International Journal of Economics and Finance; Vol. 7, No. 8; 2015
ISSN 1916-971X E-ISSN 1916-9728
Published by Canadian Center of Science and Education

Bankruptcy Prediction Using Support Vector Machines and Feature Selection During the Recent Financial Crisis

Umberto Dellepiane[1], Michele Di Marcantonio[2], Enrico Laghi[3] & Stefania Renzi[1]

[1] Department of Computer, Control and Management Engineering Antonio Ruberti, University of Rome "La Sapienza", Rome, Italy

[2] School of Statistical Science, University of Rome "La Sapienza", Rome, Italy

[3] Department of Law and Economics of Productive Activities, University of Rome "La Sapienza", Rome, Italy

Correspondence: Michele Di Marcantonio, School of Statistical Science, University of Rome "La Sapienza", Viale Regina Elena 295, 00161 Rome, Italy. Tel: 39-333-384-1228. E-mail: michele.dimarcantonio@uniroma1.it

Received: May 20, 2015 Accepted: June 4, 2015 Online Published: July 25, 2015

doi:10.5539/ijef.v7n8p182 URL: http://dx.doi.org/10.5539/ijef.v7n8p182

Abstract

This study aims at identifying an optimal set of features for predicting firms bankruptcy events in the current macroeconomic context. To this aim, among many financial features, we propose new country-specific factors which consider the macroeconomic conditions of the countries where firms operate. Our forecasting model is based on Support Vector Machines (SVMs), which are tools employed in supervised learning. Firstly, starting from a wide set of variables commonly used for bankruptcy prediction we assess the general effectiveness of SVMs also in comparison with the performances of other commonly used methods. Secondly, we try to improve the accuracy of forecasts by selecting optimal subsets of variables through a feature selection method. The results show that, in the current socio-economic context, the conjunct use of SVMs and the proposed feature selection technique significantly improves the accuracy of bankruptcy predictions compared to the performance of the other methods examined. Furthermore, we show that the proposed country-specific factors are relevant information for predicting the failure of firms and that most of the ratios proposed by Altman in 1968 are still relevant nowadays.

Keywords: bankruptcy prediction, default risk, credit scoring, support vector machines, feature selection, data mining, country-specific factors

1. Introduction

Since 1968, when Altman proposed his z-score test (Altman, 1968), many authors have been studying alternative ways in bankruptcy prediction. The reader could refer to many works in the literature. An almost complete survey on this topic is proposed in Ravi Kumar and Ravi (2007).

One of the most active researching field is focused on the use of Neural Networks (NNs) techniques. In Atiya (2001) and du Jardin (2010), the authors propose a survey of the literature on bankruptcy prediction through NNs. Recently, Cho et al. (2009) proposed an integrative model which combines some of the principal data mining techniques for bankruptcy prediction. Moreover, in Kim and Kang (2010) a method for improving the performance of classification and prediction has been proposed together with NNs. Other comparisons between well known methods in bankruptcy prediction literature are proposed by Ince and Aktan (2009) and Tseng and Hu (2010), while in Olson (2012) a review of different data mining tools is carried out, comparing the results achieved by each examined method. A comparison between the results achieved by a back-propagation NN and a Multivariate Discriminant Analysis (MDA) is described in Lee and Choi (2013). An attempt of extending the set of explanatory variables, including accounting, market and macroeconomic variables, is proposed in Tinoco and Wilson (2013).

Moreover it is important to mention authors which implemented alternative methods, such as Genetic Algorithm (GA), in order to overcome the limitations of the methods previously proposed. In Tsakonas et al. (2006) the authors applied neural logic networks by means of genetic programming. Then, through an experimentation on virtual firms, Min and Jeong (2009) proposed a binary classification method based on GA. Finally, a different approach is adopted by Jeong et al. (2012), where the goal of the work is not to improve the prediction model itself but the method used in the prediction, in this case the architecture of the NN system.

Our study contributes to the existing literature by finding a better set of features to predict bankruptcy events of firms in the recent economic context. The first step was to make a comparative analysis between the performances of methods commonly used in literature-namely the Linear Discriminant Analysis and the Logistic regression-and the performances of SVMs, that recently have gained more and more interest in the scientific community. As a benchmark for this initial analysis we considered Altman's z-score model, which still nowadays is a very common method to assess the financial capabilities of a firm. Afterwards, in order to capture the economic situation, we have provided to our models a large set of information (variables), including new macroeconomic indicators which, according to our opinion, is a noteworthy aspect. Presuming that not all the provided variables were actually useful in the forecast, we performed a feature selection. In this way we identified the variables with more attractive information content respect to today's context.

In the past, few researchers tried to use algorithms for selecting optimal subsets of predictors. In particular, in du Jardin (2010) the author attempted to analyze the impacts of correctly selecting the variables subset on the I type error, while in Tsai (2009) a comparison of well known feature selection methods is made through the employment of a Multilayer Perceptron (MLP). Recently, Zhou et al. (2014) proposed an approach for feature selection and parameters optimization used in the NN system for the prediction based on one of the most common GA techniques.

Relating to previous studies that are similar to our research, Min and Lee (2005) already used SVMs for bankruptcy prediction and Salcedo-Sanz et al. (2004), in particular, was a pioneer in the use of feature selection and SVMs for predicting insolvency of non-life insurance companies. Furthermore, Xie et al. (2011) found empirical evidence on the relevance of external market variables for bankruptcy prediction based on SVMs. In the light of their findings, in our work we propose and test the efficacy of using jointly all these elements. Furthermore, we propose the definition of new macroeconomic variables and we provide new empirical evidence on their relevance at a world-level also in the light of the recent financial crisis. Our work allows to infer whether the explanatory variables of Altman's z-score model are still relevant nowadays and if the examined macroeconomic factors (and/or other variables considered in the feature selection) have gained importance after the burst of the recent financial crisis. Finally, since we apply on the same dataset also other commonly used bankruptcy prediction methods, our results enable us to compare the performances of the examined methods in the recent economic context.

