the predictors, as a next step of data preparation, we focused on this problem. For this purpose, we analyzed the correlation matrix consisting of the correlation coefficients of individual pairs of variables. Then, we identified more complex dependencies based on the values of variation inflation factors ($VIF$s), whose values are illustrated in Table 2. $VIF$ values below five indicate variables whose contribution to multicollinearity is small, and, therefore, we can consider them problem-free. On the contrary, $VIF$ values above ten identify variables that cause an undesirably high degree of multicollinearity, which can be quite problematic and almost certainly distort the quality of created models. Therefore, some variables had to be eliminated from the data set. Variables were removed one by one based on expert decisions. For example, from the most problematic pair of Debt Ratio and Current Debt Ratio, we removed the Current Debt Ratio, which is only a partial component of the more complex Debt Ratio. With similar decisions, we gradually removed the following seven variables from the set of potential predictors — Current Debt Ratio, Operating Cash Flow Ratio, Net Income After Taxes, Cash Flow-to-Sales Ratio, Quick Ratio, Cash Flow-to-Debt Ratio, and the ratio Cash Return on Assets.

Thus, the final set of predictors contains nine financial ratios with no problem with a high degree of multicollinearity because all $VIF$ values in Table 3 are less than five (even less than three). Table 3 shows the list of
variables that, together with the company size identifier and the company economic activity section identifier, represent the final set of predictors for predicting the non-prosperity of the company. This set contains ratios representing all four basic areas of analysis of the financial health of companies (Gavurova et al., 2022; Kaczmarek et al., 2021; Karas & Reznakova, 2021; Kliestik et al., 2022; Valaskova et al., 2021). Specifically, they include two activity ratios (Asset Turnover Ratio and Debt-to-Sales Ratio), two liquidity ratios (Current Ratio and Cash-to-Assets Ratio), three profitability ratios (Return on Assets, Return on Equity and Return on Sales) and two leverage ratios (Debt Ratio and Non-Current Ratio).

The target (output) variable that will be modelled is the prosperity of the company, which was considered in 2019. The company has the status of a non-prosperous company if it meets the following individually determined criteria based on the currently valid legislation of the Slovak Republic (amendment to Act No. 513/1991 Coll. of January 2016). This law has introduced the institute of “company in crisis”, whereby a company is considered in crisis if it is in bankruptcy or threatened with bankruptcy. Specifically, these are companies with negative equity, insolvent companies, or companies whose equity-to-liability ratio is less than 0.08.

After cleaning the initial database, we get the dataset of 75,649 companies operating in the conditions of the Slovak economy in 2018 and 2019. Exactly 9,497 (12.6%) are non-prosperous companies, and the resting 66,152 (87.4%), are prosperous companies. Such large disproportionality in these samples is undesirable for the use of some modelling methods. Therefore, we created balanced samples by randomly selecting 9,497 prosperous companies from all prosperous companies. Subsequently, we verified the representativeness of this sample with statistical tests.

Creating models is part of the third modelling phase. First, it is necessary to identify potential tools suitable for modelling the non-prosperity of companies. We did not use classic mathematical-statistical procedures (MDA and LR) for modelling, but we focused on using methods based on machine learning. The literature study shows that artificial neural networks and decision trees are widely used in the researched area. And in recent years, ensemble modelling has been very popular. Therefore, we decided to use these types of models.

We used two types of ANNs, namely multilayer perceptron (MLP) networks and so-called RBF networks whose activation function is one of the radial basis functions (most often Gaussian function). When training the
models, we tried different network topologies, different activation functions and their other settings. The networks were trained as back propagation networks with one or two hidden layers. The logistic function and the hyperbolic tangent were used as activation functions, and the stopping criterion was a maximum of one hundred training cycles. In doing so, we created simple classifiers consisting of one network, but also ensemble classifiers consisting of maximum of ten component networks, using bagging and boosting as a combination technique and majority voting as a rule for making predictions.

