Journal of Scientific & Industrial Research
Vol. 78, November 2019, pp. 742-745

A Multi-Class Model to Predict the Result of the Legal Insolvency Proceedings

Daniel Pastor-Vega, Sergio M. Fernández-Miguélez, Julio Diéguez-Soto and Manuel A Fernández-Gámez*
University of Malaga, 29071, Malaga, Spain

Received 14 September 2018; revised 15 May 2019; accepted 14 August 2019

A small number of studies have been carried out to build models with which to predict the results of the insolvency proceedings. In addition, there are no models that have demonstrated a high predictive capacity for all situations in which such legal processes can end. The proposal of this study is the construction of a multi-class model that predicts with high precision the possible future situations of companies that are in legal insolvency proceedings. The results of this study reveal that using the Naive Bayes classifier the model achieves an accuracy of more than 91% and that the best predictors are variables related to the profitability, efficiency and volume of resources generated by the companies.

Keywords: Insolvency proceedings, Insolvency prediction, Business bankruptcy, Financial distress, Multi-class classifiers, Naive Bayes

Introduction

Insolvency prediction models have been successfully anticipating the future financial situation of companies, classifying them as solvent companies, in financial difficulties or bankruptcy[1,2]. However, other important aspects of corporate insolvency have been poorly modeled. Such is the case of models that predict the outcome of legal insolvency proceedings. These models are new and still little developed, and greater precision of them is necessary. Only the works of Kim[3],Staszkiewicz &Witkowski[4] and Camacho-Miñano, Pascual-Ezema & Urquía-Grande[4] have recently addressed this research topic. Kim[2]used a simple z-score insolvency model, and his results showed that current laws might be partly responsible for the difference in the z-score threshold up to the insolvency situation. Staszkiewicz& Witkowski[4] and Camacho-Miñano, Pascual-Ezema& Urquía-Grande[4] analysed the problem of mutual use the insolvency and bankruptcy variable for business failure applying logistic regression. They showed that the legal solvency criterion is adequate to classify companies in financial distress. Due to this scant evidence, the financial literature demands new research that responds to the prediction of the situation of the companies that, after entering into a bankruptcy legal proceeding, can either achieve viability through an agreement with creditors or are liquidated due to lack of viability[6]. Having detected this gap in the literature, the objective of this study is to develop a model capable of predicting the possible future situations of companies in the legal insolvency proceedings. This model, then, solves a multi-class problem where the categories to be predicted are three, in reference to solvent companies, to companies that have continuity capacity through agreements with creditors, and to companies that will be liquidated (bankruptcy). The proposed model would reduce the prediction bias, eliminating those that have a structure capable of successfully managing illiquidity problems.

*Author for Correspondence
E-mail: mangel@uma.es

Methodology

The previous literature on prediction models indicates the good results obtained by the NB classifiers[8], DT[9] and MDA[10], and also the excellent possibilities of the hybrid classification methods[11,12]. Therefore, in the present investigation, these three classification techniques are used, combined with a modification of the algorithm called C4.5 developed by Quinlan[13]. Specifically, the applied algorithm consists in testing each classifier with different sets of independent variables to find out which of these sets provides the most accurate classification. Since we have 17 independent variables, testing all possible non-empty sets of variables would involve 217 tests, which would result in a vast number. Therefore, 1000 sets of variables are selected at random, which give rise to tests that end in a reasonable time. A loop is then executed, which in each iteration examines one of those 1000 sets of variables, and uses it to classify the data set. To do this, ten repetitions are made for each set of variables, in such a way that in each repetition the set of available samples is randomly divided by 80% for training samples, 10% for validation samples and 10% for samples of testing. Once the examination of the 1000 sets of variables has been done, it is necessary to choose the set that offers the best performance. To do this, we take the set of variables that yield the maximum of classification hits on the validation set (averaged over the ten repetitions that have been made for each set). On the other hand, the average number of hits on the test set is reported (again averaged over the ten repetitions that have been made for each set). In the second phase, the aforementioned computational algorithm starts from the set of independent variables that has given better results in the previous phase. With these variables, 100 repetitions are made, taking at random, for each repetition, 80% of the data to train and 20% for the test. Once the 100 repetitions are made, the final classification percentages are obtained using the average of the correct answers on the test data set (averaged over 100 repetitions).