The paper is organized as follows. The classification methods used in this study–i.e. the Altman's z-score model, Linear Discriminant Analysis, Logistic regression and SVMs–are briefly described in section 2. Section 3 is devoted to the description of the empirical dataset used and to the procedure adopted for developing the analyses. Section 4 reports the results of our empirical study, distinguishing among the results obtained by using of Altman's z-score features, using the full set of features and using the attributes selected through a feature selection technique respectively. The final section concludes.

2. Classification Methods

2.1 Altman's Z-Score Model

Altman's model tested in this work is the well-known z-score model which was developed by Altman in (1968). Subsequent studies on this topic and extensions of Altman's original model can be found in Altman (1973, 1977), Altman *et al*. (1977), Altman *et al*. (1994), Altman and McGough (1974) and Altman and Hotchkiss (2006). The original discriminant function developed by Altman in 1968 is a linear combination of five common business ratios, weighted by their respective coefficients (equation 1):

$$Z^j = 0.012x_1^j + 0.014x_2^j + 0.033x_3^j + 0.006x_4^j + 0.999x_5^j \quad (1)$$

where, for each firm *j*, we have: Z^j is the overall index (z-score); x_1^j is (Working Capital)^j/(Total Assets)^j; x_2^j is (Retained Earnings)^j/(Total Assets)^j; x_3^j is (Earnings Before Interest and Taxes)^j/(Total Assets)^j; x_4^j is (Market Value Equity)^j/(Book Value of Total Debt)^j; x_5^j is (Sales)^j/(Total Assets)^j.

Once calculated the value of Z^j for a given company *j*, the classification of firm *j* into the solvent or the non-solvent group is performed having regard to the critical values of Z determined by Altman. In particular, a firm *j* is expected to be solvent if Z^j is greater than an upper bound Z_{up}, while it is expected to go bankrupt if Z^j is below a lower bound Z_{low}. The area between Z_{low} and Z_{up} is defined as the zone of ignorance or gray area because of the susceptibility to misclassification, hence for $Z_{low} \leq Z^j \leq Z_{up}$ the model does not classify company j into any group. The original Altman model is characterized by a fixed set of variables and fixed values of the coefficients, while the critical values of Z can vary depending on the specific characteristics of the sample of firms being analyzed. For the purposes of this work, we consider the values Z_{low} = 1.81 and Z_{up} = 2.67.

2.2 Other Classification Methods

The objective of classification methods is to classify observations into one, two or more mutually exclusive and exhaustive groups using information about a given set of variables measured for each observation.

Among these, linear classification methods are aimed at detecting one or several linear functions of the given set of variables to be used for classification. Traditionally, two of the main linear classification methods are Linear Discriminant analysis (LDA) and Logistic regression or Logit regression (Logit). Details on the LDA methodology can be found in Kolossa and Haeb-Umbach (2011), while for a description of the Logit regression structure see Boyacioglu (2009) and Tinoco and Wilson (2013). Other linear regression methods can be found in Varmuza and Filzmoser (2009).

Relating to non-linear classification method, in this work we tested the prediction accuracy of Support Vector Machines, with special reference to the problem of training a classifier able to distinguish between two sets of points. For more details about the theory or the solution approach for the SVMs learning problem applied in this study see Vapnik (1998).

3. Empirical Study

3.1 Data

The models that we present in this paper are tested using samples of companies constituted by both solvent and non-solvent firms using data collected on Bloomberg. To this aim, we selected all the 6,929 companies that were included in the equity index Market World published by Thomson Reuters-Datastream at the date of the research.

We first determined the sample of non-solvent companies (the NS-Group) by requiring the bankruptcy date to be greater than January 1st, 2007, in order to select only the firms which went bankrupt during or after the recent financial crisis. For the purposes of developing balanced analyses, we selected a random sample of solvent firms characterized by the same cardinality of the NS-Group (the S-Group). Consequently to the use of a random criterion, the selected sample covers the whole spectrum from healthy to border-line companies, thus avoiding any selection bias, as described in Atiya (2001). In order not to alter the analyses with exchange rate effects, since all the explanatory variables are ratios between monetary values or pure numbers, all monetary values were collected in their native currency. For each examined independent variable, extreme values were excluded from the group of observations through a Winsorization procedure eliminating the values that are below the 0.10% percentile and above the 99.90% percentile. Finally, the observations that did not have the necessary data to determine all the examined explanatory variables were excluded.

Data were collected on a yearly basis, at the date of December 31st of each reference year, where the reference year is specific for each company and it is determined differently for solvent and non-solvent firms. For the S-Group it was chosen 2012, while for the NS-Group we considered two cases:

1) The date of December 31st of the year before the year in which companies went bankrupt, therefore between 0 and 1 years before the bankruptcy date of each firm *j*;
2) The date of December 31st of two years before the year in which companies went bankrupt, therefore between 1 and 2 years before the bankruptcy date of each firm *j*.

This choice is aimed at examining the predictive capability of the models, one year ahead (case 1) and two years ahead (case 2) from the moment of the analyses, for each combination of set of variables and default forecasting method. The assumption made is that each year investors could apply the models at the reference date of December 31st.