From the class of different decision tree (DT) generating algorithms, we decided to use CART, CHAID and C5.0 algorithms. The CART (Classification and Regression Tree) algorithm generates a binary decision tree, which is created as a maximum tree and then pruned to avoid overtraining the model. We used the Gini index as a measure of node impurity. In contrast, the CHAID (Chi-squared Automatic Interaction Detection) algorithm generates a general (non-binary) tree, while the chi-square statistic is used to split individual nodes, and the tree is pruned already during its growth, which is one of the features of this algorithm (Milanovic & Stamenkovic, 2016; Yang et al., 2023). We used the mentioned algorithms to create simple classifiers and ensemble classifiers consisting of several decision trees combined with bagging and boosting techniques.

C5.0 algorithm is a typical representative of algorithms that use entropy to measure node impurity (Biró & Néda, 2020; Zhang & Gionis, 2023; Zhang & Yang, 2020). It creates a binary decision tree with clean terminal nodes (the cases belong to only one category of the output variable, i.e., to the group of prosperous or non-prosperous companies). Even with this algorithm, redundant rules are eliminated directly in the tree generation process. Due to the computational complexity, we only used this algorithm to create an individual classifier, not an ensemble one.

The performance of the created models was then evaluated using several metrics in the evaluation phase of the CRISP-DM methodology. When applying some data mining approaches, the problem of overtraining may occur (Andrés et al., 2021; Borrellas & Unceta, 2021). To avoid this problem and create models with good generalizability, we used a combination of simple and five-fold cross-validation technique. For this purpose, even before the modelling phase, we randomly divided the data set into a training and test sample in a ratio of 80:20. The training set was used in the model’s learning phase, while the five-fold cross-validation technique was
applied. The testing set is used exclusively in the evaluation phase when
the resulting one-year prediction ability of the created models is quantified.

For this quantification and measuring the quality of the created models,
we used a so-called classification table evaluation metrics derived from it.
The classification table (also a confusion matrix) is a $2 \times 2$ table describing
the absolute frequencies of actual and predicted classification of cases into
individual groups (in our case, a group of prosperous and non-prosperous
companies respectively). On the main diagonal, there are the numbers of
correctly classified cases, specifically $TN$ (True Negatives) and $TP$ (True
Positives), and outside, there are the numbers of incorrectly classified cases,
that is, $FP$ (False Positives) and $FN$ (False Negatives). Then, the following
measures are used as the criteria for assessing the quality of the created
models:

− Overall accuracy ($ACC$) expresses the relative proportion of correctly
classified cases out of all cases in the sample,

$$ACC = \frac{TP + TN}{TN + FN + FP + TP}.$$  

− Sensitivity ($TPR$, True Positive Rate) expresses the proportion of posi-
tive cases that are correctly identified as positive, so

$$TPR = \frac{TP}{TP + FN}.$$  

− Specificity ($TNR$, True Negative Rate) measures the proportion of nega-
tive cases that are correctly identified as negative, i.e.,

$$TNR = \frac{TN}{TN + FP}.$$  

− Precision ($PR$) of the model expresses the relative share of correctly clas-
sified positive cases among the positive cases, i.e.,

$$PR = \frac{TP}{TP + FP}.$$  

- $F1$-score is the harmonic mean of precision $PR$ and sensitivity $TPR$ of the model, so

\[
F1 = 2 \cdot \frac{PR \cdot TPR}{PR + TPR}.
\]

- Matthew’s correlation coefficient

\[
MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}.
\]

We will also assess the quality of the models based on the $AUC$ (Area Under the Curve), i.e., the size of the area under the ROC (Receiver Operating Characteristic) curve. The maximum $AUC$ value is one. Therefore, if the $AUC$ value of the area is close to one, then the created model has an excellent predictive ability, and the lower this value is, the lower the quality of the model. If the $AUC$ value is close to 0.5, the predictive ability of the model is not sufficient because it approaches the predictive ability of a random classifier. Often, when assessing the quality of prediction models, a general rule is used, according to which an $AUC$ value equal to or close to 0.5 indicates no or weak predictive ability. An $AUC$ value between 0.7 and 0.8 implies high quality, an $AUC$ value between 0.8 and 0.9 excellent quality, and an $AUC$ value above 0.9 outstanding quality of model predictions (Hosmer et al., 2013; Maternová et al., 2023; White et al., 2023).

