Predicting Financial Distress of Banks Using Random Subspace Ensembles of Support Vector Machines

Petr Hájek[1], Vladimír Olej[1], and Renata Myšková[2]

[1] Institute of System Engineering and Informatics,
Faculty of Economics and Administration, University of Pardubice,
Pardubice, Czech Republic
{Petr.Hajek,Vladimir.Olej}@upce.cz
[2] Institute of Business Economics and Management,
Faculty of Economics and Administration, University of Pardubice,
Pardubice, Czech Republic
Renata.Myskova@upce.cz

Abstract. Models for financial distress predictions of banks are increasingly important tools used as early warning signals for the whole banking systems. In this study, a model based on random subspace method is proposed to predict investment/non-investment rating grades of U.S. banks. We show that support vector machines can be effectively used as base learners in the meta-learning model. We argue that both financial and non-financial (sentiment) information are important categories of determinants in financial distress prediction. We show that this is true for both banks and other companies.

Keywords: Banks, Financial distress, Rating grade, Random subspace, Meta-learning, Support vector machines, Sentiment analysis.

1 Introduction

Avoiding over-fitting is an important issue to address when training base learners. The random subspace (RSS) method [1] represents a parallel learning algorithm generating each base learner independently. In particular, it is favourable for parallel computing and fast learning. Thus, this computing procedure alleviates the risk of local optimum trapping. This makes RSS one of the most frequently used meta-learning method.

Meta-learning algorithms (ensembles of base learners) are increasingly important in predicting financial distress of firms [2-4]. In addition to above mentioned advantages, their use in this domain is also preferable because base learners imitate decisions of individual financial experts and the final decision corresponds to a collective decision-making process.

Various meta-learning strategies have been utilized in financial distress' prediction, including boosting, bagging and stacking [5,6]. However, considerably less attention has been paid to banks and their financial performance. More specifically, previous studies have been limited to bankruptcy prediction [7]. Moreover, previous research

R. Silhavy et al. (eds.), *Artificial Intelligence Perspectives and Applications*,
Advances in Intelligent Systems and Computing 347, DOI: 10.1007/978-3-319-18476-0_14

in predicting financial distress of banks has utilized financial performance indicators only. This study seeks to remedy these problems by employing RSS method (using support vector machines (SVMs) as base learners) to predict financial distress of selected US banks. SVMs are considered state-of-the-art method for financial distress prediction mainly owing to their generalization performance [8]. We treat the financial distress prediction as a two-class problem, where investment/non-investment rating grades provided by a rating agency are used as the indicators of distress. Additionally, we argue that both financial and non-financial (qualitative) information are important determinants of the distress. We show that the qualitative information can be effectively drawn from the textual parts of banks' annual reports using sentiment analysis.

This paper has been organized in the following way. Section 2 begins by laying out the theoretical dimensions of the research, and provides justification of using information from both financial statements and textual reports of banks as the determinants of financial distress. Section 3 describes data and their pre-processing. Section 4 presents the design of prediction models and the experimental results obtained by RSS. SVMs are used as the base learners of the RSS. The results are compared with multilayer neural networks (MLPs) which have been considered as benchmark methods in previous studies [9]. We also provide a comparison with other firms in this section. The final section concludes the paper and discusses the results.

2 Theoretical Background

The late-2000s financial crisis is considered to be the worst financial crisis since the Great Depression. It was triggered by a liquidity shortfall in the U.S. banking system in 2008 [10]. For this reason, experts examine bank concentration and its impact on the effectivity and stability of banking market [11].

Another issue addressed is the size of bank capital and the relationship between bank capital and liquidity creation (see [11]). Thus far, discussion has been concerned with the size of bank capital, risk, liquidity, as well as liquid assets and liabilities in relation to liquidity [12]. However, the issue of assessing banks' financial statements regarding the sentiment has not been addressed in previous literature. The difficulty is that the behavior of stakeholders is influenced by more than financial attributes. It has been reported that psychological factors often have substantial impact on decision-making and in many cases this may result in what is considered deviation from the normative models of action [13]. An important role of voluntarily disclosed qualitative information has been reported only recently for business companies [14-16].