Once defined both the S-Group and the NS-Group(s), we selected two final samples: Sample 1, constituted by the union of the observations of the S-Group and the observations of the NS-Group *sub* case 1, and Sample 2, constituted by the union of the observations of the S-Group and the observations of the NS-Group *sub* case 2. Finally, we defined the dichotomous dependent variable indicating by 0 solvent firms and by 1 non-solvent firms. Table 1 shows descriptive statistics of all the examined variables for both Sample 1 and Sample 2.

Table 1. Descriptive statistics of the examined explanatory variables for sample 1 and sample 2

Index	Ratio	Sample 1, cardinality NS1 = 212				Sample 2, cardinality NS2 = 204			
		Mean	Dev. Std.	Min	Max	Mean	Dev. Std.	Min	Max
R1	Working Capital/Total Assets	-0,261	2.142	-27.826	0,741	-0,114	1.275	-14.531	0,741
R2	Retained Earnings/Total Assets	-2.068	8.551	-74.984	1.057	-2.750	11.823	-119.929	1.057
R3	EBIT/Total Assets	-0,125	0,553	-3.485	0,449	-0,152	0,623	-3.722	0,449
R4	Market cap./Total Liabilities	2.058	4.825	0,001	56.897	2.302	4.910	0,004	56.897
R5	Sales 12M/Total Assets	0,830	0,785	0,010	5.908	0,893	0,948	0,010	6.592
R6	Current assets/Current liabilities	1.478	1.254	0,020	10.537	1.682	2.056	0,004	25.022
R7	Current assets/Total Assets	0,367	0,238	0,018	0,976	0,386	0,237	0,018	0,976
R8	Current liabilities/Total Assets	0,628	2.179	0,012	28.766	0,500	1.270	0,012	14.590
R9	Quick ratio	0,939	0,992	0,004	8.367	1.000	1.000	0,000	8.367
R10	EBITDA/Total Assets	-0,073	0,551	-3.326	0,488	-0,104	0,617	-3.651	0,488
R11	Cash & simil./Total Liabilities	0,188	0,384	0,000	3.819	0,214	0,420	0,000	3.819
R12	Cash & simil./Total Assets	0,092	0,111	0,000	0,738	0,104	0,127	0,000	0,825
R13	Cash & simil./Net Debt	0,052	6.657	-73.988	52.904	1.021	6.652	-4.285	69.020
R14	Total Equity/Total Assets	0,016	2.210	-27.766	0,869	0,092	1.375	-13.590	0,869
R15	Net Debt/Total Equity	1.068	6.977	-9.624	85.913	-0,207	17.373	-228.278	85.913
R16	Net Debt/Total Assets	0,290	0,730	-0,728	6.993	0,297	0,975	-0,798	11.617
R17	EBITDA/Gross xed assets	-1.126	12.667	-124.718	51.987	-1.475	14.881	-157.540	51.987
R18	EBIT/Sales 12M	-2.009	10.260	-89.170	0,873	-2.049	10.418	-91.369	0,873
R19	EBITDA/Sales 12M	-1.788	9.686	-78.306	1.515	-1.862	10.148	-89.627	0,917
R20	Net Income/Total Equity	-0,391	3.149	-29.422	5.682	0,120	5.664	-45.739	52.596
R21	Net Income/Total Assets	-0,321	1.825	-25.115	0,440	-0,328	1.250	-10.759	0,496
R22	EBIT/Passive interests	1.819	275.6	-2,883.0	2.260.7	-22.8	483.3	-5,629.90	2,260.7
R23	Sales 12M/Total Liabilities	1.310	1.249	0,006	7.842	1.378	1.341	0,003	8.697
R24	Sales 12M/Total Equity	1.951	5.804	-17.482	50.362	0,812	12.823	-150.058	50.362
R25	Current assets/Sales 12M	1.686	7.403	0,042	96.215	2.181	8.873	0,042	96.215
R26	Working Capital/Sales 12M	-0,836	8.821	-122.250	15.932	-1.903	26.656	-356.671	35.747
R27	Cash & simil./Sales 12M	0,445	1.481	0,000	12.131	0,699	3.927	0,000	49.942
R28	Interests expenses/Sales 12M	0,348	2.451	0,000	33.065	0,561	3.315	0,000	33.065
R29	Sovereign rate 1Y	1.087	1.957	-0,214	11.542	1.413	2.434	-0,214	11.897
R30	Sovereign rate 10Y	2.787	1.747	0,526	12.550	3.134	2.037	0,526	12.550
R31	Sovereign rating spread	69.741	114.155	0,000	900.000	62.623	93.913	0,000	587.500

The variables from 1 to 28 shown in Table 1 are ratios derived from the accounting data reported in the annual financial reports of firms, which have been determined considering the variables most commonly used by researchers to predict bankruptcy (Du Jardin, 2010; Ravi Kumar & Ravi, 2007). The variables from 29 to 31 instead are country-specific factors, some of which have been determined following an approach that, to the extent of our knowledge, has not been used in previous studies for predicting bankruptcy. Similar research on this topic can be found in Laghi et al. (2013) and Lam (2004).

The Sovereign rate 1 Y and the Sovereign rate 10 Y are the gross sovereign rates with maturity of 1 and 10 years respectively, whose values, for a given firm *j*, have been set equal to the market value of the official gross rate of the country where company *j* operated at December 31st of its reference year.

