PREDICTING BANKRUPTCY OF MANUFACTURING COMPANIES IN EU

Václav Klepáč, David Hampel

Introduction

The best defense against existential problems of a company appears to be good financial health, based on a satisfactory financial situation. If, however, considerably weakened, the company gets into financial distress, which may turn into a financial crisis and end up in bankruptcy. The primary means, by which we would be informed about the condition of the enterprise, are bookkeeping data of a company. From there, based on financial analysis, we can identify scenarios leading to good management decisions and, consequently, to the financial health of the company.

This article deals with financial risks and prediction of bankruptcy which is heavily connected to detection of financial problems across time. In previous years, large number of authors attempt to improve techniques for classification of active and bankrupted companies and prediction possibility of this two states. Classification of companies is understood as the distribution of a given set of companies at a finite number of sub-groups, in which all of the companies are sufficiently similar in the sense of common properties. Objects that have similar characteristics form a class. Classification is performed using a classifier, which is a function that predicts class for the selected object. To build the model we need the set of correctly described objects – so-called training set.

Generally, the models of bankruptcy prediction could be divided in the three major groups: statistical techniques, artificial intelligence techniques and theoretical models. Statistical techniques are the most preliminary and the most current techniques for providing models of a financial crises predicting. Classical modeling vision has been used in these models. These models protract the signs of commercial inability which are divided into two groups of one-variable and multi-variable statistical models. Multivariate discriminant analysis (MDA), logit or probit models belong there.

Artificial intelligence techniques, similar to human logic intelligence, is a system which learns and improves its own solving problem function according to the past experiences. Techniques such as reverse algorithms: decision trees (DT), support vector machines (SVM), neural networks (NN), and genetic algorithms (GA) are among these techniques. Nowadays, the prediction of bankruptcy is not confined to mere estimation based on simple methods, but deals with so-called ensembles, collections of models. Ensembles outputs are average values of more simple models, which have a higher accuracy than the basic approaches.

Theoretical models, unlike statistical and artificial intelligence technique models, try to determine the reason of commercial inabilities. These models are naturally multi-variable and usually use statistical analysis to support theoretical issues. Original methods are gradually superseded by new ones, as the accuracy of classification and their technical availability of the application via various computing interfaces are ceaselessly growing. In the following parts, results of decision trees method will be discussed.

The main aim of this paper is to test the prediction accuracy of different classification methods with heavier concentration on decision trees methods for different lag of data connected to manufacturing sector. Practically, we want to find out if it is possible to predict bankruptcy 1-3 years ahead with solid accuracy and obtain results comparable to previous studies based on data of non-financial or manufacturing companies.

The authors of previous studies compared the degree of accuracy of classification

DOI: 10.15240/tul/001/2018-1-011

based on a distance to bankruptcy. As can be expected, the results show that the accuracy decreases. Unfortunately, there still lacks a large number of studies useful in practical decision-making dedicated for data of European small and medium-sized manufacturing companies. Focus on decision rules for determining and concretization of the decision-making process, whether the company is threatened by bankruptcy and should resist the crisis, is also missing. Studies report results which are of a technical nature, regarding the setting of classifiers, without the emphasis on practical and easy-to-handle conclusions. For this purpose, it is advisable to use a method such as DT; especially because of its clarity and possible connection with methods for the selection of key variables, which can reduce the dimension of data. In our research, we used the DT to develop EU manufacturing companies prediction model and we extended the boundary of literature reviews into the area of classification using the DT model because of the limited studies of financial bankruptcy prediction using the DT model with comparison to SVM and more advanced DT models like RF and ADA boosting (also based on DT classifiers).

The paper is organized as follows: In Section 1, the key points on relevant bankruptcy prediction is presented. Section 2 presents the whole research methodology, data used in experiment, together with metrics used for evaluation of prediction accuracy. Section 3 gives a brief description of the obtained result and shows vizualization of decision rules for the data used. Finally, Section 4 highlights the conclusions and directions for the future research and discusses the obtained results.

1. Models Used in Bankruptcy Prediction Studies

Majority of business prediction models belong to the quantitative models and its inputs are the results of the technical financial analysis. A certain degree of subjectivity in these models we can see in the determination of weighting coefficients for individual ratios or in choosing their own indicators, which are represented in the model. By the forecasting of the companies future development dealt with many experts, who compiled more or less successful prediction models. However, the prediction of bankruptcy is inherently forward directed operation; models are based on the assumption that the past can describe the present or even the future. It is therefore appropriate to mention the difficulties of accounting approaches in terms of theoretical correctness and robustness:

- Assumption of dichotomous dependent variables. As a result, in the prediction of corporate failure it is assumed that the sets of bankrupted companies and companies in good condition are well defined and clearly separated. In practice, however, this strict division into two groups is not possible.
- The selection of variables arises from models that were popular in the literature. Methodology for selecting datasets of sample companies is usually based on non-randomized data of selected companies and therefore the results can not be sufficiently generalized.
- Instability of input data in time. Their use in the context of bankruptcy prediction requires that the relationship between variables is stable over time. The problem of financial indicators volatility is the greatest for companies that are facing bankruptcy.
- Annual Accounts as a data source. It must be assumed that the annual accounts provide a true and fair view of the financial situation of the company.

One of the first authors who used these quantitative techniques in financial distress or bankruptcy prediction were Beaver (1966) and Altman (1968) who used multivariate discrimination analysis (MDA), in that he computed an individual firm's discriminant score using a set of financial and economic ratios.

Probably due to the huge demand coming from the financial sector in the beginning of 1980s more, advanced estimation methods, such as Ohlson's logit model (1980), probit model used by Zmijewski (1984) or Zavgren (1985), were employed. Compared to the MDA the logit model was easier to understand since the logistic score, taking a value between 0 and 1, was interpretable in a probabilistic way.

In 1990, the NNs technique was used into the field of corporate bankruptcy prediction. But MDA and logit analyses have remained as popular tools for financial distress prediction unless they have some issues associated to normally distributed independent variables, sensitivity to multicollienarity, equal variance-covariance matrices for distresses and non-

distressed companies (Balcaen & Ooghe, 2006; Doumpos & Zopounidis, 1996).