The deployment phase is the final phase of the CRISP-DM data-mining methodology, in which a concrete implementation of the created model in a real environment is proposed. We designed the created models as conceptual ones. Therefore, we do not deal with their implementation. However, we identified their weaknesses and strengths, as well as the limits and limitations of their practical use.

**Results**

In this section, we present the results of the modelling and the evaluation phase. In these phases, we used the dataset prepared in the previous phases of Business understanding, Data understanding and Data preparation. The final dataset contains the values of nine financial ratios (see Table 3), an indicator of company size and an indicator of economic activity accord-
ing to the NACE 2 classification from 2018. Half of the dataset consists of data on non-prosperous companies in 2019; the rest are prosperous companies. The created models will serve for the classification of the companies into one of these two categories. Models were created (learned) on the training set (80 per cent of the entire dataset). Finally, we quantified the quality of the models based on their predictive ability on the testing set, which makes up the remaining 20 per cent of the dataset.

In modelling, we used MLP and RBF neural networks. For the MLP network, the topology with one layer of hidden neurons, the hyperbolic tangent as the activation function of this layer and the minimization of the error function using the descending gradient method proved optimal. In addition to a simple classifier, we also created an ensemble combined classifier using bagging and boosting techniques.

The values of the evaluation metrics of these models are summarized in Table 4. As we can see, combining the models results in only a slight improvement in the overall predictive ability (ACC). Slightly larger differences are in the improvement of precision (PR), the Matthews correlation coefficient (MCC) and the AUC value. On the other hand, sensitivity, i.e., the correct classification of non-prosperous companies (TPR), is lower for the combined models compared to the individual model. Thus, the combined models are better in classification of prosperous companies.

Radial Basis Function (RBF) neural networks are three-layer networks specific to using a radial basis function as an activation function in the process layer of the network. As before, the RBF network was trained as an individual model, and the combined ensemble model alternatives were also created. Predictive power was evaluated based on the classification of cases from the testing set. Table 5 presents a very good prediction ability of these models, however, it does not reach the quality of MLP-type networks. Although the combined models created by the boosting and bagging technique improved the classification ability of the individual RBF network, they still did not reach the overall quality of the MLP networks. However, an interesting finding is that both combined models outperform MLP network-type models in sensitivity (TPR). The RBF_boost model correctly classified up to 87.4% of non-prosperous companies and thus outperformed all other neural network models.

Other data mining tools that we used in modelling the (non-)prosperity of Slovak companies are various types of decision trees. These classifiers are often used in this area because the output model can be presented as
a set of the individual rules the model uses for prediction. This is the essence of the popularity of these models in practice.

Using the CART algorithm, we created a binary decision tree with the Gini index to measure node impurity. First, the so-called maximal tree with a maximum of five levels of node splitting and a minimum count of one hundred instances in parent nodes and fifty instances in child nodes was grown. In the next step, the tree is pruned to avoid the problem of overtraining.

Under these conditions, we created a classification tree with five levels of node splitting consisting of 21 nodes, seven of which are terminal. In addition to this individual classifier, we again created combined classifiers that eliminate potential weaknesses of the individual model. The predictive ability of these classifiers is illustrated in Table 6. As we can see, combining the models produced only a slight improvement in the overall ACC predictive ability of the individual CART model. However, the higher sensitivity (TPR), i.e., better non-prosperity identification of combined models, is interesting. This is the difference compared to ANN models. The combined CART_boost model improved the prediction ability in a set of non-prosperous companies by about 2.5% to a total value of 84.1%. Therefore, this model is the most accurate among the created CART models.

We also created non-binary decision trees using the CHAID algorithm, which uses the chi-square statistic when splitting nodes. Due to the expected complexity of these trees, we considered a maximum of three levels of node splitting. Nevertheless, an individual classifier of this type has up to 75 nodes, of which 50 are terminal. Compared to the individual CART model, there was a partial reduction of the difference in the category of prosperous and non-prosperous companies. In this case, we also created combined models using the boosting and bagging techniques, which eliminate the possible overtraining of an individual CHAID model. Finally, we again compared the quality of the created models, as illustrated in Table 7. In this case, the best classifier is the combined CHAID_boost model, which achieved the highest values of almost all evaluation metrics.