The sovereign rating spread (SRS) was determined on the basis of the historical sovereign ratings (R) issued by Standard and Poor's (2013). Official ratings are issued in a textual form (e.g. A- or BBB+) so they must be converted to numerical values in order to be considered in the analyses. To this aim, differently than other authors who assume that one-notch movements have the same effect on credit spreads independently from the asset class (e.g. Aunon-Nerin et al., 2002), we associated different numeric values V_e with each element R_e of the S&P rating scale, hence $V_e = f(R_e)$. The values of V_e were set equal to the risk spreads-in basis points-associated with Moody's ratings estimated by Damodaran (2012) as at January of each reference year. Our only contribution to those numerals has been the association of those risk spreads, originally associated with Moodys ratings, with the corresponding ratings of the S&P scale. Table 2 shows the numeric values V_e estimated by Damodaran (2012) between 2005 and 2012 that we assigned to each element R_e of the reference S&P rating scale.

Table 2. The numeric values V_e associated with each element R_e of the S&P rating scale

		V_e							
e	R_e	2005	2006	2007	2008	2009	2010	2011	2012
1	AAA	0	0	0	0	0	0	0	0
2	AA+	50	35	35	70	30	25	25	25
3	AA	55	50	50	100	60	50	50	50
4	AA-	60	60	60	120	75	70	70	70
5	A+	60	70	70	140	90	85	85	85
6	A	80	80	80	160	105	100	100	100
7	A-	90	85	85	175	120	115	115	115
8	BBB+	110	100	100	200	160	150	150	150
9	BBB	120	115	115	225	180	175	175	175
10	BBB-	135	135	135	260	200	200	200	200
11	BB+	225	200	200	300	250	240	240	240
12	BB	270	250	250	400	300	275	275	275
13	BB-	360	300	300	525	350	325	325	325
14	B+	400	350	350	650	450	400	400	400
15	B	500	400	400	750	550	500	500	500
16	B-	600	450	450	900	650	600	600	600
17	CCC+	-	600	600	1,200	750	700	700	700
18	CCC	-	675	675	1,350	900	850	850	850
19	CCC-	-	750	750	1,500	1,000	1,000	1,000	1,000

In addition, also positive and negative outlooks issued by S&P have been considered according to the following formula (equation 2):

$$SRS_j^t = f\left(R_e^t, Outlook^t\right) = if \left\{ \begin{array}{ll} Outlook = Positive(PO^t) & \longrightarrow (V_e^t + V_{e-1}^t)/2 \\ Outlook = Stable & \longrightarrow V_e^t \\ Outlook = Negative(NO^t) & \longrightarrow (V_e^t + V_{e+1}^t)/2 \end{array} \right\} \quad (2)$$

where SRS_j^t indicates the sovereign rating spread attributed to the *j-th* firm, R_e^t is the rating at time *t* of the country where the company operates, V_e^t e is the numerical value associated with R_e^t and PO^t and NO^t represent respectively the positive or negative outlook eventually issued for the same country at time *t*.

3.2 Implementation

The analyses presented in this study were developed using different softwares for each type of default prediction model: the Altman model was implemented using Excel, the LDA and the Logit regression models were applied using Stata and SVMs were trained and tested through Weka (*Waikato Environment for Knowledge Analysis*). Relating to Weka, we used the *libSVM* classification method, which implements the SMO algorithm for kernelized support vector machines. More details about this method are available in Chang and Lin (2011).

In order to apply and test the general predictive capability of LDA, Logit regressions and SVMs, two different datasets are required: a training set, used to train the models, and a test set, used to verify the efficiency of the learning procedure. Hence, in order to guarantee the comparability of results among the examined models, although the test proposed by Altman is able to predict the default of a firm using only the information related to that firm, the tests using Altman's model were made considering only the instances within the test set. We split each sample in two parts randomly, using the proportion 2/3 and 1/3 for the training and the test set respectively. We repeated this procedure five times generating five couples of training and test sets with a different composition. This procedure was followed for both Sample 1 and Sample 2.

Regarding to the setting of the SVMs parameters we followed a strategy based on a grid search procedure. A grid search simply consists in an exhaustive searching through a manually-specified subset of possible values for the parameters. According to this strategy, we defined a certain hyperparameter space and evaluated the performances achieved by the SVMs with the different settings. At the end of these evaluations we identified the best configuration as: *i)* the Gausian Kernel with parameter $\gamma = 0.01$; *ii)* $C = 100$.

Since from a practical point of view the problem of classifying incorrectly a non-solvent firm is more serious than classifying incorrectly a solvent firm, we decided to assign different weights to the two classes, namely

$\alpha = 1.0$ if the firm is insolvent and $\alpha = 0.5$ if the firm is solvent. The best setting for the SVMs resulted to be the same for both the prediction horizons.

4. Results

4.1 Altman's Z-Score Features

In this section we analyze the results achieved by Altman's z-score test and the other examined methods using the same set of ratios of the former test. As already said, one of the strengths of this model is related to the property of realizing a prediction using only the data of a firm. On the other hand the model is characterized by two main weakness: *i)* Altman's z-score model simply consists in a linear regression with fixed values for the five regressors. In a context where the world economic scenario has changed, these values may not reflect the nowadays situation; *ii)* Altman's z-score model provides a range of values (*the gray zone*) wherein a firm is classified neither solvent nor insolvent.

Considering the previous observation, in order to compare the results obtained with Altman's model with the ones achieved using the other methods, we forced Altman's model to classify a firm as solvent or insolvent in case the score lies within the gray zone. In particular we defined two different and reasonable rules, reported in Table 3.

Table 3. The rules adopted for applying the Altman's z-score model

Rule 1		Rule 2	
$Z > 1.81$	Safe Zone	$Z > 2.24$	Safe Zone
$Z \leq 1.81$	Distress Zone	$Z \leq 2.24$	Distress Zone

According Rule 1 we consider as unique threshold value Z_{low} while Rule 2 considers as unique threshold value the midpoint between Z_{low} and Z_{up}. The results achieved by these two configurations of Altman's test are compared with the ones obtained by the LDA, the Logit and the SVM approach. In Table 4 we reported what we achieved in the prediction one year ahead, while in Table 5 the same results are reported considering the prediction two years ahead.