The NNs as data mining technique dominates the literature on business failure in 1990s, and still most frequently used as performance benchmark in recent bankruptcy prediction studies. Among other most prominent algorithms in data mining used for bankruptcy prediction are DT (there exist large portion of DT variations like CART algorithm by Breiman and Cutler (1993), C4.5 (Quinlan, 1993); support vector machine (Vapnik, 1995), k-nearest neighbors (kNN), and naive Bayes (Wu et al., 2007; Witten & Frank, 2005). Specially, the non-parametric prediction method known as decision tree (DT) or recursive partitioning has been used in an attempt to bypass the above mentioned assumptions in MDA and logit (Frydman, Altman, & Kao, 1985). Recent studies also used DT in financial distress prediction (Gepp, Kumar, & Bhattacharya, 2010; Kim, Soo, & Upneja, 2014; Li, Sun, & Wu, 2010; Huarng, Yu, & Chen, 2005).

Li, Sun and Wu (2010) demonstrated the applicability of the DT model in the area of business failure prediction and compared the performance power with four other classification methods including MDA, logit, kNN, and SVM. Many recent studies used ensemble methods to enhance the prediction accuracy of DT models, whereas Lin and McClean (2001) predicted corporate failure by using MDA, logit, DT, NN. They used 106 failed and 690 non-failed companies for their training set and 48 failed and 289 non-failed companies for their validating set. The authors showed that the DT and NN models performed better than MDA, logit.

The Ada Boosting algorithm developed by Freund and Schapire (1996) is one of the most important ensemble methods (algorithms that combines multiple models to obtain better accuracy) because it has solid theoretical evidence, accurate prediction performance, great simplicity, and wide and successful applications. Wu (2004) provided an overview of the application of data mining techniques such as DT, ANN, association rules, and factor analysis in Taiwan publicly traded companies. They indicated that the DT boosting ensemble approach showed higher accuracy and fewer type I and type II error rates than a single DT model.

Most researchers used quantitative techniques to compare the prediction performance with other techniques like MDA, SVM, NN, DT and logit for a specific data set like Min and Lee (2005) who was tested accuracy of classifiers on the set of 1,888 companies, using the classification by means of NN, MDA, logit and SVM with RBF kernel, which had accuracy around 83% in opposite to it alternatives with little below 80%.

Min, Lee and Han (2006) assessed predictive performance of SVM, NN and logistic regression on the case of prediction of bankruptcy. The measurement was carried out on real data of 614 Korean production companies, out of which 307 bankrupted in the time period 1999-2002. The highest classification accuracy 80.3% was achieved when using model SVM with 32 variables and it was the highest accuracy compared with other models. Logistic regression achieved only 68% accuracy – the same as for NN.

Ding, Song and Zen (2008) used data of bankruptcies of 250 publicly traded Chinese companies for bankruptcy prediction by 11 variables with SVM method and other methods (NN, logit and MDA) for bankruptcy prediction. These results proved that model SVM with RBF kernel to be the most advantageous. The accuracy of classification via SVM was 95.2% for training and 83.2% for testing data set. The use of neural networks led to the worst results – 76% on the testing data set.

Authors of these studies mostly used data of medium-sized or large companies with hundreds or thousands of companies in data sets and tens of variables. Review of selected key works in this field gives Aziz and Dar (2006), who have summarized large portion of studies which accommodates use of different approaches.

2. Research Methodology

Paper deals with annual accounting or financial data obtained from Amadeus database. In Tab. 1, the variables observed during the time period from 2009 to 2013 is presented. Data consists of 170 manufacturing companies from EU28 countries which gone bankrupt in 2014 and 830 active manufacturing companies. The chose of the likely variables (denominated in EUR currency) is bound to the possible data acquisition especially for the last observed year (missing values in 2013 for many of the bankrupted companies) from the source and previous empirical studies. But we use variables for capturing rentability, liquidity and

Tab. 1: Financial variables selected for the risk assessment

Variable	Variable
ROA (%) from Net Income	Operation revenue (th EUR)
Interest Cover	Sales (th EUR)
Stock Turnover	EBIT (th EUR)
Collection period	Sales/TA
Credit period	EBIT/TA
Current ratio	Sales/Fixed assets
Liquidity ratio (Cash/TA)	Shareholders/TA
Solvency ratio in % (Net Income + Depreciation)/Liabilities	Cash flow (th EUR)
Total assets (th EUR)	Bankruptcy (categorical value)

Source: own selection

leverage altogether with the selected basic univariate variables. Subsequent data analysis steps were performed in software R 3.1.1.

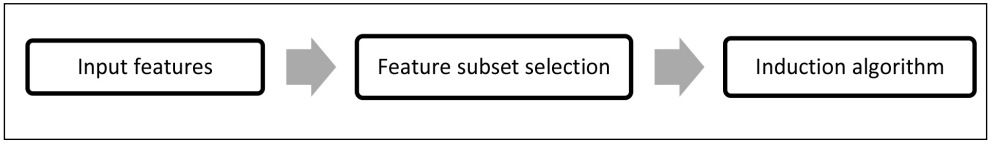
Feature selection (FS) reduces the dimensionality of feature space, improves data quality and performance of classifier by removing redundant, irrelevant, or noisy data, from this point it plays important role bankruptcy prediction because it helps to identify the best variables to capture risky profile of analysed companies. Selection techniques fall into two categories: filter and wrapper method. In this paper we concern mostly filter method.

In the case of the filter methods the subset selection procedure is independent of the learning algorithm and serves like a pre-processing step, see Fig. 1. Filters are based on some importance measure which is unlike the wrappers independent from many classification method, such as correlation between features and decision class or information gains. We use Relief algorithm and Gain ratio which are incorporated into R software package (Romanski & Kotthoff, 2015). Kononenko (1994) and Tsai (2009) describe applicability and technical details.

As long as the decision trees in SW R offers possibility for reducing number of the variables by recursive feature elimination we use also wrapper feature selection algorithm, which in contrast to filter, exists as a wrapper around the induction algorithm. According to Kohavi and John (1997) "The feature subset selection algorithm conducts a search for a good subset using the induction algorithm itself as part of the function evaluating feature subsets."