Financial ratios, on the other hand, have been used in many studies on financial distress prediction of banks. These studies can be classified into two categories, bankruptcy [17-18] and credit rating prediction [19-22]. A wide range of soft computing methods have been employed in related studies, including support vector machines, neural networks, fuzzy and rough sets, evolutionary algorithms and meta-learning algorithms, see [23,24] for reviews. The Camel model is another approach used to assess the performance of banks [25]. This model assesses, in addition to financial indicators, key banking qualitative indicators in the following areas: capital adequacy, asset quality, management, earnings, and liquidity. Based on this assessment, banks can be classified into five categories, from 1 (best) to 5 (worst). The assessment is conducted by an expert from a bank regulation or supervision organization, namely the Federal Reserve, the Office of the Comptroller of the Currency, the National Credit Union Administration, and the Federal Deposit Insurance Corporation. The resulting rating is intended to serve the top managements of banks only and, thus, it is not disclosed publicly. However, banks with a deteriorating assessment are subject to increased supervision. The above consideration suggests that banks are motivated to effectively deal with the narrative parts of their annual reports. They use the sentiment in the communication with stakeholders as a tool that illustrates the overall business position of the bank.

3 Data and Their Pre-processing

Adopting the methodology used for companies in [15], the input attributes in this study cover two main categories, financial indicators and sentiment indicators. The chosen financial indicators monitor profitability (earnings per share), financial market situation (beta coefficient, high to low stock price, std. dev. of stock price, correlation with market stock index), business situation (effective tax rate), asset structure (fixed assets / total assets), leverage ratios (market debt / total capital, book debt / total capital), dividend policy (dividend yield) and ownership structure (share of insiders' and institutional holdings). Sentiment indicators, on the other hand, refer to the qualitative assessment of business position by the management of the bank. Usually, the subjects of the assessment are business performance and risks, strategic, financial and investment policy, etc. [26].

The financial indicators were collected from financial statements (Value Line database), and the sentiment indicators were drawn from annual reports (10-Ks documents) freely available at the U.S. Securities and Exchange Commission EDGAR System. Data were collected for 126 U.S. banks selected from the Standard & Poor's database in the year 2010. In the year 2011, 89 of them were classified as investment grade (IG) and 37 as non-investment grade (NG) by the Standard & Poor's rating agency. The investment/non-investment grade position is considered important to investors due to the restrictions imposed on investment instruments. Table 1 shows the list of input and output attributes, and Fig. 1 shows boxplots of input attributes separately for IG and NG class.

The sentiment indicators were processed in the following way. First, linguistic preprocessing (tokenization and lemmatization) was carried out to obtain a set of candidate terms. Second, this set was compared with the sentiment categories from the financial dictionary provided by [16]. Then, the *tf.idf* term weighting scheme was applied to obtain the importance of terms and an average weight was calculated for each sentiment category (negative, positive, uncertainty, litigious, modal strong and modal weak) [27]. For all data, we replaced missing values by median values, and all data were standardized using the *Z*-score to prevent problems with different scales.

Table 1. Input and output attributes

	attribute		attribute
x_1	growth in earnings per share (EPS)	x_{11}	dividend yield
x_2	expected growth in EPS	x_{12}	share of insiders' holdings
x_3	beta coefficient	x_{13}	share of institutional holdings
x_4	high to low stock price	x_{14}	frequency of negative terms
x_5	std. dev. of stock price	x_{15}	frequency of positive terms
x_6	correlation with market stock index	x_{16}	frequency of uncertainty terms
x_7	effective tax rate	x_{17}	frequency of litigious terms
x_8	fixed assets / total assets	x_{18}	frequency of strong modal terms
x_9	market debt / total capital	x_{19}	frequency of weak modal terms
x_{10}	book debt / total capital	class	{IG, NG}