Before starting to illustrate the results, some explanations about the notation used in the main tables:

S: represents a solvent firm while NS represents a non-solvent firm. In particular NS/NS is the number of firms correctly classified as non-solvent, while NS/S represents the number of non-solvent firms classified as solvent. Similarly, S/NS is the number of solvent firms classified as non-solvent and S/S represents the number of solvent firms correctly classified;

cc/all: represents the percentage of instances correctly classified;

NS/aNS: represents the percentage of non-solvent firms correctly classified.

First of all we consider the results reported in Table 4 and in particular the ones reported in row Mean. In the evaluation of the performances it is possible to adopt two different points of view. The former has a theoretical nature and it is focused on evaluating the efficiency of the models according to the ratio of instances correctly classified over the cardinality of the sample. The latter gives more importance to the practical purpose of this bankruptcy test. In fact, from the point of view of investors, the misclassification of a solvent firm (S/NS) implies only the loss of an investment opportunity, while the misclassification of a non-solvent firm (NS/S) could determine a capital loss, as a consequence of investing in that bankrupt company. In this context the efficiency is evaluated by the percentage of how many instances are correctly classified as non-solvent among all the non-solvent instances in the sample. Following the first criterion, it seems that the best tool is the Logit since it reaches a mean value of 87.04%. According to the second criterion, the best method is Altman with threshold value 2.24, which obtains on average the 94.84% of correct classifications among all the non-solvent instances. At this point a question arises naturally: which is the best tool in this case? In order to answer this question, the user should express before his preference for one criterion. It is important to notice that, on the basis of the results, these two criteria seem to be in contrast. In fact while the Logit reaches an 87% of successes following the first criterion, it obtains only 77.54% following the second one. Likewise Altman 2.24 reaches 73.57% and 94.84% according to the first and the second criterion respectively.

However, the aim of this work is to find a tool which can satisfy both the criteria in a suitable manner, in other words a model that reaches a balance between the two criteria. Thus, without expressing any preference on the two criteria, we can see that the most balanced tool is the SVM, which obtains 83.66% of successes according to the first criterion and 86.67% according to the second one.

Repeating the former analysis on the results obtained in Table 5 the SVMs are again the best tool with 77.65% and 82.22% of successes. In general it can be noticed that when the prediction framework is extended, passing from one to two years ahead, the overall performances decrease.

Table 4. Features proposed by Altman in his z-score test, prediction one year ahead

Sample		Altman 1.81		Altman 2.24		LDA		Logit		SVM	
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	30	3	32	1	18	15	25	8	30	3
Sample 1A	S	12	26	19	19	1	37	2	36	8	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		78.87%	90.91%	71.83%	96.97%	77.46%	54.55%	85.92%	75.76%	84.51%	90.91%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	30	3	32	1	14	19	26	7	28	5
Sample 1B	S	15	23	22	16	0	38	2	36	6	32
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		74.65%	90.91%	67.61%	96.97%	73.24%	42.42%	87.32%	78.79%	84.51%	84.85%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	25	5	27	3	14	16	24	6	24	6
Sample 1C	S	16	25	25	16	1	40	3	38	11	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		70.42%	83.33%	60.56%	90.00%	76.06%	46.67%	87.32%	80.00%	76.06%	80.00%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	24	2	25	1	17	9	23	3	24	2
Sample 1D	S	14	31	17	28	2	43	2	43	9	36
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.46%	92.31%	74.65%	96.15%	84.51%	65.38%	92.96%	88.46%	84.51%	92.31%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	29	5	32	2	19	15	22	12	29	5
Sample 1E	S	10	27	2	23	1	36	1	36	3	34
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		78.87%	85.29%	93.22%	94.12%	77.46%	55.88%	81.69%	64.71%	88.73%	85.29%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.06%	88.55%	73.57%	94.84%	77.75%	52.98%	87.04%	77.54%	83.66%	86.67%

Table 5. Features proposed by Altman in his z-score test, prediction two years ahead

Sample		Altman 1.81		Altman 2.24		LDA		Logit		SVM	
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	22	7	25	4	9	20	17	12	23	6
Sample 2A	S	14	25	21	18	0	39	4	35	11	28
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		69.12%	75.86%	63.24%	86.21%	70.59%	31.03%	76.47%	58.62%	75.00%	79.31%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	22	6	25	3	14	14	20	8	23	5
Sample 2B	S	19	21	26	14	2	38	8	32	15	25
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		63.24%	78.57%	57.35%	89.29%	76.47%	50.00%	76.47%	71.43%	70.59%	82.14%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	24	3	25	2	20	7	22	5	26	1
Sample 2C	S	13	28	22	19	5	36	5	36	10	31
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.47%	88.89%	64.71%	92.59%	82.35%	74.07%	85.29%	81.48%	83.82%	96.30%

		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	23	4	25	2	9	18	17	10	18	9
Sample 2D	S	12	29	16	25	0	41	2	39	4	37
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.47%	85.19%	73.53%	92.59%	73.53%	33.33%	82.35%	62.96%	80.88%	66.67%
		NS	S	NS	S	NS	S	NS	S	NS	S
	NS	24	6	26	4	9	21	18	12	26	4
Sample 2E	S	14	24	22	16	1	37	1	37	11	27
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		70.59%	80.00%	61.76%	86.67%	67.65%	30.00%	80.88%	60.00%	77.94%	86.67%
Mean		cc/all	NS/Ans	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		71.18%	81.70%	64.12%	89.47%	74.12%	43.69%	80.29%	66.90%	77.65%	82.22%

4.2 Full Set of Features

As said at the beginning of this section the first goal of this work is to assess the accuracy of SVMs at their full potential. Keeping this in mind we extended the number of ratios from the 5 proposed by Altman to 31, namely those listed in Table 1. For this reason Altman's model is excluded from the following analysis and the next comparisons will be made only between LDA, Logit regression and SVMs. The results of the predictions one year and two years ahead are shown in Table 6 and Table 7 respectively.