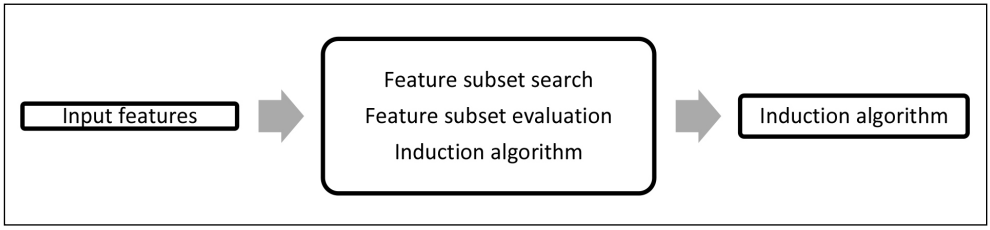
The idea behind the wrapper approach, shown in Fig. 2, is simple: the induction algorithm is considered as a black box. The induction algorithm is run on the dataset, usually partitioned into internal training and holdout sets, with different sets of features removed from the data. The feature subset with the highest evaluation is chosen as the final set on which to run the induction algorithm. The resulting classifier is then evaluated on

Fig. 1: The filter model to feature subset selection (features are filtered independently to induction algorithm)



Source: Kohavi and John (1997)

Fig. 2: **The wrapper approach to feature subset selection in training set (evaluation is performed on independent test set)**



Source: Kohavi and John (1997)

an independent test set that was not used during the search.

2.1 Classification Methods Used in Study

As a relatively new class of machine learning techniques based on statistical learning theory (Vapnik, 1995), SVM for enterprise credit risk assessment has obtained several state-of-art results in classification accuracy. In SVM, original input space is mapped into a high-dimensional dot product space called a feature space, and in the feature space the optimal hyperplane is determined to maximize the generalization ability of the classifier. The optimal hyperplane is found by exploiting the optimization theory, and respecting insights provide by the statistical learning theory. In empirical studies are often used different kernels like linear, RBF, ANOVA RBF or Hyperbolic tangent which are also supported by computational environment.

Decision trees based on CART algorithms (Breiman, 1993), which are to some degree implemented in *rpart* R software package (Therneau & Atkinson, 2015), are produced by algorithms that identify various ways of splitting a data set into branch-like segments. Indeed, the program build classification or of a very general structure using a two stage procedure; the resulting models can be represented as binary trees. According to Breiman (1993) or Jain and Zongker (1997) the tree is built by the following process: first the single variable is found which best splits the data into two groups. The data is separated, and then this process is applied separately to each sub-group, and so on recursively until the subgroups either reach a minimum size (depending on settings) or until no improvement can be made. The resultant model is, with a certainty, too complex, and the question arises as it does with all stepwise procedures of when to stop. The second stage of the procedure consists of using cross-validation to trim back the full tree. The objective of the split is to reduce the impurity of a set by creating subsets that have a greater proportion of members from one of the groups than the original set. The algorithm ends when it achieves the aim of maximizing a homogeneity or purity measure of the response variable in each of the obtained sub-groups.

First, a DT does not require any statistical assumption concerning the data in a training sample. Min, Lee and Han (2006) mention other pros and cons of decision trees and its applicability.

Random forests is an ensemble of decision trees that operate by constructing of DT and outputting final classification class which is the average of the classes of individual tree. The ensemble models grow a lot of different models, and let their outcomes be averaged across the group. Ensemble models are often robust to variance and bias. So briefly we take a large collection of individually imperfect DT, and their one-off mistakes are probably not going to be made by the rest of them. If we average the results of all these models, we can sometimes find a superior model from their combination than any of the individual parts. However the RF is hard to interpret in comparison to DT for the sake of the *black-box* nature of RF (Breiman & Cutler, 1993).

Adaptive boosting (known as AdaBoost) is one of the machine learning algorithms formulated by Freund and Schapire (1996) and further discussed in Breiman (2004). It can be used by combining other learning algorithms

to make an improved learning algorithm from the basic ones (in our study we make use of decision trees as weak classifier to obtain strong classifier). Weighted average method is used for combination. AdaBoost is often referred to as the best out-of-the-box classifier. AdaBoosting algorithm first, sets the initial distribution on the training set and then iterates it until the stopping criterion is reached by using adaptive weights. One of the main ideas of the algorithm is to maintain a distribution or set of weights over the training set. After a classifier is built, the weight of each training sample is changed according to the classification given by a classifier. The next classifier is then built using the reweighted training sample. Finally, once the training process has been completed, the single classifiers are combined into a final highly accurate classifier based on the training set. Because boosting maintains a weight for each instance, the higher the weight, the more the instance influences the learned classifier. The final classifier, therefore, usually demonstrates a high degree of accuracy (Assaad, Bone, & Cardot, 2008). Freund and Schapire (1996) offer other technical details and other informations.

2.2 Research Steps

Similarly to the papers (Ding, Song, & Zen, 2008; Lin & McClean, 2001) this one shows prediction evaluation of several classifiers. We process with the following steps:

- Dividing dataset onto training group, validation group and testing (prediction) group. We want to test accuracy for random samples with partitioning 30/30/40 for training/validation/test groups and normalization by scaling of variables between 0 and 1.
- Feature selection for at most 6 important attributes (to get simplified model) with two filter types: Gain ratio and Relief algorithm. We also use full set with all the features for comparison.
- Training and validation on data and prediction whereas target labels are Bankruptcy or Activity in year 2014. Use of different classifiers: SVM with different kernels, decision trees, adaptive boosting and random forest (both working with five hundreds basic decision trees).
- Accuracy diagnostics with Area under the Receiver operator characterics curve (AUC ROC) values and average accuracy evaluation.
- Evaluation of obtained Empirical Error I and II types values to compare with other empirical studies.
- Analysis of inducted decision trees across the time to get better recognition about the state of variables. For this part we choose the best fitting DT model according to different data partitioning.

2.3 Measuring of Prediction Accuracy

Useful method how to evaluate the results of classification/prediction is to use confusion matrix like in Tab. 2. This matrix could be possibly used for all of the data set parts (training, validation and testing sample) but we evaluate more closely just testing sample. The evaluation criteria of our experiments are adopted from the established standard measures in the fields of credit risk assessment. These measures include average accuracy, Type I error and Type II error. In this study, we prefer to use a combination of these measures, rather than a single measure, to measure the performance of enterprise credit risk models.