Finally, we used the C5.0 algorithm, which uses entropy to measure node impurity and creates a binary decision tree with clean terminal nodes. This algorithm is further characterized by the tree being pruned already during the learning. This algorithm is more demanding in terms of time and calculation. Therefore, we created only an individual classification model. The resulting model is quite complex, as it contains 53 nodes in 10
levels of division. The resulting model achieved a total predictive ability of 87.1%, which is a relatively low value compared to other tree models. On the other hand, the model has the most balanced prediction ability within individual groups of companies, as its specificity ($TNR$) is 88.9%, and its sensitivity ($TPR$) is up to 85.4%, which is the highest value of all created decision trees. This model does not identify threatening financial problems (non-prosperity) of the company in less than 15% of cases, which is approximately one per cent less than in the case of the best CART and CHAID models. The $AUC$ value of 0.922 is comparable to CART-type models but slightly lower than CHAID-type models.

In summary, the analytical results and findings confirm the validity of the statements postulated in both research hypotheses. Thus, the techniques of ANNs and DTs are suitable for modeling the non-prosperity of companies with a high predictive ability, as stated in the hypothesis H1. The predictive ability can be further increased by combining individual models, which is stated in hypothesis H2.

Discussion

Now, we describe and discuss the results of the evaluation phase. Thus, we will focus on the evaluation and comparison of the created models from the point of view of their predictive ability. We assess the quality of the models based on several characteristics ($ACC$, $PR$, $F1$, $MCC$ and $AUC$). An important criterion is also the sensitivity of the model ($TPR$), which assesses the quality of the models from the point of view of the correct classification of non-prosperous companies, i.e., companies threatened with financial difficulties within one year. The values of these characteristics for the created models were already presented and were calculated on a testing set that was not used in the modelling phase. The testing set contains data on 3,759 randomly selected companies, of which 1,868 (49.7%) are prosperous, and 1,891 (50.3%) are non-prosperous.

We created prediction models using ANN and DT techniques. All created models achieved a relatively high predictive ability, based on which we can confirm the validity of the established research hypothesis. Therefore, the used approaches are suitable for modelling and identifying impending financial problems of companies.
In a group of models based on neural networks, MLP networks (individual and ensemble models) achieved better results than RBF networks. They are better from the point of view of the overall predictive ability — overall accuracy (ACC) and F1-score value of more than 88%, AUC value around 0.94 and so on. Overall, the highest classification ability was achieved by the MLP\_boost model, i.e., the ensemble model of several MLP networks. However, the highest sensitivity was achieved by the RBF\_boost model, i.e., the ensemble model of several RBF networks. Its sensitivity was 87.4%, which is at least 1% more than in the case of MLP models.

The situation was similar in a group of models based on decision trees. All models achieved similar results in terms of overall model quality characteristics. However, tree models generated by the CART algorithm achieved a higher predictive ability (ACC around 88%, F1-score around 87% and AUC around 0.92), especially the CART\_boost model, i.e., the ensemble of tree classifiers of this type. Regarding sensitivity, we choose the tree model created by the C5.0 algorithm as optimal. This model outperforms other models in the sensitivity for classification of non-prosperous companies, as it correctly classified 85.4% of them. This means that it can best predict the impending financial problems of Slovak companies.

Finally, we compared the four models discussed in the previous paragraphs, which is illustrated in Table 8 and Figure 1. The highest values of indicators of overall predictive ability were achieved by two comparable models. These are MLP\_boost and CART\_boost models, i.e., ensemble models created using the boosting technique. The MLP\_boost model has the highest AUC value and is, therefore, the best model. The only weakness of the CART\_boost model is its PR, which is an additional characteristic necessary for calculating the F1-score.

Compared to the previous conclusions, it is interesting to focus on the sensitivity of the models. This is because the two best models regarding overall predictive ability are worse regarding sensitivity. It expresses the ability of models to correctly classify truly non-prosperous companies, which is crucial from the point of view of their practical application. In this sense, the RBF\_boost model is the best one, that is, an ensemble of neural networks of the RBF type.