Table 6. All features proposed, prediction one year ahead

Sample		LDA		Logit		SVM	
		NS	S	NS	S	NS	S
	NS	20	13	26	7	28	5
Sample 1A	S	10	28	7	31	8	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		67.61%	60.61%	80.28%	78.79%	81.69%	84.85%
		NS	S	NS	S	NS	S
	NS	15	18	22	11	28	5
Sample 1B	S	4	34	4	34	7	31
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		69.01%	45.45%	78.87%	66.67%	83.10%	84.85%
		NS	S	NS	S	NS	S
	NS	14	12	19	7	25	1
Sample 1C	S	6	35	6	35	8	33
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.06%	63.33%	83.10%	80.00%	83.10%	86.67%
		NS	S	NS	S	NS	S
	NS	14	13	21	6	23	4
Sample 1D	S	4	41	8	37	19	26
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.46%	53.85%	78.87%	73.08%	71.83%	96.15%
		NS	S	NS	S	NS	S
	NS	21	13	25	9	30	4
Sample 1E	S	2	35	2	35	8	29
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		78.87%	61.76%	84.51%	73.53%	83.10%	88.24%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		73.80%	57.00%	81.13%	74.41%	80.56%	88.15%

Table 7. All features proposed, prediction two years ahead

Sample		LDA		Logit		SVM	
Sample 2A		NS	S	NS	S	NS	S
	NS	16	13	17	12	24	5
	S	4	35	11	28	10	29
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		75.00%	55.17%	66.18%	58.62%	77.94%	82.76%
Sample 2B		NS	S	NS	S	NS	S
	NS	17	11	18	10	26	2
	S	7	33	12	28	8	32
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		73.53%	60.71%	67.65%	64.29%	85.29%	92.86%
Sample 2C		NS	S	NS	S	NS	S
	NS	18	9	21	6	24	3
	S	5	36	4	37	6	35
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		79.41%	66.67%	85.29%	77.78%	86.76%	88.89%
Sample 2D		NS	S	NS	S	NS	S
	NS	14	13	21	6	23	4
	S	0	41	7	34	11	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		80.88%	51.85%	80.88%	77.78%	77.94%	85.19%
Sample 2E		NS	S	NS	S	NS	S
	NS	18	12	21	9	26	4
	S	3	35	7	31	8	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.94%	60.00%	76.47%	70.00%	82.35%	86.67%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.35%	58.88%	75.29%	69.69%	82.06%	87.27%

We can see that, when we pass from 5 to 31 explanatory variables, the performances of all the models decrease with respect to the first criterion while increase with respect to the second. This means that, considering the performance related to the capabilities of generalization, the information within all the features generates a sort of noise or redundancy. On the other hand, the capabilities of the models to correctly classify a non-solvent firm are increased. However, similarly to what we observed with reference to the case with Altman's variables, the SVM model seems again the most balanced method for both the forecast horizons.

The results achieved in this section let us to suppose that the information carried out by the ratios used by Altman may no longer be exhaustive in the nowadays economical context. However, extending too much and without a particular logical assumption the number of ratios, the performances do not improve according both the criteria. This led us to exploit feature selection techniques. In the following section we introduce the feature selection scheme implemented and then we show the results achieved by considering only the selected variables.

4.3 Feature Selection

Before inducing a model, we have a set of information collected in some features and, most of the time, we do not know which part of it is the most significant. Theoretically, having more features should result in more discriminating power. However, practical experience with machine learning algorithms has shown that this is not always the case.

In this regard, we focus the attention on the importance of selecting the features (also called attributes) to be used in the model. Feature or attribute selection is a technique whose goal is to form a subset of the initial features of the problem aiming to improve the performance of the underlying model, both in terms of correctness and fastness. The question now is whether it is possible to discard some features, and how to select the correct subset of variables. There is not an univocal answer to the question but the underlying logic may change depending on the particular method implemented. In the previous section we underlined that extending the set of variables the performances of the models improved in one direction while worsened in the second. Using the feature selection on the SVM model our aim is to improve the prediction performance according to both the criteria and to verify if the subset of selected attributes is able to improve also the results achieved by the other two classification methods. In order to implement this kind of selection we used again the Weka software. Weka offers different kind of evaluators and search methods. Among all of them we selected as evaluation method the *CfsSubsetEval* and as search method the *BestFirst*: *CfsSubsetEval* stays for Correlation-based Feature Selection Subset Evaluation and the basic idea underlying this algorithm is to prefer subsets of features that are highly correlated with the class while having low autocorrelation; *BestFirst* is the selected search method. We used it in a forward search direction which means that it adds greedy the features starting from the empty set.