According to classes in Tab. 2 we could calculate:

- Empirical average classification accuracy as the average rate of accuracy bound

Tab. 2: Confusion matrix

		Current category	
		T (active)	**F (bankrupted)**
Predicted category	T	True positives (TP)	False Positives (FP)
	F	False Negatives (FN)	True Negatives (TN)

Source: Fawcet (2004)

both classification labels and Type I and II errors are used for more specific accuracies evaluation

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

- Empirical Error I which evaluates the number of true positives which were classified as true negatives

$$Type\ I\ Error = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}.$$

- In contrast, Empirical Error II that shows how many of true negatives were labeled false positives

$$Type\ II\ Error = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}.$$

Based on these calculated data, the analysis of prediction ability of classification methods via the construction of so-called *Receiver Operator Characteristics (ROC)* curves can be carried out. This method enables to visualise and analyse the behaviour of diagnostic systems both in medicine and in economic applications. It can be claimed to be a tool for evaluation and optimization of binary classification system which shows the relation between specificity and sensitivity of the given test or detector for all allowable values of threshold. The value AUC is the most common index describing the ROC curve, which is suitable for comparison of various classification models where the whole ROC curve is reduced into one scalar quantity with the usual value between 0.5 and 1:

- from 0.50 to 0.75 – eligible,
- from 0.75 to 0.92 – good,
- from 0.92 to 0.97 – very good,
- from 0.97 to 1.00 – perfect.

Tab. 3: Median statistic for all variables (from 2011 to 2013 for active, bankrupted and mixed data set)

Variable/Period	Bankrupted			Active			Mixed		
	2011	2012	2013	2011	2012	2013	2011	2012	2013
ROA (%)	-0.93	-3.23	-26.14	1.91	1.47	1.81	1.39	0.94	0.96
Interest Cover	0.65	-0.50	-7.35	2.35	1.93	2.26	1.83	1.45	1.55
Stock turnover (days)	4.23	4.6	5.56	7.7	6.92	7.64	6.38	6.29	7.26
Collection period (days)	116.94	126.96	109.26	83.62	85.31	83.88	90.82	90.87	86.91
Credit period (days)	87.87	101.15	131.91	49.89	49.52	47.56	54.66	56.96	52.43
Current ratio (%)	1.5	1.1	0.65	1.34	1.38	1.42	1.27	1.28	1.27
Liquidity ratio	0.65	0.63	0.44	0.93	0.93	0.98	0.86	0.86	0.86
Solvency ratio (%)	8.65	5.85	-25.72	28.75	31.09	31.53	22.82	23.70	24.15
Total assets (th EUR)	2,288.41	2,427.60	1,737.02	1,156.83	1,136.63	1,158.58	1,295.73	1,303.24	1,232.91
Operation revenue (th EUR)	1,793.59	1,579.96	1,134.41	1,248.92	1,218.85	1,176.78	1,324.10	1,294.47	1,166.77
Sales (th EUR)	1,768.76	1,515.33	1,135.63	1,224.50	1,170.99	1,164.76	1,301.75	1,263.89	1,162.02
EBIT (th EUR)	15.2	-17.18	-379.04	31.8	29.58	32.62	29.14	23.50	21.12
Sales to TA	0.83	0.78	0.73	1.12	1.9	1.7	1.6	1.00	1.1
EBIT to TA	0.01	-0.01	-0.23	0.04	0.03	0.04	0.03	0.03	0.03
Sales to FA	3.61	2.94	2.61	4.15	4.18	4.15	4.7	3.93	3.73
Shareholders to FA	0.09	0.06	-0.26	0.29	0.31	0.32	0.23	0.24	0.24
CF (th EUR)	20.5	3.19	-425.12	46.64	47.12	43.55	41.43	39.62	29.94

Source: own calculations

Tab. 4: Selected features according to its importance and selection method

Classifier/Period and feature selection method	2011		2012		2013	
	Relief	Gain ratio	Relief	Gain ratio	Relief	Gain ratio
SVM, Random forest, ADA boost	Interest Cover	EBIT	Interest Cover	Solvency ratio	Stock turnover	CF
	ROA	Shareholders/TA	Stock turnover	Liquidity ratio	TA	ROA
	TA	CF	Liquidity ratio	Sales	Operation revenue	Interest cover
	Operation revenue	ROA	Operation revenue	EBIT	Sales	Solvency ratio
	Sales	Interest Cover	Sales	Shareholders/TA	Sales/TA	EBIT
	Sales/Fixed assets	Solvency ratio	Sales/FA	CF	Sales/FA	EBIT/TA
Decision trees (inner selection procedure)	Interest cover	Interest Cover	EBIT	EBIT	Operation revenue	CF
	Operation revenue	ROA	Liquidity ratio	Liquidity ratio	Sales	Interest cover
	ROA	Solvency ratio	Sales	Sales	Sales/TA	ROA
	Sales/FA		Solvency ratio	Solvency ratio	Stock turnover	Solvency ratio
	TA		TA		TA	

Source: own calculations

The area under the ROC curve is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative example. It measures the classifiers ability in ranking a set of patterns according to the degree to which they belong to the positive class, but without actually assigning patterns to classes.

3. Results

In overall as we can see in Tab. 3 depicted by median metric, bankrupted companies differ more significantly in its financial indicators even partly two or three years before bankruptcy, when the value of assets and CF is gradually decreasing, followed with the drop of company´s liquidity and the larger drop in solvency and interest cover. It can be concluded that the financial downturn is evident for more than three years to bankruptcy. When looking for ROA, we see that already two or three years before the bankruptcy, its value is negative. The most alarming indicator of the tendency to bankruptcy becomes the subsequent decline in EBIT into negative territory, although the decline in Sales was not as high. The striking difference between defaulted and active companies can be observed for Collection and Credit periods attributes. Active companies receive funds from customers for about a month faster than bankrupted ones. Moreover, active companies also fulfill its obligations to suppliers faster than bankrupted ones.

One year before bankruptcy, companies start facing distinct financial difficulties, however their treatment is highly improbable due to higher indebtedness and very low profitability. In the opposite, active companies have been showing similar financial indicators in the field of manufacturing business for all the years.

In the case of the filter based feature selection, attributes presented in the Tab. 4 we use for the classification. Compared to the original statistical analysis, these variables describe the most important indicators that should be included in the measurement.

