Journal of Scientific & Industrial Research
Vol. 78, May 2019, pp. 269-273

Bankruptcy Prediction of Family Firms Using Combined Classifiers

M A Fernández-Gámez[1], J Diéguez-Soto[2], José António C Santos[3] and J M de la Rosa[4]
[1,2,4] University of Malaga, 29071Malaga, Spain
[3] ESGHT and CIEO, University of Algarve, Campus da Penha, 8005-139 Faro, Portugal

Received 06 June 2018; revised 12 January 2019; accepted 20 March 2019

Literature has dealt with prediction of corporate bankruptcy since some decades ago, building classification methods to predict financial firm failure in different industries, countries and regions. However, and although the previous research is profuse, no paper in this literature has developed specific models for family firms. Family firms represent sixty percent of total firms in the EU and eighty percent in the USA, and it seems essential to evaluate their particular risk of bankruptcy. In this article we used combinations of classifiers (Naïve Bayes Classifier, Algorithm C4.5, Multilayers Perceptron and Support Vector Machine) through the AdaBoost algorithm, to develop an effective model to predict the bankruptcy of family firms.

Keywords: Bankruptcy prediction, family firms, AdaBoost, classifiers combination

Introduction

Prior research on bankruptcy prediction has utilized models focused on only one industry and models centred on several industries. There have been articles devoted to predicting bankruptcy in banks[1], in agricultural firms[2], manufacturing firms[3,4], quoted companies[5], hospitality firms[6,7], or in retail firms[8]. However, to our knowledge, previous literature has never used specific samples composed only by family firms, despite family firm's unique features. Family firms are unique because not only search for the maximization of economic benefits, as non-family firms do, but they also pursue the maximization of the socioemotional wealth. Another characteristic is their unique form of acquisition, learning, transmission and management of knowledge[9]. It includes the perpetuation of family values through the business continuity[10], the ability to exercise family influence and the conservation of family dynasty[11], and the preservation of the social capital of the family firm[12], among others. These non-financial goals of family firms generate positive aspects that may impact on bankruptcy, such as higher commitment, loyalty and pride in the business[13] or long-term perspective[14]. In this article we utilized a combination of different classifiers to build a bankruptcy prediction model for family firms. This study contributes to the existing literature on bankruptcy prediction in two ways. First, by building specific models for family firms, it may improve the efficiency of bankruptcy prediction, avoiding the bankruptcy process and their negative consequences for the economy, given the importance of family firms worldwide. Second, the relationship between the evolution of certain financial variables and the survival of a firm takes on specific connotations in the case of family firms. In fact, family firms aim unique noneconomic objectives that may have a great influence on the process of bankruptcy. These are, among others, their long-term perspective[15] or their higher commitment, loyalty and pride in the business[16].

*Author for Correspondence
E-mail: jasantos@ualg.pt

Methods

The models of bankruptcy prediction try to solve a classification problem. By means of a group of explanatory variables, a binary dependent variable is estimated, representing two possible situations: bankrupt and non-bankrupt firms. As a consequence, to disentangle this classification model, we should answer at least two questions: "What do explanatory variables offer a better explanation quality?" and "Why does classification method achieve a better accuracy test for a data set?" To solve the first question raised, we have utilized six attribute selection methods available in the collection of automatic learning algorithms for data mining tasks by Waikato University. This selection of methods included three algorithms evaluating subgroups of attributes, specifically two algorithms classified as filters (*Consistency Subset Eval* and

Classifier Subset Eval), which select and evaluate the attributes independently from the learning algorithm), and an algorithm working as a Wrapper (*Wrapper Subset Eval)*, which uses the performance of some classifiers to determine the desirable result from a subgroup. Moreover, another three algorithms classified as evaluators of individual attributes, implemented besides the Ranker method, allow us to order the attributes depending on their quality (C*hi Squared Attribute Eval*, *Gain Ratio Attribute Eval* y *Info Gain Attribute Eval*). Finally, with the aim of verifying the effectiveness of variable selection, we utilized the decision tree C4.5 to compare the classification accuracy before and after the variable selection. Regarding the second question raised, the choice of a classification method, we have opted for combining classifiers' methods. To that end, we have selected four of the most important individual classifiers within the supervised classification: Naïve Bayes Classifier (NBC), Algorithm C4.5 (C4.5), Multilayers Perceptron (MLP) and Support Vector Machine (SVM). Later, we have obtained conclusions regarding the combination of them through the Adaboost algorithm[17]. Figure 1 displays our own research design implemented in this paper. The process starts with the selection of attributes or explanatory variables, choosing the subgroups of variables with better results of classification and with greater quality. Next, individual classifiers are implemented with the selected variables, their classification accuracy is obtained and is compared with that achieved through the combination of classifiers.