It seems clear now that the results achieved by training a machine learning using the subset of feature derived by a feature selection algorithm strictly depend on the composition of the sample. The higher the size of the sample, the lower this dependency. However, we are interested in finding an optimal set of ratios in order to realize a better prediction of non-solvency/solvency of a firm characterized also by a certain stability. For this reason we need to introduce a deeper approach in order to verify the performances of the SVMs. To this aim, for both Sample 1 and Sample 2, we generated five different subsets of features applying the feature selection algorithm to each of the five samples. Then, considering the *i-th* subset of variables, we trained four other SVMs, each one for the remaining four couples of training and test sets. In this way we can assess the fitness of the particular subset of feature not only for the sample on which the feature selection was made. The underlying idea is that selecting among all the subsets of features the one which on average has a better performance we obtain the group of attributes that better generalizes the bankruptcy events. Once identified the optimal subset of features, we reused the LDA and the Logit models considering only those features in order to obtain again comparable results. The attributes selected for the prediction one year ahead are reported in Table 8, while Table 9 shows the results achieved by applying the different models considering only those features.

Table 8. Features selected for the prediction one year ahead

Index	Ratio
R2	Retained Earnings/Total Assets
R3	EBIT/Total Assets
R4	Market cap./Total Liabilities
R8	Current liabilities/Total Assets
R21	Net Income/Total Assets
R30	Sovereign rate 10Y
R31	Sovereign rating spread

Table 9. Features selected by the algorithm, prediction one year ahead

Sample		LDA		Logit		SVM	
		NS	S	NS	S	NS	S
	NS	17	16	26	7	32	1
Sample 1A	S	1	37	2	36	3	35
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.06%	51.52%	87.32%	78.79%	94.37%	96.97%
		NS	S	NS	S	NS	S
	NS	13	20	27	6	25	8
Sample 1B	S	3	35	4	34	5	33
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		67.61%	39.39%	85.92%	81.82%	81.69%	75.76%
		NS	S	NS	S	NS	S
	NS	17	13	23	7	29	1
Sample 1C	S	2	39	3	38	5	36
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		78.87%	56.67%	85.92%	76.67%	91.55%	96.67%
		NS	S	NS	S	NS	S
	NS	19	7	22	4	25	1
Sample 1D	S	3	42	3	42	5	40
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		85.92%	73.08%	90.14%	84.62%	91.55%	96.15%

		NS	S	NS	S	NS	S
	NS	20	14	25	9	32	2
Sample 1E	S	0	37	1	36	1	36
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		80.28%	58.82%	85.92%	73.53%	95.77%	94.12%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.75%	55.90%	87.04%	79.08%	90.99%	91.93%

Similarly, in Table 10 are resumed the attributes selected for the prediction two years ahead and in Table 11 the results achieved by applying the different models considering only those features.

Table 10. Features selected for the prediction two years ahead

Index	Ratio
R3	EBIT/Total Assets
R14	Total Equity/Total Assets
R21	Net Income/Total Assets
R29	Sovereign rate 1Y
R30	Sovereign rate 10Y

Table 11. Features selected by the algorithm, prediction two years ahead

Sample		LDA		Logit		SVM	
		NS	S	NS	S	NS	S
	NS	10	19	18	11	25	4
Sample 2A	S	4	35	3	36	2	37
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		66.18%	34.48%	79.41%	62.07%	91.18%	86.21%
		NS	S	NS	S	NS	S
	NS	19	9	21	7	25	4
Sample 2B	S	7	33	3	37	3	36
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.47%	67.86%	85.29%	75.00%	89.71%	86.21%
		NS	S	NS	S	NS	S
	NS	18	9	23	4	26	1
Sample 2C	S	10	31	2	39	4	37
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		72.06%	66.67%	91.18%	85.19%	92.65%	96.30%
		NS	S	NS	S	NS	S
	NS	12	15	19	8	22	5
Sample 2D	S	1	40	5	36	3	38
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.47%	44.44%	80.88%	70.37%	88.24%	81.48%
		NS	S	NS	S	NS	S
	NS	15	15	23	7	28	2
Sample 2E	S	10	28	4	34	5	33
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		63.24%	50.00%	83.82%	76.67%	89.71%	93.33%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		70.88%	52.69%	84.12%	73.86%	90.29%	88.70%

It is evident that SVMs obtain definitively higher and more stable performances than the ones obtained with the other two models for the predictions both one year ahead and two years ahead. Moreover we notice that using SVM models with the appropriate information it is possible to reach a prediction accuracy higher and more balanced than the one obtained with Altman’s variables and with the full set of features.

5. Conclusions

In this work we defined a new set of features, which includes country specific macroeconomic factors, that improves the accuracy of predictions of firms' bankruptcy events in the recent economic context. Using SVMs jointly to feature selection techniques, we identified the optimal subset of variables and assessed whether the ratios proposed by Altman in 1968 are still relevant nowadays for bankruptcy prediction. Furthermore, we applied also other commonly used bankruptcy prediction methods on the same dataset and we compared their performances with the ones of SVMs.

The results show that in the current socio-economic context the conjunct application of SVMs and the proposed feature selection method significantly improves the accuracy of bankruptcy predictions compared to the examined traditional set of variables and default prediction methods. In particular, the conjunct use of these elements permits to obtain stable percentages of success around 90% for both one year and two years head predictions.

From an economic point of view, despite the use of a different model (the SVMs) and despite the deep changes of the world economic system that followed the crisis from 2007, most of the ratios proposed by Altman in 1968 result to be still relevant nowadays. Furthermore, the examined macroeconomic indicators appear to be relevant information for predicting bankruptcy. It is evident that in crisis periods, independently from individual firms characteristics, the existence of an economic stability at a country-level and the accessibility to credit have become even more important factors for the survival of a company. With reference to sovereign rating spreads, our study demonstrates that sovereign ratings, though their reliability has been widely criticized in recent years, constitute a relevant information for predicting whether a firm that operates in a given country will go bankrupt or not. However, further research should be conducted on this topic in order to develop an internal model for quantifying the risk spreads associated with each element of the reference rating scale.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589-609. http://dx.doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I. (1973). Predicting railroad bankruptcies in America. *The Bell Journal of Economics and Management Science*, *4*(1), 184-211. http://dx.doi.org/10.2307/3003144

Altman, E. I. (1977). Predicting performance in the savings and loan association industry. *Journal of Monetary Economics*, *3*(4), 443-466. http://dx.doi.org/10.1016/0304-3932(77)90015-0

Altman, E. I., & Hotchkiss, E. (2006). *Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt*. Wiley Finance. http://dx.doi.org/10.1002/9781118267806

Altman, E. I., & McGough, T. (1974). Evaluation of a company as a going concern. *Journal of Accountancy, 138*(6), 50-57.