As can be seen from the comparison of feature selection methods in Tab. 3, the classification should incorporate features like Sales, Solvency ratio, ROA, Operation revenues and EBIT including metrics. As we can see both filtering methods offers different variables to be used in the classification and Decision trees wrapper algorithm chose less

number than its competitors. For the sake of better comparison we also trained data without prior filtering procedure – so for the SVM, Random forest and ADA boost was chosen full data samples and for the decision trees there were implemented chosen variables by wrapper: EBIT, Interest Cover, Liquidity ratio, ROA and Solvency ratio in 2013. Credit period, Interest cover, Solvency ratio and Total assets in 2012. Credit period, Interest cover, Liquidity ratio, ROA, Solvency ratio and Stock turnover in 2011.

3.1 Evaluation of Prediction Accuracy

After feature selection, models are constructed and its classification accuracy is carried out for various methods. In Tab. 5, the total results of accuracy of classification together with order I and II types can be seen. They are stated via the value AUC and Accuracy for testing sample, which were based on data of companies in each year. Performing a preliminary test using linear and RBF kernel SVM, we obtained results that are not satisfactory. By the changing of the kernels (see below), better quality and relevant results are provided for discussed data.

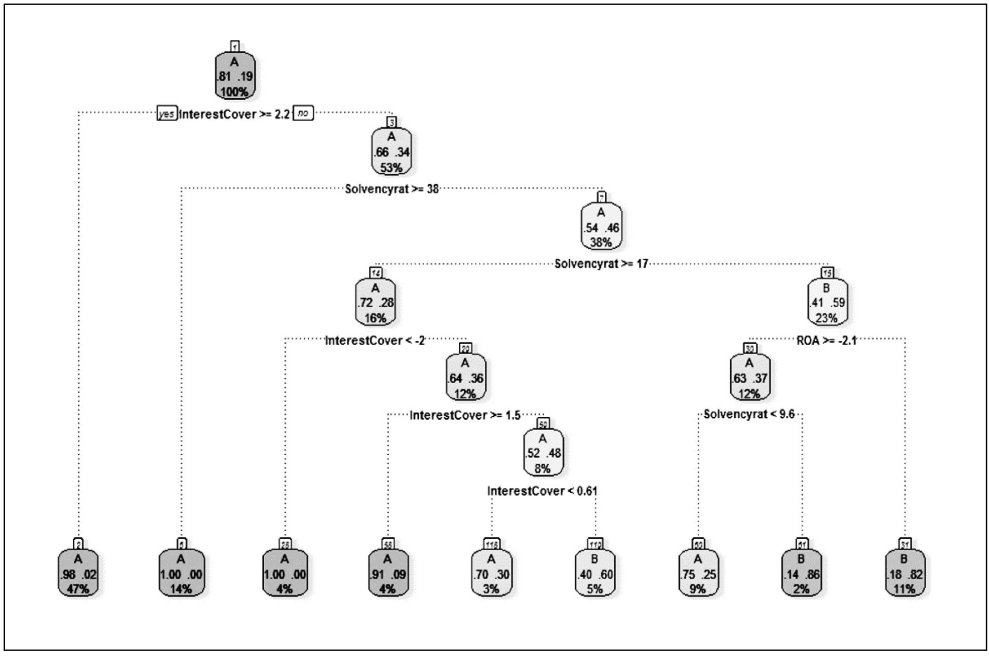
Tab. 5 illustrates that the longer time period till bankruptcy the lower the prediction ability of models is, i.e. financial indicators or real situation for previous years do not reflect the resulting financial difficulties on the satisfactory level, or companies have not faced these problems yet. The highest prediction ability can be observed no sooner than in 2013, i.e. one year before bankruptcy. Evaluation via the value AUC is also problematic. Models for lagged data with longer lags provide only low percentage of appropriately classified

Tab. 5: Evaluation of prediction accuracy according to Error I and Error II (from 2011 to 2013, Rf is Relief selection, GR means Gain ratio selection)

Classifier/Period	Total accuracy (%)			AUC (%)			Type I Error (%)			Type II Error (%)		
	2011	2012	2013	2011	2012	2013	2011	2012	2013	2011	2012	2013
Spline SVM	75.0	77.5	67.5	69.8	75.3	83.0	18.5	15.8	22.1	61.7	60.0	91.7
ANOVA SVM	86.8	88.5	93.8	87.6	78.4	98.9	7.4	8.5	5.9	46.7	34.8	9.3
HT SVM	82.5	76.8	87.8	76.2	74.6	92.0	9.4	13.8	8.8	63.3	76.7	31.7
Decision trees	83.3	87.3	94.0	83.3	85.8	89.2	11.5	8,8	3,5	46,7	35,0	20.0
Random forest	89.0	91.5	96.5	88.1	93.6	97.1	5.3	4.1	0.0	43.3	33.3	16.7
ADA Boost	87.3	91.5	91.5	88.0	95.2	99.3	7.4	4.1	0.0	43.3	31.7	31.7
Spline SVM (Rf)	77.6	84.5	81.8	57.7	77.4	49.1	14.6	1.1	5.5	65.8	96.7	90.0
ANOVA SVM (Rf)	81.2	85.0	85.0	67.7	82.9	44.9	7.3	0.0	0.0	82.9	100.0	100.0
HT SVM (Rf)	77.4	78.5	81.3	62.2	60.6	53.0	13.9	12.9	18.8	71.1	70.0	85.0
Decision trees (Rf)	76.3	81.3	78.5	64.6	80.3	63.4	16.2	13.2	9.4	66.7	50.0	90.0
Random forest (Rf)	76.3	83.3	81.8	74.8	85.2	65.6	8.2	11.8	5.0	68.3	45.0	93.3
ADA Boost with (Rf)	82.8	83.8	59.5	79.1	85.5	67.5	9.7	0.1	5.7	68.3	45.0	88.9
Spline SVM (GR)	78.0	75.0	65.0	63.9	80.0	94.2	14.4	18.8	23.5	65.0	60.0	100.0
ANOVA SVM (GR)	84.0	87.0	96.8	81.2	82.8	99.3	7.9	2.6	1.5	61.7	71.7	13.3
HT SVM (GR)	71.8	79.3	87.3	58.2	71.6	93.3	17.9	10.0	9.1	86.7	81.7	33.3
Decision trees (GR)	80.3	84.8	95.3	80.3	80.4	94.5	14.7	12.6	2.9	48.3	30.0	15.0
Random forest (GR)	85.5	61.8	96.5	84.5	87.0	98.6	8.5	8.9	1.5	48.3	57.4	15.0
ADA Boost (GR)	86.3	63.0	97.0	86.7	87.1	99.3	7.1	6.9	0.0	51.7	57.4	16.7

Source: own calculations

Fig. 3: Decision trees (based on 2011 data and bankruptcy target in 2014)



Source: own

bankruptcies; technically the differences between active companies and those which bankrupt in the future are not as obvious.