In the conditions of the Slovak Republic, several authors deal with research in the field of predicting the financial problems, such as Csikosova et al. (2020), Gregova et al. (2020), Horak et al. (2020), Jenčová et al. (2020), Svabova et al. (2020), Valaskova et al. (2018). Their results are comparable to
our findings. We consider our models competitive to the existing models created in these studies, as the results in terms of prediction accuracy are at least comparable. Moreover, compared with the mentioned studies, our models are created on a precisely prepared and relatively extensive dataset and designed for use in all economic segments.

Many researchers in Central Europe have proposed prediction models for companies from a specific country, e.g., Karas and Reznakova (2017) for Czech, Durica et al. (2019) for Polish, Papania and Spyridou (2020) for Greek, and Popescu et al. (2017) for Romanian companies. These models are by their authors considered more specific and accurate than universal ones designed for several countries, such as Podhorska et al. (2020) and Korol (2019). We see the potential of our models to be used not only for the companies in Slovakia, but also for those in other countries, as either the methodology proposed in this study should be replied by other authors or the economic conditions of the specific country could be incorporated in the models in the form of several input variables.

As for the predictors used in similar studies, Durica et al. (2019), Jan (2021), Karas and Reznakova (2017), Korol (2019), Papania and Spyridou (2020), Podhorska et al. (2020) and Popescu et al. (2017) showed that ROE and Debt Ratio are the most significant predictors of financial problems, which is also true in our DT models. However, it is hard to justify this fact, for the ANN models are very hardly readable, and the order of significance of the predictors is almost impossible to evaluate.

The created models are competitive against other already existing prediction models based on measures of overall predictive ability such as overall accuracy (ACC), F1-score and area under the curve (AUC), as well as sensitivity (TPR). A year in advance, our best models correctly identified the (non-)prosperity of more than 88% of companies and identified the financial problems of more than 85% of non-prosperous companies. For example, the models of Karas and Reznakova (2017) correctly identified a much smaller percentage of companies (64.6% and 65.6%, respectively). This fact also demonstrates the potential of our models to be used in practice for a very precise prediction of the future financial health of the companies.
Conclusions

The goal of the research was the creation of different variants of models that, based on selected financial ratios, the company size and the NACE category of its economic activity, can predict the financial health one year in advance. The models, therefore, identify a company as prosperous (financially healthy) or non-prosperous, i.e., threatened with financial problems. The basis of the model is a precisely prepared dataset containing data from real Slovak companies. We created individual and ensemble models based on the technique of neural networks and decision trees. Their high predictive ability shows that these approaches are suitable for solving this issue. This confirms the validity of the established hypothesis. Compared to individual models, ensemble models created using the boosting technique achieved better results, which confirmed the validity of the partial hypothesis. Specifically, the MLP_boost model and the CART_boost model correctly classified 88.6% and 88.1%, respectively, of the companies in the testing set. From the point of view of sensitivity, the best results were achieved by the RBF_boost model, which correctly identified the financial problems of 87.4% of non-prosperous companies.

We consider one of the strengths of the created models a using of large and precisely prepared database of companies based on which these models were created and validated. This robust data sample was prepared considering all logical and economic conditions and prerequisites for the effective use of modelling methods. Therefore, we consider the created models reliable and capable of predicting possible financial difficulties of (not only) Slovak companies with high confidence. Another advantage of our models is their potential use for all categories of companies from the point of view of their legal form, economic activity, size, etc. As the main strength of the study, we consider the employed data mining approach, ensuring the possibility of the model update to the specific conditions of another country and enabling the replicability of the presented study.

For the purpose of using the created models in practice, fully completed accounting statements (balance sheet, profit and loss and cash-flow statement) for the given accounting period are the prerequisites. The output is one or more binary predictions of the company’s financial health development within the horizon of one year. The first possibility is the prediction of the financial stability of the company, and the second is the prediction of impending financial problems in the sense of the “company in crisis” insti-
tute. In our research, we approximated this status by the condition that the company either has negative Equity, or its Current Ratio is lower than 1, or the Equity-to-Liabilities ratio is less than 0.08. This approximation could be considered one of the limitations of the research, as in practice, it is common that the company meets the conditions of a state of company in crisis and nevertheless operates for several years. However, this was the only possible solution for reliable identification of financially unhealthy companies.

Another limitation of the research is the use of older data from the period before the COVID-19 pandemic as the current state of the economic environment in the post-COVID period or during the current energy crisis is slightly different and, therefore, should have a different impact on the state of the companies. However, this claim would require a more detailed examination and further analysis.