Samples, data, and variables

In this paper we defined operatively a family firm based in two objective criteria. First, we identify family firms using a procedure based on the family name criteria, in line with earlier family-related research[18]. Second, we also demand that the family owns at least 50% of ownership[19]. Using these criteria, we have built three random samples of family firms with information for one (t–1), two (t–2) and three years (t–3) before the judicial declaration of bankruptcy, taking a sample of 102 bankrupt firms and 102 non-bankrupt firms, providing a final dataset of 204 firms. For each bankrupt firm we selected a non-bankrupt one which was comparable in terms of size and belonged to the same field of activity. Once we had the full dataset of firms (bankrupt and non-bankrupt), data were obtained from the annual accounts corresponding to the year prior to bankruptcy filing. The data was collected for the period 2005-2016 for firms whose activity was developed in Spain and Portugal. We obtained financial and corporate data from SABI (Iberian Balance Sheet Analysis System) database by Bureau van Dijk. Most literature addressing bankruptcy prediction has utilized financial variables as independent variables. In this research we have considered fifteen variables, and they have been selected because they were used in twenty or more prior studies devoted to bankruptcy prediction[20]. Most

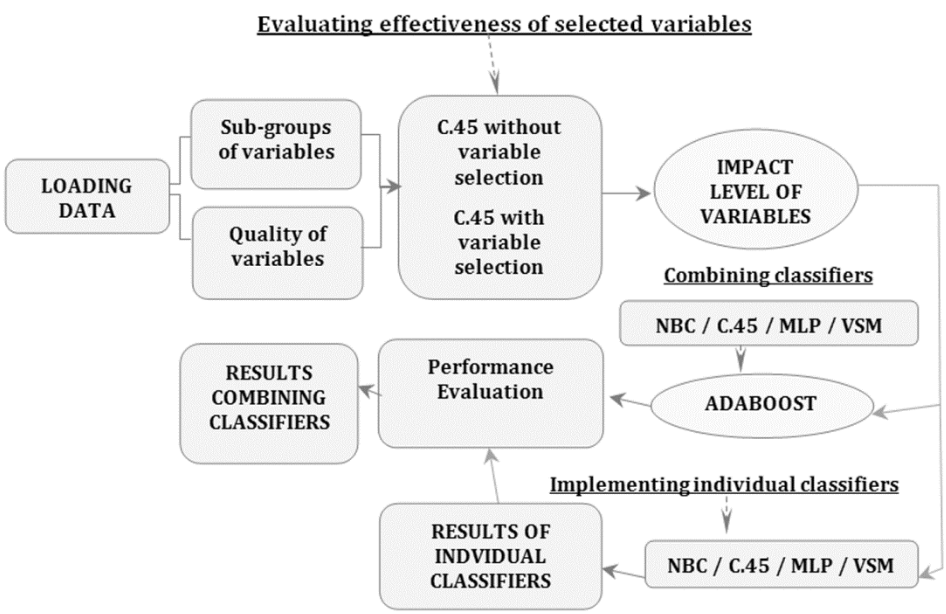

Fig. 1 —Research Design

of them are quantitative variables, and they measure profitability (V1, V8, V9), liquidity (V2, V3, V5, V7, V11), debt (V6, V10), efficiency (V4), capacity to meet debt (V12, V13) and size (V14). Moreover, we have control by sub-industry (V15). Finally, to measure dependent variables we used a dummy variable, which takes value 1 when the firm has gone bankrupt and 0, otherwise.

Empirical results

According to the methodological process proposed, firstly we have carried out the variable selection. As mentioned, to implement the selection we have used filter methods and methods based on models. Each and every evaluating method chose subgroups including from 5 to 7 variables for the three samples used. We have also performed three individual