The results show that the highest values of the accuracy of classification can be observed mostly in 2013. The difference between values can be observed mostly for AUC values, because the total accuracy in years 2011 and 2012 is biased as the models undervalue during the prediction of frequency of bankrupted companies.

It is appropriate to add that models are able to notice correct labels for active companies, however, the longer time period to bankruptcy is, the worse the results are, or the results are inconsistent for bankrupted companies which have been active till now. The spline SVM kernel is showing worse results especially when compared with ANOVA kernel. Thus we can see that results can be flawed due to the total accuracy, we need to estimate Error I type and Error II type also. Error I evaluates the number of active companies which were classified as bankrupted. In contrast, Error II type shows how many bakrupted companies were wrongly labeled as active ones. For the purposes of risk prediction, this metric is the key one. That is why the setting of the model with the lowest values is searched.

The comparison of variants leads to the clear conclusion, that the best results are provided by Random forest and Decision trees with Gain ratio based feature selection. The lowest values of type II error are achieved for decision tree model, Random forest and ADA boost. We consider more important the correct classification of distressed companies in order to eliminate the risk and provide the ability for manufacturing companies to undertake the necessary steps in time.

3.2 Decision Trees

After performing the previous steps we found empirical accuracy of prediction models. From the comparison follows that the highest accuracy achieved Decision Trees, Random forest and ADA boost – especially compared to models based on SVM. In the diagrams

below we can see the classification trees for the DT method generated using the Gain ratio approach. DT method, besides the high predictive accuracy, provides an easy to handle representation of decision rules, see figures below for all reference years and a training data set.

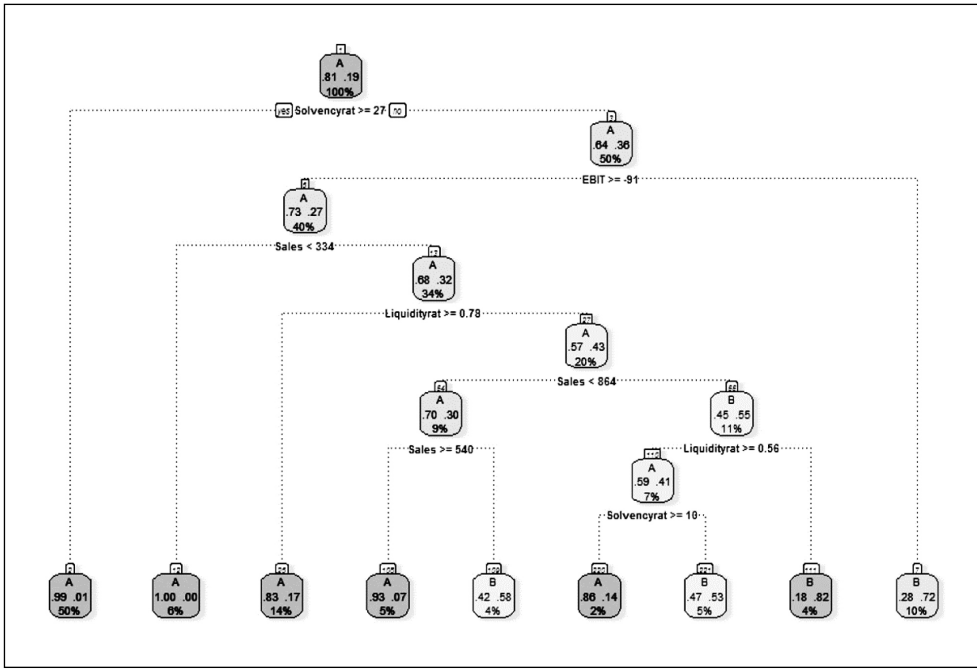
Designation "A" on Fig. 3 indicates the active company (darker color) and the designation "B" (lighter color) indicates the company bankrupted in 2014. The sum of the percent shares at the lowest level indicates the proportion of data classified by the model; overall, it is 100%. The values above percentages indicate share of data assigned to the designation class.

From the perspective of data reported by the companies in 2011, it can be argued that the primary criterions for the result of companies in 2014 are interest cover, solvency ratio, and ROA. When viewed from the top down, we see branching, which in turn separates the data according to criteria evaluation. Due to the results, the best strategy for three years to bankruptcy is to track Solvency ratio values lower than 9.6% (resp. 17% or 38%) and ROA lower than -2.1 Solvency ratio characteristics is clearly more conclusive than Interest cover characteristics, since a larger proportion of correctly classified bankrupted companies had difficulty just in this area.

As in the previous case, one of the main dividing criteria is the Solvency ratio, see Fig. 4. In this case, Solvency ratio is the primary criterion in the hierarchy – for 50% of companies it can determine whether a company goes bankrupt in the future (for the Solvency ratio less than 27%). Further we evaluate EBIT, Sales and Liquidity ratio, although the degree of companies' classification is not as high as for the Solvency ratio in this case. A robust criterion for bankruptcy seems to be a negative EBIT under -91,000 EUR, the Liquidity ratio lower than 0.78, and especially the combination of the Solvency ratio and the Liquidity ratio critical values.

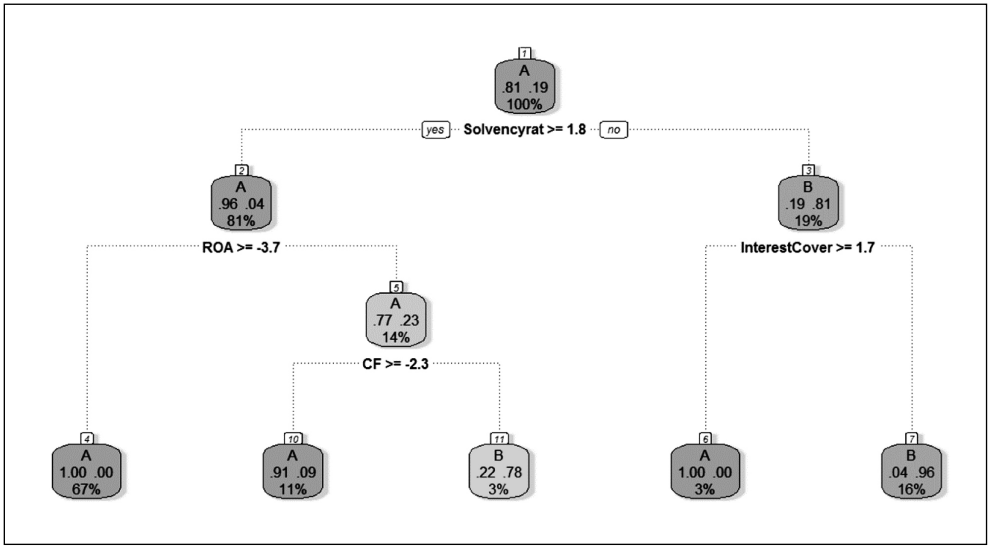
On the basis of modeling, we can clearly see that we obtained results in Fig. 5 are similar to the previous figures. However,