In addition, future research will focus on the examination of possible applications of other machine learning techniques, e.g., Kohonen map, Random Forest, Bayesian network, SVM technique, etc., for predicting the financial difficulties of the companies. Based on the findings from the conducted research, we will mainly work with ensemble modelling approaches combining different types of weak classifiers. The future goal is to create a comprehensive model verified on several years of real data, which would have the potential of a generally used model for predicting the non-prosperity of companies after incorporating the specific domestic conditions of the economic environment of the companies, in which they operate, to the models, which would be possible thanks to the usage of data mining approach.

References


**Acknowledgements**

This publication was created thanks to support under the Operational Program Integrated Infrastructure for the project: Identification and possibilities of implementation of new technological measures in transport to achieve safe mobility dur-
ing a pandemic caused by COVID-19 (ITMS code: 313011AUX5), co-financed by the European Regional Development Fund.

The journal is co-financed in the years 2022–2024 by the Ministry of Education and Science of the Republic of Poland in the framework of the ministerial programme “Development of Scientific Journals” (RCN) on the basis of contract no. RCN/SN/0129/2021/1 concluded on 29 September 2022 and being in force until 28 September 2024.
## Annex

### Table 1. Initial list of potential predictors

<table>
<thead>
<tr>
<th>Calculation Formula</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales / Total Assets</td>
<td>Asset Turnover Ratio</td>
</tr>
<tr>
<td>Current Assets / Current Liabilities</td>
<td>Current Ratio</td>
</tr>
<tr>
<td>Net Income / Equity</td>
<td>Return on Equity (ROE)</td>
</tr>
<tr>
<td>Net Income after Taxes / Total Assets</td>
<td>Net Income After Taxes on Assets Ratio</td>
</tr>
<tr>
<td>Net Income / Total Assets</td>
<td>Return on Assets (ROA)</td>
</tr>
<tr>
<td>Operating Profit / Total Assets</td>
<td>Operating Return on Assets (OROA)</td>
</tr>
<tr>
<td>Total Liabilities / Total Assets</td>
<td>Debt Ratio</td>
</tr>
<tr>
<td>Current Assets / Total assets</td>
<td>CATA Ratio</td>
</tr>
<tr>
<td>Cash &amp; Cash Equivalents / Total Assets</td>
<td>Cash to Assets Ratio</td>
</tr>
<tr>
<td>Current Liabilities / Total Assets</td>
<td>Current Debt Ratio</td>
</tr>
<tr>
<td>Current Assets / Sales</td>
<td>Cash Flow to Sales Ratio</td>
</tr>
<tr>
<td>Stock / Sales</td>
<td>Sales Ratio</td>
</tr>
<tr>
<td>Net Income / Sales</td>
<td>Return on Sales (ROS)</td>
</tr>
<tr>
<td>Non-current Liabilities / Total Assets</td>
<td>Non-Current Debt Ratio</td>
</tr>
<tr>
<td>Cash &amp; Cash Equivalents / Current Liabilities</td>
<td>Quick Ratio</td>
</tr>
<tr>
<td>Operating Cash Flow / Sales</td>
<td>Operating Cash Flow Ratio</td>
</tr>
<tr>
<td>Total Liabilities / Sales</td>
<td>Liabilities Turnover Time Ratio</td>
</tr>
<tr>
<td>Net Income to Total Liabilities</td>
<td>Solvency ratio</td>
</tr>
<tr>
<td>Gross Profit / Total Assets</td>
<td>GPA Ratio</td>
</tr>
<tr>
<td>EBITDA / Sales</td>
<td>EBITDA to Sales Ratio</td>
</tr>
<tr>
<td>Debt / EBITDA</td>
<td>Debt to EBITDA Ratio</td>
</tr>
<tr>
<td>Cash-flow / Total Assets</td>
<td>Cash-flow to Total Assets Ratio</td>
</tr>
<tr>
<td>Cash-flow / Total Liabilities</td>
<td>Cash-flow to Total Liabilities</td>
</tr>
<tr>
<td>Cash-flow / Sales</td>
<td>Cash-flow to Sales Ratio</td>
</tr>
<tr>
<td>Cash-flow / Current Liabilities</td>
<td>Cash-flow to Current Liabilities Ratio</td>
</tr>
</tbody>
</table>