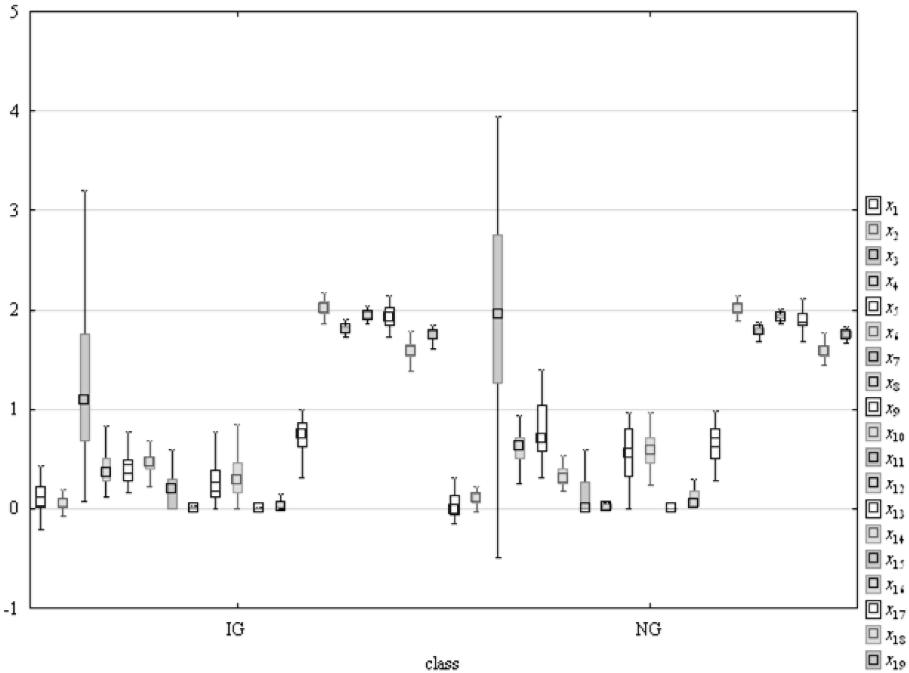

Fig. 1. Input attributes for IG and NG class. To compare the values between classes, we performed Student's t-test showing that the IG class is significantly higher for: x_6, x_{13}, x_{15}, x_{16}, and significantly lower for: x_3, x_4, x_5, x_8, x_9, x_{10} (both at p<0.1)

4 Modelling and Analysis of the Results

To predict rating grades of banks, we employed several meta-learning algorithms using SVMs as base learners. In addition, we compared the results with another base learner – MLP. The results suggest that RSS method performed significantly better when compared with the remaining meta-learning algorithms, namely multiboosting [28], adaboosting [29], bagging and dagging [30] and rotation forest [31]. Due to limited space, we report only selected results for RSS method. This method utilizes many base learners which are systematically constructed by pseudorandomly selecting subsets of components of the feature vector. Thus, it improves accuracy on testing data as it grows in complexity.

To avoid overfitting, we used 10-fold cross-validation in our experiments. The classification performance was measured by the averages of standard statistics applied in classification tasks: accuracy [%], true positives (TP rate), false positives (FP rate), and the area under the receiver operating characteristic (ROC) curve. A ROC is a graphical plot which illustrates the performance of a binary classifier system, which represents a standard technique for summarization classifier performance over a range of tradeoffs between TP and FP error rates.

Subspace size is the critical parameter in learning RSS. Therefore, we tested several settings of the subspace size to obtain the best classification performance. Fig. 2 and Fig. 3 show that the optimum subspace size was 0.4 to 0.5 for banks. The number of iterations of the RSS was set to 10. To compare the results with other dataset, we used the dataset of companies (without banks) used by [15]. This dataset describe 520 U.S. companies by 19 attributes. Similarly as in the case of the banks' dataset, 13 of the attributes are financial indicators (these are rather specific for other companies when compared with banks [32]) and 6 are sentiment indicators, see [15] for details. Although we admit that the comparison should be made with caution, it is obvious that the subspace size is more important for banks, generally requiring larger size of subspace to achieve a high accuracy. This holds true for SVMs in particular.

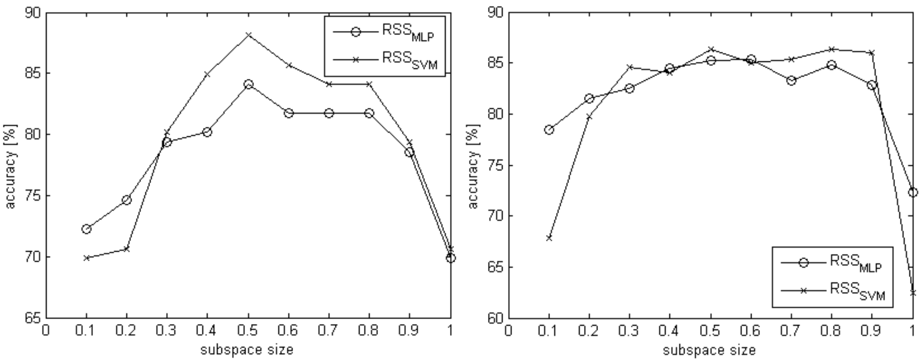

Fig. 2. Relationship between accuracy and subspace size, a) banks, b) companies. The subspace size is expressed as a ratio to the total number of attributes.