Variables and sample

In the present study, a total of 17 independent variables have been considered. Specifically, 12 of these are variables of a financial nature and have been selected from those used in 20 or more previous insolvency prediction studies[14,15,16]. In addition, other five explanatory variables have been incorporated. One of them is a binary type to gather the legal form of the company, a second one for the activity sector, and another two to inform about the type of corporate governance structure and the number of participants in eachgovernance structure. Finally, another variable to control the size of firms by the size of their assets. In addition to the explanatory variables, a multi-class variable has been used, which will be the dependent variable, to identify the firms in the sample in three categories (1, solvent companies; 2, firms that get an agreement after a legal insolvency process;and 3, companies that are liquidated after the legal insolvency process). Table 1 shows the definition of the independent variables included in the model.

The information corresponding to the selected variables has been obtained from a sample of Spanish companies that have provided data corresponding to one year (M1), two years (M2), and three years before the beginning of the legal insolvency process (M3). For this, it has been necessary to use different sources of information. The identification of the firms in the sample has been made thanks to the collaboration of a Spanish Commercial Court, which has facilitated a random selection of firms which have been solvent, in agreement, or liquidation after their corresponding insolvency proceedings. This information corresponds to the period between 2006 and 2016. The financial

Table1 — Independent variables definition

Code	Description
V1	Profit for the period/ Total Assets
V2	Currentassets/ Currentliabilities
V3	(Current assets – Currentliabilities) / Total assets
V4	(Profit for the period + Finance expenses + Income tax expense) / Total assets
V5	Sales / Total assets
V6	(Cash and cash equivalents + Trade and other receivables) / Current liabilities
V7	Non current assets / Total assets
V8	Currentassets / 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 + Amortisation – Trade and other receivables of the current year –Inventories of the current year + Trade and other receivables of the previous year + Inventories of the previous year) / Total Liabilities
V13	Ln total assets
V14	Legal form
V15	NACE Code
V16	Corporate governance structure (0, Sole administrator; 1, Joint administrator/ The board of directors)
V17	Number of members of the corporate governance structure

Table 2 — Classificationresults(%).

	NB		MDA		DT	
	Validation	Testing	Validation	Testing	Validation	Testing
M1	93,02	91,93	90,61	80,32	84,47	74,61
M2	88,01	86,76	80,27	62,83	79,83	72,73
M3	83,35	80,51	85,10	78,11	77,72	71,03

NB: Naive Bayes; MDA: Multi-Discriminant Analysis; DT: Decision Tree; M1: information of 1 year before the beginning of the legal insolvency process; M2: information 2 years before the beginning of the legal insolvency process; M3: information of 3 years before the beginning of the legal insolvency process

and corporate information of all the companies selected with the above criteria has been provided by the database of the Iberian Balance Sheet Analysis System of Bureau Van Dijk, which includes more than 850,000 Spanish companies.

Results and Discussion

A comparison of the results obtained with the multi-class classifiers used confirms that the best prediction results are obtained with NB since it has reached the maximum percentage of success in testing (91.93% for M1, 86.76% for M2, and 80.51% for M3). Secondly, MDA has also obtained excellent results (higher than 80% for M1). Last, in terms of classification power, DT has shown lower values than the other two classifiers (table 2).

On the other hand, the most significant variables selected by the highest-fit classifier (NB) appear in Table 3. The variables that have always been significant for M1, M2 and M3, are V4, V5, V9, V12 and V13. These variables refer to the profitability, efficiency and resources generated by the companies in the sample. In addition, there are other variables, that although they are not taken into consideration simultaneously for M1, M2 and M3, they are included in two of them. These variables are V7, V8 and V11 (which measure liquidity and indebtedness), and the qualitative variables V16 and V17 (which provide information on the corporate governance of companies).

Table 3 — Significant variables.

	Variables
M1	V4, V5, V7, V8, V9, V11, V12, V13, V16, V17
M2	V4, V5, V7, V8, V9, V12, V13, V16, V17
M3	V4, V5, V6, V9, V11, V12, V13, V15

M1: information of 1 year before the beginning of the insolvency legal process; M2: information 2 years before the beginning of the legal insolvency process; M3: information of 3 years before the beginning of the legal insolvency process

Our model is based on the current terminology of the Spanish Bankruptcy Law that includes the agreement and the liquidation. The previous models have used firms in procedures for suspension of payments and bankruptcy, but have been overcome by the legislative changes produced in recent times in international bankruptcy legislation. For example, Alfaro, Gámez& García[17] developed a multiclass model that included solvent companies, insolvent companies and dissolved and acquired companies. In our opinion, this classification is surpassed by the model proposed in this work, since both solvent companies, as well as insolvent companies that obtain an agreement or go to liquidation, can be subject to restructuring, global sale and sale of productive units within of the process itself, even with operating advantages over solvent companies. In addition, the judicial resolution that resolves the liquidation of a bankruptcy proceeding simultaneously involves the dissolution of the business, but there are many other cases in the mercantile legislation in which the dissolution is also proposed and do not necessarily refer to situations of insolvency.