Fig. 4: Decision trees (based on 2012 and bankruptcy target in 2014)



Source: own

Fig. 5: Decision trees (based on 2013 data and bankruptcy target in 2014)



Source: own calculation

the difference is that we have reached the exact quantification of evaluation criteria and relationships between variables, with reduced number of attributes due to the shrinking time to bankruptcy and more obvious bankruptcy profile. This also means that the values determining bankruptcy of a company have numerical levels lower than in previous years in the case of Solvency ratio and ROA. It is apparent that inappropriate theoretical value of one variable does not necessarily lead to bankruptcy, so it is better to use combinations of variables. In this case Solvency ratio with ROA and CF for the most of active companies. For bankrupted companies matters Solvency ratio under 1.8% together with Interest Cover under value 1.7.

Discussion and Conclusions

In the previous sections, we compared and analyzed the results of bankruptcy prediction models with heavier concentration on Decision trees method, with various setting applied to manufacturing companies. Such results suggest important implications. In recent times, bankruptcy of manufacturing companies rapidly increased due to the impact of the recession, which produces economic and social problems accordingly. Therefore, the need for bankruptcy prediction models is very high.

Practically, we wanted to find out if it is possible to predict bankruptcy 1-3 years ahead with the sound accuracy. Financial data acquired from Amadeus database consists of financial variables from 2011 to 2013. Our analysis uses data of 830 active manufacturing small and medium-sized companies and 170 companies which reported bankruptcy in 2014.

The important contribution of article lies in concerning statistically evaluated attributes and different classifiers for prediction of bankruptcy: filter method based on Gain ratio offers better results in terms of Type II errors than Relief algorithm. Suitable attributes as ROA, Interest cover and Operating revenues were mostly used but it also changes across the time, which are probably very obtainable. It is appropriate to add that models are able to notice correct labels for active companies, however, the longer time period to bankruptcy is, the worse the results are, or the results are inconsistent for bankrupted companies which have been active till now. Bankruptcy prediction is characterized by the fact that the results are markedly influenced by the data used and partly by setting of classification models. Therefore,

we can not clearly say whether the results are generalizable to all analogous situations. Also other authors rather refer to empirical studies with a similar focus or the number of companies in the data set.

Proposed total accuracies are on the same levels as in comparable studies for one year before bankruptcy (Li, Sun, & Wu, 2001; Ding, Song, & Zen, 2008; Min et al., 2006). But in contrast to previous studies the SVM based classifiers propose the use of different kernels to linear or RBF (Min & Lee, 2005).

The variant comparison of models with different models led us to the conclusion that SVM classifier based on spline, hyperbolic tangent and ANOVA kernel performs well for capturing total accuracy, especially for the formation 1 year ahead prediction.

However, this does not apply for evaluation of Type I and II errors – models have significant difficulties in capturing real bankruptcy or distressed profile, which holds true especially for active companies. For a longer period before bankruptcy models are not efficient enough to predict the bankruptcy – active companies are assigned with bankruptcy labels. For this sake and to treat Error II types is better to utilize Decision trees, Random forests or Adaptive boosting classifiers in comparison to SVM. Therefore, we used AdaBoosting to overcome the sensitivity problem of the DT model and to make the DT approach more replicable, as was discussed before (Alfaro, Gamez, & Garcia, 2008).

Indeed the obtained results could be further enhanced via testing of accuracy for different classifiers settings and different proportions of training/validation/testing set.

Results may also differ depend on the share of bankrupted companies, but given availability and low completeness of the data led us to analysed composition of companies. It is possible to say, that unbalanced proportion of bankrupted and well performing companies in the data set is adequate image of the real economy conditions. Despite some imbalances in the data set, the models can predict bankruptcy status several years ahead to bankruptcy; the highest accuracy was achieved for the following methods: Adaptive boosting, Decision trees and Random forest.

The paper was supported by the Internal Grant Agency FBE MENDELU under project: Testing of models for multi-variate analysis and prediction of credit risk (No. 19/2015).

References

Alfaro, E., Gamez, M., & Garcia, N. (2008). Linear discriminant analysis versus adaboost for failure forecasting. *Spanish Journal of Finance and Accounting*, *37*(137), 13-32. https://dx.doi.org/10.1016/j.dss.2007.12.002.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589-609. https://dx.doi.org/10.1016/j.sbspro.2013.12.961.

Assaad, M., Bone, R., & Cardot, H. (2008). A new boosting algorithm for improved time-series forecasting with recurrent neural networks. *Information Fusion, 9*(1), 41-55. https://dx.doi.org/10.1016/j.inffus.2006.10.009.

Aziz, M. A., & Dar, H. A. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance: The international journal of business in society*, *6*(1), 18-33. https://dx.doi.org/10.1108/14720700610649436.

Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, *4*, 71-102. https://dx.doi.org/10.2307/2490171.

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. https://doi.org/10.1016/j.bar.2005.09.001.

Breiman, L., & Cutler, A. (1993). A deterministic algorithm for global optimization. *Mathematical Programming*, *58*(1-3), 179-199. https://dx.doi.org/10.1007/BF01581266.

Breiman, L. (1993). *Classification and regression trees*. Boca Raton, Fla.: Chapman & Hall.

Breiman, L. (2004). Population theory for boosting ensembles. *The Annals of Statistics*, *32*(1), 1-11. https://dx.doi.org/10.1214/aos/1079120126.

Ding, Y., Song, X., & Zen, X. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, *34*(4), 3081-3089. https://dx.doi.org/10.1016/j.eswa.2007.06.037.