The MLP was trained using the backpropagation algorithm with momentum. The following training parameters were examined to achieve the best classification performance: the number of neurons in the hidden layer = $\{2^0, 2^1, \ldots, 2^5\}$, learning rate = {0.05, 0.1, 0.2, 0.3}, momentum = {0.1, 0.2, 0.3}, and the number of epochs = {50, 100, 300, 500, 1000}. We used grid search algorithm to find the optimum settings of these parameters.

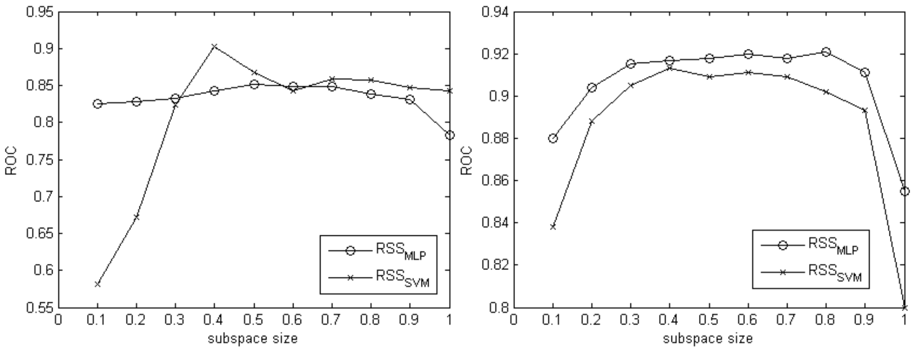

Fig. 3. Relationship between the area under ROC and subspace size, a) banks, b) companies. The subspace size is expressed as a ratio to the total number of attributes.

The SVMs was trained by the sequential minimal optimization (SMO) algorithm [33]. The best classification performance of the SVM was tested for the following user-defined parameters: kernel functions = polynomial, the level of polynomial function = 2, complexity parameter $C = \{2^0, 2^1, \dots, 2^{10}\}$, round-off error ε = 1.0E-12, and tolerance parameter = 0.001. Again, grid search algorithm was employed to find the optimum settings.

Table 2 shows the detailed classification performance on the dataset of banks. First, SVM performed significantly better when compared with MLP (using paired Student's t-test at p<0.1). Second, RSS performed better than single classifiers (significantly better in the case of MLP). Third, classification accuracy was significantly higher when incorporating sentiment indicators (x_{14} to x_{19}). Fourth, all classifiers performed better on IG class, mainly due to the imbalance of the classes.

Table 2. Best results of the analyzed methods for banks

all input attributes x_1 to x_{19}								
	MLP		SVM		RSS_{MLP}		RSS_{SVM}	
Acc [%]	81.75		**87.30**		84.13		**88.10**	
Class	IG	NG	IG	NG	IG	NG	IG	NG
TP rate	0.899	0.622	0.921	0.757	0.910	0.676	0.933	0.757
FP rate	0.378	0.101	0.243	0.079	0.324	0.090	0.243	0.067
ROC	0.824	0.824	0.839	0.839	0.852	0.852	0.868	0.868
financial indicators only (without sentiment indicators) x_1 to x_{13}								
Acc [%]	80.16		**84.13**		83.34		**86.51**	
Class	IG	NG	IG	NG	IG	NG	IG	NG
TP rate	0.876	0.622	0.899	0.703	0.921	0.622	0.955	0.649
FP rate	0.378	0.124	0.297	0.101	0.378	0.079	0.351	0.045
ROC	0.782	0.782	0.801	0.801	0.843	0.843	0.854	0.854

In Table 3, we provide the results for companies for comparison purposes. Similarly to banks, sentiment indicators lead to significantly higher accuracy. This corroborates the findings obtained by [15]. In contrast to banks, the accuracy was higher for the NG class, which can be explained by the different frequencies of classes (also imbalanced but the NG class prevailed for companies). Altogether, the total accuracy was higher for banks, but this is mainly due to the higher accuracy on the IG class.

Table 3. Best results of the analyzed methods for companies

all input attributes x_1 to x_{19}								
	MLP		SVM		RSS_{MLP}		RSS_{SVM}	
Acc [%]	84.04		**86.15**		85.38		**86.34**	
Class	IG	NG	IG	NG	IG	NG	IG	NG
TP rate	0.728	0.908	0.785	0.908	0.759	0.911	0.790	0.908
FP rate	0.092	0.272	0.092	0.215	0.089	0.241	0.092	0.210
ROC	0.898	0.898	0.846	0.846	0.920	0.920	0.909	0.909
financial indicators only (without sentiment indicators) x_1 to x_{13}								
Acc [%]	81.92		**84.62**		83.65		**84.81**	
Class	IG	NG	IG	NG	IG	NG	IG	NG
TP rate	0.708	0.886	0.779	0.866	0.744	0.892	0.744	0.911
FP rate	0.114	0.292	0.114	0.221	0.108	0.256	0.089	0.256
ROC	0.883	0.883	0.833	0.833	0.915	0.915	0.911	0.911