Table 1 — Variables selection

t-1			t-2			t-3		
SUBGROUPS OF ATTRIBUTES SELECTION								
Consistency	Classifier	Wrapper	Consistency	Classifier	Wrapper	Consistency	Classifier	Wrapper
V2	V2	V2	V2	V4	V4	V1	V4	V4
V4	V4	V4	V4	V5	V5	V4	V6	V8
V5	V6	V5	V5	V6	V8	V5	V8	V9
V8	V8	V8	V8	V8	V12	V8	V14	V12
V9	V14	V12	V9	V14	V14	V10	V15	V14
V12			V12			V12		
V14			V14			V14		
INDIVIDUAL ATTRIBUTES RANK								
Chi Squared	Gain Ratio	Info Gain	Chi Squared	Gain Ratio	Info Gain	Chi Squared	Gain Ratio	Info Gain
V8	V8	V4	V8	V8	V4	V8	V8	V4
V4	V4	V8	V5	V4	V8	V4	V4	V8
V2	V14	V14	V4	V14	V14	V12	V14	V14
V14	V5	V5	V14	V5	V5	V14	V12	V13
V5	V2	V2	V15	V6	V10	V15	V6	V10
V12	V6	V9	V12	V2	V9	V5	V2	V9
V6	V12	V12	V6	V12	V12	V6	V5	V12
V9	V9	V6	V10	V11	V6	V9	V1	1V6
V15	V10	V10	V2	V10	V2	V2	V10	V2
V10	V15	V13	V9	V15	V13	V10	V13	V5
V1	V13	V15	V1	V13	V15	V1	V15	V15
V11	V11	V11	V11	V9	V11	V7	V9	V11
V7	V7	V7	V7	V7	V3	V11	V7	V3
V13	V1	V1	V3	V3	V1	V3	V1	V7
V3	V3	V3	V13	V1	V7	V13	V3	V1
SELECTED VARIABLES								
Variables	Frequency		Variables	Frequency		Variables	Frequency	
V2	6		V4	6		V4	6	
V4	6		V5	6		V8	6	
V8	6		V8	6		V14	6	
V5	5		V14	6		V12	6	
V14	5							

Note: Variables definition. V1: Profit for the period/Total assets; V2: Current assets/Current liabilities; V3: (Current assets–Current liabilities)/Total assets; V4: Revenue/Total assets; V5: (Cash and cash equivalents+Trade and other receivables)/ Current liabilities; V6: Non-current liabilities/Total assets; V7: Current assets/Total assets; V8: (Profit for the period+finance expenses+Income tax)/Total assets; V9: Profit for the period/Equity; V10: Total liabilities/Total assets; V11: Cash and cash equivalents/Total assets: V12: (Profit for the period+Amortization)/Total liabilities; V13: (Profit for the period+Amortization–Trade and other receivables for current year–Inventories for current year+Trade and other receivables for previous year+Inventories for previous year)/Total liabilities; V14: Natural logarithm of total assets; V15: Global Industry Classification Standard Code.

Table 2 — Accuracy classification

t-1				t-2				t-3			
Accuracy classification with individual classifiers (%)											
NBC	C4.5	MLP	SVM	NBC	C4.5	MLP	SVM	NBC	C4.5	MLP	SVM
94.78	94.30	92.09	91.02	93.45	93.06	89.41	88.90	91.01	90.32	89.15	89.00
Accuracy classification with AdaBoost (%)											
98.96				95.47				92.35			

evaluators to select variables, which offer an ordered list of these variables depending on their quality. To test the variable selection performed, we implement the classifier C4.5. Previously, we checked that, before the variable selection, the accuracy rates of data classification were 90.33%, 88.41% and 85.03% for t-1, t-2 y t-3, respectively. The accuracy rates obtained implementing C4.5, after different variable selections, increased in all the cases, achieving a lower number of errors with Wrapper and Classifier methods. Later, the choice of these variables was made considering the results obtained by the different methods used. For this, we have taken into consideration the appearance frequency of the variables in the distinct subgroups. That selected subgroups by methods with lower error rates have been considered especially relevant. Next, we order the variables by selection frequency, and specifically for Ranker methods, we select those variables occupying the first positions (Table 1).

Table 2 shows the results obtained implementing individual and combined classifiers from the variable selection. For individual classifiers, the highest accuracy classification is achieved with NBC, being of 94.78%. However, and as we expected, the classification rate implementing to AdaBoost algorithm is improved in all the samples used.

The results show that it is possible to obtain an efficient model to predict bankruptcy in family firms by combining classifiers, regardless the variables used. Specifically, the AdaBoost algorithm is a stable combination method in accuracy terms, enhancing significantly the results obtained in prior literature with regard to other firm groups[3]. Secondly and according to the results, we have confirmed that there is a group of variables susceptible to being considered to predict bankruptcy in family firms. The group of variables demonstrates that liquidity, profitability, efficiency or capacity to meet debt to tackle with financial crises in family firms are essential aspects. Thus, a lower proportion of liquid assets (V2, V5), scant return on assets (V8), deteriorated efficiency levels (V5) and insufficient generation of ordinary resources to meet debt (V12) are signs of early warning. Finally, the firm size (V14) has also been considered relevant in our study in such a way that smaller family firms are more likely to go into bankruptcy.