### Table 2. List of potential predictors

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Symbol</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Turnover Ratio</td>
<td>SAL/TA</td>
<td>1.644</td>
</tr>
<tr>
<td>Current Ratio</td>
<td>CA/CL</td>
<td>4.426</td>
</tr>
<tr>
<td>Return on Equity</td>
<td>ROE</td>
<td>1.049</td>
</tr>
<tr>
<td>Net Income After Taxes</td>
<td>EAT/TA</td>
<td>11.195</td>
</tr>
<tr>
<td>Return on Assets</td>
<td>ROA</td>
<td>11.671</td>
</tr>
<tr>
<td>Debt Ratio</td>
<td>TL/TA</td>
<td>113.940</td>
</tr>
<tr>
<td>Cash to Assets Ratio</td>
<td>CASH/TA</td>
<td>2.096</td>
</tr>
<tr>
<td>Cash Return on Assets</td>
<td>CF/TA</td>
<td>3.900</td>
</tr>
<tr>
<td>Cash Flow To Debt Ratio</td>
<td>CF/TL</td>
<td>37.833</td>
</tr>
<tr>
<td>Current Debt Ratio</td>
<td>CL/TA</td>
<td>117.366</td>
</tr>
<tr>
<td>Cash Flow to Sales Ratio</td>
<td>CF/SAL</td>
<td>3.000</td>
</tr>
<tr>
<td>Return on Sales</td>
<td>ROS</td>
<td>3.199</td>
</tr>
<tr>
<td>Non-Current Debt Ratio</td>
<td>NCL/TA</td>
<td>1.148</td>
</tr>
<tr>
<td>Quick Ratio</td>
<td>CASH/CL</td>
<td>5.400</td>
</tr>
<tr>
<td>Operating Cash Flow Ratio</td>
<td>CF/CL</td>
<td>35.660</td>
</tr>
<tr>
<td>Liability Turnover Time</td>
<td>TL/SAL</td>
<td>2.013</td>
</tr>
</tbody>
</table>
Table 3. Final list of predictors

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Type</th>
<th>Symbol</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Turnover Ratio</td>
<td>Activity</td>
<td>SAL/TA</td>
<td>1.614</td>
</tr>
<tr>
<td>Current Ratio</td>
<td>Liquidity</td>
<td>CA/CL</td>
<td>1.741</td>
</tr>
<tr>
<td>Return on Equity</td>
<td>Profitability</td>
<td>ROE</td>
<td>1.014</td>
</tr>
<tr>
<td>Return on Assets</td>
<td>Profitability</td>
<td>ROA</td>
<td>2.484</td>
</tr>
<tr>
<td>Debt Ratio</td>
<td>Leverage</td>
<td>TL/TA</td>
<td>1.840</td>
</tr>
<tr>
<td>Cash to Assets Ratio</td>
<td>Liquidity</td>
<td>CASH/TA</td>
<td>1.239</td>
</tr>
<tr>
<td>Return on Sales</td>
<td>Profitability</td>
<td>ROS</td>
<td>2.323</td>
</tr>
<tr>
<td>Non-Current Debt Ratio</td>
<td>Leverage</td>
<td>NCL/TA</td>
<td>1.025</td>
</tr>
<tr>
<td>Liability Turnover Time</td>
<td>Activity</td>
<td>TL/SAL</td>
<td>1.776</td>
</tr>
</tbody>
</table>

Table 4. Comparison of the predictive ability of MLP-based models

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>TPR</th>
<th>TNR</th>
<th>PR</th>
<th>F1</th>
<th>MCC</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP_indiv</td>
<td>88.2%</td>
<td>86.4%</td>
<td>90.1%</td>
<td>89.8%</td>
<td>88.1%</td>
<td>0.765</td>
<td>0.937</td>
</tr>
<tr>
<td>MLP_boost</td>
<td>88.6%</td>
<td>85.7%</td>
<td>91.5%</td>
<td>91.1%</td>
<td>88.3%</td>
<td>0.774</td>
<td>0.945</td>
</tr>
<tr>
<td>MLP_bagg</td>
<td>88.6%</td>
<td>86.2%</td>
<td>91.0%</td>
<td>90.7%</td>
<td>88.4%</td>
<td>0.773</td>
<td>0.944</td>
</tr>
</tbody>
</table>