5 Conclusion and Discussion

The informative value of financial statements is increasingly important for stakeholders. Although the structure of these statements is given by the accounting legislation of the corresponding country, the quality of the disclosed information is subject to expert assessment. Management communicates the financial results with stakeholders in the textual parts of annual reports. This is true for both banks and companies. In this paper, we compare these two categories of economic subjects, which differ in both business activity and legislative regulation. On the other hand, the categories have one important characteristics in common – they are established in order to make profit. Thus, they are obliged to disclose true and undistorted information about their business activity.

The extent and quality of information in the report on economic activities (annual report) depend mainly on regulation and have both quantitative and qualitative character. Therefore, we aimed at using the information to accurately predict the rating grades of banks. The results showed that sentiment information hidden in annual reports should be considered an important determinant in financial distress prediction models. This finding has serious implications for stakeholders, regulators and other authorities. More specifically, the accuracy of prediction was increased by about 1.5 % using sentiment information for both banks and companies. We also reported that RSS meta-learning algorithm significantly increases the performance of base learners

SVM and MLP in case of banks in particular. Combining the contribution of sentiment information and meta-learning approach the accuracy was increased by about 4 %. This may lead to substantial savings owing to the loss associated with potential distress. Additionally, the more accurate prediction model makes it possible to better anticipate the effects of potential financial crises.

In future research, the differences in the content of annual report between industries need to be examined, because it is the connection between sentiment and corresponding subject that provides more detailed qualitative information. Thus, a more complete picture would be extracted from annual reports. We further intend to compare the dictionary approach used in this study with machine learning algorithms such as Naïve Bayes [34].

Acknowledgments. This work was supported by the scientific research project of the Czech Sciences Foundation Grant No: 13-10331S.

References

1. Ho, T.K.: The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(8), 832–844 (1998)
2. Verikas, A., Kalsyte, Z., Bacauskiene, M., Gelzinis, A.: Hybrid and Ensemble-based Soft Computing Techniques in Bankruptcy Prediction: A Survey. Soft Computing 14(9), 995–1010 (2010)
3. Hájek, P., Olej, V.: Predicting Firms' Credit Ratings Using Ensembles of Artificial Immune Systems and Machine Learning – An Over-Sampling Approach. In: Iliadis, L., Maglogiannis, I., Papadopoulos, H. (eds.) AIAI 2014. IFIP AICT, vol. 436, pp. 29–38. Springer, Heidelberg (2014)
4. Heo, J., Yang, J.Y.: AdaBoost based Bankruptcy Forecasting of Korean Construction Companies. Applied Soft Computing 24, 494–499 (2014)
5. Alfaro, E., García, N., Gámez, M., Elizondo, D.: Bankruptcy Forecasting: An Empirical Comparison of AdaBoost and Neural Networks. Decision Support Systems 45(1), 110–122 (2008)
6. Kim, M.J., Kang, D.K.: Ensemble with Neural Networks for Bankruptcy Prediction. Expert Systems with Applications 37(4), 3373–3379 (2010)
7. Shin, S.W., Lee, K.C., Kilic, S.B.: Ensemble Prediction of Commercial Bank Failure Through Diversification of Input Features. In: Sattar, A., Kang, B.-H. (eds.) AI 2006. LNCS (LNAI), vol. 4304, pp. 887–896. Springer, Heidelberg (2006)
8. Hajek, P., Olej, V.: Credit Rating Modelling by Kernel-Based Approaches with Supervised and Semi-Supervised Learning. Neural Computing and Applications 20(6), 761–773 (2011)
9. Hajek, P.: Municipal Credit Rating Modelling by Neural Networks. Decision Support Systems 51(1), 108–118 (2011)
10. Shie, F.S., Chen, M.Y., Liu, Y.S.: Prediction of Corporate Financial Distress: An Application of the America Banking Industry. Neural Computing and Applications 21(7), 1687–1696 (2012)
11. Berger, A.N., Bouwman, C.H.S.: Bank Liquidity Creation. Review of Financial Studies 22, 3779–3837 (2009)