Discussion and Conclusion

The results obtained have shown that through the combination of classifiers it is likely to achieve an efficient method to predict bankruptcy in family firms, with accuracy rates of 98.96%, 95.47% and 92.35%, using information from 1, 2 and 3 years before the bankruptcy occurred. Moreover, the models of bankruptcy prediction have confirmed that liquidity, profitability, efficiency and capacity to meet debt are the best variables to predict bankruptcy in family firms. Our findings have several implications to literature on family firms and bankruptcy prediction. It has implications for the ongoing discussion regarding the factors explaining the financial failure and also for the building of models predicting bankruptcy. We thereby extend and challenge current literature by studying a unique element that affects the financial failure and as a consequence it should be taken into account when predicting bankruptcy of family firms. This paper has also strong implications for practitioners, managers and directors of firms as they may utilize the models displayed here to detect earlier financial unbalances and to carry out more accurate diagnosis. Consequently, these models would allow them to show useful evidence to improve their decisions and the continuity and viability of their firms.

References

1 López F J, & Pastor I, Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks, *Expert Syst Appl*, 42(6) (2015), 2857-2869

2 Karas M, Reznakova M, & Pokorny P, Predicting bankruptcy of agriculture companies: Validating selected models, *Pol J Manage Studies*, **15(1**) (2017), 110-120

3 Callejón A M, Casado A M, Fernández M A, & Peláez J I, A System of insolvency prediction for industrial companies using a financial alternative model with neural networks, *Int J Comput Int Sys*, 6(1) (2013), 29-37

4 Zhang C, He W & Hao, R, Comprehensive estimation of the financial risk of iron and steel enterprise-based on carbon emission reduction, *J Sci Ind Res*, 75(3) (2016), 143-149

5 Alaminos D, del Castillo A, & Fernández M A, A Global Model for Bankruptcy Prediction. PLoS ONE, 11(11) (2016), 1-18, https://doi.org/10.1371/journal.pone.0166693

6 Fernández-Gámez M A, Cisneros-Ruiz A J, & Callejón-Gil A, Applying a probabilistic neural network to hotel bankruptcy prediction, *Tour & Manage Studies*, 12(1) (2016), 40-52

7 Cordova J P & Alberton A, Model categorisation of family hotel ventures and life cycle: A study in the cities of Itapema and Porto Belo/SC/Brazil, *Tour & Manage Studies*, 10(2) (2014), 94-103

8 Keener M, Predicting the financial failure of retail companies in the United States, *Int J Bus Econ Res*, 11(8) (2013), 373-380

9 Peñalver A J B, Santos José António C., Conesa J A B, & Santos M C, Influence of cooperation and collaboration in ICT in knowledge management, *J Sci Ind Res*, 77(6), 313-317.

10 Dressler J B, & Tauer L, Socioemotional wealth in the family farm, *Agri Fin Rev*, 75(3) (2015), 403-415

11 Baek H Y, Cho D D, & Fazio, P L, Family ownership, control and corporate capital structure: Examination of small capitalizaction public firms, *J Fam Bus Manage*, 6(2) (2016), 169-185

12 Carney M, Van Essen M, Gedajlovic E R, & Heugens, P P, What do we know about private family firms? A meta-analytical review, *ET&P*, 39 (2015), 513-544

13 Diéguez-Soto J, & López-Delgado P, Does family and lone founder involvement lead to similar indebtedness? *J Small Bus Manage*, (online) (2018), https://doi.org/10.1111/jsbm.12422

14 Rojo-Ramírez A A, & Martínez-Romero M J, Required and obtained equity returns in privately held businesses: The impact of family nature - evidence before and after the global economic crisis, *Rev Manage Sci*, (online) (2017), 1-31, https://doi.org/10.1007/s11846-017-0230-7

15 Hauck J, Suess-Reyesm J, Beck S, Prügl R, & Hermann F, Measuring socioemotional wealth in family-owned and -managed firms: A validation and short form of the FIBER Scale, *J Fam Bus Strategy*, 7(3) (2016), 133-148

16 Sageder M, Mitter C, & Feldbauer- Durstmüller B, Image and reputation of family firms: A systematic literature review of the state of research, *Rev Manage Sci*, 12(1) (2018), 335-377

17 Baig M M, Awais M M, & El-Alfy E M, AdaBoost-based artificial neural network learning, *Neurocomputing*, **248**(C) (2017), 120-126

18 López-Delgado P, & Diéguez-Soto J, Lone founders, types of private family businesses and firm performance, *J Fam Bus Strategy*, 6(2) (2015), 73-85

19 Instituto de la Empresa Familiar, La empresa familiar en España (2015). IEF, Madrid, 2015

20 Zięba M, Tomczak S K, & Tomczak J M, Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, *Expert Syst Appl*, 58 (2016), 93-101