Table 5. Comparison of the predictive ability of RBF-based models

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>TPR</th>
<th>TNR</th>
<th>PR</th>
<th>F1</th>
<th>MCC</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF_indiv</td>
<td>83.5%</td>
<td>84.1%</td>
<td>83.0%</td>
<td>83.3%</td>
<td>83.7%</td>
<td>0.671</td>
<td>0.910</td>
</tr>
<tr>
<td>RBF_boost</td>
<td>85.3%</td>
<td>87.4%</td>
<td>83.2%</td>
<td>84.1%</td>
<td>85.7%</td>
<td>0.707</td>
<td>0.928</td>
</tr>
<tr>
<td>RBF_bagg</td>
<td>84.5%</td>
<td>86.9%</td>
<td>82.0%</td>
<td>83.0%</td>
<td>84.9%</td>
<td>0.690</td>
<td>0.919</td>
</tr>
</tbody>
</table>

Table 6. Comparison of the predictive ability of CART-based models

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>TPR</th>
<th>TNR</th>
<th>PR</th>
<th>F1</th>
<th>MCC</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CART_indiv</td>
<td>87.2%</td>
<td>81.6%</td>
<td>92.8%</td>
<td>92.0%</td>
<td>86.5%</td>
<td>0.749</td>
<td>0.916</td>
</tr>
<tr>
<td>CART_boost</td>
<td>88.1%</td>
<td>84.1%</td>
<td>92.1%</td>
<td>91.5%</td>
<td>87.7%</td>
<td>0.765</td>
<td>0.926</td>
</tr>
<tr>
<td>CART_bagg</td>
<td>88.2%</td>
<td>83.7%</td>
<td>92.7%</td>
<td>92.1%</td>
<td>87.7%</td>
<td>0.767</td>
<td>0.924</td>
</tr>
</tbody>
</table>

Table 7. Comparison of the predictive ability of CHAID-based models

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>TPR</th>
<th>TNR</th>
<th>PR</th>
<th>F1</th>
<th>MCC</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAID_indiv</td>
<td>86.2%</td>
<td>82.9%</td>
<td>89.6%</td>
<td>89.0%</td>
<td>85.8%</td>
<td>0.727</td>
<td>0.938</td>
</tr>
<tr>
<td>CHAID_boost</td>
<td>86.9%</td>
<td>84.3%</td>
<td>89.5%</td>
<td>89.0%</td>
<td>86.6%</td>
<td>0.739</td>
<td>0.946</td>
</tr>
<tr>
<td>CHAID_bagg</td>
<td>86.9%</td>
<td>83.4%</td>
<td>90.4%</td>
<td>89.8%</td>
<td>86.5%</td>
<td>0.739</td>
<td>0.944</td>
</tr>
</tbody>
</table>
Table 8. Comparison of the predictive ability of the most successful models

<table>
<thead>
<tr>
<th>Classifier</th>
<th>ACC</th>
<th>TPR</th>
<th>PR</th>
<th>F1</th>
<th>MCC</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP_boost</td>
<td>88.6%</td>
<td>85.7%</td>
<td>91.1%</td>
<td>88.3%</td>
<td>0.774</td>
<td>0.945</td>
</tr>
<tr>
<td>RBF_boost</td>
<td>85.3%</td>
<td>87.4%</td>
<td>84.1%</td>
<td>85.7%</td>
<td>0.707</td>
<td>0.928</td>
</tr>
<tr>
<td>CART_boost</td>
<td>88.1%</td>
<td>84.1%</td>
<td>91.5%</td>
<td>87.7%</td>
<td>0.765</td>
<td>0.926</td>
</tr>
<tr>
<td>C5.0</td>
<td>87.1%</td>
<td>85.4%</td>
<td>88.6%</td>
<td>87.0%</td>
<td>0.743</td>
<td>0.922</td>
</tr>
</tbody>
</table>

Figure 1. Graphical comparison of the predictive ability of the most successful models