12. de Haan, L., van den End, J.W.: Bank Liquidity, the Maturity Ladder, and Regulation. Journal of Banking & Finance 37, 3930–3950 (2013)
13. Kliger, D., Levy, O.: Mood-Induced Variation in Risk Preferences. Journal of Economic Behavior and Organization 52(4), 573–584 (2003)
14. Lu, H.M., Tsai, F.T., Chen, H., Hung, M.W., Li, S.H.: Credit Rating Change Modeling using News and Financial Ratios. ACM Transactions on Management Information Systems 3(3), 14 (2012)
15. Hájek, P., Olej, V.: Evaluating Sentiment in Annual Reports for Financial Distress Prediction Using Neural Networks and Support Vector Machines. In: Iliadis, L., Papadopoulos, H., Jayne, C. (eds.) EANN 2013, Part II. CCIS, vol. 384, pp. 1–10. Springer, Heidelberg (2013)
16. Loughran, T., McDonald, B.: When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. The Journal of Finance 66(1), 35–65 (2011)
17. Chauhan, N., Ravi, V., Karthik Chandra, D.: Differential Evolution Trained Wavelet Neural Networks: Application to Bankruptcy Prediction in Banks. Expert Systems with Applications 36(4), 7659–7665 (2009)
18. Ravi, V., Pramodh, C.: Threshold Accepting Trained Principal Component Neural Network and Feature Subset Selection: Application to Bankruptcy Prediction in Banks. Applied Soft Computing 8(4), 1539–1548 (2008)
19. Chen, Y.S.: Classifying Credit Ratings for Asian Banks using Integrating Feature Selection and the CPDA-based Rough Sets Approach. Knowledge-Based Systems 26, 259–270 (2012)
20. Orsenigo, C., Vercellis, C.: Linear versus Nonlinear Dimensionality Reduction for Banks' Credit Rating Prediction. Knowledge-Based Systems 47, 14–22 (2013)
21. Gogas, P., Papadimitriou, T., Agrapetidou, A.: Forecasting Bank Credit Ratings. The Journal of Risk Finance 15(2), 195–209 (2014)
22. Bellotti, T., Matousek, R., Stewart, C.: A Note Comparing Support Vector Machines and Ordered Choice Models' Predictions of International Banks' Ratings. Decision Support Systems 51(3), 682–687 (2011)
23. Demyanyk, Y., Hasan, I.: Financial Crises and Bank Failures: A Review of Prediction Methods. Omega 38(5), 315–324 (2010)
24. Ravi Kumar, P., Ravi, V.: Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques – A Review. European Journal of Operational Research 180(1), 1–28 (2007)
25. DeYoung, R., Flannery, M.J., Lang, W.W., Sorescu, S.M.: The Information Content of Bank Exam Ratings and Subordinated Debt Prices. Journal of Money, Credit and Banking 33(4), 900–925 (2001)
26. Hajek, P., Olej, V.: Comparing Corporate Financial Performance and Qualitative Information from Annual Reports using Self-organizing Maps. In: 10th International Conference on Natural Computation (ICNC 2014), Xiamen, China, pp. 93–98 (2014)
27. Hajek, P., Olej, V., Myskova, R.: Forecasting Corporate Financial Performance using Sentiment in Annual Reports for Stakeholders' Decision-Making. Technological and Economic Development of Economy (2014) (in press), doi:10.3846/20294913.2014.979456
28. Webb, G.I.: MultiBoosting: A Technique for Combining Boosting and Wagging. Machine Learning 40(2) (2000)
29. Freund, Y., Schapire, R.E.: Experiments with a New Boosting Algorithm. In: 13th Int. Conf. on Machine Learning, San Francisco, CA, pp. 148–156 (1996)
30. Ting, K.M., Witten, I.H.: Stacking Bagged and Dagged Models. In: 14th Int. Conf. on Machine Learning, San Francisco, CA, pp. 367–375 (1997)

31. Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J.: Rotation Forest: A New Classifier Ensemble Method. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(10), 1619–1630 (2006)
32. Hajek, P., Michalak, K.: Feature Selection in Corporate Credit Rating Prediction. Knowledge-Based Systems 51, 72–84 (2013)
33. Platt, J.C.: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In: Schoelkopf, B., Burges, C., Smola, A. (eds.) Advances in Kernel Methods - Support Vector Learning, pp. 185–208. MIT Press, Cambridge (1998)
34. Kearney, C., Liu, S.: Textual Sentiment in Finance: A Survey of Methods and Models. International Review of Financial Analysis 33, 171–185 (2014)