Doumpos, M., & Zopounidis, C. (1996). A multicriteria discrimination method of the

prediction of financial distress: The case of Greece. *Multionational Finance Journal*, *3*(2), 71-101. https://dx.doi.org/10.1016/j.knosys.2013.12.006.

Fawcet, T. (2004). *ROC Graphs: Notes and Practical Considerations for Researchers*. Technical report HP Laboratories. Kluwer Academic Publishers. Retrieved October 21, 2015, from http://www.hpl.hp.com/techreports/2003/HPL-2003-4.pdf.

Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. In *Machine Learning: Proceedings of the Thirteenth International Conference* (pp. 148-156).

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. https://dx.doi.org/10.1111/j.1540-6261.1985.tb04949.x.

Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting*, *29*(6), 536-555. https://dx.doi.org/10.1002/for.1153.

Jain, A., & Zongker, D. (1997). Feature selection: Evaluation, application and small sample performance. *IEEE Transactions on Pattern Analysis Machine Intelligence, 19*(2), 153-158. https://dx.doi.org/10.1109/34.574797.

Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models. *Economic Modelling*, *36*, 354-362. https://dx.doi.org/10.1016/j.econmod.2013.10.005.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1-2), 273-324. https://dx.doi.org/10.1016/s0004-3702(97)00043-x.

Kononenko, I. (1994). Estimating attributes: Analysis and extensions of RELIEF. In F. Bergadano & De Raedt L. (Eds.), *Machine Learning: ECML-94. ECML 1994. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol. 784* (pp. 171-182). Berlin, Heidelberg: Springer. https://dx.doi.org/10.1007/3-540-57868-4_57.

Li, H., Sun, J., & Wu, J. (2010). Predicting business failure using classification and regression tree: an empirical comparison with popular classical methods and top classification mining methods. *Expert Systems with Applications*, *37*(8), 5895-5904. https://dx.doi.org/10.1016/j.eswa.2010.02.016.

Lin, F., & McClean, S. (2001). A data mining approach to the prediction of corporate failure. *Knowledge-Based Systems*, *14*(3-4), 189-195. https://dx.doi.org/10.1016/S0950-7051(01)00096-X.

Huarng, K., Yu, H., & Chen, C. (2005). The application of decision trees to forecast financial distress companies. In *Proceedings of International Conference on Intelligent Technologies and Applied Statistics, Taipei, Taiwan.*

Min, S.-H., Lee, J., & Han, I. (2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert Systems with Applications, 31*(3), 652-660. https://dx.doi.org/10.1016/j.eswa.2005.09.070.

Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications, 28*(4), 603-614. https://doi.org/10.1016/j.eswa.2004.12.008.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109-131. https://dx.doi.org/10.2307/2490395.

Romanski, P., & Kotthoff, L. (2015). *Selecting attributes. R vignette of the R-package FSelector*. Retrieved October 21, 2015, from https://cran.r-project.org/web/packages/FSelector/FSelector.pdf.

Therneau, E., & Atkinson, J. (2015). *An introduction to recursive partitioning using the RPART routines* (Technical Report 61). Retrieved from http://www.mayo.edu/hsr/techrpt/61.pdf.

Tsai, C. F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, *22*(2), 120-127. https://dx.doi.org/10.1016/j.knosys.2008.08.002.

Quinlan, J. R. (1993). *Programs for Machine Learning*. Morgan Kaufmann Publishers. https://dx.doi.org/10.1007/BF00993309.

Vapnik, V. M. (1995). *The Nature of Statistical Learning Theory*. New York: Springer. https://dx.doi.org/10.1007/978-1-4757-3264-1.

Wu, C. (2004). Using non-financial information to predict bankruptcy: a study of public companies in Taiwan. *International Journal of Management*, *21*(2), 194-201.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Liu, B. et al. (2007). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), https://dx.doi.org/10.1007/s10115-007-0114-2.

Witten, I., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.). Burlington, MA: Morgan

Kaufmann Publishers in an imprint of Elsevier.

Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. *Journal of Business Finance & Accounting*, *12*(1), 19-45. https://dx.doi.org/10.1111/j.1468-5957.1985.tb00077.x.

Zmijewski, M. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, *22*, 59-86. https://dx.doi.org/10.2307/2490859.

Ing. Václav Klepáč, Ph.D.
Mendel University in Brno
Department of Statistics and Operation Analysis
Czech Republic
xklepac@node.mendelu.cz

doc. Mgr. David Hampel, Ph.D.
Mendel University in Brno
Department of Statistics and Operation Analysis
Czech Republic
david.hampel.uso@mendelu.cz

Abstract

PREDICTING BANKRUPTCY OF MANUFACTURING COMPANIES IN EU

Václav Klepáč, David Hampel

Article focuses on the prediction of bankruptcy of the 1,000 medium-sized retail business companies in EU from which 170 companies gone bankrupt in 2014 with respect to lag of the used features. In recent times, bankruptcy of manufacturing companies rapidly increased due to the impact of the recession, which produces economic and social problems accordingly. Therefore, the need for bankruptcy prediction models is very high. From various types of classification models we chose Support vector machines method with spline, hyperbolic tangent and RBF ANOVA kernels, Decision trees, Random forests and Adaptive boosting to acquire best results. Pre-processing is enhanced with filter based feature selection like Gain ratio and Relief algorithm to acquire attributes with the best information value. As we can see both filtering methods offers different variables to be used in the classification and Decision trees wrapper algorithm chose less number than its competitors. Suitable attributes as ROA, Interest cover, Solvency ratio based on assets and Operating revenues were mostly used but it also changes across the time, which are probably very obtainable. It is apparent that inappropriate theoretical value of one variable does not necessarily lead to bankruptcy, so it is better to use combinations of these variables. From the results it is obvious that with the rising distance to the bankruptcy there drops precision of bankruptcy prediction. The last year (2013) with avaible financial data offers best total prediction accuracy, thus we also infer both the Error I and II types for better recognizance of misclassification rates. The Random forest and Decision trees offer better accuracy for bankruptcy prediction than SVM method, both method offers prediction accuracy which is comparable to previous empirical studies.

Key Words: *Bankruptcy prediction, classification, decision tree, feature selection.*

JEL Classification: *C52, C44.*

DOI: *10.15240/tul/001/2018-1-011*