Altman, E. I., Haldman, R. G., & Narayanan, P. (1977). A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance*, *1*(1), 29-54. http://dx.doi.org/10.1016/0378-4266(77)90017-6

Altman, E. I., Kahya, E., & Theodossiou, P. (1994). *A time series business failure prediction model for us firms*. Technical report: Working Paper, NYU Salomon Center.

Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks*, *12*(4), 929-935. http://dx.doi.org/10.1109/72.935101

Boyacioglu, M. A., Kara, Y., & Baykan, O. K. (2009). Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (sdif) transferred banks in Turkey. *Expert Systems with Applications*, *36*(2), 3355-3366. http://dx.doi.org/10.1016/j.eswa.2008.01.003

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2*(3). http://dx.doi.org/10.1145/1961189.1961199

Cho, S., Kim, J., & Bae, J. K. (2009). An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Systems with Applications*, *36*(1), 403-410. http://dx.doi.org/10.1016/j.eswa.2007.09.060

Damodaran, A. (2012). *Documents risk premiums of other markets as at January of the years from 2005 to 2012*. Retrieved March 30, 2014, from http://pages.stern.nyu.edu/ adamodar/, section Updated Data/Data Sets

Du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, *73*(10-12), 2047-2060. http://dx.doi.org/10.1016/j.neucom.2009.11.034

Ince, H., & Aktan, B. (2009). A comparison of data mining techniques for credit scoring in banking: A managerial perspective. *Journal of Business Economics and Management*, *10*(3), 233-240. http://dx.doi.org/10.3846/1611-1699.2009.10.233-240

Jeong, C., Min, J. H., & Kim, M. S. (2012). A tuning method for the architecture of neural network models incorporating {GAM} and {GA} as applied to bankruptcy prediction. *Expert Systems with Applications*, *39*(3), 3650-3658. http://dx.doi.org/10.1016/j.eswa.2011.09.056

Kim, M. J., & Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications*, *37*(4), 3373-3379. http://dx.doi.org/10.1016/j.eswa.2009.10.012

Kolossa, D., & Haeb-Umbach, R. (2011). *Robust Speech Recognition of Uncertain or Missing Data: Theory and Applications*. Springer. http://dx.doi.org/10.1007/978-3-642-21317-5

Laghi, E., Mattei, M., & Di Marcantonio, M. (2013). Assessing the value relevance of goodwill impairment considering country-specific factors: Evidence from EU listed companies. *International Journal of Economics and Finance*, *5*(7), 32-49. http://dx.doi.org/10.5539/ijef.v5n7p32

Lam, M. (2004). Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis. *Decision Support Systems*, *37*(4), 567-581. http://dx.doi.org/10.1016/S0167-9236(03)00088-5

Lee, S., & Choi, W. S. (2013). A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, *40*(8), 2941-2946. http://dx.doi.org/10.1016/j.eswa.2012.12.009

Min, J. H., & Jeong, C. (2009). A binary classification method for bankruptcy prediction. *Expert Systems with Applications*, *36*(3, Part 1), 5256-5263. http://dx.doi.org/10.1016/j.eswa.2008.06.073

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. http://dx.doi.org/10.1016/j.eswa.2004.12.008

Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, *52*(2), 464-473. http://dx.doi.org/10.1016/j.dss.2011.10.007

Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques a review. *European Journal of Operational Research*, *180*(1), 1-28. http://dx.doi.org/10.1016/j.ejor.2006.08.043

Salcedo-Sanz, S., DePrado-Cumplido, M., Segovia-Vargas, M. J., Pérez-Cruz, F., & Bousoño-Calzón, C. (2004). Feature selection methods involving support vector machines for prediction of insolvency in non-life insurance companies. *Intelligent Systems in Accounting, Finance and Management*, *12*(4), 261-281. http://dx.doi.org/10.1002/isaf.255

Standard & Poor's. (2013). *Sovereign rating and country t&c assessment histories*. Retrieved March 30, 2014, from http://www.standardandpoors.com/ratings/sovresearch/en/us

Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, *30*, 394-419. http://dx.doi.org/10.1016/j.irfa.2013.02.013

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

Tsakonas, A., Dounias, G., Doumpos, M., & Zopounidis, C. (2006). Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming. *Expert Systems with Applications*, *30*(3), 449-461. http://dx.doi.org/10.1016/j.eswa.2005.10.009

Tseng, F. M., & Hu, Y. C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, *37*(3), 1846-1853. http://dx.doi.org/10.1016/j.eswa.2009.07.081

Vapnik, V. (1998). *Statistical Learning Theory*. New York: Wiley.

Varmuza, K., & Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. CRC press. http://dx.doi.org/10.1201/9781420059496

Xie, C., Luo, C., & Yu, X. (2011). Financial distress prediction based on SVM and MDA methods: The case of Chinese listed companies. *Quality & Quantity*, *45*(3), 671-686. http://dx.doi.org/10.1007/s11135-010-9376-y

Copyrights