Conclusions

As a result of the results obtained, we believe that the present study contributes to business financial knowledge by developing a multi-class prediction model that discriminates with high precision between solvent and insolvent companies, distinguishing between those that obtain an agreement with their creditors or end up in liquidation. This multiclassification of the proposed model has not been offered up to now by the previous literature and can provide criteria to improve the financial situation of companies in financial distress. The above considerations make us think that the proposed model would be useful to achieve an improvement in the efficiency of the legal insolvency proceedings. The model may be useful when considering the future structures that insolvent companies must keep in the approaches of their agreements with creditors and, if this was not possible, it might be helpful to accelerate settlement procedures in the quickest and efficient possible way. Therefore, it might provide advantages such as reducing the cost of the procedure in terms of professional fees and reducing the labour cost through the advance payment of redundancy schemes. Equally, it may provide creditors with a priori knowledge of their possibilities of collection and may facilitate the firm global sale or the specific sale of the production units before undertaking the firm's cessation of activity and the abandonment of assets of the firms.

References

1 Altman E I, Iwanicz-Drozdowska M, Laitinen E K &SuvasA, Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-Score model, *J Int Fin Manag Acc*, **28**(2) (2017) 131-171.

2 Martins, A I, efficiency determinants in Portuguese banking industry: an application through fractional regression models, *Tour & Manage Stud*, **14**(2), (2018), 63-71.

3 Kim J, Bankruptcy and institutions: theory and empirical evidence from Korea and the United States, *Emerg Mark Financ Tr*, **54**(1) (2018) 219-233.

4 Staszkiewicz P&Witkowski B, Failure models for insolvency and bankruptcy. In: Jajuga K., Locarek-Junge H., Orlowski L. (eds) *Contemporary Trends and Challenges in Finance*, *Springer, Cham*, 2018.

5 Camacho-Miñano M M, Pascual-Ezema D&Urquía-Grande E, In the efficiency of bankruptcy law: empirical evidence in Spain, *Int Insolv Rev*, **22**(2013)171-187.

6 Alaminos D, del Castillo A &Fernández MA, A Global Model for Bankruptcy Prediction,*PLoS ONE*, **11**(11) (2016), e0166693.doi:10.1371/journal. pone.0166693

7 Law 22/2003 of 9 July 2003.

8 Netti K & Radhika YA, Model for accurate prediction in GeoRSSdata using naive bayes classifier, *J Sci Ind Res*, **76**(8)(2017) 473-476.

9 Mozumder A, Taking decisions the Fuzzy way! *Sci Rep*, **63**(8)(2016) 39-41.

10 Mekparyup J &Saithanu K, Prediction of groundwater quality in the east of Thailand with discriminant analysis method, *Indian J Geo-Mar Sci*, **47**(9)(2018) 1847-1850.

11 Chen M Y, A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering, *Inform Sciences*, **220**(2013) 180-195.

12 Farid D M, Zhang L, Rahman C M, Hossain M A& Strachan R, Hybrid decision tree and naive Bayes classifiers for multi-class classification tasks, *Expert Syst Appl*,**41** (2014) 1937-1946.

13 Quinlan J, *C4.5: Programs for Machine Learning*, Morgan Kaufmann, 1993.

14 Bellovary J, Giacomino D& Akers M, A Review of bankruptcy prediction studies: 1930 topresent, *J FinancEduc*, **33**(2007)1-42.

15 Fernández-GámezM A, Cisneros-Ruiz AJ & Callejón-Gil A, Applying a probabilistic neural network to hotel bankruptcy prediction. *Tour & Manage Stud*, **12**(1)(2016)40-52.

16 Fernández-Gámez, M A, Diéguez-Soto, J, Santos, José António C. & Rosa, J M, Bankruptcy prediction of family firms using combined classifiers, *J Sci Ind Res*, **78**(5), (2019) 269-273.

17 Alfaro E, GámezM&GarcíaN, Multiclass corporate failure prediction by Adaboost.M1, *Int Adv Econ Res*, **13**(3)(2007) 301-312.